



REPUBLIKA E SHQIPËRISË
UNIVERSITETI POLITEKNIK I TIRANËS
FAKULTETI I INXHINIERISË ELEKTRIKE

DISERTACION
PËR MBROJTJEN E GRADËS SHKENCORE
“DOKTOR”

**Parashikimi i konsumit të energjisë në ndërtesat
rezidenciale dhe publike me anë të metodave
të inteligjencës artificiale**

Përgatitur nga
Msc. Ing. Aulon Shabani

Udhëheqës shkencor
Prof. Dr. Orion Zavalani

Tiranë 2018



REPUBLIKA E SHQIPËRISË
UNIVERSITETI POLITEKNIK I TIRANËS
FAKULTETI I INXHINIERISË ELEKTRIKE

DISERTACION

i

paraqitur nga

Msc. Ing. Aulon Shabani

për mbrojtjen e gradës shkencore

“DOKTOR”

**Parashikimi i konsumit të energjisë në ndërtesat
rezidenciale dhe publike me anë të metodave
të inteligjencës artificiale**

Mbrohet me datë 13.02.2018 para Jurisë të miratuar nga
Rektori i UPT me shkresën Nr.1509/1 datë 19.12.2017.

- | | |
|---------------------------------|-------------------|
| 1. Prof. Dr. Aida SPAHIU | Kryetar |
| 2. Prof. Dr. Aleksandër XHUVANI | Anëtar (oponent) |
| 3. Prof. Dr. Petrika MARANGO | Anëtar |
| 4. Prof. Dr. Kozeta SEVRANI | Anëtar (oponente) |
| 5. Prof. Asoc. Flamur BIDAJ | Anëtar |

Deklaratë e Autorit mbi Originalitetin e Punimit

Me anë të kësaj deklarate, unë i nënshkruari, Aulon Shabani, deklaroj se jam autori i punimit të Doktoratës me titull:

“ Parashikimi i konsumit të energjisë në ndërtesat rezidenciale dhe publike me anë të metodave të inteligjencës artificiale”

Deklaroj në përgjegjësinë time personale, se i gjithë punimi i doktoraturës, rezultatet si dhe përfundimet shkencore janë origjinale dhe në përputhje me parimet e antiplagjaturës dhe etikës profesionale.

I gjithë informacioni i marrë nga literatura dhe i përfshirë në punim është kryer në përputhje të plotë me rregullat akademike dhe etikën profesionale, duke cituar me përpikmëri çdo referencë.

Autori

Aulon Shabani

Nënshkrimi

Përmbajta

Mirënjohje	x
Përmbledhja	xi
Abstract	xiii
Hyrje	1
1 HISTORIKU I PARASHIKIMIT	4
1.1 Sistemet e monitorimit dhe matjes së energjisë.	4
1.2 Metodatat e parashikimit	8
1.2.1 Metodatat inxhinierike	8
1.2.2 Metodatat statistikore	10
1.2.3 Rrjetat artificilale neurale	12
1.2.4 Regresioni me vektorët ndihmës	15
1.2.5 Pemët e regresionit	16
1.2.6 Metoda e arsyetimit bazuar tek rastet	17
1.3 Përmbledhje e kontributeve kryesore	18
2 NJOHURI TEORIKE MBI METODAT	20
2.1 Metodatat inxhinierike	20
2.2 Metodatat e Inteligjencës Artificiale	23
2.2.1 Mësimi i supervizuar	25
2.2.2 Regresioni linear	26
2.2.3 Rrjetat neurale artificiale	28
2.2.4 Makinat me vektorë ndihmës (MVN)	32
2.2.5 Pemët e regresionit	38

2.2.6	Pyjet e rastësishëm	40
3	ARSYETIMI I BAZUAR TEK RASTET: VËSHTRIM	
	TEORIK	42
3.1	Njohuria	42
3.2	Përcaktimet bazë	43
3.3	Përcaktimet e ngjashmërisë	45
3.3.1	Distanca e Minkowskit	46
3.3.2	Spërdredhja dinamike në kohë	48
3.3.3	Ngjashmëria sipas formës	49
3.4	Hapat e implementimit të një modeli ABR	51
3.4.1	Ndërtimi i rasteve	51
3.4.2	Përzgjedhja e rasteve	52
3.4.3	Adaptimi i rastit	54
4	PËRDORIMI I ABR NË PARASHIKIM	56
4.1	Çështje praktike	57
4.1.1	Rrjedha e funksionimit	57
4.1.2	Mjedisi eksperimental	57
4.1.3	Të dhënat dhe përpunimi i tyre	57
4.1.4	Vlerësimi i performancës	60
4.2	Implementimi i modelit	61
4.2.1	Ndërtimi i rasteve	61
4.2.2	Grupimi i rasteve të ngjashme	63
4.2.3	Metodat e ngjashmërisë	64
4.2.4	Ngjashmëria e rastit	68
4.2.5	Përcaktimi i peshës së variablave	69
4.2.6	Adaptimi i zgjidhjes	70
4.2.7	Modeli statik i parashikimit	71
4.2.8	Krahasimi i modeleve	73
4.2.9	Modeli dinamik i parashikimit	74
4.3	Përmbledhje	75
5	PARASHIKIMI I KONSUMIT NË SKENARË TË NDRYSHËM	77
5.1	Parashikimi statik “offline”	78

5.1.1	Përshkrimi i të dhënave	78
5.1.2	Modelet e përdorura	81
5.1.3	Parashtrimi i problemit-Modelimi i të dhënave	84
5.1.4	Rezultatet eksperimentale	87
5.2	Parashikimi dinamik 'online'	97
5.2.1	Rregulli i modelit adaptiv	97
5.2.2	Modeli inkremental i pemëve	100
5.2.3	Metodat e bazuara tek shembujt me të dhëna në kohë reale	101
5.2.4	ABR si model i parashikimit online	103
5.2.5	Rezultatet eksperimentale	104
5.3	Diskutime	107
6	PARASHIKIMI ME INTERVAL I KONSUMIT	109
6.1	Historiku	110
6.2	Parashikimi transduktiv	110
6.3	Parashikimi induktiv	112
6.4	Vlerësimi i jokonformitetit	113
6.5	Implementimi i metodës	115
6.6	Rezultatet eksperimentale	116
6.7	Përfundime	123
7	PËRMBLEDHJE	124
7.1	Vazhdimësia	126
Shtesa A	Parashikimi i konsumit me ABR	127
A.1	Kodi kryesor për parashikuesin ABR	127
A.1.1	Ndarja e ditëve në raste me raste prej 9 orësh	128
A.1.2	Ndërtimi i bazës së rasteve	129
A.1.3	Vlerësimi i ngjashmërisë midis variablave të rasteve	129
A.1.4	Ngjashmeria midis rasteve	130
A.1.5	Përshtatja e zgjidhejs për parashikim	130
A.2	Parashikuesi bazë i cili parashikon vlerën e konsumit sa mesatarja e gjithë pjesës së trajnimit	132
A.3	Parashikuesi bazë i cili parashikon vlerën e konsumit sa mesatarja e tre orëve të mëparshme	132

A.4	Parashikuesi bazë i cili parashikon vlerën e konsumit sa ora e mëparshme	133
A.5	Parashikimi duke përdorur rrjetat neurale	134
A.5.1	Gjenerimi i të dhënave të trajnimit	134
A.5.2	Gjenerimi i të dhënave të testimit	134
A.5.3	Testimi i rrjetës neurale	134
Shtesa B Algoritmat e parashikimit ‘offline’ për 5 ndërtesa		136
B.0.1	Gjenerimi i matricës së të dhënave nga ato origjinale	136
B.0.2	Testimi i performancës së rrjetës neurale	137
B.0.3	Testimi i performancës së pemës së regresionit	139
B.0.4	Testimi i performancës së pyllit të rastësishëm	140
B.0.5	Testimi i performancës së regresionit me vektorët ndihmës	142
B.1	Rezultatet e parashikimit ‘offline’ për skenarë të ndryshëm	145
Shtesa C Intervalet e parashikimit		150
Bibliografia		154

Figurat

1	Konsumi në shkallë globale i energjisë elektrike	2
1.1	Instalimi total i matësve inteligjent sipas tipit të ndërtesës	5
1.2	Bloku diagrama e një matësi inteligjent	5
2.1	Modeli i plotë termik për një ndërtesë	21
2.2	Modeli i thjeshtuar i simulimit	23
2.3	Motodat e mësimit të mbikëqyrur	24
2.4	Përshtatja e një modeli linear në të dhënat.	27
2.5	Struktura tipike e një rrjete neurale me shumë shtresa.	29
2.6	Modeli i një neuroni artificial.	29
2.7	Modeli i një vektori ndihmës.	33
2.8	Modeli i një transformimi me anë të funksionit Kernel.	35
2.9	Zona ϵ e një regresioni me vektorë ndihmës	36
2.10	Modeli i një peme.	38
2.11	Procesi i ndërtimit të një pylli të rastësishëm	41
3.1	Një cikël i një modeli ABR	44
3.2	Llogaritja e distancës së Minkowskit.	47
3.3	Matja e distancës fikse Euklideane dhe matja në formë të spërdredhur	48
3.4	Analiza sipas formës e dy serive kohore X dhe Y.	49
3.5	Përcaktimi i ngjashmerisë sipas kosinuset midis dy vektorëve të nënsekuencave.	50
3.6	Skema parimore e një modeli ABR.	51
3.7	Skematika e një procesi CBR	54
3.8	Rasti i mësimit për një sistem CBR	55
4.1	Bloku diagrama e një modeli ABR për parashikimin e konsumit të energjisë.	58

4.2	Të dhënat e konsumit të energjisë për të gjithë periudhën e matur.	59
4.3	Skematizimi i grupimit të rasteve të ngjashme.	64
4.4	Transformimi i distancës Euklidiane në ngjashmëri	65
4.5	Krahasimi i distancës Euklidiane midis serive kohore.	67
4.6	Paraqitja e një sekuence kohe si një grupim vektorësh	68
4.7	Transformimi në amplitudë i vlerave të konsumit të energjisë . . .	71
4.8	Performanca e modelit online ABR.	75
5.1	Konsumi javor i energjisë në secilën ndërtësë	79
5.2	Konsumi vjetor i energjisë në secilën ndërtësë	80
5.3	Skematika e parashikimit të konsumit të energjisë	87
5.4	Konsumi i parashikuar i energjisë në ndërtësën e parë për një javë.	94
5.5	Konsumi i parashikuar i energjisë në ndërtësën e dytë për një javë.	94
5.6	Konsumi i parashikuar i energjisë në ndërtësën e tretë për një javë.	95
5.7	Konsumi i parashikuar i energjisë në ndërtësën e katërt për një javë.	95
5.8	Konsumi i parashikuar i energjisë në ndërtësën e pestë për një javë.	96
5.9	Gabimet online për ndërtesën e parë	105
5.10	Gabimet online për ndërtesën e dytë	105
5.11	Gabimet online për ndërtesën e tretë	106
5.12	Gabimet online për ndërtesën e katërt	106
5.13	Gabimet online për ndërtesën e pestë	107
6.1	Paraqitja skematike e parashikimit me interval.	116
6.2	Intervalet e parashikimit në ndërtesën 1 parashikuar me RN për nivel besueshmërie 95%.	118
6.3	Intervalet e parashikimit në ndërtesën 1 parashikuar me PR_{ast} për nivel besueshmërie 95%.	118
6.4	Intervalet e konsumit në secilën ndërtesë për çdo vlerë besueshmërie.	120
B.1	Parashikimi për ndërtesën 1 me pyjet e rastësishëm	145
B.2	Parashikimi për ndërtesën 2 me pyjet e rastësishëm	146
B.3	Parashikimi për ndërtesën 3 me pyjet e rastësishëm	147
B.4	Parashikimi për ndërtesën 4 me pyjet e rastësishëm	148
B.5	Parashikimi për ndërtesën 5 me pyjet e rastësishëm	149
C.1	Intervalet e parashikimit me pyjet e rastësishëm duke përdorur jokonformitetin e normalizuar.	150

C.2	Intervalet e parashikimit me pyjet e rastësishëm duke përdorur jokonformitetin jo të normalizuar.	151
C.3	Intervalet e parashikimit me rrjetat neurale duke përdorur jokonformitetin e normalizuar.	152
C.4	Intervalet e parashikimit me rrjetat neurale duke përdorur jokonformitetin jo të normalizuar.	153

Tabelat

1.1	Kategoritë e bazave të të dhënave.	6
4.1	Përshkrimi i variablave dhe njësitë matëse	59
4.2	Modelet e rasteve, informacioni në rastin e ri dhe rastin ekzistues	62
4.3	Shmangiet specifike për secilën variabël.	66
4.4	Përshkrimi i variablave dhe njësitë matëse	70
4.5	Rezultatet e modelit offline në termat e CV-RMSE (%)	74
5.1	Modeli i vlerësimit kryq (5-fold cross validation).	81
5.2	Përshkrimi i variablave, vlerat minimale dhe maksimale të shmangies së tyre	84
5.3	Rezultatet për vektorë të ndryshëm të hyrjeje	86
5.4	Rezultatet e parashikimit për ndërtesën e parë.	89
5.5	Rezultatet e parashikimit për ndërtesën e dytë.	90
5.6	Rezultatet e parashikimit për ndërtesën e tretë.	91
5.7	Rezultatet e parashikimit për ndërtesën e katërt.	92
5.8	Rezultatet e parashikimit për ndërtesë e pestë.	93
6.1	Rezultatet e parashikimit me interval për ndërtesat 1-3	121
6.2	Rezultatet e parashikimit me interval për ndërtesat 4-5	122

Mirënjohje

Në vend të parë dua të falenderoj udhëheqësin tim Prof. Dr. Orion Zavalani i cili gjatë viteve të doktoraturës më ka dhënë një ndihmesë shumë të vyer, duke më konsultuar dhe këshilluar në mënyrë të vazhdueshme. Duke më propozuar këtë temë, krijoi një hapësirë të re kërkimi për mua, duke mësuar aftësi edhe teknika krejt të panjohura për mua.

Një hapësirë tepër të veçantë në finalizimin e kësaj teme zë edhe Prof. Dr. Eyke Hüllermeier, ndihma e të cilit pati një vlerë të paçmueshme për mua. Gjatë qëndrimit tim në grupin e Sistemeve Inteligjente, Universiteti i Paderborn më dha mundësinë e njohjes dhe përvetësimit të teknikave të parashikimit me inteligjencën artificiale. Me mirënjohjet më të thella dua të falenderoj udhëheqësit e mi Prof. Orion Zavalani dhe Prof. Eyke Hüllermeier pa ndihmën e të cilëve kjo temë nuk do të ishte e realizueshme.

Falenderime të shumta edhe për bashkpunëtorët dhe kolegët, ku vijon Adil Paul i cili dha një ndihmë shumë të madhe në rrugëtimin tim, po ashtu Radu Platon dhe Ammar Shaker. Diskutimet, ndihmesat sigurimi i të dhënave ishin një faktor vendimtar në kompletimin e këtij punimi. Po ashtu, dua të falenderoj Denis Panxhin për ndihmën dhe motivimin e tij gjatë kësaj kohe. Një tjetër falenderim qëndron për Ilda Kazanin ku bashkpunimet, diskutimet dhe këshillat e saj më kanë ndihmuar në përmirësimin e kërkimit shkencor gjatë kësaj periudhe.

Përzemërsisht të falenderoj kolegët e Universitetit Politeknik, e Fakultetit Inzhinierisë Elektrike dhe Departamentit Elektroteknikës të cilët me këshillat dhe motivimet më të mira e kanë bërë më të lehtë këtë rrugëtim. Një falenderim i veçantë edhe për të gjithë profesorët e shkollës së doktoraturës.

Në fund, por më të rëndësishmit qëndron familja time. Falenderime të panumërta shkojnë për nënën time, tim atë, Alketin dhe Klaudian. Pa ndihmën e të cilëve nuk do të kisha përkushtimin që kisha. Falenderoj Klaudian për durimin dhe motivimin gjatë qëndrimit tim në Gjermani.

Përmbledhja

Konsumi i energjisë në ndërtesa varet nga një sërë faktorësh, siç përmendim kushtet atmosferike, kushtet e funksionimit, sistemet e instaluara, struktura dhe karakteristikat, sjellja e banorëve. Kjo marrëdhënie komplekse midis tërë faktorëve e bën analizën dhe parashikimin e konsumit një problem të vështirë për t'u zgjidhur me qëllim arritjen e rezultateve të një saktësie tepër të lartë. Kjo tezë fokusohet në parashikimin e konsumit të energjisë për ndërtesa me funksione të ndryshme duke përdorur metodat e inteligjencës artificiale.

Në tezë një hapësirë e veçantë i kushtohet sistemeve të matjes dhe monitorimit në kohë reale të energjisë, pasi implementimi i metodave të parashikimit kërkon medoemos ekzistencën e të dhënave. Më tej, vijon një vëzhgim mbi teknikat më të fundit të përdorura në parashikim si, metodat inxhinierike të simulimit, metodat statistikore dhe metodat e inteligjencës artificiale. Duke qënë se në fokus kemi metodat e inteligjencës artificiale të cilat ndahen në metodat e bazuara tek modeli dhe metodat e bazuara tek shembujt, analiza teorike e aplikimit të tyre në zgjidhjen e problemeve të parashikimit vijon. Konkretisht në ndarjen e parë paraqesim rrjetat neurale (RN), regresionin me vektorët ndihmës (RVN), pemët e regresionit (PR) dhe pyjet e rastësishëm PR_{ast} duke vazhduar me grupin tjetër që përfshin arsyetimin e bazuar tek rastet (ABR).

Kontributet e kësaj teze do të grupoheshin, parashikimi i energjisë në ndërtesa sipas skenarit “offline”, parashikimi i konsumit sipas skenarit “online” dhe parashikimi i konsumit me interval. Parashikimi i konsumit “offline” fillon me përmirësimin e një modeli ABR për parashikimin e konsumit në një ndërtesë institucionale në harkun kohor të 6 orëve. Rezultatet e përfutuara me modelin demostrojnë pikat e forta për përdorimin e tij si metodë parashikimi të energjisë. Më tej parashikojmë konsumin e energjisë për një numër prej 5 godinash, duke u nisur nga vetëm dy të dhëna të matura. Këtu krahasojmë të gjithë modelet e inteligjencës artificiale si RN, RVN, PR, PR_{ast} dhe ABR për të parashikuar konsumin për orën vijuese. Rezultatet tregojnë saktësinë e lartë të pyjeve të rastësishëm në parashikim, por edhe efektivitetin e metodave të tjera referuar direktivave mbi kufijtë e gabimit.

Parashikimi “online” i konsumit është një risi që shfaqet në këtë punim, dhe një nga sfidat e tij. Modelet trajnohen me të dhëna që vijnë në kohë reale duke parashikuar konsumin e orës që vijon me metoda të inteligjencës artificiale. Parashikimi i konsumit me këto metoda tregon rëndësinë praktike që ka kjo qasje, po ashtu tepër rëndësi kanë rezultatet e arritura të cilat mbas njëfarë kohe konvergjojnë shumë afër rezultatit statik.

Në fund, një nga sfidat më të mëdha të këtij punimi është parashikimi i

konsumit me një garanci të caktuar, duke përdorur një metodë origjinale dhe krejt të re në problematikën në fokus. Për këtë përdorim parashikimin me interval, ku ndryshe nga dy qasjet më sipër parashikojmë jo një vlerë të vetme por një interval i cili garanton që vlera reale ndodhet brenda tij me një nivel të caktuar besueshmërie. Rezultatet e arritura ndodhen brenda kufirit të besueshmërisë, duke zënë një diapazon tepër të vogël nga diapazoni i konsumit.

Fjalë kyçe: Efienca e energjisë, ndërtesat, parashikimi i konsumit, inteligjenca artificiale, rrjetat neurale, regresioni me vektorët ndihmës, pemët e regresionit, pyjet e rastësishëm, arsyetimi i bazuar tek rastet, parashikimi dinamik, intervalet e besueshmërisë, parashikimi me interval

Abstract

Building energy consumption is influenced by many factors, such as outdoor weather conditions, operation mode, installed devices, building structure and characteristics, occupant behavior. This complex behavior makes the analysis and prediction of building energy consumption a difficult problem to solve gaining high accuracy results. Main focus of this thesis is to predict short term energy consumption of multi purpose buildings applying up to date machine learning algorithms.

Since the existence of historical datasets with regard to energy consumption is necessary to apply machine learning techniques, we devote a considerable space to real time energy monitoring systems and automated smart metering. Further, we review recently developed models for energy prediction, such as simplified and detailed engineering methods, statistical and artificial intelligence approaches. Given that in focus of this thesis are machine learning approaches, we have selected model based methods such as Artificial Neural Networks (ANN), Support Vector Regression (SVR), Regression Trees (RT) and Random Forests (RF). Second kind of approaches are instance based likewise Case Based Reasoning (CBR).

Main contributions of thesis are, prediction of building energy consumption in batch learning scenario, on-line learning and confidence interval (conformal prediction). Initially, we improve a CBR model to predict next 6 hours energy consumption of an institutional building. Results show significant improvements compared to existent previous model and also to successful machine learning algorithms like neural networks. Hereinafter, 5 different buildings are taken in consideration to predict next hour energy consumption with regard on only two measured attributes. A wide range of algorithms like ANN, SVR, RT, RF, and CBR are compared in this part obtaining high performance results with random forest algorithm as the most relevant among others, but nevertheless other methods have reliable and accurate results compared to international guidelines.

On-line learning is one of the main contributions in this thesis, where machine learning models are trained to predict on data streams. On-line learning has deep practical significance since nowadays and moreover in future energy building data are huge and measured on-line. Results show the importance and relevance of used approaches since the cumulative error converges to batch learning error after certain number of samples.

Last and one of novel contributions in this thesis is conformal prediction, where for a given confidence interval the framework predicts an interval of values with high probability that the true value falls within that interval. This approach

has higher impact on giving confident prior information to building operators for better control strategies and maintenance of buildings.

Keywords: Energy efficiency, buildings, energy prediction, artificial intelligence, neural networks, support vector regression, regression trees, random forest, case based reasoning, on-line learning, confidence intervals, conformal prediction

Hyrje

Konsumi i energjisë elektrike dekadat e fundit po pëson një rritje të vazhdueshme referuar statistikave sidomos në vendet në zhvillim po ashtu edhe në ato të zhvilluara. Pjesa më e madhe e konsumit shkon në ndërtesa veçanërisht ato publike, tregtare si dhe rezidenciale. Vërehet se në Europë, konsumi i energjisë në ndërtesa njeh një total prej 40% të konsumit të përgjithshëm si dhe 36% të prodhimit të CO₂ [1], po ashtu energjia elektrike zë 44% të konsumit në EU si dhe 74% në Shqipëri [2]. Në fig. 1 më poshtë tregohet ecuria e konsumit të përgjithshëm për secilin nga vendet më të mëdha konsumatore për sektorin publik dhe privat. Duket qartë se politikat e ndërmarra nga vendet e zhvilluara bëjnë që të kemi të njëjtën ecuri në konsum përgjatë viteve, ndersa në vendet në zhvillim vërejmë një rritje të ndjeshme të konsumit.

Rritja e konsumit lidhet ngushtë me dy faktorë: ekonomik dhe mjedisor. Një rritje e konsumit sjell kosto shtesë për shoqërinë, si pakësimin e burimeve natyrore, rritjen e emetimit të ndotjes në natyrë veçanërisht CO₂. Arsyet e sipërpërmendura thërrasin për reflektim të gjithë aktorët edhe subjektet të lidhur ngushtë me çështjen. Problemi që shtrohet është ulja e konsumit të energjisë në ndërtesa, duke mos cënuar komoditetin e përdoruesve të këtyre të mirave, edhe për rrjedhojë duke ndikuar në uljen e niveleve të ndotjes. Termi më kuptimplotë që përmbledh së bashku të gjitha problemet e sipërpërmendura, edhe jo vetëm ato do të ishte: “*efiçenca e energjisë në ndërtesa*”. Nëse na u desh një fjali e vetme për të emërtuar problemin, zgjidhjet që jepen në literaturë janë të panumërta. Një ndër to është parashikimi i konsumit për orët që vijnë ose parashikimi afatshkurtër. Parashikimi i konsumit kontribuon veçanërisht në uljen e ngarkesës së pikut dhe në uljen e konsumit total për ndërtesat. Kjo si pasojë e ‘armatosjes’ së operatorëve në ndërtesa me informacion paraprak mbi konsumin, dhe mbi ndryshimet që priten në sjelljen e konsumit. Dekadat e fundit, ky problem me fokus në rritje, në punën e komuniteteve shkencore ka bërë që të lindin, përshtaten dhe zhvillohen një spektër shumë i gjerë metodash parashikimi. Referuar numrit të madh të metodave vëmendja përqëndrohet në ato më të zëshmet vitet e fundit, tek metodat e inteligjencës

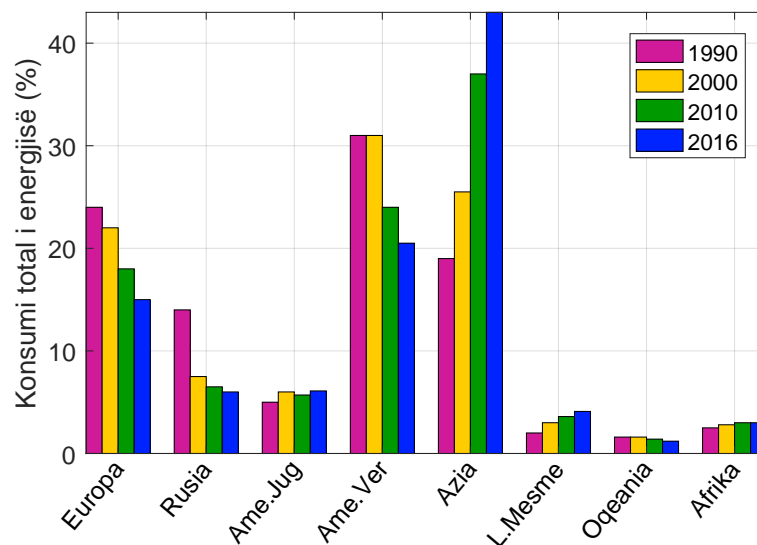


Figura 1: Konsumi në shkallë globale i energjisë elektrike [1]

artificiale, specifikisht dega e *mësimin të makinës* (machine learning).

Përse metodat e inteligjencës artificiale?

Përgjigja e pyetjes jep edhe kuptimin e të gjithë punimit, por shkurtimisht motivimi vjen në dy drejtime të cilat ndërthuren me njëra tjetrën: e para ka të bëjë me tendencën gjithmon e në rritje vitet e fundit të epokës së të dhënave ose *digitale* ku për çdo proces regjistrohen edhe zotërohen të dhëna ndërsa drejtimi i dytë ka të bëjë me bum-in e zhvillimit të metodave statistikore të cilat e kanë të domosdoshme ekzistencën e të dhënave historike për t'u implementuar.

Kjo temë mbështetet dhe zhvillohet mbi arsyetimin e mësipërm, ku theksi i punimit vendoset mbi implementimin e metodave të inteligjencës artificiale. Një hapësirë e veçantë i kushtohet mënyrës dhe teknologjive të matjes dhe ruajtjes së të dhënave me anë të smartmetërave si pjesë e pandashme e zhvillimit. Procesi i implementimit të një metode statistikore për parashikimin e konsumit kalon në dy faza: faza e trajnimit ku modeli "mëson" mbi të dhënat historike dhe faza e testimit ku modeli i mësuar do të testohet në situata të panjohura për të. Metodatat gjerësisht të përdorura në literaturë janë, regresioni linear, rrjetat neurale artificiale, regresioni me vektorët ndihmës si dhe metoda të përdorura rishtazi pemët e regresionit, pyjet e rastësishëm dhe arsyetimi mbi bazën e rasteve.

Në praktikë njihen disa tipe ndërtesash, siç janë ato të banimit ose rezidenciale, ato publike ose për zyra ku madhësia dhe forma e tyre janë nga

më të larmishmet, mund të jenë nga dhoma të vogla deri në hapësira të mëdha banimi dhe pune. Megjithatë problemi i konsumit të energjisë në ndërtesa është tejet kompleks, referuar në literaturën [3, 10, 77, 78, 79, 80] kontribuesit kryesorë në konsumin e energjisë në godinat e mëdha janë sistemet e ngrohje/ftohjes, ndriçimit dhe elementëve të tjerë ndërsa në godinat rezidenciale përfshihen edhe paisjet elektroshtëpiake. Konsumi i energjisë në një ndërtesë varet kryesisht nga këta faktorë: kushtet atmosferike, veçanërisht temperatura e mjedisit, struktura e ndërtesës, veçoritë termike dhe fizike të materialeve të përdorura, aktiviteti i personave që janë pjesë e këtyre ambienteve, dhe performanca e paisjeve të instaluara brenda një objekti.

Referuar kompleksitetit të problemit, parashikimi me një saktësi shumë të lartë has në shumë vështirësi. Vitet e fundit një sërë përpjekjesh janë bërë nga kërkues të ndryshëm për të rritur performancën në parshikimin e konsumit, duke sjellë një mori qasjesh të zgjidhjes së problemit, si propzimi i zgjidhjeve të thjeshtuara, komplekse ose kombinacion të disa metodave. Puna kërkimore e deritanishme analizon këtë problem në skenarë të ndryshëm.

Kapitulli 1

HISTORIKU I PARASHIKIMIT

1.1 Sistemet e monitorimit dhe matjes së energjisë.

Nga më sipër dërtesat konsumojnë afërsisht 40% të energjisë, kështu që diktojnë një mënyrë monitorimi dhe matjeje me qëllim që informacioni të jetë cilësor për një kuptim sa më të mirë të konsumit në ndërtesa. Mangësitë e deritanishme në matje si dhe të dhënat mbi konsumin, kanë qënë një pengesë për zhvillimin e metodave të sakta statistikore që analizojnë konsumin e përgjithshëm po ashtu edhe për ta parashikuar atë. Zhvillimi i shpejtë i teknologjisë së smartmetrave fig. 1.1, sjell letësimin e problematikës dhe ndihmon në përparimin e mëtejshëm të implementimit të metodave parashikuese.

Një *smartmetër* bën matjen në mënyrë elektronike të konsumit të energjisë sipas intervaleve kohore (1 orë, 30 min ose në intervale më të vogla), dhe transmeton informacionin e mbledhur në sisteme apo paisje të tjera. Këto sisteme shërbejnë për ruajtjen e informacionit për përdorim të mëvonshëm apo t'i japin informacion konsumatorëve dhe aktorëve të tjerë në kohë reale mbi ecurinë e konsumit. Bazuar mbi të dhënat konsumatorët (në ndërtesat e vogla) dhe operatorët (në ndërtesat e mëdha) e kanë të mundur informimin mbi konsumin si dhe vendimmarrjen më të lehtë për të rritur kursimin e energjisë. Zhvillimi i shpejtë i teknologjisë ka sjellë implementimin edhe të teknologjive të komunikimit wireless me kosto të ulët, duke mundësuar matjen e faktorëve të tjerë, sidomos atyre ambientalë. Këto sisteme implementohen brenda sistemeve më të mëdha siç janë (BAS, BMS, BEMS, etj) sisteme që shërbejnë për menaxhimin e energjisë në ndërtesa, veçanërisht në ato tregtare dhe rezidenciale të mëdha duke ofruar monitorim më të detajuar të elementëve të ndryshëm. Në fig. 1.2 tregohet skema kryesore e funksionimit të një matësi inteligjent.

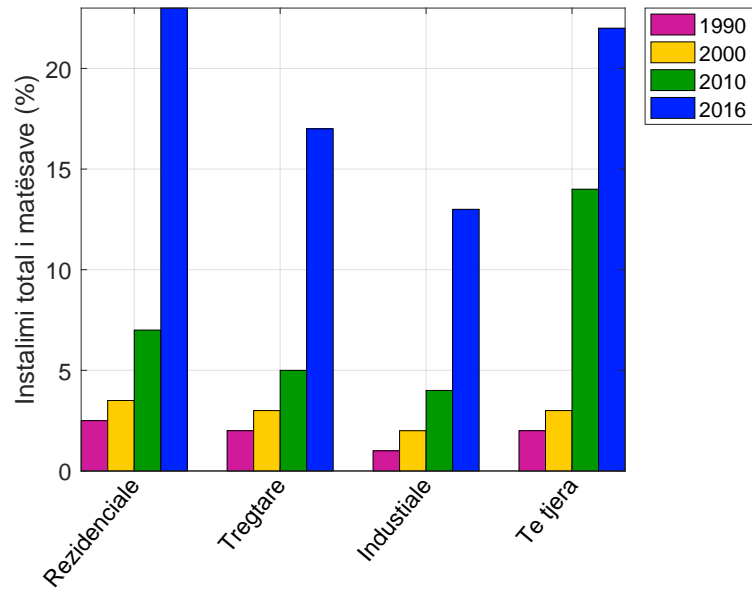


Figura 1.1: Instalimi total i matësve inteligjent sipas tipit të ndërtësës [13]

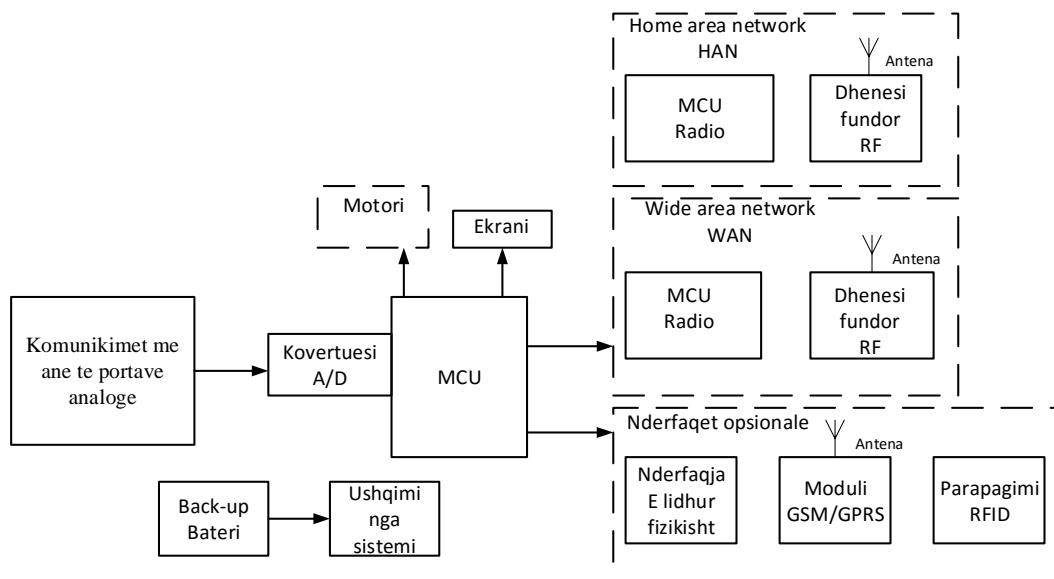


Figura 1.2: Bllok diagrama e një matësi inteligjent [14].

Meqënëse të dhënat maten në intervale kohore të caktuara siç u përmend më sipër, atëherë lind nevoja e ruajtjes së këtyre të dhënave në sisteme të ndryshme në formën e baza e të dhënave. Bazat e të dhënave të kësaj natyre kryesisht ndahen në tre kategori kryesore në funksion të madhësive që sistemi mat. Në Tabelën 1.1 po tregojmë faktorët që luajnë rol në këtë ndarje. Në kushtet e tanishme në modelet e sistemeve të matjes dhe ruajtjes së të dhënave

gjejmë tre kategoritë e treguara në Tabelën 1.1, por në të ardhmen kategoria C do të jetë ajo që do të përdoret gjerësisht pasi siguron informacione më të detajuara dhe iu jep mundësinë aktorëve të cilët punojnë në rritjen e efijçencës, për të përmirësuar politikat dhe veprimet e tyre.

Tabela 1.1: Kategoritë e bazave të të dhënave.

Lloji i bazës së të dhënave	Faktorët ndikues			
	I	II	III	+
A, Niveli 1, (bazë)	Të dhënat klimaterike, zarfi i ndërtesës, paisjet e instaluara			Ana sociale, ekonomike, ligjore
B, Niveli 2, (mesatar)	Të dhënat klimaterike, zarfi i ndërtesës, paisjet e instaluara	Punimi dhe mirëmbajtja, kushtet e brendshme mjedisore		Ana sociale, ekonomike, ligjore
C, Niveli 3, (avancuar)	Të dhënat klimaterike, zarfi i ndërtesës, paisjet e instaluara	Punimi dhe mirëmbajtja, kushtet e brendshme mjedisore	Sjellja e banorëve	Ana sociale, ekonomike, ligjore

Direktiva dhe standarte për matjen me smartmetër

Përdorimi gjithnjë e më i gjerë i sistemeve inteligjente çon në aplikimin e standarteve dhe rregulloreve të matjes. Dy janë direktivat më të fundit të lëshuara nga Komuniteti European, referuar metodave dhe paisjeve matëse [Direktiva 2010/22/EC], dhe [Direktiva 2012/32/EC] referuar efijçencës dhe përdorimit të energjisë nga konsumatorët fundor. Projeksionet në direktivën për shërbimet e energjisë (ESD) parashikojnë se në vitin 2020 rreth 80% e tregut european duhet të instalojë "sistemet e matjes inteligjente", ku vendosen edhe kushte në implementimin e metodave dhe instalimin e paisjeve. Përsa i perket synimeve të shprehura më sipër, një rëndësi të veçantë zë aplikimi i standarteve për këto paisje. Organizatat kryesore për standartizimin e smartmetrave brenda europës si CEN, CENELEC, ETSI kanë hartuar një bazë për funksionalitetet si dhe komunikimin e pasijeve inteligjente, ku sipas tyre paisjet duhet të përmbushin këto kritere:

- Siguria e të dhënave, autenticiteti, komunikimi, matja, instalimi.
- Ofrimi i të dhënave të matura për zgjidhjen e gjendjes në tregjet e energjisë elektrike.
- Shërbimi i leximit të konsumit, informacioni i klientit, monitorimi i cilësisë së energjisë.
- Monitorimi dhe kontrolli i ngarkesës, mundësi për parashikimin e kërkesës.

Përdorimi i gjerë i teknologjisë matëse kërkon aplikimin e standarteve në matje dhe përdorim, këtu përmendim Amerikën e Veriut dhe Europën si modelet lider në këtë fushë.

Amerika e Veriut përdor si standart **ANSI C12** [12] kundrejt **IEC 62056** [18] që përdoret në Europë.

Teknologjitë e ndryshme në sistemet e monitorimit online

Studimet kanë treguar se implementimi i sistemeve të matjes online dhe afishimit të të dhënave ndikon në përmirësimin e efikasitetit të energjisë. Ueno dhe Nakano në studimin e tyre [15] tregojnë se dhënia e informacionit konsumatorëve në kohë reale mbi konsumin e energjisë, reduktoi mesatarisht me 9% konsumin në 8 godina në stinën e dimrit. Përderisa sensorët janë paisje të cilat kushtëzohen nga pamundësia e ruajtjes së informacionit, ruajtja e të dhënave bëhet në një bazë të dhënash. Modeli i propozuar në [16] implementon një bazë dhënash të krijuar në MySQL me ndërfaqe të krijuar në LabVIEW. Sistemi në tërësi duke filluar nga sensorët e matjes deri tek bazat e ruajtjes së të dhënave kordinohen në një rrjet me kategori të caktuar, ku topologjitë më të përdorshme janë: Seri, Yll, Kontur, Unazë dhe Pemë.

Hapat e shpejta me të cilat zhvillohet teknologjia e sensorëve si dhe teknologjitë e tjera të lidhura ngushtë me sistemet e matjes inteligjente ka bërë që autorë të ndryshëm të krijojnë prototipe të ndryshme sistemesh matjeje. Ku secili prej tyre dallohet nga karakteristikat e veçanta të tij.

Duke filluar me BACnet, që është platforma e monitorimit online të të dhënave e propozuar nga Jang [19]. Një koordinim midis sensorëve magnetik me komunikimet hardware RS232 sjell si risi uljen e konsumit të energjisë midis nyjeve të sistemit kur sensorët nuk transmetojnë të dhëna. Kompleksiteti dhe kostoja e lartë e këtij produkti, sjell në arenë sistemin Kilavi i projektuar nga Oksa [20]. Mëshirimi i teknologjisë wireless në transmetim, si dhe mundësia e përdorimit nga celulari apo panelet e kontrollit e bën atë produkt të së ardhmes. Por kostoja e lartë e sistemeve të deritanishme sjell në skenë një modul të ri “low-cost” si Zigbee e sjellë nga Sung [21], me një ndërfaqe të ndërtuar

në LabVIEW si dhe mundësi transmetimi në frekuenca më të larta se dy të parat, gjen aplikim në shumë fusha. Sistemet e sipërpërmendura kryejnë funksionin e matjes, transmetimit dhe ruajtjes së të dhënave. Ndërsa Bayindir [22] pa të arsyeshme implementimin i analizatorit të fuqisë e mundësuar RS485 Ethernet (Modbus). Vlen të përmendet si risi edhe komunikimi me anë të sistemeve PLC me kompjuterin.

Sistemet e përshkruara më sipër shprehin rëndësinë e tyre në procesin e matjes dhe ruajtjes së të dhënave. Por ku ndihmojnë këto në implementimin e metodave të parashikimit? Metodat e inteligjencës artificiale kanë të domosdoshme ekzistencën e të dhënave historike, kështu që të dhënat e ruajtura në bazën e të dhënave do të përdoren në fazën e parashikimit të konsumit. Rritja e përdorimit të matjes inteligjente e ndihmon së tepërmi fushën e kërkimit që fokusohemi. Më poshtë analizojmë disa nga metodat e implementuara në parashikimin e konsumit të energjisë në ndërtesa.

1.2 Metodat e parashikimit

Në këtë pjesë analizojmë metodat më të përdorura deri tani për parashikimin e konsumit të energjisë në ndërtesa. Metodat mund të grupohen: metodat inxhinierike, statistikore dhe ato të inteligjencës artificiale. Metodat inxhinierike janë ato të thjeshtuara (modeli pranon pak parametra fizik nga ndërtesa) dhe komplekse (ku modeli pranon një gamë të gjerë parametrash). Metodat statistikore janë të panumërta, por fokusohemi kryesisht tek modelet e regresionit që parashikojnë seritë kohore.

Metodat e inteligjencës artificiale më të përdorshme në parashikim janë rrjetat neurale (RN), regresioni me vektorët ndihmës (RVN), po ashtu vërejmë dhe metodat e përdorura rishtazi si pemët e regresionit (PR), pyjet e rastësishëm (PR_{ast}), por dhe metodën e arsytimit bazuar tek rastet (ABR) e cila zë një hapësirë tepër të rëndësishme në punimin tonë.

Një tablo mbi teknikat e përdorura në parashikim jepet në [8] ku pasqyrohen avantazhet dhe disavantazhet e aplikimit të tyre.

1.2.1 Metodat inxhinierike

Metodat inxhinierike janë optimale për parashikim në momentet e para të funksionimit të një ndërtese, kjo për arsyen se veçoritë fizike të godinës ndryshojnë me rritjen e kohës së përdorimit si pasojë e amortizimit. Bacher dhe Madsen [23] analizojnë statistikisht vetitë fizike të ndërtesave me qëllim përshtatjen e modelit optimal për parashikimin e konsumit. Pra një kombinim i metodave inxhinierike për parashikim me ato statistikore për optimizimin e

modelit sjell suksesin e metodës sipas tyre.

Analiza e faktorëve që ndikojnë mbi konsumin e energjisë në një ndërtesë është në fokusin e autorëve [24], ku analiza e vetive fizike të ndërtesave për 6 zona të ndryshme klimatike në SHBA simulohet në AUTODESK Green Buildings Studio (GBS), i cili mbështetet tek DOE-2.2. DOE është departamenti i energjisë në Shtetet e Bashkuara të Amerikës i cili ka në fokusin e tij përmirësimin e efijencës së energjisë në ndërtesa. Ata ofrojnë një sërë simulatorësh për gjenerimin e të dhënave mbi konsumin. Në kuadrin e përmirësimit të efijencës një sërë softesh janë zhvilluar si DOE-2, EnergyPlus, BLAST, ESP-r të cilat mund të gjenden në [25]. Saktësia e këtyre simulatorëve është shumë e lartë, por vlen të theksohet se në aplikimin e të tilla metodave hasen veshtirësi, pasi implementimi i këtyre metodave lidhet ngushtë me njohuri të detajuara mbi veçoritë fizike të godinës si dhe parametrat e mjedisit. Në njërën anë këto metoda kërkojnë informacion të detajuar mbi karakteristikat e nënhapësirave të godinave të mëdha, ku ky lloj informacioni është i vështirë për t'u evidentuar dhe jo të gjithë ekspertët e disponojnë, në anën tjetër simulimi me anë të këtyre metodave kërkon njohuri të thella të ekspertëve të fushës gjë që e rrit akoma më shumë koston e aplikimit të këtyre metodave. Për arsytet e sipërpërmendura një pjesë e studiuesve kanë propozuar metoda të reja ku ndër to përmenden modelet e thjeshtuara, metodat statistikore dhe ato të inteligjencës artificiale siç paraqiten në Seksionin 2.1. Një analizë mbi përdorimin e metodave inxhinierike në parashikimin e konsumit të energjisë dhe potencialet që kanë këto metoda për t'u përdorur më tej në këtë fushë kërkimore jepen në [26]. Këtu theksohet rëndësia e analizës së konsumit të energjisë për godinat duke rritur efijencën e energjisë pa ndërhyrë në cilësinë e jetës së konsumatorit. Në këtë punim autori rishqyrton dy metoda kryesore. E para që njihet dhe si (*metoda e gradës së ditës*) është veçanërisht për të llogaritur konsumin e energjisë në ndërtesat relativisht të vogla. Metoda mbështetet në vetëm një matje mbi ecurinë e sjelljes termike të godinës në ditë të caktuara referuar një dite referencë (referuar temperaturës) që metoda tenton ta gjejë. Autori e modifikon metodën duke sjellë në dukje një faktor rregullimi të metodës së vjetër për përmirësimin e rezultateve. E dyta është metoda e frekuencës së temperaturës (*bin*) e cila simulon në disa skenarë sjelljen e ndërtesës. Kjo metodë është shumë e vlefshme për godinat e mëdha ku konsumi nuk varet nga inercia e godinës ose nga tendenca historike si në metoden e parë, por nga disa faktorë shtesë si kushtet atmosferike, popullimi etj. Një shqyrtim i detajuar në këtë punim i metodave të thjeshtuara kundrejt atyre të modeluara me kompjuter ose të detajuara vë në pah pikat e forta dhe fushat e përdorimit të tyre.

Një analizë e detajuar e simulatorëve më të përdorshëm në 50 vitet e fundit nga komuniteti i parashikimit konsumit të energjisë në ndërtesa jepet në [27]. Këtu vihen në pah aftësitë, cilësitë si dhe mundësitë e 20 simulatorëve më të përdorur për të llogaritur konsumin e energjisë në godina, si dhe të dhënat që duan paprakisht secili, po ashtu të mirat dhe të metat e secilit simulator. Një analizë mbi softin EnergyPlus jepet në [28]. Softi, një produkt i DoE, i cili ka dalë në përdorim si kombinim i dy softeve të mëparshme të prodhuara nga DoE [25], i DOE-2 si dhe BLAST, ku rezultatet e simuluarat për konsumin e energjisë janë në intervale prej 1 ore.

Një model për parashikimin e konsumit të energjisë elektrike me hark kohor prej 1 ore duke anashkaluar modelin e detajuar dhe duke përdorur modelet e thjeshtuara implemetohet nga Fumo [29]. Ata arrijnë të krijojnë modele tip për koeficientët e ndërtesës duke u nisur nga faturat e konsumit të mëparshëm të energjisë dhe i aplikojnë tek simulatori EnergyPlus. Modelet tregojnë rezultate të kënaqshme me një kufi gabimi prej 10%, vlerë brenda standarteve të vendosura nga ASHRAE [30].

Duke qënë se simulatorët për parashikimin dhe analizën e konsumit të energjisë kanë gjetur një përdorim shumë të gjerë, autorët në [31] kanë parë të arsyeshme publikimin e një punimi mbi qëllimin e përdorimit të këtyre softeve. Aty shtrohen një sërë pyetjesh thelbësore, të cilat ndihmojnë kërkuesit me informacionin se cila është mënyra optimale që mund të zgjedhim një soft për ta përdorur.

1.2.2 Metodat statistikore

Metodat statistikore bëjnë ndërlidhjen e konsumit të energjisë me variablat ndikuese në konsum duke përshtatur një ekuacion matematikor. Këto janë metoda empirike dhe mbështeten në të dhënat historike të grumbulluara paprakisht, me të cilat modeli do të trajnohet. Parashikimi i konsumit të energjisë me këto metoda mund të ndahet në tre grupime kryesore metodash. E para përfshin metodat e thjeshtuara ku metoda tenton të japë rezultate bazuar në një numër të vogël variabash p.sh në faktorët klimaterik, e dyta janë metodat që marrin në konsideratë edhe elementë të tjerë si indeksi i energjisë dhe faktorë suplementarë, e treta janë metodat komplekse ku përfshihen elementë shtesë të cilët japin informacion të plotë mbi faktorët ndikues në konsum, si p.sh sjellja termike, nënsistemet e tjera të instaluarat etj. Një kombinim i metodave statistikore me metodat inxhinierike për të parashikuar në terma afatgjatë sjelljen termike të një ndërtese publike analizohet në [32]. Metodat statistikore llogarisin marrëdhënien midis temperaturës së jashtme me atë të brendshme. Po ashtu Bauer [33] i përdor metodat statistikore si

mjet për evidentimin e faktorëve ndikues në sjelljen e sistemeve HVAC, për të realizuar optimizimin kontrollit të tyre. Zhvillimet e komunikimeve të bazuara në rrjetin web kanë çuar autorët [34] të studiojnë sjelljen e konsumatorëve individual në banesat rezidenciale bazuar në të dhënat e gjeneruara nga aksesit i tyre. Ata përdorin metodat statistikore për të parashikuar sjelljen e tyre konsumator-energji për periudha të shkurtra. Ansari [35] përdor një metodë të thjeshtë për llogaritjen e konsumit të nevojshëm për ftohje duke marrë parasysh kontributin e parametrave fizikë të ndërtesës, njësoj si në metodat inxhinierike. Ghiaus [36] aplikon një model regresioni robust për të parashikuar eksperimentalisht performancën energjitike të ndërtesave bazuar në kushtet e jashtme atmosferike. Bauer dhe Scartezini [33] përdorin metodat statistikore për të llogaritur njëkohësisht kërkesën për ngrohje edhe ftohje duke marrë në konsideratë kushtet e mjedisit brenda edhe jashtë ndërtesës. Dhar [37, 38] modeloi konsumin e energjisë me anë të transformimit Furie të serisë kohore të vetëm një variabli si temperatura e jashtme, duke evidentuar jolinearitetin që ekziston mes temperaturës dhe konsumit. Lei dhe Hu [39] tregojnë se për zonat me verë të nxehtë dhe dimër të ftohtë në Kinë konsumi i energjisë varet nga temperatura mesatare e muajit dhe ndikimi i faktorëve të tjerë mjedisorë është jo shumë ndikues. Këtu modeli i regresionit linear i varur nga një variabël është i përshtatshmi sipas autorëve. Ma [40] përdorin regresionin linear të shumfishtë të parashikojnë konsumin e energjisë në shkallë të gjerë për sektorin publik. Regresioni linear i shumfishtë si dhe auto-regresioni përdoren nga Jiménez dhe Heras [41] për të llogaritur parametrat fizikë të një ndërtese që ndikojnë mbi konsumin e energjisë dhe të parashikojnë temporeturën së jashtme. Një krahasim i metodave të thjeshta regresive dhe auto-regresionit paraqitet në [56]. Hoffman [43] përdor një model auto-regresioni për të parashikuar ngarkesën e pikut në ndërtesat komerciale. Regresioni linear i kombinuar me karakteristikat që vijnë nga pozicionimi i objektit përdoren në [50] për të parashikuar konsumin e energjisë. Një theks i veçantë u vendos nga autorët në [51] edhe mbi rolin e rezidentëve në tendencën e konsumit të përgjithshëm, analizuar kjo me një nga modelet e autoregresionit si ARIMA. Parashikimi online i konsumit për ditën e ardhshme u aplikua nga Kimbara [52], ku modeli i përdorur ishte ARIMA i modeluar duke i shtuar variabla në hyrje dhe njihet si ARIMAX. Një kombinim metodash për parashikimin e kërkesës për konsumi aplikohet në [53], ku alternimi i ARIMA me arsyetimin bazuar tek rastet (ABR) krahasohet me metodat më tradicionale si rrjetat neurale.

Sasia e të dhënave luan një rol kryesor në saktësinë e modelit dhe për këtë autorët në [42] analizojnë modelin e regresionit linear në tre situata të ndryshme kur të dhënat historike përfshijë 1 ditë deri 3 muaj. Ku për

të dhëna vetëm 1 ditore gabimi arrin deri 100%, me të dhëna 1 javore deri në 30% si dhe në 3 muaj zbret në 6%. Lam [44, 45] analizojnë varësinë e konsumit të energjisë nga faktorët klimaterik për një periudhe 27 vjeçare. Analiza e komponentëve kryesorë (PCA) përdoret për të evidentuar faktorët klimaterik me më shumë ndikim mbi rezultatet e simuluar për konsumin e energjisë në një godinë në Hong Kong. Vërehet se performanca e rezultateve të dhëna nga regresioni është e një saktësie 98%, por mund të zbresë deri në 94%. Një metodë e përdorur shpesh ndër metodat statistikore të regresionit është edhe ajo e analizës së kushtëzuar të kërkesës ose njihet si në literaturë CDA. Hapat e para në përdorimin e kësaj metode nisnin në 1980 ku Michael dhe Cynthia Parti [46] parashikojnë konsumin e përdoruesve rezidencialë duke nxjerrë të dhënat nga faturat mujore të energjisë elektrike. Teknika u aplikua më tej në [47], ku u analizua tendenca e konsumit për konsumatorët e vegjël në një periudhë 10 vjeçare, duke evidentuar faktorët ndikues në përmirësimin e efikasitetit të energjisë. Aigner [48] përshtasin 24 modele regresioni për konsumatorët e energjisë ku secili model është përafrim i konsumit për secilën orë të ditës, duke bërë edhe evidentimin e ngarkesave që përdoren pa pasur matje fizike. Aydinalp-Koxsal dhe Ugursal [49] propozojnë përdorimin e metodës (CDA) për parashikimin e konsumit në nivel kombëtar, dhe një krahasim me rrjetat neurale dhe metodat inxhinierike performanca e metodës ishte konkurruese, duke shtuar edhe aftësinë e metodës për të evidentuar faktorët social/ekonomik.

Metodat e regresionit zënë një pjesë të konsiderueshme në publikimin [54] për të bërë parashikimin afatshkurtër të konsumit të energjisë. Kombinimi i metodave të inteligjencës artificiale me seritë kohore kaotike për parashikimin e konsumit implementohet nga [55]. Ndërsa një analizë mbi ndikimet sasiore të konsumit të energjisë në ndërtesa duke pasur si analizues metodat statistikore paraqitet nga Tsanas [57].

1.2.3 Rrjetat artificiale neurale

Rrjetat neurale janë modeli më i përhapur i inteligjencës artificiale që përdoren për të parashikuar konsumin e energjisë në ndërtesa, për vetë faktin se ky model është i aftë të zgjidhë probleme të cilat kanë natyrë jolinare dhe janë efektive në analizën e modeleve komplekse. Fillimet e saj metoda i ka që në 1943 me McCulloch dhe Pitts [58], ku modeli i tyre u bazua në metodat matematikore dhe në algoritma logjike. Më tej Hebb [59] modeloi një mekanizëm të bazuar në plasticitetin neural. Farley dhe Clark [60] e përdorën modelin e Hebb në makinat "llogaritëse" dhe më tej nocioni aplikohet në [61]. Zbulimi i perceptronit nga Rosenblatt [62] e çoi në nivel tjetër përdorimin e rrjetave neurale. Minsky dhe Papert [64] në analizën e tyre mbi fuqinë dhe dobësitë e

rrjetave neurale ndihmuan në zhvillimin e mëtejshëm të metodës. Implementimi i algoritmit me përhapje të prapme nga Werbos [63] solli rritjen e efikasitetit në kohën e trajnimit. Efektiviteti u përmirësua me daljen në skenë edhe të procesimit paralel [65]. Zhvillimet e rrjetave neurale vijojnë edhe në ditët e sotme, po ashtu edhe fusha e aplikimit është zgjeruar së tepërmi.

Përdorimi i rrjetave neurale gjen aplikim edhe në parashikimin e konsumit të energjisë, ku përdoret për modelimin e rrethanave të caktuara të funksionimit të ndërtesave, parashikimin e konsumit të ngrohje/ftohjes, të ndriçimit, të pjesëve të përbërëse të nënsistemeve. Në këtë seksion do të paraqesim një analizë të modeleve RN që më të përdorshëm në literaturë.

Azadeh mbështetet në modelin e RN në disa prej punimeve të tij, në [66, 69] ai përdor RN së bashku me seritë kohore për të parashikuar konsumin e energjisë, duke vazhduar më tej me implementimin të algoritmave gjenetikë [68] dhe metodave inxhinierikë në parashikim [67]. Duke kaluar po ashtu në përdorimin e metodës në parashikimin e konsumit për konsumatorët e mëdhenj industrial të energjisë për një periudhë një vjeçare. Qasja e tyre mbështetet në një Multi Layer Perceptron (MLP) të mbykqyrur (supervised), duke theksuar përdorimin e rrjetave neurale në parashikimet afatgjata. Një RN me përhapje të përparme përdoret për parashikimin e konsumit të energjisë me 24 orë në përparim [71], ku merren parasysh metodat statistikore për të eliminuar karakteristikat me më pak kontribut në hyrje të modelit duke përmirësuar performancën e modelit. Gonazalez dhe Zamarreno në [72] propozojnë një lloj të veçantë të RN duke bërë të mundur krijimin e një modeli i cili në iteracionet që zhvillon përdor një pjesë të daljes së sistemit duke ushqyer si informacion në hyrje. Saktësia e lartë e metodës arrihet duke përdorur vetëm temperaturën e jashtme si variabël.

Kalogirou në [73] prezanton përdorimin e rrjetave artificiale në sistemet e energjisë, duke vazhduar në [74] me përdorimin e një rrjete neurale me përhapje të prapme për të parashikuar kërkesën totale të godinave për ngrohje. Modeli është trajnuar në një numër shumë të madh ndërtesash prej (225) me madhësi dhe karakteristika të ndryshme. Në [75] autorët përdorin të njëjtin model për të parashikuar konsumin e kërkuar për ngrohjen e ndërtesës, duke konsideruar karakteristikat fizike të godinës si variabla hyrjeje në modelin e tyre. Saktësia e modelit është shumë e lartë me saktësi mbi 95%. Në [76] autorët përdorin një rrjet neurale me përhapje të prapme për të parashikuar konsumin e paisjeve ftohëse në një godinë, duke bashkërenduar me një metodë optimizimi për të përcaktuar parametrat optimalë të sistemit. Duke pasur vetëm si të dhëna parashikimin e temperaturës si dhe kohën e kryerjes së matjeve autorët në [77] përdorin një rrjet neurale të përsëritur për të përcaktuar kërkesat për ngrohje dhe ftohje të një ndërtese.

Referuar varësisë të konsumit të energjisë nga kushtet klimaterike në [78] përdoret një rrjetë neurale e cila modelohet bazuar në modelin e kampionit të ditës së nxehtë dhe kampionit të ditës së ftohtë, ku këto dy modele përdoren si parametra në hyrje të rrjetës neurale. Modeli testohet në 3 zona të ndryshme atmosferike për të parashikuar kërkesën për ngrohje dhe për ftohje të ndërtesës.

Rrjetat neurale përdoren po ashtu për të analizuar dhe optimizuar sjelljen e elementëve përbërës në një ndërtesë sidomos të sistemeve të ngrohje-ftohjes pasi këta janë dhe konsumatorët më të mëdhenj brenda një ndërtese. Parashikimi i konsumit të ajrit të kondicionuar për të optimizuar kontrollin të një sistemi HVAC¹, është pjesë e punës së autorëve në [30]. Rrjetat neurale përdoren në [80] për të përcaktuar defektet dhe avaritë që ndodhin gjatë periudhës së funksionimit të ndërtesës. Metodat shmangin analizën e thellë matematikore duke i bërë tepër tërheqëse dhe konkurruese në zgjidhjen e ketyre problemeve, ku një avantazh të madh marrin nga natyra e tyre për të zgjidhur problemet jolineare. Në [81] autorët parashikojnë ftohjen e një hoteli bazuar në impaktin e popullimit të ndërtesës dhe funksionimit të saj duke ruajtur konformitetin e konsumatorit. Në [82] autorët tregojnë se rrjetat neurale mund të përdoren për të parashikuar konsumin e energjisë për paisjet elektroshtëpiake, ndriçimin si dhe ftohjen e hapësirës, parashikimi i këtyre sjell evidencimin të sjelljes socio-ekonomike të përdoruesve të energjisë në Kanada. Pra jo vetëm ana teknike, por edhe sjellja sociale si dhe popullimi i godinave mund të parashikohet me këto metoda.

Rrjetat neurale angazhohen për të përcaktuar parametrat e performancës së një ndërtese, ku në [83, 84, 85, 86] autorët përcaktojnë në punën e tyre koeficientët e humbjes së energjisë termike, kapacitetin total termik si dhe aftësinë e ndërtesës për të ruajtur inercinë termike të saj, faktorë këta shumë të rëndësishëm në përcaktimin e efikasitetit të energjisë si dhe parashikimi i nevojës për energji në ndërtesë. Këto parametra përcaktohen duke përdorur një model të thjeshtë të rrjetës neurale i trajnuar me të dhëna që përdorin diferencën e midis temperaturës jashtë dhe brenda ndërtesës. Autorët në [87] japin rezultate mbi parashikimin e konsumit në hapa kohorë prej 1 ore si dhe gjatë eksperimentimit të modeleve bazuar mbi të dhënat që ata disponojnë mundën të përcaktojnë vlerat e rezistencës termike R , si dhe kapacitetin termik C . Të dhënat për të trajnuar modelet e rrjetave neurale mund të sigurohen ose nga matjet direkte të paisjeve matëse të cilat janë të instaluar në një ndërtesë sic mund të jenë BMS (Building Management Systems), nga vëzhgime të ndryshme nga faturat e pagesës së energjisë elektrike ose nga

¹ HVAC-Heating Ventilation Air Conditioning

simulimet e ndryshme. Një rrjetë neurale me dy cikle përdoret nga autorët në [72] ku në ciklin e parë ata përdorin rrjetën për të parashikuar të dhënat klimaterike si dhe në ciklin e dytë ata përdorin të dhënat e parashikuara të ciklit të parë si hyrje të modelit së bashku me të dhëna të tjera si (momenti i matjes, dita, ngarkesa mesatare) për të parashikuar konsumin e energjisë për godinën. Më sipër përmendëm se metodat inxhinierike janë të përdorshme në parashikimin e energjisë në ndërtesa, një sërë autorësh i kanë marrë në konsideratë të dyja modelet për të bërë krahasimin se cila është më e vlefshme për t'u përdorur në situatën e tyre. Një krahasim i detajuar i modeleve inxhinierike me rrjetat neurale që përdoren për të parashikuar konsumin e energjisë elektrike paraqitet në [88], ku kemi një krahasim të avantazheve dhe disavantazheve të të dy modeleve dhe rezultatet e paraqitura ata tregojnë se të dy modelet sigurojnë saktësi të lartë, por me epërsinë e rrjetave neurale. Krahasimi i metodave inxhinierike me rrjetat neurale është fokus edhe i autorëve në [82]. Duke analizuar një sërë faktorësh që ndikojnë mbi sjelljen e konsumatorëve si ndriçimi, sistemet të ngrohjes-ftohjes, kushtet socio-ekonomike ata vënë në dukje epërsinë e rrjetave neurale. Kombinimi i rrjetave neurale me algoritmat gjenetik paraqitet në [89]. RN fillimisht trajnohen për të mësuar “sjelljen” energjitike të ndërtesës dhe me tej kombinohen me algoritmat gjenetikë për optimizimin e komoditetit termik të banorëve si dhe optimizimin e kontrollit të konsumit.

Metodat inxhinierike të simulimit përdoren për të gjeneruar të dhëna mbi konsumin në ndërtesa. Një vëzhgim i detajuar mbi përdorimin e rrjetave neurale në problemet e parashikimit të konsumit të energjisë elektrike gjendet në [90], ku një vend zë përdorimi i metodave të tjera për parashikimin e konsumit të energjisë. Ndërsa përsa i përket metodave inxhinierike dhe kombinimit të tyre me metodat e inteligjencës artificiale autorët në [91] ku tregohen më qartë potencialet e secilës metodë për t'u përdorur.

Rrjetat neurale janë në fokusin tonë për parashikimin e konsumit të energjisë në orën që vijon e traguar qartazi në Kapitullin 5 ku dy janë modelet e parashikimit që implementohen: statik i cili jepet në publikimin [5] si dhe një model i parashikimit dinamik tregohet në [9].

1.2.4 Regresioni me vektorët ndihmës

Regresioni me vektorët ndihmës (RVN) fillimisht e propozuar nga Vapnik [95] gjen përdorim si një metodë mjaft efektive për zgjidhjen e problemeve jolineare, si dhe për zgjidhjen e problemeve të parashikimit. Veçoria kryesore e metodës është përdorimi i funksioneve kernel, të cilët bëjnë transformimin e të dhënave nga një dimension në një tjetër. Fillimisht parashikimi i konsumit duke përdorur vetëm funksionin Kernel përdoret nga Brown [96] ku modelet e

propozuar në punim kanë performancë më të lartë se modeli fitues i kompeticionit të ASHRAE. Fillimet e përdorimit të RVN i gjejmë tek raporti i Solomon [97], ku studiohet parashikimi i konsumit të energjisë për ndërtesa e mëdha, zgjidhja përgjithësohet duke e aplikuar për parashikim në ndërtesa të tjera. Aplikimi i metodës vazhdon nga Dong [98], ku parashikimi i konsumit për ndërtesat në zonën tropikale tregon një performancë të lartë të SVN kundrejt rrjetave neurale. Hsu et al. [99] përdor algoritmat gjenetike (AG) për të përmirësuar zgjedhjen e parametrave për RVN në parashikimin e konsumit, ku modeli i optimizuar me (AG) tregon një performancë tejet të lartë kundrejt modeleve më të mira ekzistuese. Lai et al. [100] përdorin të dhënat e matura për periudhën 15 mujore referuar një godine rezidenciale dhe testojnë performancën e RVN në parashikim, duke analizuar ndikimin e faktorëve klimaterik në konsumin e energjisë. Krahasimi i performancës së SVM me modele të ndryshme të rrjetave neurale paraqitet në [101], ku parashikimin afatshkurtër i konsumit të energjisë për një ndërtesë zyresh publike në Kinë varet nga temperatura e jashtme, lagështia dhe rrezatimi diellor. Performanca e lartë e RVN kundrejt këtyre metodave, motivon autorët ta përdorin dhe në punimin tjetër [102], ku kryhet parashikimi i konsumit për ftohjen e ndërtesës në terma të afatshkurtër. Epoka e "Big Data" ka influencuar Grolinger [117] për ta aplikuar, ku RVN ndërthuret me metodën H2O për të reduktuar të dhënat që do të trajnohet modeli.

1.2.5 Pemët e regresionit

Kjo është një nga metodat më të lehta dhe të kuptueshme në përdorim ditët e sotme nga metodat e inteligjencës artificiale. Funksionon duke e ndarë hapësirën e të dhënave në pjesë duke krijuar degëzime dhe në fund krijon modelin e ngjashëm me një pemë [118, 119]. Më tej në problematikën e parashikimit të kërkesës për konsum përdorimet e metodës hasen vonë. Yang dhe Stenzel [120] përdorin pemët e regresionit për të parashikuar konsumin e energjisë në terma afatshkurtër. Një krahasim i tre modeleve të parashikimit implementohet nga Tso dhe Yau [121], ku implementojnë regresionin linear, pemët e regresionit dhe rrjetat neurale. Modelet krahasohen për dy stinët verë dhe dimër, ku nga rezultatet vërehet një epërsi e lehtë e pemëve të regresionit kundrejt të tjerave, ndërsa në dimër kjo epërsi humbet kundrejt RN. Pemët e regresionit përdoren në [122] për të analizuar faktorët që ndikojnë në konsumin e energjisë për një sërë rezidencash në Poloni. Yu et al. [123] përdorin pemët për parashikimin e konsumit për konsumatorët rezidencialë bazuar në ndikimin e faktorëve të jashtëm, ku performanca e metodës është 92%.

Pemët e regresionit përdoren për parashikimin e konsumit të energjisë në një kampus universitar nga autorët në [124]. Modeli përdor

të dhëna të matura në një periudhë 3 vjeçare të siguruara nga matësit dixhital të instaluar në ndërtesë, dhe performanca e modelit përmirësohet krahasuar modelet bazë me 53%. Po kështu e njëjta autore në [125] përdor këto modele për të analizuar më tej natyrën e konsumit si dhe faktorët që ndikojnë në të për një “smart-grid”. Po ashtu studimi i parashikimit të konsumit në rrjetat smart bëhet në [127], ku parashikimi i konsumit në ndërtesa tipike të mëdha me anë të modelit të pemës (mbCRT) tregon përmirësime të mëdha kundrejt modeleve të tjera duke sjellë reduktim të konsumit dhe kursime vjetore deri 45000 \$ për një ndërtesë komerciale. Një tjetër model peme ai M5 përdoret në [126] me qëllim të parashikohet piku i ngarkesës, ku modeli i pemës tejkalon performancën e rrjetave neurale dhe regresionit me vektorët ndihmës.

Po ashtu pemët e regresionit mund të angazhohen së bashku në parashikim duke krijuar metoda të bashkësive si pyjet e rastësishëm. Këto janë një risi në fushën e kërkimit tonë, ku një model i bashkësive të pemës përdoret në [128] për parashikimin në terma afatshkurtër të konsumit në ndërtesa duke arritur një përmirësim të parashikimit mesatarisht 18%.

1.2.6 Metoda e arsytimit bazuar tek rastet

Kjo është një nga metodat e përdorura rishtazi në zgjidhjen e problematikës së parashikimit të energjisë elektrike me qëllim përmirësimin e efikasitetit të energjisë. Metoda ka në qëllimin e saj kryesor zgjidhjen e problematikave të reja (të panjohura më parë) duke u bazuar në zgjidhjet të cilat ekzistojnë nga problematika të ngjashme duke modeluar problemin me anën e rasteve. Fillimet e saj metoda i ka në 1980 me profesorin e universitetit të Yale Roger Schank dhe studentët e tij mbi problemet e lidhura me memorien dinamike [134] dhe më vonë njohu përhapjë të gjerë në vitet 90 në fusha të tjera. Metoda njih një sërë realizimesh në praktikë si dhe veçanërisht në fushën e shëndetit ku mund të referohemi për aplikimet në punimet [92, 93, 94]. Në problematikën e parashikimit të konsumit të energjisë elektrike këto metoda janë metoda të përdorura së fundmi, pra janë metoda të reja, por që ofrojnë avantazhe, kur në rastin e mungesës së të dhënave historike ABR mund të funksionojë mjaft mirë në krahasim me metodat e tjera të inteligjencës artificiale. Po ashtu metoda paraqet një lehtësi në kuptimin e mënyrës se si ajo bën parashikimin si dhe mund të modelohet dhe mund të përmirësohet në cilindo hap të mundshëm të saj. Autorët në [129] hedhin hapat e para për përdorimin e kësaj metode në parashikimin e konsumit të energjisë elektrike duke e bazuar modelin e tyre në parashikimin e energjisë nga të dhëna të cilat vijnë nga monitorimi si dhe të dhëna të gjeneruara sintetikisht nga simulatori EnergyPlus i cili mbështetet në metodat inxhinierike. Punimi bën parashikimin e konsumit bazuar në disa

intervale kohore si ato të 1, 3 dhe 6 orëve në vijim. Modeli tregon rezultate të kënaqshme me të dhënat sintetike si dhe rezultate më të dobta nga të dhënat reale (të monitoruara), por të dyja rezultatet janë brenda normave të përcaktuara nga ASHRAE [30]. Rezultatet tregojnë se modeli është i përshtatshëm për të parashikuar kërkesën për konsum si dhe mund të implementohet në sistemet e funksionimit të ndërtesës. Po ashtu në [130] autorët vazhdojnë mbi të njëjtën ide për të zhvilluar modelin duke e aplikuar në ndërtesa të tjera. Në [131] autorët përdorin një model të përmirësuar nga ai të cilin përmendëm më sipër duke përmirësuar mënyrën sesi zgjidhen karakteristikat përshkruese të të dhënave duke përdorur metodën e analiza e komponentëve parësor (PCA) si dhe me anë të kësaj metoda parashikon edhe rëndësinë e secilit atribut (feature). Po ashtu një krahasim i kësaj metode bëhet me atë të rrjetave neurale ku epërsia e rrjetave neurale duket qartë, por në punimin [132] ata bëjnë një kapërcim të modelit duke e simuluar atë në mungesë të të dhënave historike pra implementojnë një minimodel online duke treguar se mund të arrihen rezultate edhe kur kemi shumë pak të dhëna. Modeli rimodelohet nga Shabani et al.[133] duke ripërpunuar teknikën e llogaritjes së ngjashmërisë midis rasteve dhe përmirësimin në algoritmin e përgjithshëm të parashikimit. Në punim autorët arrijnë rezultate me gabim CVRMSE sa gjysma e rezultateve të paraqitura në [132], po ashtu në kushtet e të dhënave të pakta metoda performon më mirë se rrjetat neurale.

1.3 Përmbledhje e kontributeve kryesore

Kontributet kryesore të kësaj teze janë si vijon:

1. Fillimisht një vëzhgim i plotë i metodave që përdoren për parashikimin e konsumit të energjisë, analizës së konsumit dhe po ashtu faktorëve kryesorë jepen duke përfshirë metodat komplekse dhe të thjeshtuar inxhinierike, metodat statistikore dhe metodat e inteligjencës artificiale.
2. Prezanton njohuri teorike për funksionimin e metodave të inteligjencës artificiale duke i ndarë në dy grupe: metodat e bazuara tek modeli (rrjetat neurale, regresioni me vektorët ndihmës, pemët e regresionit dhe pyjet e rastësishëm) dhe metodat e bazuara tek shembujt (arsyetimi bazuar tek rastet).
3. Përmirëson modelin e arsyetimit të rasteve duke parashikuar konsumin e energjisë elektrike për 6 orët e ardhshme.
4. Implementon një krahasim të gjithë metodave të sipërpërmendura duke parashikuar konsumin e orës që vijon për 5 ndërtesa të mëdha me

funksionalitete të ndryshme, duke modeluar dhe përpunuar të dhënat fillimisht për të vazhduar me parashikimin.

5. Për të njëjtat godina implementohet parashikimi dinamik i konsumit ku të dhënat vijnë hap pas hapi, metodat e përdorura janë arsyetimi i bazuar tek rastet i modeluar për parashikim “online” dhe përshtaten disa metoda të tjera mjaft të njohura për parashikime të kësaj natyre. Njohuritë teorike të të cilave jepen në kapitullin e realizimit të parashikimit.
6. Implementohet një metodë e re për parashikimin e konsumit të energjisë e njohur si intervali i parashikimit ndryshe nga më sipër ku parashikohet vetëm një vlerë.
7. Përmbledhen realizimet e kësaj pune duke lënë probleme të hapura për të ardhmen.

Kapitulli 2

NJOHURI TEORIKE MBI METODAT

2.1 Metodat inxhinierike

Metodat inxhinierike ose ndryshe të njohura si metodat e simulimeve të detajuara fokusohen në llogaritjen e nevojës për energji të ndërtesave duke u bazuar në modelimin termik të tyre, për këtë metodat përdorin veçoritë fizike të ndërtesës në procesin e modelimit. Zakonisht konsumi modelohet si në nivel global (gjithë ndërtesa) dhe në atë lokal (pjesë të saj). Këto metoda kanë njohur zhvillim të shpejtë 50 vitet e fundit. Referuar literaturës [30, 31] këto metoda mund të klasifikohen në dy nënkategori: në metodat e simulimeve të detajuara dhe në ato të simulimeve të thjeshtuara. Grupi i parë i metodave kërkon një informacion të detajuar dhe voluminoz mbi veçoritë fizike të ndërtesave për të kryer llogaritjet hap pas hapi të konsumit të energjisë për çdo element të instaluar në ndërtesë duke marrë në konsideratë informacionet shitesë mbi kushtet atmosferike, si temperatura, lagështia, shpejtësia e erës, rrezatimi, karakteristikat e ndërtesës, mënyra e funksionimit dhe paisjet e instaluara në të. Për informacione të mëtejshme mbi mënyrën e llogaritjeve si dhe zgjedhjen e modeleve mund të referohemi tek [23].

Në metodat inxhinierike përdoren zakonisht skemat RC rezistivo-kapacitive për të modeluar një sistem muresh ose të transmetimit termik të nxehtësisë dhe nisur nga kjo bëhet llogaritja e konsumit të energjisë. Në fig. 2.1 tregohet një skemë e modelimit të një ndërtese për të llogaritur konsumin e energjisë në ndërtesa.

Ku kemi temperaturat në çdo pjesë të ndërtesës dhe secila paraqet një variabël gjendjeje,

T_s : temperatura e sensorit, e cila për modelin e plotë përdoret dhe si

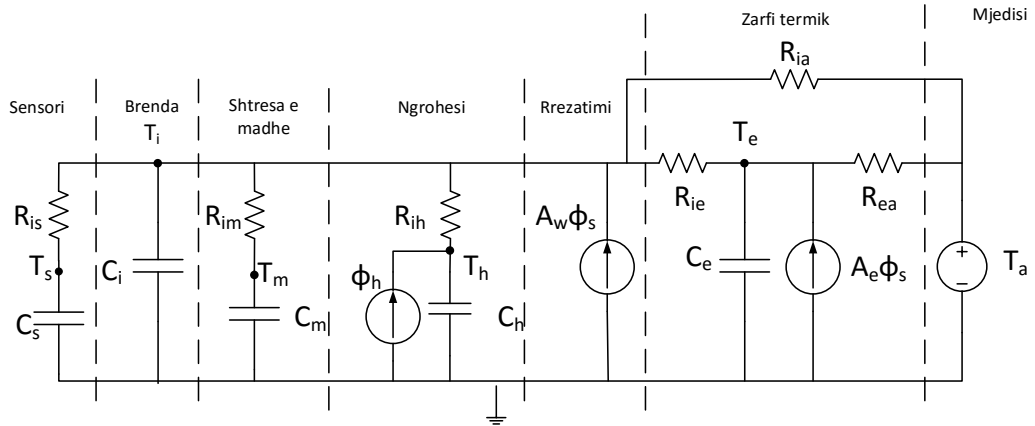


Figura 2.1: Modeli i plotë termik për një ndërtesë [23].

madhësi në dalje.

T_i : temperatura e brendshme, p.sh. ajri i brendshëm.

T_m : temperatura e mjedisit të brendshëm, p.sh. muret e brendshme si dhe mobiljet.

T_h : temperatura e paisjeve ngrohëse.

T_e : temperatura e zarfit të ndërtesës (building envelope).

Parametrat termikë të modelit paraqiten me anë të rezistencave ku:

R_{is} : rezistenca e materialit midis sensorit dhe mjedisit brenda,

R_{im} : rezistenca e materialit midis pjesës brenda si dhe mjedisit termik të brendshëm,

R_{ih} : rezistenca e materialit midis paisjeve ngrohëse dhe mjedisit të brendshëm,

R_{ia} : rezistenca e materialit midis mjedisit brenda dhe mbientit,

R_{ie} : rezistenca e materialit midis materialit brenda dhe zarfit termik,

R_{ea} : rezistenca e materialit midis zarfit termik dhe mjedisit.

Po ashtu ndërtesat janë të prirura të krijojnë kapacitete termike ku modelohen me anë të kapacitorëve.

C_s : për sensorin e temperaturës,

C_i : për brendësinë,

C_m : për muret e brendshme dhe mobiljet,

C_e : për zarfin termik.

Si dhe dy koeficientë janë përfshirë në modelim që shprehin sipërfaqen efektive të godinës e cila ka dëpërtim të energjisë së diellit që janë:

A_w : sipërfaqja efektive e dritareve për ndërtesën,

A_e : sipërfaqja efektive ku rrezatimi diellor përshkon godinën.

$$dT_s = \frac{1}{R_{is}C_s} (T_i - T_s) dt + \sigma_s d\omega_s \quad (2.1)$$

$$\begin{aligned} dT_i = & \frac{1}{R_{is}C_s} (T_i - T_s) dt + \frac{1}{R_{im}C_i} (T_m - T_i) dt + \\ & \frac{1}{R_{ih}C_i} (T_h - T_i) dt + \frac{1}{R_{ie}C_i} (T_e - T_i) dt + \\ & \frac{1}{R_{ia}C_i} (T_a - T_i) dt + \frac{1}{C_i} A_w \Phi_s dt + \sigma_i d\omega_i \end{aligned} \quad (2.2)$$

$$dT_m = \frac{1}{R_{im}C_m} (T_i - T_m) dt + \sigma_m d\omega_m \quad (2.3)$$

$$dT_h = \frac{1}{R_{ih}C_h} (T_i - T_m) dt + \frac{1}{C_h} \Phi_h dt + \sigma_h d\omega_h \quad (2.4)$$

$$dT_e = \frac{1}{R_{ie}C_e} (T_i - T_e) dt + \frac{1}{R_{ea}C_e} (T_a - T_e) dt + \frac{1}{C_e} A_e \Phi_s dt + \sigma_e d\omega_e \quad (2.5)$$

Më sipër përmendëm modelet e plota dhe të detajuara të llogaritjes së konsumit të energjisë elektrike, por në këto modele kërkohet një sasi e madhe informacioni mbi parametrat fizikë të ndërtesës. Për të thjeshtuar sasinë e llogaritjeve si dhe të evitohet mangësia apo mosgjetja e informacionit në janë propozuar modelet e thjeshtuara të cilët kanë vetëm një variabël gjendjeje T_i si dhe parametrat e tjerë si vijon:

R_{ia} : rezistenca termike e mjedisit,

C_i : kapaciteti termik i gjithë ndërtesës duke përfshirë ajrin brenda, muret e brendshme, mobiljet, etj,

A_w : sipërfaqja efektive e dritareve për ndërtesën.

Ku ekuacioni diferencial stokastik i cili përshkruan fluksin e nxehtësisë do të jetë:

$$\frac{dT_i}{dt} = \frac{1}{R_{ia}C_i} (T_a - T_i) dt + \frac{1}{C_i} A_w \phi_s + \frac{1}{C_i} \phi_h + \sigma_i \frac{d\omega_e}{dt} \quad (2.6)$$

Si dhe ekuacioni i matjes është:

$$Y_k = T_{i,k} + e_k \quad (2.7)$$

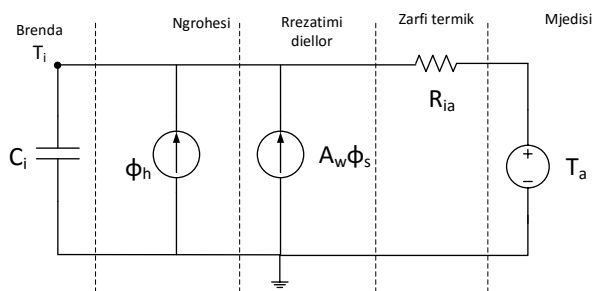


Figura 2.2: Modeli i thjeshtuar i simulimit [23].

2.2 Metodatat e Inteligjencës Artificiale

Kuptimi biologjik i “mësimit” është procesi i përsëritjes së një sjelljeje të mësuar më parë, nga shikimi, dëgjimi, përdorimi i shqisave të tjera duke imituar përsëri atë gjest. Duke kopjuar dhe më pas adoptuar një sjellje të ditur qëniet kanë arritur të mbijetojnë dhe të evolojnë. Duke tentuar adaptimin e sjelljeve biologjike për mësimin e një procesi tek makineritë ka lindur një fushë e re kërkimi që njihet si “*inteligjenca artificiale*” dhe ajo degë e njohur si “*mësimi makinë*” (machine learning). Sipas njohurive një kompjuter është i aftë të kryejë funksione komplekse.

Përse do të duhej që kompjuterat të mësonin?

Në realitet ekzistojnë probleme të cilat nuk mund të zgjidhen nga modelimi me ekuacione ekzakte si p.sh njohja e tekstit, procesimi i imazheve, parashikimi i vlerave, njohja e fytyrës, të cilat kërkojnë teknika më të avancuara duke përfshirë edhe përdorimin e procesit të mësimit. Një algoritëm mësimi që të implementohet duhet të trajnohet me eksperiencat ekzistuese të problemit dhe më tej ai testohet në situata të panjohura, të ngjashme me problemin e trajnuar.

Algoritmat e njohur përdoren për zgjidhjen e problemeve të klasifikimit dhe regresionit, të parët dallojnë midis dy ose më shumë klasave binare si, ([po, jo], [0, 1], [-1, 1], [mashkull, femër], etj...), ndërsa regresioni parashikon vlera reale si konsumin e energjisë, shpejtësinë e erës, të ardhurat, etj. Të dy tipet e metodave “mësojnë” duke u trajnuar sipas këtyre skenarëve: a) të mbikëqyrur b) pambikëqyrur c) gjysëm të mbikëqyrur d) mësimi i detyruar dhe e) mësimi online. Algoritmat e implementuar në këtë tezë do të jenë regresioni i trajnuar në mënyrë të mbikëqyrur dhe mësimi online, sipas të parit algoritmi krijon një model bazuar mbi të dhënat gjatë procesit të trajnimit ndërsa sipas të dytit algoritmi përshtat sjelljen sipas ardhjes së të dhënave. Të dhënat zakonisht janë matje të cilat vijnë në formën e serive kohore, për të qënë pajtueshëm me

algoritmat ato modelohen si vektorë me variabla hyrje dhe dalje e dëshiruar ose në formën e rasteve. Algoritmi përshtat një model bazuar në mardhënien (*hyrje, dalje*) duke e përdorur këtë model në procesin e parashikimit. Gjatë parashikimit modeli ushqehet me hyrje të panjohura më parë dhe për këto hyrje ai parashikon daljen. Metodat tradicionale të parashikimit të cilat u përmendën gjatë inspektimit në Kapitullin 1 si tregohen në fig. 2.3.

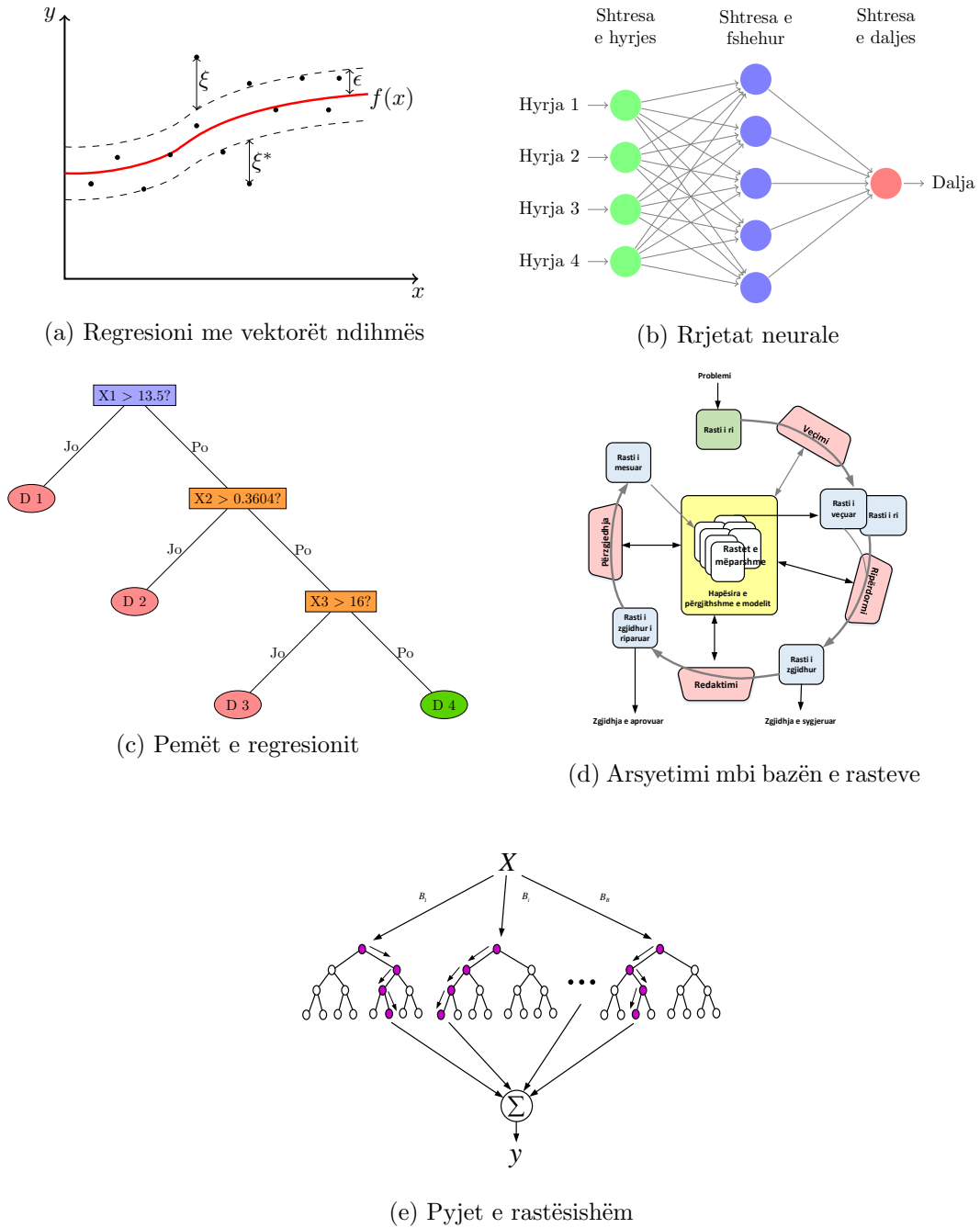


Figura 2.3: Motodat e mësimit të mbikëqyrur

2.2.1 Mësimi i supervizuar

Mësimi i supervizuar (mbikëqyrur) përfshin atë lloj algoritmesh të mëimit makinë të cilët përdorin një dataset të njohur (pjesa e trajnimit) për të bërë parashikime. Në pjesën e trajnimit përfshihen të dhënat e hyrjes dhe përgjigja në dalje. Nisur nga kjo algoritmi i supervizuar kërkon të ndërtojë një model i cili të parashikojë vlerat në dalje për hyrje nga një dataset i ri. Algoritmi mëson një *detyrë* duke përshtatur një funksion apo model tek shembujt e trajnimit. Secili shembull është një çift që përbëhet nga objekti i hyrjes (zakonisht një vektor) dhe nga një vlerë e dëshiruar e daljes. Konkretisht për një set të dhënë shembujsh trajnimi $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, një algoritëm mësimi kërkon një funksion $g : X \rightarrow Y$, ku X është hapësira e hyrjes dhe Y hapësira e daljes. Qasja e mëimit makinë është të përshtatet një model i varur nga një set parametrash:

$$y = g(x|\theta)$$

ku $g(\cdot)$ është modeli dhe θ janë parametrat e modelit. Parametrat e modeli optimizohen duke përdorur teknika të cilat minimizojnë gabimin e parashikimit.

Implementimi i një algoritmi të supervizuar mbështetet në këto hapa kryesore:

1. Përcaktohet tipi i shembujve të trajnimit. Pra fillimisht duhet fokusuar tek qëllimi i problemit që duhet zgjidhur.
2. Të sigurohen të dhënat, të cilat duhet të jenë sa më domethënëse dhe informative për zgjidhjen e problemit në fokus.
3. Përcaktohet vektori i variablave të hyrjes së modelit. Një prezantim sa më i mirë i variablave të hyrjes sjell performancën e lartë të modelit
4. Përcaktohet algoritmi më i favorshëm për zgjidhjen e problemit në fokus.
5. Eksperimentohet algoritmi, përcaktohen parametrat optimalë të tij në pjesën e të dhënave të trajnimit
6. Modeli optimal ekzekutohet me të dhëna të reja nga pjesa e testimit, dhe vlerësohet performanca e modelit.

Në praktikë ekzistojnë një mori algoritmesh mësimi të supervizuar, por nuk mund të thuhet se cili është më i miri për t'u përdorur. Secili prej tyre ka avantazhet dhe disavantazhet e tij, dhe përdorimi i tyre varet shumë nga natyra e të dhënave të problemit që duhet të zgjidhet.

2.2.2 Regresioni linear

Regresioni linear është një metodë statistikore ku në punimin tonë përdoret si model ndihmës në llogaritje, dhe jo si një model bazë parashikimi. Metoda përshtat një funksion matematikor tek të dhënat, duke marrë parasysh kontributin e të gjithë variablave në madhësinë e daljes duke i atribuar çdo variabli një peshë sipas ndikimit që ka mbi përgjigjen në dalje. Saktësia e modelit varet nga mënyra e përcaktimit të peshës së secilës variabël mbi madhësinë në dalje, në mënyrë që dalja e parashikuar të përafrohet sa më shumë me daljen reale. Vlerat në hyrje dhe dalje të modelit duhet të jenë vlera numerike. Megjithatë kjo metodë është e përshtatshme në rastet kur të dhënat (variablat) mund të veçohen nga një funksion linear pra një drejtëz, në rastin e kundërt kur të dhënat janë vështirësisht të veçueshme atëherë modele më kompleks regresioni tentojnë të përshtasin modelin, duke minimizuar gabimin midis diferencimit të variablave.

Modelimi matematik i problemit do të ishte si më poshtë. Në matematikë apo statistikë një model i regresionit linear bën të mundur që të përafrojmë një model matematik apo statistik me të dhënat që disponojmë të cilat do të shprehen me anë të parametrave që nuk njohim. Parimi i këtij modeli është i tillë që të zgjidh një problem ekuacionesh linear duke minimizuar gabimin midis të dhënave dhe modelit. Përgjithësisht do të kishim për një sistem ekuacionesh lineare:

$$\sum_{j=1}^n X_{i,j}\beta_j = y_i, \quad (i = 1, 2, \dots, m) \quad (2.8)$$

ku m numri i ekuacioneve me n të panjohura $\beta_1, \beta_2, \dots, \beta_n$ me $m > n$, problemi do të shkruhej në formën matricore si:

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{y} \quad (2.9)$$

ku më qartësisht do të shpreheshin:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & & x_{2n} \\ \vdots & & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (2.10)$$

Zakonisht sisteme të tilla janë pa zgjidhje, por qëllimi është që të gjemë koeficientët β të tillë që ju afrohen më shumë ekuacioneve

në termat e problemit të *minimizimit kuadratik*.

$$\hat{\beta} = \arg \min_{\beta} S(\beta) \quad (2.11)$$

Ku funksioni S do të shprehej si:

$$S(\beta) = \sum_{i=1}^m \left| y_i - \sum_{j=1}^n X_{ij} \beta_j \right|^2 = \|\mathbf{y} - \mathbf{X}\beta\|^2 \quad (2.12)$$

Ku zgjidhja unike do të ekzistonte dhe vlerat e koeficienteve β do të përcaktoheshin nga:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (2.13)$$

Por duke qënë se vlerat e peshave për variablat na duhen pozitive si dhe shuma e tyre të konvergjojë në 1 atëherë do të përdorim regresionin linear të kushtëzuar, dhe kushti i vetëm që vendosim është ai që nuk lejon vlera më të vogla se 0 dhe relacioni më poshtë tregon modifikimin.

$$\arg \min_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|^2, \text{ me kushtin që } X \geq 0 \quad (2.14)$$

Nocionet matematike do të formalizoheshin në fig. 2.4:

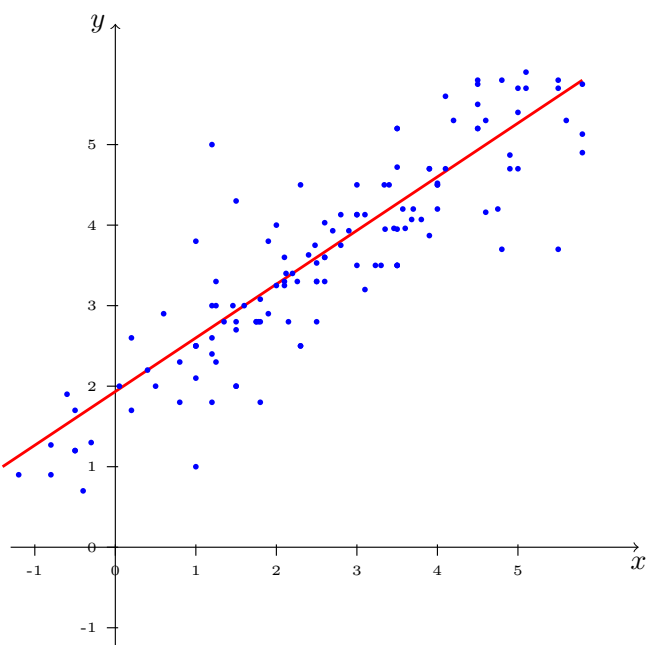


Figura 2.4: Përshtatja e një modeli linear në të dhënat.

2.2.3 Rrjetat neurale artificiale

Rrjetat neurale të cilat frymëzohen nga rrjetat neurale biologjike, operojnë si grupime të njëjësive elementare të quajtura “*neurone*” [168]. Rrjetat funksionojnë me hapa ku në fazën e parë bëhet grumbullimi i dijes eksperimentale dhe memorizimi i saj për hapat e mëtejshëm. Procesi ngjason me funksionimin e trurit të njeriut, ku *dija e cila fitohet nëpërmjet procesit mësimi, nga ndërlidhjet midis neuroneve, të njohura si ngacmime sinaptike, informacioni i mësuar ruhet në neurone për t’u përdorur si eksperiencë e mëvonshme.* Në ndryshim nga rrjetat biologjike ku mbas krijimit të rrjetës lidhjet e neuroneve shkëputen ose shtohen, në rrjetat artificiale çdo lidhjeje midis neuroneve i atribohet një peshë e cila përditësohet hap pas hapi nga algoritmi që mëson ose ndryshe nga algoritmi i optimizimit. Sinjali hyrës përparon në drejtimin e lidhjeve të krijuara deri sa arrihet dalja. Peshat përditësohen nga algoritmi i cili tenton të minimizojë gabimin midis vlerës së parashikuar nga modeli dhe vlerës reale të sinjalit.

Një tablo e strukturës së rrjetave neurale e paraqesim më poshtë duke nisur me qelizën bazë neuronin, pastaj me mënyrën e përhapjes së sinjalit dhe me algoritmin optimizues.

Struktura e rrjetave neurale

Një rrjetë neurale përbëhet nga shtresa si në fig. 2.5 të cilat ndërlidhen midis tyre me neuronet përkatëse. Rrjeta ka në qendër të saj “neuronin” ose ndryshe njësinë bazë. Zgjidhja e problemeve me rrjetat neurale u lehtësua me përdorimin e rrjetave me shumë shtresa me përhapje të përparme, ku modeli më i thjeshtë është një shtresë hyrjeje, një shtresë e fshehur me disa neurone dhe shtresa e daljes fig. 2.5. Informacioni i cili vjen në formën e vektorëve kalon nga shtresa e hyrjes pa u përpunuar ku daljet e neuroneve të hyrjes japin informacionin tek neuronet e fshehura dhe bashkë me të secilës dalje i kanë atribuar nga një peshë, informacioni përpunohet nga neuronet e fshehura duke bërë mesatarizimin e shumës së hyrjeve me peshat dhe duke përdorur një funksion shmangieje që më i përdorshmi është funksioni sigmoidal që merr vlera nga 0 në 1,

$$f(z) = \frac{1}{1 + \exp(-z)}$$

Informacioni përpunohet për të dhënë vlerën në dalje të rrjetës.

Neuroni artificial

Referuar informacionit që përshkruam më sipër po detajojmë funksionimin e neuroneve. Modeli kompleks i një neuroni natyral thjeshtohet duke

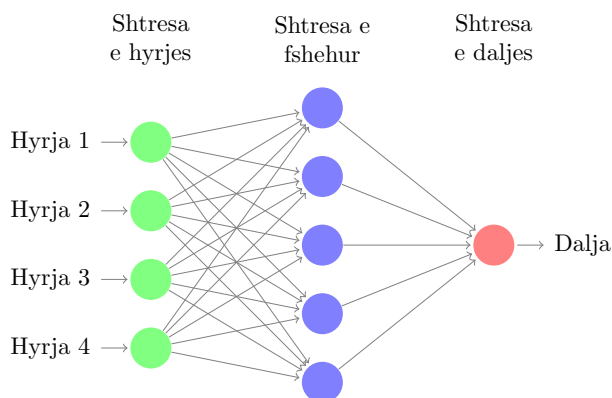


Figura 2.5: Struktura tipike e një rrjete neurale me shumë shtresa.

sjellë në vëmendje modelet matematike për modelimin e neuronit dhe të përftojme neuronin artificial “*njësia*”. Të dhënat e modeluara për problemin e caktuar ushqejnë rrjetën, ku vlerat e secilës variabël transmetohen në secilin neuron të shtresës hyrëse (neuronet e hyrjes sa atributet), informacioni në neuronet hyrëse nuk përpunohet, por ai transmetohet tek nyjet e fshehura. $a = (a_1, a_2, \dots, a_n)$ me peshat e atribuara për to $w = (w_1, w_2, \dots, w_n)$ nga që i është bashkangjitur çdo hyrjeje si dhe duke i shtuar një term konstant “*shmagjia*” që ka vlerën e barabartë me 1, duke përdorur një funksion jolinear aktivizues.

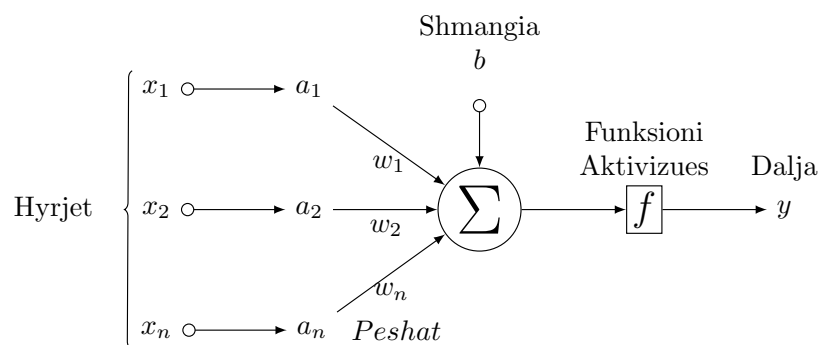


Figura 2.6: Modeli i një neuroni artificial.

$$z = b + \sum_{i=1}^N a_i w_i \tag{2.15}$$

$$a_{dalje} = f(z) \tag{2.16}$$

Përderisa dalja do të ketë vlerë reale dhe jo të transformuar atëherë për rrjetën me një dalje funksioni në dalje është shuma e produktit të hyrjeve me peshat në neuronin e daljes.

$$y = \sum_{h=0}^n v_h z_h$$

Përgjithësisht neuronet e hyrjes dhe daljes përdorin funksion aktivizues linear, ndërsa neuronet e fshehura atë jolinear.

Rrjetat me përhapje të përparme

Rrjetat neurale me përhapje të përparme janë një tip i rrjetave neurale ku informacioni lëviz vetëm sipas një drejtimi, nga hyrja drejt daljes. Multilayer perceptron (MLP) është një klas i rrjetave me përhapje të përparme e cila përbëhet minimalisht nga tre shtresa (*layers*) ku secila shtresë përbëhet nga grupimi i njësive bazë të cilat japin informacionin tek njësitë e tjera brenda shtresës, por që nuk janë të lidhura midis tyre. Shtersat kryesorë në një MLP janë tre, *shtresa e hyrjes*, *shtresa e fshehur* (një apo disa) dhe *shtresa dalëse*. Shtresa hyrëse është shtresa e cila nuk merr informacion nga asnjë shtresë tjetër, por ruan në të vektorin e vlerave të hyrjes që shërben si aktivizues për njësitë e saj. Përmendëm se shtresa e fshehur mund të jetë më shumë se një, dhe në rastin e përgjithshëm shtresa e parë e fshehur është e lidhur drejtpërdrejt me shtresën e hyrjes dhe plotësisht transmeton informacion tek shtresa pasardhëse e fshehur. Shtresa e fundit e fshehur lidhet me shtresën dalëse, ku aktivizimi i njësive në këtë shtresë shërben si dalje e rrjetës. Procesi i llogaritjes së daljes së rrjetës me përhapje të përparme llogaritet në tre hapa kryesorë.

1. Hyrjet e rrjetës shërbejnë si aktivizues i njësive hyrëse,
2. Shtersat e fshehura i llogarisin aktivizimet e tyre sipas një rregulli topologjik,
3. Shtresa dalëse llogarit aktivizimin e saj dhe e transmeton si dalje të rrjetës.

MLP-të përdoren shpesh për të përafuar funksione të panjohura nga vlerat hyrëse në ato dalëse.

Rrjetat me përhapje të prapme

Rrjetat me përhapje të prapme njihen si algoritmi më i përdorur i RN në praktikë për problemet e mbykëqyrura, me veçorinë e përhapjes së pasme të gabimeve. Peshat rregullohen duke tentuar të minimizohet shuma e katrorëve të gabimeve.

$$E = \frac{1}{2} \sum_{i=1}^N (Dalja_i - Reale_i)^2 \quad (2.17)$$

Ku *Dalja* tregon vlerën e parashikuar nga modeli ndërsa *Reale* tregojnë përgjigjet për N vlerat e madhësive në hyrje të dhëna nga përdoruesi ose modeluesi i sistemit. Në vijim ideja e funksionimit të algoritmit do të përshkruhet në mënyrë më të detajuar, këtu po japim disa nocione bazë të funksionimit të një rrjetë me përhapje të prapme. Peshat e rrjetës do të përcaktohen si një problem i minimizimit të funksionit të gabimit kundrejt peshave të rrjetës. Problemi zgjidhet duke e modeluar sipas ndryshimit negativ të gradientit,

$$-\frac{\partial E}{\partial w} \quad (2.18)$$

kundrejt ndryshimit të peshës Δw_{ij} në propocion me shkallën e mësimimit α (konstante pozitive), si dhe pjesë të ndryshimit të peshave të një hapi më parë i quajtur moment i shtohen ndryshimit të ri, me qëllim që procesi i mësimimit të përshpejtohet [169].

$$\Delta \tilde{w}_{ij} = \beta \Delta w_{ij} - \alpha \frac{\partial E}{\partial w_{ij}} \quad (2.19)$$

$$\tilde{w}_{ij} = w_{ij} + \Delta w_{ij} \quad (2.20)$$

ku \tilde{w}_{ij} është vlera e re e peshës. Qëllimi kryesor i algoritmit është gjetja e gradientit të gabimit. Marrim në analizë një rrjetë me L shtresa me topologji sipas rendit hyrja në fillim dhe dalja në fund. Llogaritja do të ndahet në tre hapa kryesorë, ku k është treguesi i shtresës U_k tregon numrin e njësive që ndodhen në këtë shtresë, si dhe shtresa ka një matricë peshash w_{ij}^k .

1. **Përhapja e përparme.** Vektori i hyrjes a kopjohet në aktivizimi a_i^1 të njësive të shtresës hyrëse. Për secilën njësi në çdo shtresë të fshehur apo të daljes llogariten potenciali z_i^k dhe aktivizimi a_i^k .
2. **Përhapja e prapme.** Në këtë hap llogaritet derivati i gabimit për shtresën e daljes në formën

$$\Delta_i^L = (Dalja a_i - a_i^L) \frac{\partial f(z_i^L)}{\partial z_i^L} \quad (2.21)$$

Ndërsa e veçanta e këtij hapi është implementimi i peshave në shtresën e fshehur duke bërë ushqimin në kah të kundërt të rrjetës me aktivizimet, pra nisemi nga $h = L - 1$ deri në $h = 2$ dhe derivati gabimit është:

$$\Delta_i^h = \sum_{j=1}^{U_{h+1}} \Delta_j^{h+1} w_{ij}^{h+1} \frac{\partial f(z_i^h)}{\partial z_i^h} \quad (2.22)$$

3. **Përditësimi i peshave.** Peshat në secilën shtresë llogariten si vijon:

$$\Delta \tilde{w}_{ij}^k = \beta \Delta w_{ij}^k + \alpha \Delta_i^{k+1} a_j^k \quad (2.23)$$

$$\tilde{w}_{ij}^k = w_{ij}^k + \Delta w_{ij}^k \quad (2.24)$$

Peshat përditësohen hap pas hapi duke llogaritur gabimin total që kontribuon tek të gjithë shembujt.

Modeli i rrjetës neurale të përdorur në këtë model është një rrjetë “multilayer perceptron” me algoritëm mësimi Levenberg-Marquardt [145, 146, 152] i zhvilluar nga Kenneth Levenberg dhe Donald Marquardt, ku jep zgjidhje për problemet me sasi të dhënash të pakta dhe mesatare për nga numri i instancave trajnuese duke bërë inimizimin e funksioneve lineare. Ky është një algoritëm i shpejtë për nga ana e ekzekutimeve dhe ka konvergencë të qëndrueshme.

2.2.4 Makinat me vektorë ndihmës (MVN)

MVN-të janë një grup metodash që përshtasin modele bazuar tek të dhëna paraprake (*hyrje, dalje*). Këto mendohet të jenë ndër metodat më konkurruese në mësimin e supervizuar në problemet e klasifikimit, regresionit, dhe mësimin gjysëm të supervizuar. Ideatori i kësaj metode njihet Vapnik i cili së bashku me bashkpunëtorët e tij në 1992 hodhi idenë bazë [156], si dhe më tej ideja e tij gjeti zhvillim në [95, 165, 167]. Gjithashtu botimi i tij [164] hedh dritë mbi ecurinë e mëtejshme të metodës. Metoda më vonë gjen shumë aplikimi si në dedektimin e sëmundjeve [158, 159], në kategorizimin e tekstit dhe imazheve [160, 161] si edhe në marrjen e vendimeve [162, 163].

Ideja bazë e këtyre metodave është minimizimi i gabimit gjatë trajnimit të modelit, duke bërë minimizimin e gabimit në kompleksitetin e modelit. Kjo ndryshon nga metodat e tjera të minimizimit të gabimit gjatë trajnimit [164]. Bazuar mbi këtë parim MVN-të arrijnë të krijojnë një performancë më të lartë në zgjidhjen e problemeve jo-lineare referuar metodave të tjera të supervizuara. Për të kuptuar parimin e funksionimit të MVN nisemi nga problemet e klasifikimit ku atributi i cili do të parashikohet ka vlera binare p.sh $[0, 1]$, $[po, jo]$, $[-1, 1], \dots$ etj. Parashtrajmë fillimisht zgjidhjen e problemit të klasifikimit për të arritur tek problemet e regresionit.

Principet e makinave me vektorë ndihmës

Klasifikimi ka si qëllim gjetjen e një plani me kufij sa më të gjerë i cili bën ndarjen midis dy klasave. Për një set trajnimi X me n kampionë ku x_i

tregon kampionin e i -të, ndërsa y_i tregon vlerën e daljes për këtë kampion që mund të jete -1 ose 1 . Qëllimi është që të gjendet plani i cili bën ndarjen e këtyre dy klasave, dhe të përshtat një model i cili bën parashikimin me kufij sa më të gjerë të dhëna të reja (të panjohura) me një saktësi sa më të lartë. Siç tregohet në fig. 2.7, plani ndarës duhet të jetë sa më larg nga klasat. Problemi për gjetjen e planit të kërkuar formulohet në këtë mënyrë,

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (2.25)$$

kushtin që $y_i(w^T x + b) \geq 1, i = 1, \dots, n$

ku $g(x) = 1$, nëse $x \geq 0$, ose $g(x) = -1$, nëse $x \leq 0$. Kështu në fig. 2.7 plani më i mirë që mund të ndajë dy klasë është $w^T x + b = 0$, kështu problemi kthehet në gjetjen e vlerave të parametrave w dhe b . Intuitivisht plani ideal ka largësinë maksimale nga shembujt e dy klasave. Kështu gjetja e planit kthehet në një problem optimizimi të parametrave. Ky është një problem optimizimi

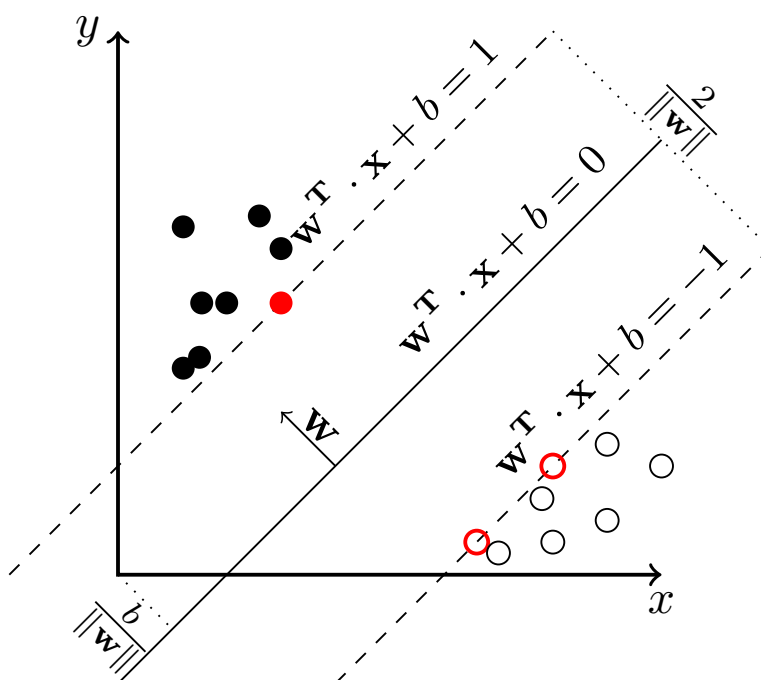


Figura 2.7: Modeli i një vektori ndihmës.

kuadratik me kushte lineare, i cili mund të zgjidhet me metodat e zgjidhjes së problemeve kuadratikë. Megjithatë problemet në praktikë janë probleme jolineare, që nuk mund të veçohen dy klasë në mënyrë lineare. Për të bërë të mundur zgjidhjen e këtyre problemeve një rregullues l_1 i shtohet

funksionit:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + \sum_{i=1}^n \xi_i$$

kushtin që

$$y_i(w^T x + b) \geq 1 - \xi_i, i = 1, \dots, n$$

$$\xi_i \geq 0, i = 1, \dots, n$$
(2.26)

Ku ξ_i është një variabël e lirë që i shtohet një kampioni i . Kjo do të bëjë të mundur që klasifikuesi t'i japë mundësi kampioneve të përjashtuar që janë klasifikuar gabim të rifuten në model. Duke futur termin rregullues në sistem atëherë zgjidhja e problemit bëhet më komplekse. Në ndihmë të zgjidhjes na vijnë koeficientët e Langrazhit të cilët japin zgjidhjen duale të këtij problemi optimizimi. Zgjidhja duale jep zgjidhje të fundme dhe na lejon mundësinë të përdorim “*dredhinë*” e Kernelit për të zgjidhur problemet jolineare. Problemi duke përdorur koeficientët e Langrazhit shtrohet:

$$L(w, b, \xi, \alpha, r) = \frac{1}{2} w^T w + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i [y_i(w^T x_i + b) - 1 + \xi_i] - \sum_{i=1}^n \xi_i r_i$$
(2.27)

Ku α, r janë koeficientët e langrazhit me kushtin që $\alpha_i, r_i \geq 0$. Në ekuacionin më sipër duke bërë derivatet e pjesshme në lidhje me w, b dhe ξ si dhe duke i barazuar këto me zero. Ekuacionet e përfutuara i vendosim tek funksioni i Langrazhit dhe merret zgjidhja duale e problemit e cila është:

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n y_i y_j \alpha_i \alpha_j < x_i, x_j >$$

kushtin që

$$0 \leq \alpha_i \leq C, i = 1, \dots, n$$

$$\sum_{i=1}^n y_i \alpha_i, i = 1, \dots, n$$
(2.28)

Ku α është vektori i n variablave i cili do të optimizohet, nëse zgjidhim α atëherë funksioni i cili përcakton planin do të jetë:

$$w^T x + b = \left(\sum_{i=1}^n y_i \alpha_i x_i \right)^T + b$$

$$= \sum_{i=1}^n y_i \alpha_i < x_i, x > + b$$
(2.29)

ku forma është e shpërndarë pasi shumica e koeficientëve α_i janë zero. Modeli kuadratik për MVN mund të shkruhet si më poshtë:

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - \sum_{i=1}^n y_i \alpha_i$$

kushtin që

$$y^T \alpha = 0$$

$$0 \leq \alpha_i \leq C, \forall i = 1, 2, \dots, n$$
(2.30)

Ku Q është një matricë semi-definitive me dimensione n me n . Ku secili element i Q ka trajtën $Q(i, j) = y_i y_j \langle x_i, x_j \rangle$, si dhe $\langle x_i, x_j \rangle$ është produkti skalar midis x_i dhe x_j . Nëse do të zëvendësohet x_i me një funksion të saj $\phi(x_i)$ dhe produkti skalar do bëhej $\langle \phi(x_i), \phi(x_j) \rangle$ atëherë në algoritëm nuk do të kishte ndryshim në zgjidhje por të dhënat e hyrjes do të transformoheshin nga një dimension në një dimension tjetër. Kuptohet se të dhëna jolinearisht të ndashme duke u transformuar në një dimension tjetër do të jepnin mundësinë e gjetjes së planit që bën ndarjen lineare të tyre siç tregohet në fig. 2.8.

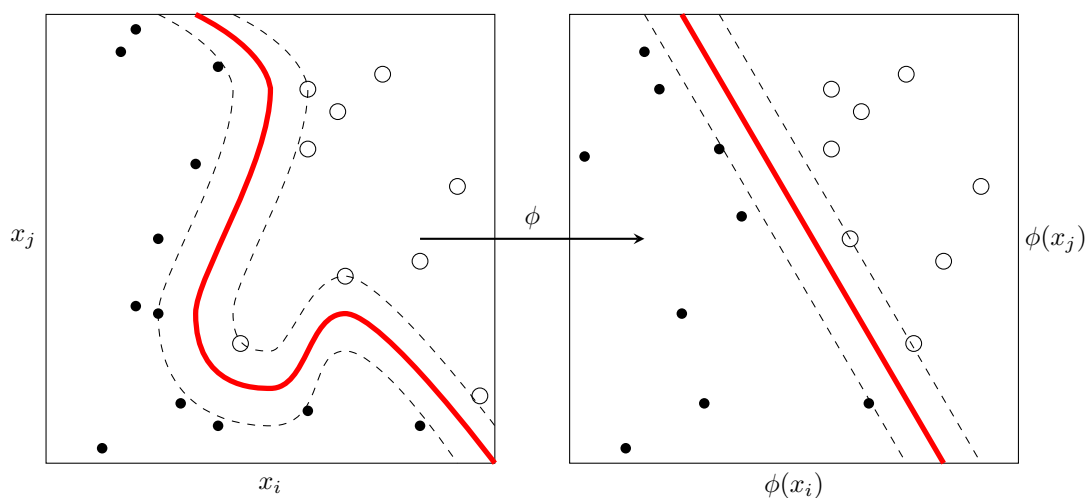


Figura 2.8: Modeli i një transformimi me anë të funksionit Kernel.

Funksionet Kernel

Nëse kemi një pozicionim të caktuar të të dhënave ϕ , funksioni kernel përcaktohet si $K(x, z) = \phi(x)^T \phi(z)$, atëherë produkti i brendshëm $\langle \phi(x_i), \phi(x_j) \rangle$ do të shkruhej si $K(x_i, x_j)$. Atëherë matrica Q do të kishte për elementë $Q(i, j) = y_i y_j K(x_i, x_j)$ dhe kjo njihet si matrica kernel. Por lind natyrshëm pyetja se cilët funksione mund të kene veçoritë e një funksioni kernel? Dhe përgjigjen e japin kushtet Mercer si më poshtë:

Teoremë 1 *Jepet $K : R^N \times R^N \rightarrow R$, që K të jetë funksion kernel është e nevojshme dhe mjaftueshme që për çdo set trajnimi $\{x_1, x_2, \dots, x_n\}, n \leq \infty$, matrica korrensponduese Q duhet të jetë simetrike, pozitive gjysëm-definitive.*

Funksionet kernel më të përdorshëm në praktikë janë:

1. Funksioni linear $K(x, z) = x^T z$,
2. Funksioni polinomial $K(x, z) = (\gamma x^T z + koef)^d$,

3. Funkzioni Radial Basis $K(x, z) = \exp(-\gamma |x - z|^2)$,
4. Funkzioni sigmoid $K(x, z) = \tanh(\gamma x^T z + koeff)$.

Regresioni me vektorët ndihmës

Këtu x dhe z përdoren për të zëvendësuar dy kampionet x_i dhe x_j . Ideja e përdorimit të funksioneve kernel nuk i përket vetëm MVN, por në çdo algoritëm tjetër i cili përdor produktin e brendshëm skalar [164].

Më sipër u diskutua për modelet klasifikuese pra ku qëllimi ishte klasifikimi midis klasave të ndryshme, por në praktikë shpesh herë na kërkohet parashikimi i vlerave numerike dhe të vazhdueshme, një shembull do të ishte parashikimi i konsumit të energjisë, problem i cili nuk mund të zgjidhet si klasifikim, por si regresion. Regresioni me vektorët ndihmës është i dizenuar për këtë qëllim [166]. Për ta bërë vlerësimin sa më të qëndrueshëm edhe sa më përfshirës do të prezantojmë funksionin ϵ të humbjes i propozuar nga Vapnik [167].

$$L(y - f(x)) = \begin{cases} 0 & \text{nëse } |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon & \text{në çdo rast tjetër} \end{cases} \quad (2.31)$$

Funksioni i humbjes përcakton një zonë përreth lakores së parashikimit siç tregohet në fig. 2.9.

Kjo nënkupton se nuk ekziston ndonjë devijim midis vlerave që duam

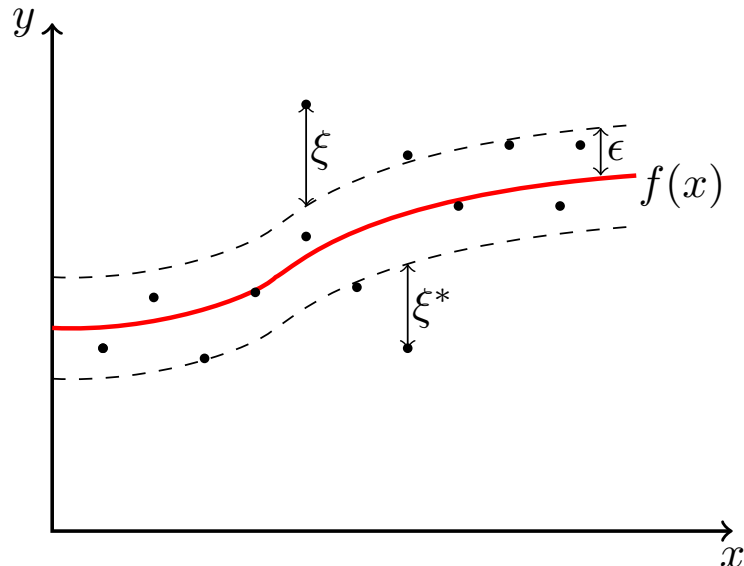


Figura 2.9: Zona ϵ e një regresioni me vektorë ndihmës

të parashikojmë kundrejt atyre të matura nëse ato ndodhen brenda zonës

të përcaktuar nga ϵ , është kjo arsyeja që në këtë pjesë prezantojmë dy variabla të lira ku në pjesën e klasifikimit kishim vetëm një. Variablat e lira janë ξ_i, ξ_i^* , $i = 1, 2, \dots, n$, është:

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^n \xi_i + \sum_{i=1}^n \xi_i^* \right) \\ \text{kushtin që} \quad & y_i - f(x_i) \geq \epsilon + \xi_i^* \\ & f(x_i) - y_i \geq \epsilon + \xi_i \\ & \xi_i^*, \xi_i \geq 0, \quad i = 1, 2, \dots, n \end{aligned} \quad (2.32)$$

ku C është një konstante rregulluese, e cila përcakton balancën midis kapacitetit të $f(x)$ si dhe pikave të cilat ndodhen jashtë zonës ϵ . Gjetja e zgjidhjes së funksionit të shprehur në (2.32), lidhet me përcaktimin e koeficientëve të Langrazhit, $\alpha, \alpha^*, \gamma, \gamma^*$, dhe funksioni i Langrazhit merr trajtën:

$$\begin{aligned} L(w, b, \xi^*, \xi, \alpha^*, \alpha, \gamma, \gamma^*) = & \frac{1}{2} \|w\|^2 \\ & + C \left(\sum_{i=1}^n \xi_i + \sum_{i=1}^n \xi_i^* \right) \\ & - \sum_{i=1}^n \alpha_i [y_i - (wx_i) - b + \epsilon + \xi_i] - \alpha_i^* [y_i - (wx_i) - b + \epsilon + \xi_i^*] \\ & - \sum_{i=1}^n n(\gamma_i^* \xi_i^* + \gamma \xi) \end{aligned} \quad (2.33)$$

Katër koeficientët e Langrazhit kënaqin kushtet $\alpha \geq 0, \alpha^* \geq 0, \gamma \geq 0, \gamma^* \geq 0, i = 1, 2, \dots, n$ nëse:

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial b} = \frac{\partial L}{\partial \xi^*} = \frac{\partial L}{\partial \xi} = 0$$

Kushtet e gjetura nga aplikimi i derivateve të pjesshme i vendosim në ekuacionin e Langrazhit, problemi kthehet në problem optimizimi ku zgjidhja është maksimumi i ekuacionit (2.33) duke kënaqur koeficientët e Langrazhit. Hapi tjetër është gjetja e α_i dhe α_i^* për të maksimizuar ekuacionin e mëposhtëm:

$$W(\alpha_i^*, \alpha_i) = \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) - \epsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) (x_i \cdot x_j) \quad (2.34)$$

Duke zëvendësuar produktin skalar midis vektorëve x_i, x_j me një funksion kernel $K(x_i, x_j)$, si dhe duke përdorur vlerat e α dhe α^* të gjetura nga ekuacioni (2.34) funksioni objektiv, përfundimisht funksioni që përshtat modelin do të jetë:

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(x_i, x) + b \quad (2.35)$$

2.2.5 Pemët e regresionit

Në këtë pjesë do të prezantojmë një klasë të modeleve jolineare të parashikimit, të cilat në dukje duken tepër të thjeshta dhe njihen si pemët e parashikimit dhe ndahen në dy modele në varësi të funksionit, në *pemët e vendimit* (decision trees) dhe *pemët e regresionit* (regression trees). Ideja bazë është tepër e thjeshtë, nëse duam të parashikojmë një përgjigje ose një klasë Y nga variablat e hyrjes X_1, X_2, \dots, X_n e bëjmë këtë duke rritur një pemë binare në secilën nyje të brendshme të pemës duke aplikuar një test të zakonshëm për një variabël X_i të çfardoshme. Në varësi të përgjigjes që na jep testi, kemi dy zgjidhje ose shkojmë në degën e majtë ose në degën e djathtë, ku praktikisht kemi arritur në një gjethe, këtu kryhet parashikimi. Në këtë hap parashikimi bëhet duke mbledhur apo berë mesataren e parashikimeve në secilën gjethe. *Pse e bëjmë këtë?* Modelet parashikuese si regresioni linear apo polinomial japin zgjidhje globale, ku një formulë e vetme aplikohet mbi të gjitha të dhënat. Ndërsa në rastet kur kemi të dhëna me karakteristika jolineare ose ku pjesa më e madhe e variablave ndërveprojnë në mënyrë të komplikuar është e vështirë të përshtatet një model i tillë.

Në rastet e problemeve komplekse me të dhëna jolineare bëhet ndarja e të dhënave në pjesë më të vogla, ku ndërveprimi i variablave është më i thjeshtë. Këtu zgjidhja bëhet në çdo nënpjesë duke bashkuar zgjidhjet *lokale* për të dhënë zgjidhjen e përgjithshme.

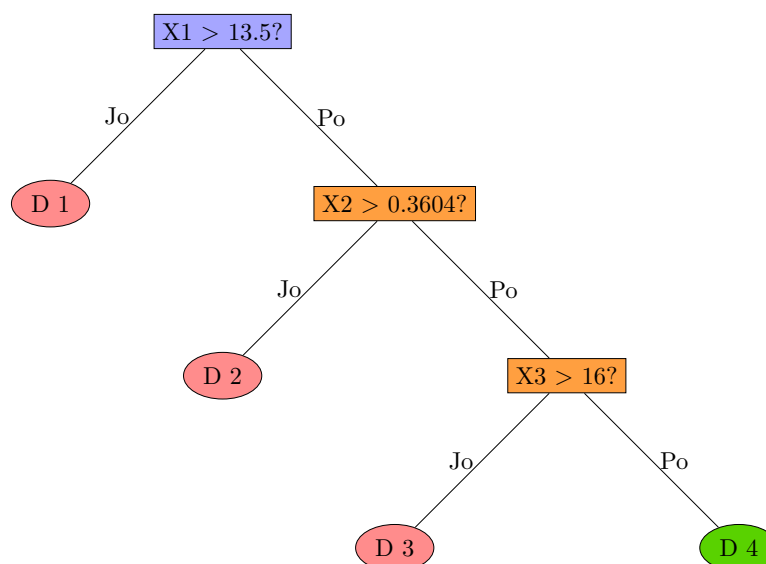


Figura 2.10: Modeli i një peme.

Përshtatja e modelit të pemës në problemin e dhënë

Duke ju referuar [170] procesi i ndërtimit të një peme kalon në dy faza kryesore:

1. Hapsira e të dhënave, X_1, X_2, \dots, X_n ndahet në J pjesë të veçanta që nuk mbivendosen R_1, R_2, \dots, R_J
2. Për secilën matje që ndodhet në zonën j , parashikimi është thjesht mesatarja e të gjitha vlerave të përgjigjeve që ndodhen në R_j

Duke analizuar më tej fazën e parë, zonat i ndërtojmë në formën e drejtkëndëshave për arsye thjeshtësie. Qëllimi është të ndahet hapësira e të dhënave në pjesë të cilat kanë gabimin më të ulët.

$$E = \sum_{j=1}^J \sum_{i \in R_j} (y_i - \hat{y}_i)^2 \quad (2.36)$$

Zakonisht qasja e përdorur për ndarjen e hapësirës njihet si ajo *lart-poshtë* (*top-down*). Ndarja e hapësirës është një proces rekursiv, kështu që për një variabël X_j të zgjedhur rastësisht merret një vlerë ndarëse e çfardoshme s e cila e ndan hapësirën e të dhënave në dy pjesë $\{X|X_j < s\}$ dhe $\{X|X_j > s\}$. Pra na u krijuan dy gjysëm plane, këtu fillojmë të kërkojmë për vlerat e j dhe s të cilat japin planet optimale.

$$E_{min} = \sum_{i: x_i \in R_1(j,s)} (y_i - \hat{y}_{R1})^2 + \sum_{i: x_i \in R_2(j,s)} (y_i - \hat{y}_{R2})^2 \quad (2.37)$$

ku vlerat e \hat{y}_{R1} është përgjigja mesatare e vlerave të ndodhura në planin R_1 , ndërsa \hat{y}_{R2} është përgjigja mesatare e vlerave të ndodhura në planin R_2 . Tani procesi përsëritet disa herë duke vazhduar në nivele më poshtë të ndarjeve si tregohet në fig. 2.10. procesi i ndarjes vazhdon deri në plotësimin e një kushti paraprak të vendosur nga përdoruesi i algoritmit, kushti zakonisht i referohet numrit minimal të variablave që duhet të jenë në një ndarje.

Një algoritëm tipik i ndërtimit të një peme T sipas [155] do të jetë si vijon.

Krasitja e pemës

Procesi i ndërtimit të pemës duke përdorur mënyrën rekursive të ndarjes, mund të jetë shumë efikas në pjesën e trajnimit, por mund të vuajë nga problemi i mbingopjes. Ky problem sjell ulje të performancës së modelit. Në praktikë problemi zgjidhet duke ndërtuar pemën e madhe T_0 dhe pastaj kjo

pemë krasitet për të përftuar një *nënpemë të re* T . Pyetja që shtrohet është se cila do të jetë pema optimale mbas procesit të “*krasitjes*” pasi nënpemët e mundshme janë një pafundësi. Kjo bëhet duke llogaritur gabimin e çdo peme të mundshme, por marrja në konsideratë e secilës pemë rrit kohën e veprimit të algoritmit. Kështu prezantohet një konstante e njohur si *konstantja e kompleksitetit të kostos së krasitjes* α . Për secilën vlerë të α do të kemi një nënpemë $T \subset T_0$ e cila:

$$E = \sum_{m=1}^{|T|} \sum_{x_i \in R_m} (y_i - \hat{y}_{R_m})^2 + \alpha|T| \quad (2.38)$$

ku vlera e gabimit do të jetë minimale. Vlera e konstantes α kontrollon kompleksitetin e nënpemës që do të prodhohet, p.sh kur $\alpha = 0$ do të thotë se nënpema është e njëjtë me pemën ose kur α shumë e madhe nënpema është tepër e thjeshtë. Por si do ta gjemë vlerën optimale të α ? Zgjidhja vjen duke gjeneruar nënpemët si funksion i α , dhe mbi këto nënpemë të krijuara vlerësojmë gabimin duke përdorur teknikën e vlerësimit kryq me pjesë. Pasi përcaktohet α gjenerojmë *nënpemën* optimale, duke e mbyllur këtu procesin e mësimin. I gjithë procesi i gjenerimit të modelit të pemës do të jepej në hapat më poshtë:

1. Për të dhënat përdorim copëtimin binar, duke rritur një pemë të madhe deri sa numri i variablave që ndodhen në një nyje është më i vogël se kushti i paracaktuar. Ndarja bëhet duke provuar se cili variabël gjeneron degë me gabim më të vogël.
2. Pemën e gjeneruar e krasisim duke provuar vlera të ndryshme α .
3. Përcaktojmë vlerën e α që gjen pemën më të mirë duke përdorur vlerësimin kryq.
4. Gjenerojmë *nënpemën* e cila i korrensopdon vlerës optimale të α .

2.2.6 Pyjet e rastësishëm

Pyjet e rastësishëm janë modele të bashkësive të cilat mbështeten mbi pemët e regresionit si modele bazë [170]. Këto modele lindën si nevojë për të eliminuar një problem i cili ekziston tek pemët e regresionit. Pemët vuajnë nga problemi i *variancës* së lartë, që do të thotë se nëse të dhënat e trajnimit do të ndaheshin në dy pjesë rezultatet e përfuara nga përshtatja e pemëve në dy pjesët veçmas do të ndryshonin shumë. Eliminimi i kësaj dobësie vjen duke prezantuar teknikën e “*bootstrapping*”, ku sipas saj nga hapësira e të dhënave gjenerojmë disa sete të tjera duke ripërsëritur një pjesë të vlerave. Duke u gjeneruar një mori setesh $1, \dots, B$ nga të dhënat e trajnimit mbi këto të dhëna mund të përshtasim

nga një model peme regresioni, dhe rezultati i secilës pemë mesatarizohet duke dhënë një model përfundimtar që njihet si pylli i rastësishëm. Modeli i gjeneruar duke përshtatur një pemë për secilin set të ri nga të dhënat do të jetë:

$$\hat{f}_b(x) = \frac{1}{B} \sum_{i=1}^B \hat{f}_i(x). \quad (2.39)$$

ku $\hat{f}_i(x)$ është modeli i gjeneruar nga pema e trajnuar në kampionin i , dhe gjithë modelet mesatarizohen. Kjo e treguar në fig. 2.11.

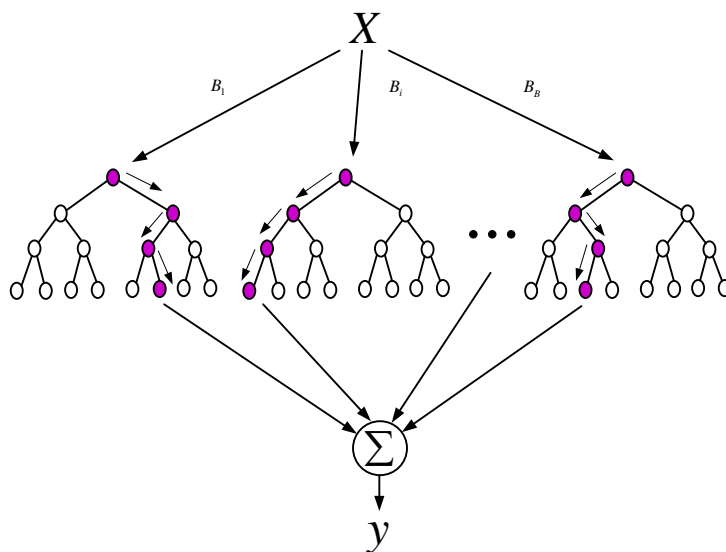


Figura 2.11: Procesi i ndërtimit të një pylli të rastësishëm

Ndryshe nga pemët e regresionit pemët e përdorura nga pyjet e rastësishëm nuk mund të 'krasiten', por këtu mund të kontrollojmë thellësinë e rritjes së një peme elementare. Procesi i trajnimit të një pylli të rastësishëm mund të përmbliidhej:

1. Përcaktojmë modelin e pyllit që do të përdorim, më tej vijojmë me përcaktimin e numrit të pemëve që do të përdoren (përdorim metodë optimizimi)
2. Për numrin e pemëve të përdorura gjenerohet një set të dhënash nga të dhënat origjinale
3. Trajnojmë modelin e pemës tek pjesa e të dhënave
4. Modeli i pyllit jepet duke mesatarizuar rezultatet e gjithë pemëve të përdorura

Kapitulli 3

ARSYETIMI I BAZUAR TEK RASTET: VËSHTRIM TEORIK

Hyrje

Në parim ABR përdoret për të zgjidhur problemet e reja bazuar në adaptimin e zgjidhjeve ekzistuese për probleme të ngjashme. Kjo është një metodë e inteligjencës artificiale e cila është lehtësisht e kuptueshme nga lexuesi pasi çdo hallkë dhe çdo veprim që ndodh në zinxhirin e zgjidhjes është lehtësisht e kuptueshme jo si në metodat e bazuara tek modeli, ku përmendim gjithë metodat e përshkruara në Kapitullin 2. Pra të gjithë hapat e një modeli ABR mund të rimodelohen duke u përshtatur tek një problem specifik.

Kjo është një ndër metodat kyçe e përdorur në këtë tezë për parashikimin e konsumit të energjisë për kategori të ndryshme godinash. Duke u bazuar në teknikat e ABR dhe duke i konsideruar të dhënat e matura për një periudhë të caktuar si “*seri kohore*” përshkruajmë hapat e implementimit të një sistemi të tillë. Nacionet e përdorura për të shpjeguar metodën janë përshtatur nga punimet [135],[136] dhe [151]. Më poshtë po tregojmë parimisht hapat kryesore në implementimin e një sistemi ABR.

3.1 Njohuria

Njohuria kuptohet si një nocion informal ku përshkruan një detyrë ose një problem të caktuar i cili mund të zgjidhet nga një qenie, sistem ose një makinë. Për të përdorur njohurinë operatori duhet të këtë akses tek ajo dhe të dijë si një njohuri mund të përdoret për të zgjidhur një problem të caktuar. Në problemet e serive të kohës njohuria shprehet në formën e strukturave të të dhënave si dhe disa informacione shtesë të cilat lejojnë modifikime të mëvonshme të strukturës. Në përgjithësi një strukturë të dhënash në ABR paraqitet në formën e *attributeve*.

Ku secili atribut jepet në këtë formë:

- i. Emërtimi V
- ii. Zakonisht zona e përshkrimit të variablit V që është e fundme, dhe
- iii. vlera e variablit x_v i cili jep vlerën e atributit në çdo hap kohe.

Për një set fundor atributesh V_i , ku $1 \leq i \leq n$, një *vektor atributesh* është një nënset (v_1, \dots, v_n) i tillë që secila v_i ndodhet në zonën e përcaktimit.

3.2 Përcaktimet bazë

Si metodë e inteligjencës artificiale një përcaktim sa më i shkurtër dhe i saktë për ABR do të ishte si më poshtë:

Gjetja e zgjidhjes për një problem të caktuar, kryhet duke kërkuar në hapësirën e eksperiencës (zgjidhjeve ekzistuese), duke i marrë ato si zgjidhje dhe ia përshtasim problemit tonë nëpërmjet adaptimit

Një nga veçoritë e sistemeve të bazuara mbi njohuritë ose ndryshe ato të inteligjencës artificiale është që ato kërkojnë medoemos ekzistencën e të dhënave të mëparshme, pra përdorimi i ekperiencave të shkuara, ku pa to implementimi i një metodologjie të tillë është i pamundur. Kjo metodë e ka thelbin e saj në paraqitjen e problemit me anë të “*rasteve*” ku secili *rast* mbart informacion në trajtën (\mathbf{X}, y) atributet dhe zgjidhjen. Ku kategoria e problemit do të përmbliidhej si:

(problem, zgjidhje)

Metoda modelohet duke ndërtuar rastet, të cilët ndahen në dy tipe, si më poshtë:

- I. *Rastet e shkuara*, ku rasti përshkruhet nga karakteristikat (*atributet*) që janë të njëjta për të dy rastet, dhe zgjidhja e dhënë për këtë rast në të shkuarën,
- II. *Rasti i ri*, ku përshkruhet nga karakteristikat që janë të njëjta, por nuk përmbajnë të njëjtat *vlera variablash* (p.sh. nëse do kishim temperaturën si variabël në të dy rastet *karakteristika* do të ishte temperatura, por vlerat janë ato të rastit të ri), si dhe zgjidhja e kërkuar që është ajo që modeli do të parashikojë.

Koncepti tjetër i rëndësishëm në këtë proces është *ngjashmëria*, hapi i dytë më i rëndësishëm pasi modelojmë problemin me anë të rasteve. Vlerësimi i *ngjashmërisë* kalon në dy etapa, *etapa e parë* përfshin ‘rastet e ngjashme’, çdo *rast* që ndërtoam mund të grupohet në *nëngrupe* më të vegjël me veçori të njëjta midis tyre (ku veçoritë përcaktohen me ekuacione ose rregulla të thjeshta), dhe *etapa e dytë* ku përdoren metoda specifike për të gjetur *ngjashmërinë* midis një *rasti të ri* dhe një *të vjetri*, ku më poshtë do të analizojmë mënyrën e implementimit.

Hapi i fundit është *adaptimi* i zgjidhjes e cila ekziston në *rastin e vjetër* tek *rasti i ri*, ku kërkohet. Adaptimi përdor teknika të ndryshme matematikore. Në mënyrë të përgjithshme një sistem i ABR për zgjidhjen e problemeve të natyrave të ndryshme jepet në fig. 3.1:

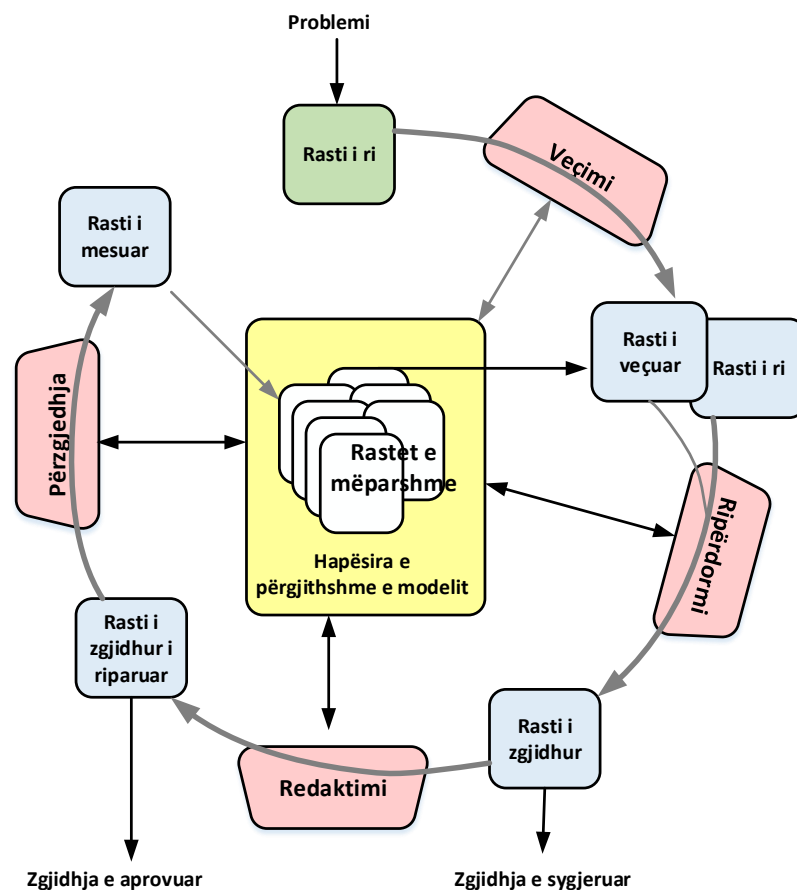


Figura 3.1: Një cikël i një modeli ABR, Aamodt dhe Plaza [136].

Një cikël i përgjithshëm implementimi do të përshkruhej në këto etapa:

1. *Veçimi* i rasteve të mëparshme të cilët i bashkon një ose disa veçori të njëjta

(p.sh. këto kanë të ngjashme ose janë të përafërta në veçori brenda një kushti të caktuar të vendosur nga projektuesi i modelit), ku mund të grupohen si *raste* me sjellje të ngjashme,

2. *Ripërdorimi* ka të bëjë me marrjen në përdorim të zgjidhjeve të cilat kanë ekzistuar në rastet e veçuara,
3. *Redaktimi* ose adaptimi i rasteve të ngjashme (ku zgjidhja e një rasti të ngjashëm do të rimodelohet sipas rastit të ri në bazë të teknikave matematikore) dhe e fundit,
4. *Përzgjedhja* e rastit ose e rasteve më të ngjashme për të dhënë zgjidhjen më të saktë.

Në shumë aplikacione praktike disa nga hapat mund të përfshihen brenda një hapi të vetëm, pasi është e vështirë të bëhet ndarja e qartë midis redaktimit dhe ripërdorimit ku ata mund të njihen si një hap i vetëm dhe mund të quhet si *adaptim*.

3.3 Përcaktimet e ngjashmërisë

Në analizën normale të *bazës së rasteve*, pyetja që shtrohet është e tillë nëse një rast i dhënë (*rast i ri*) përputhet më një rast i cili ndodhet i ruajtur në *bazën e rasteve*. Ky është formulim ekzakt dhe tregon “*përpuethshmërinë*”, ndërsa në rastet kur një rast i përafrohet një rasti tjetër, por jo përpuethshmërisht atëherë kemi konceptin e “*ngjashmërisë*”. Në problemet e *ngjashmërisë* në seritë kohore kemi: *ngjashmëri sekuenciale* ku e gjithë sekuenca duhet të plotësojë kushtet e ngjashmërisë, si dhe *ngjashmëri të pjesshme sekuenciale* ku analizohen nënsekuenca veç e veç nga e tërë sekuenca.

Zakonisht, llogaritja e ngjashmërisë bëhet duke llogaritur distancën midis sekuecave të serive kohore, dhe nga distanca kalojmë në ngjashmëri. Një shembull i thjeshtë i llogaritjes së ngjashmërisë nga distanca mund të ishte:

$$\text{sim}(x, y) = \frac{1}{1 + D(x, y)} \quad (3.1)$$

Por në metoda të ndryshme llogaritja e ngjashmërisë mund të mjaftohet vetëm me llogaritjen e distancës minimale si në rastin e *fqinjtit më të afërt* ku zgjedhja e k elementëve më të afërt bëhet duke marrë ata me distancën minimale dhe jo duke kaluar problemin në *ngjashmërinë* maksimale, por në shumë fusha përdorimi nocioni i *ngjashmërisë* është ai më i përdorshmi.

Përcaktimi i distancës midis dy objekteve ose sekuecave kohore $D(X, Y)$ duhet të plotësojë disa veçori:

1. **Simetrinë**, $D(x, y) = D(y, x)$,
2. **Kushti i vëtë-ngjashmërisë**, $D(x, x) = 0$,
3. **Pozitiviteti**, $D(x, y) = 0$ nëse $x = y$ si dhe,
4. **Mosbarazimi i trekëndëshave**, $D(x, y) \leq D(x, z) + D(y, z)$

Disa nga metodat për matjen e distancës midis serive kohore po i përmendim më poshtë:

3.3.1 Distanca e Minkowskit

Distanca e Minkowskit është një nga metodat më të vjetra dhe më efektive në analizën e serive të kohës. Për dy sekuenca kohore më të njëjtën gjatësi $X = (x_1, x_2, \dots, x_n)$ dhe $Y = (y_1, y_2, \dots, y_n)$ distanca e Minkowskit do të jetë:

$$D(X, Y) = \sqrt[p]{\sum_{i=1}^n (x_i - y_i)^p} \quad (3.2)$$

Pra duke i marrë të dy sekuenat kohore të barabarta atëherë duke e paraqitur çdo pikë të një sekuence në hapësirën Euklidiane atëherë distanca do të jepen nga $D(x, y) = (|x_1 - y_1|^p + |x_2 - y_2|^p + \dots + |x_n - y_n|^p)^{1/p}$. Për vlera të ndryshme të p do të përftonim lloje të ndryshme distancash.

(a) $p = 1$ **Manhattan**

(b) $p = 2$ **Euclidean**

(c) $p = \infty$ **Max**

Në fig. 3.2 po tregojmë qartësisht një interpretim të distancës sipas Minkowskit. Ku distanca midis dy serive kohore do të llogaritet si shumë e distancave të çdo hapi korrespondues.

Ku në figurë duket qartë që në çdo hap distanca është ndërmjet pikave respektive të dy sekuencave të kohës. Gjerësisht e përdorur në literaturë për të përcaktuar ngjashmërinë në seritë e kohës është metoda e distancës Euklidiane, ku në disa raste kur llogarisim vetëm distancën midis dy pikave ajo është sa vlera absolute e distancës $D(x, y) = |x_i - y_i|$, ndërsa në rastin e përgjithshëm është si më poshtë:

$$D(X, Y) = \sqrt{2 \sum_{i=1}^n (x_i - y_i)^2} \quad (3.3)$$

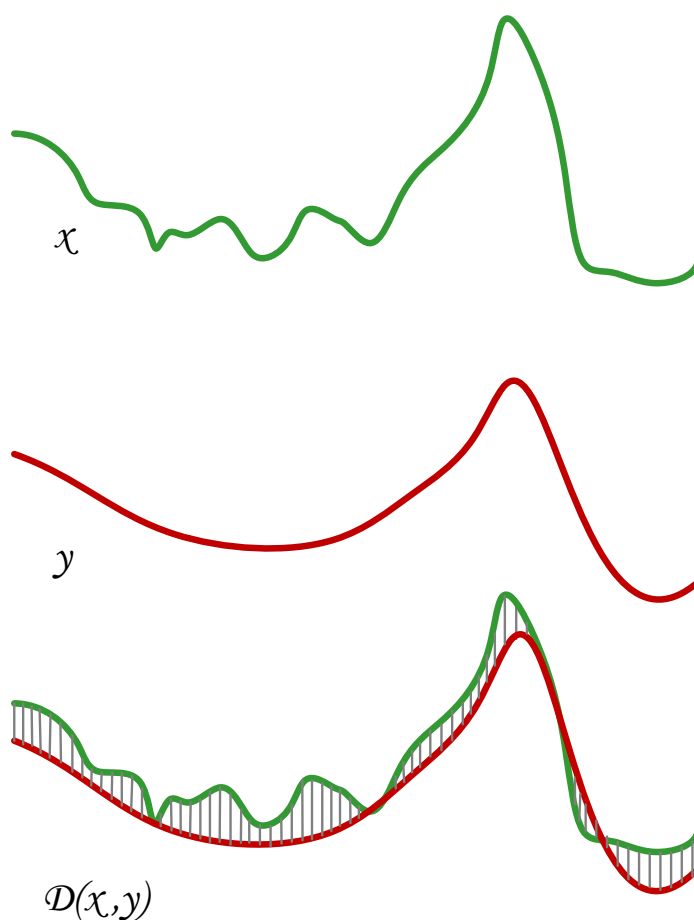


Figura 3.2: Llogaritja e distancës së Minkowskit.

Përderisa në bazën e të dhënave kemi më shumë se një *atribut* që përshkruan të dhënat atëherë është parë se kontributi që ka secila në madhësinë e daljes (zgjidhjen) nuk është i njëjtë, kështu është e nevojshme që secilit *atribut* ti bashkangjisim një “peshë” që tregon rëndësinë e saj. Nga kjo lindi dhe distanca *Euklidiane e peshuar* që është një modifikim i lehtë i asaj normale ku për çdo *atribut* kemi dhe peshën që i atribuohet relativisht problemit.

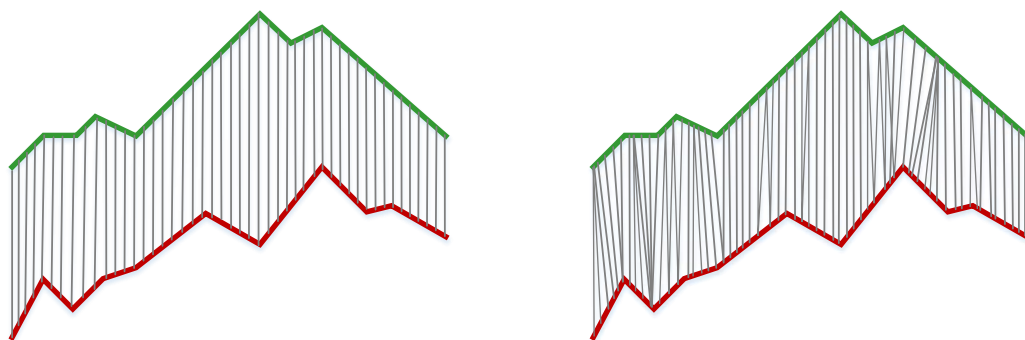
$$D_w(X, Y) = \sqrt[2]{\sum_{i=1}^n w_i (x_i - y_i)^2} \quad (3.4)$$

Ku $X = (x_1, x_2, \dots, x_n)$, $Y = (y_1, y_2, \dots, y_n)$ janë vektorët e secilës sekuençe kohore si dhe vektori i peshave $W = (w_1, w_1, \dots, w_1)$. Por problemi

më i madh në këtë pikë është përcaktimi sa me eficient i rëndësisë së attributeve “*peshave*”, ku në përcaktimin e tyre përdoren metodat statistikore dhe njohuritë e ekspertit. Në pjesën që vijon do të bëjmë një paraqitje më të detajuar të mënyrës si përcatohet rëndësia e attributeve dhe metoda që janë zhvilluar në literaturë. Megjithëse është një nga metodat që gjen zbatueshmërinë më të madhe në llogaritjen e ngjashmërisë bazuar në distancën e serive të kohës, ajo ka disa pika të dobëta sepse nuk mund të zgjidhë problemet ku seritë e kohës kanë gjatësi të ndryshme si dhe të përcaktojë diferencën mes serive që kanë largësi të njëjtë, por trend të ndryshëm. Një nga metodat e cila bën të mundur analizën e problemeve ku kemi sekueca kohe me gjatësi të ndryshme është dhe “*Spërdredhja dinamike në kohë*” që po e analizojmë më poshtë.

3.3.2 Spërdredhja dinamike në kohë

Siç përmendëm më sipër Distanca Euklidiane e ka të vështirë të merret me problem të natyrës së serive kohore me madhësi të ndryshme, ose të kërkojë në pika të ndryshme nga ajo korrespondente e njëjës seri tek tjetra. Në fig. 3.3 po paraqesim grafikisht se çfarë duam të tregojmë. Kjo metodë është më e softistikuar krahasuar me distancën Euklidiane pasi nuk përdor relacione matematike për llogaritje, por përdor programimin dinamik duke vendosur kushte për kufijtë e poshtëm e të sipërm po ashtu dhe kushte të tjera referuar problemit. Metoda është mjaft e njohur në problemet e njohjes së zërit pasi tenton të gjejë distancën minimale midis dy sekuecave. Në këtë fushë kanë punuar veçanërisht autorët në [39] që janë mjaft të njohur në aplikacionet e ditëve të sotme.



(a) Krahasimi i sekuecave “një për një”

(b) Krahasimi i sekuecave “një kundër disa”

Figura 3.3: Matja e distancës fikse Euclidean (3.3a) si dhe matja në formë të spërdredhur (3.3b).

3.3.3 Ngjashmëria sipas formës

Kur flasim për ngjashmërinë sipas formës duhet të kuptojmë se analizohen sekuenat e kohës në raport me tendencën e tyre, pra me formën që kanë kundrejt njëra-tjetrës. Një nga metodat e njohura është edhe *ngjashmeria e kosinuset*¹ ku llogaritet këndi midis vektorëve që krijohen në çdo hap.

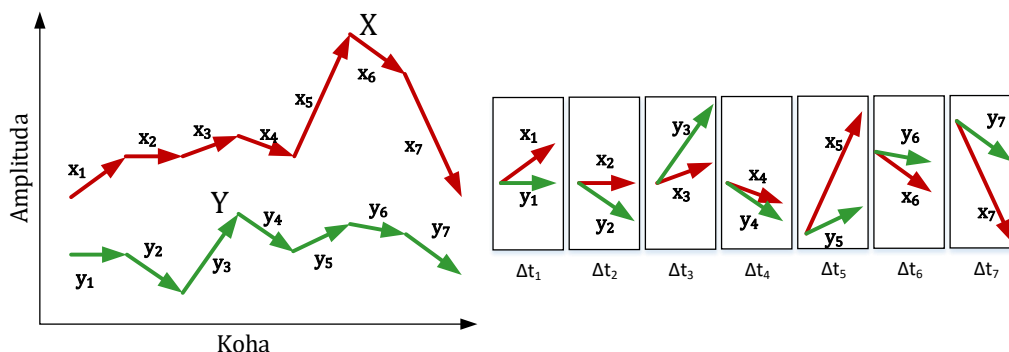


Figura 3.4: Analiza sipas formës e dy serive kohore X dhe Y.

Nëse në hapësirën 1-dimensionale kemi një seri kohore $X = ((a_1, t_1), (a_2, t_2), \dots, (a_n, t_n))$ ku a_n është vlera në një moment të caktuar kohe t_n . Nëse e shprehim më thjeshtë sekuençën kohore do të merrnim një prezantim më të thjeshtë të sekuençës kohore, $X = (x_1, x_2, \dots, x_n)$ ku x shpreh një çift (*madhësi kohë*) ndërsa nëse duam ti shprehim sekuenat si vektor atëherë do mbështetemi në diferencat e rendit të parë pra në diferencën e amplitudave midis një hapi kohe tek hapi tjetër që do të ishte:

$$\begin{aligned} \mathbf{X} &= ((x_2 - x_1), (x_3 - x_2), \dots, (x_n - x_{n-1})) \\ &= (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}) \end{aligned} \quad (3.5)$$

Një mënyrë paraqitje së një sekuence kohe të transformuar tregohet në fig. 3.4. ku në të theksuar jepen vektorët të cilët llogariten sipas diferencave në çdo hap sipas (3.5). E njëjta procedurë ndiqet edhe për sekuençën kohore Y siç tregohet më poshtë:

$$\begin{aligned} \mathbf{Y} &= ((y_2 - y_1), (y_3 - y_2), \dots, (y_n - y_{n-1})) \\ &= (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-1}) \end{aligned} \quad (3.6)$$

Ngjashmëria sipas formës për dy sekuenca kohe llogaritet si këndi ndërmjet dy vektorëve të nënsekuençave përkatëse, ndryshe kjo njihet dhe si “*ngjashmëria*

¹ Cosine similarity

e kosinuset”, dhe më poshtë do japim metodologjinë e llogaritjes. Theksojmë se për një sekuençë me gjatësi n a do të krijohen $n - 1$ vektorë të nënsekuencave.

$$sim(x_i, y_i) = \begin{cases} -1 & \text{nëse } \theta = \pi \\ \cos\theta = \frac{\mathbf{x}_i \cdot \mathbf{y}_i}{|\mathbf{x}_i| |\mathbf{y}_i|} & \text{në çdo rast tjetër} \\ -1 & \text{nëse } \theta = \pi \end{cases} \quad (3.7)$$

Ekuacioni (3.7) shpreh mënyrën e llogaritjes së ngjashmërisë, ku θ është këndi ndërmjet dy vektorëve sekuencialë. Duke qënë se kosinusi i këndit ndërmjet dy vektorëve merr vlera brenda segmentit $[-1, 1]$, në praktikë bëhet dhe transformimi i këtij relacioni ngjashmërie në $[0, 1]$ ku qëndron edhe interesi në shumicën e aplikimeve.

$$sim'(x_i, y_i) = 1 - \frac{\cos^{-1}(sim(x_i, y_i))}{\pi} = \begin{cases} 0 & \text{nëse } \theta > \pi/2 \\ \cos\theta = \frac{\mathbf{x}_i \cdot \mathbf{y}_i}{|\mathbf{x}_i| |\mathbf{y}_i|} & \text{në çdo rast tjetër} \end{cases} \quad (3.8)$$

Ndërsa ekuacioni (3.8) shpreh vlerësimin e ngjashmërisë për dy vektorë,

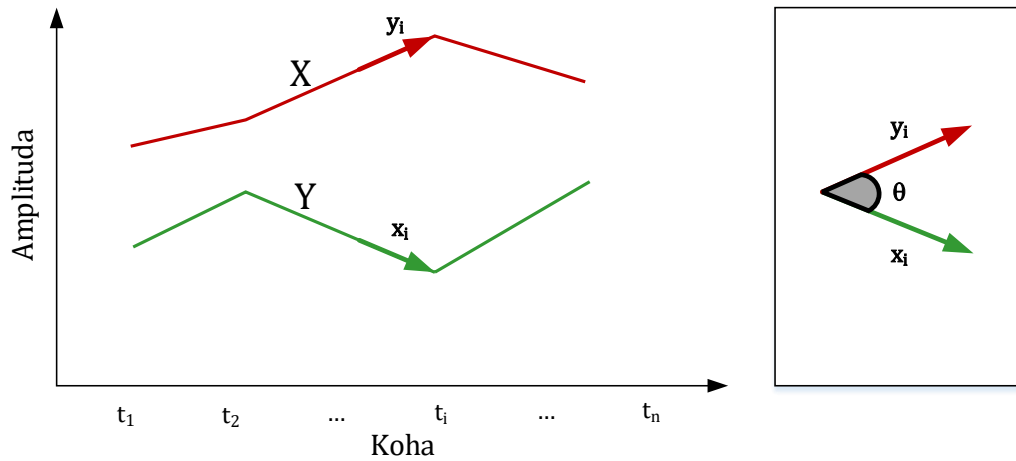


Figura 3.5: Përcaktimi i ngjashmërisë sipas kosinuset midis dy vektorëve të nënsekuencave.

atëherë lind nevoja e llogaritjes së ngjashmërisë për sekuençën e plotë e cila është shuma e mesatarizuar e ngjashmërive individuale (3.9):

$$VS = \frac{\sum_{i=1}^m sim'(x_i, y_i)}{m} \quad (3.9)$$

ku, m është numri i segmenteve në shqyrtim pra është 1 më pak se numri i hapave kohorë të një sekuençe. Metoda është shumë e përdorshme në klasifikimin e teksteve, por gjen përdorim të gjerë edhe në analizën e problemeve të shprehura me *seritë e kohës*.

3.4 Hapat e implementimit të një modeli ABR

Për implementimin e një modeli të ABR mbështetemi në sa thamë më sipër, duke i përfshirë të gjitha metodat në një sistem të vetëm. ABR si një metodë e inteligjencës artificiale e jep zgjidhjen bazuar në zgjidhjet ekzistente për probleme të ngjashme. Parimisht procesi kalon në këto etapa, *paraqitjen e problemit në formën e rasteve, grumbullimi në grupe e rasteve të ngjashme me rastin tonë, si dhe adaptimi i zgjidhjeve ekzistente tek problemi në fjalë*. Më poshtë po paraqesim një skemë të asaj që përmendëm dhe përshkruajmë në mënyrë të detajuar të gjithë hapat:

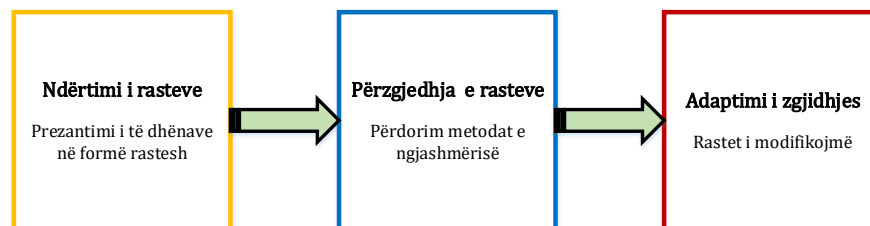


Figura 3.6: Skema parimore e një modeli ABR.

3.4.1 Ndërtimi i rasteve

Ndërtimi i rasteve lidhet me mënyrën sesi të dhënat që disponohen të paraqiten në mënyrë më efektive të mundshme, e tillë që të gjendet zgjidhja më e mirë duke përfshirë karakteristikat më domethënëse që disponojmë. Në literaturë gjenden disa nga metodat më të përdorura për të ndërtuar rastet, ku vlen të theksohet se janë të përdorshme për natyrën e problemeve që zgjidhin. Theksojmë se, ruajtja e rastit në memorje njih dy metoda kryesore:

Modeli i memories dinamike, sipas këtij modeli rastet do të paraqiten në formën e rasteve specifike ku secili nga rastet do të grupohet nën struktura me veçori të ngjashme. Kjo është një nga metodat e para të përdorura në literaturë e cila i ka fillesat e para nga modeli i Schank [134], ndërtimi i rasteve bëhet në mënyrë “episodike”, rastet me ngjashmëri të caktuara brenda një bazë rastesh të plotë do të veçohen në një strukturë më të përgjithshme ose (strukturë episodike).

Modeli kategorik dhe ekzemplar, ky model mbështetet në ruajtjen e rasteve në memorje mbi bazën e mënyrës natyrale të përshkrimit të problemit. Pjesa ku ruhen rastet ose memorja bën të tillë që rastet të formojnë një strukturë e cila lidh mes tyre *kategoritë, rastet si dhe pointerat e indeksimit të rasteve*. Ndërmjet këtyre rasteve ekziston një ndërlidhje e cila bën të mundur që të krijohen nënstruktura të cilat çojnë në evidentimin e rasteve të ngjashme kundrejt një rasti të caktuar në mënyrën sa me efektive.

Keto janë dy metodat më të përdorshme dhe më të hershme të ruajtjes së rasteve në memorje për përdorim të mëvonshëm, por e rëndësishme është të përmendet se çfarë duhet të kemi parasysh në momentin e ndërtimit të rasteve për të prezantuar problemin, dhe kemi:

1. Qëllimi i ndërtimit të një sistemi ABR. Për shembull vlen të përmendet struktura e përdorur, më e favorshme do të ishte një strukturë hierarkike e kërkimit të zgjidhjes.
2. Numri i rasteve të krijuara të ruajtura për të prezantuar një problem, sa më shumë të rritet numri i rasteve aq më kompleks dhe më i kushtueshm në terma kohorë bëhet problemi, kështu që mënyra të ndërmjetme dhe efektive do të ishin të rëndësishme të implementohen që në mënyrën e ruajtjes.
3. Numri i karakteristikave të cilat janë përdorur për të paraqitur një rast për një problem të caktuar.
4. Njohuritë e projektuesit të një sistemi ABR dhe sistemeve të inteligjencës artificiale në përgjithësi mbi fushën specifike ku ky problem kërkon zgjidhje.

Në përfundim ndërtimi i rasteve mund të përmblihet me dy etapa, e para *specifikimi i problemit* dhe e dyta *zgjidhja* e tij. E para ka të bëjë me përshkrimin e problemit më anë të karakteristikave që ne disponojmë tek të dhënat për zgjidhjen e kërkuar. E dyta ka të bëjë me parashtrimin në mënyrë të saktë të problemit në mënyrë që të arrijmë zgjidhje të dëshiruara.

3.4.2 Përzgjedhja e rasteve

Procesi i përzgjedhjes së rasteve përfshin me zgjedhjen e rasteve më të ngjashme me rastet që mbartin problemin që duhet të zgjidhim. Për të arritur një formë sa më efektive të zgjedhjes së rastit apo grup-rasteve më me vlerë për të ofruar një zgjidhje duhet të kemi parasysh vendosjen e kritereve që bëjnë evidentimin e rasteve më të përafërta. Në disa situata përzgjedhja e rasteve bëhet duke kërkuar mbi ngjashmëritë e të gjithë *attributeve*

që janë përdorur për të ndërtuar një rast ose dukë kërkuar pjesërisht tek *atributet*. Metodatat e përzgjedhjes së rasteve nga *baza e rasteve* ku janë ruajtur janë nga më të ndryshmet. Përmendim metodën e *fqinjit më të afërt*, vazhdojmë tek metodatat e agjentëve inteligjentë, të cilat përfshijnë metodatat e përcaktimit të ngjashmërisë midis *tributeve* të çdo rasti. Më poshtë paraqiten metodatat më të përdorshme për përzgjedhjen e rasteve:

1. *Përzgjedhja sipas fqinjit më të afërt*. Sipas kësaj metode rasti ose rastet më të përafërta me rastin që kërkohet zgjidhja, do të jetë ai/ata të cilët plotësojnë kushtin që shumica e ngjashmërive të gjithë *tributeve* që merren në shqyrtim është më e madhe se ajo e rasteve të tjerë të cilët u përdorën në analizë. Në përgjithësi *atributet* që janë më të rëndësishme, pra ato që luajnë një rol më të madh në përcaktimin e zgjidhjes së problemit do t'ju atribuohet një peshë më e madhe rëndësie, ku shumica e të gjithë peshave të *tributeve* do të jetë e barabartë me 1.
2. *Qasjet induktive*. Sipas kësaj qasjeje rastet do të përzgjidhen duke kërkuar sipas strukturave hierarkike, ku rëndësi ka vlera relative e *tributeve* që përshkruajnë një rast. Pra në këtë metodë kërkohet brenda strukturave dhe nënstrukturave të cilat janë përcaktuar që në ndërtimin e rasteve. Kjo çon në reduktimin e hapësirës së kërkimit.
3. *Metodat e udhëzuara nga njohja*. Kjo metodë i referohet më së shumti njohjes së fushës ku veprohet, ku përdorimi i njohurive specifike mbi fushën në të cilën zgjidhet problemi bën që të përzgjidhen *atributet* më domethënëse që mund të japin zgjidhjen më të mirë. Këtu ka rëndësi që të përmendet se ky proces mund të ndërtohet i "shtresëzuar" ku në çdo hap bëhet përzgjedhja e karakteristikave më domethënëse sipas njohurive të fushës.

Për të përcaktuar metodën e përzgjedhjes së rasteve do të mbështetemi në disa faktorë të përgjithshëm:

1. Numri i rasteve që duhet të kërkohet,
2. Sasia e njohjes mbi një fushë të caktuar,
3. Përcaktimi i peshës së secilit *atribut* që ndërtohet një rast,
4. Nëse të gjitha rastet do të përshkruhen nga të njëjtat karakteristika apo do të jetë një mozaik i ndërtimit të rasteve.

Nëse rasti apo rastet e përzgjedhura janë ata më të vlefshmit për zgjidhje ose jo, atëherë është etapa tjetër në të cilën kalojmë. Për këtë arsye bëhet një analizë dhe kalohet në adaptimin e rasteve fig. 3.7.

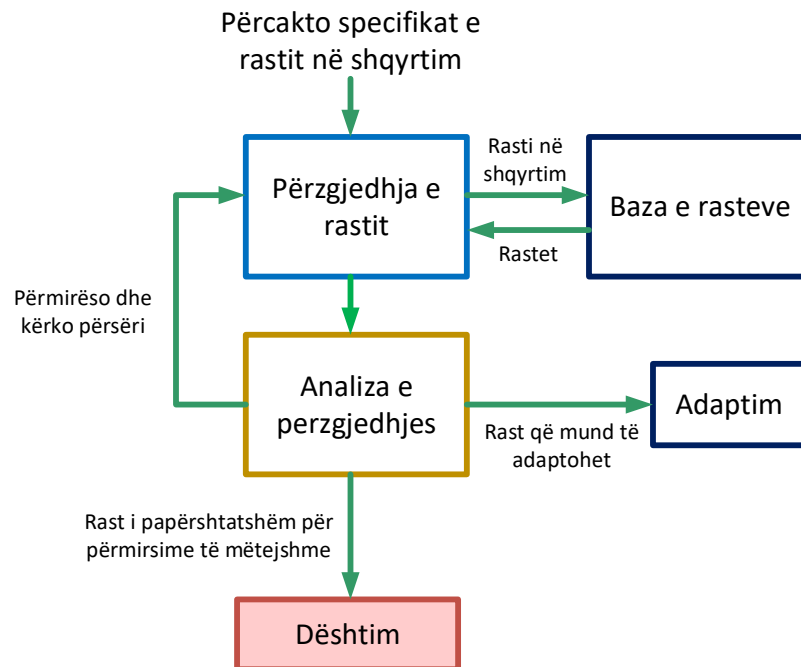


Figura 3.7: Skematika e një procesi CBR [135].

3.4.3 Adaptimi i rastit

Adaptimi i rastit është procesi i transformimit të zgjidhjes së dhënë nga përzgjedhja në një zgjidhje të përshtatshme për rastin që kemi në shqyrtim. Kërkuarit e fushës argumentojnë se kjo është etapa më e rëndësishme e një sistemi ABR pasi kjo e thekson shumë elementin inteligjent në problem dhe jo thjesht të marrë një zgjidhje ‘bruto’ si të përshtatshme.

Ekzistojnë një sërë metodash të cilat mund të përdoren për të adaptuar rastet:

1. Rasti ose rastet e përzgjedhura mund të merren siç janë pa ndonjë modifikim dhe ti përshtaten zgjidhjes, ose mund të pranohen modifikime të vogla në varësi të situatës që ne kërkojmë.
2. Ekziston metoda e transformimit të rasteve ekzistues që japin një zgjidhje të ngjashme me rastin e kërkuar. Përmendim se kjo metodë ka në fokus transformimin e zgjidhjeve që kanë ekzistuar në të kaluarën duke i përshtatur tek zgjidhja e re, ku kjo mund të bëhet duke kërkuar një operator

transformues $\{T\}$. Ky operator bën të mundur transformimin e zgjidhjeve ekzistente tek zgjidhja e re duke përdorur “*njohurinë*” që ka mbi problemin.

- Si dhe metoda e derivuar ku sipas kësaj metode bëhet e mundur që të transformohet dhe të adaptohet metoda e cila është përdorur për zgjidhjen e problemeve të ngjashme.

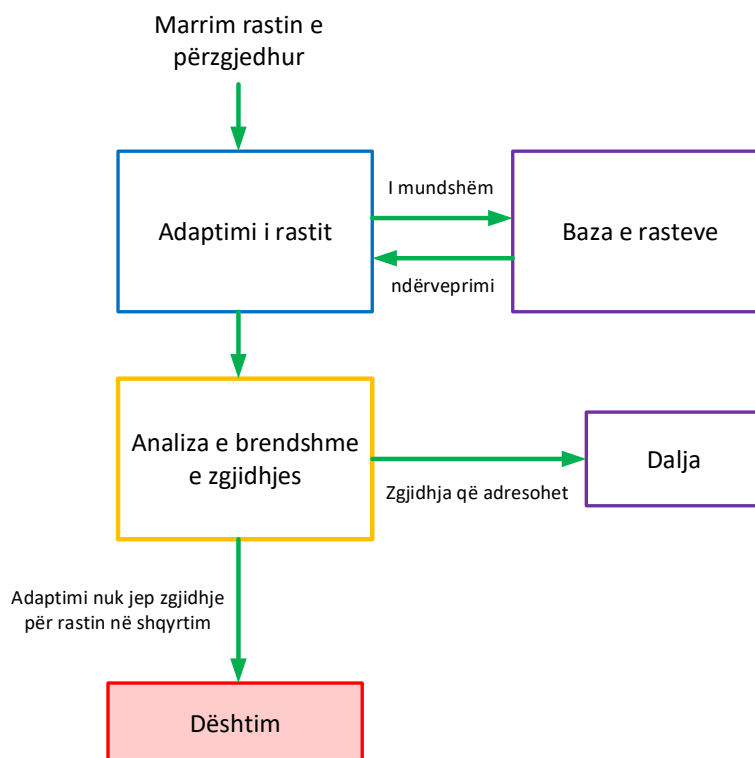


Figura 3.8: Rasti i mësimit për një sistem CBR [135].

Mbas adaptimit të rastit/rasteve është e domosdoshme të analizohet nëse ishte e përshtatshme e gjithë procedura e adaptimit apo duhet përsëritur përsëri ose rezulton negative si zgjidhje. Kjo fazë gjithashtu përkon me përfundimin e implementimit të një sistemi të mundshëm ABR, ku në fig. 3.8 jepet procesi i mësimin me një model ABR.

Kapitulli 4

PËRDORIMI I ABR NË PARASHIKIM

Hyrje

Ky kapitull analizon implementimin e një modeli të inteligjencës artificiale i cili përdoret për parashikimin e konsumit të energjisë në ndërtesa. Përdorimi metodave në analizën e konsumit të ndërtesave fokusohet kryesisht në dy drejtime, i pari ka të bëjë me parashikimin e konsumit ndërkohë i dyti ka të bëjë me evidentimin e defekteve dhe anomalive gjatë procesit të punës së një ndërtese. Të dy qasjet luajnë një rol tepër të rëndësishëm në analizën e konsumit gjatë periudhës së projektimit si dhe gjatë periudhës së funksionimit të ndërtesës.

Në këtë pjesë fokusohemi kryesisht në implementimin e një modeli për parashikimin e konsumit të energjisë elektrike në terma kohorë afatshkurtër për një ndërtesë me funksione publike. Hapsira kohore e parashikimit është për gjashtë orët e ardhshme. Qëndrueshmëria dhe saktësia e modelit të përdorur vlerësohet duke bërë krahasimin me modele të tjerë gjerësisht të përdorur për zgjidhjen problemit në fokus. Eksperimentet e zhvilluara në këtë kapitull i referohen dy skenarëve kryesorë të parashikimit, i pari është parashikimi *offline* ose statik ku për të gjithë rastet e ruajtura në bazën e të dhënave kërkohet zgjidhja, ndërsa i dyti ai *online* apo dinamik ku modeli mëson dhe parashikon sipas mënyrës të shtimit të shembujve në bazën e rasteve në mënyrë inkrementale, duke filluar nga vetëm një rast duke rritur bazën e rasteve me kalimin e kohës.

Vlerat e parashikuara nga modeli *statik* krahasohen me vlerat reale (të matura) duke përdorur metodat e vlerësimit të gabimit të shprehura në seksionin 4.1.4, ndër këto metoda zgjidhet teknika CV-RMSE (4.2) si një më të përdorshmet në literaturë. Ndërsa parashikimi sipas modelit *dinamik* përdor vlerësimin e gabimit CV-RMSE grumbullues, e thënë ndryshe vlerat e

parashikuara vlerësohen hap pas hapi duke u shtuar në bazën e parashikimit.

Implementimi i modelit në parashikimin e konsumit të energjisë për 6 orët e ardhshme analizohet hollësisht në seksionet që vijojnë.

4.1 Çështje praktike

Në këtë pjesë analizojmë çështjet praktike të implementimit të modelit ABR në problemin e parashikimit të konsumit të energjisë për një periudhë afatshkurtër. Fillimisht do të analizojmë hapat e funksionimit, më tej të dhënat që disponojmë dhe përpunimi paraprak i tyre dhe më vonë modeli që implementojmë.

4.1.1 Rrjedha e funksionimit

Metodat të cilat mbështeten në teknikën e mësimit të supervizuar, për t'u implementuar eksperimentalisht duhet që të dhënat të cilat disponohen paraprakisht të ndahen në pjesën e cila do të përdoret për trajnim si dhe pjesën për testimin e modelit. Ku pjesa e trajnimit përdoret për të trajnuar modelin me shembuj të ngjashëm nga e kaluara, duke mësuar një model në mënyrë të tillë që ai të jetë i aftë për të zgjidhur probleme të ngjashme në të ardhmen, ndërsa pjesa e testimit përdoret për të testuar modelin në parashikimin e vlerave për të ardhmen si dhe të vlerësojë performancën e tij. Vlersimi i performancës kryhet duke krahasuar vlerat reale nga të dhënat me vlerat e parashikuara të modelit për shembuj të njëjtë, metodat e vlerësimit përshkruhen me detaje në seksionin 4.1.4.

4.1.2 Mjedisi eksperimental

Pjesa hardware ku kryhen eksperimentet është një workstation me CPU 2*2.4 GHz, 1333 MHz FSB dhe 6 Gb memorje të instaluar me sistem operimi Windows 10/64-bit. Të gjitha eksperimentet janë implementuar në MATLAB[®] R2015a. Kodi i ABR Shtesa A, është implementuar i gjithë nga fillimi dhe nuk i referohet ndonjë librerie të caktuar apo ndonjë Toolbox që përdoret në këtë fushë, ndërsa modeli i rrjetës neurale që përdoret si krahasim vjen nga Neural Network Toolbox, MATLAB. Modeli është lehtësisht i riprodhueshëm në gjuhë të tjera programimi.

4.1.3 Të dhënat dhe përpunimi i tyre

Implementimi i metodave të inteligjencës artificiale kërkon medoemos ekzistencën e të dhënave paraprake të grumbulluara për një periudhë të caktuar. Ekzistenca e të dhënave është e domosdoshme veçanërisht për metodat e supervizuara pasi metoda e cila implementohet kërkon që mbi këto të dhëna

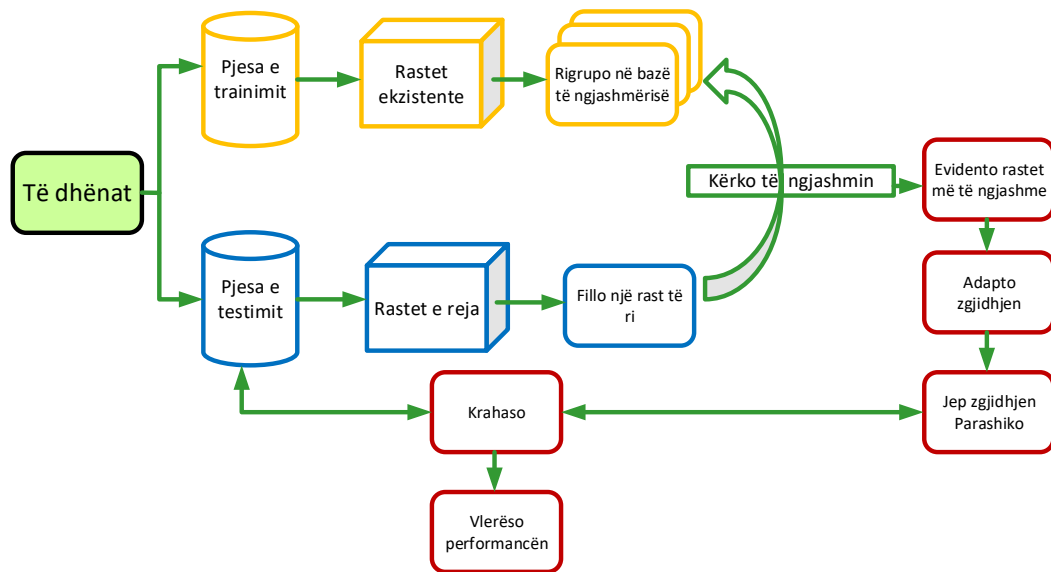


Figura 4.1: Bllokdiagrama e një modeli ABR për parashikimin e konsumit të energjisë.

ekzistuese të përshtasin një model dhe më pas këtë ta përdorin për parashikimin e vlerave që na duhen. Të dhënat e përdorura për të marrë këto rezultate janë grumbulluar nga një ndërtesë institucionale në Calgary (Alberta, Canada) për ditët e punës nga 1 Janar 2013 deri më 9 Maj 2014 (me disa të dhëna që mungojnë nga 29 Mars 2013 deri më 1 maj 2013), por mungesa e datave nuk përbën ndonjë problem për implementimin e këtij tipi modeli. Ndërtesa ka një sipërfaqe prej $16800 m^2$ dhe funksionon duke ofruar kryesisht hapësira për zyra dhe veprimtari tregtare. Pjesa më e madhe e konsumit shkon për paisjet e ngrohje ftohjes ku sistemi HVAC përbëhet nga 1 çiller dhe nga 3 bolier për ngrohjen, si dhe 5 njësi të ajrimit. Të dhënat e siguruara janë matje të konsumit total dhe për paisjet e sipërpërmendura në intervale 1 orë. Përveç madhësive të matura referuar konsumit të dhënat përfshijnë dhe dy parametra shumë të rëndësishëm për nga roli i tyre në konsumin total të cilët janë parametrat ambientale të temperaturës dhe lagështisë. Përderisa interesohemi për parashikimin e konsumit të energjisë atëherë parametrat ambientale janë ata të parashikuar për periudhën e matjes dhe që na jepen nga shërbimi meteorologjik Kanadez. Referuar statistikave mbi konsumin e energjisë në godinën në shqyrtim vërehet se konsumi pjesën më të madhe të tij e zë gjatë orëve të punës ku dhe matjet janë përqëndruar, matjet fillojnë në 7 në mëngjes dhe përfundojnë në 5 mbasditë.

Të dhënat përfshijnë vlerat e 10 variablave (*attributeve*) të matura për

periudhën që përmendëm si dhe vlerat e konsumit të energjisë në intervale prej 1 ore. Në tabelën 4.1 tregohen variablat e matura së bashku me njësitë matëse të tyre.

Tabela 4.1: Përshkrimi i variablave dhe njësitë matëse

Variabli	Njësia
V_1	Temperatura e parashikuar e ajrit jashtë (°C)
V_2	Lagështia e parashikuar e ajrit jashtë (%)
V_3	Temperatura e ajrit të ngrohtë në njësinë 2 (°C)
V_4	Temperatura e ajrit të ngrohtë në njësinë 3 (°C)
V_5	Temperatura e ajrit të ftohtë në anën perëndimore (°C)
V_6	Temperatura e ajrit të ftohtë në njësinë 4 (°C)
V_7	Temperatura në dalje të çillerit (°C)
V_8	Rrjedhja e ujit në dalje të çillerit (l/s)
V_9	Temperatura në dalje të bolierit (°C)
V_{10}	Rrjedhja e ujit në dalje të bolierit (l/s)

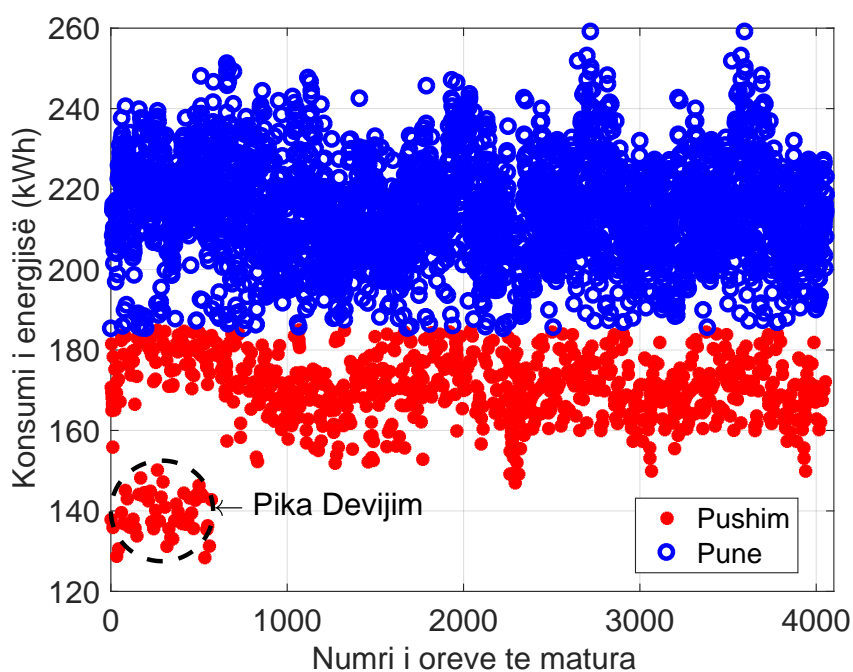


Figura 4.2: Të dhënat e konsumit të energjisë për të gjithë periudhën e matur.

Në fig. 4.2 tregohet konsumi i energjisë për gjithë periudhën e matur. Nga

figura duket lehtësisht se ekzistojnë pika që devijojnë nga lakorja normale e të dhënave, prania e të cilave është kontribut i një morie faktorësh të njohur apo të panjohur. Në literaturë [138, 139, 140] këshillohet që të bëhet një vizualizim dhe një filtrim paraprak i të dhënave. Pjesa e të dhënave e qarkuar me ngjyrë të zezë tregon se aty ka disa devijim në të dhëna dhe një filtrim shumë i thjeshtë duke zhvendosur këto të dhëna do ta ndihmojë modelin të rrisë performancën. Nga të dhënat vërehet ekzistenca e një ndarjeje midis pjesës kur godina është e populluar (regjim pune) dhe në momentin ku është jashtë funksioni kur vetëm paisjet bazë funksionojnë.

4.1.4 Vlerësimi i performancës

Metodat e vlerësimit të performancës së një modeli i cili bën parashikimin e konsumit të energjisë në ndërtesa për periudhën afatshkurtër përcaktohen nga udhëzimi i publikuar nga ASHRAE (ASHRAE guideline 14-2002) [30]. Sipas këtij udhëzimi (një nga më të përdorshmit nga kërkuesit e kësaj fushe), përcaktohen metodat që përdoren për vlerësimin e performancës së modeleve. Në udhëzim paraqiten këto metoda kryesore vlerësimi të cilat jepen më poshtë:

1. Koeficienti i variacionit të deviacionit standart (CVSTD).

$$CVSTD = \sum_{i=1}^{|N|} \sqrt{\frac{(P_i - \hat{P}_i)^2}{|N| - 1}} \bar{P} \times 100 \quad (4.1)$$

2. Koeficienti i variacionit i gabimit mesatar të rrënjës katrore (CVRMSE).

$$CV - RMSE = \frac{\sqrt{\frac{1}{|N|-p} \sum_{i=1}^{|N|} (P_i - \hat{P}_i)^2}}{\bar{P}} \times 100 \quad (4.2)$$

3. Koeficienti i vlerësimit mesatar të gabimit absolut në përqindje (MAPE).

$$MAPE = \frac{1}{|T|} \sum_{t=1}^{|T|} \left| \frac{P_t - \hat{P}_t}{P_t} \right| \cdot 100 \quad (4.3)$$

4. Gabimi absolut mesatar MAE jepet në (kW).

$$MAE = \frac{1}{|T|} \sum_{t=1}^{|T|} |P_t - \hat{P}_t| \quad (4.4)$$

5. Gabimi i normalizuar i shmangies (NMBE).

$$NMBE = \frac{1}{|N| - p} \frac{\sum_{i=1}^{|N|} (P - \hat{P}_i)}{\bar{P}} \times 100 \quad (4.5)$$

Ku P_i është vlera reale të cilën e kemi të ruajtur në pjesën e të dhënave për testim ndërsa \hat{P}_i është vlera e parashikuar sipas modelit, N është numri i parashikimeve që bëjmë, \bar{P} është vlera mesatare e matjeve reale në zonën e parashikimit si dhe p është koeficient që në rastin tonë e marrim 1. Referuar mënyrës të grumbullimit të të dhënave ekzistojnë dhe kufijtë brenda së cilës është i pranueshëm gabimi. Për matjet me përsëritje në intervalin prej 1 ore dhe janë të kalibruara sipas këtij intervali, këshillohet që CV-RMSE të ndodhet brenda kufijve të 30% ndërsa NMBE brenda kufirit 10%. Implementimi i problemit në fokus përdor vetëm CVRMSE si metodë të matjes së gabimit pasi në pjesën dërrmuese të rasteve në punimet që gjejmë kjo është metoda e përdorur dhe kjo na bën të mundur krahasimin me to.

4.2 Implementimi i modelit

Implementimi i modelit kalon në disa hapa ku gjithësecili ka rëndësinë e tij, në vijim japim një analizë të detajuar. Modeli i implementuar i referohet përshkrimit të bërë në Kapitullin 3, duke përdorur të dhënat e përshkruara në pjesën 4.1.2.

4.2.1 Ndërtimi i rasteve

Hapi i parë në implementimin e një sistemi ABR për të parashikuar konsumin e energjisë është paraqitja e problemit në formën e rasteve. Ndërtimi i rasteve, siç përmendet në Kapitullin 2 mbështetet në modelimin sipas (hyrjeve, daljes), ku hyrja përfshin *atributet* dhe dalja *konsumin e energjisë*. Intervali kohor për të cilin kërkohet të kryhet parashikimi është 6 orë. Fillimisht të dhënat ndahen në pjesën e trajnimit dhe atë të testimit siç tregohet në fig. 4.1. Për secilën pjesë të të dhënave ndërtimi i rasteve përdor të njëjtën qasje, por me ndryshime në përfshirjen e të dhënave siç shpjegohet qartazi vijim. Të dhënat ndahen në mënyrë të tillë: fillimisht matjet që i përkasin vitit 2013 do të përdoren si pjesë për trajnimin e modelit ndërsa pjesa e matjeve për vitin 2014 do të shërbejë si pjesa e testimit dhe vlerësimit të saktësisë së parashikimeve. Përveçse kësaj mënyrë ndarjeje përdoren edhe raporte të tjera ndarjesh të treguara në pjesën eksperimentale.

Rastet e përfshira në modelin e parashikimit ndahen në dy lloje, *rastet ekzistues* ku ekziston zgjidhja të cilën duam ta përshtasim tek *rasti i ri* i cili kërkon parashikim. Një rast ndërtohet sipas kësaj mënyre: treguar në tabelën 4.2

të dhënat horizontale indeksohen me V_1, \dots, V_{10} përshkruajnë variablat e treguara në tabelën 4.1, ndërsa me V' tregohen të njëjtat variabla por për rastin e ri. Ndërsa në kolonën e shënuar me P jepen vlerat në dalje ose aty ku interesohemi për të zgjidhur problemin, pra energjia e konsumuar në çdo hap kohor. Nga tabela numrat në blu tregojnë vlerat e parashikuara, numrat në gri tek rasti burim që përfshihen në modelim, por nuk përdoren në vlerësimin e ngjashmërisë, si dhe numrat në të kuq tregojnë vlerat e energjisë së konsumuar.

Tabela 4.2: Modelet e rasteve, informacioni në rastin e ri dhe rastin ekzistues

Rasti i ri												
	data dhe ora	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	P
t_6	2014-04-07 15:00	10	32.2									?
t_5	2014-04-07 14:00	9	39.9									?
t_4	2014-04-07 13:00	8	32.2									?
t_3	2014-04-07 12:00	9	33.8									?
t_2	2014-04-07 11:00	10	34.6									?
t_1	2014-04-07 10:00	11	29.4									?
t_0	2014-04-07 09:00	12	29.9	29.4	28.4	15.8	24.9	30.5	-.05	67.3	76.2	203.1
t_{-1}	2014-04-07 08:00	12	31.2	29.6	17.4	10.1	21.4	32.7	-.05	65.2	76.2	203.8
t_{-2}	2014-04-07 07:00	11	31.0	28.3	9.7	10.6	22.4	30.8	-.04	66.3	73.2	197.6

Rasti ekzistues												
	data dhe ora	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	P
t_6	2014-03-06 14:00	12	29.9	25.4	24.7	17.9	22.2	31.2	-.04	63.6	77.1	202.9
t_5	2014-03-06 13:00	12	29.9	28.4	27.4	16.7	25.2	31.2	-.04	65.8	78.1	204.6
t_4	2014-03-06 12:00	12	29.9	31.1	29.1	17.4	26.1	31.4	-.05	69.1	77.1	205.1
t_3	2014-03-06 11:00	11	20.5	27.4	27.4	19.8	22.6	29.4	-.04	71.1	76.8	202.1
t_2	2014-03-06 10:00	12	28.4	29.2	28.8	15.9	24.2	30.6	-.05	67.8	76.0	203.8
t_1	2014-03-06 09:00	11	29.3	27.4	25.4	17.8	26.9	31.5	-.05	66.3	77.2	204.6
t_0	2014-03-06 08:00	12	29.4	29.2	28.4	16.8	22.9	31.5	-.05	69.3	77.2	204.2
t_{-1}	2014-03-06 07:00	11	31.2	29.8	17.6	10.1	22.4	32.7	-.05	65.2	76.2	204.8
t_{-2}	2014-03-06 06:00	10	31.0	28.1	9.8	11.6	21.4	31.8	-.04	65.3	73.2	199.8

Secili nga rastet burim përfishin 9 orë të vlerave të matura, nga të cilat 6 janë orët që përkrijnë me orët e parashikimit dhe 3 janë orët para këtyre, 2 orët e kaluara përkatësisht t_{-1} , t_{-2} si dhe t_0 është momenti kohor nga ku do të fillojë parashikimi. E thënë në mënyrë më të thjeshtë nëse jemi të interesuar të parashikojmë energjinë e konsumuar duke nisur në orën 10^{00}

paradite atëherë një *rast burim* do të përfshinte vlerat për të 10 variablat në orën $10^{00}, 9^{00}, 8^{00}$ si dhe vlerat e parashikuara për temperaturën dhe lagështirën për orën $11^{00}, 12^{00}, 13^{00}, 14^{00}, 15^{00}, 16^{00}$ të treguara me blu në tabelë. Gjithashtu vlerat e treguara në gri për 8 variablat e mbetura përfshihen në rast, por nuk merren në shqyrtim në procesin e vlerësimit të ngjashmërisë. Diferenca mes një rasti të ri ku kërkohet zgjidhja dhe një rasti burim është se në rastin e ri për intervalin ku do të kryhet parashikimi kemi vetëm vlerat e parashikuara të temperaturës dhe lagështisë së ajrit. Vlerat që do të parashikojë modeli janë ato të treguara me (?). Në rastin e ri nuk kemi vlera të tjera të përfshira për orët e parashikimit, por kemi vlerat e gjithë attributeve të 2 orëve të shkuara dhe orës nga ku nis parashikimi. Në pjesët më poshtë shpjegohet mënyra sesi punohet me rastet për të prodhuar vlerat e kërkuara (parashikimi).

4.2.2 Grupimi i rasteve të ngjashme

Pasi ndërtohen *rastet burim* ata ruhen përkatësisht në bazën e rasteve ku kërkojmë për zgjidhjen, ndërsa *rastet e reja* ruhen në pjesën e testimit. Një rast i ri i marrë në konsideratë për t'u zgjidhur kërkon në bazën e rasteve për raste ekzistues të ngjashëm me të. Hapi i parë që implementohet tenton zvogëlimin e hapësirës së kërkimit duke grupuar rastet burim në raste të ngjashme me secilin rast në shqyrtim nga rastet e reja (në fokus). Me raste të ngjashme kuptojmë ata raste ekzistues që janë të ngjashëm me një rast të ri referuar disa kriterëve që projektuesi i sistemit vendos, këtu krijohet një situatë që për çdo rast të ri kemi një nëngrup rastesh të ngjashme nga baza e rasteve që më vonë kërkohet zgjidhja duke përshtatur atë ekzistuese tek rasti jonë në shqyrtim. Rastet grupohen në të ngjashme bazuar në dy kushte që paracaktojmë: nëse koha e matjes së një rasti burim nuk është jo më larg se një orë nga koha e matjes së rastit të ri, pra nëse parashikimi fillon në orën 11^{00} paradite atëherë raste të ngjashme do të përfshihen $11^{00}, 10^{00}$ dhe 12^{00} , si dhe kushti i dytë është që temperatura në vlerë absolute midis kohës t_0 të dy rasteve të mos jetë me e madhe se 2°C . Më poshtë paraqesim relacionet matematike për rastin:

$$\begin{cases} |t'_0 - t_0| \leq 1, & \text{kushti i kohës} \\ |V'_{1,0} - V_{1,0}| \leq 2, & \text{kushti i temperaturës} \end{cases} \quad (4.6)$$

Sipas këtyre dy kushteve rastet do të grupohen në të ngjashëm referuar çdo rasti të ri i cili do të merret në shqyrtim. Në fig. 4.3 po japim një skematizim të mënyrës së përzgjedhjes së rasteve sipas mënyrës që thamë.

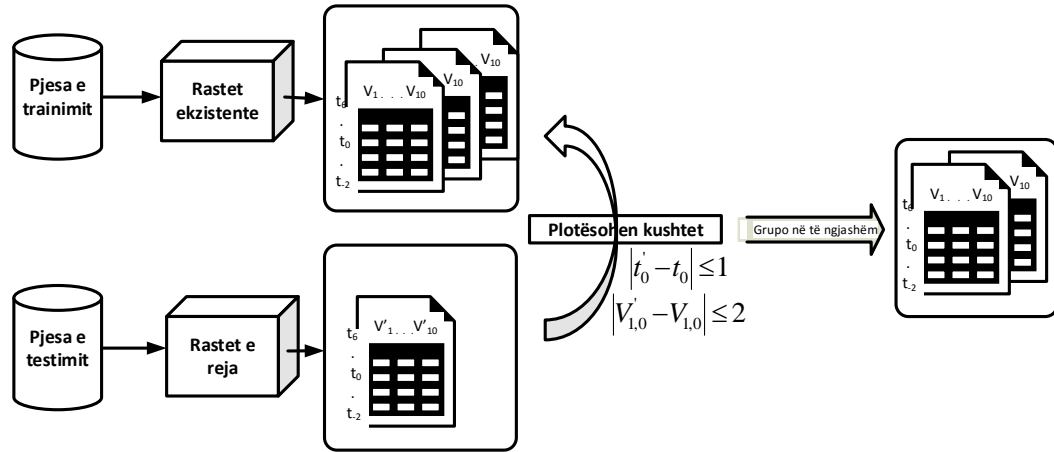


Figura 4.3: Skematizimi i grupimit të rasteve të ngjashme.

4.2.3 Metodën e ngjashmërisë

Analiza e ngjashmërisë që do të përdoret në këtë punim pranon përcaktimet e mëposhtme si pikë reference për çdo ekuacion të mëtejshëm, ku me $x_{i,j}$ do të njehësojmë të gjitha vlerat për rastin burim ku $x_{i,j}$ është vlera e variablit V_i në çastin t_j , me $y_{i,j}$ do të njohim vlerat e rastit të ri, vlerat kanë kufijtë ($1 \leq i \leq 10$, $-2 \leq j \leq 6$). Ndërsa p_j është vlera e fuqisë P në çastin e kohës t_j . Çdo rast shpreh matjet për 9 orë gjithësej sa janë dhe rreshtat në tabelë për të 10 variablat, ku me $x_{i,j}$ kemi matjet për rastin burim si dhe me $y_{i,j}$ për rastin e ri ku kërkohet zgjidhja. Ngjashmërinë midis dy rasteve e njehësojmë si kontribut të ngjashmërisë së *variablave* e cila vlerësohet për gjithë variablat e secilit rast, dhe ngjashmëria midis rasteve është shuma e normalizuar e ngjashmërive lokale.

$$CS = \sum_{i=1}^M v_i VS_i(x_i, y_i) \quad (4.7)$$

Me $v_i \geq 0$ është shënuar pesha e secilit variabël, $M = 10$ shpreh numrin e variablave, ndërsa $VS'_i(x_i, y_i)$ është ngjashmëria lokale e secilës variabël për një rast të caktuar. Matematikisht ngjashmëria e variablave jepet në relacionin (4.8),

$$VS'_i(x_i, y_i) = \begin{cases} 0 & \text{nëse } D_w(x_i, y_i) > d_{max}^i \\ \frac{D_w(x_i, y_i) - d_{min}^i}{d_{max}^i - d_{min}^i} & \text{nëse } d_{min}^i \leq D_w(x_i, y_i) \leq d_{max}^i \\ 1 & \text{nëse } D_w(x_i, y_i) < d_{min}^i \end{cases} \quad (4.8)$$

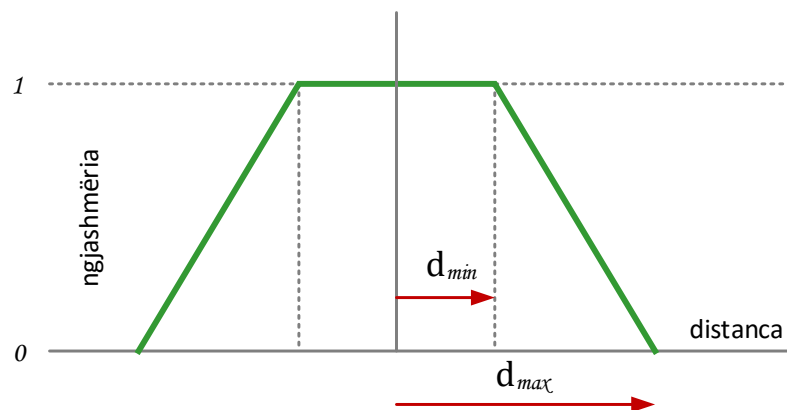


Figura 4.4: Transformimi i distancës Euklidiane në ngjashmëri

ku $D_w(x_i, y_i)$ është distanca Euklidiane e normalizuar sipas peshës specifike të çdo momenti matjeje, ku d_{min}^i dhe d_{max}^i janë shmangie specifike të përcaktuara për secilën variabël nga njohuria e ekspertit. Thënë më qartë përcaktimi i shprehur matematikisht në (4.8) do të ishte se nëse vlera e distancës Euklidiane ndërmjet vlerave respektive të variablave për dy rastet do të ishte më e vogël se shmangia minimale d_{min} atëherë ngjashmëria do të merrte vlerën maksimale pra atë 1, në të kundërt nëse distanca do të kishte vlerë më të madhe se shmangia maksimale d_{max} atëherë do të kishim vlerën minimale të ngjashmërisë 0, ndërsa kur ndodhet në vlerat midis tyre kemi një relacion ku varësia do të ishte lineare siç tregohet në fig. 4.4. Distanca Euklidiane e normalizuar sipas peshës specifike të çdo variabli jepet me relacionin (4.9),

$$D_w(x_i, y_i) = \frac{\sqrt{\sum_{j=-2}^n w_j (x_{i,j} - y_{i,j})^2}}{\sum_{j=-2}^n w_j} \quad (4.9)$$

ku n tregon numrin e matjeve, pra i bie të kemi gjithësej 9 orë me vlera të matura, si dhe pesha specifike për secilën matje përcaktohet sipas një rendi zbritës referuar kohës nga ku fillon parashikimi, thënë ndryshe sa më larg nga koha e nisjes aq më pak peshë i jepet matjes në të dy drejtimet në atë rritës dhe në atë zbritës. Matematikisht do të kishim për $j \in \{-2, -1, 0\}$ peshat do të përcaktoheshin $w_j = 1 + j/3$, ndërsa për $j \in \{1, \dots, 6\}$ peshat do të ishin $w_j = 1 - j/7$, ku influencën më të madhe e kanë vlerat në momentin t_0 . Për të bërë më të qartë se cila pjesë e ngjashmërisë lokale (variablave një për një) do të kalojë në ngjashmërinë globale (ngjashmëria midis rasteve), japim shpjegimin se për tre orët e kaluara pra e tashmja dhe dy të shkuarat që ndodhen në secilin rast bëhet krahasimi sipas distancës së normalizuar Euklidiane i secilës vlerë për

të 10 variablat e matura dhe më tej ajo transformohet në ngjashmëri sipas (4.9) referuar shmangieve minimale dhe maksimale (shiko Tabelën 4.4), ndërsa për 6 orët ku do të kryhet dhe parashikimi bëhet krahasimi vetëm për 6 vlerat e dy madhësive të parashikuara të temperaturës dhe lagështisë pasi supozohet që njohim vetëm vlerat e parashikuara të motit dhe shuma e ngjashmërive lokale të secilës variabël dhe të përkthehet në ngjashmëri sipas rasteve (*globale*) sipas (4.7). Modelimi i ngjashmërisë së rasteve vetëm sipas distancës Euklidiane

Tabela 4.3: Shmangiet specifike për secilën variabël.

Variabli		d_{min}	d_{max}
V_1	Temperatura e parashikuar e ajrit jashtë	2	6
V_2	Lagështia e parashikuar e ajrit jashtë	10	25
V_3	Temperatura e ajrit të ngrohtë në njësinë 2	2	6
V_4	Temperatura e ajrit të ngrohtë në njësinë 3	2	6
V_5	Temperatura e ajrit të ftohtë në anën perëndimore	2	6
V_6	Temperatura e ajrit të ftohtë në njësinë 4	2	6
V_7	Temperatura në dalje të çillerit	2	15
V_8	Rrjedhja e ujit në dalje të çillerit	5	30
V_9	Temperatura në dalje të bolierit	2	15
V_{10}	Rrjedhja e ujit në dalje të bolierit	5	30

është një nga metodat që përdoret shpesh në problematikat e serive të kohës, por përdorimi i kësaj ngjashmërie ka edhe pikat e dobëta. Në këtë punim sjellim një përmirësim të njehësimit të ngjashmërisë duke eliminuar pikën e dobët të ngjashmërisë Euklidiane. Ndodh që dy seri kohore të kenë të njëjtën vlerë të distancës Euklidiane, por të kenë trend të ndryshëm si tregohet në fig. 4.5. Nëse analizojmë seritë kohore A, B dhe C, ato kanë të njëjtën vlerë të distancës totale Euklidiane midis tyre duke qënë se distanca shprehet si rrenja katrore e largësisë në katror. Edhe pse kanë të njëjtën distancë nga seria referencë A, midis A dhe B vërehet një trend identik ndërsa mes C dhe A ka një trend të ndryshëm fig. 4.5. Sipas ngjashmërisë së përcaktuar në (4.8) rezultati që do të prodhohej nga krahasimi A-B dhe A-C do të ishte i njëjtë. Për të evituar këtë problematikë një metodë e re matjeje ngjashmërie përdoret si është ngjashmëria sipas formës e përcaktuar dhe në [151] ose (*cosine similarity*) e shpjeguar me detaje në seksionin 3.3.3: Një sekuencë vlerash $x_i = (x_{i,-2}, x_{i,-1}, \dots, x_{i,n})$ do të konsiderohet si një grup vektorësh në hapësirën

1-dimensionale:

$$\left\{ (\Delta x_j, \Delta t_j)^T = (x_{i,j+1} - x_{i,j})^T \mid j = -2, \dots, n-1 \right\} \quad (4.10)$$

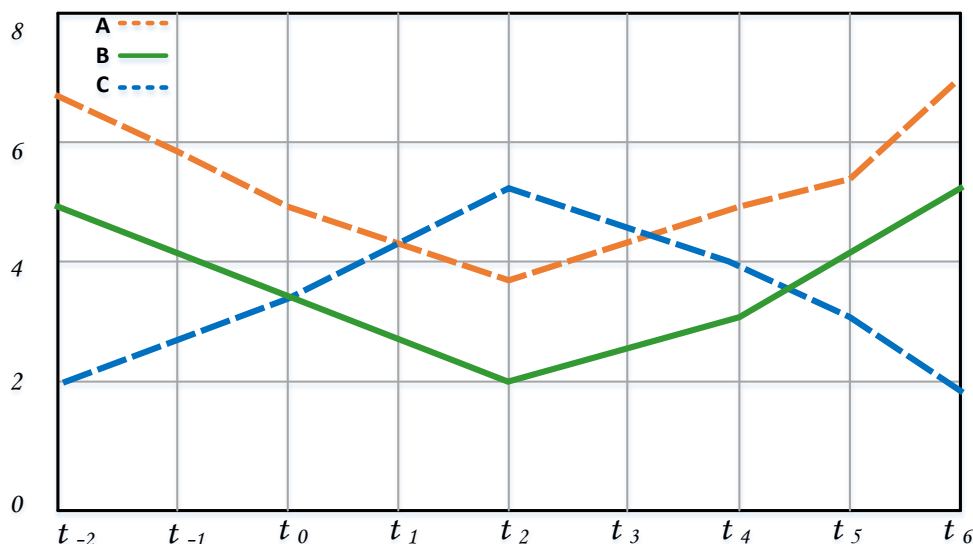


Figura 4.5: Krahasimi i distancës Euklidiane midis serive kohore.

Ngjashmëria jepet si këndi i normalizuar midis dy vektorëve, ku hapi i ndryshimit të kohës është gjithmonë 1 pasi matjet janë çdo 1 orë, elementi i dytë në të djathtë të barazimit (4.11). Ngjashmëria sipas formës për dy seri kohore jepet sipas relacionit (4.11).

$$VS''_i(x_i, y_i) = \frac{1}{\pi(n+2)} \sum_{j=-2}^{n-1} \cos^{-1} \frac{\Delta x_j \Delta y_j + 1}{\sqrt{(\Delta x_j)^2 + 1} \sqrt{(\Delta y_j)^2 + 1}} \quad (4.11)$$

Kjo metodë përdoret për të vlerësuar ngjashmërinë e dy madhësive të parashikuara ambientale, siç janë temperatura dhe lageshtia për 6 orët e parashikimit si dhe 3 orët e shkuara. Seria kohore e cila analizohet ka një gjatësi intervali 9 orë dhe përfshin vetëm variablat që kanë edhe vlerat e parashikuara jo për gjithë variablat. Ngjashmëria për çdo çift varet nga këndi që formohet midis tyre fig. 4.6, për kënd 0 radian marrim ngjashmërinë të barabartë me 1 si dhe për π radian marrim 0. Matematikisht përmbledhet në (4.11).

Vlerësimi i ngjashmërisë sipas variablave shprehet si kontribut i ngjashmërisë sipas distancës Euklidiane me ngjashmërinë sipas formës (4.12).

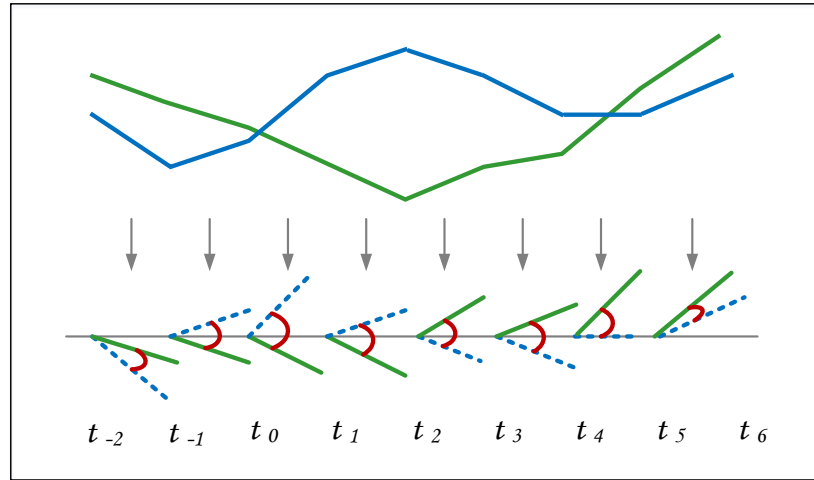


Figura 4.6: Paraqitja e një sekuence kohe si një grupim vektorësh

Mesatarja e tyre jep ngjashmërinë totale.

$$VS_i(x_i, y_i) = 0.3 VS'_i(x_i, y_i) + 0.7 VS''_i(x_i, y_i) \quad (4.12)$$

Vlerat 0.3 dhe 0.7 që kemi vendosur përcaktohen gjatë fazës së trajnimit të modelit, ku zgjidhen vlerat në të cilat arrihet performanca më e lartë.

4.2.4 Ngjashmëria e rastit

Më sipër u përcaktua ngjashmëria e variablave që ndërtojnë një rast, ndërsa për të përcaktuar ngjashmërinë globale. Kjo përcakton sa të ngjashëm janë dy raste midis tyre, këtu i referohemi relacionit në (4.7) i varur nga dy terma, njëri shpreh ngjashmërinë totale midis variablave dhe tjetri shpreh peshën specifike që secila variabël ndikon në konsumin e energjisë. Peshat e variablave do të përcaktohen sipas relacionit (4.13). Përzgjedhja e rastit ose e rasteve më të vlefshëm do të bëhet duke marrë rastet me ngjashmëritë me të larta, në literaturë nga punime të ngjashme ekziston një qasje ku zgjidhen rastet të cilat plotësojnë një kusht të caktuar për ngjashmërinë, si për shembull në [131, 132] autorët pranojnë që nëse rastet plotësojnë një kusht ku $CS > 0.8$ atëherë ato janë rastet të cilat do të merren për zgjidhjen përfundimtare, ndërsa në modelin tonë bazuar në metodën e *fqinjtit më të afërt* pranojmë si më të vlefshëm 50 rastet të cilat kanë ngjashmërinë më të lartë (*ku numri 50 optimizohet gjatë procesit të trajnimit të modelit*) dhe në rast se ky kusht nuk plotësohet, pra kemi më pak raste të ngjashëm nga baza e rasteve atëherë pranojmë ata që ndodhen.

4.2.5 Përcaktimi i peshës së variablave

Në literaturë shpesh ndodh që për thjeshtësi të modelit rëndësia e variablave pranohet e njëjtë për gjithë variablat që do të thotë se të gjitha ndikojnë njësoj tek madhësia e daljes. Rasti konkret tek [132], ku autorët pranojnë se pesha e secilit variabël është $v_i = 1/M$, ku M është numri i variablave. Megjithatë, përderisa roli i variablave të ndryshëm medoemos që është i ndryshëm në madhësinë e daljes (*energjinë e konsumuar*) atëherë del e nevojshme edhe përcaktimi i ndikimit të këtyre variablave në konsum duke atribuar pesha specifike tek secila variabël. Problemi i njehësimit të rëndësisë së variablave lidhet ngushtë me vlerësimin e ngjashmërisë në modelet ABR [141, 142, 143]. Duke u nisur nga fakti se ngjashmëria globale e shprehur në (4.7) është kombinacion linear i ngjashmërive lokale, jepet si shumë e tyre, problemin e njehësimit të peshave e formalizojmë në një problem të regresionit linear seksioni 2.2.2.

Peshat e variablave do të përcaktohen për rastet që ndodhen në pjesën e trajnimit, ku për çdo dy raste burim X dhe Z ne llogarisim ngjashmëritë lokale për çdo variabël duke marrë një marrëdhënie të tillë:

$$(s_1, \dots, s_M) = (VS_1(x_i, z_i), \dots, VS_M(x_i, z_M)), \in [0, 1]^M \quad (4.13)$$

Ku vlerat e koeficientëve s_1, \dots, s_M janë vlerat e prodhuara nga transformimi në (4.8) kjo është vetëm për vlerat e koeficientëve të variablave ndërsa për vlerën e madhësisë në dalje të kishim vlera të përcaktuara nga zgjedhja e $d_{min} = 15$ dhe $d_{max} = 35$ dhe ekuacionet që duhet të zgjidhim për përcaktimin e peshave duhet të plotësojnë kushtin.

$$\sum_{j=1}^M v_j \cdot s_j \approx s_{out} \quad (4.14)$$

Ku s_{out} janë koeficientët e prodhuar nga vlera e madhësisë dalese (*konsumit të energjisë*), përcaktimi i peshave bëhet në MATLAB duke përdorur *constrained linear least-squares*¹ dhe vlerat jepen në tabelën më poshtë.

Nga rezultatet në tabelën 4.4 duket qartë se rolin kryesor (peshën më të madhe) e luajnë madhësitë e lidhura ngushtë me kushtet atmosferike, pasi dihet qartë se temperatura dhe lagështia luajnë një rol të rëndësishëm në konsumin e energjisë ku në temperatura të ulta si në rastin në fjalë (Kanada) ndikojnë mbi sistemet e ngrohjes dhe ftohjes, po ashtu rol të madh luajnë parametrat e paisjeve që janë instaluar për sistemin ngrohës në godinë. Prandaj përcaktimi i peshave të modelit është një proces shumë me rëndësi në metodat inteligjente si ABR.

¹ <http://uk.mathworks.com/help/optim/ug/lsqlin.html>

Tabela 4.4: Përshkrimi i variablave dhe njësitë matëse

Variabli	Pesha	
V_1	Temperatura e parashikuar e ajrit jashtë	0.1961
V_2	Lagështia e parashikuar e ajrit jashtë	0.1540
V_3	Temperatura e ajrit të ngrohtë në njësinë 2	0.0001
V_4	Temperatura e ajrit të ngrohtë në njësinë 3	0.0001
V_5	Temperatura e ajrit të ftohtë në anën perëndimore	0.0009
V_6	Temperatura e ajrit të ftohtë në njësinë 4	0.1065
V_7	Temperatura në dalje të çillerit	0.1075
V_8	Rrjedhja e ujit në dalje të çillerit	0.0001
V_9	Temperatura në dalje të bolierit	0.3064
V_{10}	Rrjedhja e ujit në dalje të bolierit	0.2184

4.2.6 Adaptimi i zgjidhjes

Ky hap jep zgjidhjen përfundimtare, mbasi janë pranuar rastet më të ngjashme me secilin rast në shqyrtim (rast të ri) do të transformohen zgjidhjet ekzistuese për të dhënë zgjidhjen e re. Një nga qasjet për të finalizuar këtë hap është duke marrë zgjidhjen ekzistuese nga rastet burim ashtu siç është duke e normalizuar me ngjashmërinë e rastit dhe bëjmë parashikimin (4.15).

$$\hat{P}_j = \frac{\sum_{k=1}^K CS(X_k, Y) \cdot P_{k,j}}{\sum_{k=1}^K CS(X_k, Y)} \quad (4.15)$$

ku $P_{k,j}$ është konsumi i rastit X_k në çastin e kohës nga $j \in \{1, \dots, 6\}$ që korrespondon me intervalin ku kryhet parashikimi. Megjithatë ekzistojnë hapësira ku mund të përmirësohet ky hap i rëndësishëm që përkon me fazën e fundit në implementimin e modelit. Përmirësimi në adaptim vjen duke propozuar implementimin e një mënyrë të re të adaptimit të zgjidhjes, e njohur në literaturë si transformimi i amplitudës [144]. Më specifikisht, duke pranuar se konsumi për një rast të ri do të jetë përafërsisht i ngjashëm me konsumin e rasteve (burim) nga ku merret dhe adaptohet zgjidhja (4.15), atëherë propozojmë një faktor transformimi i cili bën që energjia e konsumuar në rastet burim do të zhvendoset proporcionalisht sipas këtij faktori para se të përdoret për parashikim. Pranojmë se vlerat e konsumit janë P_{-2} , P_{-1} , P_0 për tre orët e shkuara nga rasti i ri, ku jemi të interesuar për të parashikuar për 6 orët në vijim, janë të njohura (shiko Tabelën 4.3), si dhe le të jenë $P_{k,-2}$, $P_{k,-1}$, $P_{k,0}$

konsumi për 3 orët e shkuara i k rasteve fqinje më të afërta me rastin tonë (*ngjashmërisht më afër*). Me këto pranime bëjmë transformimin e 6 orëve të rasteve burim më të ngjashme, nga ku do të adaptohet parashikimi i konsumit, duke përmirësuar parashikimin.

$$P'_{k,j} = P_{k,j} \frac{P_{-2} + P_{-1} + P_0}{P_{k,-2} + P_{k,-1} + P_{k,0}} \quad (4.16)$$

Dhe duke bërë zëvendësimin e (4.16) tek ekuacioni (4.15) do të bëhej zgjidhja përfundimtare, pra do të jepeshin vlerat e parashikuara për energjinë e konsumuar:

$$\hat{P}_j = \frac{\sum_{k=1}^K CS(X_k, Y) P'_{k,j}}{\sum_{k=1}^K CS(X_k, Y)} \quad (4.17)$$

Për një ilustrim më të qartë jepet grafiksht në fig. 4.7 ku sekuenca e rastit burim (vija e verdhë solide) do të zhvendoset sipër sipas tregimit të shigjetës e tillë që mesatarja e vlerave të tre pikave (zona gri) të përkojë me mesataren e rastit të ri (vija e ndërprerë); dhe parashikimi do jetë vija blu.

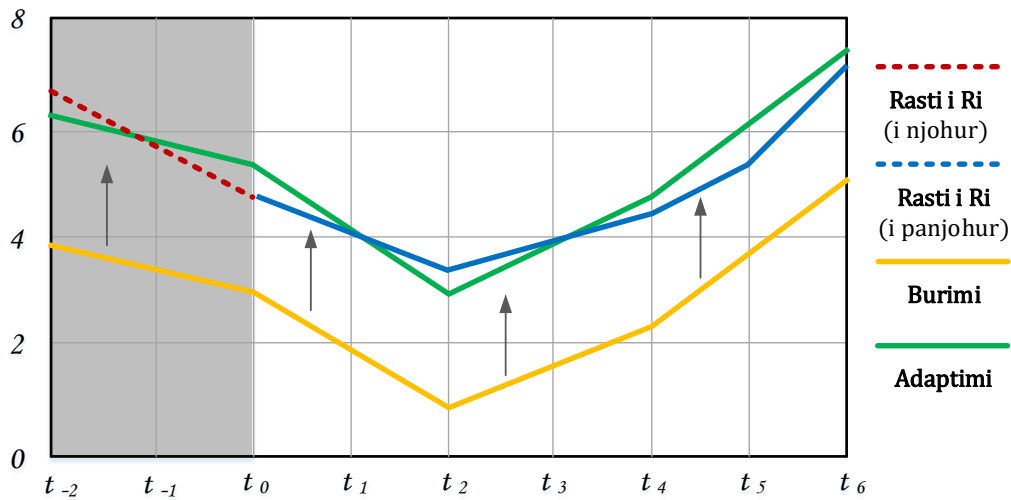


Figura 4.7: Transformimi në amplitudë i vlerave të konsumit të energjisë

4.2.7 Modeli statik i parashikimit

Modeli offline ose ndryshe modeli ku rastet kërkohen brenda një bazë rastesh që janë ruajtur më parë, ku disponohen të dhëna për një periudhë të caktuar dhe për to veprojmë si ne seksionin 3.4.1. Theksojmë se rastet e krijuara

janë raste që përmbajnë të dhëna për 9 orë dhe brenda një dite nga $07^{00} - 17^{00}$ do të krijohen më shumë se një rast. Rastet krijohen duke ju ‘mbivendosur’ njëri-tjetrit, domethënë nëse rasti i parë ka për t_{-2} orën 07^{00} pra fillimin si dhe mbarimi i bie të jetë ora 07^{00} , rasti i dytë do të kishte orën 07^{00} si pikën e nisjes dhe përkatësisht për orët e parashikimit ($10^{00}, \dots, 15^{00}$) rasti i parë dhe ($11^{00}, \dots, 16^{00}$) rasti i dytë do të kishim mbivendosje në orët ($11^{00}, \dots, 15^{00}$), përkundrazi kjo nuk është mangësi e metodës pasi i lejon përdoruesi fleksibilitet dhe arrin të japi një tablo të plotë të gjithë intervalit të matjes. Një veprim tjetër do të çonte në mangësi të modelit dhe mundësi jo përmirësimi, fleksibiliteti ofrohet në mundësinë që përdoruesi kërkon të nisë në orën 11^{00} parashikimin dhe një tjetër në 09^{00} , dhe kjo e lejon. Gjithësej numri i rasteve për vitin 2013 të formuara është 443 si dhe numri i rasteve për të dhënat që vijnë nga viti 2014 është 153. Për të parë performancën e modelit e krahasojmë atë me rrjetat neurale artificiale një model mjaft i suksesshëm në literaturë për probleme të ngjashme ku në seksionin 2.2.3 japim të detajuar mënyrën sesi funksionon ky model, si dhe kundrejt tre modele të tjera bazike ose siç njihen ndryshe ‘bazë’. Eksperimentalisht do të krahasojmë:

1. *Modeli bazë 1* metoda përfshin vlerën mesatare të konsumit energjetik për të gjithë rastet e marra si raste burim. Pra vlera e parashikuar për 6 orët në vijim do të jetë e barabartë me vlerën mesatare të konsumit nga gjithë pjesa e trajnimit.
2. *Modeli bazë 2* përfshin vlerën mesatare të konsumit për tre orëve të shkuara nga rasti ti ri. Qartësisht vlera e parashikuar për 6 orët në vijim do të jetë sa vlera mesatare e konsumit në t_{-2}, t_{-1}, t_0 pra:

$$\hat{q}_j = \frac{q_{-2} + q_{-1} + q_0}{3} \quad (4.18)$$

3. *Modeli bazë 3* parashikon vlerën e konsumit në momentin ku fillon parashikimi. Pra konsumi i 6 orëve në vijim është sa $\hat{q}_j = q_0$.

Implementimi dhe vlerësimi i performancës së modeleve “offline” do të bëhet duke ndarë të dhënat në disa raporte (trajnim, testim). Kjo për të vlerësuar përgjithësimin e modelit mbi të dhënat si dhe performancën e tij kur ekzistojnë pak të dhëna, dhe etapat janë:

1. 01-Janar-2013 deri më 01-Prill-2013
2. 01-Prill-2013 deri më 01-Qershor-2013

3. 01- Qershor -2013 deri më 01-Gusht -2013
4. 01-Gusht -2013 deri më 01- Tetor -2013
5. 01- Tetor -2013 deri më 01-Dhjetor -2013

Të gjitha këto intervale kohore që sipërpërmendëm tregojnë se sa është sasia e rasteve burim që kemi të njohur pra sa rasteve ju njohim përgjigjen. Për periudhën e parë kemi vetëm 4 muaj të dhëna që përdorim për trajnim dhe pjesa tek tjetër që mbetet deri ne fund të matjeve do të kryhet vlerësimi i modelit, pra do të bëjmë parashikimin e konsumit. Pasi kemi parashikuar konsumin dhe kemi marrë vlerat do të vlerësojmë performanën e modeleve dhe për këtë përdorim metodën e dytë të përshkruar seksionin në 4.1.4.

$$CV - RMSE = 100 \cdot \frac{\sum_{i=1}^n \sqrt{\frac{(q_i - \hat{q}_i)^2}{(n-1)}}}{\hat{q}_i} \quad (4.19)$$

Vlerat reale q_i krahasohen me vlerat e parashikuar \hat{q}_i në periudhën e testimit për n kampionet e parashikuara.

4.2.8 Krahasimi i modeleve

Performanca e modelit ABR që kemi implementuar si metodë për parashikimin e konsumit të energjisë për 6 orët e ardhshme në një godinë institucionale krahasohet me modelet e tjera si rrjetat neurale dhe modelet bazë për skenarë të ndryshëm ndarjeje të të dhënave siç u përmendën.

1. Modeli bazë 1
2. Modeli bazë 2
3. Modeli bazë 3
4. ABR
5. Rrjetat neurale

Në tabelën 4.5 tregohen rezultatet e përfuara nga eksperimentimi në situata të ndryshme për secilin model. Performanca më e mirë është shënuar me të errët. Nga rezultatet duket qartë se modeli ABR i implementuar në këtë pjesë performon më mirë në kushtet kur kemi më pak të dhëna për trajnim, kjo është një nga pikat dominante për metodat që mbështeten tek ABR. Rezultatet e arritura me këtë model janë brenda kushteve të përcaktuara nga ASHRAE të cilat përmenden në pjesën 4.1.4, po ashtu vlen të theksohet se rezultatet e arritura janë pothuajse dy herë më pak se rezultatet e marra në [132] i vlerësuar tek të njëjtat të dhëna.

Tabela 4.5: Rezultatet e modelit offline në termat e CV-RMSE (%)

Të dhënat e trajnimit	M.bazë 1	M.bazë 2	M.bazë 3	RN	ABR
01/2013 – 04/2013	9.97	9.55	8.80	8.99	7.94
01/2013 – 06/2013	10.15	9.50	8.65	8.15	7.39
01/2013 – 08/2013	10.23	9.41	8.58	7.63	7.45
01/2013 – 10/2013	10.12	9.35	8.66	7.31	7.69
01/2013 – 12/2013	10.35	9.54	8.73	6.17	6.55

4.2.9 Modeli dinamik i parashikimit

Modeli online është modeli ABR i cili zhvillohet pa pasur raste burim të ruajtura në bazën e rasteve më parë, por ajo pasurohet si pasojë e mënyrës inkrementale të përditësimit të rasteve. Rastet e ndërtuara në skenarin e parashikimit statik seksioni 4.2.7 përdoren edhe në këtë skenar. Metodatat e prezantuara për gjetjen e rasteve të ngjashme janë po ato, por ajo që ndryshon këtu është mënyra e kërkimit të zgjidhjes dhe përditësimit të rasteve. Rastet fillojnë të ruhen në bazën e rasteve hap pas hapi, duke filluar me një rast të vetëm dhe duke shtuar në secilin hap kohe nga një të ri. Në momentin që në bazën e rasteve kemi vetëm një rast të ruajtur rasti që vjen pas tij përdoret si rast i ri, pra ky kërkim të gjej vlerën e parashikimit të konsumit për 6 orët që vijojnë. *Si kryhet parashikimi?* Rasti i ri në shqyrtim kërkim mbi ngjashmërinë me rastin në bazën e rasteve, ai vlerëson ngjashmërinë dhe bën parashikimin sipas (4.17), duke marrë në konsideratë vetëm rastin ekzistues, pasi parashikojmë konsumin për një rast ai ruhet në bazën e rasteve si rast burim ku vlerat e konsumit janë ato të matura dhe jo ato të parashikuara. Me shtimin e rasteve në bazën e rasteve atëherë konsumi i energjisë është adaptimi i vlerave të rasteve më të ngjashme. Me rritjen e rasteve në bazën e rasteve modeli përmirëson performancën e tij. Vlerësimi i performancës kryhet duke përdorur gabimin grumbullues CVRMSE që nënkupton se ai llogaritet sa herë që një rast i ri parashikon vlerat e tij dhe pastaj futet në bazën e të dhënave si rast burim. Performanca e modelit krahasohet me tre modelet bazë të përdorur për parashikimin statik. Ku modelet përshaten lehtësisht, modeli bazë 1 parashikon mesataren e aq rasteve sa janë ruajtur në bazën e rasteve, ndërsa dy modelet e tjera nuk pësojnë ndryshim pasi varen nga rasti i ri. Performanca e modeleve tregohet në fig. 4.8.

Nga fig. 4.8 duket qartë se performanca e modelit online me rritjen e numrit të rasteve sa vjen e përmirësohet duke arritur nivelin asimptotik të gabimit që paraqet modeli offline. Pra sasia e rasteve në bazën e të dhënave luan rol vendimtar në performancën e modelit, sa më i lartë numri i rasteve aq më e

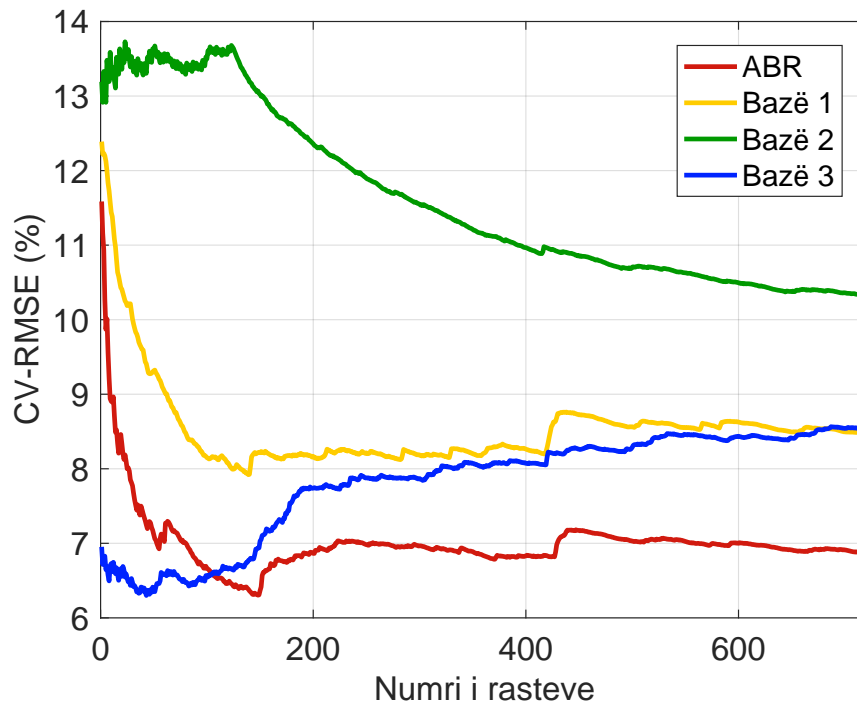


Figura 4.8: Performanca e modelit online ABR.

lartë performanca. Modeli online ka avantazhin se mund të aplikohet në godinat e reja të cilat e kanë të pamundur mbledhjen e të dhënave historike, ky model mund të integrohet me sistemet e menaxhimit të konsumit të energjisë (BMS) të instaluar në një godinë. Kjo bën të mundur që menaxherët e godinave, operatorët, pjesa teknike dhe ekonomike e njësisë të kenë mundësinë e marrjes së masave paraprake për kontrollin e konsumit, për rritjen e efijencës si dhe për zvogëlimin e pikut të konsumit total ku kjo e fundit lidhet ngushtë me politikat e kompanive shpërndarëse të energjisë si dhe me menaxhimin e rrjetit në tërësi.

4.3 Përmbledhje

Në këtë kapitull elaboruam idenë e implementimit të një modeli ABR për parashikimin e konsumit të energjisë në një ndërtesë institucionale në Kanada, parashikimi bëhet në intervalin e 6 orëve në avancë dhe rezultatet tregojnë performancë shumë të mirë të modelit krahasuar me modelet e tjera, po ashtu një avantazh të lehtë kemi edhe ndaj rrjetave neurale, një model ky mjaft i suksesshëm në literaturë [70, 72, 75] veçanërisht në situatat kur sasia e të dhënave është e limituar, po ashtu modeli jep rezultate qartësisht më të dukshme referuar udhëzimit të ASHRAE i cili karakterizon se gabimi (CV-RMSE) për të dhëna të kalibruara në varësi të orëve duhet të jetë jo më i madh se 30%, një hap tjetër

i rëndësishëm i modelit është tejkalimi pothuajse me dy-fishin e performancës të njëjtat modele të zhvilluara më parë [131, 132] për të njëjtat të dhëna si dhe modelin tjetër ABR ku ai përdor të dhëna të tjera të paraqitur në [130].

Kapitulli 5

PARASHIKIMI I KONSUMIT NË SKENARË TË NDRYSHËM

Në këtë kapitull përshkruhen rezultatet eksperimentale të arritura në këtë punim nga përdorimi i metodave të inteligjencës artificiale në të dhëna për disa godina me karakteristika dhe matje të ndryshme nga ato që përmendëm në Kapitullin 4. Kontributet kryesore të këtij kapitulli i ndajmë në dy kategori sipas teknikës së parashikimit:

1. Parashikimi i konsumit në skenarin statik “offline”,
2. Parashikimi i konsumit në skenarin dinamik “online”,

Shkurtimisht parashikimi statik kryhet duke përdorur algoritmat tradicionale të machine learning. Modeli mëson sipas teknikës së supervizuar, ku të dhënat e trajnimit atij i shfaqen njëkohësisht dhe mbi to përshtatet modeli sipas çiftit (*hyrje, dalje*).

Ndryshe nga skenari i parë, parashikimi dinamik i merr të dhënat hap pas hapi dhe mëson mbi to në mënyrë dinamike, pra modeli përditësohet sa herë që shembujt e mësimi shtohen.

Parashikimi statik kryhet duke aplikuar metodat e rrjetave neurale (RN), regresionit me vektorët ndihmës (RVN), pemëve të regresionit (PR), pyllit të rastësishëm (PR_{ast}) dhe arsytimit të bazuar tek rastet (ABR) i cili po ashtu përshtatet në parashikimin dinamik. Metodat e parashikimit dinamik përshkruhen në seksionin 5.2. Parashikimi kryhet për 5 ndërtesa me funksionalitete të ndryshme, të dhënat e disponueshme përshkruhen në seksionin 5.1.1.

5.1 Parashikimi statik “offline”

Kjo pjesë fokusohet në parashikimin e konsumit të energjisë sipas skenarit statik. Siç përmendet modeli i merr të dhënat që në fillim duke mësuar mbi to një herë të vetme, për të bërë parashikimin e konsumit. Metodatat të cilat përdorim ndahen në dy kategori: të bazuara tek modeli si RN, RVN, PR, PR_{ast} dhe të bazuara tek shembujt si ABR. Më poshtë vijojmë me përshkrimin e të dhënave, implementimin e modeleve dhe rezultatet e arritura në secilën metodë.

5.1.1 Përshkrimi i të dhënave

Të dhënat e përdorura në këtë pjesë i sigurojmë nga aksesimi i lirë që ato ofrojnë sepse janë (*të vlefshme publikisht*¹). Të dhënat e siguruar për 5 ndërtesa me funksione të ndryshme si rezidenciale, tregtare, shërbime publike përmbajnë matje për vetëm dy parametra temperatura e jashtme e ajrit dhe konsumi i energjisë. Matjet e kryera në frekuencën e 15 minutave, mbulojnë periudhën e plotë 1 vjeçare (nisur në 1 Janar 2010 deri më 1 Janar 2011). Temperatura matet në gradë Fahrenheit, ndërsa konsumi i energjisë në kW. Vlerat e matura i konvertojmë sipas madhësive në sistemin SI, temperatura e konvertuar në gradë Celsius është,

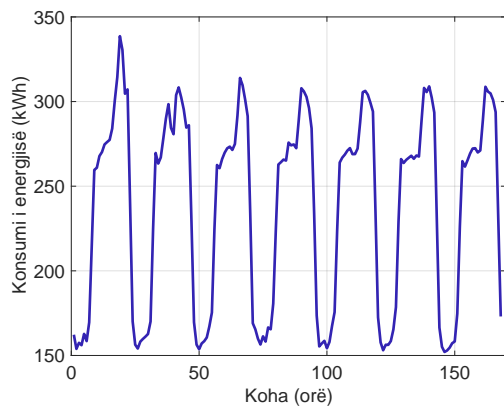
$$T(^{\circ}C) = (T(^{\circ}F) - 32) \cdot 5/9$$

Duke qënë se interesohemi në parashikimin për orën e ardhshme të konsumit, atëherë vlerat e matura sipas intervaleve 15 minutëshe i kthejmë në matje orare thjesht duke mesatarizuar sipas orës,

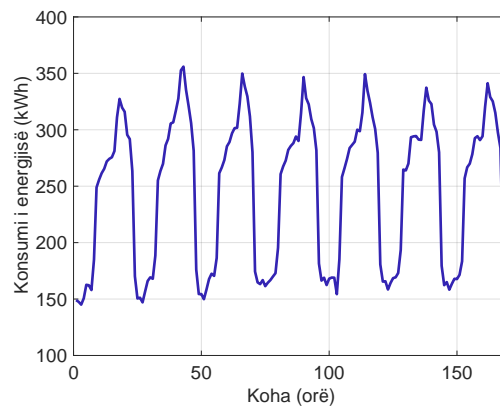
$$P_t = \frac{1}{n} \sum_{i=1}^n P_{ti}$$

ku n është numri i intervaleve dhe q_{ti} konsumi i energjisë. Zakonisht të dhënat bruto përmbajnë vlera të matura gabim ose mungojnë vlera në to, ekzistenca e këtyre vlerave ndikon ndjeshëm mbi performancën e modelit. Një metodë e thjeshtë aplikohet për të zëvendësuar vlerat zero duke marrë mesataren e dy vlerave fqinje, ndërsa kur kemi mungesë të më shumë vlerave ai interval eliminohet nga të dhënat, për shembull në ndërtesën 3 të dhënat (nga 16/08/2010 deri më 3/09/2010) zhvendosen nga baza e të dhënave pasi vlerat e tyre mungojnë. Të dhënat e konsumit pasi pastrohen dhe transformohen sipas përshkrimeve të mësipërme ku konsumi javor tregohet në fig. 5.1 dhe ai vjetor në fig. 5.2.

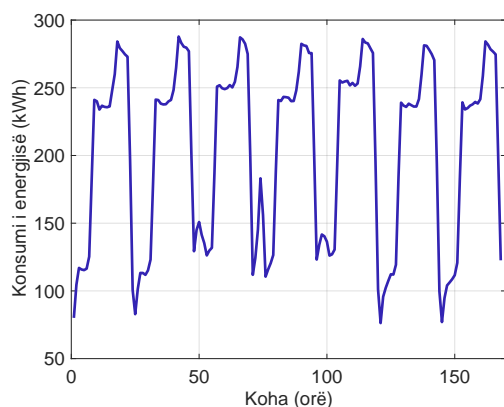
¹<http://en.openei.org/datasets>



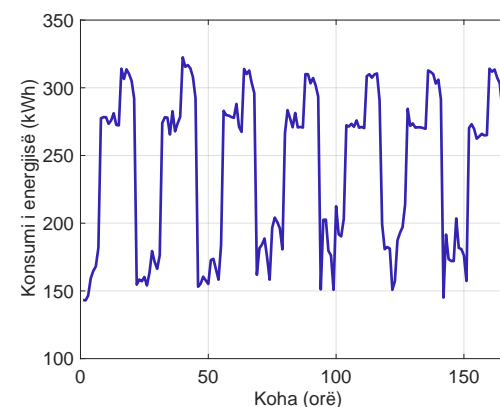
(a) Ndërtësa 1



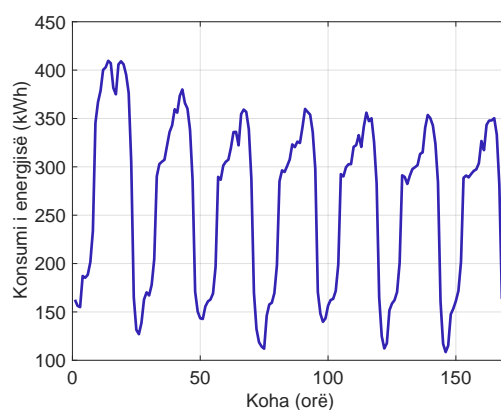
(b) Ndërtësa 2



(c) Ndërtësa 3



(d) Ndërtësa 4



(e) Ndërtësa 5

Figura 5.1: Konsumi javor i energjisë në secilën ndërtësë

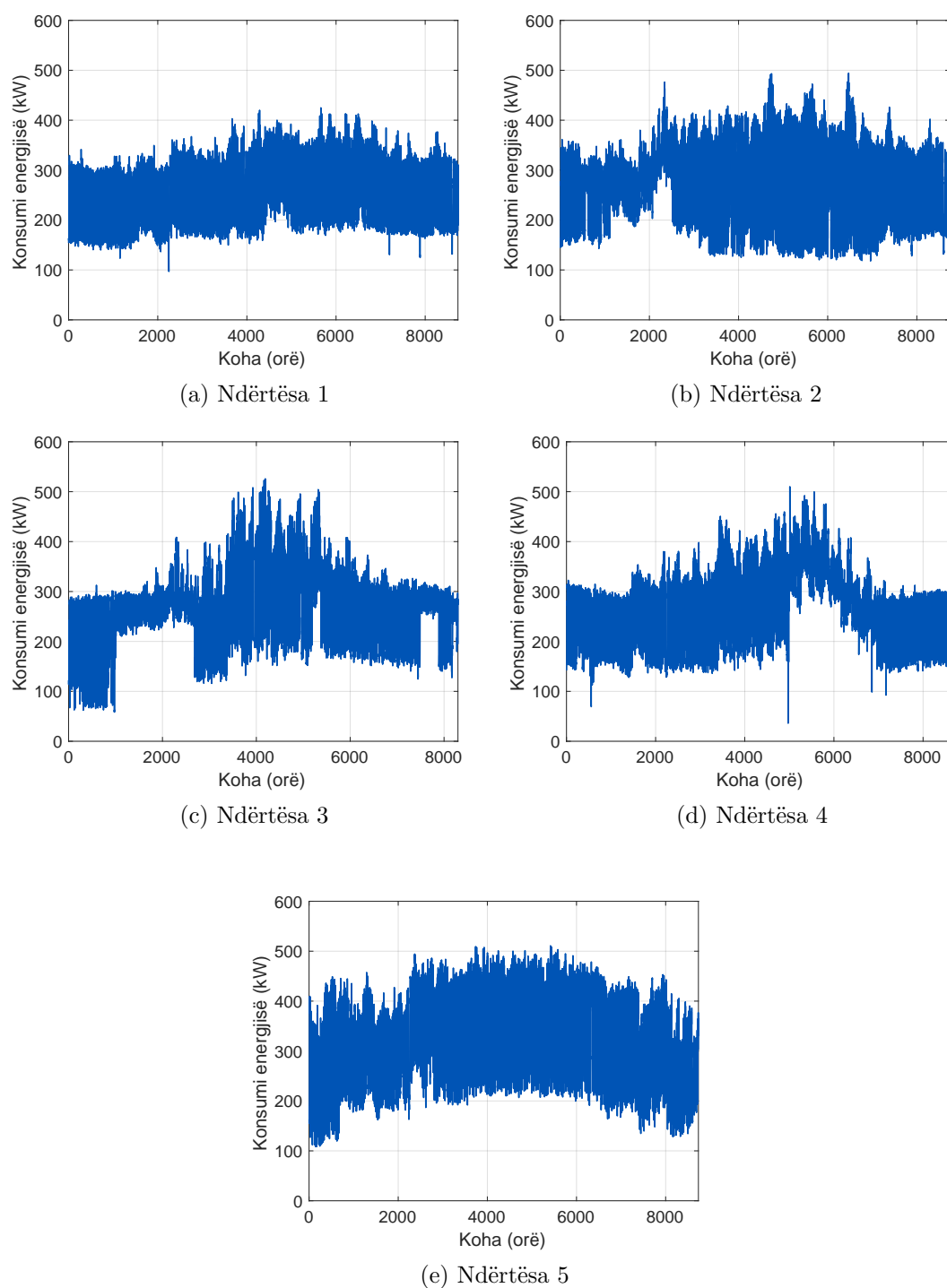


Figura 5.2: Konsumi vjetor i energjisë në secilën ndërtëse

5.1.2 Modelet e përdorura

Rrjetat Neurale

Rrjetat Neurale, të përdorura në këtë punim vijnë nga mjedisi MATLAB, Neural Network Toolbox TM. Përdorimi i rrjetave neurale i referohet rrjetave me përhapje të prapme (backpropagation) të cilat ofrojnë disa algoritma trajnimi, ku përzgjedhja e algoritmit të *mësim*it është funksion i disa veçorive, kompleksiteti i problemit, sasia e të dhënave, etj. Në rastin tonë struktura e rrjetës së përzgjedhur është një rrjetë me një shtresë të fshehur. Përzgjedhja e saj bëhet referuar eksperiencave të nxjerra nga eksplorimi i literaturës [171], ku kjo rezulton si rrjeta më efikase në zgjidhjen e problemeve të njëjta. Element tjetër i rëndësishëm është përcaktimi i numrit të neuroneve në shtresën e fshehur. Pasi përcaktojmë algoritmin e mësimiit që është BFGS Quasi-Newton (*trainbfg*), përzgjedhim numrin e neuroneve të fshehura duke përdorur vlerësimin kryq për të dhënat e hyrjes të ndara në 5 tufa (5 fold cross validation). Ekperimentimi kryhet duke ndryshuar numrin e neuroneve nga 1 në 10, dhe rezulton se me performancën më të lartë për 5 godinat e ka struktura me 10 neurone të fshehura. Në shembullin e tablës së mëposhtme 5.1 tregojmë sesi saktësisht funksionon një procedurë e vlerësimiit kryq. Fillimisht kampionet që kemi rezervuar për trajnim ndahen në 5 pjesë (tufa) të barabarta dhe pastaj për një vlerë të caktuar parametrash modeli trajnohet në 4 pjesë dhe testohet në një të prej tyre, kjo procedurë përsëritet 5 herë. Në rastin e rrjetave neurale vlerësimi kryq bëhet duke fiksuar një strukturë rrjete me një numër të caktuar neuronesh dhe më tej bëhet vlerësimi në të gjitha pjesët C_1, C_2, C_3, C_4 , mbasi mbaron kjo procedurë ndryshojmë neuronet në shtresën e fshehur dhe rivlerësojmë modelin, procedurën e përsërisim deri sa përfundojmë numrin e neuroneve të marrë në konsideratë.

Tabela 5.1: Modeli i vlerësimiit kryq (5-fold cross validation).

Të dhënat:

p_1	p_2	p_3	p_4	p_5
-------	-------	-------	-------	-------

Hapi	Trajno tek	Testo tek	Rezultati
1	p_1, p_2, p_3, p_4	p_5	x_1
2	p_1, p_2, p_3, p_5	p_4	x_2
3	p_1, p_2, p_4, p_5	p_3	x_3
4	p_1, p_3, p_4, p_5	p_2	x_4
5	p_2, p_3, p_4, p_5	p_1	x_5

Regresioni me Vektorët Ndhmës

Regresioni me Vektorët Ndhmës, i dedikohet funksionit *fitrsvm* nga Statistics and Machine Learning Toolbox, MATLAB. Aplikimi i metodës në zgjidhjen e problemeve konkretë kërkon fillimisht përcaktimin e funksionit kernel që do të përdor “mësuesi”. Funksionet kernel të sipërpërmendur në seksionin 2.2.4 janë disa llojesh dhe duke provuar secilin prej tyre, për rastin tonë më i favorshmi për t’u përdorur është *kerneli linear*, kjo vjen nga rezultatet e vlerësimit kryq me 5 pjesë. Pasi përcaktohet funksioni kernel që do të përdoret, më tej vazhdohet me përcaktimin e parametrave të vektorëve ndhmës që janë çifti C, ϵ . Parametrat optimal të C, ϵ përcaktohen duke përdorur optimizimin Bayesian, ku sipas të cilit kemi të bëjmë me një metodë optimizimi globale të funksioneve duke i pranuar ata si funksione të rastësishëm. Optimizimi Bayesian funksionon duke ndërtuar një model statistikor, i cili kërkon në hapësirën e të dhënave vlerat ku funksioni objektiv (në rastin tonë gabimi) arrin vlerën minimale. Kjo e përkthyer në problemin tonë modeli optimizues kërkon vlerat e C, ϵ për të cilat gabimi i RVN i vlerësuar në pjesën e trajnimit është minimal. Parametrat e metodës do të përzgjidhen duke përdorur optimizimin Bayesian nga Statistics and Machine Learning Toolbox, ku parametrat vlerësohen duke përdorur vlerësimin kryq me të dhënat që ndahen në 5 copa siç tregohet në tabelën 5.1 ku rezultati do të jetë çifti optimal i parametrave C, ϵ . Eksperimentet përsëriten për të gjitha skenaret e parashikimit si edhe për të gjitha ndërtesat.

Pemët e regresionit

Pemët e Regresionit, implementimi i të cilave i dedikohet funksionit *fitrtree* nga Statistics and Machine Learning Toolbox, MATLAB. Cilësuar si një nga metodat e thjeshta në literaturë për arsyen se modelet e gjeneruara janë të kuptueshëm. Implementimi i kësaj metode kërkon përcaktimin paraprak e parametrave të modelit. Ndërtimi i një peme regresioni do të përshkruhej si në hapat e mëposhtëm: fillimisht të dhënat e trajnimit ndahen duke përdorur teknikën binare (*secila pjesë ndahet më dysh*) duke marrë në konsideratë të gjitha të dhënat ato do të ndahen në copa duke krijuar modelin e një peme derisa nyja fundore të ketë më pak matje se kushti i përcaktuar. Kushti i numrit minimal të vlerave në nyjen fundore përcaktohet duke përdorur optimizimin Bayesian ashtu siç u veprua për dy modelet e sipërpërmendura me vlerësimin kryq me 5 pjesë. Pema e ndërtuar në fazën fillestare “krasitet” duke e optimizuar me vlerësimin kryq bazuar në gabimin MSE (mean squared error). Përcaktimi i hapit optimal i rritjes së pemës si dhe hapi i krasitjes janë funksione të optimizimit Bayesian të ekzekutuara në Statistics and Machine Learning Toolbox MATLAB.

Pyjet e rastësishëm

Pyjet e rastësishëm, janë asamble metodash kryesisht të ndërtuar mbi pemët e regresionit si ‘mësues’ bazë, por mund të përdorin edhe ‘mësues’ të tjerë. Kjo si metodë mbështetet në gjenerimin e të dhënave nga të dhënat kryesore me teknikën e bootstrapping, ku të dhënat origjinale të hyrjes ndahen në kampione me përsëritje. Mbi kampionet e gjeneruara nga bootstrapping modeli mëson dhe parashikon duke përdorur pemët e regresionit duke dhënë rezultatin final që shuma e mesatarizuar e rezultateve të mësuesve të veçantë. Metoda kërkon përcaktimin e numrit të pemëve që do të përdorim si mësues ku zgjedhim numrin 100, mbasi përcaktojmë numrin e pemëve, optimizimi Bayesian aplikohet për të përcaktuar sesa do të rriten pemët në pyll, si edhe numri i parashikuesve të cilët do të kampionohen në secilën nyje të pemës. Vlerësimi kryq duke i ndarë të dhënat në 5 pjesë përdoret gjatë optimizimit.

Arsyetimi i bazuar tek rastet

ABR, është i njëjti model që përdoret në kapitullin 4, por me disa ndryshime në modelimin e të dhënave. Modeli i cili implemetohet për të parashikuar konsumin e energjisë mbështetet tek të dhënat e hyrjes të treguara në ekuacionin (5.4) me ndryshimin e vetëm se kodimi i ditëve të javës bëhet ndryshe (E Hënë me 1 deri në e Dielë me 7). Të dhënat si dhe vlerat minimale edhe maksimale të shmangies së tyre tregohen në tabelën 5.2:

ABR si metodë e bazuar tek zgjidhja e problemit mbi ndërtimin e rasteve si dhe gjetjen e rasteve e ngjashme me to i ndërton rastet duke marrë variablat e hyrjes në ekuacionin (5.4) si përshkruhen në tabelën 5.2. Fillimisht rastet grupohen në raste të ngjashme referuar secilit rast tek i cili do të kryhet parashikimi, grupimi bëhet bazuar në kushtet që temperatura midis rastit konkret dhe rasteve ekzistuese të mos kalojë $2^{\circ}C$ si edhe koha e matjes të mos ndryshojë më shumë se 1 orë. Mbasi rastet janë grupuar në të ngjashme atëherë vlerësohet ngjashmëria midis variablave (5.1) duke përdorur distancën e peshuar Euklidiane.

$$VS'_i(x_i, y_i) = \begin{cases} 0 & \text{if } D_w(x_i, y_i) > d_{max}^i \\ \frac{D_w(x_i, y_i) - d_{min}^i}{d_{max}^i - d_{min}^i} & \text{if } d_{min}^i \leq D_w(x_i, y_i) \leq d_{max}^i \\ 1 & \text{if } D_w(x_i, y_i) < d_{min}^i \end{cases} \quad (5.1)$$

Pasi vlerësohet kjo atëherë llogaritet ngjashmëria sipas formës ku sequenca kohore që merret në konsideratë është ajo e temperaturës së orëve të kaluara deri në orën konkrete,

$$VS''_i(x_i, y_i) = \frac{1}{\pi(n)} \sum_{j=0}^{n-1} \cos^{-1} \left(\frac{\Delta x_j \Delta y_j + 1}{\sqrt{(\Delta x_j)^2 + 1} \sqrt{(\Delta y_j)^2 + 1}} \right) \quad (5.2)$$

Tabela 5.2: Përshkrimi i variablave, vlerat minimale dhe maksimale të shmangies së tyre

	Variabli	d_{min}	d_{max}
V_1	Indeksi i ditës së javës	1	3
V_2	Temperatura 3 orë para parashikimit	2	6
V_3	Temperatura 2 orë para parashikimit 2	2	6
V_4	Temperatura 1 orë para parashikimit	2	6
V_5	Temperatura në orën parashikimit	2	6
V_6	Konsumi i energjisë 24 orë përpara	10	40
V_7	Konsumi i energjisë 1 javë përpara	10	40
V_8	Konsumi mesatar i energjisë 24 orë përpara	10	40
V_9	Konsumi i energjisë 3 orë përpara	10	40
V_{10}	Konsumi i energjisë 2 orë përpara	10	40
V_{11}	Konsumi i energjisë 1 orë përpara	10	40

si edhe ngjashmëria midis rasteve do të vlerësohet si,

$$VS_i(x_i, y_i) = \beta VS'_i(x_i, y_i) + (1 - \beta) VS''_i(x_i, y_i) \quad (5.3)$$

ku β është pesha e ndikimit të secilës ngjashmëri tek rezultati e vlerësuar gjatë periudhës së trajnimit me vlerën 0.6. Peshat e secilës variabël llogariten duke përshtatur një model regresioni linear tek vlerat që dalin nga ngjashmëritë në (5.1) siç veprojmë në 4.2.5.

5.1.3 Parashtrimi i problemit-Modelimi i të dhënave

Të dhënat finale jepen në formën e një serie kohore e cila përmban matje me intervale prej një ore për konsumin e energjisë dhe temperaturën e mjedisit. Modelimi i problemit duke përdorur në mënyrë sa më efektive informacionin që përmbajnë të dhënat, përbën impaktin kryesor në performancën e tij. Në seritë kohore të matura në periudha relativisht të gjata (vjeçare) vërehen femomenet e përsëritjeve, ku tendenca e të dhënave ngjason në periudha të ndryshme, nëse i referohemi të dhënave në fig. 5.1 vërehet se kurbat ndjekin një sjellje pak a shumë përsëritëse brenda javës. Po ashtu në fig. 5.2 vërehet se megjithë ndryshimeve në periudha të caktuara piku dhe minimumi i konsumit nuk tregojnë ndryshime drastike. Duke patur parasysh sjelljet në konsum të sipërpërmendura përpiqemi që nga baza jonë e të dhënave të nxjerrim

informacione shtesë (*variabla*). Koha kur është kryer matja jep informacion duke indeksuar përsëritjen e sjelljeve të njëjta në periudha të ndryshme, ajo kodohet duke përdorur një qasje shumë të thjeshtë (1, ..., 24), po ashtu dita e javës të cilës i përket matja do të ishte shumë informative në parashikim pasi konsumi në ditët e punës nuk është i njëjtë me fudjavat, informacioni i marrë kodohet në mënyrë që të jetë i pajtueshëm me algoritmin (p.sh. e Hënë (1, 0, 0, 0, 0, 0, 0), ..., e Dielë (0, 0, 0, 0, 0, 0, 1)).

Dimë se në kushte normale pune konsumi i energjisë gjatë 24 orëve zakonisht shfaq sjellje periodike. Konsumi nis nga një pikë minimale duke arritur më tej një vlerën maksimale për zbritur përsëri në një vlerë minimale, kjo duket qartë tek lakoret e konsumit javor fig. 5.1 ku kjo analogji përsëritet pothuajse në secilën ditë të javës për 5 ndërtesat. E parë në një këndvështrim tjetër mund të themi se konsumi në orët pasardhëse ndjek tendencën e orëve paraardhëse, pra nuk kemi një ndryshim me hop, por vijueshmërinë e tendencës. Kështu që në modelimin e vektorit të të dhënave përfshijmë konsumin e tre orëve paraardhëse nga çasti i nisjes së parashikimit. Jo vetëm periodicitetet afatshkurtra në konsumin e energjisë mund të evidentohen, por dhe ato pak më afatgjata që mund të përmendim konsumin në të njëjtin çast kohe i ditës paraardhëse, po ashtu konsumi mesatar i ditës para do të shërbente si faktor informues mbi konsumin. Vërejmë se në fig. 5.1 kurbat për 7 ditët e javës shfaqin lakore pak a shumë përsëritëse në sjellje duke ngjasuar midis tyre. Nisur nga kjo në vektorin e modelimit të të dhënave hyrëse për të vlerësuar modelet përfshihet konsumi i një jave më parë i matur në të njëjtën kohë me parashikimin që do të bëjmë. Vektori i variablave të hyrjes për secilin model të *machine learning*, i modeluar nga sa thamë më sipër është,

$$X(t) = \{t, ID, T_{-3}, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \bar{P}_{-24}, P_{-3}, P_{-2}, P_{-1}\} \quad (5.4)$$

ku t nënkupton kohën e matjes, ID indeksi i ditës së javës, T temperatura në të njëjtin çast, konsumi i tre orëve paraardhëse P_{-3}, P_{-2}, P_{-1} , ndërkohë P_{-24} konsumin i energjisë 24 orë përpara, P_{-168} konsumi i një jave më parë dhe \bar{P}_{-24} konsumi mesatar i ditës së mëparshme.

Përzgjedhja e variablave të cilat do të “*ushqejnë*” modelet bëhet duke ju referuar njohurisë së ekspertit. Vlerësimi i rëndësisë së këtyre variablave të përzgjedhura kryhet duke përdorur teknikat statistikore. Përzgjedhja e variablave të përshkrimit të të dhënave njih kryesisht dy teknika, njëra është duke e nisur nga numri minimal i attributeve dhe duke shtuar variabla ose tjetra duke e nisur nga numri maksimal i attributeve dhe duke e zvogëluar sasinë e variablave. Procedura që ne pranojmë është e tillë që vektori (5.4) pranohet

si përshkrimi maksimal i variablave hyrëse, më tej fillojmë heqim variabla dhe vlerësojmë performancën e modelit me vektorin e ri të hyrjes. Statistikisht kjo procedurë bëhet duke vlerësuar kryq modelet e trajnuar me vektorët derivat të (5.4). Zhvendosja e të dhënave nga vektori në ekuacionin (5.4) bëhet duke hequr vlerat e temperaturës dhe konsumit të energjisë për hapat kohorë më të largët nga koha e nisjes së parashikimit, sipas rendit $t = -1$ deri $t = -3$. Nga zhvendosja e këtyre krijohen 12 vektorë hyrjeje të rinj mbi të cilët aplikojmë procedurën e vlerësimit kryq duke e ndarë hapësirën e të dhënave të hyrjes së modelit në 5 pjesë të ndryshme dhe bëjmë vlerësimin referuar gabimit CV-RMSE.

Procedura aplikohet për secilin nga modelet e marrë në konsideratë dhe vektori i cili ka performancë më të lartë do të jetë vektori i hyrjes me të cilin modeli do të mësojë të parashikojë. Në tabelën 5.3 tregohen 12 vektorët e përfutur nga eliminimi i madhësive të hyrjes të zgjedhura nga njohuria e ekspertit. Pasi kemi përfutur 12 vektorët e ndryshëm atëherë aplikojmë modelet e përshkruar në 5.1.2 ku eliminojmë vetëm ABR pasi kjo metodë mbështetet tek ndërtimi i modelit referuar shembujve, por metodat e tjera janë të mjaftueshme për të dhënë konkluzione referuar rezultateve të cilat kërkojmë.

Rezultatet e paraqitura në tabelën 5.3 shprehin vlerën mesatare të 5 pjesëve të vlerësimit kryq. Arrihet në përfundimin se të dhënat e modeluara nga eksperti (5.4) e modelojnë më mirë problemin.

Tabela 5.3: Rezultatet për vektorë të ndryshëm të hyrjeje

	Vektori i hyrjes $X(t)$	CV-RMSE			
		RN	RVN	PR	PR_{ast}
1	$t, ID, T_{-3}, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-2}, P_{-1}$	7.34	7.75	6.07	5.03
2	$t, ID, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-2}, P_{-1}$	7.52	8.99	6.46	5.05
3	$t, ID, T_{-3}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-2}, P_{-1}$	7.43	7.86	6.51	5.11
4	$t, ID, T_{-3}, T_{-2}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-2}, P_{-1}$	7.86	7.77	6.52	5.15
5	$t, ID, T_{-3}, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-2}, P_{-1}$	7.78	7.81	6.18	5.14
6	$t, ID, T_{-3}, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-1}$	7.52	7.77	6.50	5.23
7	$t, ID, T_{-3}, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-2}$	7.76	8.11	6.95	5.17
8	$t, ID, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-2}, P_{-1}$	7.56	7.86	6.55	5.14
9	$t, ID, T_{-3}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-1}$	7.95	7.87	6.60	5.19
10	$t, ID, T_{-3}, T_{-2}, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-2}$	7.66	8.01	6.96	5.13
11	$t, ID, T, P_{-24}, P_{-168}, \overline{P}_{-24}, P_{-3}, P_{-2}, P_{-1}$	7.8	8.24	6.43	5.18
12	$t, ID, T_{-3}, T_{-2}, T_{-1}, T, P_{-24}, P_{-168}, \overline{P}_{-24},$	7.96	8.14	6.93	5.50

Më sipër përmendëm vetëm modelimin e vektorit të hyrjes, por siç kemi thënë teknika e mësimit të supervizuar mëson një sjellje apo model duke u trajnuar në çiftin (*hyrje, dalje*). Ndërsa hyrja u përcaktua si (5.4), si dalje do të përdorim konsumin e energjisë që përkon me kohën e parashikimit P_t . Secili nga modelet do të trajnohet sipas $(\mathbf{X}(t), P_t)$.

Skematika kryesore e parashikimit të konsumit të energjisë tregohet në fig. 5.3. Të dhënat e një viti ndahen në të dhënat e trajnimit dhe të testimit. Pjesa e trajnimit përdoret fillimisht për të optimizuar modelet, ku duke përdorur teknikën e vlerësimit kryq me 5 pjesë, RN, RVN, PR, PR_{ast} dhe ABR gjejnë vlerat optimale për parametrat e brendshëm tv secilit model. Modelet optimale trajnohen në të dhënat e trajnimit duke përdorur $(\mathbf{X}(t), P_t)$, modeli i mësuar në këtë fazë përdoret për të parashikuar vlerat e konsumit të energjisë për kampionet e testimit duke marrë vetëm hyrjen \mathbf{X}_{test} dhe jep një dalje sipas asaj që ai kishte mësuar. Vlerat e parashikuara krahasohen me vlerat reale nga pjesa e testimit të cilat përdoren vetëm për vlerësim, duke dhënë kështu edhe performancën e modelit.

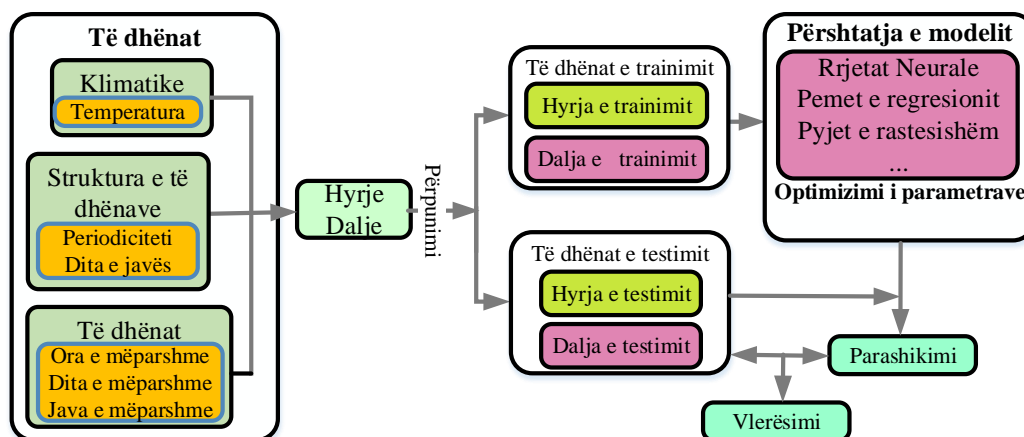


Figura 5.3: Skematika e parashikimit të konsumit të energjisë

5.1.4 Rezultatet eksperimentale

Vlerësimi i performancës së modeleve të inteligjencës artificiale bëhet duke ju referuar modeleve të implementuara siç përshkruhen me detaje në pjesën 5.1.2. Secili nga modelet optimizohet duke përcaktuar parametrat e tij për secilin skenar që do të përshkruajmë më poshtë. Modelet trajnohen dhe testohen në tre skenarë të ndryshëm: i pari që përdor 9 muaj si të dhëna trajnimi dhe seksionin e mbetur prej tre muajsh përdoret për testim, i dyti përdor 10 muaj për trajnim

kurse pjesa tjetër për testim si dhe skenari i tretë përdor 11 muaj për trajnim dhe 1 muaj testim. E thënë më qarte:

Periudha trajnim/testim

1. Trajnim: Janar-Shtator, Tesimim: Tetor-Dhjetor
2. Trajnim: Janar-Tetor, Tesimim: Nëntor-Dhjetor
3. Trajnim: Janar-Nëntor, Tesimim: Dhjetor

Modelet krahasohen midis tyre në të tre skenarët e sipërpërmendur duke përdorur tre modele vlerësimi (4.2), (4.3) dhe (4.4). Po ashtu modelet krahasohen me një model parashikimi bazë i cili parashikon mesataren e konsumit të energjisë për periudhën e trajnimit.

$$\hat{P}_t = \frac{1}{T} \sum_{i=1}^T P_i$$

Modelet të cilët trajnohen me të dhëna të modeluara njësoj sipas (5.4) për secilën nga ndërtesat, parashikojnë konsumin për energjinë për orën e ardhshme. Përshkrime mbi të dhënat e ndërtesave dhe rezultatet e përfutuara nga 6 modelet i japim të detajuara më poshtë.

Ndërtesa 1, është një ndërtesë publike tek e cila janë matur konsumi i energjisë dhe temperatura e mjedisit jashtë siç përshkruhet në pjesën 5.1.1, në intervalet e 15 min dhe më tej ato u transformuan në të dhëna në intervale ore. Një analizë e thjeshtë mbi të dhënat, tregon se statistikat për temperaturën janë: ajo arrin vlerën minimale $T_1(min) = 0.83^\circ C$, ndërsa vlera maksimale është $T_1(max) = 37.22^\circ C$ si dhe deviacioni standart mbi të gjitha matjet është $\sigma_{T_1} = 4.82^\circ C$, ndërsa konsumi minimal i energjisë prek vlerën $P_1(min) = 30.85(kW)$ si dhe maksimumin $P_1(max) = 414.3(kW)$ si dhe deviacioni standart $\sigma_{P_1} = 63.28(kW)$. Në treguesit statistikor të të dhënave treguam vetëm për temperaturën dhe konsumin pasi këto janë madhësitë nga ku rrjedhin edhe variablat e tjera. Deviacioni standart përdoret si vlerësues statisikor tek të dhënat pasi ai tregon se sa është shmangia e matjeve nga mesatarja e gjithë matjeve dhe llogaritet si:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Rezultatet e treguara në Tabelën 5.4 shprehin mbizotërimin e metodave të inteligjencës artificiale kundrejt parashikuesit bazë, po ashtu rezultatet e përfutuara ndodhen brenda kufijve shumë të kënaqshme referuar udhëzimeve

të ASHRAE [30] për të dhëna të kalibruara sipas orës CVRMSE duhet të jetë brenda 30%. Nëse krahasojmë metodat e inteligjencës artificiale vërejmë se pyjet e rastësishëm mbizotërojnë mbi të tjerat kjo referuar faktit që kjo është një model i bashkësive pra kombinon disa “mësues” që të japë rezultatin e dëshiruar, ndryshe nga “mësuesit” e vetëm që nuk alternojnë mënyra mësimi. Nëse krahasojmë metodat e tjera IA si të bazuara tek modeli vërehet një epërsi e lehtë e RN-ve në skenarët 1 dhe 2, ndjekur nga RVN dhe më tej nga PR. Kurse në skenarin e 3 kemi një mbizotërim të PR të ndjekura nga RN dhe RVN. Vërehet se ABR si metodë e bazuar tek shembujt tregon një disavantazh të lehtë krahasuar me metodat e bazuara tek modeli, por duke treguar epërsi kundrejt modelit bazë dhe kufirit të ASHRAE.

Tabela 5.4: Rezultatet e parashikimit për ndërtesën e parë.

Trajnimi	Metoda	MAE (kW)	MAPE (%)	CVRMSE (%)
Jan-Sht	Bazë	53.8710	23.3273	23.2763
	RN	11.3662	4.9127	6.7285
	PR	11.4791	4.8463	7.8878
	PR_{ast}	7.8177	3.3990	5.3535
	RVN	10.6844	4.6749	6.8907
	ABR	13.2772	5.4895	7.6383
Jan-Tet	Bazë	51.8409	23.7835	22.7315
	RN	10.6910	4.8673	6.8724
	PR	9.8470	4.5837	7.3572
	PR_{ast}	7.2681	3.3680	5.1373
	RVN	10.5786	4.8473	7.2947
	ABR	13.3049	5.6971	7.9501
Jan-Nën	Bazë	51.0902	23.7244	22.5692
	RN	11.7009	5.3404	7.1116
	PR	8.9506	3.9462	6.0569
	PR_{ast}	7.4572	3.5197	5.3435
	RVN	10.9098	5.0623	7.2026
	ABR	14.3000	6.2229	8.9893

Ndërtesa 2, është një ndërtesë publike tek e cila janë matur po ashtu konsumi i energjisë dhe temperatura e mjedisit të jashtëm siç përshkruhet në pjesën 5.1.1. Statistikat për temperaturën janë: vlera minimale e temperaturës është $T_2(min) = 4.44^\circ C$, kundrejt vlerës maksimale që është $T_2(max) = 38.75^\circ C$ si dhe deviacioni standart i temperaturës është $\sigma_{T_2} = 3.70^\circ C$, ndërsa konsumi

minimal i energjisë prek vlerën $P_2(min) = 67.50(kW)$ si dhe maksimumin $P_2(max) = 497.15(kW)$ si dhe deviacioni standart $\sigma_{P_2} = 79.52(kW)$.

Tabela 5.5: Rezultatet e parashikimit për ndërtesën e dytë.

Trajnimi	Metoda	MAE (kW)	MAPE (%)	CVRMSE (%)
Jan-Sht	Bazë	63.6917	29.6577	26.5306
	RN	13.6186	5.9058	7.4721
	PR	11.2500	4.6646	6.2570
	PR_{ast}	9.2264	3.9928	5.5695
	RVN	12.6485	5.4916	7.5802
	ABR	17.0270	7.2876	8.0872
Jan-Tet	Bazë	60.0290	27.9801	25.4045
	RN	14.1939	6.3771	8.1451
	PR	11.3278	4.9746	7.2188
	PR_{ast}	9.5929	4.3077	6.4485
	RVN	12.9374	5.7114	8.3559
	ABR	17.3063	7.5831	8.4175
Jan-Nën	Bazë	56.7514	27.2829	24.6711
	RN	13.8927	6.1982	8.0955
	PR	9.9983	4.2529	6.6793
	PR_{ast}	8.2273	3.6046	5.5033
	RVN	12.2377	5.4645	8.0990
	ABR	17.4043	7.7106	8.7804

Rezultatet e treguara në tabelën 5.5 përforcojnë përsëri idenë se PR_{ast} mbizotëron kundrejt metodave të tjera. Nëse rezultatet krahasohen me ato të përfituara në tabelën 5.4 shikohet një përkeqësim i lehtë i performancës, kjo i atribuohet edhe natyrës së të dhënave e cila duket qartë nga fig. 2.3c se luhatjet janë më të larta se tek ndërtesa 1.

Ndërsa në metodat e tjera vërehet një fitore e lehtë e PR në tre skenarët. Tre të tjerat si RN, RVN dhe ABR konkurojnë më afër njëra tjetrës. Metoda e bazuar tek shembujt e ngushton diferencën në performancë kundrejt metodave të bazuara tek modeli për faktin se evidenton më lehtë rastet e ngjashme në një situatë të tillë të dhënash.

Ndërtesa 3, një ndërtesë publike me të dhëna të matura për konsumin e energjisë dhe temperaturën e mjedisit jashtë, njësoj si tek ndërtesat e para. Nga një vëzhgim i thjeshtë mbi të dhënat, statistikat për temperaturën minimale janë $T_3(min) = -8.61^\circ C$, vlera maksimale është $T_3(max) = 38.89^\circ C$ si dhe

deviacioni standart i temperaturës është $\sigma_{T_3} = 10.70^\circ C$, ndërsa konsumi minimal i energjisë prek vlerën $P_3(min) = 32.50(kW)$ si dhe maksimumin $P_3(max) = 514.10(kW)$ si dhe deviacioni standart $\sigma_{P_3} = 78.56(kW)$.

Tabela 5.6: Rezultatet e parashikimit për ndërtesën e tretë.

Trajnimi	Metoda	MAE (kW)	MAPE (%)	CVRMSE (%)
Jan-Sht	Bazë	43.4609	21.9641	21.2198
	RN	12.1566	5.2858	6.5694
	PR	10.7120	4.8456	6.5016
	PR_{ast}	8.9823	3.9704	5.6926
	RVN	10.4174	4.5750	6.0233
	ABR	19.5186	8.3293	11.7180
Jan-Tet	Bazë	40.3865	20.9064	20.5848
	RN	13.2517	5.8228	7.2477
	PR	9.3390	4.0500	5.5672
	PR_{ast}	8.9556	4.0305	5.4267
	RVN	10.7592	4.8258	6.3766
	ABR	22.7266	9.6563	13.8474
Jan-Nën	Bazë	34.1477	16.9155	17.9709
	RN	12.2150	5.4207	6.3922
	PR	7.9593	3.6186	5.0809
	PR_{ast}	6.5739	3.0056	4.1642
	RVN	9.8168	4.3740	5.4521
	ABR	27.4874	11.0036	16.7335

Qëndrueshmëria e PR_{ast} për të performuar mbi metodat e tjera vërehet edhe në tabelën 5.6 kjo referuar arsyeve të sipërpërmendura për ndërtesat e tjera. Ndërsa “mësuesit” e bazuar tek modeli si RVN mbizotërojnë lehtësisht PR dhe RN sipas skenarit të parë. Kurse në skenarët 2 dhe 3 kemi performancë më të lartë të PR kundrejt RVN, por të dyja e mundin RN në tre skenarët.

Vërejmë se metoda e bazuar tek shembujt ABR tregon një dobësi për të mësuar tek ndërtesa e tretë kjo referuar natyrës së të dhënave fig. 5.2c, pasi ka një ndryshim midis pjesës së rasteve të trajnimit dhe testimit. Ky një problem i tejkaluar nga metodat e bazuara tek modelet.

Ndërtesa 4, është një ndërtesë publike tek e cila janë matur konsumi i energjisë si dhe temperatura e mjedisit jashtë siç përshkruhet në pjesën 5.1.1. Statistikat për temperaturën janë: temperatura minimale arrin vlerat $T_4(min) =$

$-16.67^{\circ}C$, vlera maksimale është $T_4(max) = 33.06^{\circ}C$ si dhe deviacioni standart i temperaturës është $\sigma_{T_4} = 11.68^{\circ}C$, ndërsa statistikat e konsumit janë: konsumi minimal i energjisë prek vlerën $P_4(min) = 36(kW)$ si dhe maksimumin $P_4(max) = 36(kW)$ si dhe deviacioni standart $\sigma_{P_4} = 77.91(kW)$.

Tabela 5.7: Rezultatet e parashikimit për ndërtesën e katërt.

Trajnimi	Metoda	MAE (kW)	MAPE (%)	CVRMSE (%)
Jan-Sht	Bazë	54.1087	29.7308	25.1317
	RN	13.7841	6.5871	8.8668
	PR	14.8400	7.0492	10.5728
	PR_{ast}	10.9150	5.3548	8.1802
	RVN	11.9983	5.8637	8.3335
	ABR	17.7155	8.2016	10.2558
Jan-Tet	Bazë	54.4976	30.7627	25.7734
	RN	14.4014	7.0326	9.3381
	PR	12.7455	6.5705	11.1755
	PR_{ast}	10.2348	5.1790	8.3571
	RVN	11.7421	5.7836	8.4584
	ABR	16.0949	7.4787	9.8532
Jan-Nën	Bazë	49.6963	27.5676	24.1783
	RN	14.3206	7.1435	9.3210
	PR	10.9240	5.3988	8.0715
	PR_{ast}	10.2224	5.3143	7.8764
	RVN	12.1025	6.0173	8.6716
	ABR	16.0636	7.4144	10.2128

Rezultatet e treguara në tabelën 5.7, japin tablonë se për ndërtesën e katërt performanca e modeleve shënon rënie krahasuar me performancën e tyre tre godinat e analizuar më sipër. Nëse i referohemi fig. 5.2d vërehet se konsumi ka tendencë rritëse nga janari deri afër fillimit të testimit, ku modeli i përafruar e shikon si tendencë rritëse në konsum. Por jo vetëm ky faktor, por edhe nëse analizojmë statistikat vërehet se temperatura arrin vlerat minimale të saj krahasuar me ndërtesat e tjera, po ashtu varianca e temperaturës është më e madhe se tek të tjerat.

PR_{ast} përsëri fiton mbi të tjerat e ndjekur ngushtë nga RVN në dy skenarët e parë. Në skenarin e tretë modelet e bazuar tek pemët si PR dhe PR_{ast} fitojnë mbi të tjerat. Metodat e bazuara tek modeli si rasti i PR humbasin lehtësisht kundrejt atij të bazuar në shembujtnë dy skenarët e parë.

Ndërtesa 5, është një ndërtesë publike tek e cila janë matur konsumi i energjisë si edhe temperatura e mjedisit jashtë siç përshkruhet në seksionin 5.1.1. Nga një vëzhgim i thjeshtë mbi të dhënat, statistikatat për temperaturën minimale janë $T_5(min) = 0.697^\circ C$, vlera maksimale është $T_5(max) = 34.44^\circ C$ si dhe deviacioni standart mbi të gjitha matjet është $\sigma_{T_5} = 6.21^\circ C$, ndërsa konsumi minimal i energjisë prek vlerën $P_5(min) = 108.47(kW)$ si dhe maksimumin $P_5(max) = 510.95(kW)$ si dhe deviacioni standart $\sigma_{P_5} = 93.53(kW)$.

Tabela 5.8: Rezultatet e parashikimit për ndërtesë e pestë.

Trajnimi	Metoda	MAE (kW)	MAPE (%)	CVRMSE (%)
Jan-Sht	Bazë	70.4309	31.5118	26.8713
	RN	17.5916	6.3740	7.6603
	PR	17.2060	6.3944	7.8752
	PR_{ast}	9.2393	4.0040	5.6008
	RVN	16.1717	5.9304	7.6017
	ABR	19.8551	7.3483	8.3731
Jan-Tet	Bazë	78.6203	36.1097	28.3413
	RN	17.2968	6.5624	8.2346
	PR	18.4342	7.2246	9.1007
	PR_{ast}	9.6276	4.3283	6.4794
	RVN	17.5431	6.6517	8.8892
	ABR	20.0047	7.7199	9.1831
Jan-Nën	Bazë	80.0767	39.7075	29.6764
	RN	17.4454	6.9446	8.4571
	PR	16.0865	6.7686	8.9638
	PR_{ast}	7.9421	3.4585	5.3119
	RVN	16.9802	6.8207	8.9984
	ABR	20.6197	8.3597	10.3337

Përsëri rezultatet në tabelën 5.8 shprehin mbizotërimin e metodave të inteligjencës artificiale krahasuar me mësuesin bazë. Performanca e PR_{ast} është fituese mbi të tjerat. Në skenarin 1 RVN tregon një epërsi të lehtë kundrejt RN, këto të ndjekura nga PR dhe ABR. Në skenarin e 2 dhe 3 RN ka epërsi kundrejt RVN dhe PR të ndjekura nga ABR.

Të gjitha rezultatet e paraqitura më sipër janë mesatarizim i rezultateteve të përfutuara nga ekzekutimi 10 herë i secilit model RN, RVN, PR, PR_{ast} pasi këto janë modele stokastike pra japin rezultate të ngjashme sa herë i ekzekutojmë, ndërsa ABR vetëm 1 herë pasi rezultati i dhënë nga ky model është deterministik.

Po ashtu një paraqitje grafike e rezultateve të parashikuara për periudhën 1 javore për secilën ndërtesë tregohet në fig. 5.8, kurse rezultatet e parashikuar për gjithë skenarët e parashikimit me metodën e pyjeve të rastësishëm si modeli me performancën më të lartë e tregojmë në shesën B.

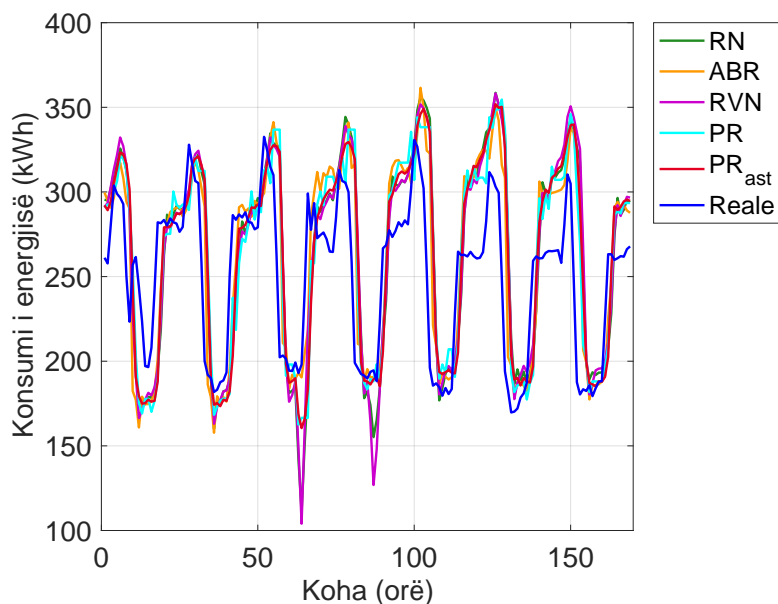


Figura 5.4: Konsumi i parashikuar i energjisë në ndërtesën e parë për një javë.

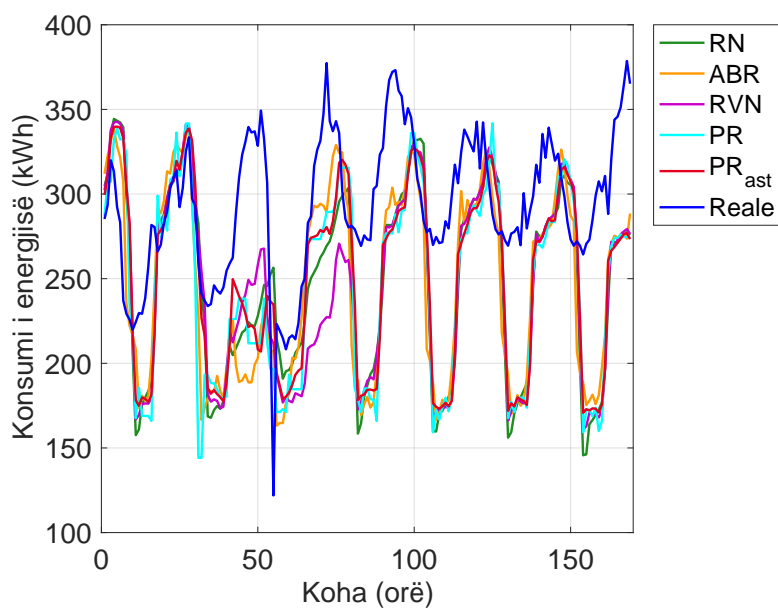


Figura 5.5: Konsumi i parashikuar i energjisë në ndërtesën e dytë për një javë.

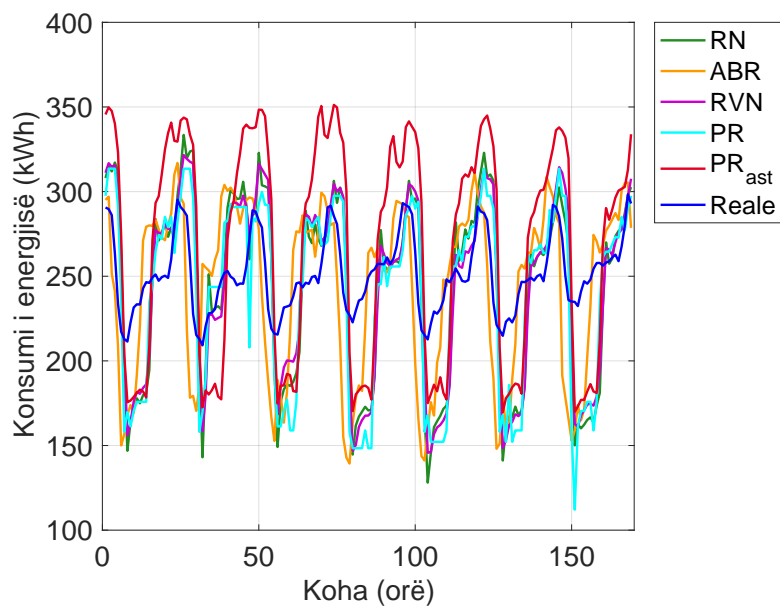


Figura 5.6: Konsumi i parashikuar i energjisë në ndërtësën e tretë për një javë.

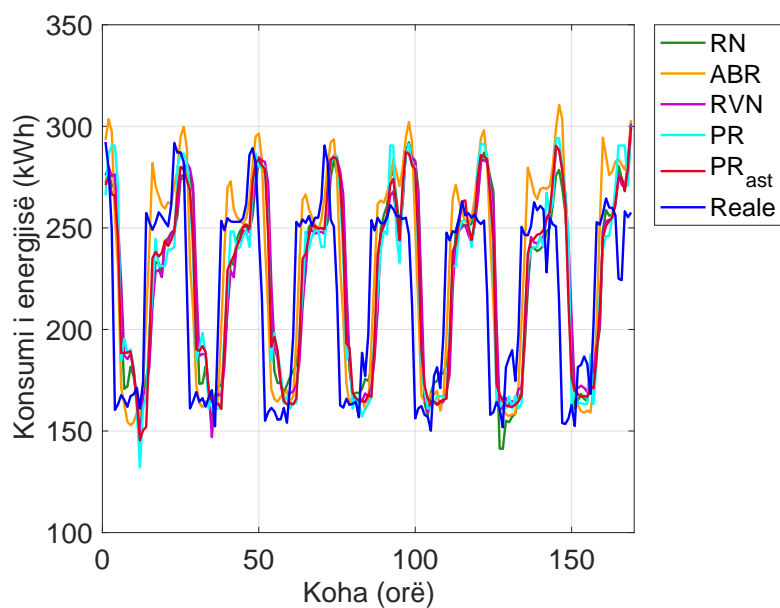


Figura 5.7: Konsumi i parashikuar i energjisë në ndërtësën e katërt për një javë.

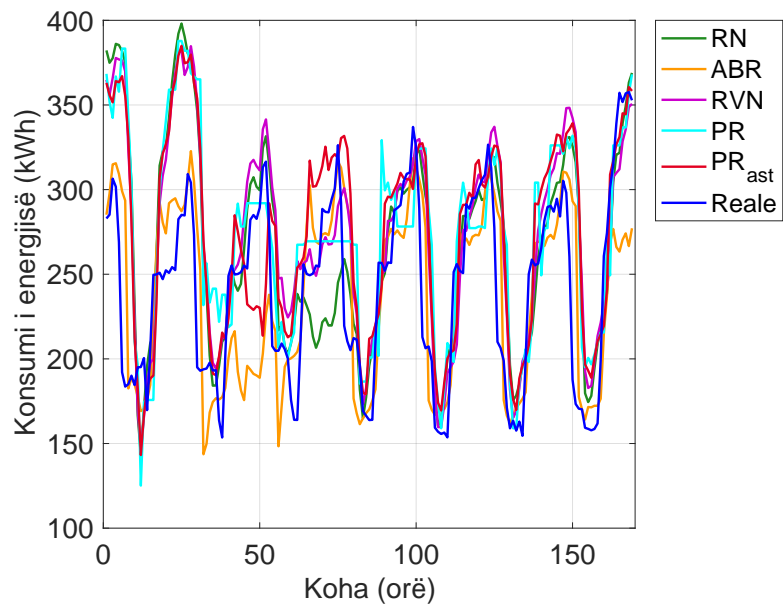


Figura 5.8: Konsumi i parashikuar i energjisë në ndërtësën e pestë për një javë.

5.2 Parashikimi dinamik 'online'

Parashikimi i konsumit online i referohet asaj dege të inteligjencës artificiale ku modeli parashikon sipas ushqimit me të dhëna në mënyrë sekuenciale. Në ndryshim nga metoda statike e të mësuarit ku modeli mëson mbi të gjithë setin e të dhënave, në skenarin dinamik të dhënat vijnë në mënyrë inkrementale dhe modeli mëson mbi një pjesë të tyre duke përditësuar dhe rritur performancën e tij me rritjen e bazës së të dhënave. Këto teknika janë tepër të përdorshme në rastet kur përdoruesi nuk është i interesuar të përdori gjithë setin e të dhënave ose kur problematika që interesohemi ofron grumbullim të dhënash në kohë reale apo modeli duhet të përditësohet me të dhënat e reja që bëhen të përdorshme. Problemi ynë përkon me tematikën e parashikimit dinamik pasi monitorimi dhe matja e të dhënave bëhet në kohë reale që do të ishte e vlefshme për operatorët e energjisë si funksion i parashikimit të konsumit. Në seksionet e mëposhtme do paraqesim disa nga teknikat të përdorura për parashikimin online të të dhënave në kohë reale si dhe do të adaptojmë këto teknika në zgjidhjen e problemit tonë.

5.2.1 Rregulli i modelit adaptiv

Modelet adaptive së bashku me dy modelet e tjera të zgjedhura si metoda parashikimi njihen në literaturë si metoda të regresionit që mësojnë në 'rrymat' e të dhënave. Më poshtë përshkruajmë disa nga çështjet e lidhura me funksionimin e algoritmit në parashikim ku pseudokodi i të cilit jepet në Algoritmin 1.

Me ardhjen e një shembulli të ri algoritmi fillon procesin e mësimit. Ky proces lidhet me vendosjen e një grupi rregullash secilit shembull. Rregulli fillestar është bosh dhe më tej ai përditësohet. Sa herë që një shembull trajnimi bëhet i vlefshëm për analizë, algoritmi kontrollon nëse shembulli mbulohet nga ndonjë rregull, duke analizuar nëse gjithë kushtet e vendosura për *atributet* që modelojnë shembullin qëndrojnë. Nëse shembulli plotëson kushtet atëherë vlerat e daljes (target) përdoren për të përditësuar statistikën e rregullit.

Duke qënë se të dhënat vijnë në mënyrë inkrementale (hap pas hapi) atëherë për secilin shembull i cili merret në konsideratë analizohet grupi i rregullave i cili e mbulon një shembull, në rast se shembulli nuk mbulohet nga asnjë rregull atëherë rregulli i paracaktuar (default) i bashkangjitet këtij shembulli. Rregulli nis me një vlerë të paracaktuar, por ato përditësohen duke u zgjeruar me rritjen e shembujve të trajnimit. Zgjerimi i rregullit bëhet duke marrë në konsideratë secilën nga variablat e shembullit X_i dhe për secilën vlerë të saj v_j llogarisimi diferencën e vlerës me mesataren e gjithë atributit, nëse kemi ndryshim brenda një rregulli të vendosur atëherë rregulli zgjerohet në të kundërt jo.

Procesi i trajnimit përfshin mësimin e një grupi rregullash nga pjesa e trajnimit. Rregullat e mësuara do të përdoren për të parashikuar vlerat optimale. Ekzistojnë dy teknika të parashikimit, rregulli i mësuar është i renditur ose i parenditur. Në rastin e parë, vetëm rregulli fillestar merret parasysh në parashikim ndërsa në të dytin pranohet mesatarja e të gjithë rregullave. Secili nga rregullat në metodë përshkruhet nga një model jolinear i cili është trajnuar duke përdorur metodën inkrementale të gradientit rënës. Fillimisht peshat që i ngarkohen modelit kanë vlera të vogla në intervalin nga -1 në 1 .

Përpara se një shembull të përshkruhet nga ndonjë rregull zbulimi i ndonjë ndryshimi testohet duke përdorur testin PH (Page-Hinckley) i cili monitoron gabimin online të secilit rregull. Nëse një ndryshim ekziston midis shembullit dhe rasteve në setin e rregullave atëherë rregulli që i atribuohet shembullit të ri do të eliminohet nga seti i rregullave ose në të kundërt seti do të zgjerohet. Zgjerimi i setit të rregullave analizohet kur plotësohet një kusht si numri minimal i shembujve të jetë sa N_{min} .

Njësoj edhe metodat e tjera vijuese përdorin testin Page-Hinckley (PH) për të monitoruar gabimin online të çdo rregulli i cili aplikohet mbi shembullin e ri apo metode e cila bën klasifikimin e një shembulli të ri nën një rregull të caktuar. Test përkon me analizën sekuenciale të vlerësimit të shembujve. Ky test është akumulativ përgjatë gjithë kohës, dhe ai konsideron një variabël akumuluese m_T e cila jepet si diferenca e akumuluar midis vlerave të matura dhe mesatares të gjithë atyre deri në momentin e caktuar të kohës:

$$m_T = \sum_{t=1}^T (x_t - \bar{x}_T - \delta)$$

ku $\bar{x}_T = \frac{1}{T} \sum_{t=1}^T x_t$ si dhe δ është amplituda e ndryshimeve të lejuara. Vlera minimale e variablit m_T llogaritet si: $M_T = \min(m_T, t = 1, \dots, T)$ dhe testi monitoron nëse ekziston ndonjë diferencë midis M_T dhe m_t më e madhe se një vlerë shmangieje (λ) dhe atëherë gjenerohet një sinjal dedektues për një ndryshim domethënës midis shembujve. Vlera e λ zgjidhet nga përdoruesi e tillë që të eliminohen alarmet false, por vlera e saj duhet të mos jetë e madhe sa të neglizhohen ndryshimet në sinjal.

Modeli realizohet eksperimentalisht duke e trajnuar në dy situata:

1. Përditësohet vetëm rregulli i parë i cili mbulon shembujt e trajnimit. Në rast se grupi i rregullave është i renditura, strategjia e parashikimit përdor vetëm rregullin e parë që mbulon shembujt e testimit.
2. Përditësohen të gjithë rregullat që mbulojnë shembujt e trajnimit. Në këtë rast grupi i rregullave është i parenditur, dhe strategjia

e parashikimit përdor shumë e mesatarizuar të gjithë rregullave që mbulojnë shembullin e testimit.

Algoritmi 1 Modeli Adaptiv **AMRules** [175]

```

1: Hyrjet:
   S: Numri i shembujve
    $N_{min}$ : Numri minimal i shembujve,  $\lambda$ : Shmangia
    $\alpha$ : Vlera e ndryshimeve të cilat lejohen
2: Fillo:
    $Nese\ RS \leftarrow \{\}$ 
    $Nese\ Rregullidefault\ L \leftarrow \{\}$ 
3: for  $\forall$  shembull  $(\mathbf{x}, y_k) \in S$  do
4:   for  $\forall$  rregull  $r \in RS$  do
5:     if  $r$  mbulon shembullin then
6:       Përditëso testin e dedektimit të ndryshimit
7:       Llogarit gabimin  $= x_t - \bar{x}_t - \alpha$ 
8:       Thirr Testin PH (gabimin,  $\lambda$ )
9:       if Dedektohet ndryshim then
10:        Hiqe rregullin;
11:      else
12:        if Numri i shembujve në  $\mathcal{L}_r \geq N_{min}$  then
13:           $r \leftarrow$  zgjero rregullin( $r$ )
14:          Përditëso statistikën e mjaftueshme për  $r$ 
15:        end if
16:        if seti i renditur then
17:          SHKËPUT
18:        end if
19:      end if
20:    end if
21:  end for
22:  if asnjë nga rregullat në RS nuk plotësohet then
23:    if Numri i shembujve në L është sa  $N_{min} = 0$  then
24:       $RS \leftarrow RS \cup RregulliZgjerimit$ 
25:      Përditëso statistikën e mjaftueshme për rregullin default
26:    end if
27:  end if
28: end for

```

5.2.2 Modeli inkremental i pemëve

Algoritmi i pemëve inkrementale ose ndryshe siç njihet (Fast Incremental Model Trees with Drift Detection) FIMT-DD mbështetet në modelin e pemëve të regresionit, por tani mësimi është dinamik. Kjo nënkupton se modeli ndryshon herë pas here me rritjen e të dhënave. Duke qënë se të dhënat e njëpasnjëshme nuk kanë ndryshime rrënjësore midis tyre atëherë modeli ushqehet me pjesë nga të dhënat, ku nënkuptojmë se marrim nja grup shembujsh dhe jo një shembull të vetëm. Me ardhjen e një pjese shembujsh algoritmi kërkon të gjejë ndarjen optimale për secilën nga atributet (referohemi tek 2.2.5). Gjetja e ndarjes optimale për secilin atribut kalon në fazën tjetër atë të rankimit të attributeve në varësi të një madhësie vlerësimi.

Fillimisht algoritmi nis me një gjethe boshe, me ardhjen e një grupi shembujsh kërkohet gjetja e ndarjes më të favorshme për secilën atribut. Pasi u gjet kushti i ndarjes për shembujt e parë, grupi tjetër i shembujve që vjen do të kalojnë poshtë në degët e pemës që u krijuan nga ndarja e parë bazuar tek kriteri fillestar i ndarjes. Secili nga shembujt e rinj do të grupohet në njërin nga grupet e krijuar me të cilët kanë ngjashmërinë më të lartë.

Në rastin kur atributet dhe vlerat e parashikuara duhet të jenë numerike kriteri i ndarjes që aplikohet është i tillë: në gjethet e pemës së regresionit për secilën atribut bashkangjitet një pemë binare. Pema binare përdoret për të mbajtur statistikën të cilat na ndihmojnë të rendisim vlerat në kohë reale. Secila nyje është një strukturë që përmban dy vektorë me nga tre elementë. Strukturat e pemës binare përditësohen në mënyrë inkrementale, duke marrë në konsideratë secilin shembull i cili ndodhet në gjethen e pemës së regresionit ku pema binare është ruajtur.

Parashikimi përfaqësohet nga tre teknika adaptimi i modelit të pemëve.

1. Strategjia e parë e përshtatjes përfshin në ndarje të mëtejshme të strukturës së vjetër duke krijuar nënpemë të reja. Megjithëse parashikimet mund të përmirësohen nga kjo teknikë, struktura e pemës mund të jetë e gabuar ose konfuze.
2. Strategjia e dytë përfshin krasitjen e pemës në ato pika ku dedektohet një ndryshim. Problemi që lind është nëse dedektimi ndodh herët atëherë nënpema e gjeneruar zotëron një numër tepër të vogël gjethesh, dhe do të sillte një performancë tepër të ulët të modelit.
3. Strategjia e tretë kërkon të eliminojë problemin që lindi tek e dyta, ajo propozon se në gjethen tek e cila dedektohet ndryshimi të nis lindja e një nënpeme të re. Procesi i ndërtimit të pemëve ndodh paralelisht deri

në momentin që nënpema e re arrin performancë më të lartë se pema mëmë. Nyjet e nënpemës së re të krijuar nuk do të analizojnë ndonjë dedektim ndryshimi deri në momentin që ajo merr vetitë e pemës mëmë.

Algoritmi 2 Modeli inkremental i pemëve FIMT-DD [172]

```

1: Hyrjet:
    $N_{min}$ : Numri minimal i shembujve
    $\delta$ : Probabilitet i përcaktuar nga përdoruesi
2: Fillo:
   Merr një gjethe boshe (RRËNJË)
3: for  $\forall$  shembull tek të dhënat do
4:   Lexo shembullin e radhës
5:   Transformo shembullin nëpërmjet pemës në Gjethe
6:   Përditëso testin në rrugën e ndjekur
7:   if Dedektohet ndryshim then
8:     Adapto modelin e pemës;
9:   else
10:    Përditëso statistikën në Gjethe
11:     $\forall N_{min}$  shembuj të parë në Gjethe
12:    Gjej ndarjen më të mirë për tipar (attribute)
13:    Rendit tiparet duke përdorur të njëjtën teknikë vlerësimi
14:    if Kriteri i ndarjes plotësohet then
15:      Bëni ndarjen në atributin më të mirë
16:      Krijohen dy degë të reja që çojnë në Gjethe boshe
17:    end if
18:  end if
19: end for

```

5.2.3 Metodatat e bazuara tek shembujt me të dhëna në kohë reale

Metodat e bazuara tek shembujt ose ndryshe IBLStreams (Instance Based Learning on Streams) mbështeten në parashikimin e vlerave duke e paraqitur problemin me anën e rasteve. Kur një shembull i ri $\langle x_0, y_0 \rangle$ vjen në mënyrë inkrementale, fillimisht ai përditësohet në bazën e rasteve. Me rritjen e numrit të shembujve në bazën e rasteve shembujt e rinj që vijnë analizohen. Analiza i referohet grupimit të rasteve në të ngjashëm midis tyre. Pra atributet e shembullit të ri $\langle x_0, y_0 \rangle$ analizohen se a kanë ngjashmëri me shembujt të cilët janë në bazë. Në rastin kur ata ngjasojnë rasti i ri grupohet në një nëngrup C

brenda bazës së rasteve.

Por si zgjidhen shembujt e ngjashëm me shembullin e ri? Ata zgjidhen duke përdorur teknikën e fqinjët më të afërt e cila analizon distancën Euklidiane midis secilës vlerë të attributeve dhe një pjesë nga të gjithë kandidatët me distancën më të vogël klasifikohen si të ngjashëm me rastin e ri. Numri i rasteve në nëngrupin C është i paracaktuar, kështu që lind nevoja që në rastin e përditësimit të bazës së rasteve me një të ri na duhet të zhvendosim rastin më pak të ngjashëm në bazë. Në analizë nuk merren shembujt e futur rishtaz për t'i dhënë kohë përshtatje në problem. Rastet e vjetër me pak të ngjashëm eliminohen nga baza e rasteve kur një rast i ri vjen.

Një nga problemet që shtrohet në implementimin online të modelit është edhe adaptimi i madhësisë së bazës së rasteve të ngjashëm k . Në konkluzion të kësaj kemi analizën nëse duhet të rritet baza e rasteve me 1 apo të zvogëlohet. Për të dhënë këtë vendim analizojmë gabimin e 100 rasteve të fundit, duke provuar gabimin në skenarët me 101 dhe 99 raste. Nëse ndonjë nga këto skenarë performon më mirë atë herë merret vendimi i ndryshimit të vlerës së k .

Në rastin e regresionit kandidati i cili duhet të eliminohet nga baza e rasteve i nështrohet rregullit të mëposhtëm: kandidatët e rinj k_{kand} nga nënseti C zgjidhen për të përcaktuar një interval besueshmërie $\left[\bar{y} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{k_{kand}}}, \bar{y} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{k_{kand}}} \right]$, ku \bar{y} është vlera mesatare në dalje për shembujt e marrë në konsideratë dhe σ është deviacioni standart dhe α është një vlerë rëndësie e cila zgjidhet nga përdoruesi.

Algoritmi funksionon i tillë që shembujt të cilët kanë vlerat e daljes së shembullit jashtë segmentit të caktuar në relacionin më sipër do të eliminohen nga skema ose do të përfshihen në bazën e rasteve. Pasi përditësohet baza e rasteve, rastet më domethënës referuar rastit të ri përdoren për të kryer parashikimin ku rezultati i parshikimit do të jetë kombinimi rezultateve të të gjithë rasteve duke bërë thjesht mesataren e rezultateve ekzistuese apo duke përdorur ndonjë metodë tjetër kombinimi të rezultateve.

Algortmi 3 Modeli i bazuar në shembuj IBLStream [174]

1: **Hyrjet:**

Baza e rasteve \mathcal{D} , shembull $e = \langle x_0, ? \rangle$

k : Numri i shembujve në konsideratë,

σ : varianca, e përdorur në rastin e peshimit Gaussian ose
exponencial,

WM : metoda e peshimit (e barabartë, distancaInverse,
lineare, kernelGaussian, kernelEksponecial),

PM : Metoda e parashikimit

2: **Fillo:**

$S = k$ fqinjët më të afërt në \mathcal{D}

3: **if** $\exists x_i \in S : x_i = x_0$ **then**

4: Jep rezultatin λ_{x_i} për shembullin x_i

5: **end if**

6: $W = \text{merrVektorin_e_Normalizuar_te_Peshave}(e, W, M, k, \sigma)$

7: Rasti: {*Regresion*}

8: **if** $PM = wKNN$ **then**

9: {*Zgjidhe si wKNN*}

10: $\hat{\lambda}_{x_0} = W^T Y$

11: **else**

12: $X = [x_i 1]_{x_i \in S}$

13: $X = [\lambda_{x_i}]_{x_i \in S}$

14: $\hat{\beta} = (X^T W X)^{-1} X^T W X$

15: $\hat{\lambda}_{x_0} = [x_0 1] \hat{\beta}$

16: **end if**

5.2.4 ABR si model i parashikimit online

Modeli ABR i përshkruar me detaje në pjesët e mësipërme 4.2.9 në këtë pjesë përdoret si model i parashikimit dinamik. Algortmi i përdorur, modelohet identikisht si ai i implementuar në pjesën 5.1.2 për parashikimin statik. Modeli funksionon duke marrë rastet e krijuara me radhë në mënyrë inkrementale. Rastet e ndërtuara pasi janë përditësuar në bazën e rasteve mbi to aplikohet kërkimi mbi ekzistencën e rasteve të ngjashme sipas dy kushteve të temperaturës dhe kohës së matjes. Daljet e rasteve të evidentuar si të ngjashëm përdoren duke bërë adaptimin e vlerave të konsumit për të marrë daljen e rastit në shqyrtim. Kur baza e rasteve përmban disa raste kushti i përdorimit të një numri të caktuar rastesh bie duke marrë rastet e ngjashe ekzistente. Konsumi i energjisë jepet me

anë të relacionit të mëposhtëm:

$$\hat{P} = \frac{\sum_{k=1}^K CS(X_k, P) \cdot P'_k}{\sum_{k=1}^K CS(X_k, Y)}, \quad (5.5)$$

ku P'_k është konsumi referuar rastit X_k në të njëjtin çast kohor i shumëzuar nga një koefiçent i adaptimit.

5.2.5 Rezultatet eksperimentale

Rezultatet eksperimentale për tre metodat e para do të realizohen në platformën MOA (Massive Online Analysis) e cila është një software open source për parashikimin dhe analizën e të dhënave të rrjedhshme. E veçanta e platformës është se ofron një sërë metodash për parashikimin online, si për shembull ofrohen pemët e regresionit të sipërpërmendura nën emërtimin (FIMT-DD). AMRules, FIMT-DD dhe IBLStreams implementohen në MOA ndërsa ABR duke qënë e implementuar në MATLAB i marrim rezultatet andej.

Rëndësi për metodat paraqet vektori i hyrjes që do të përdoret për të ushqyer secilin nga modelet. Vektori i hyrjes është i njëjti me atë të përdorur në parashikimin statik (5.4), por këtu të dhënat vijnë në mënyrë inkrementale. Secili nga modelet përveçse krahasohet me modelet e tjera të inteligjencës artificiale krahasohet edhe me tre parashikues bazë si tregohen më poshtë:

- energjia e konsumuar e parashikuar është sa mesatarja e rasteve pasardhëse
- parashikimi është sa konsumi mesatar i tre orëve të kaluara $(P_{-3} + P_{-2} + P_{-1})/3$
- parashikimi është sa konsumi i orës parardhëse P_{-1}

Parametrat e secilit model të implementuar janë si vijon,

AMRules dhe FIMTDD përdorin koefiçentin e besueshmërisë së ndarjes 0.2 dhe shmangien e nyjes 0.2. IBLStreams përdorin funksionin e regresionit linear me kernel Gaussian si funksion peshues dhe adaptimin e finjesisë k , ku parametri kernel i zgjedhur është 0.05, dhe madhësia fillestare e setit të fqinjëve më të afërt është $k = 16$ dhe madhësia e bazës është $cb = 1000$. Metoda e vlerësimit të gjithë modeleve është CV-RMSE grumbulluese, ku gabimi llogaritet mbas seclit kampion të parashikuar duke marrë parasysh edhe parashikimet deri në çastin e parashikimit.

Nga rezultatet e treguara në figurat 5.9 - 5.13 vërehet mbizotërimi i metodave të inteligjencës artificiale kundrej parashikuesve bazë. Parashikuesi bazë më i mirë është ai i cili merr si konsum energjie orën e mëparshme

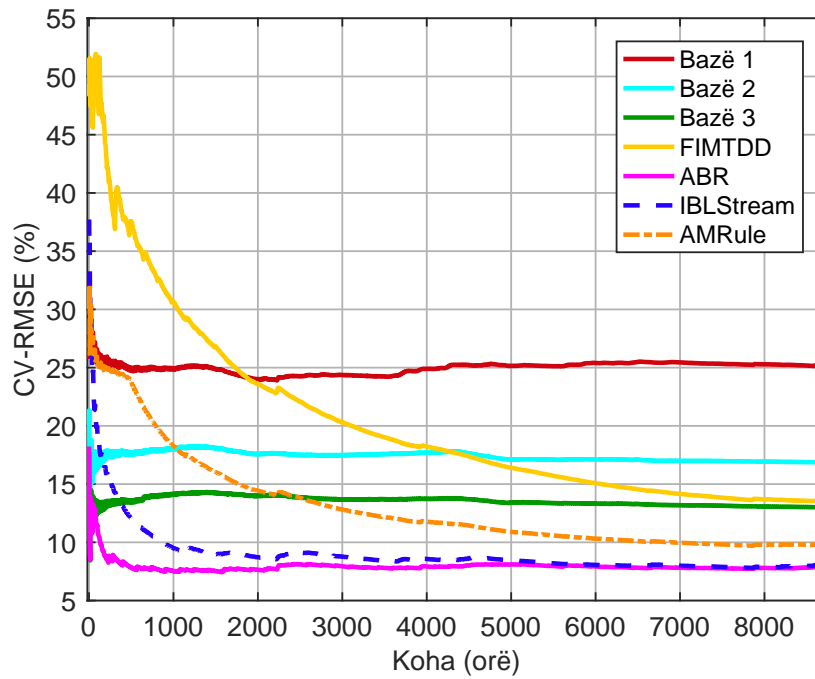


Figura 5.9: Gabimet online për ndërtesën e parë

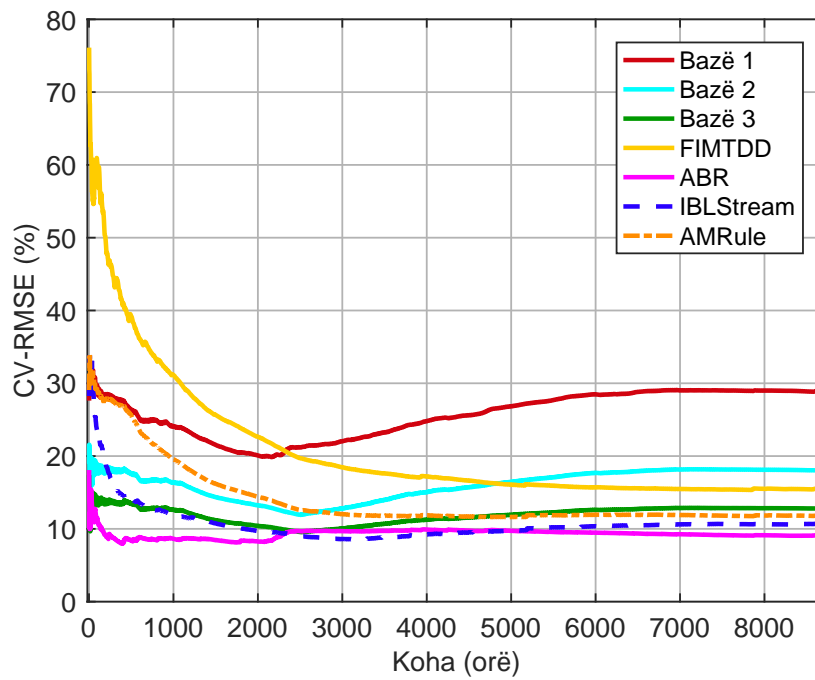


Figura 5.10: Gabimet online për ndërtesën e dytë

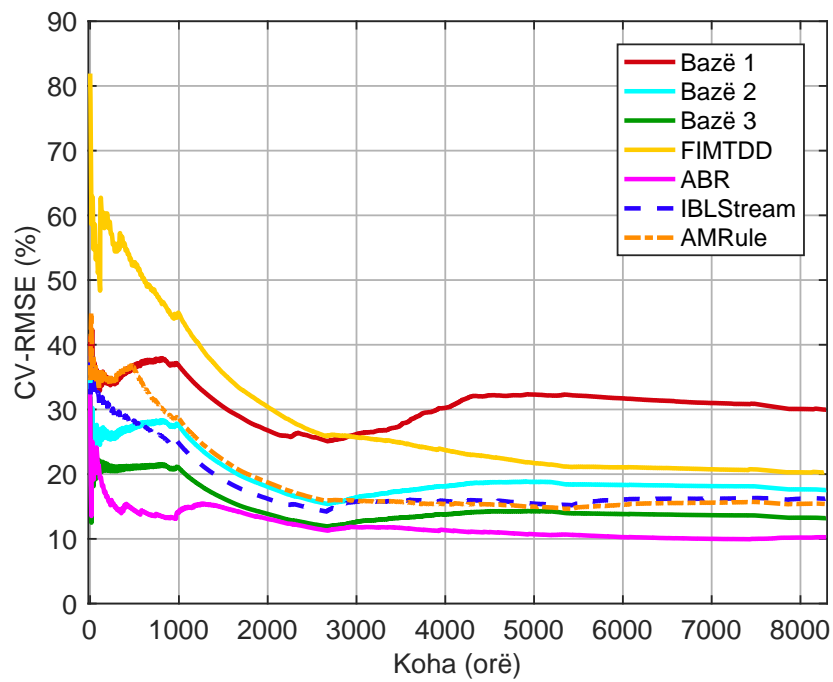


Figura 5.11: Gabimet online për ndërtesën e tretë

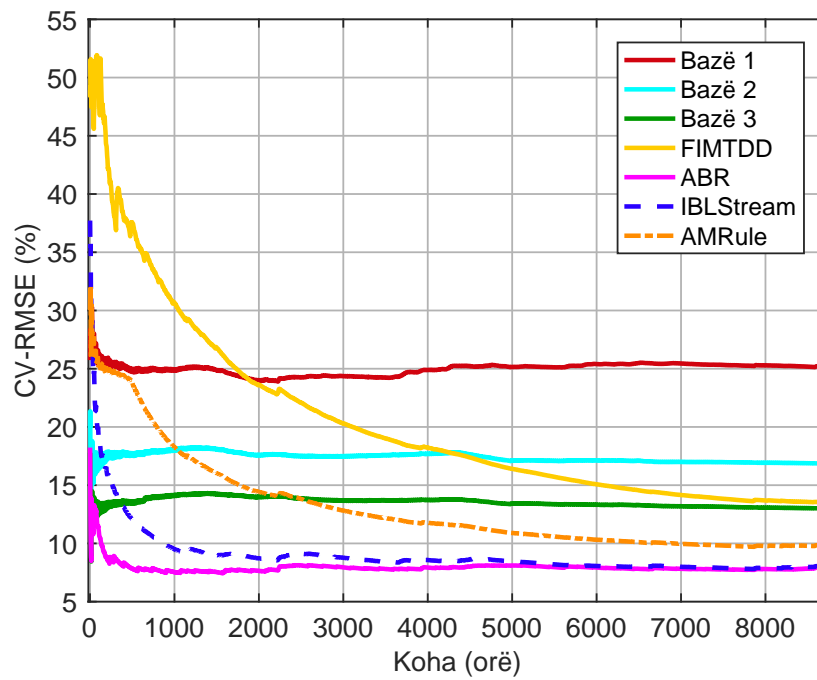


Figura 5.12: Gabimet online për ndërtesën e katërt

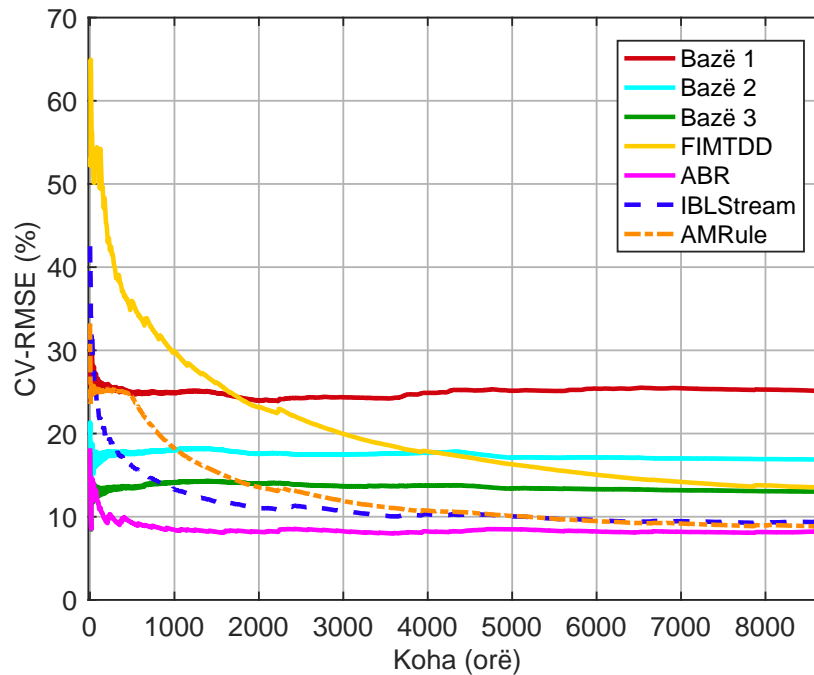


Figura 5.13: Gabimet online për ndërtesën e pestë

kundrejt dy parashikuesve të tjerë bazë. Ndërsa referuar metodave të tjera shikojmë se ABR ka një performancë më të mirë kundrejt gjithë metodave ndërsa vërehet se rregullimin më të shpejtë të gabimit e zotëron IBLStream praktikisht e ngjashme me ABR pasi edhe kjo është metodë e bazuar tek shembujt por me ndryshimin se këtu numri i rasteve që ruhen në bazën e të dhënave kufizohet. Me një performancë më të ulët se dy të sipërpërmendurat është AMRules, por që sillet ngjashëm me dy të parat. Duke qënë se parashikimin në rastet e para e ka më të vështirë se dy të tjerat gje e cila e shoqëron edhe gabimin final. Ndërsa pemët e regresionit FIMT-DD tregojnë një performancë më të ulët se metodat e sipërpërmendura të inteligjencës artificiale kjo veçanërisht për faktin se ato kërkojnë njëfarë kohe të mësojnë mbi të dhënat pastaj të parashikojnë në mënyrë dinamike.

5.3 Diskutime

Në këtë kapitull aplikuar parashikimin e konsumit të energjisë në pesë ndërtesa të ndryshme. Metodat e machine learning janë në fokusin e parashikimit, ku skenarë të ndryshëm analizohen. Dy janë qasjet që implemetuam në këtë kapitull, parashikimi statik apo offline si dhe parashikimi dinamik apo online. Të dhënat e përdorura si ushqim për metodat janë publikisht të vlefshme me vetëm dy madhësi të matura i

nështrohen procesit të ekstraktimit të karakteristikave nga këto matje, po ashtu aplikohet edhe pastrimi i të dhënave. Me të pasur të dhënat e gatshme qasja e parë aplikohet duke patur në fokus një gamë të gjerë metodash të parashikimit statik. Qëllimi i jonë nuk qëndronte tek aplikimi i një metode fikse si tek kapitulli i mëparshëm, por krahasimi dhe evidentimi i metodës më të favorshme në tre skenarë të ndryshëm trajnimi. Nga rezultatet del qartë se pyjet e rastësishëm PR_{ast} janë metoda më efikase në tre termat vlerësues të aplikuar nga ne, kjo për faktin se kjo është një metodë e cila kombinon disa mësues. Rezultatet janë relativisht shumë mirë krahasuar me standartet e vendosur nga ASHRAE [77], po ashtu krahasuar me parashikuesin bazë.

Qasja e dytë i referohet parashikimit online të konsumit të energjisë, ku metodat e përdorura janë ato më të përdorshmet në parashikimin online të problemeve me të dhëna të vijueshme. Ne aplikojmë këto metoda për faktin se kjo përbën një interes të madh praktik, ku konsumi i energjisë është marrëdhënie e vijueshme e shumë faktorëve dhe po ashtu kërkohet të njihen gjithmon vlerat e parashikuar të tij. Metodatat e aplikuar janë kryesisht të bazuara tek shembujt si ABR, AMRules dhe IBLStreams ndërsa FIMT-DD është metodë e bazuar tek modeli. Vërehet se metodat si ABR dhe IBLStreams të ngjashme mes tyre kanë performancën më të lartë. Nga rezultatet e arritura vërejmë se qasja online është e aplikueshme dhe me rezultate shumë të mira të cilat konvergjojnë tek rezultati i modelit statik.

Kapitulli 6

PARASHIKIMI ME INTERVAL I KONSUMIT

Shpesh herë kur parashikojmë një vlerë \hat{y} lind pyetja se sa i saktë është parashikimi, apo sa të sigurt jemi që vlera e parashikuar i përafrohet asaj reale $\hat{y} \approx y$. Përgjigja e këtyre pyetjeve vjen nga një fushë kërkimi e inteligjencës artificiale e njohur si parashikimi me interval ose ndryshe e njohur në literaturë intervalet e besueshmërisë apo *conformal prediction*. Sipas kësaj qasje parashikimi kryhet në trajtën e një intervali dhe jo të një vlere fikse, i cili ka pritshmëri shumë të larta që të përmbajë vlerën reale për një nivel të paracaktuar besueshmërie. Parashikimi me interval përdor njohurinë ekzistuese mbi problemin për të përcaktuar nivelet e besueshmërisë së parashikimeve. E shprehur më thjeshtë për një nivel besueshmërie 95% një metodë parashikimi tradicional (një nga të sipërpërmendurat kapitulli 2) parashikimi me interval do të prodhojë intervale parashikimi të cilat kanë afërsisht 95% shance që të përmbajnë vlerën reale brenda tyre. Metoda tradicionale e “machine learning” parashikon një vlerë \hat{y} dhe rreth kësaj vlere ndërtohet një interval $\Gamma^{0.05}$ nga platforma e parashikimit me interval, ku vlera 0.05 është vlera e shmangies nga besueshmëria $1 - 0.95$.

Parashikimi me interval mund të përdoret me këdo nga metodat e inteligjencës artificiale si RN, RVN, PR, PR_{ast} etj. Nisur nga parashikimi i një pike fikse me metodat tradicionale siç kemi bërë në Kapitullin 5, vlerësojmë “*jokonformitetin*” që tregon se sa ndryshon shembulli konkret nga shembujt e mëparshëm. Algoritmi konformal e kthen këtë rezultat në një zonë vlerash të parashikuara. Në literaturë njihen dy modele të parashikimit me interval që njihen si induktiv dhe transduktiv. Ndryshimi midis tyre është se në modelin induktiv nga të dhënat e trajnimit nxirret një rregull i përgjithshëm për të gjithë shembujt dhe zbatohet tek shembujt e

rinj ndërsa në modelin transduktiv nuk gjenerohet ndonjë rregull i përgjithshëm, por për secilin shembull aplikohet algoritmi për të përfituar rregullin. Kostot llogaritëse të qasjes transduktive janë tepër të larta kështu që gjerësisht gjen përdorim qasja induktive që do të përshkruhet imtësisht më poshtë.

6.1 Historiku

Propozimi i parë i parashikimit me një nivel besueshmërie vjen nga Gammerman, Vapnik dhe Vovk [176] të cilët përpunon të modifikojnë algoritmin kyç të Vapnik për klasifikimin me suport vektorët (MVN) [177]. Kjo ide ndiqet nga Saunders, Gammerman dhe Vovk [178] duke përmirësuar më tej versionin e parashikimit me interval transduktiv. Ndërkohë që parashikimi me interval transduktiv haste vështirësi në zgjidhjen e problemeve të klasifikimit dhe regresionit një zgjidhje vjen nga Nouretdinov [179] ku propozohet një mënyrë e re e përcaktimit të intervaleve të besueshmërisë. Ideja përpunohet më tej nga Papadopoulos [180] i cili del në versionin e ri të parashikimit me interval induktiv (Inductive Conformal Prediction), ku veçoria kryesore e algoritmit është në efikasitetin llogaritës krahasuar me qasjen transduktive. Aplikimet e qasjes induktive gjejnë një përdoshmëri shumë të madhe në literaturë ku edhe aplikimi ynë i dedikohet kësaj qasje. Siç u theksua më sipër parashikimi me interval përdor një mësues bazë dhe këtu vërejmë se aplikimi i tij duke përdorur mësues bazë RN është objekt i Papadopoulos [181] ku themelet teorike dhe praktike hidhen, ndërkaq aplikimi i tij në problemet e regresionit qartësohet më tej në punimet [182, 183]. Po ashtu Papadopoulos përdor teknikën e fqinjëve më të afërt për të parashikuar me interval [185]. Pyjet e rastësishëm dhe qasja induktive janë në fokusin e Johansson [184], ku teknika të ndryshme të matjes së jokonformitetit testohen. Publikimet e mësipërme i dedikohen më së shumti evidentimit të përdorimit të metodës me të dhëna të ndryshme duke e analizuar në shkallë të gjerë, ndërsa aplikimi në fusha të veçanta gjendet në [186] ku parashikimi i shpejtësisë së erës me interval aplikohet. Po ashtu, në [188] problemet e klasifikimit dhe regresionit në fushën e kimisë zgjidhen. Për informacione shtesë mbi aplikimin e parashikimit me interval referohemi tek punimet teorike [189, 190].

6.2 Parashikimi transduktiv

Hapat e përgjithshëm për implementimin e një algoritmi për parashikimin me interval sipas qasjes transduktive janë si më poshtë:

E gjithë baza e të dhënave ndahet në pjesën e trajnimit $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$ dhe të testimit $\{(\mathbf{x}_{l+1}, y_{l+1}), \dots, (\mathbf{x}_n, y_n)\}$; ku $\mathbf{x}_i \in \mathbb{R}^d$

dhe $\mathbf{y}_i \in \mathbb{R}$. Qëllimi është që për një shembull të ri x_{l+g} nga pjesa e testimit të gjejmë intervalin e mundshëm \hat{Y} i cili e përmban me një siguri të caktuar vlerën e vërtetë. Një algoritëm standard i mësimin makinë (që mund të jetë nga të sipërpërmendurit) trajnohet tek pjesa e të dhënave të trajnimit duke përshtatur një model të caktuar si në (6.1) mbi marrëdhënien hyrje-dalje.

$$y_i \approx f(\mathbf{x}_i) \quad (6.1)$$

Duke marrë një shembull të ri (x_{l+g}) nga pjesa e testimit parashikojmë vlerën e mundshme duke përdorur modelin e mësuar me një algoritëm bazë. Qëllimi është i tillë që vlera e parashikuar \tilde{y} nga algoritmi të jetë sa më afër vlerës reale të shembullit x_{l+g} . Gjetja e vlerës sa më të afërt të parashikuar kryhet duke i caktuar një vlerë jokonformiteti $\alpha_i^{\tilde{y}}$ (sqarohet me detaje në pjesën 6.4) secilit shembull (x_i, y_i) . Vlera e jokonformitetit tregon mospajtueshmërinë midis shembujve në pjesën e trajnimit dhe vlerave të parashikuara nga modeli i trajnuar, kjo llogaritet duke parashikuar vlerat e daljes \hat{y}_i për secilin shembull x_i nga trajnimi duke përdorur modelin e mësuar.

$$\hat{y}_i \approx f(\mathbf{x}_i) \quad (6.2)$$

Vlera e jokonformitetit llogaritet thjesht si diferenca midis vlerës reale y_i dhe asaj të parashikuar \hat{y}_i , ndërsa për shembullin e ri x_{l+g} nga pjesa e testimit vlera e jokonformitetit llogaritet si ndryshim midis vlerës së parashikuar dhe një vlere të supozuar si e vërteta (pasi vlerat reale nuk mund t'i përdorim). Vlerat e supozuara në rastin e regresionit janë vlera reale që kanë një interval shumë të gjerë (të pafundëm), gjë e cila përbën një probleme tepër të madh në kostot llogaritëse. Vlerat e jokonformitetit $\alpha_{l+g}^{\tilde{y}}$ të llogaritura për shembujt e testimit krahasohen me të gjithë vlerat e jokonformitetit të cilat i janë caktuar secilit shembull në pjesën e trajnimit që të gjendet se sa është ndryshimi midis saj dhe vlerave të tjera. Krahasimi i referohet funksionit (6.3):

$$p(\tilde{y}) = \frac{\#\{i = 1, \dots, l, l+g \mid \alpha_i^{(\tilde{y})} \geq \alpha_{l+g}^{(\tilde{y})}\}}{l+g} \quad (6.3)$$

sipas të cilit llogaritet numri i shembujve të cilët ndryshojnë ose janë njësoj jokonform kundrejt (X_{l+g}, \tilde{y}) . Dalja e këtij funksioni ndryshe njihet vlera (p) e cila merr vlera nga $\frac{1}{l+1}$ deri në 1, ku për çdo vlerë të besueshmërisë $\forall \delta \in [0, 1]$, probabiliteti që vlera reale të ndodhet brenda intervalit me besueshmëri δ është sa ose më shumë se vlera e besueshmërisë.

$$p(\tilde{y}) = \frac{\#\{i = 1, \dots, l, l+g \mid \alpha_i^{(\tilde{y})} \geq \alpha_{l+g}^{(\tilde{y})}\}}{l+g} \geq \delta \quad (6.4)$$

Por pamundësia për të llogaritur secilën vlerë të mundshme të supozuar \tilde{y} në rastin e regresionit çon ekspertët e fushës tek zgjidhje të tjera siç është parashikimi me interval induktiv, ku një përshkrim gjendet në seksionin më poshtë.

6.3 Parashikimi induktiv

Në këtë seksion japim një përshkrim të detajuar të aplikimit të platformës induktive së parashikimit me interval. Në problematikën e regresionit, për një shembull të dhënë x_i algoritmi parashikon intervale në termat (a,b) me një probabilitet $P(a < y_i < b) \geq 1 - \delta$ që vlera e vërtetë y_i ndodhet brenda intervalit për një vlerë të përcaktuar besueshmërie (δ) nga përdoruesi, ku $1 - \delta$ përbën nivelin e besueshmërisë. Për një shembull të ri intervali prodhohet si pasojë e trajnimit të *algoritmit bazë*, i cili parashikon një vlerë fikse të cilës më tej do t'i bashkangjitet intervali i parashikuar nga platforma. Fillimisht të dhënat ndahen në pjesën e trajnimit, dhe të testimit, ku pjesa e trajnimit ndahet në dy nënpjesë atë të trajnimit dhe të kalibrimit si tregohet e detajuar më poshtë.

Për një pjesë të dhënë trajnimi $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$; ku $\mathbf{x}_i \in \mathbb{R}^d$ është vektori i atributëve për shembullin e i -të dhe $y_i \in \mathbb{R}$ është përgjigja për këtë shembull, të dhënat totale ndahen në dy pjesë, në pjesën e trajnimit $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$ dhe pjesën e testimit, ku pjesa e trajnimit ndahet përsëri në dy nënpjesë: ajo e tranimit të algoritmit 'mësues' $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_m, y_m)\}$, ($m < l$) si dhe pjesa e kalibrimit $\{(\mathbf{x}_{m+1}, y_{m+1}), \dots, (\mathbf{x}_l, y_l)\}$, $j = 1, \dots, q$. Për të gjitha çiftet e trajnimit (\mathbf{x}_i, y_i) një model parashikimi gjenerohet duke u bazuar në trajnimin e një algoritmi bazë (mësimi):

$$\hat{y}_i \approx f(\mathbf{x}_i) \quad (6.5)$$

ku \hat{y}_i është dalja (parashikimi) i algoritmit bazë. Modeli i mësuar do të përdoret për të kryer parashikimin e vlerave fikse si në metodikën standarte që kemi shpjeguar dhe implementuar në kapitujt e mësipërm.

Për secilin shembull nga pjesa e kalibrimit modeli i mësuar nga (6.5) do të përdoret për të parashikuar \hat{y}_j duke përdorur \mathbf{x}_j si hyrje për modelin. Mospajtimi midis vlerës reale y_j dhe asaj të parashikuar \hat{y}_j në të dhënat e kalibrimit njihet si rezultati i jokonformitetit, ku në rastin e regresionit llogaritet thjeshtë si diferenca absolute e vlerave si në (6.6).

$$\alpha_j = |y_j - \hat{y}_j| \quad (6.6)$$

Vlerat e jokonformitetit të prodhuara nga llogaritjet e mësipërme renditen sipas rendit zbritës për secilin shembull nga pjesa e kalibrimit \mathbf{x}_j dhe për

secilin kampion marrim vektorin me vlerat e jokonformitetit të renditura $\mathbf{A} = \{\alpha_1, \dots, \alpha_q\}$.

Qëllimi përfundimtar i algoritmit është të parashikojë intervale për një shembull të ri testimi \mathbf{x}_k , $k \in \{l+1, \dots, n\}$ nga pjesa e testimit, ku përgjigjet e shembujve të testimit të gjenden brenda intervali për një nivel besueshmërie. Shembulli i testimit \mathbf{x}_j ushqen 'mësuesin' dhe parashikon pikën fikse \hat{y}_k . Duke marrë parasysh secilin target të mundshëm \tilde{y}_k , një vlerë probabiliteti (p-value) llogaritet duke përdorur jokonformitetin $\alpha_k^{\tilde{y}}$ në vlerat e jokonformitetit α_j nga pjesa e kalibrimit të ruajtura në A .

$$p(\tilde{y}) = \frac{\#\{j = m+1, \dots, l, k | \alpha_j \geq \alpha_k^{\tilde{y}}\}}{l - m + 1} \quad (6.7)$$

Nëse $p(\tilde{y}) < \delta$, shancet e \tilde{y} të jenë brenda intervalit me vlerë sigurie δ të paracaktuar nga përdoruesi janë më të vogla, kështu vlerat që plotësojnë këtë kusht eliminohen si mundësi. Në rastin e regresionit është e vështirë që të llogariten vlerat (p) për secilën target të mundshëm pasi ato janë të pafundme kështu zgjidhja vjen nga Nourtdinov [179], i cili thotë se në vend që të kërkohen zgjidhjet e mundshme \tilde{y} ndryshojmë vlerat e jokonformitetit $\alpha_k^{\tilde{y}}$ për secilin shembull \mathbf{x}_k , ku $\alpha_k^{\tilde{y}} = \alpha_j$, në mënyrë që të plotësohet kushti i besueshmërisë. Është e mundur që të gjendet vlera kufi ose minimale e jokonformitetit $\alpha_j \in A$ e cila kënaq relacionin (6.8),

$$p(\tilde{y}) = \frac{\#\{j = m+1, \dots, l, k | \alpha_j \geq \alpha_k^{\tilde{y}}\}}{l - m + 1} \geq \delta \quad (6.8)$$

kjo nënkupton se të gjithë shembujt të cilët plotësojnë kushtin $\alpha_j \geq \alpha_k^{\tilde{y}}$ përfshihen në interval, në të kundërt eliminohen. Pasi u llogaritën marrëdhëniet e mësipërme intervali i parashikimit do të llogaritet si:

$$\hat{Y}_k^\delta = \hat{y}_k \pm \alpha_s(\delta) \quad (6.9)$$

ku $\alpha_s(\delta)$ është vlera më e vogël e jokonformitetit për një nivel të caktuar sigurie (δ) që plotëson kushtin (6.8).

6.4 Vlerësimi i jokonformitetit

Matja e jokonformitetit siç u sipërpërmend është një nga hapat kryesorë në implementimin e parashikimit me interval. Jokonformiteti shpreh diferencën që ekziston midis vlerave të parashikuara nga modeli i mësuar me algoritmin bazë dhe vlerave reale ekzistuese. Marrim në konsideratë një sasi të dhënë shembujsh $\{z_1, \dots, z_n\}$, ku $z_i = (\mathbf{x}_i, y_i)$ me ndihmën e algoritmit

bazë (6.5) krijohet një model parashikimi $\mathcal{M}_{\{z_1, \dots, z_n\}}$ i cili për secilin shembull \mathbf{x}_j parashikon një vlerë \hat{y}_j . Diferenca midis vlerave të parashikuara dhe atyre reale na jep vlerën e jokonformitetit $\alpha_k = |y_k - \hat{y}_k|$, ku vlerat e jokonformitetit në rastin tonë llogariten për pjesën kalibruese të të dhënave $\hat{y}_j = \mathcal{M}_{\{z_1, \dots, z_n\}}(x_j)$. Atëherë, për secilin shembull në pjesën e kalibrimit caktojmë një vlerë jokonformiteti (6.6) dhe më tej secila vlerë jokonformiteti e renditur ruhet në vektorin $\mathbf{A} = \{\alpha_1, \dots, \alpha_q\}$.

Për një vlerë të caktuar besueshmërie δ intervalet e parashikuara nga platforma do të kenë të njëjtën gjerësi për secilën shembull nga pjesa e testimit, kjo pasi vlera e jokonformitetit α e zgjedhur si e përshtatshme dhe plotëson kushtin (6.8) përdoret si vlerë e kufirit për të gjithë shembujt e testimit për atë vlerë fikse të paracaktuar besueshmërie. Megjithatë ky përbën një mangësi për algoritmin pasi intervalet që do të prodhohen kanë të njëjtën gjerësi për të gjithë shembujt dhe nuk merret parasysh vështirësia e secilit shembull. Eliminimi i kësaj mangësie përmirësohet nga Papadopoulos [191] i cili prezanton konceptin e jokonformitetit të normalizuar, ku sipas të cilit vlerat e jokonformitetit pjestohen me gabimin e algoritmit bazë gjatë trajnimit.

$$\alpha_j = \frac{|y_j - \hat{y}_j|}{\sigma_j} \quad (6.10)$$

Intervalet e parashikimit të prodhuara për shembujt e testimit duke përdorur (6.10) do të kenë gjerësi të ndryshme pasi vlera e jokonformitetit (6.6) pjestohet me një vlerë σ_j në (6.10), i cili varet nga shembujt e rinj. Vlerat e σ janë gabimet e algoritmit bazë gjatë trajnimit që përftohen duke trajnuar një algoritëm më të thjeshtë se algoritmi bazë. Pasi trajnohet modeli bazë (6.5), për të llogariten gabimet absolute $|\hat{y}_i - y_i|$ dhe më tej një algoritëm më i thjeshtë si regresioni linear trajnohet duke përdorur si shembuj çiftin $(\mathbf{x}_i, \ln |\hat{y}_i - y_i|)$ dhe një model i ri i cili parashikon gabimin prodhohet $\mathcal{M}'_{\{\mathbf{x}_1, \ln |\hat{y}_1 - y_1|, \dots, \mathbf{x}_i, \ln |\hat{y}_i - y_i|\}}$, dhe për një shembull të ri \mathbf{x}_j nga pjesa e kalibrimit gabimi i parashikuar σ_j nga modeli i ri është:

$$\sigma_j \approx \mathcal{M}'_{\{\mathbf{x}_1, \ln |\hat{y}_1 - y_1|, \dots, \mathbf{x}_i, \ln |\hat{y}_i - y_i|\}}(\mathbf{x}_j)$$

Normalizimi duke përdorur saktësinë e modelit të trajnuar përmirëson gjerësinë e intervalit të prodhuar, duke parashikuar intervale më të gjera për shembujt më të vështirë dhe më të ngushta për shembujt më të lehtë. Intervali i parashikimit për një nivel të caktuar besueshmërie δ tek shembulli i ri \mathbf{x}_k merr trajtën,

$$\hat{Y}_k^\delta = \hat{y}_k \pm \alpha_s(\delta) \cdot \sigma_k \quad (6.11)$$

ku $\alpha_s(\delta)$ është vlera e normalizuar e jokonformitetit e cila plotëson kushtin për

vlerën (p) (6.4), si dhe σ_k është gabimi i parashikuar për shembullin e ri \mathbf{x}_k duke përdorur modelin \mathcal{M}' .

6.5 Implementimi i metodës

Të dhënat e përdorura për implementimin e metodës janë të njëjtat me ato të përdorura në Kapitullin 5 të përshkruara në pjesën 5.1.1. Siç u përmend më sipër implementimi i platformës së parashikimit me interval kërkon përdorimin e algoritmave standard për parashikimin sipas pikave fikse të cilët trajtohen me detaje në Kapitullin 5, ku në fokusin tonë në këtë kapitull janë pyjet e rastësishëm PR_{ast} si metoda më e sukseshme e aplikuar tek këto të dhëna dhe RN si metoda më e përdorshme në probleme të ngjashme dhe rezulton të jetë shume konkurruese e aplikuar tek të dhënat tona. Fillimisht të dhënat ndahen rastësisht në dy pjesë: *trajnim* 80% dhe *testim* 20%. Më tej pjesa e trajnimit ndahet në *pjesën e duhur* 70% dhe *pjesën e kalibrimit* 30%.

Pjesa e duhur shërben si pjesa ku algoritmi ‘bazë’ trajnohet siç shprehet në ekuacionin (6.5) për të përshtatur një model \mathcal{M} tek të dhënat. Nëse vlerësimi i jokonformitetit të përdorur është ai i normalizuar (6.10) atëherë numri i algoritmave që përdoren është dy. Algoritmi i parë trajnohet tek $D_i = \{\mathbf{x}, y\}_i^{|T|_{pr}}$, ku i tregon numrin e shembujve në pjesën e duhur $|T|_{pr}$, dhe algoritmi i dytë trajnohet tek $D_i = \{\mathbf{x}, \ln|y - \hat{y}|\}_i^{|T|_{pr}}$, ku $\ln|y - \hat{y}|$ tregon gabimin e algoritmit të parë gjatë trajnimit. Në rastin e përdorimit të jokonformitetit sipas (6.6), atëherë vetëm përdorimi i një algoritmi është i mjaftueshëm.

Pjesa e kalibrimit përdoret për të llogaritur vlerat e jokonformitetit, ku jokonformiteti jo i normalizuar llogaritet si $\alpha_j = |y_j - f(\mathbf{x}_j)|$ tek $D_j = \{\mathbf{x}, y\}_j^{|T|_{cal}}$. Nga ana tjetër vlerat e normalizuara të jokonformitetit llogariten sipas (6.10), ku σ_j janë vlerat e parashikuar të cilat prodhohen nga algoritmi i dytë i trajnuar tek gabimet e algoritmit të parë me model \mathcal{M}' . Vlerat e llogaritura të jokonformitetit renditen sipas rendit zbritës $\mathbf{A} = \{\alpha_1, \dots, \alpha_j, \dots, \alpha_q\}$ për të gjithë shembujt q në pjesën e kalibrimit.

Pjesa e testimit shërben si mjedis për parashikimin e intervaleve me një nivel të caktuar besueshmërie. Algoritmi bazë i trajnuar në pjesën e duhur me model të përfituar \mathcal{M} parashikon për secilin shembull të ri \mathbf{x}_k nga $D_k = \{\mathbf{x}, y\}_k^{|T|_{test}}$ një vlerë fikse të konsumit të energjisë. Intervali që do të parashikohet merr parasysh vlerat e jokonformitetit për një nivel besueshmërie të paracaktuar. Qartësisht nga pjesa e renditur e jokonformiteteve α vlerat të cilat kanë vlerën (p) më të madhe se δ përfshihen në intervalin e parashikuar, përndryshe eliminohen nga intervali.

Marrim një shembull të thjeshtë, nëse në pjesën e kalibrimit ndodhen 99

shembuj për një nivel besueshmërie 95% vlera e jokonformitetit e cila zgjidhet si kufiri ndarës midis vlerave të cilat do të përfshijmë në interval dhe tyre që do të eliminohen do të jetë ajo e ndodhur në në pozicionin e 5-të. E thënë ndryshe vlerat e jokonformitetit të cilat kanë vlera më të vogla se vlera kufi α_{kufi} përfshihen në interval ndërsa ato më të mëdha nuk përfshihen.

Metoda e implementuar parimisht mbështetet në algoritmin e treguar si në fig. 6.1.

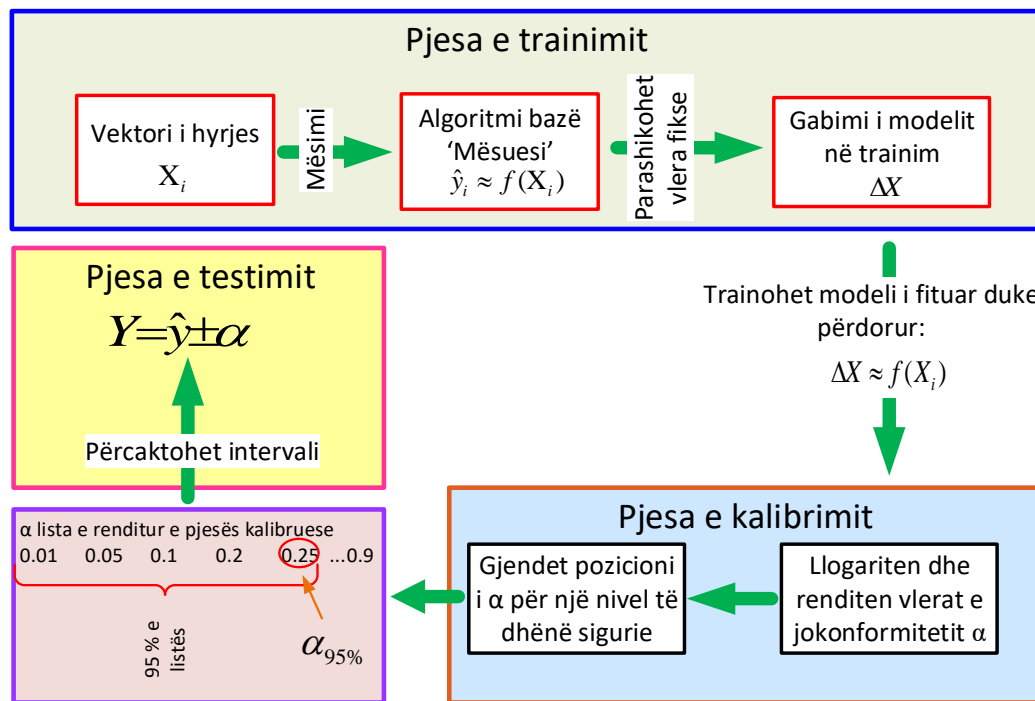


Figura 6.1: Paraqitja skematike e parashikimit me interval.

6.6 Rezultatet eksperimentale

Platforma e përshkruar më sipër për parashikimin e kosnumit të energjisë me interval mbështetet në një model përgjithësues të parashikimit me interval dhe jo ndonjë modeli fikse si ata që jepen në literaturë [182, 183, 185, 184, 186] sidomos për mënyrën e përzgjedhjes së pjesëve të të dhënave, ku shumica e punimeve përdorin një numër fikse për përzgjedhjen e pjesës së kalibrimit $q = 100n - 1$, $n \in N$.

Dy janë algoritmat bazë të përdorur për parashikimin me interval, rrjetat neurale RN dhe pyjet e rastësishëm PR_{ast} . Përpara ekzekutimit të eksperimenteve, parametrat optimal për secilin model zgjidhen. Rrjeta neurale e zgjedhur është një rrjetë me algoritëm mësimi me përhapje të prapme Levenberg–Marquardt, ku përdoret një shtresë e fshehur me 10 neurone

të fshehura dhe funksion aktivizimi sigmoidal. Numri i neuroneve të fshehura zgjidhet duke përdorur një vlerësim kryq me 5-pjesë. Po ashtu, neuronet në shtresën e hyrjes janë 18 sa numri i variablave ndërsa në dalje kemi një neuron. Po ashtu metoda e pyllit të rastësishëm *TreeBagger* përdor si algoritëm bazë 100 pemë regresioni, ku thellësia e pemëve dhe numri i variablave në secilën gjethe përcaktohet duke përdorur optimizimin Bayesian nga Statistics and Machine Learning Toolbox MATLAB [192]. Numri minimal i gjetheve në secilën nga pemët është 10 dhe numri i variablave që kampionohen në secilën degë është 15.

Kur përdorim vlerësimin e jokonformitetit të normalizuar, në rastin e RN algoritmi i dytë i përdorur për parashikimin e gabimit të modelit është një rrjetë neurale lineare ndërsa në rastin e PR_{ast} përdorim një model më të thjeshtë se algoritmi bazë që është një *TreeBagger* me 50 pemë regresioni dhe parametra të paracaktuar (default).

Vlerësimi i metodës për dy algoritmat bëhet për tre vlera të ndryshme të besueshmërisë përkatësisht 90%, 95%, 99% për pjesë të përzgjedhura rastësisht nga të gjithë të dhënat, dhe ku rezultatet e përfituara janë mesatrizim i përsëritjes të eksperimenteve 20 herë. Rezultatet e treguara në tabelat 6.1, 6.2 tregojnë tre vlerësime të ndryshme. Vlerësimi i parë tregon sa nga vlerat reale ndodhen jashtë intervalit të parashikuar dhe shprehen në (%), i dyti tregon sa është gjerësia e intervaleve të parashikuar ky një tregues tepër i rëndësishëm pasi kriteri më i rëndësishëm është ngushtësia e intervalit, si dhe vlerësimi i tretë tregon gjerësinë e intervalit duke eliminuar 20% të intervaleve praktikisht 10% të sipërm dhe 10% të poshtëm, ky një tregues statistikor.

Interpretimi i rezultateteve do të bëhet duke ju referuar rezultateve të treguara në tabelat 6.1 dhe 6.2, dhe do përshkruajmë domethënien e tyre në zgjidhjen e problemit tonë. Për të kuptuar më qartë metodikën e përshkruar intervalet e parashikuar për ndërtesën e parë me nivel besueshmërie 95% tregohen në figurën 6.2 për rrjetat neurale dhe në 6.3 për pyllin e rastësishëm.

Domethënia e gjerësisë mesatare të intervaleve vlerësohet si raport i saj me diapazonin e ndryshimit të konsumit të energjisë për secilën ndërtesë (Ndërtesa 1-Ndërtesa 5). Në rastin kur algoritmi bazë janë rrjetat neurale dhe vlerësimi i jokonformitetit është i panormalizuar, raportet për nivel besueshmërie 99% janë respektivisht 42.20%, 34.13%, 33.88%, 29.65% dhe 32.00%. Gjerësia e intervalit mbulon përafërsisht një të tretën e diapazonit, ndërsa normalizimi i jokonformitetit na jep 34.70%, 28.84%, 30.45%, 28.59% dhe 29.03%. Rezultatet tregojnë se normalizimi tregon një përmirësim në vlera të përafërta me 1% – 8%. Kuptohet qartë se rezultatet e përfituara nga përdorimi i normalizimit janë ato më të mirat, pra me intervale më të ngushta përfitojmë saktësi (treguesi 1) brenda

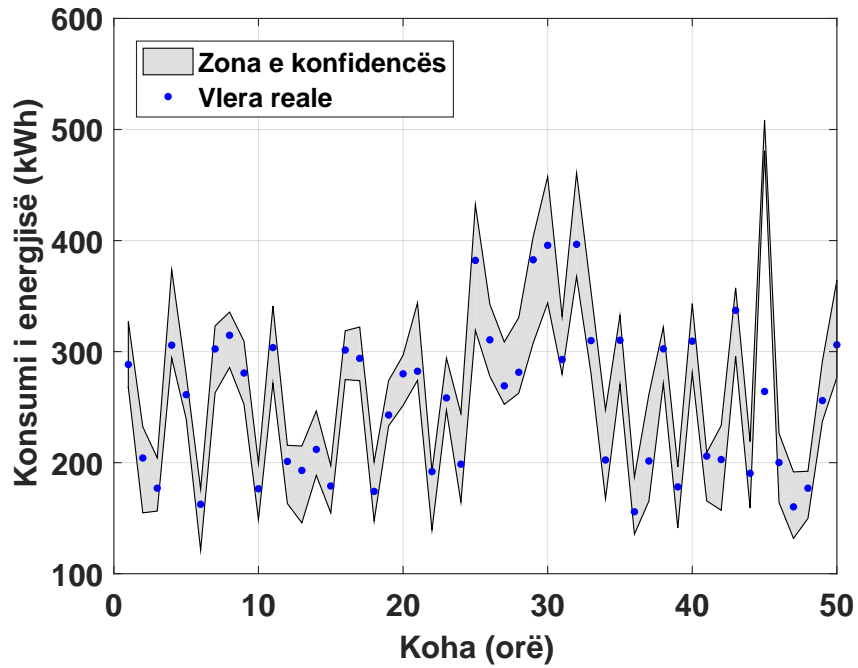


Figura 6.2: Intervalet e parashikimit në ndërtesën 1 parashikuar me RN për nivel besueshmërie 95%.

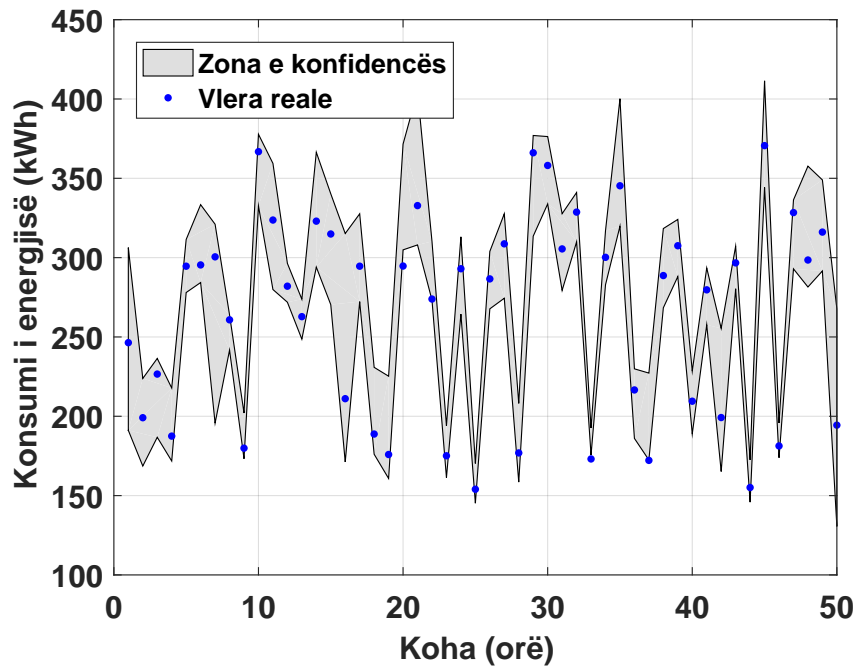


Figura 6.3: Intervalet e parashikimit në ndërtesën 1 parashikuar me PR_{ast} për nivel besueshmërie 95%.

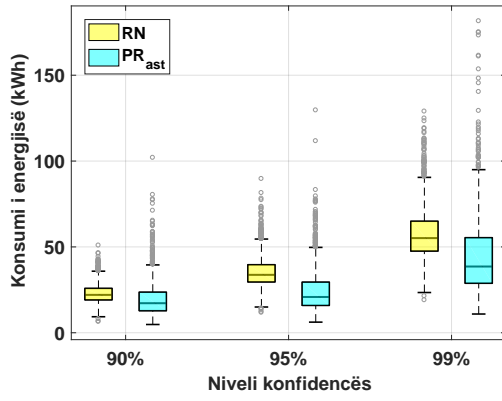
nivelit të besueshmërisë (konfidencës).

Ndërsa për nivelin e besueshmërisë 95% raportet më të mira të gjerësisë së intervaleve me relacionin (6.10) janë 21.03%, 15.52%, 17.90%, 14.53% dhe 15.47%. Krahasuar me rezultatet e marra me nivelin e besueshmërisë 99% duket qartë se gjerësia e intervalit zvogëlohet përafërsisht 50%.

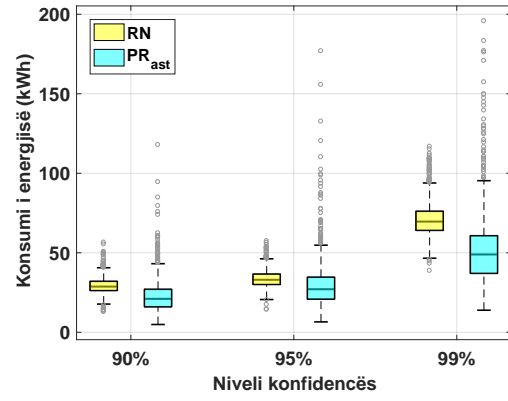
Më tej tregohen rezultatet e arritura nga përdorimi i pyllit të rastësishëm si algoritëm bazë dhe krahasimi i dy algoritmave përshkruhet. Përdorimi i jokonformitetit (6.6) tregon se domethëniet e gjerësisë së intervaleve për secilën ndërtesë janë respektivisht, 31.49%, 25.27%, 31.15%, 25.32%, 24.71%, gjerësia e intervaleve mbullon nga çereku i diapazonit në një të tretën. Eksperimentimi me jokonformitet të normalizuar (6.10) jep rezultatet 23.62%, 22.89%, 19.89%, 21.95% dhe 16.52%, ku përmirësimet janë nga 2.5% në ndërtesën 2 deri në 11% ndërtesën 3. Krahasojmë rezultatet e përfituara për nivel besueshmërie 99% dhe jokonformitet (6.10) me rezultatet 12.80%, 12.82%, 11.40%, 10% dhe 10.43% të fituar për nivel besueshmërie 95%, vërejmë se përmirësimi është afërsisht me 50%, por fakti më domethënës është se intervali mbulon vetëm 10% të diapazonit.

Krahasimi i dy algoritmave të cilët përdorin (6.10) për tre nivelet e besueshmërisë tregohet në fig. 6.4 si dhe rezultatet e parashikuara për një hapësirë kohore 50 orësh me të dy algoritmat sipas secilit nivel besueshmërie tregohen në shtesën C. Nga krahasimi i këtyre algoritmave referuar tre treguesve të cilësisë vërehet se pylli i rastësishëm performon më mirë se rrjetat neurale, kjo për faktin se PR_{ast} kanë rezultate më të sakta në parashikimin e vlerave fikse e treguar në Kapitullin 5 po ashtu edhe algoritmi i dytë i mësimi ka performancë më të lartë krahasuar me rrjetat lineare.

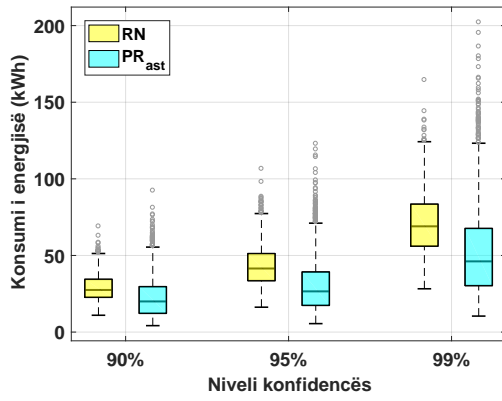
Rezultatet e treguara tek kjo pjesë forcojnë idenë e aplikimit të metodës në problemin e parashikimit të konsumit të energjisë.



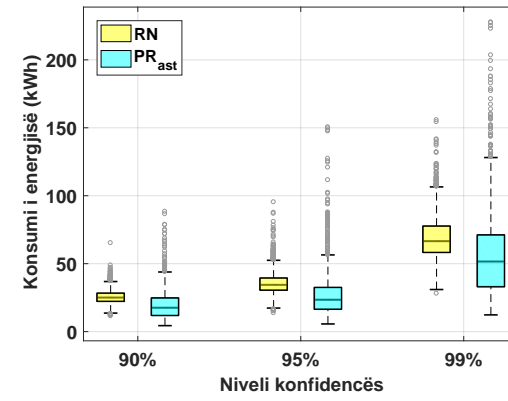
(a) Ndërtesa 1



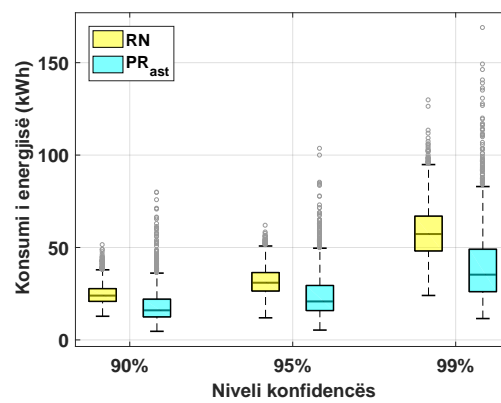
(b) Ndërtesa 2



(c) Ndërtesa 3



(d) Ndërtesa 4



(e) Ndërtesa 5

Figura 6.4: Intervalet e konsumit në secilën ndërtesë për çdo vlerë besueshmërie.

Tabela 6.1: Rezultatet e parashikimit me interval për ndërtesat 1-3

Metoda	Jokonformiteti	Vlera jashtë (%)			Gjerësia mesatare			Gjerësia mesatare e 80%		
		90%	95%	99%	90%	95%	99%	90%	95%	99%
Ndërtesa 1										
RN	(6.6)	7.6131	3.6634	0.3434	47.8195	69.7406	138.4386	47.8195	69.7406	138.4386
	(6.10)	7.8420	2.9765	1.2021	44.1345	68.9755	110.5730	45.1097	70.4241	113.0988
PR_{ast}	(6.6)	9.0441	4.4648	0.9731	40.0887	55.2372	103.2996	40.0887	55.2372	103.2996
	(6.10)	8.0710	3.8351	0.9159	35.0431	41.9757	77.4924	36.7066	45.4640	83.2582
Ndërtesa 2										
RN	(6.6)	7.6131	3.3200	0.7441	58.2399	74.4990	146.9100	58.2399	74.4990	146.91000
	(6.10)	8.1855	3.8924	0.5724	55.0081	66.6880	123.6514	53.7256	67.1926	125.7512
PR_{ast}	(6.6)	10.4751	3.6634	0.8014	44.1536	62.4069	108.5916	44.1536	62.4069	108.5916
	(6.10)	9.4448	4.7510	0.8014	42.0603	55.0642	98.3760	43.2686	55.6496	98.7080
Ndërtesa 3										
RN	(6.6)	8.7402	3.7372	0.7233	61.7908	93.0443	158.3180	61.7908	93.0443	158.3180
	(6.10)	8.4991	3.4358	0.7836	55.8767	83.5997	142.8474	57.3840	85.0700	144.7602
PR_{ast}	(6.6)	10.8499	4.0989	0.5425	46.9481	73.0981	145.4686	46.9481	73.0981	145.4686
	(6.10)	9.8855	4.4002	0.9042	40.0169	53.2253	92.9085	42.4491	57.1716	98.9483

Tabela 6.2: Rezultatet e parashikimit me interval për ndërtesat 4-5

Metoda	Jokonformiteti	Vlera jashtë (%)			Gjerësia mesatare			Gjerësia mesatare e 80%		
		90%	95%	99%	90%	95%	99%	90%	95%	99%
Ndërtesa 4										
RN	(6.6)	8.9296	3.3772	0.5152	59.9833	88.5120	140.5556	59.9833	88.5120	140.5556
	(6.10)	8.2427	4.1786	0.8014	50.5577	68.9144	135.2960	51.0471	70.0757	138.1007
PR_{ast}	(6.6)	8.1282	5.6669	1.4310	46.1708	62.2739	120.0175	46.1708	62.2739	120.0175
	(6.10)	8.2999	4.0641	0.6869	35.1317	47.3615	104.0059	37.0176	49.2690	106.1845
Ndërtesa 5										
RN	(6.6)	9.6165	4.4076	0.4007	49.6093	65.1782	128.7755	49.6093	65.1782	128.7755
	(6.10)	7.3268	3.4917	0.6297	48.2966	62.2894	116.8168	48.9939	63.5437	118.3655
PR_{ast}	(6.6)	8.5289	5.2662	1.3738	39.4999	54.1678	106.1590	39.4999	54.1678	106.1590
	(6.10)	10.3034	3.7779	1.0876	32.2903	41.9840	70.9676	34.6108	45.3651	76.1107

6.7 Përfundime

Në këtë kapitull aplikuaam një teknikë të re për parashikimin e konsumit të energjisë në ndërtesa. Metoda sjell në vëmendje parashikimin me interval dhe jo të një vlere fikse si haset në metodat tradicionale. Rezultatet e marra shprehin efikasitetin për përdorimin e metodës në dy aspekte kryesore: aspekti i parë tregon se për një nivel të caktuar besueshmërie metoda arrin të parashikojë intervale të cilat përmbajë vlerat reale brenda këtij niveli besueshmërie si dhe aspekti i dytë tregon se gjerësia e këtyre intervaleve është shumë domethënëse pasi përfshin pjesë relativisht të vogla të diapazonit të energjisë. Po ashtu, implementimi i dy teknikave të vlerësimit të jokonformitetit thekson rëndësinë që ka në performancë veçanërisht në gjerësinë e intervaleve të parashikuara. Rezultatet e marra përforcojnë motivimin për aplikimin e kësaj metode në parashikim, pasi parashikimi me interval ka një garanci tepër të lartë krahasuar me parashikimin në një pikë fikse.

Kapitulli 7

PËRMBLEDHJE

Parashikimi i konsumit të energjisë është pjesë shumë e rëndësishme që përmirëson procesin e projektimit, mirëmbajtjes dhe funksionimit të ndërtesave. Një projektim shumë i mirë nuk do të thotë sjellje eficiente në aspektin energjistik, kështu që përmirësimi duke u fokusuar tek procesi i funksionimit është thelbësor në kursimin e konsumit dhe në ruajtjen e mjedisit (ulja e prodhimit të CO_2). Kjo tezë u fokusua në aplikimin e algoritmave më në zë momental të inteligjencës artificiale për parashikimin e konsumit të energjisë në ndërtesa në terma afatshkurtër.

Fillimisht, një përshkrim i motivimit dhe historikut të zhvillimit të disiplinës në zgjidhjen e problemit jepet duke u pasuar nga një pasqyrim i detajuar i punëve ekzistuese në parashikimin e konsumit të energjisë. Metodatat e shqyrtuara janë metodatat inxhinierike, metodatat statistikore dhe ato të inteligjencës artificiale si: rrjetat neurale, regresioni me vektorët ndihmës, pemët e regresionit, pyjet e rastësishëm dhe arsyetimi i bazuar tek rastet. Nuk mungon një përshkrim mbi metodatat e matjes inteligjente pasi ekzistenca e të dhënave përbën bërthamën e implementimit të metodave. Punimet ekzistuese të analizuarat të cilat parashikojnë konsumin, përpiqen të optimizojnë metodatat e përdorura apo të hedhin hapa të rinj në metoda të reja. Ku në këtë tezë përmbledhim pikat e forta dhe të dobëta të metodave, por pa thënë se cila metodë ka përparësi sepse metodatat nuk janë testuar në të njëjtat kushte dhe me të njëjtat të dhëna. Midis të gjithë metodave të analizuarat ato të inteligjencës artificiale janë edhe më të përdorshmet nga literatura. Theksojmë se analiza e metodave inxhinierike zë një hapësirë në këtë tezë pasi ajo është metodë tepër e vlefshme në simulimin e konsumit energjistik të ndërtesës për të analizuar në hapat e para të funksionimit, ku kemi mungesën e të dhënave të matura. Gjenerimi i të dhënave sintetike për implementimin e metodave të inteligjencës artificiale është një avantazh i këtyre metodave.

Më tej, tentojmë të japim konceptet teorike për të gjitha metodat e inteligjencës artificiale të analizuar në këtë punim. Përshkrimin e ndajmë në dy kapituj, përkatësisht Kapitulli 2 që përfshin gjithë metodat e inteligjencës artificiale të bazuara tek modeli dhe Kapitulli 3 që përfshin metodën e bazuar tek shembujt. Metodat Përshkruhen me detajet e nevojshme për aplikimin në problemet e regresionit dhe më tej neve i përshtasim sipas problemit tonë.

Kapitulli 4 analizon implementimin e ABR si metodë për parashikimit e konsumit të energjisë. Ku të dhënat e matura për periudhën kohore më shumë se një vit përmbajnë kryesisht vlerat e konsumit gjatë orëve të punës për një ndërtesë institucionale. Metoda e implementuar parashikon konsumin për 6 orët në vijim duke përmirësuar ndjeshëm performancën, kjo referuar punimeve të publikuara më parë. Kontributi në këtë punim qëndron në propozimin e një metode të re ngjashmërie dhe në përmirësimin e adaptimit të zgjidhjeve që vijnë nga shembujt ekzistues. Një kontribut i rëndësishëm po ashtu është edhe aplikimi i parashikimit online të konsumit me metodën tonë.

Analiza e qëndrueshmërisë së një numri më të gjerë të metodave duke arritur në konkluzione më domethënëse është në fokusin e Kapitullit 5. Aplikimi i metodave shumë të përdorshme si rrjetat neurale dhe regresioni me vektorët ndihmës krahasohen me metoda më pak të gjendura në literaturë si pemët e regresionit, pylli i rastësishëm dhe me arsyetimin e bazuar tek rastet. Implementimi i metodave përdor të dhëna të matura për 5 ndërtesa, ku vetëm dy janë parametrat e matura. Nga të dhënat e matura nxjerrim informacion shtesë i cili përmirëson performancën e modeleve. Metodat e inteligjencës artificiale aplikohen në dy skenarë kryesorë: i pari ai i parashikimit statik “offline” dhe ai i dytë ai i parashikimit dinamik “online”.

Sipas parashikimit “offline” kemi tre skenarë të ndryshëm të zgjedhur sipas ndarjes së të dhënave. Nga rezultatet e marra arrijmë në përfundimin se pylli i rastësishëm mbizotëron mbi metodat e tjera si dhe metodat e bazuara tek modeli fitojnë kundrejt asaj të bazuar tek shembujt. Skenari i dytë motivohet nga matja e të dhënave në kohë reale, ku tre nga metodat e përdorura njihen si metoda të bazuara tek shembujt përkundrejt njëres të bazuar tek modeli. Theksojmë se implementimi i metodave tradicionale të aplikuara në skenarin statik në literaturë është shumë i pakët. Rezultate e arritura na motivojnë se implementimi ka sens praktik ku gabimi tenton të konvergjojë drejt gabimit statik me rritjen e të dhënave.

Kontributi tjetër i kësaj teze është parashikimi me interval, ku nga përdorimi i parashikimi të konsumit të energjisë për orën që vijon arrijmë të prodhojmë intervale parashikimi të cilët përmbajnë vlerën e matur

brenda një niveli besueshmërie (konfidence) të paracaktuar. Rezultatet e marra nga implementimi i dy algoritmave standard të inteligjencës artificiale prodhojnë intervale të cilët janë domethënës kundrejt diapazonit të të dhënave. Rezultatet motivojnë përdorimin e metodës në parashikimin e konsumit të energjisë.

7.1 Vazhdimësia

Aplikimi i parashikimit të konsumit të energjisë në ndërtesa, rezulton me shumë pyetje shkencore që mbeteten të hapura në përfundimin e këtij punimi. Kërkimi i ardhshëm do të tentojë të fokusohet në këto aspekte kyçe:

1. Zhvillimi i një modeli ABR i cili parashikon konsumin tek ndërtesat që nuk kanë të dhëna të matura duke transferuar modelin e mësuar tek një ndërtesë tek një ndërtesë tjetër e ngjashme, ndryshe njihet si transferim i modelit të mësuar (transfer learning).
2. Analiza e teknikave të tjera për përdorimin e karakteristikave të të dhënave duke përdorur intervale të tjera matjeje si (15 min, 30 min, etj).
3. Ndërthurja e algoritmave të parashikimit me sistemet e menaxhimit të energjisë duke parë realisht përmirësimet që mund të sjellë në konsum metoda.
4. Gjetja dhe grumbullimi i të dhënave të tjera për konsumin e energjisë në terma afatmesëm dhe afatgjatë.
5. Optimizimi dhe përdorimi i parashikimit me interval në skenarin online dhe me algoritma të tjerë mësimi.
6. Analiza e ndërthurjes së konsumit të energjisë me faktorë të tjerë si popullimi dhe mënyra e punës së ndërtesave.
7. Parashikimi i konsumit të energjisë për konsumatorët rezidencialë në stinë të ndryshme duke u nisur nga faturat e energjisë elektrike.

Shtesa A

Parashikimi i konsumit me ABR

A.1 Kodi kryesor për parashikuesin ABR

```
1 %Deklarimi i parametrave
2
3 alpha = 0.3;    % Perdoret per vleresimin e peshes te
   seciles ngjashmeri
4 counter = 0;
5 errorSum = 0;
6
7 % Leximi i te dhenave
8
9 load('data\cbFinal2013.mat');
10 load('data\cbFinal2014.mat');
11 kp=443;        % Koeficient i cili tregon numrin e rasteve
   te perzgjedhura per trainim ...
12                % Ndryshimi i tij ndryshon
   raportin trainim/testim
13 ks=kp+1;
14 cbFinal=[cbFinal2013 cbFinal2014];
15 errors = [];
16
17 CB = cbFinal(1:kp); % baza e rasteve
18 trues = [];
19 predictions = [];
20 trueAll = [];
21 predictionAll = [];
22 for i = ks:length(cbFinal)
```

```

23     presentCase = cbFinal{i};
24     simCases = similarCasesOnline(presentCase, CB);
25     if length(simCases) ~= 0
26         prediction = predictSplit(presentCase, CB, alpha);
27     else prediction=repmat(presentCase(3,17),6,1);
28     end
29     true = presentCase(4:9,17);
30     true1=presentCase(4:9,17)';
31     trues=[trues; true1];
32     trueAll = [trueAll; true];
33     predictionAll = [predictionAll; prediction];
34     predictions = [predictions; prediction'];
35     error = CVRMSE(true, prediction);
36     errors = [errors, error];
37     CV_Error=mean(errors);
38     counter=counter+1;
39 end

```

A.1.1 Ndarja e ditëve në raste me raste prej 9 orësh

```

1  cbFinal2014 = struct([]);
2  counter = 1;
3  for i=1:size(cbDays2014,2)
4      nextCase = cbDays2014(i);
5      nextCaseSize = size(nextCase{1},1);
6      if nextCaseSize == 9
7          cbFinal2014{counter} = nextCase{1};
8          counter = counter + 1;
9      elseif nextCaseSize == 10
10         cbFinal2014{counter} = nextCase{1}(1:9,:);
11         counter = counter + 1;
12         cbFinal2014{counter} = nextCase{1}(2:10,:);
13         counter = counter + 1;
14     end
15     if nextCaseSize == 11
16         cbFinal2014{counter} = nextCase{1}(1:9,:);
17         counter = counter + 1;
18         cbFinal2014{counter} = nextCase{1}(2:10,:);
19         counter = counter + 1;
20         cbFinal2014{counter} = nextCase{1}(3:11,:);

```

```
21         counter = counter + 1;
22     end
23 end
```

A.1.2 Ndërtimi i bazës së rasteve

```
1 function createCaseBase()
2 S = load('data/dataabr.mat');
3 M = S.dataabr;
4     CaseBase2014 = struct([]);
5     for i=1:size(M,1)-8
6         nextCase = M(i:i+8,:);
7         CaseBase2014{i} = nextCase;
8     end
9 return
```

A.1.3 Vlerësimi i ngjashmërisë midis variablave të rasteve

```
1 function VS = variableSimilarity(presentCase, pastCase,
2     inputNo, d_min, d_max)
3 if inputNo == 3 || inputNo == 4 % Vlerat e parashikuara te
4     ajrit jashte
5     weights = [6/7 5/7 4/7 3/7 2/7 1/7]; %pesha qe
6     vendosen per rendesine e kohes se matjes per 6 oret
7     e ardhshme
8     d = sqrt( (weights(1)*(presentCase(4,inputNo+shift)-
9         pastCase(4,inputNo+shift)))^2 ...
10         + (weights(2)*(presentCase(5,inputNo+shift)-
11         pastCase(5,inputNo+shift)))^2 ...
12         + (weights(3)*(presentCase(6,inputNo+shift)-
13         pastCase(6,inputNo+shift)))^2 ...
14         + (weights(4)*(presentCase(7,inputNo+shift)-
15         pastCase(7,inputNo+shift)))^2 ...
16         + (weights(5)*(presentCase(8,inputNo+shift)-
17         pastCase(8,inputNo+shift)))^2 ...
18         + (weights(6)*(presentCase(9,inputNo+shift)-
19         pastCase(9,inputNo+shift)))^2 ...
20         ) / sum(weights);
21 else
```

```

13     weights = [2/3 1/3]; % pesha qe vendosen per rendesine
        e kohes se matjes
14     d = sqrt( (weights(1)*(presentCase(2,inputNo+shift)-
        pastCase(2,inputNo+shift)))^2 ...
15             + (weights(2)*(presentCase(1,inputNo+shift)-
        pastCase(1,inputNo+shift)))^2 ...
16             ) / sum(weights);
17 end
18 if d <= d_min
19     VS = 1;
20 elseif d_max <= d
21     VS = 0;
22 else
23     VS = (d_max-d)/(d_max-d_min);
24 end
25 return

```

A.1.4 Ngjashmeria midis rasteve

```

1 function CS = caseSimilarity(presentCase, pastCase, alpha)
2 variableWeights = [ 0.1961 0.1540 0.0001 0.0001 0.0009
        0.1065 0.1075 0.0001 0.3064 0.1284];
3 d_min = [ 2, 10, 2, 2, 2, 2, 2, 5, 2, 5];
4 d_max = [ 6, 25, 6, 6, 6, 6, 15, 30, 15, 30];
5 CS = 0;
6
7 for i=1:10
8     CS = CS + (variableSimilarityFirstOrder2(presentCase,
        pastCase, i, d_min(i), d_max(i), alpha) *
        variableWeights(i));
9 end
10 CS = CS/sum(variableWeights);
11 return

```

A.1.5 Përshtatja e zgjidhejs për parashikim

```

1 function prediction = predictSplit(presentCase, CB, alpha,
        K)
2 % find similar cases
3 sc = similarCasesOnline(presentCase, onlineCB);

```

```
4 similarities = zeros(1,size(sc,2));
5 % initialize prediction
6 prediction = [];
7 counter = 0;
8 denominator = 0;
9 for i=1:size(sc, 2)
10     nextCase = sc{i};
11     similarities(i) = caseSimilarity(presentCase, nextCase
        , alpha);
12 end
13 [~,I] = sort(similarities, 'descend');
14 similaritiesSorted = similarities(I);
15 sc1 = sc(I);
16 if size(sc1,2) >= 50
17     sc1 = sc1(1,1:50);
18 else sc1 = sc1;
19 end
20 for i=1:size(sc1, 2)
21     nextCase = sc1{i};
22     nextSim = caseSimilarity(presentCase, nextCase,
        alpha);
23
24     x1=mean(nextCase(1:3,17));
25     x2=mean(presentCase(1:3,17));
26     adaptFactor=((x2-x1)/(x1+x2));
27     powerPastC = nextCase(4:9,17);
28     predictionTemp = powerPastC +adaptFactor*
        powerPastC;
29
30     counter = counter + 1;
31     prediction = [prediction, predictionTemp*nextSim];
32     denominator = denominator + nextSim;
33 end
34
35 nominator = sum(prediction,2);
36 prediction = nominator/denominator;
37 return
```

A.2 Parashikuesi bazë i cili parashikon vlerën e konsumit sa mesatarja e gjithë pjesës së trajnimit

```
1 X=mean(TrainingData(1:trainIndex,19); % llogarit mesataren
    e te gjithë te dhenave te perdorura per training
2 K=1698; % numri i vlerave te cilat do te parashikohen
3 prediction = repmat(X, K, 1);
4 true2 = temp1; % vlera reale sa vlera momentale
5 CVRMSE(prediction, true2)
```

A.3 Parashikuesi bazë i cili parashikon vlerën e konsumit sa mesatarja e tre orëve të mëparshme

```
1 % path to the test case base
2 S = load('data/cbFinal2014.mat');
3 Q = load('data/cbFinal2013.mat');
4 cb2014 = S.cbFinal2014;
5 cb2013=Q.cbFinal2013;
6 %
7 cbfinal= cb2014;
8 counter = 0;
9 errorSum = 0;
10
11 temp1b = [];
12 temp2b = [];
13
14 for i = 1:size(cbfinal,2)
15     presentCase = cbfinal{i};
16     X=mean(presentCase(1:3,17));
17     prediction = repmat(X,6,1);
18     if (~isempty(prediction))
19         counter = counter + 1;
20         true = presentCase(4:9,17);
21         error = CVRMSE(true, prediction);
22         temp1b = [temp1b>true];
```



```
23         temp2b = [temp2b; prediction];
24         errorSum = errorSum + error;
25     end
26 end
27 finalError = errorSum/counter
```

A.4 Parashikuesi bazë i cili parashikon vlerën e konsumit sa ora e mëparshme

```
1 % Lexohen te dhenat nga baza e te dhenave
2 S = load('data/cbFinal2014.mat');
3 cb2014 = S.cbFinal2014;
4 %
5 counter = 0;
6 errorSum = 0;
7
8 cbfinal= cb2014;
9 temp1b = [];
10 temp2b = [];
11
12 for i = 1:size(cbfinal,2)
13     presentCase = cbfinal{i};
14     Y=presentCase(3,17);
15     prediction = repmat(Y,6,1);
16     if(~isempty(prediction))
17         counter = counter + 1;
18         true = presentCase(4:9,17);
19         error = CVRMSE(true, prediction);
20         temp1b = [temp1b;true];
21         temp2b = [temp2b;prediction];
22         errorSum = errorSum + error;
23     end
24 end
25 finalError = errorSum/counter
```

A.5 Parashikimi duke përdorur rrjetat neurale

A.5.1 Gjenerimi i të dhënave të trajnimit

```

1 load( '\data\cbFinal2013.mat' );
2 NNTrainingInputs2013 = [];
3 NNTrainingOutputs2013 = [];
4 for i=1:size(cbFinal2013,2)-286
5     nextExample = cbFinal2013{i};
6     inputs = nextExample(1:3,7:end);
7     inputs = reshape(inputs,[1,33]);
8     NNTrainingInputs2013 = [NNTrainingInputs2013;inputs];
9     outputs = nextExample(4:9,17)';
10    NNTrainingOutputs2013 = [NNTrainingOutputs2013;outputs
11                               ];
11 end

```

A.5.2 Gjenerimi i të dhënave të testimit

```

1 S=load( '\data\cbFinal2014.mat' );
2 Q=load( '\data\cbFinal2013.mat' );
3 cb2013 = Q.cbFinal2013;
4 cb2014test = S.cbFinal2014;
5 cbFinal2014=[cb2013(158:end) cb2014test];
6 NNInputs2014 = [];
7 NNOutputs2014 = [];
8 for i=1:size(cbFinal2014,2)
9     nextExample = cbFinal2014{i};
10    inputs = nextExample(1:3,7:end);
11    inputs = reshape(inputs,[1,33]);
12    NNInputs2014 = [NNInputs2014;inputs];
13    outputs = nextExample(4:9,17)';
14    NNOutputs2014 = [NNOutputs2014;outputs];
15 end

```

A.5.3 Testimi i rrjetës neurale

```

1 %Inicializimi i rrjetes neurale
2
3 net = fitnet(10); % numri i neuroneve ne shtresen e
   fshehur

```

```

4 net.trainParam.epochs = 1000;
5 % net.trainParam.showWindow = 0;
6 net.trainParam.max_fail = 1000;
7
8 % Trainimi i rrjetes
9
10 NNTrainingInputs2013=NNTrainingInputs2013(1:250,:);
11 NNTrainingOutputs2013=NNTrainingOutputs2013(1:250,:);
12 [net,~] = train(net,NNTrainingInputs2013',
    NNTrainingOutputs2013');
13
14 %% Testimi i rrjetes me te dhenat e reja
15
16 testOutputs = net(NNInputs2014');
17 trues = NNOutputs2014'; % vlerat e parashikuara te
    konsumit
18
19 %% Llogaritja e gabimit CVRMSE
20
21 cvrmseError = 0;
22 for i=1:size(testOutputs,2)
23     cvrmseError = cvrmseError + CVRMSE(testOutputs(:,i),
        trues(:,i));
24 end
25 cvrmseError = cvrmseError/size(testOutputs,2)

```

Shtesa B

Algoritmat e parashikimit 'offline' për 5 ndërtesa

B.0.1 Gjenerimi i matricës së të dhënave nga ato origjinale

```
1 function [X, dates, labels] = genmatrix(Data, term)
2 dates=Data.Date+(Data.Time-1)/24;
3 % create matrix of input variables
4
5 prev24HrAveLoad = filter(ones(1,24)/24, 1, Data.Power);
6 prevDaySameHourLoad = [NaN(24,1); Data.Power(1:end-24)];
7 prevWeekSameHourLoad = [NaN(168,1);Data.Power(1:end-168)];
8 prevhourLoad = [NaN(1,1);Data.Power(1:end-1)];
9 prevTwoHoursLoad = [NaN(2,1);Data.Power(1:end-2)];
10 prevThreeHoursLoad = [NaN(3,1);Data.Power(1:end-3)];
11
12 if strcmpi(term, 'long', 4);
13     % Long Term Forecast Predictors
14     X = [Data.Time Data.Day Data.Dayofweek Data.Temp_F_
15         daily5dayLowAve ];
16     labels = {'Hour', 'Day', 'Dayofweek', 'Temeperature',
17             'daily5dayLowAve'};
18 else
19     % Short Term
20     X = [Data.Time Data.Indicator1 Data.Indicator2 Data.
21         Indicator3 Data.Indicator4 ...
22         Data.Indicator5 Data.Indicator6 Data.Indicator7
23         Data.Temperature prevWeekSameHourLoad
24         prevDaySameHourLoad ...
25         prev24HrAveLoad prevhourLoad prevTwoHoursLoad
26         prevThreeHoursLoad];
27     labels = {'Hour', 'Indicator1', 'Indicator2', '
28             Indicator3', 'Indicator4', ...
```

```

22         'Indicator5', 'Indicator6', 'Indicator7', '
           Temperature', 'PrevWeekSameHourLoad', '
           prevDaySameHourLoad', ...
23         'prev24HrAveLoad', 'prevhourLoad', '
           prevTwoHoursLoad', 'prevThreeHoursLoad'}];
24
25 end
26 return

```

B.0.2 Testimi i performancës së rrjetës neurale

```

1 A=load('Final_DATA_CSV\Building_1.mat');
2 NN_Train=A.Building1(:,1:end-1);
3 NN_Out=A.Building1(:,end);
4 trainind=6550;           %indeksi i trainimit per 9 muaj
5 % trainind=7294;       %indeksi i trainimit per 10 muaj
6 % trainind=8014;       %indeksi i trainimit per 11 muaj
7 %%
8 errors=[];
9 MAPEs=[];
10 predicteds=[];
11 MAEs=[];
12 NNtrainInputs=NN_Train(1:trainind,:);
13 NNtrainOutputs=NN_Out(1:trainind,:);
14 NNtestInputs=NN_Train(trainind+1:length(NN_Train),:);
15 NNOutputs=NN_Out(trainind+1:length(NN_Train),:);
16
17 net = fitnet(9);
18 net.trainParam.epochs = 1000;
19 net.trainParam.max_fail = 1000;
20 [net,~] = trainbfg(net,NNtrainInputs', NNtrainOutputs');
21 testOutputs =sim (net,NNtestInputs');
22 predicted=testOutputs';
23 trues = NNOutputs;
24 m=size(predicted,2);
25 error = zeros(1, m);
26 for t = 1: m
27     error(t) = ((sqrt(sum((predicted(:, t) - trues(:, t))
           .^2)/(size(predicted,1)-1)))/(mean(predicted(:, t))
           ))*100;

```

```

28 end
29
30 err = trues-predicted;
31 y=mean(predicted);
32 n=length(predicted);
33 RMSE = ((sqrt((sum(trues-predicted).^2)))/(n));
34 CVRMSE = (RMSE/y)*100;
35 errpct = abs(err)./trues*100;
36 MAE = mean(abs(err));
37 MAPE = mean(errpct(~isinf(errpct)));
38
39
40 if trainind==6550
41 filename='ANNoutput_9_Months.mat';
42 else if trainind==7294
43 filename='ANNoutput_10_Months.mat';
44     else
45         filename='ANNoutput_11_Months.mat';
46     end
47 end
48
49 errors =error;
50 MAPEs=MAPE;
51 predicteds=predicted;
52 MAEs=MAE;
53
54 if exist(filename,'file')
55     load(filename)
56
57 errors = [errors , error];
58 MAPEs=[MAPEs,MAPE];
59 predicteds=[predicteds , predicted];
60 MAEs=[MAEs,MAE];
61 save(filename,'errors','MAEs','MAPEs','predicteds');
62 else
63
64 save(filename,'errors','MAEs','MAPEs','predicteds');
65 end

```

B.0.3 Testimi i performancës së pemës së regresionit

```

1 A=load('Final_DATA_CSV\Building_1.mat');
2 Data_Train=A.Building1(:,1:end-1);
3 Data_Out=A.Building1(:,end);
4 trainind=6550;           %indeksi i trainimit per 9 muaj
5 % trainind=7294;       %indeksi i trainimit per 10 muaj
6 % trainind=8014;       %indeksi i trainimit per 11 muaj
7 %%
8 errors=[];
9 MAPEs=[];
10 predicteds=[];
11 MAEs=[];
12
13 trainX=Data_Train(1:trainind,:);
14 trainY=Data_Out(1:trainind,:);
15 testX=Data_Train(trainind+1:length(Data_Train),:);
16 testY=Data_Out(trainind+1:length(Data_Train),:);
17
18 %% THERRET HIPERPARAMETRAT
19 if trainind==6550
20 load('Tree_HyperParameters_9_months.mat');
21 else if trainind==7294
22 load('Tree_HyperParameters_10_months.mat');
23     else
24     load('Tree_HyperParameters_11_months.mat');
25     end
26 end
27
28 %%
29 minleaf=Mdl.ModelParameters.MinLeaf;
30 minpar=Mdl.ModelParameters.MinParent;
31 mdlTree = fitrtree(trainX,trainY,'MinParentSize',minpar,'
    MinLeafSize',minleaf,'prune','on');
32 Y = predict(mdlTree,testX);
33 %
34 err = testY-Y;
35 y=mean(Y);
36 n=length(Y);

```

```

37 errpct = abs(err)./testY*100;
38 MAE = mean(abs(err));
39 MAPE = mean(errpct(~ isinf(errpct)));
40 %%
41 error = ((sqrt(sum((Y(:) - testY(:)).^2)/(size(Y,1)-1)))/(
    mean(Y(:))))*100;
42
43 if trainind==6550
44 filename='Tree_Output_9_Months.mat';
45 else if trainind==7294
46 filename='Tree_Output_10_Months.mat';
47     else
48     filename='Tree_Output_11_Months.mat';
49     end
50 end
51
52 errors =error;
53 MAPEs=MAPE;
54 predicteds=Y;
55 MAEs=MAE;
56
57 if exist(filename, 'file')
58     load(filename)
59
60 errors = [errors, error];
61 MAPEs=[MAPEs,MAPE];
62 predicteds=[predicteds, Y];
63 MAEs=[MAEs,MAE];
64 save(filename, 'errors', 'MAEs', 'MAPEs', 'predicteds');
65 else
66 save(filename, 'errors', 'MAEs', 'MAPEs', 'predicteds');
67 end

```

B.0.4 Testimi i performancës së pyllit të rastësishëm

```

1 % Lexohen te dhenat nga baza
2 A=load('Final_DATA_CSV\Building_1.mat');
3 Data_Train=A.Building1(:,1:end-1);
4 Data_Out=A.Building1(:,end);
5 trainind=6550;           %indeksi i trainimit per 9 muaj

```



```

6 % trainind=7294;           %indeksi i trainimit per 10 muaj
7 % trainind=8014;           %indeksi i trainimit per 11 muaj
8 %%
9  errors = [];
10 MAPEs = [];
11 predicted = [];
12 MAEs = [];
13
14 trainX = Data_Train(1:trainind, :);
15 trainY = Data_Out(1:trainind, :);
16 testX = Data_Train(trainind+1:length(Data_Train), :);
17 testY = Data_Out(trainind+1:length(Data_Train), :);
18 %% Ngarkohen hiperparametrat
19 if trainind==6550
20 load('Tree_HyperParameters_9_months.mat');
21 else if trainind==7294
22 load('Tree_HyperParameters_10_months.mat');
23     else
24     load('Tree_HyperParameters_11_months.mat');
25     end
26 end
27 %%
28 minleaf = Mdl.ModelParameters.MinLeaf;
29 mdlTreeBag = TreeBagger(100, trainX, trainY, 'method', '
    regression', ...
30 'oobpred', 'on', 'MinLeafSize', minleaf);
31 Y1 = predict(mdlTreeBag, testX);
32 %%
33 err = testY - Y1;
34 y = mean(Y1);
35 errpct = abs(err) ./ testY * 100;
36 MAE = mean(abs(err));
37 MAPE = mean(errpct(~isinf(errpct)));
38 error = ((sqrt(sum((Y1(:) - testY(:)).^2) / (size(Y1,1) - 1)))
    / (mean(Y1(:)))) * 100;
39 %%
40 if trainind==6550
41 filename = 'Bag_Tree_Output_9_Months.mat';
42 else if trainind==7294

```

```

43 filename='Bag_Tree_Output_10_Months.mat';
44     else
45         filename='Bag_Tree_Output_11_Months.mat';
46     end
47 end
48
49 errors =error;
50 MAPEs=MAPE;
51 predicteds=Y1;
52 MAEs=MAE;
53
54 if exist(filename, 'file')
55     load(filename)
56
57 errors = [errors, error];
58 MAPEs=[MAPEs,MAPE];
59 predicteds=[predicteds, Y1];
60 MAEs=[MAEs,MAE];
61 save(filename, 'errors', 'MAEs', 'MAPEs', 'predicteds');
62 else
63
64 save(filename, 'errors', 'MAEs', 'MAPEs', 'predicteds');
65 end

```

B.0.5 Testimi i performancës së regresionit me vektorët ndihmës

```

1 % Lexohen te dhenat nga baza
2 A=load('Final_DATA_CSV\Building_1_mod.mat');
3 X=A. Building1(:,1:end-1);
4 [X, mean_X, std_X] = standardize(X);
5 Data_Train=X;
6 Data_Out=A. Building1(:,end);
7 trainind=6550;           %indeksi i trainimit per 9 muaj
8 % trainind=7294;         %indeksi i trainimit per 10 muaj
9 % trainind=8014;         %indeksi i trainimit per 11 muaj
10 %%
11
12

```

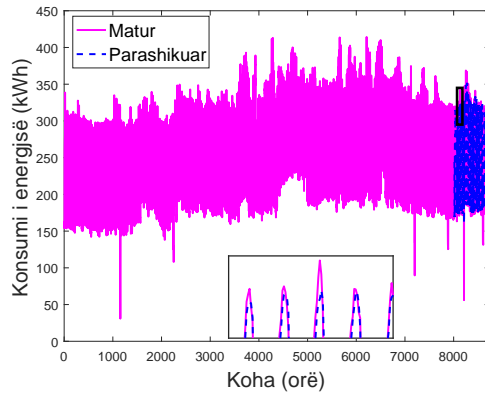
```

13 trainX=Data_Train(1:trainind ,:);
14 trainY=Data_Out(1:trainind ,:);
15 testX=Data_Train(trainind+1:length(Data_Train),:);
16 testY=Data_Out(trainind+1:length(Data_Train),:);
17 %%
18 sigma=Mdl.ModelParameters.KernelScale; % parametreat e
    SVR
19 e=Mdl.ModelParameters.Epsilon; % parametreat e
    SVR
20 C=Mdl.ModelParameters.BoxConstraint; % parametreat e
    SVR
21 SVMmdl = fitrsvm(trainX,trainY,'KernelFunction','linear',
    'KernelScale','auto');
22 Ypred= predict(SVMmdl,testX);
23
24 %%
25 err = testY-Ypred;
26 y=mean(Ypred);
27 n=length(Ypred);
28 errpct = abs(err)./testY*100;
29 MAE = mean(abs(err));
30 MAPE = mean(errpct(~isinf(errpct)));
31 error = ((sqrt(sum((Ypred(:) - testY(:)).^2)/(size(Ypred
    ,1)-1)))/(mean(Ypred(:))))*100;
32 %save('SVM_Output_11_Months.mat', 'error', 'MAE', 'MAPE',
    Ypred')
33
34
35 if trainind==6550
36 filename='SVM_Output_9_Months.mat';
37 else if trainind==7294
38 filename='SVM_Output_10_Months.mat';
39 else
40 filename='SVM_Output_11_Months.mat';
41 end
42 end
43
44 errors =error;
45 MAPEs=MAPE;

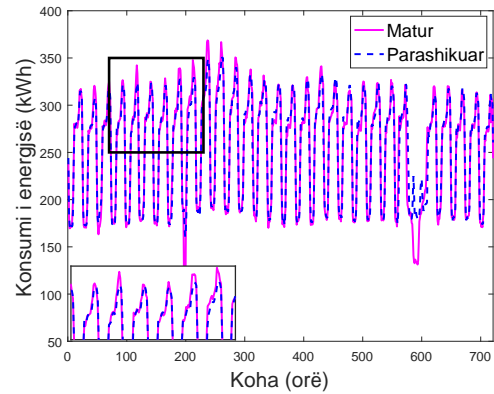
```

```
46 predicteds=Ypred;
47 MAEs=MAE;
48
49 if exist(filename,'file')
50     load(filename)
51
52 errors = [errors, error];
53 MAPEs=[MAPEs,MAPE];
54 predicteds=[predicteds,Ypred];
55 MAEs=[MAEs,MAE];
56 save(filename,'errors','MAEs','MAPEs','predicteds');
57 else
58
59 save(filename,'errors','MAEs','MAPEs','predicteds');
60 end
```

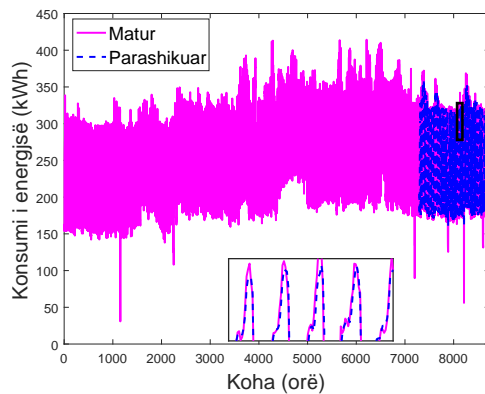
B.1 Rezultatet e parashikimit 'offline' për skenarë të ndryshëm



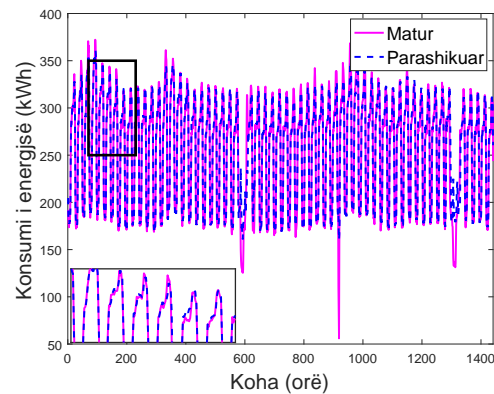
(a) Parashikimi 1 muaj



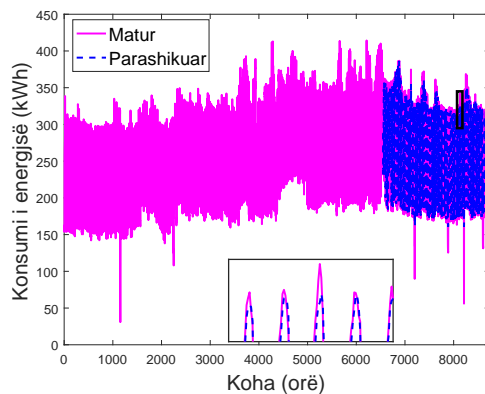
(b) Parashikimi 1 muaj detajuar



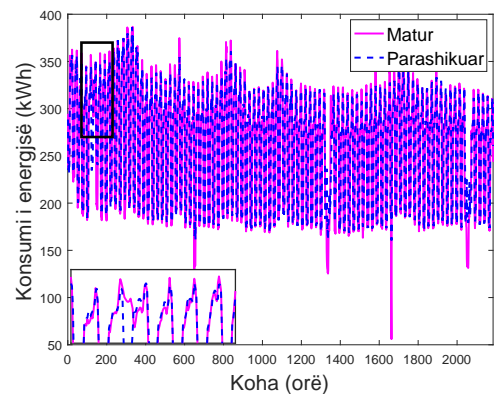
(c) Parashikimi 2 muaj



(d) Parashikimi 2 muaj detajuar

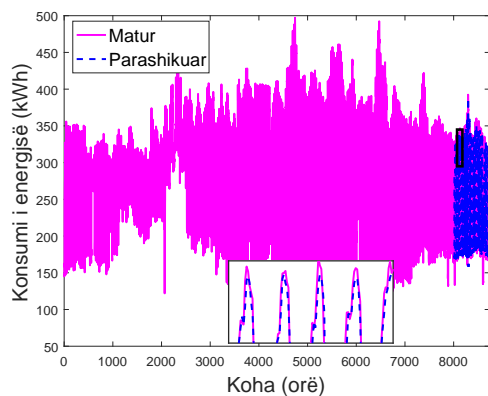


(e) Parashikimi 3 muaj

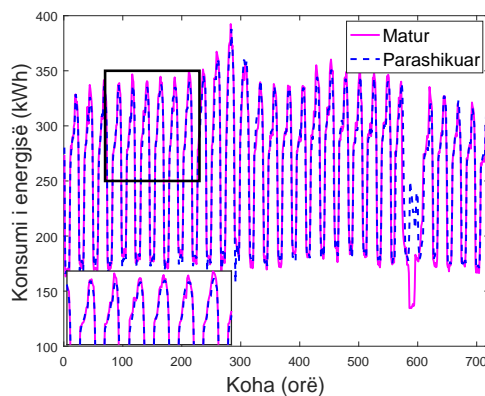


(f) Parashikimi 3 muaj detajuar

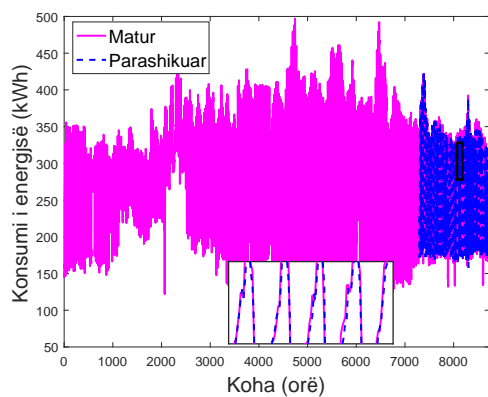
Figura B.1: Parashikimi për ndërtesën 1 me pyjet e rastësishëm



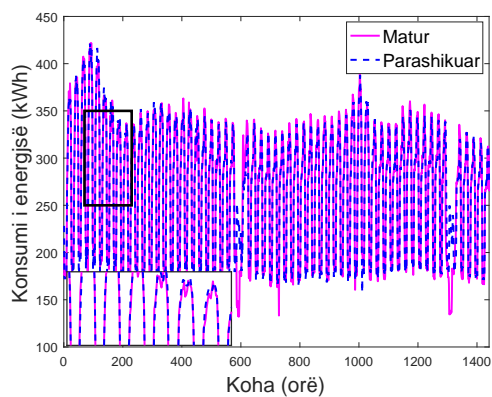
(a) Parashikimi 1 muaj



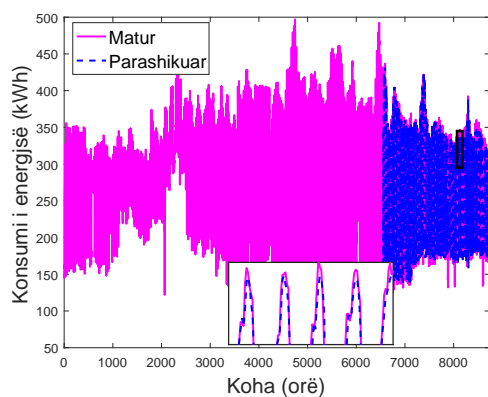
(b) Parashikimi 1 muaj detajuar



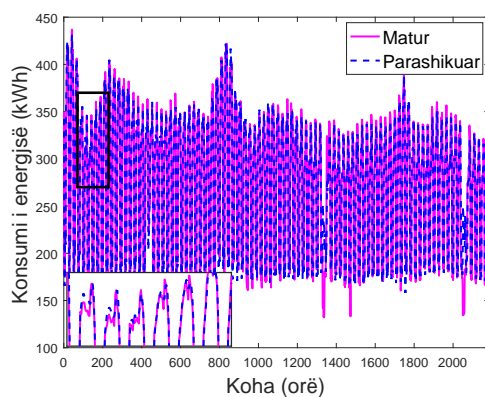
(c) Parashikimi 2 muaj



(d) Parashikimi 2 muaj detajuar

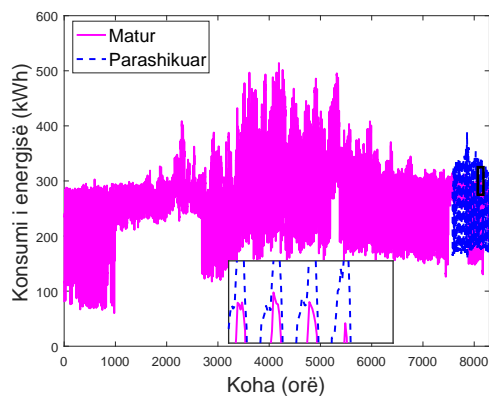


(e) Parashikimi 3 muaj

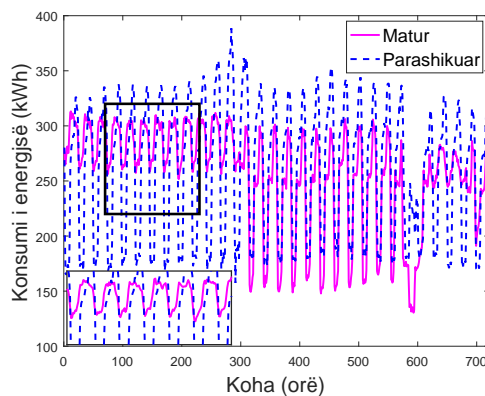


(f) Parashikimi 3 muaj detajuar

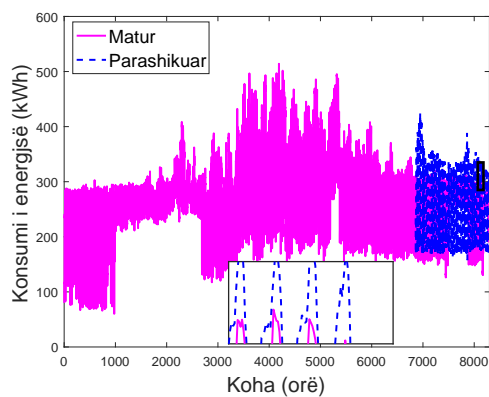
Figura B.2: Parashikimi për ndërtesën 2 me pyjet e rastësishëm



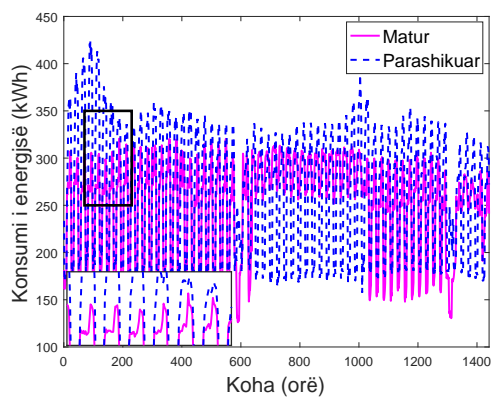
(a) Parashikimi 1 muaj



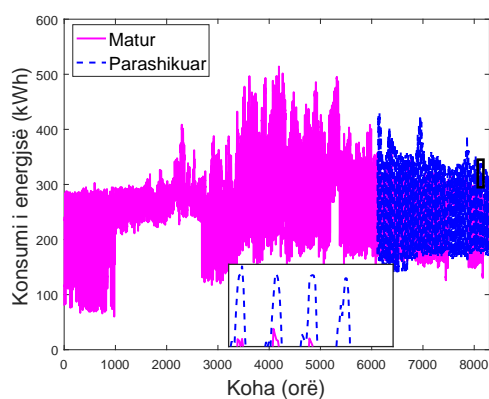
(b) Parashikimi 1 muaj detajuar



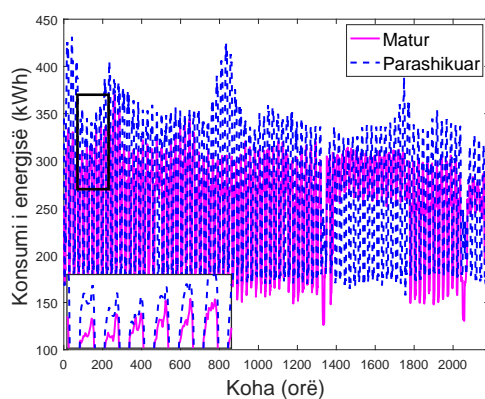
(c) Parashikimi 2 muaj



(d) Parashikimi 2 muaj detajuar

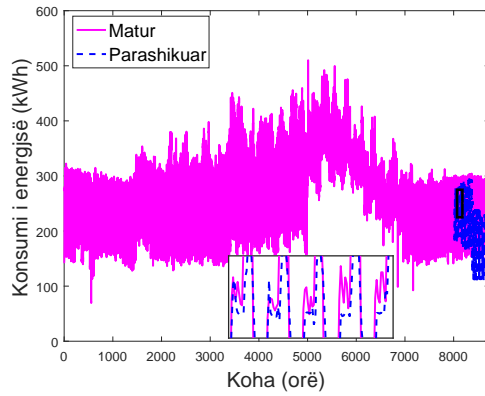


(e) Parashikimi 3 muaj

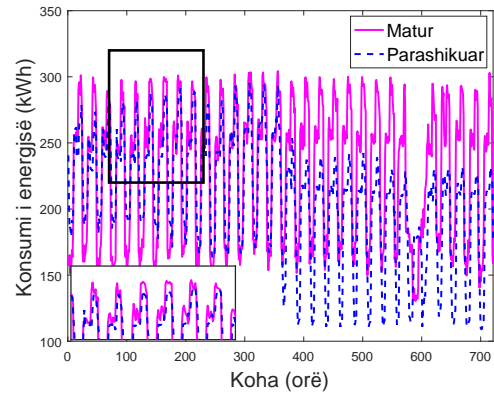


(f) Parashikimi 3 muaj detajuar

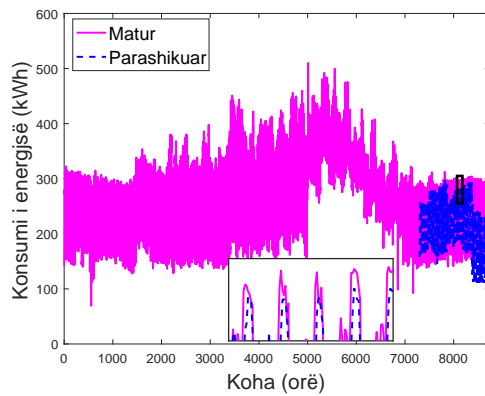
Figura B.3: Parashikimi për ndërtesën 3 me pyjet e rastësishëm



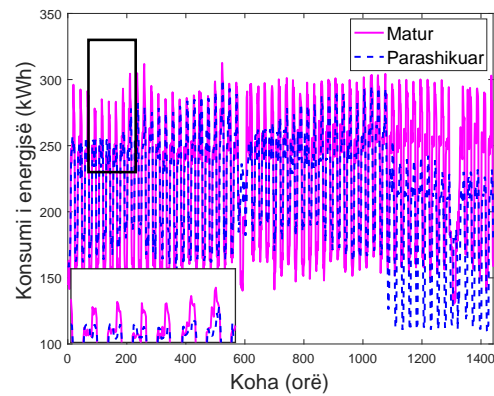
(a) Parashikimi 1 muaj



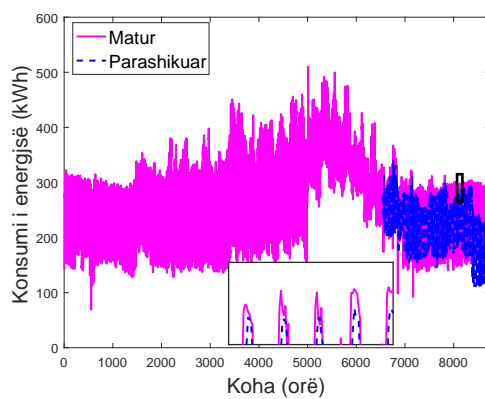
(b) Parashikimi 1 muaj detajuar



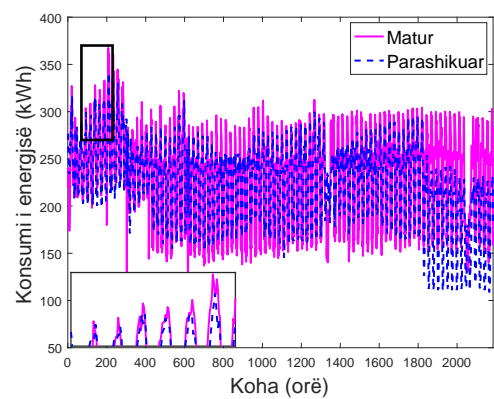
(c) Parashikimi 2 muaj



(d) Parashikimi 2 muaj detajuar

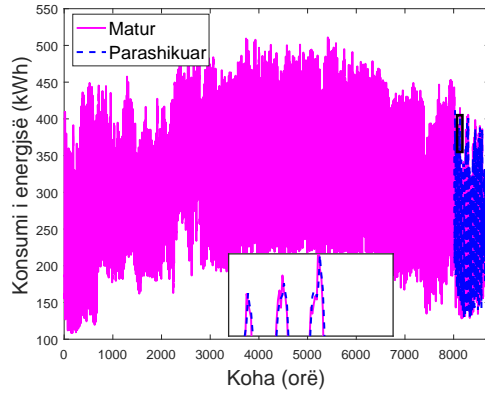


(e) Parashikimi 3 muaj

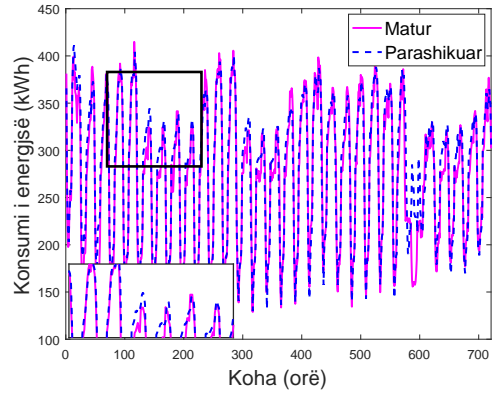


(f) Parashikimi 3 muaj detajuar

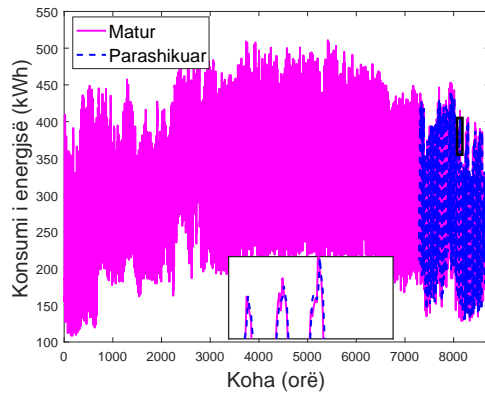
Figura B.4: Parashikimi për ndërtesën 4 me pyjet e rastësishëm



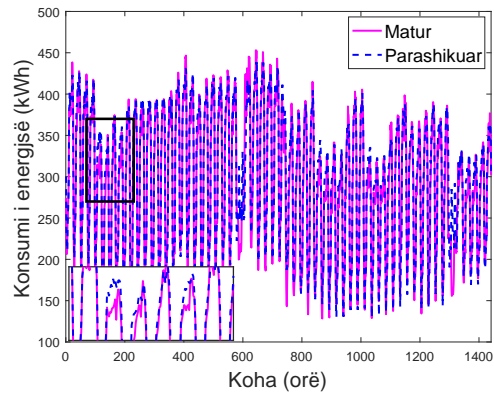
(a) Parashikimi 1 muaj



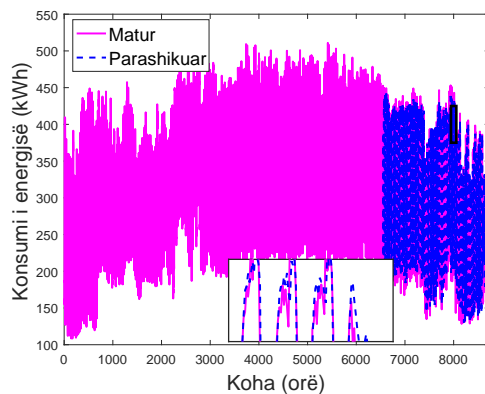
(b) Parashikimi 1 muaj detajuar



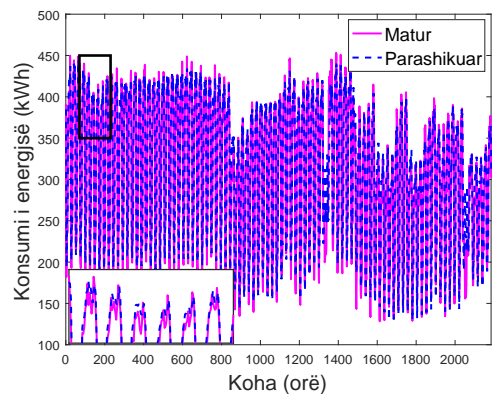
(c) Parashikimi 2 muaj



(d) Parashikimi 2 muaj detajuar



(e) Parashikimi 3 muaj



(f) Parashikimi 3 muaj detajuar

Figura B.5: Parashikimi për ndërtesën 5 me pyjet e rastësishëm

Shtesa C

Intervalet e parashikimit me pyjet e rastësishëm duke përdorur jokonformitetin e normalizuar.

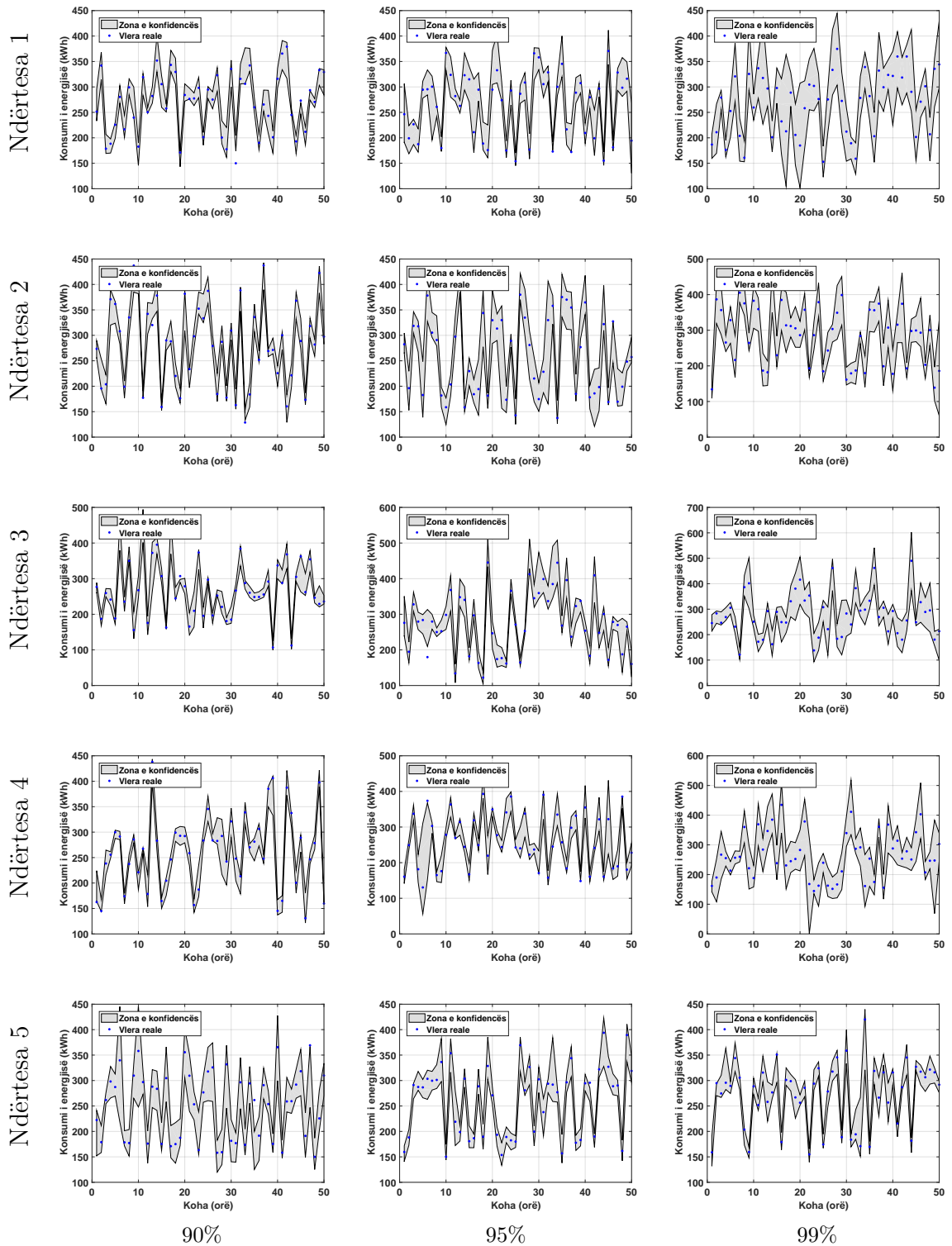


Figura C.1: Intervalet e parashikimit me pyjet e rastësishëm duke përdorur jokonformitetin e normalizuar.

Intervalet e parashikimit

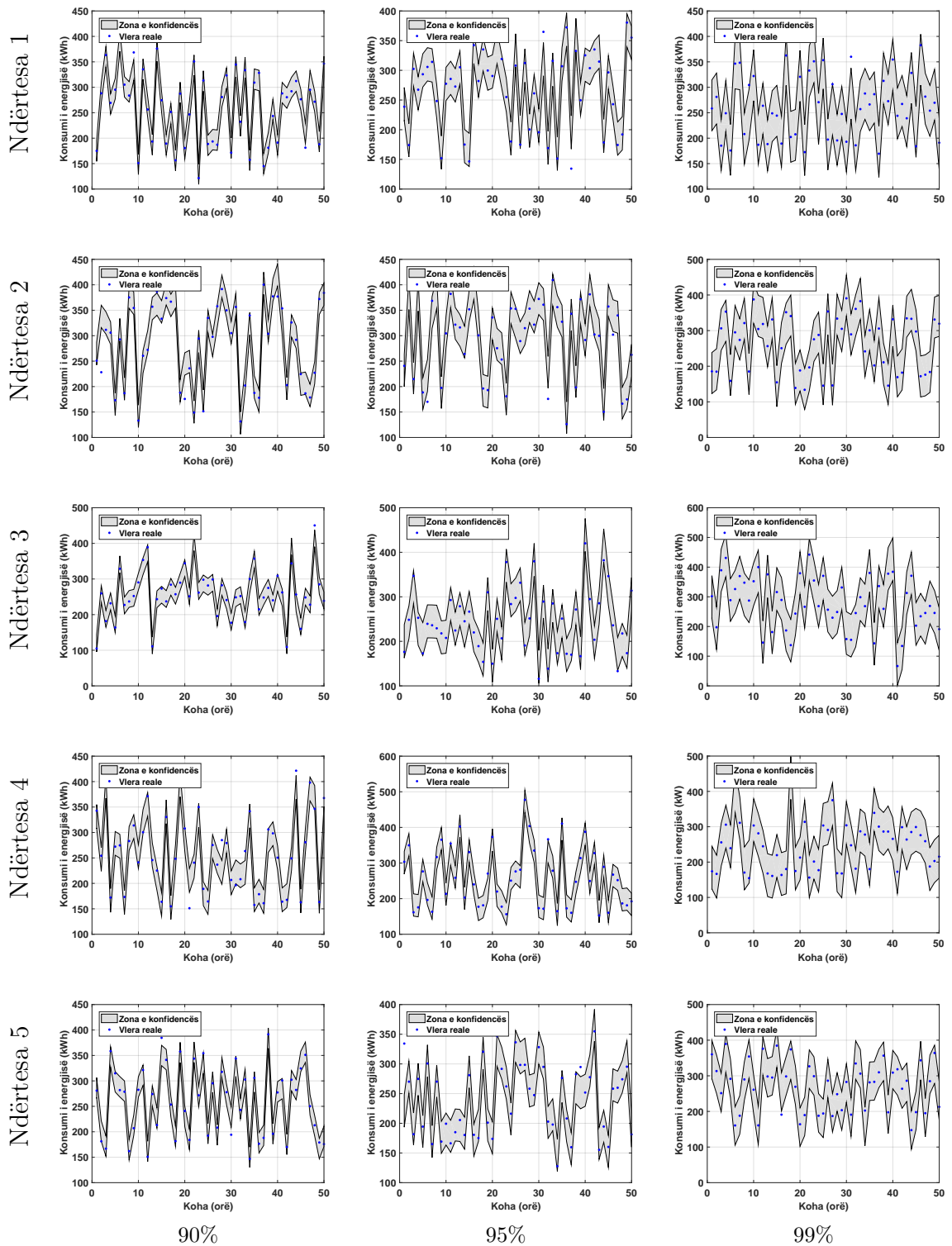


Figura C.2: Intervalet e parashikimit me pyjet e rastësishëm duke përdorur jokonformitetin jo të normalizuar.

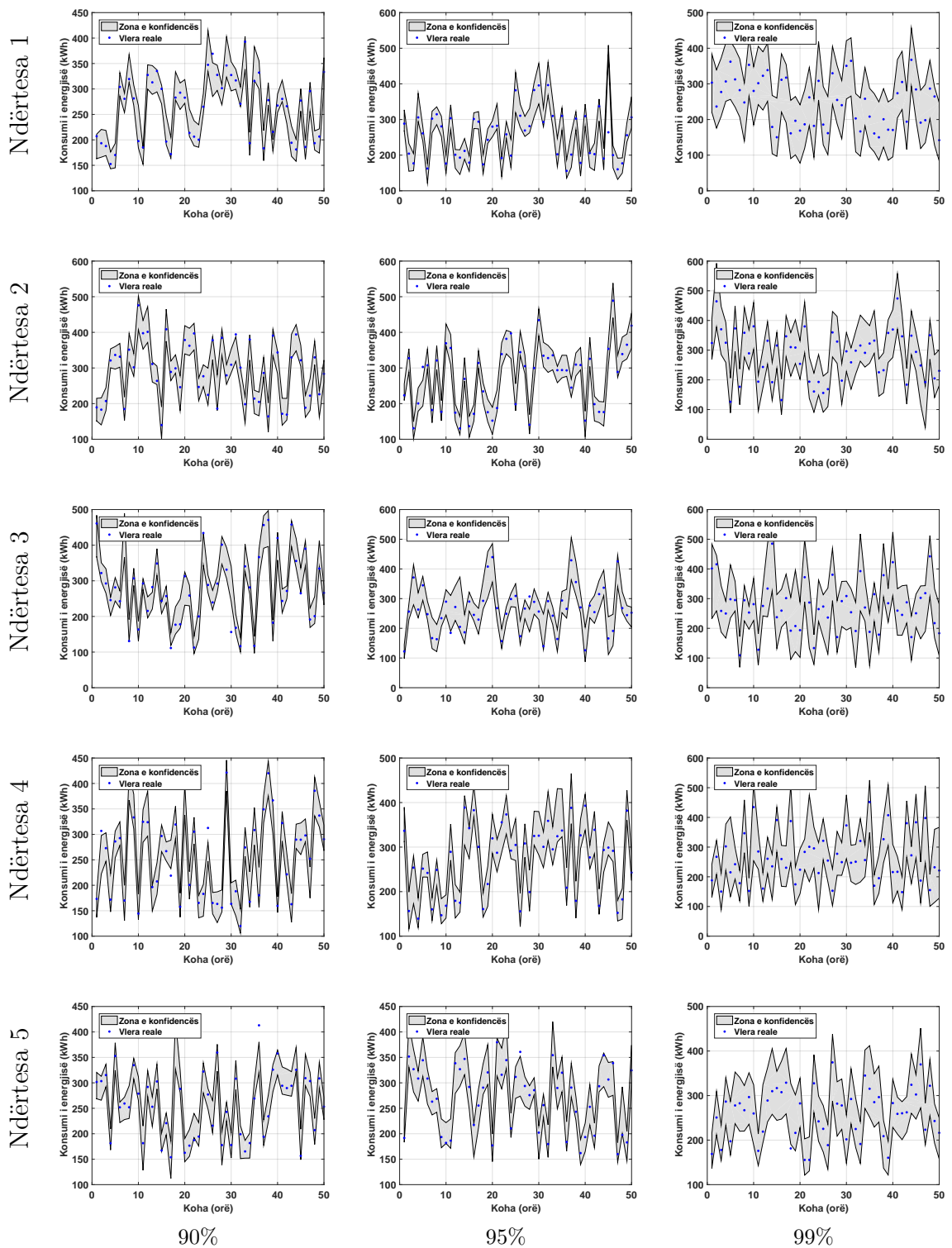


Figura C.3: Intervalet e parashikimit me rrjetat neurale duke përdorur jokonformitetin e normalizuar.

Intervalet e parashikimit

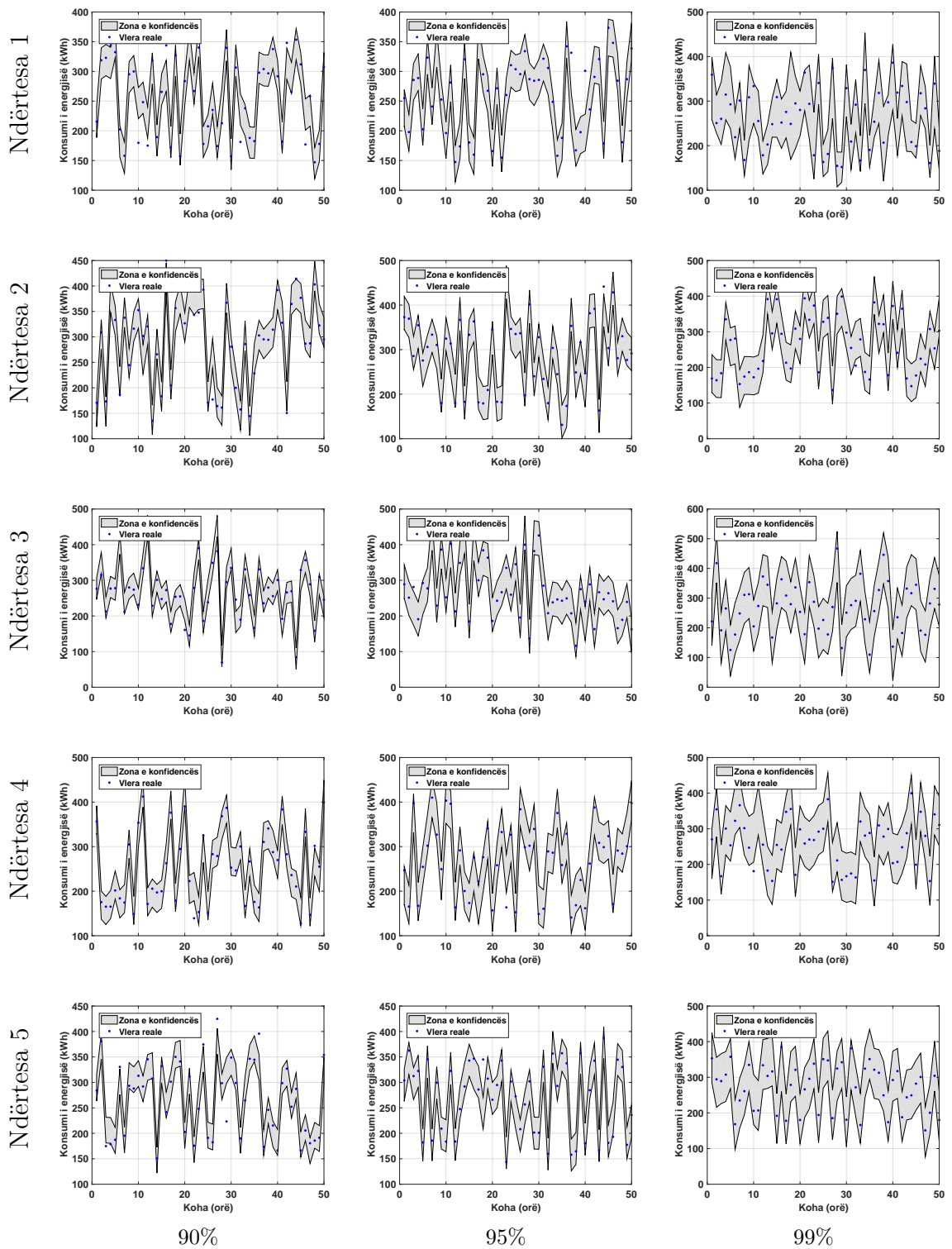


Figura C.4: Intervalet e parashikimit me rrjetat neurale duke përdorur jokonformitetin jo të normalizuar.

Bibliografia

- [1] IEA, "World energy outlook 2015. International Energy Agency; September 2015".
- [2] Zavalani O and Luga Y, "Energy and Water Saving Possibilities in Public Facilities in Albania," in Fourth UKSim European Symposium on Computer Modeling and Simulation (EMS), Pisa, Italy, 2010.
- [3] Zavalani O, "Reducing energy in buildings by using energy management systems and alternative energy-saving systems," in 8th International Conference on the European Energy Market (EEM-11), Zagreb, 2011.
- [4] Aulon Shabani, Orion Zavalani, and Denis Panxhi. "State of art identification and monitoring methods for electric loads in commercial and residential buildings", 5th International Conference of Ecosystem, ICE 2015 - Essay on Ecosystems and Environmental Research, pp.680-686, ICE-2015.
- [5] Aulon Shabani, Denis Panxhi, and Orion Zavalani. "Short term building energy consumption forecasting using artificial neural network model". 6th International Conference of Ecosystems - Proceedings Book, pp.194 - 198, ICE-2016.
- [6] Aulon Shabani, Ilda Kazani, Orion Zavalani, Genti Guxho, S. Kasami. "An innovative electro-conductive textile prototype sweater". 6th International Textile Conference, Tirana, Albania - Book of Proceedings, pp. 123-130, Polytechnic University of Tirana, 2014.
- [7] Aulon Shabani, Ilda Kazani, Orion Zavalani. "A review on improving energy efficiency in buildings using textile materials", 7-th International Textile Conference, Book of Proceedings, pp. 440-449, Polytechnic University of Tirana, 2016.

- [8] Aulon Shabani, and Orion Zavalani. “Predicting Building Energy Consumption using Engineering and Data Driven Approaches: A Review”. *European Journal of Engineering Research and Science* 2, no. 5 (2017): 44-49.
- [9] Aulon Shabani, and Orion Zavalani. “Hourly Prediction of Building Energy Consumption: An Incremental ANN Approach”. *European Journal of Engineering Research and Science* 2, no. 7 (2017): 27-32.
- [10] Orion Zavalani, Aulon Shabani , Aida Spahiu, Lindita Dhamo. “Innovation, Technology Transfer and Automation Systems in Albania”. *HCTL Open International Journal of Technology Innovations and Research (IJTIR)*, Volume 14, April 2015, eISSN: 2321-1814, ISBN (Print): 978-1-62951-946-3.
- [11] Aulon Shabani, and Orion Zavalani. “Surveying the Capabilities of Building Intelligent Systems”. *HCTL Open International Journal of Technology Innovations and Research (IJTIR)*, Volume 25, Issue 1, October 2017, pp. 1-11, e-ISSN: 2321-1814.
- [12] A. C12.20-2010, ”American National Standard for Electricity Meters 0.2 and 0.5 Accuracy Classes,” American National Standards Institute, Inc. , Philadelphia, 2010.
- [13] EEI, *Smart Meters and Smart Meter Systems: A Metering Industry Perspective*, Edison Electric Institute, Washington, 2011.
- [14] Bourns, *Monitoring and Protecting Smart Meter Circuitry and Communications*, Bourns, 2014.
- [15] T. Ueno and Y. Nakano, Effectiveness of Displaying Energy Consumption Data in Residential Buildings: To Know Is to Change, *ACEEE Summer Study on Energy Efficiency in Buildings*, 7-264–7-277,2006.
- [16] Q. Han, S. Mehrotra, Application-aware integration of data collection and power management in wireless sensor networks, *Journal of Parallel and Distributed Computing* 67(9): 992-1006, 2007.
- [17] Directive 2004/22/EC of the European Parliament and of the Council of 31 March 2004 on measuring instruments.
- [18] International standard IEC 62056-21 Electricity metering – Data exchange for meter reading, tariff and load control.

-
- [19] W. Jang and W. Healy, Wireless sensor networks as part of a web-based building environmental monitoring system, *Automation in Construction* 17(6): 729-736, 2008.
- [20] P. Oksa and M. Soini, Kilavi platform for wireless building automation, *Energy and Buildings* 40(9): 1721-1730, 2008.
- [21] W.-T. Sung and Y.-C. Hsu, Designing an industrial real-time measurement and monitoring system based on embedded system and ZigBee, *Expert Systems with Applications* 38(4): 4522-4529, 2011.
- [22] R. Bayindir and E. Irmak, Development of a real time energy monitoring platform, *International Journal of Electrical Power and Energy Systems* 33(1): 137-146, 2011.
- [23] P. Bacher and H. Madsen, Identifying suitable models for the heat dynamics of buildings, *Energy and Buildings*, vol. 43, 1511 – 1522, 2011.
- [24] J. Lee and Y. Ham, Impact Analysis on the Variations of the Thermo-physical Property of Building Envelopes and Occupancy in Building Energy Performance Assessment, *Procedia Engineering*, vol. 145, 556–564, 2016.
- [25] Building Energy Software Tools, US Department of Energy, 2011. [Online]. Available: <http://www.buildingenergysoftwaretools.com/>.
- [26] M. S. Al-Hamoud, Computer-aided building energy analysis techniques, *Building and Environment*, vol. 36, pp. 421–433, 2001.
- [27] D. B. Crawley, J. W. Hand, M. Kummert and B. T. Griffith, Contrasting the capabilities of building energy performance simulation programs, *Building and Environment*, vol. 43, pp. 661–673, 2008.
- [28] D. B. Crawley, L. K. Lawrie, F. C. Winkelmann, W. Buhl, J. Y. Huang, C. O. Pedersen, R. K. Strand, R. L. Liesen, D. E. Fisher, D. J. Witte and J. Glazer, EnergyPlus: creating a new generation building energy simulation program, *Energy and Buildings*, vol. 33, 319–331, 2001.
- [29] N. Fumo, P. Mago and R. Luck, Methodology to estimate building energy consumption using EnergyPlus Benchmark Models, *Energy and Buildings*, vol. 42, no. 12, 2331–2337, 2010.
- [30] ASHRAE, Measurement of Energy and Demand Savings, ASHRAE, Atlanta, Georgia, Guideline 14-2002.

- [31] T. Hong, S. K. Chou and T. Y. Bong, Building simulation : an overview of developments and information sources, *Energy and Environment*, vol. 35, 347–361, 2000.
- [32] J. Pfafferott, S. Herkeland and J. Wapler, Thermal building behaviour in summer: long-term data evaluation using simplified models, *Energy and Buildings* 37(3): 844–852, 2005.
- [33] M. Bauer and J.L. Scartezzini, A simplified correlation method accounting for heating and cooling loads in energy-efficient buildings, *Energy and Buildings* 27(2):147–154, 1998.
- [34] K. E. Westergren, H. Högberg and U. Norlén, Monitoring energy consumption in single-family houses, *Energy and Buildings* 29(3): 247–257, 1999.
- [35] F. A. Ansari , A. S. Mokhtar, K. A. Abbas and N. M. Adam, A Simple Approach for Building Cooling Load Estimation, *American Journal of Environmental Sciences* 1 (3): 209-212, 2005.
- [36] C. Ghiaus, Experimental estimation of building energy performance by robust regression, *Energy and Buildings* 38(6): 582-587, 2006.
- [37] A. Dhar, T. A. Reddy and D. E. Claridge, Modeling hourly energy use in commercial buildings with Fourier series functional form, *ASME Journal of Solar Energy Engineering* 120: 217–23, 1998.
- [38] A. Dhar, T. A. Reddy and D. E. Claridge, A Fourier series model to predict hourly heating and cooling energy use in commercial buildings with outdoor temperature as the only weather variable, *Journal of Solar Energy Engineering* 121: 47–53, 1999.
- [39] F. Lei and P. Hu, A Baseline Model for Office Building Energy Consumption in Hot Summer and Cold Winter Region, In *Proceedings of International Conference on Management and Service Science*, 1 – 4, 2009.
- [40] Y. Ma, J. q. Yu, C. y. Yang and L. Wang, Study on Power Energy Consumption Model for Large-Scale Public Building, In *Proceedings of the 2nd International International Workshop on Intelligent Systems and Applications*, Wuhan, 1-4, 2010.
- [41] M.J. Jiménez and M.R. Heras, Application of multi-output ARX models for estimation of the U and g values of building components in outdoor testing, *Solar Energy* 79(3): 302-310, 2005.

- [42] S. H. Cho, W. T. Kim, C. S. Tae and M. Zaheeruddin, Effect of length of measurement period on accuracy of predicted annual heating energy consumption of buildings, *Energy Conversion and Management* 45(18–19): 2867-2878, 2004.
- [43] A. J. Hoffman, Peak demand control in commercial buildings with target peak adjustment based on load forecasting, *International Conference on Control Applications*, Trieste, Italy, September 1998, 1292-1296.
- [44] J. C. Lam, K. W. Wan, K. L. Cheung and L. Yang, Principal component analysis of electricity use in office buildings, *Energy and Buildings* 40(5): 828-836, 2008.
- [45] J. C. Lam, K. W. Wan, S. L. Wong and T. N. T. Lam, Principal component analysis and long-term building energy simulation correlation, *Energy Conversion and Management* 51(1): 135-139, 2010.
- [46] M. Parti and C. Parti, The total and appliance specific conditional demand for electricity in the household sector, *The Bell Journal of Economics* 11:309-21, 1980.
- [47] G. Lafrance and D. Perron, Evolution of Residential Electricity Demand by End-Use in Quebec 1979-1989: A Conditional Demand Analysis, *Energy Studies Review* 6(2): 164-173, 1994.
- [48] D. J. Aigner, C. Sorooshian and P. Kerwin, Conditional Demand Analysis for Estimating Residential End-Use Load Profiles, *The Energy Journal* 5(3):81-98 , 1984.
- [49] M. Aydinalp-Koksal and V. Ismet Ugursal, Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector, *Applied Energy* 85(4): 271-296, 2008.
- [50] T. Catalina, V. Iordache and B. Caracaleanu, Multiple regression model for fast prediction of the heating energy demand, *Energy and Buildings* 57: 302-312, 2013.
- [51] G. R. Newsham and B. J. Birt, Building-level occupancy data to improve ARIMA-based electricity use forecasts, In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*: 13-18, New York, 2010.

- [52] A. Kimbara, S. Kurosu, R. Endo, K. Kamimura, T. Matsuba, and A. Yamada. On- line prediction for load profile of an air-conditioning system. *ASHRAE Transactions*, 101(2):198 – 207, 1995.
- [53] T. Matsuba, H. Tsutsui and K. Kamimura, *Building Simulation 6*: 909-916, Kyoto, Japan, 1999.
- [54] A. Singh, Ibraheem, S. Khatoon and M. Muazzam, An Overview of Electricity Demand Forecasting Techniques, *National Conference on Emerging Trends in Electrical, Instrumentation and Communication Engineering 3(3)*: 38-48, 2013.
- [55] S. Karatasou, M. Santamouris and V. Geros, Prediction of energy consumption in buildings with artificial intelligent techniques and Chaos time series analysis, *International Workshop on Energy Performance and Environmental Quality of Buildings*, 1-6, Greece, 2006.
- [56] P. Dagnely, T. Ruetter, T. Tourwé, E. Tsiorkova and C. Verhelst, Predicting Hourly Energy Consumption. Can Regression Modeling Improve on an Autoregressive Baseline?, *Data Analytics for Renewable Energy Integration: Third ECML PKDD Workshop*, Porto, Portugal, 2015.
- [57] A. Tsanas and A. Xifara, Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools, *Energy and Buildings* 49: 560-567, 2012.
- [58] W. McCulloch and P. Walter Pitts A Logical Calculus of Ideas Immanent in Nervous Activity, *Bulletin of Mathematical Biophysics* 5 (4): 115–133, 1943.
- [59] D. Hebb, *The Organization of Behavior*, New York: Wiley, 1949.
- [60] B.G. Farley and W.A. Clark , Simulation of Self-Organizing Systems by Digital Computer, *IRE Transactions on Information Theory*. 4 (4): 76–84, 1954.
- [61] N. Rochester, J.H. Holland, L.H. Habit and W.L. Duda, Tests on a cell assembly theory of the action of the brain, using a large digital computer”. *IRE Transactions on Information Theory*. 2 (3): 80–93, 1956.
- [62] F. Rosenblatt, The Perceptron: A Probabilistic Model For Information Storage And Organization In The Brain, *Psychological Review* 65 (6): 386–408, 1958.

-
- [63] P.J. Werbos, *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, 1975.
- [64] M. Minsky and S. Papert, *Perceptrons: An Introduction to Computational Geometry*, MIT Press, 1969.
- [65] D.E. Rumelhart and J. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge: MIT Press, 1986.
- [66] A. Azadeh, S.F. Ghaderi and S. Sohrabkhani, Forecasting electrical consumption by integration of Neural Network, time series and ANOVA, *Applied Mathematics and Computation* 186(2): 1753-1761, 2007.
- [67] A. Azadeh, S.F. Ghaderi, S. Tarverdian and M. Saberi, Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption, *Applied Mathematics and Computation* 186(2): 1731-1741, 2007.
- [68] A. Azadeh and S. Tarverdian, Integration of genetic algorithm, computer simulation and design of experiments for forecasting electrical energy consumption, *Energy Policy* 35(10): 5229-5241, 2007.
- [69] A. Azadeh, S.F. Ghaderi and S. Sohrabkhani, A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran, *Energy Policy* 36(7): 2637-2644, 2008.
- [70] A. Azadeh, S. Ghaderi and S. Sohrabkhani, Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors, *Energy Conversion and Management*, vol. 49, no(8), 2272 – 2278, 2008.
- [71] S. Karatasou, M. Santamouris and V. Geros, Modeling and predicting building's energy use with artificial neural networks: Methods and results, *Energy and Buildings*, vol. 38, no. (8), 949 – 958, 2006.
- [72] P. González and J. Zamarreño, Prediction of hourly energy consumption in buildings based on a feedback artificial neural network, *Energy and Buildings*, vol. 37, no. (6), pp. 595 – 601, 2005.
- [73] S. Kalogirou, Applications of artificial neural-networks for energy systems, *Applied Energy* 67(1-2): 17-35, 2000.

- [74] S. Kalogirou and M. Bojic, Artificial neural networks for the prediction of the energy consumption of a passive solar building, *Energy* 25(5): 479 -491, 2000.
- [75] B. Ekici and U. Aksoy, Prediction of building energy consumption by using artificial neural networks, *Advances in Engineering Software*, vol. 50, no. 5, 356 – 362, 2009.
- [76] R. Yokoyama, T. Wakui and T. Satake, Prediction of energy demands using neural network with model identification by global optimization, *Energy Conversion and Management*, vol. 50, no. 2, p. 319—327, 2009.
- [77] J. Kreider, D. Claridge, . P. Curtiss, . R. Dodier, J. Haberl and M. Krarti, Building energy use prediction and system identification using recurrent neural networks, *Journal of Solar Energy Engineering*, vol. 117, no. 3, pp. 161–166, 1995.
- [78] C.-w. Yan and J. Yao , Application of ANN for the prediction of building energy consumption at different climate zones with HDD and CDD, in *Proceedings of the 2nd international conference on future computer and communication*, 2010.
- [79] Z. Hou , Z. Lian, Y. Yao and X. Yuan , Cooling-load prediction by the combination of rough set theory and an artificial neural-network based on data-fusion technique, *Applied Energy*, vol. 83, no. 9, pp. 1033–1046, 2006.
- [80] Gerretzen, J., Szymańska, E., Jansen, J. J., Bart, J., van Manen, H.-J., van den Heuvel, E., Buydens, L. M. , Simple and Effective Way for Data Preprocessing Selection Based on Design of Experiments, *Analytical Chemistry*, 87(24), 12096–12103, 2015.
- [81] J. W. Moon, K. Kim and H. Min, ANN-Based Prediction and Optimization of Cooling System in Hotel Rooms, *Energies*, vol. 32.pp. 10775–10795, 2015.
- [82] M. Aydinalp, V. Ugursal and A. Fung , Modeling of the appliance, lighting, and spacecooling energy consumptions in the residential sector using neural networks, *Applied Energy*, vol. 71, no. 2, pp. 87–110, 2002.
- [83] T. Olofsson and S. Andersson, Analysis of the interaction between heating and domestic load in occupied single-family buildings, 5th symposium on building physics in the Nordic countries, 473–480, 1999.

-
- [84] T. Olofsson and S. Andersson, Overall heat loss coefficient and domestic energy gain factor for single-family buildings, *Building and Environment* 37(11):1019–1026, 2002.
- [85] M. Lundin, S. Andersson and R. Östin, Validation of a neural network method for estimation heat loss and domestic gain in buildings. In: *Proceedings of the 6th symposium on building physics in the Nordic countries*, 325–332, 2002.
- [86] M. Lundin, S. Andersson and R. Östin, Development and validation of a method aimed at estimating building performance parameters. *Energy and Buildings* 36(9):905–914, 2004.
- [87] J. Kreider, D. Claridge, P. Curtiss, R. Dodier, J. Habe and M. Krarti, Building energy use prediction and system identification using recurrent neural networks, *Journal of Solar Energy Engineering*, pp. 161–166, 1995.
- [88] A. Neto and F. Fiorelli , Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, *Energy and Buildings*, vol. 40, no. 12, pp. 2169–2176, 2008.
- [89] M. Laurent and H. Fariborz , Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network, *Building and Environment*,, vol. 45, no. 3, pp. 739–746, 2010.
- [90] H.-x. Zhao and F. Magoulès, A review on the prediction of building energy consumption, *Renewable and Sustainable Energy Reviews*, vol. 16, pp. 3586–3592, 2012.
- [91] A. Dounis and C. Caraiscos, Advanced control systems engineering for energy and comfort management in a building environment—a review, *Renewable and Sustainable Energy Reviews* , vol. 13, no. 6–7, pp. 1246 – 1261, 2009.
- [92] M. Bill , Case-Based Reasoning for Autoclave Management, *Proceedings of the Case-Based Reasoning Workshop*, 1989.
- [93] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, Ning and M. Fol, Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments, *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 41, no. 4, p. 421–434, 2011.

- [94] N. Trung , C. Mary and L. Dan , COMPAQ QuickSource: Providing the Consumer with the Power of Artificial Intelligence, in Fifth Annual Conference on Innovative Applications of Artificial Intelligence, Washington, 1993.
- [95] V. Vapnik, S. E. Golowich and A. Smola, Support vector method for function approximation, regression estimation, and signal processing, Annual Conference on Neural Information Processing Systems (NIPS): 281-287, 1996.
- [96] M. Brown, C. B-Leigh and Z Brown, Kernel Regression for Real-Time Building Energy Analysis, Journal of Building Performance Simulation, 5 (4): 263-276,1993.
- [97] D. Solomon, R. Winter, A. Boulanger, R. Anderson and L. Wu, Forecasting Energy Demand in Large Commercial Buildings Using Support Vector Machine Regression, Columbia University Computer Science Technical Reports CUCS-040-11, 2000.
- [98] B. Dong, Ch. Cao and S. E. Lee, Applying support vector machines to predict building energy consumption in tropical region, Energy and Buildings 37(5): 545-553, 2005.
- [99] C. Hsu, C. Wu, S. Chen and K. Peng, Dynamically Optimizing Parameters in Support Vector Regression : An Application of Electricity Load Forecasting, Proceedings of the 39th Hawaii International Conference on System Sciences, 1-8, 2006.
- [100] F. Lai, F. Magoulès and F. Lherminier, Vapnik's learning theory applied to energy consumption forecasts in residential buildings, International Journal of Computer Mathematics 85(0): 1563-1588, 2008.
- [101] Q. Li, Q. Meng, J. Cai, H. Yoshino and A. Mochida, Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks, Energy Conversion and Management 50(1): 90-96, 2009.
- [102] Q. Li, Q. Meng, J. Cai, H. Yoshino and A. Mochida, Applying support vector machine to predict hourly cooling load in the building, Applied Energy 86(10): 2249-2256, 2009.
- [103] Y.-M. Zhang and W.-G. Qi, Interval Forecasting for Heating Load Using Support Vector Regression and Error Correcting Markov Chains, Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, 1106-1110, 2009.

- [104] H.-x. Zhao and F. Magoulès, Feature Selection for Predicting Building Energy Consumption Based on Statistical Learning Method, *Journal of Algorithms and Computational Technology* 6(1): 59-78.
- [105] H.-x. Zhao and F. Magoulès, Parallel Support Vector Machines Applied to the Prediction of Multiple Buildings Energy Consumption, *Journal of Algorithms and Computational Technology* 4(2): 231-249, 2010.
- [106] H.-x. Zhao and F. Magoulès, New parallel support vector regression for predicting building energy consumption, *IEEE Symposium Series on Computational Intelligence (SSCI 2011)*, Apr 2011, Paris, France.
- [107] Q. Li, P. Ren and Q. Meng, Prediction Model of Annual Energy Consumption of Residential Buildings, *International Conference on Advances in Energy Engineering Prediction*, 223-226, 2010.
- [108] L. Jinhu, L. Xuemei, D. Lixing and J. Liangzhong, Applying Principal Component Analysis and Weighted Support Vector Machine in Building Cooling Load Forecasting, *International Conference on Computer and Communication Technologies in Agriculture Engineering*, 109-112, 2010.
- [109] R. Edwards, J. New and L. E. Parker, Predicting future hourly residential electrical consumption : A machine learning case study, *Energy and Buildings* 49: 591-603, 2012.
- [110] Z. Hu, Y. Bao and T. Xiong, Electricity Load Forecasting Using Support Vector Regression with Memetic Algorithms, *The Scientific World Journal*, 1-10, 2013.
- [111] L. Xuemei, D. Lixing, L. Jinhu, X. Gang and L. Jibin, A Novel Hybrid Approach of KPCA and SVM for Building Cooling Load Prediction, *Third International Conference on Knowledge Discovery and Data Mining*, 522-526, 2010.
- [112] L. Hussain, M. S. Nadeem, S. A. A. Shah, Short Term Load Forecasting System Based on Support Vector Kernel Methods, *International Journal of Computer Science and Information Technology (IJCSIT)* 6(3): 93-102, 2014.
- [113] R. K. Jain, K. M. Smith, P. J. Culligan and J. E. Taylor, Forecasting energy consumption of multi-family residential buildings using support vector regression : Investigating the impact of temporal and spatial monitoring granularity on performance accuracy, *Applied Energy* 123: 168-178, 2014.

- [114] S. Paudel, P. H. Nguyen, W. L. Kling and M.Elmitri, Support Vector Machine in Prediction of Building Energy Demand Using Pseudo Dynamic Approach, The 28th International Conference On Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems 2015, France.
- [115] J. F. M. Pessanhaa and N. Leonb, Forecasting long-term electricity demand in the residential sector, *Information Technology and Quantitative Management* 55: 529-538, 2015.
- [116] H. Son and C. Kim, Forecasting Short-term Electricity Demand in Residential Sector Based on Support Vector Regression and Fuzzy-rough Feature Selection with Particle Swarm Optimization, *Procedia Engineering* 118: 1162-1168, 2015.
- [117] K. Grolinger, M. Capretz, and L. Seewald, Energy Consumption Prediction with Big Data: Balancing Prediction Accuracy and Computational Resources, *Electrical and Computer Engineering Publications*, Paper 90, 2016.
- [118] J.R. Quinlan, Induction of decision trees, *Machine Learning* (1986).
- [119] J. Han, M. Kamber, *Data mining concepts and techniques*, Elsevier Inc., San Francisco, 2006.
- [120] J. Yang and J. Stenzel, Short-term load forecasting with increment regression tree, *Electric Power Systems Research* 76: 880–888, 2006. Short-term
- [121] G. Tso and K. Yau, Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks, *Energy* 32(9): 1761-1768, 2007.
- [122] B. Gladyszd and D. Kuchta, Application of regression trees in the analysis, *Operations Research and Decisions* 4: 19-28, 2008.
- [123] Zh. Yu, F. Haghghat, B. C.M. Fung and H. Yoshino, A decision tree method for building energy demand modeling, *Energy and Buildings* 42(10): 1637-1646, 2010.
- [124] S. Aman, Y. Simmahan and V. K. Prasanna, Improving energy use forecast for campus micro-grids using indirect indicators, *Proceedings - IEEE International Conference on Data Mining*, 389-397, 2011.

-
- [125] S. Aman, M. Frincu, Ch. Chelmiss, M. U. Noor, Y. Simmhan and V. K. Prasanna, Empirical Comparison of Prediction Methods for Electricity Consumption Forecasting, 1-12, 2014.
- [126] S. S. Abdelkader, K. Grolinger and M. A. M. Capretz, Predicting energy demand peak using M5 model trees, Electrical and Computer Engineering Publications, Paper 72, 2015.
- [127] M. Behl, A. Jain and R. Mangharam, Data-Driven Modeling, Control and Tools for Cyber-Physical Energy Systems, ACM/IEEE 7th International Conference on Cyber-Physical Systems, April 2016.
- [128] Z. Wang and R. Srinivasan, Homogeneous Ensemble Model for Building Energy Prediction : A Case Study Using Ensemble Regression Tree, ACEEE Summer Study on Energy Efficiency in Buildings, 1-12, 2016.
- [129] D. Monfet, M. Corsi, D. Choinière and E. Arkhipova, Development of an energy prediction tool for commercial buildings using case-based reasoning, Energy and Buildings, vol. 81, pp. 152–160, 2014.
- [130] E. Arkhipova, D. Choiniere and D. Monfet, Evaluation of a case-based reasoning energy prediction tool for commercial buildings, Energy Systems Laboratory, 2013.
- [131] R. Platon, J. Martel and V. D. Dehkordi, Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis, Energy and Buildings, vol. 92, pp. 10–18, 2015.
- [132] R. Platon, J. Martel and M. Zoghلامي, CBR model for predicting a building's electricity use: On-line implementation in the absence of historical data, in ICCBR, Frankfurt, 2015.
- [133] A. Shabani, A. Paul, R. Platon, and E. Hüllermeier, Predicting the Electricity Consumption of Buildings: An Improved CBR Approach, pages 356–369. Springer International Publishing, Cham, 2016.
- [134] R. Schank, Dynamic Memory: A Theory of Learning in Computers and People (New York: Cambridge University Press, 1982).
- [135] S. K. Pal and S. C. K. Shiu, Foundations of soft Case-Based Reasoning, John Willey.

- [136] A. Aamodt and E. Plaza, Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, *AI Communications*, vol. 7, no. 1, pp. 39–59, 1994.
- [137] E. Keogh and C. A. Ratanamahatana, Exact indexing of dynamic time warping, *Knowledge and Information Systems* , vol. 7, no. 3, p. 358–386, 2005.
- [138] J. Gerretzen, E. Szymańska, J. J. Jansen, J. Bart, H.-J. van Manen, E. R. van den Heuvel and L. M. Buydens, Simple and Effective Way for Data Preprocessing Selection Based on Design of Experiments, *Analytical Chemistry*, vol. 87, no. 24, pp. 12096–12103, 2015.
- [139] I. Krcmar, P. Matic and S. Krunić , Influence of data preprocessing on prediction of complex valued load time series, *INFOTEH-JAHORINA*, vol. 13, pp. 1089–1093, 2014.
- [140] N. Mohd Nawi, W. H. Atomi and M. Z. Rehman, The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks, *Procedia Technology*, vol. 11, pp. 32–39, 2013.
- [141] T. Gabel and E. Godehardt, Top-down induction of similarity measures using similarity clouds, in *ICCBR 2015, 23rd International Conference on Case-Based Reasoning*, Frankfurt am Main, Germany, 2015.
- [142] A. Stahl, Learning similarity measures: A formal view based on a generalized CBR model, in *ICCBR-05, International Conference on Machine Learning*, Chicago, USA, 2005.
- [143] A. Stahl and T. Gabel, Optimizing similarity assessment in case-based reasoning, in *AAAI-06, 21st National Conference on Arti*, 2006.
- [144] K. W. Chu and K. a. M. Hon Wong, Fast time-series searching with scaling and shifting, in *18th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, New York , 1999.
- [145] K. Levenberg, A method for the solution of certain problems in least squares, *Quarterly of Applied Mathematics* , vol. 5, p. 164–168, 1944.
- [146] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, *SIAM Journal on Applied Mathematics*, vol. 11, no. 2, p. 431–441, June 1963.

-
- [147] M. T. Hagan and M. Menhaj, Training feedforward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks*, vol. 5, no. 6, p. 989–993, 1994.
- [148] M. S. Gul and S. Patidar, Understanding the energy consumption and occupancy of a multi-purpose academic building, *Energy and Buildings*, vol. 87, pp. 155-165, 2015.
- [149] D. Whitely, T. Starkweather and C. Bogart, Genetic algorithms and neural networks: optimizing connections and connectivity , *Parallel Computing*, vol. 14, pp. 347–361, 1990.
- [150] Y. Dong, Z. Sun and H. Jia, A cosine similarity-based negative selection algorithm for time series novelty detection, *Mechanical Systems and Signal Processing*, vol. 20, no. 6, p. 1461–1472, 2006.
- [151] E. Hüllermeier, *Case-Based Approximate Reasoning* , vol. 44, Springer Netherlands: Springer Netherlands, 2007.
- [152] H. Yu and B. M. Wilamowski, Levenberg-Marquardt training, *Industrial Electronics Handbook, Intelligent Systems*, pp. 12.1–12.18, 2011.
- [153] Tom Mitchell, *Machine Learning*, McGraw-Hill, New York, 1997.
- [154] Blackwell, David and M. A. Girshick , *Theory of Games and Statistical Decisions*, New York: Wiley,1954.
- [155] Berk Richard A., *Statistical Learning from a Regression Perspective*, Springer Series in Statistics, New York: Springer-Verlag, 2008.
- [156] B. E. Boser, I. M. Guyon, and V. N. Vapnik, A Training Algorithm for Optimal Margin Classifiers, 5th Annual Workshop on Computational Learning Theory, Pittsburgh, Pennsylvania, USA, 1992
- [157] N. E. Ayat, M. Cheriet, and C. Y. Suen, Automatic model selection for the optimization of SVM kernels, *Pattern Recognition* 38: 1733-1745, 2005.
- [158] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, Gene Selection for Cancer Classification using Support Vector Machines, *Machine Learning* 46: 389-422, 2002.
- [159] J. Zhang and Y. Liu, Červical Cancer Detection Using SVM Based Feature Screening, 7th Medical Image Computing and Computer-Assisted Intervention, Saint-Malo, France, 2004.

- [160] T. Joachims, Text Categorization with Support Vector Machines: Learning with Many Relevant Features, 10th European Conference on Machine Learning, Chemnitz, Germany, 1998.
- [161] G. Guo, S. Z. Li, and K. L. Chan, Support vector machines for face recognition, *Image and Vision Computing* 19: 631-638, 2001.
- [162] B. Baesens, T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens, and J. Vanthienen, Benchmarking state-of-the-art classification algorithms for credit scoring, *Journal of the Operational Research Society* 54: 627-635, 2003.
- [163] S. Viaene, B. Baesens, T. Van Gestel, J. A. K. Suykens, D. Van den Poel, J. Vanthienen, B. De Moor, and G. Dedene, Knowledge discovery in a direct marketing case using least squares support vector machines, *International Journal of Intelligent Systems* 16: 1023-1036, 2001.
- [164] V. N. Vapnik. The nature of statistical learning theory. Springer-Verlag New York, Inc., New York, NY, USA, 1995.
- [165] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik. Support vector regression machines. *NIPS*, 155—161, 1996.
- [166] B. Schölkopf, J. C. Platt, J. C. Shawe-Taylor, A. J. Smola, and R. C. Williamson, Estimating the support of a high-dimensional distribution, *Neural Computation*, 13:1443–1471, July 2001.
- [167] O. Chapelle and V. Vapnik, Model Selection for Support Vector Machines. In *Advances in Neural Information Processing Systems*, Vol 12, 1999.
- [168] Simon S. Haykin, *Neural networks and learning machines*, Pearson, Upper Saddle River, N.J, 2009.
- [169] Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature*. 323 (6088): 533–536
- [170] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated, 2014.
- [171] A. Shabani, and O. Zavalani. Predicting Building Energy Consumption using Engineering and Data Driven Approaches: A Review. *European Journal of Engineering Research and Science* 2(5): 44-49, 2017.

-
- [172] Elena Ikonomovska, João Gama, and Sašo Džeroski. Learning model trees from evolving data streams. *Data mining and knowledge discovery*, 23(1):128–168, 2011.
- [173] H. Mouss, D. Mouss, N. Mouss, and L. Sefouhi. Test of page-hinckley, an approach for fault detection in an agro-alimentary production system. In *2004 5th Asian Control Conference (IEEE Cat. No.04EX904)*, volume 2, pages 815–818 Vol.2, July 2004.
- [174] Ammar Shaker and Eyke Hüllermeier. Iblstreams: a system for instance-based classification and regression on data streams. *Evolving Systems*, 3(4):235–249, 2012.
- [175] Ezilda Almeida, Carlos Ferreira, and Joao Gama. Adaptive model rules from data streams. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 480–492. Springer, 2013.
- [176] A. Gammerman, V. Vovk, and V. Vapnik, *Learning by Transduction*, *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 1998, pp:148–155,
- [177] Cortes .C, Vapnik .V, *Support-vector networks*, *Machine Learning*, 1995;20(3):273–297.
- [178] Saunders C, Gammerman A, Vovk V. *Transduction with confidence and credibility*. In: *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI'99)*. vol. 2; 1999. p. 722–726.
- [179] Nouretdinov I, Melluish T, Vovk V. *Ridge Regression Confidence Machine*. In: *Proceedings of the Eighteenth International Conference on Machine Learning. ICML '01*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2001. p. 385–392.
- [180] Papadopoulos H, Proedrou K, Vovk V, Gammerman A. In: Elomaa T, Mannila H, Toivonen H, editors. *Inductive Confidence Machines for Regression*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2002. p. 345–356.
- [181] Papadopoulos H. *Inductive conformal prediction: Theory and application to neural networks*. INTECH Open Access Publisher; 2008.
- [182] Papadopoulos H, Haralambous H. *Reliable prediction intervals with regression neural networks*. *Neural Networks*. 2011;24(8):842 – 851. *Artificial Neural Networks: Selected Papers from ICANN 2010*.

- [183] Papadopoulos H, Haralambous H. Neural networks regression inductive conformal predictor and its application to total electron content prediction. In: International Conference on Artificial Neural Networks. Springer; 2010. p. 32–41.
- [184] Johansson U, Boström H, Löfström T, Linusson H. Regression conformal prediction with random forests. *Machine Learning*. 2014;97(1):155–176.
- [185] Papadopoulos H, Vovk V, Gammerman A. Regression conformal prediction with nearest neighbours. *Journal of Artificial Intelligence Research*. 2011;40:815–840.
- [186] Ji GR, Dong Z, Wang DF, Han P, Xu DP. Wind speed conformal prediction in wind farm based on algorithmic randomness theory. In: 2008 International Conference on Machine Learning and Cybernetics. vol. 1; 2008. p. 131–135.
- [187] Balasubramanian V, Chakraborty S, Panchanathan S, Ye J. Kernel Learning for Efficiency Maximization in the Conformal Predictions Framework. In: 2010 Ninth International Conference on Machine Learning and Applications; 2010. p. 235–242.
- [188] Norinder U, Carlsson L, Boyer S, Eklund M. Introducing Conformal Prediction in Predictive Modeling. A Transparent and Flexible Alternative to Applicability Domain Determination. *Journal of Chemical Information and Modeling*. 2014;54(6):1596–1603.
- [189] Shafer G, Vovk V. A tutorial on conformal prediction. *Journal of Machine Learning Research*. 2008;9(Mar):371–421.
- [190] Balasubramanian V, Ho SS, Vovk V. *Conformal Prediction for Reliable Machine Learning: Theory, Adaptations and Applications*. Newnes; 2014.
- [191] Papadopoulos H, Gammerman A, Vovk V. Normalized nonconformity measures for regression conformal prediction. In: Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA 2008); 2008. p. 64–69.
- [192] MATLAB. version 9.10.0 (R2016b). Natick, Massachusetts: The MathWorks Inc.; 2016.