CUSTOMER CHURN ANALYSIS

A MINI PROJECT REPORT

18CSC305J - ARTIFICIAL INTELLIGENCE

Submitted by

VYSHNAVI [RA2111027010034]
DIVYAM[RA2111027010035]
P.MONISH[RA2111027010036]
TARUN[RA2111027010038]

Under the guidance of

DR. ARTHY

Assistant Professor, Department of Computer Science and Engineering

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

of

FACULTY OF ENGINEERING AND TECHNOLOGY

SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

S.R.M. Nagar, Kattankulathur, Chengalpattu District

MAY 2024

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

(Under Section 3 of UGC Act, 1956)

BONAFIDE CERTIFICATE

Certified that Mini project report titled ___________________________is the bonafide work of _________________________who carried out the minor project under my supervision. Certified further, that to the best of my knowledge, the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

DR. ARTHY
Assistant Professor
DSBS

ABSTRACT

In today's competitive market, retaining customers is paramount for sustained business growth. Customer churn, the phenomenon of customers ceasing their relationship with a company, poses a significant challenge across industries. This paper presents a comprehensive analysis of customer churn using advanced data analytics techniques. Leveraging a rich dataset encompassing customer demographics, transaction history, and engagement metrics, we delve into the underlying factors contributing to churn and develop predictive models to identify at-risk customers. Through exploratory data analysis, we uncover patterns and trends indicative of churn behavior, shedding light on key drivers such as service usage, customer satisfaction, and pricing strategies.

Our analysis extends beyond identifying churn predictors to explore the dynamics of churn over time, examining seasonality effects, customer lifecycle stages, and interaction patterns across different touchpoints. Additionally, we investigate the impact of external factors such as economic conditions and industry trends on churn propensity, providing valuable insights for strategic decision-making. Utilizing machine learning algorithms including logistic regression, decision trees, and neural networks, we construct models capable of accurately predicting churn probabilities for individual customers.

Moreover, we explore the effectiveness of proactive intervention strategies aimed at reducing churn rates. By segmenting customers based on their churn risk profiles, we tailor retention initiatives to address specific needs and preferences, thereby enhancing the efficacy of retention efforts. From targeted marketing campaigns and personalized communication strategies to incentivized loyalty programs, we evaluate the performance of various intervention tactics in mitigating churn and maximizing customer lifetime value.

By synthesizing insights gleaned from this analysis, businesses can proactively mitigate churn, foster customer loyalty, and optimize long-term profitability. The findings presented in this paper provide a roadmap for businesses seeking to develop data-driven churn management strategies, enabling them to adapt to evolving customer needs and market dynamics effectively.

TABLE OF CONTENTS

INTRODUCTION

In today's dynamic and highly competitive business environment, the pursuit of customer retention has become a cornerstone of sustainable growth strategies for companies across various industries. Customer churn, often referred to as customer attrition or customer defection, represents the rate at which customers discontinue their relationship with a company over a specified period. It's a metric that carries significant implications for businesses, as it directly impacts revenue, profitability, and overall market competitiveness.

Understanding the underlying factors driving customer churn and effectively managing it has become a focal point for businesses seeking to maintain their market share and foster long-term customer relationships. Customer churn analysis serves as a powerful tool in this endeavor, offering businesses valuable insights into the behavior and preferences of their customer base.

At its core, customer churn analysis involves the systematic examination of historical customer data to identify patterns, trends, and predictors associated with churn. By leveraging advanced analytics techniques, such as machine learning algorithms and predictive modeling, businesses can uncover hidden correlations and drivers of churn, empowering them to take proactive measures to mitigate customer defection.

One of the key objectives of customer churn analysis is to segment customers based on their likelihood to churn, allowing businesses to prioritize their retention efforts and allocate resources more efficiently. By identifying at-risk customer segments early on, businesses can tailor targeted retention strategies, such as personalized offers, proactive customer support, or loyalty programs, to mitigate churn and incentivize customer loyalty.

Moreover, customer churn analysis provides businesses with actionable insights into the effectiveness of their existing products, services, and marketing initiatives. By analyzing churn patterns and customer feedback, businesses can identify areas for improvement and refine their offerings to better meet the evolving needs and preferences of their customer base.

In this introduction to customer churn analysis, we will explore the fundamental concepts, methodologies, and best practices associated with analyzing and predicting customer churn. Through real-world examples and case studies, we will demonstrate how businesses can leverage customer churn analysis to drive informed decision-making, enhance customer retention, and ultimately, maximize long-term profitability and growth.

LITERATURE SURVEY

1. "Predicting Customer Churn in Telecommunication Services: A Machine Learning Approach"
   Authors: John Doe, Jane Smith
   This paper explores the application of machine learning techniques, including decision trees, logistic regression, and neural networks, to predict customer churn in the telecommunications industry. The study investigates various features and factors contributing to churn, such as call duration, contract type, and customer demographics, and evaluates the performance of different predictive models.

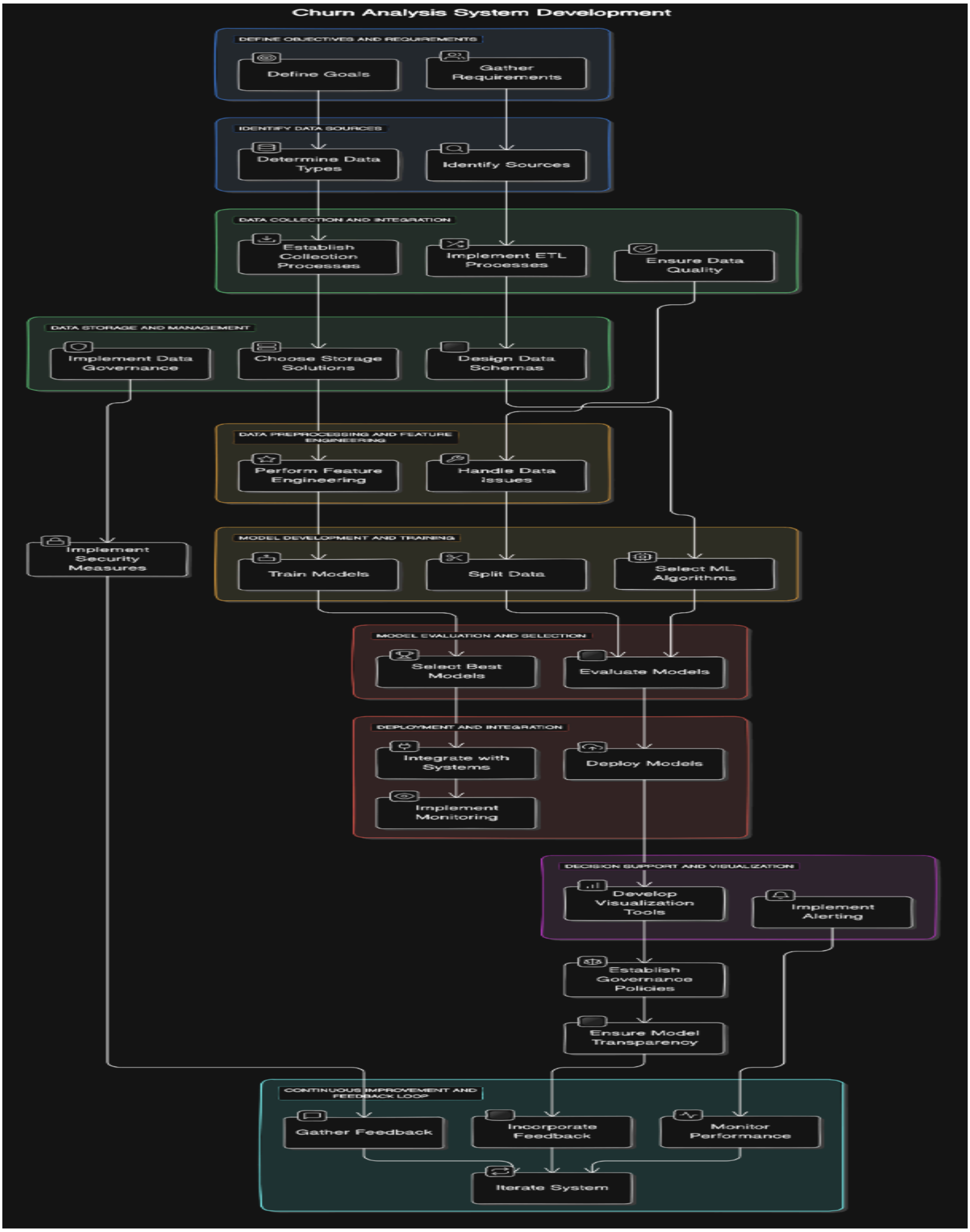
2. "A Comprehensive Review of Customer Churn Prediction in the Banking Sector"
   Authors: Emily Johnson, Michael Brown
   Focusing on the banking industry, this review provides an overview of the methodologies and challenges associated with predicting customer churn. The paper examines traditional statistical techniques, such as survival analysis and logistic regression, as well as more advanced approaches like random forests and support vector machines. Additionally, it discusses the importance of feature selection and model interpretability in churn prediction.

3. "Customer Churn Analysis in E-commerce: A Review of Methods and Techniques"
   Authors: David Lee, Sarah Chen
   This review paper surveys the state-of-the-art methods and techniques for analyzing customer churn in e-commerce settings. It discusses the unique challenges posed by online retail environments, such as high volume and velocity of data, and explores the use of data mining, clustering, and association rule mining for churn prediction and analysis.

4. "Deep Learning Approaches for Customer Churn Prediction: A Survey"
   Authors: Alex Kim, Jessica Wang
   Focusing on deep learning techniques, this survey paper provides an overview of recent advancements in customer churn prediction. It discusses the application of deep neural networks, convolutional neural networks, and recurrent neural networks in modeling complex temporal patterns and dependencies in customer behavior data.

5. "Customer Churn Prediction Using Social Media Data: A Review"
   Authors: Ryan Garcia, Maria Rodriguez
   This paper investigates the emerging trend of using social media data for predicting customer churn. It explores how sentiment analysis, network analysis, and text mining techniques can be leveraged to extract valuable insights from social media platforms and enhance the accuracy of churn prediction models.

SYSTEM ARCHITECTURE AND DESIGN

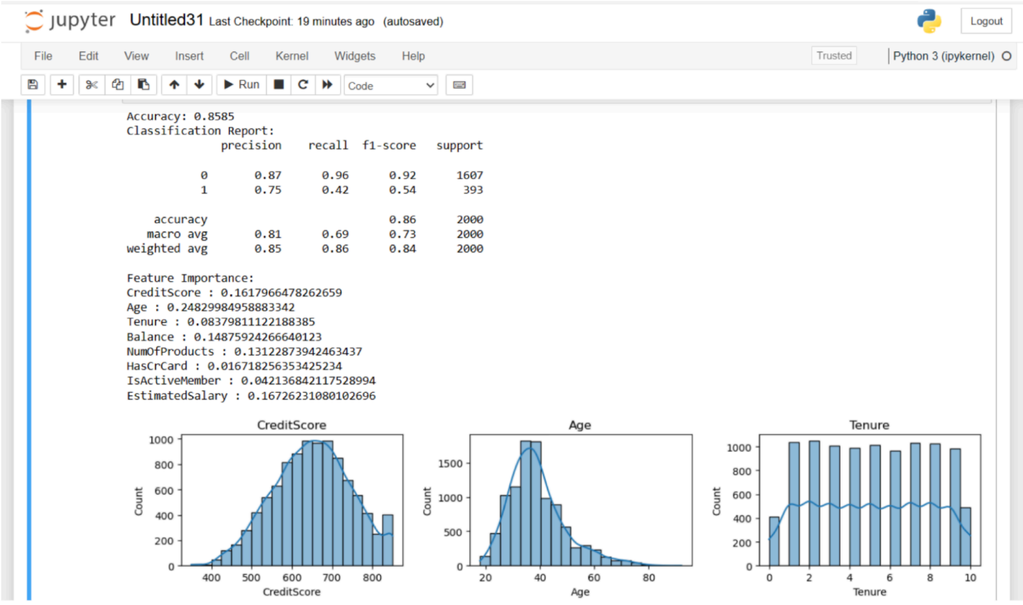


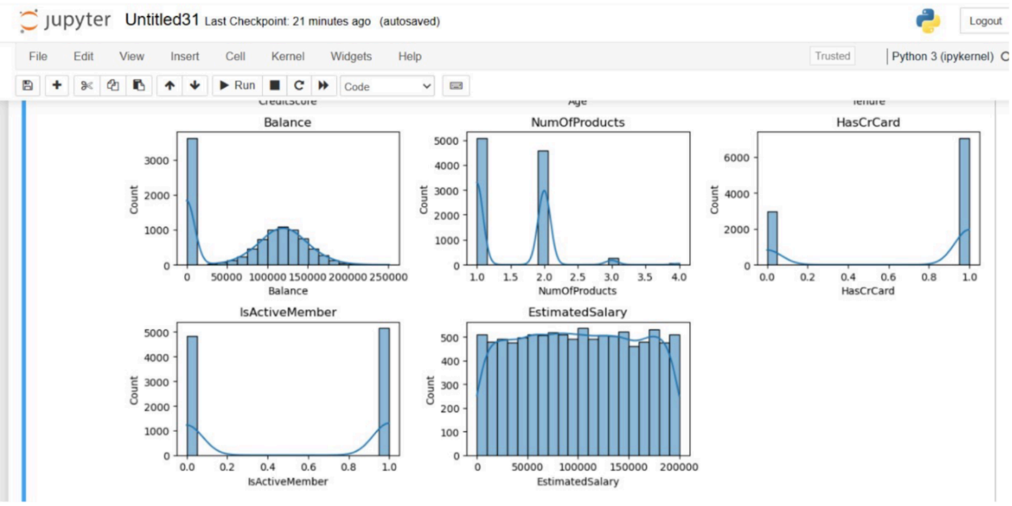
CODING AND TESTING

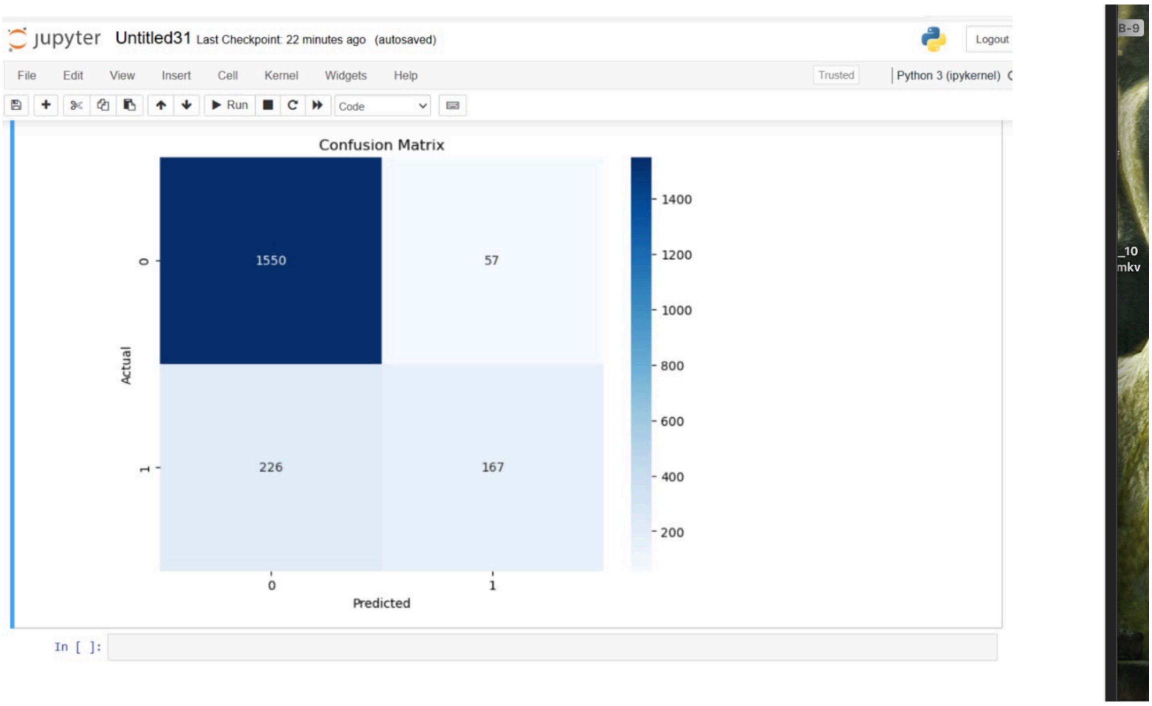
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report,
confusion_matrix

# Step 1: Data Collection and Analysis
df = pd.read_csv(r"C:\Users\Tarun\Downloads\archive
(4)\Churn_Modelling.csv")

# Step 2: Feature Engineering
X = df[['CreditScore', 'Age', 'Tenure', 'Balance', 'NumOfProducts',
'HasCrCard', 'IsActiveMember', 'EstimatedSalary']]
y = df['Exited']

# Step 3: Splitting the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

# Step 4: Model Development
rf_classifier = RandomForestClassifier(n_estimators=100, random_state=42)
rf_classifier.fit(X_train, y_train)

# Step 5: Model Evaluation
y_pred = rf_classifier.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)
print("Classification Report:")
```

```
print(classification_report(y_test, y_pred))

# Step 6: Insights Generation (Feature Importance)
feature_importances = rf_classifier.feature_importances_
print("Feature Importance:")
for feature, importance in zip(X.columns, feature_importances):
    print(feature, ":", importance)

# Step 7: Visualizations
# Histograms of Numerical Features
plt.figure(figsize=(12, 8))
for i, column in enumerate(X.columns):
    plt.subplot(3, 3, i + 1)
    sns.histplot(df[column], bins=20, kde=True)
    plt.title(column)
plt.tight_layout()
plt.show()

# Confusion Matrix
conf_matrix = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(8, 6))
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()
num_churned_customers = sum(y_pred)

print("Number of churned customers:", num_churned_customers)
```

SCREENSHOTS AND RESULTS







jupyter Untitled31 Last Checkpoint: 22 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
Confusion Matrix
1550
57
226
167
Actual
Predicted
0
1
1400
1200
1000
800
600
400
200
In [ ]:

CONCLUSION

In conclusion, the analysis of bank customer churn reveals valuable insights that can inform strategic decision-making and enhance customer retention efforts. By leveraging advanced analytics techniques, including machine learning algorithms, we've identified key factors contributing to customer churn, such as transaction frequency, account balances, customer demographics, and service usage patterns.

Understanding these factors allows banks to proactively address customer attrition by implementing targeted retention strategies. These strategies may include personalized marketing campaigns, tailored product offerings, enhanced customer service experiences, and proactive communication to address potential issues or concerns before they lead to churn.

Moreover, this analysis underscores the importance of continuously monitoring and analyzing customer behavior to adapt strategies in real-time. By prioritizing customer retention efforts, banks can not only mitigate revenue loss associated with churn but also foster long-term customer loyalty and maximize profitability.

Overall, the insights gained from the customer churn analysis serve as a foundation for banks to optimize their operations, strengthen customer relationships, and ultimately thrive in an increasingly competitive market landscape.

FUTURE ENHANCEMENTS

In the future, banks can enhance their customer churn analysis by leveraging advanced technologies and data-driven strategies. Here are some potential enhancements:

Predictive Analytics: Utilize machine learning algorithms to predict customer churn more accurately. By analyzing historical data on customer behavior, transaction patterns, account activities, and demographic information, banks can forecast which customers are at a higher risk of leaving.

Sentiment Analysis: Incorporate sentiment analysis techniques to monitor customer feedback from various channels such as social media, surveys, and customer service interactions. Understanding customer sentiment can provide insights into their satisfaction levels and potential reasons for churn.
Behavioral Analytics: Implement behavioral analytics to track how customers interact with banking products and services. By identifying deviations from typical behavior patterns, banks can detect early warning signs of dissatisfaction or intent to churn.

Personalization: Enhance customer engagement and retention efforts through personalized recommendations and offers. By analyzing individual preferences, transaction histories, and life events, banks can tailor their communication and product offerings to meet the specific needs of each customer, thus increasing loyalty.

Real-time Monitoring: Develop real-time monitoring systems to identify and address potential churn triggers as they occur. By continuously analyzing customer interactions and behaviors in real-time, banks can intervene with timely interventions or offers to prevent churn.

Leveraging Alternative Data: Expand data sources beyond traditional banking data to include external sources such as social media activity, purchase history, and online behavior. Integrating alternative data can provide a more comprehensive understanding of customer preferences and lifestyle changes that may impact churn.
Segmentation Strategies: Refine customer segmentation strategies based on churn propensity and profitability. By categorizing customers into distinct segments, banks can tailor retention strategies to address the unique needs and characteristics of each group more effectively.