Zero Shot Time Series Forecasting Using Kolmogorov Arnold Networks

Abhiroop Bhattacharya
Department of Electrical Engineering
École de Technologie Supérieure, Montreal, Canada.
abhiroop.bhattacharya.1@ens.etsmtl.ca

Nandinee Haq
Hitachi Energy Research
Montreal, Canada.
nandinee.haq@hitachienergy.com

Abstract

Accurate energy price forecasting is crucial for participants in day-ahead energy markets, as it significantly influences their decision-making processes. While machine learning-based approaches have shown promise in enhancing these forecasts, they often remain confined to the specific markets on which they are trained, thereby limiting their adaptability to new or unseen markets. In this paper, we introduce a cross-domain adaptation model designed to forecast energy prices by learning market-invariant representations across different markets during the training phase. We propose a doubly residual N-BEATS network with Kolmogorov Arnold networks at its core for time series forecasting. These networks, grounded in the Kolmogorov-Arnold representation theorem, offer a powerful way to approximate multivariate continuous functions. The cross domain adaptation model was generated with an adversarial framework. The model's effectiveness was tested in predicting day-ahead electricity prices in a zero shot fashion. In comparison with baseline models, our proposed framework shows promising results. By leveraging the Kolmogorov-Arnold networks, our model can potentially enhance its ability to capture complex patterns in energy price data, thus improving forecast accuracy across diverse market conditions. This addition not only enriches the model's representational capacity but also contributes to a more robust and flexible forecasting tool adaptable to various energy markets.

1 Introduction

The increasing competitiveness of electricity markets has driven significant advancements in electricity price forecasting. Accuracy of the forecasts drive the bids for buying and selling electricity in the day-ahead market and hence reliable price forecasts are essential for market participants such as suppliers and traders. In markets where data is scarce or training could be costly, domain adaptation based machine learning techniques could offer solutions to generate forecasts for electricity prices in a zero-shot fashion without the requirement of training the model on the target market. While domain adaptation has seen successful applications in the field of computer vision, applying these methods to time series forecasting requires considerations to the temporal dynamics and local patterns inherent to time series [1].

In the past years, several different time series specific models have emerged that focuses on learning the temporal patterns for forecasting. The Neural Basis Expansion Analysis model, N-BEATS [2] is one such model that has shown superior performance for domain-specific forecasting tasks. N-BEATS model architecture comprises of two main Multi-Layer Perceptron (MLP) based components: the backcast stack which processes historical data, and the forecast stack which predicts future values. In this paper, we propose an adversarial domain adaptation based framework for zero shot forecasting of electricity prices with an architecture based on N-BEATS model. Inspired by the Kolmogorov-Arnold representation theorem [3], we propose integrating Kolmogorov-Arnold

NeurIPS 2024 Workshop on Time Series in the Age of Large Models.

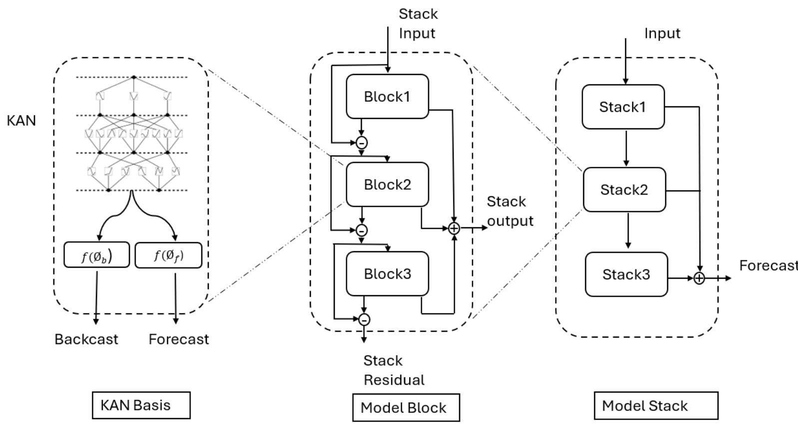

Figure 1: Line Schematic showing the model architecture consisting of KAN layers stacked together using residual connections inspired by the N-BEATS architecture [2]

Networks (KANs) within the doubly residual architecture of the N-BEATS model for generalized feature extraction. KANs have emerged as a promising alternative to MLPs, which, unlike MLPs, utilize learnable activation functions on the edges by replacing the linear weights with univariate functions parametrized as splines. This architecture enables KANs to outperform MLPs in terms of accuracy and interpretability, achieving better results with fewer parameters and providing more intuitive visualizations. Our proposed model, built by integrating KANs with an N-BEATS-like architecture and trained with adversarial technique using a gradient reversal layer [4], ensures that the initial stack captures generalizable features useful for extracting domain-invariant representations, while deeper stacks focus on domain-specific features.

The key contributions of this paper are two-fold. This is the first work that leverages a combination of Kolmogorov Arnold networks and doubly residual connection networks like N-BEATS for time series forecasting. Building on this, we further propose an adversarial domain adaptation based framework for zero shot forecasting of energy prices by creating a generalized representation.

2 Model Architecture

In this work, we use a deep stack of Kolmogorov Arnold networks (KANs) with doubly residual connections. The network decomposes the time series into local projections by using univariate function parameters along the edge of the network. The Kolmogorov-Arnold representation theorem states that any multivariate continuous function can be decomposed into a finite sum of compositions of univariate functions. This allows KAN to model complex interactions in high-dimensional data through compositions of simpler univariate functions. KAN applies adaptive, learnable activation functions on the edges between nodes. These functions are parameterized as B-spline curves, which adjust dynamically during training to better capture the underlying data patterns.

Since the univariate functions are piecewise polynomials with specific degrees and global smoothness, they exhibit excellent approximation behavior relative to their degrees of freedom. The doubly residual principle, inspired by the N-BEATS architecture [2], is used between stacks. The time series is sequentially decomposed by subtracting the predicted backcast $\hat{y}^b_{i,j}$ from the original series to obtain the next series $y^b_{i,j+1}$. The output of each forecast $\widehat{y}$ is obtained through hierarchical aggregation of each block's forecast and the last backcast derived by a residual sequence of blocks which serves as an input to the next stack. Fig. 1 shows the proposed model architecture where the model is composed of three sequential stacks to generate the overall forecasts. Each stack has three sequential blocks of neural networks, and each block consists of KAN layers that generate the backcast and forecast estimates, which are then fed onto the next block.

This model architecture is used as the backbone for creating a generalized representation using a domain adaptation approach. We take day ahead prices from two established markets with significant historical data to generate the domain generalized model, and learn the generalized representation

Table 1: Comparison of zero shot performance for the Nord Pool Market.

	MAE			SMAPE		
Primary	KAN	N-BEATS	Proposed	KAN	N-BEATS	Proposed
FR	3.0649	2.7416	2.5056 ± 0.10	0.1069	0.0932	0.0854 ± 0.003
PJM	3.1942	3.2545	2.5697 ± 0.12	0.1110	0.1044	0.0862 ± 0.002
BE	3.1409	2.5904	2.5144 ± 0.09	0.1094	0.0869	0.0857 ± 0.004

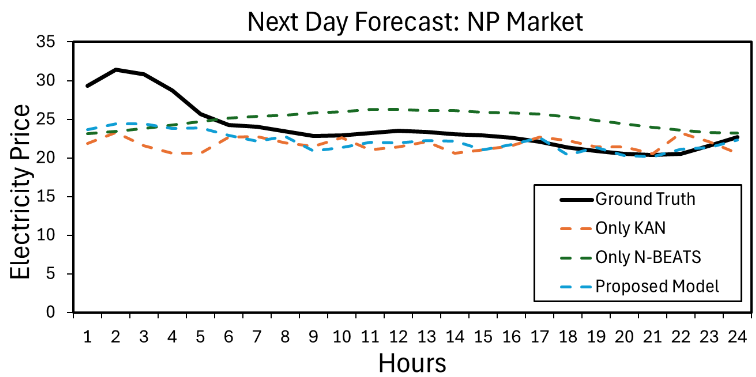

Figure 2: The next day forecast presents a comparison between the KAN, N-BEATS and Proposed model for the NP Market. As indicated, the N-BEATS model produces a smooth forecast while the proposed model uses the flexibility of the B-Spline along with the power of N-BEATS model to produce the best forecast.

by using the supervised forecasting error on the primary market and a notion of feature distance between the primary and secondary market prices. This setup enables the model to learn the domain invariant features, which if given to a classifier, the classifier should not be able to predict which domain or market the features are originating from. In addition to the domain invariant features, we use the supervised training approach to learn domain or market specific features. To implement the adversarial training between the forecasting model and the domain classifier, we use a gradient reversal layer proposed by Ganin *et.al.* [4].

3 Dataset

We train and evaluate our model's forecasting capabilities using day ahead electricity prices from major power markets. Day ahead hourly electricity prices from the Nord Pool electricity market (NP), which corresponds to the Nordic countries exchange was taken as the target or unseen market for our experiments. The test period was from 1st January, 2018 to 24th December, 2018. One full year of test data was used to capture errors across all seasons. Hourly electricity price data from three different markets were considered when training the domain-generalized models. The first train dataset is from the Pennsylvania-New Jersey-Maryland (PJM) market in the United States, which contains data from 1st January, 2013 to 24th December, 2018. The remaining two market prices are obtained from the integrated European Power Exchange (EPEX). The Belgium (BE) and French (FR) market data spans from 9th January, 2011 to 31st December, 2016.

4 Results

To present a comprehensive set of results, we conduct a series of experiments considering the Nord Pool market as the test market. The dataset is split into training and test subsets as per the method defined earlier. The hyperparameters for the proposed model and the optimization settings are optimized using a Bayesian optimization method. This method uses a tree-structured Parzen estimator to explore the hyperparameter space [5]. All the results are reported on the NP market as the unknown or new market in a zero-shot manner. For comparison, we also repeated the experiments with standard

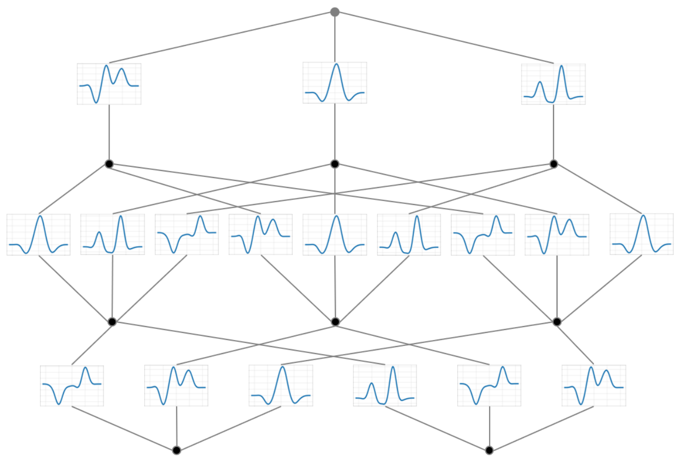

Figure 3: This representative example shows some of the functions learned by the KAN network when we do zero shot forecasting on the Nord Pool market, using France and Belgium as the primary and secondary markets repsectively.

N-BEATS architecture [2] and the standard KAN architecture [6]. Table 1 presents a comparison between the KAN, N-BEATS and the proposed model. For each set of experiments, we consider each of the markets from the train set (FR, PJM and BE) as the primary market, and use the remaining markets from the train set as secondary. The values are averaged over different models, each time with a different market as the secondary market. For all the cases, zero-shot forecasts were generated for the NP market prices. We observe an improvement of around 13% and 24% in accuracy for the proposed model compared to the N-BEATS and KAN model respectively. The performance of our proposed model, in terms of forecasting errors, is within the same order of magnitude as reported in literature [7]. Furthermore, Fig. 2 shows 24-hour ahead multi step forecast generated by our proposed model, where it can be observed that the N-BEATS model tends to produce a smoother forecast while the proposed model can leverage the flexibility of the spline curve to align with the shape of the distribution.

In a lot of real world applications, it is important that the model used for forecasting time series data has a low inference time and is easy to scale on low resource environments. Moreover, it is important that the model is easy to understand. Since KAN uses smooth functions to approximate the time series data, KAN based architectures make it easier for the users to interpret the forecast [8] as opposed to the large foundation models whose inner workings are difficult to understand. Fig. 3 shows a sample representation of the learned functions for our proposed multi layer KAN architecture with Nord Pool as the test market.

5 Conclusion

In this paper, we present a domain adaptation framework which uses the Kolmogorov Arnold Network with a doubly residual structure as the backbone for forecasting electricity prices for new markets. It builds on the N-BEATS like architecture with KAN at its core. Mainly composed of KAN layers the architecture is relatively light weight and fast to optimize, and has better interpretability than the MLP based models. We propose an adversarial training with two market data to generate a domain-generalized model that can then be applied to forecast prices from unseen markets in a zero-shot manner. We show the performance of the proposed method using a set of benchmark datasets from the electricity price forecasting domain and demonstrate that the proposed model outperforms the baseline model for zero shot forecasting. Although, the current model can be directly applied for several domain adaptation tasks across markets we believe that there is scope to further improve the model by incorporating external factors like weather parameters to augment the univariate features. It would be also interesting to extend the framework to allow multiple secondary markets to create a more generalized feature representation.

References

[1] Yuan Li, Jingwei Li, Chengbao Liu, and Jie Tan. Time series forecasting model based on domain adaptation and shared attention. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pages 215–225. Springer, 2023.

[2] Boris N Oreshkin, Dmitri Carpov, Nicolas Chapados, and Yoshua Bengio. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint arXiv:1905.10437*, 2019.

[3] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y Hou, and Max Tegmark. KAN: Kolmogorov-Arnold Networks. *arXiv preprint arXiv:2404.19756*, 2024.

[4] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The journal of machine learning research*, 17(1):2096–2030, 2016.

[5] James Bergstra, Dan Yamins, David D Cox, et al. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, volume 13, page 20. Citeseer, 2013.

[6] Cristian J Vaca-Rubio, Luis Blanco, Roberto Pereira, and Màrius Caus. Kolmogorov-Arnold networks (KANs) for time series analysis. *arXiv preprint arXiv:2405.08790*, 2024.

[7] Kin G Olivares, Cristian Challu, Grzegorz Marcjasz, Rafał Weron, and Artur Dubrawski. Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with NBEATSx. *International Journal of Forecasting*, 39(2):884–900, 2023.

[8] Kunpeng Xu, Lifei Chen, and Shengrui Wang. Kolmogorov-Arnold Networks for Time Series: Bridging Predictive Power and Interpretability. *arXiv preprint arXiv:2406.02496*, 2024.