Forecasting Day-Ahead Energy Prices

by

Didier Iyamuremye

A master's thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Data Engineering.

Prof.Stefan Kettemann
Supervisor

Prof.Dr.Ing.Hendro Wicaksono
Co-reviewer

Date of Master's thesis submission: August 2, 2020

—Jacobs University Bremen —

Statutory Declaration

Family Name,Given/First Name	Iyamuremye, Didier
Matriculation number	30002209
What kind of thesis are you submitting: Bachelor-, Master- or PhD-Thesis	Master

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

This document was neither presented to any other examination board nor has it been published.

German: Erklaerung der Autorenschaft (Urheberschaft)

Ich erklaere hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprueft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukuenftigen Arbeiten verglichen werden zu koennen. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfaeltigung.

Diese Arbeit wurde noch keiner anderen Pruefungsbehoerde vorgelegt noch wurde sie bisher veroeffentlicht.

Signature

02.08.2020
Place, Date

Abstract

Real world of now is well known to be very interested in what should happen tomorrow,especially in technological aspect, as of it's advancing every day. Forecasting day ahead energy price is one of the useful way of reducing the drama of how the tomorrow energy price should be by the help of past and today's prices. Guessing using human's brain not enough,it seemed to have shown the limitations, where thinkers came with the idea of using advanced technologies in the help of computers for big data analysis by saving time and to the extend of what the nature of human's capacity can do. This master's thesis presents a comparison of performance of different techniques applied to daily time series data to develop better strategies to treat seasonality and trends with respect to error field amplitude, error correlation matrix and heat maps to detect seasonal patterns in order to improve model performance for machine learning algorithms or advanced deep learning neural network to automatically used for complex mapping from inputs to outputs forecast.

There are several tools and models found in machine learning and deep learning which are promising to be used to forecast a certain time series; however, it is really not always clear which model is appropriate for selection, depending on data the models are suited,main key is to understand the nature of data you have and put it into suitable form for analysis.

The aim of this masters thesis research is to try different machine/deep learning models and test around with their algorithms using well known statistical tools to know its behaviours and be able to differentiate which model does well in the performance of forecasting daily step ahead of energy price specifically in French energy market. Models performed well and give very promising results,The models implemented in this study are convolutional neural network,Seasonal auto regression and moving average,Vector regression,Vector auto regression and moving average,Recurrent neural network which act on sequences of inputs and long and short term memory model.

Keywords:Seasonal auto regression and moving average,Vector auto regression,Convolutional neural network,Recurrent neural network,French energy market,time series forecasting,deep learning,machine learning.

Contents

1 Chapter 1

Introduction

Machine learning and deep learning are nowadays on top in big data of petabytes analysis in our world, starting from easy and heavy data. Roughly saying modern computers for today, have equipped humans with talent and cleverness to process information in ways unimaginable when they were first widely profitable a few decades ago. Indeed, even today's high technologies specifically in using advanced models to predict by the target of the future outcomes.

In this masters thesis project, to be able to get the data to analyze in order to achieve my goal in thesis, I have used the open source which is the platform designed to provide data to use for analysis, I have managed to download all the specific data for France energy market from the year range of 2018 to april 2020. Then I have targeted only 6 features from the entire data in the wind and solar energy area which are: The actual total load, day ahead price, day-ahead total load forecast, Generation-solar day ahead and Generation-wind onshore day ahead to use according to the main purpose of the project of applying vulnerable machine learning and deep learning models to do forecasting of day-ahead prices.

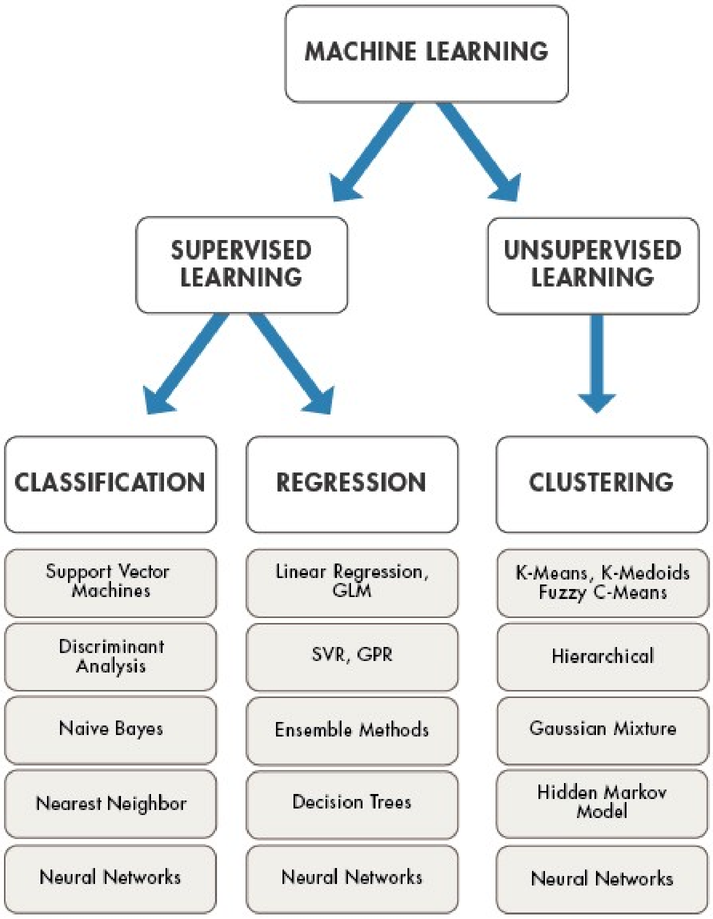


Figure 1: Machine Learning Algorithm [1]

I have used Machine Learning concepts in my master thesis, The figure above,Figure 1, shows the summary of modelling using machine learning.The algorithm is obvious accordingly, but I want to be more specific for my data, I have supervised learning data,with regression whereby I have used machine learning and deep learning models to forecast the day-ahead price of French Energy Market in energy translation. As I have mentioned from above, I used regression models for my scenario I am having supervised learning which are ARIMA,SARIMA,SARIMAX,VAR for machine learning, and RNNs,CNNs and long and short term memory for deep learning.

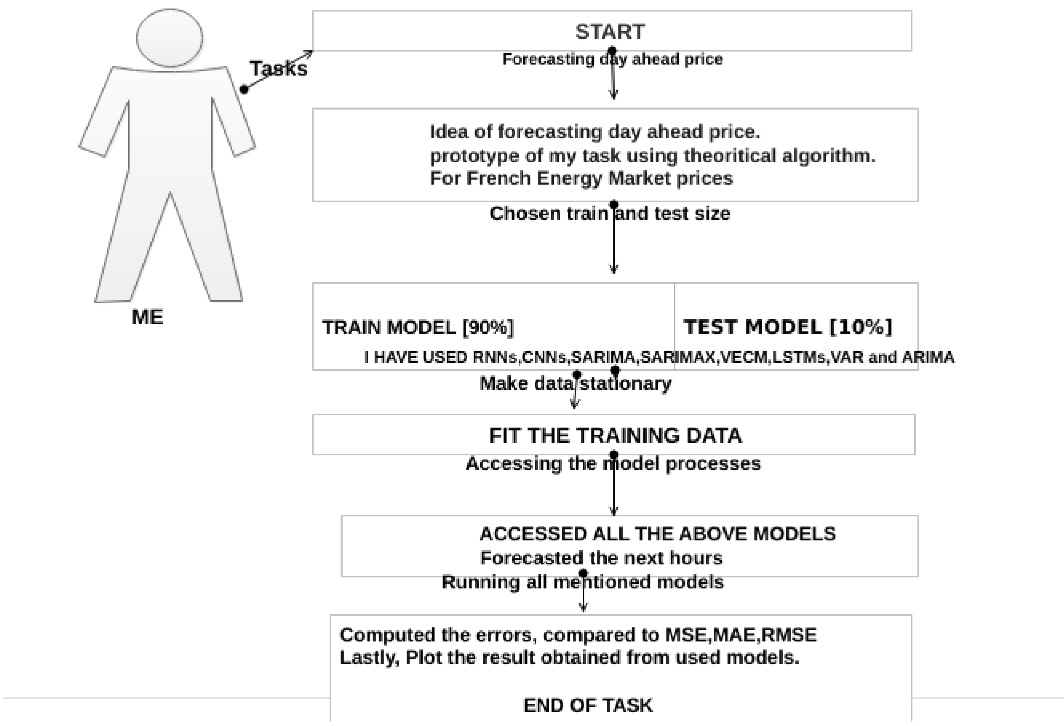


Figure 2: Machine learning and deep learning algorithm

Machine Learning defined as a type of data analysis that uses algorithms that learn from data. And also taken as a study of algorithms and statistical models that computer systems use to perform a specific task without using any explicit instructions, but relying on patterns and interested in inference or prediction instead.[2] Machine Learning is a subset and a blanch of artificial intelligence based on the idea that systems can learn from data, identify patterns and also make decisions with less human intervention and is typically categorized as supervised or unsupervised learning. For supervised learning technique you have to train the machine using data which is well labeled where the output data patterns are known to the system,for the supervised learning algorithms we have different models like linear regression,k nearest neighbour,Random Forest etc. But,In contrast, unsupervised learning technique is another machine learning technique where in this case you do not need to supervise the model. It helps us to find all kind of unknown patterns in data, it uses different methods as well to do clustering like Convolutional neural network,auto vector regression,k-means clustering algorithms,is known to be a complex method compared to supervised learning.[3]

1.1 French energy market

This image below is showing the French Energy Market, energy consumption and trade balance;in 2012, France's primary energy consumption (PEC) reached 259 Mtoe. More than 40 percent of gross consumed is derived from nuclear power. Figure is showing the percentage of different energies in French energy market. where the renewable energies amounted to 8.0 percent of the total energies produced in 2012. figure(3) [4]

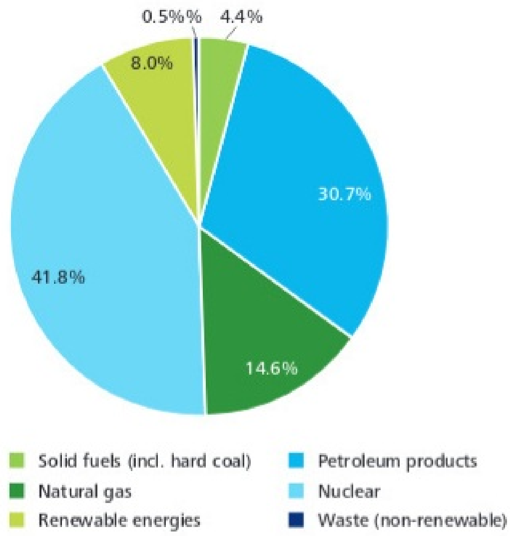


Figure 3: French energy market

1.2 Research question and motivation

Starting from the previous projects I did under supervision of Prof.Stefan Kettemann, Firstly, the first advanced project only made focus on heat maps and doing correlation analysis among wind and solar power for French Energy Market. For the second advanced project, I only used 2 different ML models by applying them on the same data I had used for preceding advanced project,the model performance was good for predicting total load by using VAR and day ahead price using ARIMA model. But not really perfect and I had mind to raise up the performance of the models in forecasting anything but here I mean forecasting day ahead price in French Energy. But this master research became motivated to use other different models to improve result performance of models in order to get a better accuracy of the results than the previous ones for advanced project II, I have done all models required in my master thesis and performed very good to my data of forecasting . Energy for French Energy Market, I only used 2 machine learning models which are ARIMA and VAR models for prediction . BUT now the target of this reasearch was to go beyond the models I have used to predict French Energy Market, and used other recent models to do the forecasting of day ahead price for again French Energy Market. Those machine and deep models are SARIMAX, VARIMAX, CNNs, RNNs and LSTMs.

My movitation came actually after finishing advanced project II, whereby I had a projection of doing forecasting using more than two models I used in previous advanced academic project, I was eager to know how the other machine learning and deep learning techniques can perform besides ARIMA and VAR models.

1.3 Problem Statement of project

There is a trend and seasonality problem in the time series of data we have which causes bad results in forecasting. I will use different machine learning techniques such as ARIMA,SARIMAX,VAR and deep learning algorithms such as Recurrent Neural Networks and convolutional Neural Networks. I also use heat maps first for data visualization and to detect seasonal patterns and use it to improve those machine learning and deep learning models to solve the issue and get a better forecast result.

1.4 Research Purpose and objectives

The main research purpose was to do method development by trying several machine learning and deep learning techniques to be able to get a good accuracy in daily ahead price forecasting for French Energy Market. By developing the better strategies to Treat Seasonality, Trends. I have to optimise ARIMA and VAR models with respect to error field amplitude and error correlation matrix by using ARIMAX and regARIMA models.

In addition to that, in order to develop better strategies to treat seasonality and trends, I have used daily averaged time series as additional features to those ones I have downloaded from open source provider for data to use for analysis.

In order to be able to detect seasonal patterns along side my data, I have used heat maps in order to easily see the gaps from the data and then be able to get ideas to improve the performance of my forecasting accuracy. The heat maps i have used, it has some missing patterns in their heat map results. Furthermore, within the objectives of knowing very well how deep learning and other machine learning models besides Auto Regression and Moving Average (ARIMA) and Vector Auto Regression (VAR) on different data in the help of forecasting anything but specifically Forecasting Day-Ahead Energy Prices in French Energy Market.

I have used convolutional NN for time series analysis and then forecast the daily ahead price but by each hourly ahead after doing auto arima model.

Last but not least in my objectives I also used Convolutional Neural Networks and Recurrent Neural Networks to forecast day ahead price by careful considering the correlation analysis and I have added also seasonal time series by averaging over daily. as additional features that are used as input into Recurrent Neural Networks and Convolutional Neural Networks

1.5 Thesis overview

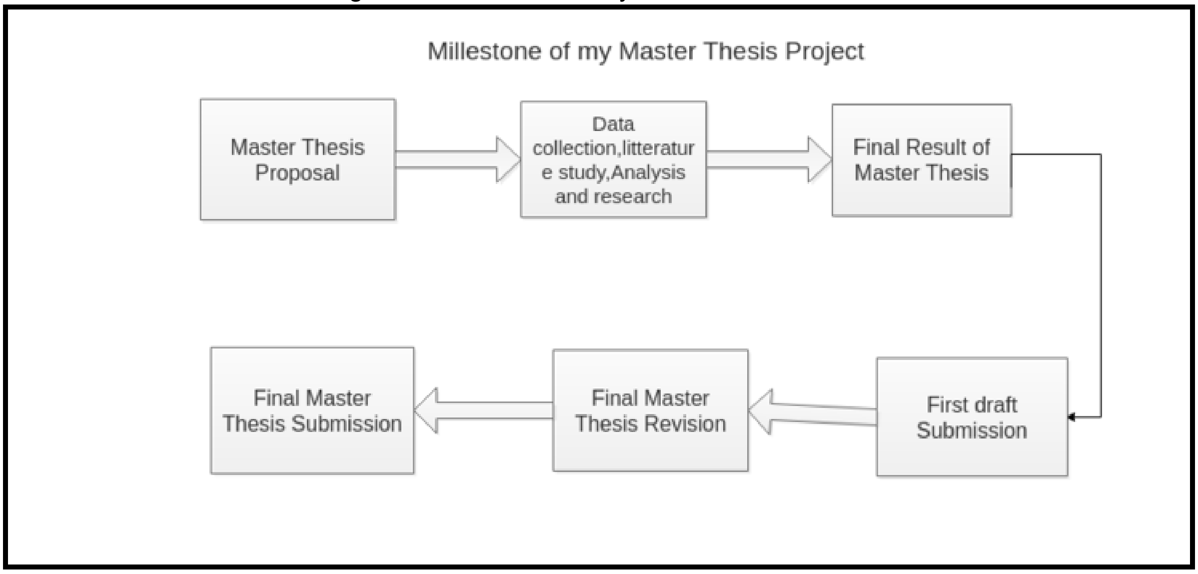
We have reviewed so many various notions related to forecasting day ahead prices for French Energy Market using various models either machine learning or deep learning techniques. this master thesis project is mainly structured as follows:

It starts with chapter 1 which presents all the necessarily information regarding the task of project in theory and mathematical implementations. Chapter 2, gives the results I got from all models I have used. Chapter 3 introduces the discussion and conclusions about the project ,Chapter 4 proposed the Future work and lastly in chapter 5,the conclusion of this master thesis belongs in chapter 5. the limit of the current work are summarized and future reflection are outlined.

1.6 Millestone of my Master Thesis

During my master thesis , I have followed lifecycle methodology to finish the project, that's the reason why I am showing all the ways I have passed through to be able to finish the project. This is the millestone of my journey till the last step of my project of forecasting day ahead price using machine learning and deep learning techniques.(figure 4)

Figure 4: Millestone of my Master Thesis.



1.7 Used sources

This master thesis partially uses results reported previously in my independent advanced projects. Some for Chapter 2 and Chapter 3 are from my independent advanced projects [5] completed in Fall session 2019.

Others are the contributions of the authors that I have documented to be able to do theoritical framework and coding part I have used research from different online sites related to what I was eager to do. In coding side of mathematical formulas, the experiments of this master thesis, we use different tools, then for coding I have used Jupyter notebook,google colab to be able to get results of the models. I have used ,and for this scientific report, I have used LaTeX through my sublime text editor though LaTeX has its own set up. I have used different libraries such as Keras and inside keras we have so many other libraries like ,TensorFlow,etc.

Those software package and tools are used for collection of python modules with the purpose of helping anyone who is interested to do forecasting or other purposes relevant to what the software does.

1.8 Delimitation

I had the task to develop better strategies to treat seasonality and trends, by doing performance of different models in several ways either applying machine learning techniques or deep learning models.

I had first trying machine learning ARIMAX,SARIMAX in machine learning models and VAR. However, I have used them but I had also a task to find out which better way can treat seasonal and trends I had in my data in orderto get a perfect performance. I had chosen to use convolutional neural network (CNN), Recurrent Neural Networks(RNNs) and Long and Short Term Memory (LSTMs).

Due to the fact that, after plotting heat maps I was supposed to use for helping me to detect the missing patterns in the result in order to easily use it to perform again on models to increase the performance of the previous model results.

Data Description

During this stage of knowing the data to be selected, I have managed to get the 6 data attributes which are :

1. Day-ahead Total Load Forecast
2. Actual Total Load
3. cet-timestamp
4. Day-ahead-Price
5. Generation-wind onshore Day ahead
6. Generation-solar Day ahead
7. Daily average price.

But over the open source platform was allowed to select whatever data you want to use,depending on the task you have. And there is an option to download the data in any available formats it has, for my side, I have chosen csv format for all datasets.

Table 1: Data Description table.

Data Attributes	Data Description	Data types	Null values
Day-ahead Total Load Forecast	total load forecast in MH/h	float64	20
Actual Total Load	actual total load in MW/h	float64	4
cet-timestamp	time format :DD.MM.YY HH:MM	object	0
Day-ahead-Price	day ahead price in EUR	float64	4
Generation-wind onshore Day ahead	wind onshore day ahead in MW/h	float64	4
Generation-solar Day ahead	solar day ahead in MW/h	float64	20
Daily average price	daily average price in EUR	float64	20

Preprocessing Data

Data processing is a data mining technique that involves transforming raw data into an understandable format. By the fact that real-world data is often incomplete and inconsistent or lacking in certain trends or behaviors and is likely to contain many errors. so from this stage of preprocessing the data,after downloading the data sets via an open source platform[6] ,generally the data was not consistent either by its time series formats or even missing values,smooth noisy data it was required to remove some outliers and then resolve inconsistencies before starting processing data in order to try machine learning methods. I did data cleaning by filling the missing values,identity and remove the outliers,even because the format of Time series was string it was not easy to manage them,in addition Machine Learning has steps for Data Proprocessing.

1. Import Libraries
2. Importing the Dataset
3. Taking care of the missing data in dataset
4. Encoding categorial data
5. Splitting the dataset into training and test set
6. Feature extraction/scaling

Then from there apply the Machine learning models to the data. Additionally, I have first changed its format to timestamp as shown in the table below.
I have shown only 2 features among all other features I have in the dataset, the time shown in

Table 2: Original Data Format table.

MTU(CET)	Day-ahead Price [EUR/MWh]	Actual Total Load [MW]	Data types
01.01.2018 00:00-01.01.2018 01:00	132	937	csv
01.01.2018 01:00-01.01.2018 02:00	809	1045	csv
01.01.2018 02:00-01.01.2018 03:00	1622	1055	csv
01.01.2018 03:00-01.01.2018 04:00	2154	1065	csv
01.01.2018 04:00-01.01.2018 05:00	2318	1076	csv

table 2 has the string format, and there was much time redundancy or duplicates which was not necessary which also caused some problems on the stage of data visualization for forecasting. I have shown only 2 different categories of renewable energies among the one I have downloaded, 5 first features range time with corresponding amounts of Actual Total Load [MW], day ahead price [EUR/MWh] with also 5 amounts of data in megawatts during the year 2018.

Table 3: New Data Format Table.

MTU(CET)	Day-ahead Price [EUR/MWh]	Actual Total Load [MW]	Data types
01.01.2018 01:00	132	937	csv
01.01.2018 02:00	809	1045	csv
01.01.2018 03:00	1622	1055	csv
01.01.2018 04:00	2154	1065	csv
01.01.2018 05:00	2318	1076	csv

This table 3 shows the new data I have used after changing the time format from string to timestamp, I have shown only two different amounts which are generation solar in mega watts and generation wind onshore in mega watts with the csv data formats during the year of 2018. Gathering and processing data which are unstructured data into a suitable format for further steps to be able to do time series forecasting depend on the task and purpose in general.

1.9 Time Series Analysis

Before explaining time series analysis, let me ask myself why we need time series in forecasting the daily ahead price. I would like to first explain how the data behaves with time series, while we do download of data over that open source I have mentioned from above [6]. we choose whatever data we want, with different data format but comes with series of time depend on the chosen years. Time series analysis is the analysis we do in order to be able to know the exact time we have to use either for training, testing or validating the data.

For time series we have seasonality,trend and noise that you might also consider to be able to explore the models on the data chosen.

1.9.1 Stationary and non-stationary time series

For applying the model you must be sure that the time series of the data you have are stationary otherwise you have to use differentiation concepts to change non-stationary data into stationary. The main reason is that while the time series of data is stationary it would be easy to continue for further steps like model evaluation.[7]

1.9.2 Seasonal decomposition

SARIMA defined as Seasonal Auto Regression and Moving Average Model but in which it makes focus on seasonality that the data has.we consider very well the seasonalities which is shown in time series forecasting. Generally, we have additive and multiplicative seasonal decomposition with trend, seasonal and residual for time series analysis. Actually saying seasonality, I am talking about periodic movement.[5]

Data Acquisition

The data I have used to do analysis of my project are found from entsoe transparency platform. [6] . Briefly entsoe transparency platform is an open source platform defined as a central collection and publication of electricity generation,consumption and transportation data but only limited to the pan-European countries. However, It requires to create an account in order to be able to export the data you want in different data formats such as csv,xlsx,xml, such data are for day ahead, total load, price for wind and solar generated within the different years but for this master thesis I have selected from 2018- till april 2020, I was able to export in different ways just depends on which data you want. I have only selected 5 features which are :

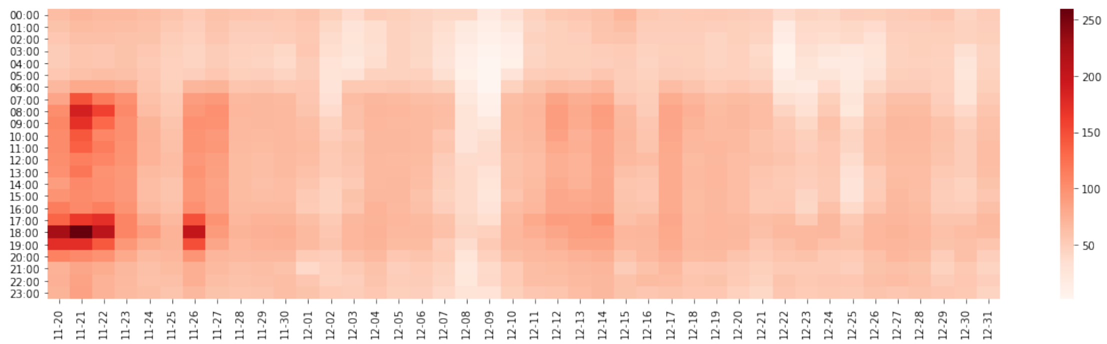
1. Day ahead price
2. Generation Forecast -Day ahead
3. Day-ahead total load forecast
4. Actual total load
5. aily average of day ahead price

But this one I had to use mathematical formula to find it from day ahead price data. The mathematical formula is as follows:

1.10 Heat Map results

Heat map is generally defined as one way of representing data but in the form of a map or diagram in which data values are represented by colors during a period of time.

Figure 5: Heat map result.



Theoritical framework

1.11 Data science

1.11.1 Time series forecasting

During the time we access the models, we use time series in order to be able to do forecasting in general.It is very important in machine learning to consider the time series during the forecasting process because you know the specific time to use for past values in order to do forecasting. Time series is the time range taken as past values in order to be able to do prediction of future values while you are using different models. [7]

1.11.2 Optimal parameters

optimal parameters are defined as model parameters that are supposed to be optimized.In machine learning, hyperparameter optimization or tuning is the problem of choosing a set of optimal hyperparameters or a learning algorithm. Whereby hyperparameter is a parameter whose value is used to control the learning process.[8]

1.11.3 Strategies for model improvement

I would say that there are so many ways to use models to improve the performance, but again as it is not that easy to confirm that this or that model is very good in result of forecasting a day ahead price or in other forecasting purposes. It can Only be confirmed depending on the accuray results. However, there is no guarantee that it's the first better model that can perform well to the scenario of data for analysis.
As everyone uses his own strategy to get the final result in forecasting anything in different areas, for me specifically I had a task to improve the model performance using model development to forecast a day head price. For instance, before trying modelling using machine learning or deep learning, it is very recommended to first understand the behaviours/nature of your data if it's supervised, unsupervised or reinforcement learning in order to know easily where to start by doing the model performance on data set.
But mostly we have to find a way to develop the strategies to treat seasonality and trends for scenario we have. For me personally, after managing to get the 5 features to use, I have also added a feature of daily average time series for the day ahead price as additional features. Both machine learning and deep learning,I have used different techniques by testing and trying with their parameters,using auto arima for machine learning but for deep learning I have used with optimal paramaters to be able to get the best accuracy for the day ahead forecast.

1.12 Machine learning models for day-ahead price

1.12.1 Auto regression and moving average

Auto Regressive Integrated Moving Average is one of the machine learning model which is abbreviated as 'ARIMA' which is actually a class of models that explains a given time series based on its own past values,roughly meaning that model uses lags and the lagged forecast errors so that the equation can be used to forecast future values.[3]

Mathematically,an ARIMA model can be created using the statistical model library as follows:
1.Define the model by calling ARIMA() function and then passing in the p, d, and q parameters.
2.The ARIMA model is prepared on the training data by calling the fit() function.
3.Predictions can be made by calling the predict() function and specifying the time to be predicted or the index of time.

An ARIMA model is characterized by 3 terms: p, d, q where:
p is the order of the AutoRegressive term
q is the order of the Moving Average term
d is the number of differencing required to make the time series stationary

Then,If a time series, has seasonal patterns, then you need to add seasonal terms and it becomes SARIMA, short for 'Seasonal ARIMA'. More on that once we finish ARIMA. The mathematical formula for the AutoRegressive and Moving Average models,first of all a pure Auto Regressive model is one where Yt depends only on its own lags,means that Yt is a function of the lags of Yt. [3]

$$\mathbf{Y(t)}=\alpha+\beta 1Y(t-1)+\beta 2Y(t-2)+..+\beta pY(t-p)+\theta 1\epsilon(t-1)+\theta 2\epsilon(t-2)+..+\theta q\epsilon(t-q) \quad (1)$$

where, $Y(t-1)$ is the lag1 of the series, $\beta 1$ is the coefficient of lag1 that the model estimates and α is the intercept term, also estimated by the model.

Likewise, again for a pure Moving Average (MA only) model is one where Yt depends only on the lagged forecast errors. The formula above is the final equation of ARIMA model by combining AR and MA models. ARIMA model in full detail, Predicted Y(t) = Constant + Linear combination Lags of Y (up to p lags) + Linear Combination of Lagged forecast errors (up to q lags).

The objective, therefore, is to identify the values of p, d and q. As I said before d stands for differenciation, so if the time series is not stationary you have to apply differencing in order to make it stationary, my actual entire data was stationary(DF test results find it on page 12) but because ARIMA model need to split the data into small portion of recent observations and after doing that, the new observation data becomes non-stationary, I have done the stationarity process to make it stationary otherwise it would have affected the model parameters. From the time series where is stationary, the value of d=0, but once you split the data into small portion of recent observations and data becomes non-stationary, if you make the time series stationary again,you get the value of p,d,q and the value of d will also change,from there you get everything needed to find the ARIMA model where you only have to pass the ARIMA() function the order which has the values of p,d,q and gives you the observations of the Akaike Information Critera (AIC) which is a widely used as a measure of a statistical model. Also it basically quantifies the goodness of fit and the simplicity of the model into a single statistic, Bayesian Information Criteria(BIC). [9]

Steps for the methods used to check if the time series of the data is stationary:
By checking the series if is stationary we have to make sure that variance,mean and covariance are constant with time if they are changing with time that means that series is not stationary. Then you might follow these steps to make the series stationary. Mathematically, we can use statistical test properties to check if a given series are not constant with time, which is the condition for stationary time series. Suppose we have a time series:

$$\textbf{Y(t)} = a * Y(t-1) + \epsilon(t)$$

(2)

where $Y(t)$ is the value at the time instant t and $\epsilon(t)$ is the error term. In order to calculate $Y(t)$ we need the value of $Y(t-1)$ which is :

$$\textbf{Y(t-1)} = a * Y(t-2) + \epsilon(t-1)$$

(3)
Then, if we do that for all observations, the value of $Y(t)$ will come out to be :

$$\textbf{Y(t)} = a^n * Y(t-n) + \sum_{t=1}^{n}(\epsilon(t-i) * a^i)$$

(4)

If the value of a is 1 which is a unit in the above equation, then the predictions will be equal to the $Y(t-n)$ and sum of all erros from t-n to t, which means that the variance will increase with time. This is known as unit root in a time series. We know that for a stationary time series, the variance must not be a function of time. The unit root tests check the presence of unit root in the series by just checking if the value of a=1. Below are the two of the most commonly used unit root stationary tests:[10] **ADF(Augmented Dickey Fuller)**Test. The Augmented Dickey Fuller test is one of the most popular statistical tests. Which can be used to determine the presence of unit root in the series, and hence help us undestand if the series is stationary or not. The null and alternate hypothesis of this test are follows: Null Hypothesis: The series has a unit root with the value of a=1
Alternative Hypothesis: The series has no unit root

Additionally, if we fail to reject the null hypothesis, we can immediately say that the series is non-stationary, this tells us that the series can be linear or difference stationary.
Test for stationary: If the test statistic is less than the critical value, we can reject the null hypothesis means that the series is stationary. When the test statistic is greater than the critical value, we fail to reject the null hypthoses which means that the series is not stationary.

Generally, we have different types of stationarity,which are follows:

1.Strict Stationary: this strict stationary series satisfies the mathematical definition of a stationary process. Then for a strict stationary series the mean,covariance and variance are not the function of time. The main aim is to convert a non-stationary series into a strict stationary series for making predictions.
2.Trend Stationary: It's a series that has no unit root but displays a trend is referred to as a trend stationary series. Once the trend is removed, the resulting series will be strict stationary.
3.Difference Stationary: This type of stationary it's a time series that can be made strict stationary by differencing falls under difference stationary. ADF test is also known as a difference stationarity test according to what i have explained early.
Differencing as one way used to make series stationary, mathematically it can be written as:

$$\textbf{Y(t)} = Y(t) - Y(t-1)$$

(5)
Where $Y(t)$ is the value at a time t

But for seasonal Differencing, instead of calculating the difference between consecutive values, we calculate the difference between an observation and a previous observation from the same season, mathematically it can be written as:

$$\textbf{Y(t)} = Y(t) - Y(t-n)$$

(6)

where $Y(t)$ is the value at a time t for previous observation, and $Y(t-n)$ is the value at a time t and n(number of observations) for an observation.

For figure[7], I have used the C[-1,1] as the correlation function with the intervals between -1 and 1 on the y axis and the day ahead price[EUR/Mwh] data but here shows the size of the laf between the elements of the time series on the x axis,AutoCorrelation Function(ACF) is the residual used to determine ARIMA model selection.

Figure 6: Differencing Recent Observation for Day-ahead Price [EUR/MWh].

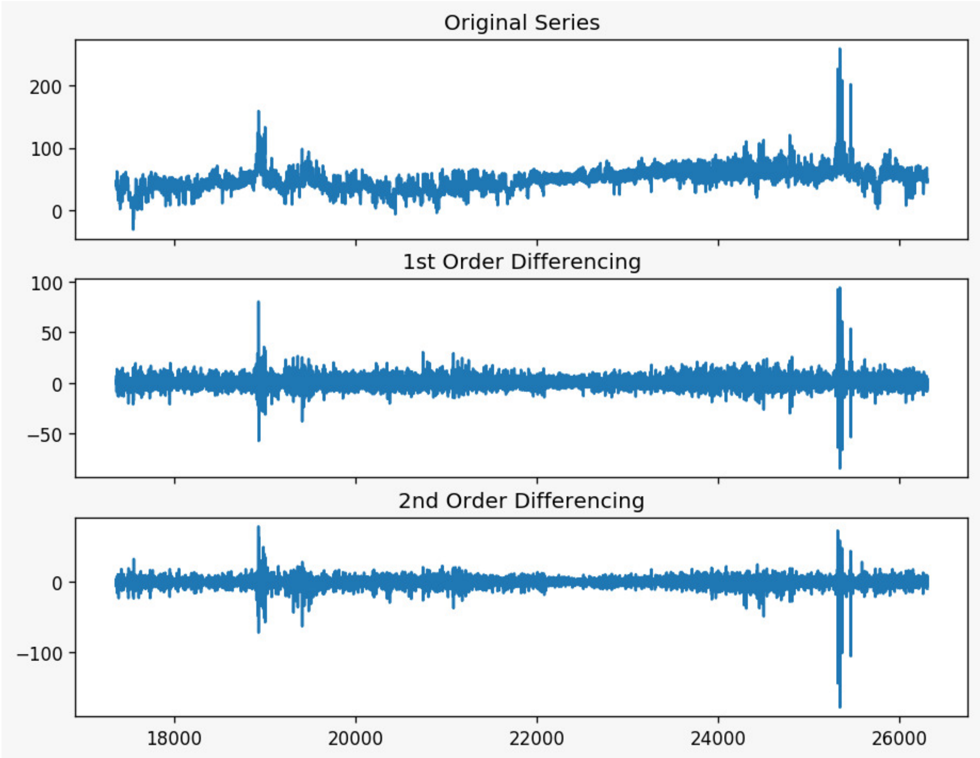
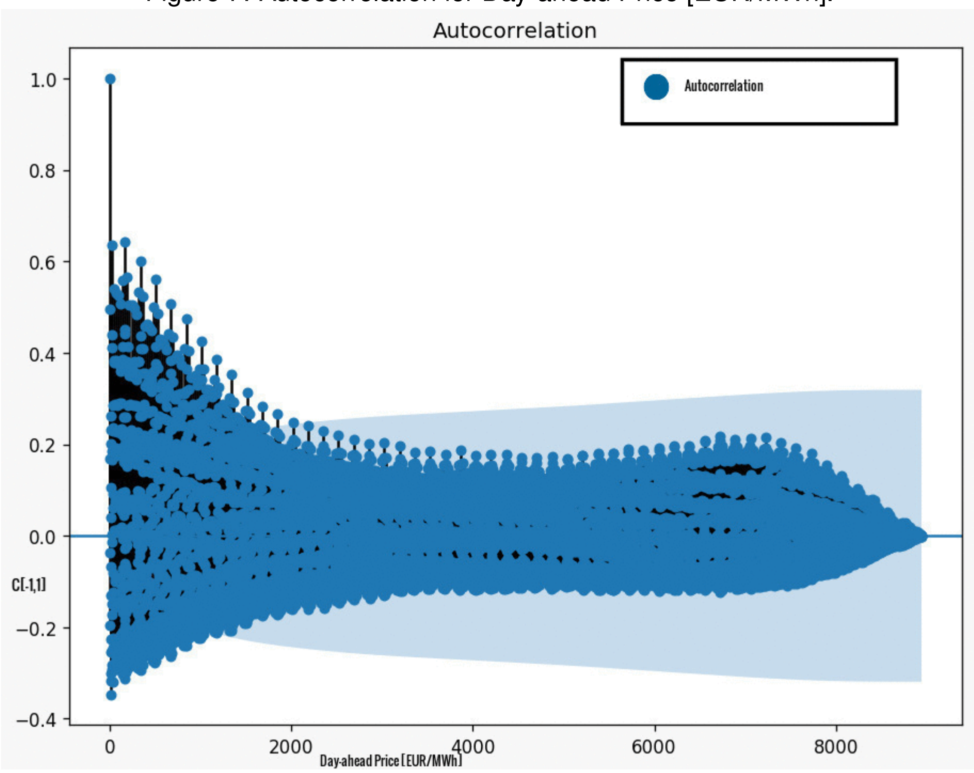
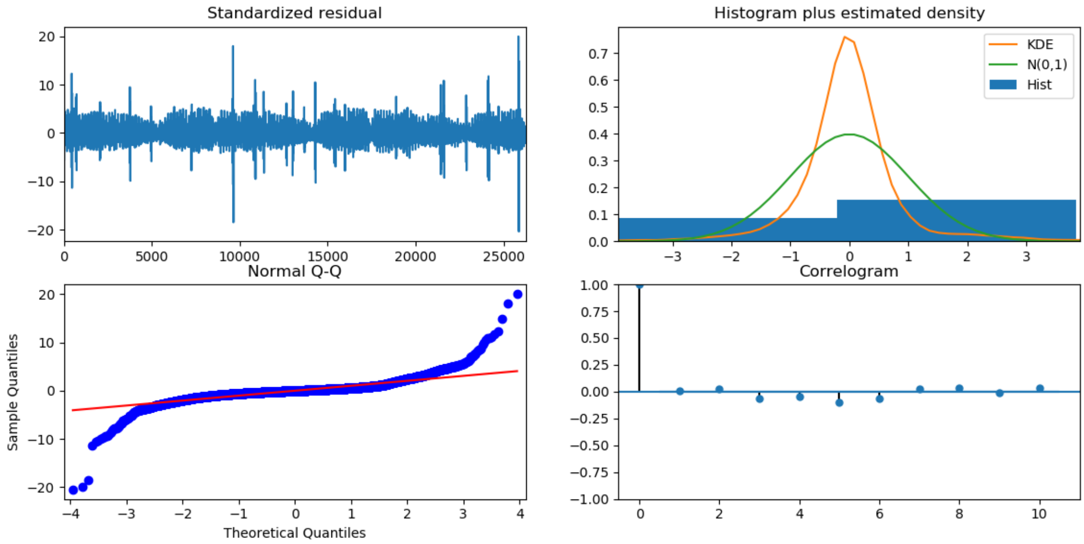


Figure 7: Autocorrelation for Day-ahead Price [EUR/MWh].



In Seasonal ARIMA, the model parameter selection is carrried out through two plots named ACF and PACF. ACF stands for Auto-correlation and PACF stands for partial auto-correlation function. Autc-correlation function is used to find how correlated lags (points) are with one another whereas

Figure 8: MODEL DIAGNOSTICS.



Partial Auto-correlation function determines the correlation between two lags that are separated by a larger periods. ACF plot is used to select the Moving Average part which are q and Q and PACF is used to select the Autoregressive part of the model namely p,P. The blue rectangle line is known as the significance threshold which is essential for understanding which lags are important. The figure 9 shows the ACF and PACF plot for the non-stationary dataset. We observe that the dataset needs to be differenced. After taking the first-difference, the ACF and PACF plot is plotted again which is presented in figure 10. In this figure, sinus like behavior is clearly presented which indicates that strong seasonality exists in the data. In order to remove the seasonal effects, we apply seasonal differencing with the period of 24. The seasonal period is highlighted in figure 10 where lag 24, 48 etc. are the highest. In figure 11 the effect of seasonal differencing on the ACF and PACF plots are visualized. Further differencing is not needed anymore since the linear and sinus-like behavoir in ACF and PACF plots have been removed. The last step is to select the rest of the model parameters p,q and P, Q. The non-seasonal parameters p, q are selected by looking at the first lags of ACF and PACF. We see that for both plots, only one lag 2 expands beyong the significance line before it drops again. Hence, p, d, q is select to equal to 1,1,1 respectively. For parameters P and Q, we look at the seasonal circles for the 24 hour period which means that in ACF only lag 24 is significant before it drops. In PACF, we observe six lags that are significant during for 24 seasonal period. However, the significance of the lags reduce over this period indicating to select a lower value than six. The highest consecutive lags remain to be 24, 48 and 72 and therefore P is selected to be 3 and Q to be 1. Finally, SARIMA model parameters are selected to be SARIMA(1,1,1)(3,1,1,24).

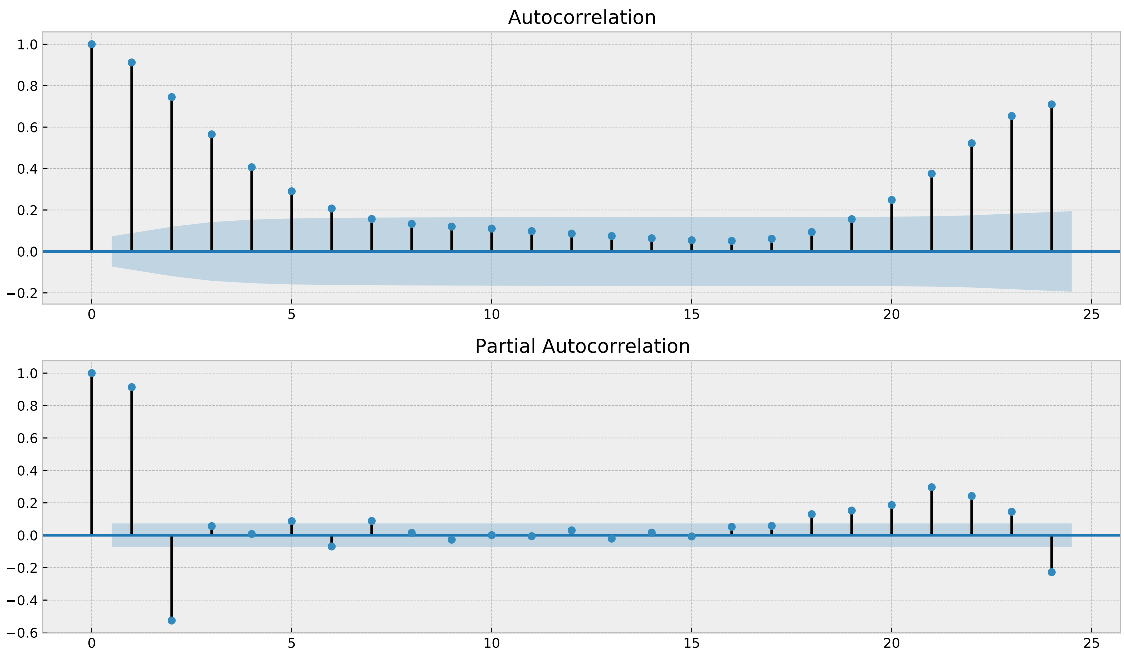


Figure 9: This figure dipslays the ACF and PACF plot for the non-stationary plot.

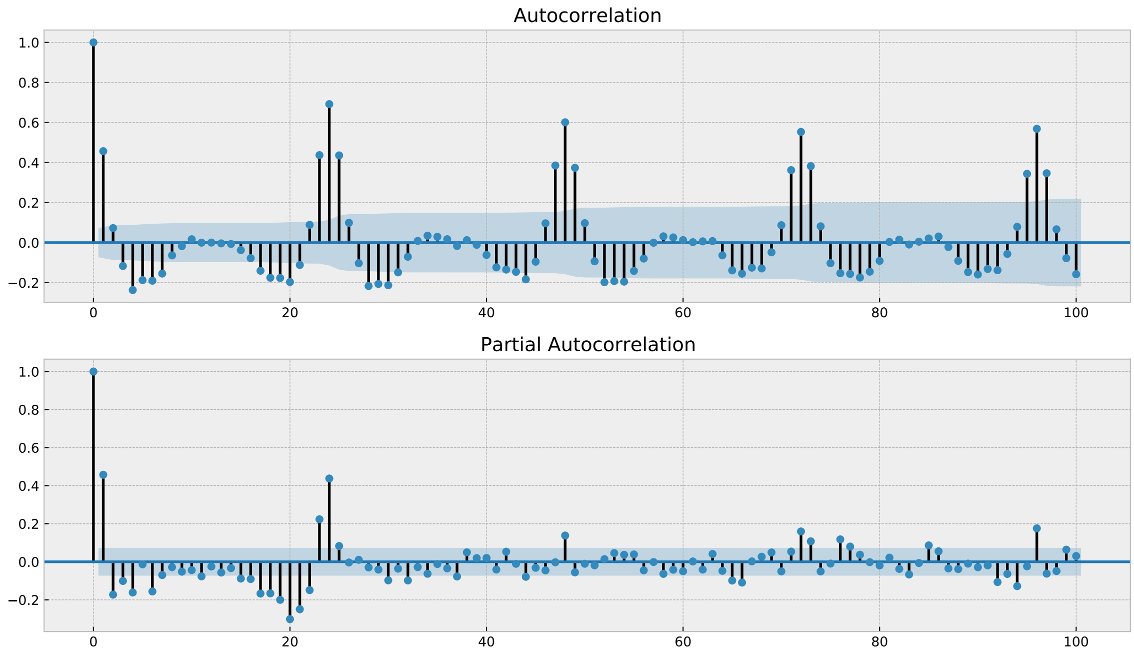


Figure 10: This figure dipslays the ACF and PACF plot after taking the first difference.

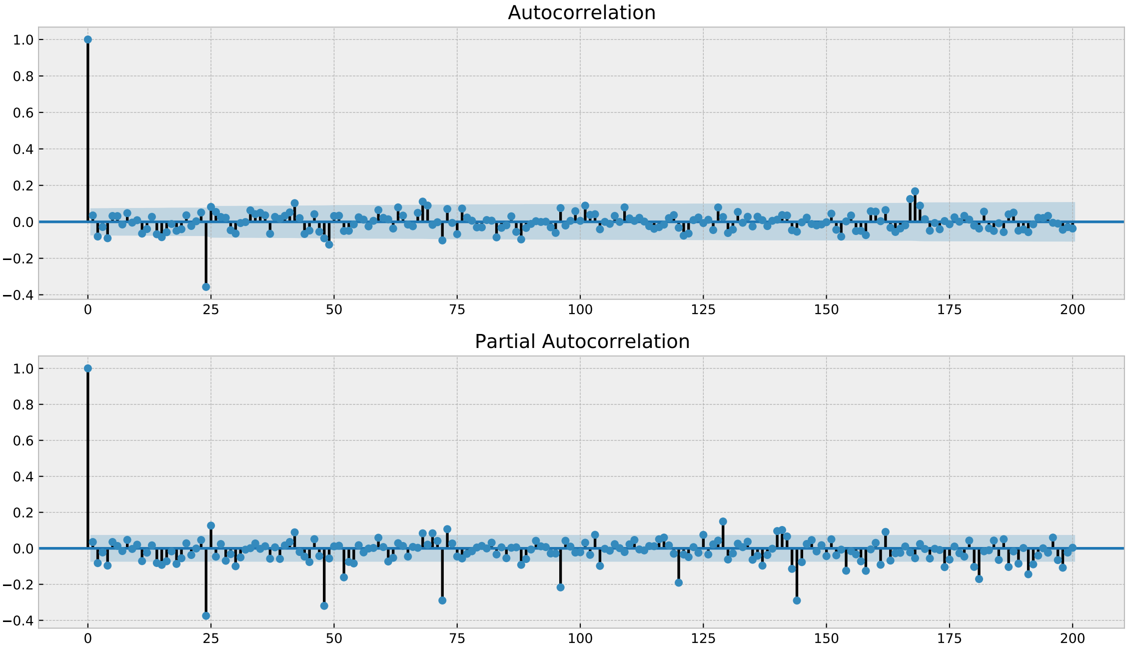


Figure 11: This figure dipslays the ACF and PACF plot after taking the first and seasonal difference.

In order to select the ARIMA Model order,to get the measure of Accuracy I have used the Mean Squared error(MSE),Mean Absolute Error(MAE),Mean Percentage Error(MAPE) and the Root Mean Squared Error(RMSE), these are their mathematical formula:

Mean squared error MSE = $\frac{1}{n}\sum_{t=1}^{n} e_t^2$

Root mean squared error RMSE = $\sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$

Mean absolute error MAE = $\frac{1}{n}\sum_{t=1}^{n} |e_t|$

Mean absolute percentage error MAPE = $\frac{100\%}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right|$

Where MSE=Mean Squared Error,n is the number of samples, t is the time,e is the estimator MAPE=Mean Absolute Percentage Error is a measure of prediction accuracy of a forecasting method in statistics like in trend estimation and also used as a loss function for regression problems in machine learning. MAE=Mean Absolute Error, refers to the results of measuring the difference between two continuous variables but in machine learning is a model evaluation metric often used with regression models. RMSE= Root Mean Squared Error is just a squared root of the Mean squared error. Due to the fact that ARIMA model does not performing well on the large amount of data sets, after dividing the time series data into small portion, I have done Cross Validation(CV)in order to test the performance of ARIMA model after using only few selected recent observations.[2]

Cross validation uses this mathematical formula:

$$\textbf{CV} = 1/T \sum_{t=1}^{n} [et/(1-ht)]^2$$

(7)
Where et is the residual obtained from fitting the model to all T observations. Meanwhile when

computing the cross validation statistic, it is not needed to actually fit T separate models.[2]

1.12.2 Seasonal Auto regression and moving average model

SARIMA(X) defined as actual known ARIMA and I have explained the mathematical formulas that ARIMA uses. But SARIMA considering the Seasonality which is in time series analysis. the ending "X" defined as Exogenous which is used in SARIMAX model to consider the train size with all features except one used by Exogenous in forecasting using SARIMAX.

1.12.3 Vector auto regression model

Vector Autoregression(VAR) is one of the machine learning methods used in a multivariate forecasting algorithm that can take place when two or more time series influence each other. It is called Autoregressive because each variable here I mean time series is modeled as a function of the past values, that is the predictors are nothing but the lags(time delayed value) of the series. In addition to that, VAR model is bi-directional not uni-directional as I said from above the reason is that the variables influence each other.
VAR model formula, each variable is modeled as a linear combination of past values of itself and the past values of other variables in the system. Since you have multiple time series that influence each other, it is modeled as a system of equations with one equation per variable(Time series).
For example, if you have 4 time series that influence each other, we will also have a system of 5 equations. Generally, the equation is framed like this:
Let's suppose that you have two variables(Time series) Y1 and Y2, from there you need to forecast the values of these variables at time(t).
To calculate Y1(t), Vector Autoregression model will use the past values of both Y1 as well as Y2. In the same manner, to computer Y2(t), the past values of both Y1 and Y2 be used.
The system of equations for a VAR(1) model with two time series(variables Y1 and Y2) is as follows:

$$\textbf{Y(1,t)} = \alpha 1 + \beta 11, 1Y(1, t-1) + \beta 12, 1Y(2, t-1) + \epsilon(1, t)$$

(8)

$$\textbf{Y(2,t)} = \alpha 2 + \beta 21, 1Y(1, t-1) + \beta 22, 1Y(2, t-1) + \epsilon(2, t)$$

(9) Where, $Y(1, t-1)$ and $Y(2, t-1)$ are the first lag of time series Y1 and Y2 respectively. The above equation is referred to as a VAR(1) model,because each equation is of order 1, and it contains up to one lag of each of the predictors(Y1 and Y2). So what I can add to that is that, since the Y terms in the equations are interrelated, the Y's are considered as endogenous or produced variables, rather than as exogenous predictors. Likewise, the second order VAR(2) model for two variables would include up to 2 lags for each variable (Y1 and Y2).

$$\textbf{Y(1,t)} = \alpha 1 + \beta 11, 1Y(1, t-1) + \beta 12, 1Y2, (t-1) + \beta 11, 2Y(2, t-2) + \beta 12, 2Y(2, t-2) + \epsilon(1, t)$$

(10)

$$\textbf{Y(2,t)} = \alpha 2 + \beta 21, 1Y(1, t-1) + \beta 22, 1Y(2, t-1) + \beta 21, 2Y1(1, t-2) + \beta 12, 2Y(2, t-2) + \epsilon(2, t)$$

(11)
Now, you can imagine what a second order VAR(2) model with 3 variables(Y1,Y2,Y3) would look like with 3 different equations and so on till n order VAR(n) model.[11]

1.13 Deep learning techniques for day ahead network models

1.13.1 Multivariate long short term memory neural network

Theoritically, Long Short Term Memory is one of deep learning which is also used for forecasting future data, I am adding multivariate because it's not using only one feature for me though it can also used as univariate in time series but output can be set up in a similar way, but it will be a single vector instead of a sequence, It can use multiple features at time in forecasting the day ahead price for my case after checking dependence, correlation analysis and added other feature of daily average of day ahead price. But though I had only 5 features, I have added average of daily price as additional feature besides Day-ahead Total Load Forecast, Actual Total Load, cet-timestamp, Day-ahead-Price Generation-wind onshore ,Day ahead Generation-solar as all being used as input for Long Short Term Memory Neural Network model as a better strategy to increase the performance of the model itself.[12]

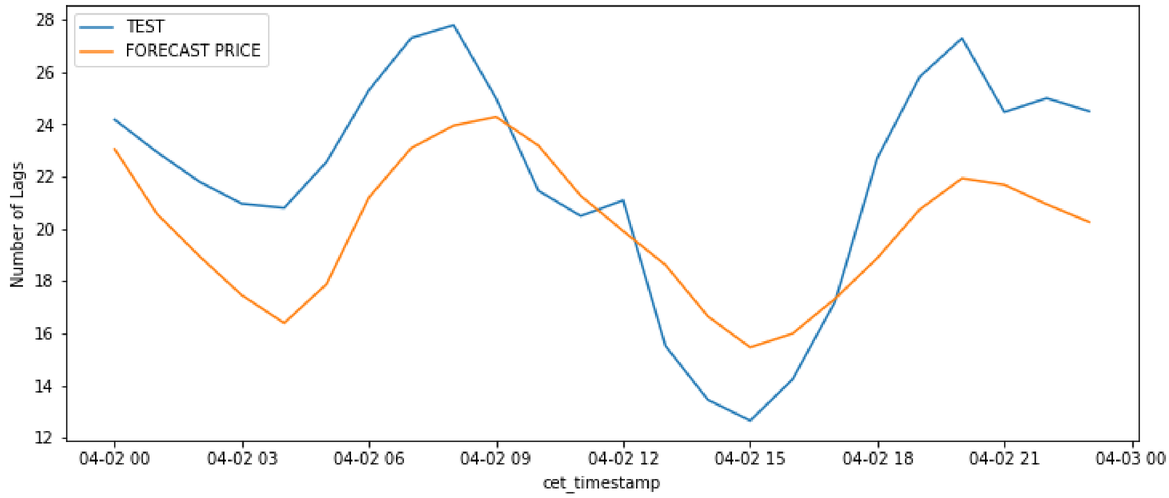


Figure 12: LSTM Result

1.13.2 Convolutional Neural Network for day ahead Price

It's univariate CNNs because it used only one feature at time to do forecasting for day ahead of prices. you might be ensure that the model does not gain the information from future time periods. We will set T=10, this means that the input for each sample is a vector of the previous 10 hours of the day ahead price, in addition to that, the choice of T=10 was just arbitrary but should be selected through experimentation. But the forecast here is just one step hourly ahead. [13]

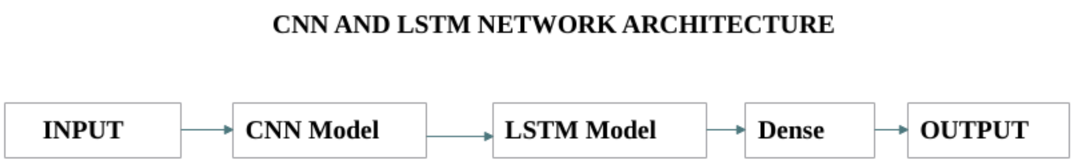


Figure 13: CNN and LSTM Network Architecture

1.13.3 Mathematical theory for CNN

Convolutional neural network has different layers where each layer can either be :
1. convolutional layer -CONV- and being followed with an activation function
2.Pooling layer -POOL-
3.Fully connected layer -FC- it is a layer which is similar to one from a feedforward neural network.

by doing deep into convolutional layer then as we have discribed from above,we use convolutional products by using so many filters whereby on the input followed by an activation function called "". In precision of the model, It regards at the 1th layer , we denote : [14]

More preciously, at the $\mathbf{l^{th}}$ **layer**, we denote:

- `Input` : $a^{[l-1]}$ with size $(n_H^{[l-1]}, n_W^{[l-1]}, n_C^{[l-1]})$, $a^{[0]}$ being the image in the input
- `Padding` : $p^{[l]}$, `stride` : $s^{[l]}$
- `Number of filters` : $n_C^{[l]}$ where each $K^{(n)}$ has the dimension: $(f^{[l]}, f^{[l]}, n_C^{[l-1]})$
- `Bias` of the n^{th} convolution: $b_n^{[l]}$
- `Activation function` : $\psi^{[l]}$
- `Output` : $a^{[l]}$ with size $(n_H^{[l]}, n_W^{[l]}, n_C^{[l]})$

And we have:

$\forall n \in [1, 2, ..., n_C^{[l]}]$:

$$conv(a^{[l-1]}, K^{(n)})_{x,y} = \psi^{[l]}(\sum_{i=1}^{n_H^{[l-1]}} \sum_{j=1}^{n_W^{[l-1]}} \sum_{k=1}^{n_C^{[l-1]}} K_{i,j,k}^{(n)} a_{x+i-1,y+j-1,k}^{[l-1]} + b_n^{[l]})$$
$$dim(conv(a^{[l-1]}, K^{(n)})) = (n_H^{[l]}, n_W^{[l]})$$

Thus:

$$a^{[l]} =$$
$$[\psi^{[l]}(conv(a^{[l-1]}, K^{(1)})), \psi^{[l]}(conv(a^{[l-1]}, K^{(2)})), ..., \psi^{[l]}(conv(a^{[l-1]}, K^{(n_C^{[l]})}))]$$
$$dim(a^{[l]}) = (n_H^{[l]}, n_W^{[l]}, n_C^{[l]})$$

With:

$$n_{H/W}^{[l]} = \left\lfloor \frac{n_{H/W}^{[l-1]} + 2p^{[l]} - f^{[l]}}{s^{[l]}} + 1 \right\rfloor ; s > 0$$
$$= n_{H/W}^{[l-1]} + 2p^{[l]} - f^{[l]}; s = 0$$
$$n_C^{[l]} = number\ of\ filters$$

The `learned parameters` at the l^{th} layer are:

- **Filters** with $(f^{[l]} \times f^{[l]} \times n_C^{[l-1]}) \times n_C^{[l]}$ parameters
- **Bias** with $(1 \times 1 \times 1) \times n_C^{[l]}$ parameters (broadcasting)

Figure 14: CNN mathematical theory
[14]

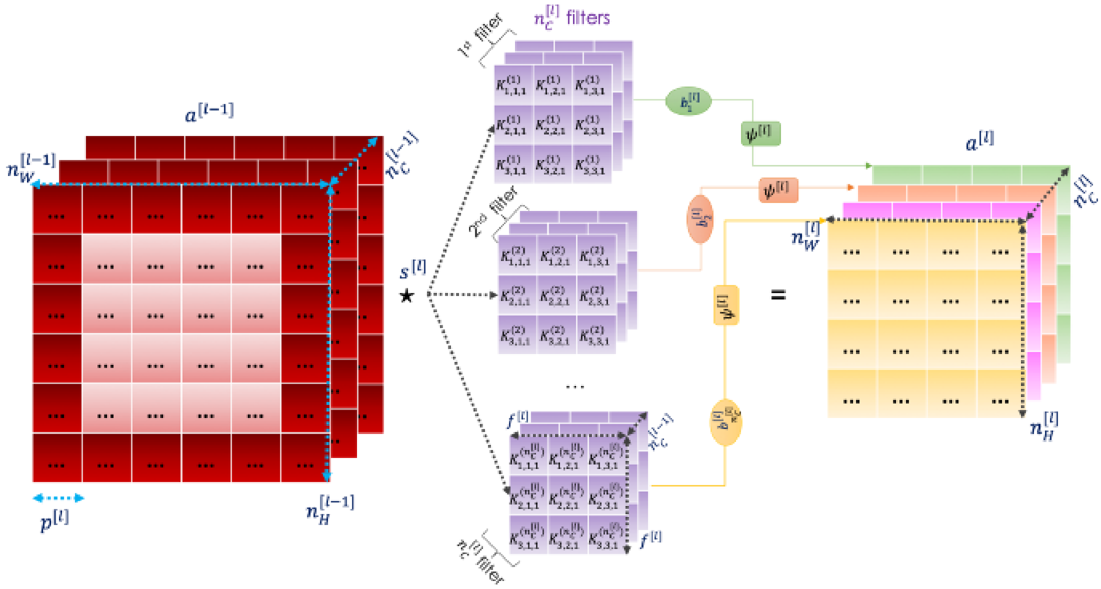


Figure 15: Sum up of Convolutional layer
[14]

convolution puts the input images through a set of convolutional filters, each of which activates certain features from the images Rectified linear Unit(ReLU) this allows for faster and more effective training by mapping negative values to zero and maintaining positive values. this is sometimes referred to as activation,the reason is that only the activated features are carried forward into the next layer for the pooling stage, it simplifies the output by performing nonlinear downsampling and recuding the number of parameters that the network needs to learn.

These operations are repeated over tens or hundreds of layers ,with each layer learning to identify different features.(figure 16). [15]

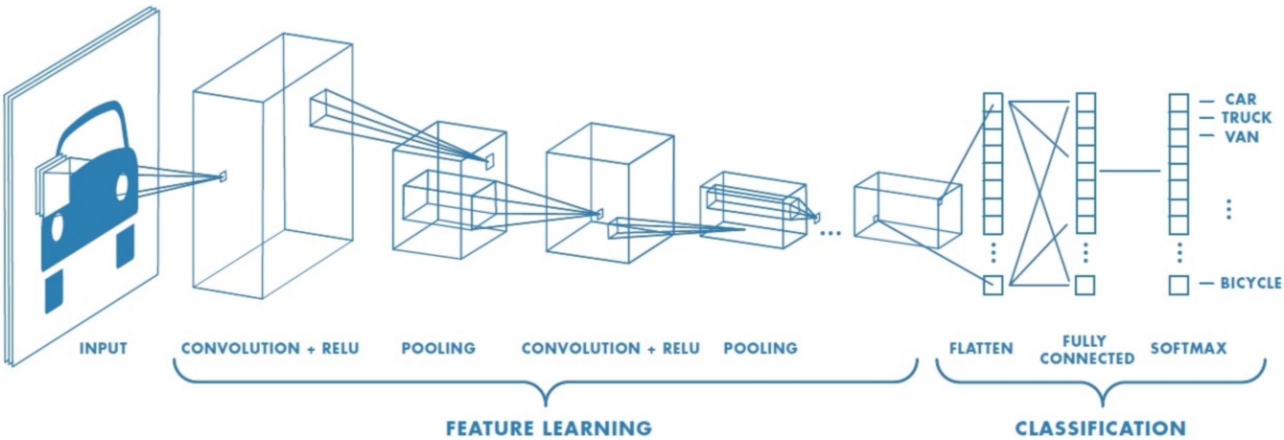


Figure 16: CNN model
[15]

1.13.4 Recurrent Network for day ahead price neurons

Recurrent neural network is a neural network model which is on high level in forecasting in general, I used to predict the future prices by considering the inputs of past values of the prices in the years between 2018 till april 2020 in time series forecasting. It is a model that utilizes standard sequence to sequence recurrent neural network architecture.s In general, RNNs needs less number of epochs compared to CNNs in batch size of 64.[16]

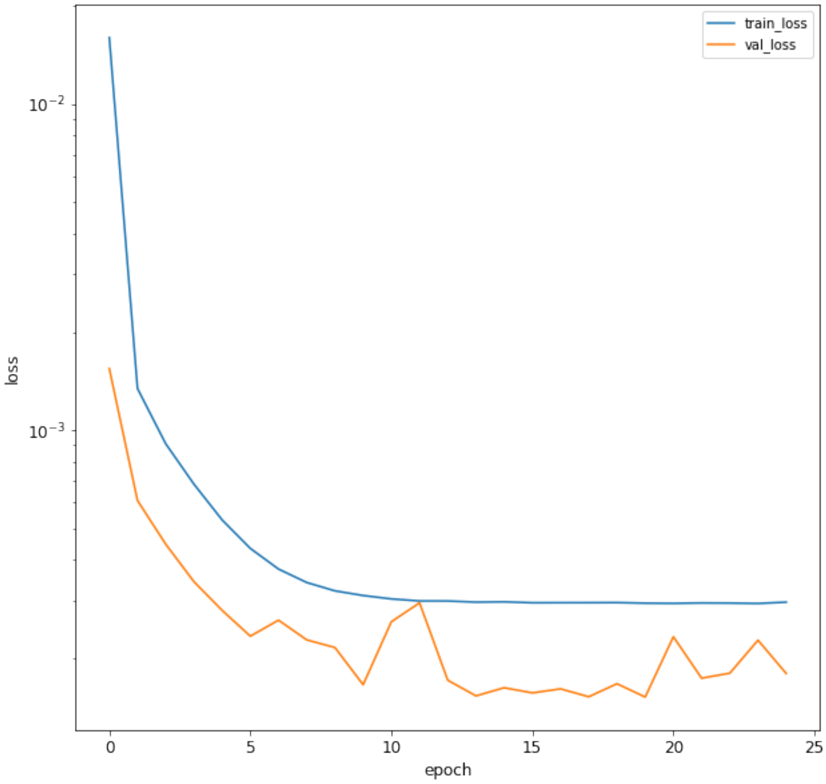


Figure 17: TRAIN AND LOSS VALIDATION FOR RNN

1.13.5 Hyperparameters tuning

In order to get a perfect accuracy in forecasting day ahead price using deep learning such as RNNs, CNNs you need to understand how the neural networks are characterized by different turnable parameters via numerical experiments. Then to evaluate/test how well the models perform, you have to do parameter turning in order to get a better results in forecasting the price. Most neural network in their natural functions ,have parameters to be tuned to be able to increase the accuracy of the results. So this implies that we need to set up the experiments in a way that the tuned parameters can be fairly compared to untuned ones because even neural network by default has the turning parameters, but you need to keep turning them in order to evaluate the performance of models. During the tuning parameters phase, for my scenario, I have tuned the following parameters :

1. Number of epochs
2. The batch size in range of 32,64 or 128
3. Kernel size
4. Dense
5. Specify the dimensionality
6. Set T (time)

The core advice to understand these neural network is to go first with simple model that does not risk to overfit the model either in multivariate or univariate time series of data.

1.14 Development tools

Since my master project requires machine learning and deep learning with statistical tools, I have used different tools which are :

1.14.1 Python, keras and TensorFlow

I have used some libraries which helped me to achieve this task, I have used python as programming language and used keras and tensorFlow libraries and inside those libraries there are other I have accessed and helped me to finish the project.

1.14.2 LaTeX

Latex is known as a scientific text editor that helps us to do the scientific report and so easy to write some mathematical formulas compared to other comparable editors,I can recommend anyone who is willingly to write a scientific report so easily to use it.

1.14.3 Jupyter notebook

JUpyter notebook is one of the tools used for data analysis and process them till you get the final result for what you are intended to do. So it's among the statistical tools I have used for this master thesis.

1.14.4 Sublime text editor

It is a text editor that is designed to help people to write codes or scientific report by using Latex commands and other commands.

1.14.5 Edraw Max

This is used to draw any kind of database,Entity relationship diagram or other designs so easily and it depends on which design you want to do.

Model assessment

From above design, I was explaining how the model works starting from scrach , Here are the steps in details for what I have designed:

1. Train model with TRAIN
2. Fit() the model
3. Forecast the next hourly,daily,monthly or yearly.
4. we put real value of testing hour
5. computed the errors like MSE,RMSE,MAE,MAPE in order to be able to easily understand how the actual prices is correlating with the forecasted ones.

1.15 Model Performance -Day ahead price

Table 4: table of Model performance results.

Models	MSE Result	MAE Result
SARIMAX	13.5	1.8
CNNs	13.4	1.73
RNNs	14.5	2.3
VARIMAX	15.5	3.5
SARIMA	15.5	3.5
LSTM	15.5	3.5

1.16 Model Performance comparaison

During this journey of trying different models, I have approached machine learning and deep learning models and performed differents over the data. The deep learning models such as RNNs and CNNs performed very well compared to the machine learning models whereby for RNNs the absolute mean error was 2.3 and for CNNs is close to that. But also the performance of machine learning models was quite good for their model accuracy. Specifically for machine learning I have tried different models such as SARIMAX, ARIMA, VAR and showed their result in the report as well. For deep learning techniques it requires to use big data tools such as tensorflow and keras while for machine learning techniques we only use simple tools.

Chapter 2

2 Results

2.1 ARIMA Model Result

In this chapter, I am showing all results found in research of this master thesis project. Best ARIMA (2,1,3) MSE= 2.228 The result of best ARIMA after taking only price feature and applying the model. We know that ARIMA Model usually used to perform by using only one feature during the process of accessing the model.(figure 18). In this figure, I am showing the result of ARIMA Model while taking few data and forecasted daily ahead price starting from 16th december 2019. The model performance is quite good as shown on the figure the correlation result between the past values and forecasted values obtained from forecasting is well correlated.(Figure 18).

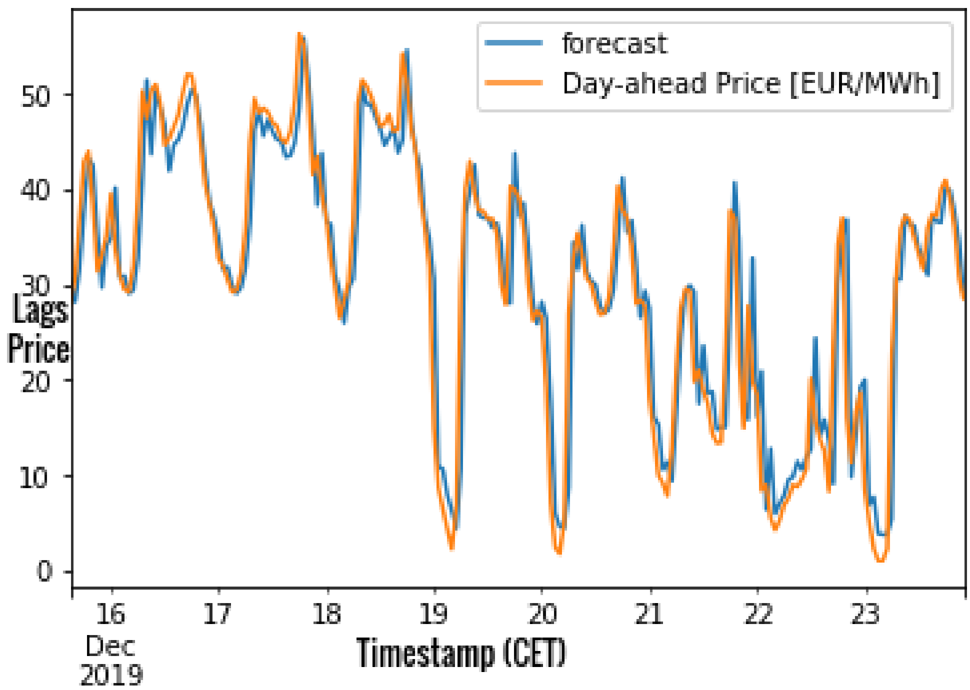


Figure 18: ARIMA model forecasted for only 100 data result

2.2 SARIMA Model Result

The SARIMA(1,1,1)(3,1,1,24) model was forecasted in a similar manner as ARIMA. In figure 19 the forecast is displayed which resulted in a forecast accuracy of 2.54 MAE (mean absolute error) which means that the models average forecast error is 2.5.

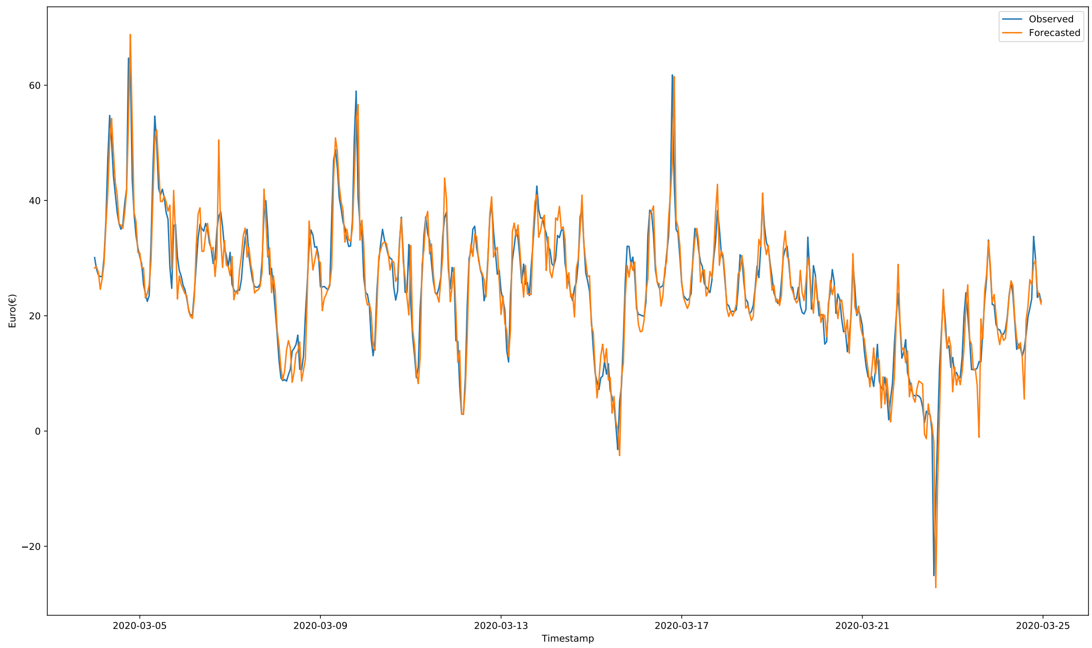


Figure 19: SARIMA model result

2.3 CNN Model Result and Heat-Map

In this figure, I am showing the result of CNN prediction of one step ahead for Day-ahead price[EUR/Mwh] starting from 04 march 2020 till 25 march 2020 while taking the 80 number of epochs. And the prediction result is quite good obviously from the figure.It's good to use deep learning models such as CNN for predicting the price. It does well during the prediction step.(Figure 20)

Here I am showing the train and validation loss for Convolutional Neural Network model with only 120 epochs how the correlation it was between train and validation data(figure 21)

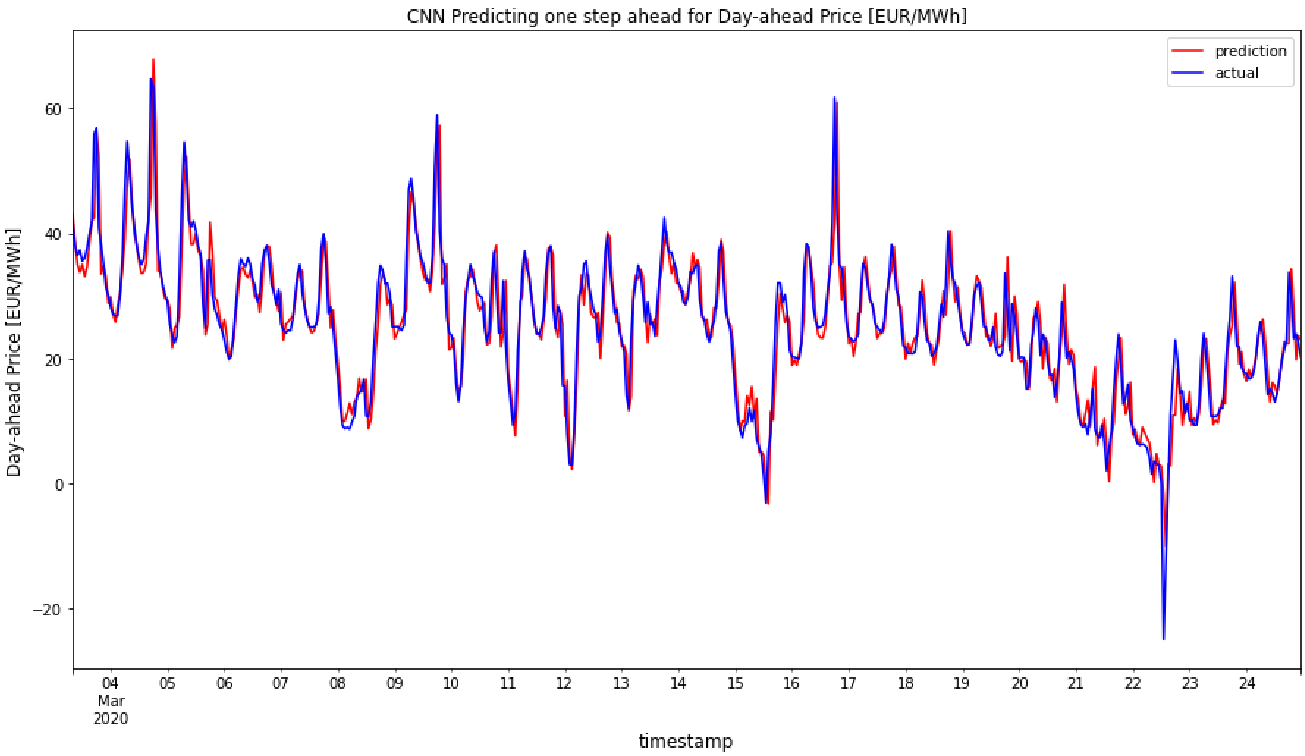


Figure 20: CNN model forecasted result

As you can see, the blue color represents the train data,red color represents the test data and the green color represents the validation data that I used for models.

Here I am showing the train and validation loss for Recurrent Neural Network model with only 25 epochs how the correlation it was between train and validation data(figure 9)

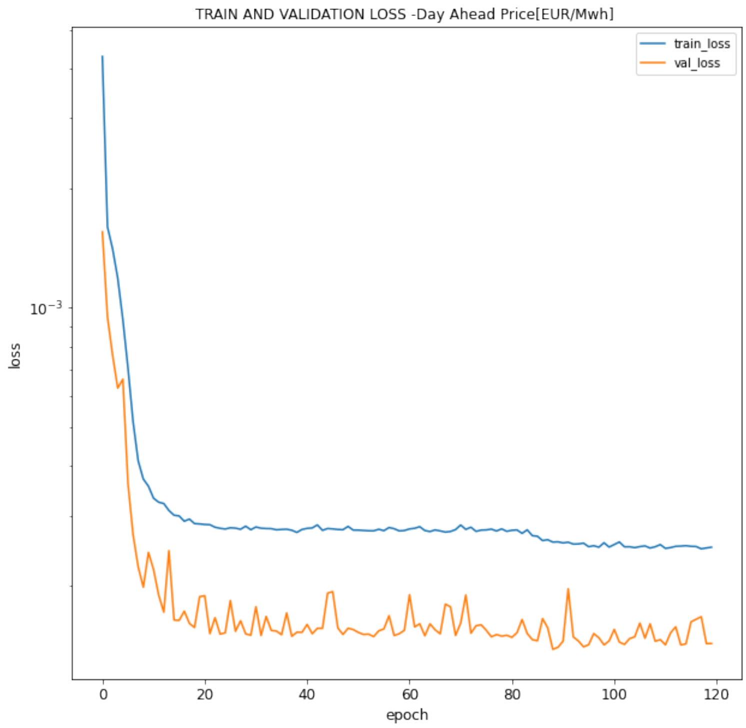


Figure 21: CNN TRAIN,TEST AND VALIDATION LOSS ERROR

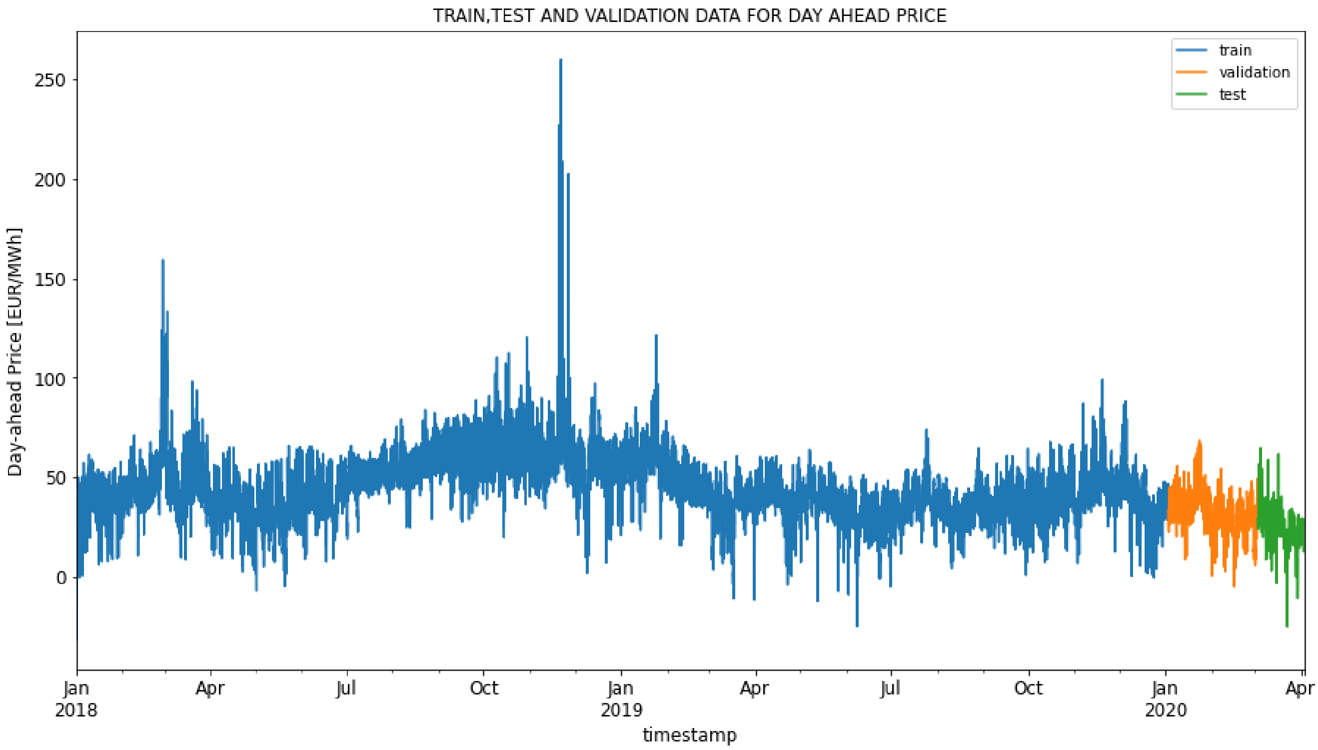


Figure 22: TRAIN AND TEST AND VALIDATION DATA result

In order to forecast using the help of heat maps by taking them as images and applying convolutional neural network to forecast using images instead of actual univariate or multivariate time series of data. As I have tried by using the shown images and found myself having prediction of image pixels in between 0 to 255 of the heat maps not the prediction of the prices in EUR

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv1d_1 (Conv1D)            (None, 10, 5)             15
_________________________________________________________________
conv1d_2 (Conv1D)            (None, 10, 5)             55
_________________________________________________________________
conv1d_3 (Conv1D)            (None, 10, 5)             55
_________________________________________________________________
flatten_1 (Flatten)          (None, 50)                0
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 51
=================================================================
Total params: 176
Trainable params: 176
Non-trainable params: 0
_________________________________________________________________
```

Figure 23: CNN model summary for sequencial model

represented by such heat maps.Which ended up confusing a bit. These two heat maps (figure 25 and figure 26) I used them to analyse the heat maps in order to know if it's necessary to detect the seasonality in heat map while I used them to do prediction. the last heat map (figure 24) is showing the result of prediction by image pixels.

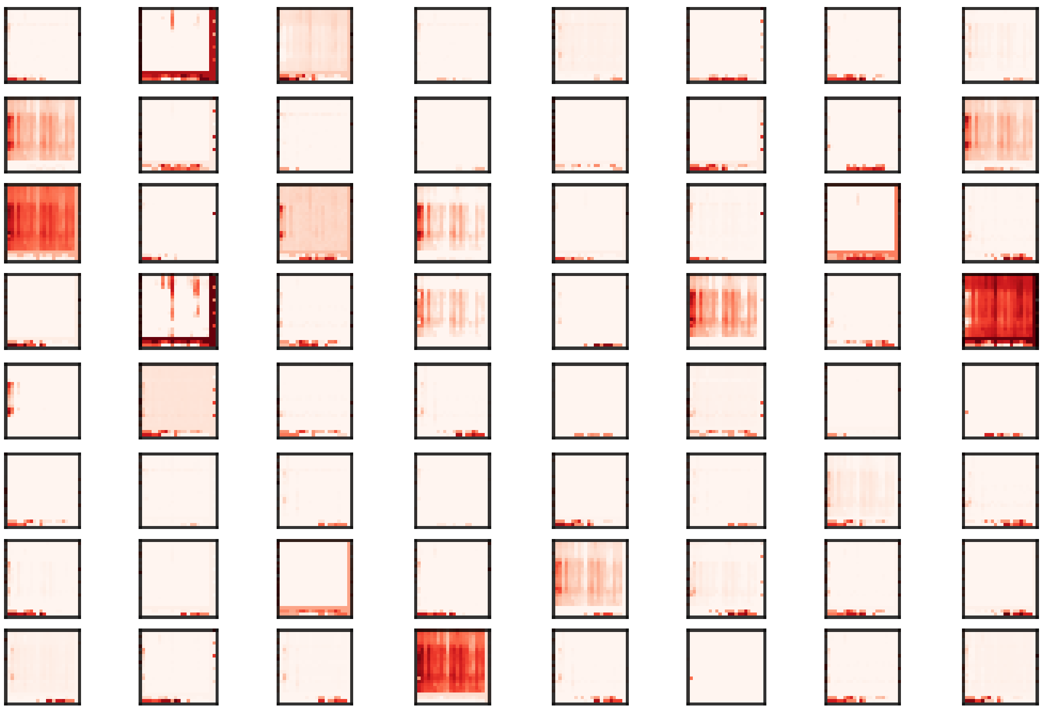


Figure 24: one single image of predicted heat map

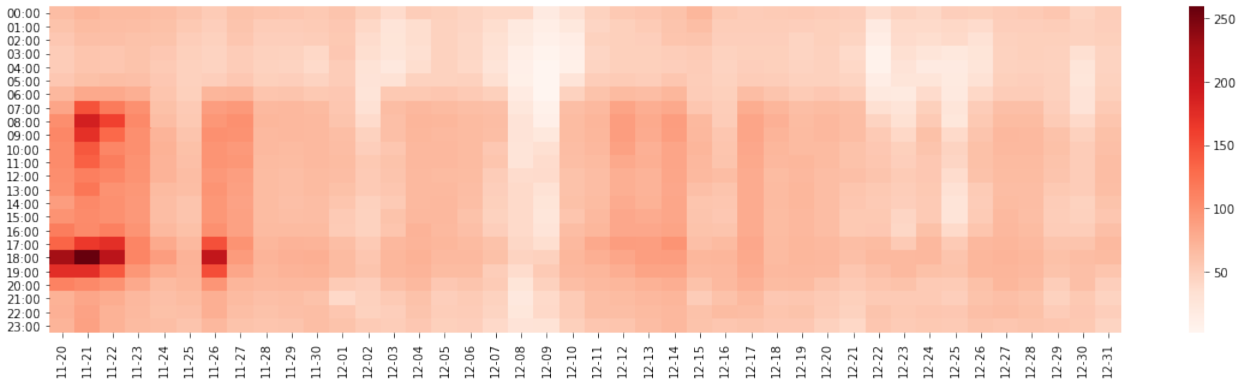


Figure 25: HEAT MAP RESULT

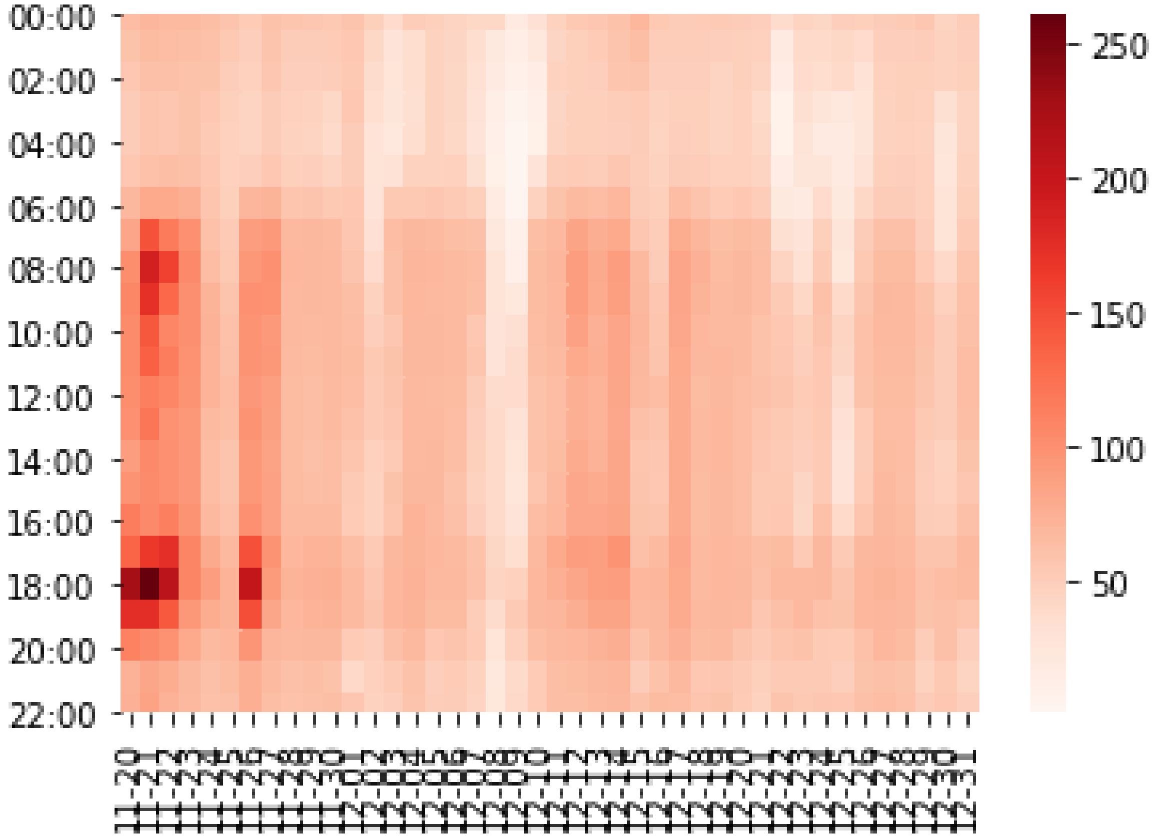


Figure 26: HEAT MAP RESULTS

2.4 RNNs Model Result

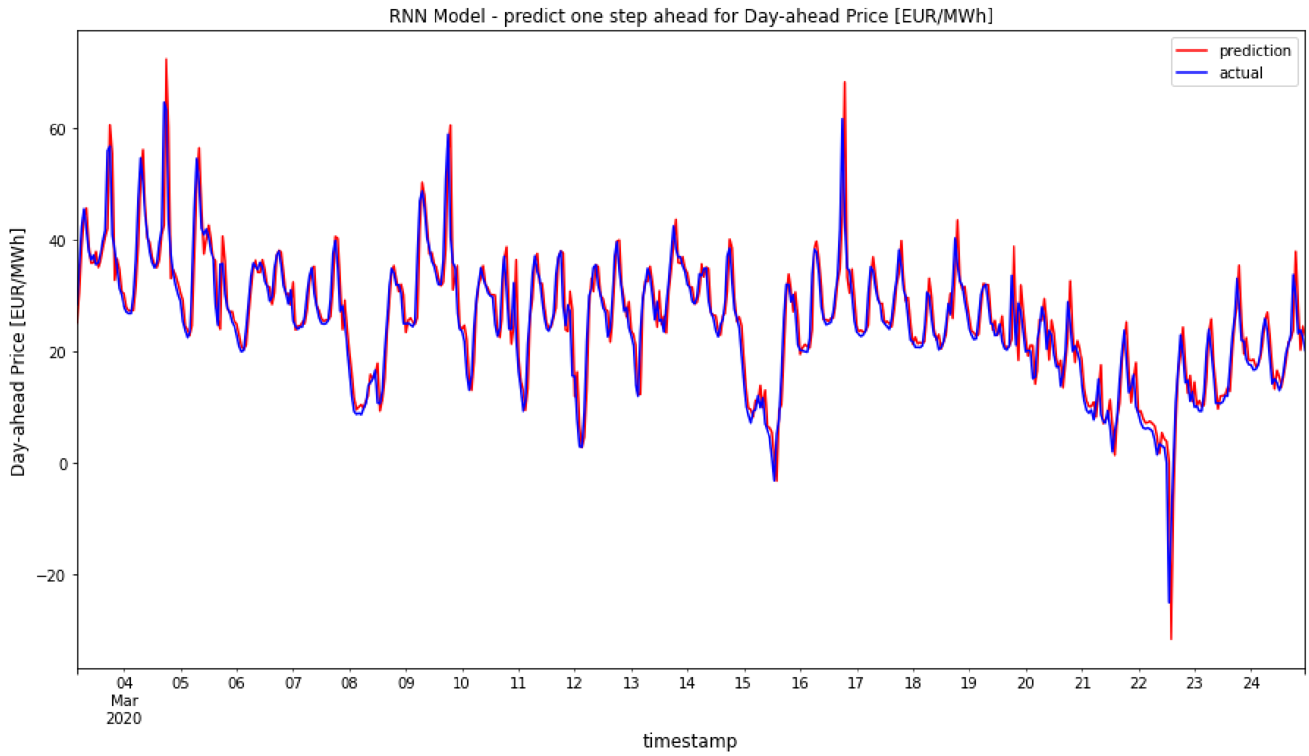


Figure 27: RNNs model forecasted result

In this figure, I am showing the result of RNN prediction of one step ahead for Day-ahead price[EUR/Mwh] starting from 04 march 2020 till 25 march 2020 while taking the 80 number of epochs. And the prediction result is quite good obviously from the figure. It is good to use deep learning models such as RNN for predicting the price. It does well during the prediction step.(Figure 14)

Chapter 3

3 Discussion and conclusions

3.1 Discussion

During this research I have tried different model approach to achieve the task I had for forecasting the day ahead price for french energy I have tried my best and exploring the data I had, but as limitation I didn't use all recommended approach from my supervisor due to the fact that some are not working to the data I had and others are in future work for other scientific research when necessarry. but more than 80 percent I have used the recommendation approaches from my supervisor. I have found the result to solve the tasks I had and I am showing them here in the report as well. I also talked about the limitation and future reflections.

3.2 Conclusions

As conclusion, I have explored so many models for the target of reaching the perfect forecasting of the price, I have gotten the good accuracy for sarimax for 2.5 Mean Absolute Error as one of machine learning model development I have used, and for the deep learning I have gotten the best result compared to the best ones I got for machine learning where for Convolutional neural network the forecasting price using the past values of prices was perfect and recurrent neural networks which I came to conclude that among all model development I have tried, the deep learning performed well compared to machine learning models for my data. But still I got good result also for machine learning as I explained very early.

Having such experience for exploring the models in forecasting is amazing whereby the promising result shows you that the forecasting is the best thing in the future especially in the area of energy translation where you come to understand the analysis of the prices of energy due to the actual prices and other features used to do data analysis. As I have also used the other feature of average daily price also helped me to be able to get more data precisely to try deep learning techniques with more data which also increseas the chance of the model performance.

Concusively I fully enjoyed to do forecasting of day ahead price plus learning the big data tools which helped me to get the chance to undestand more about the neural networks. I will continue enjoying the model development in my future but also in different domains.
The interest came while I was finishing with my advanced project where I only explored the application of machine learning in the domain of analysis the load and price for France Energy Market and as for master thesis I thought of exploring then the advanced models and advanced big data tools in order to train my brain to be able to forecast very well the future prices.
I have never tried any deep learning model before, But because of the task I had I did deep research to be able to explore them and then use it after understanding their algorithm in my task to get perfect forecasting of the day ahead price.

3.3 Potential model improvements

In this project, In order to improve the potential of the model,we know that the model for improvement provides a framework for developing, testing and implementing changes that lead to improvement. It is based in scientific method moderates the impulse to take immediate action with the wisdom of careful study.

Chapter 4

4 Future work

4.1 CNNs model challenge

Actually It is very possible to use Machine Learning methods to the heat map of the data, I have done it, Since I had only 3 different heat maps taken as images in CNN. I thought that It would be either way possible to train the heat maps data taken as images data using some features like standard deviation,skewness,mean and so on by applying either random forest model to be able to deal with some missing patterns in their results, using linear regression model or other suitable model. In order to use heat maps to detect seasonal patterns and use it to improve ARIMA, VAR and other models in general.

Here from above, I have results for the predicted heat maps images that I thought that I did the prediction for the price using heat maps, but later I realized that the CNN focused on the image pixels 0 to 255 of the heat maps not the numerical number of prices instead. which means it does the kernel and multiplying with matrix in image processing and data visualization that represented by the heat maps. I do believe that once I have more than like 200 heat map images I can easily manage to do prediction by taking heat maps as image and apply convolutional neural network.

4.2 Future reflection

First of all, due to the fact that I fully enjoyed to work on this project, I will continue for further research in this enjoyable field of doing forecasting by the aim of updating my mind about high technology in my career and also study other models for the helping the world when it meets the problem related to the model I know can help to resolve the matter. Not so easy as humans to know what tomorrow holds but this high technology at least predicts the near guessing of the reality that should happen in different situations of our daily life.

Though I am done with the required task for my master thesis, I am still eager to know how it would be to forecast using the help of heat maps taking them as images and applying convolutional neural network to forecast using images instead of actual univariate or multivariate time series of data. As I have tried it and found myself having prediction of image pixels in between 0 to 255 of the heat maps not the prediction of the prices in EUR represented by such heat maps.Which ended up confusing a bit.

Chapter 5

5 Appendix

5.1 Pearson's correlation coefficient

Pearson's correlation defined as : [17]

$$r = \frac{N\sum XY - (\sum X \sum Y)}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}} \tag{12}$$

where N=number of pairs of scores
Sum of the products of paired scores

$$\sum XY \tag{13}$$

Sum of X scores

$$\sum X \tag{14}$$

Sum of Y scores

$$\sum Y \tag{15}$$

Sum of squared X scores

$$N\sum X^2 \tag{16}$$

Sum of squared Y scores

$$N\sum Y^2 \tag{17}$$

5.2 Software Used

I have used :
Python 2.7 version
Jupyter notebook
Pandas
Tensorflow
Keras
EdrawMax for drawing any kind of database or Entity relationship diagram.
I have used latex and sublime text editor to do this report

5.3 Others

Operating System: kali linux
Hard disk : 1TB
RAM : 8GB
Processor: i3

You can find codes of my project here

List of Figures

List of Abreviation

Auto Regression and Moving Average (ARIMA)
Seasonal Auto Regression and Moving Average (SARIMA)
Vector Auto Regression (VAR)
Convolutional Neural Networks (CNNs)
Recurrent Neural Networks (RNNs)
Long and Short Term Memory (LSTM)
Mean Absolute Error (MAE)
Mean Squared Error (MSE)
Mean Absolute Percentage Error (MAPE)
Root Mean Squared Error (RMSE)
Cross Validation (CV)
Auto Regression (AR)
Partial AutoCorrelation Function (PACF)
AutoCorrelation Function (ACF)
Augmented Dickey Fuller (ADF)

References

[1] Machine learning algorithm. https://in.mathworks.com/help/stats/machine-learning-in-matlab.html?w.mathworks.com.

[2] Forecasting:principles and practice. https://otexts.com/fpp2.

[3] Saptarsi Goswami Prapanna Mondal, Labani Shit. Study of effectiveness of timeseries modeling(arima) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.4, No.2*, April 2014.

[4] European energy market reformcountry profile: France. https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Energy-and-Resources/gx-er-merket-reform-france.pdf.

[5] Iyamuremye. Advanced project ii. https://github.com/DidierG33k/Didier_Projects/blob/master/AdvancedProjectII.pdf, December 2019.

[6] entose transparency platform. https://transparency.entsoe.eu/, April 2014.

[7] Jason Brownlee. What is time series forecasting? https://machinelearningmastery.com/time-series-forecasting/, December 2016.

[8] Hyperparameter optimization. https://en.wikipedia.org/wiki/Hyperparameter_optimization.

[9] Helmut Luetkepohl. Vector autoregressive models. *EUI Working Paper ECO2011/30*, 2011.

[10] He X Luo L, Zhang X. Hospital daily outpatient visits forecasting using a combinatorial model based on arima and ses models. *BMC Health Serv Res 17*, July 2017.

[11] Correlation analysis of time series. https://nwfsc-timeseries.github.io/atsa-labs/sec-tslab-correlation-within-and-among-time-series.html.

[12] How to develop lstm models for time series forecasting. https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/.

[13] Thomas Wiatowski and Helmut B olcskei. A mathematical theory of deep convolutionalneural networks for feature extraction. https://arxiv.org/pdf/1512.06293.pdf.

[14] Convolutional neural networks' mathematics. https://medium.com/swlh/convolutional-neural-networks-mathematics-1beb3e6447c0.

[15] Deep learning. https://www.mathworks.com/solutions/deep-learning/convolutional-neural-network.html.

[16] Time series forecasting with rnns. https://towardsdatascience.com/time-series-forecasting-with-rnns-ff22683bbbb0.

[17] Pearson correlation coefficient: Formula, example significance. https://study.com/academy/lesson/pearson-correlation-coefficient-formula-example-significance.html.