Technical Indicators and Prediction for Energy Market Forecasting

***Abstract*—Machine learning usage for forecasting is popular in financial trading, particularly for stock price prediction and this is often combined with technical indicators to extract key predictive indicators from large time series trading datasets. Energy market trading data have similar characteristics to financial trading data, therefore deriving technical indicators specifically for electricity prices will help predict future prices and reduce trading costs. We have derived eight technical indicators for the Integrated Single Electricity Market (ISEM) energy market in Ireland using hourly electricity price data over the period February 2019 until November 2019. Technical indicator based models were obtained by using machine learning regression algorithms (Extreme Gradient [XG] Boost, Random Forest, and Gradient Boosting) trained with the proposed novel technical indicators. The results of the technical indicator models were compared against the baseline model (raw price data only) to see if using technical indicators as inputs improves model performance. We conclude that electricity prices can be accurately predicted using the proposed technical indicators.**

Keywords—Electricity Price Forecasting, Short-term, Time Series, Technical Indicators, Machine Learning

I. Introduction

Energy market prices are difficult to predict as the data exhibit nonlinear and nonstationary characteristics [1]. Price forecasting is necessary in the energy market to optimise purchasing in response to demand and supply fluctuations, and this can be managed through time series models which observe previous prices and help forecast future market prices. Time series models are algorithms that analyse the pattern of a series. Future values can be predicted by training the model with historical data and accounting for previous market trends to optimise trading. These models often perform best when the forecasting window for prediction is small, controlling volatility as the relationship between actual and predicted values is stronger over a short timeframe [2]. Therefore, short-term time series models are promising for electricity price forecasting. This research focusses on day-ahead energy price forecasting with the aim of building an innovative system that helps electricity suppliers to reduce their electricity generation purchase costs in the long term.

Electricity price forecasting is influenced by two approaches: fundamental (economic variables) and technical (derived from raw historical data) [3]. Technical indicators have been widely applied as inputs to financial trading market data to determine relationships when predicting future stock prices and trends [4]. Previous literature on financial trading has demonstrated that for short-term price forecasting, technical approaches are more appropriate than raw market data [5]. Here, development of technical indicators specifically for electricity pricing and application of these indicators as inputs in day-ahead forecasting models is used to demonstrate how historical behaviours of electricity price can be used to forecast future values. Existing literature on technical indicators suitable for day-ahead electricity price is relatively limited [6], therefore this paper presents new progressive findings for electricity price forecasting technical indicators.

On developing new technical indicators, we examine those that contribute to short-term price prediction using machine learning models with technical indicators as model inputs and actual electricity price as model output. The technical indicators are further utilised to train machine learning prediction models to forecast day-ahead electricity price and model performance for all techniques is evaluated. For comparison, and to set in context the accuracy of the derived technical indicators, additional machine learning models including only raw price data as input were analysed and are referred to as persistence models.

This paper is structured as follows: Section II outlines the proposed technical indicators for the energy market, detailing how each is calculated in terms of electricity price and Section III describes the proposed modelling techniques. The results are presented and discussed in Section IV highlighting the accuracy of each model and the significance of each technical indicator. Section V concludes with a summary of the key findings and possible future work in predicting electricity prices.

II. Technical Indicators

In stock market trading, an appropriate tool for finding information on upcoming share price movement is technical analysis, which builds indicators from the raw price data to capture trends over time [7]. The Integrated Single Electricity Market (ISEM) is a new advancement that gives energy traders in the whole of Ireland greater flexibility and control over electricity prices. This has led to the design of novel technical indicators for ISEM electricity price forecasting to help energy traders observe market price trends and predict when to buy or sell electricity. Deciding on parameters optimisation is a key problem with respect to technical analysis [7]. When considering optimisation of parameters, the sliding window size, corresponding to the number of historical values required to calculate each of the technical indicators from the raw data, is of key interest [8]. There are various types of technical analysis to observe price patterns, but the main indicators can be split into three types: (i) trend, (ii) oscillator, and (iii) momentum [9].

This work presents eight novel technical indicators specific to energy trading derived from, but not identical to, the standard indicators applied to financial trading. There are multiple technical indicators frequently used in financial forecasting, however this work is solely interested in price indicators which improve day-ahead accuracy as that is the typically requirement from the energy market in which we wish to make predictions. The calculation for each of the individual technical indicators is listed below:

1. Percentage Price Change Moving Average (PPCMA): A trend indicator in time-series that, for the energy market, we calculate price change as the difference between the current price (Hour *n*) and the price from the same time period the day before (Hour *n* Lag 24), all divided by the price at Hour *n* Lag 24. The moving average percentage price change was calculated for is a rolling 24-hour window:

$$PPCMA = \sum_{24} [PPC] \quad (1)$$

where

$$PPC = \frac{Price_{Hour\ n} - Price_{Hour\ n\ Lag\ 24}}{Price_{Hour\ n\ Lag\ 24}} * 100 \quad (2)$$

2. Moving Average Deviation (MAD): A trend indicator that utilises the PPCMA indicator to calculate the deviation rate of the current electricity price from PPCMA:

$$MAD = \frac{Price_{Hour\ n} - PPCMA}{PPCMA} \quad (3)$$

3. Percentage Range (PR): An oscillator indicator that finds a relationship between current electricity price and the highest/lowest prices over a 24-hour window. This indicator oscillates between 0 and 100, with a value above 80 determined to indicate energy units are oversold and a value below 20 indicating that energy units are overbought:

$$PR = \left[\frac{HighPrice_{n\ to\ Lag\ 24} - Price_{Hour\ n}}{HighPrice_{n\ to\ Lag\ 24} - LowPrice_{n\ to\ Lag\ 24}}\right] * 100 \quad (4)$$

4. Average True Range (ATR): A trend indicator measuring price volatility. Over a 24-hour window there are three different values calculated: highest price over the 24-hour period minus lowest price over the 24-hour period; highest price over the 24-hour period minus starting electricity price; and lowest price over the 24-hour period minus starting electricity price. The maximum value from these three values is selected for each trading hour and averaged over a rolling 24-hour window.

$$ATR = \sum_{24} MAX\ [\ A, B, C\ ] \quad (5)$$

$$A = HighestPrice_{n\ to\ Lag\ 24} - LowestPrice_{n\ to\ Lag\ 24} \quad (6)$$

$$B = |\ HighestPrice_{n\ to\ 24} - Price_{Hour\ n\ Lag\ 24}\ | \quad (7)$$

$$C = \left|\ LowestPrice_{n\ to\ 24} - Price_{Hour\ n\ Lag\ 24}\ \right| \quad (8)$$

5. Relative Strength Index (RSI): An oscillator indicator that compares recent price gains to recent price losses. This indicator oscillates between 0 and 100, with a value over 70 determined to indicate that energy units are overvalued and a value below 30 indicating that energy units are undervalued. Price Up is the average of the previous 24 hours when price difference increased. Price Down is the average of the previous 24 hours when price difference decreased.

$$RSI = 100 - \left[\frac{100}{D}\right] \quad (9)$$

where

$$D = \left(1 - \frac{\sum_{24} Price\ Up\left[Price_{n} - Price_{Hour\ n\ Lag\ 24}\right]}{\sum_{24} Price\ Down\left[Price_{n} - Price_{Hour\ n\ Lag\ 24}\right]}\right) \quad (10)$$

6. Average Directional Movement Index (ADX): A trend indicator measuring the strength of the trend, grouping the two directional movement indexes depending whether price change, calculated as current electricity price minus previous 24-hour price, is grouped as a Price Up (positive) change or Price Down (negative) change. The two indexes are combined and smoothed with a moving average.

$$ADX = \left[\frac{\sum_{24} DX\ Up(a) - \sum_{24} DX\ Down(b)}{\sum_{24} DX\ Up(a) + \sum_{24} DX\ Down(b)}\right] * 100 \quad (11)$$

where

$$a = \frac{\sum_{24} Price\ Up[Price_n - Price_{Hour\ n\ Lag\ 24}]}{ATR} \quad (12)$$

$$b = \frac{\sum_{24} Price\ Down[Price_n - Price_{Hour\ n\ Lag\ 24}]}{ATR} \quad (13)$$

7. Moving Average Convergence/Divergence (MACD): An oscillator indicator that considers the strength, direction, and duration of the trend as well as price momentum through moving averages of previous price values with rolling window sizes of 12 and 24.

$$MACD = \sum_{12} Price\ MA_{Hour\ n\ Lag\ 1\ to\ Lag\ 12} - \sum_{24} Price\ MA_{Hour\ n\ Lag\ 1\ to\ Lag\ 24} \quad (14)$$

8. Price Momentum (PMOM): A momentum indicator that measures the power of the market by observing the current electricity price with the previous trading value (1 hour before).

$$PMOM = Price_{Hour\ n} - Price_{Hour\ n\ Lag\ 1} \quad (15)$$

For all technical indicators that include a moving average, this was derived from the previous *n* hours (with hourly prices) for the required sliding window size and does not include the current hour's value.

III. Modelling Methodology

Technical indicator models were obtained using three machine learning algorithms, implemented using SkLearn, and trained with the novel technical indicators. A Random Forest regression algorithm trains multiple decision trees, split at nodes into partitions, so that no individual tree observes the full training data [10]. Due to the setup of the Random Forest it is an efficient non-linear technique that avoids over-fitting [11]. It also shows transparency as a tuning parameter manages the amount of input features in each node by selecting when to divide the input data to create a new classifier [12]. Training multiple trees allows an overall ranking of feature importance which improves in accuracy as more trees are included in the Random Forest [13].

Boosting algorithms, through sequential learning [14], create strong learners with an error rate close to zero by combining weak learner models and converting them to a strong learner model [15]. The Gradient Boosting regression algorithm builds strong learners to minimize error residuals (difference between actual and predicted) by optimising a loss function from a collection of weak learners [11]. Extreme Gradient Boosting (XGBoost) is another regression algorithm based on decision trees, similar to Random Forest and Gradient Boosting algorithms, but with additional features [16]. XGBoost applies the framework of Gradient Boosting, but includes Newton boosting as a method for approximations and an extra randomisation parameter to reduce correlation among trees [15]. XGBoost is a popular and advanced machine learning algorithm that often performs better than other algorithms; this is due to its speed, multiple tuning parameters, and ability to train well on large datasets [17].

IV. Results

Hourly electricity price data from the ISEM day-ahead market were retrieved from the Single Electricity Market Operator (SEMOpx) website [18] and data ranged from 1st February 2019 until 30th November 2019. The technical indicators outlined in Section II were calculated using the raw electricity price data over a sliding window the size of which is dependent on the indicator, generally set to 24 hours since we are dealing with day-ahead electricity price forecasting. However, for weekly or monthly predictions the calculation window size could be increased to 168 or 744 hours respectively.

The calculated technical indicators were used as training inputs for the machine learning models to forecast future electricity prices. The data used to obtain the technical indicators for the machine learning models were split 85% for training (04th February 2019 until 16th October 2019) and 15% for model testing (17th October 2019 to 30th November 2019). When training all machine learning models, the input variables were the eight technical indicators and the output was the actual electricity price aligned with the input values at time T. First the persistence model T+24 was created using the raw price data as input with the same train/test split and examined as a baseline to predict the test set historical electricity price data. A persistence model predicts future values on the assumption that conditions from the current time to the future time stay unchanged [19]. For the persistence machine learning models, the input variable was price at time T and the output variable was price at time T+24. The summary results of both training and testing from the persistence models obtained using Random Forest, Gradient Boosting and XGBoost are shown in Table I. The model accuracy of all the persistence models during training and testing ranged between 73% and 80%.

TABLE I. SUMMARY RESULTS FOR PERSISTENCE MODELS

Model	Training Accuracy	Training RMSE	Testing Accuracy	Testing RMSE
Gradient Boosting	74.21%	12.55	75.44%	13.18
Random Forest	80.06%	7.29	73.21%	16.30
XGBoost	75.95%	10.60	75.02%	14.39

The Root Mean Square Error (RMSE) can be used to evaluate overall model accuracy for the test set, in which the closer the value is to zero the better the model performance. This is computed in terms of electricity price as the difference in the actual electricity price and the predicted electricity price value. The Random Forest model generated the lowest RMSE value of 7.29 during model training (Figure 1) and Gradient Boosting provided the lowest RMSE value of 13.18 during model testing with unseen data (Figure 2). Both Figure 1 and Figure 2 plot the actual price and model predicted prices. The figures demonstrate that the model predicted price closely follows the same trend as the actual prices however, the prediction, particularly for model testing, is not very accurate.

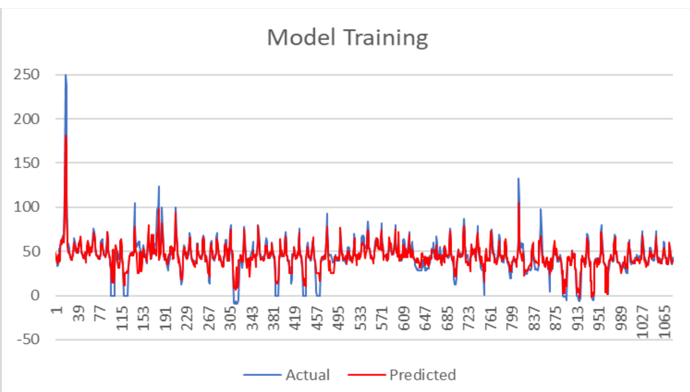

Fig. 1. Random Forest Persistence Model Training

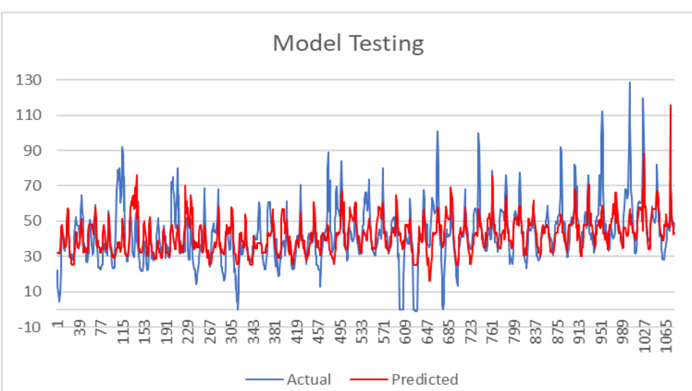

Fig. 2. Gradient Boosting Persistence Model Testing

With the baseline as a reference point, the machine learning algorithms were trained using the technical indicators as inputs and actual price as output. Predictive performance was then compared using the testing data. Summary results for both training and testing and all three machine learning models are presented in Table II. Overall model accuracy ranged from 84% to 93%. During the training stage the RMSE was below 3 for both Random Forest and XGBoost. Using testing data, all models had a RMSE of approximately 6.8 or less, which is much lower than the persistence model findings and hence demonstrates improved performance. The Random Forest algorithm performed best when observing the training results; however, testing is the key stage with unseen data and even though Random Forest had higher testing accuracy, XGBoost had the lowest RMSE and therefore is considered the optimal algorithm in the comparison.

TABLE II. SUMMARY RESULTS FOR TECHNICAL INDICATORS MODELS

Model	Training Accuracy	Training RMSE	Testing Accuracy	Testing RMSE
Gradient Boosting	84.82%	5.89	86.90%	6.66
Random Forest	93.22%	2.34	91.57%	6.77
XGBoost	92.68%	2.45	89.70%	5.34

Both the model training output (Figure 3) and the model testing output, with unseen data (Figure 4) exhibited a well-fitted model compared to the baseline model. The visual output shows a close fit between the actual price and the predicted price. The RMSE also demonstrates the overall accuracy of the models using technical indicators where the Random Forest at the training stage had a RMSE of 2.34 and XGBoost model testing stage had a RMSE of 5.34.

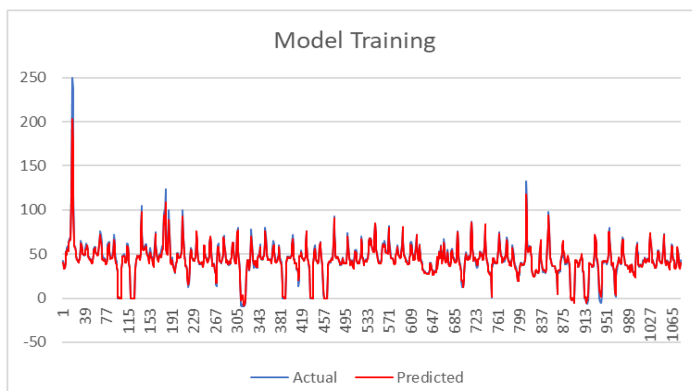

Fig. 3. Random Forest Technical Indicator Model Training

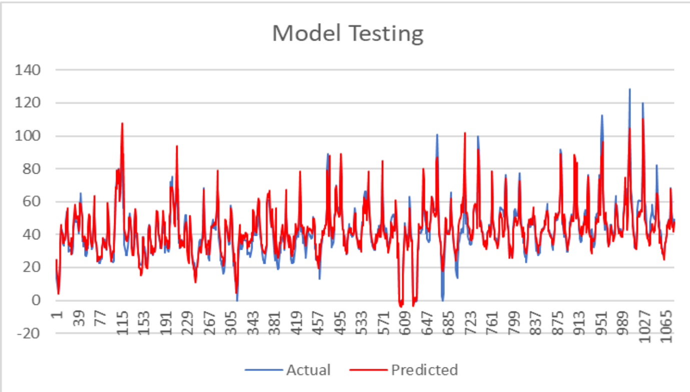

Fig. 4. XGBoost Technical Indicator Model Testing

The XGBoost model showed promising results during model testing therefore we tested the model further using additional technical indicators input data to predict the next day electricity price values (01st December 2019). The results of the prediction are illustrated in Figure 5. The forecasted price has a good fit with the actual price resulting in an overall RMSE of 5.75.

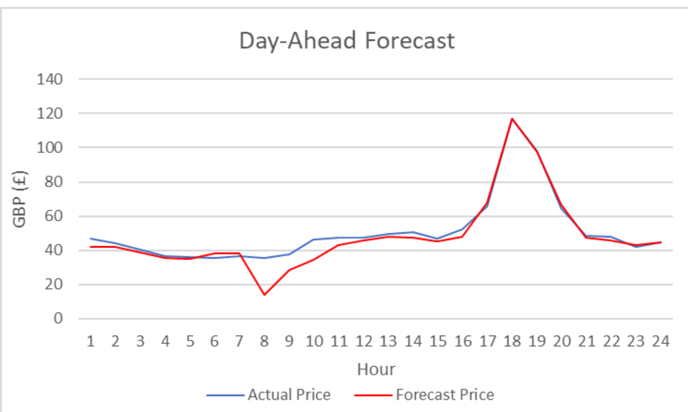

Fig. 5. XGBoost 01st December 2019 prediction

Finally, we analysed the significance of each of the technical indicators through feature importance which is an inbuilt function in SkLearn for regression trees with output ranging from 0 to 1; the closer the value is to 1 the greater the relevance the feature is to predicting the target price variable. Table III displays the feature importance score (coefficient value) of each individual technical indicator for all three machine learning models. From the results, the feature importance score for each of the technical indicators varies in terms of the model's performance. From observing the feature importance output we have considered any score above 0.1 to be important to the price prediction. For Gradient Boosting there were three significant indicators above 0.1 (MAD, PPCMA, and Percentage Range). Random Forest had three significant indicators above 0.1 (Percentage Range, MAD, and ATR). There were three significant indicators above 0.1 (Percentage Range, Price Momentum, and MAD) for XGBoost. MAD and Percentage Range were important for all three techniques suggesting that both are good indicators for electricity price prediction. ADX, MACD, and RSI scores were not important for any of the three techniques and therefore these technical indicators were considered insignificant.

TABLE III. FEATURE IMPORTANCE OF TECHNICAL INDICATORS

Technical Indicator	**Gradient Boosting**	**Random Forest**	**XGBoost**
PPCMA	0.1891	0.0751	0.0625
MAD	0.3055	0.1639	0.1215
Percentage Range	0.1538	0.4229	0.4016
ATR	0.0989	0.1085	0.0798
RSI	0.0777	0.0613	0.0852
ADX	0.0368	0.0225	0.0459
MACD	0.0474	0.0518	0.0651
Price Momentum	0.0906	0.0940	0.1382

V. CONCLUSION

We have developed novel technical indicators specifically derived for energy trading to predict day-ahead electricity prices using various transparent machine learning models. Raw electricity price data from 2019 were collected and the eight technical indicators (PPCMA, MAD, Percentage Range, ATR, RSI, ADX, MACD, and Price Momentum) where calculated over this dataset.

Machine learning algorithms (Gradient Boosting, Random Forest, XGBoost) were implemented to forecast short-term electricity price. Model data were split 85% for training and 15% for testing. First the algorithms only included raw price data as a baseline model (the persistence model). For all three algorithms, the persistence model accuracy was 73% or greater. Random Forest performed the best during model training and Gradient Boosting performed the best during model testing.

The models generated using the technical indicators presented performance accuracy results of at least 84% accuracy showing significantly improved model performance. Again, Random Forest had the lowest RMSE during model training but during model testing XGBoost performed the best when predicting electricity prices. The machine learning models including technical indicators accurately matched predicted prices compared with the persistence models, with both the model training and model testing stages displaying close predictions in comparison to the actual historical price. With promising testing output using the technical indicators, the XGBoost model was then tested to predict next day electricity prices (01st December 2019) and the forecasted results were able to match the pattern of the actual values very accurately outputting a RMSE of 5.75.

Finally, the feature importance for each of the machine learning algorithms was examined to find which technical indicators were most relevant to prediction for the energy price market. From the model output, five technical indicators (PPCMA, MAD, Percentage Range, ATR, and Price Momentum) were determined to be significant for at least one machine learning technique. Both MAD and Percentage Range were significant for all techniques. ADX, RSI, and MACD were considered insignificant as for all techniques their scores were below 0.1.

To conclude, technical indicators for electricity price are beneficial in forecasting future day-ahead prices and should be considered for energy trading to help follow market trends and, over time, lower price costs. Further work will explore technical indicators in more detail, in particular to include other energy factors (wind, temperature, load, etc.) to determine if model accuracy can be further improved. Another area to consider going forward is having a separate model for each of the 24 hours when calculating the technical indicators as it would be helpful in capturing market trends.

REFERENCES

[1] H. Mosbah and M. El-Hawary, "Hourly Electricity Price Forecasting for the Next Month Using Multilayer Neural Network," in *Canadian Journal of Electrical and Computer Engineering*, 2016, vol. 39, no. 4, pp. 283–291.

[2] N. Amjady and M. Hemmati, "Energy price forecasting: Problems and proposals for such predictions," in *IEEE Power and Energy Magazine*, 2006, vol. 4, no. 2, pp. 20–29.

[3] N. Pandey and K. G. Upadhyay, "Different price forecasting techniques and their application in deregulated electricity market : A comprehensive study," in *International Conference on Emerging Trends in Electrical , Electronics and Sustainable Energy Systems (ICETEESES)*, 2016, pp. 1–4.

[4] E. A. Gerlein, et al., "Evaluating machine learning classification for financial trading: An empirical approach," in *Expert Systems with Applications*, 2016, vol. 54, pp. 193–207.

[5] Y. Shynkevich, "Computational intelligence techniques for forecasting stock price movements from news articles and technical indicators," Thesis, 2016.

[6] S. Demir, K. Mincev, K. Kok, and N. G. Paterakis, "Introducing technical indicators to electricity price forecasting: A feature engineering study for linear, ensemble, and deep machine learning models," *Appl. Sci.*, vol. 10, no. 1, 2020.

[7] J. Diego, J. Ignacio, J. Francisco, and L. Jose, "Multiobjective Optimization of Technical Market Indicators," in *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference*, 2009, pp. 1999–2004.

[8] Y. Shynkevich, et al., "Forecasting price movements using technical indicators: Investigating the impact of varying input window length," in *Neurocomputing*, 2017, vol. 264, pp. 71–88.

[9] M. T. Ã and S. Tokuoka, "Adaptive use of technical indicators for the prediction of intra-day stock prices," in *Statistical Mechanics and its Applications*, 2007, vol. 383, pp. 125–133.

[10] L. Khaidem, S. Saha, and S. R. Dey, "Predicting the direction of stock market prices using random forest," vol. 00, no. 00, pp. 1–20, 2016.

[11] Q. Qin, Q.-G. Wang, J. Li, and S. S. Ge, "Linear and Nonlinear Trading Models with Gradient Boosted Random Forests and Application to Singapore Stock Market," *J. Intell. Learn. Syst. Appl.*, vol. 05, no. 01, pp. 1–10, 2013.

[12] K. Mulrennan, J. Donovan, D. Tormey, and R. Macpherson, "A data science approach to modelling a manufacturing facility's electrical energy profile from plant production data," *Proc. - 2018 IEEE 5th Int. Conf. Data Sci. Adv. Anal. DSAA 2018*, pp. 387–391, 2018.

[13] S. Jansen, *Hands-On Machine Learning for Algorithmic Trading*. Packt Publishing Ltd, Chapter 10, 2018.

[14] A. Paul *et al.*, "A real-time iterative machine learning approach for temperature profile prediction in additive manufacturing processes," *Proc. - 2019 IEEE Int. Conf. Data Sci. Adv. Anal. DSAA 2019*, pp. 541–550, 2019.

[15] R. Gandhi, "Gradient Boosting and XGBoost," 2018. [Online]. Available: https://medium.com/hackernoon/gradient-boosting-and-xgboost-90862daa6c77. [Accessed: 06-Jan-2020].

[16] S. Dey, Y. Kumar, S. Saha, and S. Basak, "Forecasting to Classification : Predicting the direction of stock market price using Xtreme Gradient Boosting Forecasting to Classification : Predicting the direction of stock market price using Xtreme Gradient Boosting," no. October, pp. 1–10, 2016.

[17] M. Pathak, "Using XGBoost in Python," 2019. [Online]. Available: https://www.datacamp.com/community/tutorials/xgboost-in-python. [Accessed: 06-Jan-2020].

[18] SEMOpx, "Day-Ahead Market data," 2019. [Online]. Available: https://www.semopx.com/market-data/market-results/. [Accessed: 10-Dec-2019].

[19] H. T. C. Pedro and C. F. M. Coimbra, "Assessment of forecasting techniques for solar power production with no exogenous inputs," *Sol. Energy*, vol. 86, no. 7, pp. 2017–2028, 2012.