

# 学習データの選別による将棋盤面評価値の改善に関する研究

1230314 門木 天馬 【高度プログラミング研究室】

## 1 はじめに

ゲーム AI の発展は将棋や囲碁などの二人完全情報ゲームにおいて、トッププレイヤーに勝利するほど強くなっている。近年では強いだけでなく、人を楽しませる AI の研究も行われている。将棋 AI は対局以外では棋譜の解析や対局の解説などで人間らしさを求められる場合がある。しかし対局の解説に将棋 AI が用いられた場合、人間には指せないような手が示されることがある。これは将棋 AI が最善のやりとりをした場合の探索結果に基づいた評価値であるため、人間にとっては高度すぎるのである。よって人間プレイヤーの実感や実際の対局結果に近い評価値が必要である。

小川と池田による研究 [1] では局面のみを入力とした教師あり学習と、局面とレートを入力とした教師あり学習の学習結果を比較することでレートを入力に加えることで予測結果が正確になることを示そうとした。研究結果は勝率予測、着手予測ともに有意な差が得られなかった。

本稿では学習データをレートごとに選別することで学習にどのような影響が出るのかを調査する。

## 2 提案手法

学習データの棋譜データを高レート、低レートの棋譜データに選別しそれぞれの学習モデルを作成、比較することで学習データのレートの値が学習に及ぼす影響を調査する。

ネットワークの構造、学習オプションは山岡による書籍 [2] を参考に作成した。学習は局面の勝率予測と着手確率分布の予測をするマルチタスク学習を行っている。訓練データは高レート、低レートそれぞれ 63,000 棋譜を使用し、テストデータには 700 棋譜を使用した。入力特徴は駒ごとの配置の配列、手番ごとの占有座標の配列、手番ごとの持ち駒の配列によるビットボードから作成した。入力特徴では、駒の種類ごとにチャンネルを分け、各チャンネルを駒の座標を表す 2 値画像とし、先手と後手それぞれ別のチャンネルとしている。持ち駒には種類ごとの最大枚数分のチャンネルを割り当てるので、合計チャンネル数は 104 となっている。中間層のフィルター枚数は 256、1 層目は畳み込み層、2 層目以降は Residual Network を 5 ブロック重ねている。

出力形式は、勝率予測は 2 値分類問題とし、着手確率予測では駒の移動方向と移動先の座標の組み合わせからなる 2187 ラベルを分類する多クラス分類問題とする。オプティマイザは SGD を使用した。

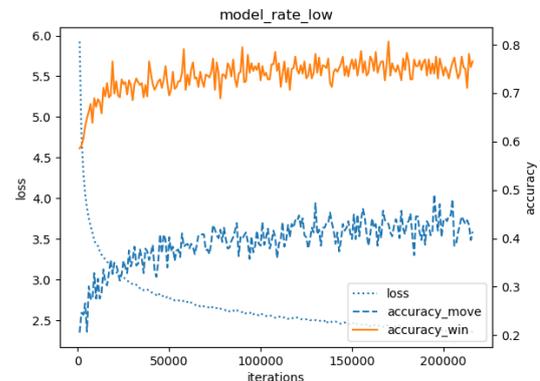


図1 低レートの棋譜データを使用した学習の様子

## 3 実験結果

図1は低レートの棋譜データを学習した様子であり、損失平均結果は 2.6712、勝率一致確率結果は 0.7513、着手一致確率結果は 0.4292 となった。高レートの棋譜データを学習した学習モデルでは、損失平均結果は 2.7966、勝率予測結果は 0.7004、着手一致確率結果は 0.4072 となった。

低レートの棋譜データを学習したモデルの方が損失平均が小さく、勝率一致確率、着手一致確率、共に高かった。これは高レートの対局では実力が拮抗しやすく、予測が難しくなったためだと考えられる。低レートの対局では反対に一度対局の有利、不利ができてしまうと逆転するのが難しいため、勝率、着手の予測が簡単になったためだと考えられる。また学習棋譜データを揃えたことで、学習局面数に差ができてしまい、低レート帯の学習局面数が多くなってしまったことで学習に差ができてしまったことが考えられる。

## 4 まとめ

本稿では学習データを高レートのデータと低レートのデータに選別して学習を行い、二つのモデルを比較することでレートが学習に与える影響を調査した。調査した結果に有意な差があるか評価し、なぜそのような差ができたのかを考察した。

## 参考文献

- [1] 小川竜欣, 池田心, “対局状況をより正確に表現するための盤面評価”, 第 26 回ゲームプログラミングワークショップ論文集, pp.28-33, 2021.
- [2] 山岡忠夫 “将棋 AI で学ぶディープラーニング”, マイナビ出版, 2017.