Data Science and Analytics Project

Packages and libraries : numpy, joblib, webbrowser, os, sklearn

Main goal of this project; **Will they move to new job ?** And what features effected this decision and how they effected (as %) ? Basically I tried to extract this prediction from dataset.

First of all normalized the excel (dataset) with KNIME then we read excel files and show it on web page for easier project and presentation.

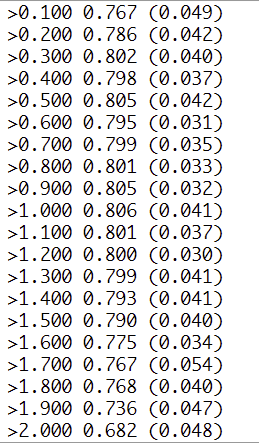
Then, used labelencoder on strings. Gave variables to x and y. Train model is x and test model is y. Set them as %30 test size and %70 train data set size. Used fit regression model for train and saved this model as dumb so we can use this model later.

Then, Loaded the trained model created with train\_model.py and created a numpy array based on the model's feature importances, sorted the feature labels based on the feature importance rankings from the model, then printed each feature label to show results.

Finally applied 5 different classifiers to this model. These are ;

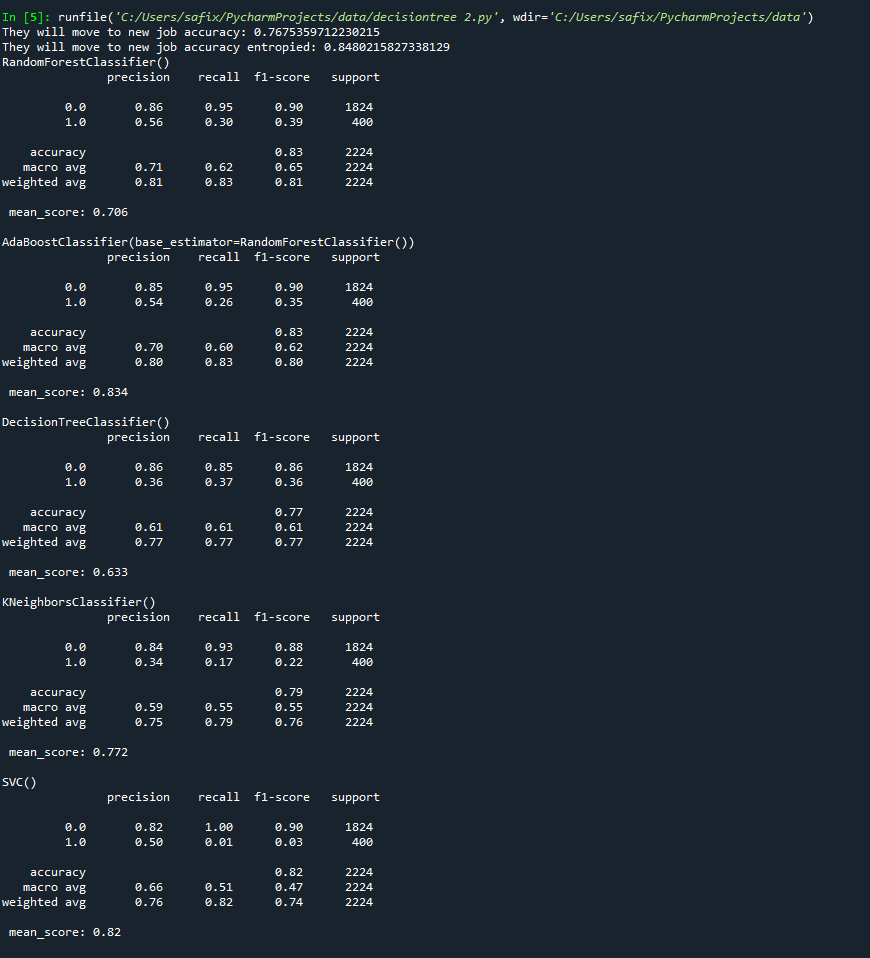
* RandomForestClassifier
* AdaBoostClassifier
* DecisionTreeClassifier
* KNeighborsClassifier
* SVC

Most successful accuracy came from Adaboost Classifier. AdaBoost algorithms can be used for both classification and regression problem. Adaboost helps you **combine multiple “weak classifiers” into a single “strong classifier”.** Accuracy from adaboost almost top value, cause after certain decision trees accuracy going to be less like these graphic sample;



Real-world data(dataset we used from real-world) includes some patterns that are linear but also many that are not. Switching from linear regression to ensembles of decision stumps(aka AdaBoost) allows us to capture many of these non-linear relationships, which translates into better prediction accuracy on the problem of interest,

Searched about my results, it may be overfitting or underfitting. We trained our dataset just fined and didn’t cause overlearning or underfitting. Dataset size is fine, not so small for overlearning.



Safi CENGİZ