National University of Sciences and Technology

School of Electrical Engineering and Computer Science

Department of Computing

CS 405 Deep Learning

Fall 2023

# 

# Lab 11

**Open-Ended Lab**

**Text Generation Via Transformers:**

**Story Generation based on a Given Prompt**

**Name: Alina Nasir**

**Qalam ID: 342350**

# Date: 25th Nov 2023

Data Exploration and Loading:

The dataset was downloaded from Kaggle using the link provided in the manual. In order to familiarize with the dataset, the head of the dataset with four prompts in training and validation dataset was checked.

Data-Preprocessing:

The writing prompts dataset contained the separate files for the target and source for each of the train, validation and test data. However, the model I decided to use that is GPT-2 required one-line data for training. Therefore, in order to make it one-line data I processed the dataset for stories and prompts into a single file. This pre-processing helped to create three files namely train.wp\_combined, validation.wp\_combined, test.wp\_combined that could be passed to the model. Moreover, the text data contained starting and end tokens, which were also removed before passing the data to the model.

Model Selection:

The GPT-2 model was selected for training GPT-2 is pre-trained on a massive corpus of diverse text data. This pre-training helps the model learn general language patterns, grammar, and world knowledge. This large-scale pre-training allows GPT-2 to capture a broad range of contextual information. Moreover, GPT-2 is a generative language model, meaning it can generate coherent and contextually relevant text. This makes it suitable for tasks that involve generating new text based on given prompts, as is the case with prompts and stories.

Fine-Tuning:

The dataset was then fine-tuned on GPT-2 model. The dataset used was the validation dataset due to the limitations of processing capabilities. Moreover, the number of epochs was kept one at a time to due to constraints of Colab environment.

Evaluation:

The model was evaluated on the test dataset where the perplexity achieved was 22.2. This value of perplexity is considered good for GPT-2.

Story Generation:

The model was then put into the evaluation state for generating the stories. The maximum length of stories was kept of 300 words. The results of stories are displayed in the notebook.