

The overall performance of the roberta-large model output results in par with the results within the research paper "Trillion Dollar Words", particularly for the features "time_label" and "certain_label." The model achieved an accuracy of 78.57% and 82.14%, respectively, with strong F1 scores of 0.78 and 0.84. However, for the feature "sentiment_label," the accuracy and precision (61.23% and 67.13%) suggest that the model requires more fine-tuning to better capture the nuances of the dataset and fully reproduce the findings in the research paper. Additional optimization of hyperparameters and training epochs could minimize the gap between the current results and the benchmarks set by the research.