CUSTOMER CHURN PREDICTION

ABSTRACT

Customer churn prediction refers to the pra ctice of using data analysis and predictive modeling techniques to forecast which cust omers are likely to stop using a product or service, often referred to as "churning" or " churned customers." Churn prediction is a valuable business strategy, especially for s ubscription-based services, telecom compa nies, e-commerce platforms, and other bus inesses that rely on customer retention and loyalty.

PROBLEM DEFINITION

The project involves using IBM Cognos to predict customer churn and identify factors influencing customer retention. The goal is to help businesses reduce customer attrition by understanding the patterns and reasons behind customers leaving. This project includes defining analysis objectives, collecting customer data, designing relevant visualizations in IBM jCognos, and building a predictive model.

DESIGN THINKING

ANALYSIS OBJECTIVES

ANALYSIS OBJECTIVES

Define the specific objectives of predicting customer churn, such as identifying potential churners and understanding the key factors contributing to churn.

1.Identify Potential Churners

2.Early Detection

3.Reduce Churn Rate

1.The primary objective of churn prediction is to identify customers who are at risk of churning. This can be done by developing a predictive model that assigns a churn probability score to each customer

2.Aim to detect potential churners as early as possible. Early detection allows for proactive measures to be taken, such as targeted marketing campaigns or personalized incentives, to retain these customers.

3.Set a specific target for reducing the churn rate. This objective could be framed as a percentage reduction in churn over a specified time period (e.g., reduce churn by 10% in the next quarter).

4.Segmentation

5.Feature Analysis

6.Customer Lifetime Value (CLV)

4.Segment the customer base based on churn probabili ty and other relevant factors. This allows for tailored ret ention strategies for different customer groups. For exa mple, high-value customers may receive different retent ion efforts compared to low-value customers.

5.Understand the key factors contributing to churn. Con duct feature importance analysis to identify which custo mer attributes, behaviors, or interactions with the comp any have the most significant impact on churn.

6.Calculate CLV for each customer and analyze how it correlates with churn. The objective may be to increase the CLV of customers at risk of churning

7.Model Performance

8.Actionable Insights

9.Monitoring and Iteration

7.Set performance benchmarks for your churn prediction model. This includes metrics such as accuracy, precision, recall, and F1-score. Aim to achieve a certain level of model accuracy in predicting churn.

8.The ultimate goal is to provide actionable insights to the business. Ensure that your churn prediction analysis translates into specific actions that can be taken to retain customers. These actions may include sending targeted offers, improving customer service, or enhancing product features.

9.Implement a system for continuous monitoring of churn and model performance. Establish a process for regular model retraining and refinement to adapt to changing customer behaviors and market conditions.

10.Cost Reduction

11.Customer Feedback Integration:

12.Benchmarking

10.Evaluate the cost of customer acquisition compared to the cost of retaining customers. The objective may be to reduce the cost of retention efforts while maximizing their effectiveness.

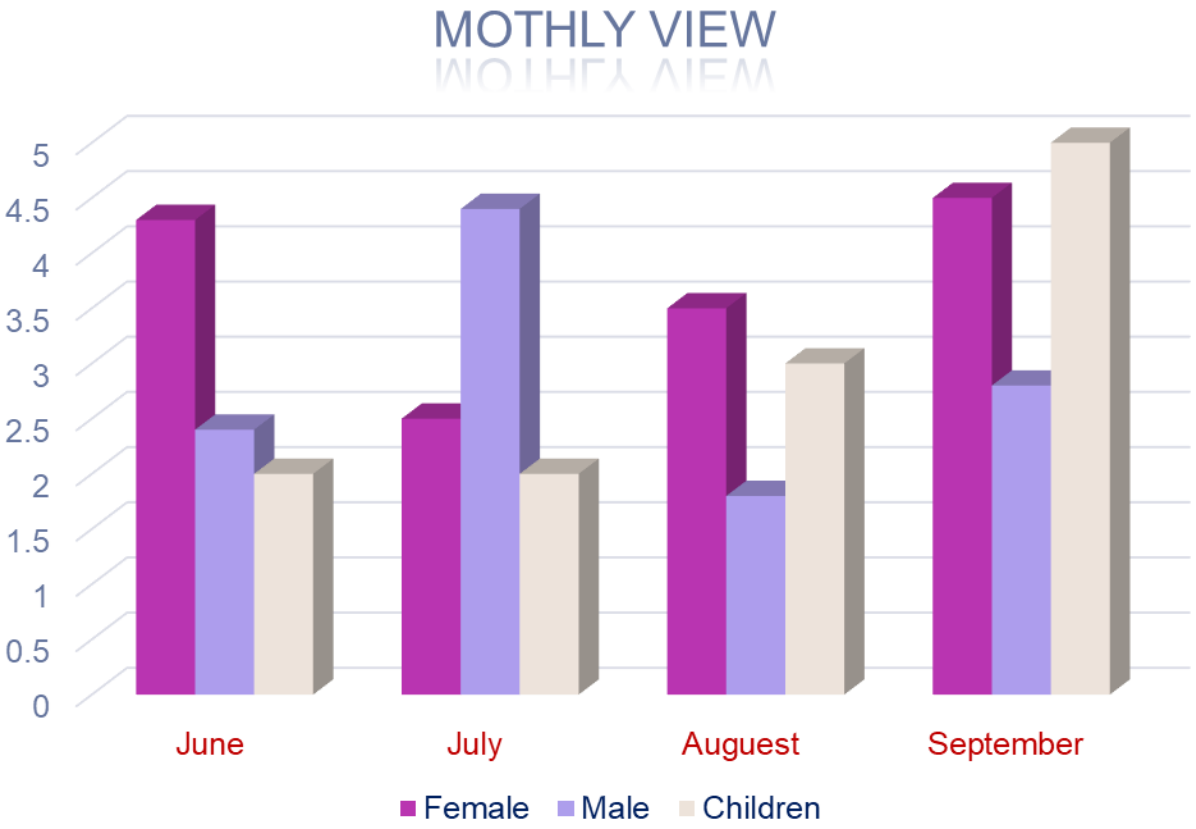
11.Integrate customer feedback into the churn prediction process. Identify the sentiment of customer feedback from potential churners and use it to refine retention strategies.

12.Compare your churn prediction and retention efforts with industry benchmarks or competitors to assess your performance and identify areas for improvement

DATA COLLECTION

Data Collection

Determine the sources and methods for collecting customer data, including customer demographics, usage behavior, and historical interactions.


MOTHLY VIEW
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
June
July
Auguest
September
Female
Male
Children

Methods for Collecting Customer Data:

1. Data Mining
2. Machine Learning Models
3. Third-party Data

1. Use data mining techniques to extract valuable insights from large datasets. This can help identify patterns and factors that contribute to customer churn.

2. Implement predictive models like logistic regression, decision trees, or neural networks to analyze historical data and predict future churn based on customer behavior and demographics.

3. Consider using external data sources, such as market data or industry benchmarks, to enhance your analysis and gain a broader perspective on customer behavior.

Visualization strategy

Visualization Strategy

Plan how to visualize the in sights using IBM Cognos, s howcasing factors affecting churn and retention rates fo r customer churn prediction project

1.Understand the Data

2.Choose the Right Visualizations

1.Start by thoroughly understanding your dataset and the variables that may customer churn and retention. Identify key features and potential predictors.

2.Select appropriate visualization types for different types of data. For Example: Use line charts to visualize trends in churn and retention rates over Time. Create bar charts or pie charts to represent categorical variables like product usage, demographics, or subscription Type. Scatter plots can be useful to explore relationships between variables.

DATASET

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	Strea
2	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	No	No	No	No
3	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No
4	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No
5	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	No	Yes	Yes	No	No
6	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	No	No	No	No	No
7	9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes
8	1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	Yes	No	No	Yes	No
9	6713-OKOMC	Female	0	No	No	10	No	No phone service	DSL	Yes	No	No	No	No	No
10	7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	No	Yes	Yes	Yes	Yes
11	6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL	Yes	Yes	No	No	No	No
12	9763-GRSKD	Male	0	Yes	Yes	13	Yes	No	DSL	Yes	No	No	No	No	No
13	7469-LKBCI	Male	0	No	No	16	Yes	No	No	No internet servic	No internet servic	No internet servic	No internet servic	No internet servic	No ir
14	8091-TTVAX	Male	0	Yes	No	58	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes
15	0280-XJGEX	Male	0	No	No	49	Yes	Yes	Fiber optic	No	Yes	Yes	No	Yes	Yes
16	5129-JLPIS	Male	0	No	No	25	Yes	No	Fiber optic	Yes	No	Yes	Yes	Yes	Yes
17	3655-SNQYZ	Female	0	Yes	Yes	69	Yes	Yes	Fiber optic	Yes	Yes	Yes	Yes	Yes	Yes
18	8191-XWSZG	Female	0	No	No	52	Yes	No	No	No internet servic	No internet servic	No internet servic	No internet servic	No internet servic	No ir
19	9959-WOFKT	Male	0	No	Yes	71	Yes	Yes	Fiber optic	Yes	No	Yes	No	Yes	Yes
20	4190-MFLUW	Female	0	Yes	Yes	10	Yes	No	DSL	No	No	Yes	Yes	No	No
21	4183-MYFRB	Female	0	No	No	21	Yes	No	Fiber optic	No	Yes	Yes	No	No	Yes
22	8779-QRDMV	Male	1	No	No	1	No	No phone service	DSL	No	No	Yes	No	No	Yes
23	1680-VDCWW	Male	0	Yes	No	12	Yes	No	No	No internet servic	No internet servic	No internet servic	No internet servic	No internet servic	No ir
24	1066-JKSGK	Male	0	No	No	1	Yes	No	No	No internet servic	No internet servic	No internet servic	No internet servic	No internet servic	No ir
25	3638-WEABW	Female	0	Yes	No	58	Yes	Yes	DSL	No	Yes	No	Yes	No	No
26	6322-HRPFA	Male	0	Yes	Yes	49	Yes	No	DSL	Yes	Yes	No	Yes	No	No
27	6865-JZNKO	Female	0	No	No	30	Yes	No	DSL	Yes	Yes	No	No	No	No
28	6467-CHFZW	Male	0	Yes	Yes	47	Yes	Yes	Fiber optic	No	Yes	No	No	Yes	Yes
29	8665-UTDHZ	Male	0	Yes	Yes	1	No	No phone service	DSL	No	Yes	No	No	No	No

DATA PREPROCESSING

VISUALIZATION

Check for missing values in each columns and decide how to handle them

Handle data types appropriately(eg.convert the ‘date’ column to datetime)

Ensure data consistency and correctness, such as checking that percentages are within valid Ranges(0-100%)

Develop informative and visually appealing charts And graphs

Consider creating interactive visualization for Online sharing or presentations

Ensure that your visualizations are well labled And easy to interpret

PREDICTIVE MODELING

Algorithms to predict customer churn prediction such as ensemble techniques

1. SVM - SVM or Support Vector Machine
2. Ridge Classifier
3. Random Forest
4. XG boost

About the algorithms

SVM - SVM or Support Vector Machine is a supervised machine learning technique used for classification and regression. Finding a hyperplane in an N-dimensional space that classifies the data points is the goal of the SVM method. The number of features determines the hyperplane's size.

Ridge Classifier - Ridge classification is a metho d used in machine learning to assess linear discri minant models. In order to prevent overfitting, thi s type of normalization limits model coefficients.

Random Forest - Random Forest is a classificatio n algorithm that uses multiple decision trees on s maller sets of the input dataset and averages the r esults to enhance the dataset's prediction accurac y.

XG Boost - Formally speaking, XGBoost ma y be described as a decision tree-based ens emble learning framework that uses Gradien t Descent as the underlying objective functio n. It offers excellent flexibility and efficiently uses computation to produce the mandated results.

Conclusion

In conclusion, customer churn prediction plays a pivotal role in helping businesses retain their customers. By leveraging data-driven models a nd analytics, companies can identify potential c hurners and take proactive measures to retain t hem. This not only helps in maintaining revenu e but also enhances customer satisfaction and l oyalty.