الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Churn Prediction Project

By: G-1 Team

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

G-1 Team Members

- Ammar Alhawashem
- Maram Alzahrani
- Noof Alsafi
- Taif Alzahr
- Sara
- Noura Alajmi
- Hibah Sindi
- Shahad Ali
- Reema Almeshal
- Raya

Outlines:

- Project Overview
- Data Analysis and Visualization
- Clustering
- Predictive Model
  - SVM model
  - RF model
  - XG Boost model
  - Multi Layer perceptrons
- Model Deployment
- Conclusion

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Project Overview

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Problem Statement

The project involves predicting employee churn using HR data, conducting exploratory data analysis, preprocessing, clustering analysis and employing classification algorithms, and deploying the best model via Streamlit

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Goal

Develop a robust predictive model leveraging HR data to accurately forecast employee churn, enabling companies to proactively address retention strategies and reduce attrition rates

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Dataset

The dataset comprises various HR-related attributes for company employees, including satisfaction levels, performance evaluations, project assignments, monthly working hours, and tenure

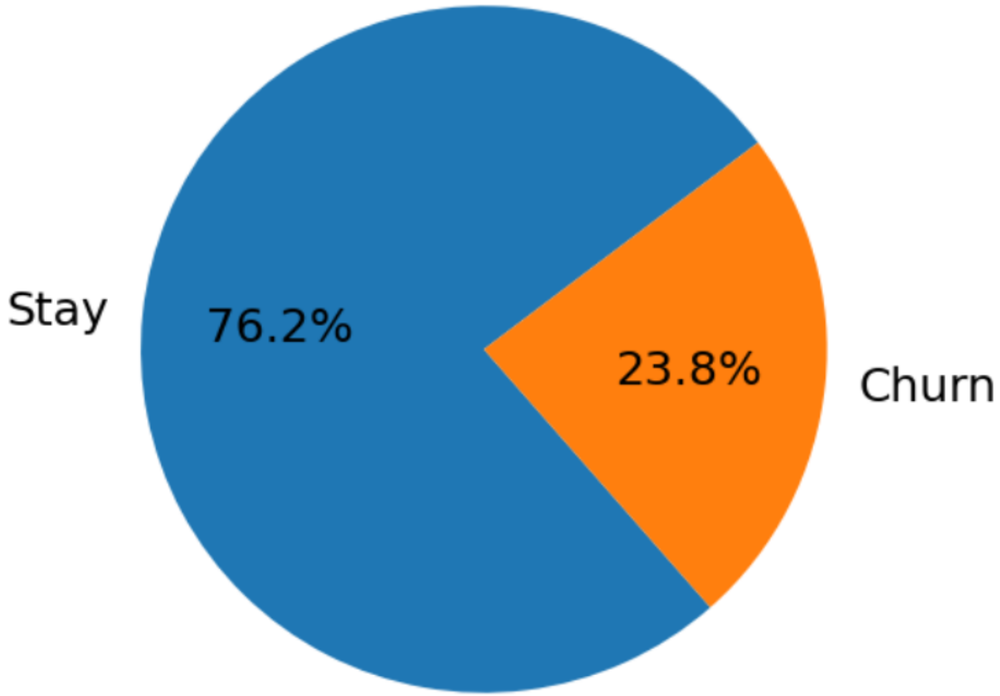
ıtisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	left	promotion_last_5years	Departments	salary
0.38	0.53	2	157	3	0	1	0	sales	low
0.80	0.86	5	262	6	0	1	0	sales	medium
0.11	0.88	7	272	4	0	1	0	sales	medium
0.72	0.87	5	223	5	0	1	0	sales	low
0.37	0.52	2	159	3	0	1	0	sales	low

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

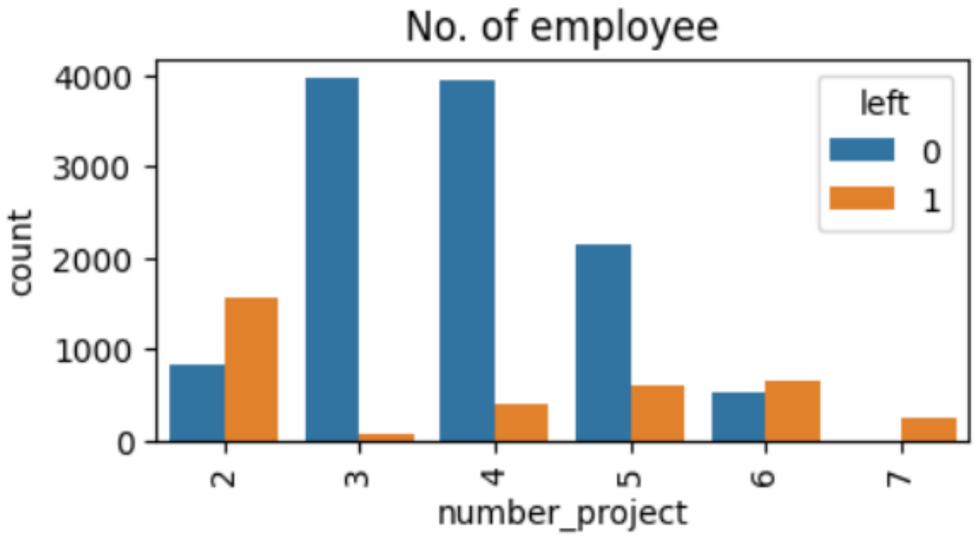
CLARUSWAY
WAY TO REINVENT YOURSELF

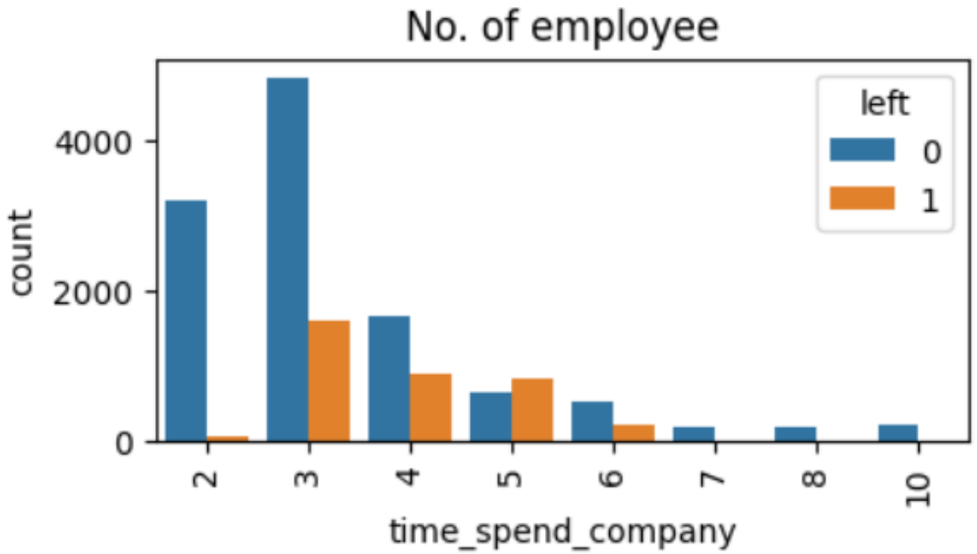
Data Analysis and Visualization

Employees churn distribution

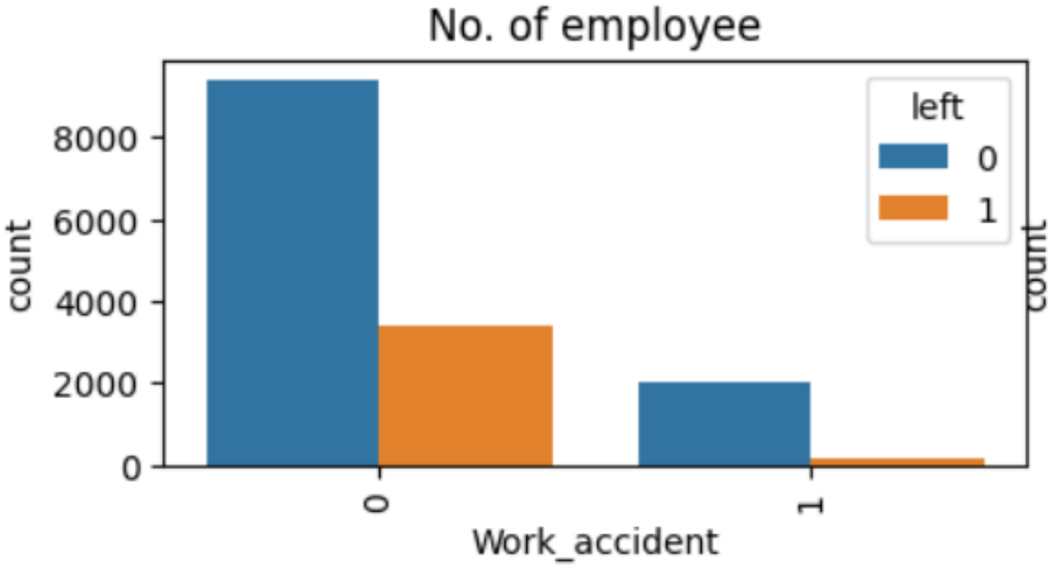


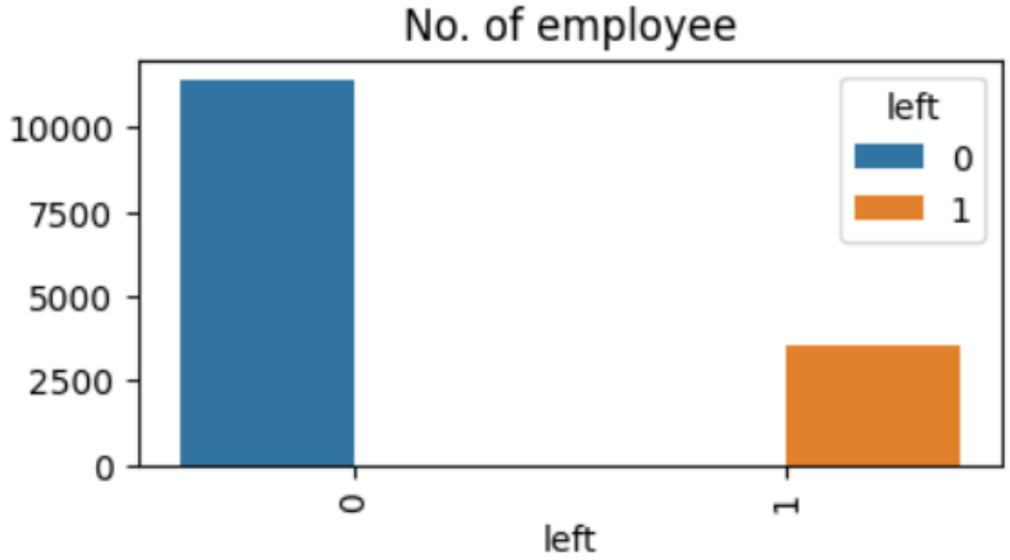
Employee Departure Analysis Across Factors



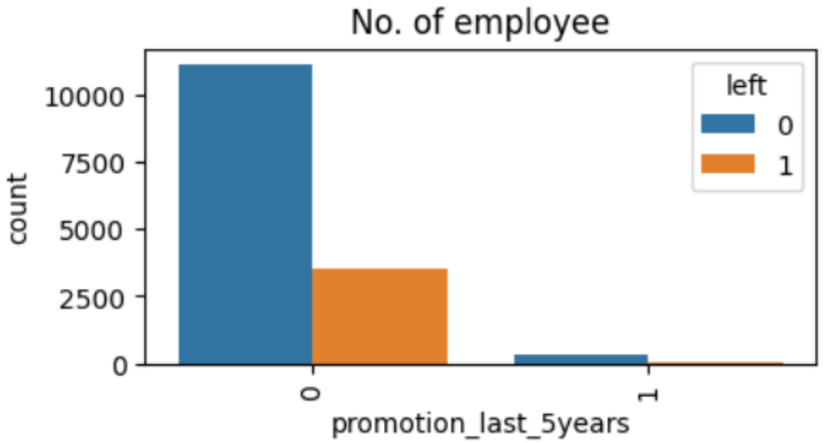


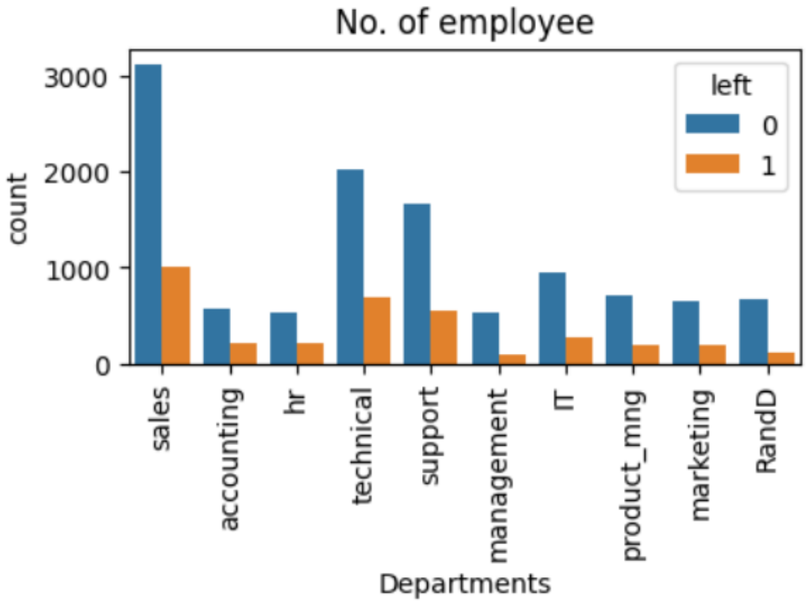
Employee Departure Analysis Across Factors



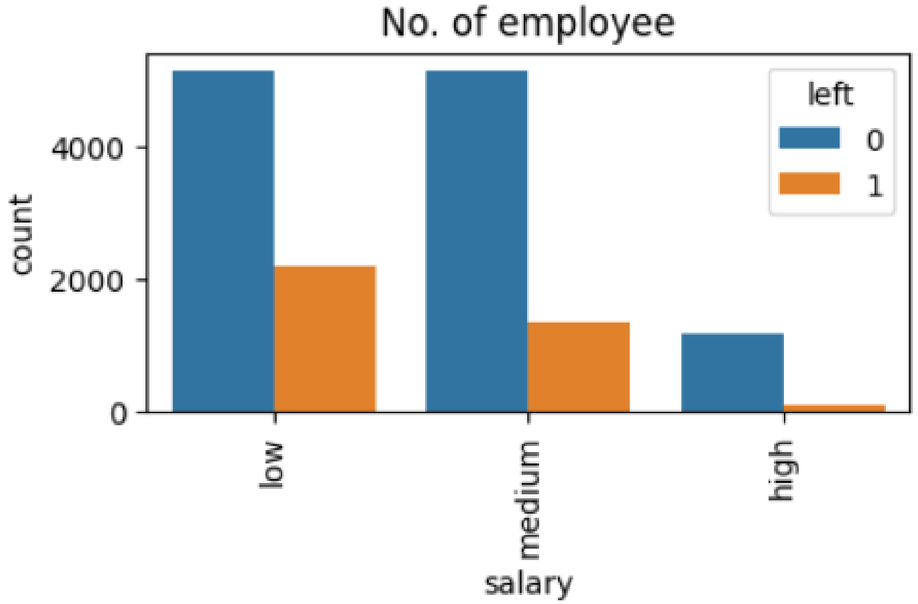


Employee Departure Analysis Across Factors





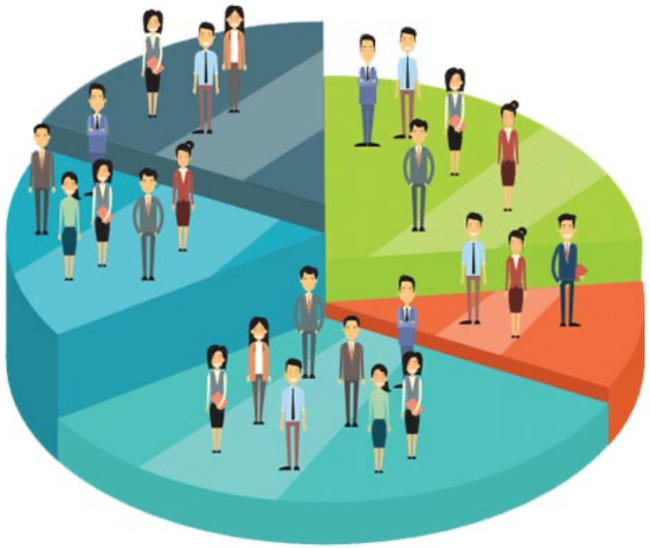
Employee Departure Analysis Across Factors



الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

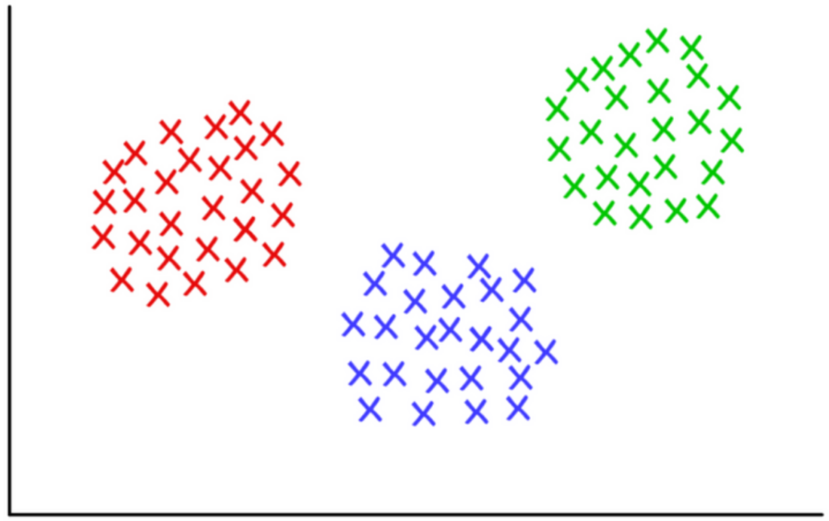
CLARUSWAY
WAY TO REINVENT YOURSELF

Clustering

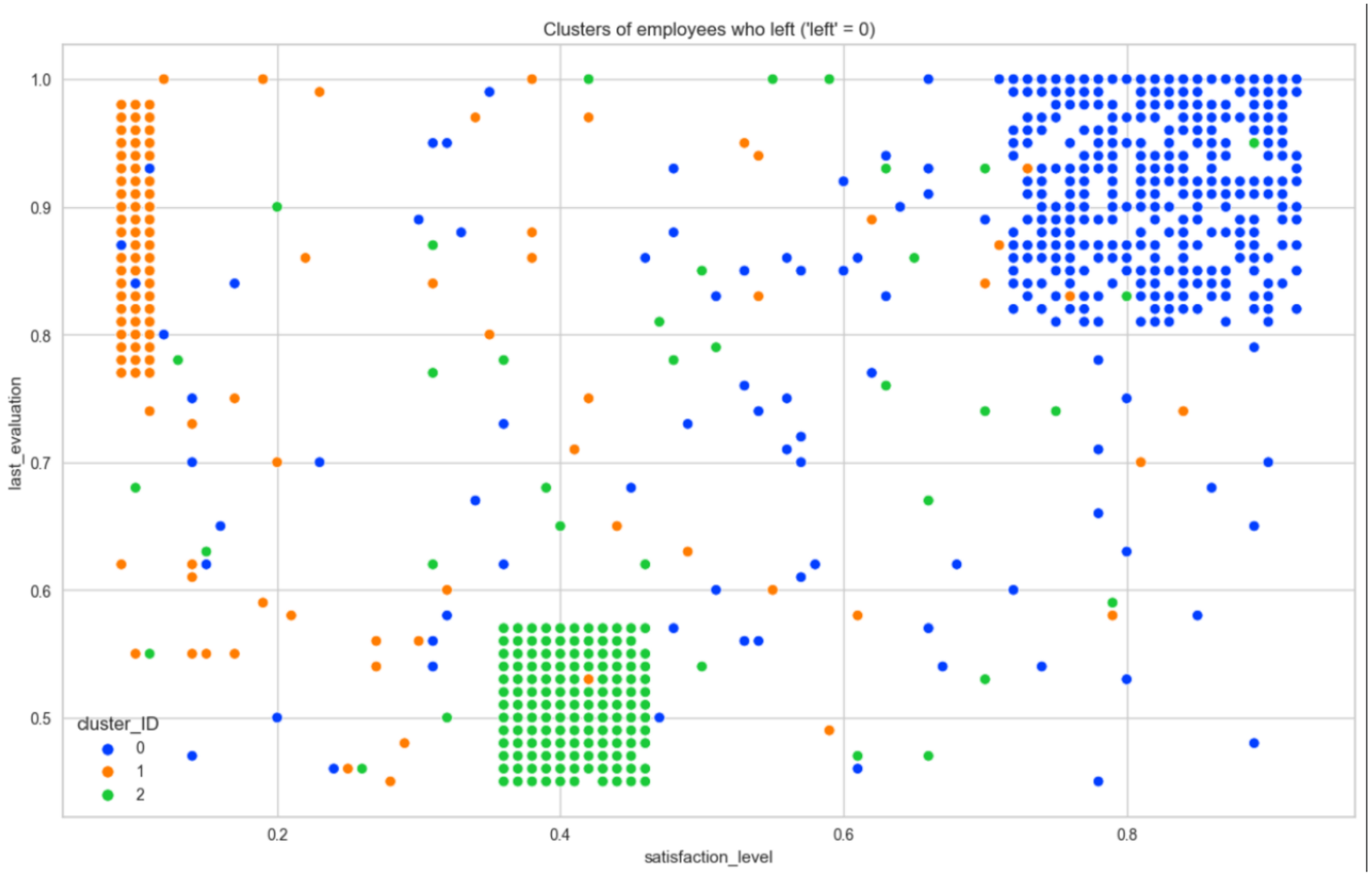


What is clustering?

Grouping similar data points together based on their inherent characteristics or patterns.



clustering results



الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Predictive Model

Predictive Model Building

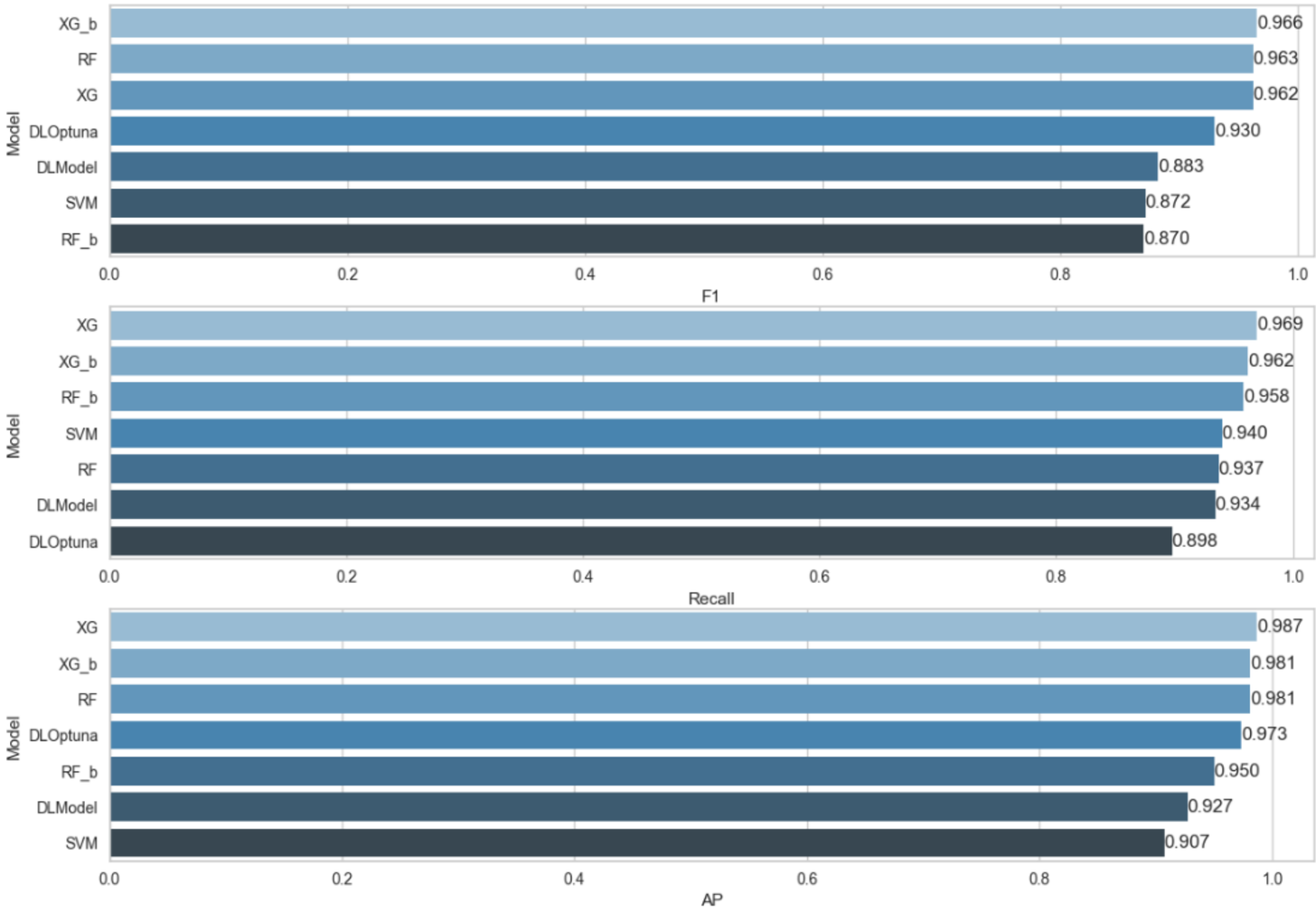
Classification algorithms we applied:

1. Support Vector Machine (SVM)
2. Random Forest (Bagging)
3. XGBoost (Boosting)
4. Multi Layer perceptrons (MLP)

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

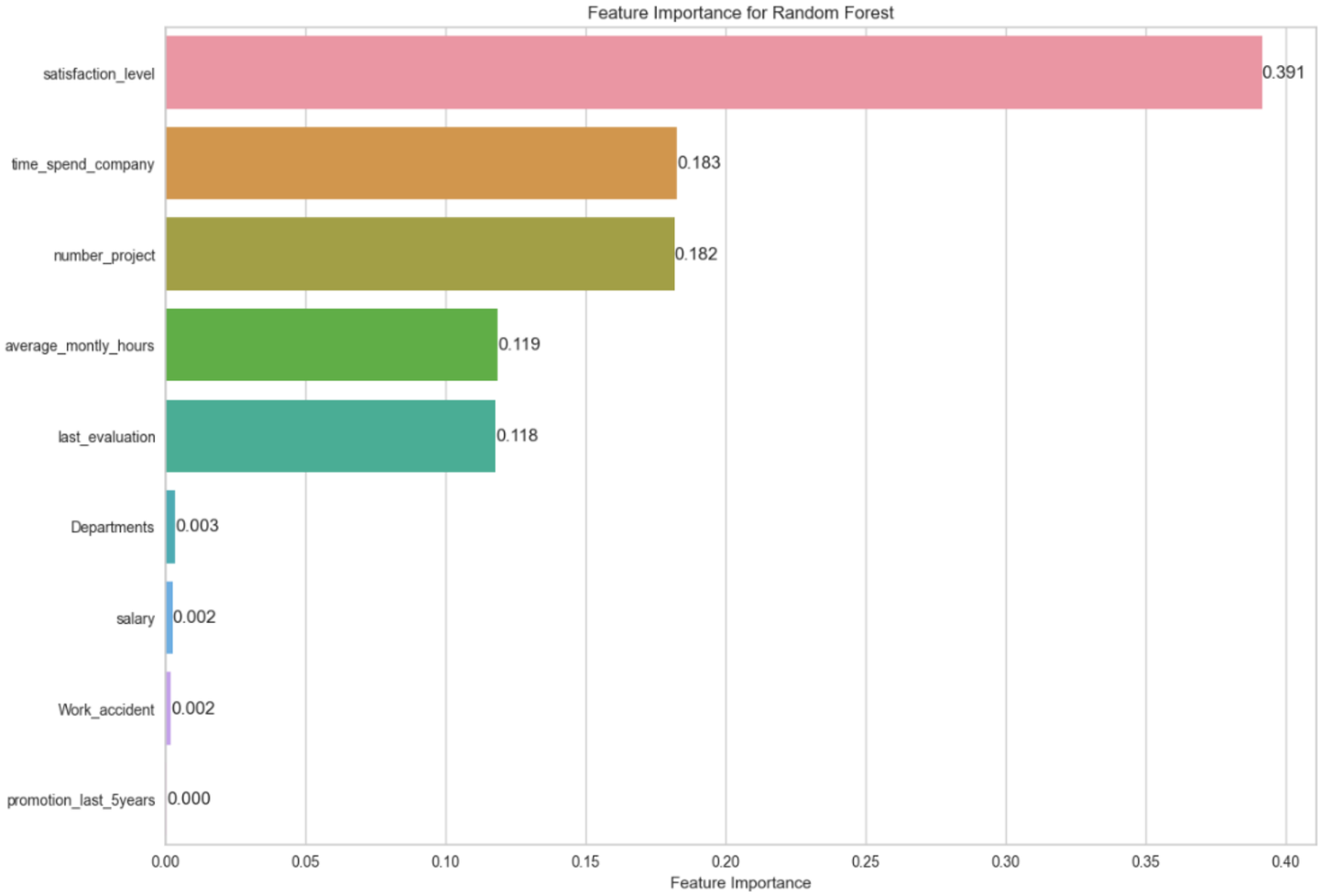
Predictive Models Results Comparison



الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Features Importance for the Best Model



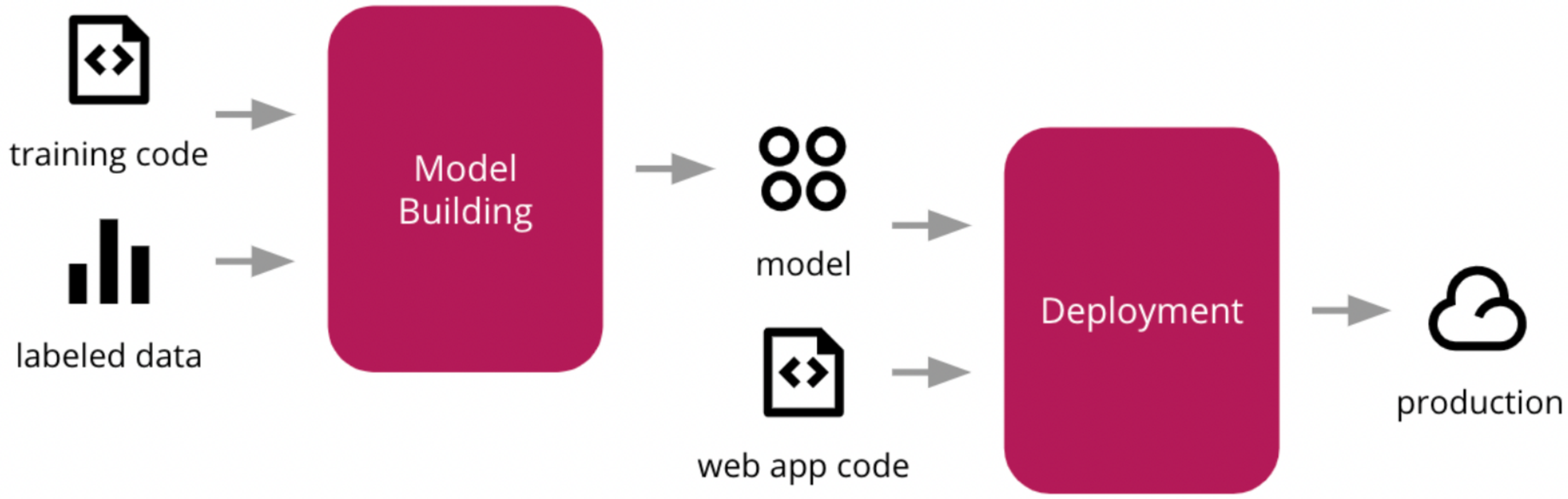
الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

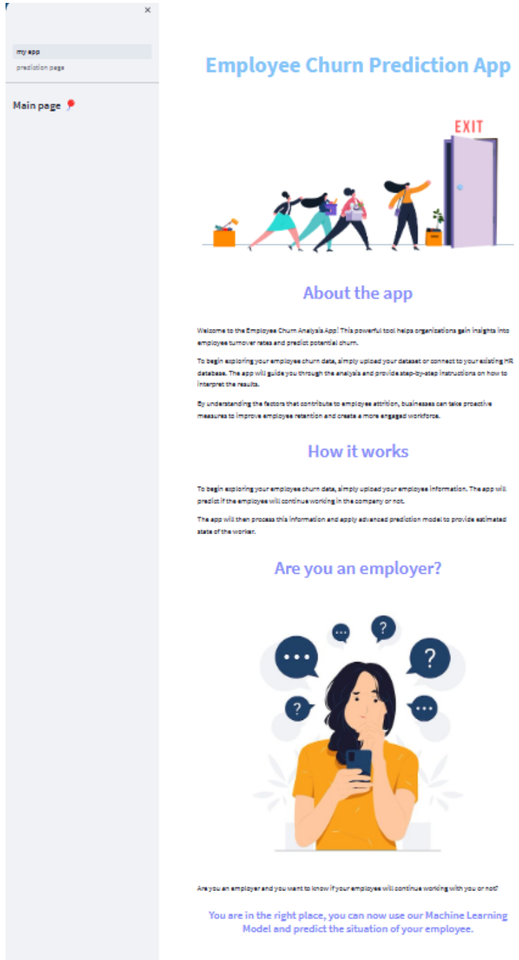
Model Deployment

What is Model Deployment?

Deployment means putting a machine learning model into action in a real business setting to make decisions based on data. It's one of the last steps in the machine learning process. In this project we deployed our ML model with Streamlit on AWS.



Employee Churn Prediction App

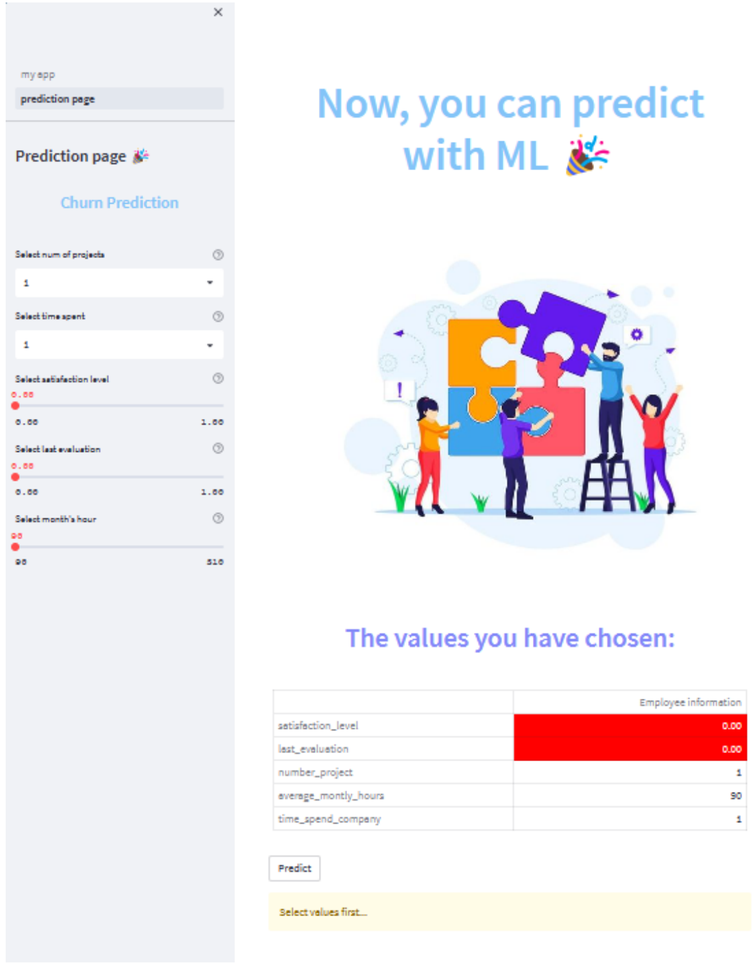


Page 1

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Employee Churn Prediction App



Page 2

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Visit Our App

Demo

URL

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

Conclusion

الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY

CLARUSWAY
WAY TO REINVENT YOURSELF

In this project we deploy a Streamlit app that predicts employees churn based on user inputs. The app loads a trained Random Forest classifier model and applies it to the user input data to predict whether or not the employee is likely to churn.

Thanks for Listening