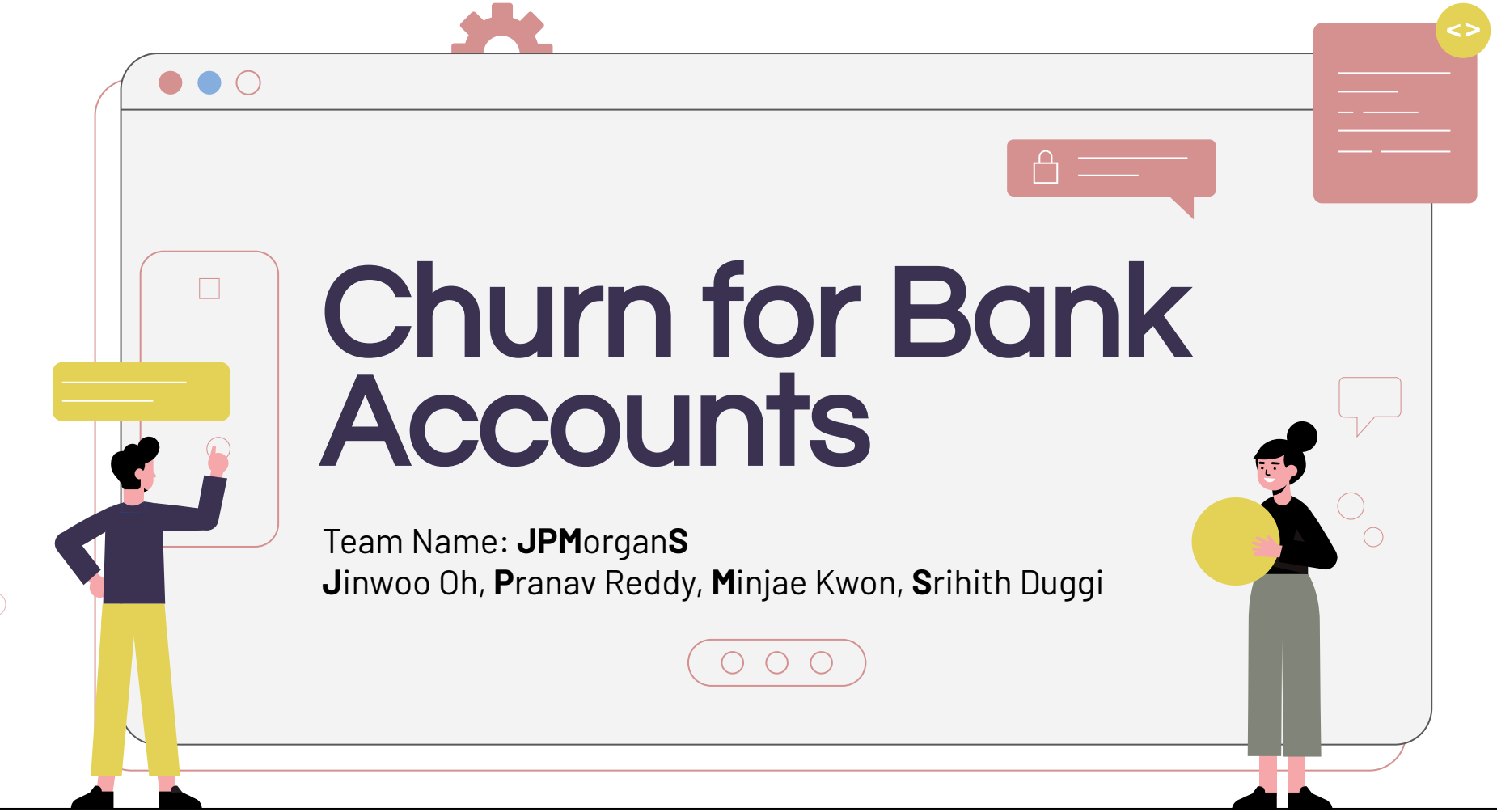


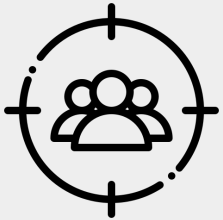
Churn for Bank Accounts

Team Name: **JPM**organ**S**
Jinwoo Oh, **P**ranav Reddy, **M**injae Kwon, **S**rihith Duggi

Goal / Audience / Purpose / Context

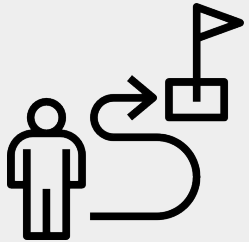
Goal

Predicting if customers' likely to turn over or not based on their features



Audience

Banks that are experiencing decreasing number of customers



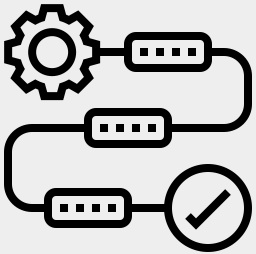
Purpose

Better for the banks to know why the customers are leaving their bank

Context

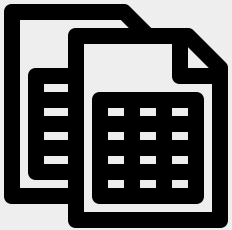
Companies are constantly looking for ways to reduce churn rate

Method & Dataset



Method

- Data curation, cleaning, ETL, visualisation
- Model coding and selection : MLP Classifier
- Model fairness assessment and interpretation (LIME)
- Evaluate the result and interpret it so that it can be used to help with the decisions of various banks



Dataset

- From kaggle.com (https://www.kaggle.com/datasets/mathchi/churn-for-bank-customers)
- 14 columns / about 10,000 rows
- The three types of dataset: Numerical, Categorical, and Binary
- Mainly using: CreditScore, Age, EstimatedSalary, Balance, Gender, IsActiveMember, and Exited

Dataset

RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited
1	15634602	Hargrave	619	France	Female	42	2	0	1	1	1	101348.88	1
2	15647311	Hill	608	Spain	Female	41	1	83807.86	1	0	1	112542.58	0
3	15619304	Onio	502	France	Female	42	8	159660.8	3	1	0	113931.57	1
4	15701354	Boni	699	France	Female	39	1	0	2	0	0	93826.63	0
5	15737888	Mitchell	850	Spain	Female	43	2	125510.82	1	1	1	79084.1	0
6	15574012	Chu	645	Spain	Male	44	8	113755.78	2	1	0	149756.71	1
7	15592531	Bartlett	822	France	Male	50	7	0	2	1	1	10062.8	0
8	15656148	Obinna	376	Germany	Female	29	4	115046.74	4	1	0	119346.88	1
9	15792365	He	501	France	Male	44	4	142051.07	2	0	1	74940.5	0
10	15592389	H?	684	France	Male	27	2	134603.88	1	1	1	71725.73	0

Progress & Challenges

Progress	• Determining which data set to use • Determining which features we are going to use : Credit Score, Gender, Age, Tenure, Balance, NumOfProducts, etc • Determining which column is going to be our label : Excited • Determining which Classifier we are going to use: MLP Classifier
Challenges	• We tested the MLP Classifier's accuracy and got 79.6 %, which is pretty low. Need to think about how to increase accuracy. • Need to identify what prejudices can arise. • Need to make a test data set so that we can predict for a single instance