

Assessing Feature with Random Forests

- Use the RandomForestClassifier implemented in Python scikit-learn to evaluate the importance of features in gender dataset we provide.
- Gender dataset format:

Fields	Description
obvious_laryngeal_prominence	Value=1 : With obvious larynx Value=0: Without obvious larynx
make_up	Value=1 : With makeup habits Value=0: Without makeup habits
like_coffee	Value=1 : Love to drink coffee Value=0: Don't like coffee
label	Value=1 : male Value=0: female

- The gender dataset is a .pkl file. You can use **pandas.read_pickle** to read the dataset.
- Note that for Python 2.7, please read the gender_27.pkl file, but read the gender_36.pkl for Python 3.6.
- Random forest parameter settings:
 - random_state=2
 - Use default setting for other parameters

Result

