TFG Structure - Machine Learning Based Predictive Modeling of Energy Prices

Acknowledgements

Dedication (to… mom dad gf grandparents my teacher etc)

Index

1. Introduction
   1. Context
   2. Motivation
   3. Objectives
   4. Summary of the results
2. Background
   1. Description of the different blocks
   2. ML
   3. Energia
3. Proposal
   1. Theoretical system description
4. Experimentation
   1. Data Description
   2. Set Up Explanation
   3. Results
   4. Low level discussion
5. Conclusions
   1. Recap
   2. Revisit Objectives
   3. Future Work
6. **Introduction**
   1. **Context and Inspiration**

Energy price prediction

My signals are energy prices

I have a time-series of prices, with an hourly resolution.

My first attempt will be to predict the price at 14:00H for the next day (the day after tomorrow since prices are set a day ahead of time).

Why at 14h – It’s a valley time slot, so we suppose it will be fairly stable

* 1. **Motivation**

**App**

* 1. **Objectives**

Figure out how to create an algorithm to predict decently accurately energy prices for a given hour slot

* 1. **Summary of the results**

1. **Background**
   1. **Description of the different blocks**
   2. **ML**
   3. **Energia**
2. **Proposal**
   1. **Theoretical system description**

We initially discussed various ways to attack the problem at hand. We thought about different approaches such as Time-Series Forecasting, about Neural Networks but settled on a plan to explore first established method of linear regression. This was a natural continuation of what was learnt in the third year subject *Modern Theory of Detection and Estimation.*

Since the original ideal was to create an algorithm to predict hourly prices, we decided to simplify the problem. The fist idea we developed was to separate out the project in 24 blocks, one for each hour block to predict. The reason behind splitting the data in 24 blocks was simple, we assumed that data between the same time slot across neighbor days, would be similar, or would follow a certain trend.

This kind of granular breaking up of data is similar to what a Random Forest does in order to absorb more details and create higher resolution predictions. We made this assumption based on the idea that it would simplify the project substantially. This forced us to generalize our approach, creating a system to analyze and evaluate the data in 24 distinct blocks, figuring out the best matrix dimensions for our linear regression matrix.

1. **Experimentation**
   1. **Data Description**

The data used in the study was obtained from the public domain

* 1. **Set Up Explanation (overview)**

The development environment was built with Conda, specifically miniconda. A environment

* 1. **Results**
  2. **Low level discussion (of the code? Of the different steps?)**

1. **Conclusions**
   1. **Recap**
   2. **Revisit Objectives**
   3. **Future Work**