

CPSC 8430 Deep Learning

Danish Bhatkar (C15563212) – dbhatka@clemsn.edu

[GitHub Repository Link](#)

HW3 – Extractive Question Answering

Problem Statement

To develop a question-answering system using the BERT (Bidirectional Encoder Representations from Transformers) architecture. This model:

1. Accepts a text-based question as input
2. Processes the question using BERT's transformer-based architecture
3. Generates a concise textual response that directly answers the input question

The functionality relies on BERT's ability to understand context and relationships within text, using the pre-trained model. This system aims to provide short, accurate answers to a wide range of questions.

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 separate JSON files (test and train).

Method

A BERT sequence model is used for text-to-text generation, processing input text to produce output. The workflow begins with pre-processing steps such as adding answer end markers, context processing, and lowercasing. This is followed by tokenization, where the text is converted into tokens for the model. Next, relevant information is retrieved to formulate answers. The model is then trained on the dataset before being evaluated through testing on a designated test set. This streamlined process enables the BERT model to effectively generate concise and relevant answers to textual questions.

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