IV. IMPLEMENTATION

1. Model Training and Validation: We offer a comprehensive explanation of the process of dividing the dataset into training and validation sets, several cross-validation techniques, and the procedures for training models for each methodology.

2. Model evaluation: This section examines the assessment metrics, namely accuracy, precision, recall, F1-score, and ROC-AUC, that are employed to gauge the performance of each model. In addition, we employ confusion matrices and ROC curves to visually represent the outcomes.

3. Model Integration: We combine sentiment scores with other features and evaluate their effect on churn prediction performance. We utilise feature significance measures to choose the features that have the most predictive power for churn.

V. RESULT ANALYSIS:

The models were trained and tested using a real-time dataset obtained from Kaggle. The dataset includes various features that influence customer churn, such as customer activity, purchase history, and sentiment analysis from customer reviews. The performance of each model was evaluated using standard metrics: Accuracy, Precision, Recall, F1 Score, and the Area Under the Receiver Operating Characteristic curve (AUC). Each metric provides insights into different aspects of the model's prediction capabilities.

Logistic Regression showed an accuracy of 77.88%, indicating a relatively high rate of correct predictions. The high precision (96.22%) suggests that when it predicts churn, it is very likely to be correct. However, its recall is lower at 86.75%, indicating some true churn cases were missed. The F1 score, which balances precision and recall, is at 71.79%, and the AUC is quite high at 98.52%, indicating excellent model discrimination ability.

Naive Bayes achieved an accuracy of 83.39% and had the lowest precision of 71.02%. Its recall is high at 94.73%, which means it can identify most of the actual churn cases, but it also has a higher false positive rate. The F1 score is at 93.82%, and the AUC is 70.05%, suggesting moderate discrimination ability.

The SVM model had a better balance with an accuracy of 91.06%, precision of 75.24%, and recall of 80.16%. With an F1 score of 81.01% and an AUC of 74.22%, the results outperform Naive Bayes, although the AUC is still lower than that of Logistic Regression.

The accuracy of the Random Forest ensemble model is 83.77%, while its precision and recall are 79.82% and 77.77%, correspondingly. With a high discriminating ability and a decent balance between precision and recall, the F1 score is 86.58% and the AUC is 92.59%.

It is noteworthy that, while Support Vector Machines (SVM) have the highest accuracy, Logistic Regression has the highest AUC value, indicating that it is the most effective model for

differentiating between non-churners and churners. Logistic regression is especially helpful in situations when there is a large cost associated with false negatives, or failing to identify a churner. The particular business context and cost-benefit analysis may also play a role in selecting the optimal model. For instance, a model with a better precision, such as logistic regression, can be favored if the cost of false positives—predicting churn when none exists—is large. Alternatively, a high recall model such as Naive Bayes can be more appropriate if losing out on possible churners is more expensive.

CONCLUSION:

This study investigated the incorporation of sentiment analysis into machine learning algorithms to forecast client attrition in the E-commerce industry. We evaluated the effectiveness of Logistic Regression, Naive Bayes, SVM, and Random Forest models using a comprehensive Kaggle dataset. We measured their performance using measures such as Accuracy, Precision, Recall, F1 Score, and AUC. The results of our study demonstrate that Logistic Regression achieved outstanding performance in predicting churn, particularly in terms of AUC, which indicates its high discriminatory power. Naive Bayes demonstrated a high recall rate, SVM exhibited a well-balanced performance, and Random Forest exhibited robustness with good precision and recall rates. The use of sentiment analysis improved the forecast precision, highlighting the influence of consumer emotions on churn. The study emphasises the efficacy of these models in predicting customer turnover in E-commerce, with Logistic Regression being particularly notable for its high level of accuracy. These insights can assist firms in formulating more effective client retention strategies. Subsequent investigations could enhance these discoveries by delving into more intricate algorithms and amalgamating diverse models to get enhanced precision.