Customer Churn Prediction System

# Project Overview

Customer churn prediction system ek predictive analytics tool hai jo businesses ko help karta hai apne customers ke churn (service ya product ko chhodne) ko predict karne mein. Is system ke zariye, businesses customers ko retain karne ke liye effective strategies develop kar sakte hain.

# 1. Data Collection

Description:

Data collection ke stage mein, relevant data gather kiya jata hai jo customer behavior aur churn se related hota hai.

# Sources:

- Publicly available datasets (e.g., Kaggle's Telco Customer Churn dataset).

Key Features:

- Customer ID

- Gender

- Age

- Services availed (e.g., internet, phone)

- Monthly charges

- Churn status (1 = Yes, 0 = No)

## 2. Data Preprocessing

Description:

Data preprocessing ke process mein raw data ko clean aur transform kiya jata hai taaki machine learning algorithms ke liye use kiya ja sake.

Steps:

- \*\*Data Cleaning:\*\* Missing values aur duplicates ko remove karein.

- \*\*Feature Encoding:\*\* Categorical features (jaise gender, services) ko encode karein.

- \*\*Normalization:\*\* Numerical features ko normalize karein taaki model performance improve ho.

## 3. Exploratory Data Analysis (EDA)

\*\*Description:\*\*

EDA mein data ke patterns aur insights ko samajhne ke liye visualizations aur statistical techniques ka istemal kiya jata hai.

\*\*Techniques:\*\*

- Histograms, box plots, aur scatter plots ka istemal karke churn rates ka analysis karein.

- Demographic analysis ke liye bar graphs aur pie charts banayein.

## 4. Feature Engineering

\*\*Description:\*\*

Feature engineering ke stage mein new features create kiye jaate hain jo model ki predictive power ko enhance karte hain.

\*\*Key Features:\*\*

- Average tenure

- Total charges

- Services availed count

## 5. Model Development

\*\*Description:\*\*

Model development mein machine learning algorithms ka istemal karke churn prediction model banaya jata hai.

\*\*Steps:\*\*

- \*\*Model Selection:\*\* Logistic Regression, Random Forest, ya Gradient Boosting algorithms ka selection karein.

- \*\*Model Training:\*\* Data ko training aur testing sets mein divide karein aur model ko train karein.

- \*\*Evaluation Metrics:\*\* Accuracy, Precision, Recall, F1-score, aur ROC-AUC curves ka istemal karke model ki performance evaluate karein.

## 6. Database Integration

\*\*Description:\*\*

Database integration ke through, model predictions ko ek SQL database mein store kiya jata hai taaki unhe easily access kiya ja sake.

\*\*Steps:\*\*

- SQL database create karein.

- Model predictions ko SQL database mein insert karne ke liye SQL queries likhein.

## 7. Deployment

\*\*Description:\*\*

Deployment ke stage mein web application banaya jata hai jahan users churn predictions dekh sakte hain.

\*\*Steps:\*\*

- \*\*Web Application Development:\*\* Flask ya Django ka istemal karke web application develop karein.

- \*\*Azure Deployment:\*\* Application ko Azure Web App ya Azure App Service par deploy karein.

## 8. Visualization Dashboard

\*\*Description:\*\*

Dashboard create karne se users ko churn predictions aur customer insights ko visualize karne mein madad milti hai.

\*\*Tools:\*\*

- Power BI ya Tableau ka istemal karke dashboard banayein.

\*\*Features:\*\*

- Interactive filters aur charts provide karein taaki users data ko analyze kar sakein.

Technology Stack

#### 1. Project Overview

\*\*Project Name:\*\* ChurnGuard

\*\*Description:\*\* Yeh project customers ke churn ko predict karne ke liye AI/ML techniques ka istemal karta hai. Isse businesses ko apne customers ko retain karne mein madad milegi.

#### 2. Technology

- \*\*Programming Language:\*\* Python

- \*\*Libraries:\*\*

- Pandas

- NumPy

- Scikit-learn

- Matplotlib

- Seaborn

- SQLAlchemy

- \*\*Database:\*\* SQLite / PostgreSQL / MySQL

- \*\*Web Framework:\*\* Flask / Django

---

#### 3. Data Collection

\*\*Description:\*\*

Yeh step data ko collect karne aur cleaning karne ka hota hai.

\*\*Tools/Technologies:\*\*

- \*\*Pandas:\*\* Data manipulation ke liye.

\*\*Implementation Example:\*\*

```python

import pandas as pd

# Data load karna

df = pd.read\_csv('customer\_data.csv')

```

---

#### 4. Data Preprocessing

\*\*Description:\*\*

Isme data ko clean kiya jata hai, missing values ko handle kiya jata hai, aur features ko encode kiya jata hai.

\*\*Tools/Technologies:\*\*

- \*\*Pandas\*\*

- \*\*Scikit-learn\*\*

\*\*Implementation Example:\*\*

```python

# Missing values ko handle karna

df.fillna(0, inplace=True)

# Categorical features ko encode karna

from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()

df['Category'] = le.fit\_transform(df['Category'])

```

---

#### 5. Exploratory Data Analysis (EDA)

\*\*Description:\*\*

Data ka visualization aur insights generate karna.

\*\*Tools/Technologies:\*\*

- \*\*Matplotlib\*\*

- \*\*Seaborn\*\*

\*\*Implementation Example:\*\*

```python

import seaborn as sns

import matplotlib.pyplot as plt

sns.countplot(x='Churn', data=df) # Churn status ka count plot

plt.show()

```

---

#### 6. Feature Engineering

\*\*Description:\*\*

Existing features se naye features create karna.

\*\*Implementation Example:\*\*

```python

# Total charges ka feature create karna

df['TotalCharges'] = df['MonthlyCharges'] \* df['tenure']

```

---

#### 7. Model Development

\*\*Description:\*\*

Machine learning model ko train aur evaluate karna.

\*\*Tools/Technologies:\*\*

- \*\*Scikit-learn\*\*

\*\*Implementation Example:\*\*

```python

from sklearn.model\_selection import train\_test\_split

from sklearn.ensemble import RandomForestClassifier

X = df.drop('Churn', axis=1) # Features

y = df['Churn'] # Target variable

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

model = RandomForestClassifier()

model.fit(X\_train, y\_train)

```

---

#### 8. Database Integration

\*\*Description:\*\*

Model predictions ko database mein store karna.

\*\*Tools/Technologies:\*\*

- \*\*SQLAlchemy\*\*

\*\*Implementation Example:\*\*

```python

from sqlalchemy import create\_engine

engine = create\_engine('sqlite:///churn\_data.db') # Database connection

df.to\_sql('churn\_predictions', con=engine, if\_exists='replace', index=False)

```

---

#### 9. Deployment

\*\*Description:\*\*

Web application banakar predictions ka access dena.

\*\*Tools/Technologies:\*\*

- \*\*Flask\*\*

\*\*Implementation Example:\*\*

```python

from flask import Flask, request, jsonify

app = Flask(\_\_name\_\_)

@app.route('/predict', methods=['POST'])

def predict():

data = request.json # JSON data receive karein

prediction = model.predict(data)

return jsonify(prediction)

```

---

#### 10. Visualization Dashboard

\*\*Description:\*\*

User-friendly dashboard create karna jahan predictions dikhaye ja sakein.

\*\*Tools/Technologies:\*\*

- \*\*Dash / Tableau / Power BI\*\*

---

#### Conclusion

\*\*ChurnGuard\*\* project ka aim hai customers ke churn ko accurately predict karna aur businesses ko insights provide karna. Is documentation se aapko project ke har aspect ko samajhne mein madad milegi.

---

### Notes

- Documentation ko regularly update karein jaise project mein naye features add hote hain.

- Example codes ko apne project ke context ke hisaab se modify karein.

Is tarike se aap apne project ki documentation bana sakte hain. Agar aapko aur koi specific section add karna ho, toh mujhe batayein!