

TASK 22

COMPULSORY TASK 2

QUESTION 1

Read up on any innovative technology using NLP (by companies such as google or IBM for instance) and write a brief summary about the technology, what it achieves / does and an overview of how it works (250-500 words)

ANSWER

In this particular use-case, I have researched and learnt about the GOOGLE BERT(Bidirectional Encoder Representations from Transformers mode) NLP algorithm which has become a super useful tool in modern technology because of its impact on NLP in general.

The primary objective of BERT is to capture the context of words in a sentence by considering words that come before and after the word in question. Unlike previous NLP models that uses unidirectional approach, BERT uses a bidirectional approach, allowing it to better understand the nuances and meaning of words within a sentence.

BERT as a pre-trained language representation, refers to how BERT is first trained on large sources of text such as wikipedia. The training results can then be applied to other NLP tasks such as question answering and sentiment analysis etc.

BERT and AI makes NLP model training so much faster and efficient. BERT achieves its goal through a two-step process.

1st: pretraining : during this stage, BERT is exposed to a large corpus of unlabelled text such as books, wikipedia, websites and learns to predict missing words within sentences. It does this by using random words in the input and training to predict the correct words based on the surrounding context.

2nd: fine-tuning : BERT is then fine-tuned after pretraining to specific downstreamed tasks such as text classification, sentiment analysis, question answering or named entity recognition. During the fine-tuning phase, BERT is trained to optimise performance for targeted tasks by training it on labelled dataset specific to each task.

BERT can also process multilingual NLP as google recently released a multilingual version which can effectively handle multiple languages enabling more inclusivity and comprehensive language processing capabilities.

BERT's ability to understand the relationship and dependencies between words in a sentence more comprehensively sets it apart from other models because it is able to

handle complex language tasks including ambiguous, syntactic structures or long-range dependencies.

BERT have revolutionised NLP because of its significance in advanced NLP applications and its accuracy and effectiveness in tasks such as sentiment analysis, named entity recognition, document classification and many more. By leveraging BERT developers can enhance their ability to understand and generate natural language, leading to improved chatbots, language translation systems and search engines.

In Summary, Google BERT is a revolutionary NLP technology that has changed the industry when it comes to language understanding and processing tasks. By adopting the bidirectional approach and leveraging large-scale pretraining, BERT is able to capture the contextual understanding of words better than other NLP's. Its ability to generate high-quality contextual embeddings has revolutionised the NLP application leading to significant improvement in NLP classification. It has become a cornerstone technology for the NLP field pushing more boundaries to enhance NLP capabilities.

REF: <https://cloud.google.com/ai-platform/training/docs/algorithms/bert-start>,
[https://en.wikipedia.org/wiki/BERT_\(language_model\)](https://en.wikipedia.org/wiki/BERT_(language_model)),
(<https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>)