Translated from Greek to English - www.onlinedoctranslator.com

HAROKOPIO UNIVERSITY
DEPARTMENT OF INFORMATION & TELEMATICS

Report 1st Assignment: Machine Learning and Applications

Manousos Linardakis, it22064

The code can be executed by running:

```
python3test_lr.py
```

This run gives the answers to questions 3.1, 3.2 and 3.3.

Indicative execution result:

```
Exercise 3_1 Answer ======== :
0.7284008391518103
======================
Exercise 3_2 Answer ======== :
Average RMSE: 0.8103139007144605
Standard Deviation of RMSE: 0.3133107528343484
======================
Exercise 3_3 Answer ======== :
Sklearn linear regression:
Average RMSE: 0.7272591467976938
Standard Deviation of RMSE: 0.009588274075152071
======================
```

In the archive `linear_regression.py` is the implementation of the LinearRegression class (as asked in question 2).

3_3 Question: Comparison - Commenting on Results:
Running query 2's LinearRegression 20 times and sklearn's LinearRegression another 20 times and comparing the mean and standard deviation of the RMSE, we notice that sklearn has better predicted the true values of y. This is first shown by the RMSE of sklearn, which after 20 trials equals 0.7272591467976938 which is smaller (so has less "errors" in the predictions) than that of the LinearRegression class of the 2nd query, which equals

0.8103139007144605. The same is demonstrated by the standard deviation of the RMSE, which after 20 trials equals 0.009588274075152071 in sklearn while in the LinearRegression class of the 2nd query it equals 0.3133107528343484.