**Sandya Rani Prasadam HW3**

**Github:** <https://github.com/sandya-33/Deep-Learning_HW3>

**Model**: I used a pre-trained model “bert-base-uncased” model. 'bert-base-uncased', has 12 transformer layers and uses a basic uncased vocabulary of 30,522 tokens. The model was pre-trained on a large corpus of text data and can be fine-tuned for various natural language processing (NLP) tasks such as question-answering, text classification etc.,

**Tokens**:

Bert Tokenizer

BERT tokenizer is responsible for converting raw text into a sequence of tokens that can be understood by the BERT model. It breaks the input text into individual words or sub words and then maps each word or sub word to a unique ID.

**Parameters:**

Batch size = 16

Max len = 387

Num epochs = 10

Learning rate = 2e-5

AdamW optimizer

Loss: 0.1446