Suicide and Emotion Prediction

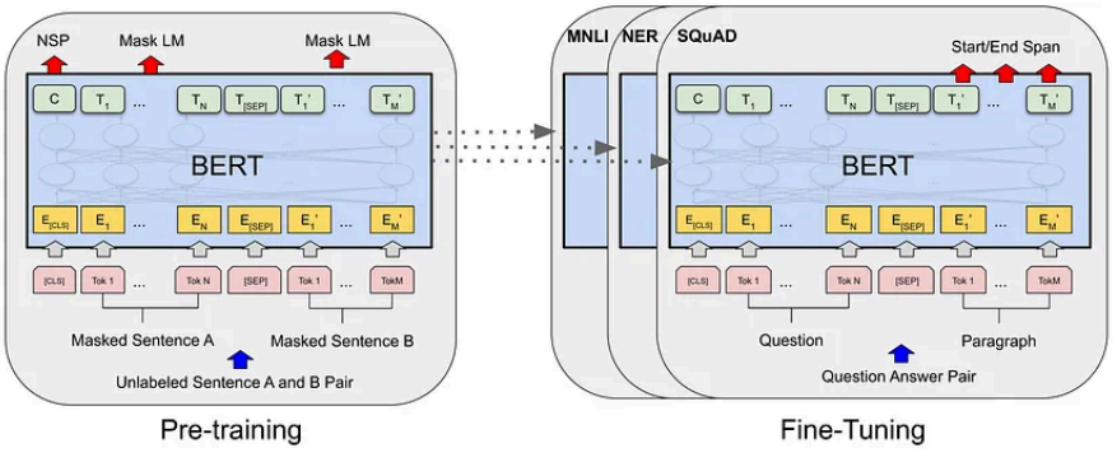
We finetuned RoBERTa to predict the suicidal thought probability or emotion associated with particular text. RoBERTa, which stands for "Robustly Optimised BERT Approach," is a variant of the BERT (Bidirectional Encoder Representations from Transformers) model, developed by Facebook AI researchers. Like BERT, RoBERTa is a transformer-based language model that leverages self-attention to process input sequences and create contextualised word representations within sentences.

A key distinction between RoBERTa and BERT lies in its training approach. RoBERTa was trained on a significantly larger dataset—160GB of text, over 10 times the size of BERT's training data. Additionally, RoBERTa employs a dynamic masking technique during training, allowing it to generate more robust and generalizable word representations.

RoBERTa has consistently outperformed BERT and other top models across various natural language processing tasks, such as language translation, text classification, and question answering. It has also served as the foundation for numerous successful NLP models, gaining popularity in both research and industry applications.

In summary, RoBERTa is a highly effective and influential language model that has advanced the field of NLP and played a pivotal role in the development of a wide range of applications.

RoBERTa was introduced by Facebook and Washington University researchers, with the goal of optimising BERT's training process to reduce pre-training time.



Fine-Tuning BERT

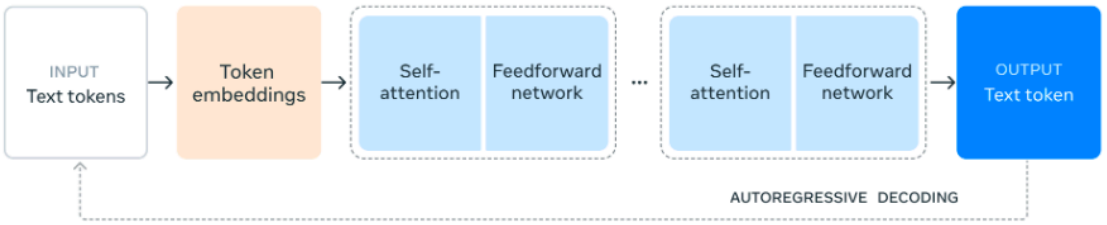
Conversation Chatbot

Llama 3.1 is a new state-of-the-art model from Meta available in 8B, 70B and 405B parameter sizes.For this release, Meta has evaluated the performance on over 150 benchmark datasets that span a wide range of languages. In addition, Meta performed

extensive human evaluations that compare Llama 3.1 with competing models in real-world scenarios. Meta's experimental evaluation suggests that our flagship model is competitive with leading foundation models across a range of tasks, including GPT-4, GPT-4o, and Claude 3.5 Sonnet. Additionally, Meta's smaller models are competitive with closed and open models that have a similar number of parameters.

Category / Benchmark	Llama 3.1 8B	Gemma 2 9B IT	Mistral 7B Instruct	Llama 3.1 70B	Mixtral 8x22B Instruct	GPT 3.5 Turbo
General MMLU (0-shot, CoT)	73.0	72.3 (5-shot, non-CoT)	60.5	86.0	79.9	69.8
MMLU PRO (5-shot, CoT)	48.3	-	36.9	66.4	56.3	49.2
IFEval	80.4	73.6	57.6	87.5	72.7	69.9
Code HumanEval (0-shot)	72.6	54.3	40.2	80.5	75.6	68.0
MBPP EvalPlus (base) (0-shot)	72.8	71.7	49.5	86.0	78.6	82.0
Math GSM8K (8-shot, CoT)	84.5	76.7	53.2	95.1	88.2	81.6
MATH (0-shot, CoT)	51.9	44.3	13.0	68.0	54.1	43.1
Reasoning ARC Challenge (0-shot)	83.4	87.6	74.2	94.8	88.7	83.7
GPQA (0-shot, CoT)	32.8	-	28.8	46.7	33.3	30.8
Tool use BFCL	76.1	-	60.4	84.8	-	85.9
Nexus	38.5	30.0	24.7	56.7	48.5	37.2
Long context ZeroSCROLLS/QuALITY	81.0	-	-	90.5	-	-
InfiniteBench/En.MC	65.1	-	-	78.2	-	-
NIH/Multi-needle	98.8	-	-	97.5	-	-
Multilingual Multilingual MGSM (0-shot)	68.9	53.2	29.9	86.9	71.1	51.4

Model Architecture



References

Meta Llama3.1 blog
BERT Explained
RoBERTa