

SelfAttention を用いた麻雀における山牌予測

高木 俊太郎[†] 鈴木 海友[†] 松澤 智史[‡]

1. はじめに

近年、完全情報ゲームでは人間のトッププレイヤーと同等かそれ以上の実力を持つ AI が開発されている。一方、不完全情報ゲームでは不確定の情報を把握できないことなどから AI の苦手な分野とされていた。しかし近年では、ポーカー [1] や麻雀 [2] のような不完全情報ゲームの AI も急速な発展を遂げている。

現在、麻雀 AI の実装方法として「打牌モデル」や「副露モデル」といった複数のモデルを組み合わせる手法が有力である。麻雀 AI 開発ではこれらの部分モデルの精度を高める取り組みがなされている。したがって、麻雀 AI の部分モデルである「山牌予測モデル」が高精度になれば麻雀 AI の発展に寄与できる。

2. 研究目的

麻雀において勝利するために必要な技術として「山牌予測」がある。これは河や副露などの公開情報から非公開情報である山にどの牌が何枚残っているかを予測する技術である。麻雀で勝利するためには、他プレイヤーより多く点数を所持している必要がある。そのため点数を獲得する機会を多く作り、より多く点数を獲得することが重要である。山牌にどの牌が何枚残っているかという情報を得ることができれば自分の手を効率よく進めることができ、得点の機会となる。よって山牌を高い精度で予測することは勝利するために重要である。本研究では、山牌の予測を行うことを目的とする。

3. 関連研究

青木ら [3] は各プレイヤーの捨て牌情報と副露の情報を用いて山牌を推定する機械学習モデルを提案した。この研究では、他プレイヤーが立直を行った時点での山牌を提案モデルで推定した。評価には各牌の残り枚数の正答率と各牌が枚数を問わず残っているかどうかの二つを用い、ランダム予測との比較実験を行った。結果として、どちらの評価もランダム予測よりも高い評価が得られた。

青木らの研究では、捨て牌情報と副露の情報を用いて予測を行ったが、特徴量として「牌が捨てられた順番」が存在しない。そこで本研究では「牌が捨てられた順番」を考慮するモデルを構成し、予測を行う。

4. 提案手法

4.1 概要

本研究では打牌データや副露データから山牌を予測する手法を提案する。「牌が捨てられた順番」を考慮するために BiLSTM [4] を用いてモデルを構成する。また、SelfAttention [5] を用いることにより、捨て牌データのどの部分に注目するかを踏まえて山牌を予測する。

なお青木らの研究では牌の残り枚数の確率を出力し、正答率を評価している。しかし、正答率の評価では予測枚数の誤差の差を考慮できない。よって誤差の大小関係を考慮するため、本研究では予測した枚数そのものを出力するモデルとする。

4.2 入出力データ

本研究の入力データは主入力と補助入力の 2 種類を用意した。

主入力は全プレイヤーの捨て牌を時系列として扱う入力である。特徴量は、捨てたプレイヤーを表す 4 次元、手出しか自摸切りかを表す 1 次元、牌の種類を表す 34 次元の合計 39 次元である。時系列データとしての時間長は局の開始から打牌が 40 回行われるまでである。

補助入力は視点となるプレイヤーから見える牌を扱う入力である。特徴量は、視点プレイヤーを表す 4 次元、各牌の見えた枚数 34 次元の合計 38 次元である。

また、出力データは正解ラベルとして牌 34 種類に対して実際に山に残っている牌の枚数を与える。

4.3 ニューラルネットワークの構成

提案モデルの構造を図 1 に示す。図 1 の Main Input が主入力を渡す層であり、Sub Input が補助入力を渡す層である。入力された主入力は BiLSTM 層に渡され、次に SelfAttention 層に渡される。その後補助入力と連結するために Flatten 層に渡される。Flatten 層の出力は補助入力と連結され、最終的に 6 層の全結合層に渡される。この全結合層の中間層のユニット数は 256、活性化関数は ReLU [6] とした。最終層の活性化関数は linear 関数、損失関数は平均二乗誤差とした。

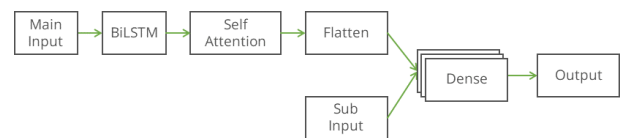


図 1: 提案モデルの構造

5. 実験

5.1 データセット

実験で用いるデータセットはオンライン麻雀サイト天鳳[§]より取得した牌譜のうち、上位 0.7% のプレイヤーの牌譜を用いた。学習には 360,000 局のデータを用いた。

5.2 実験 1

本実験では、モデルが予測した各牌の山牌の枚数を式 (1) で評価する。

$$\text{loss} = \sqrt{\frac{1}{34} \sum_{i=1}^{34} (y_i - \hat{y}_i)^2} \quad (1)$$

[†]東京理科大学 理工学研究科 情報科学専攻[‡]東京理科大学 理工学部 情報科学科[§]<https://tenhou.net/>

$loss$ は誤差平均, i は各牌に対するインデックス, y_i は牌 i が山牌に残っている枚数, \hat{y}_i は牌 i を予測した枚数である. この評価では誤差平均が低いほど高い精度で予測できているといえる.

なお予測モデルとの比較のために, 捨て牌と副露と手牌から見えている牌を除いた牌の中から単純に確率で予測した際の誤差平均と一般プレイヤーの予測の誤差平均も算出した. 一般プレイヤーの予測は, 被験者 1 人に視点プレイヤーの手牌とプレイヤーの捨て牌データ, 副露の情報を実際の麻雀の盤面通り与え, 34 種類の牌が何枚山牌に残っているかを整数で予測させた.

各予測の結果を表 1 に示す.

表 1: 山牌予測の誤差平均

	テストデータ数	誤差平均
ランダム	40000	1.117
一般プレイヤー	100	0.981
予測モデル	40000	0.767

本研究の予測モデルは, ランダム予測と一般プレイヤーの予測よりも良い精度を得ることができた. (表 1)

5.3 実験 2

本実験ではモデルが予測した各牌の山牌の残り枚数を用いて, 青木らの山牌予測モデルとの比較を行う. 比較には, 青木らと同様に牌 34 種が山牌の中に残っている枚数の正答率と, 牌 34 種が枚数を問わず残っているか否かの残存正答率を用いた.

また青木らは 0 から 24 枚, 25 から 48 枚, 49 から 82 枚といった形で山牌残り枚数ごとに分けて正答率を算出している. 本研究での山牌残り枚数は牌の枚数 136 枚から打牌された牌 40 枚と 4 人のプレイヤーの手牌 52 枚を除いた 44 枚である. したがって比較は 25 から 48 枚のときの評価値を用いる.

なお本研究の予測モデルでは山牌に残っている枚数が小形で出力されるので, 出力された値を四捨五入し, それを予測枚数として評価を算出した.

山牌に残っている枚数の正答率の予測の結果を表 2 に, 残存正答率の結果を表 3 に示す.

表 2: 山牌の中に残っている枚数の予測の評価

	山牌残り枚数	残り枚数正答率
青木らのモデル	25 から 48 枚	0.5471
本研究のモデル	44 枚	0.5185

表 3: 枚数を問わず残っているか否かの予測の結果の評価

	山牌残り枚数	残存牌正答率
青木らのモデル	25 から 48 枚	0.8261
ランダム	25 から 48 枚	0.6890
本研究のモデル	44 枚	0.8519

各種残り枚数正答率では本研究のモデルの正答率が青

木らのモデルの正答率より低くなり, 山牌残り枚数では正答率が高い結果となった. (表 2)(表 3)

6. 評価と考察

実験 1 ではランダム予測と比べ提案モデルの誤差平均が低いことから正しく学習できた. また一般プレイヤーの予測と比べても低い誤差平均となっており, 一般プレイヤーよりも高い精度で予測ができた.

実験 2 では各種残り枚数正答率において, 本研究のモデルが青木らのモデルよりも正答率が低くなる結果となった. これは, 山牌予測を行う条件の違いに原因がある. 本研究の条件では山牌残り枚数は 44 枚であり, 青木らの条件である 25 から 48 枚の中では多い部類に属する. 山牌残り枚数が少ないほど不確定情報が増えるので, 本研究では予測が難しくなった.

残存牌正答率では本研究のモデルが青木らのモデルの正答率を上回った. これは本研究のモデルが BiLSTM 及び SelfAttention を用いたことによる優位性があったといえる.

7. おわりに

本研究では捨て牌や副露などの公開情報から非公開情報である山牌の予測を行った. 結果として, 実験 1 ではランダム及び一般プレイヤーの予測の精度を上回った. 実験 2 で青木らのモデルと比較をしたところ, 残り枚数正答率では評価値を下回ったが, 残存牌正答率では評価値が高いモデルを作成することができた.

参考文献

- [1] N. Brown and T. Sandholm “Superhuman AI for multiplayer poker”, Science, vol. 365, no. 6456, pp. 885–890, 2019.
- [2] Junjie Li, Sotetsu Koyamada, Qiwei Ye, Guoqing Liu, Chao Wang, Ruihan Yang, Li Zhao, Tao Qin, Tie-Yan Liu, Hsiao-Wuen Hon, “Suphx: Mastering Mahjong with Deep Reinforcement Learning”, CoRR, vol. abs/2003.13590, 2020.
- [3] 青木 龍汰, 永井 秀利, 中村 貞吾, “麻雀におけるプレイヤの捨て牌に基づく山牌推定”, 電気関係学会九州支部連合大会講演論文集, vol. 2020, p. 233, 2020.
- [4] Mike Schuster, Kuldeep K Paliwal, “Bidirectional recurrent neural networks”, IEEE Transactions on Signal Processing, vol. 45, Issue 11, pp. 2673–2681, 1997.
- [5] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin, “Attention is All you Need”, CoRR, vol. abs/1706.03762, 2017.
- [6] Vinod Nair, Geoffrey E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines”, ICML, pp. 807–814, 2010.