CPSC 8430 Deep Learning

Danish Bhatkar (C15563212) – dbhatka@clemson.edu

GitHub Repository Link

HW3 – Extractive Question Answering

Problem Statement

Create a model that generates answers for an input question text on BERT architecture. The input is a question in the form of text. The output will also be a sequence of text that answers the question in short.
It is achieved by BERT architecture which is based on Transformers.

Data set

The dataset provided as part of the homework is used to train the model. The dataset contains 37,111 question-answer pairs for training & 5351 for the purpose of testing. The data is in their respective JSON files.

Method

The model in this homework is BERT sequence model that takes text as input and generates text. The process goes through the steps like pre-processing(adding answer_end, context processing, lowercasing, etc), tokenization, answer retrieval, training and testing.

Data fields:-

“paragraphs” – text data about a topic

“context” – some information about a paragraph

“question” – question based on the paragraph

“answers” – some answers to that question

Results

Epochs - 3

Training Loss – 2.238900

Validation Loss – 3.138374

F1 Score – 74