

Opis Problema

U projektu se pravi veliki jezički model koji vrši klasifikaciju filmskih recenzija na pozitivne i negativne na osnovu tekstualnog sadržaja recenzije.

Skup Podataka

Skup podataka se sadrži od ukupno 100.000 (po 25.000 za treniranje i testiranje i 50.000 za nesupervizijsko učenje) ulaza u formatu:

```
{  
  "label": 0,  
  "text": "Goodbye world2\n"  
}
```

gde label uzima vrednosti iz skupa {0,1}, gde 0 predstavlja negativnu recenziju a 1 pozitivnu. Polje 'text' označava tekstualni sadržaj recenzije.

Skup podataka se može preuzeti na [linku](#)

Metoda za resavanje problema

Model koristi encoder-only arhitekturu, adaptiranu u odnosu na dati transformer i prilagođen zadatku klasifikacije. Ulazni tekst se najpre tokenizuje i mapira u vektorske reprezentacije pomoću embedding sloja, nakon čega se dodaju pozicione informacije kako bi se očuvao redosled tokena u sekvenci. Dobijene reprezentacije se zatim prosleđuju kroz više slojeva enkodera koji koriste mehanizam self-attention za modelovanje zavisnosti između reči u okviru cele recenzije.

Nakon prolaska kroz encoder slojeve, reprezentacija cele sekvence se agregira u jedan vektor (npr. pomoću srednje vrednosti vektora tokena ili korišćenjem specijalnog klasifikacionog tokena). Ovaj vektor se zatim prosleđuje linearnom klasifikacionom sloju koji proizvodi izlazne logike za dve klase: pozitivni i negativni sentiment.

Proces učenja je u potpunosti nadgledan. Kao funkcija greške koristi se cross-entropy loss, koji meri odstupanje između predikcija modela i stvarnih oznaka u skupu podataka. Optimizacija parametara modela vrši se pomoću algoritma kao što je Adam ili AdamW, uz primenu standardnih tehnika regularizacije (npr. dropout) radi smanjenja preučenja.