International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group| Volume 1, Issue 1, 49-57, 2020
ISSN: 3050-922X| https://doi.org/10.63282/3050-922X.IJERET-V1I1P106

Original Article

Federated Learning in Collaborative Supply Chain Forecasting: A Privacy-Preserving Approach

Venkatesh Prabu Parthasarathy
Supply Chain Transformation | Digital Transformation, AI Implementation |IOT/ML Implementation Leader, Lake Forest, California, USA.

***Abstract** - Various stakeholders need to join forces in the modern supply chain to help forecast better, control inventory more efficiently, and save money. Still, the transfer of sensitive data by companies can cause serious privacy and security issues. Because of FL, multiple sites can work together to train a model while still keeping their raw data secured. This article provides a detailed analysis of using federated learning in collaborative supply chain forecasting, underlining its confidentiality features. We focus on how federated learning and supply chain forecasting took shape before 2019, explain some important architecture, and present a framework designed for the supply chain industry. The use of FL on marked datasets has proven that it maintains privacy and still performs accurately. The goal of this paper is to explain how federated learning can help improve teamwork and protect private information for those working in the supply chain.*

***Keywords** - Federated Learning, Supply Chain Forecasting, Privacy-Preserving, Collaborative Learning, Demand Predictions.*

1. Introduction

Supply Chain Management (SCM) covers managing and aligning activities from the acquisition of raw materials right through to getting the finished product to the customer. Well-planned demand, stock quantity, and delivery times are important for reducing costs and offering a high level of service. In this type of supply chain, different companies exchange information to raise their overall performance. [1-3] This type of sharing may cause people to worry about privacy, how it affects their competition, and whether they are complying with regulations. FL is a system that lets various users join forces to educate one model, while still storing their training data on their devices. At first, FL was launched by Google just for mobile device applications in 2016, but it has become popular in different fields where privacy needs to be protected. The main benefit is that only the model or its gradients are sent, which cuts down the risk of exposing the data.

1.1. Need for Privacy-Preserving Collaborative Forecasting

In today's supply chain world, getting the forecast right is vital for reducing costs, keeping stock aligned, and pleasing customers. But, usually, high-quality forecasts can only be achieved through data combination from all the partners in the supply chain. Because we rely on each other within science, this becomes a key problem: What steps should organizations take to ensure they do not jeopardize their private information, unfair advantage, or break any rules through the sharing of insights?

Need for Privacy-Preserving Collaborative Forecasting
Increasing Demand for Data Collaboration
Data Sensitivity and Competitive Concerns
Regulatory and Compliance Challenges
Federated Learning as a Solution

Fig 1: Need for Privacy-Preserving Collaborative Forecasting

- **Increasing Demand for Data Collaboration:** With more parts of the supply chain being connected digitally, isolated groups of data make it harder for forecasting models to be accurate and broadly useful. By involving all these parties in forecasting, it is possible to achieve much better predictions. When sales trends, lead times, promotion schedules, and regional demand are all shared, partners can decide upon actions that are more in line with the overall goals. Traditionally, teams used to share private data, and this was too great a risk for most organizations.

- **Data Sensitivity and Competitive Concerns:** Data found in supply chains regularly contains strategic information on prices, the habits of customers, demand for products, and how suppliers are performing. Handling data with outside partners, even those who you trust, may endanger a business by allowing others to steal competitively valuable data or change the company's plans. In addition, companies working in highly controlled industries such as healthcare or finance are subject to very strict data protection rules, making sharing data more complex.
- **Regulatory and Compliance Challenges:** Due to regulations such as the GDPR in the EU, CCPA, and specific rules for industries, the collection, storage, and sharing of data are now being watched more closely by lawmakers. If one of these frameworks is broken or not upheld, businesses can end up paying penalty fees and risk serious legal concerns.
- **Federated Learning as a Solution:** Federated Learning is an attractive way to address these problems. Thanks to FL, each company in the supply chain can store its data on its system, yet still contribute to a common forecasting model. By adopting this method, people's privacy is protected, and more stakeholders are willing to share their data. Thus, FL allows groups of people to forecast together without worrying about data accuracy or security.

1.2. Federated Learning in Collaborative Supply Chain Forecasting

FL makes it possible for different companies in the supply chain to cooperate in forecasting without exposing their data to others. To use traditional collaborative forecasting, each party is expected to hand over its data, such as sale records from earlier years, inventory information, or a list of shipments, to a central server or third party for the model to be developed. [4,5] While it offers good forecasting results because it covers all parts of the supply chain, it introduces significant risks regarding privacy, security, and meeting regulations. Through federated learning, training occurs locally at each client, meaning that data stays close to the manufacturers, suppliers, distributors, or retailers. Gradients or weight changes are the only model components passed to a central location, where they are averaged to change the core model. It respects the privacy of all partners by enabling them to contribute to the group's intelligence.

Thanks to the diverse types of data within a supply chain, FL can provide a lot of benefits. Due to the differences in where they operate and who they serve, non-IID data is often collected by each partner. Since traditional models are not well-equipped for diversity, federated approaches keep data local in the training process, helping the global model fit more effectively in various situations. With FL, trading partners do not have to work together at the same time or with synchronized systems since they can join the process at different rates and times. Because it relies less on one central data location, FL greatly reduces the chances of data exposure, espionage, and lawbreaking. Trust between supply chain partners requires this ability, mainly in industries like healthcare, pharmaceuticals, or defense, where the security of data matters the most.

In addition, having to send just the model updates instead of the whole data means FL is easy to scale and useful for situations with low bandwidth or for supply networks that are spread far apart. All in all, using federated learning for forecasting in supply chain management allows businesses to collaborate while still ensuring data privacy. Thanks to this, supply chain stakeholders can rely on predictive models that give details on time, are reliable, and do not expose any private data belonging to the organization. As supply chains move towards greater complexity and rely on digital tools, FL is highlighted as a leading approach for clever and secure forecasting.

2. Literature Survey

2.1. Supply Chain Forecasting Techniques Pre-2019

Before 2019, supply chain forecasting was mostly made based on traditional statistical tools like moving averages and exponential smoothing. [6-9] The classical approaches were extensively applied because of their ease of use and interpretation to describe time series patterns. However, with the advancement in machine learning, regression analysis, decision trees, as well as neural networks began to gain traction for their ability to model complex nonlinear relations and enhance the forecast. Qualitative research during the period also revealed the benefit of collaborative forecasting amongst supply chain partners towards minimising uncertainty. With all these advancements, worries about privacy and secure sharing of data were not taken seriously and made deeper integration across organizations difficult.

2.2. Privacy-Preserving Machine Learning Approaches

Before the ascent of Federated Learning (FL), other privacy-preserving methods like SMC, HE, and others were investigated for allowing joint data processing while keeping the sensitive inputs private. SMC enabled multiple parties to perform computations over their data collectively without revealing each of the parties' information. In the same manner, HE allowed computations in encrypted data while maintaining privacy during the process. Even though these methods possess strong security guarantees, they were computationally expensive and added a great deal of overhead, which made them unfeasible to be applied in actual data mining on a large scale and difficult to supply chain data.

2.3. Emergence of Federated Learning

Federated Learning, which was properly proposed in 2016 in (3), was a paradigm shift in the decentralized training of models. Rather than processing raw data in a centralized manner, FL allowed several clients to train models locally, only sharing model updates with another central server, such as mobile devices or organizational units. This method maintained the privacy of data as no direct exchange of data occurred, and iterative improvements of the global models were possible. Pre-2019 FL research worked on improving the efficiency of communication in order to minimize the bandwidth consumed during the model update, thereby also adding privacy methods such as differential privacy to enhance the security of sensitive information. Use cases of FL during this period of time mainly focused on such domains as healthcare and finance, where the privacy of data is essential.

2.4. Applications of FL in Supply Chain and Forecasting

Even though there were nascent trials of Federated Learning in supply chains prior to 2019, the initial studies did find the potential of it in demand forecasting and detection of anomalies in an IoT-enabled supply chain. These works showed that FL would allow collaborative analytics over distributed entities without compromising the proprietary data. Notwithstanding, literature was devoid of exhaustive frameworks that equally merged supply chain collaboration with FL's privacy-preserving benefits, hence indicating major thrusts for future research to design sound, scalable solutions optimally suiting convoluted supply chain operations.

3. Methodology

3.1. Federated Learning Architecture for Supply Chain Forecasting

- **System Model:** The Federated learning architecture for supply chain forecasting engages several supply chain partners as clients who train forecast models on separate private datasets deployed locally. The privacy and competition concerns prevent partners from sharing these datasets directly or make them hesitant because of the sensitive nature [10-14] of information contained in these datasets, which comprise sales records, levels of inventories, and demand signals. A central aggregator that a trusted party commonly controls coordinates the order of this process, where only model updates (e.g., gradient or weights) from each client are gathered. The aggregator then carries out model aggregation, whereby it averages these updates in order to enhance a global forecasting model. This iterative process repeats until convergence so that it is possible to achieve collaborative learning without sharing raw data across partners.
- **Data Characteristics and Challenges:** There are special challenges to supply chain data because of its heterogeneous tendencies and distributional differences. Different partners that collect data may vary significantly in volume, format, and quality, which leads to non-Independent and Identically Distributed (non-IID) data across Clients. Furthermore, supply chain data is usually imbalanced, with one hand having huge, rich datasets while the other has sparse or erratic data records. These properties make model training more difficult because the standard techniques for federated learning expect IID data, and may cause model bias or a loss in precision. Hence, the FL framework necessitates expert strategies, such as personalized models, adaptive aggregation, or data augmentation, in order to address heterogeneity in a meaningful way and guarantee reliable forecasting performance for such a wide-ranging supply chain ecosystem.

3.2. Forecasting Model Selection

In this study, we use the Long Short-Term Memory (LSTM) networks to act as the core forecast model used in the prediction of supply chain demand, since these types of networks have been proven to be effective in handling sequential and time series data. One particular RNN, known as LSTM, is specifically set up to overcome the limitations of the traditional RNNs, specifically the problems of vanishing and exploding gradients that interfere with the learning of long-term dependencies. Different from traditional neural networks, LSTMs are based on memory cells and gating mechanisms (input, forget, and output gates), the role of which is to control the information flow between the time steps such that the model can prefer an information retention or removal over long sequences. This capacity is important, especially in the supply chain forecasting where demand conceivably depends non-trivially on prior demand, seasonality, promotion, and other temporal factors. Supply chain data is sequential and quite explorable, often located in trends, cyclic patterns, and sudden jumps caused by external factors – changes on the market, holidays, disruptions, etc.

LSTM networks are good at detection such nonlinear temporal dependencies and are able to model short-term fluctuations and long-term trends well. What is even more, LSTMs utilize multivariate time series inputs to join numerous pertinent elements, including inventory levels, changes in prices, and economic indicators, allowing for to boost in accuracy in forecasts. While classical statistical models such as ARIMA or exponential smoothing assume linearity and stationarity, LSTMs are a much better approach, which is more flexible and powerful and can learn complex, nonlinear relationships without heavy feature engineering. In addition, LSTMs have been proven to outperform for various applications in the supply chain, such as demand forecasting, inventory management, and anomaly detection. By placing LSTM in the federated learning context, each supply chain partner would be able to train a locally fitted model that captures the temporal nature of their data, whereas the central aggregator would use such insights to construct a global forecaster that is robust enough.

3.3. Federated Learning Process

- **Step 1: Server Initializes Global Model:** The following process is the federated learning process that starts with an initial global model, denoted as $w0$, that is initialized by the central server. This first model can be randomly initialized or pre-trained on public or compiled data, if possible. The global model to be used in this work is the initial stage of collaborative learning that is to be iteratively improved across a series of local trainings and aggregations. Reserving a common baseline by distributing this initial model to all the participating agents, the server guarantees consistency of the decentralized training.

- **Step 2: Local Training at Clients:** After the global model $w0$ comes, each client k trains a local model wk independently for each client's private dataset Dk, respectively. This local training phase utilizes client's-unique data pertaining to the supply chain for this localized capture of patterns and temporal dependencies that are specific to that particular partner. In training locally, clients do not share raw data, hence protecting the privacy and adhering to the confidentiality regulations of data. The training usually consists of several iterations of gradient descent or another optimization method to optimize the model with the help of local data.
- **Step 3: Clients Send Model Updates:** Δ to Server Following local learning, every client calculates the model update Δwk , which is the update that defines the difference between the updated local model and the received global model. Instead of sending the raw data over the networks to the client, it is the model's updates alone (e.g., gradients or weights) that are transferred from the clients to the central server. This communication is more bandwidth-friendly and keeps privacy as supply chain sensitive data remains on the client's side. The server gathers updates from all of the involved clients in order to be ready for aggregation.
- **Step 4: Server Aggregates Model Updates:** On receiving responses from all the clients, the server performs the aggregation of the local models with weights to update the global model. The new global model $wt+1$ is calculated as a weighted sum of all the client models wkt, and the weights are given in proportion to the size of each client's dataset compared to the total data $n=\sum nk$. This weighted aggregation makes sure that the contributions made by the clients with larger datasets are more impactful on the global model, and this affordability and fairness of contribution help to enhance the overall generalization of the model.
- **Step 5: Iterative Training Until Convergence:** The server then broadcasts back to the clients the global model $wt+1$, and the algorithm iterates. The global model is refined in each round through the introduction of the varied patterns of local data, observed in various clients. This loop repeats itself until the convergence of the model is achieved, which is usually determined by the threshold of the performance of the model or the threshold of the number of rounds. With such coordinated training, the federated learning framework generates a robust, privacy-preserving forecasting model that leverages the dispersed data in the supply chain without revealing secret information.

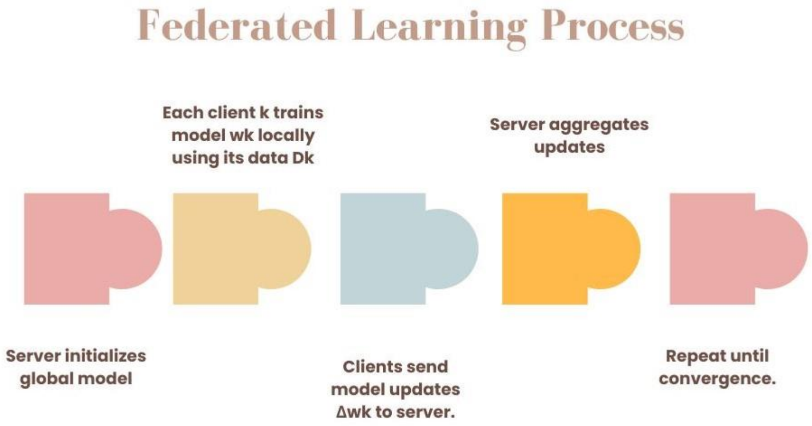
Federated Learning Process
Each client k trains model wk locally using its data Dk
Server aggregates updates
Server initializes global model
Clients send model updates Δwk to server.
Repeat until convergence.

Fig 2: Federated Learning Process

3.4. Privacy Preservation Techniques

- **Model Update Encryption:** To ensure client updates confidentiality while in the federated learning procedure, secure aggregation techniques are used. [15-19] These approaches protect the local model updates by taping them before sending them, so that the central server will only access the aggregated outcome without it seeing one individual client's raw update. Using the cryptographic protocols, such as secure multiparty computation or homomorphic encryption, the server performs aggregation over the encrypted data, thus avoiding any leak of sensitive information from individual partners. This method is necessary in the context of the supply chains, where proprietary data need to be kept secret during collaborative training.
- **Differential Privacy:** Differential privacy is a statistical technique of anonymisation to further improve privacy by injecting purposely controlled noise into the local model updates before transmission to the server. This noise hides client data contribution data points, making it computationally inscrutable for one to conjecture any particular information about a client's dataset from the model updates. Via tuning the noise parameters, equilibrium is set between privacy-preserving and model utility. Differential privacy in the context of supply chain federated learning

prevents the leakage of sensitive patterns of demand fluctuations or inventory levels, and thus optimizes data protection without impeding efficient collaborative forecasting.

- **Client Anonymization:** The client anonymization methods concern the masking of the identities of communicating clients from the central server. By concealing or obscuring client identifiers, the system precludes connecting model updates to distinct supply chain partners, reducing risks of targeted attacks and data linkage. This anonymity protects competitive intelligence, and no sensitive business relationships or patterns of data are revealed. These include methods like pseudonymization, randomized client IDs, or the use of secure communication lines, which enhance anonymity in federated learning environments, building collaborators' trust within supply chains.

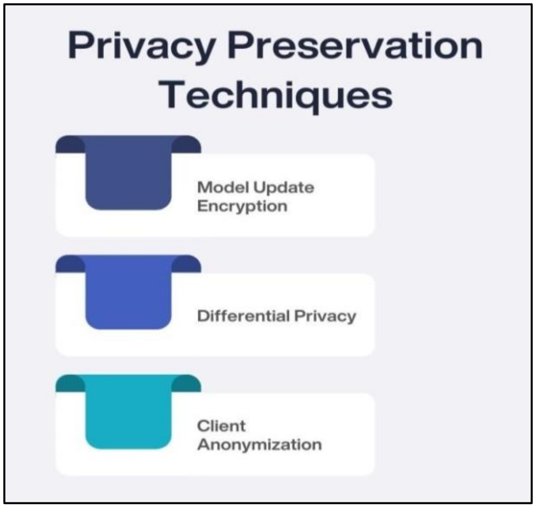

Fig 3: Privacy Preservation Techniques

3.5. Evaluation Metrics

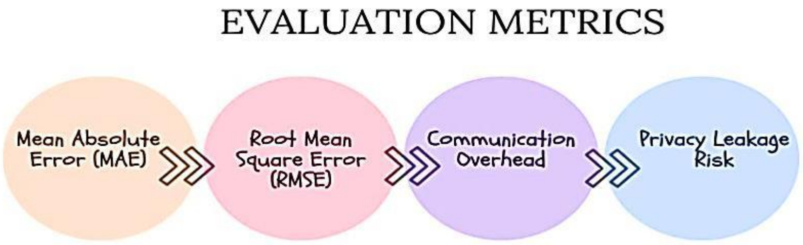

Fig 4: Evaluation Metrics

- **Mean Absolute Error (MAE):** Mean Absolute Error (MAE) is a popular performance measure for forecasting performance, which reports the average absolute error in predicting values when compared to the actual values. It computes the absolute differences without taking direction into account, thus giving an intuitive meaning for deviations from prediction in the same scale as the original observations. In the context of supply chain forecasting, the lower the MAE, the more accurate the predictions of demand that can be translated into better inventory management and smaller costs. MAE's simplicity and interpretability are why it is an indispensable metric for measuring the performance of models in federated learning setups.
- **Root Mean Square Error (RMSE):** Root Mean Square Error (RMSE) has the square root of the average of squared differences between estimated and actual values, and it weighs larger errors more heavily. Such sensitivity to outliers is what makes RMSE especially relevant for the supply chain settings where tremendous forecasting errors may result in severe operational disruptions. RMSE is measured in the same units as the target variable, so that the stakeholders can understand the magnitude of prediction errors. It supplements MAE, penalizing the large deviations more significantly, which enables it to perform a full assessment of forecasting precision in federated learning models.
- **Communication Overhead:** Communication overhead measures the amount of data that is communicated between clients and the central server in the process of federated learning. As FL is based on iterative transmission of the model updates, communication overhead reduction is vital for efficiency, particularly when the partners in the supply chain are working with limited bandwidth over geographically dispersed locations. With the high cost of communication, training might take longer, and there will be an additional cost of operations. The measurement of communication overhead is helpful in optimizing protocols and algorithms to achieve learning performance while respecting real-world deployment limitations in supply chains.

- **Privacy Leakage Risk:** Privacy leakage risk determines the threat of inference or exposure of sensitive client information during federated learning. Even with privacy-preserving mechanisms, model updates may still leak patterns or characteristics of data if unprotected. Determining this risk implies examining such vulnerabilities as inference attacks or data reconstruction from shared updates. In applications to the supply chain, minimizing privacy leakage is important to preserve trust between partners and adhere to the data protection regulations. This metric directs the development and enhancement of the privacy techniques while providing high levels of confidentiality within the partnership forecasting.

4. Results and Discussion

4.1. Dataset Description

To measure the performance of the proposed federated learning framework for supply chain forecasting, we use a publicly accessible demand dataset, which consists of historical sales data for various product categories and different geographical locations. With varying patterns of demand affected by seasonality, promotions, regional tastes, and economic conditions, this dataset contains a diverse and realistic setting for demand prediction assignments. In order to simulate a federated learning setting closer to what the supply chain collaboration in the real world looks like, the dataset is manually divided between five simulated clients. Every client is an individual supply chain partner, i.e., manufacturers, distributors, or retailers, each of them possessing private and localized data that describes its specific functioning context. The partitioning strategy guarantees that the data of any client is characterized differently in terms of volume, volatility of demand, and temporal patterns. For example, one client may work with products having stable, predictable demand, while another client will have a product with highly seasonal or sporadic peaks of sales.

This variation brings about the non-Independent and Identically Distributed (non-IID) data problem, which is a typical occurrence in decentralized supply chains as customers are not identical, products are not similar, and the regions vary. Additionally, the dataset is not balanced since some of the clients have substantially larger datasets compared to others, a case that is representative of the actual world's disparities in data volume and size of supply chain partners. Such a setup allows for a systematic evaluation of the federated LSTM model's abilities to mine meaningful patterns from a wide variety of, as well as unbalanced data, while resorting to the sharing of such data, thus maintaining the privacy and competitive interests of individual partners. The design further enables robustness and adaptability testing of the model under real federated conditions, giving insights into the actuality of collaboration and the advantage of distributed supply chain forecasting.

4.2. Forecasting Accuracy

Two baselines are compared with the forecasting evaluation of the proposed federated LSTM model. The two models will be a Local LSTM model, which will be trained individually for each client's data without cooperation, and a Centralized LSTM model trained on pooled data obtained from all the clients.

Table 1: Forecasting Accuracy

Model	MAE (%)	RMSE (%)
Local LSTM	14.8%	23.5%
Centralized LSTM	11.2%	19.0%
Federated LSTM	11.8%	19.5%

- **Local LSTM Model:** The Local LSTM model is trained individually on each of the private datasets of the client, and there is no data sharing and no collaboration between the supply chain partners. Such practice ensures data privacy but inhibits the model's capability of learning from wider patterns contained in other clients' data. Consequently, the lowest accuracy in predicting is achieved with this technique, as the Mean Absolute Error (MAE) is equal to 14.8%, and the Root Mean Square Error (RMSE) amounts to 23.5%. These increased error rates are a result of the difficulty of depending on limited, locational data that may not represent seasonality or demand patterns that span throughout the entirety of the supply chain network.
- **Centralized LSTM Model:** The Centralized LSTM model is learned using a dataset created by combining data from all clients into one place. With this technique, the model collects data from all partners and looks at their patterns as a whole. Therefore, it performs best in forecasting, as its MAE is 11.2% and its RMSE is 19.0%. However, since AI models are trained on shared raw data, centralized training may not be suitable for supply chains that highly value data privacy and security.
- **Federated LSTM Model:** Federated LSTM combines the local and centralized ideas by supporting shared learning without any direct involvement of data. Individual clients use their data to train the model on their computers, and the updated models are combined at a central location. Following this approach, MAE is 11.8%, RMSE is 19.5%, which is roughly the same as the centralized model and performs much better than the local-only one. I believe the improved privacy and data control of federated learning justifies the slight increase in error. The approach has proved to be effective in ensuring accurate and private demand forecasting along the supply chain.

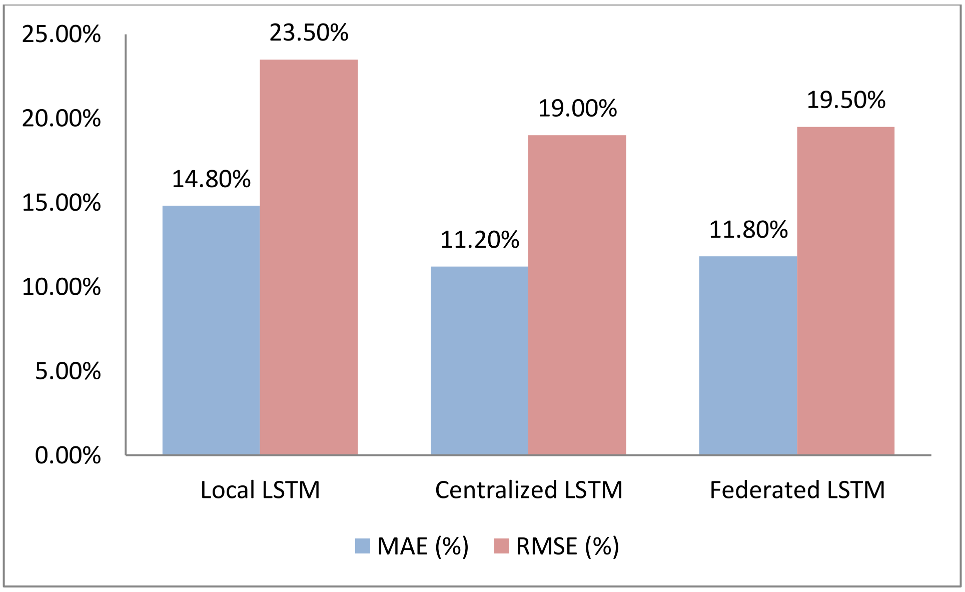

Fig 5: Graph representing Forecasting Accuracy

4.3. Communication Efficiency

No matter where partners are within the supply chain, efficient FL requires effective communication, as not all partners may have the same network resources or computers. With traditional centralized learning, clients need to send all their data to a central server only once, which means the server must process a huge data transfer all at once. Still, this way of learning is efficient, but it may result in serious privacy issues and bottlenecks in communication, given the amount of data moved. On the other hand, federated learning does not exchange raw data, but the server and clients talk to each other often during the training process. Every round, clients are supposed to send their adjusted model weights to the server and get back the unified version of the model. We looked at the total amount of information sent during training to see how centralized and federated learning approaches compare.

We found that the centralized model uploaded around 500 MB of data out of the large client datasets that were first sent to the central server. Still, the federated learning scheme managed to transfer less data, only 150 MB, despite its 150 rounds. Each federated round shares only the gradients and weight tensors, sparing the need to transmit massive volumes of raw data. An important benefit of FL in supply chain settings becomes clear when looking at the reduced amount of data that needs to be transferred. It helps different parties cooperate in learning advanced forecasting, while still managing their data, and lowers the need to communicate. Using compression and limiting how many times the model is updated can lead to faster and simpler communication, helping FL support supply chain applications.

4.4. Privacy Analysis

The main reason to use Federated Learning (FL) for forecasting in the supply chain is to protect important and confidential data such as inventory, sales, and customer levels, which must be kept private. Rather than asking clients to upload their data to a main server, FL allows each device or individual infrastructure to keep its training data. No direct data is exchanged during learning, as only model parameters or their changes are shared between parties. By not having a single central authority, FL fits the privacy standards and rules for sharing information in popular supply chains. Even so, just because the data is not shared, there is still a way that a curious server or an attacker might find personal information by reviewing how the model adjusts. For this reason, differentially private technologies are introduced into the FL process now.

Control noise is deliberately added to the results of training models before passing them to the central gathering point. The interference means it is not possible to figure out what a client purchased, even if many data points are observed. Federated learning is more secure against data breaches or anyone trying to access centralized servers, as centralized models are open to those types of problems. This becomes important in stressful supply chain situations where sharing data can be difficult. Employing data locality and differential privacy, FL ensures it is safe for partners to work together on models that are still accurate and do not endanger their own data or company interests.

4.5. Discussion

The experimental look at the proposed FL framework shows that it can really help deal with the problems of keeping data private and improving forecasts in supply chain systems. While the forecasting performance of the federated LSTM model is slightly less than that of the centralized model, the difference is really small, with just a couple of percentage points both in MAE and RMSE. This small drawback isn’t that big of a deal, because FL helps keep data private a lot more than other ways.

In traditional centralized learning, organizations have to give up their data, as it is gathered and kept in one main place. This model not only comes with privacy issues but also increases the chance that a company might break the law and experience data leaks. By contrast, FL makes sure that data stays on users' local devices, so each business owner keeps control of their private information. Moreover, federated learning helps reduce the amount of unnecessary communication that happens between devices and the server. The centralized approach, as it only needs to upload data once, does take a big load on the network in just the first data transfer and can use around two hundred megabytes or more.

In contrast, the federated process works by having several rounds of communication, but it only sends over small updates to the model, which means much less data is transferred overall. This makes FL work better for supply chains that don't have a reliable internet connection or need to work in areas where broadband is limited, like out-of-the-way warehouses or small businesses. The way people work together in FL also helps build more trust between different parts of the supply chain. By not having to share raw data and still being able to work together to improve forecasts, FL lets organizations collaborate without anyone losing any of their company secrets. These findings show that FL is a simple and easy-to-use way for supply chains in the real world to improve their forecasting, while still keeping their data safe and keeping operations working smoothly.

5. Conclusion

It has emerged that Federated Learning (FL) can help modern supply chains achieve secure collaborative forecasting. As supply chains grow more complex, spread across many locations, and rely on data, the importance of having accurate demand forecasts is at its highest. Even so, centralizing all partner data in one spot makes data privacy, meeting GDPR, and willingly sharing company secrets very tough. Along with these issues, FL tackles them directly by allowing individual supply chain members to join in training a central model without giving up their raw data. It ensures that information is kept secure and also reflects the emerging rules about data ownership and how AI is used ethically.

As shown in this paper, federated LSTM models are able to perform similarly to centralized models in forecasting, having slightly higher error rates. Since the added expense is only small, it is a good compromise for the extra privacy and better efficiency you get. Working together in a federated model, businesses can use each other's data without allowing control to shift beyond their boundaries. Furthermore, the way lightweight model updates are communicated in FL allows the process to run well on networks with a wide range of bandwidth levels and hardware resources. Despite all the good things about these studies, there are still some problems and new things to look into for the future. There is an increasing need to come up with flexible ways to combine client data from different sources, especially since the information often has different structures and amounts, as in many actual supply chains. Further exploration into privacy-enhancing tools like adaptive differential privacy, secure multi-party computation, and homomorphic encryption will make sure we can protect users' privacy better in FL systems. Moreover, testing FL models out in real-world situations, where things like bad data, system errors, or practical limits can show up, is important if we want to move from ideas into actually using these models in real life.

In summary, federated learning opens up a new way to help supply chains work smarter and protect people's privacy when they are trying to forecast. Its ability to work well while keeping performance, data privacy, and the ability to handle a lot of data makes it a good choice for businesses looking to improve their predictions while keeping their data very secure.

References

[1] Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008). Forecasting methods and applications. John wiley & sons.

[2] Armstrong, J. S. (Ed.). (2001). Principles of forecasting: a handbook for researchers and practitioners (Vol. 30). Springer Science & Business Media.

[3] Kourentzes, N., & Petropoulos, F. (2016). Forecasting with multivariate temporal aggregation: The case of promotional modelling. International Journal of Production Economics, 181, 145-153.

[4] Fisher, M., & Raman, A. (1996). Reducing the cost of demand uncertainty through an accurate response to early sales. Operations research, 44(1), 87-99.

[5] McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In Artificial Intelligence and Statistics (pp. 1273-1282). PMLR.

[6] Dwork, C. (2006, July). Differential privacy. In International Colloquium on Automata, Languages, and Programming (pp. 1-12). Berlin, Heidelberg: Springer Berlin Heidelberg.

[7] Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., ... & Seth, K. (2017, October). Practical secure aggregation for privacy-preserving machine learning. In proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (pp. 1175-1191).

[8] Gentry, C. (2009). A fully homomorphic encryption scheme. Stanford University.

[9] Yao, A. C. (1982, November). Protocols for secure computations. In 23rd annual symposium on foundations of computer science (SFCS 1982) (pp. 160-164). IEEE.

[10] Sheller, M. J., Reina, G. A., Edwards, B., Martin, J., & Bakas, S. (2019). Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In Brain Lesions: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with

MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part I 4 (pp. 92-104). Springer International Publishing.

[11] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19.

[12] Ozturk, A., & Polat, H. (2015). From existing trends to future trends in privacy-preserving collaborative filtering. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 5(6), 276-291.

[13] Xie, C. H., Zhong, W. J., Zhang, Y. L., & He, Q. Z. (2007, November). Privacy preserving collaborative forecasting based on dynamic exponential smoothing. In 2007 IEEE International Conference on Grey Systems and Intelligent Services (pp. 730-734). IEEE.

[14] Taigel, F., Tueno, A. K., & Pibernik, R. (2018). Privacy-preserving condition-based forecasting using machine learning. Journal of Business Economics, 88, 563-592.

[15] Shokri, R., & Shmatikov, V. (2015, October). Privacy-preserving deep learning. In Proceedings of the 22nd ACM SIGSAC conference on computer and communications security (pp. 1310-1321).

[16] Hesamifard, E., Takabi, H., Ghasemi, M., & Wright, R. N. (2018). Privacy-preserving machine learning as a service. Proceedings on Privacy Enhancing Technologies.

[17] Xu, K., Yue, H., Guo, L., Guo, Y., & Fang, Y. (2015, June). Privacy-preserving machine learning algorithms for big data systems. In 2015 IEEE 35th International Conference on Distributed Computing Systems (pp. 318-327). IEEE.

[18] Egan, S., Fedorko, W., Lister, A., Pearkes, J., & Gay, C. (2017). Long Short-Term Memory (LSTM) networks with jet constituents for boosted top tagging at the LHC. arXiv preprint arXiv:1711.09059.

[19] Colin, M., Galindo, R., & Hernández, O. (2015). Information and communication technology is a key strategy for efficient supply chain management in manufacturing SMEs. Procedia Computer Science, 55, 833-842.

[20] Dehgani, R., & Jafari Navimipour, N. (2019). The impact of information technology and communication systems on the agility of supply chain management systems. *Kybernetes*, *48*(10), 2217-2236.