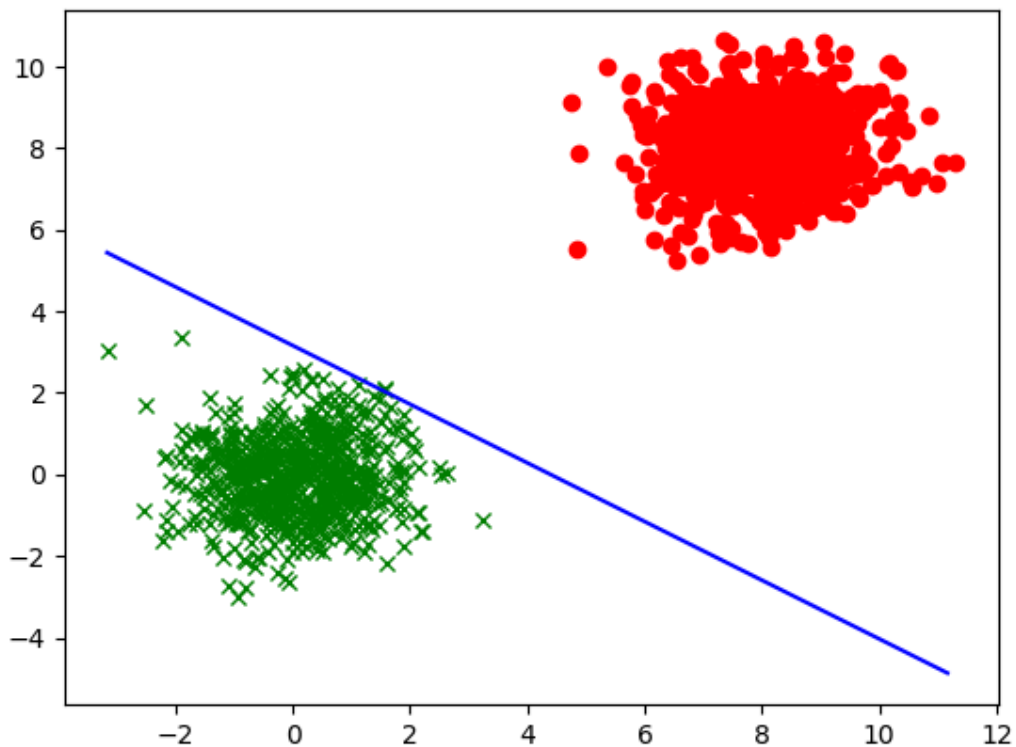


Μηχανική μάθηση - Λογιστική παλινδρόμηση

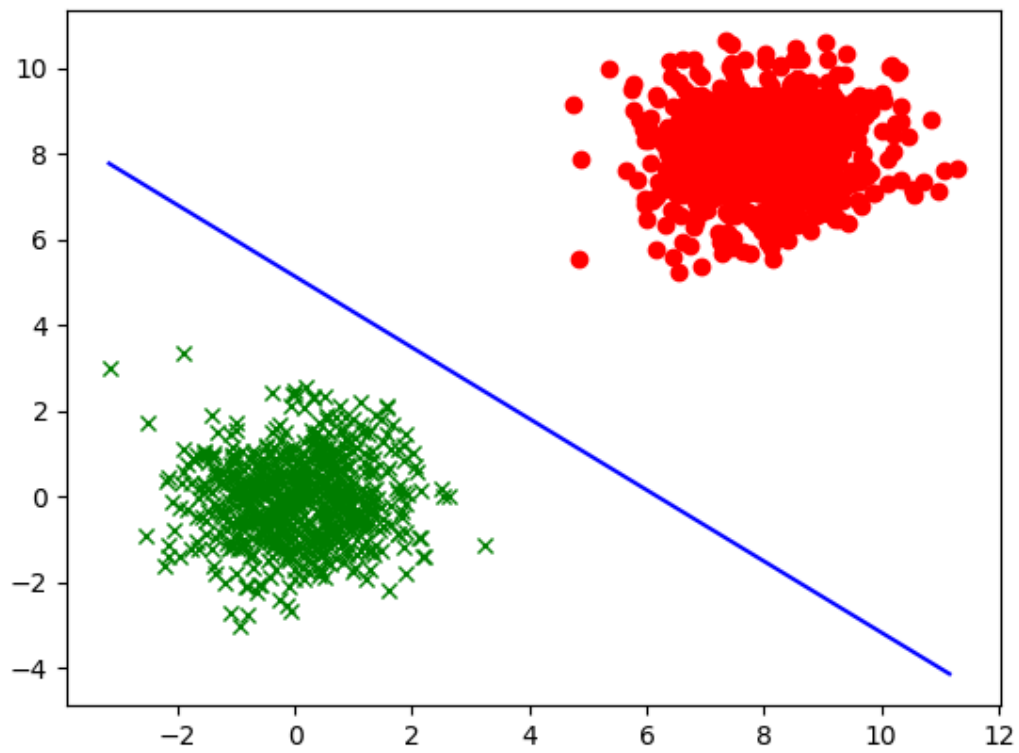
Αλεξίου Δημήτριος - it2021002

- Πείραμα 1: Συνθετικά δεδομένα:

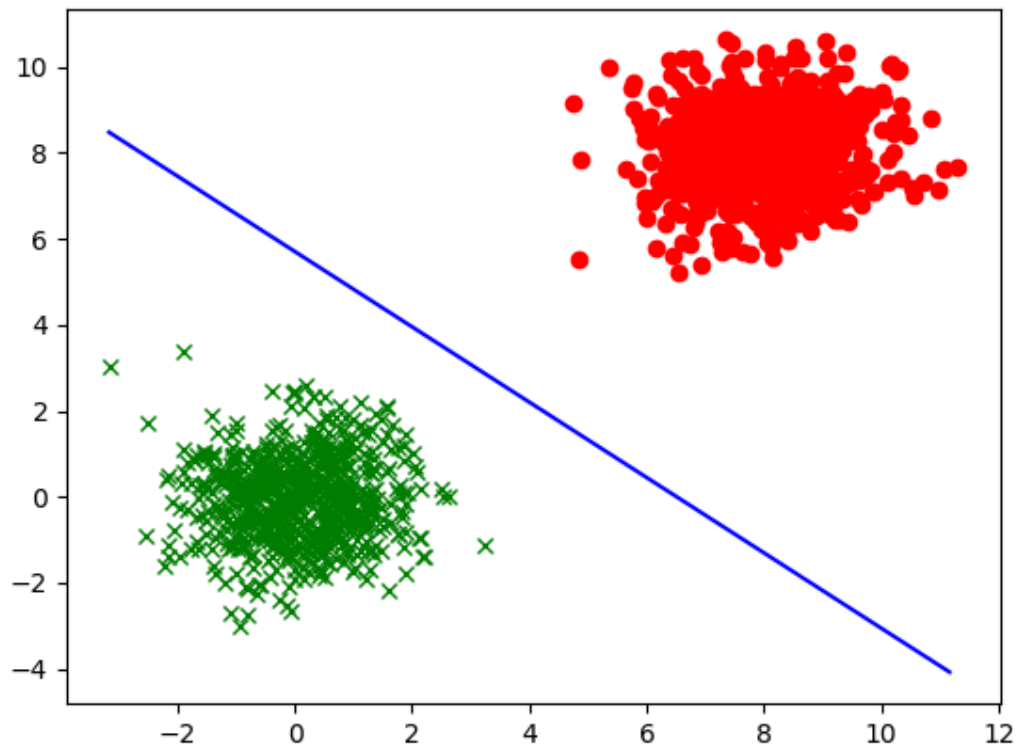
1ο Iteration:



50 Iteration:



100 Iteration:



Παρατηρούμε πως σε κάθε Iteration η γραμμή προσαρμόζεται ώστε να διαχωρίζει καλύτερα τα δεδομένα των δύο κατηγοριών. Επομένως, το μοντέλο με κάθε επανάληψη βελτιώνει τις προβλέψεις του και η συνάρτηση απώλειας μειώνεται όλο και περισσότερο.

Για $C = ([0, 8], [0, 8])$

```
Iteration 1000: Loss = 0.11915577089246226
Iteration 2000: Loss = 0.06657094802204874
Iteration 3000: Loss = 0.04580251429980661
Iteration 4000: Loss = 0.0348702118686495
Iteration 5000: Loss = 0.028159141572267357
Iteration 6000: Loss = 0.02362866542651718
Iteration 7000: Loss = 0.020366870463417744
Iteration 8000: Loss = 0.017906868752778
Iteration 9000: Loss = 0.015985404549108784
Iteration 10000: Loss = 0.014442943749042775
Accuracy: 1.0
```

Για $C1 = ([0, 5], [0, 5])$ τα κέντρα είναι πιο κοντά οπότε θα είναι δυσκολότερα διαχωρίσιμο

```
Iteration 1000: Loss = 0.14412736296341083
Iteration 2000: Loss = 0.08477799229111978
Iteration 3000: Loss = 0.06061071898835958
Iteration 4000: Loss = 0.047564986305979114
Iteration 5000: Loss = 0.039388041803596965
Iteration 6000: Loss = 0.03376950547036209
Iteration 7000: Loss = 0.02966166436709636
Iteration 8000: Loss = 0.02652110304974035
Iteration 9000: Loss = 0.02403785152435555
Iteration 10000: Loss = 0.022022097737034205
Accuracy: 0.9983333333333333
```

Παρατηρούμε λοιπόν πως η συνάρτηση απώλειας είναι μεγαλύτερη και το μοντέλο δεν έχει 100% ευστοχία.

Για $C2 = ([0, 3], [0, 3])$ παρατηρούμε πως η διαφορά είναι ακόμη μεγαλύτερη

```
Iteration 1000: Loss = 0.2030718395362548
Iteration 2000: Loss = 0.14584543653511547
Iteration 3000: Loss = 0.12019433573833795
Iteration 4000: Loss = 0.10543789295699098
Iteration 5000: Loss = 0.09575293802716296
Iteration 6000: Loss = 0.08885977989670615
Iteration 7000: Loss = 0.08367743498597986
Iteration 8000: Loss = 0.07962457655600887
Iteration 9000: Loss = 0.07635940631028522
Iteration 10000: Loss = 0.07366711032530303
Accuracy: 0.9766666666666667
```

● Πείραμα 2: Breast cancer dataset

Τα πειράματα εκτελούνται σε υπολογιστή με τα εξής χαρακτηριστικά:

Επεξεργαστής: Intel Core i5-8600k, 3.6 GHz

Μνήμη: 16GB DDR4, 2400MHz

```
Mean accuracy: 0.9786549707602339
Standard deviation of accuracy: 0.008921722983855171
Execution time: 6.052183151245117
```

Οπότε το πείραμα εκτελείται σε περίπου 6 δευτερόλεπτα.

Ο μέσος όρος της ευστοχίας είναι 0,978 και η τυπική απόκλιση 0,008

- Πείραμα 3: Breast cancer dataset (scikit-learn)

```
Mean accuracy: 0.9769005847953215  
Standard deviation of accuracy: 0.01132073724129002  
Execution time: 0.15776467323303223
```

Το πείραμα εκτελείται σε λιγότερο από 1 δευτερόλεπτο (0,15 sec).
Ο μέσος όρος της ευστοχίας είναι 0,976 και η τυπική απόκλιση 0,011

Συμπεραίνουμε λοιπόν πως το δικό μας μοντέλο, ενώ έχει μεγαλύτερο χρόνο εκτέλεσης από το μοντέλο του scikit-learn, έχει πιο υψηλό μέσο όρο ευστοχίας και μικρότερη τυπική απόκλιση. Συνεπώς, το δικό μας μοντέλο είναι πιο αποτελεσματικό.