STAR MODEL: classifying interviews for accurate feedback

Andrew Warkentin

What is STAR?

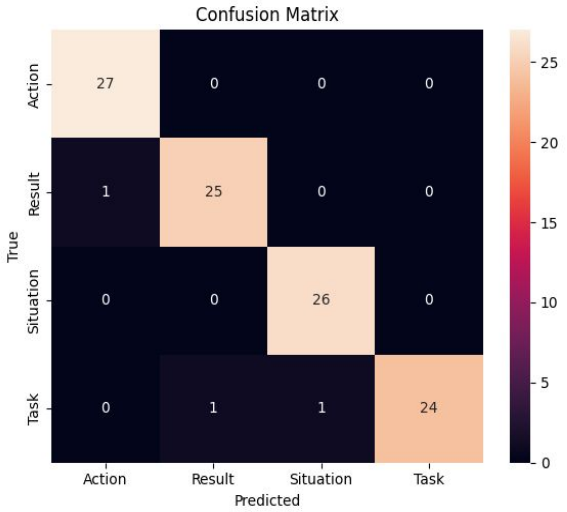
- Behavioral interviewees frequently utilize the STAR (Situation, Task, Action, Result) method as a framework to structure responses.
- Issue is evaluating completeness, both in terms of time spent on each section and time spent on each.
- BERT-based classifier with sentence-level sequence labeling.

How does BERT classifier work?

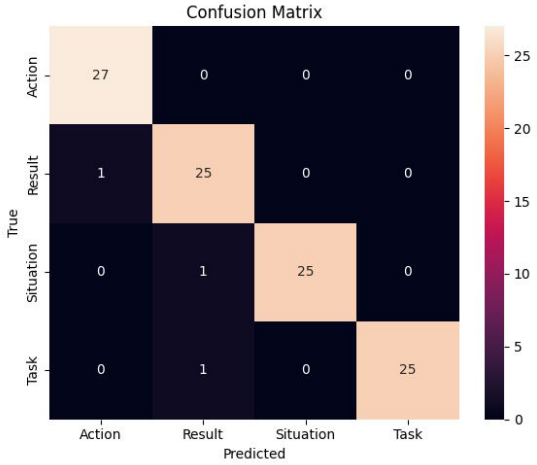
- Pre-trained on masked-language modeling → rich bidirectional context
- Add a task head: `[CLS]` embedding → linear layer → soft-max over STAR labels
- Fine-tune end-to-end on < 1 k labeled sentences (low-data friendly)
- Self-attention lets every token “see” every other token in the sentence

Fine-tuned Models

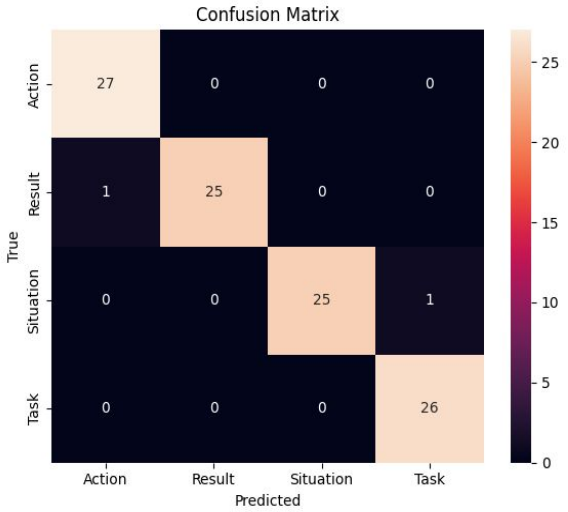
Bio-bert



Distilbert



bert-base-uncased

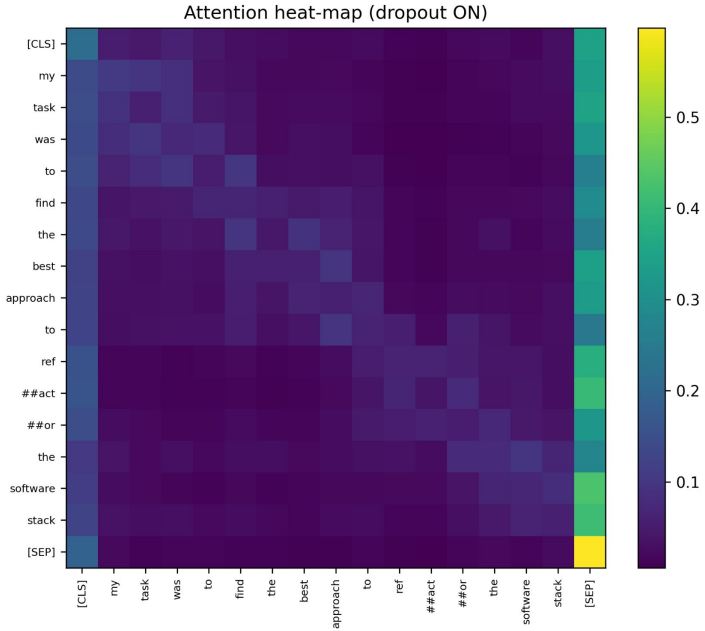


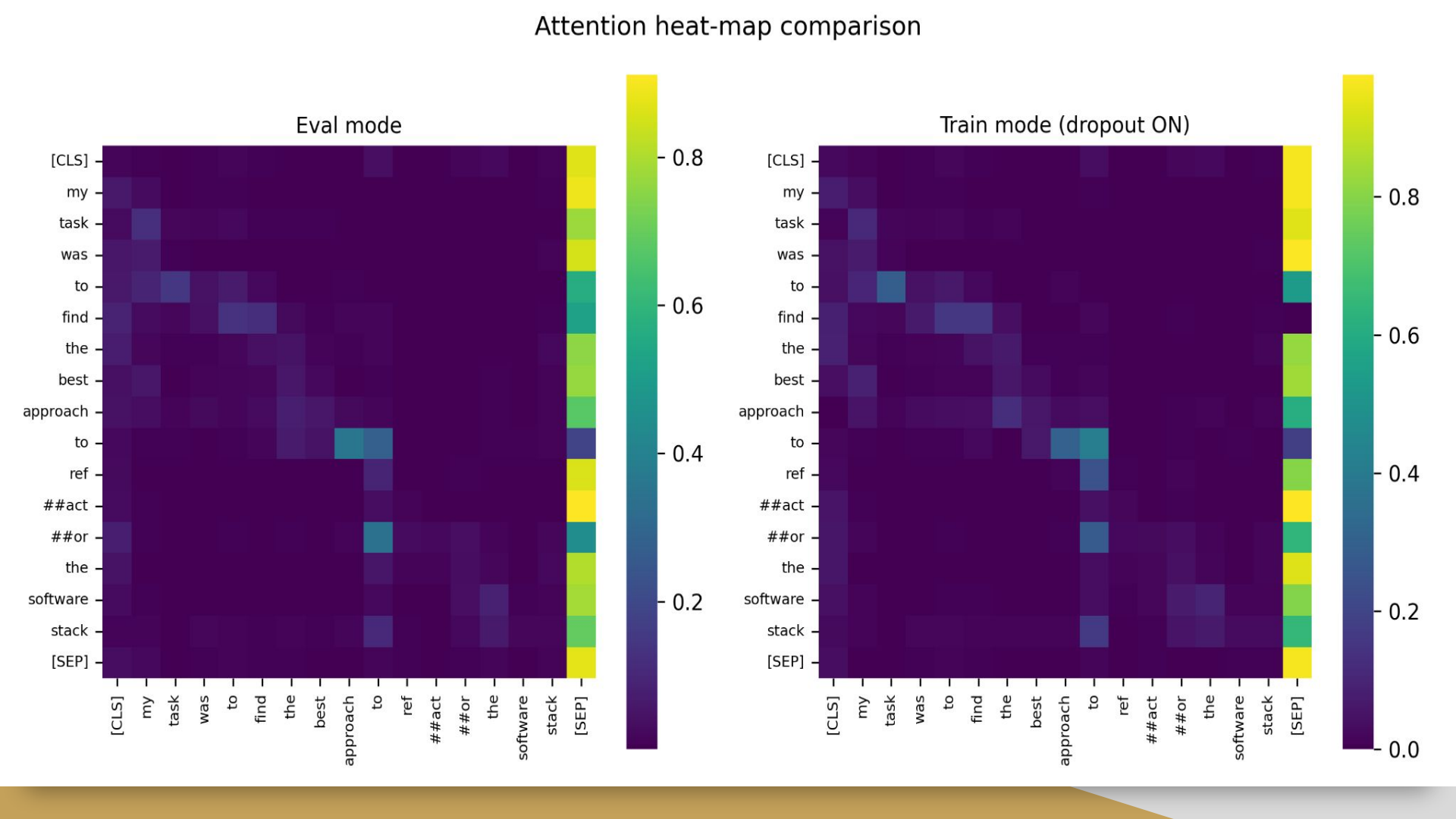
What is AdamW Optimization?

- Scale the learning rate by an accumulated gradient.
- Faster convergence on small batches; no loss oscillations, less over-fitting

How does dropout help?

- Randomly zeroes 10 % of activations each forward pass
- Forces learning **redundant, distributed** features
- Results: more robust attention maps (less peaky) around certain words like “the”

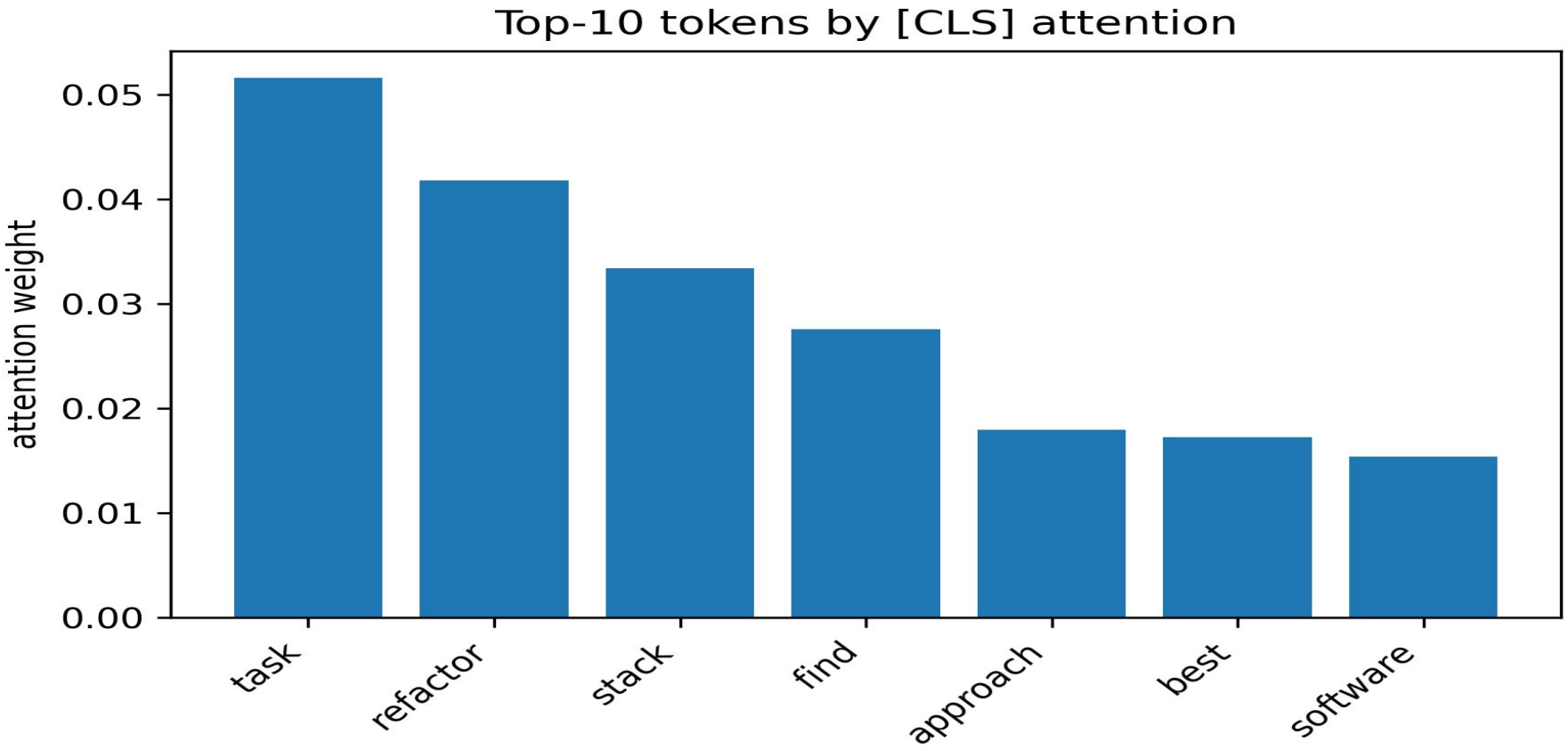



Attention heat-map comparison
Eval mode
Train mode (dropout ON)
[CLS]
my
task
was
to
find
the
best
approach
to
ref
##act
##or
the
software
stack
[SEP]
0.0
0.2
0.4
0.6
0.8

How does Attention help?

- **Scaled dot-product** compares every token pair in parallel
- Multi-head setup → different heads learn syntax, co-reference, cue words
- Enables long-range dependencies (e.g., “because … so the result…”)

“My task was to find the best approach to refactor the software stack” -> Task

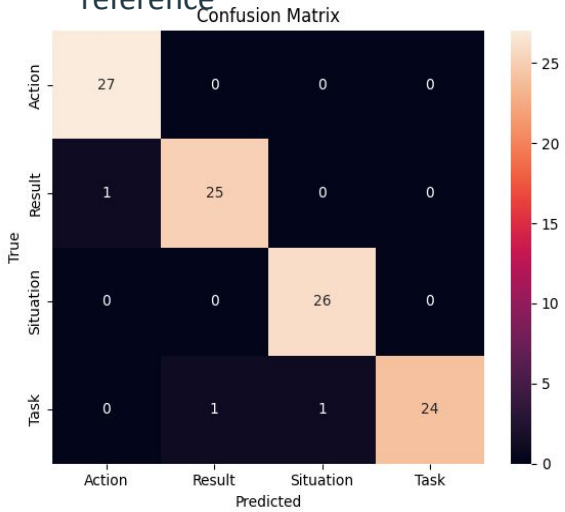


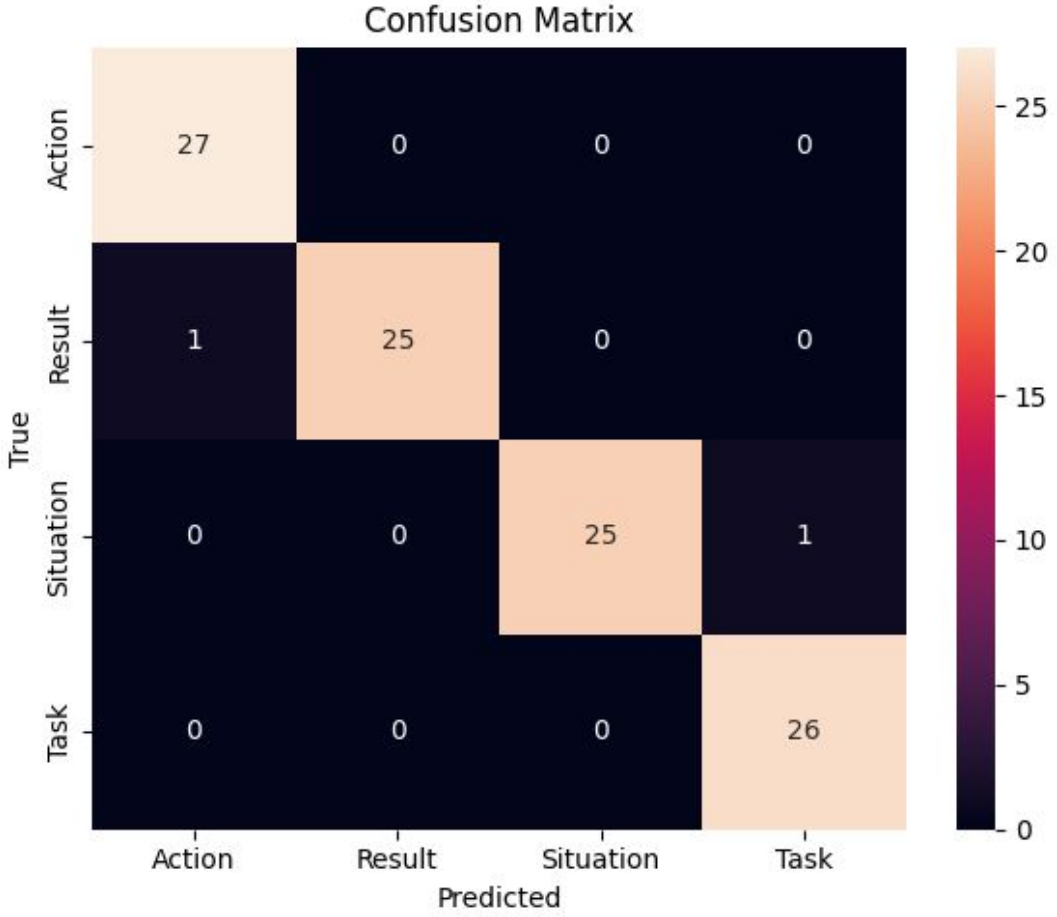
How can we verify the results?

- **Confusion matrix** + per-class precision/recall to spot skewed labels
- **Macro-F1 & loss curves** → watch for over-fit / under-fit crossover
- **Qualitative error review**: inspect top-attention tokens on mis-predictions

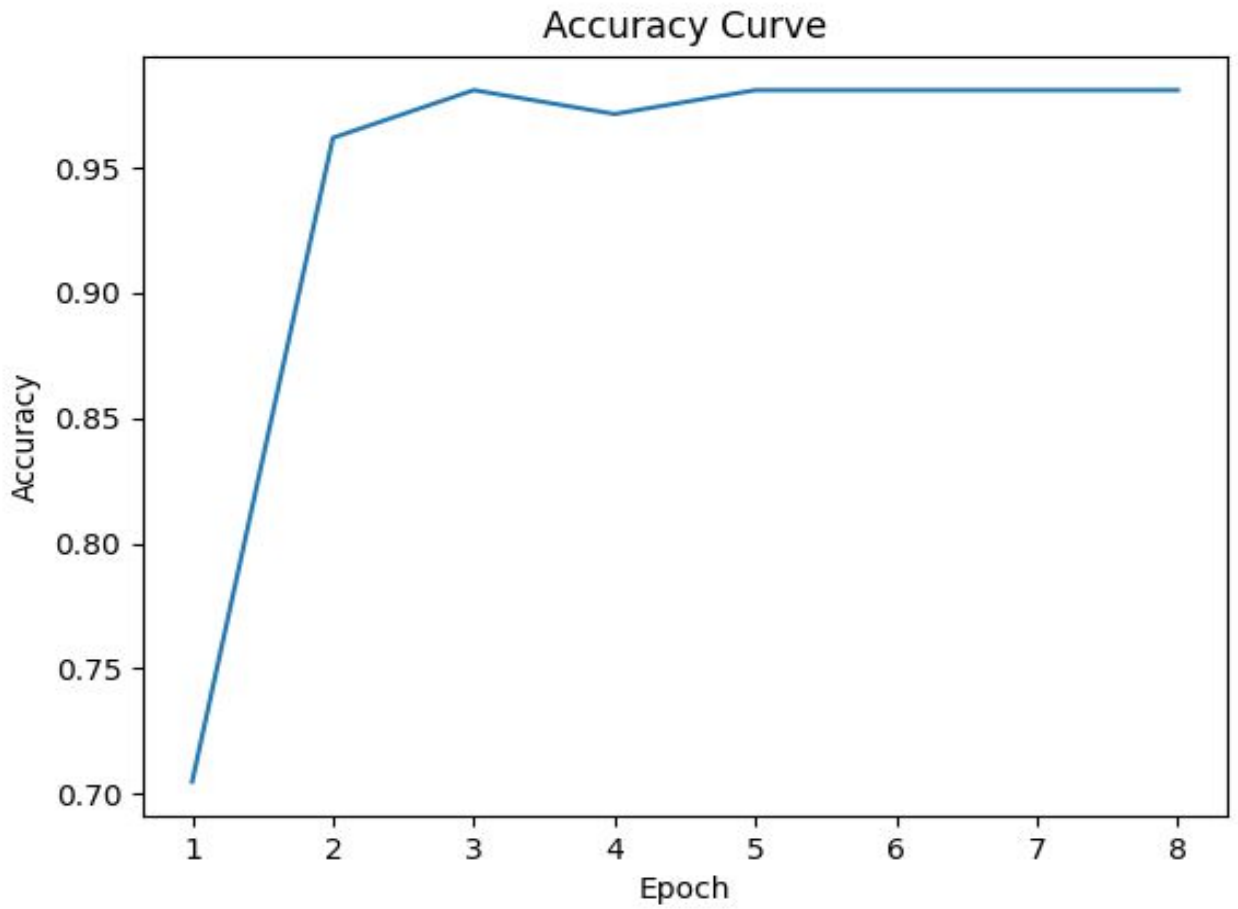
Confusion Matrix (bert-base-uncased)

BIO-BERT for reference

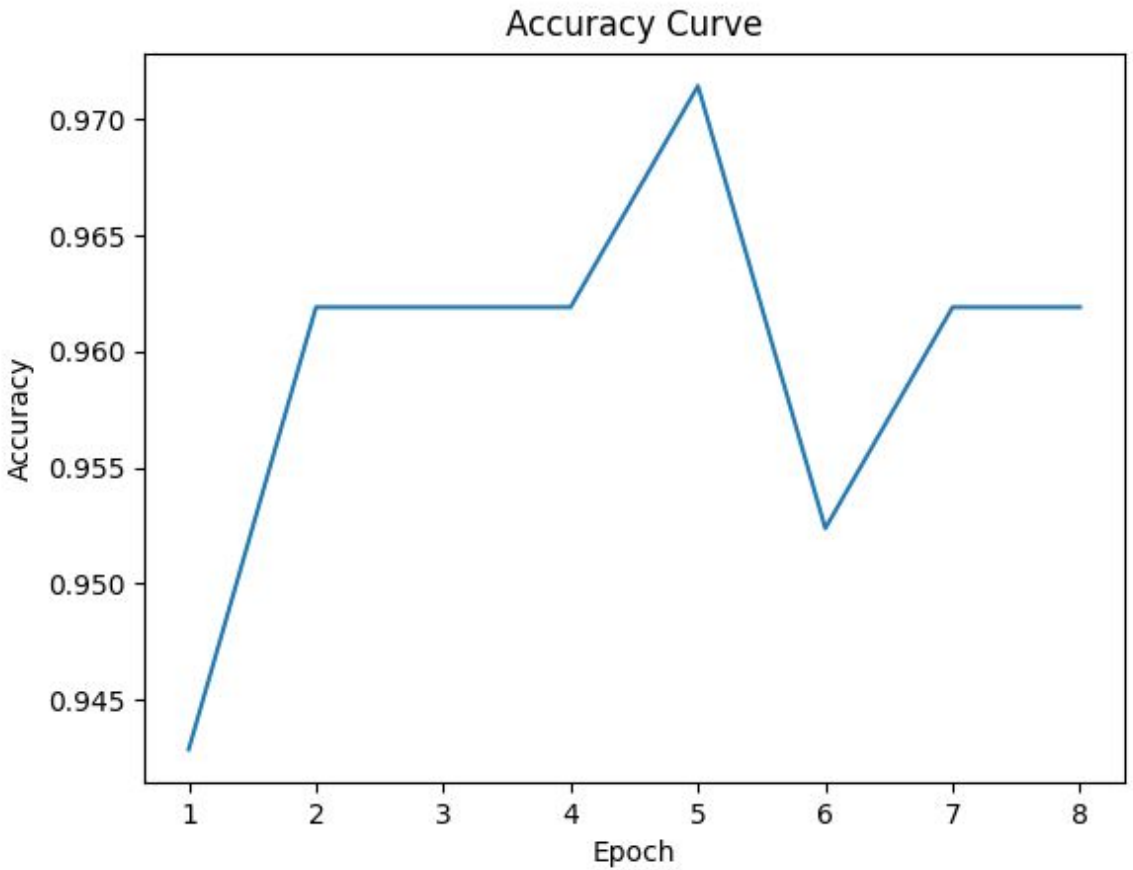




Accuracy (bert-base -uncased)

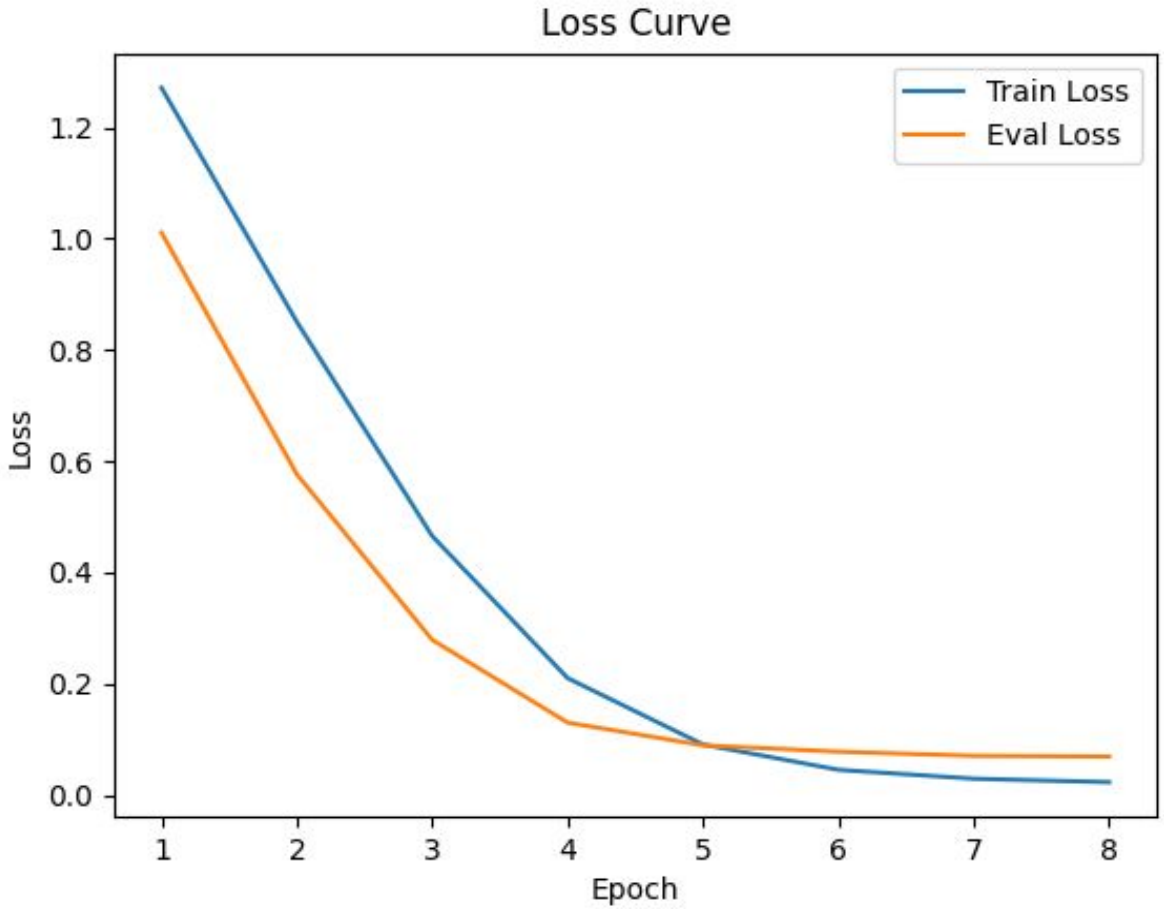


Comparison (bio-bert)



Cross-Entropy Loss

Bert-base-uncased



Thanks for watching!