Dimitrios Roussis

Evaluation of deep learning architectures for short-term price forecasting in the Spanish electricity market

Data Science & Information Technologies
Deep Neural Networks
Fall 2019-2020

<u>Instructors:</u>
Haris Papageorgiou &
Athanasia Kolovou

HELLENIC REPUBLIC
National and Kapodistrian
University of Athens

Scope and Contents

The purpose of this study is to evaluate and compare the performance of 5 different deep learning models for the prediction of the next hour's electricity price in the Spanish market (XGBoost is used as the baseline).

We also compare **(i)** Univariate vs. Multivariate implementations of the models
(ii) Different number of previous values (3, 10, 25) that are fed into the models

- Case Study
- Data Analysis (3 slides)
- Evaluation Framework
- Deep Learning (2 slides)
- Results (2 slides)
- Future Research Questions

Case Study

We analyze 2 datasets which contain **35,064** hourly observations from 01/01/2015 up to 31/12/2018 (4 years):

- ‘**weather_features.csv**’: Contains information about the temperature, humidity, pressure, wind speed, etc. for 5 major cities of Spain.

- ‘**energy_dataset.csv**’: Contains information about the electricity price, the total load of the energy grid and the amount of electricity generated by different energy sources (e.g. nuclear, solar, wind, biomass, oil, gas, etc.).

Weather information for 5 major cities of Spain: **Madrid, Barcelona, Seville, Valencia & Bilbao**

The power sector has been deregulated since the enactment of Electric Power Act 54/1997

Data Analysis (1/3)

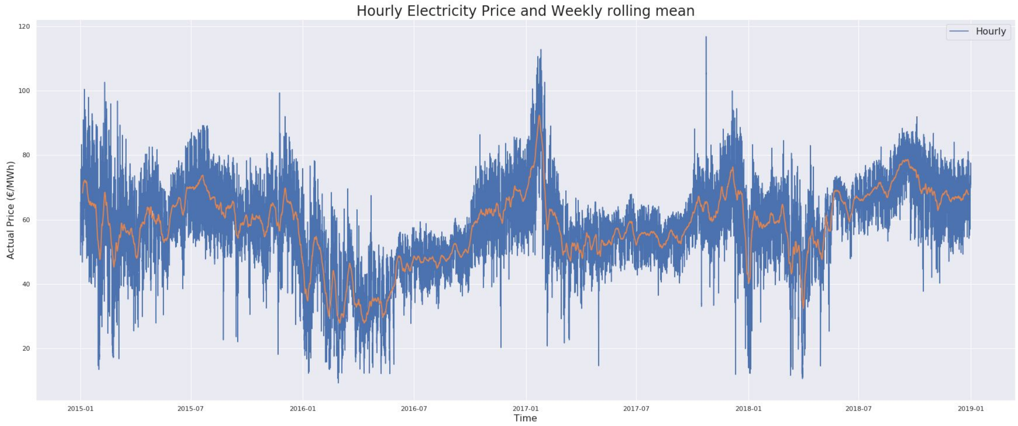
Important cleaning and preprocessing steps that we did for both datasets:

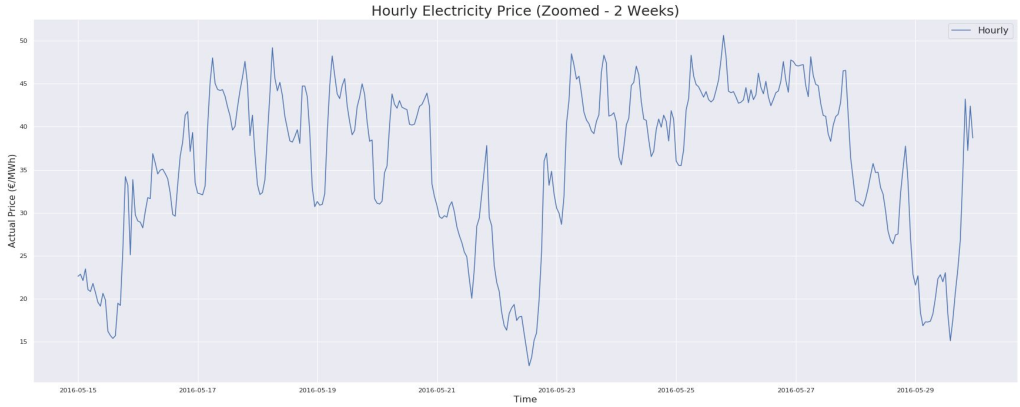
- Parse the dates correctly.
- Remove **duplicate** rows.
- Drop the inconsistent features (e.g. qualitative descriptions of the weather).
- **Fill NaNs** with the use of linear interpolation.

Important analysis steps that we did for the electricity price time-series:

- Visualize the series and **detect seasonal patterns**.
- Check for **stationarity** (Augmented Dickey-Fuller and KPSS tests).
- Plot Autocorrelation and Partial Autocorrelation functions in order to detect **which previous time-lags** of the electricity price are more important in forecasting the next hour's price.

Data Analysis (2/3)





The time-series of the electricity price reveals various seasonal patterns.
If we zoom into weekly scales (**start:** Sunday 15/06/2016; **end:** Sunday 29/06/2016), we can notice that it is:

- Higher during business days and lower during weekends.
- Higher during the day and lower during the night.
- Dropping for a few hours during midday, probably due to the “siesta” (13:30 to 16:30).

Data Analysis (3/3)

Important feature engineering steps that we did for the multivariate forecasting framework:

- **Feature generation:** e.g. Business hours; Month; Weekend; Temperature range
- **Feature selection:** In order to choose which features to use, we computed
  (i) the **Pearson correlation** between each feature and the electricity price and
  (ii) the **F-score** (feature importance) given by **XGBoost** algorithm.
- **Automatic outlier removal:** We use **11 selected features** as inputs, after removing all data points which are **4 standard deviations away from the mean**.

Finally, in order to prepare the dataset for the artificial neural networks, we did the following:

- **Convert** time-series to a **supervised learning** problem, i.e. create multiple sequences of past values (X) and the next hour's electricity price (y).
- **Normalize** the data (X and y) to a range between 0 and 1.

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Evaluation Framework

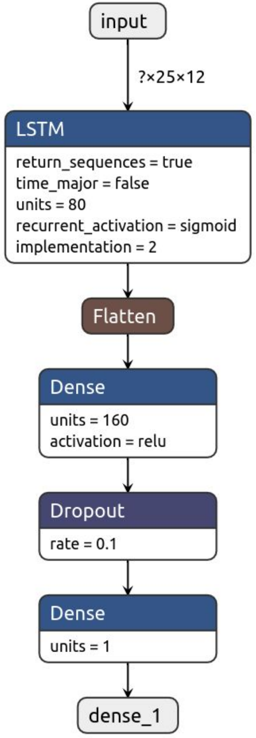
Train set: First 27,048 rows (1,461 days)
Validation set: Next 4,008 rows (167 days)
Test set: Last 4,008 rows (167 days)

- 5 different neural network architectures:
  Multilayer Perceptron, CNN, LSTM, Stacked LSTMs, Hybrid CNN-LSTM
- 2 modeling frameworks:
  Univariate (Only the previous values of the electricity price are given to the models) &
  Multivariate (The previous values of the 11 selected features are given as well)
- 3 different time-lags of previous values:
  We use the **3**, **10** and **25** previous values (time-lags) of the price / features as inputs.

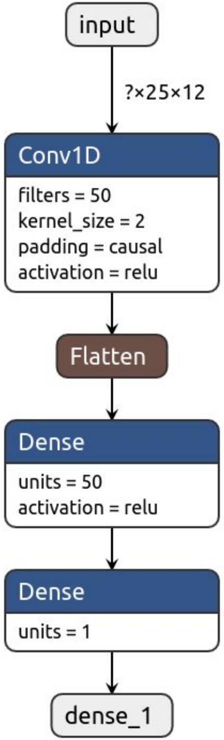
We calculate the *Root Mean Squared Error* (RMSE) between the real and the predicted values of the electricity price.

Persistence model (“naive” forecast) : **3.111** RMSE (€/MWhr)
XGBoost (multivariate) for each of the 3 time-lag choices

Deep Learning (1/2)



LSTM



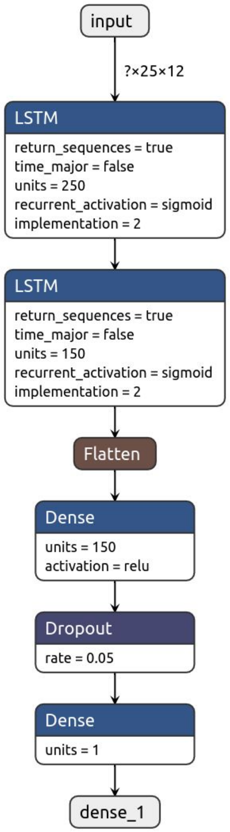
CNN

All the neural network architectures were implemented with Keras on top of TensorFlow 2.

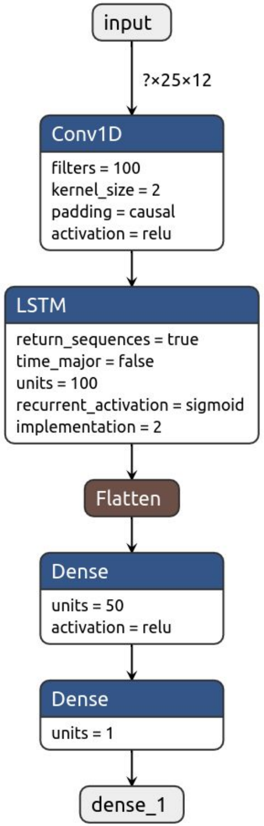
Common parameters and configurations:

- **Optimizer**: Adam (AMSGrad variant)
- **Loss**: Mean Squared Error
- **Batch Size**: 32 samples (845 steps / epoch)
- **Early Stopping**: 10 epochs (max 120)
- **Model Checkpoint** (saves the best model)
- Define auxiliary model which makes use of the **Learning Rate Scheduler** (gradual increase from 0.0001 to 10) and choose the learning rate in the flat plateau region (stable val. loss)

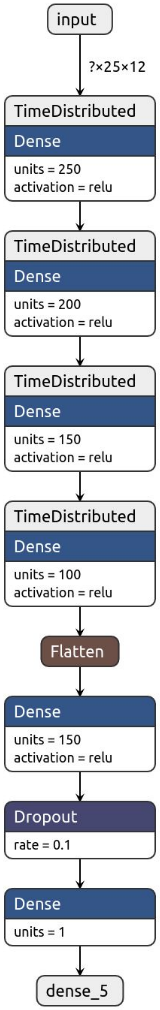
Deep Learning (2/2)



Stacked LSTMs



CNN - LSTM



Multilayer Perceptron (Time Distributed)

Results (1/2)

(i): All models outperform the persistence model (RMSE: 3.111)
(ii): XGBoost outperforms all the univariate models for 3 & 10 lags.
(iii): We get better performance when we use more time-lags; **but** multivariate models perform better with 3 lags than with 10.
Exceptions: CNN-LSTM & MLP
(iv): In each time-lag configuration, a multivariate model always has the best performance.
Best: **CNN** (25) - **1.955** RMSE
(v): More information does not lead to better performance necessarily!

Lags	Framework	LSTM	Stacked LSTM	CNN	CNN - LSTM	MLP	XGBoost
3	Univariate	2.566	2.561	2.556	2.523	2.532	2.373
	Multivariate	2.159	**2.148**	2.170	2.264	2.351	
10	Univariate	2.478	2.370	2.402	2.384	2.400	2.349
	Multivariate	2.277	2.219	2.263	**2.099**	2.253	
25	Univariate	2.052	1.981	1.990	2.003	2.001	2.181
	Multivariate	2.047	2.100	**1.955**	2.034	2.024	

RMSE of the electricity price forecast for all experiment configurations

The models with the best performance for each different configuration of time-lags are highlighted in **bold** and the best ones for each framework are underlined.

Results (2/2)

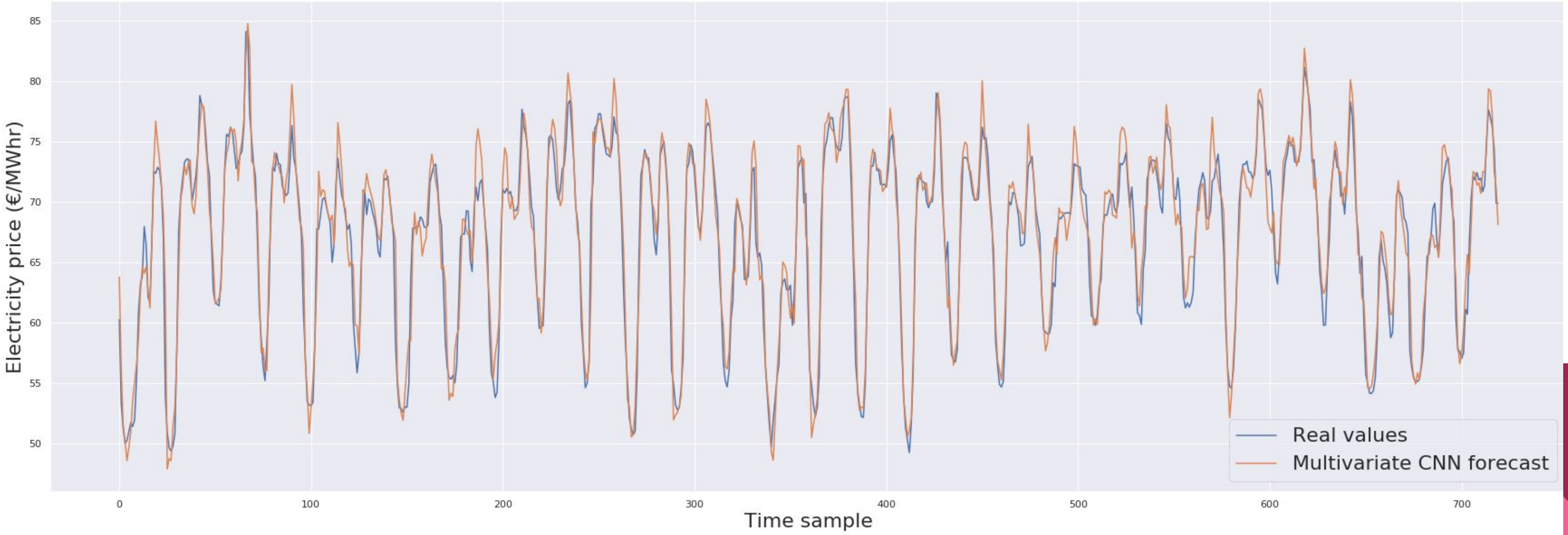
Best performance:

Multivariate CNN with 25 time-lags

1.955 RMSE (€/MWhr)

Below: Predictions for the **last month** (12/2018)

A real-time forecasting model can prove valuable for electric power companies, investors, public institutions and more generally, to all actors / stakeholders involved in the energy market.



Future Research Questions

- **A:** More information does not lead to better results necessarily.
  Q: Select specific past values of the features based on their importance?

- **A:** Complex model architectures do not achieve remarkably better results.
  Q: Automated fine-tuning of their hyper-parameters? Better preprocessing?

- **A:** Evaluation of performance only on the last months of the time-series.
  Q: Multiple train-test splits? The golden standard: Walk forward validation?

- **A:** Prediction of the electricity price just for the next hour.
  Q: Implement multi-step forecasting models?

Thank you for your time!

Questions? Comments? Ideas?

Do not hesitate to contact me:
droussis@outlook.com.gr