

Dokumentacija projekta „Generator tekstova pjesama“

Dragan Mirković, SW41-2019
Softversko inženjerstvo i informacione tehnologije
Osnove računarske inteligencije
Novi Sad, jul 2022. godina

1. Upotreba rješenja

Rješenje se sastoji od dvije implementacije. Jedna implementacija koristi „Tensorflow“-ovu „CuDNNLSTM“ mrežu, koja predstavlja „Long short-term memory“ mrežu koja računanje obavlja na grafičkoj kartici. Generisanje rezultata se vrši na osnovu proslijeđene osnove, „seed“-a. Drugo rješenje je implementirano upotrebom „textgenrnn“ ekstenzije za Python, ono takođe koristi „CuDNNLSTM“, ali samo inicijalizovanje mreže i generisanje rezultata je skriveno unutar ekstenzije. Prva implementacija se nalazi unutar „main.py“, a druga unutar „implementation_textgenrnn.py“ datoteke. Metrika za ocjenjivanje rezultata nalazi se unutar „rouge_metric.py“ datoteke. Promjena rješenja za koje se računa metrika izvodi se izmjenom putanje do datoteke generisanog teksta unutar koda. Pokretanje se obavlja jednostavnim pozivom „python/3 [naziv_datoteke]“. Zbog različitih verzija paketa između dvije implementacije savjetuje se upotreba virtuelnog okruženja.

2. Potrebne biblioteke

Za obje implementacije potrebno je imati „Python 3.9“, „CUDA Toolkit 11.2“, „cuDNN SDK 8.1.0“, „Pandas“, „Nltk“, „Numpy“ i „Mathplotlib“, „Tensorflow 2.5“ i „Keras 2.9/2.4.3“. Za pokretanje metrike potrebno je imati instalirano „Datasets“ „rouge-score“ i „Strawberry-Perl“.