

GBDT を用いた麻雀における危険牌予測手法

2018SC031 各務裕貴

指導教員:河野浩之

1 麻雀 AI 開発についての現状

不完全情報ゲームの予測は完全情報ゲームとそれと比べ難解である。なかでも麻雀はすべてのプレイヤーのツモや相手プレイヤーの打牌など運の要素が互いに関係しあう。さらに麻雀牌の種類は 40 種類を超えているためゲーム木が複雑になってしまう。

Microsoft により Suphx[1] が開発され麻雀プラットフォーム天鳳 [2] で AI として初めて十段を達成する快挙を成し遂げるなど近年では不完全情報ゲームの研究も活発となっている。しかしこのような計算機はコストが高く初学者が簡単に手を出せるものではない。そこでゲーム全体を表現するのではなくプレイヤーの打牌と上がり牌に注目してモデルを作ることにより、計算コストを抑えながら強力な予測モデルを作ることを目指す。

2 麻雀の打牌提案システムに関する関連研究

麻雀の打牌提案を行うためには麻雀の試合のデータを学習させる必要がある。試合データを学習させるために先行研究でどのように学習させているかについて先行研究ごとの特徴を表 1 に示す。

表 1 打牌提案システムの先行研究

著者	学習方式	特徴
我妻ら [3]	SVM	・人間とシステムの回答の完全一致率:13.4%
清水ら [4]	強化学習報酬予測	・平均順位が 1.5 ・点棒状況による押し引きの概念を付与

3 決定木を用いたロン牌予測システム

ロン牌予測システムの構成を図 1 に示す。本提案手法では上がったプレイヤーの打牌と上がり牌の関係について学習させた後、学習に用いていない試合データを使用して予測を行う。

図 1 の麻雀試合データはプレイヤーの行動について記述されている牌譜を利用する。天鳳と呼ばれる麻雀サイトから牌譜をダウンロードし、学習用データとして打牌とロン牌についての記述を抽出する。そのうちトレーニング用として打牌とロン牌をセットとして GBDT に学習させ、予測用の打牌データを用いて予測を行えるか検証を行う。

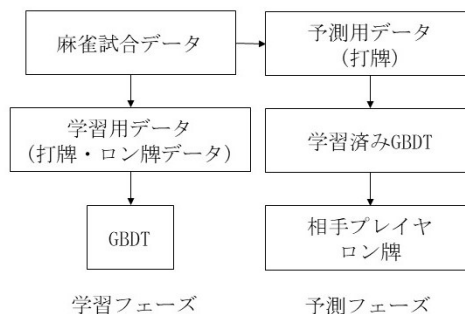


図 1 ロン牌予測システムのフロー

4 使用するデータと学習手法

4.1 学習データ:麻雀サイト「天鳳」の牌譜

教師データとしてオンライン麻雀サイト「天鳳」の牌譜と呼ばれる試合ログを利用した。天鳳ではプレイヤーの実力によって階級が分けられ、近い実力のプレイヤー同士で試合を行っている。本論文ではこのうち最上位の実力層である天鳳卓の試合にのみ牌譜を限定し、2014 年のデータを利用する。またプレイヤー数は四人と三人の場合があるが今回は四人のものを利用する。

4.2 機械学習器:GBDT

ロン牌予測システムには GBDT(勾配ブースティング決定木) という、勾配降下法と決定木をブースティングで強学習器とした手法を用いた。勾配降下法とは、とある点 x の傾きを計算して 0 になる場所 (最小値) を探すというアルゴリズムである。初期点 x_1 を定め x_1 での傾きを求め 0 であるか確認し、0 でない場合新たな点 x_2 を傾き $f'(x)$ と学習率 μ により求める、…というように傾きが 0 に等しい点 x_k を見つけるまで繰り返すが、初期点の定め方によっては局所解に陥る可能性がある。

決定木とは条件分岐をしてクラス分けを行うことで与えられた説明変数から目的変数を予測するものである。データを分割したとき異なるクラス同士がどれくらいの割合で混ざっているかという指標である不純度が 0 に近づくように決定木を成長させる。

ブースティングとは学習器の学習による誤差を引き継ぎ、新たな学習器の学習データとする手法である。与えられた学習データに対して最初の予測モデルが構築されたあと不正解のサンプルを重視するように次の予測モデルを構築して複数のモデルの予測値を組み合わせることで性能を向上させる。

これらの手法を使い一つの強学習器としたものがGBDTであり、図2のように組み合わせている。まず目的変数の平均値を計算しそれぞれの真値との誤差 ε を求め、それぞれの目的変数が分類される最下層葉ノードの重みとする。これが最下層葉ノードの +2.0 などの数値であり次の決定木の作成にも用いる。二目目以降の決定木において、誤差 ε を計算し決定木の最下層葉ノードの重みとするという処理を決められた回数で実行したあと各決定木の数値をあわせることで予測値となる。

図2において学習率が 0.1 で学習データにおけるロン牌番号の平均値が 17 であるとする。ある未知のデータにおいて決定木 1 では重み +2.0 ノードに、決定木 2 では重み +0.45 ノードに分類されたとき、予測値 p は $p = 17 + 0.1 \times 2.0 + 0.1 \times 0.45 + \dots$ と表される。

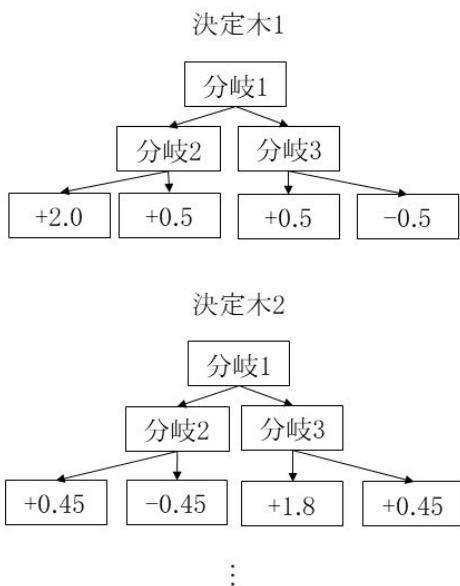


図2 GBDTの手法組み合わせイメージ

5 実験と考察

ロン牌予測システムを構築する実験について説明する。学習設定を表2のように設定し、葉の数のみを変更しながら実験する。ロン牌予測システムでは学習するデータに1局でなされた打牌数が10のものを使用している。

実験結果について説明する。表3に葉の数ごとの予測精度を、表4に葉の数を5としたときのboosting roundごとの予測精度を示す。葉の数を変えながら実験したところ表3のようになり、トレインデータとテストデータのlogloss平均値から葉の数が5の時に最も精度の高い予測を行うモデルを構築できることがわかった。また、葉の数を5に固定したところ表4のようになりboosting roundが366のときにearly stoppingした。このとき、ロン牌予測システムでは予測精度が6.505%程度であり先行研究[3]で得られた人間の回答との完全一致度13.4%を超えることができなかった。理由は2つ考えられ、1つ目は同じ

カラムに格納された牌番号の重みの違いから、2つ目は特徴量の少なさからだと考えられる。

表2 実験の学習設定

手法	GBDT
出力	多クラス
入力	34
学習率	0.1

表3 適した葉数の検証

葉数	正解率	適合率	再現率
4	0.06593	0.05364	0.05161
5	0.06497	0.05220	0.05159
6	0.06504	0.05240	0.05116
10	0.06441	0.05208	0.05026
20	0.06336	0.04724	0.04877

表4 適したboosting roundの検証

boost round	正解率	適合率	再現率
100	0.06741	0.05591	0.05115
200	0.06679	0.05374	0.05171
300	0.06571	0.05391	0.05164
366	0.06509	0.05274	0.05133

6 まとめ

本論文ではGBDTを用いて麻雀の危険牌予測手法を提案した。ロン牌予測システムを構築し予測精度を検証した結果、先行研究[3]で得られた予測精度を超えることができなかった。

参考文献

- [1] Japan News Center, “麻雀 AI Microsoft Suphx が人間のトッププレイヤーに匹敵する成績を達成”, <https://news.microsoft.com/ja-jp/2019/08/29/190829-mahjong-ai-microsoft-suphx/>, Aug. 29, 2019.
- [2] 有限会社シー・エッグ, オンライン対戦麻雀 天鳳, <https://tenhou.net/sc/raw/>, 参照 July. 18, 2021.
- [3] 我妻敦, 原田将旗, 森田一, 古宮嘉那子, 小谷善行, “SVRを用いた麻雀における捨て牌の危険度の推定,” 研究報告ゲーム情報学, 2014-GI-31 巻, 12 号, pp.1-3, Mar. 10, 2014.
- [4] 清水大志, 田中哲郎, “深層強化学習を用いた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2020論文集, 2020 巻, pp147-154, Nov. 2020.