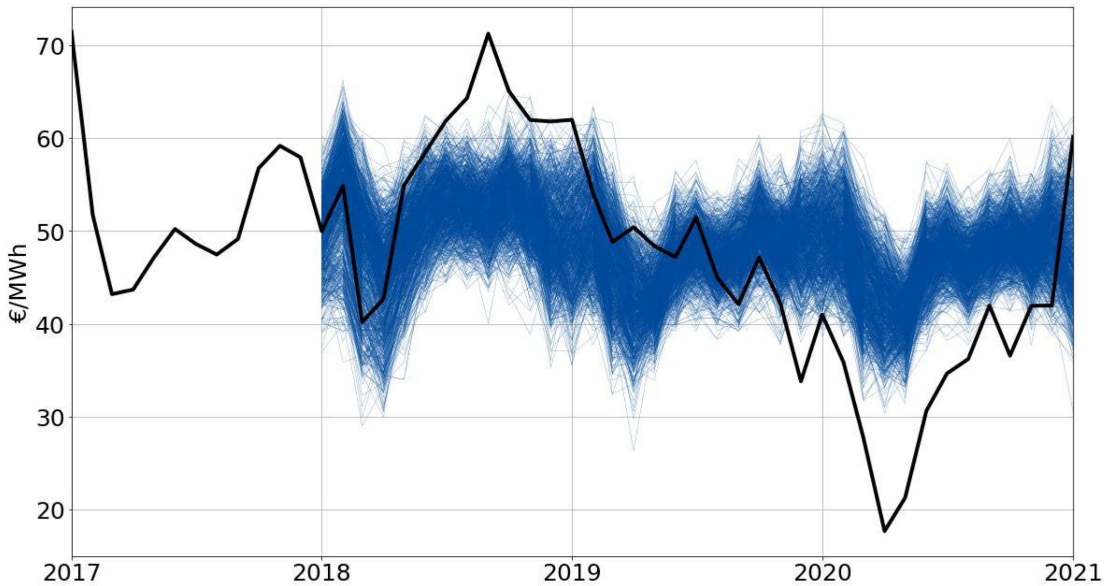
ELECTRICITY PRICE PREDICTION

Applied Data Science Phase-2 Document

Problem statement:



To develop a model that can predict electricity prices for a specified future period based on historical data and relevant features, in order to enable stakeholders like utility companies, regulators, and consumers to make informed decisions.

Problem definition:

- Leverage data science techniques to forecast electricity prices based on historical and contextual data, aiming to improve the decision-making processes for utility providers, consumers, and policymakers.
- Predict the average wholesale prices of electricity for the next quarter/year, assisting utility companies in their long-term procurement strategies and contract negotiations.

Design Thinking:

Applying design thinking to the electricity price prediction problem means
focusing on a human-centered approach to address the challenges and needs
associated with forecasting electricity prices.

Data Source:

For the problem statement of electricity price prediction, several data sources are crucial to ensure accurate, reliable, and holistic forecasts. These data sources encompass both historical and real-time data, as well as various auxiliary factors influencing electricity prices.

Data preprocessing:

Data preprocessing is a crucial step in any predictive modeling
project. For electricity price prediction, the data collected might be vast and
varied, and it's essential to ensure that this data is clean, consistent, and ready
for modeling.

Feature Engineering:

Feature engineering is a pivotal step in the predictive modeling
process. It involves creating new features from the existing ones, capturing

additional information or patterns that can enhance the model's predictive
performance.

Model selection:
Model selection for electricity price prediction should take into account the
nature of the data (time series), the complexity of the problem, and the
desired forecast horizon (short-term vs. long-term)

Model Training:
Model training for electricity price prediction involves taking the
preprocessed data and the chosen model (or models) to learn the underlying
patterns within the data

Evaluation:
Evaluation is a critical phase in the model development process, especially for
a problem like electricity price prediction. Proper evaluation not only
measures the performance of your model but also gives insights for potential
improvement

New innovation:

Statement:

Consider exploring more advanced time series forecasting techniques like Prophet or deep learning models for improved accuracy in predicting future electricity prices.

Innovating the Approach:

Certainly! Forecasting electricity prices is a challenging problem due to the complex interplay of various factors, such as demand and supply, generation mix, regulations, and unexpected events. Here's a quick dive into two advanced time series forecasting techniques: Prophet and deep learning models.

1. Prophet:

Prophet is a forecasting tool developed by Facebook. It's designed for forecasting at scale and is especially robust in the presence of missing data and outliers.

2. Deep Learning Models:

Deep learning models, particularly **Recurrent Neural Networks (RNNs)** and **Long Short-Term Memory networks (LSTMs)**, have been used successfully in time series forecasting

.

3.State Space Models & the Kalman Filter:

.

- **Description**: State space models are a powerful class of time series models. They break down the observation into its underlying states using the Kalman Filter.
- **Use Cases**: When there's a need to account for regime shifts or sudden structural changes in the data.

.

4.ARIMA-X:

.

- **Description**: An extension of the ARIMA (AutoRegressive Integrated Moving Average) model, ARIMAX models include external variables (X).
- **Use Cases**: When there are known external factors (like weather, special events) influencing electricity prices.

.

5.Vector Autoregression (VAR) & Vector Error Correction Model (VECM):

.

- **Description**: VAR is used for multivariate time series forecasting. VECM is an extension that considers long-run equilibrium relationships.
- **Use Cases**: When there are multiple interrelated time series data to consider. For example, forecasting electricity prices based on both demand and supply time series.

.

6.Exponential Smoothing State Space Model (ETS):

.

- **Description**: Considers error, trend, and seasonality components. Holt-Winters is a popular variant.
- **Use Cases**: When the data has clear trends and seasonality.

.

7.GARCH (Generalized AutoRegressive Conditional Heteroskedasticity):

.

- **Description**: Used primarily for modeling financial time series data where volatility is important.
- **Use Cases**: If the volatility of electricity prices is of concern, GARCH can capture those dynamics.

.

8.Gaussian Processes (GP):

.

- **Description**: GPs are a non-parametric method that provides a distribution over possible functions given the data.
- **Use Cases**: When there's uncertainty in the predictions and you want to quantify it.

.

9.Hybrid Models:

.

- **Description**: Combines two or more forecasting techniques. For example, ARIMA residuals can be modeled with a GARCH model (ARIMA-GARCH).
- **Use Cases**: When a single model doesn't capture all the nuances in the data.

10.Ensemble Methods:

- **Description**: Combine multiple forecasting models to generate a final forecast, typically yielding better accuracy.
- **Use Cases**: When you have multiple models and want to benefit from the strengths of each.

For electricity price prediction, incorporating external data (e.g., weather data, demand-side data, etc.) is often beneficial. These exogenous variables can help account for the drivers of electricity prices that pure time series models might miss. Proper feature engineering, understanding the problem's domain, and cross-validation for time series (e.g., time series split) are crucial to getting the best results.

Conclusion:

In conclusion, while challenges abound, the benefits of accurate electricity price prediction are manifold. Harnessing the power of modern data science and machine learning techniques, combined with domain expertise and continuous iteration, will pave the way for robust and efficient solutions that cater to the evolving needs of the energy sector.

.