
進捗報告

1 今週やったこと

- パラメータの最適化
- 実験結果の分析

2 パラメータの最適化

CLS トークンを使うモデルと使わないモデル両方でもパラメータをチューニングしたけど、CLS トークンを用いた場合はどう調整しても使わない場合より精度が低い。

今回は Tranformer レイヤーの数、一つブロックの次元数、multi head attention のヘッド数をチューニングした。三つのパラメータを一緒にチューニングした結果は納得出来なかったのので、三つのなかの二つを固定して、別々でまたチューニングした。

結果としては、三つのパラメータを一緒にチューニングした場合は、レイヤー数 7, 次元数 112, ヘッド数 7。

別々でチューニングした結果は、次元数は 104 から 128 が最適、128 以上はそんなに変わらない。レイヤー数とヘッド数は結果への影響が小さい、

3 実験結果の分析

最初から黙テンを予測出来なかったのと回し打ちする時のテンパイを予測出来なかったのが多い。

意外と「降りる」に対しての予測はできている。

捨て牌と合わせて結果を分析したら、モデルは何で判断しているのかわかるかもしれない。

四人モデルと一人モデルを比較したい。