

Capstone Project - NLP

Compulsory Task 1

NLP applications	Use - cases
Text classification	A model that allocates which mail folder an email should be sent to (work, friends, promotions, important), like Gmail's inbox tabs.
Automatic summarisation and text classification	A model that helps decide what grade to award to an essay question. This can be used by a university professor who grades a lot of classes or essay competitions.
Question answering and sentiment analysis	A model that provides assistive technology for doctors to provide their diagnosis. Remember, doctors ask questions, so the model will use the patients' answers to provide probable diagnosis for the doctor to weigh and make decisions.

Compulsory Task 2

Google's BERT (Bidirectional Encoder Representations from Transformers) model is an innovative technology that has garnered considerable attention in recent years in the field of natural language processing (NLP). It is a pre-trained deep learning model that leverages context to enhance language comprehension.

The primary achievement of BERT is its capability to comprehend the context of words within a sentence, which is crucial in deciphering the meaning of the text. For instance, the meaning of the word "bank" in a sentence could vary based on the context in which it is used. BERT can comprehend the context of words in a sentence, providing precise interpretations, even in intricate sentence structures.

The functioning of BERT involves a transformer-based neural network architecture that has been pre-trained on a vast corpus of unlabeled text. Through this training, the model can grasp the connections between words and sentences and their application within context. The training process entails two tasks: masked language modeling and next sentence prediction.

During the masked language modeling task, BERT selects some words at random within a sentence and trains the model to forecast the masked words based on the sentence's context. This activity facilitates the model in comprehending the relationships between words within a sentence.

In the next sentence prediction task, BERT learns to forecast whether a particular sentence follows another sentence. This process helps the model in comprehending the interrelation between sentences and how the context of adjacent sentences can alter a sentence's meaning.

BERT's key advantage is its proficiency in comprehending the intention behind a user's query. Google has implemented BERT to refine the pertinence of search outcomes by interpreting the user's query context and presenting more precise search results. Furthermore, BERT has been employed to enhance the performance of machine translation and chatbots.

To summarize, BERT is a ground-breaking technology that has vastly enhanced natural language processing by enabling machines to interpret the context of words and sentences in text. Its proficiency in understanding the intention behind a user's query has made it a potent instrument in enhancing the relevance of search outcomes, machine translation, and chatbots.