

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



**Hipóteses das limitações antes dos resultados das aplicações  
considerando os estudos**

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

**São Paulo - SP**  
**Outubro de 2023**

**Necessidade de Dados Adequados:**

O BERT exige um grande volume de dados rotulados para ajuste fino em uma tarefa específica. Se não houver um conjunto de dados adequado de diálogos com crianças disponíveis, a coleta e rotulação de dados podem ser demoradas e trabalhosas.

**Idioma e Dialeto:**

O BERT foi treinado principalmente em texto em inglês. Se o seu projeto envolve diálogos em outros idiomas ou dialetos, o desempenho do BERT pode ser limitado, a menos que seja ajustado fino em dados nesse idioma específico.

**Compreensão Contextual Limitada:**

Embora o BERT leve em conta o contexto à esquerda e à direita das palavras em uma frase, ele ainda pode ter dificuldade em compreender contextos complexos, especialmente em diálogos infantis, que podem incluir gírias, linguagem figurativa ou falta de clareza gramatical.

**Disponibilidade de Hardware:**

Treinar e executar modelos BERT pode ser intensivo em termos de recursos de hardware, como GPUs. Não ter acesso a hardware poderoso, pode limitar a escala e a eficiência do treinamento e inferência do modelo.

**Interpretação de Contexto:**

O BERT pode não ser capaz de interpretar completamente o contexto emocional ou social de uma conversa com crianças. A detecção de ironia, sarcasmo e nuances emocionais pode ser desafiadora.

**Generalização Limitada:**

O BERT pode ter dificuldade em generalizar bem para casos fora do conjunto de dados em que foi treinado. Se as interações com crianças forem muito diferentes dos dados usados no pré-treinamento, o desempenho do modelo pode ser prejudicado.

**Privacidade e Ética:**

A coleta e o uso de dados de crianças envolvem preocupações éticas e legais significativas. É importante garantir o consentimento adequado dos pais ou responsáveis e proteger os dados pessoais das crianças.

**Feedback Limitado:**

O feedback que as crianças fornecem pode ser limitado, especialmente em interações automatizadas. Isso pode tornar desafiador melhorar continuamente o sistema de diálogo com base no feedback do usuário.

**Complexidade do Domínio:**

Se o projeto envolver tópicos altamente especializados ou vocabulário técnico, o BERT pode não ser tão eficaz, pois não foi treinado especificamente neste domínio.