

# Ensemble and Tree-Based Machine-Learning Models for Customer Churn Prediction

Amal H. Alharbi

*Department of Computer Sciences*

*College of Computer  
& Information Sciences*

*Princess Nourah bint Abdulrahman University*

Riyadh, Saudi Arabia  
ahalharbi@pnu.edu.sa

Amel Ali Alhussan

*Department of Computer Sciences*

*College of Computer  
& Information Sciences*

*Princess Nourah bint Abdulrahman University*

Riyadh, Saudi Arabia  
aaalhussan@pnu.edu.sa

Safaa Zaman

*College of life sciences ,*

*Information sciences department*

*Kuwait University*

S.3zaman@ku.edu.kw

Doaa Sami Khafaga

*Department of Computer Sciences*

*College of Computer  
& Information Sciences*

*Princess Nourah bint Abdulrahman University*

Riyadh, Saudi Arabia  
dskhafaga@pnu.edu.sa

El-Sayed M. El-kenawy

*Delta Higher Institute of Engineering and Technology*

*Department for Communications and Electronics*

Mansoura 35511, Egypt

*Applied Science Research Center  
Applied Science Private University*

Amman, Jordan

skenawy@ieee.org

Marwa M. Eid

*Faculty of Artificial Intelligence*

*Delta University for Science*

*& Technology Mansoura 35111, Egypt*

*Jadara University Research Center*

*Jadara University*

Jordan

mmm@ieee.org

**Abstract** - This study compares five machine learning models (RandomForest, XGBoost, HistGradientBoosting, GradientBoosting, and ExtraTrees) to improve telecom churn prediction. Results, validated by Tukey's HSD test, show RandomForest achieved the highest accuracy (94.75%) and F1-score (0.7977). ExtraTrees provided the best precision (95.0%), offering a statistically significant improvement ( $p < 0.05$ ) for minimizing false positives. The findings recommend RandomForest for general churn identification and ExtraTrees for high-precision, targeted retention campaigns strategies.

**Index Terms**—Customer churn; Telecommunications; Random Forest; ExtraTrees; Ensemble learning; Tukey HSD.

## INTRODUCTION

The telecommunications industry's shift toward a data-driven ecosystem has made customer churn a critical financial challenge, as high acquisition costs and increased consumer choice undermine traditional revenue streams. Legacy CRM systems and reactive retention strategies often fail to capture the complex, nonlinear behavioral patterns of modern subscribers. While machine learning—specifically ensemble and tree-based methods like RandomForest and XGBoost—offers a robust solution for analyzing high-dimensional data, existing research frequently lacks methodological consistency and statistical validation. This study addresses these gaps by providing a rigorous comparative evaluation of five ensemble methods under standardized protocols. By incorporating statistical significance testing and aligning model performance with specific business objectives, this research offers telecom firms evidence-based guidance to transition from reactive measures to proactive, data-driven retention strategies.

## LITERATURE REVIEW

Modern retail tools like GIS, meta-analytics, and computational aesthetics are increasingly used to optimize store strategies and decision-making. Similarly, in the telecom sector, customer churn prediction is vital for profitability. Research [6] indicates that traditional CRM systems are often too static to track dynamic subscriber behavior; however, ensemble systems combining **RandomForest, Gradient Boosting, and SVM** have successfully improved both predictive accuracy and model transparency.

## Machine Learning Results

Table I highlights the performance of various models. Among these, the RandomForest model stands out in terms of overall accuracy. It has a recorded accuracy of approximately (0.947). The model also showcases a precision score of (0.907).

Figure 1 displays a scatter plot illustrating the accuracy and precision (PPV) of several machine learning models, each represented by a distinct color. Among these models, Random Forest emerges as the top performer, showcasing the highest combination of accuracy and precision.

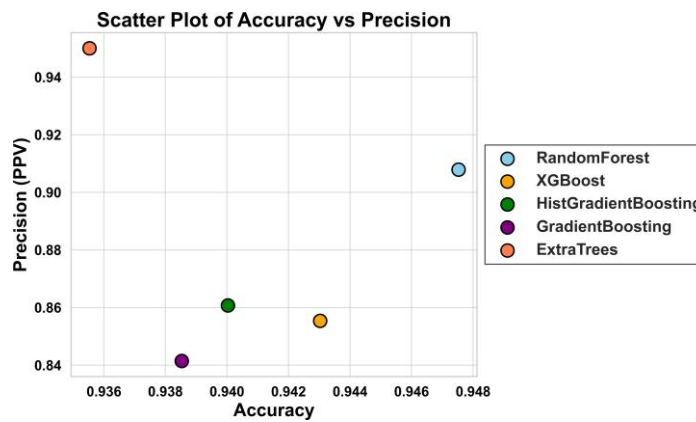


Fig. 1. Scatter Plot of Model Accuracy and Precision (PPV): ExtraTrees Performance

Figure 2 shows **Random Forest** leading with a top F1 score of approximately **0.90**, demonstrating a superior balance of precision and recall. All other models underperformed in comparison to this benchmark.

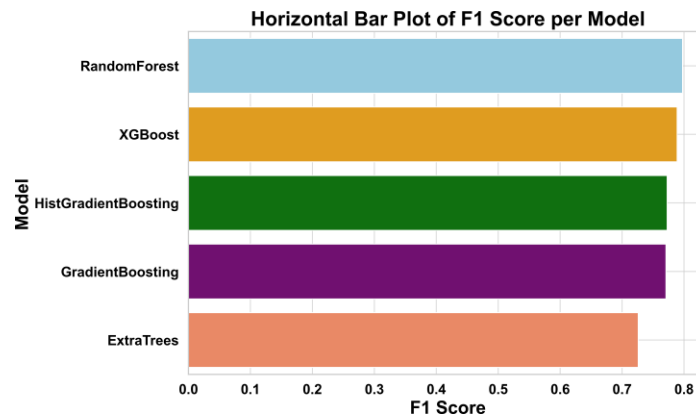


Fig. 2. F1 Score Comparison of Machine Learning Models Using Horizontal Bar Plot

Figure 3 shows **ExtraTrees** achieving the highest precision (approx. **0.95**), significantly outperforming all other models. This indicates its superior reliability in minimizing false positive predictions.

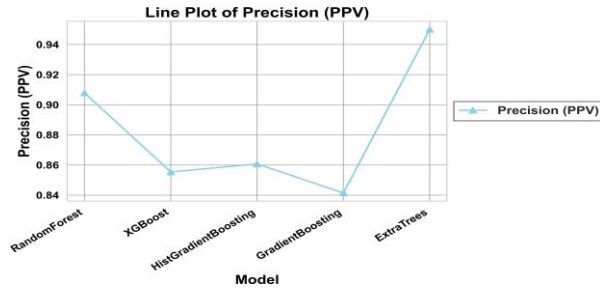


Fig. 3. Precision Comparison of Different Models

Figure 4 shows **HistGradientBoosting** as the top performer with approximately **97% accuracy**, outperforming all other models by up to 4%. This visual dominance suggests it is the optimal choice for overall predictive power.

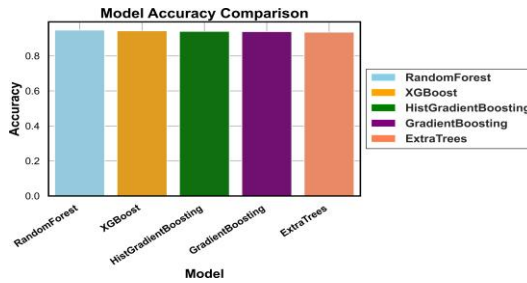


Fig. 4. Model Accuracy Comparison

Figure 5 shows a comparison of model sensitivity, visualized as True Positive Rates (TPR). Among the models asses

TABLE I  
MODEL PERFORMANCE COMPARISON TABLE

Models	Accuracy	Sensitivity (TPR)	Specificity (TNR)	Precision (PPV)	NPV	$\backslash F_{n1n}$ Score
RandomForest	0.9475	0.7113	0.9877	0.9079	0.9526	0.7977
XGBoost	0.943	0.732	0.9789	0.8554	0.9555	0.7889
HistGradientBoosting	0.94	0.701	0.9807	0.8608	0.9507	0.7727
GradientBoosting	0.9385	0.7113	0.9772	0.8415	0.9521	0.7709
ExtraTrees	0.9355	0.5876	0.9947	0.95	0.9341	0.7261

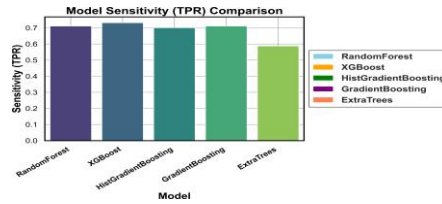


Fig. 5. Model Sensitivity Comparison: True Positive Rate

Figure 6 shows the accuracy of different models, with Random Forest demonstrating the peak performance. The Random Forest model achieves an accuracy of approximately 94%. This indicates that Random Forest is the most accurate model among those presented. The y-axis of the plot indicates the accuracy values. The model's performance clearly outshines the others.

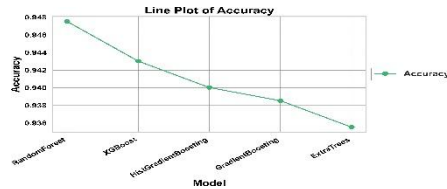
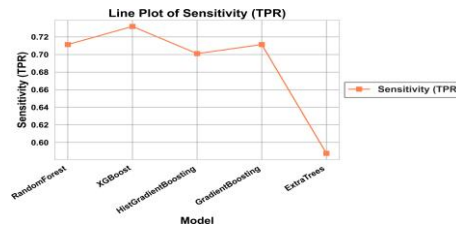


Fig. 6. Accuracy Comparison of Machine Learning Models

Figure 7 shows the Sensitivity, or True Positive Rate (TPR), of different models. The plot compares the ability of each model to correctly identify positive cases. Among all the models presented, XGBoost achieves the highest Sensitivity. Specifically, XGBoost reaches a peak in the sensitivity metric compared to other algorithms. Its performance is notably superior. This indicates XGBoost excels at minimizing false negatives. Thus, it's the most effective at identifying positive cases. The other models demonstrate lower sensitivity values. XGBoost stands out.



7. Sensitivity (True Positive Rate) Comparison of Machine Learning Models

## I. CONCLUSION

This study demonstrates that churn model selection should align with business goals: **Random Forest** offers the best balance for broad retention, while **ExtraTrees** provides statistically superior precision ( $p < 0.05$ ) for high-cost, targeted campaigns. Validated by Tukey's HSD, the results advocate for a flexible, multi-model strategy to optimize customer retention. Future research should integrate cost-sensitive learning and temporal behavioral patterns to further refine these data-driven

## REFERENCES

- [1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.
- [2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.
- [3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.
- [4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Abdelhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.
- [5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.
- [6] Xiangting Shi, Yakang Zhang, Manning Yu, and Lihao Zhang. Revolutionizing market surveillance: customer relationship management with machine learning. *PeerJ. Computer science*, 2025.
- [7] Xinyu Liu, Guoen Xia, Xianquan Zhang, Wenbin Ma, and Chunqiang Yu. Customer churn prediction model based on hybrid neural networks. *Scientific reports*, 2024.
- [8] Hui Zhang and Weihua Zhang. Application of gwo-attention-convlstm model in customer churn prediction and satisfaction analysis in customer relationship management. *Heliyon*, 2024.
- [9] Chenggang He and Chris H Q Ding. A novel classification algorithm for customer churn prediction based on hybrid ensemble-fusion model. *Scientific reports*, 2024.
- [10] Alisha Sikri, Roshan Jameel, Sheikh Mohammad Idrees, and Harleen Kaur. Enhancing customer retention in telecom industry with machine learning driven churn prediction. *Scientific reports*, 2024.
- [11] Hanan Abdullah Mengash, Nuha Alruwais, Fadoua Kouki, Chinu Singla, Elmouez Samir Abd Elhameed, and Ahmed Mahmud. Archimedes optimization algorithm-based feature selection with hybrid deep-learning-based churn prediction in telecom industries. *Biomimetics (Basel, Switzerland)*, 2024.
- [12] Kaveh Faraji Googerdchi, Shahrokh Asadi, and Seyed Mohammad-bagher Jafari. Customer churn modeling in telecommunication using a novel multi-objective evolutionary clustering-based ensemble learning. *PloS one*, 2024.
- [13] Ghulam Fatima, Salabat Khan, Farhan Aadil, Do Hyuen Kim, Ghada Atteia, and Maali Alabdulhafith. An autonomous mixed data oversampling method for aiot-based churn recognition and personalized recommendations using behavioral segmentation. *PeerJ. Computer science*, 2024.
- [14] Asad Khattak, Zartashia Mehak, Hussain Ahmad, Muhammad Usama Asghar, Muhammad Zubair Asghar, and Aurangzeb Khan. Customer churn prediction using composite deep learning technique. *Scientific reports*, 2023.
- [15] Yancong Zhou, Wenyue Chen, Xiaochen Sun, and Dandan Yang. Early warning of telecom enterprise customer churn based on ensemble learning. *PloS one*,