

Τεχνητή Νοημοσύνη II

Εργασία 2

REPORT - PROJECT DOCUMENTATION

Φουκανέλης Χρήστος-Γεώργιος, 1115201900204

Δεκέμβριος 2022

Περιγραφή υλοποίησης

Η εργασία υλοποιήθηκε με βάση τις οδηγίες της εκφώνησης και του φροντιστηρίου. Για την εύρεση των καλύτερων υπερπαραμέτρων των μοντέλων, χρησιμοποιήθηκε η βιβλιοθήκη *optuna*. Έχουν υλοποιηθεί συναρτήσεις οι οποίες αρχικοποιούν το μοντέλο με τις υπερπαραμέτρους του και το αξιολογούν με αυτόματο τρόπο. Ενδεικτικά παραθέτονται 3 παραδείγματα με διαφορετικές αρχιτεκτονικές δικτύου και επιλογές υπερπαραμέτρων. Για τα word embeddings χρησιμοποιήθηκαν vectors 200 διαστάσεων καθώς ήταν αρκετά αποδοτικά. Όσον αφορά τον καθαρισμό των δεδομένων έχουν εφαρμοστεί οι ίδιες τεχνικές που περιγράφονται στο report της 1ης εργασίας. Οι συναρτήσεις *objective*, *build model*, *train and evaluate* έχουν υλοποιηθεί με βάση το *documentation* της *optuna* και εκπαιδεύουν και αξιολογούν ένα μοντέλο με βάση τις παραμέτρους που θα λάβουν οι οποίους έχουν παραχθεί κατά την εξερεύνηση του χώρου αναζήτησης των υπερπαραμέτρων μέσω της *create study*. Οι συναρτήσεις αυτές είναι υλοποιημένες 2 φορές με τις δεύτερες να μην λαμβάνουν την παράμετρο *trial* οπότε και φτιάχνουν μοντέλο με τις δοθέντες παραμέτρους. Η *model test* αξιολογεί ένα εκπαιδευμένο μοντέλο πάνω σε ένα test dataset.

Υπερπαραμετροί και σχολιασμός

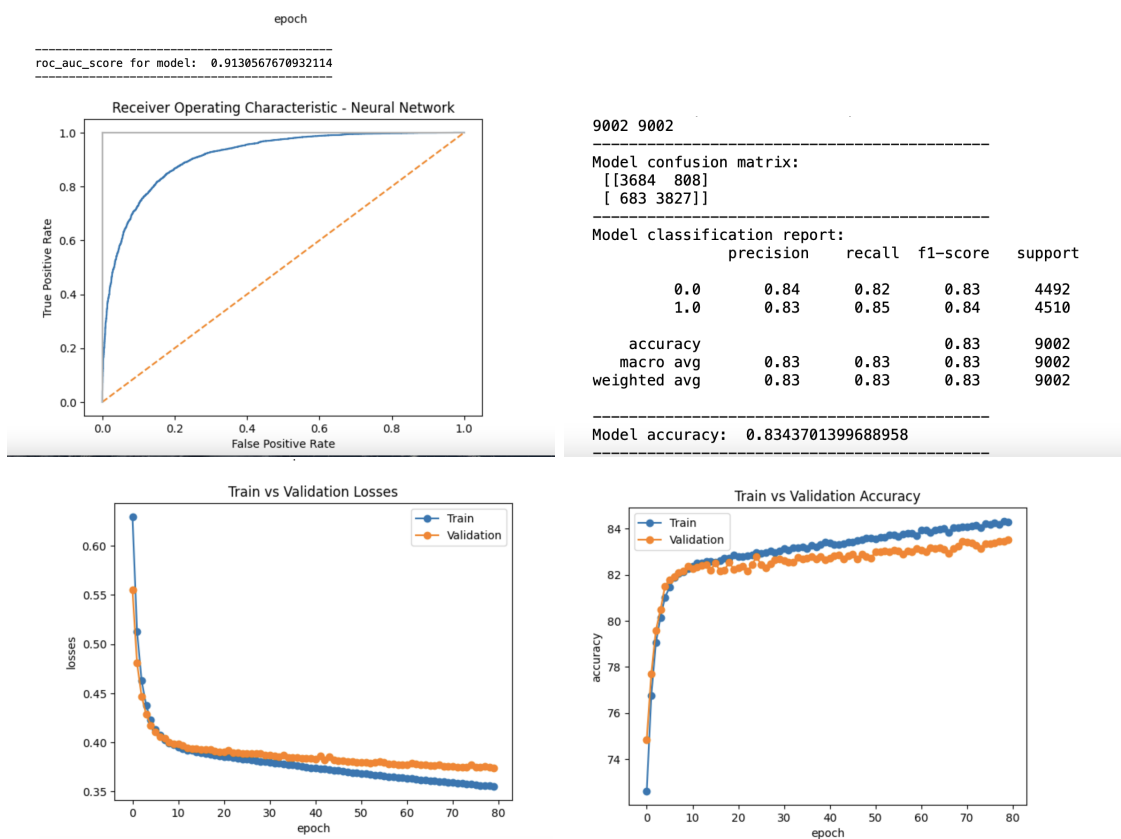
Η *study* της *optuna* διασχίζει ένα χώρο αναζήτησης με πολλές υπερπαραμέτρους.

Ειδικότερα, δοκιμάστηκαν παραλλαγές για το *learning rate*, τον *optimizer*, τη *loss function*,

την *activation function* στα ενδιάμεσα *layers*, το *batch size*, των αριθμό των *epochs*, την χρήση ή μη *dropout layers* αλλά και το πλήθος των ενδιάμεσων *layer* και των *nodes* κάθε *layer*. Δεν χρησιμοποιήθηκε *batch normalisation*.

Η *optuna* επέστρεψε τις παραμέτρους με τις οποίες το μοντέλο είχε το περισσότερο *accuracy*. Πολλές είναι οι πιθανές παράμετροι καθώς διάφοροι συνδυασμοί δίνουν αποτέλεσμα πολύ κοντά σε αυτό του βέλτιστου. Στα *learning curves* βλέπουμε ότι παρόλο που μεγαλύτερο *learning rate* έδωσε σε ένα παράδειγμα ελάχιστα καλύτερο *accuracy*, οι μεταβολές του *accuracy* είναι μεγάλες από ένα *epoch* σε άλλο.

Ενδεικτικές μετρικές ενός μοντέλου



Σημειώσεις - Παρατηρήσεις:

Όπως φαίνεται, το μοντέλο πετυχαίνει μια καλή ακρίβεια της τάξης του 0,83. Στο παραδοτέο συμπεριλαμβάνονται: Το *ipynb* αρχείο, το *readme* και το *.tex* αρχείο (ενδέχεται να υπάρχουν errors καθώς έγινε compile σε online editor), τα αποθηκευμένα trained μοντέλα. ***Το project στο τελευταίο κελί έχει ένα testing pipeline όπου αρκεί να εισαχθεί το path για ένα csv αρχείο, και θα εκτελεστεί η αξιολόγηση του μοντέλου. Ενδεικτικά αποθηκεύεται ένα μοντέλο ώστε να μην χρειαστεί να ξαναγίνει train. ***