HUMAN RESOURCES ANALYTICS

T5 BOOTCAMP DATA SCIENCE PROJECT

Done by:

ARWA ALOLYANI ,
GHAIDAA ALTHOBAITI

Submit to:

MR.ALI EL-KASSAS

PROJECT GOAL

Predict the probability of a candidate looking for a new jobAt that stage, I analyzed and explored the data in detail about ways to review the features affecting the goal and extract information from it, and that was the first stage.

Data Processing Sequence

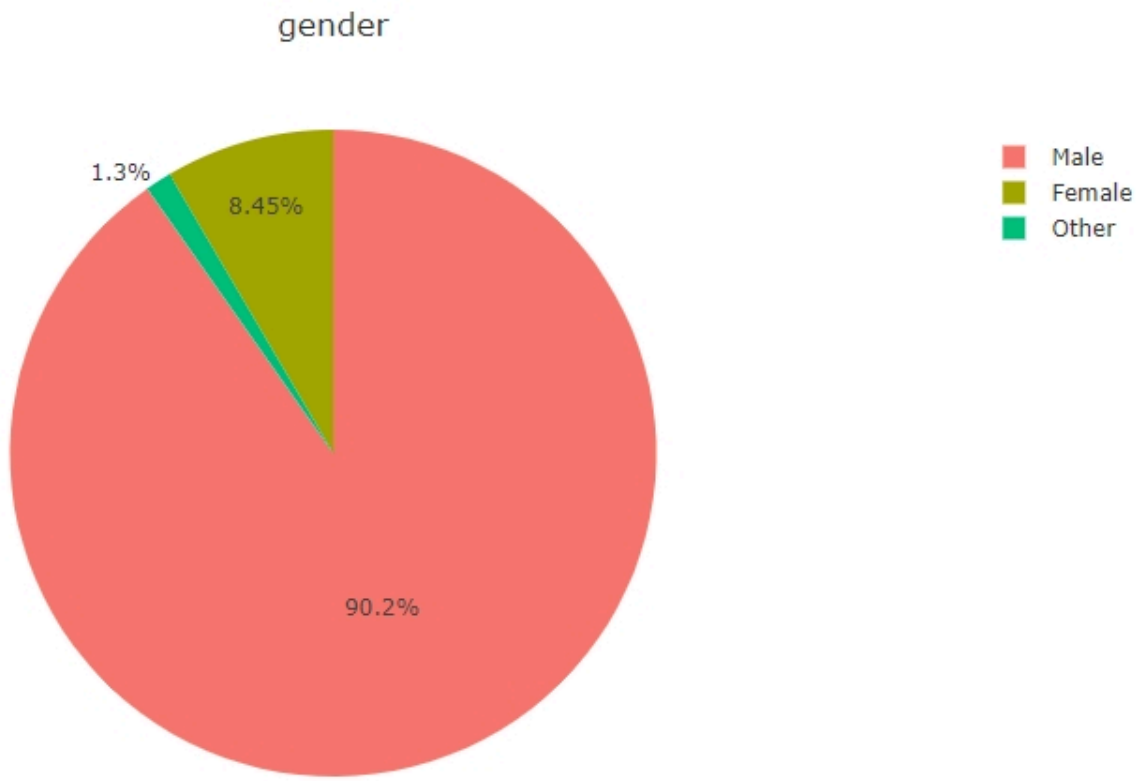
Clean data

Data visualization

Features Engineering

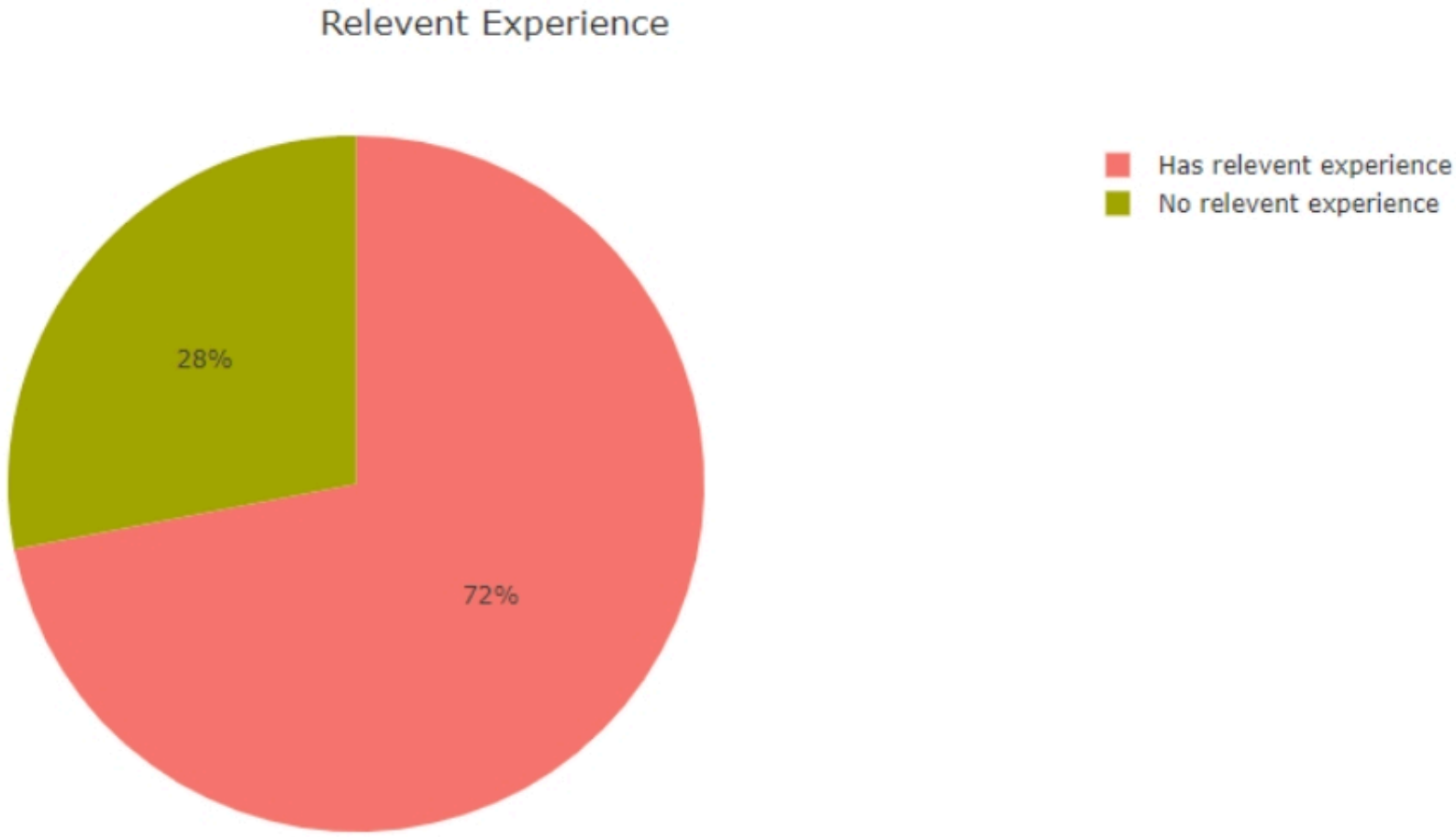
Models and results

Data Visualization

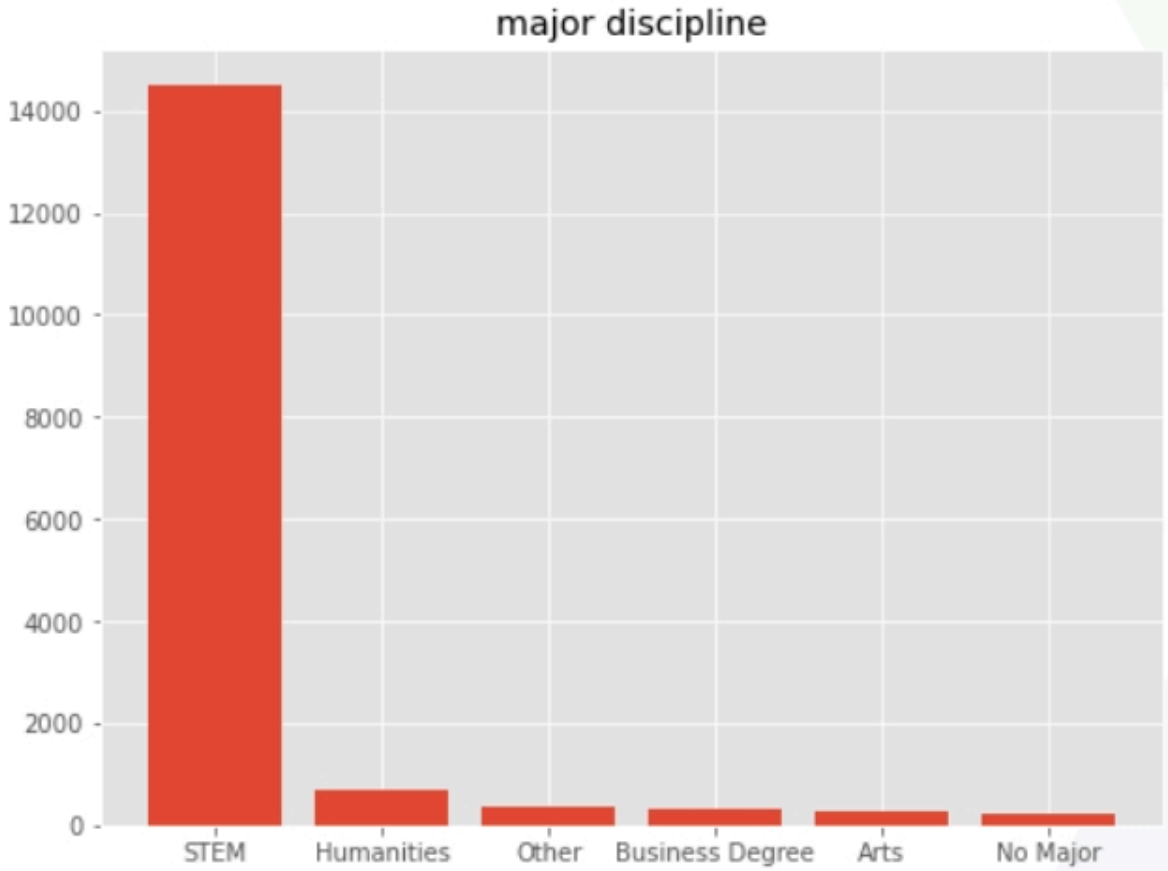


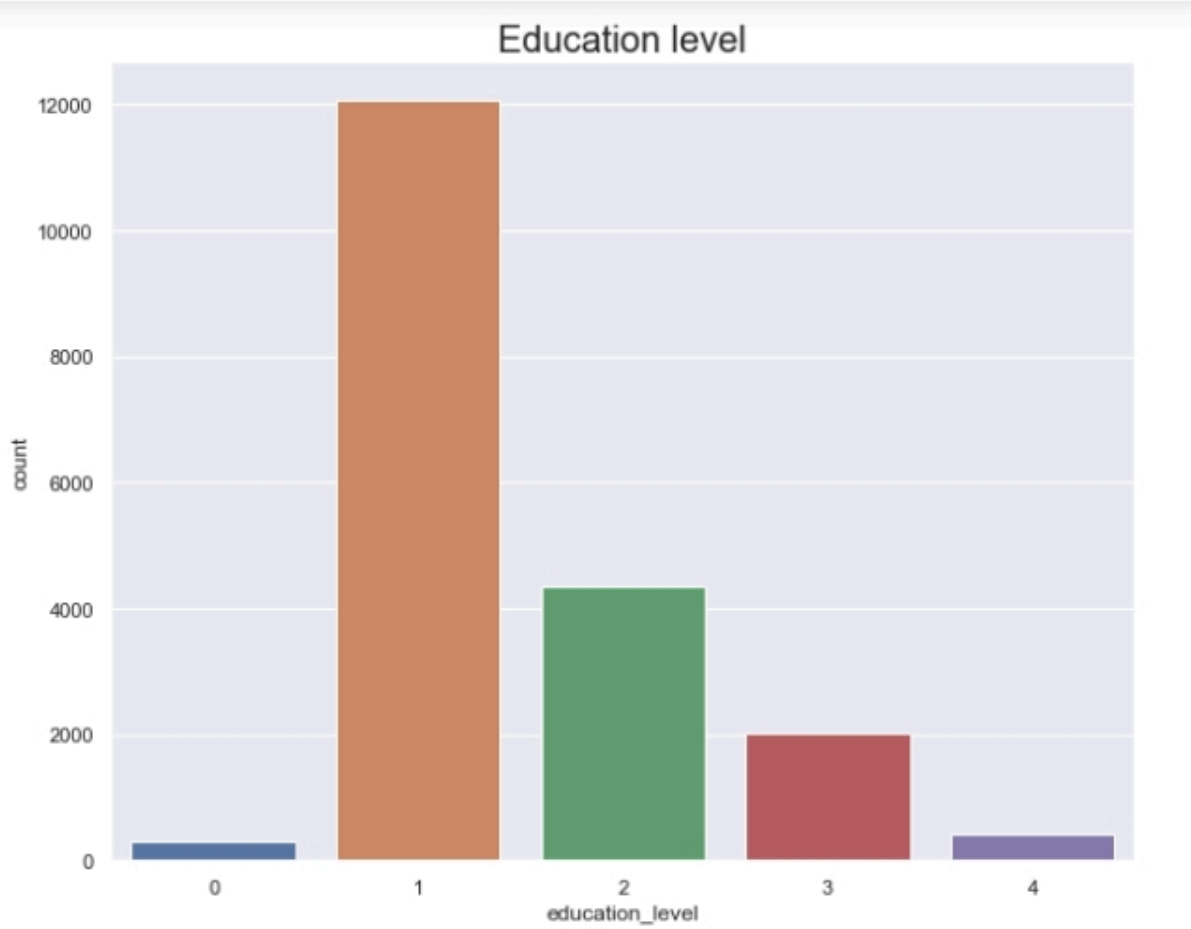
	Missing
city_development_index	0.000000
gender	0.235306
relevent_experience	0.000000
enrolled_university	0.020148
education_level	0.024011
major_discipline	0.146832
experience	0.003393
company_size	0.309949
company_type	0.320493
last_new_job	0.022080
training_hours	0.000000

Data Visualization

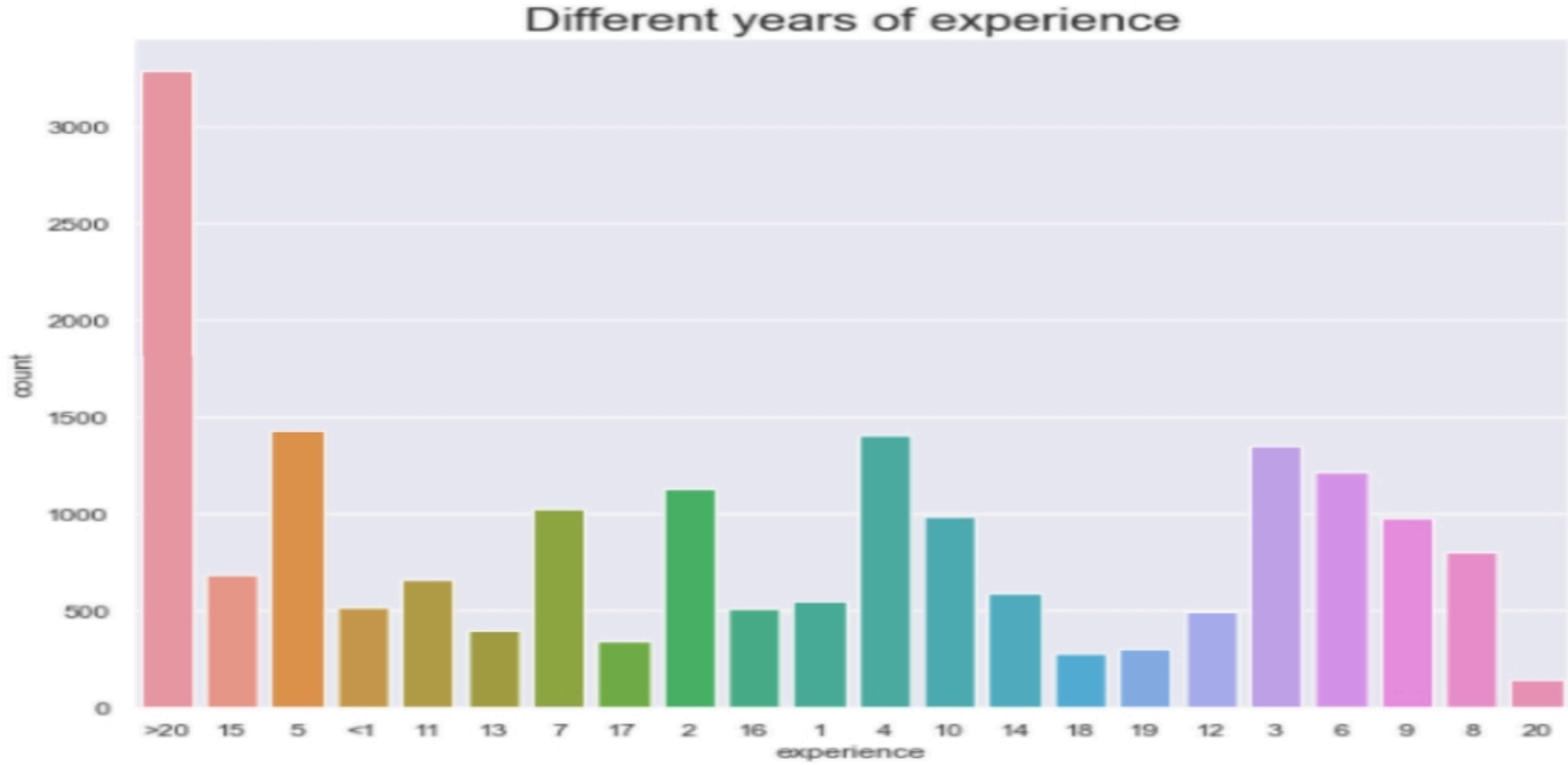


Data Visualization

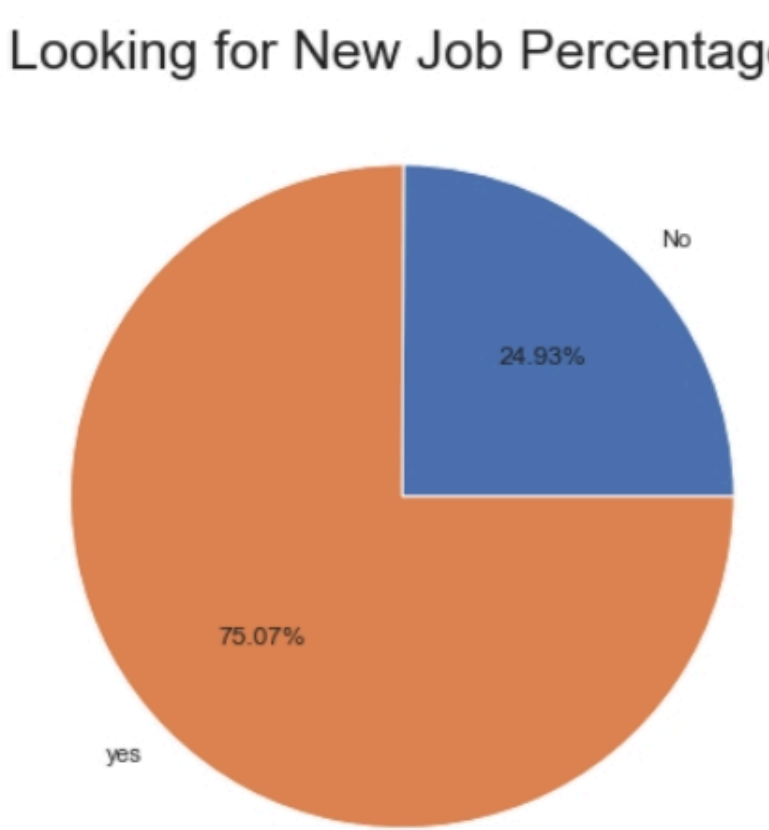


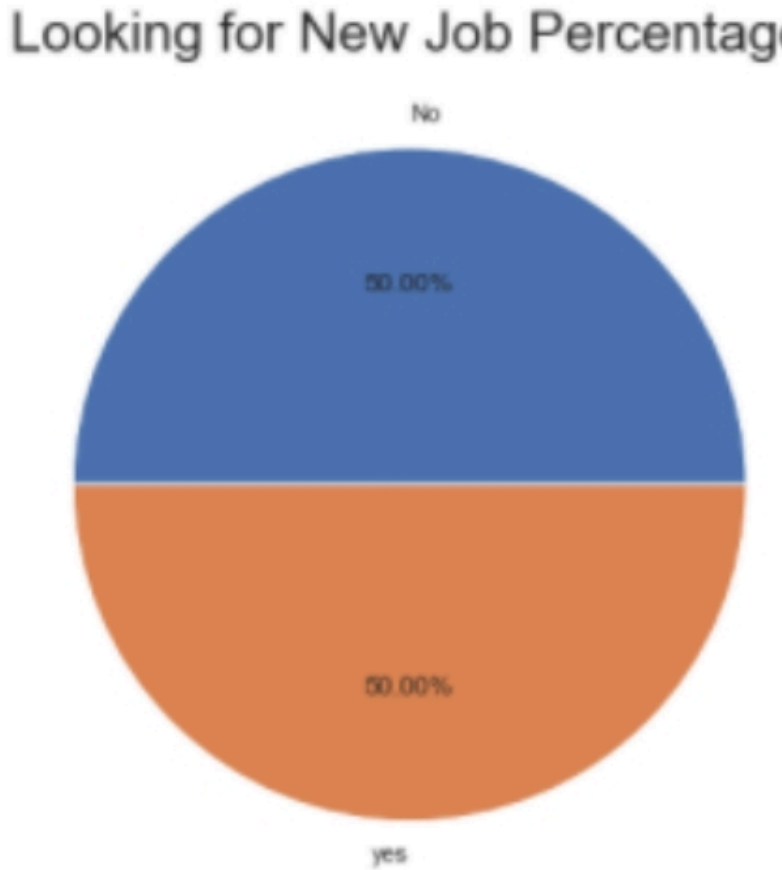


Data Visualization



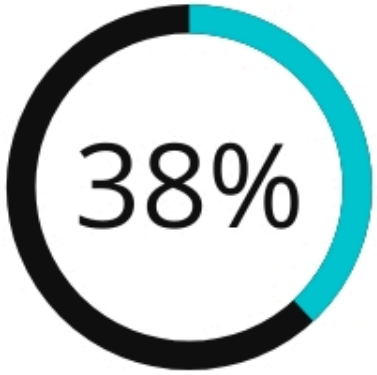
Class imbalance (by : SMOTE)





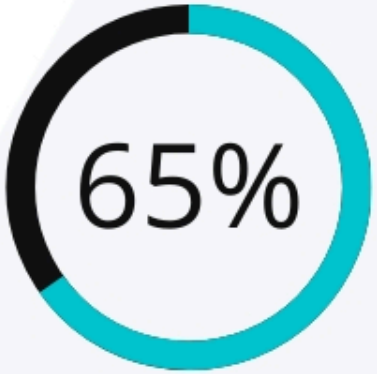
Model and Evaluation

We applied four different rating models and compare IF scores for each model.

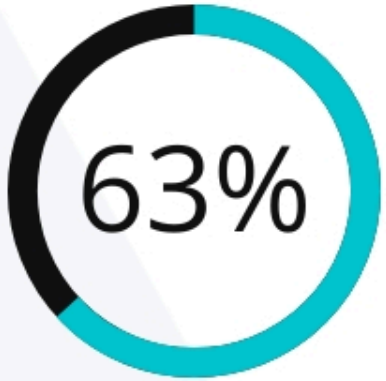


KNeighborsClassifier

LogisticRegression

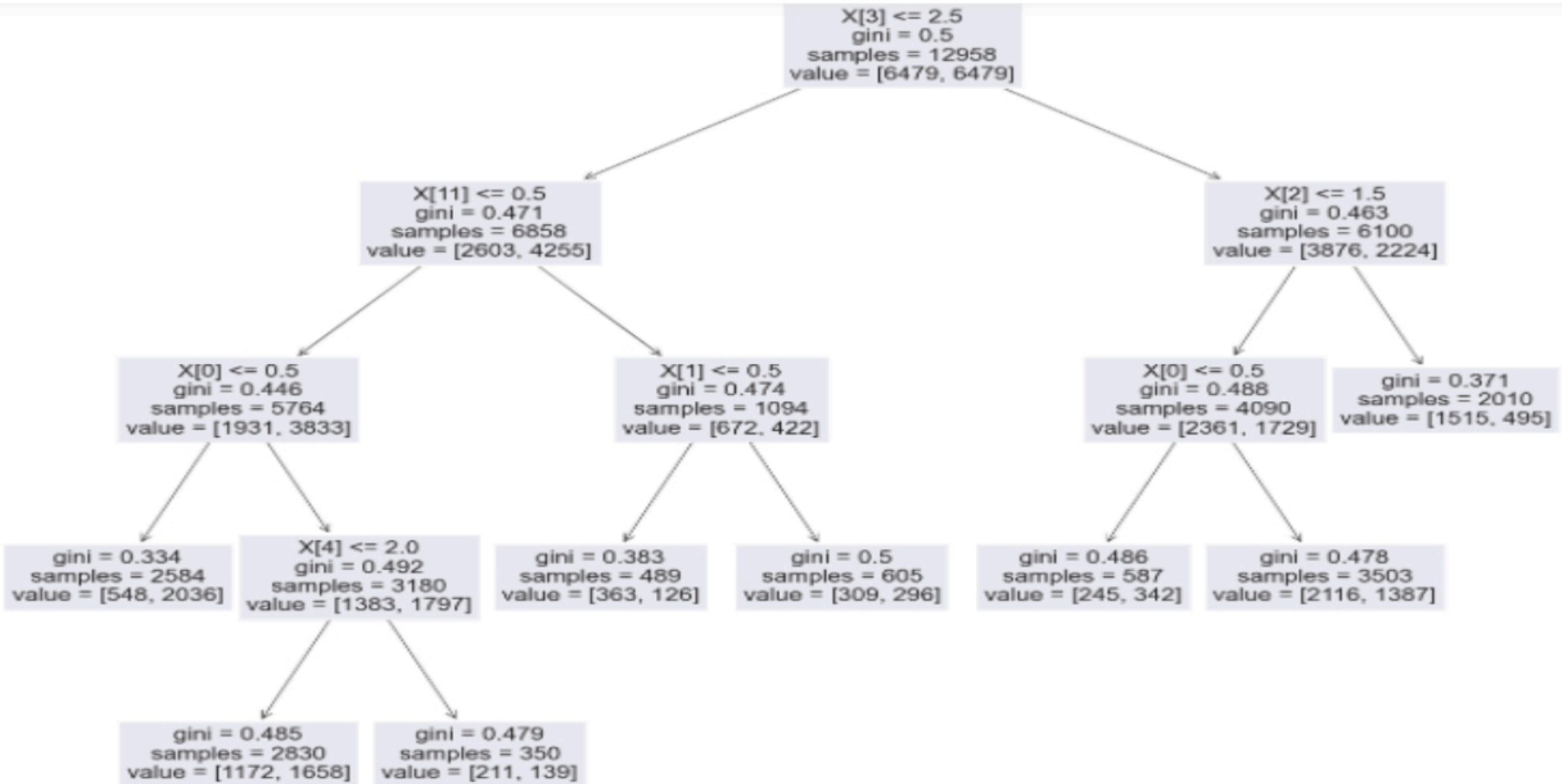


DecisionTreeClassifier



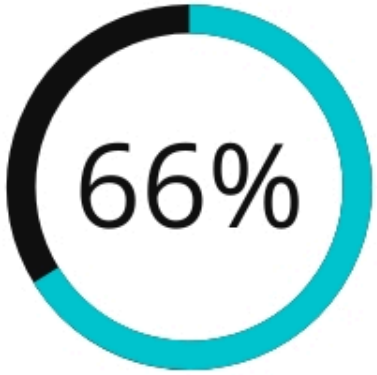
RandomForestClassifier

Exporting Decision Tree

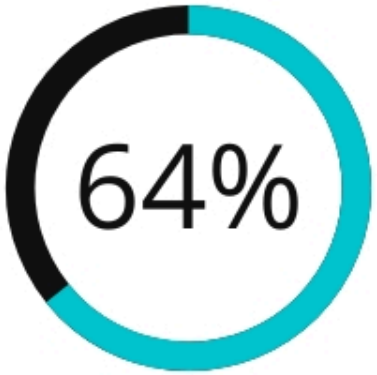


Ensemble Learning

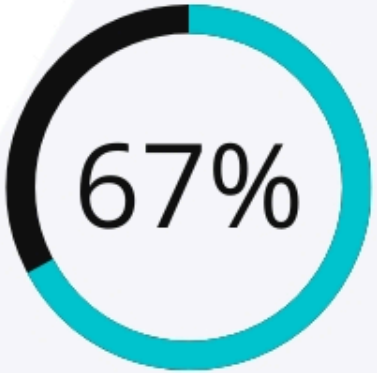
We applied most popular Ensemble methods, including bagging, boosting, stacking



VotingClassifier



BaggingClassifier



AdaBoostClassifier

StackingClassifier