

機械学習を用いた麻雀戦術における状況予測手法の提案

第 1 グループ 青野 義樹

1. はじめに

近年、機械学習や深層学習の手法に基づく、ゲームをプレイする人工知能プログラムが盛んに開発されている。本研究では不完全情報ゲームの 1 種である麻雀を対象とした。

麻雀では、あるプレイヤーが捨てた牌の種類や捨てた順番などから、そのプレイヤーが現在手持ちになっている牌の状態の予測が可能である。この予測は戦術構築の際に重要であり、一般的に上級者ほどその予測精度は高いと考えられている。水上らの研究 [1] では、ロジスティック回帰と線形回帰を用いたモデルが人間の上級者と同程度の予測精度を得ているが、モデルの単純さや特徴量の設計などの面から、局面中の複雑な特徴を十分に捉えて予測しているとはいえない。そこで本研究では、Recurrent Neural Network (RNN) の一種である Long Short-Term Memory (LSTM) によるニューラルネットワークモデルを用いて、捨てた牌の時系列的な特徴を捉えることで予測精度の向上を試みる。

2. 麻雀の用語とゲーム進行について

以下に麻雀のゲーム進行と用語について説明する。

麻雀は 4 人のプレイヤーによってプレイされる。1 ゲームは局と呼ばれる単位によって区切られている。局の始めに、各プレイヤーはそれぞれ 13 枚の牌を手持ちにしており、1 枚手持ちに加える行為と、1 枚手持ちから捨てる行為を繰り返すことで牌を入れ替えていく。牌の組合せで 3 枚の組 4 セットと 2 枚の組 1 セットという特定の条件を満たすことでアガリとなり、アガったプレイヤーが点を得て局が終了する。局を複数回 (通常 8 回程度) 繰り返すことで 1 ゲームが終了する。1 ゲームの終了時に最も得点の多いプレイヤーの勝利となる。また、局の最中、プレイヤーは他プレイヤーがどの種類の牌を何枚手持ちしているか直接知ることはできず、この点から麻雀には不完全情報性があるといえる。

- 鳴き
他プレイヤーが捨てた牌を利用して自身の手持ち牌をアガリへ近付ける行為。
- テンパイ
アガリの一つ手前の状態。他プレイヤーの捨てた牌や鳴いた牌の情報などからそのプレイヤーの手持ち牌がテンパイ状態にあるかの予測をテンパイ予測という。

3. 数値実験

以下に、テンパイ予測の実験方法を示す。用いたデータはオンライン麻雀サイトの対戦ログファイルから、特定の局面のデータを取り出したものである。学習には 2.8×10^6 個のデータを、テストには 6.0×10^4 個のデータをそれぞれ用いた。

実験は局面中のある 1 人のプレイヤーを対象にした。上記の局面のデータの内、そのプレイヤーの捨てた牌、鳴いた牌、リーチをかけているかなどの情報から学習に用いる入力の特徴ベクトルと出力のラベルを生成した。具体的には、捨てた牌 1 枚を 19 次元の特徴量ベクトルで表した。捨てた牌全体は、ルール上の捨てた牌の最大枚数に合わせて 19×24 次元の行列で表した。鳴いた牌も同様に、1 セットを 14 次元の特徴量ベクトルで表し、鳴いた牌全体を 14×4 次元の行列で表した。入力の素性ベクトルには、これらの行列と局面中のその他の情報を表すベクトルを用いた。出力のラベルには、対象プレイヤーがテンパイしているかしていないかを表す 2 クラスを用いた。

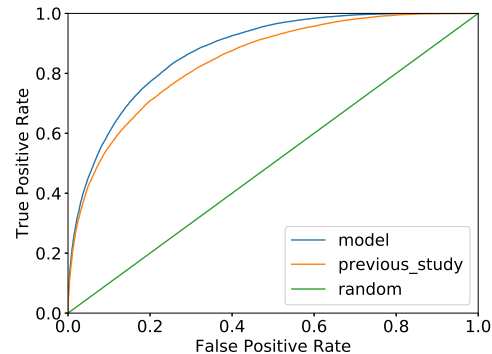


図 1: テンパイ予測の ROC 曲線

表 1: ROC 曲線の曲線下面積

モデル	曲線下面積
model	0.8761
previous_study	0.8430
random	0.5000

用いたモデルの構成を以下に示す。まず入力層の次の層で、捨てた牌のデータと鳴いた牌のデータをそれぞれ異なる 2 つの LSTM 層で処理する。その次の層で、LSTM 層の 2 つの出力ベクトルとその他の入力情報ベクトルを結合する。得られたベクトルを Multi-Layer Perceptron (MLP) の入力とし、MLP の出力をモデルの出力とする。

3.1. 実験結果と考察

図 1 に実験の結果得られた ROC 曲線を示す。表 1 に図 1 中の各曲線の曲線下面積を示す。各図表中の model は本研究で作成したモデル、previous_study は従来研究の論文中的モデルを本研究と同じデータを用いて再現したモデル、random は一様乱数によってテンパイと非テンパイを出力するランダムなモデルである。表 1 で示した曲線下面積は 1 に近いほどモデルが高精度であることを示す。本研究のモデルは従来研究のモデルおよびランダムなモデルと比較して高精度であることがわかる。

また、本研究で用いた特徴量は 514 次元であり、従来研究の特徴量の 6,888 次元と比較して約 7.5% 程度であるため、特徴量ベクトルの次元数を大幅に削減できたといえる。

4. まとめと今後の課題

本研究では、LSTM によるニューラルネットワークモデルを用いて、従来よりも高精度な他プレイヤーのテンパイ予測モデルを構築した。今後の課題としては、今回構築したモデルの予測結果を踏まえた上で最適な捨て牌を決定するモデルの構築が挙げられる。

参考文献

- [1] 水上直紀, 鶴岡慶雅他. 牌譜を用いた対戦相手のモデル化とモンテカルロ法によるコンピュータ麻雀プレイヤーの構築. ゲームプログラミングワークショップ 2014 論文集, 2014:48-55, 2014.