Advanced Generative Chatbot Study

Jonathan Agustin Eric Barnes Massimillano Repupilli

Chatbots are Everywhere!

Generative AI has become ubiquitous

Many Fortune 500 companies use chatbots

Rapid development of AI models

Project Overview: 3 Models & 1 Dataset

Leveraged HuggingFace transformers

Compared BERT, DistilBERT, RoBERTa

Used 2nd version of Stanford Question Answering Dataset (SQuADv2)

Our Project is All Open Source

Repository

https://github.com/aai520-group6/project

BERT

https://huggingface.co/aai520-group6/bert-finetuned-uncased-squad_v2

DistilBERT

https://huggingface.co/aai520-group6/distilbert-finetuned-uncased-squad_v2

RoBERTa

https://huggingface.co/aai520-group6/roberta-finetuned-uncased-squad_v2

Background Research
The Fascinating World of Transformers

Innovative self-supervised learning strategy

Masks words in sentence and guesses them

Predict if two sentences relate

Deeper understanding of language

Stanford Question Answering Dataset (SQuAD)

The second version of SQuAD has **unanswerable questions**
(answer not provided to model)

Analogy: Telling someone to meet at the secret hiding spot
without telling them where it is

100K+ answerable / 50K+ unanswerable questions

Preprocessing required

Model Initialization, Preprocessing, & Training

Initialized BERT, DistilBERT, RoBERTa

Used HuggingFace transformers

Finetuned on SQuADv2

Trained on Google Colab

Compared metrics

Evaluation

Validated models on SQuADv2

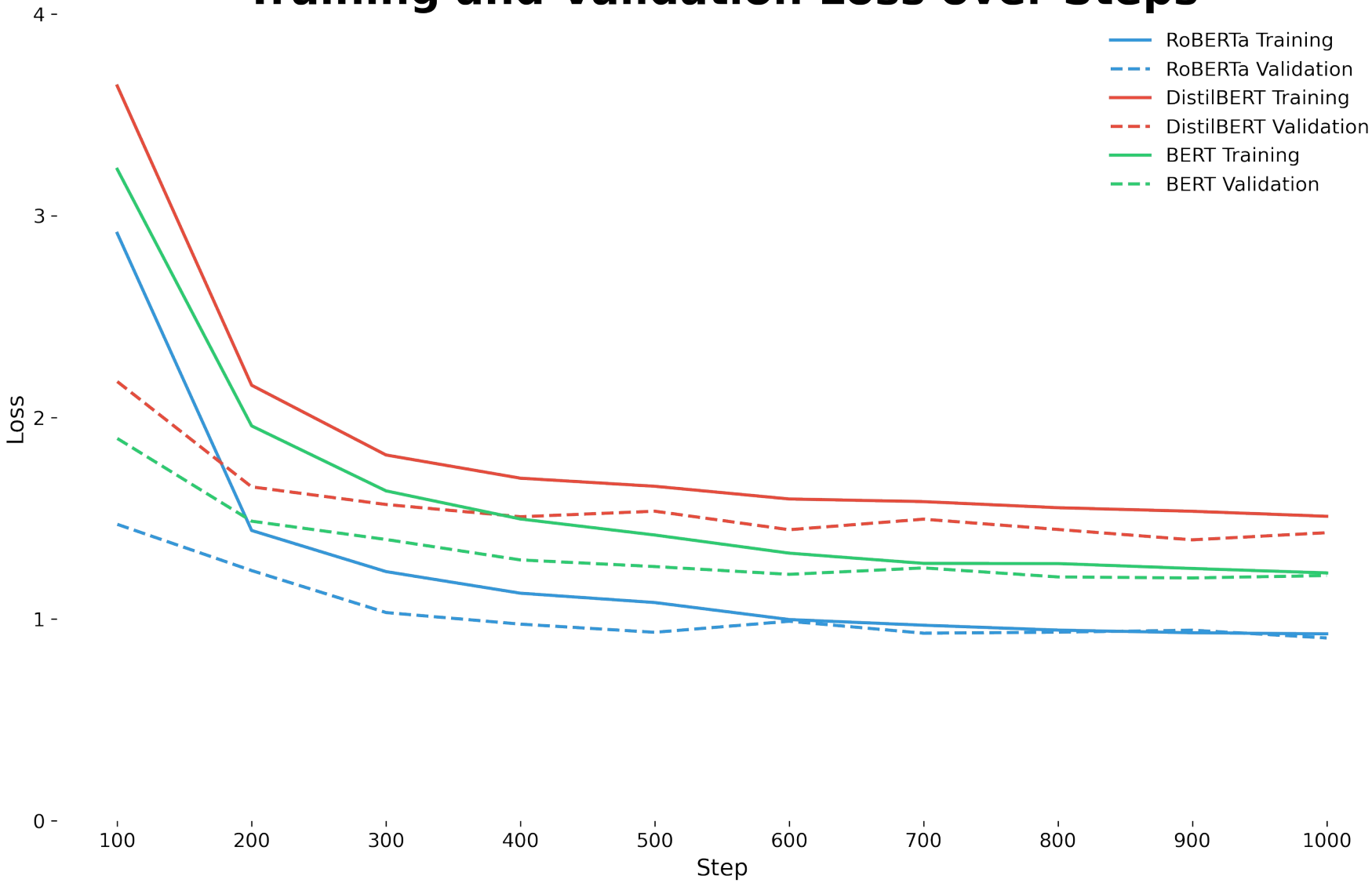
Measured over a dozen metrics

Focused on Exact Match and F1 Score

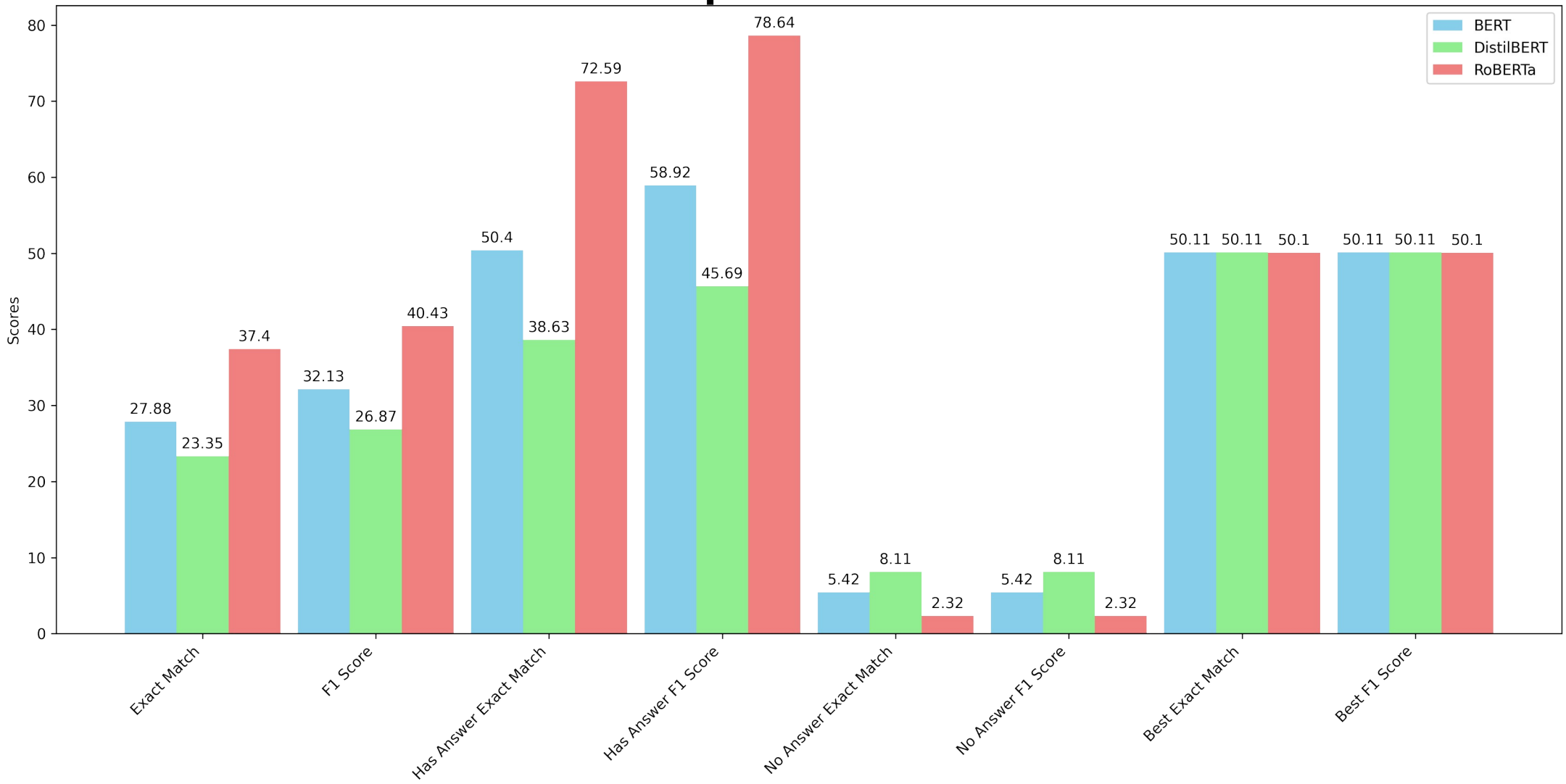
Analyzed answerable questions

Evaluated unanswerable questions

Training and Validation Loss over Steps



Fine-Tuned Model Comparison across Different Metrics



Discussion

Successfully developed 3 Q&A chatbots

Published code & models on GitHub and HuggingFace

High performance on answerable questions

Poor performance on unanswerable questions

Out-of-Memory problems mitigated with checkpointing

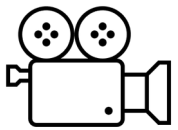
Conclusion and Recommendations

Low scores for unanswerable questions

Continue work on answerability detection

Enable explainability for model introspection

Contributions

Jonathan Agustin	Eric Barnes	Massimillano Repupilli
		
Video	Report	PowerPoint

Each team member independently trained and evaluated all models.