

修士論文の和文要旨

研究科・専攻	大学院 情報理工学研究科 情報・ネットワーク工学専攻 博士前期課程		
氏名	三留 大明	学籍番号	1731152
論文題目	副露した他家の聴牌を予測する研究		
要旨	<p>本論文では、多人数不完全情報ゲームの一つである麻雀を題材とし、機械学習を用いて人工知能 (AI) プレイヤの実力を向上させることを目標に行った研究内容について述べる。</p> <p>多人数不完全情報ゲームにおいて、他プレイヤーのみが知る情報を推定することは、強い AI プレイヤを作成することに繋がる。そこで、本研究では副露 (フォーロ) した他家が聴牌 (テンパイ) しているか否かという非公開情報の推定に着目した。麻雀は、最終的な持ち点を競うゲームであるため、和了 (ホーラ) 即ちあがりにより得点を得るだけでなく、放銃などによる失点を防ぐこともすべきであると一般に考えられている。そのような麻雀の性質から、失点を避けるために、副露した他家の聴牌を予測することは、強い AI プレイヤを作成するうえで重要であると考えられる。</p> <p>聴牌確率の推定精度が向上する要因を調査するために行った実験では、次のような結果を得た。</p> <p>1 つ目の実験では、ゲーム状況を表す特徴の数を 22 から 3,000 程度に増やすことにより、推定精度が向上することが明らかになった。</p> <p>2 つ目の実験では、打牌数、副露数で場合分けして推定することにより、推定精度が向上することが明らかになった。</p> <p>3 つ目の実験では、ロジスティック回帰ではなく、3 層ニューラルネットワークを用いて推定することにより、推定精度が向上することが明らかになった。</p> <p>本研究ではまた、聴牌確率推定の精度向上が麻雀 AI プレイヤの強さに寄与するかを検証するための対戦実験も行った。上記の実験結果に基づき、打牌数、副露数で場合分けし、多くの特徴からロジスティック回帰を用いて作成したモデルを既存の麻雀 AI プレイヤ に組み込み、対戦実験を行った。約 27,000 回の対戦の結果、平均順位は 2.503 で既存の AI プレイヤより低い平均順位であったが、有意な差ではなかった。</p>		

電気通信大学 情報理工学研究科
情報・ネットワーク工学専攻 修士学位論文

副露した他家の聴牌を予測する研究

2019年3月4日

学籍番号 1731152

三留大明
指導教員 保木邦仁

目次

1	はじめに	4
2	基礎知識	6
2.1	麻雀のルールと用語	6
2.1.1	ゲームの流れ	6
2.1.2	牌の種類と面子	7
2.1.3	和了形と聴牌	8
2.1.4	麻雀用語の説明	8
2.1.5	天鳳で認められている役の一覧	10
2.2	二項ロジスティック回帰と全結合ニューラルネットワーク	11
2.2.1	二項ロジスティック回帰	11
2.2.2	全結合ニューラルネットワーク	11
2.2.3	交差エントロピー誤差	12
2.2.4	勾配降下法と確率的勾配降下法	13
2.3	ROC 曲線と AUC 値	14
2.3.1	ROC 曲線	14
2.3.2	AUC 値	18
3	先行研究	19
3.1	栗田らの放銃・放銃点の確率推定	19
3.2	水上らの他家の待ち牌・放銃点の確率推定	20
3.3	水上らの降りるべき局面の認識	20
4	目的	22
5	実験方法	23
5.1	訓練データとテストデータに用いた牌譜	23
5.2	訓練データとテストデータのラベル付け	23
5.3	聴牌確率推定に用いた特徴	24
5.4	聴牌確率推定モデルの構成	25

6	実験結果	26
6.1	ゲーム状況を表す特徴が聴牌確率推定精度に与える影響の検証	27
6.2	打牌数と副露数での場合分けが聴牌確率推定精度に与える影響の検証	29
6.3	3層ニューラルネットワークが聴牌確率推定精度に与える影響の検証	31
7	対戦実験	32
8	おわりに	33
9	謝辞	33

1 はじめに

ゲーム人工知能における研究目標の一つとして、強いプレイヤーの作成が挙げられる。

二人完全情報ゲームにおいては、既にトップレベルの人間プレイヤーと同等かそれ以上の実力を持つコンピュータプレイヤーが作られている。チェスでは IBM 社が開発した DeepBlue [1] が 1997 年に世界王者カスパロフに勝利し、将棋では山本らが開発した Ponanza が 2017 年に佐藤天彦名人に勝利した。囲碁では Google DeepMind 社が開発した AlphaGo [2] が 2017 年に世界トップ棋士カ・ケツ九段に勝利した。

一方、多人数不完全情報ゲームに関する研究は、二人完全情報ゲームほどには進んでいない。これは、複数のプレイヤーがいるためにゲーム状況が複雑であること、ゲーム状況に関する全ての情報がプレイヤー間で共有されていないことなど、二人完全情報ゲームにはない難しさがあるためである。

本研究では、多人数不完全情報ゲームにおいて、他プレイヤーのみが知る情報を推定し、強いプレイヤーを作成することを目指す。題材は、多人数不完全情報ゲームの一つである麻雀とする。特に、本研究では、麻雀において、副露した他家が聴牌しているか否かという非公開情報を先行研究より正確に推定するモデルの作成を目標とする。本研究の題材である麻雀は、最終的な持ち点を競うゲームであるため、和了することにより得点を得るだけでなく、放銃（聴牌している他家に持ち点を奪われる行為）などによる失点を防ぐこともすべきであると一般に考えられている。また、水上らは愚直に和了を目指すプレイヤーと状況により和了を放棄するプレイヤーとを比較して、後者は平均順位が 0.23 高くなった [3] という結果から、和了を目指すだけでなく、放銃を避けるよう行動を選択することの重要性を示している。このような麻雀の性質から、放銃による失点を避けるために、副露した他家が聴牌しているか否かを知ることは重要であると考えられる。ここで、本研究にて副露した他家の聴牌に着目した理由は、立直している他家が聴牌していることは明らかであり、副露した他家は立直も副露もしていない他家と比較して聴牌確率が高いためである。

他家の聴牌確率推定に関する研究は既に行われており、水上らの多くの特徴からロジスティック回帰を用いた他家の聴牌確率推定 [4] や栗田らの打牌数、副露数などで場合分けし、少ない特徴からロジスティック回帰を用いた他家の聴牌確率推定 [5] などがある。これら 2 つの手法は筆者の知る限り現在最も強い麻雀 AI プレイヤーに用いられていることから、聴牌確率推定において、これらの手法は一定の成功を収めていると言える。水上らの聴牌確率推定が成功している理由は、多くの特徴から聴牌確率推定を行っているためであると考えた。また、栗田らの聴牌確率推定が成功している理由は、打牌数、副露数で場合分けしているためであると考えた。

そこで、本研究では、水上らと同様に多くの特徴を用い、栗田らと同様に打牌数、副露数で場合分けを行うことで更なる精度向上を目指す。また、先行研究ではロジスティック回帰を用いて聴牌確率推定を行っていたが、本研究では、3 層ニューラルネットワークを用いた

推定精度も検証する。提案手法を組み込んだAIプレイヤーの強さの検証は、対戦実験にて行う。

本論文の構成について述べる。第2章では、麻雀のルールと用語を説明する。また、本研究で用いた分析手法と評価指標について説明する。第3章では、本研究の聴牌確率推定に関連した先行研究を紹介する。第4章では、本研究の目的について述べる。第5章では、本研究における聴牌確率推定モデルの作成方法について説明する。第6章では、4つの聴牌確率推定手法の推定精度を比較した結果を述べる。第7章では、提案手法を組み込んだAIプレイヤーの対戦実験結果について述べる。最後に第8章で本研究を総括し、今後の課題を述べる。

2 基礎知識

本章では、本研究に関する基礎知識を説明する。

2.1 節では、麻雀のルールと用語について説明する [6]。麻雀には、様々なローカルルールがあるが、本研究ではインターネット雀荘天鳳 [7] のルールを採用した。天鳳では、教師あり学習を行ううえで必要な上級者の牌譜が公開されている。

2.2 節では、本研究で扱う二項ロジスティック回帰と全結合ニューラルネットワークについて説明する。二項ロジスティック回帰は一般化線形モデルの 1 種であり、2 値の目的変数を扱う場合に用いられる。全結合ニューラルネットワークとは、人間の脳神経系のニューロンを数理モデル化したものであり、隣接層のユニット同士がすべて結合しているニューラルネットワークである。

2.3 節では、Receiver Operating Characteristic (ROC) 曲線と Area Under Curve (AUC) 値について説明する。ROC 曲線は、偽陽性率と真陽性率の関係をグラフにしたものであり、分類器の性能を評価するために用いられる。AUC 値は、分類器の性能を表す評価尺度の一つである。

2.1 麻雀のルールと用語

麻雀とは、4 人のプレイヤーが各々牌を組み合わせて和了形 (特定の手牌の組合せ) を作り、和了 (ホーラ) することにより、点数を獲得し、その点数を競うゲームである。

2.1.1 ゲームの流れ

対局とそのゲーム状況を図 1 に示す。

局開始時に、自家と他家 (下家, 対面, 上家) に手牌としてそれぞれ 13 枚の牌が配られる。各プレイヤーは終局まで、反時計回りに自摸 (ツモ) と呼ばれる壁牌 (ピーパイ) から牌を 1 枚引く行為か副露 (フーロ) と呼ばれる他家の捨て牌を取得する行為を行った後、打牌 (ダハイ) と呼ばれる牌を 1 枚捨てる行為を行う。

通常は、いずれかのプレイヤーが和了するか壁山がなくなるまでが 1 局である。約 8 局対局し、最終的な点数を競う。

基本的に各プレイヤーが得ることのできる情報は、自分の手牌、各プレイヤーの捨て牌、ドラ表示牌、副露牌などである。他家の手牌については、副露牌のみ知ることができる。

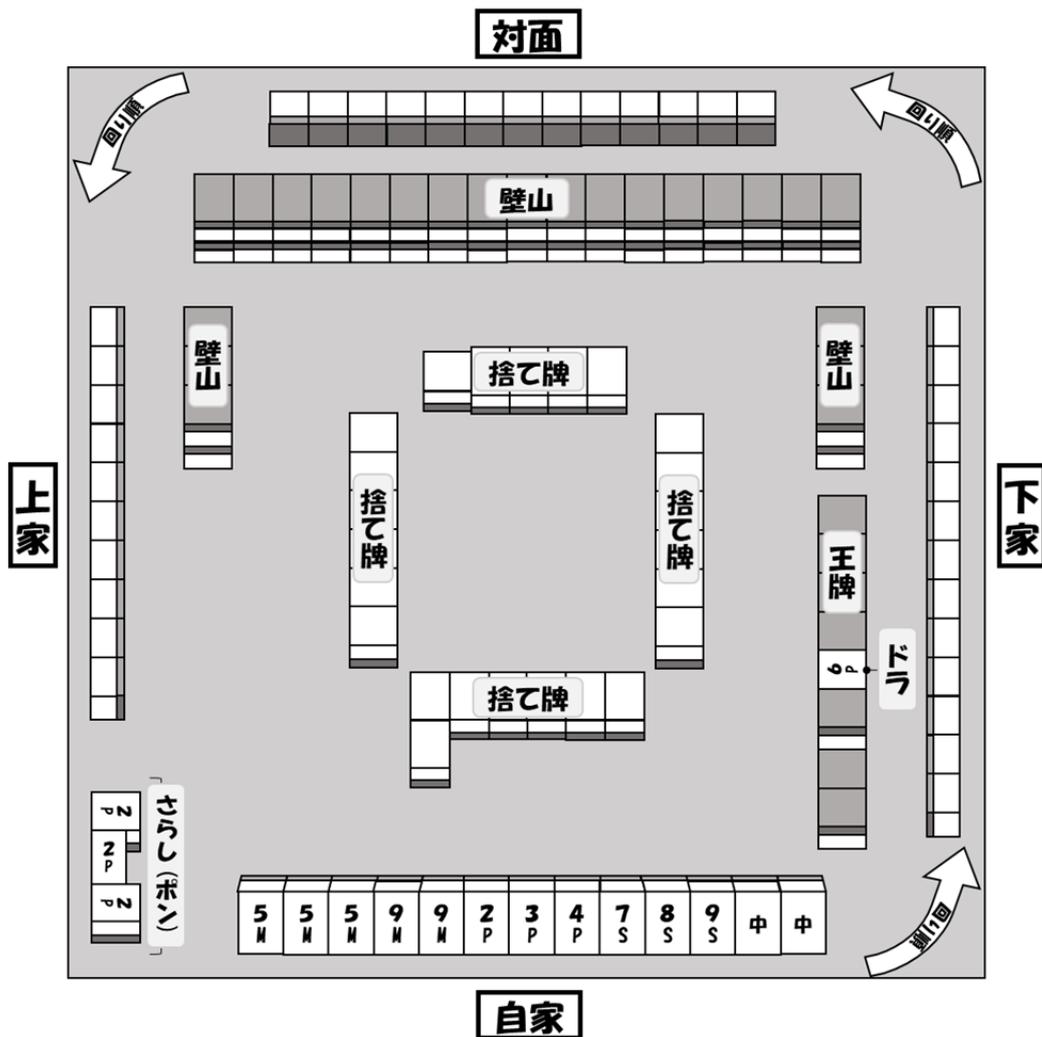


図1 対局とそのゲーム状況

2.1.2 牌の種類と面子

麻雀で使用される牌は34種類各4枚の合計136枚である。

牌の種類は、数牌（シュウパイ）と字牌（ツーパイ）の2つに大別される。数牌には、萬子（マンズ）、筒子（ピンズ）、索子（ソーズ）の3色がある。各色は1から9まであり、数牌は計27種である。字牌は、東、南、西、北の4種からなる風牌（フォンパイ）と白、発、中の3種からなる三元牌の計7種である。また、数牌の1と9と字牌は么九牌（ヤオチュウハイ）と呼ばれる。

牌の組合せのことを面子（メンツ）といい、面子には完成面子と未完成面子の2種類がある。完成面子には、同じ牌種を3つ集める刻子（コーツ）、同じ牌種を4つ集める槓子（カンツ）、同色の数牌で連続した牌を3つ集める順子（シュンツ）がある。未完成面子には、同じ牌種を2つ集める対子（トイツ）などがある。

2.1.3 和了形と聴牌

和了とは、手牌を公開して、点数を獲得する行為である。和了するための条件は、手牌を和了形に揃えることであり、和了形には3種類ある。1つ目は、特定の牌の組合せか特定の状況により役を成立させ、4組の完成面子と1組の対子に手牌を揃えることである。2つ目は、牌種の異なる7つの対子に手牌を揃えることである。3つ目は、么九牌13種14枚に手牌を揃えることである。

聴牌とは、基本的には和了形となるために必要な牌の枚数が残り1枚であり、かつその和了に必要な牌種を既に手牌に4枚持っていない状況のことである。しかし、偶然役を除く役が成立する見込みはないが、残り1枚で4組の完成面子と1組の対子が揃う状況も聴牌と呼ばれる。このような聴牌は、特に形式聴牌と呼ばれる。偶然役とは、2.1.5の表2にて示している役のことである。また、本研究では、形式聴牌以外の聴牌を役有聴牌とする。

和了の方法には、自摸牌で和了する自摸和了と他家の捨て牌で和了する栄(ロン)和了がある。自摸和了した場合、和了点を他家3人で分割し、各他家から点数を貰う。栄和了した場合、栄和了された他家のみから和了点を貰う。栄和了されて失点しないために、他家が役有聴牌しているか否かを予測することが重要である。

2.1.4 麻雀用語の説明

一般的な麻雀用語を表1に示す。

表1 麻雀用語

和了(ホーラ)	和了形の手牌を公開して、点数を獲得する行為
和了形	以下のいずれかの条件を満たす手牌 1 役を成立させ、4組の完成面子と1組の対子を揃える 2 牌種の異なる7つの対子を揃える 3 么九牌を13種14枚揃える
自摸和了(ツモホーラ)	自摸牌で和了する行為
栄和了(ロンホーラ)	他家の捨て牌で和了する行為
放銃(ホウジュウ)	自分の打牌に対して他家が栄和了する行為
降り	他家に栄されないように打牌する行為
役	和了した時の手牌の特定の組合せ、 および、和了した時の特定の状況
偶然役	偶然の状況により成立する役
染め手	萬子、筒子、索子のいずれか1種類のみか、 萬子、筒子、索子のいずれか1種類と字牌で 手牌が構成された時に成立する役
喰いタン	副露した状況で断么九を成立させること
喰い下がり	副露した場合、翻数が1つ下がること
ドラ	和了時手牌にある場合、1枚につき翻数が1つ上がる牌

赤牌	赤く着色された5の数牌のこと (和了時手牌にある場合、1枚につき翻数が1つ上がる)
自風牌 (ジカゼハイ)	自分の風に該当する風牌 (東家は字牌の東、南家は字牌の南)
場風牌 (バカゼハイ)	現在の局における場と同じ風牌 (東場ならば字牌の東、南場ならば字牌の南)
自摸 (ツモ)	積まれた牌から牌を1枚引く行為
手出し	自摸牌を手牌に加え、自摸牌以外を捨てる行為
自摸切り	自摸牌を手牌に加え、そのまま捨てる行為
副露 (フーロ)	他家の捨て牌を取得することにより面子を完成させる行為
碰 (ポン)	対子を刻子にする副露
吃 (チー)	塔子を順子にする副露
槓 (カン)	刻子を槓子にする行為 4つの牌を自分で揃える暗槓 (アンカン)、 他家の捨て牌を取得することにより手牌の刻子を 槓子にする明槓 (ミンカン)、 碰することで完成した刻子に 自摸牌を加える加槓 (カカン)がある
門前 (メンゼン)	副露していない状況
面子 (メンツ)	特定の牌の組合せ 完成面子と未完成面子の総称
完成面子	刻子、槓子、順子の総称
刻子 (コーツ)	同じ牌種を3つ集めた組
槓子 (カンツ)	同じ牌種を4つ集めた組
順子 (シュンツ)	同色の数牌で連続した数字の牌を3つ集めた組
未完成面子	あと1枚で完成面子となる牌の組
対子 (トイツ)	同じ牌種を2つ集めた組
塔子 (ターツ)	後1枚で順子になる数牌2枚の組
聴牌 (テンパイ)	基本的には、以下の2つの条件を満たす状況 1 和了に必要な牌の枚数が残り1枚である 2 和了に必要な牌を既に手牌に4枚持っていない
形式聴牌	偶然役を除く役が成立する見込みはないが、 残り1枚で4組の完成面子と1組の対子が揃う状況
振聴 (フリテン)	聴牌しているが、和了形を揃えるために必要な牌種を 既に打牌してしまっているような状況
空聴 (カラテン)	聴牌しているが和了牌が全て見えており、 和了の望みがない状況
么九牌 (ヤオチュウハイ)	1、9の数牌と字牌の総称
中張牌 (チュンチャンパイ)	2~8の数牌の総称

2.1.5 天鳳で認められている役の一覧

天鳳で認められている役を表2に示す。

表中の役判定とは、本研究において形式聴牌であるか否かの判定に使用したことを指す。

表2 天鳳で認められている役の一覧

役	飜数	門前のみ	食い下がりあり	偶然役	役判定
門前清自摸和 (メンゼンチンツモホー)	1	○			
立直 (リーチ)	1	○			
槍槓 (チャンカン)	1			○	
嶺上開花 (リンシャンカイホウ)	1			○	
海底摸月 (ハイテイラオユエ)	1			○	
河底撈魚 (ホウテイロン)	1			○	
役牌 (ヤクハイ)：門風牌・場風牌・三元牌	1				○
断么九 (タンヤオチュー)	1				○
一盃口 (イーペーコー)	1	○			
平和 (ピンフ)	1	○			
混全帯么九 (ホンチャンタイヤオチュー)	2		○		○
一気通貫 (イツキツウカン)	2		○		○
三色同順 (サンショクドウジュン)	2		○		○
両立直 (ダブルリーチ)	2	○		○	
三色同刻 (サンショクドウコー)	2				○
三槓子 (サンカンツ)	2				○
対々和 (トイトイホー)	2				○
三暗刻 (サンアンコー)	2				○
小三元 (ショウサンゲン)	2				○
混老頭 (ホンロートー)	2				○
七対子 (チートイツ)	2	○			
純全帯么九 (ジュンチャンタイヤオチュー)	3		○		○
混一色 (ホンイーソー)	3		○		○
二盃口 (リャンペーコー)	3	○			
清一色 (チンイーソー)	6		○		○
天和 (テンホー)	役満	○		○	
地和 (チーホー)	役満	○		○	
大三元 (ダイサンゲン)	役満				○
四暗刻 (スーアンコー)	役満	○			
四暗刻単騎 (スーアンコータンキ)	役満	○			
字一色 (ツーイーソー)	役満				○
緑一色 (リュウイーソー)	役満				○
清老頭 (チンロートー)	役満				○
国士無双 (コクシムソウ)	役満	○			
国士無双13面	役満	○			
大四喜 (ダイスーシー)	役満				○
小四喜 (ショウスーシー)	役満				○
四槓子 (スーカンツ)	役満				○
九蓮宝燈 (チュウレンポウトウ)	役満	○			
純正九蓮宝燈 (ジュンセイチュウレンポウトウ)	役満	○			

2.2 二項ロジスティック回帰と全結合ニューラルネットワーク

本節では、本研究で扱う二項ロジスティック回帰と全結合ニューラルネットワークについて説明する。二項ロジスティック回帰は、一般化線形モデルの1種であり2値の目的変数を扱う場合に用いられる。全結合ニューラルネットワークとは、人間の脳神経系のニューロンを数理モデル化したものであり、隣接層のユニット同士がすべて結合しているニューラルネットワークである。

本節の内容は、書籍 [8] [9] から本実験に用いた手法についての記述を抜粋したものである。

2.2.1 二項ロジスティック回帰

二項ロジスティック回帰は、一般化線形モデルの1種であり、2値の目的変数を扱う場合に用いられる。ここでは、ある現象 A が説明変数 \mathbf{x} の下で生起するか否かを予測するという問題について考える。その説明変数 \mathbf{x} は以下のように表す。

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

ただし、 $x_i (= 1, \dots, n)$ は実数値を表す。現象 A が生起する条件付き確率 $p(A|\mathbf{x})$ を以下のように表す。

$$p(A|\mathbf{x}) = \frac{\exp(g(\mathbf{x}))}{1 + \exp(g(\mathbf{x}))} \quad (1)$$

以降、 $p(A|\mathbf{x})$ は $p(\mathbf{x})$ と略記する。 $g(\mathbf{x})$ は、 $p(\mathbf{x})$ のロジット、対数オッズなどと呼ばれる。 $g(\mathbf{x})$ は、 $p(\mathbf{x})$ で以下のように表すことができる。

$$g(\mathbf{x}) = \log\left(\frac{p(\mathbf{x})}{1-p(\mathbf{x})}\right) = \text{logit } p(\mathbf{x}) \quad (2)$$

二項ロジスティック回帰モデルは、以下のように表すことができる。

$$\text{logit } p(\mathbf{x}) = w_1x_1 + w_2x_2 + \dots + w_nx_n + b \quad (3)$$

ここで、 $\mathbf{w} = (w_1, \dots, w_n)$ はパラメータ、 b はバイアスである。

2.2.2 全結合ニューラルネットワーク

全結合ニューラルネットワークとは、隣接層のユニット同士がすべて結合しているニューラルネットワークのことである。

ネットワークを構成する各層のユニットは、複数の入力から総入力を計算し、活性化関数を通した値を出力する。例えば、第1層のユニットを添え字 $i = 1, \dots, I$ 、第2層のユニットを $j = 1, \dots, J$ で区別し、第1層のユニット i の入力を x_i 、第2層のユニット j の総入力を u_j 、出力を p_j とすると、全結合層1つ分の計算手順は以下ようになる。

$$u_j = \sum_{i=1}^I w_{ji}x_i + b_j \quad (4)$$

$$p_j = f(u_j) \quad (5)$$

式(4)で, x_i は入力, w_{ji} はパラメータ, b_j はバイアスである. これらをベクトルと行列を用いて表記すると以下のようになる.

$$\mathbf{u} = \mathbf{W}\mathbf{x} + \mathbf{b} \quad (6)$$

$$\mathbf{p} = \mathbf{f}(\mathbf{u}) \quad (7)$$

ただし, 各ベクトルと行列は以下のように定義する.

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_j \\ \vdots \\ u_J \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_J \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_j \\ \vdots \\ b_J \end{bmatrix}, \mathbf{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_j \\ \vdots \\ p_J \end{bmatrix},$$

$$\mathbf{W} = \begin{bmatrix} w_{11} & \cdots & w_{1I} \\ \vdots & \ddots & \vdots \\ w_{j1} & \cdots & w_{jI} \end{bmatrix}, \mathbf{f}(\mathbf{u}) = \begin{bmatrix} f(u_1) \\ \vdots \\ f(u_j) \end{bmatrix}$$

ここで, 活性化関数とは, ニューラルネットワークにおいて, 線形変換した後に適用する非線形関数もしくは恒等関数のことである. ユニットが持つ活性化関数には, 通常単調増加する非線形関数を用いられる. 本研究で使用した活性化関数を以下に示す.

- ・ロジスティックシグモイド関数

$$f(u) = \frac{1}{1 + \exp(-u)} \quad (8)$$

- ・正規化線形関数 (ReLU)

$$f(u) = \max(u, 0) \quad (9)$$

2.2.3 交差エントロピー誤差

交差エントロピー誤差とは, 二項ロジスティック回帰モデルやニューラルネットワークの出力と訓練データとの近さを示す尺度の 1 つである. この値は, 訓練データに対するパラメータ \mathbf{w} の尤度を求め, その対数を取り, さらに符号を反転させることで求めることができる.

訓練データとは, 入力 \mathbf{x} とこれに対する正解ラベル $d \in \{0,1\}$ の組の集合であり, 以下のように表すことができる.

$$D = \{(\mathbf{x}_1, d_1), (\mathbf{x}_2, d_2), \dots, (\mathbf{x}_N, d_N)\}$$

各 $n(= 1, \dots, N)$ の組 (\mathbf{x}_n, d_n) に対して, 入力 \mathbf{x}_n を与えたときの二項ロジスティック回帰モデルやニューラルネットワークの出力 $p(\mathbf{x}_n)$ が, d_n に近づくように \mathbf{w} を調整することはモデルやニューラルネットワークの学習と呼ばれる.

訓練データに対する \mathbf{w} の尤度は, 以下の式で与えられる.

$$L(\mathbf{w}) = \prod_{n=1}^N \{p(\mathbf{x}_n)\}^{d_n} \{1 - p(\mathbf{x}_n)\}^{1-d_n} \quad (10)$$

この尤度の対数を取り, さらに符号を反転したことで得られる値は交差エントロピー誤差と呼ばれ, 以下の式で求められる.

$$E(\mathbf{w}) = - \sum_{n=1}^N [d_n \log p(\mathbf{x}_n) + (1 - d_n) \log(1 - p(\mathbf{x}_n))] \quad (11)$$

2.2.4 勾配降下法と確率的勾配降下法

学習の目的は、与えられた訓練データに対する誤差関数 $E(\mathbf{w})$ を、パラメータ \mathbf{w} について最小化することである。しかし、大域的な最小値を求めることは通常困難である。そこで代わりに局所的な極小点を求めることを考える。局所的な極小点は、何らかの初期値 $\mathbf{w}^{(0)}$ を出発点に \mathbf{w} を繰り返し更新する反復計算によって求めることができる。勾配降下法はそのような手法の1つである。

勾配降下法は、現在の \mathbf{w} を負の勾配方向 $(-\nabla E)$ に少し動かし、これを何度も繰り返す。1ステップ分のパラメータの更新は、現在のパラメータを $\mathbf{w}^{(t)}$ 、更新後のパラメータを $\mathbf{w}^{(t+1)}$ とすると、以下の式で表すことができる。

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} - \epsilon \nabla E \quad (12)$$

ϵ は学習係数または学習率と呼ばれ、 \mathbf{w} の更新量の大きさを定める。本節では、 \mathbf{w} を $\mathbf{w} = (w_1, \dots, w_N)^\top$ とする。初期値 $\mathbf{w}^{(0)}$ を適当に決め、式(12)を $t = 0, 1, \dots$ に対し繰り返し計算することで $\mathbf{w}^{(1)}, \mathbf{w}^{(2)}, \dots$ を得る。ここで勾配は、以下のように表すことができる。

$$\nabla E \equiv \frac{\partial E}{\partial \mathbf{w}} = \left[\frac{\partial E}{\partial w_1}, \dots, \frac{\partial E}{\partial w_N} \right]^\top \quad (13)$$

勾配降下法には、訓練データに冗長性がある場合、計算効率が悪くなる問題がある。この問題を解決する手法として、確率的勾配降下法がある。勾配降下法では、全訓練データ D に対して計算される誤差関数 $E(\mathbf{w})$ を最小化することを考えた。これに対して、確率的勾配降下法では、一部の訓練データに対して計算される誤差関数 $E_n(\mathbf{w})$ を最小化することを考える。ここで、訓練データが入力 \mathbf{x}_n と正解ラベル $d_n \in \{0, 1\}$ の組であるとする、 $E_n(\mathbf{w})$ は以下のように表すことができる。

$$E_n(\mathbf{w}) = -d_n \log p(\mathbf{x}_n) - (1 - d_n) \log(1 - p(\mathbf{x}_n)) \quad (14)$$

確率的勾配降下法は、勾配降下法と同様に現在の \mathbf{w} を負の勾配方向 $(-\nabla E)$ に少し動かし、これを何度も繰り返す。1ステップ分のパラメータの更新は、現在のパラメータを $\mathbf{w}^{(t)}$ 、更新後のパラメータを $\mathbf{w}^{(t+1)}$ とすると、以下の式で表すことができる。

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} - \epsilon \nabla E_n \quad (15)$$

本研究では、 n の値は、 t を N で割ったときの余りに1を足した値とする。また、データの並びに規則性を無くするため、 N 回パラメータの更新を行う度に訓練データの添え字を付け替えてシャッフルする。勾配降下法と同様に初期値 $\mathbf{w}^{(0)}$ を適当に決め、式(15)を $t = 0, 1, \dots$ に対し繰り返し計算することで $\mathbf{w}^{(1)}, \mathbf{w}^{(2)}, \dots$ を得る。ここで、勾配は以下のように表すことができる。

$$\nabla E_n \equiv \frac{\partial E_n}{\partial \mathbf{w}} = \left[\frac{\partial E_n}{\partial w_1}, \dots, \frac{\partial E_n}{\partial w_N} \right]^T \quad (16)$$

2.3 ROC 曲線と AUC 値

ROC 曲線は、偽陽性率と真陽性率の関係をグラフにしたものであり、これは分類器の性能を評価するために用いられる。ROC 曲線を用いることで、クラス間のデータ数の偏りに関係なく分類器の性能を評価することができる。AUC 値とは、分類器の性能を表す評価尺度として用いられる値であり、ランダムな分類器の場合 0.5、完全な分類器の場合 1 となる。

本節の内容は、書籍 [10] から本実験に用いた手法についての記述を抜粋したものである。

2.3.1 ROC 曲線

ROC 曲線は、偽陽性率と真陽性率の関係をグラフにしたものであり、2 クラス分類の性能を評価するために用いられる。偽陽性率と真陽性率については後述する。

あるデータがあるクラス C に属していると分類器が判断する場合を p 、属していないと分類器が判断する場合を n とする。また、あるデータが実際に C に属している場合を p^* 、実際に C に属していない場合を n^* とする。したがって、分類器が分類したデータは、表 3 のように 4 つに場合分けされる。また、このようにクラス分類の結果をまとめた表のことを混同行列と呼ぶ。

表3 混同行列

		分類したクラス	
		p	n
実際のクラス	p^*	True Positive (TP)	False Negative (FN)
	n^*	False Positive (FP)	True Negative (TN)

この表から、以下のような性能評価値を計算することができる。

(1) 偽陽性率 (false positive rate) :

$$FPR = \frac{N_{FP}}{N_{FP} + N_{TN}} \quad (17)$$

(2) 真陽性率 (true positive rate) :

$$TPR = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (18)$$

(3) 正答率 (accuracy) :

$$p_A = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FN} + N_{FP} + N_{TN}} \quad (19)$$

ここで、 N_{TP} , N_{FN} , N_{FP} , N_{TN} は、それぞれ場合 TP, FN, FP, TN に属するデータ数を表す。偽陽性率とは、あるデータが実際には C に属していないにもかかわらず、分類器が C に属していると分類した割合のことである。真陽性率とは、あるデータが実際に C に属しており、分類器も C に属していると判断した割合のことである。正答率とは、あるデータが C に属しているか否かを正しく分類することができた割合のことである。

正答率を分類器の評価指標として用いたときにクラス間でデータ数に偏りがある場合、機能しなくなるという問題がある。例えば、 $N_{p^*} = 90$, $N_{n^*} = 10$ とする。このとき、分類器が常に C に属していると判断した場合、以下のような結果となる。

表4 データが場合 p^* に偏っている例

		分類したクラス	
		p	n
実際のクラス	p^*	$N_{TP} = 90$	$N_{FN} = 0$
	n^*	$N_{FP} = 10$	$N_{TN} = 0$

上記より、常に C に属していると判断する分類器の正答率を式 (19) から求めることができる。

$$p_A = 0.9$$

このようにクラス間でデータ数の偏りがあると、データの多いクラスに常に属していると判断する分類器でも高い正答率を示してしまう。そのため、クラス間のデータ数の偏りに依存しない指標が必要となる。

ROC 曲線は、閾値を変化させたときの偽陽性率と真陽性率の関係をグラフにしたものである。偽陽性率は実際に C に属していないデータのみから、真陽性率は実際に C に属しているデータのみから計算されるため、クラス間のデータ数に大きな違いがあったとしても、ROC 曲線は影響を受けることはないという性質を持っている。

具体例を用いて ROC 曲線の描き方を説明する。各データに対して、分類器が出力する予測スコアが以下ようになったとする。ここで、分類器が出力する予測スコアとは、あるデータが C に属している確率のことである。

表5 分類器のスコア (書籍 [10] より抜粋)

実際のクラス	予測スコア
p*	0.90
p*	0.80
n*	0.70
p*	0.55
p*	0.45
n*	0.40
n*	0.34
p*	0.30
n*	0.20
n*	0.10

この予測スコアが 0.55 以上の場合を p, 0.55 未満の場合を n というように閾値を決めて分類すると, 混同行列は表 6 のようになる.

表6 簡単な具体例

		分類したクラス	
		p	n
実際のクラス	p*	$N_{TP} = 3$	$N_{FN} = 2$
	n*	$N_{FP} = 1$	$N_{TN} = 4$

上表より, 偽陽性率と真陽性率を式 (17), (18) から求めることができる.

$$FPR = 0.2$$

$$TPR = 0.6$$

各予測スコアを閾値として変化させたときに, 各予測スコアに対する偽陽性率と真陽性率の関係をまとめたものが表 7 になる.

表7 予測スコアに対する偽陽性率と真陽性率 (書籍 [10] より抜粋)

実際のクラス	予測スコア	偽陽性率	真陽性率
p*	0.90	0.00	0.20
p*	0.80	0.00	0.40
n*	0.70	0.20	0.40
p*	0.55	0.20	0.60
p*	0.45	0.20	0.80
n*	0.40	0.40	0.80
n*	0.34	0.60	0.80
p*	0.30	0.60	1.00
n*	0.20	0.80	1.00
n*	0.10	1.00	1.00

上記の表を真陽性率と偽陽性率ごとにプロットすると，図2のようなグラフになる。

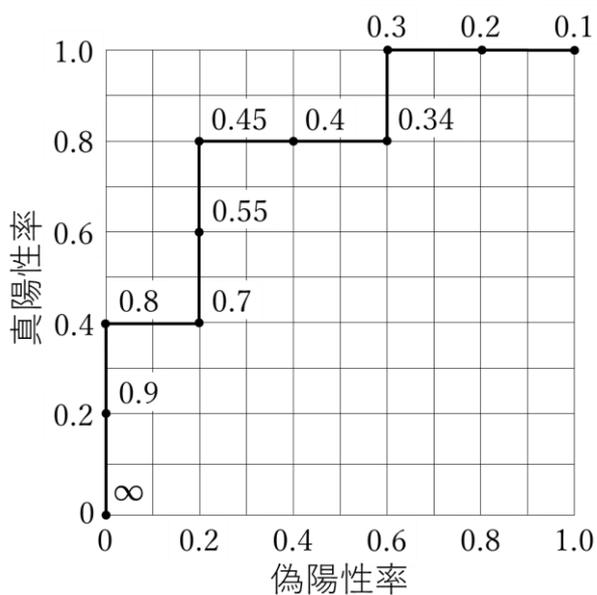


図2 分類器のスコアから描かれた ROC 曲線 (書籍 [10] より抜粋)

このように，閾値を変化させたときの偽陽性率と真陽性率による各点を結んだものが ROC 曲線である。

2.3.2 AUC 値

AUC 値とは、ROC 曲線の下部の面積のことであり、これは分類器の性能を表す評価尺度として用いられる。

完全な分類器の ROC 曲線は、原点から $(0, 1)$ を通り、 $(1, 1)$ を結ぶ折れ線となり、AUC 値は 1.0 となる。ランダムな分類器の ROC 曲線は原点と $(1, 1)$ を結んだ直線となり、AUC 値は 0.5 となる。したがって、通常 AUC 値は 1.0 と 0.5 の間にあり、大きい値ほど性能が良いことになる。図 2 の AUC 値は 0.8 であり、この分類器はランダムな分類器より性能が良いことが分かる。

3 先行研究

本章では、本研究と関連する先行研究を紹介する。ここで述べる先行研究は、本研究において聴牌確率推定を行ううえで参考にしたものである。

3.1 栗田らの放銃・放銃点の確率推定

不完全情報ゲームである麻雀では、通常他家の手牌を知ることができない。栗田らは、このような非公開情報を予測する研究を行った [5]。この研究では、他家の聴牌、他家の手牌または待ち形、放銃点の3つの非公開情報を予測した。

この研究の方法は、次のようである。まず、他家の聴牌確率と手牌または待ち形確率の推定を行い、それらの確率に基づき、放銃確率を推定した。他家の聴牌確率は、打牌数、副露数、染め手有無で場合分けをし、ロジスティック回帰を用いて、捨て牌の種類数など2つか4つの特徴から推定した。手牌確率は、他家が2副露以上している場合、栄可能なおおよそ全ての手牌に対して推定した。待ち形確率は、他家が副露していない場合、おおよそ全ての待ち形に対して推定した。副露していない他家への放銃点の確率は、翻と符がその他家の捨て牌によらない定常的な確率分布に従うという仮定に基づいて推定した。

この研究で得られた結果は、次のようである。放銃確率推定精度の検証では、打牌数10の場合、予測値と実測値の差が0.01程度であった。また、副露していない他家への放銃点の確率推定では、表8の結果を得た。低い翻での放銃を予測した場合、推定精度が高かったが、高い翻での放銃を予測した場合、推定精度は低かった。これは、高い翻数は他家の手牌に役牌やドラ牌が含まれていることに多くが起因し、予測が困難なためであると考えられる。

表8 放銃時の翻の予測値と実測値の分布 (文献 [5] より抜粋)

		予測			
		1翻	2翻	3翻	4翻以上
実測	1翻	0.231	0.105	0.025	0.010
	2翻	0.126	0.190	0.029	0.012
	3翻	0.034	0.112	0.028	0.005
	4翻以上	0.012	0.052	0.012	0.013

本研究との共通点は、打牌数、副露数で場合分けをして、他家の聴牌確率推定を行ったことである。本研究との相違点は大きく分けて2つある。1つ目は、本研究では染め手の有無で場合分けを行わずに聴牌確率推定を行っていることである。2つ目は、先行研究では、聴牌確率を推定するために、2つか4つの特徴からロジスティック回帰を用いていたが、本研究では、3,000程度の特徴から3層ニューラルネットワークも用いたことである。

3.2 水上らの他家の待ち牌・放銃点の確率推定

水上らは、他家の聴牌、他家の待ち牌、放銃点の3つの非公開情報を予測した [4].

この研究の方法は、次のようである。まず、他家の聴牌確率の推定を行った。次に、他家が聴牌している時、他家の待ち牌確率を推定した。さらに、その待ち牌を打牌して放銃した時の放銃点を推定した。他家の聴牌確率は、ロジスティック回帰を用いて、立直の有無や捨て牌の数などの6,888の特徴から推定した。待ち牌確率は、ロジスティック回帰を用いて、麻雀の牌の種類数である34種に対して、立直時に切った牌やドラの種類など31,416の特徴から推定した。放銃点は、重回帰分析を用いて、確定している役や見えているドラの枚数など26,889の特徴から推定した。

この研究で得られた結果は、次のようである。他家の聴牌確率推定精度は、上級者と同等程度であった。ここでの上級者とは、麻雀サイト天鳳において、鳳凰卓でプレイできるプレイヤーのことを指す。鳳凰卓でプレイできるプレイヤーは、全プレイヤーの内上位0.1%ほどである。他家の待ち牌推定精度は、ランダムに選択した場合より高く、上級者より低かった。放銃点の推定精度は、上級者を超える実力を示した。

本研究では、水上らが聴牌確率推定を行うために使用していた特徴を参考にした。本研究との相違点は大きく分けて2つある。1つ目は、本研究では打牌数、副露数で場合分けし、聴牌確率推定を行ったことである。2つ目は、先行研究では、聴牌確率を推定するためにロジスティック回帰のみ用いていたが、本研究では、3層ニューラルネットワークも用いたことである。

3.3 水上らの降りるべき局面の認識

コンピュータ麻雀プレイヤーの実力は、人間プレイヤーの平均レベルにも届いていない。この要因の1つは、多人数ゲームである麻雀では他家の状況によりプレイヤーの行動が変化してしまうからである。そこで、水上らは多人数性を削除した1人麻雀を考え、1人麻雀と4人麻雀の差を解析した。その結果、最も大きな差は降りであると結論づけ、降りる局面を機械学習により分類できるようにした [3].

この研究の方法は、次のようである。まず、1人麻雀プレイヤーを作成し、人間の上級者プレイヤーとの差を解析した。次に、最も大きな差であると考えられる降りる局面の認識を行った。ここでの上級者とは、麻雀サイト天鳳において、鳳凰卓でプレイできるプレイヤーのことを指す。1人麻雀プレイヤーを作成するため、学習には牌譜との一致を目指した平均化パーセプトロンを用いた。また、北川らの論文 [11] を参考に、向聴数や各字牌の枚数など37,488の特徴を用いた。1人麻雀と4人麻雀との差を解析するため、実際のある局面で上級者が打牌した牌と1人麻雀プレイヤーがその局面で最も打牌すべきとした牌を比較した。降りるべき局面は、Support Vector Machine を用いて、自分の手牌や他家の副露数など34の特徴から推定した。

この研究で得られた結果は、次のようである。1人麻雀プレイヤーの和了率は、上級者より

やや低く、平均プレイヤーより高かった。ここでの平均プレイヤーとは、麻雀サイト天鳳において上位 50%程度のプレイヤーのことである。実際のある局面で上級者が打牌した牌と 1 人麻雀プレイヤーがその局面で最も打牌すべきとした牌を比較した結果、一致率は 53%であり、打牌が一致しなかった局面のうち約 4 割が降りるべき局面であった。1 人麻雀プレイヤーと降りる局面の認識を組み込んだ 1 人麻雀プレイヤーをそれぞれ天鳳の一般卓で 500 戦以上対戦させた結果、後者の平均順位は前者より 0.23 高くなった。また、プレイヤーの実力を示すレーティングにおいて、後者は 1,507 と平均的なプレイヤーと同程度の実力を示した。

降りるべき局面とは、他家が聴牌しており放銃する可能性のある局面などであると考えられる。そのことから、水上らが行った降りるべき局面の認識は、間接的に聴牌確率推定も行っているのではないかと考え、本研究では、降りるべき局面の推定に用いていた特徴を参考にした。

4 目的

本研究の目的は、先行研究より正確に副露した他家の聴牌確率推定を行うことにより、より強い麻雀 AI プレイヤを作成することである。

既存の聴牌確率推定手法としては、水上らの 6,888 の特徴からロジスティック回帰を用いた手法や栗田らの打牌数、副露数などで場合分けし、22 の特徴からロジスティック回帰を用いた手法がある。水上らのこの聴牌確率推定を用いている AI プレイヤ (爆打) は、鳳凰卓でプレイするための条件である七段レーティング 2,000 を達成しており、人間の上級者と同等の実力があると考えられる。栗田らの上記の聴牌確率推定を用いている AI プレイヤ (Ako_Atmarshi) は、麻雀 AI プレイヤ対戦サーバ floodgate for mahjong [12] にて、爆打 1 体、AI プレイヤ (Manue) 2 体と対局を行った結果、最も平均順位が良かったことから、爆打と同等程度以上の実力があり、人間の上級者に匹敵すると考えられる。筆者の知る限り現在最も強い AI プレイヤに上記の 2 つの手法が用いられていることは、聴牌確率推定において、これらの手法は一定の成功を収めていると言える。水上らの聴牌確率推定が成功している理由は、多くの特徴から聴牌確率推定を行っているためであると考えた。また、栗田らの聴牌確率推定が成功している理由は、打牌数、副露数で場合分けしているためであると考えた。

そこで、本研究では、水上らと同様に多くの特徴を用い、栗田らと同様に打牌数、副露数で場合分けを行うことで更なる精度向上を目指す。また、先行研究ではロジスティック回帰を用いて聴牌確率推定を行っていたが、本研究では、3 層ニューラルネットワークを用いた推定精度も検証する。これらの手法を評価するために、本研究では正答率と AUC 値を用いる。最後に、推定精度の向上が AI プレイヤの強さに寄与するか検証するため、現在最も強いとされる AI プレイヤの 1 つである Ako_Atmarshi 3 体と Ako_Atmarshi に本研究の聴牌確率推定モデルを組込んだ AI プレイヤ 1 体とで対戦を行う。

5 実験方法

本章では、本研究における3つの聴牌確率推定モデルの作成方法について説明する。1つ目のモデルは、打牌数、副露数で場合分けを行い、他家の捨て牌などの特徴からロジスティック回帰を用いて聴牌確率推定を行う。2つ目のモデルは、打牌数、副露数で場合分けをせずに、1つ目のモデルと同様の特徴に打牌数と副露数を追加し、ロジスティック回帰を用いて聴牌確率推定を行う。3つ目のモデルは、打牌数、副露数で場合分けを行い、1つ目のモデルと同様の特徴から3層ニューラルネットワークを用いて聴牌確率推定を行う。

5.1節にて、聴牌確率推定モデルを作成する上で必要な牌譜について述べる。5.2節にて、その牌譜を用いてどのように訓練データとテストデータのためのラベル付けを行ったかを述べる。5.3節にて、聴牌確率推定に用いた特徴について述べる。5.4節にて、聴牌確率推定モデルの構成について述べる。

5.1 訓練データとテストデータに用いた牌譜

聴牌確率推定を行うモデルを作成するためには、多くの訓練データが必要となる。また、そのモデルの汎化性能を調べるためにテストデータが必要となる。訓練データとテストデータとして、麻雀サイト天鳳の鳳凰卓の牌譜を用いる。鳳凰卓でプレイできるプレイヤーは、全プレイヤーの内上位0.1%ほどであるため、鳳凰卓の牌譜は、麻雀上級者の牌譜のみの質の高いデータであると考えられる。データ数は全部で約72万半荘であり、全プレイヤー数は5,513人である。最も対戦回数の多いプレイヤーのデータ数は、1.4%以下である。

5.2 訓練データとテストデータのラベル付け

牌譜に記録されている打牌1つに対して、その打牌を行ったプレイヤーが聴牌しているか否かのラベルを1つ付ける。1つの牌譜から得られるラベルの数は、その牌譜に記録されている打牌の数と同じである。役有聴牌している場合に聴牌しているとラベル付けをし、そうでなければ聴牌していないとラベル付けをする。役有聴牌するためには、以下の3つの条件を満たす牌種 H が必要となる。

1. H と手牌で4面子1雀頭になる
2. H の牌が手牌に4枚ない
3. H と手牌で役が1つ以上成立する

ここでの役とは、2.1.5の表2で示した役判定に丸印のある役のことである。

上記の条件を設定する理由は、形式聴牌のような放銃する可能性がない聴牌を予測する必要はないと考えたからである。放銃する可能性のない聴牌として、形式聴牌以外に振聴や空聴があるが、上級者がこれらの状況になることは稀であるため、栗田らの聴牌確率推定と同様に振聴や空聴の場合は聴牌しているとラベルを付ける。

5.3 聴牌確率推定に用いた特徴

聴牌確率推定を行うために用いた特徴を表 9 に示す。表中の栗田らが用いた特徴とは、栗田らが聴牌確率推定に用いた特徴である。また、本研究にて用いた特徴は、水上らの聴牌確率推定と降りるべき局面認識に用いられた特徴を参考に設計したものに加えて、筆者が聴牌確率推定を行ううえで重要だと考えたものである。ここで、打牌数、副露数のフラグは、打牌数、副露数で場合分けしない聴牌確率推定モデルでのみ用いる。

表9 聴牌確率推定に用いた特徴

		特徴数
栗田らが用いた特徴	捨てた字牌の種類数	1
	捨てたある色の牌の種類数	3
	ある色の牌と字牌以外の打牌が最も連続した数	3
	ある色の牌と字牌以外の打牌が最も連続した数 × 最後の手出しがある色の牌か字牌のフラグ	3
	ある色の牌と字牌以外の打牌が最も連続する前に ある色の牌を打牌しているかのフラグ	3
	ある色の牌と字牌以外の打牌が最も連続する前に 字牌を打牌しているかのフラグ	3
	手出しの回数	1
	他家の立直後に手出しした回数	1
	捨てた么九牌の種類数	1
	捨てた中張牌の種類数	1
	本研究で用いた特徴	n 順目の捨て牌のフラグ
n 順目が手出しであるかどうかのフラグ		n
n 順目に行った副露の種類数のフラグ		$n \times 157$
n 順目の打牌時に他家が 立直しているかどうかのフラグ		n
局開始時のドラ牌のフラグ		34
赤ドラを捨てているかどうかのフラグ		1
自風と場風のフラグ		4
打牌数のフラグ ※場合分けをしない時のみ使用		18
副露数のフラグ ※場合分けをしない時のみ使用		4

6 実験結果

本章では、各手法の聴牌確率推定の精度を正答率と AUC 値で比較する。ここで、精度の比較の対象として次の手法 4 も導入する。

表11 栗田らの聴牌確率推定モデルの構成

手法	特徴数	重みの数	打牌数、副露数での場合分け	モデル
4	22	29	有	二項ロジスティック回帰モデル

6.1 節では、ゲーム状況を表す特徴が聴牌確率推定精度に与える影響を検証するために、特徴数が多い手法 1 と特徴数が少ない手法 4 を比較する。6.2 節では、打牌数と副露数での場合分けが聴牌確率推定精度に与える影響を検証するために、手法 1 と手法 2 を比較する。6.3 節では、3 層ニューラルネットワークが聴牌確率推定精度に与える影響を検証するために、手法 1 と手法 3 を比較する。

ここで、正答率の 95%信頼区間は以下の手順で見積もる。聴牌有無のラベル n 個が独立に生起したと考える。このデータに対する分類器の正答率を p とすると、標準誤差 E は以下のように求めることができる

$$E = \sqrt{\frac{p(1-p)}{n}} \quad (20)$$

よって、正答率の 95%信頼区間は $p \pm 1.96 \times E$ と見積もることができる。

6.1 ゲーム状況を表す特徴が聴牌確率推定精度に与える影響の検証

ゲーム状況を表す特徴が聴牌確率推定精度に与える影響を検証するために、多くの特徴を用いる手法1と少ない特徴を用いる手法4について、副露数ごとの正答率とAUC値を比較し、表12, 13, 14に示す。

表12 副露数1の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法4	手法1	手法4
打牌数 4	95.3±0.2	95.3±0.2	0.70	0.65
打牌数 5	91.9±0.2	91.9±0.2	0.69	0.64
打牌数 6	87.5±0.2	87.5±0.2	0.68	0.63
打牌数 7	82.8±0.3	82.8±0.3	0.67	0.64
打牌数 8	78.0±0.3	78.0±0.3	0.67	0.64
打牌数 9	73.4±0.3	72.8±0.3	0.69	0.64
打牌数 10	70.4±0.3	69.0±0.3	0.70	0.65
打牌数 11	69.5±0.3	66.5±0.3	0.72	0.67
打牌数 12	70.2±0.3	65.3±0.3	0.75	0.68
打牌数 13	71.5±0.3	65.6±0.3	0.77	0.69
打牌数 14	73.1±0.3	66.2±0.3	0.79	0.71
打牌数 15	75.7±0.3	68.3±0.3	0.82	0.73
打牌数 16	77.1±0.3	69.2±0.3	0.83	0.74
打牌数 17	78.3±0.3	71.0±0.3	0.84	0.75
打牌数 18	79.8±0.3	72.1±0.3	0.85	0.76

表13 副露数2の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法4	手法1	手法4
打牌数 4	82.7±0.3	82.7±0.3	0.64	0.59
打牌数 5	75.4±0.3	75.3±0.3	0.63	0.60
打牌数 6	68.3±0.3	67.7±0.3	0.64	0.60
打牌数 7	63.3±0.3	60.8±0.3	0.64	0.61
打牌数 8	61.3±0.3	57.2±0.3	0.65	0.61
打牌数 9	62.3±0.3	58.1±0.3	0.67	0.62
打牌数 10	64.5±0.3	60.6±0.3	0.69	0.63
打牌数 11	67.4±0.3	62.3±0.3	0.71	0.63
打牌数 12	70.2±0.3	63.8±0.3	0.73	0.63
打牌数 13	72.2±0.3	64.3±0.3	0.76	0.65
打牌数 14	73.8±0.3	64.2±0.3	0.78	0.66
打牌数 15	75.0±0.3	65.0±0.3	0.80	0.68
打牌数 16	76.0±0.3	65.9±0.3	0.82	0.70
打牌数 17	76.8±0.3	66.0±0.3	0.84	0.71
打牌数 18	77.6±0.3	66.9±0.3	0.85	0.72

表14 副露数3の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法4	手法1	手法4
打牌数 8	80.6±0.8	80.5±0.8	0.70	0.65
打牌数 9	83.6±0.8	83.4±0.8	0.72	0.67
打牌数 10	85.4±0.7	85.1±0.8	0.73	0.68
打牌数 11	86.6±0.7	86.1±0.7	0.75	0.69
打牌数 12	87.1±0.7	86.1±0.7	0.75	0.69
打牌数 13	87.2±0.7	86.1±0.7	0.77	0.70
打牌数 14	87.0±0.7	84.6±0.8	0.80	0.71
打牌数 15	86.7±0.7	83.6±0.8	0.82	0.72

表 12 より，副露数 1 の打牌数 10~18 において，特徴数を増やすことにより有意に正答率が向上し，AUC 値も向上した．また，打牌数が多いほど推定精度が高い傾向があった．これは，本研究の聴牌確率推定モデルの特徴は，打牌数が増えるごとに増えるためであると考えられる．表 13, 14 より，副露数 2, 3 でも多くの打牌数において正答率と AUC 値が向上した．以上の結果から，水上らの聴牌確率推定が成功した要因と考えられた多くの特徴を用いることで，聴牌確率推定精度が先行研究 [5] の手法より向上することを確認できた．

6.2 打牌数と副露数での場合分けが聴牌確率推定精度に与える影響の検証

打牌数と副露数での場合分けが聴牌確率推定精度に与える影響を検証するために，モデルの学習を打牌数，副露数で場合分けして行う手法 1 とモデルの学習を場合分けせずに行う手法 2 について，副露数ごとの正答率と AUC 値を比較し，表 15, 16, 17 に示す．ここで，推定精度の比較のため，テストデータは打牌数，副露数で場合分けをする．また，手法 2 の訓練データには，立直しているプレイヤーのデータも含まれている．

表15 副露数1の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法2	手法1	手法2
打牌数 4	95.3±0.2	95.2±0.2	0.70	0.69
打牌数 5	91.9±0.2	91.5±0.2	0.69	0.67
打牌数 6	87.5±0.2	87.5±0.2	0.68	0.66
打牌数 7	82.8±0.3	82.1±0.3	0.67	0.67
打牌数 8	78.0±0.3	76.7±0.3	0.67	0.67
打牌数 9	73.4±0.3	72.4±0.3	0.69	0.68
打牌数 10	70.4±0.3	69.3±0.3	0.70	0.70
打牌数 11	69.5±0.3	68.1±0.3	0.72	0.71
打牌数 12	70.2±0.3	69.3±0.3	0.75	0.72
打牌数 13	71.5±0.3	70.6±0.3	0.77	0.74
打牌数 14	73.1±0.3	71.6±0.3	0.79	0.77
打牌数 15	75.7±0.3	73.8±0.3	0.82	0.78
打牌数 16	77.1±0.3	74.9±0.3	0.83	0.80
打牌数 17	78.3±0.3	75.8±0.3	0.84	0.79
打牌数 18	79.8±0.3	77.4±0.3	0.85	0.80

表16 副露数2の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法2	手法1	手法2
打牌数 4	82.7±0.2	82.6±0.3	0.64	0.65
打牌数 5	75.4±0.3	75.3±0.3	0.63	0.64
打牌数 6	68.3±0.3	68.2±0.3	0.64	0.64
打牌数 7	63.3±0.3	62.7±0.3	0.64	0.65
打牌数 8	61.3±0.3	60.8±0.3	0.65	0.65
打牌数 9	62.3±0.3	61.5±0.3	0.67	0.65
打牌数 10	64.5±0.3	64.0±0.3	0.69	0.67
打牌数 11	67.4±0.3	67.0±0.3	0.71	0.69
打牌数 12	70.2±0.3	69.3±0.3	0.73	0.72
打牌数 13	72.2±0.3	71.7±0.3	0.76	0.74
打牌数 14	73.8±0.3	73.0±0.3	0.78	0.75
打牌数 15	75.0±0.3	74.1±0.3	0.80	0.76
打牌数 16	76.0±0.3	75.0±0.3	0.82	0.77
打牌数 17	76.8±0.3	75.9±0.3	0.84	0.79
打牌数 18	77.6±0.3	76.4±0.3	0.85	0.80

表17 副露数3の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法2	手法1	手法2
打牌数 8	80.6±0.8	80.6±0.8	0.70	0.66
打牌数 9	83.6±0.8	83.5±0.8	0.72	0.67
打牌数 10	85.4±0.7	85.1±0.8	0.73	0.67
打牌数 11	86.6±0.7	86.2±0.7	0.75	0.69
打牌数 12	87.1±0.7	86.6±0.7	0.75	0.71
打牌数 13	87.2±0.7	87.0±0.7	0.77	0.72
打牌数 14	87.0±0.7	86.5±0.7	0.80	0.74
打牌数 15	86.7±0.7	85.4±0.7	0.82	0.75

表 15 より，副露数 1 の打牌数 7~18 において，打牌数，副露数で場合分けすることにより有意に正答率が向上し，AUC 値も向上した．また，打牌数が多いほどより推定精度が高い傾向があった．これは，打牌数で場合分けすることによって推定精度が向上した結果ではないかと考える．表 16, 17 より，副露数 2, 3 でも多くの打牌数において正答率と AUC 値が向上した．また，副露数 1 の場合と副露数 3 の場合の結果を比較して，副露数 3 の場合の方がより推定精度が向上している傾向が見られた．これは，副露数で場合分けすることによって推定精度が向上した結果ではないかと考える．

6.3 3 層ニューラルネットワークが聴牌確率推定精度に与える影響の検証

3 層ニューラルネットワークが聴牌確率推定精度に与える影響を検証するために，ロジスティック回帰を用いる手法 1 と 3 層ニューラルネットワークを用いる手法 3 について，副露数ごとの正答率と AUC 値を比較し，表 18, 19 に示す．

表 18 副露数 1 の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法3	手法1	手法3
打牌数 7	82.8±0.3	82.8±0.3	0.67	0.72
打牌数 10	70.4±0.3	72.0±0.3	0.70	0.74
打牌数 13	71.5±0.3	72.6±0.3	0.77	0.79
打牌数 16	77.1±0.3	77.4±0.3	0.83	0.84

表 19 副露数 2 の場合の聴牌確率推定精度比較

	正答率(%)		AUC値	
	手法1	手法3	手法1	手法3
打牌数 7	63.3±0.3	65.2±0.3	0.64	0.71
打牌数 10	64.5±0.3	67.2±0.3	0.69	0.74
打牌数 13	72.2±0.3	73.8±0.3	0.76	0.80
打牌数 16	76.0±0.3	77.2±0.3	0.82	0.83

表 18 より，副露数 1 の打牌数 10, 13 において，3 層ニューラルネットワークを用いることで正答率が有意に向上し，AUC 値も向上した．また，表 19 より，副露数 2 の打牌数 7, 10, 13, 16 においても，3 層ニューラルネットワークを用いることで正答率が有意に向上し，AUC 値も向上した．以上の結果から，ロジスティック回帰ではなく，3 層ニューラルネットワークを用いることで推定精度が向上することを確認できた．

7 対戦実験

本章では、聴牌確率推定精度の向上が麻雀 AI プレイヤの強さに寄与するか検証するために行った対戦実験の結果について説明する。

対戦実験は、現在最も強いとされる AI プレイヤの 1 つである Ako_Atatarashi 3 体と Ako_Atatarashi に本研究の聴牌確率推定モデル (手法 1) を組込んだ AI プレイヤ (Ako_Atatarashi+) 1 体とで、麻雀 AI プレイヤ対戦サーバ Mjai [13] にて行った。対戦ルールは、東風戦 (約 4 局)、赤牌有、喰いタン有である。

26,833 回対戦した結果を表 20 に示す。平均順位が 2.503 となり、Ako_Atatarashi より低い平均順位であったが、有意な差ではなかった。

表20 Ako_Atatarashi+の対戦成績

1位率	0.253 ± 0.006
2位率	0.246 ± 0.006
3位率	0.246 ± 0.006
4位率	0.255 ± 0.006
平均順位	2.503 ± 0.014
和了率	0.219 ± 0.003
放銃率	0.134 ± 0.002

ここで、平均順位の 95%信頼区間は、以下の手順で見積もる。 n 個の試合結果が、それぞれ独立に生じたと仮定する。1 位となった試合数を N_1 、2 位となった試合数を N_2 、3 位となった試合数を N_3 、4 位となった試合数を N_4 とすると、平均順位 μ 、分散 σ^2 、標準誤差 E は以下のように求めることができる。

$$\mu = \frac{N_1 + 2N_2 + 3N_3 + 4N_4}{n}$$

$$\sigma^2 = \frac{N_1(1 - \mu)^2 + N_2(2 - \mu)^2 + N_3(3 - \mu)^2 + N_4(4 - \mu)^2}{n}$$

$$E = \sqrt{\frac{\sigma^2}{n}}$$

そして、平均順位の 95%信頼区間は、 $\mu \pm 1.96 \times E$ と見積もることができる。また、1 位率、2 位率、3 位率、4 位率、和了率、放銃率の 95%信頼区間は、式 (20) を用いて見積もった。

8 おわりに

本研究では、副露した他家の聴牌確率推定を先行研究より正確に行うため、水上らと同様に多くの特徴を用い、打牌数、副露数で場合分けする手法を提案し、栗田らの聴牌確率推定モデルより高い精度でこの推定を行うことができた。また、特徴数を増やすこと、打牌数、副露数で場合分けすること、3層ニューラルネットワークを用いることによって、推定精度が向上することを確認した。以上の結果から、特徴数をさらに増やし、より大規模なネットワークを構築することで、さらに高精度な聴牌確率推定が可能になるのではないかという感触を得た。

対戦実験は、Ako_Atarashi と Ako_Atarashi+ とで行った。対戦1回には約15分を要した。聴牌確率推定精度の向上が麻雀AIプレイヤーの強さに寄与することを期待したが、AIプレイヤーの強さに与える影響は小さく、これを約27,000回の対戦実験では認識することはできなかった。

9 謝辞

本研究を進めるにあたり、指導教員の保木邦仁准教授には、研究の方向づけから論文の書き方の基本に至るまで、数多くのご助言、ご指導をいただきました。心より感謝申し上げます。

HEROZ株式会社 栗田萌氏には、麻雀AIを提供していただくとともに、研究のご指導もいただきました。心より感謝申し上げます。

最後になりましたが、保木研究室の皆様には、大学院生活を共に過ごし、支えていただきましたことに深く感謝致します。

参考文献

- [1] Murray Campbell, Arthur Joseph Hoane, and Feng hsiung Hsu. Deep blue. *Artif. Intell.*, Vol. 134, No. 1-2, pp. 57-83, January 2002.
- [2] David Silver, Aja Huang, Christopher J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, Vol. 529, pp. 484-503, 2016.
- [3] 水上直紀, 中張遼太郎, 浦晃, 三輪誠, 鶴岡慶雅, 近山隆. 降りるべき局面の認識による1人麻雀プレイヤーの4人麻雀への適用. ゲームプログラミングワークショップ 2013 論文集. pp. 1-7, 2013.
- [4] 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. pp. 48-55, 2017.
- [5] 栗田萌, 保木邦仁. 麻雀における他家の手牌と待ちの予測に基づく放銃確率推定. 情報処理学会研究報告. GI, Vol. 2017-GI-38, No. 5, pp. 1-8, 2017.
- [6] 二階堂亜樹. 麻雀を始めたい人のために. 成美堂出版, 2017.
- [7] 角田伸吾. 天鳳. <https://tenhou.net/> (2019/1/27 accessed)
- [8] 丹後俊郎, 山岡和枝, 高木晴良. ロジスティック回帰分析—SAS を利用した統計解析の実際. 朝倉書店, 1996.
- [9] 岡谷貴之. 深層学習. MLP 機械学習プロフェッショナルシリーズ. 講談社, 2015.
- [10] 平井有三. はじめてのパターン認識. 森北出版株式会社, 2012.
- [11] 北川竜平, 近山隆. 麻雀の牌譜からの打ち手評価関数の学習. ゲームプログラミングワークショップ 2007 論文集. pp. 76-83, 2007.
- [12] 水上直紀, 亀甲博貴, 万代悠作, 横山秀. floodgate for mahjong (仮). www.logos.ic.i.u-tokyo.ac.jp/~mizukami/slide/majong_server.pdf#search=%27http%3A%2F%2Fwww.logos.t.u-tokyo.ac.jp%2Fmjlog%2F%27/ (2019/1/27 accessed)
- [13] Hiroshi Ichikawa. Mjai. <https://github.com/gimite/mjai/> (2019/1/27 accessed)