International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Advanced Customer Churn Prediction using CNN-LSTM Hybrid Model and Ensemble Voting Mechanism

K. Pavani[1], Ch. Satyanarayana Reddy[2], P. Mary Kamala[3]
#1 Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada.

#2 Assistant Professor in the Department of MCA, SRK Institute of Technology, Vijayawada.

#3 Student in the Department of MCA, SRK Institute of Technology, Vijayawada.

Abstract: In the telecom sector, where keeping current clients is more economical than finding new ones, customer churn prediction is crucial. In order to efficiently capture both spatial feature patterns and temporal relationships in consumer behavior data, this research suggests an advanced hybrid architecture that integrates Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks. To improve prediction accuracy and resilience, an ensemble Voting Classifier that integrates Boosted Decision Trees and Extra Trees is also used. SMOTE-based data balancing strategies are used to solve the problem of class imbalance, guaranteeing equitable and dependable model performance. Additionally, a Flask-based web application is created to offer a user-friendly interface for safe access and real-time churn prediction. In terms of accuracy, generalization, and stability, experimental assessment shows that the suggested hybrid technique performs noticeably better than standalone deep learning models and conventional machine learning.

***Index terms* - —** *Customer Churn Prediction, CNN-LSTM, Ensemble Learning, Voting Classifier, SMOTE, Class Imbalance, Deep Learning, Telecommunication Industry, Flask*

1. INTRODUCTION

Over the past few decades, the telecommunications sector has grown rapidly, creating fierce rivalry among service providers. Customer retention has become a significant difficulty in such a competitive climate since customers may quickly move between service providers owing to alluring offers and better services. The loss of current customers, or customer churn, has a direct effect on telecom firms' earnings and profitability. Therefore, creating successful retention tactics requires an accurate prediction of client attrition.

For churn prediction, conventional machine learning methods like Support Vector Machines, Random Forest, and Logistic Regression have been extensively employed. However, these techniques frequently fail to capture the temporal patterns and intricate, non-linear correlations seen in consumer data. More complex models have been developed as deep learning has progressed, but many of them are still unable to efficiently learn spatial and sequential relationships at the same time.

This research suggests a hybrid deep learning and ensemble architecture for customer churn prediction in order to overcome these difficulties. The suggested method combines Long Short-Term Memory (LSTM) networks to capture temporal relationships in consumer behavior with Convolutional Neural Networks (CNN) for feature extraction. Additionally, an ensemble Voting Classifier that combines Extra Trees and Boosted Decision Trees is used to increase prediction resilience and accuracy.

Additionally, SMOTE-based data balancing approaches are used to address the problem of class imbalance, which is prevalent in churn datasets. A Flask-based web application is created to improve usability, enabling users to anticipate churn in real time via an interactive interface. In comparison to current methods, the suggested system seeks to provide more accuracy, improved generalization, and useful deployment capabilities.

2. LITERATURE SURVEY

a) Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector::

Customer churn prediction (CCP) is recognized by many organizations as an essential process for retaining customers, and customer retention is a major issue in many business fields. CCP has grown to be an essential need in the telecom sector as the number of service providers has expanded. Recently, machine learning (ML) and deep learning (DL) approaches have been used to construct effective CCP models. The ISMOTE-OWELM model for CCP, a novel improved synthetic minority over-sampling technique (SMOTE) with optimal weighted extreme machine learning (OWELM), is presented in this paper. The model that is being described consists of preprocessing, classification, and balancing the unequal dataset. The multi-objective rain optimization algorithm (MOROA) is used for both WELM parameter modification and the optimal sample rate of SMOTE. The client data is first subjected to class labeling and data normalization. ISMOTE is then used to manage the imbalanced dataset, and the rain optimization algorithm (ROA) is used to determine the optimal sample rate. Lastly, the WELM model is used to establish the class labels of the applied data. A lot of testing is done in order to validate the ISMOTE-OWELM model against the CCP Telecommunication dataset. The simulation results showed that the ISMOTE-OWELM model outperformed other models, with accuracy values of 0.94, 0.92, and 0.909 on the applied datasets I, II, and III, respectively.

b) New Approach to Telecom Churn Prediction Based on Transformers:

It costs telecom firms more to retain current customers than to acquire new ones. Churn prediction has therefore emerged as a critical issue for leading service providers worldwide. As a result, large quantities of money are increasingly being spent on developing innovative churn-fighting strategies. Machine

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

learning is one of the current initiatives in this field. We have created a new technique for precisely forecasting churn in this study. The foundation of our concept is the conversion of data into radar chart pictures, followed by the use of a transformer architecture for categorization. A satisfactory accuracy score of 81% was attained by our suggested model.

c) Data Visualization and Prediction for Telecom Customer Churn:

Due to the telecom industry's continuous reform and heightened competition, the customer turnover rate of telecom companies is continuously increasing. The capacity to predict and effectively reduce customer attrition is closely linked to the survival and expansion of telecom companies. In order to effectively deal with unbalanced classification and improve the accuracy of high-value customer churn prediction in the telecom industry, this paper uses a telecom customer data set from the Kaggle platform to analyze people's use of telecom services, assist telecom operators in identifying the reasons behind customer churn, and develop a churn prediction model to lower customer churn rate. In this study, the data set is imported before the data visualization analysis is carried out. The GBDT, SVM, and random forest models are then displayed for comparison. Experiments show that random forest improves the accuracy of high-value customer churn prediction and performs better than other classification methods.

d) Arithmetic Optimization With Ensemble Deep Learning SBLSTM-RNN-IGSA Model for Customer Churn Prediction:

Businesses in a variety of industries utilize the customer churn prediction (CCP) approach to keep their current customers. Insurance companies need to be able to forecast churn in order to improve the efficacy and usefulness of deep learning approaches. Deep learning techniques have a big influence on forecasting and improving client retention. Numerous studies employ deep learning and standard machine learning approaches to increase client retention, despite the fact that they have a number of accuracy issues. In response to this need, this paper describes the development of an RNN model and stacked bidirectional long short-term memory (SBLSTM) for the Arithmetic Optimization Algorithm (AOA) in CCP. The proposed AOA-SBLSTM-RNN model aims to precisely forecast the frequency of client attrition in the insurance industry. The AOA model first performs pre-processing to transform the original data into a format that may be utilized. Additionally, the SBLSTM-RNN model is used to differentiate between churning and non-churning users. In order to improve the CCP outcomes of the SBLSTM-RNN model, this study optimizes the hyperparameter tuning process using the Improved Gravitational Search Optimization Algorithm (IGSA). In this article, four sets of experiments were conducted, and three health insurance datasets were used to evaluate performance. The metrics of genuine churn, fake churn, specificity, precision, and accuracy are used to assess the efficacy of the proposed approach. The experimental findings show that the Ensemble Deep

Learning model AOA-SBLSTM-RNN with IGSA performs better than all other models in terms of prediction, with accuracy values of 97.89 and 97.67 on datasets 1 and 2.

e) Early warning of telecom enterprise customer churn based on ensemble learning:

Businesses may save a substantial amount of money on operating costs while providing customized services and targeted marketing strategies by employing machine learning algorithms to analyze client traits and provide an early warning of customer churn. Using the Python programming language, preliminary processes including data purification, oversampling, data standardization, and others are applied to the personal information and historical behavior data set of 900,000 telecom users. The appropriate model parameters were used in the construction of the Back Propagation Neural Network (BPNN). The Adaboost dual-ensemble learning model, which employed Random Forest (RF) as the foundation learner, was introduced together with the two conventional ensemble learning models, Random Forest (RF) and Adaboost. These four models, along with the other four conventional machine learning models—decision tree, naive Bayes, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM), in that order—were used to analyze the customer turnover data. In terms of recall rate, accuracy rate, F1 score, and other metrics, the RF-Adaboost dual-ensemble model outperforms the other four models, according to the results. On positive samples, the recall rates of the BPNN, RF, Adaboost, and RF-Adaboost dual-ensemble models are 79%, 90%, 89%, and 93%, respectively; the accuracy rates are 97%, 99%, 98%, and 99%; and the F1 scores are 87%, 95%, 94%, and 96%. The RF-Adaboost dual-ensemble model performs best, with the three indicators being 10%, 1%, and 6% higher than the reference, respectively. With the aid of the customer churn prediction statistics, telecom companies may reduce customer turnover and implement appropriate retention strategies for pre-churn consumers.

3. METHODOLOGY

i) Proposed Work:

The proposed system introduces a hybrid customer churn prediction framework that combines deep learning and ensemble learning techniques to improve prediction accuracy and reliability. Initially, the input telecom dataset undergoes preprocessing and class balancing using SMOTE-based techniques to handle data imbalance effectively. The core of the system utilizes a CNN-LSTM model, where the Convolutional Neural Network (CNN) extracts important spatial features from customer data, and the Long Short-Term Memory (LSTM) network captures temporal dependencies and sequential behavior patterns. This hybrid architecture enables the model to learn complex relationships and patterns that traditional machine learning models fail to capture.

In addition to the deep learning model, an ensemble Voting Classifier is integrated to further enhance prediction performance by combining the outputs of multiple models such as Boosted Decision Trees and Extra Trees. This ensemble approach improves robustness and reduces overfitting by leveraging the

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

strengths of different algorithms. Furthermore, a Flask-based web application is developed to provide a user-friendly interface for real-time churn prediction, allowing users to input customer data and obtain instant results. Overall, the proposed system ensures high accuracy, better generalization, and practical usability in real-world telecom environments.

ii) System Architecture:

The system architecture begins with dataset collection, where three benchmark datasets—IBM Telco, Churn-in-Telecom, and UCI Churn datasets—are gathered to ensure diversity and robustness in training. The collected data is then passed through a data preprocessing stage, which includes removing irrelevant attributes, handling missing values, and applying encoding techniques such as label encoding and one-hot encoding to convert categorical data into numerical form. After preprocessing, the data undergoes data balancing using advanced techniques like SMOTE, SMOTEEN, and SMOTETomek to address class imbalance and improve model fairness.

Following data preparation, the processed dataset is fed into the hybrid prediction framework, where the CNN-LSTM model extracts spatial and temporal features, and the ensemble Voting Classifier enhances prediction accuracy by combining multiple model outputs. The system performance is validated using 10-fold cross-validation to ensure reliability and generalization. Finally, the model is evaluated using multiple performance metrics such as accuracy, precision, recall, F1-score, AUC, and MCC. The entire system is integrated with a Flask-based web interface, enabling users to perform real-time churn prediction through an interactive and user-friendly platform.

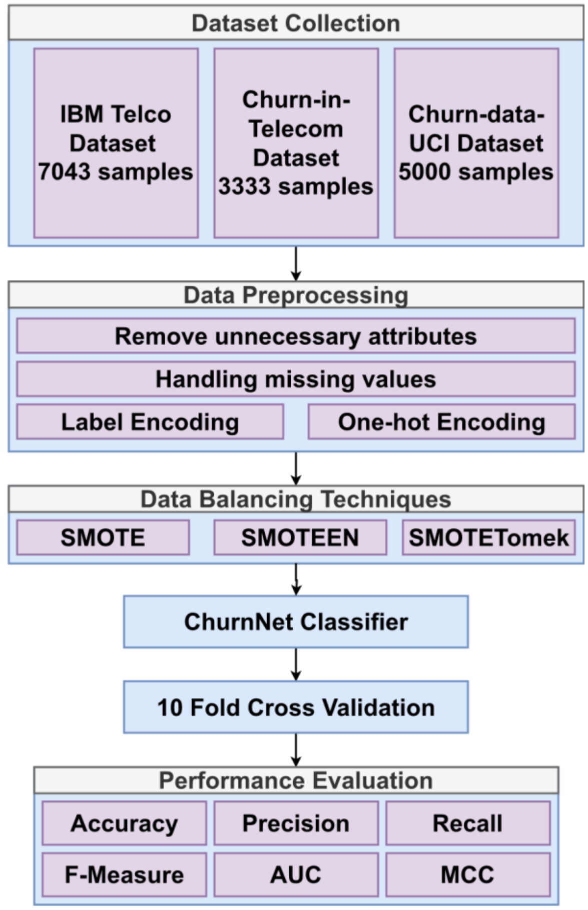

Fig1 proposed architecture

iii) Modules:

1. Dataset Collection

This module is responsible for gathering telecom customer datasets such as IBM Telco, Churn-in-Telecom, and UCI datasets. These datasets contain customer demographics, service usage, and churn labels. Using multiple datasets improves the diversity, reliability, and generalization capability of the model.

2. Data Preprocessing

In this module, raw data is cleaned and prepared for model training. Unnecessary attributes are removed, missing values are handled, and categorical data is converted into numerical format using label encoding and one-hot encoding. This step ensures the data is consistent and suitable for machine learning models.

3. Data Balancing

This module addresses the class imbalance problem commonly found in churn datasets. Techniques such as SMOTE, SMOTEEN, and SMOTETomek are applied to balance the number of churn and non-churn samples. This improves model fairness and prevents biased predictions.

4. Feature Learning using CNN-LSTM

This module uses a hybrid deep learning model combining CNN and LSTM. CNN extracts important feature patterns from the data, while LSTM captures temporal and sequential dependencies in customer behavior. Together, they help in learning complex relationships effectively.

5. Ensemble Voting Classifier

This module enhances prediction performance by combining multiple models such as Boosted Decision Trees and Extra Trees. The final prediction is made based on majority voting, which improves accuracy and reduces overfitting by leveraging multiple model strengths.

6. Model Training and Validation

In this module, the model is trained using the processed dataset and evaluated using 10-fold cross-validation. This ensures that the model performs consistently across different data splits and improves generalization to unseen data.

7. Performance Evaluation

This module evaluates the model using various metrics such as accuracy, precision, recall, F1-score, AUC, and MCC. These metrics provide a comprehensive understanding of model performance and help compare it with existing approaches.

8. Web Application (Flask Interface)

This module provides a user-friendly web interface using Flask. Users can input customer data and get real-time churn predictions. It also ensures secure access and easy interaction with the system for practical deployment.

iv) Algorithms:

a. Decision Tree

Decision Tree is a supervised learning algorithm that splits the dataset into multiple branches based on feature conditions to make predictions. In this project, it helps identify key factors influencing customer churn by providing clear and interpretable decision paths. Its simplicity allows easy visualization of decision rules, enabling telecom providers to understand important variables affecting customer retention.

b. Bagging Classifier

Bagging (Bootstrap Aggregating) improves model performance by training multiple models on different subsets of data and combining their predictions. In this system, Bagging is applied with Decision Trees to reduce variance and overfitting. By aggregating outputs from multiple learners, it provides more stable and reliable churn predictions.

c. AdaBoost

AdaBoost is an ensemble learning technique that combines multiple weak classifiers into a strong classifier by focusing on misclassified instances. In this project, it iteratively adjusts weights to improve prediction accuracy. This helps the model better identify difficult churn cases and enhances overall performance.

d. AdaBoost + MLP

This hybrid approach combines AdaBoost with a Multi-Layer Perceptron (MLP) to improve learning capability. AdaBoost enhances the training of the neural network by emphasizing hard-to-classify samples. As a result, the model captures complex patterns in customer data more effectively, leading to improved churn prediction accuracy.

e. Bagging + MLP

Bagging combined with MLP improves model stability by training multiple neural networks on different data subsets and averaging their predictions. This reduces overfitting and increases generalization ability. In this project, it helps in capturing complex relationships in customer behavior while maintaining robust performance.

f. CNN (Convolutional Neural Network)

CNN is a deep learning model used for extracting important features from structured data. In this system, CNN identifies hidden patterns and relationships within customer attributes using convolutional layers. This improves the model's ability to detect churn-related patterns and enhances prediction accuracy.

g. LSTM (Long Short-Term Memory)

LSTM is a type of recurrent neural network designed to capture long-term dependencies in sequential data. In this project, it analyzes customer behavior over time, identifying trends and usage patterns. This temporal learning capability significantly improves churn prediction by considering historical data.

h. GRU (Gated Recurrent Unit)

GRU is a simplified version of LSTM that also captures sequential dependencies but with fewer parameters. It processes time-based customer data efficiently and reduces training time. In this project, GRU helps in fast and effective churn prediction while maintaining good accuracy.

i. ChurnNet (CNN with Residual Layer)

ChurnNet is an advanced deep learning architecture combining CNN with residual connections. The residual layers help overcome the vanishing gradient problem and improve feature learning. This model captures complex patterns in customer data, resulting in higher prediction accuracy and better performance compared to traditional methods.

j. Voting Classifier (Boosted Decision Tree + Extra Tree)

The Voting Classifier combines predictions from multiple models and produces the final output based on majority voting. In this system, Boosted Decision Trees and Extra Trees are used to improve prediction robustness. This ensemble method reduces bias and enhances overall churn prediction reliability.

k. CNN + LSTM Hybrid Model

The CNN + LSTM model integrates feature extraction and sequence learning into a single framework. CNN extracts spatial features from customer data, while LSTM captures temporal behavior patterns. This combination enables the system to learn both static and dynamic relationships, providing highly accurate churn predictions.

4. EXPERIMENTAL RESULTS

The proposed customer churn prediction system was implemented using a Flask-based web application, where users can upload test datasets and obtain real-time predictions. The system successfully classifies customers as churn or non-churn based on input features, as shown in the prediction output screen. Multiple test instances were evaluated, and the model demonstrated consistent and accurate predictions across different customer records. This confirms the effectiveness of the hybrid CNN-LSTM and ensemble Voting Classifier approach in capturing complex customer behavior patterns.

To evaluate model performance, several metrics such as Accuracy, Precision, Recall, and F1-Score were analyzed and compared across different models and data balancing techniques. The performance graph shows that the CNN-LSTM model with SMOTEEN balancing achieved the highest accuracy and overall performance, outperforming other configurations like SMOTE and SMOTETomek. In contrast, models without data balancing showed significantly lower performance, highlighting the importance of handling class imbalance. Overall, the experimental results demonstrate that the proposed system provides high accuracy, robustness, and improved generalization, making it suitable for real-world telecom churn prediction applications.

Accuracy: A test's accuracy is its capacity to distinguish healthy from ill cases. Find the percentage of instances with genuine positives and negatives to assess test accuracy.

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Precision: Classification accuracy or positive cases constitute precision. The formula for accuracy is:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\Pr e\,cision = \frac{TP}{(TP + FP)}$$

Recall: A model's recall measures its ability to recognize all appropriate machine learning class instances. The ratio of accurately predicted positive observations to total positives indicates a model's class instance detection skill.

$$Recall = \frac{TP}{(FN + TP)}$$

mAP: Mean Average Precision ranks quality. It considers the number and order of relevant ideas. Calculating MAP at K uses the arithmetic mean of each user or query's Average Precision (AP).

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{ the AP of class } k$$

$$n = \text{ the number of classes}$$

F1-Score: A high F1 score suggests an accurate machine learning model. Integrating recall and precision improves model correctness. Accuracy measures how often a model predicts a dataset correctly.

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr e\,cision)}{(Recall + \Pr e\,cision)}$$

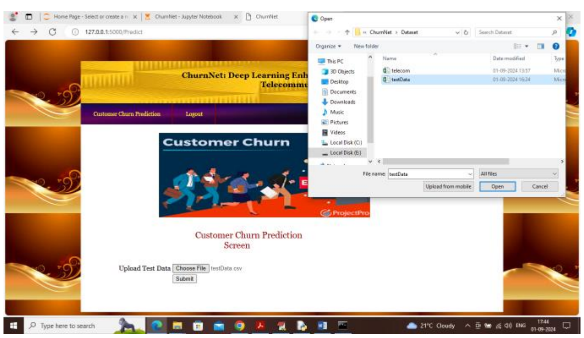

Fig 2 uploading test data file

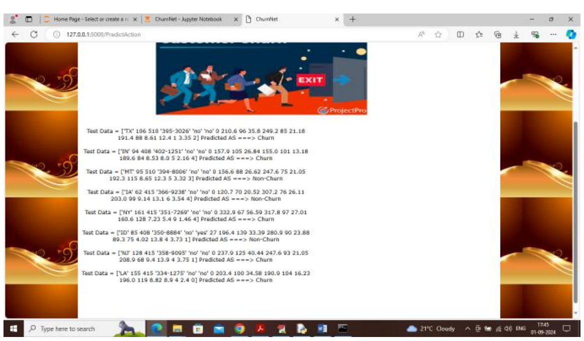

Fig 3 predicted results.

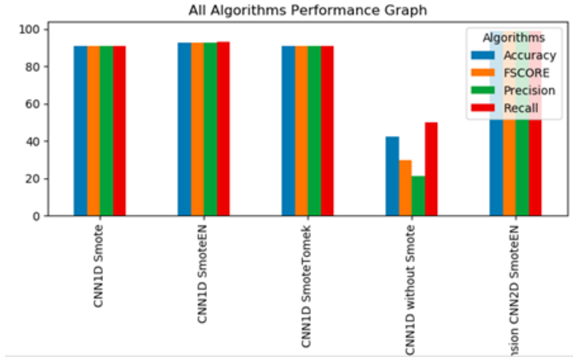

Fig6 Performance Comparison of Forecasting Algorithms

5. CONCLUSION

This paper presents a hybrid customer churn prediction framework that integrates CNN-LSTM deep learning with an ensemble Voting Classifier to improve prediction accuracy in the telecommunication industry. The proposed system effectively captures both spatial and temporal patterns in customer data, while SMOTE-based data balancing techniques address the issue of class imbalance. The experimental results demonstrate that the model outperforms traditional machine learning approaches in terms of accuracy, robustness, and generalization.

Furthermore, the development of a Flask-based web application enables real-time prediction and practical usability of the system. The proposed approach provides a reliable solution for identifying potential churn customers, allowing telecom providers to take proactive retention measures. Future work can focus on incorporating Explainable AI (XAI) for better model interpretability and federated learning for decentralized data environments, enhancing scalability and real-world deployment.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

6. FUTURE SCOPE

The proposed churn prediction system can be further enhanced by integrating Explainable Artificial Intelligence (XAI) techniques to improve model transparency and help telecom providers better understand the reasons behind churn predictions. Additionally, the system can be extended using federated learning, allowing multiple telecom operators to collaboratively train models without sharing sensitive customer data, thereby improving privacy and scalability.

Future improvements may also include the use of advanced deep learning architectures such as Transformer-based models for better sequence modeling and prediction accuracy. The system can be deployed on cloud and big data platforms to handle large-scale real-time data processing. Furthermore, incorporating real-time streaming data analysis and personalized recommendation systems can help telecom companies take proactive actions to retain customers more effectively.

REFERENCES

[1] I. V. Pustokhina, D. A. Pustokhin, P. T. Nguyen, M. Elhoseny, and K. Shankar, ‘‘Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector,’’ Complex Intell. Syst., vol. 9, no. 4, pp. 3473–3485, Aug. 2023.

[2] J. Rabbah, M. Ridouani, and L. Hassouni, ‘‘New approach to telecom churn prediction based on transformers,’’ in Proc. Int. Conf. Artif. Intell. Comput. Vis. Cham, Switzerland: Springer, 2023, pp. 565–574.

[3] P. Yang, ‘‘Data visualization and prediction for telecom customer churn,’’ Highlights Sci., Eng. Technol., vol. 39, pp. 1080–1085, Apr. 2023.

[4] N. Jajam, N. P. Challa, K. S. L. Prasanna, and C. H. V. S. Deepthi, ‘‘Arithmetic optimization with ensemble deep learning SBLSTM-RNN-IGSA model for customer churn prediction,’’ IEEE Access, vol. 11, pp. 93111–93128, 2023.

[5] Y. Zhou, W. Chen, X. Sun, and D. Yang, ‘‘Early warning of telecom enterprise customer churn based on ensemble learning,’’ PLoS ONE, vol. 18, no. 10, Oct. 2023, Art. no. e0292466.

Author Profiles

Ms. K. Pavani Working as Assistant & Head of Department of MCA, in SRK Institute of technology in Vijayawada. She has done MCA, M. Tech in Computer Science. Her area of interest includes Machine Learning with Python and DBMS.

Mr. Ch. Satyanarayana Reddy Completed his MCA, BCA. He has been System Administrator, Networking Administrator

and Oracle Administrator. He is also a web developer and python developer, currently working as an Assistant Professor in the department of MCA at SRK Institute of Technology, Enikepadu, NTR District. His area of interest includes Artificial Intelligence and Machine Learning.

Ms. P. Mary Kamala is an MCA Student in the Department of Computer Applications at SRK Institute of Technology, Enikepadu, Vijayawada, NTR District. She has Completed Degree in B.Sc. (Computer Science) from Maris Stella College,Vijayawada. Her area of interests are DBMS and Machine Learning with Python.