

Forecasting Electric Prices in Spain

Magnus Bigelow



Problem Statement

Predict electricity prices in Spain for each hour of the upcoming day more accurately than estimates provided by the Spanish transmission agent and operator.

Use information available during the 2pm-3pm window the previous day during which generators in Spain submit their bids.

Understanding the Problem

Wholesale Electricity

Understand how electric markets work at a high level

Spanish market overview

Spanish market auction

Why is predicting prices important?

Process

Data & data sources

Processing data

Key variables

Modelling techniques

Results

Summary statistics

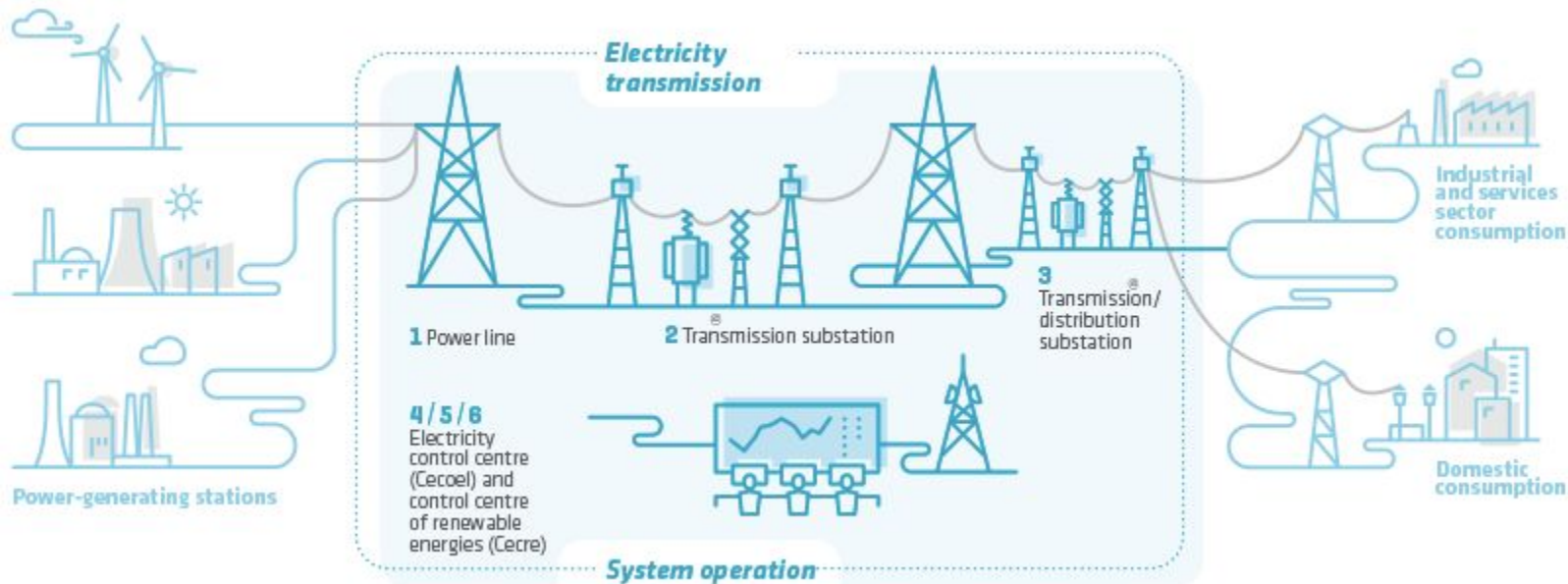
Best models

Next steps

Wholesale Electric Market:

What is it and why do
we need to predict
tomorrow's prices?

How do Wholesale Electric Markets Function?



Generators:
Produce electricity

**Transmission & System
Operator (TSO):**
Manage the grid

Consumers:
Use electricity

How are Electric Prices Determined?

Day Ahead Market

- TSO forecasts electric demand
- Generators submit bids indicating how much electricity they are willing to produce and at what price
- TSO clears the market by accepting the cheapest bids until demand is satisfied
- All cleared bids are paid the clearing price

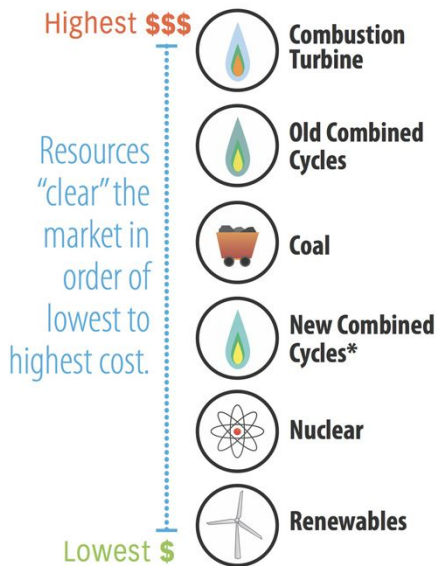
Intraday Market

- Electricity often traded in time increments as low as 15 minutes
- Ensures that generation meets demand throughout the day

Relative Cost of Electricity Generation

No Free Lunch

- Cheaper fuel sources generally have drawback
 - Variable nature: Renewables
 - Hazardous waste: Nuclear
- More expensive sources have advantages
 - Short time to operational



*New combined cycles are more fuel efficient.

Spanish Electric Market

Generators

- 3 players have ~50% market share
 - Grupo Endesa, Iberdrola & Gas Natural Fenosa
- Over 50 active electric companies

TSO

- RED Electrica Espana
- Part state owned and part privately owned
- TSO for the national grid

Renewables

- Plans to be 100% renewable by 2050
 - One of the world leaders in wind power with 23 GW of installed capacity
 - 20 GW of hydro capacity

Spanish Day Ahead Market

- TSO provides load forecasts for the next day
- Bids submitted during the 2pm - 3pm window locally
 - Bids submitted for all 24 hours
- Matching and publication of results at 3:07pm

Why is it important to predict electric prices?

Short Term:

- Managing day to day operations
- Maximizing profits
- Crucial for demand response

Medium Term:

- Balance sheet calculations
- Risk management
- Derivatives pricing

Long Term:

- Investment profitability analysis
- Long term planning

Process:

Steps taken to solve
the problem

Gathering Data

Kaggle - Energy

- 29 variables of hourly data from Spain covering 2014 - 2018
 - 35064 observations
- <https://www.kaggle.com/nicholasjhana/energy-consumption-generation-prices-and-weather>

Oil Prices

- Daily prices in euros
- Scraped from the web
- <https://www.exchangerates.org.uk/commodities/OIL-EUR-history.html>

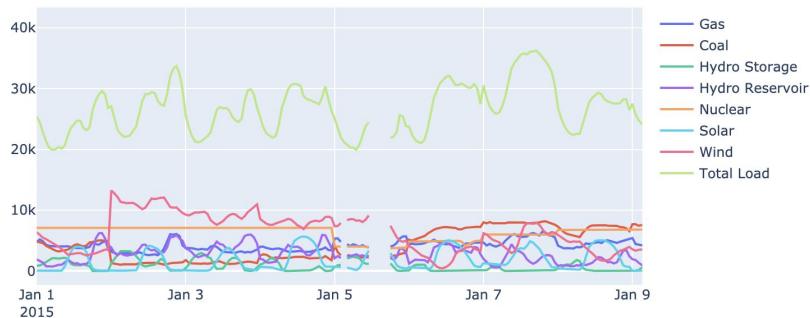
Kaggle - Weather

- Hourly weather data for 5 largest cities in Spain 2014 - 2018
 - Madrid, Barcelona, Sevilla, Bilbao, Valencia
- Subject to future inclusion

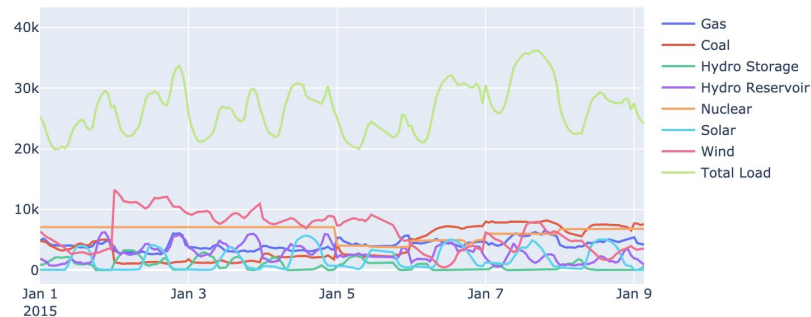
Cleaning

Almost no missing or clearly wrong values. All missing data was filled through linear interpolation which essentially draws a line between the surrounding known data points.

Top Generation Sources

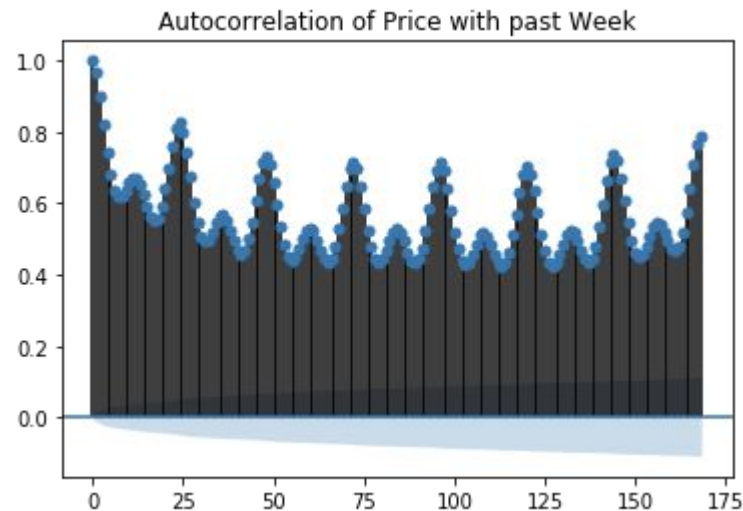
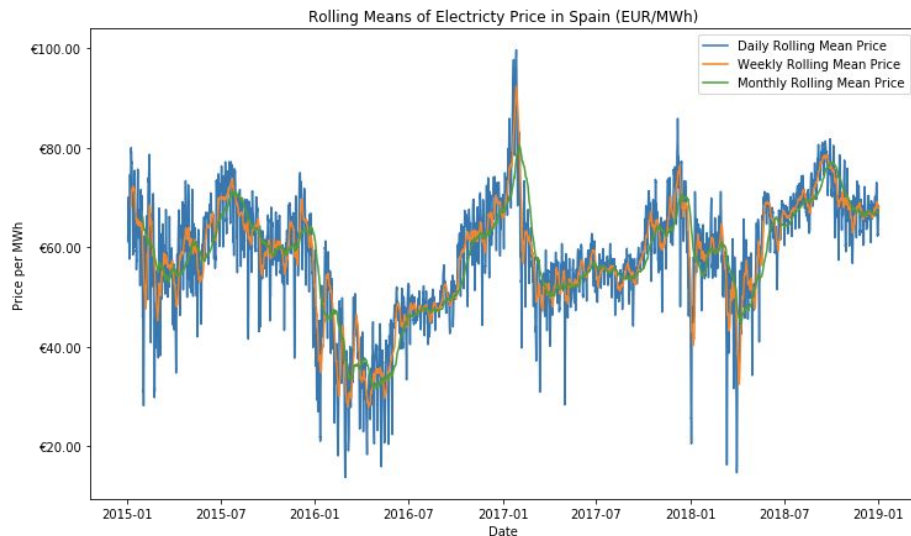


Top Generation Sources



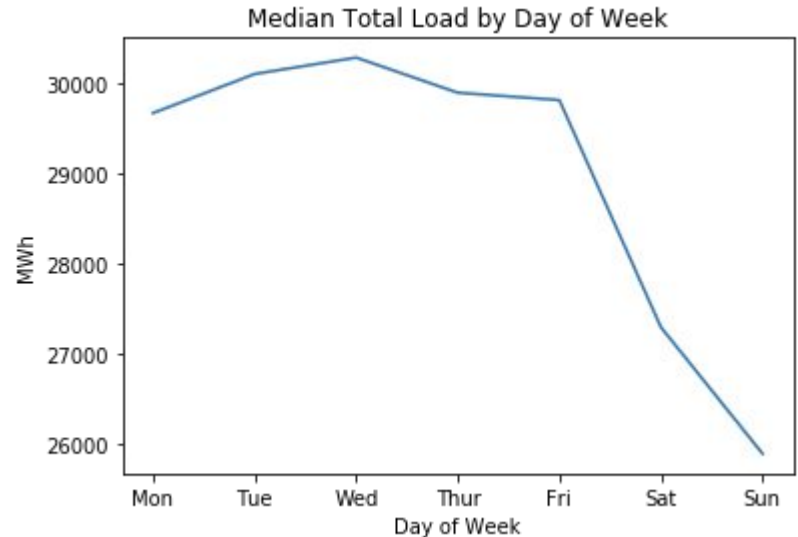
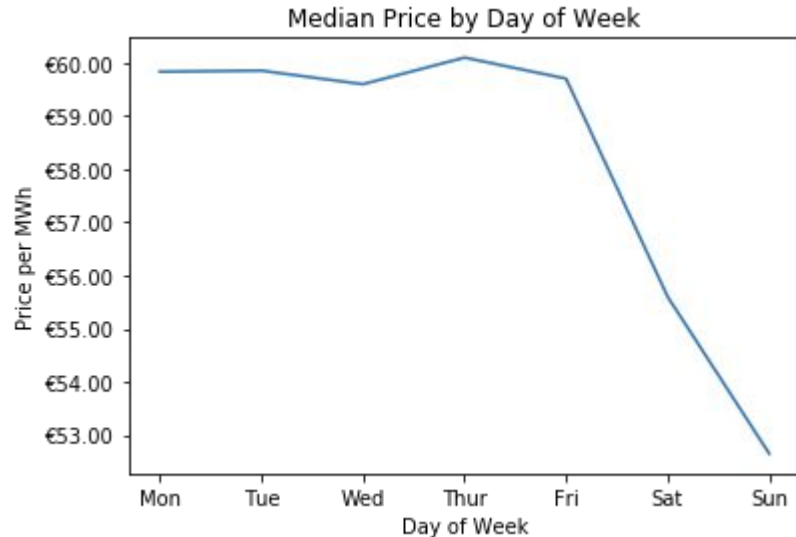
Exploratory Analysis – Prices

Looking at the whole data set there are no clear trends but when we zoom in we can see some clear trends on an hourly and weekly basis



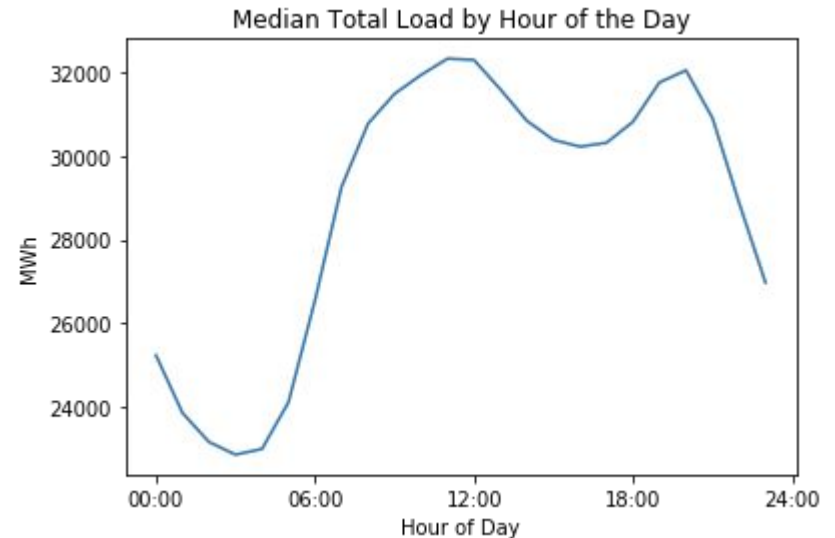
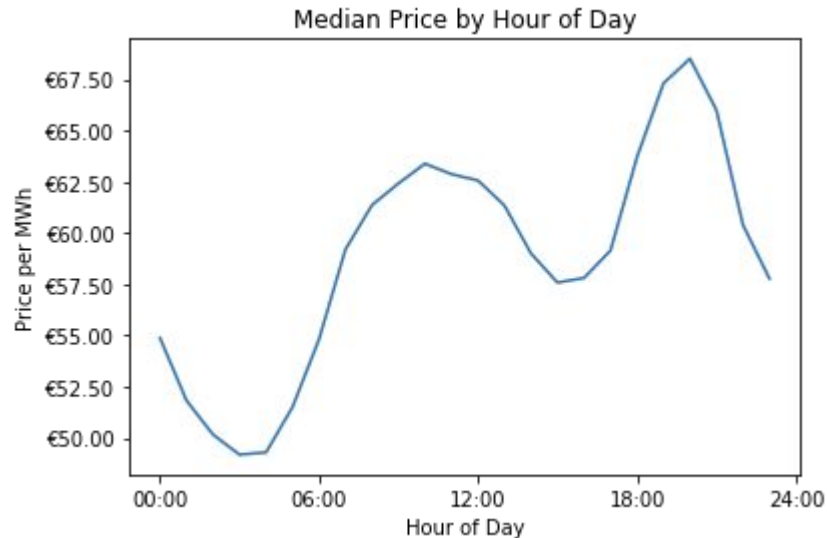
Price & Load by Weekday

Both load and price are relatively consistent throughout the week before dropping off on Saturday and even further on Sunday



Price & Load by Hour of Day

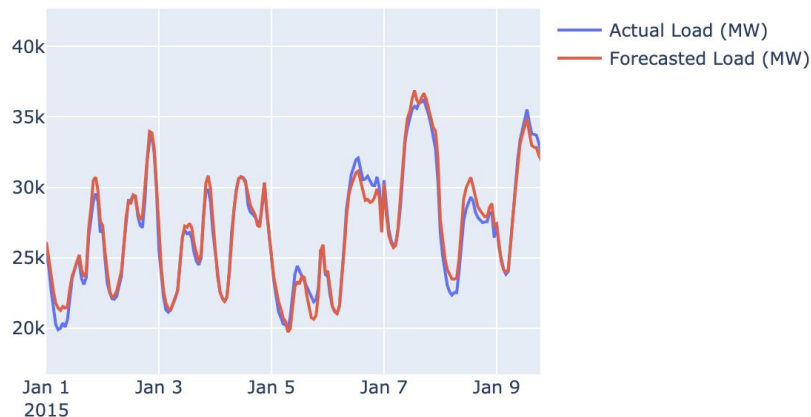
Price and load both have similar camel shaped profiles with lows while most people are asleep and peaks during the late morning and then also the evening



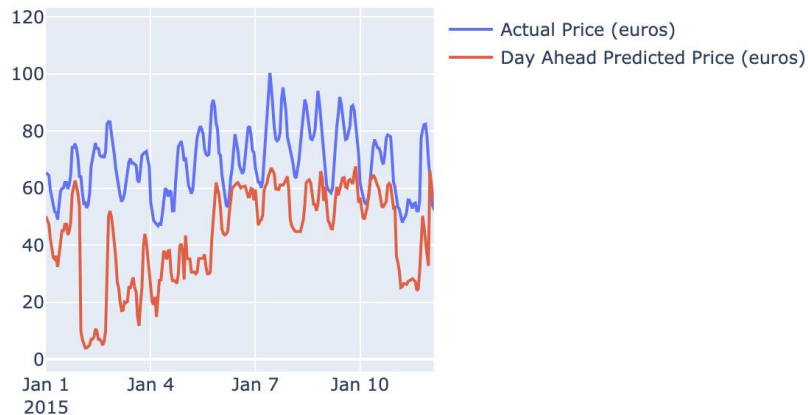
Exploratory Analysis – Predictions

The load forecasts provided by the TSO are extremely accurate while the predicted prices were an average of 13.4% lower than actual prices

Actual & Predicted Load in Spain



Actual & Predicted Energy Prices in Spain



Transformation For Modeling

The data was transformed such that the relevant X and y variables were the same row of the data frame in order to facilitate modeling. This created a dataset in which in each day had a single row with the following data points:

- Actual, hourly electric prices for the next day
- Projected, hourly total load and wind generation for the next day
- Actual, hourly electric prices from the beginning of the previous day up to and including the 2pm-3pm time slot
- Projected, hourly prices for the current day from 3pm - midnight
- Actual crude oil price for the current day

Model Types

Standard Regressors

- 6 different models from the sklearn library
- Classic approaches to regression problems
- Easy to tune hyperparameters

Vector Auto Regressors (VAR)

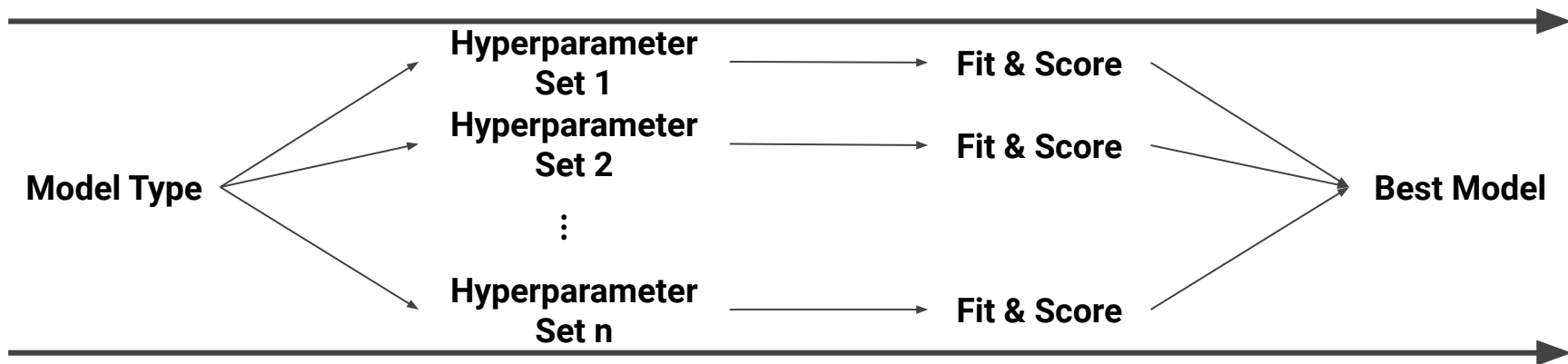
- Used for time-series data and rely on past values to predict future values
- Implemented using the statsmodels library

Neural Networks

- Set of algorithms that loosely resemble the human brain designed to recognize patterns
- Implemented using the tensorflow keras library

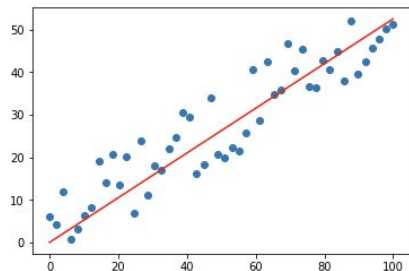
Modeling Process

For each model we set up a process that tests different hyperparameters (i.e. model settings) and compare how well they perform to find the optimal model of that type

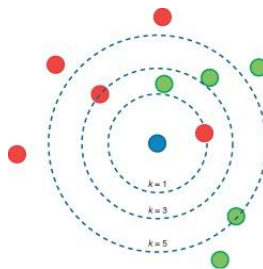


Standard Regressors

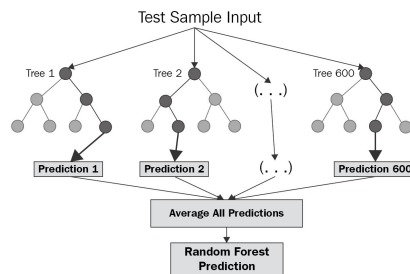
Linear Regression



K-Nearest Neighbors



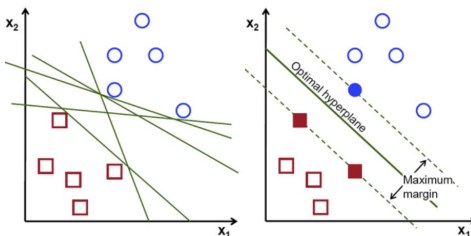
Random Forest



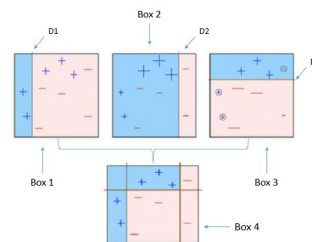
Elastic Net

Linear regression with regularization

Support Vector

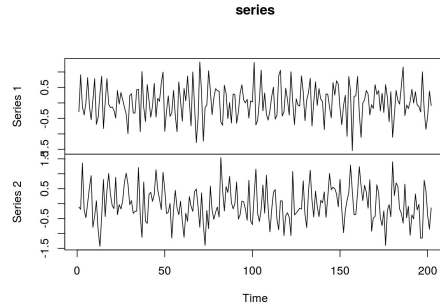


AdaBoost



Vector Auto Regressors

VAR

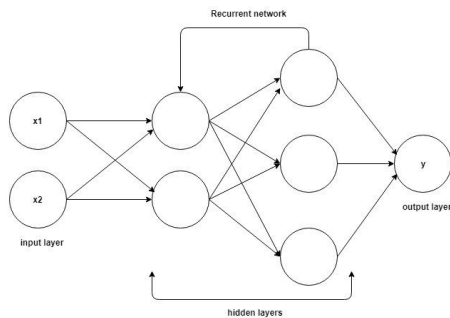


VARMAX

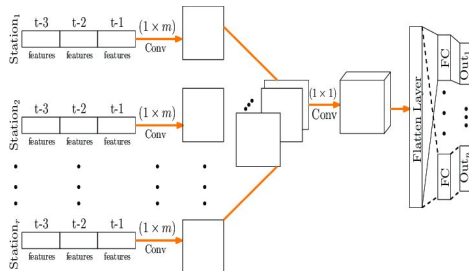
VAR with exogenous
variables

Neural Networks (NN)

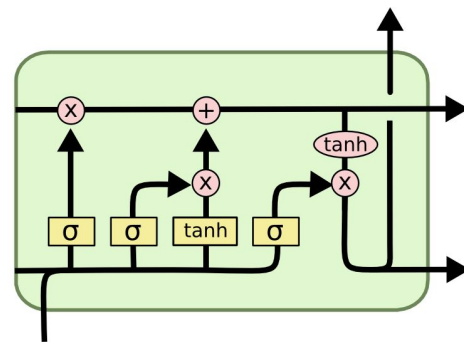
Recurrent NN



Convolutional NN



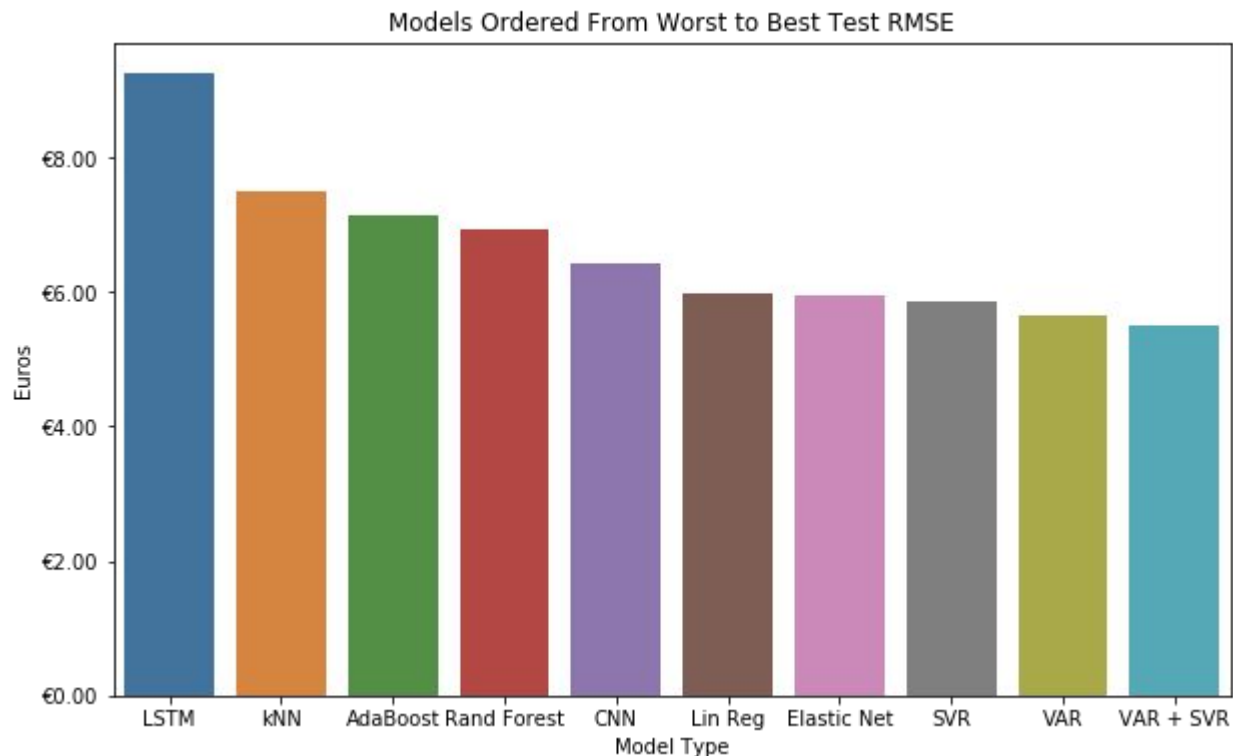
LSTM NN



Results:

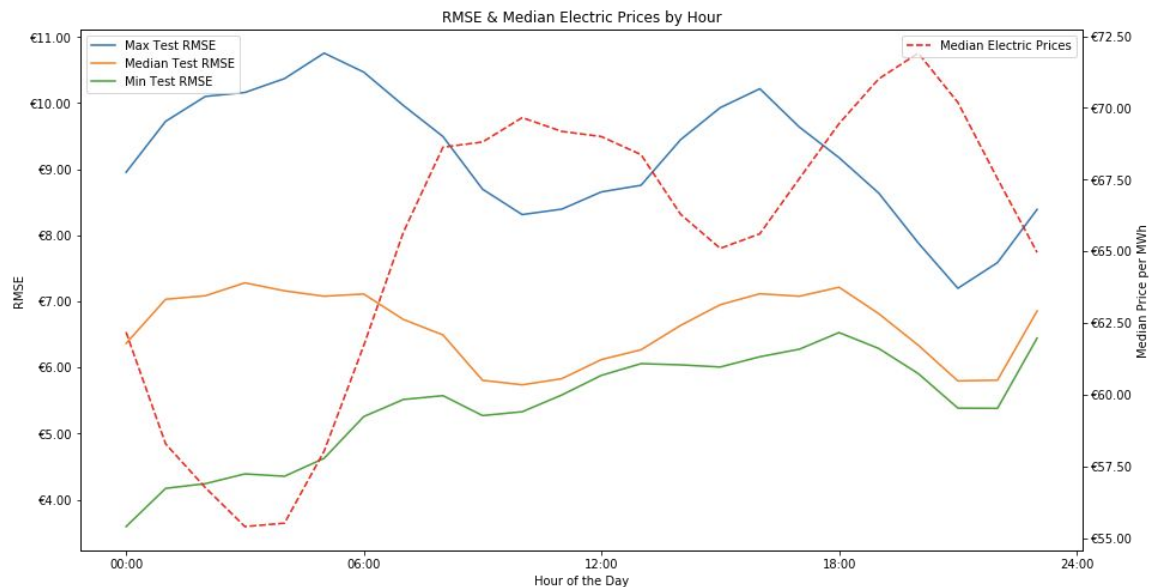
Summary statistics,
best models, next
steps

Summary Statistics



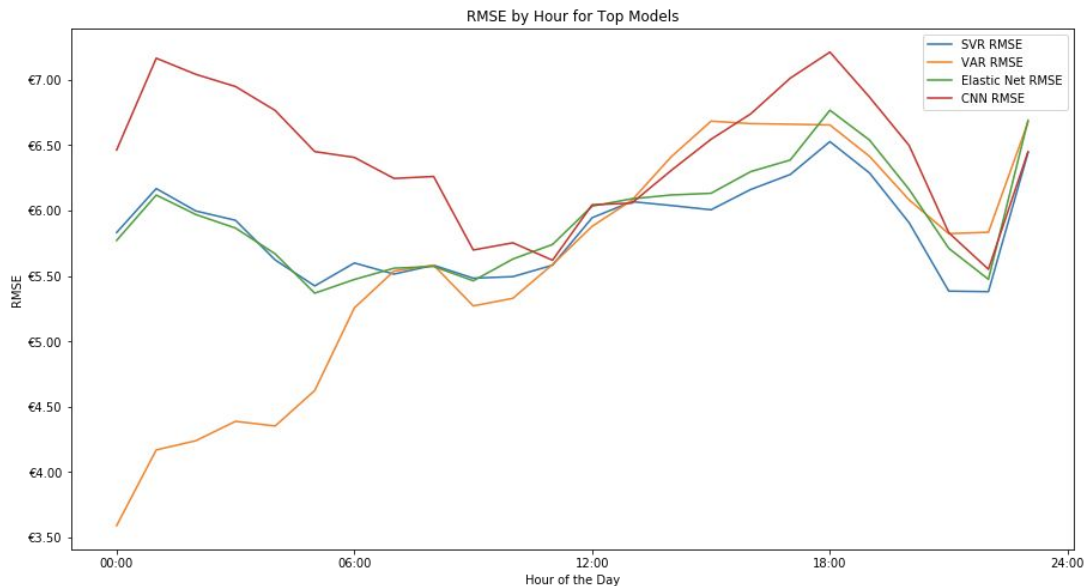
Performance by Hour

Interestingly our models generally performed worse (higher root mean square error (RMSE)) during times of the day when median electric prices were lower



Best Models

The VAR model performed very well early in the day but got progressively worse while the SVR was relatively steady and was slightly better than other models later in the day



Current Status

- Best model that combines VAR (first 12 hours of the day) and SVR (next 12 hours) is significantly better than the baseline
- While having significantly improved upon the baseline our best model is still not at a level where it could be used in production

Model	Baseline	SVR	VAR	VAR-SVR
Full Test RMSE	€13.25	€5.87	€5.65	€5.49
Full Test r2	0.130	0.769	0.786	0.798

Next Steps

Continued Research

- I only started looking deeply at this market recently, more research into electric markets & modeling techniques is needed

Model & Feature Improvements

- Weather & additional fuel data
- Incorporate connected grids
- Continued iteration & testing of models

Online Learning, Production Code & App / Website

- Continuously learning models
- Jupyter notebook - pure python code with connected data feeds

Real-time, Medium-term & Long-term forecasting

- Develop forecasting on other relevant timelines