

Main files are in Release folder.

Libraries required:

- * pandas
- * numpy
- * scikit-learn

LinearRegression Folder:

Data Files:

- * linear-regression-train.csv
- * linear-regression-test.csv

Code File:

- * linearRegression.py

How to Run:

- * If folder "output" does not exist in the LinearRegression folder, please create one before running the code
- * Run linearRegression.py file
- * The file runs linear regression with all 6 parameters one after the other. While testing one part of your code, you may comment out the other calls to LinearRegression() and lm.predict() to avoid spam from missing lines of other parts.

The output text files storing predicted y values will be in the "output" folder with names "linear-regression-output_betatype_zscore".

Additional Information:

- 1) You can change the input for an algorithm in two ways
 - a) You can paste your file in the folder, and change the name of input file in code.
 - b) You can change the name of your file to the default name given to input file and replace the default input file with your file in the folder.

2) DataFrame documentation:

<https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.html>

3) Numpy reference: <https://docs.scipy.org/doc/numpy-1.13.0/reference/index.html>

LogisticRegression Folder:

Code File:

- * logisticRegression.py

How to Run:

- * Fill in the missing code sections

- * run python logisticRegression.py

- * Example:

- * python ./logisticRegression.py

DecisionTree Folder:

Data Files:

- * house-votes-84.data

Code File:

- * DecisionTree.py

How to Run:

- * python3 DecisionTree.py data [0/1, 0-Information Gain, 1-Gain Ratio, default: 0]

- * Example:

- * python3 ./DecisionTree.py house-votes-84.data 0