输入特征

1. content_id
2. answered_correctly
3. part
4. prior_question_elapsed_time
5. prior_question_had_explanation
6. lag_time1 - convert time to seconds. if lag_time1 >= 300 than 300.
7. lag_time2 - convert time to minutes. if lag_time2 >= 1440 than 300 (one day).
8. lag_time3 - convert time to days. if lag_time3 >= 365 than 365 (one year).

I found lag time split to different time format boosting score around 0.003.

Transformer

Encoder Input

- question embedding
- part embedding
- position embedding
- prior question had explanation embedding

Decoder Input

- position embedding
- reponse embedding
- prior elapsed time embedding
- lag_time1 categorical embedding
- lag_time2 categorical embedding
- lag_time3 categorical embedding
- Note that I have tried categorical and continuous embedding in prior elapsed time and lag time. The performance of categorical embedding is better than continuous embedding.

参数

- max sequence: 100
- d model: 256
- number of layer of encoder: 2
- number of layer of decoder: 2
- batch size: 256
- dropout: 0.1
- learning rate: 5e-4 with AdamW