

MAIS 202 - Project Deliverable 1: Interview Practice Bot

Project Proposal by Luai Abuelsamen

Problem Statement

The project's goal is to create an Interview Practice Bot that uses the Stanford Question Answering Dataset (SQuAD) to provide a realistic interview practice environment. By employing advanced natural language processing, this bot aims to help users enhance their interview skills through practice and feedback.

Dataset

The SQuAD 2.0 dataset, blending over 150,000 questions, was preprocessed by selecting interview-relevant questions and modifying answers for conversational realism. Unanswerable questions were also included to teach the bot to handle ambiguous queries. This approach ensures the Interview Practice Bot is a practical tool for simulating real-life interviews, enhancing user preparedness and confidence.

Model

A transformer-based model like BERT was chosen for its proficiency in NLP tasks, implemented using PyTorch or TensorFlow alongside Hugging Face's Transformers library for ease of use and flexibility. The model architecture, optimized with techniques like dropout and AdamW, and hyperparameters such as batch size and learning rate, were fine-tuned based on validation set performance to prevent overfitting and ensure the model's effectiveness. The training and validation splits were derived from train-v2.0.json and dev-v2.0.json, respectively, to balance learning and evaluation. Challenges related to computational resources and model tuning were addressed through cloud-based solutions and iterative testing with the evaluate-v2.0.py script to assess performance metrics like Exact Match and F1 score. This comprehensive approach, emphasizing careful dataset preparation, model selection, and rigorous validation, underpins the project's aim to develop a reliable and user-friendly tool for interview preparation.