Master Thesis

-

Master Thesis Report

Roberto Rodríguez

Aalborg University
Electronics and IT

Here you can write something about which tools and software you have used for typesetting the document, running simulations and creating figures. If you do not know what to write, either leave this page blank or have a look at the colophon in some of your books.

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Scheduling of home appliances based on adaptive user optimization and diverse forecasting models.

Theme:
Energy, machine learning

Project Period:
Spring Semester 2020

Project Group:
Roberto Rodríguez

Participant(s):
Thomas Dyhre Nielsen

Supervisor(s):
1

Copies: May 17, 2021

Page Numbers: 42

Date of Completion:

Electronics and IT
Aalborg University
http://www.aau.dk

Abstract:

In United States the residential and commercial buildings consume 73% of the electricity. The Smart Grid implementations have grown boosting concepts such as: Demand Side Management (DSM), Advanced Metering (AM), Demand Response (DR) and Scheduling and Forecasting (SF). The renewable energy sources as wind turbines and photovoltaics (PV) behave uncertainly, therefore there is a gap between the supply and demand energy. To tackle the imbalances, many studies have proposed solutions based on DR strategies to reschedule the load energy. From this perspective to accomplish energy efficiency at household level, it is necessary to use the flexibility concept to adjust the supply demand gap. This project proposes to get the possible energy loads that can be rescheduled as flexible consumption descriptions (flex-offers). This work focuses on wet devices (washing machine, dishwasher) because they can change the behaviour to fit in the RES production energy and they represent 30% of household consumption.
In Demand Side Management, the pricing mechanisms are designed to encourage the consumers to change their behaviour, for example the time-of-use pricing sets different prices during the day, hence the consumer change the demand to off-peak hours. In this context, to schedule the consumer loads, we have to apply the best machine learning models to get the best results.

Contents

Preface **ix**

1 Introduction **1**
1.1 Dataset: Description and Analysis 2
1.1.1 Washing machine 3
1.1.2 Dishwasher 6
1.1.3 Dryer 6
1.1.4 Oven 8
1.1.5 House size in square footage 11
1.1.6 Features 11

2 Flex Offer **13**
2.1 Flexibility and Flex offers 13
2.2 Predictions 15
2.2.1 Prediction of Appliance activation 15
2.2.2 Probabilistic Flex- Offers 15

3 Machine Learning Algorithms **19**
3.1 Supervised Learning Algorithms 19
3.2 Bayesian Network (BN). 19
3.2.1 Naive Bayes Classifier 19
3.3 Linear Regression 21
3.4 Neural Networks 23
3.5 Decision Trees 24

4 Adaptive user utility models **27**
4.1 User utility 27
4.2 User Flexibility 28
4.3 Estimation of User Flexibility 28

5 Scheduling appliances **31**
5.1 Spot Market Savings 31
5.2 User utility by scheduling 32

6 Experiments **33**
6.1 Dataset 33
6.2 Evaluating activation device prediction 34

7 Conclusions **39**

Bibliography **41**

Preface

Aalborg University, May 17, 2021

Roberto Rodríguez
<rdro17@student.aau.dk>

Chapter 1

Introduction

In 2015, residential and commercial buildings consumed 73% of the electricity in the U.S. [8] The Smart Grid implementations have grown boosting concepts such as: Demand Side Management (DSM), Advanced Metering (AM), Demand Response(DR) and Scheduling and Forecasting(SF)[2]. The renewable energy sources as wind turbines and photovoltaics (PV) behave uncertainly, therefore there is a gap between the supply and demand energy. To tackle the imbalances, many studies have proposed solutions based on DR strategies to reschedule the load energy. From this perspective to accomplish energy efficiency at household level, it is necessary to use the flexibility concept to adjust the supply demand gap [1]. The proposed aim is to get the possible energy loads that can be rescheduled as flexible consumption descriptions (flex-offers). This project focuses on wet devices (washing machine, dishwasher) because they can change the behaviour to fit in the Renewable Energy Production (RES) production energy and they represent 30% of household consumption [12].

In DSM, the pricing mechanisms are designed to encourage the consumers to change their behaviour, for example the time-of-use pricing sets different prices during the day, hence the consumer changes the demand to off-peak hours [13]. In this context, to schedule the consumer loads, it is necessary to get real time and predicted future data, maximum device flexibility, device usage preferences and manual device operation scheduling [5]. In the DSM environment the consumer interaction is necessary to shift the loads, however the new consumer requirement to participate in decisions can lead consumers to get tired of rescheduling appliances. This phenomenon is known as “response fatigue” [11].

Based on the time horizon the energy load forecasting can be classified in: Very Short Term forecasting (VSTF), Short Term Forecasting (STF), Medium Term Forecasting (MTF) or Long Term Forecasting (LTF). Considering that the Locational Marginal Price (LMP) is given by Day-Ahead or Real-Time pricing, in both cases the prices are released for the next day. This project is narrowed down to focus

Column name	type	Description
localminute	date	Date and hour of the measure
dataid	int64	The unique identifier for the home
use	float64	The total home consumption
clotheswasher1	float64	The clothes washer consumption
dishwasher1	float64	The dishwasher consumption
drye1	float64	The dryer consumption
oven1	float64	The oven consumption

Figure 1.1: Data dictionary

only at STF, specifically to estimate daily the next two consecutive hour-level and day-ahead wet-device activations (the details will be discussed in Section 4) [6]. The hour to make the predictions daily will be related with the smaller activity of each device at the end of the day, giving time to the user to override the new schedule.

1.1 Dataset: Description and Analysis

The dataset used in this Project is from Pecan Street organization. In dataport web site that manages the Pecan Street datasets, provides access to data on consumer energy and water consumption behaviour. From Dataport it can analyse, visualize and create custom reports from appliance-level consumer. The dataset provides measures circuit-level usage and generation from approximately 1000 homes that are located mainly in Austin Texas. The measures can be obtained minutely or hourly. As this is a large dataset is important be careful to select the right data. This study selected the columns shown in Figure1.1:

The number of houses from which the data could be collected is inversely proportional to the period of time that it selects for the query. i.e. It is possible to get cured data from 20 houses in a period of one year, however it can get the data from 11 houses in two years. This means the shorter period of time, the more information can be obtained. The selected data period is from January 2017 until December 2018 and this period has been chosen because it was the most recent data available from Dataport.

As mentione in the introduction this project focuses on wet devices, such as dishwashers and washing machines, they are responsible for approximately of 30% of the household energy consumption.[4]. To find relations between the wet devices and other devices, it has been selected two more appliances that are related

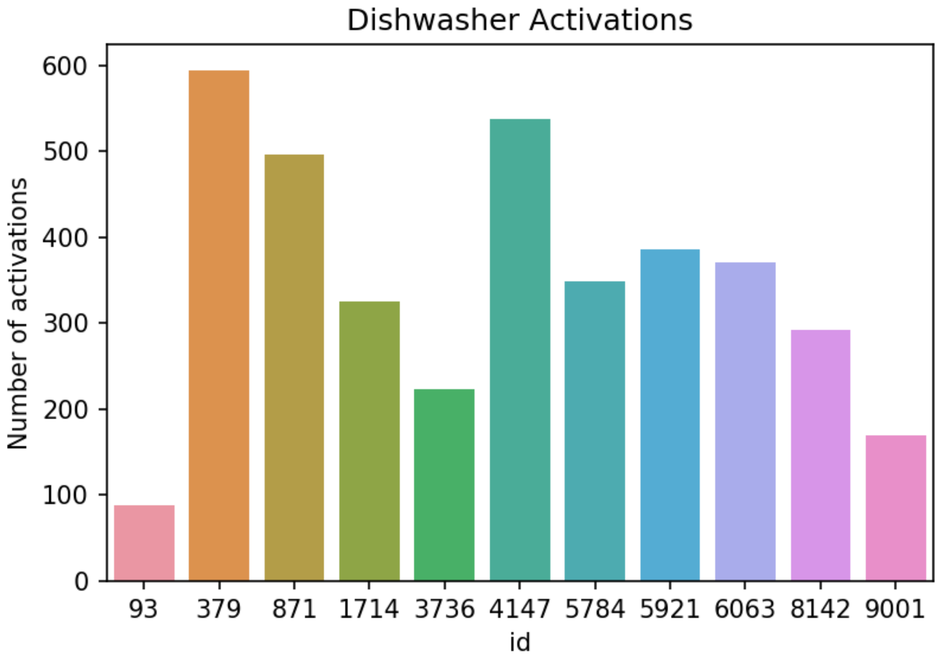

Figure 1.2: Total Activations for House:Washing Machine

with the wet devices, the oven and the dryer.

To analyse appliances behaviour along the time, it has been created graphs with different features that we are showing above:

1.1.1 Washing machine

The total activations for houses along the two years for the washing machine is shown in the Figure1.2,there are three houses(379,4147,871), whose values are from 500 to 600 activations. Five houses have from 300 to 400 activations. And finally 3 houses under 200 activations.

The total activations for weekdays are shown in the Figure1.3. The weekdays are shown from Monday to Sunday and as it supposed, Sunday is the most popular day to make laundry, whereas Tuesday is the second most popular one. The number of activations of these two days is around 600, meanwhile the least popular weekday, Thursday has around 400 activations.

The total activations of all houses per hour is shown in the Figure1.4. It is visible that, the most frequent hours to activate the washing machine are the morning hours from 9:00 to 12:00.

The washing machine activations per house, group by weeekday is shown in the Figure **??**

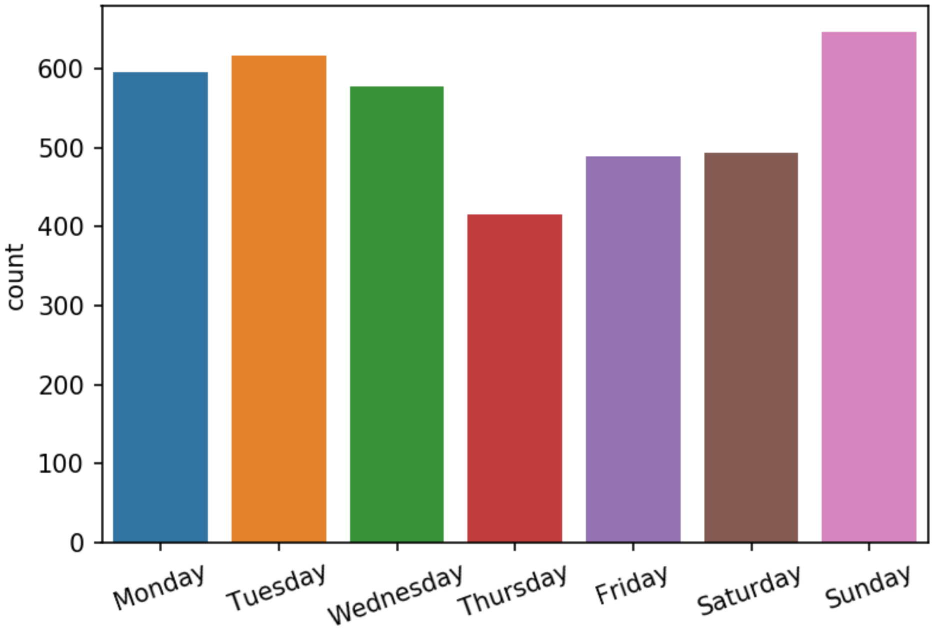

Figure 1.3: Total Activations for Weekday: Washing Machine

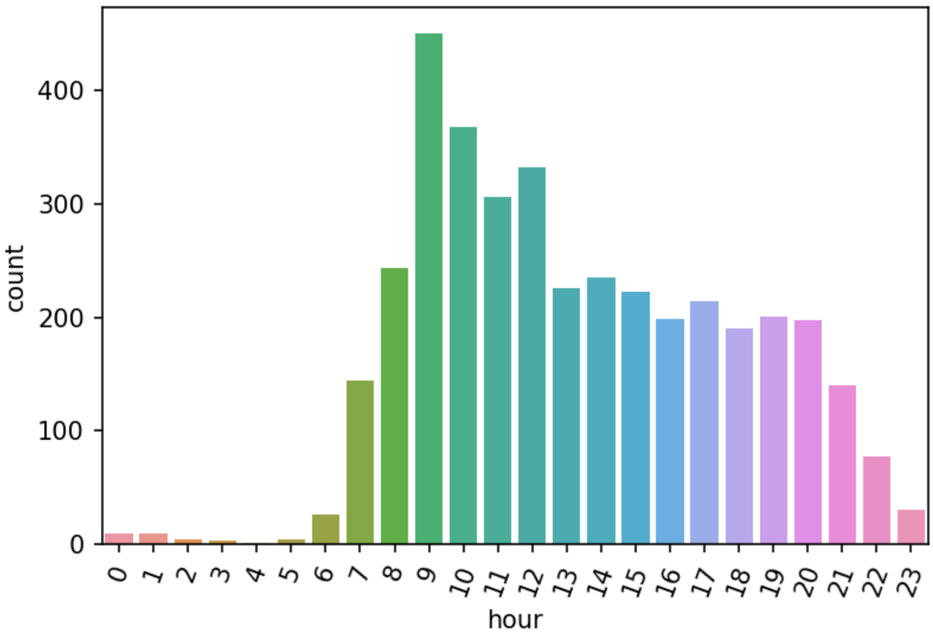

Figure 1.4: Total Activations for hour: Washing Machine

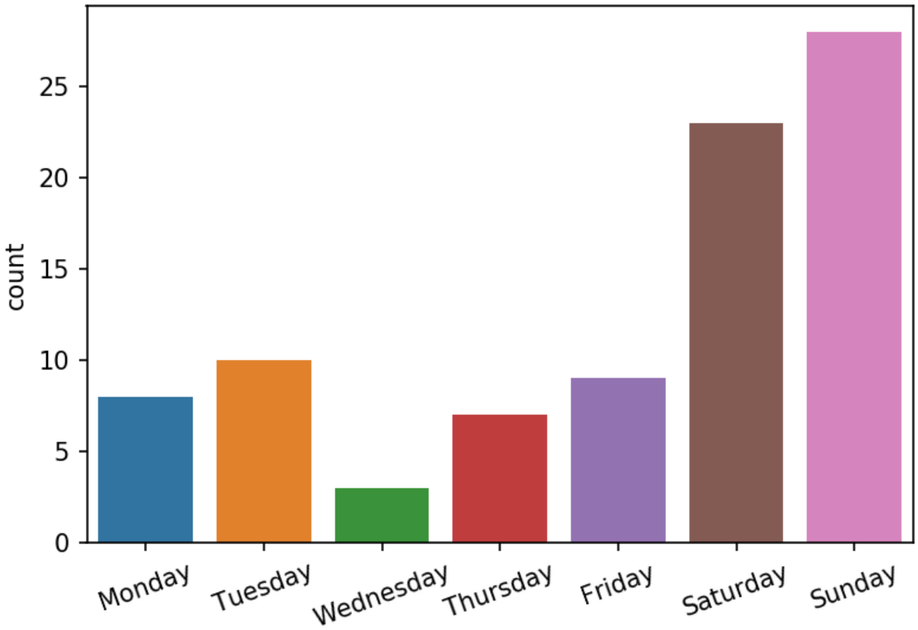

Figure 1.5: House 93

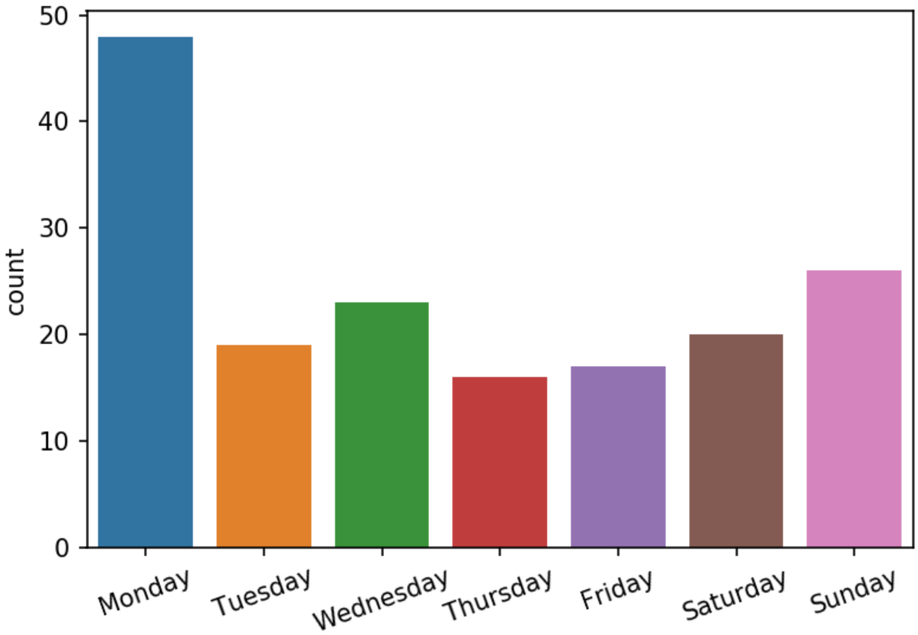

Figure 1.6: House 9001

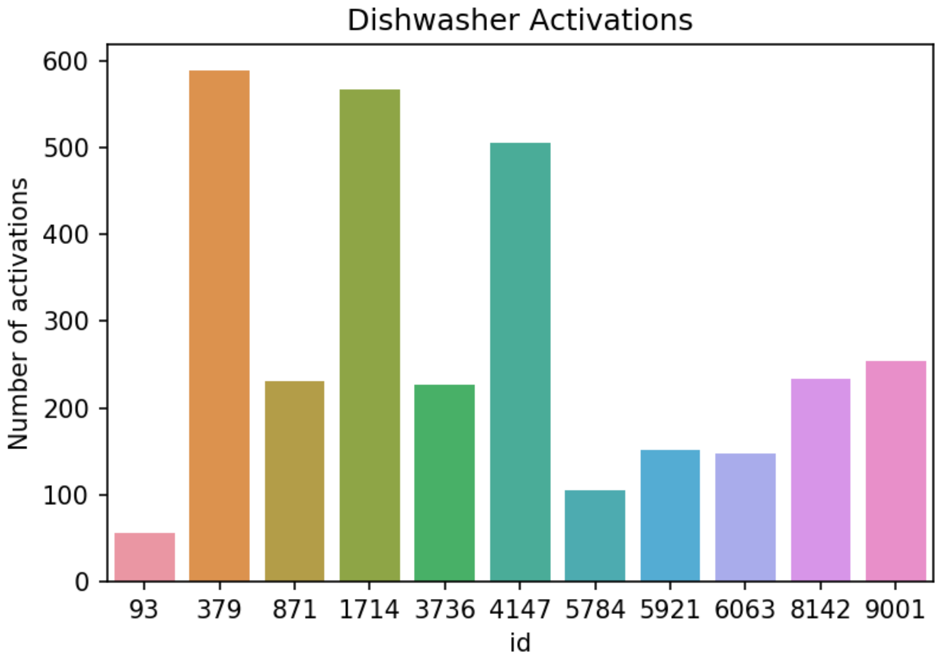

Figure 1.7: Total Activations for House: Dishwasher

1.1.2 Dishwasher

The total dishwasher activations for houses is shown in the Figure1.7. Houses(379, 1714, 4147) have between 500 to 600 activations. The houses 379 and 4147 have higher activations number with respect to the washing machine too, where the house 93 has the lowest value, similar to washing machine activations.

The activations number for weekdays are shown in the Figure1.8. The highest number of activation is on Sunday, with more than 500 activations. However it does not differ much from the lowest number of activations corresponding to Friday having 400 activations.

The total activations for every hour is shown in the Figure1.9. As it is visible, it has a substantial number of activations in two ranges: from 6:00 to 12:00 and from 18:00 to 21:00.

1.1.3 Dryer

Figure1.10 shows the total dryer activations for houses. The highest numbers of activations are in houses 379, 4147 and 871. As we observed in the Figure1.2 these are the three houses with more activations for washing machine too.

The total dryer activations for weekdays is shown in the Figure1.11. It has

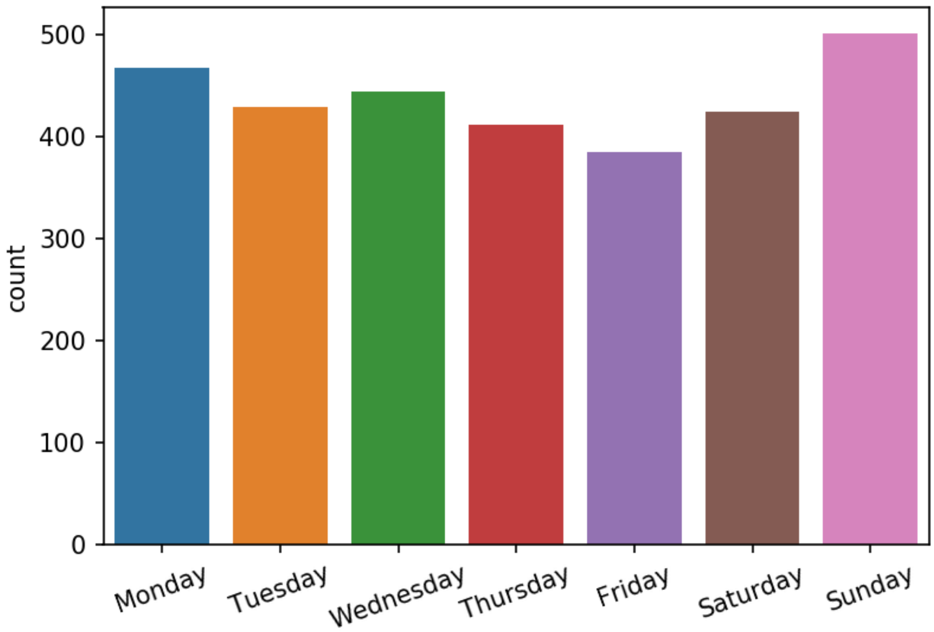

Figure 1.8: Total Activations for Weekday: Dishwasher

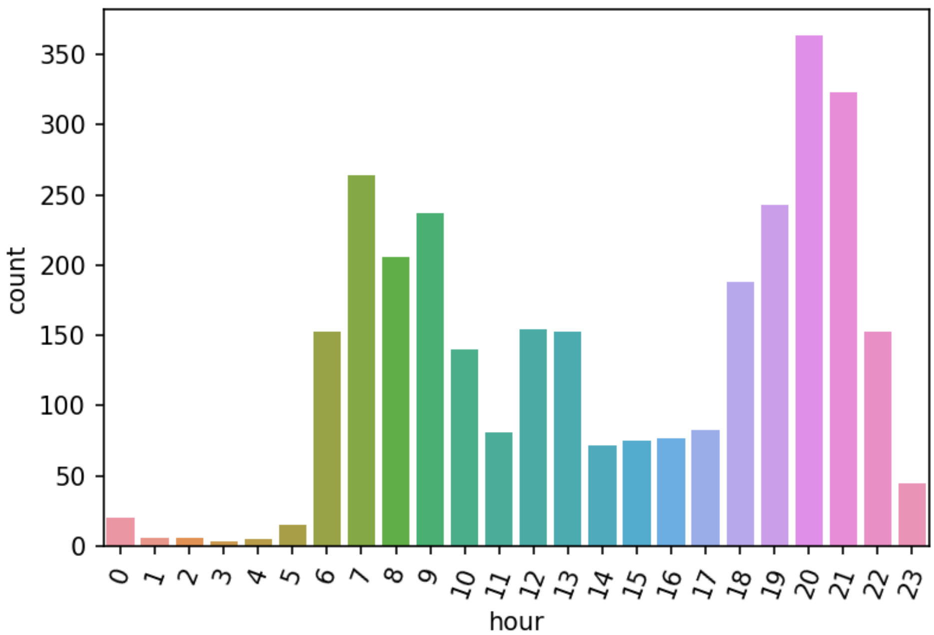

Figure 1.9: Total Activations for Hour: Dishwasher

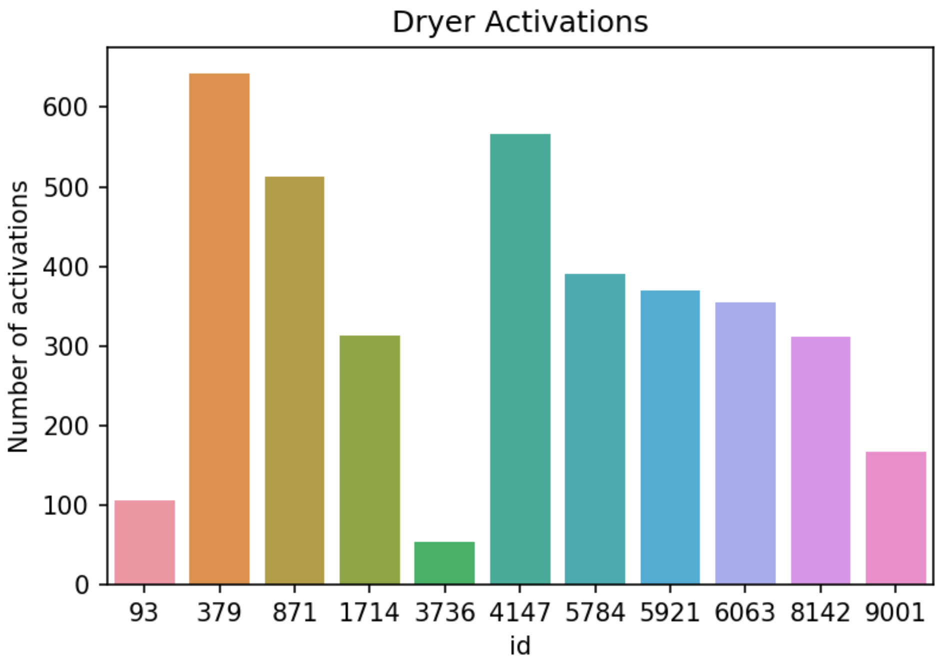

Figure 1.10: Total Activations for House: dryer

been observed that the dryer has similar behavior with the washing machine with respect to the number of activations.

The total activations for an hour is shown in the Figure1.12. There is a considerable dryer activity from 8:00 to 22:00, however The peak hours are from 10:00 to 13:00.

1.1.4 Oven

The oven activations for houses is shown in the Figure1.13. It is notable that there are two houses(5784,8142) with very high values of activations. Oven data has been chosen to find out a correlation with the dishwasher.

The oven activations for weekdays is displayed in the Figure1.14. As all the appliances the most popular day to use it is Sunday, and the smaller value belongs to Thursday.

The total oven activations along the day are shown in the Figure1.15. The peak hours are in three different ranges: 7:00 to 8:00, 11:00 to 12:00 and 17:00 to 18:00

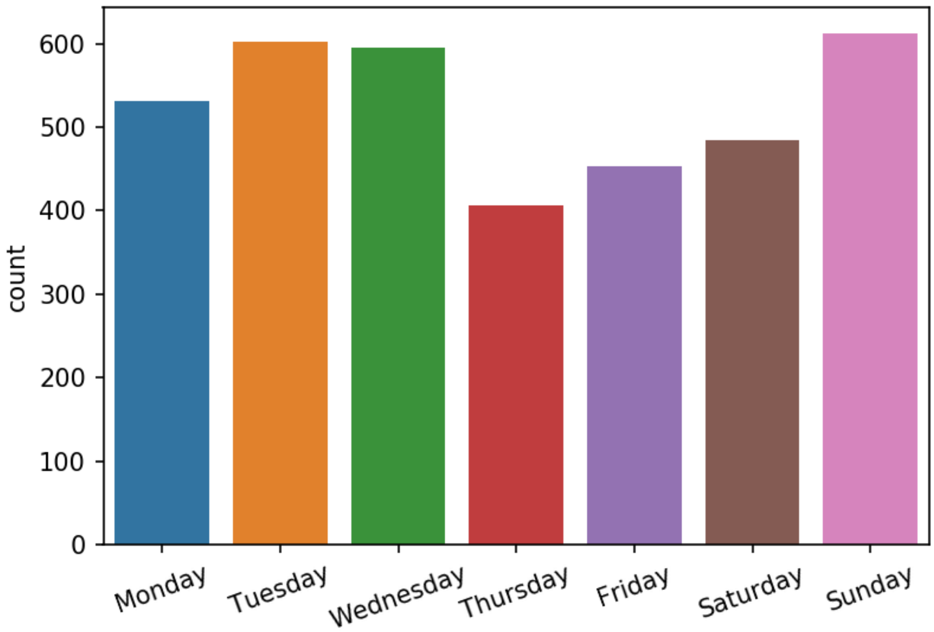

Figure 1.11: Total Activations for Weekday: dryer

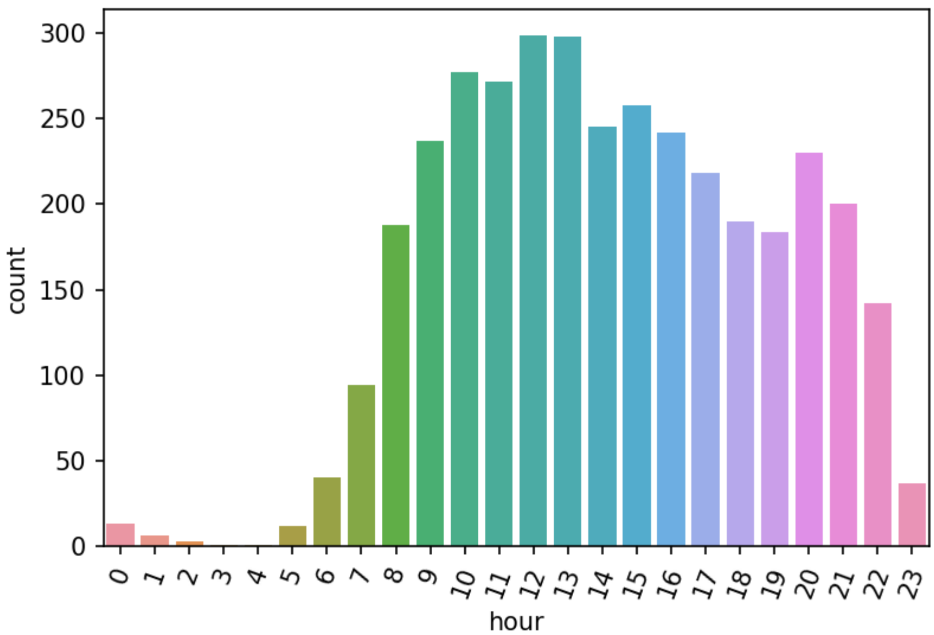

Figure 1.12: Total Activations for Hour: Dryer

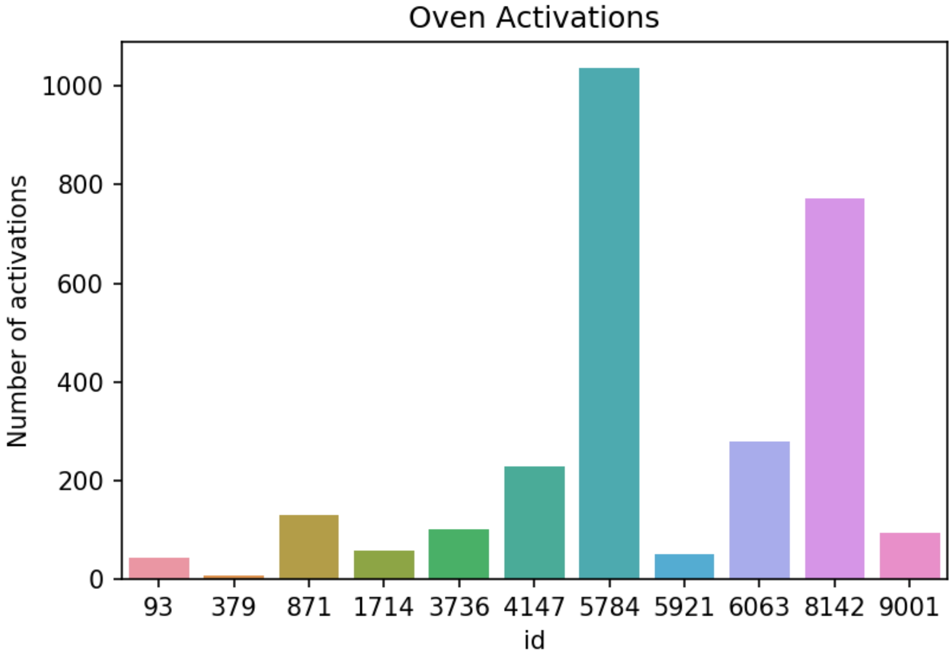

Figure 1.13: Total Activations for House: oven

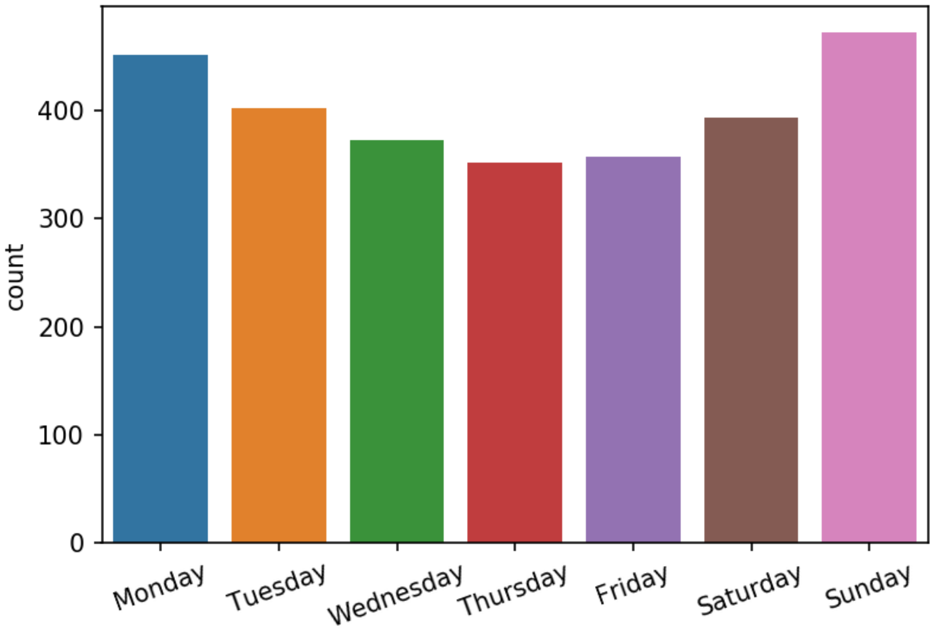

Figure 1.14: Total Activations for Weekday: oven

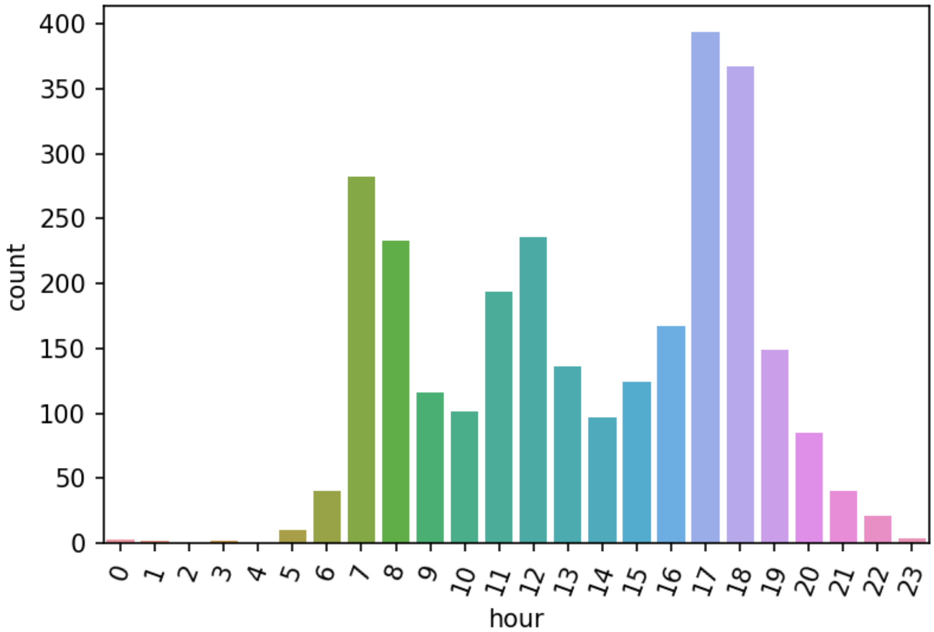

Figure 1.15: Total Activations for Hour: Oven

1.1.5 House size in square footage

One of the factors that could be considered as a predictor is the number of people in the house, however this information is not availableon the dataset. Althought it provides the information of the size in square footage to check the correlation with the appliances activations. The results are shown in the Figure1.16 where is visible that the house 379 second biggest size, has the highest number of activations regarding to the washing machine, dishwasher and dryer. Yet the rest of the houses do not have any relationship between size house and activations. For example the house 93 is one of the biggest house, however it has the smallest value in activations in most of the appliances.

1.1.6 Features

As it could observe in the previous graphics about the data analysis. Base on the high correlation between the features and the response, it has been selected the following predictors:

- Day of the week
- Day of the month

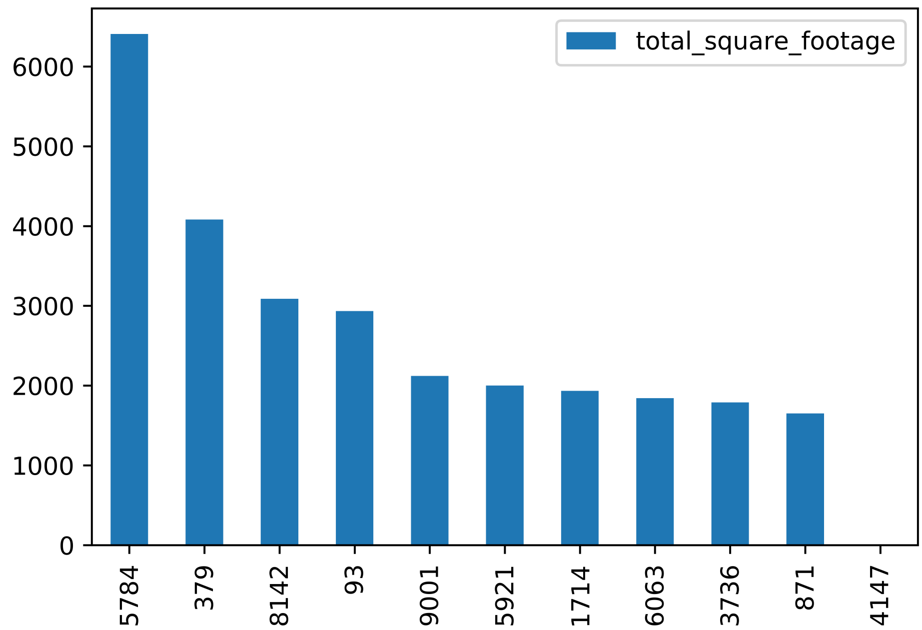

Figure 1.16: house size

- Month
- Year
- hour

Additionally it will be tested eigth more predictors:

- Time elapsed, between the current activation and the previous five activations. It means that there will be five new features.
- The mean among the time elapsed features.
- The mean among all the time elapsed for each house.
- The standard deviation among all the time elapsed for each house.

Chapter 2

Flex Offer

2.1 Flexibility and Flex offers

To manage the energy demand, this project proposes to use the flexibility devices which means shifting load-devices from a planned timestamp to other timestamp with utility benefits. The flexible energy demand can be represented in two dimensions[9].

- The first dimension is the time flexibility that depicts the possibility to reschedule the load-device.
- The second dimension is the amount flexibility that represents the energy consumption at specific time.

Definition 1. A flex-offer f is a tuple

$$f = ([t_{es}, t_{ls}], p)$$

, where

$$[t_{es}, t_{ls}]$$

is the time interval during which to trigger the Activate action and p is the energy profile. p is a sequence of slices

$$< s_1, ..., s_d >$$

, where a slice sd is a continuous range

$$[e_{min}, e_{max}]$$

defined by the minimum e_{min} and maximum e_{max} energy bounds, and d is the number of slices in p[13]

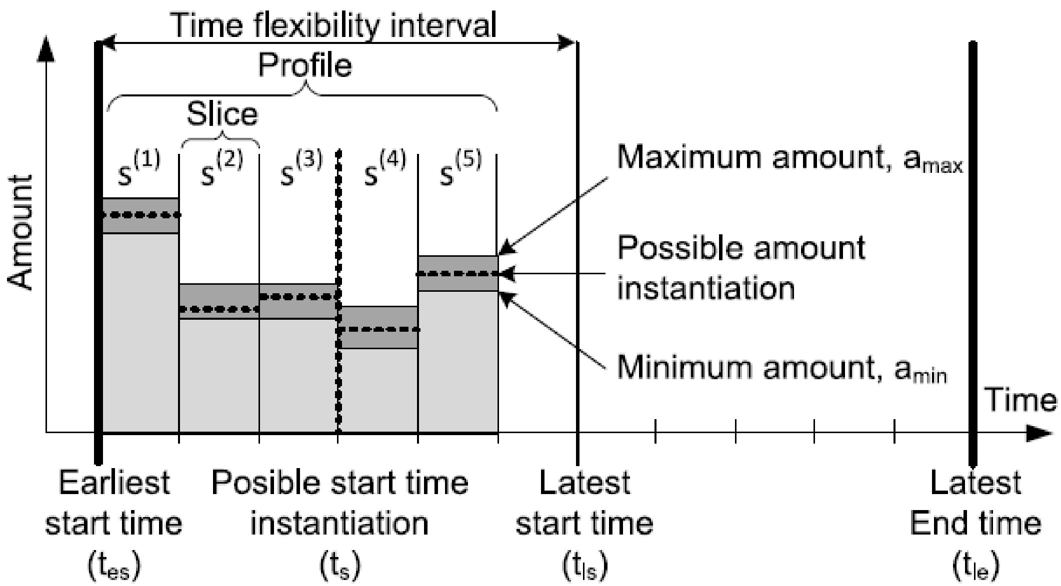

Figure 2.1: Flex Offer on two dimensions [13]

The latest end time of the device operation is calculated as

$$t_{le} = t_{ls} + d$$

that is shown in the figure 2.1 [9]. The flex-offer works on both time and amount flexibility, this project only focuses on the time dimension of flex-offers, as we focus on wet-devices and they do not allow amount postponement, but only for time shifting[1], such that

$$e_{min} = e_{max}$$

While we are working on uncertain environment on demand and supply in the energy market, Flex-offers can be a very important aid to the energy market players such as Balance Responsible Parties (BRPs) and Distribution System Operators (DSOs). BRPs could get benefit of Flex-Offers, using it to schedule demands that minimize their market deviations. In this context the prosumers (individual who both consumes and produces) provide flexibilities to the market players, who can exploit it for financial and share part of the benefits to the prosumers.[1]

2.2 Predictions

2.2.1 Prediction of Appliance activation

The Home Energy Management System(HEMS) can supply Appliance Load Monitoring (ALM) in costumers sites, in this project we use Dataport dataset described in 1.1 that works like Non-intrusive Load Monitoring (NILM). This type of load monitoring allows us to get disaggregated household electrical load measured at a single appliance [14] Dataport supply the load measurements in terms of Kw/h, therefore we just take the initial timestamp of each appliance to transform this information in activations per appliance. We analyzed the behavior in the appliances(washing machine and dishwasher), and we concluded that the average operation duration is 70 minutes, it means that even there are energy load absense in this 70 minutes, the machine still works in the same task. Device activations are described like:

$$A = [a_1, a_2, ..., a_n]) \tag{2.1}$$

where a_1 is the first activation and a_n the last activation in the period of two years. It is important to mention that the period between a_1 and a_2 is going to be a feature to predict the activations.

Then we use the activation set per each appliance to predict an hour-level predictions of the next two activations . Thus we can produce the earliest start time for the first activation and the latest start time for the second activation, of a flex offer.

Given the random variables T_{es}, the timestamp of the earliest start time, and T_{le}, the timestamp of the latest end time, and given the dataset 1.1 with the data series, with the granularity of the chosen aggregation level, in this project is hourly granularity. We will predict T_{es} probability, given an evidence set $P(T_{es} \mid e)$ in this case the set is the dataset from Dataport, with the selected features described in 1.1.6. Then we predict

$$P(T_{le}, T_{es}|e) = P(T_{le}|T_{es}, e)P(T_{es}|e) \tag{2.2}$$

Being the probability of the time of the earliest start time and the latest end time, conditioned by the probability of T_{es}.

2.2.2 Probabilistic Flex- Offers

As we described in section 2.1 the Flex Offer model represent the range flexibility between two timestamps, although our model represents the T_{es} and the T_{le} like random variables, therefore it is necessary creating the new definition of probabilistic flex offer [3]. A probabilistic flex offer is a tuple

$$f = ([T_{es}, T_{le}], p)$$

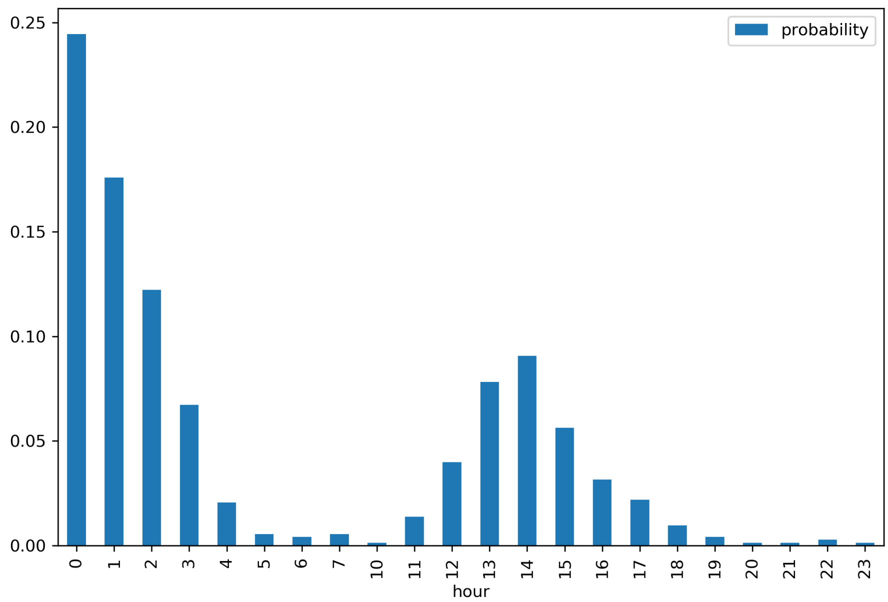

Figure 2.2: Earliest Start Time

where T_{es} and T_{le} are discrete random variables to represent the earliest start time and the latest end time respectively and p is the energy profile of activation. $[T_{le}, T_{es}]$ defines a set of $|T_{le}|$ x $|T_{es}|$ possible flexible intervals where the activation can be reschedule, described by the tuple $<[t_{es}, t_{le} - |p|], P(T_{es} = t_{es}, T_{le} = t_{le})>$ where $t_{es} \in T_{es}$ and $t_{le} \in T_{le}$, $|p|$ is the length of operation and $P(T_{es} = t_{es}, T_{le} = t_{le})$ is the interval probability define in 2.2 The probabilistic flex offer considers all the intervals $[t_{es}, t_{ls}]$ where $t_{ls} = t_{le} - |p|$ on the period between the earliest start time and the latest end time. The figure 2.2 shows the earliest start hour distribution for our dataset, and the figure 2.3 shows the distribution for the latest start hour. In 5 it will be detailed how to use flex-offers to get an accurate schedule flexible demand.

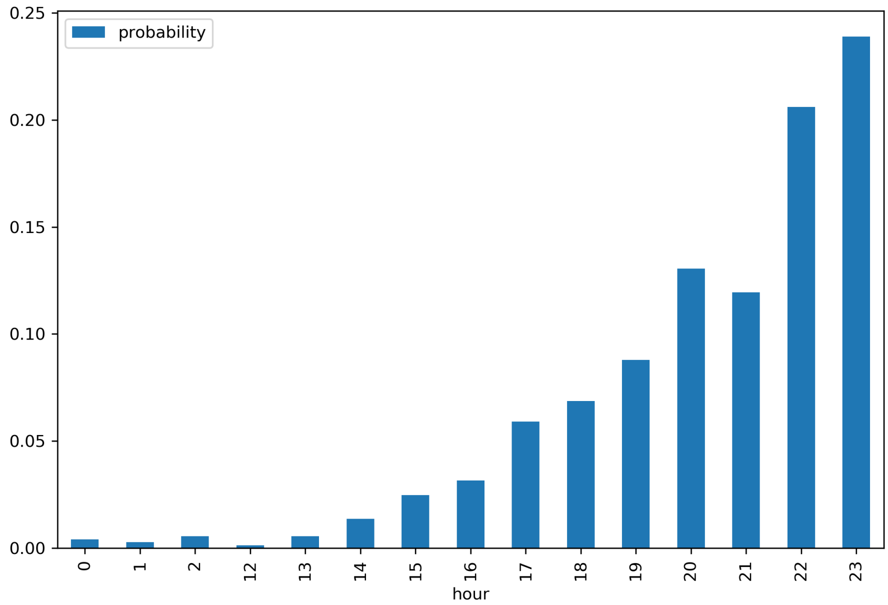

Figure 2.3: Latest Start Time

Chapter 3

Machine Learning Algorithms

3.1 Supervised Learning Algorithms

Inductive machine learning is the process of deducing a set of rules from training set examples. It means building a classifier algorithm to use for generalizing unseen instances [7]. The figure 3.1 shows the process to solve a real problem by supervised learning

3.2 Bayesian Network (BN).

A BN is a probabilistic graphical model that represents the relationships among a set of variables (features X) and their conditional dependencies by a directed acyclic graph (DAG) S. The nodes in the graph S are in one-to-one correlation with variables X. The arcs among the nodes are the effects between the nodes, whilst the nodes that are not connected represent variables that are conditionally independent of each other. To learn a BN is necessary to develop two parts: first the DAG learning and second the probabilistic parameters resolution. We can represent the probabilistic parameters via set of tables, one for each variable, as it is shown in the figure 3.2 [7]

3.2.1 Naive Bayes Classifier

Naive Bayesian networks (NB) basically are Bayesian networks that are made of DAG with only one parent (the unobserved node) and several children (the observed nodes) with a strong assumption of independence among child nodes respect to their father. In this way, the independence model (Naive Bayes) is based on estimating the equation 3.1

$$R = \frac{P(i|X)}{P(j|X)} = \frac{P(i)P(X|i)}{P(j)P(X|j)} = \frac{P(i)\prod P(X|i)}{P(j)\prod P(X|j)} \tag{3.1}$$

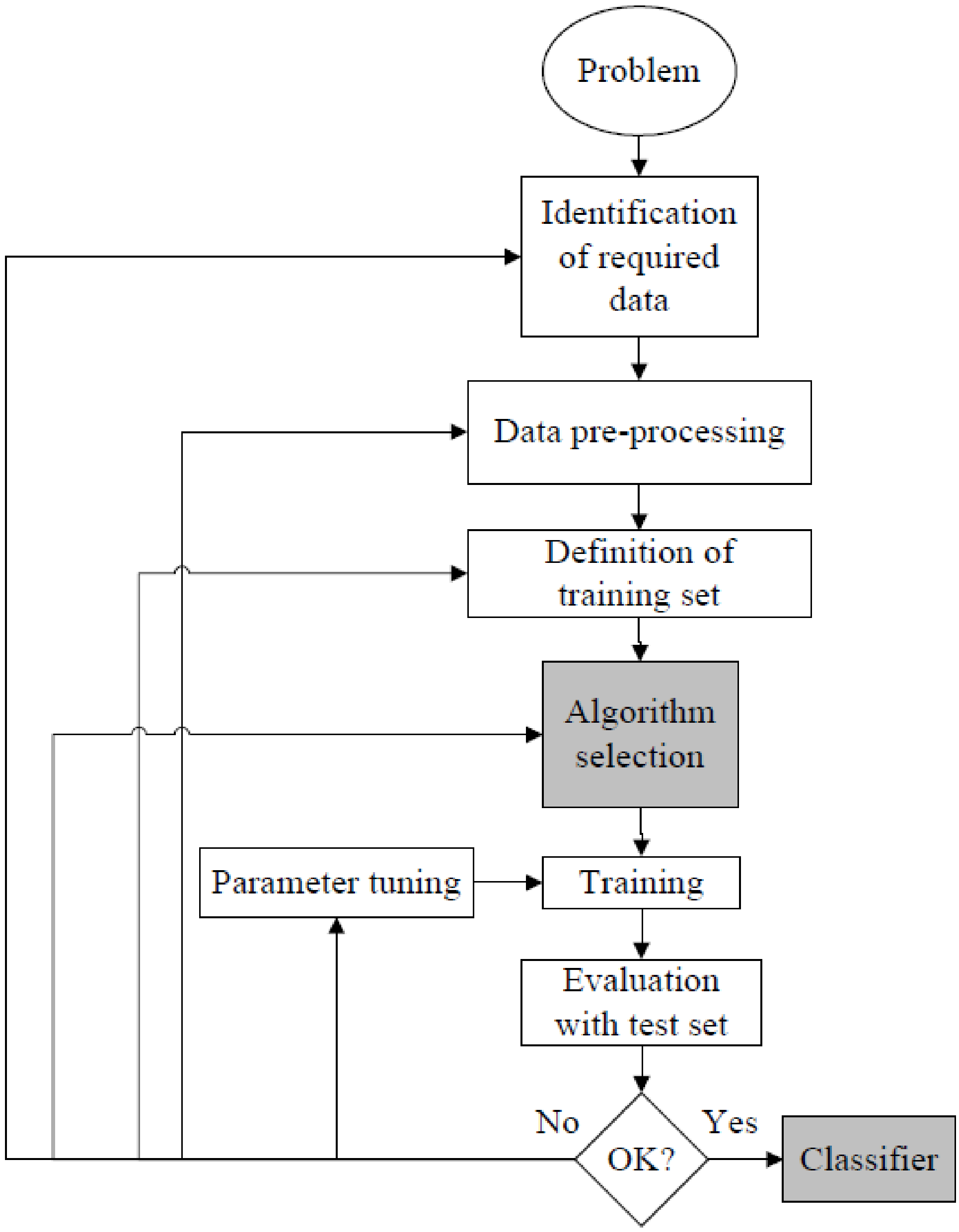

Figure 3.1: Supervised Learning process to solve a problem

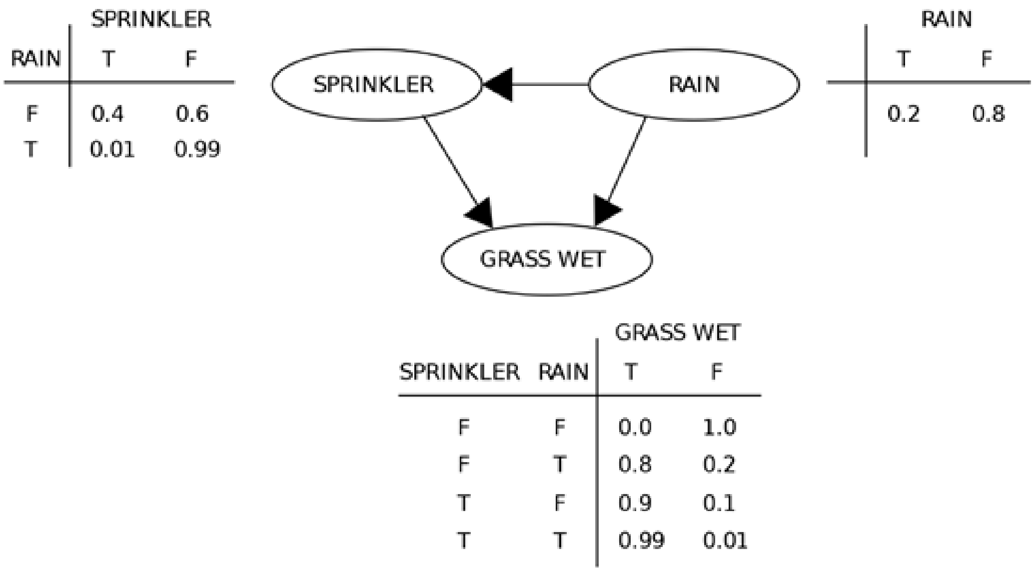

Figure 3.2: Bayesian Network

It compares these two probabilities

$$P(i|X)$$

and

$$P(j|X)$$

and the longer probability specifies which class label that is more likely to be the real label (if R>1: predict i; else predict j). Naive Bayes classification is prone to being extremely affected by 0's probabilities, if you have no occurrences of a class label and a certain attribute value together then the frequency-based probability estimate will be zero, NB uses a product operation to compute the probabilities P(X, i). This can be avoided by using Laplace estimator or m-esimate, by adding one to all numerators and adding the number of added ones to the denominator [7] . The assumption of independence among child nodes is constantly erroneous therefore other sophisticated algorithms like ANNs are more accurate than NB.

3.3 Linear Regression

The concept of regression and the methods to research the relationships between two variables have been studied since 100 years ago. There are two types of linear regression, the simple regression and the multiple linear regression. We will focus on the multiple linear regression.

Suppose that there are multiple input features (predictors) X1,. . . ,Xn are all numeric and there is one target Y. The linear function of the input must be like the

$$Y^{\bar{w}}(e) = w_0 + w_1 * X_1(e) + ... + w_n * X_n(e) = \sum_{i=0}^{n} w_i * X_i(e) \tag{3.2}$$

where

$$\bar{w} =< w_0, w_1, ..., w_n > \tag{3.3}$$

is a tuple of weights, to make easier to manipulate the equation, we create X0, with value 1.

We say that E is a set of data, The sum of square errors on examples E for target Y is

$$error(E, \bar{w}) = \sum_{e \epsilon E} (Y(e) - Y^{\bar{w}}(e))^2 = \sum_{e \epsilon E} (Y(e) - \sum_{i=0}^{n} w_i * X_i(e))^2 \tag{3.4}$$

The weights proposed can be calculated analytically. Other approach to use for wider classes of functions is computing the weights iteratively [10].

Gradient descent is an iterative approach to find the function's minimum. Gradient descent for minimizing errors begins with a set of weights; in each step, it decreases each weight in proportion to its partial derivative like 3.5 shows.

$$w_i := w_i - \eta * \frac{\partial}{\partial w_i} error(E, \bar{w}) \tag{3.5}$$

where

$$\eta$$

, is the learning rate. The learning rate, the features and the data, is given as input to the learning algorithm. The partial derivative defines how much a little change in the weight would alter the error. The sum-of-squares error is convex for a linear function with only one local minimum, which is the global minimum. As gradient descent with small step size will converge to a local minimum, therefore the algorithm will converge to the global minimum [10].

Consider minimizing the sum-of-squares error. The partial derivative of the error in the equation in 3.5 with respect to weight wi.

$$\frac{\partial}{\partial w_i} error(E, \bar{w}) = \sum_{e \epsilon E} -2 * \delta(e) * X_i(e) \tag{3.6}$$

where

$$\delta(e) = Y(e) - \bar{Y}^{\bar{w}}(e) \tag{3.7}$$

Gradient descent will renew the weights after sweeping through all examples. An alternative is to renew each weight after each example. Each example e can update each weight wi using the equation 3.8

$$w_i := w_i + \eta * \delta(e) * X_i(e), \tag{3.8}$$

where the constant 2 has been deleted, because we expect it is consumed by

$$\eta$$

3.4 Neural Networks

The neural networks are inspired by the brain's neurons but they do not replicate neurons. The number of neurons in the Artificial neural networks are much less than the 1011 neurons that there are in the human brain, and the artificial neurons, called units, are much simpler than the real biological neurons. Neural networks have had considerable success in different machine learning applications such as for image interpretation, speech recognition and machine translation[10]. The main reason is that they are very flexible and can create new features. In this project, we will use feed-forward neural networks. Feed-forward networks are the simplest artificial neural network (ANN), in this type of ANN, the data flow just in one direction. This ANN can be seen like a hierarchical organization chart, but with more connections. A regular architecture is shown in the figure 3.3 where we can appreciate three layers, with multiple units (neurons) in each layer. In the first layer on the left are the input neurons for the input features (predictors), the second layer in the middle is the hidden layer, could there be more than one, where are the hidden neurons, that are features that we never observed, but are convenient for predictions. The last layer on the right are feed by the hidden neurons and is the output layer that are the predictions of this architecture.

Between the input layer and the Hidden Layer we have a complete linear layer, where each output Oj is a linear function of the input values Vi to the layer (as in linear regression, we added X0 =1) described as

$$O_j = \sum i W_{ji} V_i \tag{3.9}$$

for weights W_{ji} that are learned. For each input-output, there is a weight. In the figure 3.3 there is a weight for every arc for the linear functions. Every linear function is affected by an activation f thus: O_i = f (V_i). Examples of activation function are: sigmoid an relu.

Backpropagation.

Back-propagation is an algorithm for training feedforward neural networks, that implements stochastic gradient descent, for all weights. As we saw in3.3 stochastic gradient descent updates each weight w with

$$\frac{\partial}{\partial w_i} error(e) \tag{3.10}$$

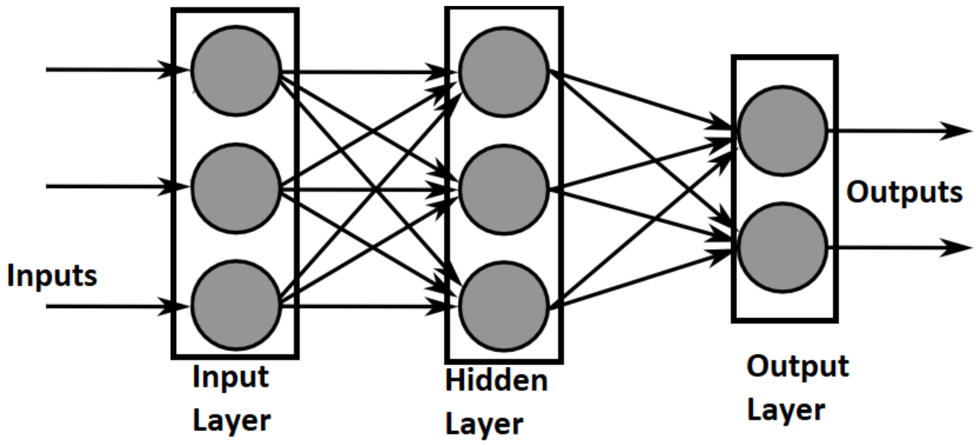

Figure 3.3: Neural Network Architecture

for each example e

There are two properties that back-propagation applies like algorithm:

• Linear rule: the derivative of a linear function is given by:

$$\frac{\partial}{\partial w}(aw + b) = a \tag{3.11}$$

therefore the derivative is just the number that is multiplied by w in the linear function.

• Chain rule: if g is a function of w and function f , that is independent on w is applied to g(w), then

$$\frac{\partial}{\partial w}(f(g(w))) = f'(g(w)) * \frac{\partial}{\partial w} g(w) \tag{3.12}$$

where f' is the derivative of f

The learning process has two activities through the network for each example:

• Prediction: calculate the value for the outputs of the layer. • Back-propagation: go backwards through the layers to update all of the weights of the network to reduce the error.

3.5 Decision Trees

Decision tree learning is one of the simplest useful techniques of predicting algorithms. We assume there is a single discrete target feature called the classification. Each component in the domain of the classification is a class. A decision tree or a classification tree is a tree in which

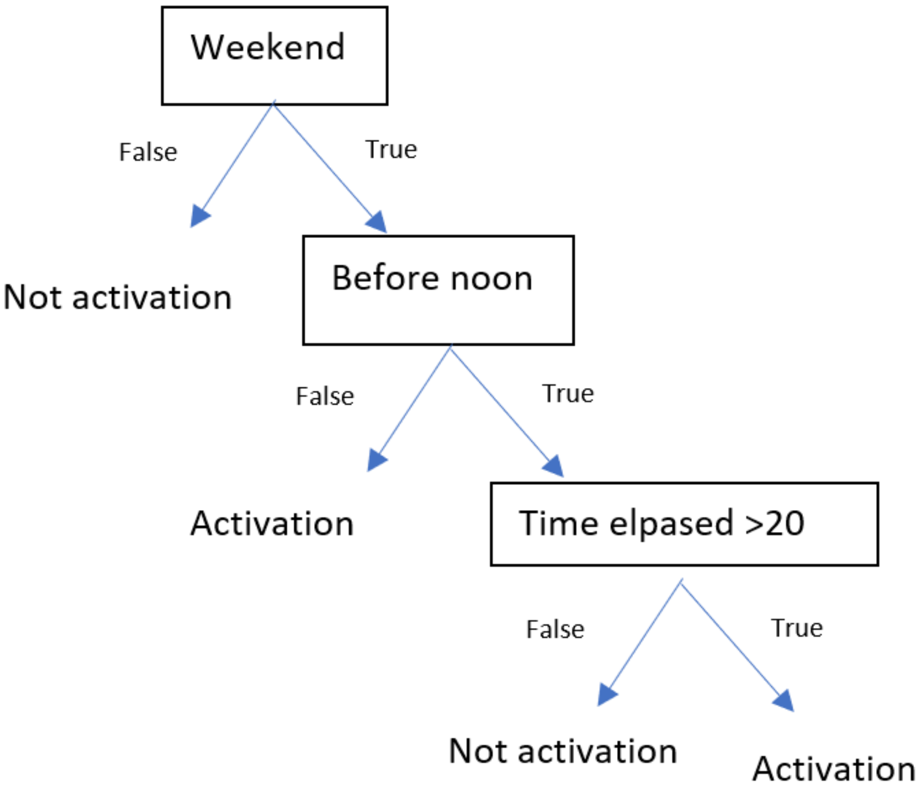

Figure 3.4: Decision Tree Example

In the decision tree the middle nodes, the branches, represent the solutions, and the leaves represent the predictions.

A decision tree is a tree where:

Each middle node (non-leaf) node is labeled with a condition. Each middle node has two children, true and false. Each leaf (prediction) of the tree is labeled with a estimation of the class.

The middle nodes (conditions) is evaluated and the arc corresponding to the result (true or false) is followed. When a leaf is reached, the classification corresponding to that leaf is returned. A decision tree can be represented as a nested if-then-else structure in coding. In the figure 3.4 we can see how to classify if there is going to be an activation of the washing machine or not. We have three conditions, based on the day, time and time elapsed since the last activation. For example if is not a weekend day inmediately is goig to classify like not activation.

Chapter 4

Adaptive user utility models

In this section as [6] proposes it will describe how to model the acceptance of rescheduling by flex-offers 2.2.2 to get user utility that is a combination of financial profits and user flexibility.

4.1 User utility

In this DR scheme, the reschedules appliances have to being approved by the users based on their requirements, otherwise the involved parties will lose financial benefits.The user flexibility is defined by the flex offers described in 2.1. Regarding to the users, they are free to make decisions based on their own profit energy, regardless other external factors such as other users, market, etc. The utility model is a combination between financial interests and device interests, as each user has different approaches, each user wiil have different utility function. Hence the function will incorporate the specific user-device pair, the time of schedule and the financial benefit. Based on these assumptions [6] proposes an user assumption: "Users are flexible in regards to their devices being rescheduled in return for financial benefits, as long as the schedule matches their preferred device behavior. Their acceptance of the device schedule is positively correlated with the financial interests and negatively correlated to the amount of delay by the schedule"

The user has the control of devices activity. In this context the scheduling of a device activation can be represented by a quid pro quo between financial benefits(higher flexibility) and device interests(lower flexibility). The User utility will be defined for a user u and device d, whose operation o starting at time t_{es} has been scheduled to time t, as:

$$E[U_{u,d}(t \mid t_{es},t_{le},A)] = G(A{=}a,t,t_{es}) \,.\, P(A{=}a \mid t,t_{es},t_{le}) + G(A{=}r,t,t_{es}) \,.\, P(A{=}r \mid t,t_{es},t_{le}) \quad (4.1)$$

where $G(A=a,t,t_{es})$ is the financial benefit when the new schedule is approved by the user, A=a, delaying the activation of o from t_{es} to t, while $G(A=r,t,t_{es})$ is the financial benefit by rejecting the new schedule, A=r. $P(A=r \mid t,t_{es},t_{le})$ is the probability that the user would either approve, P(A=a), or reject, P(A=r). the new schedule, with respect to the delay $t-t_{es}$ in the interval $[t_{es},t_{el}]$ and the user flexibility. The financial benefit $G(A=a,t,t_{es})$ = Price(0,t) - Price(o,t_{es}. For example if the new schedule is rejected $G(A=r,t,t_{es})$ = Price(0,t_{es}) - Price(o,t_{es}) = 0 Consequently 4.1 simplifies in

$$E[U_{u,d}(t \mid t_{es},t_{le},A)] = G(A=a,t,t_{es}) \,.\, P(A=a \mid t,t_{es},t_{le}) \quad (4.2)$$

4.2 User Flexibility

The user flexibility is the rate of approval of schedule by the user. The aim is understanding the user preference to maximize the probability $P(A=a \mid t,t_{es},t_{le})$ then the expected user utility.In [6] uses a model based on exponential distributions where it describes the probability distribution of the time intervals between events(ready actions) in a stochastic process where the events occur at a constant average rate. Let T be a random variable depicting the distance between two ready device(d) events, and $t-t_{es}$ be the delay on device activation o, dictated by the new schedule. Let us assume $t_{es} <= t <= t_{le} - \mid p \mid$, then $P(A=a \mid t,t_{es},t_{le})$ simplifies into $P(A=a \mid t,t_{es})$.

Therefore,

$P(A=r \mid t,t_{es}) = P(T<=t-t_{es})$,

"where $P(T<=t-t_{es})$ is the cumulative distribution function of the probability that a user will need the device ready before the proposed time t"[6]. Alternatively,

$P(A=r \mid t,t_{es}) = 1 - P(T<=t-t_{es})$, will be the probability of new schedule user acceptance. The figure 4.1 shows the exponential distribution of the time elapsed between two consecutive activations.

4.3 Estimation of User Flexibility

[6] proposes a data-driven user flexibility model, based just on the device-level activations. The aim is estimating the function $1 - P(T<= t - t_{es})$, therefore it has calculate the distribution of T, i.e. the distribution of time between two ready events.

Let T be the random variable that depicts the distance in hours(granularity) between two ready actions following an exponential distribution. Therefore the

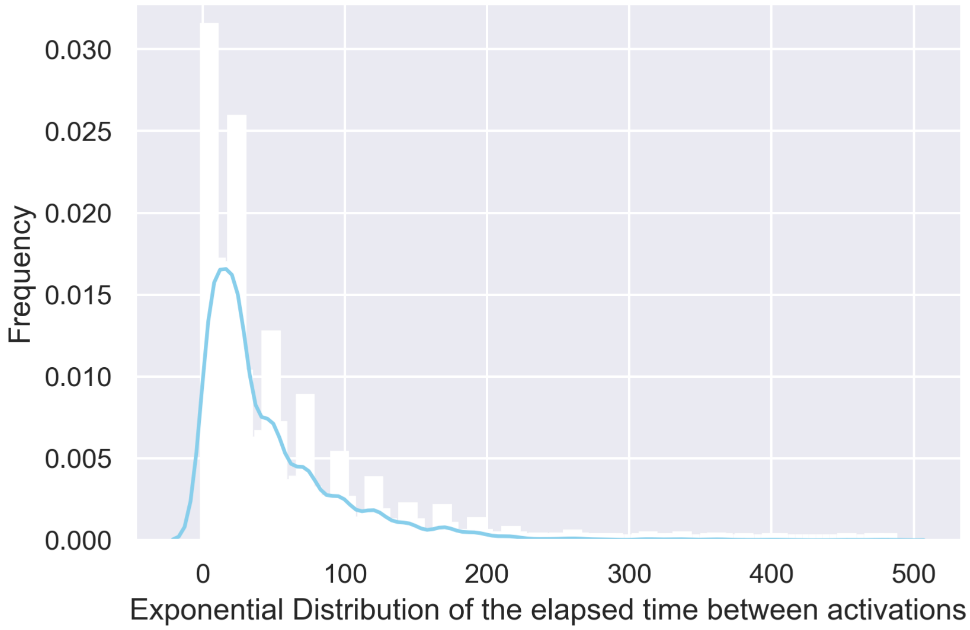

Figure 4.1: Exponential Distribution of Elapsed Time

cumulative distribution function of T is represented as:

F(t′,λ) $= \{1 - \lambda e^{-\lambda t'}, if t' >= 0$
$0, otherwise$
(4.3)

where t′ $\in$ T is a time interval, and λ is the rate parameter, describing the frequency of an interval of t′ that separates two activations. In this implementation λ parameter will be calculated by μ where

$$\lambda = \frac{1}{\mu} \tag{4.4}$$

μ will be calculated based on the historical data, but it will be updated by the new device-level load consumption data.

Chapter 5

Scheduling appliances

The User utility 5.2 is the basis to schedule flexible demand of appliances. The two factors used to calculate the user utility are:

- Financial gain from spot market
- User flexibility 4.2

The user utility depicts the quid pro quo between maximizing financial gain and minimizing loss of user-perceived quality of service, then presents an efficient schedule based on these two factors.

The demand scheduling depends on the financial benefit that the flexibility can produce from the spot market. This project will evaluate the savings that can be obtained by the spot market.

5.1 Spot Market Savings

Spot market savings are the financial benefits of energy demands and the corresponding flex offers at the spot market for the predicted device activation 2.2. To maximize this factor, an appliance activation is rescheduled so that the cost of the energy required for the appliance functioning is minimized. The hourly spot prices between the earliest start time and the latest start time is represented by

$$Spot = [o^s(te_e s), ..., o^s(te_l e)] \tag{5.1}$$

The total energy consumption cost of the device operation is the product of energy demand and the respective spot price represented as:

$$S(x) = \sum_{i=0}^{|p|-1} e_i . o^s(x+i) \tag{5.2}$$

where |p| is the duration in hours of the device functioning that begins at timestamp x, and e_i is the demand. Thus the savings by spot market rescheduling the device activation from $t_e s$ to t is:

$$\Delta S = S(t) - S(t_{es}) \tag{5.3}$$

$$= \sum_{i=0}^{|p|-1} e_i.o^s(t+i) - \sum_{i=0}^{|p|-1} e_i.o^s(t_{es}+i) \tag{5.4}$$

5.2 User utility by scheduling

The reschedule of a device activation o described by a probabilistic flex offer f maps the start time of activation t_{es} to a new timestamp t, producing a postponement of $t - t_{\mathrm{es}}$ in the device activation. The new timestamp will be added between the probabilistic time flexibility interval $[T_{\mathrm{es}}, T_{\mathrm{le}}]$ of the flex offer and the latest start time is delineated as $t_{\mathrm{ls}} = t_{\mathrm{le}} - |p|$. We will conclude the user utility definition described by 4.1 replacing $G(A, t, t_e s)$ with the spot market savings, of equation 5.4 , given by:

$$\varepsilon[U_{u,d}(t|t_{es}, t_{el}, A)] = (\Delta S).P(A = a|t, t_{es}, t_{el}) \tag{5.5}$$

Although, the scheduling depends on $P(t_{es}, t_{le}|e)$, i.e., the flexibility interval $[t_{es}, t_{le}]$ is correct. consequently, the objective function to schedule and operation o, considering the uncertainty of the flexibility intervals, is defined by:

$$\varepsilon[t] = \sum_{t_{es}, t_{el}} \varepsilon[U_{u,d}(t|t_{es}, t_{el}, A)].P(t_{es}, t_{le}|e) \tag{5.6}$$

where $P(t_{es}, t_{el}|e)$ is the probability of flex offer interval $[t_{es}, t_{le}]$.As T_{es} and T_{le} are are discret random variables. Then the 5.6 get the summation over the discrete values in $[t_{es}, t_{le}]$ rather than integration. Finally, the scheduling function for the operation o selects the t that maximizes the expected utility $\varepsilon[t]$ is given by:

$$Sched(o) = argmax\varepsilon[t]$$

Chapter 6

Experiments

It has been developed and tested several experiments to check the performance of the DR model. First it will be presented the flex offer scheduling evaluation. Second, it will be showed how the adaptive user flexibility can satisfy the user confort. And finally, we show how the prediction of probabilistic flex offers affect on the scheduling process.

6.1 Dataset

It was used the Dataport dataset described in 1.1, we selected 11 houses, because of the filtered data that contains. Each house has the consumption records of washing machine and dishwasher. The granularity selected of the time series is 1 hour. It means we will consider the demand at hourly resolution. Based on the information of activations per house, it was determined that the dataset is extremely unbalanced, with an average number of activations of 1.7%. Showing an average of 0.42 activations per day. For evaluation proposes, the dataset was splitted in 67% for training and 33% for test. For the first level prediction that utilizes Naive Bayes model to forecast the appliance activation for the next 24 hours, it is used the columns: day of the month, day of the week, hour, month and year. It was added a new features: the time that has passed since the last five activations until the current activation that are named HourDiff, HourDiff2, HourDiff3, HourDiff4 and HourDiff5; the mean time among the 5 last activations that is called HourMean; the mean among all the activations for house that is called HouseMean and the last feature is the standard deviation among the time elapsed of the activations for each house. The target is a binary class that contains 1 if there is an activation and 0 if there is not an activation within the 24 next hours, this new target increased the percentage of the value 1 in the target to 31% . I tried to tackled the problem of unbalanced dataset by oversampling, that means it will generate extra data from the minority class, to overcome its shortage of data. The Synthetic Minority over-

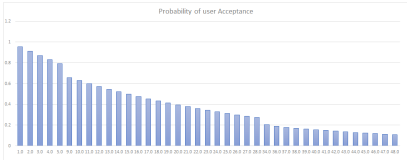

Figure 6.1: Probability of user acceptance

sampling Technique (SMOTE) [3] is one of the main methods to obtain this extra sample generation. It is based on generating examples on the lines connecting a point and one its K-nearest neighbors.

For the Financial evaluation, it was used the The Electric Reliability Council of Texas (ERCOT) that can get from Dataport, the same provider of the dataset consumption appliances.

For the device activation prediction, Neural Network was implemented for the first layer, to predict if will be there an activation the next 24 hours. And the second layer was the same as the first experiment with NB.

Evaluating device operation scheduling

The schedule's approval by the user and the utility from the financial spot market i given by the accuracy of the probabilistic flex-offer and the modeling of the use flexibility. As we described before as much as rigid is the user flexibility is lower the approval and vice versa. Therefore, it emphasizes that a rescheduling must perform an arrangement between user acceptance and financial benefits. In the figure **??** we can appreciate the probability of the user acceptance prediction related to the delay in hours. As we said before, as longer as the delay, the probability acceptance will decrease.

Based on the User acceptance predictions we could get saves of 13% from the spot market.

6.2 Evaluating activation device prediction

The efficacy of the flex offer prediction depends mainly on the accuracy of activity device prediction. To observe the features with biggest influence in the target, it was created the correlation matrix, that we can observe in 6.2,and we see that the most correlated features are Hourdiffs, means and the standard deviation. To predict the if there is an activation the next 24 hours (first layer), two models were evaluated, the first one the Naive Bayes model, and the second one was Neural Networks, using the library Scikit-learn and Keras from Python. The model was evaluated at dayly and hourly level. In spite of SMOTE was applied, the results for the first layer were not much better than the results without SMOTE, as is shown in the confusion matrix 6.3 Therefore, it decided working without over-sampling.

Additionally, in the second layer Linear Regression was evaluated.

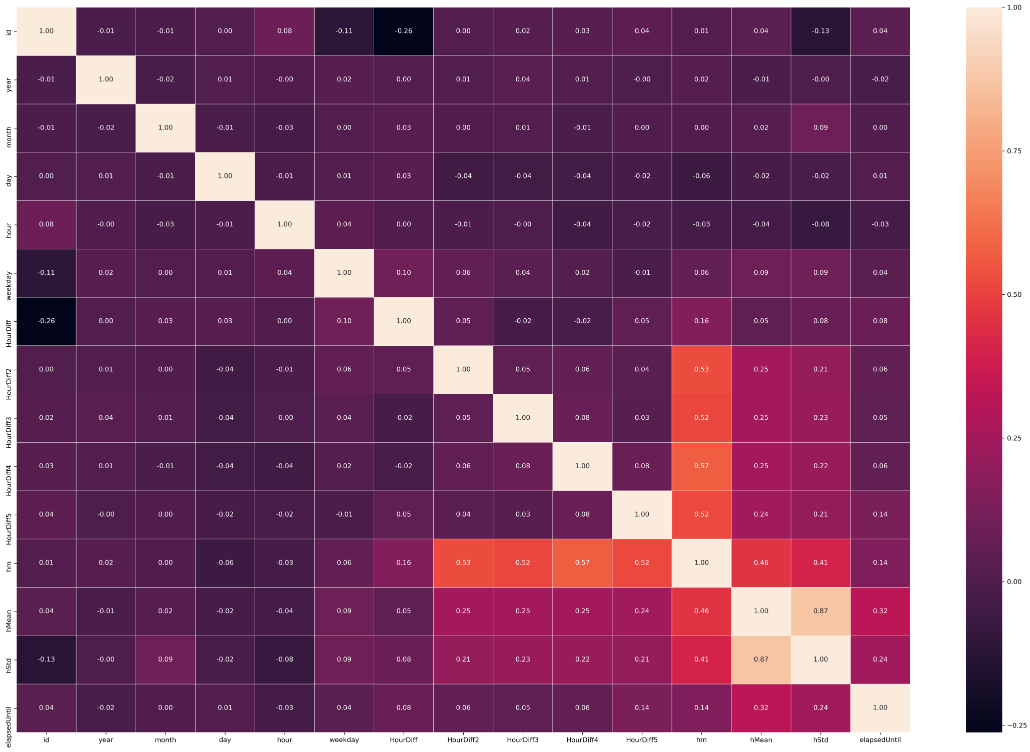

Figure 6.2: Features Correlation

The results are shown in the confusion matrix 6.5 where we can calculate the specificity that is 0.7 and the specificity that is 0.25, concluding that NN model is almost the double efective related to NB taking account the specificity. The results were verified by the loss in the test data vs the loss in the validation data shown in the 6.4 It was concluded that the model fits adequately.

Finally we can see the results in 6.6 the on the second layer, for both NB and NN the seconda layer with Linear Regression has the same setup. We can appreciate that after NN the second layer has a much better performance.

All the results were evaluated in Jupyter using Python, and is in the repository https://drive.google.com/drive/u/1/folders/1wzRoqi2F6IvOvTRmzqsonSlKXpvCTl9g

	Without SMOT		
	Actual Value		
Predicted	13695	18600	Sensitivity = 0.42
Value	7177	17876	Specificity = 0.71

	With SMOT		
	Actual Value		
Predicted	19836	20203	Sensitivity = 0.50
Value	14175	26013	Specificity = 0.65

Figure 6.3: SMOTE NB results

Figure 6.4: Train loss vs Test Loss

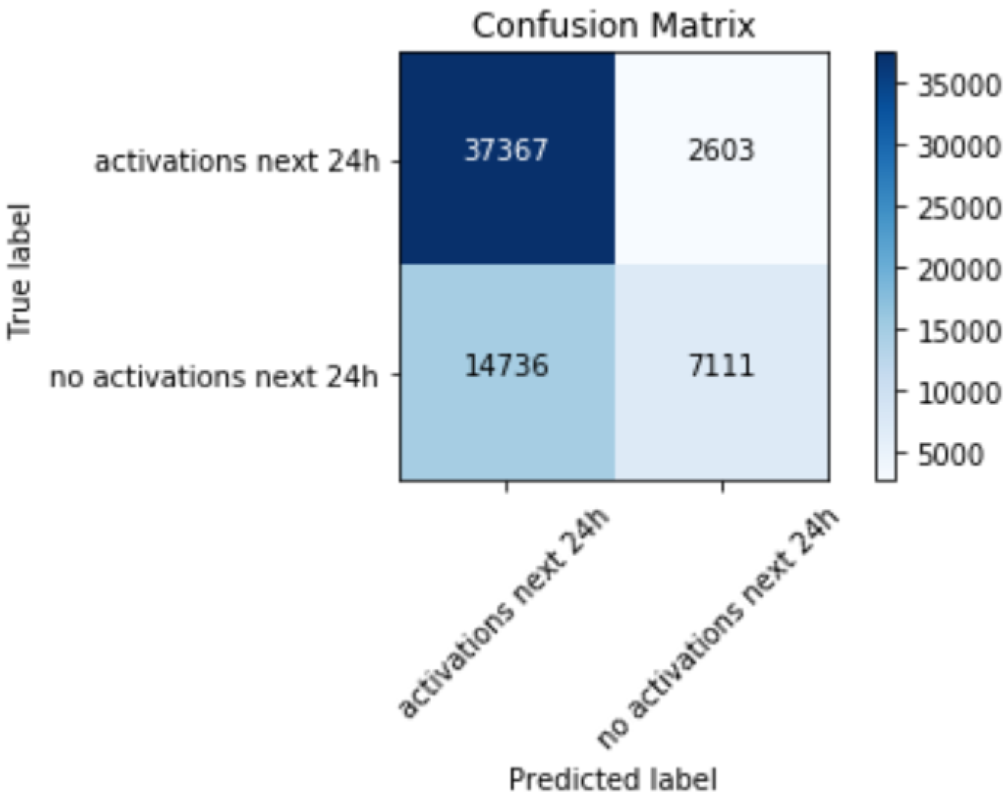

Figure 6.5: Confusion Matrix NN

Second Layer Linear Regression RMSE	
With NN in first layer	7.45
With NB in first layer	50.4281667

Figure 6.6: RMSE Second Layer Linear Regression

Chapter 7

Conclusions

In this project was presented a Demand Response model to predict the household appliances behavior. Several factors such as flexible flex offers, user utility and user flexiblity were analyzed to understand how we have to proceed to obtain a model to maximize the utility for consumers. The collected data was meticulously filtered and new features were created to achieve better results in the machine learning models tested. It is important to tackle the problem of imbalanced dataset as one of the most important features to get the best results is the prediction of activations. However, in some cases like this the oversampling method does not improve the results. It is important to review the models if they are overfitting, like the Neural Network. The best results that it can get are from Neural Networks in the first Layer. We conclude that if we use more advanced models such as NN, Convolutional NN, etc, we could get better results than Naive Bayes. To get better results is important the size of the dataset, in this project the number of rows and other features affected directly the different models performance.

Bibliography

[1] Hunt Allcott and Michael Greenstone. "Is There an Energy Efficiency Gap?" In: *Journal of Economic Perspectives* 26.1 (2012), pp. 3–28. DOI: 10.1257/jep.26.1.3. URL: http://www.aeaweb.org/articles?id=10.1257/jep.26.1.3.

[2] M. Amer et al. "Smart home energy management systems survey". In: *International Conference on Renewable Energies for Developing Countries 2014*. 2014, pp. 167–173. DOI: 10.1109/REDEC.2014.7038551.

[3] H. Cao et al. "Integrated Oversampling for Imbalanced Time Series Classification". In: *IEEE Transactions on Knowledge and Data Engineering* 25.12 (2013), pp. 2809–2822. ISSN: 2326-3865. DOI: 10.1109/TKDE.2013.37.

[4] A. Druckman and T. Jackson. "Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model". In: *Energy Policy* 36.8 (2008), pp. 3177 –3192. ISSN: 0301-4215. DOI: https://doi.org/10.1016/j.enpol.2008.03.021. URL: http://www.sciencedirect.com/science/article/pii/S0301421508001559.

[5] Filipe Fernandes et al. "Dynamic load management in a smart home to participate in demand response events". In: *Energy and Buildings* 82 (2014), pp. 592 –606. ISSN: 0378-7788. DOI: https://doi.org/10.1016/j.enbuild.2014.07.067. URL: http://www.sciencedirect.com/science/article/pii/S0378778814006161.

[6] Davide Frazzetto et al. "Adaptive User-Oriented Direct Load-Control of Residential Flexible Devices". In: *CoRR* abs/1805.05470 (2018). arXiv: 1805.05470. URL: http://arxiv.org/abs/1805.05470.

[7] S. Kotsiantis. "Supervised Machine Learning: A Review of Classification Techniques". In: *Informatica (Slovenia)* 31 (2007), pp. 20–268.

[8] M. Manic et al. "Building Energy Management Systems: The Age of Intelligent and Adaptive Buildings". In: *IEEE Industrial Electronics Magazine* 10.1 (2016), pp. 25–39. ISSN: 1941-0115. DOI: 10.1109/MIE.2015.2513749.

[9] T. B. Pedersen, L. Šikšnys, and B. Neupane. "Modeling and Managing Energy Flexibility Using FlexOffers". In: *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. 2018, pp. 1–7. DOI: 10.1109/SmartGridComm.2018.8587605.

[10] David Poole and Alan Mackworth. *Artificial Intelligence: Foundations of Computational Agents*. 2nd ed. Cambridge, UK: Cambridge University Press, 2017. ISBN: 978-0-521-51900-7. URL: http://artint.info/2e/html/ArtInt2e.html.

[11] M. Shafie-Khah and P. Siano. "A Stochastic Home Energy Management System Considering Satisfaction Cost and Response Fatigue". In: *IEEE Transactions on Industrial Informatics* 14.2 (2018), pp. 629–638. ISSN: 1941-0050. DOI: 10.1109/TII.2017.2728803.

[12] J. Stragier, L. Hauttekeete, and L. De Marez. "Introducing Smart grids in residential contexts: Consumers' perception of smart household appliances". In: *2010 IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply*. 2010, pp. 135–142. DOI: 10.1109/CITRES.2010.5619864.

[13] Jagruti Thakur and Basab Chakraborty. "Demand side management in developing nations: A mitigating tool for energy imbalance and peak load management". In: *Energy* 114 (2016), pp. 895 –912. ISSN: 0360-5442. DOI: https://doi.org/10.1016/j.energy.2016.08.030. URL: http://www.sciencedirect.com/science/article/pii/S0360544216311331.