TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Hatef Palizgar 184600IVEM11/18

Energy Consumption in Smart Homes: Analysing User Behaviour using Combined Vector Machine and Neural Network Techniques

Master's thesis

Supervisor: Yannick Le Moullec
Professor

Tallinn 2020

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Hatef Palizgar 184600IVEM11/18

Energiatarbimine arukates kodudes: kasutaja käitumise analüüsimine kombineeritud vektormasina ja närvivõrgu tehnikate abil

magistritöö

Juhendaja: Yannick Le Moullec

Professor

Tallinn 2020

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Hatef Palizgar

03.01.2021

Abstract

A smart home, or smart house, is a collection of sensors and services on a home network to improve life quality. These homes have an evident ability to make life easier and increase the well-being of the people and save energy. Improving energy consumption and resource management provides excellent and efficient use of energy. Although so far different ways have been used to reduce energy consumption and improve the use of resources in smart homes, the challenge of relatively high energy consumption remains open. Recent progress in machine learning has paved the way for addressing such a challenge, among others. In this thesis, we use a combination of support vector machines and neural networks for analysing user behaviour with the longer-term perspective to optimize energy consumption.

The proposed method starts with a preliminary clustering step and is followed by the proposed combined neural network and support vector machine approach. This proposed method is simulated using a real-life dataset. The four selected main performance indicators show better performance for the proposed approach as compared to the individual support vector machine and neural network algorithms alone, i.e., prediction accuracy, precision, and recall are higher by up to 8.44 percentage points and the error rate is lower by down to 3.51 percentage points. The results pave the way for optimizing energy consumption in smart homes based on the residents' behaviour.

This thesis is written in English and is 92 pages long, including 6 chapters, 31 figures and 10 tables.

Annotatsioon
Energiatarbimine arukates kodudes: kasutaja käitumise analüüsimine kombineeritud vektormasina ja närvivõrgu tehnikate abil

Nutimaja on koduvõrgus olevate andurite ja teenuste kogum elukvaliteedi parandamiseks. Nendel kodudel on selge võime elu hõlbustada ja inimeste heaolu suurendada ning energiat säästa. Energiatarbimise ja ressursside haldamise parandamine tagab energia hea ja tõhusa kasutamise. Hoolimata asjaolust, et seni on nutikodudes energiatarbimise vähendamiseks ja ressursside paremaks kasutamiseks kasutatud erinevaid viise, on suhteliselt kõrge energiatarbimise väljakutse endiselt lahtine. Masinaõppe hiljutised edusammud on sillutanud teed teiste probleemide lahendamiseks. Selles töös kasutatakse tugivektorimasina ja närvivõrkude kombinatsiooni, et energiatarbimise optimeerimiseks analüüsida kasutaja käitumist pikemas perspektiivis.

Kavandatud meetod algab esialgse klastrite sammuga ja sellele järgneb kombineeritud närvivõrgu ja tugivektori masina lähenemine. Seda pakutavat meetodit simuleeritakse reaalse andmekogumi abil. Neli peamist toimivusnäitajat on kavandatava lähenemisviisi jaoks paremad kui ainult üksikute tugivektorimasinate ja närvivõrgu algoritmide puhul, st prognoosimise täpsus, täpsus ja tagasikutsumine on kõrgemad kuni 8.44 protsendipunkti võrra ja veamäär on madalam alla 3.51 protsendipunkti. Tulemused sillutavad teed elanike käitumisest lähtuvalt nutikodude energiatarbimise optimeerimiseks.

Lõputöö on kirjutatud Inglise keeles ning sisaldab teksti 92 leheküljel, 6 peatükki, 31 joonist, 10 tabelit.

List of abbreviations and terms

ANN	Artificial Neural Network
BACnet	Building Automation and Control Networks
BAS	Building Automation System
BMS	Building Management System
BP	back-propagation
CCTV	Closed-circuit Television
CDF	Cumulative Distribution Function
DBSCAN	Density-based Spatial Clustering of Applications with Noise
EIB	European Installation Bus
EMS	Energy Management System
ERP	Enterprise Resource Planning
FNN	Feed-forward Neural Networks
GUI	Graphical User Interface
HMI	Human-Machine Interface
HVAC	Heating, ventilation, and air conditioning
IoT	Internet of Things
IrDA	The Infrared Data Association
IVR	Interactive Voice Response
LON	Local Operating Network
MLP	Multi-Layer Perceptron
NN	Neural Network
PEV	Plug-in Electric Vehicle
PLC	Programmable Logic Controller
PM	Preventive Maintenance
PSO	Particle Swarm Optimization
PV	Photovoltaic
RBF	Radial Basis Function
RF	Radio Frequency
RNN	Recurrent Neural Networks
SLP	Single-Layer Perceptron
SVM	Support Vector Machine

Table of Contents

List of Figures

List of Tables

1 Chapter 1

1.1 Research Overview

This chapter first examines the main subject of the research. Afterwards, it discusses the innovation and the objectives of the study.

Recent advancements in technology have created new ubiquitous computing environments that generate large amounts of data to take improved design decisions and provide enhanced services. They may contain general information, can benefit businesses to compete more, offer better services, or even reduce costs. Ubiquitous computing benefits not only the business but also research. One of the most important fields of ubiquitous computing is the smart home. A smart home is equipped with sensors and actuators to collect information and manage residents' activities, respectively.

A smart home is a set of technologies and services on a home network to improve life quality. Moreover, these homes have an evident ability to make life easier, increase well-being, and save energy. Energy management means cost-effective use. The type of consumption is different from the kind of high-consumption electrical appliances in a building. Smart homes' advantage in reducing energy consumption is installing sensors on computer networks and controlling the system without the need for a high-cost wiring system. Thanks to this, by installing e.g., metering sensors, the amount of energy consumption can be measured. Due to the rapid development of various digital sensors and digital monitoring devices, identifying activities based on the sensors has become widespread in recent years. To better understand behavioural patterns, key features are extracted using sensors, and patterns are identified and described accordingly.

Depending on, e.g., the amount of energy consumption, light, and outside temperature, such systems could select the energy consumption management method with low cost and energy savings. These systems are recognized as intelligent energy management systems in the building, e.g., they prevent over-consumption of energy by setting an hourly or a periodical functionality. In addition to saving energy consumption from an economic perspective, other factors such as high population growth and higher energy demand, limited energy resources, incorrect energy consumption behaviours, lack of systems for recycling energy, as well as increasing amount of greenhouse gases and acid rain all highlight the importance of smart homes and similar systems.

On the other hand, the current energy demand and environmental crisis have fastened electric vehicles ownership and renewable energy applications, including solar and wind power [1]. Charging of electric vehicles generating renewable energy is always intermittent. If it is not controlled, there may be a significant impact on the power grid and overload of the system, which also contributes to the energy management challenges.

Overall, the motivation behind this thesis is to investigate methods to identify residents' behaviour in smart homes with a view to save electricity therein. In this context, researchers' have developed methods that build upon, among others, fuzzy logic [2], Markov algorithms [3], and artificial intelligence algorithms. While previous works helped to save energy to some extent, this dissertation proposes an approach that combines a support vector machine algorithm and a neural network algorithm with the aim to further reduce energy consumption. Support vector machine (SVM) is one of the relatively recent methods that has shown good performance in recent years compared to older perceptron neural networks [4] [5]. Briefly, the SVM algorithm operates based on the linear classification of data. The data's linear classification is selected to create more confidence intervals for these samples and more boundaries between classes. Furthermore, artificial neural network (ANN), also known as the neural network (NN), is a mathematical model based on the biological system. The NN employs neurons to organize its structural components to perform its calculations much faster. This type of network has a natural tendency to store empirical knowledge and make it accessible for use.

Due to their individual benefits, according to research, it seems that by combining the algorithm, a new algorithm could be introduced to further reduce power consumption in smart homes.

1.2 Importance and necessity of research

The smart home management system's main tasks are managing and controlling different building states in cooling, heating, and lighting. It is also responsible for establishing a logical connection between these subsystems. The energy crisis on the planet and global warming is a significant challenge that has affected human life in this century.

However, sometimes the financial cost of installing an intelligent energy management system is much higher than the amount of financial savings from efficient consumption. Therefore, it is better to make an overall estimate between the installation costs used and other factors before installing the system.

The objectives of implementing smart homes can be summarized as follows:

- Saving energy consumption in the building;
- Manage, control, and display the status of the system;
- Reduce maintenance costs;
- Increase the useful life of devices;
- Increase the efficiency and performance of equipment;
- Connecting to other control systems and providing a comfortable environment for residents;
- In addition to the associated costs with economic burden, several factors increase the need and necessity of reducing energy consumption:
    - Increase population growth and increase energy demand;
    - Limited renewable energy sources;
    - Increased growth in energy consumption due to the wrong pattern;

The mentioned objectives are the primary motivations to optimize energy consumption. Intelligent systems in the building lead to optimal energy consumption in various forms, resulting in a drastic reduction in cost besides other advantages. According to the mentioned reasons, algorithms are needed to save energy; therefore, the algorithm to reduce energy consumption is required. For this purpose, as mentioned earlier, this thesis investigates a combination of SVM and NN in a collective intelligence system.

1.3 Research objectives

The most important objectives of this research are:

- Presenting the main concepts related to smart homes, especially from an energy consumption perspective;

- Gaining an understanding of the general background of supervised and non-supervised learning methods;
- Proposing an approach that combines SVM and NN (including a preliminary clustering step);
- Evaluating the proposed SVM-NN approach on a real-life dataset and comparing the results with the individual SVM and NN algorithms alone.

1.4 Dissertation structure

The rest of this thesis is organized as follows.

In the second chapter, some essential research in various fields and the research subject are examined. In that chapter, a general introduction to the theory and main concepts of research related to the algorithm of variables and unknowns of study and other necessary items are provided.

In chapter 3, the full description of the proposed approach is given, including a flowchart description of the proposed method. In general, in that chapter, the proposed design's architectural flowchart is examined, and the steps of implementation and simulation of the proposed method are discussed.

In chapter 4, the database used for the experimental part of this thesis is discussed, and the simulation steps of the proposed method and related software are comprehensively reviewed. Finally, comparison and evaluation with other methods are described.

In the fifth chapter, the research findings, conclusions, and suggestions according to the research hypotheses and objectives are presented.

2 Chapter 2 – Smart Home Energy Efficiency

2.1 Smart home and building energy management system

A smart home is a building in which all its control subsystems are predefined in a practical framework. These subsystems have been designed and implemented following the uses of the building. The control process in the smart home is done by different control systems that are also intelligent and communicate and interact directly. Ultimately, these systems meet the user's needs or that of the designer to achieve the purpose of creating a smart home. A smart home is a building that has a dynamic and cost-effective environment. Such an environment results from integrating four main elements: systems, structure, services, and management, and the relationship between them.

Generally, a smart home is a building equipped with a robust infrastructure of communication that continuously responds to changes in the status of the environment and adapts to these changes. The smart home also allows its residents to use the available facilities more efficiently and increase the residents' security and tranquillity. The smart home design can save a considerable amount of energy consumption, and on the other hand, it is straightforward to manage. Computer systems are used significantly in this relation. These systems are known by various names, which are referred to below.

- Building Management System (BMS): Refers to buildings where a central intelligent system controls all electrical equipment. The software and hardware set to integrate monitoring and control of essential and vital parts in the building are called BMS.
- Building Automatic System (BAS): It is possible to control the building remotely using its automatic system.
- Energy Management System (EMS): the primary purpose of using the EMS in a building is energy storage, correct and optimal use of facilities. As a result, the initial investment was spent implementing the BMS return and energy storage [6].

BMS controls different parts of the building and makes the environmental conditions suitable. This system's capabilities include intelligent control of temperature, automatic security control of the building, smart management of fire distinguishing, etc. control and access to the BMS is possible using the associated software inside the building and even outside of it (via the internet). Intelligent control systems have high flexibility that makes them quickly adapt to various needs.

This system's tasks are to continuously monitor various parts of the building and apply commands to these parts so that their performances in the building are balanced with each other. This should also be done in optimal conditions and reduce unwanted consumption, and energy resources should be allocated only to spaces during exploitation [6].

A BMS may include all electrical, mechanical, and security services of the building, as shown in Figure 1. These services include a variety of items, including heating, chilling, air conditioning, elevator, emergency power station, escalator, lighting controls, CCTV cameras, fire alarm and extinguishing, access control, etc.

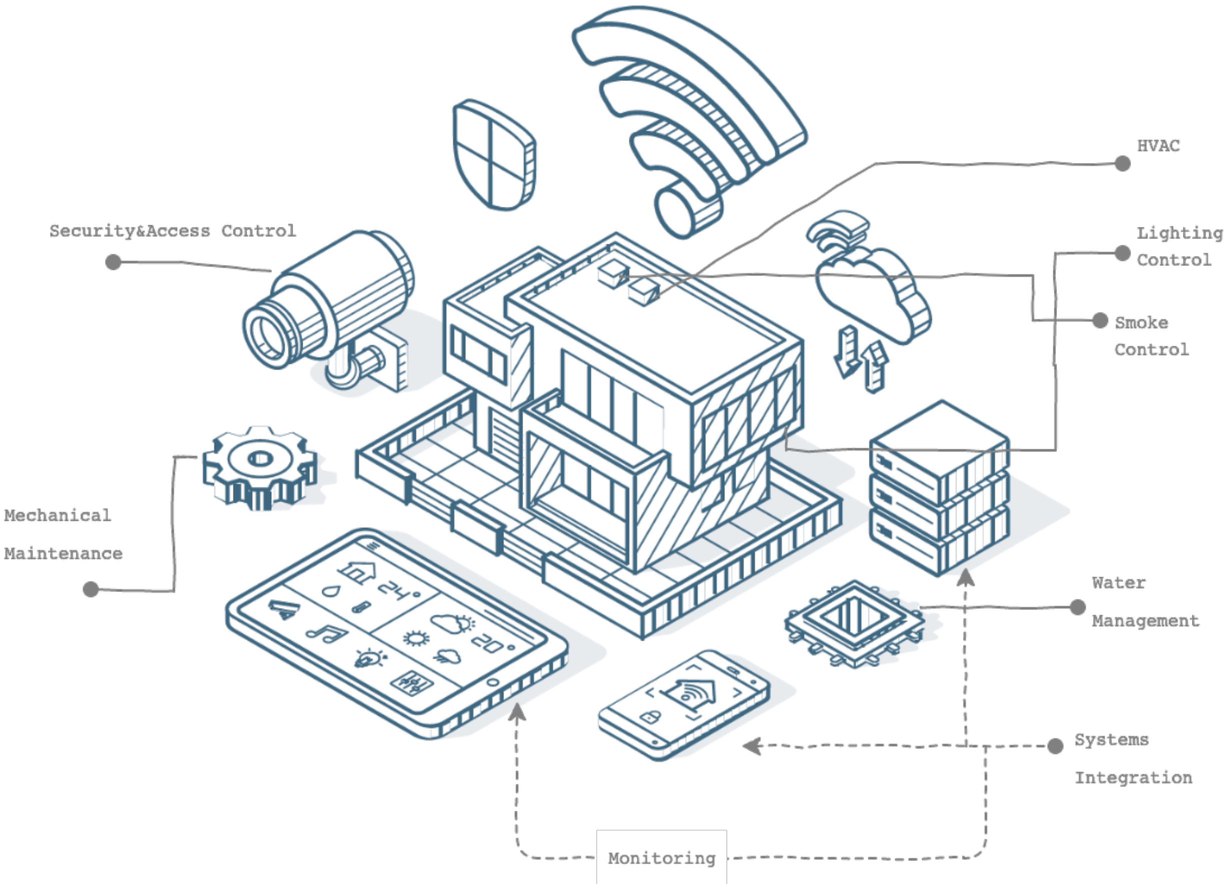

Figure 1. Components of a BMS (inspired by [7]).

Today, smart homes are considered as one of the essential parts of the BMS. Smart homes are evolving rapidly, intending to create mort comfort and security for users. Nowadays, various techniques are deployed to extensive building automation, including lighting control, temperature control, doors and windows control, curtain control, security systems, CCTV, etc. Moreover, other techniques can be implemented in smart homes, such as multimedia control, plant watering, and feeding pets.

2.1.1 Temperature control

When a person is travelling, s/he can activate the home installation system and set the desired temperature using a telephone, mobile phone, or internet. Heating and cooling installation start working at the configured time and adjust the temperature to the desired level. Adjusting the home's temperature and activating the heating and cooling installation creates a favourable environment and prevents energy loss.

The energy consumed in the building is usually spent on lighting, cooling, and heating systems. To save energy, two issues must be considered. First, it is necessary to reduce energy usage during the high consumption hours. Second, it is essential to eliminate the energy wasted when the building is uninhabited. The primary consumption of electrical energy is related to lighting and fan coils.

Figure 2 presents a chart for monthly energy consumption of a regular house which was equipped with BMS since February 2019 showing the total energy consumption divided by the difference between 20° C and the mean temperature of each month [8]. The effectiveness of such BMS system is apparent in data of March and April 2019. Once the BMS established, when the residents leave their unit, the fan coil and all lighting fixtures are turned off. The local controller of that unit controls the fan coils of each unit. Also, to save energy, fan coils receive temperature programs from the central software [9].

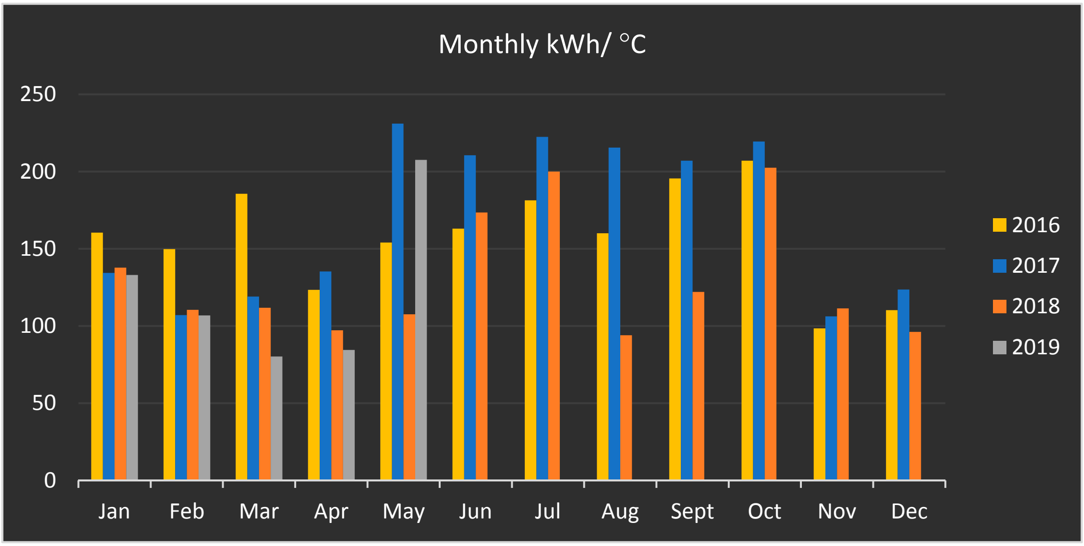

Figure 2. Monthly energy consumption for a regular house equipped with a BMS, from October 2016 till May 2019 [8]

2.1.2 Lighting control

In this method, both natural light and light produced by light sources are used to provide the desired amount of light, and this process is automatic. In general, these systems implement the desired strategy using sensors to detect motion, amount of environment light, and keys embedded in each indoor space. Obviously, in automatic mode, the light sources will be exposed according to the sensors of the amount of light and the desired amount.

2.1.3 Smart home audio and video systems

Equipping the smart home with audio systems, capabilities such as independent internal radio with the ability to store existing music playback stations for each part of the smart home, the ability to receive audio signals through audio methods available in the central unit, the possibility of two-way audio or video communication, alarms and audiobooks would be activated.

2.1.4 Smart home fire alarm systems

Another feature of the smart home is the fire alarm and extinguishing system. That is a system that announces the existence of a fire and also seeks to extinguish it. Today, fire alarm systems are widely used in smart homes and residential and industrial areas to minimize fire damage.

In fire management, timely detection of fire and its exact location is crucial. In all units, sensors related to smoke and heat detection are installed, depending on the unit's uses. These sensors are connected to the local controller of the unit (see Figure 3). When the sensors detect any smoke or heat that indicates a fire, the local controller alerts and notifies the central software. The controller also contacts the fire department and informing the unit owner. The photos of the fire scene are displayed and saved. The exact location of the accident is also indicated on the building map on the primary monitor. Suppose the unit is equipped with a fire extinguishing system and is in automatic control mode or controlled from the control room. In that case, firefighting will be done automatically (sprinklers and capsules are used automatically).

Figure 3. Smoke and heat detection sensors (source: UNIFORE[1])

2.1.5 Smart home security and safety systems

With home automation systems, images of all security cameras can be visible on wireless touch screens. One can also strengthen home security system by using motion sensors. The integration of the security system with other parts will give it more capabilities. For example, when a motion sensor detects a stranger, all the house lights can be turned on and off, except for the usual siren sound.

Each residential unit is equipped with a local controller that controls the digital identification system's entry and exit. People must enter a password or other identification mechanisms to log in. If the password is entered incorrectly, the system will

---

[1] https://www.burglaryalarmsystem.com

inform the central software via the local network, and the alarm sounds. The control room software also shows the exact location of the problem on the building map. The system will also contact the police if needed, and videos and images of the location will be displayed and stored in the control system.

2.1.6 Smart home access and remote control

A smart home can keep one updated about the status of the building and control it remotely. This mechanism can be enabled by, e.g., SMS, Interactive Voice Response (IVR) or the Internet.

2.2 Advantages of an Intelligent Management System in the Building

Owners of enterprise buildings negotiate with various companies to provide fire detection, fire extinguishing, energy saving, saving solar fuel consumption, and CCTV cameras. The number of contractors impose extra costs, including multiple wiring for different systems and experts' need to operate these systems. Integration in a smart home management system eliminates many disadvantages, reduces costs, and increases efficiency. Furthermore, the costs of maintenance and troubleshooting of integrated systems are much lower.

Building management's principal purposes benefit from economic advantages, reducing energy consumption, and creating a safe and quiet environment. Advantages and results of operation of building systems are:

- Creating a favourable environment for residents;
- Optimal use of the equipment and increase their useful life;
- Providing control to schedule performance time;
- Significant reduction in maintenance costs;
- Optimization and energy saving;
- No need for a permanent building contractor;
- Monitoring and control of all points using a mobile, computer, and the Internet;
- Due to the management integration of various facilities and systems in the building, all equipment works in a coordinated manner and eliminates the possibility of conflicting due to lack of coordination;
- Get statistical reports of all equipment and their performance to optimize consumption and performance.

In general, the benefits of smart homes could be mentioned as follows:

Comfort: A smart home using automation makes it easier for residents to do some repetitive tasks. On the other hand, to create a comfortable environment in a smart home, just one hint is enough: Scenarios are responsible for the precise temperature adjustment. Using this efficient software with a multilingual and straightforward user interface to control all the equipment is another situation that makes living in a smart home easy. Using this software does not require any special training.

Safety: In critical situations, such as fire, flooding, and theft, a smart home notifies warnings that could prevent damages or decrease them. The alarm's unique feature in the zoning of the covered spaces, the use of an accurate sensor for detecting a fingerprint's presence, and digital image recording control significantly increase homes' safety.

Flexibility: Flexibility in implementation and use is one of the hallmarks of smart technology. Using the tools provided by this technology, in most cases, there is no need to re-wiring and replace the equipment in the building to add these features to existing homes.

Energy-saving: The energy consumption in a building is a cumulative result of consumption in various building's equipment and areas (See Figure 4). Energy consumption management in smart buildings has a significant impact on energy saving. Activating lighting and ventilation systems based on an individual's presence and planning the rooms temperature during the day are examples of this energy management. Studies show that by using the right control logic, energy consumption can be reduced by up to 40% [10].

Cost reduction: As long as energy consumption is completely controlled, the right parts' price is close to the amount of useful energy consumed. This means that the user pays only the actual cost of the energy required

Integrated control: The intelligent system enables integrated control of the whole building. The control elements in the farthest parts of the building can be easily monitored and controlled, and there is no need to visit the control point.

The most well-known methods employed by BMS designers to reduce energy consumption are: Shutting down the devices according to schedule, utilization of equipment in necessary times, exploiting the minimum permissible capacity in the operation of equipment, monitoring the condition of equipment by trained operators, utilization of the data in troubleshooting equipment failures and evaluating their effective performance.

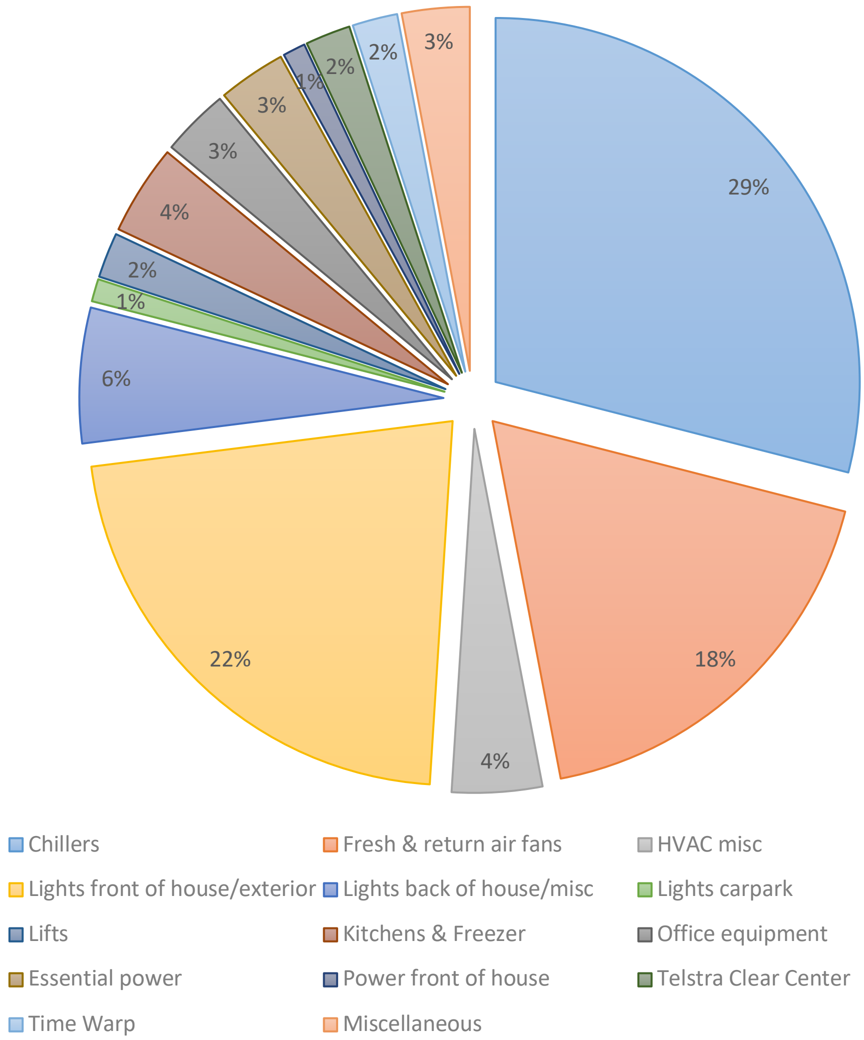

Figure 4. Energy Consumption Distribution in a Building (% of Total) [11]

2.3 BMS in Buildings

BMS is employed in buildings, skyscrapers, commercial, residential buildings, or industrial complexes to control various sectors, such as lights, cooling and heating, traffic control systems, camera surveillance, measuring equipment, security, and protection systems, fire alarm, and elevators. However, these systems link to all mentioned systems in the top and form an integrated control-circuit for all controllable components using standard and architecture-based protocols based on recognized information. Implementing such a comprehensive system in a building will turn it into a secure and smart building. Studies show that the application of BMS reduces 30 % in energy consumption in buildings maximally. The utilization of the integrated systems has a higher 10% energy-saving capability than singular systems [12]. Making smart is carried out using several sensors, which data is collected and sent to the central system, and analysed data are sent to operators. In general, there are three methods for making smart buildings:

1. **Use of the BUS wiring separately to transfer information:**

The most well-known and valued automation standards, such as the EIB, BACnet, LON, S - Bus are examples of this system. Information is transferred according to the standard using 1, 2, or 4 twisted pair wires, entirely isolated from the electrical wires. In this way, the sensors send data to a control centre (see Figure 5 for an example). The control centre sends a command to operators, mainly at the same control centre, and these operators cut-off or connect the power supply.

Figure 5. Example of a Central Controller Unit (source: 75F[1])

2. Use electrical cables as a platform to transport and transmit data (PLC)

In this system, the most widely used standard is 10X. In this method, the information is transferred without separate wiring on the existing electrical wires. In this system, some modules are installed at each control point, which receives data from the control centre, responds to it, and turns off the connected device. This system's control centre performs the scenarios' storage task, receiving codes from the remote control, and eventually sending it to the desired module.

[1] www.75f.io

3. Use wireless systems to transport and transmit data

A different method that has been studied nowadays is the usage of signals as a communication base of the smart system. In this method, the commands are sent and received via wireless signals. The essential standards of this method are Z-Wave and Zigbee [13].

In addition to the recipient's role, it also resumes its retransmission to expand the system's average area. In this way, in addition to the high speed of information transmission, there is no need to change in the building's wiring and, therefore, easily applicable to the cells that are not allowed to change in the wiring system.

2.4 The Architecture of Smart BMS

The architecture of the system is usually graded at three levels (See Figure 6). Level zero includes tools, equipment, sensors, and the final components of the control. The Monitoring and Engineering (M&E) systems are in this sector and are transmitted to the integrated controllers through inputs and outputs. This transfer may take place directly or through designed boards. Components remain after the system is set off and include I/O systems, controllers, and Level 2[1] communication software, and all control and logic algorithms are executed at this level.

BMS control software holds various capabilities. The software belongs to level 2 and is installed on appropriate servers. They usually have a suitable and straightforward graphical environment for ordinary users of the library, a variety of solutions and programs to design and develop the system in the future, provide Preventive Maintenance (PM) without the need for extra PM software, ability to define latitude and longitude to automatically adjust sunrise and sunset conditions and control energy consumption, ability to define accessible security layers for different users, ability to define security

[1] Level 2 or the monitoring level consists of information management tools, including Human-Machine Interfaces (HMIs), servers, storage equipment, and the workstations of operators and engineers who must be connected to the BMS system.

layers for the users of subsystems including Access, HVAC, and Lighting, the possibility of storing software information in databases, preparation of adjustment and comparison of various operating diagrams, including power consumption diagrams in different trends, straightforward communication of graphic software and layers of the information storage system, tracking of trees and topology of networks using BACnet, by online software [9].

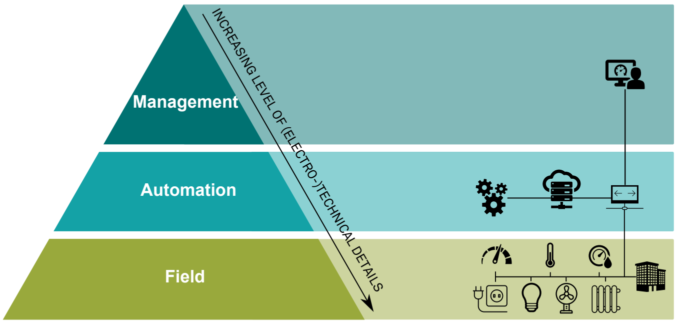

Figure 6. Automation Hierarchy according to Kastner et al. [14]

2.4.1 Electronic equipment controlled in smart homes

The smart home offers various possibilities that each part of its interest and taste could be used only in part. Using this technology, all the home equipment is displayed and managed in terms of location, features, and users' needs to the screen.

Lighting: About 19 % of the world's energy consumption is used to light residential buildings [15]. However, the power is used the lighten some spaces, such as the staircases of a storeroom and parking lot, which is sometimes a user in that space whether it passes through it, see Figure 7.

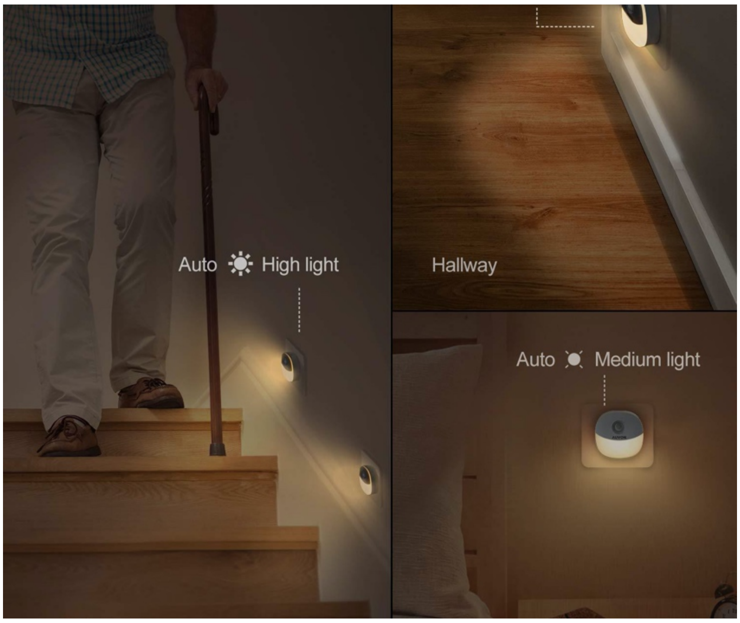

Figure 7. Partial presence in stairways, corridors and parking lot (source: AUVON[1])

Security system: The smart home could manage the alarm system of closed-circuit television (CCTV) cameras and biometric systems. The main advantages are the high-control ability to send SMS over the cellular zoning to cover and detect smoke. Doors, window, shutters, and canopies are easily controllable. When leaving the building, one can look at the screen to be aware of the closed or opened state of all the doors and windows.

Home entertainment: In a smart home, it is possible to use music and movie archives according to personal tendency and separately from each room. Music and movie files, are now used in all building rooms without moving the device or even the desired music or film files

Video door-phone: When the doorbell rings at home, the image would be displayed on all screens or the desired screen.

Automatic watering system: Flower and plant watering and scheduling artificial waterfalls could be easily controlled through the screen control. Smart homes could also control flower and plant watering in the courtyard or indoors automatically according to

[1]https://www.iauvon.com/

a predetermined schedule. For instance, when the humidity is below a specific level at each sunset, the grasses are automatically watered.

Controlling the facilities of the pool, sauna, and jacuzzi: In smart homes, the lighting, heating of sauna, finding out if the desired temperature is provided in the sauna, Automation of the purification of pool water and jacuzzi is all easily possible. Communication systems are one of the features of smart homes. Besides, in the covered environment, in-house videotelephony could be employed.

Building electrical appliances: In a smart home, it is possible to find out and control the condition of all electronic devices; for example, after the washing machine is finished, it is informed that when using the vacuum cleaner, the video door phone rings automatically.

2.4.2 How to access and control facilities in the central control panel of the smart home

A central control panel (See Figure 8) enables people to access intelligent control of all building devices and be aware of their operating conditions. The screen is equipped with touch screen technology. This technology makes it easier to use the control panel and does not require another input device.

Figure 8. Central Control Panel [16]

Controlling via radio waves: Wireless control of the building is possible using this technology. The control device can be a PC, tablet, smartphone or a simple radio remote control. The main advantage of using RF over IrDA (technology used in TV controls) is a long-distance and no direct vision.

Remote control: smart home enables users to be informed of the current situation and control the building remotely. They can take advantage of remote control via SMS, Recall Centre, and the internet.

Scenario: Scenario definition is that a set of actions can be done with the push of a button or touch on the control panel. The primary purpose of using scenarios is more speed and convenience in creating an appropriate space.

Automation: The smart home takes control of some of the day-to-day and repetitive activities, and as a result, brings more comfort and convenience to the residents. Automatic system operation can include predetermined activities or using different sensors such as temperature, light, wind, rain, hygrometry, and presence detection sensors, optimally adjust the conditions. Automatic operation can also announce the necessary warnings [17].

2.4.3 How to save costs by owners

The building's actual price is calculated according to the amount of investment and construction and commissioning cost. These costs include building valuation, equipment value, investment value, and value of procedural activities. Reducing the cost of commissioning a building is a result of using a building management system. A building management system can reduce costs using the methods that listed below:

Start and end planning: building a management system can prevent the building installation from being active 24 hours a day, seven days a week, or in non-consumable situations.

Economic model: using the building management system provides several possibilities. For example, when an office is closed, all facilities are turned off by this system. As another example, if the office temperature when it is empty is about 22°C, this system can increase the temperature to 28°C as soon as people enter.

Optimal starting and ending: changing the operating hours of machines depends on the weather conditions inside and outside. Once the ambient temperature has reached the level set for people, there is no reason for the machines to be active. Turning on machines and air conditioners depends on the ambient temperature, so they should operate with the least possible time.

Adaptability control: the building management system uses microprocessor controls that are intelligent in themselves. This allows microprocessors to learn the proper operation of a system.

Abilities of information technology: many of the systems mentioned have been used successfully in building design for many years. Nowadays, the advantages of information technology have allowed us to build and maintain buildings more efficiently.

See Table 1 for global advanced building efficiency statistics in an 8-year period.

Table 1. Global Building Efficiency Revenue (million USD) [18].

Subsegment	2011	2012	2013	2014	2015	2016	2017	2018 estima
Building Design	11,537	13,039	13,932	15,544	17,457	19,021	21,180	23,60
Building Envelope	13,017	14,006	15,855	18,668	26,187	44,960	25,676	28,96
HVAC	44,383	49,613	53,923	57,962	62,337	65,820	71,489	78,59
District Energy and CCHP	2,229	2,787	3,023	2,950	3,458	3,875	4,291	4,834
Water Heating	1,467	1,612	1,773	2,090	2,237	2,438	3,182	4,774
Lighting	41,329	47,212	52,770	103,613	116,498	124,782	133,561	139,47
Appliances	266	613	800	1,200	1,708	3,018	4,676	7,481
Demand Response & Enabling IT	3,827	4,957	5,262	6,200	6,612	7,675	9,291	10,75
Total	118,055	133,839	147,339	208,228	236,494	271,588	273,345	298,47

3 Chapter 3 – Machine Learning Methods

3.1 Introduction

In this chapter, some of the most relevant research in the field of machine learning are reviewed. The chapter also generally reviews the literature and presents details about the algorithms, variables, and unknowns of the research issues.

3.2 Different machine learning algorithms

Algorithm architecture well defines instructions for expressing functions. Algorithms are used to calculate automatic processing and independence. In simpler terms, an algorithm is a step-by-step method for calculating [19]. Machine learning can be categorized into several levels, which are discussed below.

3.2.1 Supervised learning

Supervised learning is classified into two categories: classification and regression. From a general point of view, classification and regression are the two main types of issues in which classification is used to predict discrete and nominal values, while regression is used to predict continuous values. Supervised learning is responsible for inferring a function from training data labels. In supervised learning, each input is a pair of an input value and an output value. A supervised learning algorithm analyses training data and draws an inference function, which could be used to map new samples. An optimal function is a function that could specify a class label for test data. Linear and logical regression are some of the regression algorithms. Moreover, among the most important and widely used classification algorithms are the decision tree, the SVM, and the K-nearest neighbour [20].

3.2.2 Unsupervised learning

Unsupervised learning is also a machine learning method that provides a function to describe the hidden structure of unlabelled data. Since unlabelled samples are provided to learners, there is no reward or punishment for evaluating a potential solution, making a difference to supervised learning. Among these algorithms, we can mention clustering algorithms such as K-mean and C-mean [20].

3.2.3 Semi-supervised learning

This method usually involves a mixture of a small amount of labelled data and a massive amount of unlabelled data. Semi-supervised learning is between unsupervised learning and supervised learning. Various machine learning researchers have concluded that unlabelled data could significantly improve learning accuracy when used in combination with a small amount of labelled data [21].

3.2.4 Reinforcement learning

Reinforcement learning is an essential type of machine learning inspired by the psychology of behaviourism. In this approach, the steps to maximize the concept of cumulative value are trained. This type of learning is also studied in many other disciplines, such as game theory, control theory, operational research, information theory, simulation-based optimization techniques, multifactorial systems, congestion intelligence, statistics, and genetic algorithms [21].

Due to a large amount of existing supervised learning techniques and the type of data in this study, only the most relevant supervised methods are described in what follows.

3.3 Machine learning methods

Whenever data has set of features that are not directly derived from other features, but there is a dependency between them, that specific feature could be identified by discovering a specific model based on the relation with other features. Suppose that there are details of some patients in a database that have already been identified and grouped into two types of diseases using a specific clinical test. Here, no one can have both diseases, or be healthy (having none of the diseases), or suffer from another unknown disease, which means that the categories separate the problem's space efficiently. There is a specific record for each patient in such database, including the patient's symptoms, the patient's name and disease label. A data miner decides to discover a method to diagnose disease without testing and only from the symptoms. This decision may have been made for any reason, such as a lack of facilities. What needs to be performed is called *classification operation*. The purpose of the categorization is to map the inputs X to outputs Y where $y \in \{1, \dots\dots, C\}$ and C specifies the number of classes. If C = 2, it is the binary classification otherwise If C > 2, this type of classification is called multi-class classification [22].

Data classification is a two-step process. The first step is to develop the model and the second step is the class prediction through the developed model. One must divide the data into two data sets: training data and test data. Training data is used to create a classifier based on which unlabelled data can be placed in their respective categories.

The classifier's performance is measured using test data (randomly selected from the data), and the model is run through them to check the accuracy of the classifier's prediction. Thus, a model with the appropriate accuracy is chosen for data classification.

In classification, learning is carried out by samples, and the label of each data set is defined clearly. The samples are then assigned to a predefined category according to their features.

3.3.1 Common groups of Machine Learning methods

There are many ways for classification, including the following:

- ANN
- Decision trees
- Bayesian networks
- K-nearest neighbourhood
- Support vector machine
- Rule-based methods

3.3.1.1 Artificial Neural Networks

The study of artificial neural networks is largely a matter of natural learning systems, in which a complex set of interconnected neurons is involved in learning. The human brain approximately consists of 10^{11} neurons, each of which is associated with approximately 10^{4} neurons. The transfer rate of neurons is about 10^{-3} second, which is very low compared to computers (10^{-10} second). However, one can identify an image of a human in 0.1 second. This extraordinary power must have come from parallel processing distributed across a large number of neurons [23].

ANN perform learning based on samples and training data sets. Therefore, they work like a human being. Another advantage is that these networks could inherently generalize; meaning that they could recognize patterns that resemble the samples they have already learned, rather than just recognizing patterns exactly like the trained data sets [24].

ANN is a practical method used to learn various functions such as real-valued functions, discrete-valued functions, and vector-valued functions. A single neuron can only be used to identify linearly separable functions, and since functions are generally not linearly separable in real data sets, a network of neurons is required.

There are several types of neural networks including supervised learning, unsupervised learning, and reinforcement learning used to solve problems. NNs are divided according to the type of connections, including Feed-forward Neural Networks (FNN) and Recurrent Neural Networks (RNN). FNN is the most common type of NN which has various applications. The first and last layers are named input and output layers respectively and any number of layers between these two are called middle or hidden layers. The idea behind this naming is because practically one only deals with the input and output of the NN. The NN performs as a black box and direct access to the middle layers is not possible. NNs have directed cyclic graphs in their communication structure. This means that it is possible to return to the previous and initial nodes by following the connections between nodes. Due to the structure of RNN, they have complex dynamics, and this makes the training of these networks very complex.

FNNs with more than one hidden layer are called MLP (Multi-Layer Perceptron), and FNNs with one hidden layer are called SLP (Single-Layer Perceptron), in which the output of neurons in each layer is a non-linear function of the output of the previous layers. The number of input and output layer neurons is constant. The number of input layers is equal to the characteristic space, and the number of output layers is determined according to the number of classes. In MLP (see Figure 9), nodes are usually arranged in layers. Each node receives inputs only from the previous layer and provides a function of inputs.

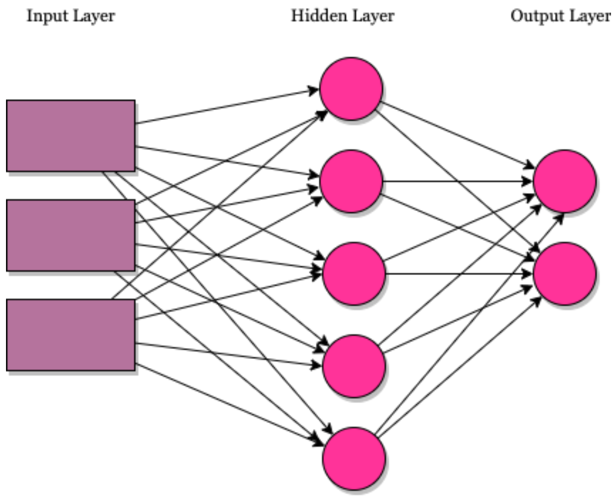

Figure 9. MLP architecture [25]

Each unit emits an output that is a non-linear function of the input values [26]. f is an activation function applied to the sum of the weights multiplied by each node's input. The most popular activation function used in neural networks is recalled the Sigmoid or Logistic function. In this function, based on Equation (1), the NN's behaviour is determined according to its weights' values.

$$f = \sigma(WX) = \frac{1}{1+e^{-wx}} \tag{1}$$

The NN learns the best values of the weights and biases according to the available data set. Neural network learning involves adjusting weights and biases until one reaches certain conditions (See Figure 10). Weights should be adjusted to reduce the error rates between the desired output and the neural network output.

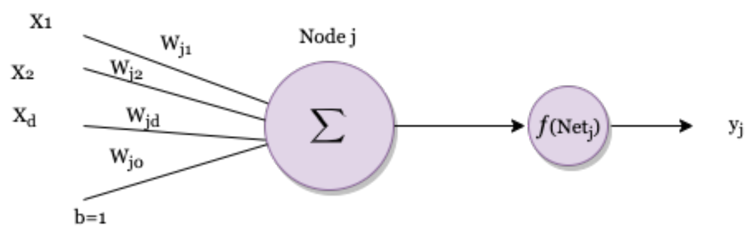

Figure 10. Activation function

FNN training (determining weights and biases) is done in two ways, including classic methods such as back-propagation (BP) algorithm and intelligent optimization methods such as genetic algorithm and particle swarm optimization (PSO) algorithm.

The BP method is created according to the descending gradient in the error space with local search capability. Modification of NN's weights is carried out to reduce both the error between the desired output and the neural network's output. This error is defined according to Equation (2):

$$E = \frac{1}{2}\sum_{n \,\epsilon\, training}(t^n - y^n)^2 \tag{2}$$

Therefore, the error is calculated for a set of training samples with n members. Further, t is the optimal output, and y is the output of the neural network. The BP algorithm could calculate the effective error for each hidden unit. Finally, each of the weights in the $m + 1$ cycle change from Equation (3) to Equation (4):

$$w(m+1) = w(m) + \Delta w \tag{3}$$

$$\Delta w = \eta\delta x \tag{4}$$

In Equation (4), η is the learning rate, and δ is the difference between the desired output and the output of the NN.

The convergence to a minimum is time-consuming in methods that operate based on the descending gradient of BP. Moreover, in these methods, if there are several local minimums at the error level, there is no guarantee that it could find the absolute minimum [27].

Evolutionary methods were employed to prevent local optimum trap and increase generalization, which was one of the disadvantages of the descending gradient algorithm for NN training. In these methods, initial populations are predetermined or randomly determined. Each member is one of the potential solutions in which the desired evolutionary algorithm searches the problem space during different periods and moves the population to the optimal point of action [28].

3.3.1.2 Decision trees

Top-down decision making is one of the common classification algorithms [29]. The transparency and interpretability of this algorithm is the reason for the widespread usage of it. Another advantage is the presence of powerful implementations such as C4.5 algorithm [30]. Decision tree algorithms operate by constructing a top-down algorithm by selecting a feature at any time and classifying data according to their feature values. The most important features are considered the root of the tree and the rest of the group are placed next in order of priority so that the nodes showing the coefficient of data access and class label are located near the root.

Figure 11 shows how to build a decision tree for Table 2.

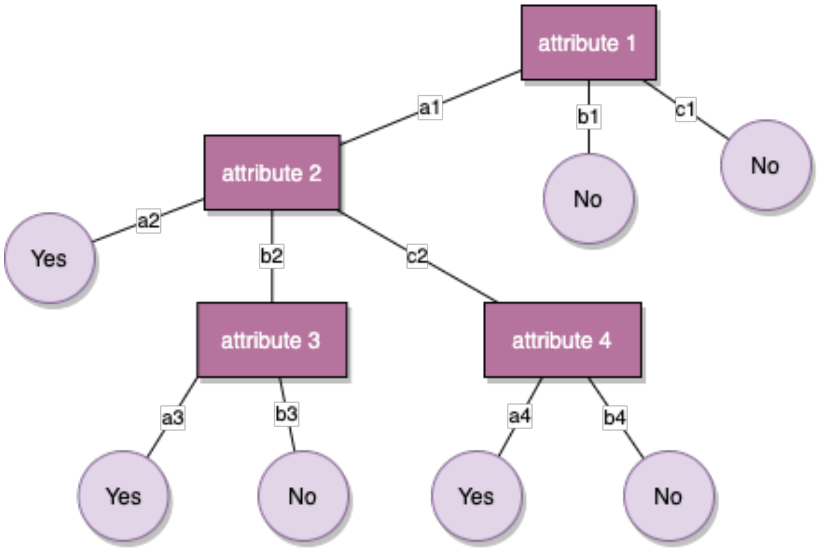

Figure 11. Decision Tree

Table 2. Training data sets

First feature	**Second feature**	**Third feature**	**Fourth feature**	**Class**
a1	a2	a3	a4	Yes
a1	a2	a3	b4	Yes
a1	b2	a3	a4	Yes
a1	b2	b3	b4	No
a1	c2	a3	a4	Yes
a1	c2	a3	b4	No
b1	b2	b3	b4	No
c1	b2	b3	b4	No

To increase the tree's interpretability, it is required to reduce the tree's size, which reduces stability. Various methods were employed to determine the optimal structure in classification problems. When we want to use the decision tree for extensive data sets, these algorithms' instability becomes more apparent since accessing all the data at once and creating a unique decision tree is not practical.

3.3.1.3 Bayesian networks

Unlike other classifications, statistical classification methods show a sample membership to each class with a probability. Bayesian Network is the most common, simple, and effective statistical method. In this method, the algorithm learns the conditional probability of a given error by the class label of training data. Afterward, the classification is performed by applying Bayesian rules to calculate the given sample class's probable value with high accuracy. This is usually carried out by estimating the probabilities of each possible combination of features. However, this is not possible when the number of features is too high. Thus, a robust independent assumption is made that all features are independent, given the class feature's value is determined. According to this assumption, it is necessary to calculate only the marginal probabilities of each class feature. However, this assumption is unrealistic, and Bayesian networks do not consider the implicit modelling of dependencies between its features [31].

The learning of a Bayesian network structure is described by having an instruction set $A = \{u_1, u_2, \dots, u_n\}$ of n samples of u; we should find a network that best matches A. The most common method is to introduce a goal function based on which any network is evaluated considering the training data [32]. The critical optimization challenges are selecting the goal function and determining the search flow for the best network.

The Bayesian networks provide an intuitive representation of the relationships linking heterogeneous sets of variables, which we can use for qualitative and causal reasoning. The main role of the network structure is to express the conditional independence relationships among the variables in the model through graphical separation [33].

Each variable represents a node in the Bayesian network. A variable, knowing its parents, is independent of non-child nodes. Suppose we attempt to turn on our computer, but the computer does not start (observation/evidence). We would like to know which of the possible causes of computer failure is more likely. We assume only two possible causes of this misfortune: electricity failure and computer malfunction [33]. The corresponding directed acyclic graph is depicted in Figure 12.

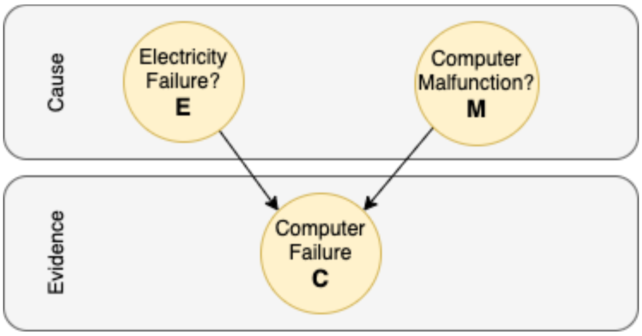

Figure 12. Directed acyclic graph representing two independent possible causes [33].

Equation (5) is used to express the *conditional probability distribution* of cause given the observed evidence using the converse conditional probability of observing evidence given the cause.

$$P\ [Cause \mid Evidence] \ = \ P\ [Evidence \mid Cause] \cdot \frac{P\ [Cause]}{P\ [Evidence]} \quad (5)$$

Table 3 shows an example setting for a computer failure (C) with respective probability

E	M	C	
		T	**F**
F	F	0	1
F	T	0.5	0.5
T	F	1	0
T	T	1	0

values for electricity failure and computer malfunction. We suppose that electricity failure, denoted by E, occurs with probability 0.1, $P\ [E \ = \ true] \ = \ 0.1$, and computer malfunction, denoted by M, occurs with probability 0.2, $P\ [M \ = \ true] \ = \ 0.2$.

E	M	C	
		T[1]	**F**[2]
F	F	0	1
F	T	0.5	0.5
T	F	1	0

[1] True
[2] False

T	T	1	0

Table 3. An example setting for an evidence C and causes E and M

In this setting, the probability of computer failure $P\ [C\ =\ \text{true}]$ can be calculated as:

$$P\ [C = true]\ = \sum_{E,M} P[C = true, E, M] = \sum_{E,M}(P[C = true|E, M] \times P[E]P[M]) = 0.19$$

3.3.1.4 K-nearest neighbourhood

The K-nearest neighbour algorithm is an example of sample-based learning in which a training dataset is employed to generate a classification model. Classification for an uncategorized sample may be found simply by comparing it to the most similar samples in the training data sets. For each new sample, the algorithm is such that comparing it with the K nearest training samples determines the target class [34]. Therefore, it is necessary to specify the distance between the samples. Table 4 could be employed to determine the distance between two samples $x = (x_1, x_2, \ldots\ldots, x_n)$ and $y = (y_1, y_2, \ldots\ldots, y_n)$:

Table 4. Distance functions for x and y samples

Distance function	**Equation**
Euclidean	$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$
Hamming	$d(x,y) = \sum_{i=1}^{n} \lvert x_i - y_i \rvert$
Chebyshev	$d(x,y) = \max_{i=1,2,\ldots.n} \lvert x_i - y_i \rvert$
Minkowski	$d(x,y) = \sqrt[p]{\sum_{i=1}^{n} (x_i - y_i)^p}$ p>0
Canberra	$d(x,y) = \sum_{i=1}^{n} \frac{\lvert x_i - y_i \rvert}{x_i + y_i}$

Obtaining distance criteria is easy for numerical data; however, group variables require special mechanisms for distance. The K-nearest neighbour method's computational time increases exponentially from all points, so the algorithm is computationally expensive. Whilst using the decision tree and neural network is much faster.

3.3.1.5 Support Vector Machine

SVM is one of the kernel methods used for machine learning classification. Vladimir Vapnik[1] introduced SVM in 1992, which is based on the statistical theory of learning. This is one of the important algorithms in supervised learning that is employed for classification and regression. This pattern simultaneously maximizes geometric margins and minimizes the empirical error of classification and is also referred to as Maximum Margin classification [35]. For a classification problem with two classes, there may be infinite lines by which the classification is carried out, but only one of these lines provides the maximum separation and classification. Among linear classifications, the separation that maximizes the margin of training data will minimize the generalization error. Data points may not necessarily exist in R^2 space but they may exist in R^n space. Linear classification may very well implement these properties, but SVM is looking for a classifier that provides a maximum classification for classes.

According to Figure 13, hyperplanes (represented by the oblique lines) that are closely related to the training data are sensitive to error. Moreover, they are less probable to have good generalizability for data outside the training datasets.

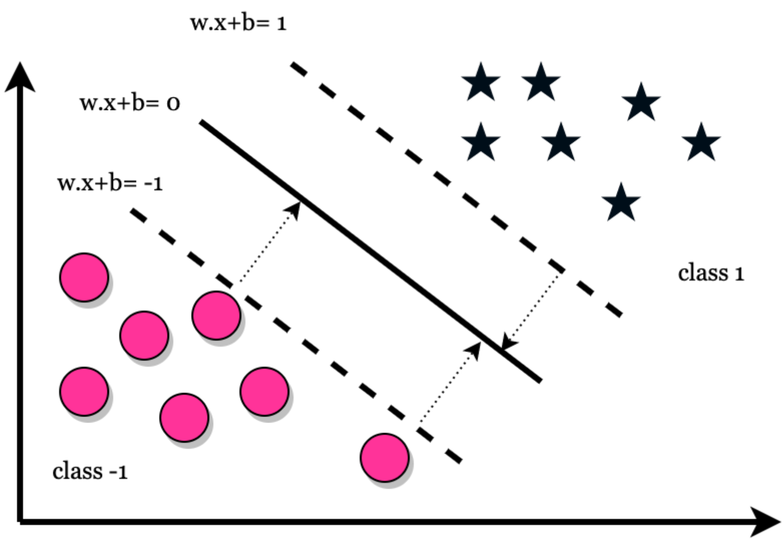

Figure 13. SVM Classifier [36]

In contrast, it seems a plane farthest to all training samples provides the most appropriate generalization capability. The closest training data to the hyperplane classifier is called

[1] https://en.wikipedia.org/wiki/Vladimir_Vapnik

Support Vector (SV) [37]. If the dataset is displayed on the page as $\{(X_1, Y_1), (X_2, Y_2), \ldots\ldots, (X_n, Y_n)\}$, Y_i could take the value of -1 and 1, which are used to determine the sets of X_i points, each of which is an n-dimensional vector. Once one has the training data classified into the right classes, the support vector classifies them into hyperplanes and places them in separate classes so that the $W^T X + b = 0$ and the W vector classifies the vertical points and b specifies the margin. Parallel hyperplanes could be defined as $w \cdot x + b = 1$ and $w \cdot x + b = -1$.

If the training data is linearly separable, hyperplanes could be selected so that there are no samples between them, and then try to maximize their distance. For each sample i, we have the following equations:

$$(w.x_i + b) \geq 1 \; (w.x_i + b) \leq 1$$

$$y_i(w.x_i + b) \geq 1, 1 \leq i \leq n$$

The distance between the two hyperplanes could be obtained through geometric analysis with $\frac{2}{\|w\|}$. Thus, the optimization problem would be as follows:

$$\max \frac{2}{\|w\|} \, or \, min \frac{1}{2} \|w\|^2$$

It could be imagined that the SVM draws a plane between the two data sets and separates the data on both sides of this plane. This hyperplane is positioned so that the two vectors first move away from each other and move so that each vector approaches its first closest data. Furthermore, the plane drawn between the middle of these two vectors will have the maximum distance from the data and is the optimal classifier.

So far, we used SVM assuming that the training samples are linearly classifiable. As is known, in practice, the distribution of data of different classes may not be easily classifiable and may interfere [35]. In this case, the classification may result in a poor generalization.

One solution is to accept some errors in classification. This is performed by introducing a slack variable ($\boldsymbol{\xi_i}$) representing the samples that are evaluated as an error by the $w^t x + b = 0$ function. This method, known as soft-margin SVM, allows some samples to be placed in the wrong region and then fines them, so unlike hard-margin SVM, this method could be employed for cases where training samples are not linearly classifiable (See Figure 14).

By introducing the slack variable ($\boldsymbol{\xi_i}$), the previous constraints are simplified, and the Equation (6) changes as follows:

$$\boldsymbol{y_i(w.x_i + b) \geq 1 - \xi_i, \xi_i \geq 0} \quad (6)$$

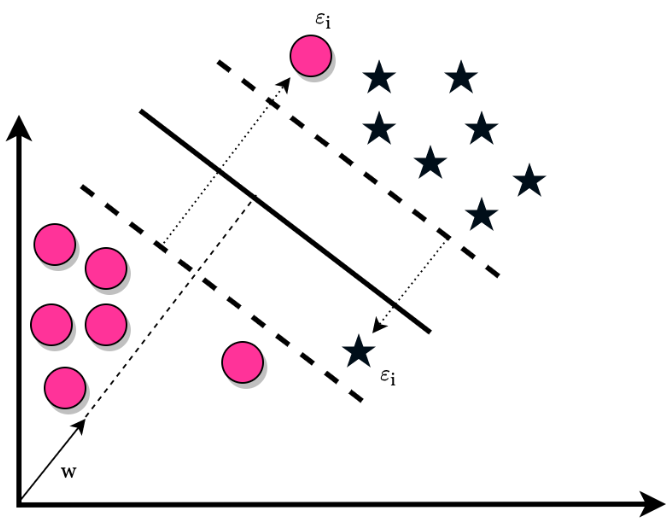

Figure 14. Soft-margin SVM Classifier [36]

In this case, the optimization problem becomes finding a w to minimize the equation below:

$$min\left\{\frac{1}{2}\|w\|^2 + c\sum_i \xi_i\right\}$$

The soft-margin SVM tries to keep ξ_i at zero while maximizing the margins of the classifier. SVM does not reduce the number of incorrectly classified samples but tries to minimize the total distances from hyperplanes' margins [35]. Large values of c cause the equation to act differently compared to the hard-margin method. The soft-margin SVM always finds a solution and is resistant to data sets that have an outlier member.

3.4 Unsupervised learning techniques

Unsupervised learning is also one of the machine learning categories. If learning to find hidden patterns is made on unlabelled data, it is called unsupervised [38]. Some unsupervised learning methods include clustering algorithms, some artificial neural networks, and the Markov model.

3.4.1 Clustering algorithms

Segmenting data into meaningful sections or cluster is known as clustering. In these clusters, two conditions must be met; on the one hand, each cluster's contents must have similar properties, and on the other hand, these contents must have different properties to the objects in the other clusters. This algorithm is used for cases in which the number of features and the data set is large.

In some cases, clustering may not make sense at first glance, but it should be considered that more in depth analysis is needed to understand the precision of this algorithm [39]. There are several types of this algorithm, some of which are described below.

3.4.1.1 K-Means clustering algorithm

This algorithm is one of the most important and widely used clustering algorithms. This algorithm works in an iterative manner and is used for classification of data into several user defined clusters. The K-Means algorithm is a straightforward and logically scalable algorithm that is often used in cluster segmentation. This algorithm can also be easily modified to be used in various methods, such as pseudo-counselling learning. Advances and general principles of the K-Means basic algorithm gradually increase its effectiveness. In this algorithm, the parameter K, which is the number of determined clusters, is considered as input. The K-Means algorithm divides a set of n objects into k clusters so that on the one hand, the level of similarity inside the cluster is high, and on the other hand, the level of similarity outside the cluster is low. The similarity of each cluster is measured in accordance with its centre [39]. This algorithm is carried out in different steps, which are discussed below. In the first step, the desired k points are randomly selected from the available n points as the initial clusters' centres. In the next step, according to each object's similarity to the clusters' centres, it is placed in a cluster. The average of the objects in each cluster is calculated; the clusters' centres are updated. In the last step, it returns to the second step and gets repeated. This step is repeated until there is no change in the cluster due to the clusters' new centres, in which case, the algorithm is literally terminated [40].

This algorithm works well when the clusters are compact and separate from each other. Therefore, this method recognizes only spherical clusters and is not applicable for detecting clusters with complex shapes, especially convex one. This method often leads to local optimization, not a global one [40].

Different methods have been developed to solve the problems in the K-Means algorithm which differ in strategies for defining cluster centres, selecting initial centres, and calculating dissimilarity. To define initial centres, the hierarchical algorithm is employed to find a proper cluster among the data, which is considered as an extension to this algorithm and the obtained clusters are then used as the data for the first step of the K-Means algorithm [39].

Another method developed for the K-Means algorithm is the K-Mode method, which is used for ordinal data. In this algorithm, cluster centres have been substituted by cluster modes. This algorithm is a method based on frequency. Thus, it uses a new equation to measure the dissimilarity for nominal or ordinal data. Compared to eccentric data, the compensation of sensitivity deficiency is one of the advantages of the mentioned method because the median is not affected by large amounts. It should be noted that the problem with this algorithm is that often, the representative of the clusters is not selected from the objects [40].

3.4.1.2 Hierarchical clustering

Hierarchical clustering is one of the clustering methods that itself includes two types of clustering:

Single linkage: this method is also known as the Bottom-Up method. In this method, each data is first considered as a cluster. It then uses an algorithm to merge clusters with similar properties each time. This step continues until the desired number of separate clusters is obtained (see Figure 15). The problem with this method is the high memory consumption and noise sensitivity [41].

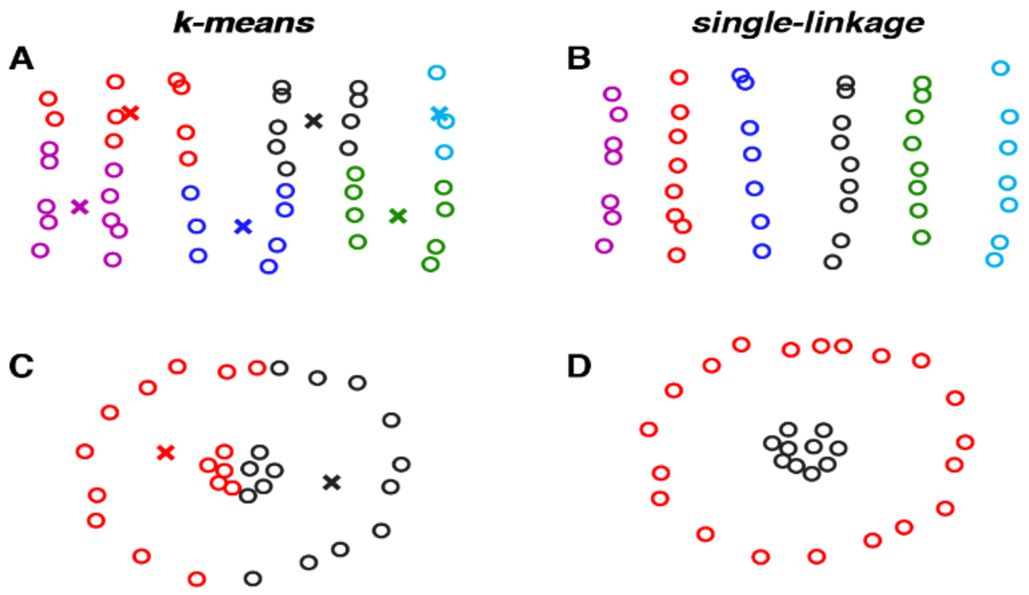

Figure 15. A sample single linkage clustering [42]

Complete linkage: this method is also known as the Top-Down method. In this method, at first, all data are considered as a cluster, like the single linkage method. Using an iterative algorithm, separate clusters are obtained by the data that has the least similarity to other data. This stage continues until one or more single-member clusters are obtained (see Figure 16). It is notable that, in this method, the noise problem is solved [41].

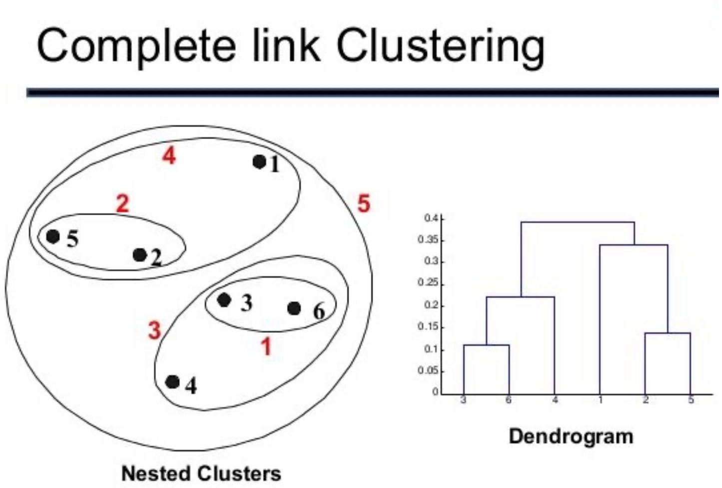

Figure 16. A sample complete linkage clustering [42]

3.4.1.3 Density-based clustering

The basis of this clustering method is the density difference, in that the clusters of higher density regions and lower-density regions are separated. One of the essential algorithms in this field is the DBSCAN algorithm [43]. This algorithm operates such that each cluster's data is accessible to a density of the same cluster's data; however, there is no access to the density of different clusters (see Figure 17). This method's advantages include determining the number of clusters automatically and the high ability to detect noise [39].

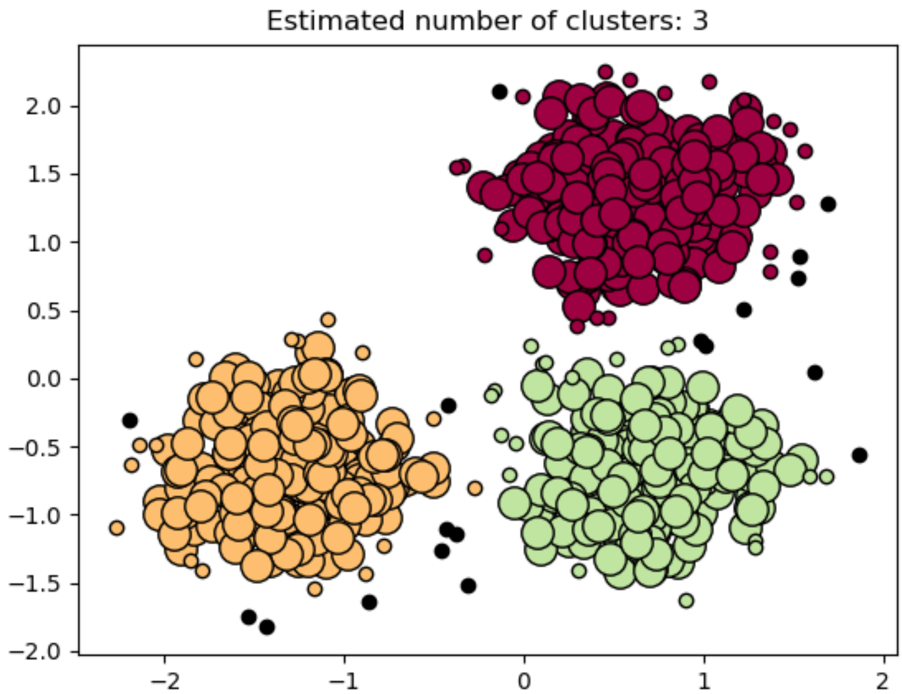

Figure 17. Demo of DBSCAN clustering [44]

This chapter has given an overview of some of the most relevant machine learning methods. The next chapter describes the proposed method that builds upon SVM and NN for application to analysis of smart home residents' behaviour analysis use-case.

4 Chapter 4 – Proposed Method

4.1 Introduction

Several methods have been proposed by researchers to reduce energy consumption in smart homes [45] [46] [47] [48]. Despite various suggested methods, still energy consumption is of high impact and challenging to deal with. In this thesis, we analyse users' behaviour by focusing on machine learning techniques (NNs and SVM) and, consequently, their activities at home, and based on these activities, we would then be able to better manage and control energy consumption.

We initially present a flowchart of the proposed method's actions and implementation to analyse residents (users) behaviour with the view to reduce smart homes energy consumption. The flowchart describes all steps in details. Then we evaluate the proposed method. At the end of the chapter, evaluation criteria for comparing the proposed method against other methods are provided.

4.2 Flowchart of the proposed method

In this section, we would present the proposed method and the steps involved in detail according to Figure 18.

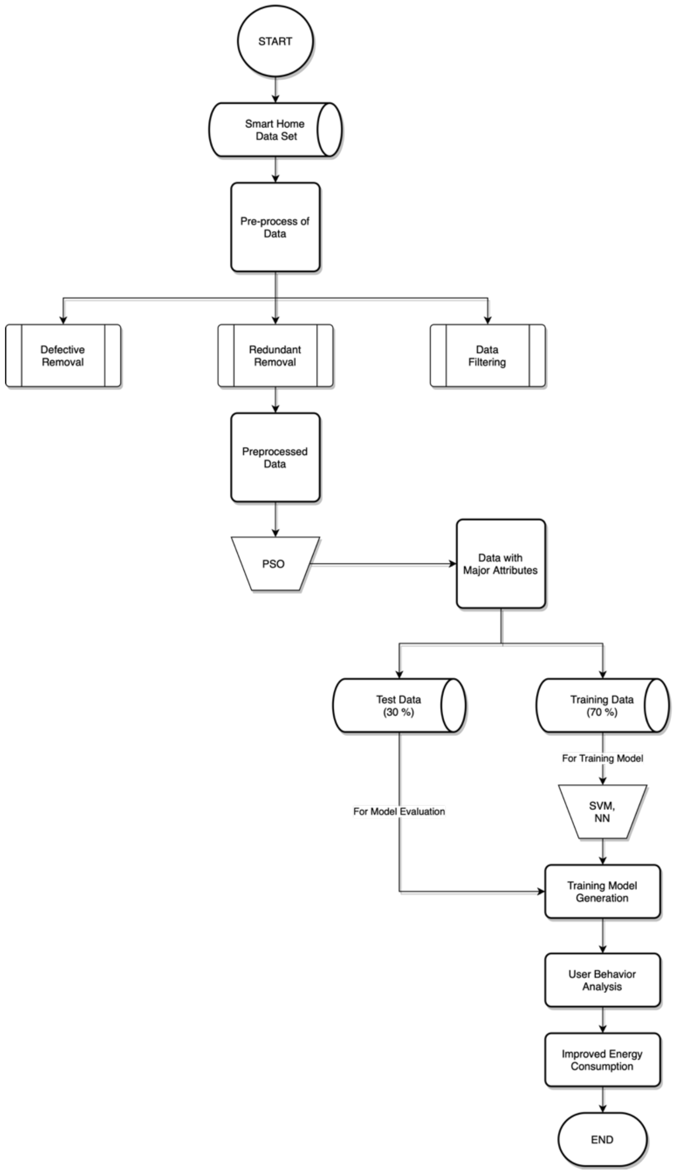

Figure 18. Proposed Method Flowchart

According to the flowchart steps, the following is a complete description of each phase in Figure 18. The steps in the proposed method to reduce energy consumption by analysing the behaviour of users in the smart home are as follows:

1. Data collection

Databases are used to collect the data for smart homes. These data include users' actions and behaviours and each sensor's position in the smart home. Evaluations and tests are performed on the provided databases. In this thesis, a smart home's energy consumption data set for three years has been utilized, i.e., the REFIT Smart Home dataset which is available online [49]. The total sample size was 2457. The number of points studied that represent the dataset features was 47 pieces of a house or features. These points hold a number that indicates the amount of energy consumption of that point from the beginning to the end of the day.

2. Pre-processing

Before designing and executing the proposed method, data is pre-processed so that samples with incomplete information are not propagated to the model design. After performing pre-processing, the data volume is decreased, and the desired data is provided by analysing user behaviour. As a result, reducing energy consumption could be provided with better speed and higher accuracy.

The most important reason for the pre-processing of data before applying the proposed algorithm, which is described below, is that the outlier data is eliminated, and the algorithm detection accuracy increases. Besides, the nature of the algorithms used in this research is generally in machine learning algorithms that incomplete examples could not be processed, or their precision would be reduced if processed. Therefore, by pre-processing, outliers would be removed, and only consistent data remains. In general, the data pre-processing process was presented as follows.

- Analysis of input data
- If the input data contain outliers, the following steps must be taken:
- Identify the record of the dataset that has outliers;
- Consider the item that has a large number of outliers;

- Monitor all of the same features of outliers from other records and dataset records that have more similarity with outliers, and average the feature values;
- Finally, replace the mean data as the final data for our record feature, which has the outliers.
- The final data is sent back to the next stage.

Therefore, in this process, data pre-processing is performed, and in the end, so-called standard and consistent data would be provided to the feature selection algorithm.

3. Selection of optimal features using PSO algorithm

After the relevant data is pre-processed and the standard was created without outliers to increase the accuracy, we reduce the data dimensions using the PSO algorithm. To this end, all pre-processed data is imported as input to the PSO algorithm. The algorithm also applies to the data and weighs the features. Finally, those features that have more weight and impact the data are selected as prominent features. This step's output is used in the SVM support vector to analyse user behaviour in the smart home. It is interpreted that the preferred input data is the choice of the PSO feature of the raw data as pre-processing. According to this data, the feature selection process is performed using the PSO algorithm, and the selected properties are entered as the output of this SVM. For the PSO algorithm's final output, some examples consist only of the features selected by this algorithm. The exact relationship and steps and the PSO feature selection algorithm are described in Section 4.3.1 later in this Chapter.

4. Classification of training and testing data

After selecting the PSO feature selection algorithm's optimal features to produce the relevant models and evaluate the proposed method, it is necessary to classify part of the final training and testing data. In this study, 70% of the data are classified as testing samples and 30% of the data as training samples for final evaluation.

5. Creating a model from SVM

Data collection is pre-processed, and finally, samples are tested and trained on them using the PSO algorithm and sent as input and backup vector algorithm. The SVM algorithm generates the relevant model based on the samples and finally evaluates the accuracy of identifying people to reduce energy consumption by analysing the behaviour of running users based on data and test samples. This step is carried out by Rapid Miner data mining tool.

6. Analysis

After developing a model to reduce energy consumption by analysing residents' behaviour in the smart home, we analyse the results using Excel and present them in various diagrams.

7. Testing and evaluating the results

Finally, after the proposed method has been designed and modelled, it should be tested and evaluated. Required assessments are carried out using the test, data mining, and other relevant software. In the next section, the different parts of the proposed method are described and examined in detail.

4.3 Description of the proposed method

In the previous section, pre-processing and classification of training and experimental data were described. In this section, the PSO and SVM procedures, which are among the most important machine learning algorithms, are examined in general.

4.3.1 PSO application

PSO is one of the most widely used feature selection algorithms employed to reduce the size of data and select prominent features from a data set. PSO is an optimization method developed by Kennedy and Ebrahart in 1995 [50]. The PSO simulates living organisms' social behaviour, such as flocks of birds, to describe a system that evolves automatically.

Each candidate solution is "a single bird from the flock", i.e., one particle in the search space; therefore, each particle of its memory and knowledge, in general, is used to obtain the best solution [51].

All particles have proportion values that are evaluated by a proportion function for optimization and have velocities that guide the particles' motion. During the motion, each particle adjusts its position according to its own experience, as well as according to the experience of an adjacent particle, resulting in useful data, which in most cases is very time-consuming. If the tuning is great, the motion of the particles will also be great, causing the algorithm to weaken soon, so a set of useful properties could not be obtained. Therefore, setting the appropriate parameter increases particle optimization to increase the performance of feature selection. For SVM, the correct parameter setting is critical since various parameters are involved. This could have a profound effect on the results. For different classes of problems, different parameters must be set for SVM. Two factors, kernel function parameter γ and penalty parameter C are especially important. Poor parameter settings harm classification accuracy [52]. The following is a more detailed description of the PSO algorithm.

4.3.1.1 Features selection based on PSO and optimize parameters

The RBF kernel function (defined by the following equation) is employed to implement our proposed method.

$$k(x_i, x_j) = \exp\left(-\frac{1}{\sigma^2}\|x_i - x_j\|^2\right)$$

$$k(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right)$$

To classify SVM since the RBF kernel function could analyse higher dimensional data, only two parameters γ and C need to be defined [53] [54]. When the RBF kernel is selected, the C and γ and the features used as input features should be optimized using the proposed PSO-SVM system [55]. Therefore, the particle consists of three parts: the mask features (discrete value), C (continuous value), and γ (continuous value) when the RBF kernel is selected.

In Table 5, particle i_e is shown with dimensions $n_F + 2$, where n_F is the number of features that differ from different datasets.

Table 5. Three particle features

Particle i is comprised of three parts: input feature mask, C, and γ

	Particle type						
	Input feature mask (discrete)					C (continuous)	γ(continuous)
Representation	$x_{i,1}$	$x_{i,2}$	...	$x_{i,d}$	...	x_{i,n_F}	$x_{i,n_F} + 1$	$x_{i,n_F} + 2$

One of the features is the Boolean mask, in which "1" indicates that this feature has been selected, and "0" indicates that this feature has not been selected.

4.3.1.2 Strategies for establishing the inertia criterion and the PSO learning process

In this study, the fixed, reductions, and random strategies used to establish the inertia criterion are shown below [35]:

1. **Fixed**: The inertia criterion is set to 1.0
2. **Reduction**: Linear reduction of inertia criterion

$$w = max - \frac{I_{current}}{I_{max}}(wmax - wmin)$$

$I_{current}$: the current iteration, I_{max} is the predefined maximum iteration, and $wmax$ and $wmin$ are minimum and maximum predefined Inertial weights.

3. **Stochastic**: Inertial weight is defined as follows:

$$w = 0.5 + \frac{rnd()}{2.0}$$

$rnd()$: a stochastic number in the range [0,1].

4.3.1.3 Feature selection steps using PSO

First, all the problem data after pre-processing and preparation is assigned to this algorithm. One of the most significant reasons for using the PSO algorithm is that the number of features is relatively high, and it is required to obtain the number of features that have the highest effect on reducing energy consumption by analysing the behaviour of residents in the smart home. To this end, we have used the PSO feature selection algorithm. The PSO feature selection algorithm selects all feature subsets according to its kernel and calculates each subset's corresponding accuracy.

This process is shown in Figure 19.

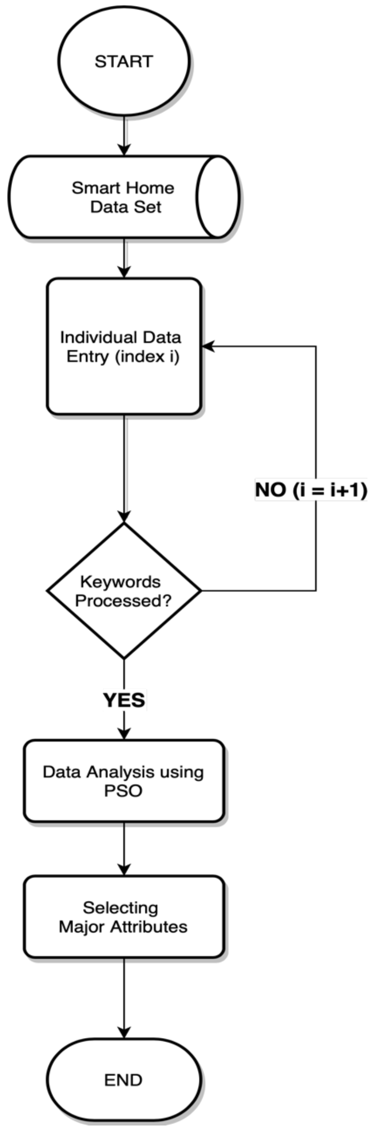

Figure 19. PSO algorithm flowchart

According to Figure 19, the PSO algorithm's input is the output of the pre-processing section. In other words, the hashing output, which is the hashed data, inputs to the PSO algorithm; the following PSO feature selection process is performed on the data, and the prominent data is selected. Finally, a subset of prominent features is selected with the most precision and influence on the output. Therefore, to select the features with the most prominent influence on determining and reducing energy consumption by analysing residents' behaviour in the smart home, the PSO has been employed, which has many applications in this field. The way such an algorithm works is if, for example, we have five features and five data, the algorithm selects a single set of features and determine the accuracy. Afterward, it selects two to five features and checks their accuracy, and finally selects the set with the highest accuracy by comparing the accuracy. This process continues until the best solution is obtained.

Features: (p1, p2, p3, p4, p5)

Obtaining the series of single features and their precise calculation:

$$(p_1)\ (p_2)(p_3)(p_4)(p_5)$$

Obtaining the series of dual features and their precise calculation:

$$(p_1, p_2)(p_1, p_3)(p_1, p_4)(p_1, p_5)(p_2, p_3), \ldots.$$

Obtaining the series of multiple features up until all features:

$$(p_1\ , p_2, p_3, p_4, p_5)$$

Finally, the PSO process compares each and selects the feature of the set with the highest accuracy. By calculating each step's accuracy and the fit of each iteration according to the classification accuracy, the selection process performs the superior feature selection. In this research, the PSO considers the features as particles and seeks to find better particles.

4.3.2 SVM application

As mentioned, the SVM algorithm is used to classify linear and nonlinear data. Briefly, SVM is an algorithm that converts the initial training data to a higher dimension using a nonlinear mapping. In this new dimension, we search for an optimal hyperplane, which classifies one class' instances from another class linearly. The main task in SVM is to select the best classifier hyperplane from multiple hyperplanes. This approach was employed to determine the largest marginal page, for instance, the largest margin for more accurate classification. SVM could be employed as a classification measure to reduce energy consumption by analysing user behaviour in various applications. Otherwise, the two-tier method may sometimes be employed as the class' algorithm, so that only the positive data set would be considered a class and the user behaviour that was analysed and identified as another class. Figure 20 is the flowchart of the SVM algorithm.

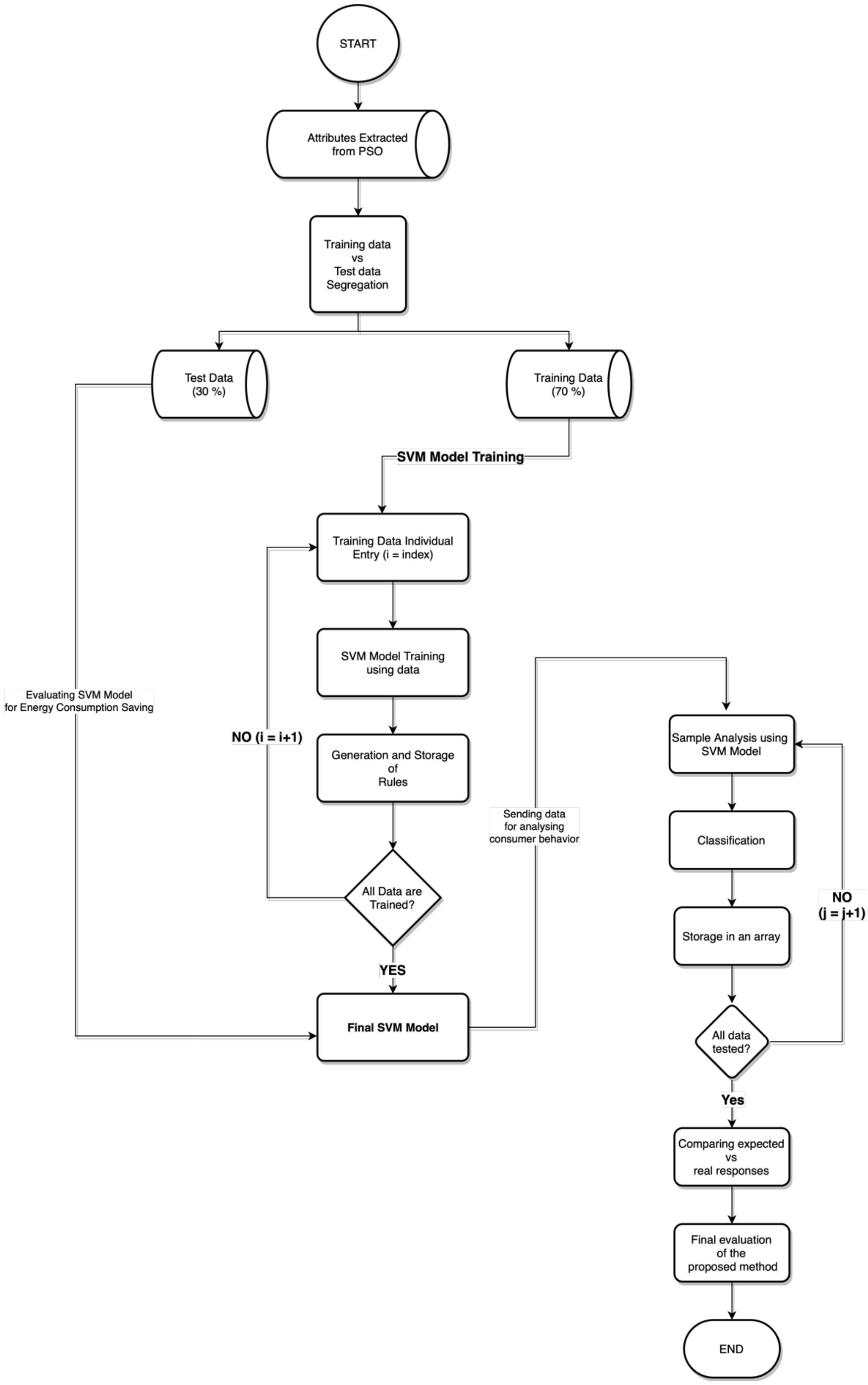

Figure 20. SVM functional flowchart

According to Figure 20**Error! Reference source not found.**, the PSO algorithm's output data is first registered as input to the SVM algorithm. These data are the prominent data selected from the principal features. In order to run the SVM algorithm, the data must be classified into two classes. Training data make up 70%, and test data make up 30% of the total samples. The reason for picking this ratio is because the proposed method is going to be compared with another approach [56] which incorporates only one-year data for the training, thus it sounds logical for us to pick 30% of the three-year period for our own training data set. Training data is utilized for training the SVM algorithm, and test data is utilized to test and estimate the SVM's performance. After the data is classified into two classes, we first use the SVM algorithm's training data. This algorithm processes and trains data one by one. Finally, a model of this algorithm is created. Test data is registered into the created model to estimate efficiency. This process continues until all test samples are examined and tested. Finally, each solution provided by the SVM is stored in an array. After all the samples have been tested, the predictions are compared with the original samples, and the accuracy, precision, recall, and error are calculated. In general, to reduce energy consumption by analysing residents' behaviour in the smart home, the library kernel or Lib kernel related to the SVM has been employed.

4.3.3 NN application

The proposed method uses a combination of NN and SVM algorithms to predict and improve energy consumption. Figure 21 shows the overall procedure:

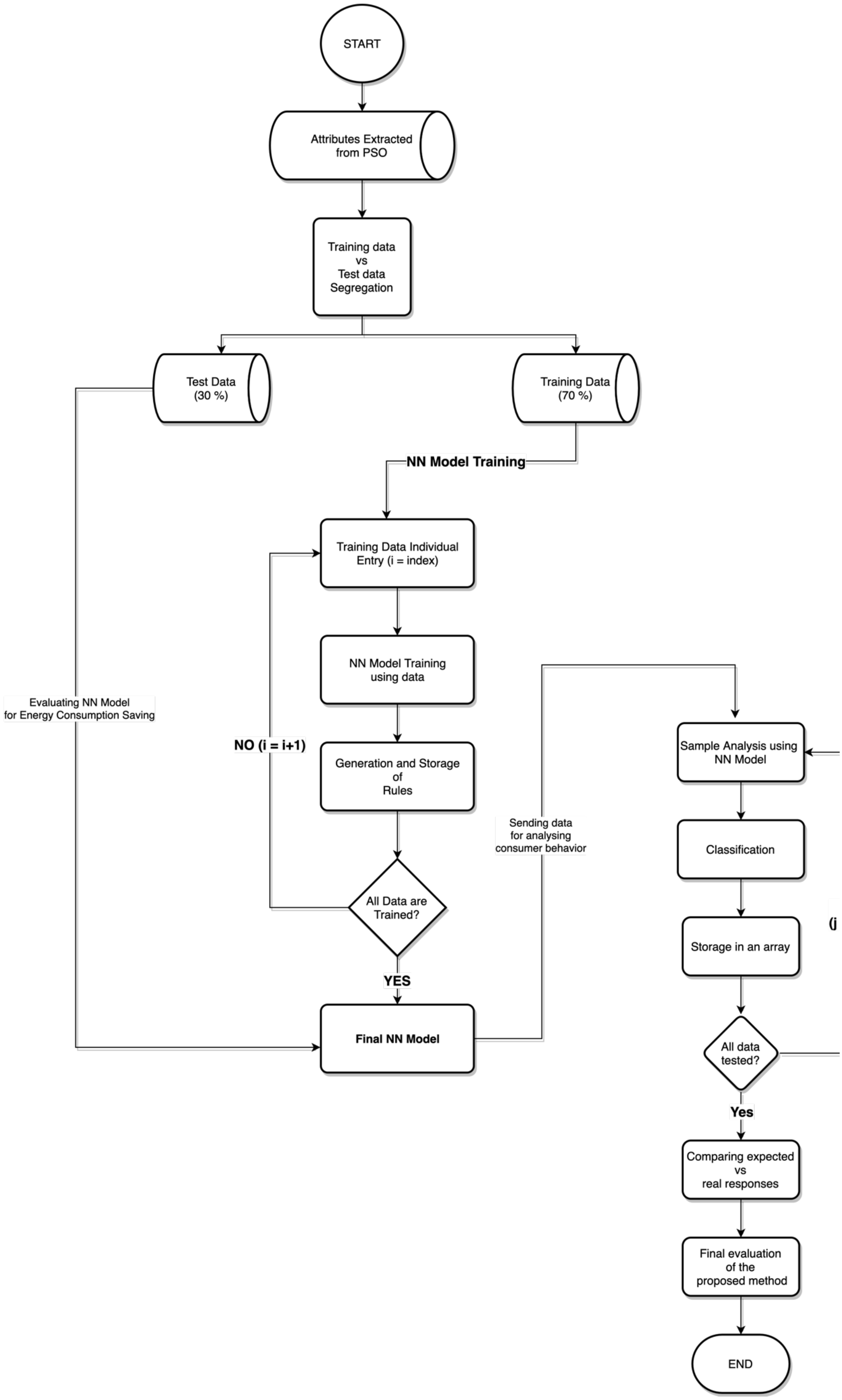

Figure 21. NN flowchart

As it is shown in Figure 21, data output from PSO algorithm is first fed into NN. These data are the major selected feature selections. Training data make up 70%, and test data make up 30% of the total samples. The reason for picking this ratio is because the proposed method is going to be compared with another approach [56] which incorporates only one-year data for the training, thus it sounds logical for us to pick 30% of the three-year period for our own training data set. After the data is classified into two classes, we first use the NN algorithm's training data. This algorithm processes and trains data one by one. Finally, a model of this algorithm is created. Test data is registered into the created model to estimate efficiency. This process continues until all test samples are examined and tested. Finally, each solution provided by the NN is stored in an array. After all the samples have been tested, the predictions are compared with the original samples, and the accuracy, precision, recall, and error are calculated.

In general, after the residents' behaviour is analysed and trained based on the sensors' position in the smart home using PSO, SVM and NN, the results of both NN and SVM algorithms were combined and ultimately residents' behaviour is predicted. Accordingly, energy consumption could be improved by adapting the energy management according to the residents' behaviour.

The process of combining the results is based on maximum decisions. The process of checking the maximum decision is such that both the NN and SVM solutions are combined. If the response of both algorithms is positive, the final response is positive, and if the response of both is negative, the final response is also negative. It is worth noting that if the response of one of the algorithms was positive and the other was negative, the final response of the algorithm which has higher accuracy is transferred to the output.

4.4 Evaluation criteria in the proposed method

In classification models, three common performance criteria are applied: precision, recall, and accuracy. The precision measures true positive to the true positive and false negative ratio. Recall measures true negative ratio to true negative and false positives, and accuracy measures the accuracy of the model's overall classification performance as the correct classification ratio (TP+TN) to all classifications, either true or false (TP, TN, FP, FN).

$$Precision = Sensitivity = \frac{TP}{TP + FP}$$

$$ReCall = Specifity = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Confusion Matrix legend:

$$\begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix}$$

TP[1]: Items that are true and have been correctly identified by the model.

FP[2]: Items that are false and misidentified by the model.

FN[3]: Items that are true but incorrectly identified by the model.

TN[4]: Items that are false and correctly identified by the model.

---

[1] True Positive
[2] False Positive
[3] False Negative
[4] True Negative

5 Chapter 5 – Simulation and Evaluation of Results

5.1 Introduction

For running simulations, RapidMiner was used in this research [57]. In this chapter, the relevant data source is introduced, and its related characteristics are described. After simulating the proposed method, all the results and findings are described and presented in various diagrams. Finally, the obtained results are evaluated and compared with other methods that have been proposed. Indeed, in this chapter, we compare our research results with that of [56]. In that article, a statistical technique was used to manage smart homes' energy resources. One of the most important criteria considered in their work was the amount of predictive error in managing energy resources in the smart home. Therefore, in our study, we compare the ultimate error criteria of our proposed combined machine learning methods (NN and SVM), with the article's prediction error.

5.2 Simulation Environment

RapidMiner is a data mining software produced by RapidMiner company [57]. This software is used in machine learning, data mining, text mining, forecast analysis, and business analysis. This software is used in both industrial and academic environments. It provides 99% of advanced no-code analytics solutions for data mining [58]. Figure 22 shows the environment of this software.

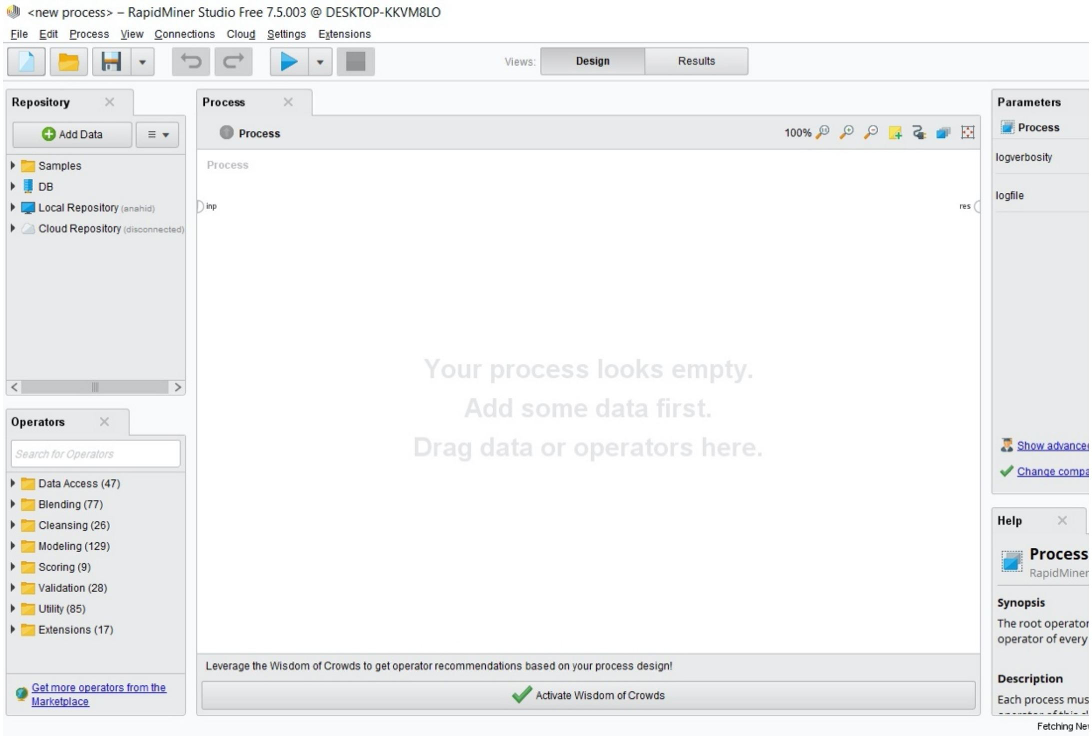

Figure 22. RapidMiner Version 7 GUI.

The environment shown in Figure 22 is a snapshot of RapidMiner V.7. This software is compiled in Java programming language. A GUI is used to perform all operations in RapidMiner. Workflow operation in RapidMiner is known as a process, which consists of several operators. Each operator performs a task in the process. This software can connect to software and data mining tools such as R and Weka [57]. Thus, according to this software's capabilities, it is employed to simulate and obtain different results in this research.

5.3 Simulation and Experimental Results

The proposed method was performed using the RapidMiner simulator. Table 6 shows the system's specifications in which the proposed method was implemented, and the results were also evaluated.

Table 6. System specifications for simulation and evaluation of results.

Hardware/Software	**Specification**
Operating system	Win7
Type of OS	32 bits
Ram	4Gb - 3.06 Gb accessible
Processor	Intel, Q720- Core i7- 1.60 GHz, 1.60 GHz

5.4 Data Set

In this study, a total of three years of energy consumption data from a smart home has been used. These points have a value that indicates the amount of energy consumed from the beginning of the day to the end of the day [49]. Table 7 contains several examples of data with different sensors in the house.

Table 7. A clipping of the dataset used in the research (all values are in Joules).

No.	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7	Sensor 8	Sensor 9	Sensor10	Sensor n	Load
1	176	422	204	599	961	16	172	394	263	599	…	57631
2	195	430	203	593	975	84	183	399	260	599	…	57998
3	274	434	203	591	966	115	180	404	257	598	…	58819
4	361	439	189	589	949	81	175	409	252	601	…	59628
5	439	442	175	593	940	76	216	412	244	602	…	60122
6	550	459	164	583	927	75	268	419	232	603	…	60449
7	670	469	146	580	915	104	395	427	219	606	…	60656
8	754	476	127	577	913	76	462	433	205	611	…	60664
9	808	484	117	574	917	83	588	442	189	613	…	60978
10	797	484	96	576	925	68	630	453	175	616	…	60819
11	825	485	93	578	948	108	646	463	161	613	…	60762
12	857	484	84	580	958	84	660	476	152	615	…	60865
13	905	490	73	577	937	54	744	482	138	615	…	60979
14	812	490	58	581	949	17	819	488	123	619	…	60786
15	696	493	53	577	1001	45	861	492	110	618	…	60193

16	579	499	43	575	1077	36	802	500	101	619	…	60044
17	504	493	43	579	1056	22	603	501	100	616	…	58945
18	320	488	30	584	989	17	383	503	93	618	…	57550
19	215	519	12	553	995	263	69	502	37	619	…	60875
20	251	521	26	553	1100	265	57	506	24	620	…	60965
21	258	526	16	551	1163	250	57	510	19	615	…	61241
22	193	528	22	546	1097	121	168	510	16	612	…	61577
23	139	526	28	547	980	9	143	505	9	614	…	60504
24	139	531	21	538	845	13	231	505	6	616	…	59691
25	159	532	42	539	787	84	272	506	1	615	…	60119
26	190	539	49	531	756	77	302	511	12	614	…	60927
27	121	537	56	533	775	17	205	512	16	612	…	61091
28	71	541	74	528	788	21	80	511	23	609	…	60298
29	21	544	78	523	800	175	15	510	29	611	…	60053
30	110	544	74	528	852	286	5	505	29	617	…	60118

This research uses 47 sensors in total but for legibility reasons, in the table above only data for 10 sensors over a 30-hours operation is shown. The values show the consumption reported by each sensor, which is a smart home light bulb. The unit of each sensor is in Joule. For example, Sensor 1 reports a consumption of 176 J of energy in the first cell. This also considers the behaviour of users, for example Sensor 1, which is located in the hallway of the house has a relatively balanced consumption compared to other sensors. This indicates that family members movement in this place is less than compared to other places. Thus, based on data mining knowledge and the use of machine learning algorithms such as NN and SVM, it is possible to predict and control the smart home's energy consumption based on user performance.

According to this data, each sample represents the smart home's amount of energy consumption on the same day. With the help of these examples, we can predict the energy consumption during upcoming days and with the help of it, we can improve energy consumption. In the following, the results obtained from the proposed method are described.

5.5 A Brief Description of Applying the Proposed Method

In this research, we seek to predict energy consumption in smart homes and attempted to improve it using the combination of the PSO and machine learning methods such as NN and SVM. In the first step, we pre-process the desired dataset, which we have described in the previous section and eliminate the outliers. Pre-processing steps are explained in Chapter 4. The original dataset, which consists of 47 records, was entered into the PSO algorithm to reduce the features and obtain those that significantly impact energy consumption. The final results are shown in Table 8, after applying data with 46 features to the PSO algorithm.

Table 8. PSO results (all values are in joules).

No.	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7	Sensor 8	Sensor 9	Sensor1	Sensor n	Load
1	176	422	204	599	961	16	172	394	263	599	…	57631
2	195	430	203	593	975	84	183	399	260	599	…	57998
3	274	434	203	591	966	115	180	404	257	598	…	58819
4	361	439	189	589	949	81	175	409	252	601	…	59628
5	439	442	175	593	940	76	216	412	244	602	…	60122
6	550	459	164	583	927	75	268	419	232	603	…	60449
7	670	469	146	580	915	104	395	427	219	606	…	60656
8	754	476	127	577	913	76	462	433	205	611	…	60664
9	808	484	117	574	917	83	588	442	189	613	…	60978
10	797	484	96	576	925	68	630	453	175	616	…	60819
11	825	485	93	578	948	108	646	463	161	613	…	60762
12	857	484	84	580	958	84	660	476	152	615	…	60865
13	905	490	73	577	937	54	744	482	138	615	…	60979
14	812	490	58	581	949	17	819	488	123	619	…	60786
15	696	493	53	577	1001	45	861	492	110	618	…	60193
16	579	499	43	575	1077	36	802	500	101	619	…	60044

Among 46 sensors, only 10 sensors with higher workload per day have been selected as the PSO algorithm's prominent features. After the PSO algorithm extracted the data, these samples should be classified into test and training data subsets. We utilize the training data subset to train NN and SVM algorithms and we use the test data subset to evaluate the algorithms' performance.

The proposed algorithm's final output is represented in Table 9 after applying them to the data mined by the PSO algorithm. As can be noticed, the effect of the PSO algorithm is to extract sensors with useful data from 46 sensors. The important column here is the 'Predicted Load' column since it is the outcome of the method.

Table 9. Some SVM-NN results (all values are in Joules).

No.	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7	Sensor 8	Sensor 9	Predicted Load	Load
1	176	422	204	599	961	16	172	394	263	57600	57631
2	195	430	203	593	975	84	183	399	260	57900	57998
3	274	434	203	591	966	115	180	404	257	58810	58819
4	361	439	189	589	949	81	175	409	252	59628	59628
5	439	442	175	593	940	76	216	412	244	60120	60122
6	550	459	164	583	927	75	268	419	232	60451	60449
7	670	469	146	580	915	104	395	427	219	60600	60656
8	754	476	127	577	913	76	462	433	205	60660	60664
9	808	484	117	574	917	83	588	442	189	60970	60978
10	797	484	96	576	925	68	630	453	175	60812	60819
11	825	485	93	578	948	108	646	463	161	60762	60762
12	857	484	84	580	958	84	660	476	152	60860	60865
13	905	490	73	577	937	54	744	482	138	60979	60979
14	812	490	58	581	949	17	819	488	123	60780	60786
15	696	493	53	577	1001	45	861	492	110	60191	60193
16	579	499	43	575	1077	36	802	500	101	60042	60044

According to the above results, the predicted consumption load based on the two neural network methods' performance and the SVM is close to the actual value, and these results are adequate.

5.6 Simulation Results of the Proposed Method

In this section, we first attempt to classify the data using the X-Means clustering algorithm. This creates an overall vision of high- and low-consumption days. Considering that the X-Means algorithm has the optimal clustering capability of k, this algorithm is used in this section. As described in Chapter 3, the algorithm does not need to calculate the number k representing the number of clusters and calculates the number of the optimal k by itself.

In the first step of the proposed method, data is presented to the X-Means algorithm and is clustered. The classification procedure is represented in Figure 23**Error! Reference source not found.** to include daily consumption in the corresponding classes.

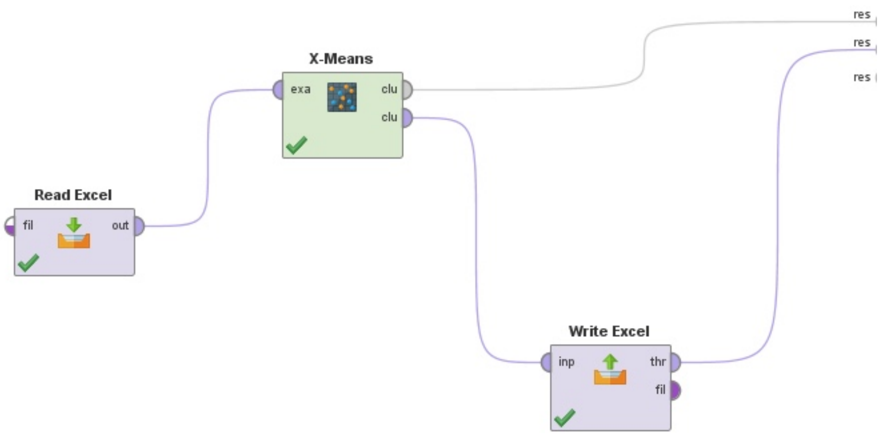

Figure 23. Illustration of the X-Means Clustering Model created in RapidMiner.

According to Figure 23, the data is entered into the X-Means by Read Excel and the output is stored using Write Excel. After the execution of this algorithm, 4 clusters are determined as clusters covering all samples. The number of samples in each cluster is shown in Figure 24.

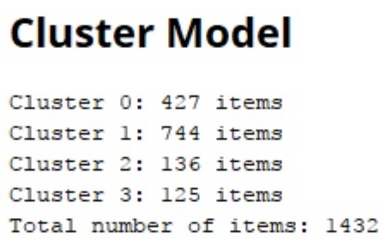

Figure 24. Samples in each cluster after applying the X-Means algorithm.

As shown above, there are 427, 744, 136, and 125 samples in cluster 0, 1, 2, and 3, respectively. Figure 25 represents the number of items within each cluster.

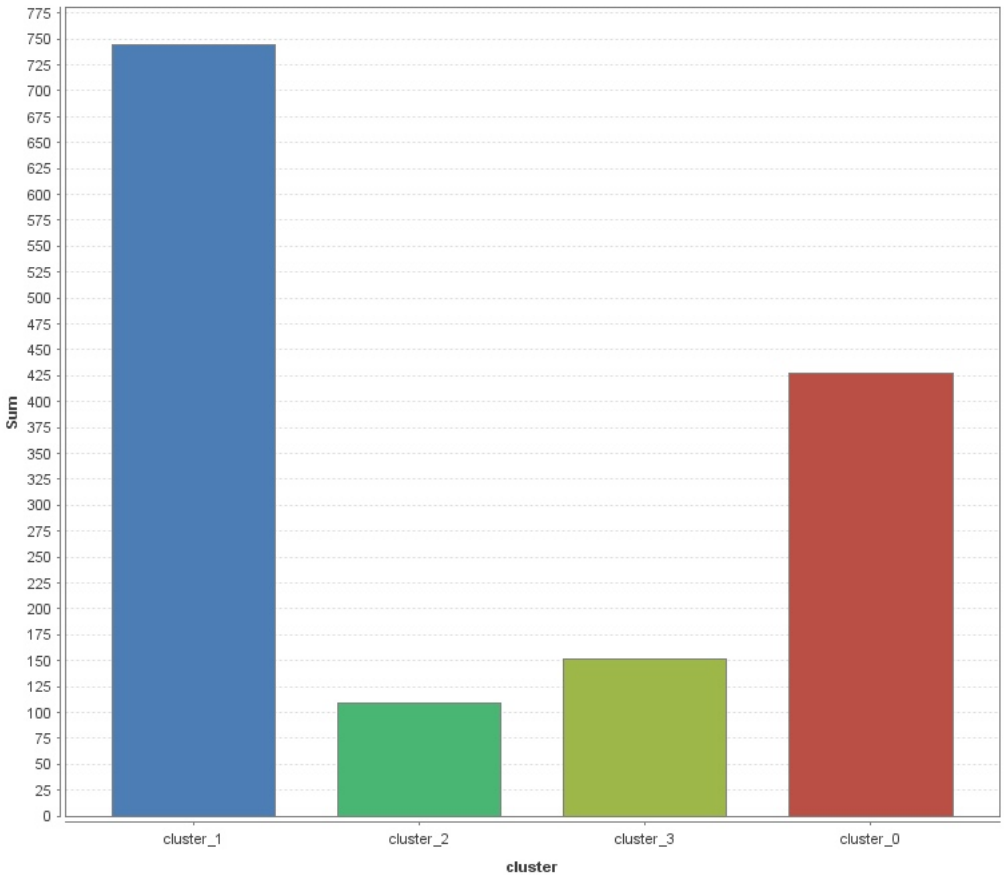

Figure 25. Number of items (denoted as sum) within each cluster of total energy consumption for each cluster.

By analysing the corresponding amount of energy consumption in each cluster, the total energy consumption for three years at the smart home was obtained as 219009959 W (≈ 219 MW). The analysis is shown in Figure 26 in more details.

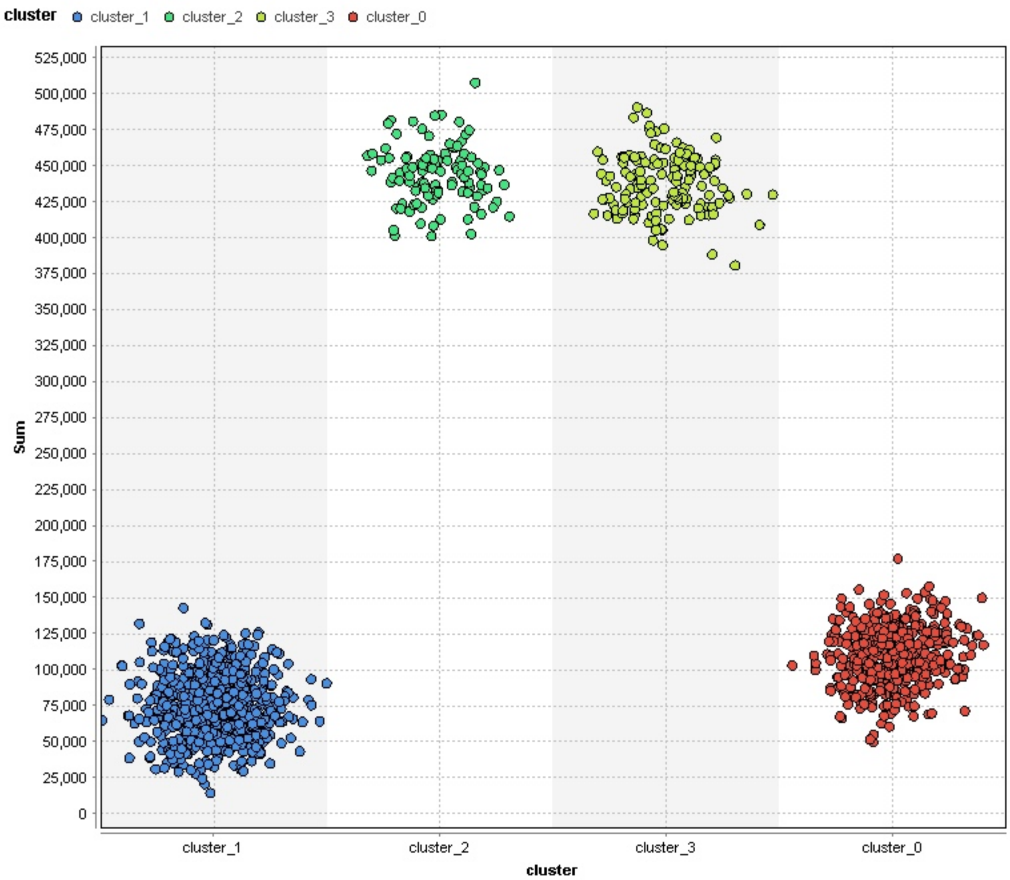

Figure 26. 3-year period clustering results.

The total energy consumption is reflected on the vertical axis, while each cluster is shown on the horizontal axis. Cluster 1, shown in blue, has the most extensive lighting density in the smart home but has less energy consumption than the others. Cluster 2, shown in green, has the lowest lighting density and maximum energy consumption. Thus, with this simple analysis, we found that we can focus on the residents' behaviour at these points and provide solutions that lead to less energy consumption. Furthermore, NN and SVM algorithms have been applied to these clustering results to predict the optimum daily consumption ratio for high-consumption points.

After the implementation of the proposed method, it has been observed that by providing an efficient scheduling during residents' sleep time or while they are not in the house, if the lights are put into standby mode and get activated instantly as soon as the user is present, there can be a considerable saving in energy consumption. Therefore, in the next sections, the accuracy, error, recall, and precision for predicting residents' behaviour using the proposed method is addressed.

Note that Appendix 2 shows the simulation results generated by RapidMiner; the next subsections contains graphical representations thereof.

5.7 Comparison of Predictive Accuracy in the Proposed Combined Method and Individual Methods

The accuracy criterion is one of the most important criteria that are considered in many issues and issues related to prediction. This criterion could measure the accuracy and precision of the proposed method. When a method's accuracy is large and close to 100%, the method could be trusted and used in the desired problem. Therefore, in this study, the accuracy criterion is used to measure the proposed method's accuracy for predicting the residents' behaviour. In Figure 27, the accuracy of the proposed SVM-NN method is compared to other methods (NN and SVM separately).

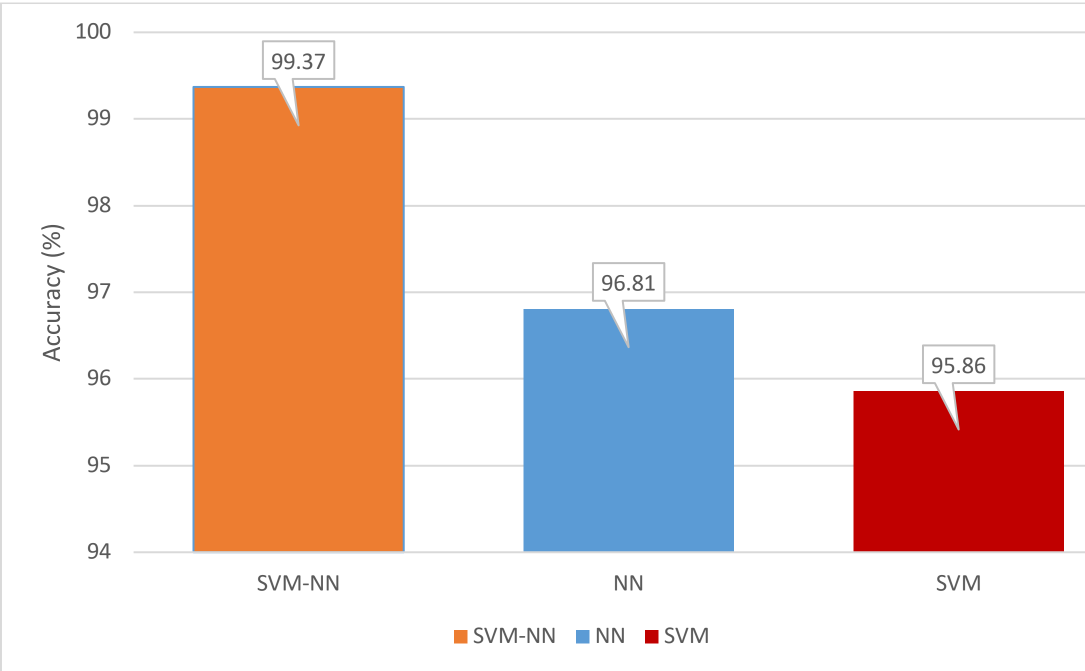

Figure 27. Comparison of Predictive Accuracy Using SVM, NN, SVM-NN.

As shown in Figure 27, the accuracy of predicting the optimum energy consumption using SVM, NN, and SVM-NN algorithms are 95.86%, 96.81%, and 99.73%, respectively. The proposed SVM-NN's precision is 3.51 and 2.56 percentage points higher than that of SVM and NN, respectively. In the next section, the predictive precision of the proposed method is compared to other methods.

5.8 Comparison of Prediction Precision in the Proposed Method and Other Individual Methods

Precision measures the degree to which a model or algorithm and method for solving a problem could be correct. This criterion evaluates the precision of predicting the desired energy consumption in the smart home. Using this criterion, it could be found to what extent the proposed method is useful in predicting the smart home's desired energy consumption. Accordingly, in this study, the precision criterion is used to evaluate the proposed method. In Figure 28, the precision of the proposed SVM-NN method is compared to other methods (NN and SVM separately).

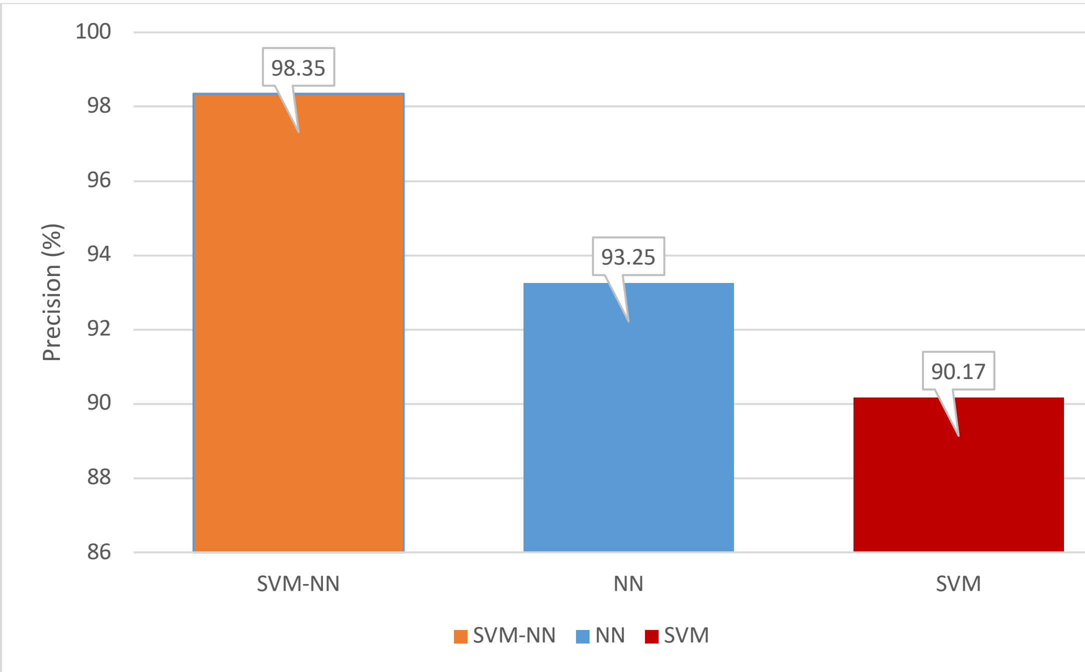

Figure 28. Comparison of Predictive Precision Using SVM, NN, SVM-NN.

According to Figure 28, the precision of predicting the desired energy consumption using the SVM, NN, and SVM-NN algorithms are 90.17%, 93.25%, and 98.35%, respectively. SVM-NN's predictive precision is about 8.18 and 5.1 percentage points higher than SVM and NN, respectively.

The next section compares and evaluates the recall rate of the proposed method compared to other methods.

5.9 Comparison of Prediction Recall Rate in The Proposed Method and Other Methods

The recall is also one of the most important criteria used to calculate the ability of a method and model in classification and prediction problems. Using this criterion, it can be concluded to what extent the prediction was correct and on the other hand, it is directly related to speed. Accordingly, in this section, the recall rate of the proposed method is compared with other methods. In Figure 29, the recall rate of the proposed SVM-NN method is compared to other methods (NN and SVM separately).

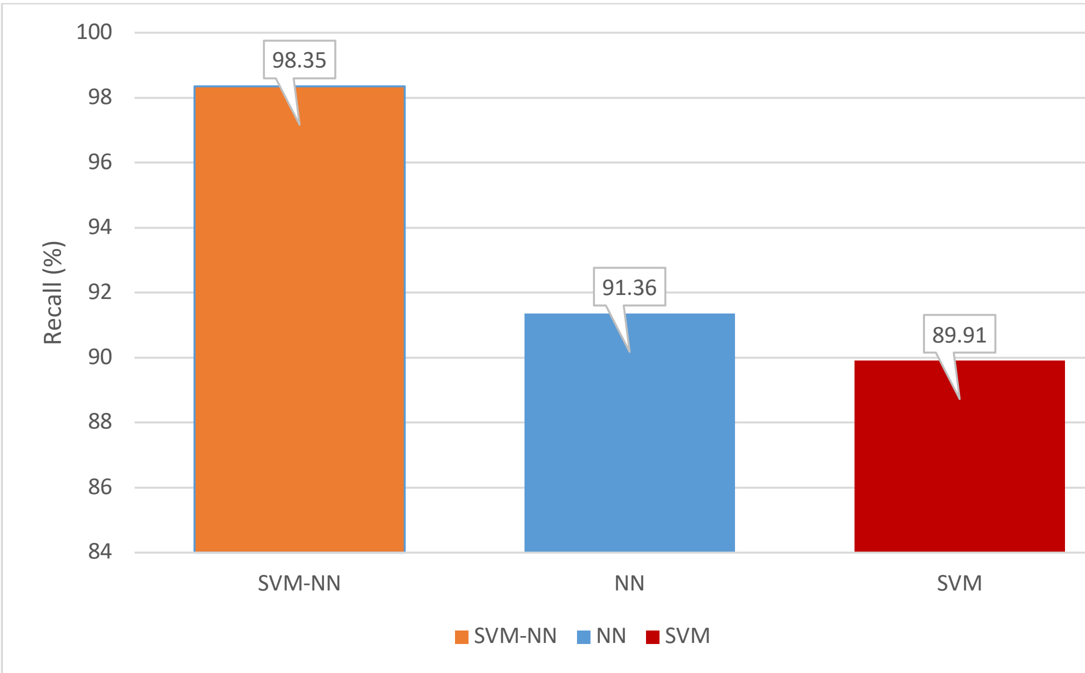

Figure 29. Comparison of Prediction Recall Rate Using SVM, NN, SVM-NN.

According to Figure 29, the recall rate of prediction of desired energy consumption using the SVM, NN, and SVM-NN algorithm was 89.91%, 91.36%, and 98.35%, respectively. SVM-NN's prediction recall is 8.44 and 6.99 percentage points higher than SVM and NN, respectively.

In the next section, we evaluate the error rate of the proposed method compared to other methods.

5.10 Comparison of Forecasting Errors in the Proposed Method and Other Individual Methods.

Error criterion is one of the most important criteria that is considered in issues related to prediction. Using this criterion, one can measure the accuracy and precision of the proposed method. The more the error rate is close to zero, the method could be trusted and used in the desired problem. Thus, the error criterion is used to measure the proposed method's accuracy for predicting the optimum energy consumption (see Figure 30).

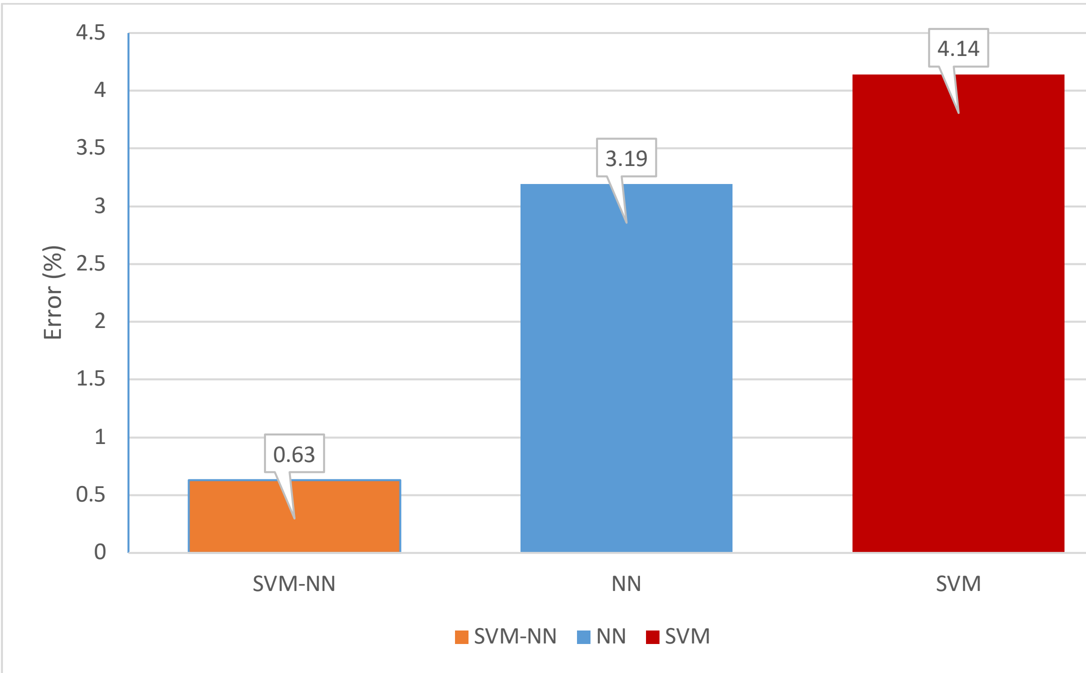

Figure 30. Comparison of Forecasting Errors Using SVM, NN, SVM-NN.

As shown in Figure 30, the desired energy consumption prediction error rate at the smart home using the SVM, NN, and SVM-NN algorithm was 4.14%, 3.19%, and 0.63%, respectively. SVM-NN's forecasting error is 3.51 and 2.56 percentage points lower as compared to the SVM and NN algorithms, respectively.

5.11 Comparison of The Proposed Method with Base Paper

One of the most critical parameters that most articles in the scientific literature use to calculate their proposed method's accuracy is the predictive error criterion. In this study, we compared our proposed method's results to reduce energy consumption in the smart home with the article [56] presented in 2016.

During this article, Xiaohua Wu et al. proposed a Stochastic Dynamic Programming (SDP) framework for the optimal energy management of a smart home with Plug-in Electric Vehicle (PEV) energy storage and Photovoltaic (PV) array. The article concludes that the smart home with PEV energy storage and PV array under optimal control can bring significant cost savings for customers [56]. In this paper, one of the essential criteria calculated is the mean of errors.

Figure 31 compares the error rate in predicting energy consumption in smart homes based on the analysis of residents' behaviour in the proposed method compared to other methods.

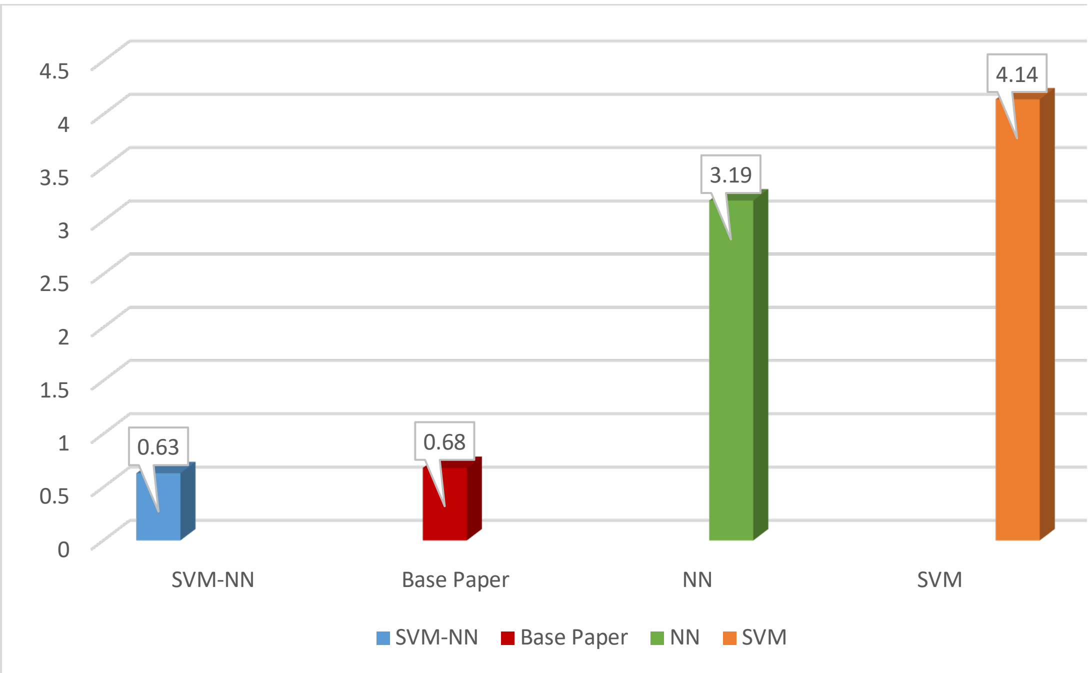

Figure 31. Comparison of Prediction Using SVM, NN, SVM-NN and base paper.

According to Figure 31, the prediction error of energy consumption in the smart home is 0.63% with the proposed SVM-NN method. This is 3.51 and 2.56 percentage points lower as compared to SVM and NN, respectively. Furthermore, this is 0.05 percentage point lower as compared to the reference work [56].

5.12 Summary of the results

This section summarizes the comparison of different criterion of the proposed method with other methods for predicting the optimum energy consumption at a smart home. The criterions that are compared are accuracy, precision, recall, error. After simulating the proposed method and base method, results were obtained. These results are compiled in Table 10.

Table 10. Summary of the simulation results comparison.

Method	**Error [%]**	**Recall [%]**	**Precision [%]**	**Accuracy [%]**
SVM	4.14	89.91	90.17	95.86
NN	3.19	91.36	93.25	96.81
Proposed SVM-NN	0.63	98.35	98.35	99.37

As observed, accuracy, precision, and recall in the proposed SVM-NN method are better than other methods, especially the error which is significantly lower than for SVM and NN separately. We could rely on the proposed method and practice it in predicting the optimum energy consumption at a smart home.

6 Chapter 6 - Conclusion and Future Suggestion

6.1 Summary

This thesis' primary purpose was to provide a solution to reduce energy consumption in smart homes using machine learning methods. After presenting the background of supervised and un-supervised learning approaches, this thesis proposed a combined SVM-NN method (that includes a preliminary clustering step based on PSO). This proposed method has then been simulated using a real-life dataset, and the results have been compared to the individual SVM and NN methods. The following main observations have been made:

The accuracy of predicting the smart home's optimum energy consumption using the proposed SVM-NN method is 99.73%, which is 3.51 and 2.56 percentage points higher than that of SVM and NN, respectively.

The precision of predicting the smart home's optimum energy consumption using the proposed SVM-NN method is 98.35%, which is 8.18 and 5.1 percentage points higher than SVM and NN, respectively.

The recall rate of the prediction of optimum energy consumption at smart homes using the proposed SVM-NN method is 98.35%, which is 8.44 and 6.99 percentage points higher than SVM and NN, respectively.

The error rate of the prediction of optimum energy consumption at the smart home using the proposed SVM-NN method is 0.68, which is 3.51 and 2.56 percentage point lower than SVM and NN, respectively.

Therefore, it can be concluded that by employing clustering (PSO) first, we could analyse the residents' behaviour in the smart home and identify the high-consumption points in the day, and that by applying combined SVM-NN next, the optimum power peaks at peak times can be identified to reduce energy consumption.

Thus, according to the results, the proposed method could be trustworthy and used in many applications. The resulting predictions would allow us to create better management plans to reduce energy consumption.

6.2 Suggestions for Future Work

Here are some suggestions to improve the performance of the proposed method in this thesis:

- The application of other data mining algorithms including decision tree, Naive Bayes classifier and regression algorithm instead of NN and SVM can be studied further to see if they promise any enhancement.
- Using a mixture of bagging and boosting methods instead of neural network algorithms and SVM algorithms and comparing the results with this study's results.
- Using feature selection algorithms such as genetic algorithm, bee colony, and ant colony to select prominent features and compare the results obtained with the results of this study.
- Using optimization algorithms such as DragonFly algorithm, Cat Swarm optimization, and Whale optimization algorithm etc. to select optimum consumption in the proposed method in smart homes at the consumption peaks and compare the results obtained with this study's results.

References

[1] S. Habib, M. Kamran and U. Rashid, “Impact analysis of vehicle-to-grid technology and charging strategies of electric vehicles on distribution networks–a review,” *Journal of Power Sources,* vol. 277, pp. 205-214, 2015.

[2] Y. Y. Ghadi, M. Rasul and M. M. Khan, “Design and development of advanced fuzzy logic controllers in smart buildings for institutional buildings in subtropical Queensland,” *Renewable and Sustainable Energy Reviews,* vol. 54, pp. 738-744, 2016.

[3] B. Ai, Z. Fan and R. X. Gao, “Occupancy estimation for smart buildings by an auto-regressive hidden Markov model,” in *American Control Conference*, Portland, 2014.

[4] E. Romero and D. Toppo, “Comparing Support Vector Machines and Feed-forward Neural Networks with Similar Parameters,” in *Intelligent Data Engineering and Automated Learning -- IDEAL 2006*, Berlin, Heidelberg, Springer Berlin Heidelberg, 2006, pp. 90-98.

[5] PureAI, “PURE AI,” 15 April 2020. [Online]. Available: https://pureai.com/articles/2020/04/10/ml-techniques.aspx. [Accessed 3 November 2020].

[6] S. Sripriya, V. Divya and T. Babu, “Building management system in metro rail,” in *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)*, Chennai, 2017.

[7] “Building Management System (BMS),” Leater Professional Solutions, 2020. [Online]. Available: https://leater.com/en/services/building-management-system-bms.html. [Accessed 12 December 2020].

[8] B. Mataloto, J. C. Ferreira and N. Cruz, “LoBEMS—IoT for Building and Energy Management Systems,” *Electronics,* vol. 8, no. 7, p. 763, 2019.

[9] M. G. J. W. M. Ann Piette, “Intelligent building energy information and control systems for low energy operations and optimal demand response,” *IEEE Design and test of computers,* vol. 29, no. 4, 2012.

[10] M. Rousselot, *Energy Efficiency Trends in Buildings in Europe,* ODYSSEE, https://www.odyssee-mure.eu, 2018.

[11] U.S. Energy Information Administration, “2015 Residential Energy Consumption Survey,” [Online]. Available: https://www.eia.gov/energyexplained/use-of-energy/electricity-use-in-homes.php.

[12] J. King and C. Perry, “Smart Buildings: Using Smart Technology to Save Energy in Existing Buildings,” American Council for an Energy-Efficient Economy (ACEEE), Washington DC, 2017.

[13] L. Babun et al., “Z-IoT: Passive Device-class Fingerprinting of ZigBee and Z-Wave IoT Devices,” in *IEEE International Conference on Communications (ICC)*, Dublin, 2020.

[14] W. Kastner, G. Neugschwandtner, S. Soucek and H. Newmann, “Communication systems for building automation and control,” *Proceedings of the IEEE,* vol. 93, no. 6, pp. 1178-1203, June 2005.

[15] Zissis G., “Energy Consumption and Environmental and Economic Impact of Lighting: The Current Situation,” in *Handbook of Advanced Lighting Technology*, Springer, Cham, 2016.

[16] Control4, “Your Advanced Home,” 2006. [Online]. Available: http://www.youradvancedhome.com/Control4-WirelessLCDKeypad.pdf. [Accessed 10 12 2020].

[17] Innocenti, Laura,; Hertz, Ellen; Dubuis, Claudia, *The smart home as a place of control and security: an analysis of domestication of smart technologies for the making of a home,* Zürich: Institut d’ethnologie Faculté des lettres et des sciences humaines, 2017.

[18] Navigant Research, “Advanced Energy Now 2019 Market Report,” Advanced Energy Economy, Washington DC, 2019.

[19] Hajighorbani, M., & Hashemi, S. M.-B, “A Review of Some Semi-Supervised Learning Methods,” *International Journal of Mechatronics, Electrical and Computer Technology (IJMEC),* vol. 2, no. 4, pp. 250-259, 2016.

[20] Romero, Cristobal & Ventura, Sebastian, “Educational data mining: A survey from 1995 to 2005,” *Expert Systems with Application,* vol. 33, no. 1, pp. 135-146, 2007.

[21] K. P. Murphy, in *Machine Learning: A Probabilistic Perspective*, London, The MIT Press, 2012, pp. 271-273.

[22] S. Haykin, Neural Networks: A Comprehensive Foundation, Englewood Cliffs NJ: Prentice Hall, 2007.

[23] P. Benardos and G. Vosniakos, “Optimizing feedforward artificial neural network architecture,” *Engineering Applications of Artificial Intelligence,* vol. 20, no. 3, pp. 365-382, April 2007.

[24] G. P. Zhang, “Neural networks for classification: a survey,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 30, no. 4, pp. 451-462, 2000.

[25] G. Ciaburro and B. Venkateswaran, Neural networks with R: smart models using CNN, RNN, deep learning, and artificial intelligence principles, Birmingham, UK: Packt Publishing, 2017.

[26] A. Mahanipour and H. Nezamabadi-pour, “Improved PSO-based feature construction algorithm using Feature Selection Methods,” in *2nd Conference on Swarm Intelligence and Evolutionary Computation (CSIEC)*, Kerman, 2017.

[27] J. R. Quinlan, C4.5: Programs for Machine Learning, San Francisco: Morgan Kaufmann Publishers, 1993.

[28] Heckerman, David, and John S. Breese, “Causal independence for probability assessment and inference using Bayesian networks,” *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans,* vol. 26, no. 6, pp. 826-831, 1996.

[29] F. Yang, “An Extended Idea about Decision Trees,” in *International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, 2019.

[30] Sanjeev and Seema Sharma, J. Agrawal, “Classification Through Machine Learning Technique: C4. 5 Algorithm based on Various Entropies,” *International Journal of Computer Applications,* vol. 82, pp. 28-32, 2013.

[31] e. a. R. O. Duda, Pattern Classification, 2nd Edition, John Wiley & Sons Inc., 2000.

[32] C. M. Bishop and N. M. Nasrabadi, Pattern recognition and machine learning, New York: Springer, 2006.

[33] M. Horný, “Bayesian Networks,” Boston University, School of Public Health, Department of Health Policy and Managemenet, Boston, 2014.

[34] R. M. Sneider et al., “Using the ADAP Learning Algorithm to Forcast the Onset of Diabetes Mellitus,” *Proceedings - Annual Symposium on Computer Applications in Medical Care,* vol. 10, 1988.

[35] A. A. AlJarullah, “Decision Tree Discovery for the Diagnosis of Type II Diabetes,” in *International Conference on Innovations in Information Technology*, 2011.

[36] “Using Neural Networks To Predict the Onset of Diabetes Mellitus,” *Journal of Chemical Information and Computer Sciences,* vol. 36, no. 1, pp. 35-41, January 1996.

[37] Lin, Wanling Liu and Weikun Wu and Yingming Wang and Yanggeng Fu and Yanqing, “Selective ensemble learning method for belief-rule-base classification system based on PAES,” *Big Data Mining and Analytics,* vol. 2, no. 4, pp. 306-318, 2019.

[38] E. Guresen, G. Kayakutlu and T. U. Daim, “Using artificial neural network models in stock market index prediction,” *Expert Systems with Applications,* vol. 38, no. 8, pp. 10389-10397, 2011.

[39] P. Vora and B. Oza, “A Survey on K-mean Clustering and Particle Swarm Optimization,” *International Journal of Science and Modern Engineering (IJISME),* vol. 1, no. 3, pp. 24-26, 2013.

[40] F. Murtagh and P. Contreras, “Algorithms for hierarchical clustering: an overview,” *WIREs Data Mining and Knowledge Discovery,* vol. 2, no. 1, pp. 86-97, 2011.

[41] J. Majumdar, A. Mal and S. Gupta, “Heuristic model to improve Feature Selection based on Machine Learning in Data Mining,” in *6th International Conference - Cloud System and Big Data Engineering (Confluence)*, Noida, 2016.

[42] “Alex Williams,” [Online]. Available: alexhwilliams.info.

[43] Ester, M., Kriegel, H. P., Sander, J., & Xu, X., “A density-based algorithm for discovering clusters in large spatial databases with noise,” *Kdd,* vol. 96, no. 34, pp. 226-231, 1996.

[44] “scikit-learn Machine Learning in Python,” [Online]. Available: Scikit-learn.org.

[45] M. Zehnder, H. Wache, H. Witschel, D. Zanatta and M. Rodriguez, “Energy saving in smart homes based on consumer behavior: A case study,” in *IEEE First International Smart Cities Conference (ISC2)*, 2015.

[46] C. Reinisch, M. Kofler, F. Iglesias and W. Kastner, “ThinkHome Energy Efficiency in Future Smart Homes,” *EURASIP Journal on Embedded Systems,* vol. 2011, no. 1, p. 104617, 2011.

[47] V. Dolce, C. Jackson, S. Silvestri, D. Baker and A. De Paola, “Social-Behavioral Aware Optimization of Energy Consumption in Smart Homes,” in *14th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, New York, 2018.

[48] S. Mennicken, J. Vermeulen and E. Huang, “From today's augmented houses to tomorrow's smart homes,” in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, 2014.

[49] “REFIT Smart Home dataset,” Loughborough University, 20 June 2017. [Online]. Available: https://figshare.com/articles/REFIT_Smart_Home_dataset/2070091. [Accessed 06 12 2020].

[50] Kennedy, J.; Eberhart, R., “Particle Swarm Optimization,” in *IEEE International Conference on Neural Networks*, Perth, Australia, 1995.

[51] J. S.-S. G. Venter, “Particle Swarm Optimization,” in *43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Denver, 2002.

[52] C-J. Tu, L-Y. Chuang, J-Y. Chang, C-H. Yang, "Feature selection using PSO-SVM," *IAENG International Journal of Computer Science,* vol. 33, 2007.

[53] C.W. Hsu, C.C. Chang, C.J. Lin, "A practical guide to support vector classification," 2003. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.

[54] H.T. Lin, C.J. Lin, "A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods," 2003. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/tanh.pdf.

[55] C-L. Huang, J-F. Dun, "A distributed PSO-SVM hybrid system with feature selection and parameter optimization," *Applied Soft Computing,* vol. 8, no. 4, pp. 1381-1391, 2008.

[56] Xiaohua Wu, X. Hu, S. Moura, X. Yin, V. Pickert, "Stochastic control of smart home energy management with plug-in electric vehicle battery energy storage and photovoltaic array," *Journal of Power Sources,* vol. 333, pp. 203-212, 2016.

[57] "RapidMiner," [Online]. Available: https://www.rapidminer.com.

[58] D. Norris, "RapidMiner – a potential game changer," Bloor Research, 15 November 2013. [Online]. Available: https://www.bloorresearch.com/2013/11/rapidminer-a-potential-game-changer/. [Accessed 10 12 2020].

Appendix 1 – Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I Hatef Palizgar

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis “Energy Consumption in Smart Homes: Analysing User Behaviour using Combined Vector Machine and Neural Network Techniques”, supervised by Yannick Le Moullec
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons’ intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

03.01.2021

1 The non-exclusive licence is not valid during the validity of access restriction indicated in the student’s application for restriction on access to the graduation thesis that has been signed by the school’s dean, except in case of the university’s right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.

Appendix 2 – Simulation Results in RapidMiner

Figure A shows the results of applying the NN algorithm.

PerformanceVector

```
PerformanceVector:
accuracy: 96.81% +/- 1.25% (mikro: 96.80%)
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       5       0
3:      0       364     0       68
0:      0       0       1582    0
2:      0       94      0       393
classification_error: 3.19% +/- 1.25% (mikro: 3.20%)
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       5       0
3:      0       364     0       68
0:      0       0       1582    0
2:      0       94      0       393
weighted_mean_recall: 91.36% +/- 3.36% (mikro: 91.10%), weights: 1, 1, 1, 1
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       5       0
3:      0       364     0       68
0:      0       0       1582    0
2:      0       94      0       393
weighted_mean_precision: 93.25% +/- 1.79% (mikro: 91.19%), weights: 1, 1, 1, 1
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       5       0
3:      0       364     0       68
0:      0       0       1582    0
2:      0       94      0       393
```

Figure A. The results of applying the NN algorithm.

According to Figure A, the NN method's accuracy is equal to 96.81% to predict the optimum energy consumption. The accuracy, recall, and prediction error are equal to 93.25%, 91.36%, and 3.19%, respectively.

Next, the results of the SVM algorithm are shown in Figure B.

PerformanceVector

```
PerformanceVector:
accuracy: 95.86% +/- 2.17% (mikro: 95.88%)
ConfusionMatrix:
True:   1       3       0       2
1:      2675    0       0       0
3:      0       277     0       8
0:      31      0       1544    0
2:      0       175     0       481
classification_error: 4.14% +/- 2.17% (mikro: 4.12%)
ConfusionMatrix:
True:   1       3       0       2
1:      2675    0       0       0
3:      0       277     0       8
0:      31      0       1544    0
2:      0       175     0       481
weighted_mean_recall: 89.91% +/- 5.11% (mikro: 89.63%), weights: 1, 1, 1, 1
ConfusionMatrix:
True:   1       3       0       2
1:      2675    0       0       0
3:      0       277     0       8
0:      31      0       1544    0
2:      0       175     0       481
weighted_mean_precision: 90.17% +/- 9.25% (mikro: 92.14%), weights: 1, 1, 1, 1
ConfusionMatrix:
True:   1       3       0       2
1:      2675    0       0       0
3:      0       277     0       8
0:      31      0       1544    0
2:      0       175     0       481
```

Figure B. The results of applying the SVM algorithm.

According to Figure B, this method's accuracy is equal to 95.86% to predict the optimum energy consumption. The accuracy, recall, and prediction error are equal to 90.17%, 89.91%, and 4.14%, respectively.

Finally, the results of the proposed SVM-NN algorithm are shown in Figure C.

PerformanceVector

```
PerformanceVector:
accuracy: 99.37% +/- 0.50% (mikro: 99.37%)
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       2       0
3:      0       453     0       26
0:      0       0       1585    0
2:      0       5       0       435
classification_error: 0.63% +/- 0.50% (mikro: 0.63%)
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       2       0
3:      0       453     0       26
0:      0       0       1585    0
2:      0       5       0       435
weighted_mean_recall: 98.35% +/- 1.34% (mikro: 98.29%), weights: 1, 1, 1, 1
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       2       0
3:      0       453     0       26
0:      0       0       1585    0
2:      0       5       0       435
weighted_mean_precision: 98.35% +/- 1.29% (mikro: 98.34%), weights: 1, 1, 1, 1
ConfusionMatrix:
True:   1       3       0       2
1:      2711    0       2       0
3:      0       453     0       26
0:      0       0       1585    0
2:      0       5       0       435
```

Figure C. The results of applying the proposed SVM-NN algorithm.

According to Figure C, this method's accuracy is equal to 99.37% to predict the optimum energy consumption. The accuracy, recall, and prediction error are equal to 98.35%, 98.35%, and 0.63%, respectively.