**Assignment\_7**

1. Explain the architecture of BERT

**Ans: BERT, which stands for Bidirectional Encoder Representations from Transformers, is based on Transformers, a deep learning model in which every output element is connected to every input element, and the weightings between them are dynamically calculated based upon their connection.Explain Masked Language Modeling (MLM).**

1. Explain Next Sentence Prediction (NSP)

**Ans: In the BERT training process, the model receives pairs of sentences as input and learns to predict if the second sentence in the pair is the subsequent sentence in the original document.**

1. What is Matthews evaluation?

**Ans: Matthew's correlation coefficient, also abbreviated as MCC. MCC is a statistical tool used for model evaluation. Its job is to gauge or measure the difference between the predicted values and actual values and is equivalent to chi-square statistics for a 2 x 2 contingency table.**

1. What is Matthews Correlation Coefficient (MCC)?

**Ans: Matthews defined VM, known as the Matthews coefficient, as the crystal volume per unit of protein molecular weight, and showed that VM bears a straightforward relationship to the fractional volume of solvent in the crystal**.

1. Explain Semantic Role Labeling

**Ans: A semantic role is the underlying relationship that a participant has with the main verb in a clause. Discussion: Semantic role is the actual role a participant plays in some real or imagined situation, apart from the linguistic encoding of those situations.**

1. Why Fine-tuning a BERT model takes less time than pretraining.

**Ans: the authors recommend only 2-4 epochs of training for fine-tuning BERT on a specific NLP task (compared to the hundreds of GPU hours needed to train the original BERT model or a LSTM from scratch!).**

1. Recognizing Textual Entailment (RTE)

**Ans: Textual entailment recognition is the task of deciding, given two text fragments, whether the meaning of one text is entailed (can be inferred) from another text (see the Instructions tab for the specific operational definition of textual entailment assumed in the challenge).**

1. Explain the decoder stack of GPT models.

**Ans: GPT-3, or the third generation Generative Pre-trained Transformer, is a neural network machine learning model trained using internet data to generate any type of text. Developed by OpenAI, it requires a small amount of input text to generate large volumes of relevant and sophisticated machine-generated text.**