

Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação
Depto. de Ciências de Computação - SCC - ICMC - USP



Estudo do Modelo BERT
(Bidirectional Encoder Representations from Transformers)

Yazid Alamou Bouari - No. USP: 13551034
Orientador(a): Profª Drª Roseli Aparecida Francelin Romero

São Paulo - SP
Outubro de 2023

1. O que é BERT:

O BERT (Bidirectional Encoder Representations from Transformers) é um dos modelos mais avançados no campo de Processamento de Linguagem Natural (PLN). Ele se destacou por sua capacidade de compreender e gerar texto com base no contexto, revolucionando muitas tarefas de PLN.

1.1. Desempenho do BERT

O BERT demonstrou desempenho excepcional em várias tarefas de PLN, alcançando resultados de estado-da-arte em tarefas como classificação de texto, análise de sentimento, tradução automática e muito mais.

2. Como o BERT Funciona:

2.1. Pré-treinamento e ajuste fino:

O BERT é pré-treinado em grandes conjuntos de dados não rotulados, permitindo-lhe aprender representações bidirecionais de palavras. Em seguida, ele pode ser ajustado em tarefas específicas, como ajustar um modelo pré-treinado para trabalhar com diálogos de crianças.

2.2. Bidirecionalidade

A principal inovação do BERT é a sua capacidade de processar palavras bidirecionalmente, levando em conta o contexto à esquerda e à direita de cada palavra em uma frase.

2.3. Tokens e atenção

O BERT divide o texto em "tokens", unidades de texto, e utiliza uma arquitetura de atenção multi-cabeça para ponderar a importância das palavras em relação umas às outras.

3. Aplicações do BERT no seu Projeto:

3.1. Interpretação de Linguagem Infantil

O BERT pode ser treinado para compreender e interpretar a linguagem usada por crianças em diálogos, capturando erros gramaticais comuns e entendendo a intenção por trás das declarações infantis.

3.2. Geração de Respostas

Através do ajuste fino, o BERT pode ser treinado para gerar respostas apropriadas e relevantes em conversas com crianças, criando sistemas de chatbots interativos para crianças.

3.3. Detecção de Sentimentos

O BERT pode ser aplicado para detectar sentimentos nas falas das crianças, identificando emoções como felicidade, tristeza ou raiva em diálogos.

4. Técnicas de Ajuste Fino com BERT:

4.1. Transfer Learning

A técnica de ajuste fino do BERT envolve a transferência de conhecimento do modelo pré-treinado para tarefas específicas, economizando tempo e recursos de treinamento.

4.2. Conjunto de Dados

Para ajuste fino do BERT em diálogos com crianças, é fundamental ter um conjunto de dados adequado e rotulado que representam as interações desejadas (terapia, estudos, jogos...)

5. Ferramentas e Recursos:

5.1. Bibliotecas de PLN (uso do BERT)

Existem várias bibliotecas que oferecem implementações do BERT, como a Hugging Face Transformers, spaCy, NLTK, entre outras. A Hugging Face Transformers é especialmente popular e oferece acesso a modelos pré-treinados e ferramentas para ajuste fino.

5.2. Hardware e Software

5.2.1. Hardware:

- **Unidade de Processamento Gráfico (GPU):**
Uma GPU de alto desempenho é altamente recomendada para treinamento eficiente de modelos BERT.
- **Memória GPU:**
Memória GPU suficiente para acomodar o modelo BERT, os dados de treinamento e os cálculos necessários.

5.2.2. Software:

- **Ambiente Virtual Python 3.x:**
Uso do ambiente virtual Python para isolar as dependências do projeto e evitar conflitos entre bibliotecas. Uso de `virtualenv` ou `conda` para criar e gerenciar ambientes virtuais.
- **Framework de Deep Learning:**
Uso de Framework TensorFlow ou PyTorch, para implementar e treinar modelos BERT.