Decentralized Smartphone Recommendation with Privacy Preservation

[1] NALLURI VENKATA SARANYA, [2] MANUKONDA LAKSHMI THANMAI, [3] CHILUKOTI NAGA SAI,
[4] NIMMAGADDA CHANDRA SEKHAR
[1,2,3] B. Tech Students, Department of CSE, RVR & JC College of Engineering, Chowdavaram, Guntur, A.P, India.
[4]Assistant Professor, Department of CSE, RVR & JC College of Engineering, Chowdavaram, Guntur, A.P, India,
E-mail:[1] y22cs127@rvrjc.ac.in, [2] l23cs210@rvrjc.ac.in, [3] l23cs203@rvrjc.ac.in, [4] nimmagadda65@gmail.com

---***---

***Abstract*—**With the rapid growth in mobile phone usage, online platforms now contain a huge number of customer reviews. While these reviews are helpful, they can also overwhelm consumers, making it difficult to choose the right product. To address this issue, this study introduces a new approach for classifying mobile phone ratings using a recommendation system based on federated learning and TF-IDF features.For this research, we created a new dataset by collecting more than 13,000 mobile phone reviews from Flipkart. The proposed method uses a Federated Deep Neural Network (FDNN) to classify ratings. The process includes data cleaning, handling imbalanced data, extracting features using TF-IDF, and making predictions through federated learning.The system is designed with two client models and one central server, and experiments were conducted over three rounds. The results show that the model achieved an accuracy of 96.68% at the server level while ensuring that user data remains secure on local devices.This approach can help customers make better purchasing decisions and can also be applied to other e-commerce platforms with large volumes of reviews.

Index Terms—Flipkart smartphone rating, recommendation system, novel dataset, deep learning, federated learning.

I. INTRODUCTION

Recommendation systems have become a fundamental component of modern e-commerce platforms, leveraging data mining and artificial intelligence techniques to analyze user behavior and predict consumer preferences. These systems aim to provide personalized recommendations, thereby improving customer satisfaction, increasing sales, and minimizing cart abandonment rates.In contemporary online marketplaces such as Flipkart and Amazon, customer reviews play a critical role in influencing purchasing decisions. These reviews provide valuable insights into product quality and user experience. Additionally, organizations utilize such feedback to refine their products and services by identifying customer needs and concerns.

Sentiment analysis, a subfield of natural language processing (NLP), is widely used to analyze textual data and determine the polarity of customer opinions. Recent advancements in machine learning and deep learning have significantly improved the ability to extract meaningful insights from large volumes of textual data. Despite these advancements, traditional centralized machine learning approaches raise significant concerns regarding data privacy and security. Federated learning has emerged as an effective alternative, enabling collaborative model training across multiple devices without requiring the sharing of raw data. This decentralized approach ensures that user data remains on local devices while still contributing to the global model. In the context of Flipkart, federated learning provides a promising solution for enhancing the reliability of mobile phone rating and review systems. By leveraging decentralized data, it is possible to improve prediction accuracy while maintaining strict privacy standards. The primary objective of this research is to develop a federated learning-based model capable of accurately predicting mobile phone ratings from user reviews. The proposed approach offers a novel contribution by combining federated learning with TF-IDF feature extraction within a deep learning framework.

This study not only improves rating prediction accuracy but also addresses critical challenges related to data

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 05 | May – 2026 DOI:10.55041/ISJEM06927
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

privacy and security. Moreover, the proposed framework can be generalized to other domains where large-scale user-generated data is involved.

This study makes the following contributions:

- A novel dataset of over 13,000 Flipkart mobile phone reviews is constructed and made publicly available for research purposes.
- A federated deep learning-based approach is proposed for mobile phone rating classification, addressing privacy concerns associated with centralized models.
- TF-IDF is utilized for effective feature extraction from textual reviews.
- The proposed model is evaluated using standard performance metrics, demonstrating high accuracy and robustness.
- The framework is scalable and can be extended to other e-commerce and review-based systems.

II. LITERATURE REVIEW

The existing literature on mobile phone rating and review systems emphasizes the critical role of accurate ratings and user feedback in influencing customer purchasing decisions. With the rapid growth of e-commerce platforms such as Amazon, Flipkart, and Snapdeal, online reviews have become a key factor in shaping consumer behavior and driving sales. Several studies have explored the application of machine learning and deep learning techniques to enhance the accuracy and reliability of these systems [11], [17], [23]. For instance, Biswas et al. [24] utilized a feed-forward Artificial Neural Network (ANN) to analyze how previous customer reviews impact the decisions of potential buyers. Their findings highlight the importance of leveraging user-generated content to understand customer satisfaction. In practice, many successful e-commerce platforms rely heavily on such feedback to evaluate product performance and support decision-making processes, including inventory management and product improvement.

Similarly, Kumar et al. [25] proposed a machine learning-based approach for predicting mobile phone ratings using customer reviews. Their study employed Random Forest (RF) and Support Vector Machine (SVM) classifiers to model rating predictions. The results indicated that their approach achieved higher accuracy compared to several existing methods. In another study, Guo et al. [26] introduced a deep learning framework for sentiment analysis of mobile phone reviews. Their model combined Convolutional Neural Networks (CNN) with Bidirectional Long Short-Term Memory (BiLSTM) networks to capture both spatial and sequential features in text data. Experimental results demonstrated that this hybrid model outperformed traditional approaches in terms of accuracy.

More recently, Federated Learning has gained significant attention as a decentralized machine learning paradigm that enables collaborative model training without sharing raw data. This approach has been successfully applied across various domains, including healthcare, finance, and recommendation systems, due to its strong privacy-preserving capabilities. In the context of mobile phone rating systems, Federated Learning offers a promising solution to address data privacy concerns while still enabling effective analysis of user reviews.

Li et al. [27] proposed a federated deep learning approach for mobile app recommendation systems, where personalized recommendations were generated while maintaining user data privacy. Their results showed that federated learning outperformed traditional centralized approaches in both recommendation accuracy and privacy preservation. Similarly, Zhang et al. [28] applied federated learning to hotel recommendation systems using a federated matrix factorization model. Their findings also confirmed improvements in both accuracy and data security compared to centralized models.

Compared to conventional centralized machine learning approaches, Federated Learning offers several key advantages. First, it significantly reduces the risk of data breaches, as sensitive user data remains on local devices rather than being transferred to a central server. Second, it enables platforms like Flipkart to leverage distributed user data collectively, thereby improving the overall accuracy of rating and recommendation systems. Third, it reduces computational overhead and training time by distributing the learning process across multiple clients.

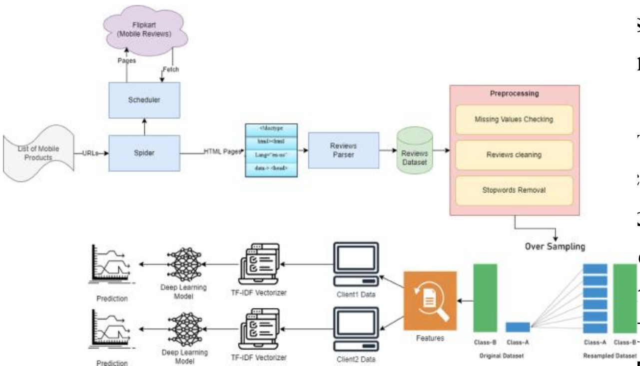
In summary, the proposed federated learning-based approach provides a more secure and efficient solution for mobile phone rating classification. It not only ensures data privacy and confidentiality but also achieves high prediction accuracy. Furthermore, this approach is highly scalable and can be extended to other domains where privacy-preserving data analysis is essential, making it a valuable contribution to the field of machine learning and recommendation systems.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 05 | May – 2026 DOI:10.55041/ISJEM06927
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

III. PROPOSED APPROACH

The proposed framework consists of several sequential stages, including dataset creation, data cleaning and preprocessing, data balancing, feature extraction, and model prediction using a federated learning approach. An overview of the complete methodology is illustrated in Fig. 1. The primary objective is to develop a predictive model capable of accurately estimating mobile phone ratings on Flipkart based on user reviews.

In the initial stage, data is collected from Flipkart using a web crawler. The crawler is designed to extract essential information such as product name, brand, price, rating, and customer reviews. However, the raw data obtained from web scraping often contains inconsistencies, missing values, and duplicate entries. Therefore, a data cleaning and preprocessing step is performed to ensure data quality. This includes removing duplicate records, correcting inconsistencies, handling missing values, and converting the data into a standardized format suitable for analysis.



s are real- be hich ome class d to ta is erm Frequency–Inverse Document Frequency (TF-IDF) method. This technique transforms textual review data into numerical feature vectors by assigning weights to words based on their importance. As a result, TF-IDF helps in identifying the most relevant keywords in customer reviews that contribute to predicting product ratings.

In the final stage, the extracted TF-IDF features are used to train a machine learning model within a federated learning framework. Unlike traditional centralized approaches, federated learning enables model training across multiple client devices without transferring raw data to a central server. This approach ensures that sensitive user data remains on local devices while still contributing to the global model. In this research, federated learning is employed to collaboratively train the model using distributed datasets. Each client device trains a local model using its own data and shares only the model updates (such as weights or gradients) with a central server. The server then aggregates these updates to form a global model, which is redistributed to the clients for further training. This iterative process continues until the model achieves satisfactory performance. The use of federated learning offers significant advantages in terms of privacy and security. Since raw data is not shared, the risk of data breaches is minimized, and user confidentiality is preserved. This is particularly important in the context of analyzing customer reviews, which may contain sensitive or personal information. By adopting this decentralized learning approach, the proposed system ensures compliance with privacy requirements while maintaining high predictive accuracy.

Overall, the proposed methodology effectively integrates data collection, preprocessing, feature extraction, and federated learning to build a robust and privacy-preserving model for mobile phone rating prediction. Furthermore, this approach is scalable and can be extended to other e-commerce platforms and applications where data privacy is a critical concern.

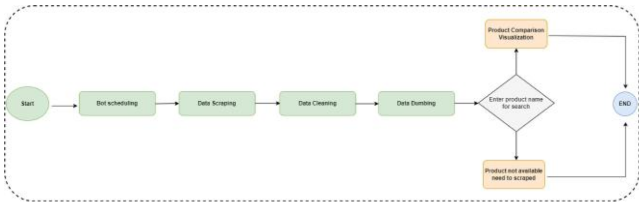


Fig. 1. Proposed approach for smartphone rating classification.

A. Dataset Selection

Flipkart provides a large collection of customer ratings and reviews for various mobile phones. These user-generated reviews serve as a valuable resource for training predictive models that can estimate how customers rate different mobile devices. In this study, a dedicated dataset was developed to enable detailed analysis and accurate prediction of user opinions on mobile phones. The dataset was created by scraping data from the website Flipkart.com using a web crawler. The crawling process was carried out over a period of ten months, from May 2022 to March 2023, resulting in the collection of more than 13,000 mobile phone reviews. A scheduling mechanism was implemented to ensure that the crawling process remained efficient, adhered to website policies, and adapted to any structural changes in

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 05 | May – 2026 DOI:10.55041/ISJEM06927
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

the website over time. The overall data crawling architecture is illustrated in Fig. 2. The web crawler (or spider) was specifically designed to extract relevant product information, including product name, brand, price, rating, and customer reviews. After processing, the final dataset consisted of 13,589 records with 9 distinct features, as detailed in Table I.

The distribution of products in the dataset, as shown in Fig. 3, indicates that the "vivo T1 44W (Starry Sky, 128 GB)" model appears most frequently, suggesting higher representation in the collected data. In contrast, the "POCO C31 (Shadow Gray, 32 GB)" model appears the least frequently. This variation reflects the relative availability or popularity of these devices within the collected reviews.

A new dataset was created in this study because existing datasets did not provide the required level of detail, coverage, or timeliness for device-specific analysis. By directly scraping data from Flipkart, the study ensured that the collected information was current, relevant, and reflective of real user experiences. Additionally, constructing a custom dataset allowed greater control over the data structure and inclusion of specific attributes necessary for the research.

The dataset was further analyzed to understand user sentiment and rating patterns. It was observed that higher ratings, particularly 4 and 5, occurred more frequently, indicating generally positive customer feedback for many products. This analysis highlights the usefulness of the dataset in capturing customer satisfaction trends and supporting predictive modeling tasks.

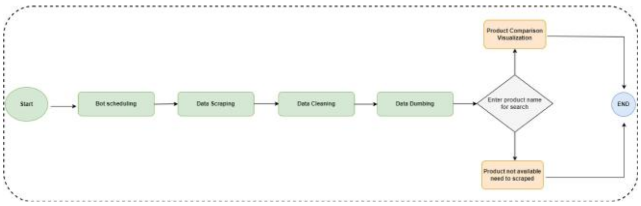


Fig. 2. Data crawling architecture.

Overall, the development of this dataset played a crucial role in achieving the objectives of the study, enabling comprehensive analysis and accurate prediction of customer attitudes toward mobile devices.

TABLE I

DATA FEATURES DESCRIPTION

Attribute	Description
product_id	A unique product id
product_title	Mobiles name on which customer gives the reviews
rating	Ratings given by the customer
summary	Reviews summary or details are given by the customer
review	Reviews given by the customer
location	Location of customer
date	Date on which customer gives the review
upvotes	Upvotes given by the customer
downvotes	Down votes given by the customer

B. ***Proposed Workflow Algorithm***

The proposed approach for mobile phone rating prediction on Flipkart is summarized in *Algorithm 1*. The algorithm takes Flipkart review data as input and produces predicted mobile phone ratings as output. It is organized into three main functions: CreateDataset, TrainModel, and PredictRatings, which are integrated within a Main function to form the complete prediction pipeline. The CreateDataset function is responsible for data collection and preprocessing. It employs a web crawler (web spider) to extract relevant product information from Flipkart, including product name, brand, price, ratings, and customer reviews. The Web Spider Process (WSP) represents the aggregation of all extracted information. Since raw data may contain inconsistencies, this function performs preprocessing steps such as removing duplicate entries, correcting errors, and converting the data into a standardized format.

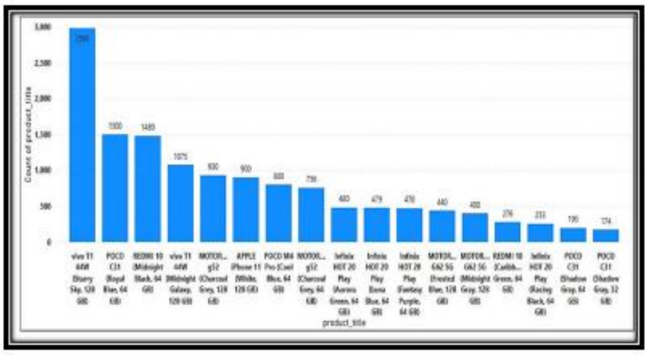


Fig. 3. **Distribution of products in the dataset**.

To ensure consistency across features, the data is normalized using a min–max scaling technique, which transforms values into a range between 0 and 1. This step helps prevent features with larger numerical ranges from dominating the model. Additionally, data balancing techniques are applied to address class imbalance by either oversampling minority classes or undersampling majority classes. Textual data, particularly customer reviews, are transformed into numerical representations using the TF-IDF vectorization technique. In this context, TF-IDF assigns weights to terms based on their

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 05 | May – 2026 DOI:10.55041/ISJEM06927
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

importance within a document and across the dataset. Specifically, term frequency (TF) measures how often a word appears in a document, while inverse document frequency (IDF) evaluates how rare or significant that word is across all documents. This transformation enables the deep learning model to effectively process textual inputs. Finally, the dataset is partitioned into two clients to support the federated learning framework. The function returns the processed feature set XXX and corresponding labels yyy. The TrainModel function is responsible for training the deep learning model using the preprocessed data. The model architecture begins with an input layer of size 50 and an embedding layer with a dimension of 20. This is followed by a Bidirectional LSTM layer with 70 units (configured to return sequences), and an LSTM layer with 100 units to capture sequential dependencies in the data. A dense output layer with 5 units and softmax activation is used for multi-class classification of ratings.

The model is compiled using the sparse categorical cross-entropy loss function and the Adam optimizer. Training is performed with a batch size of 200, a validation split of 20%, and over 5 epochs. Upon completion, the trained model is returned. The PredictRatings function uses the trained model to predict ratings for new mobile phones based on their processed review data. It takes the trained model and feature set XXX as input and generates predicted rating outputs. Finally, the Main function integrates all components into a unified system. It begins by collecting data from Flipkart, then invokes the CreateDataset function for preprocessing and client partitioning. The TrainModel function is subsequently called to train the federated deep learning model. Lastly, the PredictRatings function is used to generate predictions for new mobile phone ratings. The system outputs the predicted ratings, completing the overall prediction workflow.

C. *Data Pre-Processing*

The preprocessing stage plays a crucial role in improving data quality and ensuring that the dataset is suitable for model training. In this study, missing values were first removed from the *location* and *summary* attributes, specifically eliminating 95 and 11 missing entries, respectively, to ensure dataset completeness. Following this, several text preprocessing techniques were applied to clean and standardize the textual data. Punctuation marks such as commas, periods, and question marks were removed to reduce noise and lower the dimensionality of the dataset. Additionally, characters with undefined or non-informative representations were eliminated, as they do not contribute meaningful information.

To simplify the dataset structure, the *Review* and *Summary* columns were combined into a single textual feature. This consolidation helps in capturing the complete context of user feedback while reducing redundancy. Furthermore, contractions were expanded into their full forms (e.g., "won't" to "will not," "can't" to "cannot") to ensure consistency in textual representation. Stop words, including commonly used terms such as "no," "nor," and not," were removed to focus on meaningful words that contribute to sentiment and rating prediction. Additionally, special patterns such as "//" were replaced with whitespace to eliminate artifacts introduced during the web scraping process.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

To standardize numerical features, a normalization technique was applied using the Min-Max scaling method. This approach transforms feature values into a range between 0 and 1, ensuring that all features contribute equally to the model and preventing bias caused by varying scales. These preprocessing steps collectively enhance data quality, reduce noise, and prepare the dataset for effective feature extraction and model training.

Algorithm 1 Proposed Approach for Mobile Rating Prediction on Flipkart With Additional Equations

Require: *Flipkart data*
Ensure: *Predicted mobile phone ratings*
function CREATEDATASET(*Flipkart data*)
  Retrieve relevant information: $WSP = \sum_{i=1}^{n} P_i$
  Perform data preparation
  $X_{undersampled} = X_{majority} \cup sample(X_{minority})$
  $X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}$
  $TF - IDF(t, d) = TF(t, d) \times IDF(t)$
  return $X,\ y$
function TRAINMODEL($X,\ y$)
  Start model training
  Set input shape of (50)
  Add a 70-unit bidirectional LSTM layer and a 100-unit LSTM layer
  Add 5 units, dense layer, and softmax activation
  Use oversampling during training: $X_{oversampled} = X_{majority} \cup resample(X_{minority})$
  Early stopping to prevent overfitting: $\frac{\partial L}{\partial w} < \epsilon$
  return *model*
function PREDICTRATINGS(*model*, X)
  Start model prediction
  return $y_{predicted}$
procedure MAIN
  Flipkart data ← Collect mobile phone data from Flipkart's website
  $X,\ y \leftarrow$ CreateDataset(*Flipkart data*)
  model ← TrainModel($X,\ y$)
  $y_{predicted} \leftarrow$ PredictRatings(*model*, X)
  Calculate the results of the model
  return $y_{predicted},\ accuracy$

Data Balancing:

The collected dataset exhibited class imbalance, where certain rating categories were more frequent than others. To address this issue, an upsampling technique was applied to the minority classes. This approach involves generating additional samples by introducing slight variations to existing data points, thereby increasing the representation of underrepresented classes. Balancing the dataset is essential to prevent the model from becoming biased toward majority classes. By ensuring a more uniform class distribution, the model can achieve better generalization and improved prediction performance across all rating categories.

D. *Feature Extraction*

Feature extraction was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) technique, a widely used method in natural language processing for evaluating the importance of words in textual data. TF-IDF assigns higher weights to words that frequently appear in a specific document but are relatively rare across the entire corpus, making them more informative for classification tasks. In this study, customer reviews were treated as individual documents. The TF-IDF vectorizer was applied to transform textual data into numerical feature vectors, capturing the most significant terms in each review. These feature vectors represent the input to the machine learning model.

The TF-IDF score for a word is computed as the product of term frequency (TF) and inverse document frequency (IDF). Here, TF measures how often a word appears in a document, while IDF evaluates its rarity across the dataset. Words with higher TF-IDF scores are considered more relevant for distinguishing between different rating classes

$$Tf\text{-}IDF(w,d) = TF(w,d) * IDF(w) \qquad (2)$$

where $TF(w, d)$ is the term frequency of word w in docu- ment d, and $IDF(w)$ is the inverse document frequency of word w in the corpus. The IDF is computed in equation 3

$$IDF(w) = \log \frac{N}{n_w} \qquad (3)$$

E. *Federated Learning Modeling*

Federated Learning (FL) is a distributed deep learning paradigm that enables multiple clients to collaboratively train a shared model without sharing their raw data, thereby preserving data privacy. In the proposed approach, two clients participate in the training process, each utilizing TF-IDF vectorization to extract meaningful features from their local datasets. Although federated learning systems can involve a larger number of clients during preprocessing, it is not strictly necessary. The number of participating clients may vary depending on the system architecture, available resources, and specific application requirements. In this work, the architecture is limited to two clients due to computational constraints and resource limitations.

The training process begins with feature extraction at each client using the TF-IDF vectorizer. These features are then used as input to a deep learning model designed with a sequential architecture, comprising an input layer, an embedding layer, two LSTM layers, and a dense output layer with a softmax activation function. Initially, each client processes its local data and contributes to the training by interacting with a central server. The server aggregates the information received from the clients and updates the global model weights accordingly. The updated model is then redistributed to the clients, allowing them to continue training on their respective local datasets. This iterative process of local training and global aggregation continues until the model converges.

The model is compiled using the sparse categorical cross-entropy loss function, the Adam optimizer, and accuracy as the performance metric. Training is conducted for five epochs with a batch size of 200, and a validation split of 20% is employed to monitor model performance during training. This federated learning framework enables effective collaboration among clients while ensuring the privacy of their local data. The final trained model is subsequently used to predict ratings for new product reviews.

F. *Evaluation Metrics*

In this study, the performance of the proposed model is evaluated using standard metrics such as accuracy, precision, recall, F1-score, and confusion matrix. These metrics collectively provide a comprehensive assessment of classification performance. Accuracy measures the proportion of correctly classified instances and is defined as:
Accuracy = (TP + TN) / (TP + FP + TN + FN) Precision evaluates the correctness of positive predictions: Precision = TP / (TP + FP)

Recall measures the model's ability to identify actual positive instances: Recall = TP / (TP + FN) The F1-score represents the harmonic mean of precision and recall: F1-score = 2 × (Precision × Recall) / (Precision + Recall) In addition, the confusion matrix provides a detailed summary of classification outcomes, consisting of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The diagonal elements represent correctly classified instances, while the off-diagonal elements indicate misclassifications. These evaluation measures help in identifying the strengths and limitations of the model, thereby facilitating further performance improvement.

IV. Experimental Analysis and Results

This section presents the experimental results along with a detailed discussion of the findings. The dataset was divided into three parts: 50% was used for training and development, 20% for validation, and the remaining 30% for testing. The deep learning model was designed using an embedding layer, followed by a bidirectional LSTM layer and a dense output layer. The model was compiled using the sparse categorical cross-entropy loss function and optimized with the Adam optimizer to achieve effective learning. To evaluate the performance of the model, several metrics were employed, including accuracy, precision, recall, F1-score, and the confusion matrix, providing a comprehensive analysis of its predictive capability.

A. Server-Side Training and Data Logging

The federated learning (FL) system consists of a central parameter server and two participating clients. During the training phase of the federated deep neural network (FDNN), the server is responsible for coordinating the learning process by maintaining communication with both clients and collecting their model updates. The server aggregates the updates received from the clients to refine the global model iteratively. Additionally, it records relevant training information in the form of data logs, which capture details such as client updates, model performance metrics, and communication rounds. These logs play a crucial role in monitoring the training process and analyzing system performance.

TABLE II
Experimental Results of Client 1 Using FDNN Model

Experiments	Accuracy%	Precision%	Recall%	F1-Score%
Round 1	95.10	95.02	95.02	95.01
Round 2	96.02	96.05	96.01	96.03
Round 3	**96.36**	**96.10**	**96.03**	**96.02**

server-side FDNN model training process. Logs are records that can be retrieved from a server and used to identify a specific process. Our trials are evaluated thrice, once for each of the three rounds specified by the servers for each client. In the first three minutes after the FL system has started, we perform the FDNN evaluation. Each iteration consists of two phases: the fitting and evaluating phases. There is a fitting phase when the client transmits training results to the server and an evaluation phase where both clients transmit testing results to the server. After that, we add up all the data. Using the procedures mentioned above, we finished the experiment in 26 minutes. When the server aggregates data from all clients, the FDNN model achieves an accuracy of 96.68%. The results demonstrate the efficacy of the FDNN model on the Flipkart dataset.

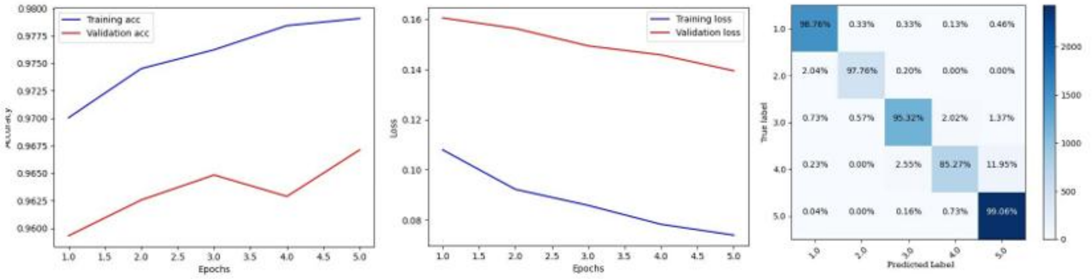
B. Experimental Results of Client 1

Table II presents the experimental results obtained for Client 1 using the FDNN model. The model's performance was evaluated using standard metrics, including accuracy, precision, recall, and F1-score. The results indicate a consistent improvement across training rounds. In Round 1, the model achieved an accuracy of 95.10%, with precision, recall, and F1-score values of 95.02%, 95.02%, and 95.01%, respectively. In Round 2, the performance improved further, with accuracy increasing to 96.02%, and precision, recall, and F1-score reaching 96.05%, 96.01%, and 96.03%, respectively. The best performance was observed in Round 3, where the model achieved an accuracy of 96.36%, along with precision, recall, and F1-score values of 96.10%, 96.03%, and 96.02%, respectively. These results demonstrate a steady enhancement in model performance over successive training rounds.

In addition to numerical evaluation, graphical analyses were conducted to better understand the model's behavior. Figure 4 illustrates these results, where Figure 4(a) presents the training and validation accuracy, and Figure 4(b) shows the corresponding loss curves. The model achieved a high training accuracy of 97.75% and a validation accuracy of 96.30%, indicating strong predictive capability on both seen and unseen data. The small gap between training and validation accuracy suggests that the model generalizes well and does not suffer significantly from overfitting. This observation is further supported by the low loss values, with training loss recorded at 0.09% and validation loss at 0.15%, reflecting stable and effective learning. Figure 4(c) presents the confusion matrix for Round 3, which shows that the majority of instances in the testing set were correctly classified.

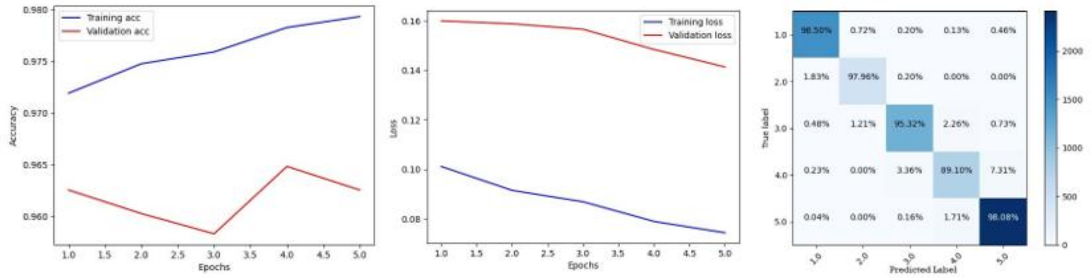
International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 05 | May – 2026 DOI:10.55041/ISJEM06927
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Overall, the results demonstrate that the model's performance improves progressively with each communication round. This highlights the effectiveness of the federated learning approach in enabling collaborative model training across distributed datasets while preserving data privacy.



(a) Training and validation accuracy graph plot (b) Training and validation loss graph plot (c) Confusion matrix of FDNN model

Fig. 4. Graphical Visualization of the Client 1 in Round 3 results.



(a) Training and validation accuracy graph plot (b) Training and validation loss graph plot (c) Confusion matrix of FDNN model

Fig. 5. Graphical visualization of the Client 2 in Round 3 results

Table III summarizes the performance of Client 2 using the FDNN model over three training rounds, evaluated using accuracy, precision, recall, and F1-score. The results show a clear improvement across rounds. In Round 1, the model achieved an accuracy of 94.02%, with precision, recall, and F1-score values close to this level, indicating good initial performance. In Round 2, the model improved significantly, reaching an accuracy of 96.05%, with all other metrics around 96%. The best performance was observed in Round 3, where the model achieved an accuracy of 96.73%, along with similarly high precision, recall, and F1-score values.

The consistently close values of precision and recall across all rounds indicate balanced model performance without class bias. The steady improvement in metrics demonstrates the model's ability to learn effectively over time. Although Client 2 achieved strong results, the overall best performance was recorded by Client 1 in Round 3. Additionally, graphical analyses such as the confusion matrix, accuracy, and loss curves were generated for Client 2 in Round 3 to further illustrate model performance. Figure 5 illustrates the performance of the proposed model through graphical analysis. Figure 5(a) shows the training and validation accuracy, while Figure 5(b) presents the corresponding loss curves. The model achieved high accuracy on both datasets, with 97.60% training accuracy and 96.40% validation accuracy. The small difference between these values indicates good generalization and a well-balanced model.

The training loss (0.09%) and validation loss (0.15%) are both low, suggesting that the model effectively minimized prediction errors on both training and validation data. Figure 5(c) presents the confusion matrix for Client 2 in Round 3. The presence of high values along the diagonal and low values elsewhere indicates that the model correctly classified most instances, with minimal misclassifications.

V. **Conclusion**

This study proposed a mobile rating classification approach for Flipkart using a federated deep learning framework. A new dataset was constructed by extracting data from the Flipkart website using a web scraping technique, followed by preprocessing steps such as duplicate removal, error correction, and data standardization. The textual data were transformed into numerical representations using the TF-IDF vectorizer, and a Federated Deep Neural Network (FDNN) model was employed for classification. The use of federated learning enabled collaborative model training across distributed clients while preserving data privacy. Experimental results demonstrated that the proposed approach achieved high classification performance and effectively predicted mobile phone ratings.

However, the current approach may require further optimization to handle larger datasets and diverse product categories. Future work will focus on exploring advanced Natural Language Processing (NLP) techniques, such as word embeddings and topic modeling, to improve feature representation. Additionally, incorporating advanced federated learning mechanisms, including differential privacy and secure aggregation, could further enhance data security and model robustness.

ACKNOWLEDGMENT

The authors would like to thank Prince Sultan University for their support.

REFERENCES

[1] Y. Himeur et al., "Blockchain-based recommender systems: Applications, challenges and future opportunities," *Comput. Sci. Rev.*, vol. 43, Feb. 2022, Art. no. 100439.

[2] M. Arafeh, P. Ceravolo, A. Mourad, E. Damiani, and E. Bellini, "Ontology based recommender system using social network data," *Future Gener. Comput. Syst.*, vol. 115, pp. 769–779, Feb. 2021.

[3] M. F. Bashir, A. R. Javed, M. U. Arshad, T. R. Gadekallu, W. Shahzad, and M. O. Beg, "Context aware emotion detection from low resource urdu language using deep neural network," *Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, pp. 1–30, May 2023.

[4] S. Lee and D. Kim, "Deep learning based recommender system using cross convolutional filters," *Inf. Sci.*, vol. 592, pp. 112–122, May 2022.

[5] P. O'Connor and G. Assaker, "Examining the antecedents and effects of hotel corporate reputation on customers' loyalty and citizenship behavior: An integrated framework," *J. Hospitality Marketing Manage.*, vol. 31, no. 5, pp. 640–661, Jan. 2022.

[6] N. N. Chaker, E. L. Nowlin, M. T. Pivonka, O. S. Itani, and R. Agnihotri, "Inside sales social media use and its strategic implications for salesperson-customer digital engagement and performance," *Ind. Market. Manage.*, vol. 100, pp. 127–144, Jan. 2022.

[7] I. Abosag, Z. B. Ramadan, T. Baker, and Z. Jin, "Customers' need for uniqueness theory versus brand congruence theory: The impact on sat- isfaction with social network sites," *J. Bus. Res.*, vol. 117, pp. 862–872, Sep. 2020.

[8] S. Ding, F. Feng, X. He, Y. Liao, J. Shi, and Y. Zhang, "Causal incre- mental graph convolution for recommender system retraining," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 16, 2022, doi: 10.1109/TNNLS.2022.3156066.

[9] Z. B. Ramadan, I. Abosag, and V. Zabkar, "All in the value: The impact of brand and social network relationships on the perceived value of cus- tomer endorsed Facebook advertising," *Eur. J. Market.*, vol. 52, nos. 7–8,
pp. 1704–1726, 2018.

[10] O. S. Itani, M. T. Krush, R. Agnihotri, and K. J. Trainor, "Social media and customer relationship management technologies: Influencing buyer-seller information exchanges," *Ind. Market. Manage.*, vol. 90, pp. 264–275, Oct. 2020.

[11] N. Sharm, T. Jain, S. S. Narayan, and A. C. Kandakar, "Sentiment anal- ysis of Amazon smartphone reviews using machine learning & deep learning," in *Proc. IEEE Int. Conf. Data Sci. Inf. Syst. (ICDSIS)*, 2022,
pp. 1–4.

[12] A. R. Javed et al., "Future smart cities requirements, emerging tech- nologies, applications, challenges, and future aspects," *Cities*, vol. 129, Oct. 2022, Art. no. 103794.

[13] M. U. Khan, A. R. Javed, M. Ihsan, and U. Tariq, "A novel category detection of social media reviews in the restaurant industry," *Multimedia Syst.*, vol. 29, pp. 1825–1838, Oct. 2020.

[14] O. S. Itani, A.-N. Kassar, and S. M. C. Loureiro, "Value get, value give: The relationships among perceived value, relationship quality, customer engagement, and value consciousness," *Int. J. Hospitality Manage.*, vol. 80, pp. 78–90, Jul. 2019.

[15] O. S. Itani, F. Jaramillo, and B. Paesbrugghe, "Between a rock and a hard place: Seizing the opportunity of demanding customers by means of frontline service behaviors," *J. Retailing Consum. Services*, vol. 53, Mar. 2020, Art. no. 101978.

[16] E. Martis, R. Deo, S. Rastogi, K. Chhaparia, and A. Biwalkar, "A proposed system for understanding the consumer opinion of a prod- uct using sentiment analysis," in *Proc. Sentiment Anal. Deep Learn.*, 2023, pp. 555–568.

[17] D. Shankar, S. Narumanchi, H. Ananya, P. Kompalli, and K. Chaudhury, "Deep learning based large scale visual recommendation and search for e-commerce," 2017, *arXiv:1703.02344*.

[18] J. Jin et al., "An agent-based traffic recommendation system: Revisiting and revising urban traffic management strategies," *IEEE Trans.*

Syst., Man, Cybern., Syst., vol. 52, no. 11, pp. 7289–7301, Nov. 2022.

[19] P. Rakshit, P. K. Srivastava, M. Afjal, and S. K. Srivastava, "Sentimental analytics on Indian big billion day of Flip Kart and Amazon," *SN Comput. Sci.*, vol. 2, pp. 1–8, Apr. 2021.

[20] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. Artif. Intell. Stat.*, 2017, pp. 1273–1282.

[21] P. Kairouz et al., "Advances and open problems in federated learn- ing," *Found. Trends® Mach. Learn.*, vol. 14, nos. 1–2, pp. 1–210, 2021.

[22] S. B. Atitallah, M. Driss, and H. B. Ghezala, "FedMicro-IDA: A federated learning and microservices-based framework for IoT data analytics," *Internet Things*, vol. 23, Oct. 2023, Art. no. 100845.

[23] Z. Singla, S. Randhawa, and S. Jain, "Sentiment analysis of customer product reviews using machine learning," in *Proc. Int. Conf. Intell. Comput. Control (I2C2)*, 2017, pp. 1–5.

[24] B. Biswas, M. K. Sanyal, and T. Mukherjee, "Feedback analysis for digital marketing in India: Empirical study on Amazon.in, Flipkart, and Snapdeal," *Int. J. Online Market.*, vol. 11, no. 1, pp. 78–88, 2021.

[25] V. Kumar, A. Pandey, and R. Srivastava, "Prediction of mobile phone ratings using machine learning techniques," in *Proc. 11th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, 2020, pp. 1–5.

[26] H. Guo, Y. Zhang, C. Yu, and H. Lin, "A deep learning approach to sentiment analysis of mobile phone reviews," *IEEE Access*, vol. 8, pp. 223357–223367, 2020.

[27] Y. Li, X. Chen, Y. Li, H. Li, J. Li, and Y. Li, "Federated learning for mobile app recommendation systems," in *Proc. 3rd Int. Conf. Comput. Sci. Appl. Eng.*, 2020, pp. 181–185.

[28] X. Zhang, X. Guo, and H. Liu, "Federated learning for hotel recom- mendation system," in *Proc. 2nd Int. Conf. Artif. Intell. Mach. Learn.*, 2020, pp. 355–360.

[29] A. L. M. Feng, L. J. Hong, S. Y. Ying, and A. H. Omar, "Opinion analysis based on TNF (textual noise fixing) algorithm," *Multidiscipl. Appl. Res. Innov.*, vol. 3, no. 2, pp. 84–91, 2022.

[30] M. U. Rehman et al., "A novel chaos-based privacy-preserving deep learning model for cancer diagnosis," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 6, pp. 4322–4337, Nov./Dec. 2022.

[31] W. Boulila, M. K. Khlifi, A. Ammar, A. Koubaa, B. Benjdira, and I. R. Farah, "A hybrid privacy-preserving deep learning approach for object classification in very high-resolution satellite images," *Remote Sens.*, vol. 14, no. 18, p. 4631, 2022.