

進捗報告

]

1 今週やったこと

- シャンテン数差の実装
- データを増やして平均化
- 三人同時予測モデル

2 シャンテン数差の実装

LOSS に catagorical crossentropy とシャンテン数差 mean squared error の和を使った. 精度は少し上がった.

表 1: 新 loss シャンテン数予測

| | precision | recall | f1 | 元値 | support |
|------------|-----------|--------|--------|--------|---------|
| 0(テンパイ) | 0.6710 | 0.5229 | 0.5878 | 0.5621 | 6735 |
| 1(1 シャンテン) | 0.5270 | 0.4699 | 0.4968 | 0.4847 | 19631 |
| 2(以外) | 0.7956 | 0.8621 | 0.8275 | 0.8296 | 43204 |
| acc | | | 0.7186 | 0.7174 | 69570 |

3 データを増やして平均化

テンパイと 1 シャンテンのデータを増やして (データ数は 1:1:1, 実際にリーチデータを除いたら, 1:1.67:1.67), シャンテン予測モデルとテンパイ予測モデルで実験した.

シャンテン数予測モデルに対して, テンパイと 1 シャンテンの精度上がった代わりに, 2 シャンテン以上の精度が下がった. テストデータは実際の麻雀をやる時の配分なので, 2 シャンテン以上のデータ数が多いため, 全体的精度が下がった. どの部分が重要なのかは不明である, ROC みたいな評価手法が必要である. 一方, テンパイ予測モデルを平均化したデータ (1:3.34) で学習した結果は,

表 2: データ平均化したシャンテン数予測

| | precision | recall | f1 | support |
|------------|-----------|--------|--------|---------|
| 0(テンパイ) | 0.5296 | 0.6937 | 0.6006 | 6735 |
| 1(1 シャンテン) | 0.4645 | 0.6288 | 0.5343 | 19631 |
| 2(以外) | 0.8755 | 0.6925 | 0.7733 | 43204 |
| acc | | | 0.6746 | 69570 |

表 3: 元のシャンテン数予測

| | precision | recall | f1 | support |
|------------|-----------|--------|--------|---------|
| 0(テンパイ) | 0.7031 | 0.4778 | 0.5690 | 6735 |
| 1(1 シャンテン) | 0.5366 | 0.4444 | 0.4862 | 19631 |
| 2(以外) | 0.7844 | 0.8848 | 0.8316 | 43204 |
| acc | | | 0.7211 | 69570 |

表 4: テンパイ予測モデル

| | precision | recall | f1 | 元値 | support |
|------------|-----------|--------|--------|--------|---------|
| 0(テンパイ) | 0.9504 | 0.9711 | 0.9606 | 0.9582 | 6735 |
| 1(1 シャンテン) | 0.6614 | 0.5272 | 0.5867 | 0.5026 | 19631 |
| acc | | | 0.9281 | 0.9229 | 69570 |
| ROC 面積 | | | 0.9234 | 0.9191 | |

4 三人同時予測モデル

まず、簡単な 4 人分の打牌を入力したモデルを作った。鳴きをした時飛ばされた順目を表す部分を実装。簡単に実験して、精度が非常に低いである、3 人予測は 0.52、一人だけ予測は 0.82。予測目標の手牌を強調する手法を探している。3 人同時予測は三つの transformer モデルが必要と考えられる。一方、モデルが膨大になったので、学習進むのは非常に遅い。