AIT 636: Final Project

Name: Shubham Khaladkar

A screenshot of a computer

Description automatically generated with medium confidence

Data Cleaning – For training data (Checking ‘na’ or ‘?’ values)

Table

Description automatically generated with low confidence

Data Cleaning – For testing data

Table

Description automatically generated

Feature engineering – For training data

Converting categorical data to numerical

Text

Description automatically generated

Removing noise data (irrelevant data).

A screenshot of a computer

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

Feature engineering – For training data

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

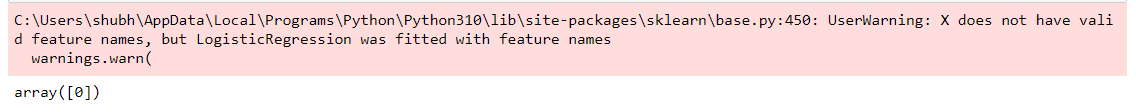
Classification Models

Training model using training data and predicting the salary on test data.

A picture containing timeline

Description automatically generated

Predicting the label output for the variation [0,0,2,4,3,1,1].



Logistic Regression



Perceptron



Decision Tree Classifier



Diagram, timeline

Description automatically generated

KNN Classifier



MLP without PCA



Linear SVC



Non-linear SVC



Regression Models

Decision Tree Regressor: Root Mean Square Error (RMSE)



Diagram

Description automatically generated

KNN Regressor: Root Mean Square Error (RMSE)



MLP classification model (using PCA)

Text

Description automatically generated

Text

Description automatically generated

Text

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated Graphical user interface

Description automatically generated

Text

Description automatically generated Text

Description automatically generated

A picture containing text

Description automatically generated Table

Description automatically generated

MLP: Find best subset selection

For training data

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated with medium confidence

For testing data

A picture containing text

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence