


Customer Churn Prediction

The background is a dark blue gradient. It is decorated with various geometric elements: thin white vertical lines of varying lengths, small squares in teal, orange, and pink, and larger squares in teal and orange. Some of these shapes are solid, while others are outlined in white. They are scattered across the frame, creating a modern, tech-oriented aesthetic.



Table of Contents

Introduction	Overview
Data Visualization	Charts and Graphs
Modeling	Algorithms & Evaluating the Models
Next Steps	Future Outlook
Streamlit App	App Demo
Thank You	Q&A Session



Introduction

Overview

01

Goal

- Develop a model for the bank to determine customer churn
- Utilize the customer's credit history
- Train classification algorithms and determine performance
- Create a simple website application



Overview

A machine learning project that uses classification algorithms to predict customer attrition for a bank.

- Kaggle
- The model is trained on features that include but is not limited to...
 - Age - Customer's Age
 - Months on Books - How many months with the bank
 - Total Transaction Amount - Total transactions made in the last 12 months

Data Visualization

Gain insights of the data

02

Tasks



Revolving Balance &
Education

Anastasiya



Marital Status Distribution

Karandeep



Card Category vs.
Customer Age

Nisha



Heatmap

Iskandar



Distribution of Total
Transactions

Felix

Modeling

Key metrics and evaluation
techniques

03

Process

Cleaned data, handled missing values, and converted categories into numbers.

**Data
Prep**

**Model
Training**

Used the cleaned data to train three models:

- Logistic Regression
- Decision Tree
- Random Forest

Tested each model to see how well they predict customer churn.

**Model
Testing**

**Model
Analysis**

Used confusion matrices and classification reports to identify strengths and weaknesses.

Key Metrics (Accuracy)

92.8%

Decision Tree

Asks a series of yes/no questions to determine the best splits.

88.3%

Logistic Regression

Calculates the probability that a customer will stay or leave.

95.9%

Random Forest

Trained on many decision trees to improve predictions by taking the majority vote.

Key Metrics (Confusion Matrix)

	Attrited	Existing
Attrited	384	112
Existing	50	2493

Next Steps

Future outlook

04

UPCOMING GOALS

Solve Class Imbalance

Top Priority



Find a solution to solve the imbalance of class.

Train a LightGBM

Medium Priority



Train a different model and examine the performance.

Feature Engineering

Low Priority



Discovering other features that impact the model.

ht

App Demo

<https://mldschnprediction.streamlit.app/>



Do you have any questions?

THANK YOU