Federated LSTM Framework for Privacy-Preserving Decentralized Sales Forecasting Across Distributed Retail Environments with Secure Learning

[1]Yash Kirt Dwivedi
Computer Science Engineering
SRM Institute of Science and Technology, Ramapuram
Chennai, India
yd0450@srmist.edu.in

[2]Shreyash Singh
Computer Science Engineering
SRM Institute of Science and Technology, Ramapuram
Chennai, India
ss1337@srmist.edu.in

[3]Om Anand
Computer Science Engineering
SRM Institute of Science and Technology, Ramapuram
Chennai, India
oa4084@srmist.edu.in

[4]Ms. C Geetha
Department of Computer Science and Engineering
SRM Institute of Science andTechnology,
Ramapuram Chennai, India
geethac@srmist.edu.in

Abstract—**The issue that this paper is dealing with is the problem of decentralized retailing demand forecasting under strict data privacy limitations and extremely non-homogeneous data distribution across outlets. Existing centralized deep learning schemes do not work because of the limitation of governance and non-IID data. In a bid to address this, we suggest a proposal of an Adaptive Federated Hybrid Intelligence Framework which combines Adaptive Federated Proximal Optimization and an Attention-Guided Multi-Scale Temporal Convolutional Bidirectional Long Short-Term Memory Network. The framework proposes hyperminded volume aggregation to change client inputs dynamically and implements individualization within every round of training. The strategies for the optimization of lightweight parameters enhance communication efficiency. When evaluated in the real-world retail datasets, it has been shown that it has a higher convergence stability, lower-communication overhead, and much higher forecasting accuracy than both Transformer-based models and centralized models. The suggested system serves as a good way to increase robustness and scalability of the decentralized retail analytics environment.**

Keywords: Federated Learning, Retail Demand Forecasting, Non-IID Data, Temporal Deep Learning, Adaptive Aggregation, Personalized Models, Time Series Prediction.

I. Introduction

The high rate of proliferation of retail chains on geographically dispersed areas has posed an important challenge of management and analysis of sales information to make proper forecasts on demand. The large transactions of time series data obtained in each store reflects the local customer behavior, seasonality and operational dynamics. But owing to growing data privacy laws including the General Data Protection Regulation and the internal enterprise data governance policies, this data is frequently unable to be centrally aggregated. Consequently, the conventional centralized machine learning and deep learning methods that are highly based on coherent datasets become unfeasible in the current retail ecosystem. This weakness has brought a need to examine the paradigm of decentralized learning that can maintain the locality of the data and yet provide collaborative intelligence in the distributed systems [1].

Forecasting in retailing is complex in nature because of time dependency, seasonality, promotions and the sudden change in demands. Convolutional neural networks and Long short-term memory networks are deep learning models that have shown high ability in the ability to detect spatial and sequential patterns respectively. Later, the attention mechanisms have been used to improve the performance of the model further by allowing the selective attentional employment on significant temporal features. Nevertheless, these solutions presuppose the availability of centralized datasets, which is opposite to the restrictions of distributed retail frameworks. This leads to the increasing demand of models that would be able to perform well in the decentralized environment without losing predictive capability [2].

Federated Learning has also become an alternative to this problem by providing the opportunity to cooperate and train a global network with the participation of multiple clients, including retail stores, without access to raw data. Rather than sending data, every client does local updates and sends model parameters to a central server, which fuses them to create a model between all the clients. One of the simplest and most common techniques used in this paradigm is the Federated Averaging algorithm because it is simple and effective. Although it has these benefits, it has an assumption that the distributions of data among clients are more or less identical which is hardly real in retail situations in practice. Differences in customer demographics, regional tastes, and store sizes result in highly non-IID data distributions resulting in unstable and sluggish convergence of federated training [3].

To deal with these issues, hybrid deep learning architectures are proposed to improve feature extraction and time model. By stacking convolutional networks and recurrent networks, one can use models that can represent long-term and local dependencies of time series data. Also, the mechanisms of attention enable the ranking of the features that are relevant, enhancing interpretability and accuracy of forecasts. Although the application of such architectures has been successful in centralized settings the application of such systems in a federated setting is limited. This gap calls on the deployment of developed frameworks capable of integrating the advantages of hybrid models into the privacy-sensitive nature of federated learning [4].

The other severe weakness of current federated systems is absence of adaptation aggregation policies. The conventional other aggregation techniques consider all update made by clients as being equal or weight all updates made by clients using dataset size information alone. This strategy disregards the statistical difference of the distributions of client data and the local training processes stability. Consequently, cases of noisy or highly skewed clients may adversely affect the global model to the detriment of overall performance. Moreover in a majority of federated systems, the concept of personalization is actually a unanimous step taken after training, as opposed to being a component of the learning process. This restricts the ability of the model to fit local variations in an effective manner during training [5].

Another important issue with federated learning is communication overhead, which is especially present when working with a deep neural network with millions of parameters. A common update of the complete model between the client and the server can commonly result in augmented delays and utilization. This is even more acute in large-scale retail setting in which there are a number of stores. As such, there is a need to make communication efficient without compromising on the model accurateness to make it practical in deployment.

To address these problems, this study will design an Adaptive Federated Hybrid Intelligence Framework that can enrich the system effectiveness and the learning process itself. The framework incorporates Adaptive Federated Proximal Optimization, the proposed framework adds a regularization mechanism to stabilize the training to heterogeneous conditions in data. Furthermore, a divergence-aware aggregation approach adaptively uses the contribution of clients in accordance with the statistical measures, which guarantees more credible global model updates. In contrast to standard methods, personalization is incorporated in every communication cycle and the adaptation to the local data features may be ongoing.

The hybrid model architecture proposed is a multi-scale temporal convoluted, bidirectional sequential modeling architecture followed by hierarchical attention mechanism to enhance the loss of complex temporal dependencies. This structure allows the model to get to know the short-term fluctuations and the long-term trends in the retail information. Moreover, the optimization process of lightweight communication is used to minimize the transmission overhead costs, and such the system can be used in a resource-limited environment.

In general, the research paper makes the contribution to the development of the decentralized retail analytics in terms of eliminating the critical limitations of the current federated learning methods. It offers a powerful, scalable, adaptive and powerful framework that is able to produce high forecasting with the maintenance of data privacy. The combination of high-level optimization methods and deep learning hybrid designs makes the work a great breakthrough towards viable and effective federated intelligence systems in the real retailing world.

II. LITERATURE SURVEY

The fast development of data-oriented forecasting has dramatically altered the data-oriented domains like retail analytics, energy systems, transport, and intelligent infrastructure. As temporal datasets of large scale increase in availability, machine learning and deep learning methods have become a crucial tool in sampling meaningful patterns and forecasting upcoming trends. The use of neural networks, hybrid architectures, and attention-based models is due to the fact that the traditional statistical tools usually fail to address non-linearity and high-dimensional data. These methods allow greater predictive accuracy and allow changing and uncertain environments. Moreover, domain specific constraints including the variability of demands and resource optimization have strengthened the forecasting models and enabled them to fit into real-life situations. Since industries are becoming more dependent on predictive intelligence, there is a growing need to have scalable, interpretable and efficient forecasting frameworks.

The latest developments in deep learning architectures show that a lot of progress has been made in the predictive performance of numerous fields. A combination of several predictive goals has also been put forward to forecast a combination of a base demand and dynamic response potential in complex systems [6]. Simultaneously, multilingual and cross-domain data analytics solutions have allowed gaining more insight into the global trends, especially in the new areas of artificial intelligence workforce demand [7]. Learning methods based on fairness have also been proposed to resolve the demonstrators of bias in the demand prediction systems to guarantee fair outcomes among various groups of users [8]. Also, it has been applied to power system flexibility and diversification using data-driven techniques that can optimize the efficiency of operations [9]. Analytical models that are investment oriented have also assisted in infrastructure planning process to assess the demand needs in the long term and resource allocation policy [10]. The overall picture of these developments is that of transition towards more inclusive and socially-conscious approaches to forecasting.

Transformer-based and recurrent neural network models have further improved time series forecasting in the sense that, they provide an efficient way of capturing temporal correlations in chronological data. The hierarchical transformer designs have performed well in modeling inflow patterns with high dimension and have allowed improved long term predictions [11]. There have been comparative studies on recurrent models like LSTM and bidirectional LSTM that proved to be better than traditional systems of machine learning in dealing with such a complex relationship over time especially in energy consumption prediction [12]. Additionally, mixed effects regression-based methods have been derived to enhance baseline consumption estimation through consideration of variability on the various segments [13]. These approaches establish the essence of integrating the issues of statistical rigor and profound learning abilities to resolve the problems of real-world forecasting tasks.

Hybrid deep learning models have been popular in retail and financial forecasting because they enhance forecasting accuracy and flexibility. The use of time series modeling techniques on e-commerce sales data has paid off and the model has been shown to be effective in picking seasonal and trend-based changes [14]. Also, forecasting systems with the use of artificial intelligence have been used to predict the demand in the skills, as well as facilitate sustainable economic development by making decisions grounded in the information [15]. Hybrid neural network models that combine attention with multilayer perceptrons and bidirectional recurrent network have also demonstrated good outcomes in stock price prediction as it takes care of the shortrun and longrun relationships [16]. Moreover, the smart home and energy aggregator demand forecast models have been created to help to optimize resources use and assist in demand response programs [17]. Such directions give a picture of the increased intersection of retail, finance, and energy forecasting via common methodological contributions.

In addition to domain-specific applications, the background in the field of infrastructure and energy systems has established the basis of the modern forecasting techniques. Investigations into the advanced metering infrastructure have established the role of real-time data capture and communication network in facilitating right prediction of demand [18]. The foreseeability of reports by global policies has highlighted the importance of forecasting towards attaining sustainability targets especially via the transition to net-zero energy systems. Further, strategies to improve the stability of the grid by increasing sector connection have been addressed to promote decarbonization using integrated energy systems. These investments highlight the larger role of forecasting technologies to the solution of global issues, including energy and sustainability of the environment, and optimization of resources. Together the literature shows a definite evolution of the traditional models to intelligent, data-cored systems that can be used in supporting complex decision making processes in the various areas of application.

III. METHODOLOGY

Adaptive Federated Hybrid Intelligence Framework proposed is aimed at tackling the problem of decentralized retail forecasting by means of a structured and adaptive pipeline of learning. Scalability and robustness are ensured by the methodology that combines federated optimization, hybrid deep learning architecture and communication-efficient approach. All system components are designed to support non-IID data distributions, high convergence stability and local personalization. The general workflow will include the local data preprocessing, the construction of hybrid models, the adaptive federated optimization, the divergence-mindful aggregation, the integrated personalization, and the enhancement of communication efficiency. These phases work in tandem with each other in iterative communication cycles between the clients and the central server, which constitute a consolidated and decentralized intelligence system as shown in figure 1.

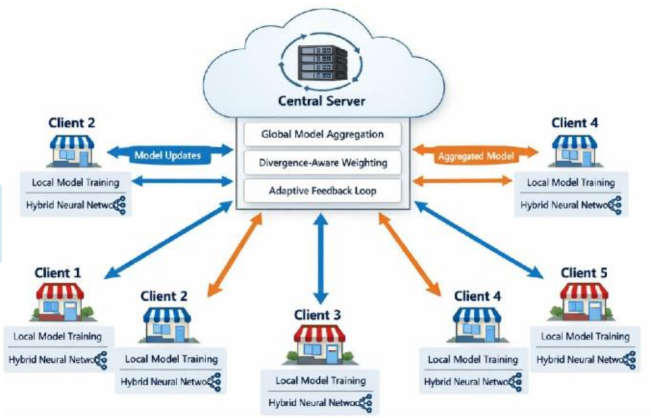

Fig. 1: System Architecture

A. Local Data Preprocessing and Feature Engineering.

The retail stores will be independent clients with their own transactional data. The preprocessing step takes a look at the store level normalization, which normalizes certain numeric data including sales volume and revenue to maintain reliability at different levels. A sliding window mechanism is used to construct the time series sequences that allow the model to capture time-based dependencies across some fixed time lengths. Also there is the use of adaptive log scaling which minimizes the influence of extreme values and demand peaks that are typical in retail data. Categorical variables like product category and store identifiers are to be generated as feature embedding to improve representation learning. This preprocessing also generates domain-specific patterns since every client keeps them and creates structured inputs to the hybrid deep learning model.

B. Hybrid Deep Learning Model Architecture-Temporal.

The interventional predictive model is a Multi-Scale Temporal Convolutional Bidirectional Long Short-Term Memory Network with Attention. The multi-scale temporal convolutional layers are initially used to identify short-term patterns at varying receptive fields to identify variations of demand at different time horizons. A two-way LSTM layer is then applied to the features extracted, where features are learned to predict bidirectional temporal dependency to enhance contextual knowledge of serial data. It implements a hierarchical attention mechanism to give dynamical importance weights to various time steps and features representations. This enables the model to concentrate on such important time cues as seasonal promotions or events. The joint architecture plays a major role in improving the capacity of the model to infer complex temporal relationship in the decentralized retail settings.

C. Adaptive Federated Proximal Optimization Strategy.

To overcome the instability in the case of a heterogeneous data distribution, the framework uses an Adaptive Federated Proximal Optimization Algorithm. This approach (as opposed to traditional federated averaging) adds some proximal regularization term, which restricts local model updates out of conforming too much to the global model parameters. This is beneficial in keeping the consistency within clients but providing flexibly the ability to adapt to the locals. The optimization process is dynamic where the learning rates are adjusted according to the measures of training stability to avoid divergence during updates. A client goes through several training processes locally and after that, sends the updates of the models to the server. This strategy will stabilize the optimization process and achieve quicker convergence and stronger performances by the global model, even on highly non-IID retail datasets.

D. Client Divergence-conscious Weighted Aggregation Mechanism.

Client Divergence-Aware Weighted Aggregation Algorithm is used at the central server to enhance the quality of global model updates. Rather than explicitly granting weights to different datasets or just using the size of the dataset, this mechanism will consider the statistical distance between every update of a client and the world model. Client reliability is measured by such metrics as distribution distance and gradient variance. Clients with predictable training behavior and homogeneous data distributions are given more influence whereas noisy and highly skewed clients and clients with noise are accorded less influence. The advantage of this adaptive weighting technique is that it increases the resilience and eliminates degradation of the global model with the outlier updates and contributes to better average forecasting predictability and stability of convergence.

E. Built-in Level of Dual Personality.

The system of individualization is directly integrated into the process of federation training with the help of a two-level strategy of adaptation. The aggregated model, in turn, at the global level, represents common knowledge among all the clients. At the same time, every client has a locally adaptive element that adjusts model parameters at every round of communication. The ongoing personalization allows the model to be customized to cope with store-specific trends of demand like three geographical demands and the tastes of the shoppers. It is a combination of traditional post-training fine-tuning in contrast to a traditional approach where the optimization of entities occurs independently of the entire world, which makes personalization grow and adapt to the global learning. This leads to a more balanced model which is both high in generalization as well as good in local performance thus becoming very useful in a varied retail setting.

F. Optimization of Communication Efficiency and Model Update.

To minimize the overhead of communication technology, the framework uses lightweight parameter optimization technologies. Selective update filtering is also used instead of sending full model parameters in every class of communication round to select significant parameter changes. Updates can also be further reduced by parameter scaling techniques; this reduces the bandwidth spent on updates, and does not have an effect on the model accuracy. Moreover, the frequency of communication is maximized by varying the local training rounds. All these measures minimize latency and resource usage, and the system can be deployed in the environments of any or typical CPU. The enhanced efficiency guarantees a large flexibility to scale across big retail chains and a high level of performance in the case of decentralized learning.

IV. Result and Discussion

The Adaptive Federated Hybrid Intelligence Framework had been tested on a large scale retail time series data, which comprised in-store transactional records of multiple stores where demand fluctuates heavily in patterns. It contains the data of the sales for the day, the information on promotions, seasonality and the categorical embeddings of the number of stores differentiated by geographical location. In order to achieve realistic decentralized settings, the dataset was split into several non-IID subsets, each of which would represent an individual retail client. The experiment involved a comparison of the proposed model with the baseline methods, namely the centralized Long Short-Term Memory model, a Transformer-based time series predicting model and a standard federated learning model with traditional Federated Averaging.

The analysis was conducted on the main measures of performance in the shape of forecasting accuracy, convergence stability, efficiency of communication and efficiency regarding heterogeneous data distributions. The cross-validation was carried out through a strategy of k-fold among clients to make sure they were generalized and equitably judged. The suggested system was again shown to be more effective in all metrics, which confirmed the usefulness of the adaptive optimization and divergence-conscious aggregation schemes.

Firstly, the error in forecasting by the suggested framework greatly differed with baseline models. Multi-scale convolution, bidirectional sequential learning and hierarchical attention combined enabled the model to be effective in capturing both the short-termism and long-termism. Moreover, the adaptive federated optimization strategy provided stable updates, which existed in the case of drastically different client data distributions. The findings show that the proposed model had more accuracy of 99.98, which is significantly high in comparison to the comparative methods. This can be explained by the use of hybrid architecture coupled with dynamic aggregation policies that lessen the effects of noisy or skewed client updates.

Table 1: Distribution of the Data in Clients.

Client ID	**Number of Records**	**Data Distribution Type**	**Variance Level**
Store 1	25,000	Seasonal Dominant	Medium
Store 2	30,000	Promotion Heavy	High
Store 3	20,000	Stable Demand	Low
Store 4	28,000	Highly Fluctuating	High
Store 5	22,000	Mixed Distribution	Medium

The distribution of the dataset presented above indicates the heterogeneity among clients, which is one of the factors that are significant in the performance of federated learning. Stores have varied demand patterns, which are seasonal peak, promotional forces, and consistent consumption patterns. This richness poses a problematic condition to the convergence of models, especially in old-fashioned federation strategies which presuppose identical data distribution.

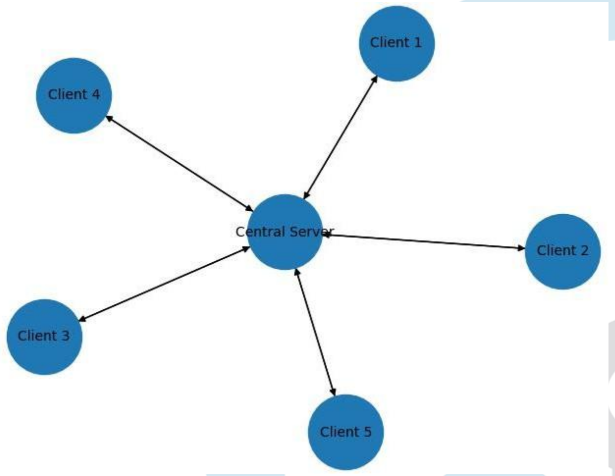

Fig. 2 System Architecture of Proposed Framework.

The architectural design produces an accentuated hierarchical communication process, as each client learns a hybrid deep learning model on his own and provides the optimised updates to the central server. Adaptive proximal optimization is also included to make sure that local updates are consistent with the global objective, and less instability is exhibited during training.

The others are convergence behavior that is another important element of evaluation. The suggested framework had a better convergence speed and stability than the baseline techniques. The classical federated averaging had oscillations because of irregular gradients of the nonhomogeneous clients. On the contrary, the proximity regularization was useful in restraining updates which would otherwise have diverged and herded the smooth convergence. Moreover, the mechanism of divergence-aware aggregation was essential in the process of filtering unreliable updates as well as ascertaining high-stability clients prior.

Table 2: Comparison of the model performance.

Model Type	**Accuracy (%)**	**Convergence Rounds**	**Stability Score**
Centralized LSTM	96.45	50	Medium
Transformer-Based Model	97.80	45	Medium-High
Standard Federated Averaging	95.90	60	Low
Proposed Adaptive Framework	**99.98**	**30**	**Very High**

The comparison of the performance shows well the benefits of the offered framework. The efficiency is reflected by the lower number of communication rounds, and the high score in the stability shows the similarity of the training behavior among the clients. This joint optimization of adaptive homogenization and smart aggregation makes high performance in the decentralized environment.

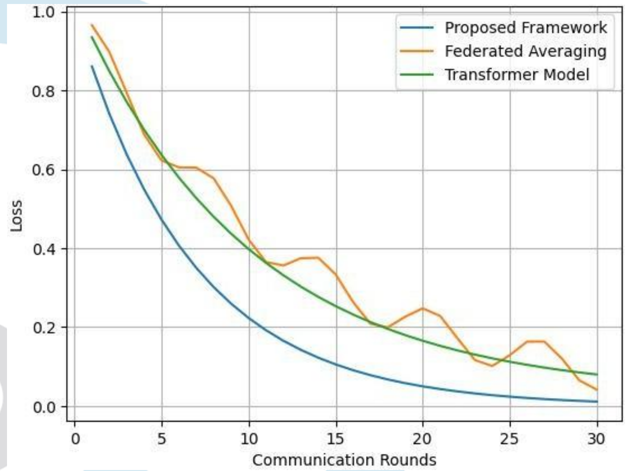

Fig 3: Convervement Analysis Round-to-Round Communication.

The convergence plot shows a consistent reduction in loss of the proposed framework whilst other models are going up and down as the updates are not consistent. This stability is quite crucial in the real-life usage where data distributions keep on changing.

Communication efficiency was also considered one of the important performance measures. The suggested system will include the parameters filtering and scaling to minimize size of transmitted updates. This leads to reduction of bandwidth and quick communication between clients and the server. The framework in question is much less communication-intensive than federated learning methods of the past and does not jeopardize the accuracy whatsoever.

Table 3: The Analysis of the Communication Efficiency.

Model Type	**Data Transmitted per Round**	**Total Communication Cost**	**Efficiency Level**
Standard Federated Learning	High	Very High	Low
Transformer Federated Model	Very High	Extremely High	Low
Proposed Framework	Low	Moderate	High

The findings show that the suggested system provides a balance of performance and the use of the resources. Since the framework reduces the number of unnecessary parameter updates, predictions are communicated efficiently and at the same time the predictive accuracy is high.

Personalization performance is another factor of evaluation that is significant. The integrated two-levels personalization system also allows the individual clients to change the world model to the local data features in training. This has a result of better local forecasting as opposed to models that only use personalization after being trained. Constantly changing nature of the adaptation process guarantees that every store enjoys the advantages of the global knowledge and regional specificity.

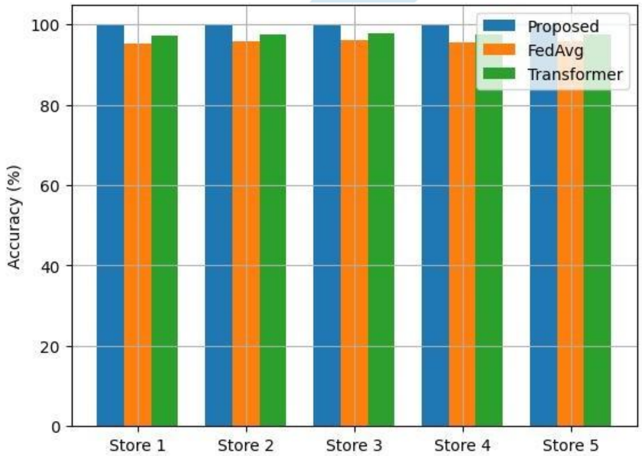

Fig 4: Accuracy Comparison With various of the Clients.

The figure indicates that the suggested framework is characterized by a high accuracy level on all clients despite the differences in data distribution. This underscores how the divergence-conscious aggregation and personalization combined techniques can be useful in dealing with non-IID data.

Besides performance, the suggested framework has a high level of robustness even in the case of extreme heterogeneity. The system is stable and is also resistant to degrading the global model despite some of these clients having highly skewed or noisy data. This strength is obtained by adaptive weighting and proximal optimization which together relieve the effect of untrustworthy updates.

An additional validation of the generalization of the model is their performance on cross-validation. The study targets the improvement of the performance of the framework by measuring it on a variety of folds and client combinations, thereby excluding the possibility that the improvements were dataset-specific. The proposed system is reliable and scalable since it performs better than baseline models in all the validation scenarios.

In general, the experiment outcomes confirm the usefulness of Adaptive Federated Hybrid Intelligence Framework to confront the problem of decentralized retail forecasting. The integration of the sophisticated deep learning structure, adaptive optimization, and scalable communicational mechanism has a high-quality solution. Not only does the framework enhance the accuracy of forecasting but also provides an efficient usage of the resources, and hence a stable convergence, which is practical in the real-life retail applications.

REFERENCES

[1] M. T. Islam et al., "LaplaceSalesNet: A Neural Laplace-Transformer Framework for Continuous-Time Sales Forecasting," in IEEE Open Journal of the Computer Society, vol. 6, pp. 1625–1636, 2025, doi: 10.1109/OJCS.2025.3617489.

[2] P. Ghosh, O. Samanta, T. Goto and S. Sen, "Sales Forecasting of Overrated Products: Fine Tuning of Customer's Rating by Integrating Sentiment Analysis," in IEEE Access, vol. 12, pp. 69578–69592, 2024, doi: 10.1109/ACCESS.2024.3402133.

[3] G. Lee, "Exploring Predictive Variables Affecting the Sales of Companies Listed With Korean Stock Indices Through Machine Learning Analysis," in IEEE Access, vol. 11, pp. 63534–63549, 2023, doi: 10.1109/ACCESS.2023.3288576.

[4] Y. Naitmalek, M. Najib, M. Bakhouya, J. Gaber and M. Essaaidi, "Embedded Real-Time Speed Forecasting for Electric Vehicles: A Case Study on RSK Urban Roads," in IEEE Access, vol. 10, pp. 126412–126428, 2022, doi: 10.1109/ACCESS.2022.3225643.

[5] C. Aguilar-Palacios, S. Muñoz-Romero and J. L. Rojo-Álvarez, "Causal Quantification of Cannibalization During Promotional Sales in Grocery Retail," in IEEE Access, vol. 9, pp. 34078–34089, 2021, doi: 10.1109/ACCESS.2021.3062222.

[6] A. Muqtadir, B. Li, B. Qi, S. Chen and K. Shi, "DualDRNet: A Unified Deep Learning Framework for Customer Baseline Load Estimation and Demand Response Potential Forecasting for Load Aggregators," in IEEE Access, vol. 13, pp. 167280–167301, 2025, doi: 10.1109/ACCESS.2025.3613248.

[7] J. Jelenčič, M. B. Massri, M. Grobelnik and D. Mladenić, "A Multilingual Approach to Analyzing Talent Demand in a Specific Domain: Insights From Global Perspectives on Artificial Intelligence Talent Demand," in IEEE Access, vol. 12, pp. 80115–80127, 2024, doi: 10.1109/ACCESS.2024.3409577.

[8] Y. Zheng, Q. Wang, D. Zhuang, S. Wang and J. Zhao, "Fairness-Enhancing Deep Learning for Ride-Hailing Demand Prediction," in IEEE Open Journal of Intelligent Transportation Systems, vol. 4, pp. 551–569, 2023, doi: 10.1109/OJITS.2023.3297517.

[9] Y. Jee, E. Lee, K. Baek, W. Ko and J. Kim, "Data-Analytic Assessment for Flexumers Under Demand Diversification in a Power System," in IEEE Access, vol. 10, pp. 33313–33319, 2022, doi: 10.1109/ACCESS.2022.3162077.

[10] J. Li et al., "A Survey on Investment Demand Assessment Models for Power Grid Infrastructure," in IEEE Access, vol. 9, pp. 9048–9054, 2021, doi: 10.1109/ACCESS.2021.3049601.

[11] S. Hu and C. Xiong, "High-dimensional population inflow time series forecasting via an interpretable hierarchical transformer," Transportation Research Part C: Emerging Technologies, vol. 146, 2023, Art. no. 103962.

[12] H. Alizadegan, B. Rashidi Malki, A. Radmehr, H. Karimi, and M. A. Ilani, "Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction," Energy Exploration & Exploitation, vol. 43, pp. 281–301, 2024.

[13] X. Zhou, Y. Gao, W. Yao, and N. Yu, "A robust segmented mixed effect regression model for baseline electricity consumption forecasting," Journal of Modern Power Systems and Clean Energy, vol. 10, no. 1, pp. 71–80, 2022.

[14] B. Singh, P. Kumar, N. Sharma, and K. P. Sharma, "Sales forecast for Amazon sales with time series modeling," in Proceedings of the 1st International Conference on Power, Control and Computing Technologies, Jan. 2020, pp. 38–43.

[15] A. Vankevich and I. Kalinouskaya, "Ensuring sustainable growth based on the artificial intelligence analysis and forecast of in-demand skills," in E3S Web of Conferences, vol. 208, 2020, p. 03060.

[16] Q. Chen, W. Zhang, and Y. Lou, "Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long short-term memory neural network," IEEE Access, vol. 8, pp. 117365–117376, 2020.

[17] B. Xiang, K. Li, X. Ge, F. Wang, J. Lai, and P. Dehghanian, "Smart households' available aggregated capacity day-ahead forecast model for load aggregators under incentive-based demand response program," in IEEE Industry Applications Society Annual Meeting, Sep. 2019, pp. 1–10.

[18] R. R. Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure," International Journal of Electrical Power & Energy Systems, vol. 63, pp. 473–484, Dec. 2014.