1. **BERT Base Uncased**:
   * "BERT Base" refers to the smaller version of the BERT model, which has fewer layers and parameters compared to "BERT Large."
   * "Uncased" indicates that the model's vocabulary does not distinguish between uppercase and lowercase letters. In other words, all words in the input text are converted to lowercase before being tokenized.
   * BERT Base Uncased has 12 transformer layers, 768 hidden units in each layer, and 110M total parameters.
2. **BERT Base Mean NLI Tokens**:
   * "Mean NLI Tokens" refers to the specific tokenization strategy used for fine-tuning BERT on natural language inference (NLI) tasks.
   * In NLI tasks, the goal is to determine the relationship between two given sentences, typically categorized as entailment, contradiction, or neutral.
   * The "Mean" tokenization strategy involves taking the mean (average) of the embeddings of the first and second sentences, which are passed through the BERT model during fine-tuning for NLI tasks.
   * This tokenization strategy is different from the standard tokenization used during pre-training or general text processing.