

DATA ANALYST WITH COGNOS

TEAM MEMBER

411421104054: S.VIJAYALAKSHMI

Phase - 2 :Document Submission

Project : Customer churn prediction in telecommunication



Introduction :

Data analytics help the telecom industry boost profits through improved network services, security and customer experience. It helped the telecom companies to take sudden actions to prevent these behaviors in the market. Predictive Analytics helps in detecting telecommunications defects in real-time problem.. Predicting the future of telecommunications is essential for businesses, policymakers, and researchers to make informed decisions and investments. Implementing a telecommunication strategy to reduce churn involves a combination of data analysis, customer engagement, and targeted actions. The main collection of these strategies to reduce churn by Data collection and Analysis, Segmentation, Churn prediction Models, Customer feedback and surveys, customer retention offers, quality of service, Documentation and Team reporting, Training and team developments. Remember that reducing churn is an ongoing effort that requires a holistic approach involving data analysis, customer engagement, and continuous improvement of your telecommunication services. By implementing these strategies, we can work to retain more customers and build stronger, lasting relationships with them. Implementing a telecommunication prediction models involves various techniques from data science, Machine learning Models, Natural language Processing (NLP), Clustering and Segmentation, Deep reinforcement learning and Statistical analysis To forecast trends and outcomes in the telecommunication industry.

1. Data Source :

Telecommunication companies collect a data from a wide range of sources includes a networks equipment, call detail records, customer relationship management system, billing system, network monitoring tools .

Dataset link : (<https://www.kaggle.com/datasets/blastchar/telco-customer-churn>)



customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines
7590-VHVEG	Female	0	Yes	No	1	No	No phone service
5575-GNVDE	Male	0	No	No	34	Yes	No
3668-QPYBK	Male	0	No	No	2	Yes	No
7795-CFOCW	Male	0	No	No	45	No	No phone service
9237-HQITU	Female	0	No	No	2	Yes	No
9305-CDSKC	Female	0	No	No	8	Yes	Yes
1452-KIOVK	Male	0	No	Yes	22	Yes	Yes
6713-OKOMC	Female	0	No	No	10	No	No phone service
7892-POOKP	Female	0	Yes	No	28	Yes	Yes
6388-TABGU	Male	0	No	Yes	62	Yes	No
9763-GRSKD	Male	0	Yes	Yes	13	Yes	No
7469-LKBCI	Male	0	No	No	16	Yes	No
8091-TTVAX	Male	0	Yes	No	58	Yes	Yes
0280-XJGEX	Male	0	No	No	49	Yes	Yes
5129-JLPIS	Male	0	No	No	25	Yes	No
3655-SNQYZ	Female	0	Yes	Yes	69	Yes	Yes
8191-XWSZG	Female	0	No	No	52	Yes	No
9959-WOFKT	Male	0	No	Yes	71	Yes	Yes
4190-MFLUW	Female	0	Yes	Yes	10	Yes	No
4183-MYFRB	Female	0	No	No	21	Yes	No
8779-QRDMV	Male	1	No	No	1	No	No phone service
1680-VDCWW	Male	0	Yes	No	12	Yes	No
1066-JKSGK	Male	0	No	No	1	Yes	No
3638-WEABW	Female	0	Yes	No	58	Yes	Yes
6322-HRPFA	Male	0	Yes	Yes	49	Yes	No
6865-JZNKO	Female	0	No	No	30	Yes	No
6467-CHFZW	Male	0	Yes	Yes	47	Yes	Yes
8665-UTDHz	Male	0	Yes	Yes	1	No	No phone service
5248-YGIJN	Male	0	Yes	No	72	Yes	Yes

2. Model Evaluation and Prediction :

Model evaluation is a critical step in the telecommunication prediction process. It helps assess the performance and accuracy of the prediction models we've developed. It can be splitting the data, Selecting evaluation metrics and Some models are logistic regression, Decision Trees, Support vector machine (SVM), Neural Networks, Model Training and Validation, Visualization and plots, Comparative Analysis and Communication results. Effective model evaluation ensures that your telecommunication prediction models are accurate, reliable, and aligned with the goals of your organization.

Model Prediction Predictive modeling in telecommunication can be used for various purposes, such as predicting customer churn, forecasting network traffic, optimizing resource allocation, and improving customer satisfaction. It allows telecom companies to



make data-driven decisions and respond proactively to industry changes and customer needs.

3.Model 1 : Logistic Regression

Logistic regression is a widely used model for binary classification problems, such as customer churn prediction. It models the probability of a customer churning based on input features. Evaluate the logistic regression model's performance on the testing dataset using appropriate metrics. Common metrics for binary classification include accuracy, precision, recall, F1-score, and the ROC-AUC curve.

Model interpretation : Logistic regression models are interpretable, and you can examine the coefficients associated with each feature to understand their impact on the prediction. Positive coefficients indicate that an increase in the corresponding feature value increases the log-odds of churn, while negative coefficients decrease the log-odds of churn.

Deployment and Retraining: If the logistic regression model meets your performance criteria, we can deploy it in our environment to predict churn for new customers. Keep the model up-to-date by periodically retraining it with new data as customer interactions and churn events occur.

4.Model 2 : Decision Trees

Implementing a decision tree model for customer churn prediction involves several steps, including data preparation, model training, evaluation, and potentially deployment. Decision trees are interpretable models that can be used effectively for this task.

Model Evaluation : Evaluate the decision tree model's performance on the testing dataset using appropriate metrics for binary classification. Common metrics include accuracy, precision, recall, F1-score, and the ROC-AUC curve.

Model Interpretation: Decision tree models are interpretable, and you can visualize the tree structure to understand how the model makes decisions.

Deployment and Retraining: If the decision tree model meets your performance criteria, you can deploy it in your production environment to predict churn for new customers. Keep the model up-to-date by periodically retraining it with new data as customer interactions and churn events occur.

5.Model 3 : Neural networks

Implementing neural networks for customer churn prediction involves several steps, including data pre processing, model design, training, evaluation, and deployment.

Model Evaluation : Evaluate the neural network's performance on the testing data set using appropriate metrics for binary classification. Common metrics include accuracy, precision, recall, F1-score, and the ROC-AUC curve. Plot training and validation curves to assess over fitting and convergence.

Deployment and Retraining : If the neural network model meets your performance criteria, you can deploy it in our production environment to predict churn for new customers. Keep the model up-to-date by periodically retraining it with new data as customer interactions and churn events occur.



Model Interpretation : Deep learning models are often considered "black boxes" due to their complexity. To gain insights into feature importance, you can use techniques like SHAP (Shapley Additive explanations) to explain individual predictions. Implementing neural networks for customer churn prediction can be a powerful approach when you have a large and complex Data set .However, it also requires careful data pre processing, model design, and hyper parameter tuning to achieve

optimal results. Additionally, regular monitoring and retraining are essential to ensure that the model remains accurate over time.

Conclusion :

Customer churn prediction in the telecommunications industry is a critical and complex task that requires a combination of data analytics, machine learning, and a customer-centric approaches. In this prediction in the telecommunications industry is a multifaceted process that combines data analysis, machine learning, and a deep understanding of customer behavior. By using this techniques the telecommunications companies can reduce churn, improve customer satisfaction, and thrive in a highly competitive market.



