

機械学習を用いた麻雀戦術モデルの構築

1 はじめに

近年、機械学習や深層学習の手法に基づき、ゲームをプレイする人工知能 (Artificial Intelligence, AI) が数多く開発されている。本研究では、代表的な不完全情報ゲームである麻雀を対象とする。

麻雀では、一般に 4 人でするゲームであり、他のプレイヤーが捨てた牌の種類や順番などから、プレイヤーが現在手持ちにし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予測精度は高いと考えられる。

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考えると、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。青野らの研究 [2] 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を向上させたが、精度が十分とではない。

本研究では、ViT [3] を参考にして、Transformer のモデルを利用して新しいモデルを提案し、捨てた牌間の関係性を捉えることで予測精度の向上を試みる。

2 麻雀とは

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ゲームである。1 ゲームは局と呼ばれる単位によって区切られている。局のはじめに、各プレイヤーはそれぞれ 13 枚の牌を手持ちにしておき、1 枚牌を加えると、手持ちから牌 1 枚捨てる行為を繰り返すことで牌を替えていく。牌の組み合わせで三枚の組 4 セットと二枚の組 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。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ような牌を手持ちにしているか直接知ることはできない。この点から麻雀は不完全情報ゲームである。

- 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である。
- テンパイ：アガリの一つ前の状態。他プレイヤーの捨てた牌や鳴いた牌の情報などから手持ち牌がテンパイ状態であるかの予測をテンパイ予測と呼ぶ。
- リーチ：手持ち牌がテンパイしている場合、リーチをかけることで得点は増える。しかしリーチをかけるとアガリ牌以外の牌は全部捨てなければならない。
- アタリ牌：テンパイしている人に対して、アガリになる牌をアタリ牌と呼ぶ
- 放銃：放銃とは自分の打牌に対して他家がロンしてしまうことである。アガった人の点数は放銃した人一人で負担するため、損失が非常に高い。
- 有効牌：有効牌というのはシャンテンを進める牌と受け入れが広がるような牌。
- 牌効率：牌効率は聴牌に至る確率や早さのこと。有効牌の多い打牌を選択したら牌効率も高くなる。

3 要素技術

3.1 Transformer

Transformer は、2017 年に発表された深層学習モデルであり、主に自然言語処理 (NLP) の分野で使用される。Transformer は時系列データを逐次処理する必要がないという特徴がある。図 1¹に Transformer のエンコーダー部分を示す。Multi Head Attention layer はモデルの核心部分で各部分の関係性を捉えることができる。また Multi Head Attention layer で使ったマスクは精度の向上に繋がっている。

3.2 ViT

図 2²に ViT のモデルを示す。ViT [3] は画像処理の分野で CNN を利用せず Transformer のみでモデルを構築したものである。

¹文献 [3] の図 1 (右) 参照

²文献 [3] の図 1 (左) 参照

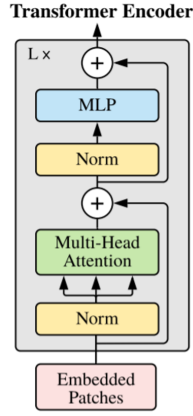


図 1: Transformer のエンコーダー部分 [3]

ViT のモデルの構造は一般的な Transformer [4] のエンコーダー部分と同じである。Transformer の Input をシークエンスデータとして受け取る必要があるため、二次元である画像データをパッチごとに一次元のシークエンスデータに変換したうえで、線形射影する。そのうえにシークエンスデータの先頭に「CLS」トークンを追加する。パッチが画像のどこにあるかを識別するために、各パッチに位置埋め込みも付与する。そして Transformer のエンコーダー部分に入力し、「CLS」トークンの出力を識別器に入れることで最終的な予測を出す。

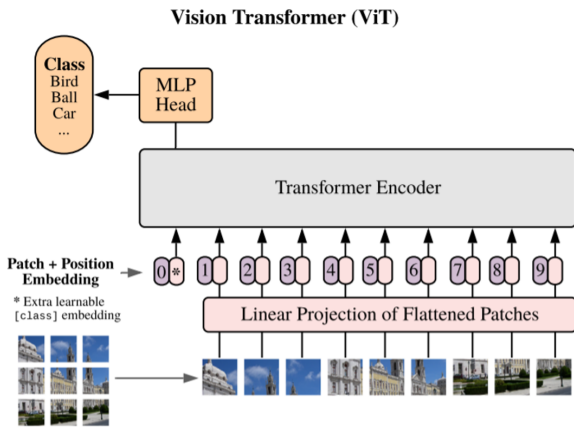


図 2: ViT のモデル [3]

4 提案手法

4.1 入力データ

本研究では天鳳 [5] の 2013 年から 2015 年までの牌譜をデータセットとして用いる。牌譜から局の流れを

表 1: 入力データ

情報	次元数	説明
自風	1	自風であるか
場風	1	場風であるか
色	4	萬 筒 索 字
数	9	
ドラ	1	ドラであるか
ドラ隣	1	ドラの隣であるか
ドラ同色	1	ドラと同じ色であるか
赤	1	赤ドラであるか
行動の種類	3	手出し 自摸切り 鳴き
鳴きの種類	5	チー ポン カン
鳴き牌の色	4	萬 筒 索 字
ドラ	1	鳴き牌はドラであるか
19	1	19 牌であるか
自風	1	自風であるか
場風	1	場風であるか
三元	1	三元であるか
鳴きの番号	9	具体的な鳴き牌
リーチ	1	リーチしているかどうか
残り牌	25	目標の周りの牌の残り枚数

再現して、一回の打牌後の局面を入力データとする。Transformer のモデルに入力するために、まずはデータを時系列化しなければならない。本研究では一回の打牌を一つの行動としてパッチ化する。表 1 に入力データのパッチ構造を示す。テンパイモデルの場合、前半 21 次元は今回捨てた牌の情報で、後半 15 次元は今回の鳴き牌の情報である。もし今回の行動が鳴いてからの打牌ではない場合、後半 15 次元の要素を 0 とする。入力データの中身は自由に変えることができる。例えばアタリ牌予測モデルは 9 次元の鳴き牌の番号と 1 次元のリーチ状態も入れている。牌山予測モデルでは 25 次元の目標牌の周りの牌の残り枚数を入れている。

一回の捨て牌の行為は一つのパッチに対応し、一人は最大 24 枚の捨て牌があるため、足りない部分に「PAD」トークンを追加し、24 パッチの入力データを作る。

4.2 モデル

本研究のモデルは ViT を参考にして、Transformer のエンコーダー部分を利用して、麻雀 Transformer モデルを構築した。図 3 に麻雀 Transformer のモデルを示す。本研究のモデルはすべてこのモデルを基づいて

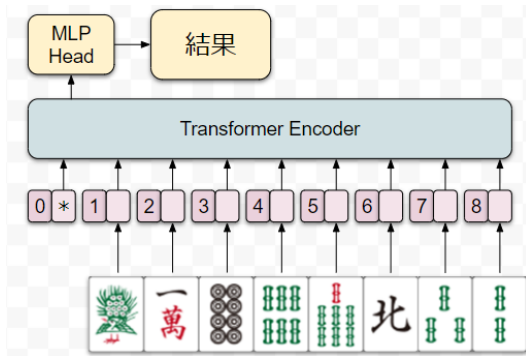


図 3: 麻雀 Transformer のモデル

構築した．入力データに位置エンコーディングを加算し，先頭に「CLS」トークンを追加し，エンコーダーに入れる．エンコーダーの出力を線形層に入れて結果を出す．エンコーダーの中にある Multi head attention layer にかけるマスクによって，「PAD」トークンは結果に影響を与えない．

5 実験

本研究では，Transformer モデルに基づいて，牌山予測モデルと当たり牌予測モデルを構築した．天鳳の牌譜を用いて，二つの実験をして，麻雀戦術モデルに対する Transformer の性能を検証した．

5.1 実験 1 牌山予測

5.1.1 概要

麻雀の牌は 1 種類 4 枚があって，3 つの状態が存在する．まのずは捨てられたのか鳴きで取られたという状態である．この状態の牌はすでに使われておりプレイヤーに公開されているが手牌では使用できない．残り 2 種類はプレイヤーが手に入れる可能性がある未公開の牌に関するものである．1 つは他の人の手牌に存在する状態である．この状態はプレイヤーは確認も使用もできないため，プレイヤーにとって好ましくない．もう 1 つは牌山に残っている状態である．この状態の牌は実際に手に入れる可能性がある牌で，プレイヤーにとって有効な状態である．プレイヤーは自分が欲しい牌は多めに牌山に残ってほしいと考えられる．

だが，プレイヤーは見えない牌の場所を知らないため，正しい判断はできない．このモデルは，他のプレイヤーの捨て牌から持っている手牌を予測し，牌山に残っている牌を計算する．予測結果によってより正しい判断

ランキング	予測	点数
1	2	1/2
2	1	1
3	4	3/4
4	3	1
5	5	1

$$\text{TOP3点数} = (1/2 + 1 + 3/4) / 3 = 0.75$$

図 4: 評価手法の例

が期待される．1 種類の牌に対して 1 つのモデルを作り，合わせて 34 個のモデルを学習した．

5.1.2 テスト方法

一つの局面を取って，違う牌を捨てた場合の有効牌の種類を計算し，この時点の場面の状況から牌山予測モデルを使って，有効牌が牌山にある枚数を予測する．違う牌を捨てた場合は，有効牌の枚数も違うため，枚数によって捨て牌にランキング付け，実際のランキングと比べて，評価手法によって点数を付ける．この点数がモデルの精度を表す．

5.1.3 評価手法

推薦システムでよく使われている mAP (mean Average Percision) を参考にして，麻雀の状況に応じた評価手法を考えた．麻雀はよく何枚かの牌を同時に着目するため，この手法は全体の精度を重視している．予測ランキングは実際のランキングより下の場合は 1 以下の点数をつける，予測ランキングが高い場合は 1 を付ける．図 4 に評価手法の例を示す．TOP3 点数を計算する場合は TOP3 の点数の平均値を計算する．

5.1.4 実験 1 結果

表 2 に実験 1 の結果を示す．Baseline は概要で紹介した，見えない牌の数で計算したランキングの点数である．この評価手法で計算した点数は高めになるが，Baseline より点数が大幅に向上した．より正しい判断ができると言える．

表 2: 牌山予測の結果

	ACC	mAP TOP1	mAP TOP3
Baseline	0.6604	0.8227	0.8319
提案手法	0.6575	0.8527	0.8583

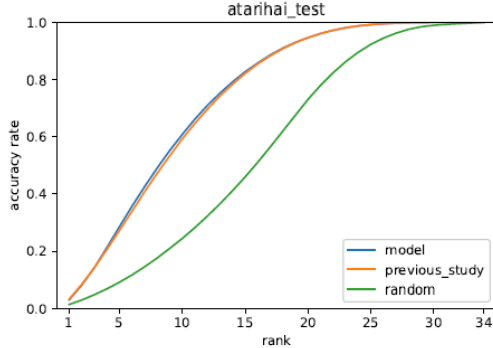


図 5: 曲線の例

5.2 実験 2 アタリ牌予測モデル

5.2.1 概要

実験 2 は捨て牌などの情報から目標のアタリ牌を予測する。学習はテンパイしているデータのみ使った。

5.2.2 テスト方法

一つの場面に対して、予測モデルにより 34 種類の牌がアタリ牌である確率を予測する。確率が高い牌から n 番目までの牌を選択し、すべてのアタリ牌が含まれている場合は正解とする。すべてのテストデータに対して上記の処理をして、 n を 1 から 34 まで変化させた際の正解率の推移を曲線で表す。この曲線の下面積 (Area Under the Curve: AUC) をモデルのスコアとする。面積が大きいほどモデルの精度が高いと言える。

図 5 に曲線の例を示す。横軸はアタリ牌を当てる際を選ぶ牌の枚数。縦軸はアタリ牌を当てる確率である。例えば、横軸が 34 の場合は 34 種類の牌を選んだら、アタリ牌が含まれている確率は 100 パーセント、横軸が 1 の場合は、1 枚だけ選んでアタリ牌を当てる確率はかなり低いことを意味する。

5.2.3 実験 2 結果

表 3 に実験 2 の結果を示す。従来手法より提案手法の精度が向上した。また、平均最低順位は 10 以下に

表 3: アタリ牌予測の結果

	AUC	平均最低順位
従来手法 LSTM	24.74	
従来手法線形回帰	24.66	
提案手法	24.79	9.74

なった。つまり安全牌の枚数は平均 24 枚があり、実際の場面でも使用できると考えられる。

6 まとめと今後の課題

本研究は ViT を参考にして、Transformer のモデルに基づいた新しい牌山予測モデルを構築し、高い精度を達成した。また、アタリ牌予測モデルを構築し、従来研究の回帰モデルと LSTM 時系列モデルより精度が向上することを確かめた。

今後の課題として、アガリ点数予測モデルを構築し、前期発表会で発表テンパイ予測モデルと牌山予測モデル、アタリ牌モデルを合わせた初心者サポートシステムを作りたいと考えている。

参考文献

- [1] 水上直紀, 鶴岡慶雅. 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の構築. 2014.
- [2] 青野義樹. 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. 大阪府立大学卒業論文, 2018.
- [3] Alexey Dosovitskiy and Others. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021.
- [4] Ashish Vaswani and Others. Attention is all you need, 2017.
- [5] 天鳳. <https://tenhou.net>.