Feature engineering

1. load trainLabels.csv to pandas

2. Use the first mode value of column 'workclass' to replace ' ?'

3. Use the first mode value of column 'occupation' to replace ' ?'

4. Use the first mode value of column 'native-country' to replace ' ?'

5. Use pandas.get_dummie() function to Convert categorical variable into indicator variables

6. Save features in taining data

7. load testLabels.csv to pandas

8. Use the first mode value of column 'workclass' to replace ' ?'

9. Use the first mode value of column 'occupation' to replace ' ?'

10. Use the first mode value of column 'native-country' to replace ' ?

11. Use pandas.get_dummie() function to Convert categorical variable into indicator variables

12. Save features in testing data

13. Add features to testing data assigned the default value 0 if some features are in training date, but not in testing data