



ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΤΗΛΕΜΑΤΙΚΗΣ

Αναφορά 1ης Εργασία: Μηχανική Μάθηση και Εφαρμογές

Μανούσος Λιναρδάκης, it22064

Ο κώδικας μπορεί να εκτελεστεί τρέχοντας:

```
python3 test_lr.py
```

Η εκτέλεση αυτή δίνει τις απαντήσεις στα ερωτήματα 3.1, 3.2 και 3.3.

Ενδεικτικό αποτέλεσμα εκτέλεσης:

```
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  
=====
```

Στο αρχείο `linear_regression.py` βρίσκεται η υλοποίηση της κλάσης `LinearRegression` (όπως ζητείται στο ερώτημα 2).

3 3 Ερώτημα: Σύγκριση - Σχολιασμός Αποτελεσμάτων:

Τρέχοντας 20 φορές το `LinearRegression` του ερωτήματος 2 και άλλες 20 φορές το `LinearRegression` του `sklearn` και συγκρίνοντας το μέσο όρο και τυπική απόκλιση του RMSE, παρατηρούμε ότι το `sklearn` έχει προβλέψει καλύτερα τις αληθινές τιμές των y . Αυτό φαίνεται αρχικά από το RMSE του `sklearn`, το οποίο μετά από 20 πειράματα ισούται με 0.7272591467976938 που είναι μικρότερο (άρα έχει λιγότερα "σφάλματα" στις προβλέψεις) από αυτό της κλάσης `LinearRegression` του 2ου ερωτήματος, το οποίο ισούται με 0.8103139007144605. Το ίδιο αποδεικνύεται και από τη τυπική απόκλιση του RMSE, η οποία μετά από 20 πειράματα ισούται με 0.009588274075152071 στο `sklearn` ενώ στη κλάση του `LinearRegression` του 2ου ερωτήματος ισούται με 0.3133107528343484.