

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) プログラムが数多く開発されている。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する。

麻雀では、ある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言えない。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、精度が十分とは言えない。

本研究では、ViT [3] のを参考にして、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り、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ず、この点から麻雀に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。
- シャンテン数：テンパイまで必要な有効牌の数はシャンテン数と呼ばれる。例えば、テンパイまで後 1 枚有効牌が必要な時は 1 シャンテンと呼ばれる。
- 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すべて負担するため、損失が非常に高い。

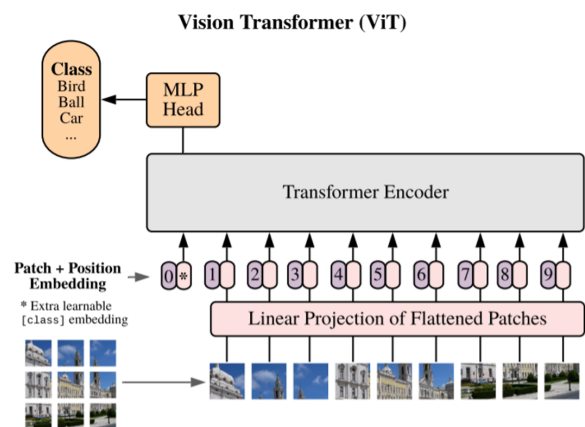


図 1: ViT のモデル [3]

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 2 に 2020 に発表した

時系列データに対する Tranformer の適用方法 [4] を示す。このモデルは回帰問題や分類問題などを処理することができる。

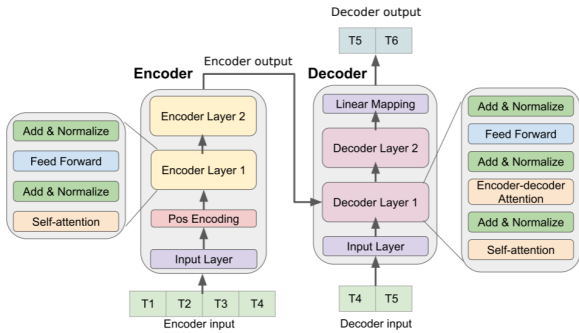


図 2: Transformer の時系列モデル [4]

3.2 ViT

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。

ViT のモデルの構造は一般的な Transformer [5] 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 [class] 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、[class] 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。Transformer のエンコーダー部分の Multi Head Attention layer は各部分の関係性を捉えることができる。

3.3 PR 曲線

PR 曲線 (Precision-Recall 曲線) は Precision と Recall からなる曲線である。Recall を横軸に、Precision を縦軸にして、上位の値から順に折れ線で描かれる。PR 曲線は歪んだ分布を持つデータセット (陽性例より陰性例の方が圧倒的に多いデータセット) に対して、ROC 曲線に比べてより高いパフォーマンスを発揮する。

4 提案手法

4.1 入力データ

本研究では天鳳 [6] の 2013 年から 2015 年の牌譜をデータセットとして用いている。牌譜から局の流れを再現して一回の打牌後の局面を入力データとして用いている。Transformer のモデルに入力するために、まずはデータをシークエンス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 に入力データの一つパッチの構造を示す。前 21 次元は今回の捨てた牌の情報で後 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だった場合、鳴き牌の情報を後 15 次元に入れる、そうではない場合は 0 で埋め込む。

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	m p s 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	m p s 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

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のを利用し、一人と四人テンパイ予測モデルと、テンパイタイミング予測モデルを構築した。

4.2.1 一人モデル

一人モデルでは、二つモデルを構築した。一つ目は ViT のように「CLS」を使用し、言語モデルのように attention layer でマスクかけて、「CLS トークンの出力を分類器に入れて予測する。二つ目は「CLS」トークンを使用せず、すべてのパッチの出力を分類器に入れて予測する。図 3 に二つのモデルを示す

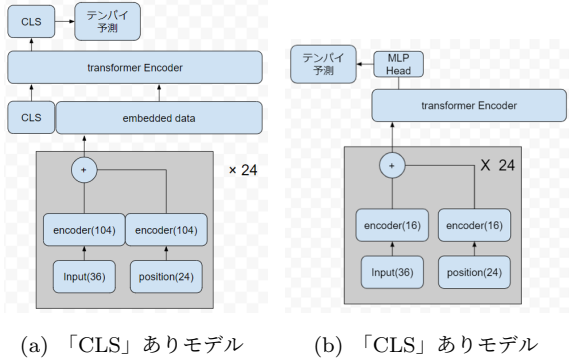


図 3: 一人モデル

4.2.2 四人モデル

図 4 に四人モデルを示す。四人モデルは四人分操作を入力し、四人に対応している四つの「CLS」トークンを使用し、「CLS」トークンの出力をそれぞれ違う分類器に入れて四人のテンパイ状態を予測する。attention layer では特別のマスクをかけている。

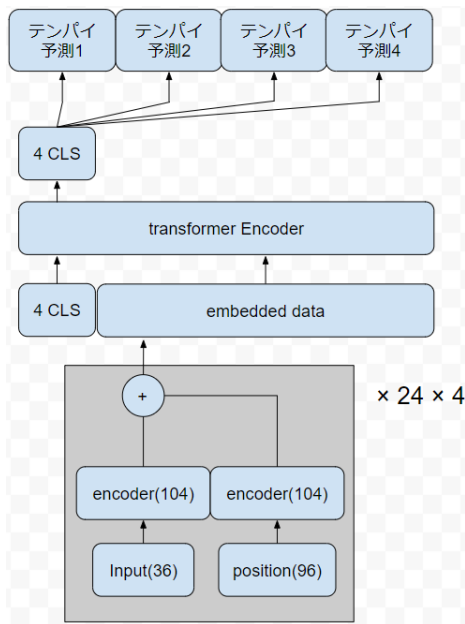


図 4: 四人モデル

4.2.3 テンパイタイミングモデル

図 5 に四人モデルを示す。Tranformer の時系列データに対する適応方法を参考にして、改造したモデルである。現在の操作をする前の状態を考慮しながら、現在の操作を重視しているモデルである。エンコーダー部分の「CLS」トークンの出力は現在の操作をする前の状態を予測し、デコーダーの出力は操作をした後の状態を予測する。

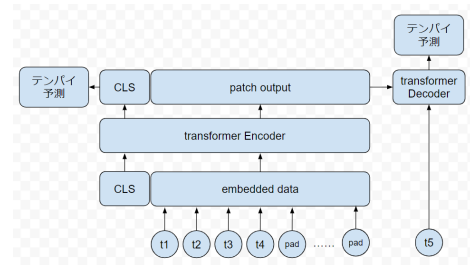


図 5: テンパイタイミングモデル

5 実験

本研究では、Transformer モデルに基づいて、テンパイとシャンテン数傾向予測モデルを構築し、天鳳の牌譜を用いて、三つの実験を行って、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 テンパイ予測

日本麻雀はリーチ麻雀であり、放銃の損失が高いため、日本麻雀の中に攻防転換が一番重要な戦術である。他のプレイヤーの捨て牌や鳴き牌などの情報を見てテンパイしているかどうかを判断して、自分の攻防戦術を選ぶ。そのためテンパイ予測は重要なモデルであると考えられる。本研究では、二つの一人モデルと一つ四人モデルを構築し、従来手法の LSTM と比較した。

表 2 に実験 1 の結果を示す。構築したモデルの中で、最も性能高いのは「CLS」を使用していない一人モデルである、他のモデルの性能が低い原因は、パッチのサイズは小さく、全ての情報は一つの「CLS」トークンに入れないことと、四人モデルでは意味のある情報が少ないことだと考えられる。いずれにしても、従来手法の LSTM より精度が高いことが分かる。これ以降は、より大きいパッチサイズと、より多い学習データを用いて実験する。

表 2: テンパイ予測の結果

	precision	recall	F1 値	AUCPR
4 人-東	0.6124	0.5807	0.5962	0.6436
4 人-南	0.6514	0.5258	0.5819	0.6270
4 人-西	0.6699	0.4867	0.5638	0.6385
4 人-北	0.7012	0.4596	0.5553	0.6336
1 人-CLS	0.6284	0.5731	0.5995	0.6356
1 人-noCLS	0.6424	0.5733	0.6070	0.6520
LSTM	0.6953	0.3935	0.5026	0.5996

5.2 実験 2 シャンテン数傾向予測モデル

実験 1 でのテンパイ予測の正答率ある程度あるけど、テンパイするタイミングの予測はずれている。Transformer はテンパイするタイミングを正確に予測する性能があるかどうかを検証するために、シャンテン数傾向予測モデルを構築した。テンパイは 1 シャンテンから 0 シャンテンになることなので、シャンテン数下がるタイミングを予測できるなら、テンパイするタイミングもある程度予測できると考えられる。

シャンテン数傾向予測モデルは「CLS」トークンを使用する一人モデルをそのまま使用し、ラベルをシャンテン数が下がったかどうかに変えた。表 3 に実験 2 の結果を示す。結果を見ると、一人モデルはある程度シャンテン数下がるタイミングを予測できているので、テンパイ予測の時にタイミングを予測できる性能はもっていると考えられる。

表 3: シャンテン数傾向予測の結果

シャンテン数	precision	recall	F1 値
下がる (手牌進む)	0.8078	0.9842	0.8873
下がってない	0.9546	0.5872	0.7271
acc	0.8405(bl = 0.6381)		

5.3 実験 3 テンパイタイミングモデル

実験 3 はテンパイタイミングモデルを用いて聴牌予測の実験をした。モデルの形の制限で、最低でも捨て牌は 2 枚が必要だが、実践的な麻雀では一巡目でテンパイが起こるのは非常に小さい確率なので、影響がないとしている。表 4 に実験 3 の結果を示す。結果を見ると精度は高いとは言えず、二つの予測が互いに影響を与えている可能性がある。.. 今後は同時に学習するのではなく、順番に学習してこのモデルを改善する。

表 4: テンパイタイミングモデルの結果

一つ前の状態	precision	recall	F1 値
ノーテン	0.9557	0.9807	0.9600
テンパイ	0.6793	0.4731	0.5784
AUCPR			0.6206
現在の状態	precision	recall	F1 値
ノーテン	0.9391	0.9754	0.9569
テンパイ	0.7053	0.4819	0.5726
AUCPR			0.6070

6 まとめと今後の課題

本研究は ViT を参考して、Transformer のモデルに基づいて、麻雀のテンパイ予測モデルとシャンテン数傾向予測モデルを構築し、テンパイするタイミングについても検討した。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、シャンテン数傾向予測モデルを利用し、テンパイタイミングの予測について更なる精度の向上を目指す。そして、麻雀に合致する Transformer モデルを用いて、他の幅広い戦術モデルを試みる。

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ヤの構築. 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] Neo Wu, Bradley Green, Xue Ben, and Shawn O'Banion. Deep transformer models for time series forecasting: The influenza prevalence case. 2020.
- [5] Ashish Vaswani and Others. Attention is all you need, 2017.
- [6] 天鳳. <https://tenhou.net>.