Electric load demand forecasting on Greek Energy Market using lightweight neural networks

Nikitas Maragkos[1], Maria Tzelepi[1], Nikolaos Passalis[1], Apostolos Adamakos[2], and Anastasios Tefas[1]
[1]*Department of Informatics*, Aristotle University of Thessaloniki, Thessaloniki, Greece
[2]*Public Power Corporation*, Athens, Greece
Email: {maragkosn, mtzelepi, passalis}@csd.auth.gr, a.adamakos@dei.com.gr, tefas@csd.auth.gr

***Abstract*—In this paper, we deal with the Electric Load Demand Forecasting (ELDF) problem, considering the real case scenario of Greek Energy Market. ELDF constitutes a critical task accompanied by many applications, e.g., power systems operations and planning. In order to address the specific objectives and requirements of the Greek Energy Market, we propose a lightweight model with a novel loss function. We evaluate the effectiveness of the proposed model in terms of mean absolute percentage error, while we also evaluate its efficiency in terms of training/inference time, complexity, required memory, etc. The proposed model accomplishes the specific objectives dictated by the Greek Public Power Corporation, while it also achieves superior performance as compared to the successful baseline model ResNetPlus.**

Keywords*—*Electric Load Demand Forecasting, Greek Energy Market, Time Series Forecasting, Load Forecasting, Deep Learning

I. Introduction

Electric Load Demand Forecasting (ELDF) refers to the task of forecasting the aggregated expected electricity demand. ELDF constitutes a vital task since the inception of electric power industry, and it is accompanied by many applications including power systems operations and planning, energy trading, etc. [1]. As energy market becomes increasingly competitive, it is of crucial importance to develop methods that allow for accurate ELDF. This has fueled the research interest over the recent years [2].

Firstly, statistical models, such as Auto-Regressive Moving Average (ARMA) [3] and Auto-Regressive Integrated Moving Average (ARIMA) [4] were used. The utilization of such models is accompanied by several shortcomings. For example, it prevents the exploitation of useful external factors such as the weather [5], while it is also observed that noise can significantly affect the robustness of the predictions [6].

Surveying the relevant literature, we also come across several machine learning algorithms for addressing the ELDF task. For example, Support Vector Regression (SVR) has been proposed for predicting the load demand [7], and even though the training process is time-consuming, with an appropriate parameter tuning, it achieves promising results considering very-short load demand forecasting problems. A random forest based ensemble system is proposed in [8] achieving considerable performance in terms of accuracy and stability.

Subsequently, motivated by the exceptional performance of Deep Learning (DL) algorithms in various problems, e.g., image classification and retrieval [9], [10], [11], several recent studies introduced DL for addressing a variety of time series forecasting tasks, ranging from retail demand forecasting [12] to financial time series forecasting [13]. Besides, DL approaches have been proposed for addressing the ELDF problem [14], [15], [16], [17], [18], [19] providing very promising results.

Furthermore, several studies have focused on the *Greek Energy Market* in the recent literature [20], [21], [22], [23]. For example, Long Short Term Memory (LSTM) models have been proposed for the ELDF problem, achieving considerable performance, since they feedback connections which increase the accuracy on sequential data [24], [25], while a fuzzy-based ensemble model that uses hybrid deep learning neural networks for load forecasting is proposed in [26] accomplishing notable performance.

In the vast majority of the aforementioned methods two assumptions are made, considering the ELDF problem. First, it is considered that any data before the target day (the day whose load demand we want to predict) are available and can be utilized. Furthermore, real weather information of target day is considered available.

In this paper, we deal with ELDF problem on Greek Energy Market,in hourly basis, considering a *real case scenario*. That is, we use real data provided by the Greek Public Power Corporation (PPC), without the unconditional availability of prior knowledge before the target day. More specifically, as it is illustrated in Fig. 1, in the real case scenario there is an information gap of 4-10 days between the target day and the past load data, while as it was previously mentioned the typical setup for energy forecasting assumes that all the previous load data are available. It should be noted, that we retain the assumption of the weather information on the target day, since its solution relies on another task known as air temperature forecasting [27].

The objective of this work is to provide effective models which can accomplish the actual target dictated by the Greek PPC, that is 99% of predictions to have mean absolute percentage error below 10%. To achieve this goal, considering the real case scenario described in the subsequent Section, we propose a lightweight neural network with a novel loss function. To the best of our knowledge, there is no other work that uses data of Greek Energy Market on a realistic setup. Overall, our approach can achieve remarkable performance, according to metrics introduced by PPC with an ensemble of lightweight

neural networks.

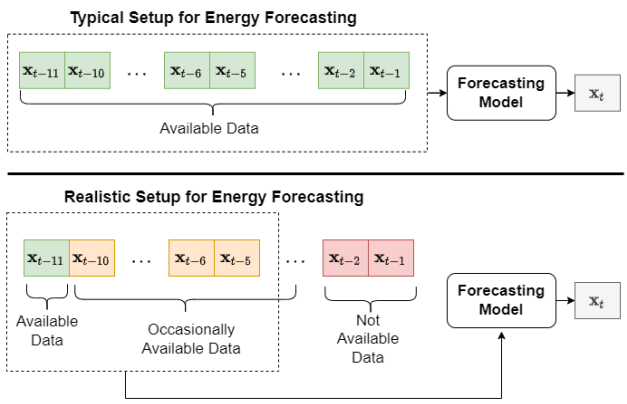

Fig. 1: Comparison between the typical setup and a realistic approach regarding the accessibility of previous load data. Available load data are printed in green, occasionally available load data in orange, and unavailable load data in red.

The remainder of the manuscript is organized as follows. Section II presents in detail the proposed method, including a description of the real scenario of Greek Energy Market, the proposed lightweight model, and the novel loss function. Subsequently, in Section III the experiments conducted to evaluate the proposed method are provided, and finally the conclusions are drawn in Section IV.

II. Proposed Method

In this paper, we propose a novel approach for electric load demand forecasting on Greek Energy Market. As it has already been mentioned, two major assumptions are made in the vast majority of relevant works [16], [28], [29], [30]. First, it is considered that every load data, which chronologically is before the target day, could be available. Second, weather information for the target day is also considered as known. In our real case scenario, we make only the second assumption, since its solution relies on another task known as air temperature forecasting [27].

Thus, we first utilize the successful model, called ResNetPlus [29], which also makes the aforementioned assumptions, and we adapt it on our scenario so as to use it as baseline model. Despite its effectiveness, ResNetPlus is a highly complex and hence computationally and memory demanding, as well as time-consuming model. However, in our case, both training and inference speed are of utmost importance, since the final implementation will be used in production where flexibility is essential. Therefore, in this paper, a more lightweight model is proposed in order to meet the specific requirements of our scenario. More specifically, we propose a lightweight Multi-Layer Perceptron (MLP) model, trained with a novel loss function. The proposed model uses a variety of load, weather and calendar features, and, by making only the second assumption, achieves remarkable performance. In the subsequent Sections, we describe in detail the real scenario of Greek Energy Market, the proposed lightweight model, as well as the proposed loss function.

A. Greek Energy Market - Real Scenario

In this paper, we deal with the real-case scenario of Greek Energy Market. In this scenario, previous week's energy data are being published each Thursday. This delayed availability creates a gap of 4-10 missing days on the data, which should be filled before moving on to the final predictions of the target day. To address this issue, we create a model per missing day, plus one more for the target day, leading to a total of 11 different models, where each one will be trained individually. The input data to each of the models are presented in Table I.

Since each of these models have the same architecture we should clarify the difference of their inputs. This can be achieved by explaining in more detail a specific case. Suppose that today is Wednesday 11th of March, we have a gap of 10 days and we want to predict the load demand for 12th of March. The available data are all temperature information until 11th of March and all load information until 1st of March. So, the first step is to fill the load data from 2nd to 11th of March. For March 2 we use Model-0 that has inputs all the original features of Table I. For March 3 we use Model-1 that has same inputs as Model-0 with only difference that L^{day} will be the output of Model-0. Respectively, with the same method March 8 will be predicted by Model-6 having as L^{day} the output of Model-5. Now, for 9-11 of March we have to replace L^{week} with the output of the appropriate model. Having done this procedure, we can now predict 12 of March which was our initial target day. Here, Model-10 replaces L^{day} with Model-9's output, L^{week} with Model-3's output, and the only feature that is not available is the temperature, T. In this work, we make the assumption that we know next day's temperatures and so we can use them to produce our final results.

B. Proposed Lightweight Model

The proposed architecture is an MLP consisting of two hidden layers with 300 and 100 neurons respectively. The input layer consists of 171 neurons and the output layer of 24, one for each hour of the target day. As activation function, the Rectified Linear Unit (RELU) activation is used. The input data that are used to predict the 24 loads of the Target Day (TD) have mainly been selected according to historical data related to this TD. To be more precise, these days are chronologically 1, 7, and 28 days before TD. In this way, we aim to capture useful information of previous day, week and month respectively. The rest of input data concern the TD. That is, we use two booleans indicating if TD is Weekend (W) or Holiday (H), an integer indicating what Day (D) of the week TD is and an array of Temperatures (T) for TD.

TABLE I: Input Data

Name	Size	Description
L^{month}	24	Load of the day that is chronologically 28 days before TD
L^{week}	24	Load of the day that is chronologically 7 days before TD
L^{day}	24	Load of the day that is chronologically 1 day before TD
T^{month}	24	Corresponding temperature for L^{month}
T^{week}	24	Corresponding temperature for L^{week}
T^{day}	24	Corresponding temperature for L^{day}
T	24	Corresponding temperature for target day
D	1	Indicator for which day of the week is the target day
W	1	Indicator if target day is weekend
H	1	Indicator if target day is holiday

C. Proposed Loss Function

In this paper, we propose a novel loss function to train the proposed models. The proposed loss function combines the Mean Absolute Percentage Error (MAPE):

$$\mathcal{L}_{MAPE}(y, \hat{y}) = |\frac{y - \hat{y}}{y}|, \quad (1)$$

where y is the ground truth label and $\hat{y}$ is the prediction from the neural network, with a new metric that aims to address the needs of Greek Energy Market. The proposed metric is based on the focal loss [31] adapted for regression tasks. With the use of the proposed loss function:

$$\mathcal{L} = a * \mathcal{L}_{MAPE} + b * max(0, \mathcal{L}_{MAPE} - c) * \mathcal{L}_{MAPE}, \quad (2)$$

if the training model does not achieve MAPE lower than c it will get higher penalty. Note that indices are omitted in (2) to avoid cluttering the notation. The first addend of this loss function is critical for overcoming the under-fitting [32]. If first addend is omitted, under-fitting will occur for high values of c, where the model produces wrong predictions but still returns zero error; hence model's weights will not update. Parameters a, b are real numbers between 0 and 1, indicating the weight of each addend. Since in Greek Energy Market the target is a MAPE lower than 10, any c below 10 is suitable in our scenario.

III. Experimental Evaluation

In this Section, we present the evaluation metrics, the implementation details of the proposed method, as well as the experimental results that validate the effectiveness of the proposed method.

A. Evaluation Metrics

The main target of this research is to create an efficient model to predict next day's load demand in Greece. The effectiveness of such a model can be easily calculated by creating three groups of MAPE errors. Errors below 10% ($C1$), errors between 10-15% ($C2$) and errors above 15% ($C3$). If a model achieves, on average, 99% of its errors to be in $C1$, then it is considered successful according to our scenario. Besides we also use MAPE to evaluate the proposed model, since it is generally the main metric error considering time-series.

Furthermore, since whole processes will be executed on the production environment of PPC, we need fast responses and as low as possible computational cost. Thus, we also evaluate the efficiency of the proposed method, providing the corresponding times for a forward pass, a complete training phase which means the time needed for the training of all 11 models that sequentially will produce the predictions for the desired TD, the complexity (training cost) in terms of Floating Point Operations Per Second (FLOPs), the model parameters, and required memory, for a single instance of both baseline and proposed model.

B. Experimental Setup

In this paper, we evaluate the performance of the proposed lightweight model, on two scenarios, that are the common scenario used in the literature (accepting both assumptions presented in II), as well as the real-case scenario of Greek Energy Market. Apart from the actual objective dictated by the Greek PPC (that is 99% of predictions to have MAPE below 10%), which is accomplished as it is experimentally shown, we use the successful and powerful ResNetPlus model to perform comparisons.

Both, proposed and ResNetPlus models are implemented on Keras/Tensorflow. As in [29], we use the mean prediction of 15 trained models with different initial weights. This has an enrichment on the output, since it is more robust to hidden data noise and it produces a more generalized model [33].

For model training, we used 5 years of data (2012-2016) and 1 year for validation (2017). For testing, we kept 1 year (2018) as well. The normalization that we used is the division with the maximum element of each column, and the total number of epochs had been set equal to 2000. In most cases, a smaller number of epochs were needed, since early stopping technique was used.

C. Experimental Results

As mentioned previously, we used 5 years for training, 1 for validation and 1 for testing. The parameters in eq. (2) were selected using grid search as follows: $a = 1$, $b = 0.4$ and $c = 2$. These parameters are revealing some interesting remarks. For instance, despite the fact that our main target is a MAPE below 10, the optimal parameter c was found to be 2; indicating that we should be more strict throughout the training, in order to achieve our desired metrics. Furthermore, the optimal parameter b is equal to 0.4 and in combination with parameter a that is set to 1, we observe that the first addend of (2), which is MAPE, had a more strong impact to the loss function.

Two scenarios have been tested. That is, the most observed scenario through relevant literature that makes both assumptions of not having any missing data and knowing real weather data for target day, and the real case scenario. First, we provide MAPE errors for all the considered cases, since it is the main metric error considering time-series in general. The experimental results for both the considered scenarios are presented in Table II. As demonstrated, the proposed model achieves smaller errors compared to the baseline model on both the cases.

TABLE II: Energy forecasting evaluation (MAPE is reported for different setups)

Model	MAPE (%)	MAPE (%) Real Scenario
ResNetPlus	1.80	2.77
Lightweight	1.74	2.52

Subsequently, we evaluate the effectiveness of the proposed method considering our production metrics, as previously described. The experimental results for each month of a year, as well as the average results, for the common scenario, are presented in Tables III and IV for the ResNetPlus and the proposed model, respectively. Better results are achieved based on our production metrics with 99.51% of our proposed model's errors to be in $C1$, as opposed to the percentage of ResNetPlus errors which is 99.47%.

TABLE III: ResNetPlus

Month	<10	[10,15)	>=15
January	100	0	0
February	100	0	0
March	98.79	1.21	0
April	100	0	0
May	100	0	0
June	100	0	0
July	99.06	0.67	0.27
August	100	0	0
September	100	0	0
October	99.06	0.4	0.54
November	100	0	0
December	96.77	1.61	1.61
Average	99.47	0.32	0.21

TABLE IV: Proposed Lightweight Model

Month	<10	[10,15)	>=15
January	99.87	0.13	0
February	100	0	0
March	99.06	0.94	0
April	99.86	0.14	0
May	100	0	0
June	100	0	0
July	99.06	0.27	0.67
August	100	0	0
September	100	0	0
October	99.06	0.27	0.67
November	100	0	0
December	97.18	2.75	1.08
Average	99.51	0.29	0.20

It is worth mentioning that the ensemble of 15 models is indeed a valuable component for achieving even better results. The 15 lightweight models have a mean MAPE of 1.92 and a standard deviation of 0.05, while the ensemble model achieves an even smaller MAPE of 1.74.

The experimental results considering the real case scenario are presented in Tables V and VI for the ResNetPlus and the proposed lightweight model, respectively. In this case, according to our production metric, ResNetPlus achieves slightly better score in $C1$ over the proposed model. Nevertheless, both models achieve a percentage above 99% in $C1$ which is the main objective of our work.

TABLE V: ResNetPlus - Real Scenario

Month	<10	[10,15)	>=15
January	100	0	0
February	99.7	0.3	0
March	99.06	0.94	0
April	98.58	0.42	0
May	100	0	0
June	100	0	0
July	99.06	0.94	0
August	100	0	0
September	100	0	0
October	99.06	0.4	0.54
November	100	0	0
December	94.49	3.9	1.61
Average	99.24	0.58	0.18

Finally, the evaluation of the efficiency of the proposed method against the baseline model is provided in Table VII. As it can be observed, the proposed model is significantly more efficient as compared to the baseline model. Considering the complexity (training cost), model parameters, required memory, training time, and inference time. Especially considering the inference time, which is of great importance, the proposed model is 74.5 times faster. Finally, it is also noteworthy that ResNetPlus requires more than 2 days for training, while the proposed model can be trained in less than a day. That is, the proposed model achieves better performance considering the MAPE metric and competitive performance considering our production metric (also better in some cases), but it is significantly more efficient and faster.

TABLE VI: Proposed Lightweight Model - Real Scenario

Month	<10	[10,15)	>=15
January	99.33	0.67	0
February	100	0	0
March	99.6	0.4	0
April	98.75	0.97	0.28
May	99.87	0.13	0
June	100	0	0
July	99.06	0.67	0.27
August	99.87	0.13	0
September	100	0	0
October	98.79	0.54	0.67
November	100	0	0
December	94.22	2.82	2.96
Average	99.12	0.53	0.35

TABLE VII: Evaluation of Efficiency

Metric	ResNetPlus	Proposed
Parameters	131,704	84,124
Memory	0.54 MB	0.31 MB
Complexity	260464 FLOPs	167824 FLOPs
Training Time	55 h	5 h
Inference Time	88 s	1.19 s

IV. Conclusions & Future Research Directions

In this paper, we proposed a lightweight neural network with a novel loss function in order to address the electric load demand forecasting problem considering the real case scenario of the Greek Energy Market. The proposed model achieves the specific target imposed by the Greek PPC, while it is also computationally efficient. The proposed model is also compared with the successful ResNetPlus, achieving competitive performance, while also being significantly more efficient in terms of training and inference time, complexity, and required memory.

Future research directions include the incorporation of the weather forecasting to the final model that predicts the electric load demand of the target day. Furthermore, more sophisticated ways [34] for realizing the ensemble of the 15 models could be studied, in order to reduce the existent noise.

Acknowledgment

This work is co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH - CREATE - INNOVATE (project code: T2EDK-03048).

References

[1] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.

[2] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, "Electricity load forecasting: a systematic review," *Journal of Electrical Systems and Information Technology*, vol. 7, no. 1, pp. 1–19, 2020.

[3] S. S. Pappas, L. Ekonomou, P. Karampelas, D. Karamousantas, S. Katsikas, G. Chatzarakis, and P. Skafidas, "Electricity demand load forecasting of the hellenic power system using an arma model," *Electric Power Systems Research*, vol. 80, no. 3, pp. 256–264, 2010.

[4] J. W. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *Journal of the Operational Research Society*, vol. 54, no. 8, pp. 799–805, 2003.

[5] H. Nie, G. Liu, X. Liu, and Y. Wang, "Hybrid of arima and svms for short-term load forecasting," *Energy Procedia*, vol. 16, pp. 1455–1460, 2012.

[6] E. Chodakowska, J. Nazarko, and Ł. Nazarko, "Arima models in electrical load forecasting and their robustness to noise," *Energies*, vol. 14, no. 23, p. 7952, 2021.

[7] A. Setiawan, I. Koprinska, and V. G. Agelidis, "Very short-term electricity load demand forecasting using support vector regression," in *Proceedings of the 2009 International Joint Conference on Neural Networks*, 2009, pp. 2888–2894.

[8] Y.-Y. Cheng, P. P. Chan, and Z.-W. Qiu, "Random forest based ensemble system for short term load forecasting," in *Proceedings of the 2012 international conference on machine learning and cybernetics*, vol. 1, 2012, pp. 52–56.

[9] M. Tzelepi and A. Tefas, "Efficient training of lightweight neural networks using online self-acquired knowledge distillation," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2021, pp. 1–6.

[10] ——, "Deep convolutional learning for content based image retrieval," *Neurocomputing*, vol. 275, pp. 2467–2478, 2018.

[11] M. Tzelepi, N. Passalis, and A. Tefas, "Efficient online subclass knowledge distillation for image classification," in *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021, pp. 1007–1014.

[12] J. Huber and H. Stuckenschmidt, "Daily retail demand forecasting using machine learning with emphasis on calendric special days," *International Journal of Forecasting*, vol. 36, no. 4, pp. 1420–1438, 2020.

[13] A. Tsantekidis, N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Forecasting stock prices from the limit order book using convolutional neural networks," in *Proceedings of the 2017 IEEE 19th conference on business informatics (CBI)*, vol. 1, 2017, pp. 7–12.

[14] Z. Wen, L. Xie, Q. Fan, and H. Feng, "Long term electric load forecasting based on ts-type recurrent fuzzy neural network model," *Electric Power Systems Research*, vol. 179, p. 106106, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779619304250

[15] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," in *Proceedings of the IEEE 26th International Symposium on Industrial Electronics (ISIE)*, 2017, pp. 1483–1488.

[16] W. He, "Load forecasting via deep neural networks," *Procedia Computer Science*, vol. 122, pp. 308–314, 2017.

[17] N. Passalis and A. Tefas, "Global adaptive input normalization for short-term electric load forecasting," in *Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020, pp. 1–8.

[18] S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," *Energies*, vol. 10, no. 1, p. 3, 2016.

[19] B. Liu, J. Nowotarski, T. Hong, and R. Weron, "Probabilistic load forecasting via quantile regression averaging on sister forecasts," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 730–737, 2015.

[20] A. N. Adamakos and M. K. Titsias, "Short-term load forecasting using a cluster of neural networks for the greek energy market," in *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, 2016, pp. 1–6.

[21] I. P. Panapakidis, T. Perifanis, and A. S. Dagoumas, "The effect of dimensionality reduction methods in short-term load forecasting performance," in *Proceedings of the 2018 15th International Conference on the European Energy Market (EEM)*, 2018, pp. 1–5.

[22] N. Andriopoulos, A. Magklaras, A. Birbas, A. Papalexopoulos, C. Valouxis, S. Daskalaki, M. Birbas, E. Housos, and G. P. Papaioannou, "Short term electric load forecasting based on data transformation and statistical machine learning," *Applied Sciences*, vol. 11, no. 1, p. 158, 2021.

[23] G. Sideratos, A. Ikonomopoulos, and N. D. Hatziargyriou, "A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks," *Electric Power Systems Research*, vol. 178, p. 106025, 2020.

[24] K. Stergiou and T. Karakasidis, "Application of deep learning and chaos theory for load forecasting in greece," *Neural Computing and Applications*, vol. 33, no. 23, pp. 16 713–16 731, 2021.

[25] A. Stratigakos, A. Bachoumis, V. Vita, and E. Zafiropoulos, "Short-term net load forecasting with singular spectrum analysis and lstm neural networks," *Energies*, vol. 14, no. 14, p. 4107, 2021.

[26] V. S. Kodogiannis and I. Petrounias, "Power load forecasting using adaptive fuzzy inference neural networks," in *Proceedings of the 2012 6th IEEE International Conference Intelligent Systems*, 2012, pp. 238–243.

[27] J. Cifuentes, G. Marulanda, A. Bello, and J. Reneses, "Air temperature forecasting using machine learning techniques: a review," *Energies*, vol. 13, no. 16, p. 4215, 2020.

[28] I. P. Panapakidis, G. C. Christoforidis, N. Asimopoulos, and A. S. Dagoumas, "Combining wavelet transform and support vector regression model for day-ahead peak load forecasting in the greek power system," in *Proceedings of the 2017 IEEE International Conference on Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, 2017, pp. 1–6.

[29] K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, and J. He, "Short-term load forecasting with deep residual networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3943–3952, 2018.

[30] M. Afrasiabi, M. Mohammadi, M. Rastegar, L. Stankovic, S. Afrasiabi, and M. Khazaei, "Deep-based conditional probability density function forecasting of residential loads," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3646–3657, 2020.

[31] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[32] S. Narayan and G. Tagliarini, "An analysis of underfitting in mlp networks," in *Proceedings of the 2005 IEEE International Joint Conference on Neural Networks, 2005.*, vol. 2. IEEE, 2005, pp. 984–988.

[33] X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, "A survey on ensemble learning," *Frontiers of Computer Science*, vol. 14, no. 2, pp. 241–258, 2020.

[34] C. Zhang and Y. Ma, *Ensemble machine learning: methods and applications*. Springer, 2012.