Homework 8

Anshul Sawant

March 22, 2025

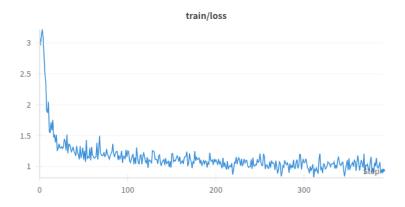
Contents

1	Q 1.1 1.1 B]
2	Q 1.3: Training and Evaluation 2.1 A: Training Loss	6 4 6 4
3	Q 1.4 Cross-encoding 3.1 A	
1 1.	Q 1.1 1 B	

It is sounder to use last-token pooling with decoder-only models because decoders process tokens left to right and a decoder would not have processed all of the input before the last token.

2 Q 1.3: Training and Evaluation

2.1 A: Training Loss



2.2 B: Mean Reciprocal Rank

The mean reciprocal rank achieved is 0.3182. This implies that on an average the first "correct" answer is at third position.

3 Q 1.4 Cross-encoding

3.1 A

```
def cross_encoder(query, passage):
    reprs = encoder(query + '[SEP]' + passage)
    similarity = Linear(reprs)
    return similarity
```

3.2 B

Advantage: Higher accurary because jointly considering query and passages allows for more nuanced semantic relationships. **Disadvantage**: Each query-passage pair requires a forward pass of the encoder making it more expensive. A bi-encoder will requires |Q| + |P| forward passes whereas, a cross-encoder will require $|Q| \times |P|$ forward passes.

3.3 C: Strategy to use both

A candidate set can be retrieved using bi-encoder and then this smaller set can be re-ranked using cross-encoder.