

BERT FINE TUNING FOR IMDB SENTIMENT ANALYSIS

PROJECT DOCUMENTATION

Η υλοποίηση της εργασίας (1^ο ερώτημα) έγινε με βάση τις υποδείξεις της εκφώνησης και των φροντιστηρίων. Χρησιμοποιήθηκε το bert-base-uncased μοντέλο σε συνδυασμό με 2 dense layers. Optimizer χρησιμοποιήθηκε ο AdamW ειδικά για bert. Το project τρέχει σε cuda για να μειωθεί ο χρόνος training. Τα βήματα της εργασίας διαφαίνονται από τα σχετικά σχόλια στον κώδικα. Το μοντέλο πιάνει ένα μέτριο accuracy το οποίο δεν υπήρχε ο χρόνος να διερευνηθεί τρόπος να αυξηθεί. Στο τελευταίο κελί του notebook υπάρχει ένα testing pipeline αρκεί να εισαχθεί το path στην μεταβλητή *test_csv_path*. Συμπεριλαμβάνεται και αποθηκευμένο ένα trained μοντέλο σε μορφή link προς κατέβασμα λόγω του μεγάλου μεγέθους του.

Model_weights:

<https://drive.google.com/file/d/1XhxoFZzDucJp0VOqh9-9f1870F8xsZC1/view?usp=sharing>

model_state:

<https://drive.google.com/file/d/1ofGWA-7fzwHiN99NspVbDcGGeogXpVFA/view?usp=sharing>