

# 機械学習を用いた麻雀の役予測手法の提案

尾崎 大夢      ライエル グリムベルゲン

**概要** : AlphaGo Zero の登場によってゲーム固有の特徴量を用いなくて強いゲーム AI を作ることが可能になった。一方で人間らしい動きをする AI や不完全情報ゲームにおける AI 研究の分野は強い AI こそ完成したものの発展途上であると考えられる。本研究では不完全情報ゲームの麻雀をテーマとして、捨て牌などの局情報をもとに相手の狙っている役の予測をニューラルネットワークで予測することを目的とする。牌譜から抽出したデータを入力データとして LSTM モデルによる予測を行わせ、役ごとに確率を出力させる。実験の結果に有意差は確認できず、今後の課題を残すこととなった。今後の課題として実験結果の原因を調査することになる。

**キーワード** : 麻雀, LSTM, 予測モデル

## A Machine Learning Method for Predicting Mahjong Tile Combinations

Daimu Ozaki      Reijer Grimbergen

**Abstract**: Since AlphaGo Zero, it has become possible to make strong game playing programs without the need for game specific features. On the other hand, creating human-like game play or strong game playing programs for imperfect information games are research topics that are on the rise. In this work, our goal is to build a neural network that can predict which Mahjong tile combinations other players are aiming for by using the information of discarded tiles. Using the data from Mahjong games of human players as input to the network, we used the LSTM model for prediction, leading to Mahjong tile combinations and their probabilities as the output of the network. However, in our current experiments we have not been able to confirm a significant difference in the probabilities. Finding the cause for these results will be our next aim.

**Keywords**: Mahjong, LSTM, Prediction Model

### 1. はじめに

2017 年に登場した AlphaGo Zero [1] はゲーム固有の特徴量を用いずに人間のトッププレイヤーを大きく上回る実力を獲得することに成功した。これによって、完全情報ゲームにおける強い AI の研究には一つの区切りがついたと言える。一方で不完全情報ゲームでは 2010 年代に入ってから盛んに研究されてきており、人狼ゲーム [2] や麻雀、ポーカーといったゲームは多くの研究報告がなされてきた。

現在では、ポーカーにおいて人間のトッププレイヤーを超える実力を有するゲーム AI [3] も開発されており、他のゲームにおいても人間のトッププレイヤーに比肩する実力を備えたゲーム AI の開発報告 [4] があげられている。

本研究では、不完全情報ゲームの一つである麻雀のうち、公開されている情報から相手の狙っている役を機械学習で予測することを目的とする。

ゲームの最中では、相手の狙っている役の状況の回答を得ることができず、終了後に評価することになる。したがって、本研究では麻雀においてゲームの経過を記録したログである牌譜を用いて

性能評価を行う。

### 2. 先行研究

水上ら [4] の研究では、教師付き学習による鳴きと降りの機能の導入手法が提案されている。上級者の牌譜を用いた 1 人麻雀プレイヤーを作成し、そこに降りるべき局面についての入力データの拡張や鳴きに関して特徴量を追加するという手法で 4 人麻雀プレイヤーを作成した。

降りと鳴きに関してはそれらを導入しなかった場合と比較して有意差があり、それらが正しく機能していたことを示していた。

同時に、1 人麻雀プレイヤーの段階での放銃率の改善、回し打ちの認識などを課題として挙げている。

萩原ら [5] の研究では、特徴量のグルーピングと組み合わせによって複雑化されたモデルを用いて相手の和了点数の予測を行った。

複雑化されたモデルは既存研究のモデルよりも予測精度が向上したとされ、特徴量や学習局面数の増加を課題として挙げている。

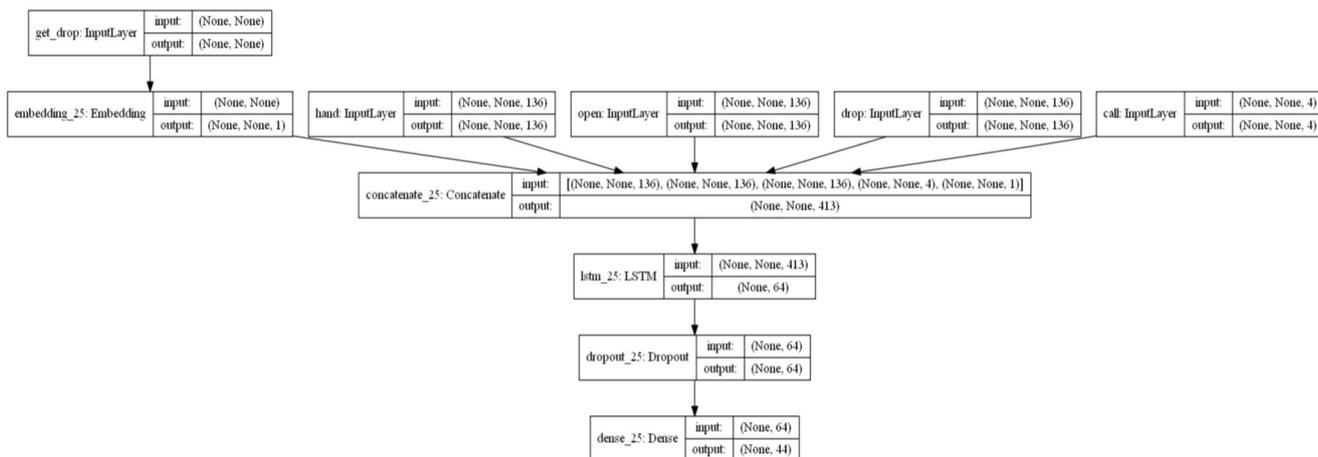


図 1: ネットワーク構成

### 3. 提案手法

本研究では、牌譜をもとに AI に役の予測を行わせ、その正答率を求めることを目的とする。

#### 3.1 AI の説明

本研究で用いる役予測 AI の構成は以下の通りとする。

1. 盤面の捨て牌をもとに予測を行うため、時系列データ処理として考えて LSTM (Long Short-Term Memory) [6] モデルによって予測を試みる。
2. 後述の牌譜を教師例として教師あり学習を行い、局面における牌推移をもとにして各役ごとに確率を出力する。
3. 明確に役が付いて上がった局の牌譜を用いて出力された確率の個別検証を行う。

明確に役が付いて上がった局とする基準は次のとおりとする。

1. その役だけで 3 翻以上になる役であること
2. ある程度の出現率を確保できること

この 2 つの理由を設定した理由として、ある程度大きな役であればその役を狙う際に捨て牌に何らかの特徴が出ると考えたためである。役満などの極めて大きな役にしなかった理由としては出現率の問題があり、あまりに出現率が低いと判別するのに必要なデータが集まらなると判断したためである。

今回の実験ではこの条件にあてはまるものとして清一色を採用することとした。オンライン麻雀対戦サイト「天鳳」[7]の鳳凰卓における 2020 年 10 月段階での出現率を表 1 に示す。3 翻以上になる可能性がある役でかつ役満でないものは表にある 4 つとなる。このうち二盃口と純全帯幺九は出現率が低くデータ不足になる可能性があるため除外した。混一色は副露時に 2 翻となること、単体翻からほとんど副露したうえで成立した役である

ことから今回の条件のうち翻数の条件に合わないと考えた。清一色は単体で満貫を確定させるほど高い役であるのに対し出現率も安定しており、狙っているか否か(3 種類の数字が書かれた牌のうち 1 種類のみを集める役のため、他の 2 種類及び字牌は不要な牌になる)の判別も行きやすいだろうと考えた。そのため、清一色の正答率から判別できているかの確認を行うこととした。

役	割合 (%)	単体翻	複合翻	%*翻	%*複
清一色	0.50783	5.17	6.25	2.62	3.17
二盃口	0.08464	3	5	0.25	0.42
純全帯幺九	0.38087	2.22	3.44	0.85	1.31
混一色	3.7664	2.08	4.06	7.83	15.28

表 1 天鳳鳳凰卓における条件を満たす役の出現率

盤面のあるプレイヤーから見える情報を入力データとして用い、終了後に成立した役を入力データにラベル付けする。

この時に入力データとして用いるものは次のとおりとする。

- 捨てた牌が手牌から切ったかつモ切りか
- あるプレイヤーの手牌
- その局面で公開されている牌とその数 (捨て牌, 鳴き牌, ドラなど)
- 役の予測を行う対象となるプレイヤーが捨てた牌
- 各プレイヤーが立直しているか

ネットワーク構成は図 1 のとおりとする。

今回のネットワークでは入力データとして 5 つ

のデータを用いるため5つの入力層を設け、このうち1つは予測対象となるプレイヤーにのみかわる情報のため3次元空間に埋め込む必要があったため埋め込み層を追加する必要があった。これらの入力データを結合して学習層に埋め込み、全体の20%をドロップアウトさせて分類器で役の判別を行い、最終出力をさせる。活性化関数はsoftmax, 損失関数は交差エントロピーを使用する。学習時の出力はその役が付いていた場合は1, そうでない場合は0になるようにし、ラベル付けされた実際に付いた役を正解として学習を進める。

ドロップアウト率を20%と設定した理由は、事前のネットワークの調整段階で相対的に最もいい数字が出たためにそのまま本実験用のネットワークに流用する形を取った。

これらの情報を各順目ごとに記録し、局ごとに独立した時系列データとして扱う。

今回の研究では局面の推移を考慮に入れて役の予測を行うようにするため時系列データ処理の一種として捉え、RNN (Recurrent Neural Network) によって処理する。

教師あり学習においては、訓練時の正答率の数値及び検証時の成立した役に対する予測率がなるべく大きくなるようにネットワークを調整する。

### 3.2 実験手法

実験に用いる牌譜は、オンライン対戦サイトの「天鳳」にて配布されているものを使用する。

この時に用いる牌譜データは以下のとおりとする。

1. 天鳳にて配布されている牌譜データのうち、2018年に鳳凰卓にて行われた対局の物を用いる。鳳凰卓は天鳳の全プレイヤーのうち0.1%ほどのプレイヤーしか入ることができない非常にレベルの高い卓である。
2. 2018年に鳳凰卓で配布されたデータのうち80%を学習用データ、10%を検証用データとし、残りの10%をテストデータとする。さらに、テストデータより先に挙げた役の付いた局の牌譜をいくつか抽出して個別に役の確率を出力させる。
3. データの総数は約21万半荘、局数にすると約120万局となる。
4. 検証用データから抽出する局については特定の役が付いたものを抽出する。同一の役が付いた局の牌譜がテストデータ内に複数ある役を選択し、これらを用いて特定の役を識別できているかについても検証する。
5. この時、入力データの数値に間隔を設けるために一定値の定数倍を行い、上がった役の大きさや局全体の推移によって重

要度に差を設けるようにする。

6. 流局になった局の牌譜は今回の実験意図と反することになるので用いない。

テストデータのうち、特定の役が付いたために抽出した牌譜の扱いとしては現時点では正答率の検証には用いないようにする。通常のテストデータと重複して用いることによるネットワークへの悪影響を避けるためである。

## 4. 実験結果

本実験の結果を以下に示す。

表2は節3.1で説明したネットワークを用いて異なる盤面を清一色の有無で区別し各10局ずつ取り、その平均を示した。

	清一色なし	清一色あり
1	0.1609073	0.17669743
2	0.18414617	0.15701741
3	0.15706629	0.17773646
4	0.17662048	0.15701127
5	0.15597248	0.15693024
6	0.1861684	0.1842567
7	0.15701741	0.16302589
8	0.1611459	0.15701017
9	0.16077742	0.17749631
10	0.15609434	0.18539712
平均	0.165591619	0.1692579

表2 清一色の予測率結果

ウェルチのt検定を行った結果、t値は0.5前後の数値を取り95%の信用水準に達しておらず、誤差の範囲内となっていることが確認された。

この原因についてだが、学習に用いるデータの数を変更しつつ数値の確認を繰り返しても結果に大きな変化が見られなかったため、ネットワークの構造上の問題と思われる。結果についてはつい実験等を行って確認していくことを構想している。

## 5. おわりに

今回はLSTMを用いて牌情報から相手の役の予測精度の分析を行った。

現状、実験結果が思うようなものになっておらず、その原因についての調査をしている段階である。原因及び追実験に関しては今後の追記として載せていくことを目標としたいと考えている。

## 参考文献

[1] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian

Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel & Demis Hassabis, "Mastering the game of Go without human knowledge", *Nature*, pp.354-359 (2017)

[2]大川貴聖, 吉仲亮, 篠原歩「深層学習を用いて役職推定を行う人狼知能エージェントの開発」ゲームプログラミングワークショップ2017論文集, pp. 50-55(2017)

[3] Noam Brown, Tuomas Sandholm "Superhuman AI for multiplayer poker", *Science* **365**, pp.885-890(2019)

[4] 水上直紀, 中張遼太郎, 浦晃, 三輪誠, 鶴岡慶雅, 近山隆「多人数性を分割した教師付き学習による4人麻雀プログラムの実現」情報処理学会論文誌 55, pp. 2410-2420(2014)

[5]萩原涼太, 山田渉央, 佐藤直之, 池田心「麻雀における相手の和了点数予測法の性能評価」情報処理学会研究報告 11, pp. -18(2016)

[6]<https://jp.mathworks.com/discovery/lstm.html> (2020年7月21日閲覧)

[7] <https://tenhou.net/> (2020年7月21日閲覧)