CLC____________ Number____________

UDC____________ Available for reference Yes ☐ No ☐

Southern University of Science and Technology

Undergraduate Thesis

Thesis Title: **Personal Federated Learning Framework for Household Load Prediction**

Student Name: **Shibo Zhu**

Student ID: **11911618**

Department: **Department of Computer Science and Engineering**

Program: **Computer Science and Technology**

Thesis Advisor: **Associate Prof. Xuan Song**

Date: 2023.06.02

Letter of Commitment for Integrity

1. I solemnly promise that the paper presented comes from my independent research work under my supervisor's supervision. All statistics and images are real and reliable.
2. Except for the annotated reference, the paper contents no other published work or achievement by person or group. All people making important contributions to the study of the paper have been indicated clearly in the paper.
3. I promise that I did not plagiarize other people's research achievement or forge related data in the process of designing topic and research content.
4. If there is violation of any intellectual property right, I will take legal responsibility myself.

Signature:

Date:

Personal Federated Learning Framework for Household Load Prediction

Shibo Zhu

(Department of Computer Science and Engineering Tutor: Xuan Song)

[Abstract]: In recent years, with the development of smart meter and the promotion of green energy concepts, as well as the impact of the energy crisis, people pay more attention to household electricity consumption. Fine grained load prediction of household appliances will help to plan household electricity plans reasonably, save energy, and develop green and energy-saving smart homes. At the same time, accurate power consumption prediction also provides reliable data and decision-making support for the energy allocation task of smart home energy management systems. More accurate power consumption prediction also provides powerful reference indicators for power grid load planning. However, due to the close correlation between electricity consumption behavior and household life, residential electricity load forecasting often faces issues such as limited data volume for individual households and data privacy. In this paper, we use federated learning technology to provide data privacy and data islanding solutions for household electricity load forecasting. At the same time, when facing the problem of different household electricity consumption behaviors and different types of electrical appliances, we introduce federated mutual learning methods to solve the problem of data heterogeneity and provide personalized electricity consumption prediction models for different households. We have also made certain improvements to the proposed seq2seq-FML framework, such as introducing a global encoder model to better extract features, and designing a dedicated linear layer for each specific consumer during downstream prediction tasks. These improvements have been proven effective in our ablation experiments. Compared with traditional federated learning methods, the seq2seq-FML framework proposed in this paper achieved approximately 5.04% performance improvement.

[Keywords]: Load Prediction; seq2seq-FML; Federated Learning; Personal

[摘要]： 近年来，随着智能电表的发展和绿色能源理念的推广，以及能源危机的影响，人们更加注重家庭用电消耗情况。细粒度的家庭用电器负载预测将有助于合理规划家庭用电计划，节约能源，发展绿色节能的智能家庭，同时，精准的功耗预测也为智能家居能源管理系统的能源分配任务提供了可靠数据和决策支持。更加精准的电力消耗预测，也为电网负载规划提供有力的参考指标。然而，由于用电行为与居民家庭生活密切相关，居民用电负载预测往往面临单个家庭数据量少，以及数据隐私的问题。在这篇文章中，我们使用联邦学习技术，为家庭用电负载预测提供数据隐私、数据孤岛方面的解决方案。同时，在面临着不同家庭用电行为不同，用电器持有种类不同的问题时，我们通过引入联邦互学方法来解决其中面临的数据异质性问题，为不同家庭提供个性化的用电预测模型。我们还对提出的 seq2seq-FML 框架进行了一定的改进，例如引入全局编码器模型以更好地提取特征，在下游预测任务期间为每个特定消费者设计专用的线性层等。这些改进已经在我们的消融实验中被证明是有效的。与传统的联邦学习方法相比，本文提出的 seq2seq-FML 框架取得了约 5.04%的性能提升。

[关键词]： 负载预测；seq2seq-FML；联邦学习；个性化

Table of Content

1. Introduction

1.1 Background

In recent years, with the development of smart meters and the promotion of green energy concepts, as well as the impact of energy crises, people pay more attention to household electricity consumption. Fine-grained household appliance load forecasting will help to plan household electricity consumption reasonably, save energy, develop smart homes that are green and energy-saving, and at the same time, accurate power consumption prediction also provides reliable data and decision support for the energy allocation task of the smart home energy management system [10]. More accurate power consumption prediction also provides a powerful reference index for power grid load planning [2]. However, due to the close relationship between electricity consumption behavior and residents' household life, household electricity load prediction often faces the problem of small amount of data for individual households and data privacy [26]. In practical applications, federated learning can provide solutions to data privacy and data island problems for household electricity load prediction, but it also faces the problem of different electricity consumption behaviors and different types of appliances held by different households. To solve the above problems, we hope to use the federated mutual learning [21] method to solve the Data Heterogeneity and Objective Heterogeneity problems faced by it.

1.2 Related Work

1.2.1 Smart home energy management system

Due to the huge impact of urban buildings [14], especially residential areas [1], on energy consumption and emissions, research on smart home energy systems [28][25] is constantly advancing. At the same time, in order to make HEMS [28] more flexible in managing and controlling smart appliances, load forecasting of energy consumption end and appliances is very important.

1.2.2 Household Electricity Forecasting

Compared to aggregated loads, load forecasting for individual consumers is prone to non-stationary and stochastic features.[23] However, fine-grained load forecasting can provide better suggestions for power scheduling. According to the predicted time granularity, electricity consumption prediction can be divided into four types: very short-term load forecasting (VSTLF), short-term load forecasting (STLF), medium-term load forecasting (MTLF) and long-term load forecasting (LTLF) [18][13][9], with corresponding usage scenarios at different time granularities. On the other hand, different artificial intelligence models are used to predict future electricity consumption based on historical energy usage data [18], such as CNN [12], LSTM [23], and hybrid deep learning approach [4][19][3]. In this article, we mainly focus on very short-term load forecasting (VSTLF) based on LSTM model.

1.2.3 Federated Learning

Household electricity data always faces privacy issues. Power companies generally do not disclose their customers' electricity usage information, and customers are unwilling to share their electricity usage data because it would largely reveal their behavioral information. Federated learning [16] is one way to solve this problem and provide a powerful global model [22][7][5][20]. However, due to the diversity of user behavior, the power consumption data for each client is not independently and identically distributed, and it is necessary to personalize the model for each client [15] [24][22].

1.3 Motivations and Contribution

The development of new energy technologies such as photovoltaics and wind power has brought more clean energy to the world. However, due to the large impact of uncertain factors such as the environment on these power generation methods, more instability has also been introduced into the power system. In order to maintain power supply balance, avoid damage to the power grid system caused by excessive power generation, and ensure the electricity consumption needs of users, it is very important to predict the production capacity of new energy electricity and the energy consumption of the user side. Good power prediction will help power-related departments better plan power allocation policies, balance power supply and demand, and maintain stable operation of the power grid. In terms of electricity consumption at the consumer end, fine-grained power consumption prediction at the level of electrical appliances will also provide better guidance for HEMS [28]. However, in terms of residential electricity consumption prediction, we often face problems such as data privacy and highly personalized household electricity consumption behavior, which makes it difficult to use a single model to describe. In order to solve the problems of data privacy and model personalization in household energy consumption prediction, we hope to use personalized federated learning to solve this problem. Fortunately, in the field of computer vision, FML [21] frameworks have been proposed to solve similar problems in image classification. The FML [21] framework combines federated learning and DML [27] methods, using knowledge distillation technology to further enhance the performance of personalized models. We hope to improve on this framework by adjusting the DML [27] method and combining the LSTM model to support personalized federated forecasting for household electricity load forecasting. The key contributions of this paper are the following:

a) A unified framework for very short-term load prediction (VSTLF) of household appliances has been established.

b) It solves problems such as low data volume and data privacy faced in household level power consumption forecasting by using federated learning technology.

c) The improved seq2seq-FML framework is used to personalize the prediction of electricity consumption behavior in different households.

2. Personal Federated Learning for Load

2.1 Problem Definition

The load forecasting of household electrical appliances can be abstracted as a typical time series forecasting problem.

For a specific household H, we observe the historical power consumption data of its electrical appliances, and use this to predict the future power consumption of electrical appliances. Specifically, the historical power consumption of electrical appliances in a household can be expressed as $A = \{(p_1^t,\ p_2^t, \dots ,\ p_m^t) \mid t \in [1, 2, \dots, T_{obs}\ ]\}$, where p_m^t represents the power of the m-th selected consumer in the household at time t, T_{obs} is the length of time observed. Similarly, we define the data to be predicted as $B = \left\{(p_1^t,\ p_2^t, \dots ,\ p_m^t) \mid t \in \left[T_{obs} + 1,\ T_{obs} +\ 2, \dots, T_{obs} + T_{pred}\right]\right\}$, T_{obs} is our predicted step size.

2.2 Framework

The overview of seq2seq FML Framework is given by (Figure 1). Overall, it consists of two parts: local model and global model. Each family has its own local model that utilizes local data for training, and the data used by users for training is not shared with other families; The global model is a multi-family shared LSTM encoder.

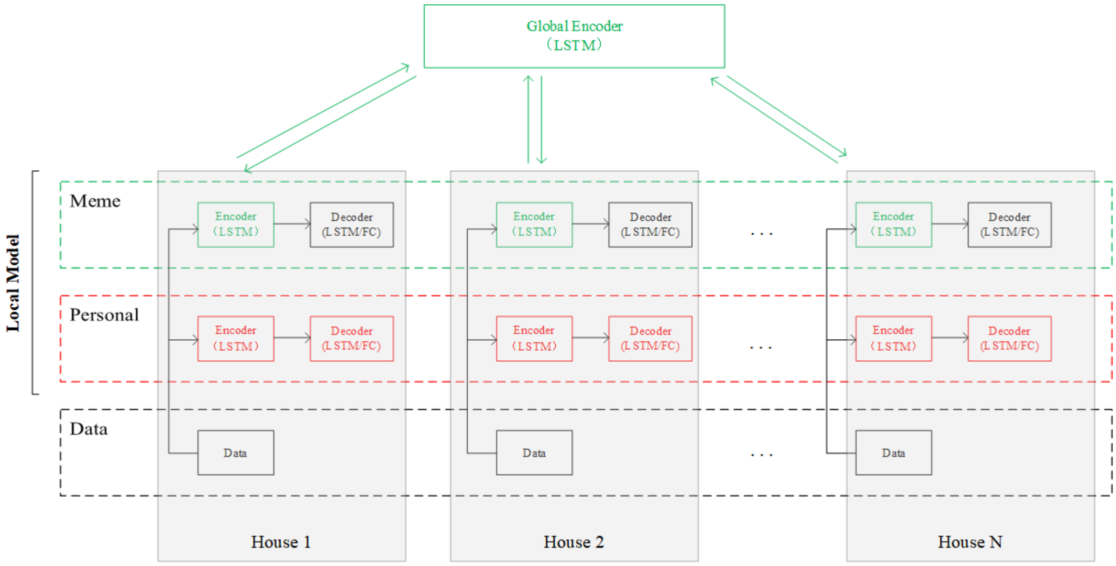


Figure 1: The seq2seq FML Framework

Local model. For a specific family, its local model consists of two parts, Meme model and personalized model, which have the same framework but different parameters. Taking the Local model as an example, it consists of two parts, an Encoder for feature extraction and a Decoder for prediction (Figure 2). The Encoder uses the LSTM model, which can well capture the long-term and short-term relationships in time series and convert the captured information into vectors output from the hidden layer. Decoder will be more complex, consisting of LSTM and two fully connected layers. The Meme model and the personalized model share the same network structure, but do not share parameters. Besides, the Encoder part of the Meme model is updated in collaboration with the global model, while the Decoder part of the Meme model only plays a role in local training. The specific training process in the local model will be given in the Loss section.

Global model. The global model plays a role in sharing information among multiple families, and consists of an LSTM network. Note that the LSTM model here is not an end-to-end prediction model, but an Encoder for feature extraction. Inspired by semi supervised learning and large-scale pre training models, we believe that only the Encoder part is shared, allowing it to learn on a larger dataset, then the extracted features will be beneficial for downstream tasks (power consumption prediction). We have also demonstrated through experiments that this design has better prediction effects than directly sharing the end-to-end prediction models. Relevant experiments are given in the experimental section.

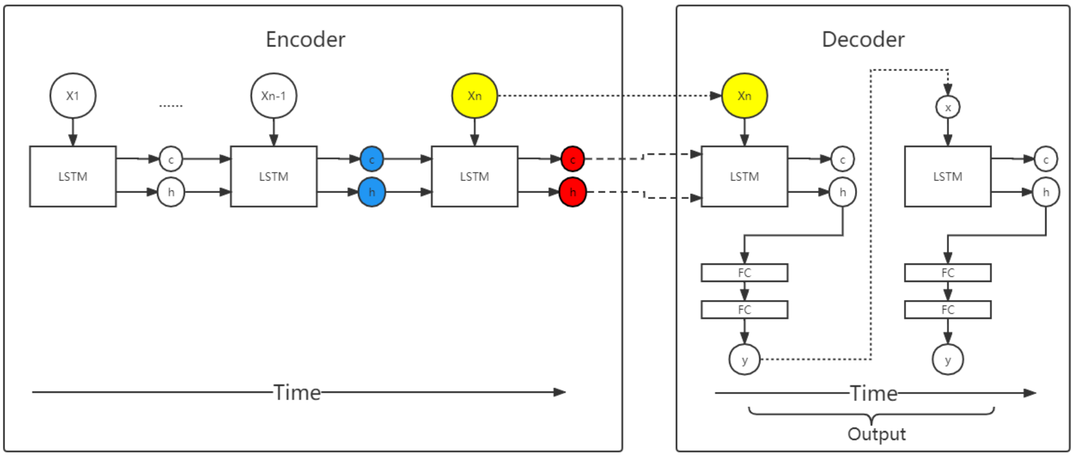


Figure 2: Encoder and Decoder

2.3 Implementation Details

2.3.1 Training Process

The training process for seq2seq FML consists of two parts, local training, and global training updates. Among them, local training uses the ideas in FML. We train two models with the same structure to each other under a small knowledge gap. Global training uses the fed-avg method to average the model parameters shared with the global model in each local model and update the global model. The specific parameter settings in the training will be given in the experiment section.

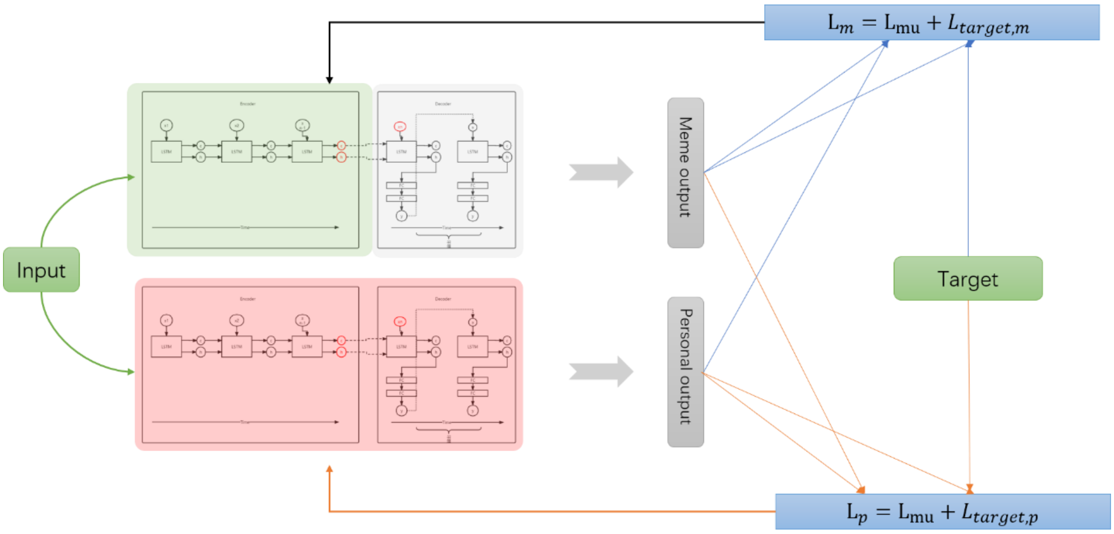


Figure 3: Loss

2.3.2 Loss Function

Knowledge distillation [8] is a common technique used for model compression, and it also plays an important role in mutual learning. In our seq2seq-FML framework, mutual learning also provided great contribution to load prediction. During the mutual learning processing, the selection of knowledge distillation methods is very important as it will help personalized models better learn the global knowledge brought by meme models in fitting, thus obtaining personalized load prediction models with stronger generalization.

In the target selection of knowledge distillation, there are generally two approaches: hard target and soft target [11]. In the FML image classification task, the author used the KL divergence between models as the Soft Target to train the model for virtual learning. However, as a time series prediction problem, KL divergence cannot well describe the gap between the two models. Based on this, we explore the effect of using different targets (including soft target and hard target) and corresponding loss function to conduct mutual learning. These loss functions are given below.

$$L_m = L_{mu} + L_{target,m} \quad (1)$$

$$L_p = L_{mu} + L_{target,p} \quad (2)$$

$$L_{mu} = MSE(O_m, O_p) \quad (3)$$

$$L_{target,m} = MSE(O_m, Y) \quad (4)$$

$$L_{target,p} = MSE(O_p, Y) \quad (5)$$

Here, L_m and L_p represents the loss function of the Meme and Personal models, respectively. O_m and O_p is the output of the two models during the training process. Y is the target sequences.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(O_i - Y_i)^2 \quad (6)$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{Y_i - O_i}{Y_i}\right| \quad (7)$$

In the selection of loss function, we choose MSE rather than the MAPE, which commonly used in the prediction of total household power consumption [2]. This is because at the level of household appliances, due to the randomness and short duration of individual appliance usage, there are many 0 values or sustained low power consumption values in training and prediction data (appliances not started or in standby state). For such data, due to the small denominator, MAPE [2] cannot accurately reflect the deviation between the predicted value and the true value of the model, making it difficult for the model to converge. However, MSE loss can effectively avoid such problems.

In order to select the best mutual learning loss that suitable for seq2seq-FML, we constructed different L_{mu} to compare the impact of hard target and different soft target choices on model performance. These results will be given in the ablation experiment section.

3. Experiments

In this section, we use real household appliance load data to conduct experiments to verify the practicality of our proposed seq2seq-FML framework.

3.1 Experimental Settings

3.1.1 Experimental Environment

In our experiments, the hardware device used is a sever with NVIDIA Titan RTX graphics cards. All programs are written based on Python 3.8 and pyTorch 11.2.

3.1.2 Dataset

The publicly available REFIT [17] dataset was used to test the effectiveness of the proposed framework. This dataset contains electrical power consumption records from 20 households in the UK over two years.

3.1.3 Data Pre-processing

Most of the data in REFIT [17] are sampled at 6-8 seconds intervals. In order to obtain data that can be used for LSTM model training, we first eliminated the obviously unreasonable data (the power of the distributor is greater than the total power), filled in the latest effective power with a time interval of 1 second as the minimum interval, and then resampled the data to obtain data with a time interval of 10 seconds and 1 minute, respectively. Before inputting the data into the model for training, we also performed a min-max normalization process on the data. The final predicted value of the model will be inversely normalized accordingly to obtain the predicted power of the consumer in watts.

In the experiment, we selected three typical households from the REFIT [17], all with typical household appliances from the UK (Fridge-Freezer, Washing Machine, Dishwasher, Microwave, and Kettle). Considering the large amount of data as a

VSTLF problem, we selected data from one month in 2014 and data from six months in 2014 to model the prediction problem for a 10s-time interval and a 1min time interval, respectively.

3.1.4 Hyper-parameter

For all experiments conducting federated learning, we set the local training round E=5, the federated training round R=10 (i.e., for a single model in a single family, the total number of training rounds is E * R=50). The learning rate is 0.003. Batch size is 256. The ratio of training set, validation set, and test set is 6:2:2. The models validated on the final test set are all the optimal models on the validation set.

For all non-federated learning methods, the total training rounds are set to E*R=50, while other parameters are consistent with federated learning methods.

3.2 Baseline

In order to more intuitively demonstrate the difficulty of very short-term load forecasting for household and the effect of our seq2seq-FML framework, we have built the following baselines.

Naive model. This is the simplest intuitive prediction model, and we use this naive model to calibrate the difficulty of predicting the power consumption of household appliances. This simple model predicts future power consumption as

$$\hat{B} = \{(p_1^{T_{obs}}, p_2^{T_{obs}}, \ldots, p_m^{T_{obs}}) |\ t\ \in [T_{obs} + 1, T_{obs} + 2, \ldots, T_{obs} + T_{pred}]\ \}$$

That is, the model predicts the power consumption of each consumer in the future from $T_{obs} + 1$ to T_{pred} as the same observation value of T_{obs} moment.

Local LSTM model. This is a basic LSTM network consisting of an LSTM cell and two fully connected layers (in order to maintain consistency with the model we used in seq2seq FML). The hidden layer dimension is 128, and the dimensions of the two linear layers are (128, 32) and (32,5), respectively.

Fed-avg. Based on the basic LSTM model, we used the fed-avg method commonly used in federated learning to combine data from different households to test the usefulness of federated learning methods in predicting household electrical energy consumption. Specifically, the training process for fed-avg is as follows: Each family (client) is now training a certain number of rounds E on the local dataset. After the training for each family (client) is completed, the obtained model parameters will be shared with the global model, averaged, and distributed to the models in each family. This iteration involves a total of R rounds.

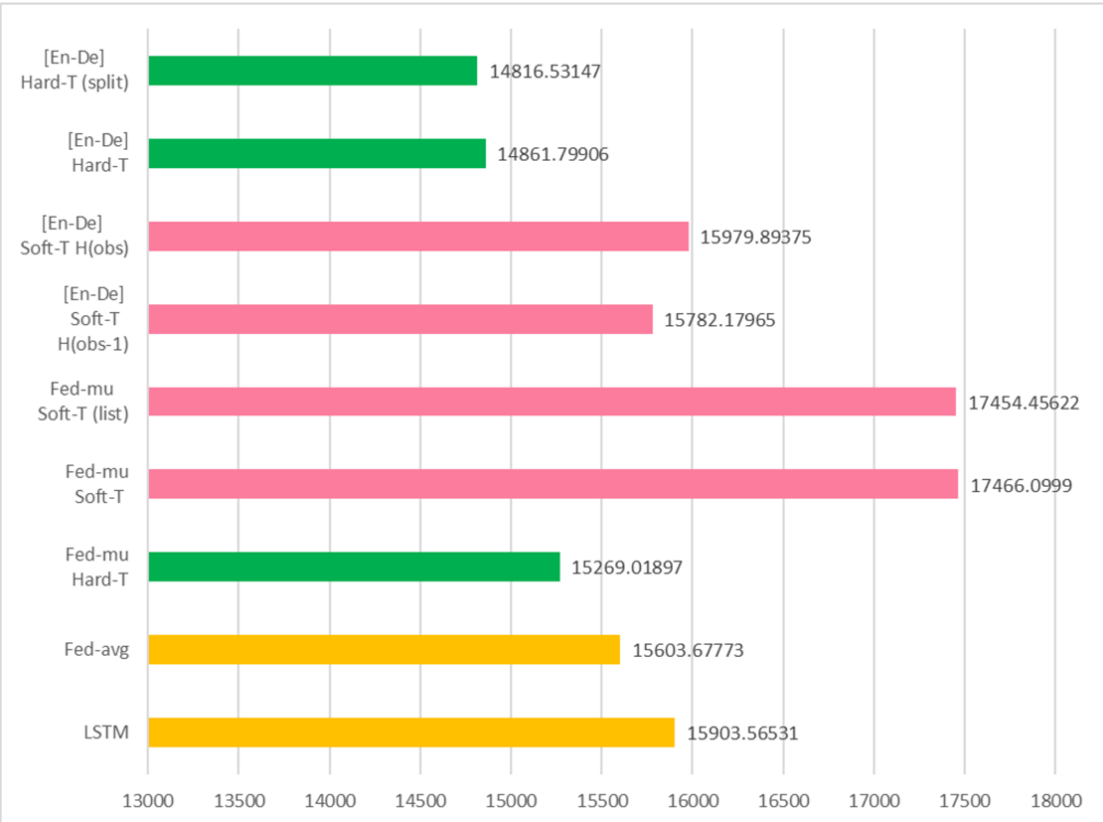


Figure 4: Model ablation experiment (MSE)

3.3 Experiments Result

Figure 4 gives the average MSE of all appliances in tested houses in time interval of 1 min, L_{obs} = 99, L_{pred} = 6. We can see that our proposed method achieves better performance than baselines under three designs. To more intuitively demonstrate the effect of the model, we plot the percentage of performance improvement in Figure 5 (based on Fed-avg).

3.4 Ablation Experiment

In this section, we hope to conduct some ablation experiments to verify the effectiveness of the model while also attempting to further improve the model.

3.4.1 Hard Target and Soft Target

Here, we attempted to investigate the performance impact of different target selection methods in knowledge distillation on the model. For Hard Target, we have both the Meme model and the Personal model output predictions for future power consumption O_m and O_p and use formula (5) to calculate L_{mu}. The result of MSE is given in (Figure 4, **Fed-mu Hard-T)**.

On the other hand, for Soft Target, we have tried various soft target selection methods. Firstly, we choose the hidden layer outputs of LSTM as the soft target to train our model, because it contains the sequence information prior to its time point. Specifically, we use the last hidden layer output $H_{m,last}$ and $H_{p,last}$ to calculate a new L_{mu} as a part of the total loss function. The MSE result of this way is given in (**Fed-mu Soft-T**, Figure 4). We also tested collecting the hidden layer outputs throughout the entire prediction process as soft target, but the effectiveness was not significantly different from using the output of the last step directly, **Fed-mu Soft-T(list)** (Figure 4). The selection of soft targets in Encoder-Decoder is vary slightly due to changes in model structure, and will be explained in detail in the following Encoder-Decoder section.

3.4.2 Encoder-Decoder

Encoder-Decoder [6] is a method of using a model to encode a sequence of symbols into a fixed length vector representation, and then using another model to decode this fixed length vector into the target symbol sequence. We hope to integrate this technology into the seq2seq-FML model to obtain a better encoder for feature extraction, while also enabling decoders in various households to better adapt to the distribution characteristics of local data.

The Encoder-Decoder structure is shown in Figure 2. The change in the structure of this model does not have a significant impact on the calculation of hard target loss. We directly use the output of two FC layers in Decoder to calculate the L_{mu} by formula (3). However, for Soft Target, we design two new methods to calculate L_{mu}. In the first way, the hidden layer output from Decoder and the last observation data are chosen as the Soft Target (red and yellow part in Figure 2), since the former contains temporal information of the observed data, while the latter provides anchor points for the current data values. Another method uses hidden layer information from the previous moment (blue and yellow part in Figure 2), to avoid carrying too much duplicate information in the output of the hidden layer, as the observed values in the last step have already been used as part of the feature for calculating L_{mu}. The performance of these two methods is shown separately in the figure 4, (**[En-De] Soft $\boldsymbol{H_{(obs)}}$** and **[En-De] Soft $\boldsymbol{H_{(obs-1)}}$**).

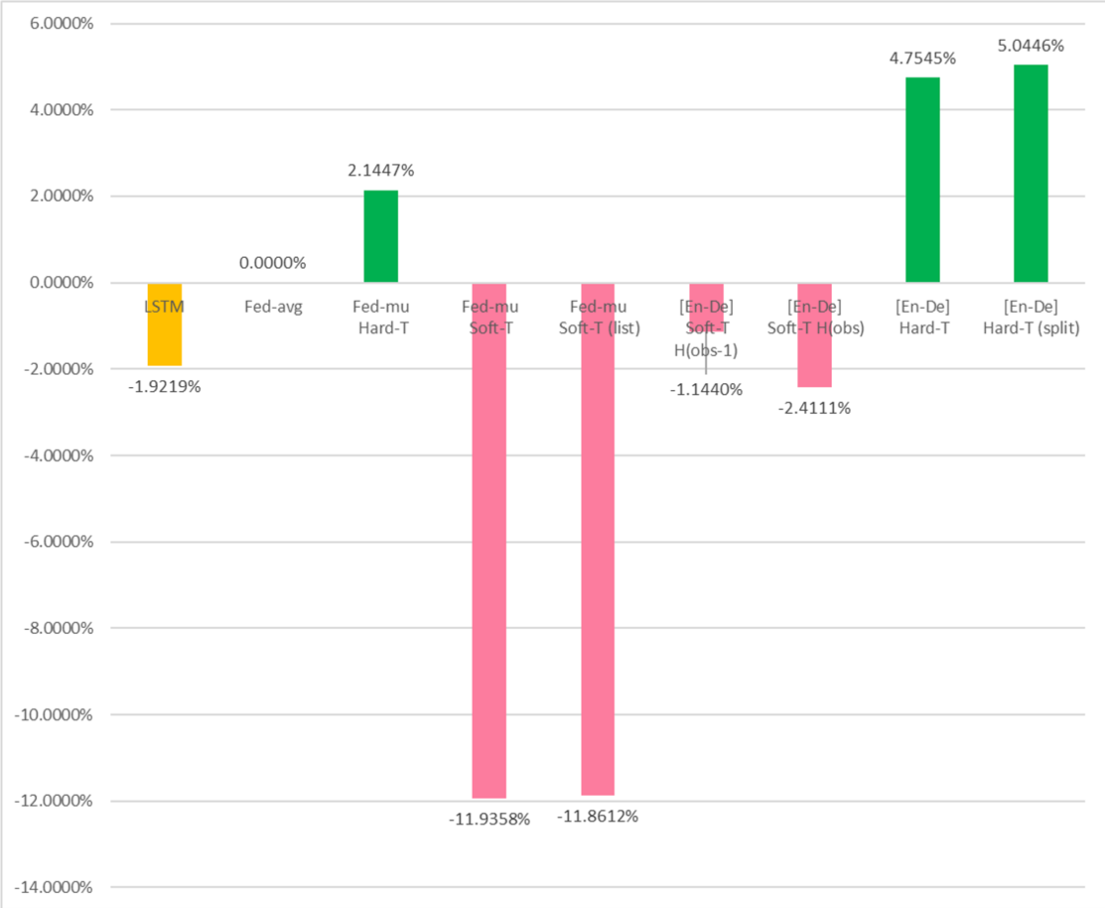


Figure 5: Model ablation experiment (%) (Base on Fed-avg)

3.4.3 Split Linear Layer Output.

During the visual analysis of the waveform predicted by the experiment, we found that our model seems to have learned too much about the interaction between different electrical appliances, but this seems to be false most of the time. To reduce the tendency of our model, we decoupled the linear layer of the final output of our model, which allowing each electrical appliance to have its own linear layer output. Experiments have shown that this method achieves a slight performance improvement on the existing basis, resulting in a total improvement of 5.0446% (Figure 5).

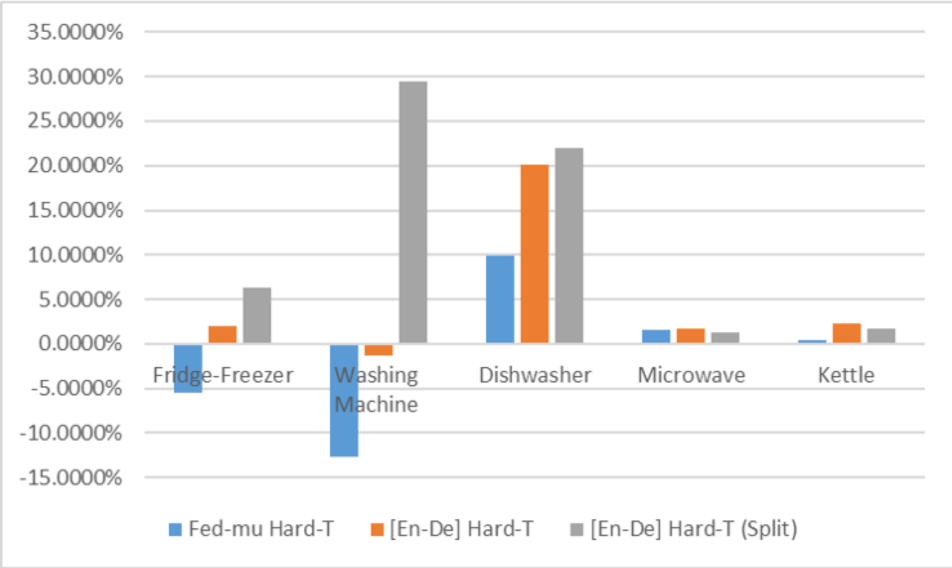
3.5 Result and Analyze

It can be seen, in Figure 4, the seq2seq FML framework has achieved a 2.1447% improvement compared to the fed avg method. After adding the Encoder-Decoder model, the performance has further improved to 4.7545%.

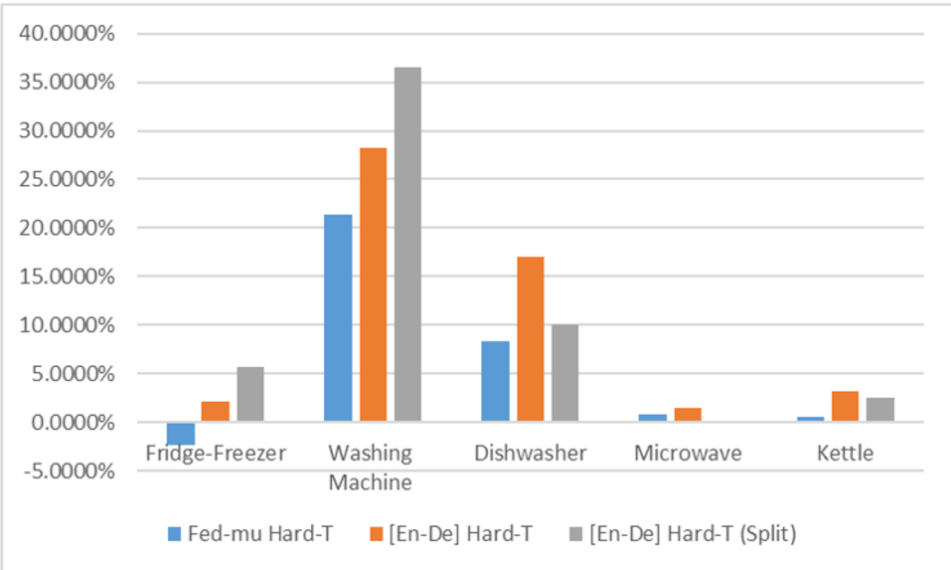
In Figure 4 and Figure 5, We also showed the results of using soft targets as the mutual loss function. Although this attempt did not have a positive effect, even so, comparing to add the Encoder-Decoder method, the performance achieved by the two training methods has also been improved, which also proves that the Encoder-Decoder structure that we designed can better predict the power consumption of household appliances by sharing the Encoder to better extract features only.

In order to analyze the performance of the model in more detail, we compared the performance of different household appliances on different models that have improved performance on average (Figure 6). In the figure, we can see that in different households, the prediction accuracy of these five types of electrical appliances has mostly improved, and the former three types (Fridge-Freezer, Washing Machine, Dishwasher) are dominated. This may be due to the better periodicity and longer usage time of the first three types of electrical appliances, which means that their electrical characteristics will be better captured in the same time span of data. For the latter two types of electrical appliances, their use may be more random and have a shorter duration, but our model has also achieved some results.

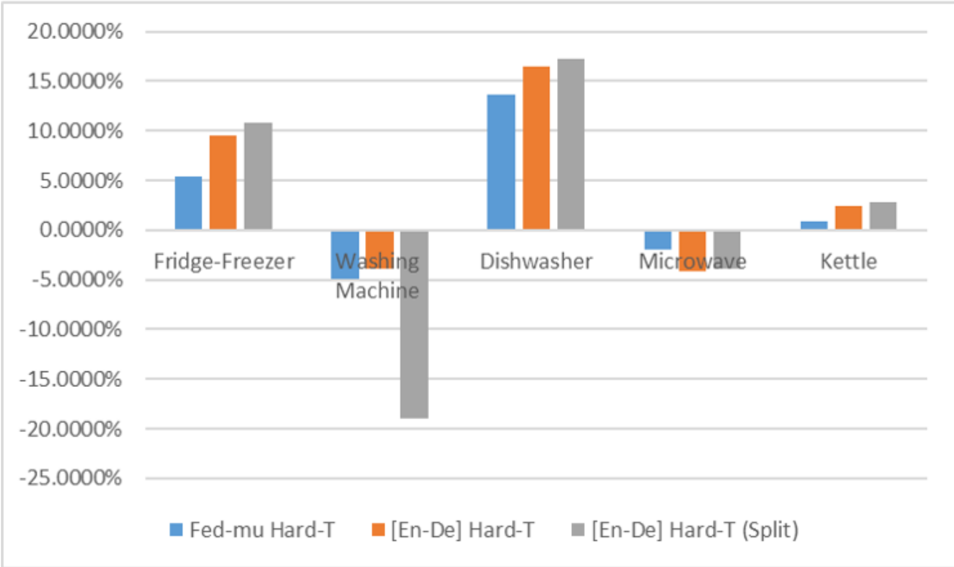
In addition, we have noticed that in House 5, all the three models have had opposite effects in predicting washing machines and microwave ovens. It is preliminarily suspected that the low frequency of use of these two types of electrical appliances in the household has resulted in the loss of too much effective information in our model during the personalization process, or made it difficult for local decoders to learn the mapping characteristics of these two types of electrical appliances, However, further experiments are still needed to prove this.



(a) House 2



(b) House 3



(c) House 5

Figure 6: Performance of different household appliances on various models

4 Conclusion

This paper studies the personalized load forecasting problem of household appliances, revealing a series of challenges and providing a new solution for load forecasting at the appliance level.

First, the federal learning mode is used to train the power consumption data of a variety of electrical appliances from different families, which to a large extent ensures the privacy and safety of users. In the context of the popularity of smart meter, it is possible to promote the power consumption prediction technology on a large scale.

Secondly, in response to the common problem of data heterogeneity in federated learning, we introduced the FML framework, which utilizes mutual learning between two models in local models to better adapt the local personalized

model to the local data distribution and achieve better prediction results. At the same time, a special adaptation has been made to the time series prediction problem, proposing a seq2seq-FML framework suitable for predicting the power consumption of multiple households and appliances.

Finally, we conducted extensive ablation experiments on the seq2seq-FML model, demonstrating the effectiveness of the Encoder Decoder structure and the Hard-Target knowledge distillation method in this model. We also fine-tuned the model structure to improve its prediction accuracy on test data by approximately 5.04% compared to the traditional Fed-avg method.

5 Shortcomings and Future Work

Despite achieving these results, the robustness and generalization of the model still need to be tested on larger datasets due to limitations in dataset size. In addition, in the future, we also hope to explore the performance under unbalanced tasks, such as the effectiveness of the seq2seq model in different households with different electrical appliances, or further improve the model to not only provide better personalized performance in data distribution but also in user usage.

References

[1] ADMINISTRATION, U. E. I. Annual Energy Outlook 2011: With Projections to 2035. Government Printing Office, 2011.

[2] AI, S., CHAKRAVORTY, A., AND RONG, C. Household power demand prediction using evolutionary ensemble neural network pool with multiple network structures. Sensors 19, 3 (2019), 721.

[3] ALHUSSEIN, M., AURANGZEB, K., AND HAIDER, S. I. Hybrid cnn-lstm model for short-term individual household load forecasting. Ieee Access 8 (2020), 180544–180557.

[4] AOUAD, M., HAJJ, H., SHABAN, K., JABR, R. A., AND EL-HAJJ, W. A cnn-sequence-to-sequence network with attention for residential short-term load forecasting. Electric Power Systems Research 211 (2022), 108152.

[5] BRIGGS, C., FAN, Z., AND ANDRAS, P. Federated learning for short-term residential load forecasting. arXiv preprint arXiv:2105.13325 (2021).

[6] CHO, K., VAN MERRIËNBOER, B., GULCEHRE, C., BAHDANAU, D., BOUGARES, F., SCHWENK, H., AND BENGIO, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078 (2014).

[7] FEKRI, M. N., GROLINGER, K., AND MIR, S. Distributed load forecasting using smart meter data: Federated learning with recurrent neural networks. International Journal of Electrical Power & Energy Systems 137 (2022),

107669.

[8] GOU, J., YU, B., MAYBANK, S. J., AND TAO, D. Knowledge distillation: A survey. International Journal of Computer Vision 129 (2021), 1789–1819.

[9] HAMMAD, M. A., JEREB, B., ROSI, B., DRAGAN, D., ET AL. Methods and models for electric load forecasting: a comprehensive review. Logist. Sustain. Transp 11, 1 (2020), 51–76.

[10] HAN, D.-M., AND LIM, J.-H. Smart home energy management system using ieee 802.15. 4 and zigbee. IEEE Transactions on Consumer Electronics 56, 3 (2010), 1403–1410.

[11] HINTON, G., VINYALS, O., AND DEAN, J. Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531 (2015).

[12] KHAN, N., HAQ, I. U., KHAN, S. U., RHO, S., LEE, M. Y., AND BAIK, S. W. Db-net: A novel dilated cnn based multi-step forecasting model for power consumption in integrated local energy systems. International Journal of Electrical Power & Energy Systems 133 (2021), 107023.

[13] KUMAR, C. J., AND VEERAKUMARI, M. Load forecasting of andhra pradesh grid using pso, de algorithms. Int J Adv Res Comput Eng Technol (2012).

[14] LIOR, N. Sustainable energy development: the present (2009) situation and possible paths to the future. Energy 35, 10 (2010), 3976–3994.

[15] MANSOUR, Y., MOHRI, M., RO, J., AND SURESH, A. T. Three approaches for personalization with applications to federated learning. arXiv preprint arXiv:2002.10619 (2020).

[16] MCMAHAN, B., MOORE, E., RAMAGE, D., HAMPSON, S., AND Y ARCAS, B. A. Communication-efficient learning of deep networks from decentralized data. In Artificial intelligence and statistics (2017), PMLR, pp. 1273–1282.

[17] MURRAY, D., STANKOVIC, L., AND STANKOVIC, V. An electrical load measurements dataset of united kingdom households from a two-year longitudinal study. Scientific data 4, 1 (2017), 1–12.

[18] NTI, I. K., TEIMEH, M., NYARKO-BOATENG, O., AND ADEKOYA, A. F. Electricity load forecasting: a systematic review. Journal of Electrical Systems and Information Technology 7, 1 (2020), 1–19.

[19] SAJJAD, M., KHAN, Z. A., ULLAH, A., HUSSAIN, T., ULLAH, W., LEE, M. Y., AND BAIK, S. W. A novel cnn-gru-based hybrid approach for short-term residential load forecasting. Ieee Access 8 (2020), 143759–143768.

[20] SAVI, M., AND OLIVADESE, F. Short-term energy consumption forecasting at the edge: A federated learning approach. IEEE Access 9 (2021), 95949–95969.

[21] SHEN, T., ZHANG, J., JIA, X., ZHANG, F., HUANG, G., ZHOU, P., KUANG, K., WU, F., AND WU, C. Federated mutual learning. arXiv preprint arXiv:2006.16765 (2020).

[22] TAÏK, A., AND CHERKAOUI, S. Electrical load forecasting using edge computing and federated learning. In ICC 2020 - 2020 IEEE International Conference on Communications (ICC) (2020), pp. 1–6.

[23] WANG, Y., GAN, D., SUN, M., ZHANG, N., LU, Z., AND KANG, C. Probabilistic individual load forecasting using pinball loss guided lstm. Applied Energy 235 (2019), 10–20.

[24] WANG, Y., GAO, N., AND HUG, G. Personalized federated learning for individual consumer load forecasting. CSEE Journal of Power and Energy Systems (2022).

[25] YOUSEFI, M., HAJIZADEH, A., AND SOLTANI, M. Energy management strategies for smart home regarding uncertainties: State of the art, trends, and challenges. In 2018 IEEE International Conference on Industrial Technology (ICIT) (2018), IEEE, pp. 1219–1225.

[26] ZHANG, Y., TANG, G., HUANG, Q., WANG, Y., WU, K., YU, K., AND SHAO, X. Fednilm: Applying federated learning to nilm applications at the edge. IEEE Transactions on Green Communications and Networking (2022).

[27] ZHANG, Y., XIANG, T., HOSPEDALES, T. M., AND LU, H. Deep mutual learning. In Proceedings of the IEEE conference on computer vision and pattern recognition (2018), pp. 4320–4328.

[28] ZHOU, B., LI, W., CHAN, K. W., CAO, Y., KUANG, Y., LIU, X., AND WANG, X. Smart home energy management systems: Concept, configurations, and scheduling strategies. Renewable and Sustainable Energy Reviews 61 (2016), 30–40.

Acknowledgment

This is my first attempt at complete academic paper writing. Every step of my undergraduate thesis cannot be separated from the great help of my teachers. Here, I particularly thank Dr. Xiaodan Shi, Prof. Xuan Song, and Prof. Haoran Zhang, for their strong support and assistance. I will also thank the hardware equipment support provided by the Computer Department of the Southern University of Science and Technology.