TELECOM CUSTOMER CHURN PREDICTION USING MACHINE LEARNING ALGORITHMS

BIA 5402 –Machine Learning and Programming 2

Submitted to : Professor Sarama Shehmir

THE TEAM

- AMARACHI EZEJI N01510367
- APARAJITA ROY N01511087
- AYESHA KARADIA N01514568
- GBENGA ADEBOWALE N01513109
- IVIN ALEXANDER N01474172

TABLE OF CONTENTS

INTRODUCTION

The goal of this project is to create a churn prediction system for a well-known telecommunications company. Our goal is to analyze past customer data using data analytics and machine learning to find trends and signals that point to possible churn behavior. The brand can cultivate long-lasting customer relationships and ensure sustainable growth and ongoing success in a market that is constantly changing by taking a proactive approach to churn prediction and retention.

DATA DESCRIPTION

- 'Senior_Citizen' : A binary variable showing if the consumer is a senior citizen
- 'Partner' and 'Dependents': Binary variables showing if the client has a partner or dependents
- 'tenure': The number of months the customer has been with the firm
- 'Phone_Service', 'Paperless_Billing': Binary variables showing if the client has phone service and paperless billing, respectively
- 'Monthly_Charges', 'Total_Charges': The monthly and total costs incurred by the customer, respectively
- 'gender': A binary variable showing if the customer is a male or female
- 'InternetService' : The type of internet service a customer is using
- 'MultipleLines', 'OnlineSecurity', 'OnlineBackup', 'DeviceProtection', 'TechSupport', 'StreamingTV', 'StreamingMovies': categorical variables indicating whether or not a consumer is using particular services
- 'Contract': The type of contract made between the customer and the firm

DATA PREPARATION

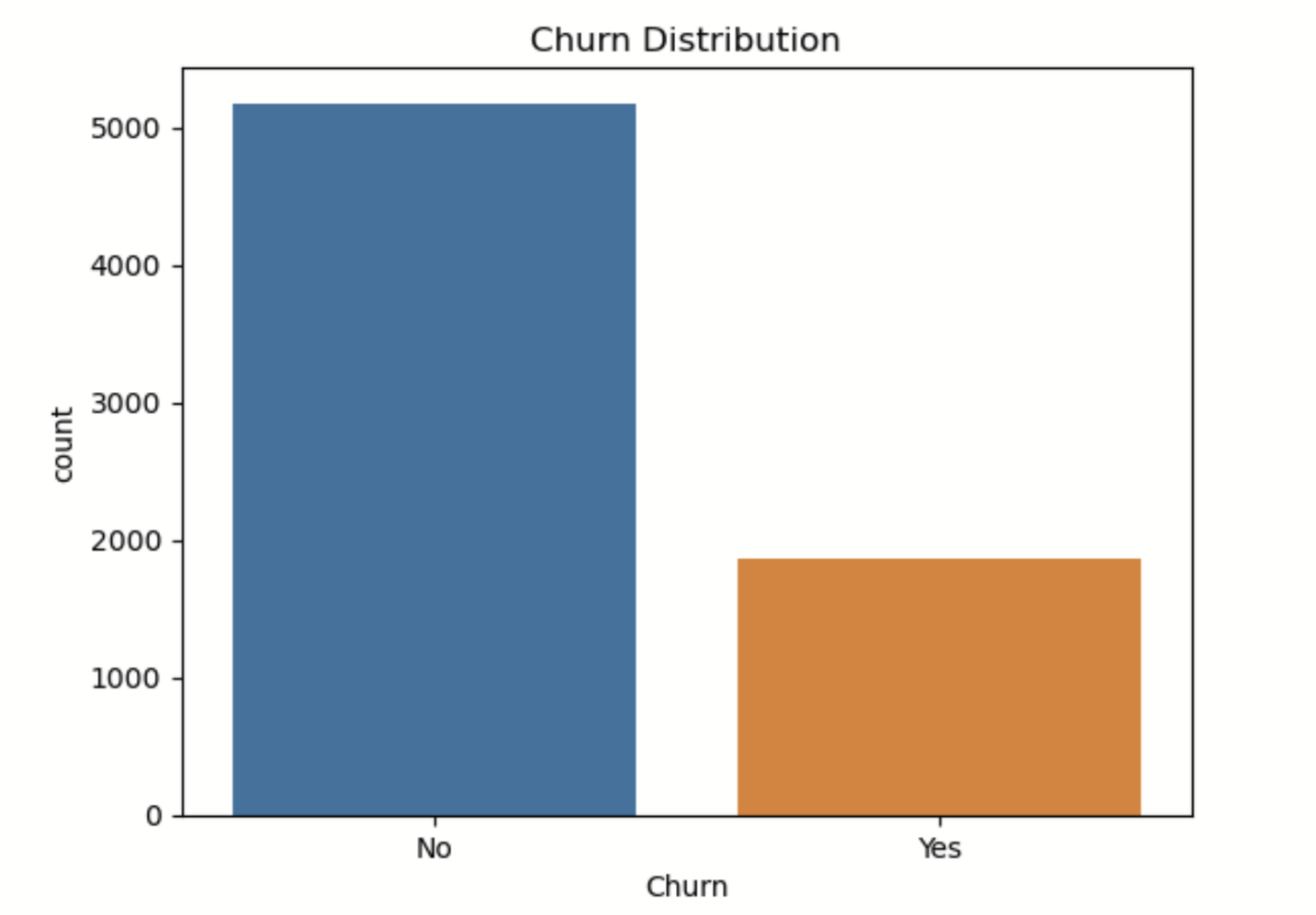
- Missing Values: The “TotalCharges” Column and “AvgMonthlySpending” Column in the dataset both had missing values and we replaced the values with the calculated means.
- Encoding Categorical Variables: One-hot encoding was used to encode categorical variables as binary or dummy variables.
- Feature Scaling: To guarantee that all features had comparable sizes, numerical features (such as "tenure," "MonthlyCharges," and "TotalCharges" were standardised using the StandardScaler.
- Data Splitting and Modeling: The dataset was split into two core components: the feature set and the target variable , a fundamental step in building an effective machine learning model.

EDA & VISUALIZATION INTERPRETATION

- EDA was undertaken to acquire insights into the data and find trends connected to churn.
- Visualizations were utilized to better comprehend the distribution of the target variable and correlations between various characteristics and churn.
- Customers with month-to-month contracts left the company at a rate that was 7-8 times higher than those with one- or two-year contracts.

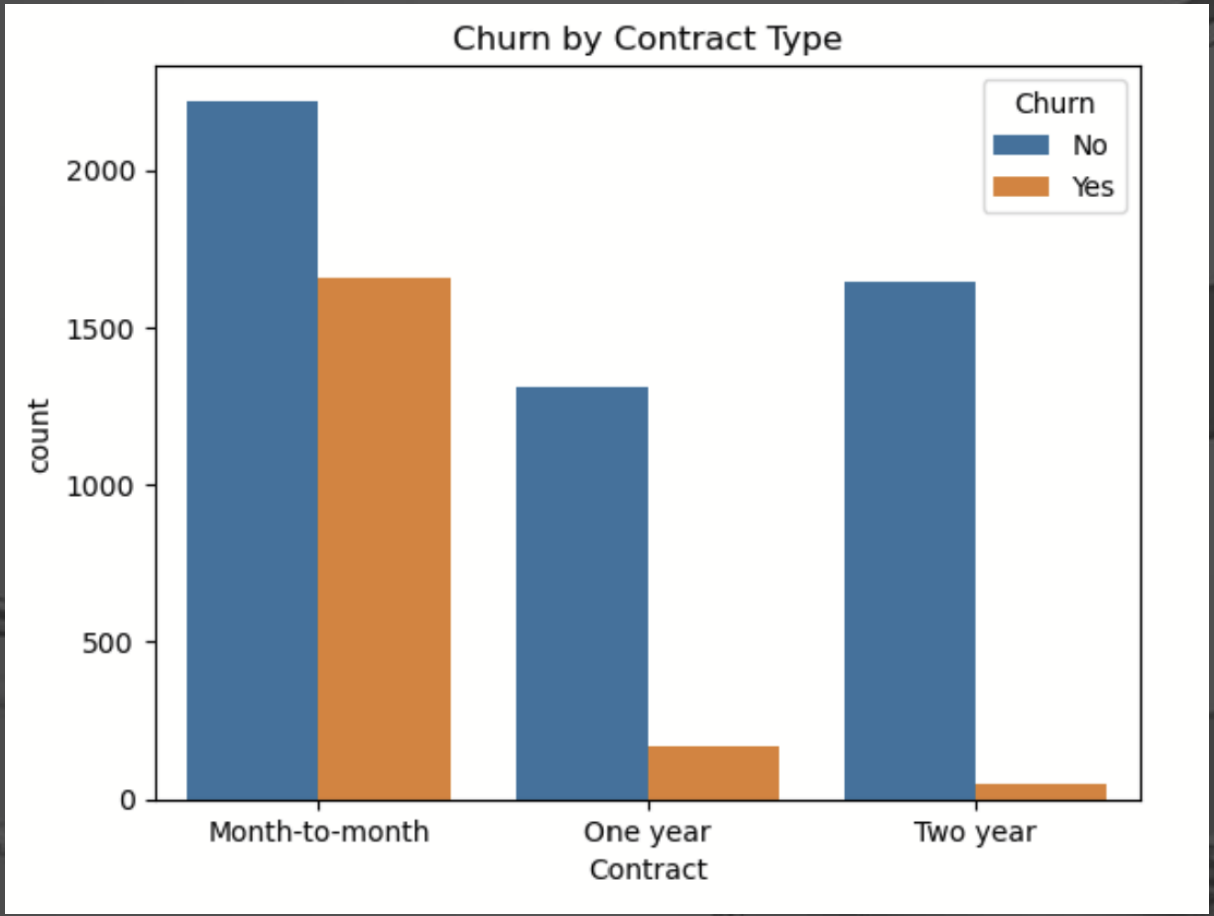
CHURN DISTRIBUTION

The graphic demonstrated that there was a class imbalance, with a greater proportion of non-churned consumers compared to churned customers.



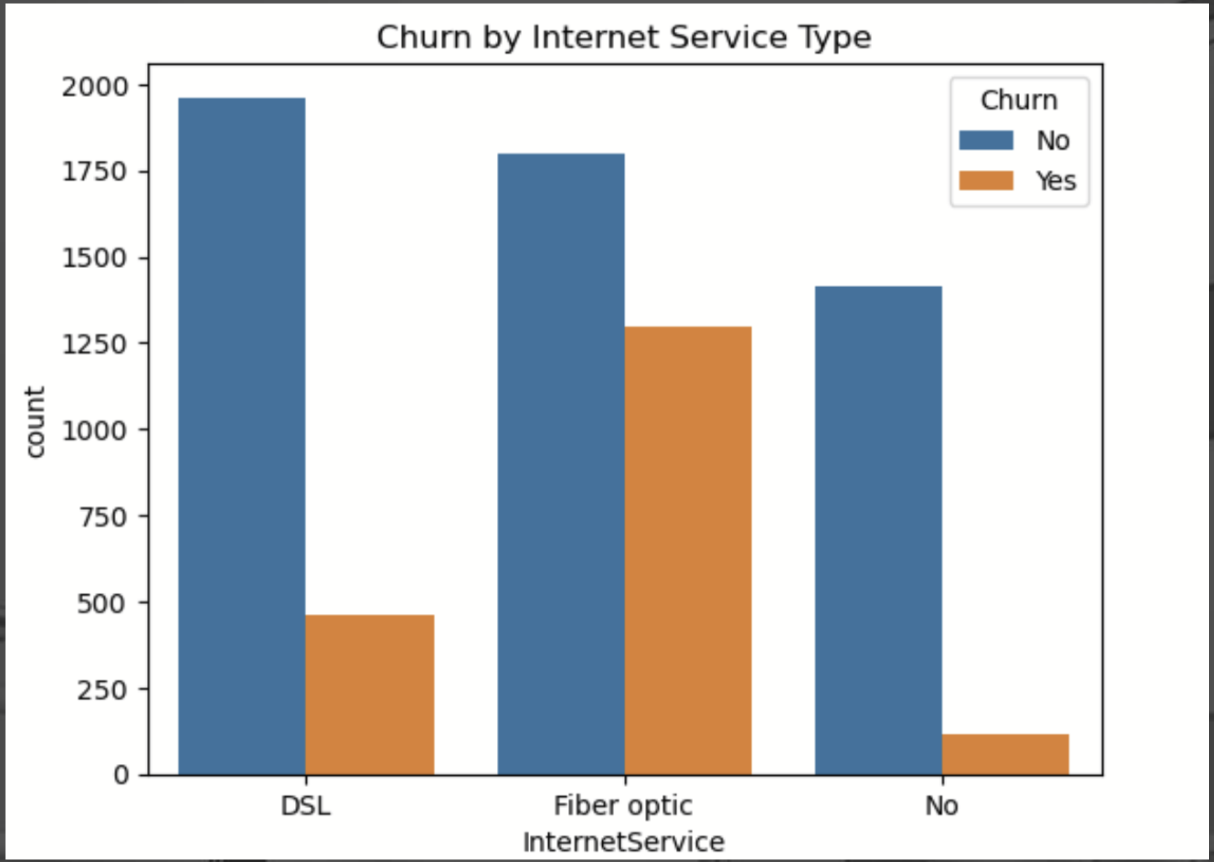
CHURN BY CONTARCT TYPE

Customers with shorter contract periods are more likely to churn, as shown in the above figure. Customers with month-to-month contracts left the company at a rate that was 7-8 times higher than those with one- or two-year contracts.



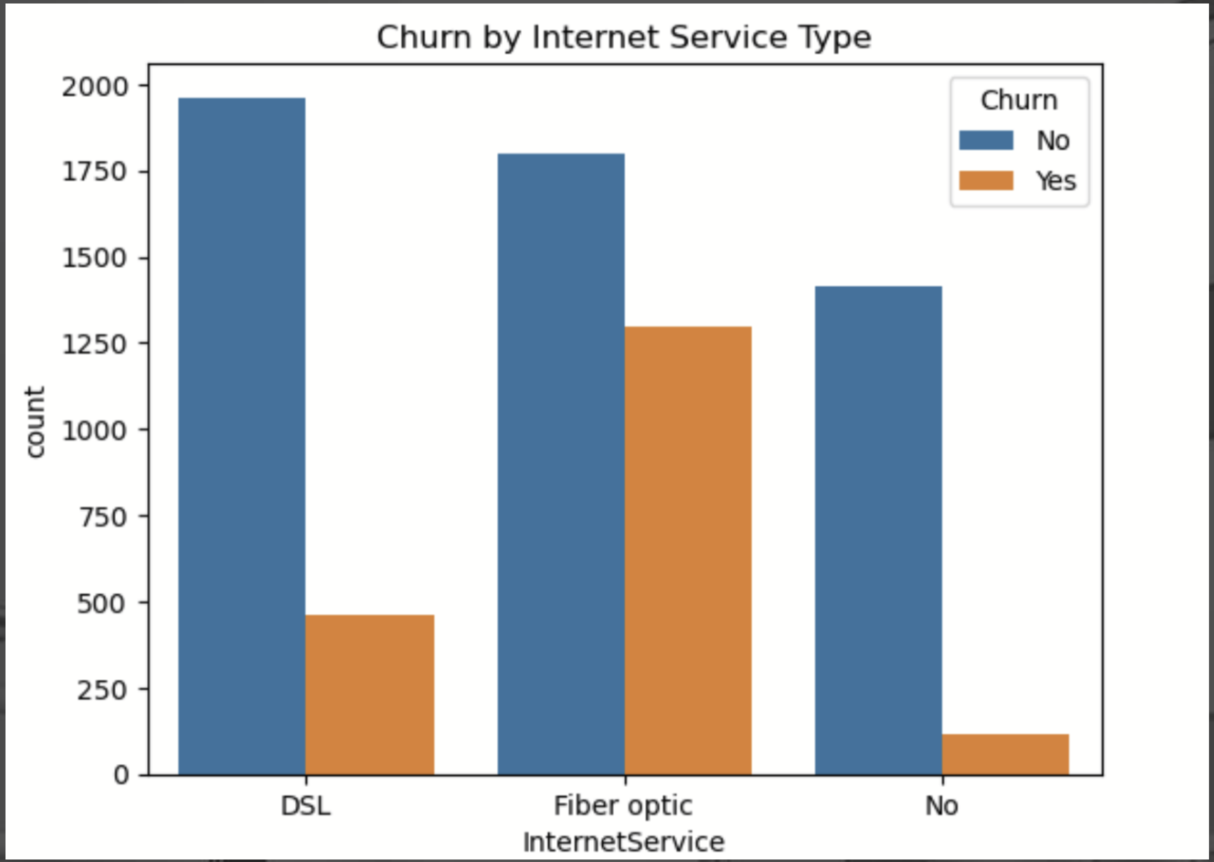
CHURN BY INTERNET SERVICE TYPE

We deduced from the analysis of the aforementioned visualisation that consumers with fibre optic internet service were more likely to leave than those with DSL or no internet access. When compared to customers with DSL or those without internet connection, customers with fibre optic internet service had a churn rate that was four times higher.



CHURN BY PAYMENT METHOD

Customers that pay by electronic cheque have a higher churn rate, as seen in the aforementioned statistic. Those who pay with an electronic cheque have a higher churn rate, which is five times higher than those who pay in another way.



MODEL BUILDING & EVALUATION

- Neural Network Classification
- Random Forest
- Association Rule Mining

NEURAL NETWORK CLASSIFICATION

- In this analysis, a predictive model for customer churn was developed using a Multi-Layer Perceptron architecture, a form of artificial neural network
- The MLP architecture was utilized for its ability to capture complex relationships within the data
- The model was trained using a range of hyperparameters that significantly influence its performance
- The performance evaluation of the neural network revealed its predictive prowess in identifying customer churn
- Overall, the MLP neural network demonstrated its capacity to make accurate predictions and offer valuable insights into customer churn behavior, which holds potential for informed decision-making in the business context

RANDOM FOREST

- Developed a predictive model for customer churn using Random Forest Classifier.
- Chose Random Forest for its capability to handle intricate data relationships.
- Leveraged ensemble learning to employ decision tree classifiers collectively.
- Utilized Random Forest architecture to address complex data interactions.
- Trained the model using an array of hyperparameters, significantly shaping its performance.
- Hyperparameters included size of hidden layers, activation functions, optimization algorithms, training iterations, and batch size.
- Performance assessment of the Random Forest model unveiled its predictive strength in detecting customer churn.
- Model's ability to make accurate predictions showcased its proficiency in understanding churn dynamics.

ASSOCIATION RULE MINING

- In this study, we used Association Rule Mining to explore the intricate aspects of customer churn from a different perspective
- By focusing on relevant categorical features, the dataset was meticulously prepared for association rule mining
- Ultimately, the Association Rule Mining approach provided a novel perspective that contributed to a more comprehensive understanding of the intricate factors influencing customer churn behavior

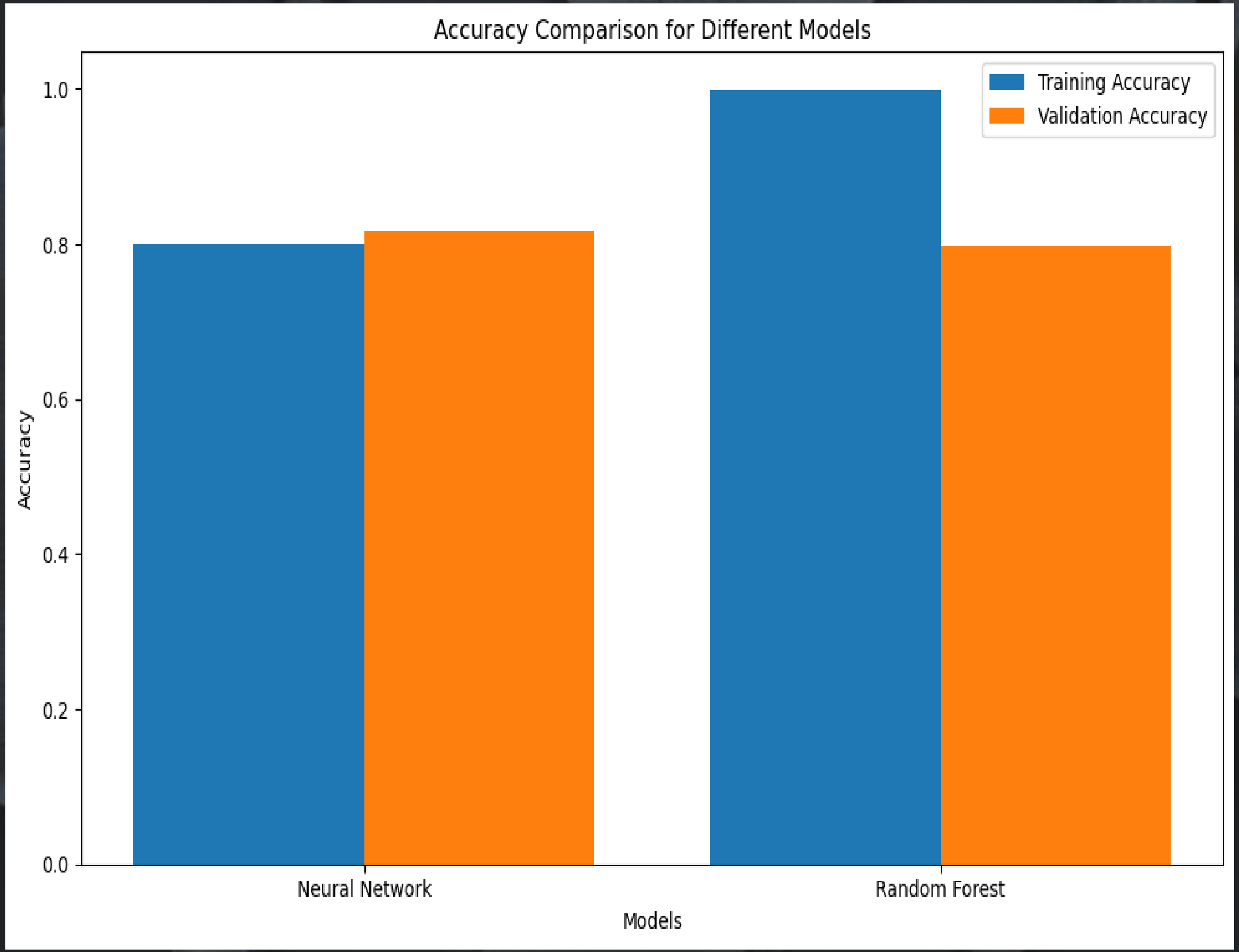
EVALUATION METRICS

To increase model performance, we may study different machine learning methods like Decision Trees, and XGBoost

Additionally, feature engineering ,feature importance and hyperparameter tweaking might be further researched to boost model accuracy and interpretability

Based on the insights acquired from the investigation, the telecom operator may execute targeted retention measures to lower churn rates

COMPARATIVE ANALYSIS



Evaluation Metrics	Neural Network	Random Forest
Training Accuracy	80.33%	99.8%
Validation Accuracy	81.26%	79.7%
Precision	68.73%	71.30%
Recall	53.62%	45.30%
F1-Score	60.24%	55.40%

CONCLUSION

- In this research, we successfully created a telecom churn prediction model employing a neural network classifier, random forest classifier and association rule mining
- The model demonstrated promising results in forecasting customer attrition, and the association rules gave useful insights into consumer behavior
- By employing this predictive model, the telecom business may take proactive efforts to retain consumers and enhance overall customer happiness

FUTURE WORK

- Future work may comprise researching additional machine learning models, undertaking feature significance analysis, and implementing the final model in a production setting.
- Regular monitoring and retraining of the model with fresh data will be crucial to maintain its continuous accuracy and efficacy in minimizing churn.

REFERENCES

- 1] Y. -b. Chen, B. -s. Li and X. -q. Ge, "Study on Predictive Model of Customer Churn of Mobile Telecommunication Company," *2011 Fourth International Conference on Business Intelligence and Financial Engineering*, Wuhan, China, 2011, pp. 114-117, doi: 10.1109/BIFE.2011.112.
- 2] P. Bhuse, A. Gandhi, P. Meswani, R. Muni and N. Katre, "Machine Learning Based Telecom-Customer Churn Prediction," *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, Thoothukudi, India, 2020, pp. 1297-1301, doi: 10.1109/ICISS49785.2020.9315951.