2η Εργασία Τεχνητής Νοημοσύνης

Θεόδωρος Αλεξέγιεβ, p3200004

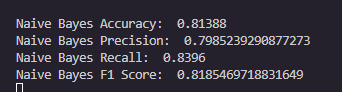
Κωνσταντίνος Ζαμπετάκης, p3200049

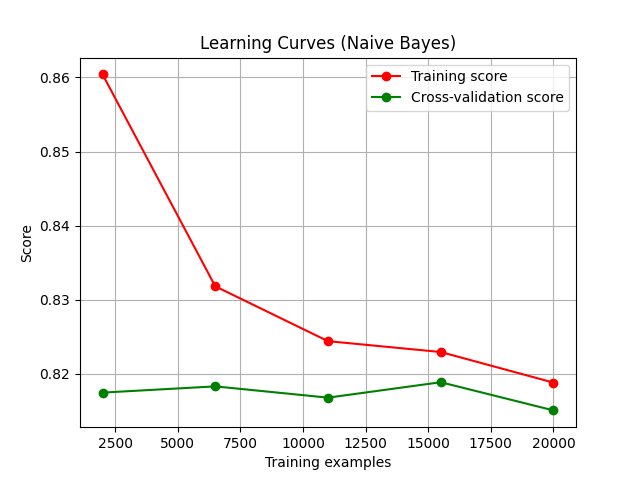
Ανδριάνα Καραγιάννη, p3200062

Έχουμε υλοποιήσει τους αλγορίθμους: 1. Αφελή ταξινομητή Bayes, 2. Random Forest και 3. AdaBoost. Η κάθε υλοποίηση βρίσκεται σε ξεχωριστό αρχείο με το αντίστοιχο όνομα. Τα αρχεία ‘testMyBayes.py’, ‘testRandomForest.py’ και ‘testMyAdaboost.py’ περιέχουν τον κώδικα για να μπορούμε να πάρουμε τα δεδομένα από την imdb βάση και να χρησιμοποιούμε τις δικές μας υλοποιήσεις των αλγορίθμων για να εμφανίζονται τα αποτελέσματα της ακρίβειας κάθε αλγορίθμου και οι καμπύλες εκμάθησής τους. Στα αρχεία ‘testNaiveBayes.py’, ‘testRandomForest\_sklearn.py’ και ‘testAdaboost\_sklearn.py’ περιέχεται ο κώδικας που παίρνει δεδομένα από την imdb βάση, όμως τώρα χρησιμοποιεί τις διαθέσιμες υλοποιήσεις των αλγορίθμων από το Scikit-learn.

Μέρος Α:

1. Αφελής ταξινομητής Bayes:

Για τον αλγόριθμο Naive Bayes έχουμε τις υπερπαραμέτρους n = 10 m = 1000 και k = 10 δίνουν την βέλτιστη ακρίβεια.



1. Random Forest:

Για τον αλγόριθμο Random Forest έχουμε τις υπερπαραμέτρους

n\_estimators και max\_depth. Βρήκαμε ότι οι τιμές, n\_estimators = 50 και max\_depth = none δίνουν την βέλτιστη ακρίβεια.

Εικόνα που περιέχει κείμενο, γραμματοσειρά, στιγμιότυπο οθόνης, αριθμός

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

1. AdaBoost:

Για τον αλγόριθμο adaboost έχουμε τις υπερπαραμέτρους n\_estimators και learning\_rate. Βρήκαμε ότι οι τιμές, n\_estimators = 50 και learning\_rate = 0.01 δίνουν την βέλτιστη ακρίβεια.

Εικόνα που περιέχει κείμενο, γραμματοσειρά, στιγμιότυπο οθόνης

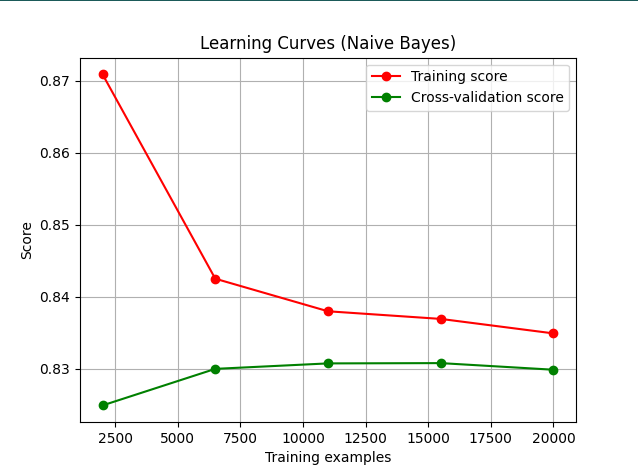
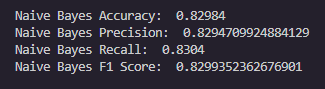
Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμμή, γράφημα

Περιγραφή που δημιουργήθηκε αυτόματα

Μέρος Β:

1. Αφελής ταξινομητής Bayes:

Παρατηρούμε ότι η δική μας υλοποίηση μας δίνει accuracy = 81.38% ενώ η υλοποίηση του Scikit-learn δίνει accuracy = 82.98%.

1. Random Forest:

Παρατηρούμε ότι η δική μας υλοποίηση μας δίνει accuracy = 54.18% ενώ η υλοποίηση του Scikit-learn δίνει accuracy = 81.53%

Εικόνα που περιέχει κείμενο, γραμματοσειρά, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα

1. AdaBoost:

Παρατηρούμε ότι η δική μας υλοποίηση μας δίνει accuracy = 53.44% ενώ η υλοποίηση του Scikit-learn δίνει accuracy = 65.54%.

Εικόνα που περιέχει κείμενο, γραμματοσειρά, στιγμιότυπο οθόνης

Περιγραφή που δημιουργήθηκε αυτόματα

Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, διάγραμμα, γραμμή

Περιγραφή που δημιουργήθηκε αυτόματα

Μέρος Γ:

Υλοποιήσαμε τον αλγόριθμο MLP sliding window. Παρατηρούμε ότι έχει accuracy 83%.