Communication Papers of the 18th Conference on Computer Science and Intelligence Systems pp. 245–250
DOI: 10.15439/2023F8389
ISSN 2300-5963 ACSIS, Vol. 37

Ensemble-based versus expert-assisted approach to carbon price features selection

Bogdan Ruszczak, Katarzyna Rudnik
0000-0003-1089-1778, 0000-0002-0653-6610
Opole University of Technology ul. Prószkowska 76, 45-758 Opole, Poland
Email: {b.ruszczak, k.rudnik}@po.edu.pl

***Abstract*—The paper comments on two main issues. First, on a model for estimating the carbon price using multi-year market data. And second, on the consideration of two approaches to feature set exploitation. On the one hand, two ensemble machine-learning models with randomly selected feature sets are employed. On the other hand, a hybrid feature selection strategy follows domain expertise on which features should be explored. This minimizes the number of feature set combinations to be tested. The additional information for the predictions was the data from other commodity contracts, which could be easily introduced into the collection, as too many of them do not necessarily improve the estimates. The results of the experiments are promising: for the model based on SVR, the MAPE obtained was 2.09% and 5.6% for the following day and week price forecasts, respectively.**

I. Introduction

THE European Union Emissions Trading Scheme (EU ETS) was established in 2005 to promote the cost-effective and economically efficient reduction of greenhouse gas emissions. According to the International Energy Agency, global CO_2 emissions reached record highs in 2021 (over EUR 60 per ton) and the price is still volatile. The costs of European Union Allowance (EUA) is increasing, not only for the environment but also for the European economy. For this reason, understanding the problem of carbon price volatility and being able to predict it has become essential for profitable business decisions in companies that emit CO_2 and are obliged to buy carbon credits, as well as in companies that are considering switching to renewable energy sources. The literature analysis shows the breadth of the scope of the topic and the potential correlation of EUA price volatility with many factors [1], [11]. Therefore, there is a need to identify a carbon price prediction model using determinants that have a particular impact on the EUA price forecasting in a dynamically changing environment.

The article is a continuation of the recent research presented by the authors in [16]. This paper describes the day-ahead carbon price prediction based on a wide range of fuel and energy indicators traded on the Intercontinental Exchange market. In the proposed approach, by combining the Principal Component Analysis (PCA) method and various methods of supervised machine learning, the possibilities of prediction in the period of rapid price increases are shown. The PCA method reduced the number of variables from 37 to 4, which were inputs for predictive models, so it reduced the complexity of models but did not improve the prediction errors [16].

Following these considerations, in this paper we propose a hybrid approach for feature selection and identification of the carbon price prediction model. In this approach, we combine a wrapper and an embedded method using different supervised machine learning methods and different time horizons of EUA price forecasting. We attempted to employ the wrapper methods, which would potentially increase the predictive power of the model, and alternatively tested the embedded approach, which allows for automatic reduction of the feature space. However, after some initial testing for the described forecasting case, we combined both approaches and supplemented them with domain expertise on which features are valuable, thus helping to reduce the number of feature set combinations.

II. Literature Review

For the purpose of testing the considered feature selection approaches, we decided to run some real data experiments. This has been performed for the field of European Carbon Emission Allowance Futures (EUA) price forecasting, and machine learning methodology has been exploited for the required estimation generation.

Different approaches to EUA price forecasting can be found in the literature. According to [27], carbon price forecasting models can be divided into the following types of models: econometric prediction model, artificial intelligence algorithms and combined prediction model.

An approach proposed in [3] employed a non-parametric method to estimate carbon prices and found that the method could reduce the prediction error by about 15% compared to linear autoregression models. In [10] a hybrid model combining the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model and a long-term memory network was presented, while in [1] the authors proposed the GARCH models and the k-nearest neighbor models. There are many approaches to EUA price prediction using machine learning methods. In [28], the authors proposed a novel paradigm of multiscale nonlinear ensemble learning, involving empirical mode decomposition and a least squares support vector machine with a kernel function prototype. An extreme learning machine optimized by the Kidney algorithm with

This work was supported by National Science Centre, Poland funds awarded to Katarzyna Rudnik under grant no. 2021/05/X/ST6/01693.

Topical area: Information Technology for Business and Society

a coefficient of proportionality and cooperation is proposed in [9].

Due to the non-linearity and non-stationarity of EUA prices, the authors in [22] developed a system consisting of an analytical module and a forecasting module. There are relatively few publications that use determinant analysis in EUA price prediction models. In [11], a theoretical model was developed and presented that combines the energy sector (crude oil, natural gas, coal, electricity prices, and the share of fossil fuels in electricity generation), economic activity, and the market for CO_2 emission allowances. In [1], the authors suggested that Brent oil, coal, and electricity can be used to forecast the volatility of coal futures. In this paper, our research fills the gap in the related literature, where we take into account a wider range of data from the fuel and energy sector in order to perform feature selection and identify the EUA price prediction model in the short term. The short term prediction is especially important for traders and other market participants who, when buying EUA prices on a regular basis, follow price trends to buy EUAs at the cheapest price [15]. Prediction over a longer period of time can be useful in making strategic decisions for market participants and carbon-based companies. However, due to the nonstationary, nonlinear, and irregular EUA price, it is a particularly difficult issue that requires a more sophisticated approach to analysis. It seems that the solution in this respect could be models using the enoising procedure [5] and deep learning [4]. This will be the subject of further research.

III. Methodology

Fig. 1 presents a general overview of the applied research methodology that has been followed in this paper. We initially collected and prepared the data, which is discussed in section IV. As a result, we obtained time series for the EUA price and 16 factors related to the contracts for fuel and energy products (current values and values of the last change), which gives a total of 33 market-driven features. Since the data acquisition for the analyzed case could be costly, we tested several ways of feature selection that could lead to finding those that are really necessary and contribute to the final model performance. On the one hand, we used ensemble machine learning models with randomly selected feature sets. Providing the model with many features at the beginning, hoping that for some models this would be beneficial for the final result [14], [24]. Such a process was automated using so-called ensemble learning models, such as random forests or extra trees, which are reported by many to work well for datasets with many features [18]. These models could return even better estimates, outperforming Support Vector Machine based prediction [8]. And on the other hand, we used selected machine learning models and iteratively added successive model features, which were selected by an expert. For this part of the experiments aimed at investigating the importance of the features, as it has been advised [7], we applied linear models, support vector machines with linear kernel and simple linear regression models with additional regularization.

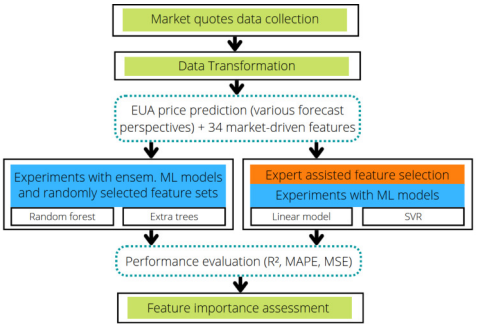

Fig. 1. The experiment evaluation procedure.

All the planned experiments were conducted under a recommended cross-validation regime [26], which allows for credible performance analysis. To evaluate the tested models, we employed several widely adopted metrics in the following experiments: the coefficient of determination (denoted as R^2), which is commonly used to compare the performance of different models [19], [21], the mean square error (MSE) or its' square root [20], the mean absolute percentage error (MAPE) [21] which should provide a good intuition on the relative scale average model's prediction error. For each tested model and each generated price prediction $p_{i,pred}$ and the actual reference prices $p_{i,ref}$ we checked the statistics according to the formulas:

$$MAPE = \frac{1}{n}\sum_i \left|\frac{p_{i,ref} - p_{i,pred}}{p_{i,ref}}\right| \quad (1)$$

$$R^2 = 1 - \frac{\sum_i (p_{i,ref} - p_{i,pred})^2}{\sum_i (p_{i,ref} - \overline{p_i})^2} \quad (2)$$

Finally, it is worth mentioning the software packages we used. We used the `Scikit-learn` framework [13] to develop the necessary machine learning models and to perform all the planned experiments. For the visualization of the results and the representation of the features, we used the `Matplotlib` library [2] and the `Seaborn` package [25], all in the `Python 3.10` environment.

IV. Market data collection and transformation

The research was carried out upon data collection that gathers daily carbon futures of the EU ETS from the Intercontinental Exchange market over a long period of time. The analyzed delivery date is December of the same year (for trade dates from January to October) and December of the next year (for trade dates from November to December). The data set comes from the Fixed Income Trading Analytics web portal [6] (accessed on 8 August 2022). To be specific it spans from 2013-10-22 to 2020-12-16. For all working days on which transactions were quoted during this period, we put a total number of 1842 rows into our data set in a standardized format. Missing data were replaced with the average factor prices for the last three days.

For the purpose of the EUA price modeling, we supplemented the collection with more than a dozen of additional

factors. We acquired data reflecting other fuel and energy factors from the Intercontinental Exchange, for the same time period as the target that could potentially provide useful information. Finally, the main series - f_1: "EUA Future" that shows the previous value of the modeled EUA prices, has been concatenated with the following series [6]:

- f_2: "AFR-Richards Bay Coal Future",
- f_3: "ATW-Rotterdam Coal Future",
- f_4: "M-UK Natural Gas NBP Future",
- f_5: "N-New York Harbor Unleaded Gasoline Future",
- f_6: "NCF-Newcastle Coal Future",
- f_7: "O-New York Harbor Heating Oil Future",
- f_8: "T-West Texas Intermediate Light Sweet Crude F.",
- f_9: "UBL-UK Power Baseload Future (Gregorian)",
- f_{10}: "G-Gasoil Future (Low Sulphur Gasoil Futures from February 2015 contract month)",
- f_{11}: "DPB-Dutch Power Base Load Futures",
- f_{12}: "TFM-Dutch TTF Natural Gas Base Load Futures",
- f_{13}: "B-Brent Crude Future",
- f_{14}: "CRF-CFR South China Coal Futures",
- f_{15}: "BPB-Belgian Power Base Load Futures",
- f_{16}: "GER-German GASPOOL Futures",
- f_{17}: "GNM-German NCG Futures".

The above factors are related to the contracts for fuel and energy products such as natural gas (f_4, f_{12}, f_{16}, f_{17}), coal (f_2, f_3, f_6, f_{14}), power (f_9, f_{11}, f_{15}), crude (f_8, f_{13}), heating oil (f_7), unleaded gasoline (f_5) and gasoil (f_{10}).

Besides the above-mentioned basic values of all commodities (f_1 ... f_{17}), the collection has been supplemented with the values of the last change of all these indices. These derivative values denote as follows: f_{18} for the last change in f_1, f_{19} for f_2, ..., and f_{34} reflects the last change in f_{17}.

V. Experiments

In order to provide reliable and low error forecasts, we conducted several experiments, starting with ensemble machine learning models and ending with linear models, operating on narrowed sets of covariates.

When modeling the prices of such commodities, in addition to the required highest possible performance, the narrowed data set in the sense of a smaller number of necessary collection features would be an advantage, and thus these experiments were focused on reducing the model inputs.

A. Experiments with ensemble models

Since the ensemble machine learning models are often reported to handle complex datasets successively [23], we tested their performance against investigated price collection. To estimate EUA prices using 34 market-driven features, we first utilized the random forest and extra trees algorithms.

We have trained models for various forecast horizons, for the next day's price, over the next several days' data, and up to ten days in advance each time. We expected lower performance for a longer forecast perspective but wanted to check the exact increase in the estimation error to get at least an approximation of how the model would perform for such several days-long predictions.

TABLE I
THE STATISTICS FOR RANDOM FOREST (RF) AND EXTRA TREES (ET) ALGORITHMS FOR 10 FORECAST PERSPECTIVES. WE INDICATE IF THE METRIC SHOULD BE MINIMIZED(↓) OR MAXIMIZED(↑).

Model	**MAPE (↓) for various forecast perspectives**									
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
ET	10.64	10.94	11.54	11.49	12.83	13.24	13.93	14.65	14.36	13.78
RF	3.920	5.595	6.889	9.852	10.83	11.77	12.33	13.09	13.31	13.91
	R^2 (↑) for various forecast perspectives									
	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
ET	-0.049	-0.033	-0.160	-0.090	-0.393	-0.533	-0.456	-0.800	-0.567	-0.453
RF	0.928	0.854	0.782	0.567	0.475	0.372	0.324	0.250	0.232	0.164

It is worth noting, that during the cross-validation training runs, we allowed both models to select the best-performing configuration for the number of estimators used. The models of lowest errors in this experiment were most often configured for the maximum number of 25 estimators (although they may have used as many as 500).

The results for this investigation are provided in Table I. We have denoted the resulting metrics (MAPE and R^2) for the forecasts for the following 10 days (d_1, d_2, ..., d_{10}).

We tested two various ensemble algorithms because they approach feature selection differently, hoping that one of them would manage to find a profitable subset of columns and return good forecasts. It is clear from Table I that random forest performed better in this round of experiments. However, the measured errors for these forecasts were not impressive. The mean absolute percentage error for the next day's forecast was 3.9%, and it was lower than 10% only up to the fourth day in advance. The coefficient of determination was less than 0.5 for the random forest models from the fifth day on. And for the extra trees based algorithms, all models have negative R^2 indicating very poor performance. The better results of the random forest were probably related to the applied sample bootstrapping applied, but still, both of those architectures left room for improvement for other models.

B. Expert assisted feature selection experiment

Since the experiment with the automated approach to feature selection resulted in a rather disappointing forecasting performance we decided to test another one. For this experiment instead of testing models against all feature combinations, we trained models on the feature subsets that were recommended by another research. We tested 5 different setups.

The first one was based on prices only, and the second one utilized mainly price derivatives. And other subsequent subsets (S_3, S_4, and S_5) that used specifically indicated features. The sets S_1 and S_2 could be rephrased as simply working on all prices or working on all price derivatives. The latter three sets reflect the top-ranked features reported in the study [16], but to provide a deeper understanding of which features are really beneficial for final estimation, we tested the top five, top ten, and best fifteen features from that ranking to build S_3, S_4,

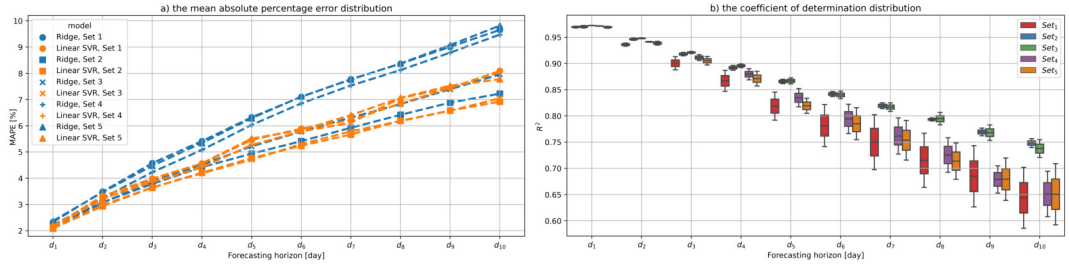

Fig. 2. A comparison of the results: (a) MAPE for two model architectures and (b) R^2 for various feature subsets.

and S_5 respectively. The detailed configuration of these subsets was as follows:

- Set 1, utilized all current prices, so the set denotes: $S_1 \in \{f_1, f_2, ..., f_{17}\}$,
- Set 2, consisted of the current EUA price value (f_1) and all features reflecting changes in the investigated prices, thus: $S_2 \in \{f_1, f_{18}, f_{19}, ..., f_{34}\}$,
- Set 3 was built on the top 5 recommended features, thus: $S_3 \in \{f_1, f_{11}, f_{15}, f_{14}, f_9, f_6\}$,
- Set 4 was built on the top 10 recommended features, thus: $S_4 \in \{f_1, f_{11}, f_{15}, f_{14}, f_9, f_6, f_3, f_2, f_4, f_5, f_{16}\}$,
- and lastly Set 5, which gathers the top 10 recommended features, with another 5, totaling the following 15 features: $S_5 \in \{f_1, f_{11}, f_{15}, f_{14}, f_9, f_6, f_3, f_2, f_4, f_5, f_{16}, f_{17}, f_{12}, f_{10}, f_8, f_7\}$,

In Fig. 2 we have plotted the results of the models trained on these 5 feature subsets. We used two various model architectures. The linear model was supported with L_2 regularization (known also as ridge regression), and the Support Vector Machine for regression with a linear kernel-based model (SVR) was also has been regularized. The comparison of the results is shown in the upper part of Fig. 2(a), and both approaches are also compared in detail in Table II. As can be seen, the models using the SVR architecture produced slightly lower errors for all subsets tested and for almost all forecast perspectives.

But the main reason for this experiment was related to the feature subsets. As we found out, in most cases the second and third subsets returned the lowest errors. This is clearly visible in the right part of Fig. 2, where the blue and green boxes represent the better coefficient of determination for all of the tested forecasting perspectives. The difference in performance for models based on S_2 and S_3 is so small that it requires another look at the results Table II. When using the R^2 as the decision criterion, the better choice would be the third subset, which has the best R^2 for 9 out of 10 forecast perspectives.

TABLE II

THE RESULTS FOR VARIOUS FEATURE SUBSETS AND FOR THE FOLLOWING DAYS' FORECASTS. THE BEST RESULTS ARE **BOLDFACED** AND WE INDICATE IF THE METRIC SHOULD BE MINIMIZED (↓) OR MAXIMIZED (↑).

Model	**Set**	**MAPE (↓) for various forecast horizons**									
		d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
Linear	S_1	2.329	3.489	4.577	5.415	6.323	7.101	7.774	8.365	9.010	9.656
	S_2	2.227	3.080	3.785	4.424	4.935	5.414	5.924	6.417	6.883	7.224
	S_3	2.142	3.067	3.874	4.522	5.213	5.777	6.313	6.828	7.394	7.960
	S_4	2.203	3.246	4.229	5.080	6.032	6.853	7.536	8.123	8.790	9.475
	S_5	2.373	3.466	4.476	5.343	6.288	7.111	7.763	8.375	9.079	9.817
SVR	S_1	2.166	3.320	3.988	4.560	5.459	5.890	6.216	6.849	7.442	8.092
	S_2	**2.085**	**2.937**	**3.642**	4.211	4.768	**5.236**	**5.668**	6.189	6.580	**6.919**
	S_3	2.086	2.974	3.650	**4.187**	**4.709**	5.300	5.787	**6.182**	**6.573**	7.047
	S_4	2.132	3.226	3.853	4.419	5.269	5.783	6.116	7.054	7.457	8.047
	S_5	2.192	3.197	3.962	4.567	5.513	5.834	6.407	7.068	7.529	7.774
		R^2 (↑) for various forecast perspectives									
		d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
Linear	S_1	0.968	0.933	0.888	0.847	0.792	0.741	0.697	0.664	0.627	0.586
	S_2	0.968	0.944	0.915	0.887	0.862	0.837	0.814	0.790	0.762	0.740
	S_3	0.972	0.947	0.919	0.892	0.861	0.833	0.809	0.783	0.753	0.721
	S_4	0.971	0.942	0.906	0.869	0.817	0.766	0.727	0.692	0.653	0.608
	S_5	0.967	0.935	0.897	0.857	0.805	0.755	0.716	0.679	0.639	0.592
SVR	S_1	0.971	0.939	0.913	0.887	0.845	0.822	0.802	0.767	0.743	0.702
	S_2	**0.973**	**0.949**	0.921	0.896	0.870	**0.847**	**0.825**	0.796	0.777	**0.757**
	S_3	**0.973**	**0.949**	**0.923**	**0.899**	**0.872**	**0.847**	**0.825**	**0.807**	**0.783**	0.755
	S_4	0.972	0.941	0.917	0.890	0.852	0.822	0.796	0.758	0.705	0.694
	S_5	0.971	0.943	0.913	0.885	0.834	0.816	0.791	0.748	0.720	0.709

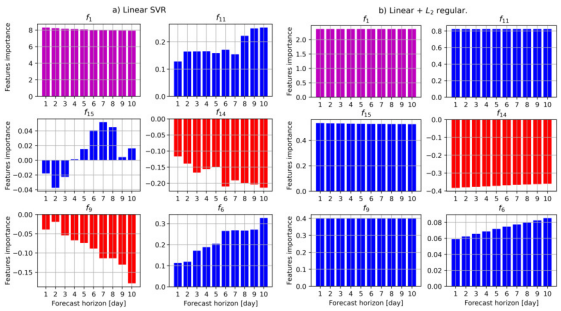

Fig. 3. The feature importance comparison for different forecast perspectives: (a) for SVR, and (b) for the linear model with L_2 regularization.

C. Features close-up

The interesting fact regarding the best-performing variant of the tested models is that it is based on the smallest subset of features. The subset S_3 consists of only 6 features. We took a closer look at the contribution of these features to the final prediction. We compared two tested variants of the linear models we trained and depicted the feature importance for both approaches (see Fig. 3). For the slightly better SVR models, the features appear to be dynamic over time dimension. Besides the most important first feature (the last denoted EUA price), other SVR features (Fig. 3a) radically change their values for the next day's forecasts. The linear model utilizes the features in a more stable manner, and only

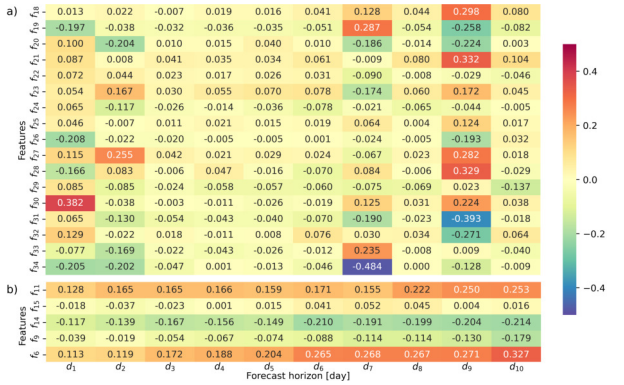

Fig. 4. The feature importance for linear SVR models and: (a) subset S_2, and (b) subset S_3. The first feature was omitted to improve chart readability.

for the last one (f_6) does the feature importance change by more than a few percent.

A closer look at the feature importance of the other feature set reveals an even more dynamic character for the set S_2 (see Fig. 4). The first feature for both sets was omitted because its importance is very dominant and would reduce the readability of the graph. The comparison of the two best-performing models using set S_2 and set S_3 which is presented there shows that the SVR-based model differs from variant to variant (when comparing variants for d_1, d_2, and the following).

The above analysis indicate that the most important are factors from two groups of individual products: power and coal. In particular, when using the SVR-based model, the current prices of the following futures have the greatest influence on the short-term forecast: “NCF-Newcastle Coal Future”, "DPB-Dutch Power Base Load Futures", "CRF-CFR South China Coal Future". It can also be seen that as the forecast horizon increases (from d_1 to d_{10}), the importance of these factors also increases. This proves that EUA price is the most sensitive to changes caused by energy markets. EUA price will largely reflect the demand for a given type of fuel in this sector. If all factors are used in the SVR-based model in the form of their last change, it is difficult to clearly rank the importance of individual factors. Their importance depends on the horizon of the forecast. In the case of the day-ahead forecast d_1, the importance of the "B-Brent Crude Future" price change is emphasized, for d_7 - "GNM-German NCG Futures", and for d_9 - "CRF-CFR South China Coal Futures". However, it is noticeable that in the case of a forecast 9 days in advance, the significance of most of the analyzed factors increases.

VI. Conclusion

As reported for the presented experiments, expertise-based feature selection could lead to better model results in some situations. For the analyzed case of EUA prices, it resulted in a lower prediction error than a more automated approach that was based on ensemble machine learning models. Such a piece of expert advice on which features to focus on could save a lot of time that would otherwise be spent experimenting with a potentially large number of different feature sets.

The apparent limitation of such an approach might be a lack of information about which covariates to focus on during modeling. This should not usually be a real concern when training a price prediction model for a relatively popular asset.

Further work should hit the time series based forecasting techniques, which would possibly lead to the ultimate performance improvement. These could be, reported to be effective for price modeling NBEATSx [12] or Temporal Convolutional Networks [17]. However, as these methods can be more time consuming, the precisely selected feature set as commented by this paper should be considered a strong asset.

References

[1] Byun, S.J., Cho, H.: Forecasting carbon futures volatility using garch models with energy volatilities. Energy Economics **40**, 207–221 (2013). https://doi.org/10.1016/j.eneco.2013.06.017

[2] Caswell, T.A., et al.: matplotlib/matplotlib: Rel: v3.7.0 (2 2023). https://doi.org/10.5281/zenodo.7637593

[3] Chevallier, J.: Nonparametric modeling of carbon prices. Ener. Econom. **33**(6), 1267–1282 (2011). https://doi.org/10.1016/j.eneco.2011.03.003

[4] Duong, V.T., et al.: Comparative study of deep learning models for predicting stock prices. In: Annals of Computer Science and Information Systems. vol. 33, pp. 103–108 (2022). https://doi.org/10.15439/2022R02

[5] E, J., Ye, J., He, L., Jin, H.: A denoising carbon price forecasting method based on the integration of kernel independent component analysis and least squares support vector regression. Neurocomputing **434**, 67–79 (2021). https://doi.org/https://doi.org/10.1016/j.neucom.2020.12.086

[6] Fixed Income Trading Analytics: https://www.theice.com/products (8 2022), www.theice.com, visited on 2022-08-02

[7] Ghosh, S., et al.: A study on support vector machine based linear and non-linear pattern classification. In: 2019 ICISS. pp. 24–28 (2019). https://doi.org/10.1109/ISS1.2019.8908018

[8] Ho, W.K., Tang, B.S., Wong, S.W.: Predicting property prices with machine learning algorithms. Journal of Property Research **38**(1), 48–70 (2021). https://doi.org/10.1080/09599916.2020.1832558

[9] Huang, Y., et al.: Carbon price forecasting with optimization prediction method based on unstructured combination. Science of The Total Env. **725**, 138350 (2020). https://doi.org/10.1016/j.scitotenv.2020.138350

[10] Huang, Y., et al.: A hybrid model for carbon price forecastingusing garch and long short-term memory network. Applied Energy **285**, 116485 (2021). https://doi.org/10.1016/j.apenergy.2021.116485

[11] Lovcha, Y., et al.: The determinants of co2 prices in the eu emission trading system. Applied Energy **305**, 117903 (2022). https://doi.org/10.1016/j.apenergy.2021.117903

[12] Olivares, K., et al.: Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with nbeatsx. Int. J. of Forec. **39** (2023). https://doi.org/10.1016/j.ijforecast.2022.03.001

[13] Pedregosa, F., et al.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)

[14] Raschka, S., et al.: Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2 (2019)

[15] Rudnik, K., Hnydiuk-Stefan, A., Li, Z., Ma, Z.: Short-term modeling of carbon price based on fuel and energy determinants in eu ets. Journal of Cleaner Production **417**, 137970 (2023). https://doi.org/10.1016/j.jclepro.2023.137970

[16] Rudnik, K., et al.: Forecasting day-ahead carbon price by modelling its determinants using the pca-based approach. Energies **15**(21) (2022). https://doi.org/10.3390/en15218057

[17] Ruszczak, B.: Evolutionary configuration of the temporal convolutional network for the electricity price prediction. GECCO ’23, ACM (2023). https://doi.org/10.1145/3583133.3596429

[18] Ruszczak, B., et al.: The detection of Alternaria solani infection on tomatoes using ensemble learning. Journal of Ambient Intell. and Smart Env. **12**(5), 407–418 (2020). https://doi.org/10.3233/AIS-200573

[19] Sun, W., et al.: A carbon price prediction model based on secondary decomposition algorithm and optimized back propagation neural network. Journal of Cleaner Production **243**, 118671 (2020). https://doi.org/10.1016/j.jclepro.2019.118671

[20] Sun, W., et al.: An ensemble-driven long short-term memory model based on mode decomposition for carbon price forecasting of all eight carbon trading pilots in china. Energy Science & Engineering **8**(11), 4094–4115 (2020). https://doi.org/10.1002/ese3.799

[21] Sun, W., et al.: Factor analysis and carbon price prediction based on empirical mode decomposition and least squares SVM optimized by improved particle swarm optimization. Carbon Management **11**(3), 315–329 (2020). https://doi.org/10.1080/17583004.2020.1755597

[22] Tian, C., et al.: Point and interval forecasting for carbon price based on an improved analysis-forecast system. Applied Mathematical Modelling **79**, 126–144 (2020). https://doi.org/10.1016/j.apm.2019.10.022

[23] Trajanoska, M., Gjorgovski, P., Zdravevski, E.: Application of diversified ensemble learning in real-life business problems: The case of predicting costs of forwarding contracts. In: Ganzha, M., Maciaszek, L.A., Paprzycki, M., Slezak, D. (eds.) Proceedings of the 17th Conference on Computer Science and Intelligence Systems, FedCSIS 2022, Sofia, Bulgaria, September 4-7, 2022. Annals of Computer Science and Information Systems, vol. 30, pp. 437–446 (2022). https://doi.org/10.15439/2022F297, https://doi.org/10.15439/2022F297

[24] Trajanoska, M., et al.: Application of diversified ensemble learning in real-life business problems: The case of predicting costs of forwarding contracts. In: FedCSIS 2022. vol. 30, pp. 437–446 (2022). https://doi.org/10.15439/2022F297

[25] Waskom, M.L.: seaborn: statistical data visualization. Journal of Open Source Soft. **6**(60), 3021 (2021). https://doi.org/10.21105/joss.03021

[26] Xie, M., et al.: Efficient cross validation for svr based on center distance in kernel space. In: IRCE 2019. pp. 128–132 (2019). https://doi.org/10.1109/IRCE.2019.00033

[27] Xu, H., et al.: Carbon price forecasting with complex network and extreme learning machine. Physica A: Statistical Mechanics and its Applications **545**, 122830 (2020). https://doi.org/10.1016/j.physa.2019.122830

[28] Zhu, B., et al.: A novel multiscale nonlinear ensemble leaning paradigm for carbon price forecasting. Energy Economics **70**, 143–157 (2018). https://doi.org/10.1016/j.eneco.2017.12.030