

## MVP

# HR Analytics: Job Change of Data Scientists



The dataset here is related to the HR analytics of a company that is active in Big Data and Data Science, to estimate the candidates looking for a job change, and the factors influencing that.

We have main keys points to note in this dataset:

- Most of the data is categorical
- There is an imbalance in the classes
- There are lots of missing values

The goal of this Model is to predict the probability of a candidate work for the company.

The MVP notebook summarizes the following:

- Exploratory Data Analytics and Data Preprocessing
- Drop any unnecessary features
- Checking independency
- Handling Imbalance in Data(use SMOTE to oversample the imbalanced class)
- I tried two models (LogisticRegression and RandomForestClassifier to get the best correlation, but both getting the same result got 77%.
- I will be exploring the data deeply and changing the features to enhance the models performance, and try extra model.