Employee attrition prediction

Melvin Jeffrey Jacob

Objective

- To predict employee attrition using MLP, KNN, Random Forest classifiers and choose the best predictor

Approach to solving the problem

- Following are the steps involved in generating models to predict Employee attrition,
  - **Initial analysis** to look for missing data and outliers
  - **Data preprocessing** to separate training and test datasets
  - **Pipelining** to handle missing values in both numerical and categorical columns
  - **Model preparation**: Generate models using MLP, KNN and Random Forest classifiers
  - **Model evaluation**: Evaluate model accuracy using AUC, classification report and confusion matrix

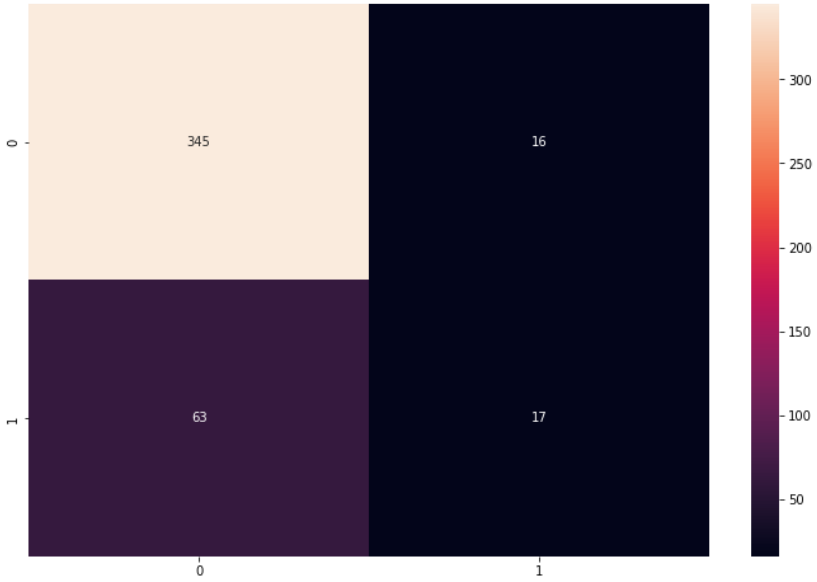
Model summary

- Classification algorithms used:
  - Multi-layer Perceptron classifier (MLP)
  - KNeighborsClassifier
  - RandomForestClassifier
- EDA:
  - Univariate analysis
  - Bivariate analysis
  - Correlation analysis
- Methods for evaluation of model accuracy:
  - AUC
  - Classification report
  - Confusion matrix

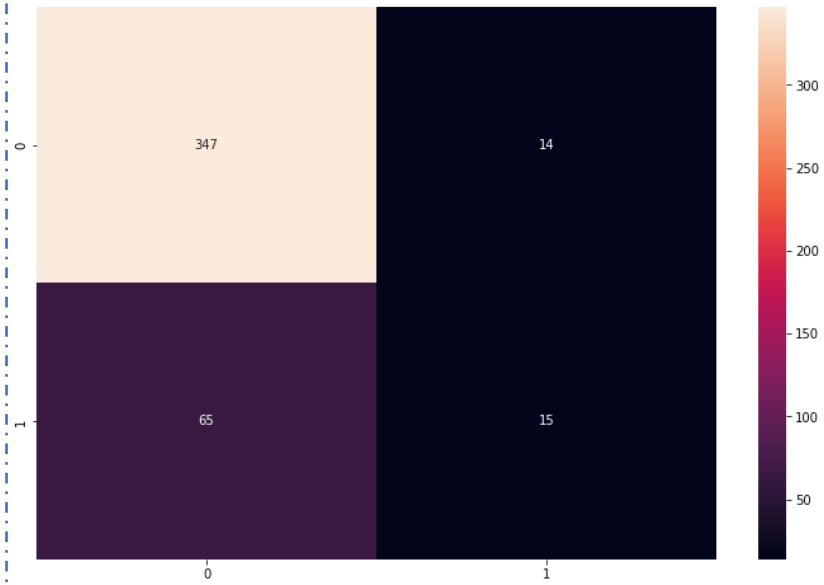
Results

Confusion matrix

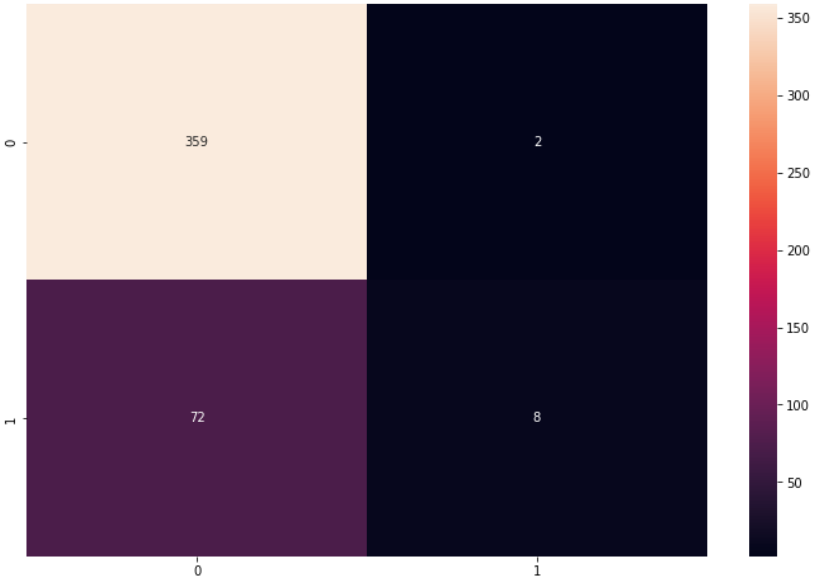
KNN classifier



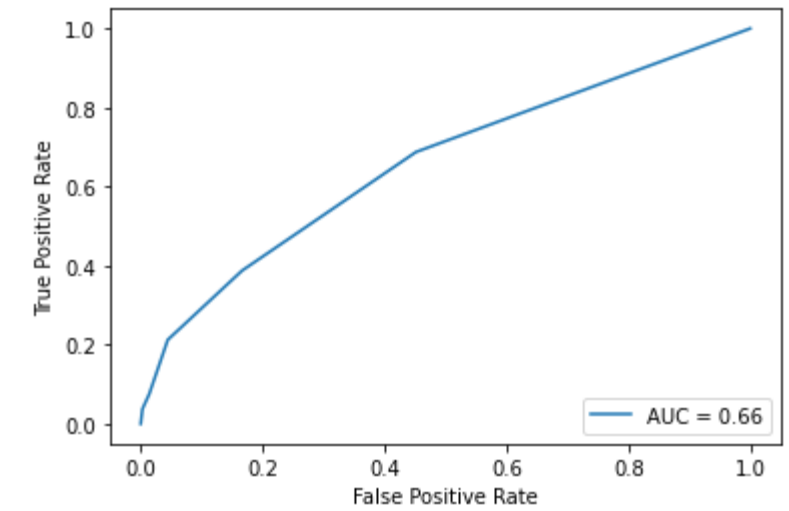
Random forest classifier

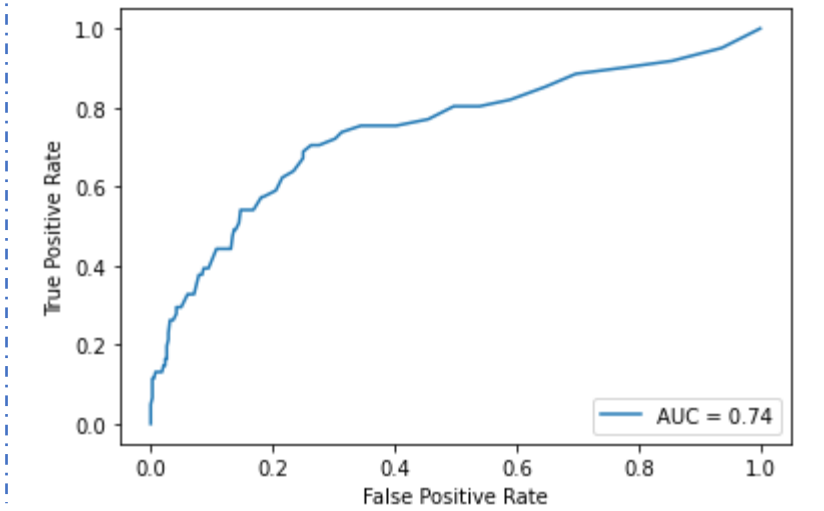


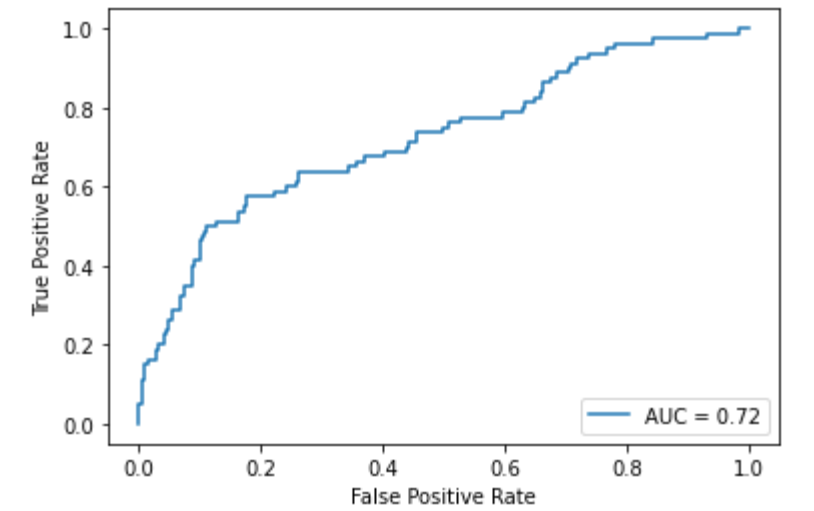
MLP



AUC







Results

Accuracy

```
{'KNN classifier': '86.49173955296405%', 'Random Forest classifier': '100.0%', 'MLP': '84.25655976676384%'}
```

Classifier report

KNN classifier Report :

	precision	recall	f1-score	support
No	0.85	0.96	0.90	361
Yes	0.52	0.21	0.30	80

Random Forest classifier Report :

	precision	recall	f1-score	support
No	0.87	0.96	0.92	377
Yes	0.44	0.19	0.26	64

MLP Report :

	precision	recall	f1-score	support
No	0.87	0.97	0.91	377
Yes	0.41	0.14	0.21	64

Inference

- All 3 models have more than 85% accuracy
- Random Forest classifier is the best out of the 3 classifiers with,
  - Highest AUC (Area Under Curve)
  - Highest accuracy
  - Highest F1 score

References

- Employee attrition prediction Jupyter notebook: https://github.com/MelvinJeffrey/melvinjeffrey/blob/195c3b93a1e146547a28196dcf5415d4bc6cab16/EmployeeAttritionPrediction.ipynb