An AI for a Modification of Dou Di Zhu

By

Xuesong Luo

Advisor

Dr. Chris Pollett

Dr. Mark Stamp

Dr. Fabio Di Troia

San Jose State University

May 13th, 2020

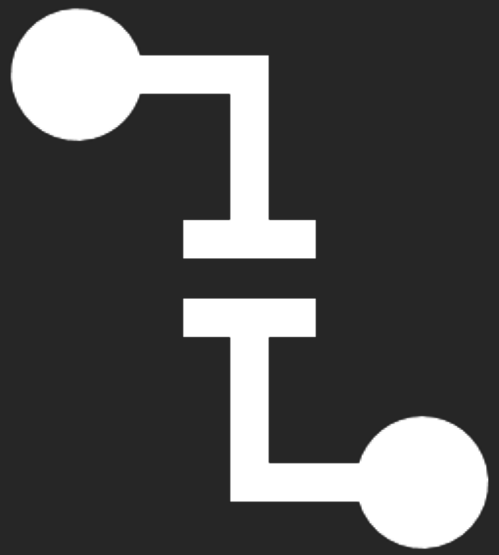
Outline

1. Introduction
2. Dou Di Zhu
3. Design
4. Experiments
5. Conclusion

Introduction

- This project is the implementation of AIs for the Chinese game Dou Di Zhu.
- Dou Di Zhu is a popular card game in China, and there are almost 1 million Dou Di Zhu players online.
- We design and implement a Deep Q-learning Neural Network (DQN) and Q-learning algorithm to play the Dou Di Zhu.
- Re-implement an exist rule-based model to compare with our model.

Related Work

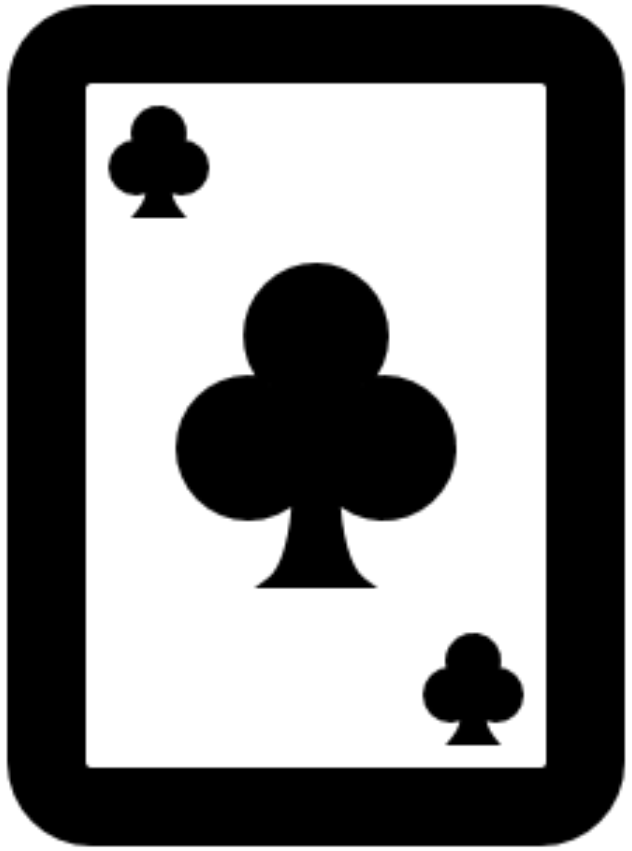


- Playing Dou Di Zhu by using the Rule-based model or Decision Tree model.

- Renzhi Wu, Shuai Liu, Shuqin Li, Meng Ding, “The design and implementation of a computer game algorithm of Dou Dizhu”, 2017
- Zhennan Yan, Xiang Yu, Tinglin Liu, Xiaoye Han, “Fight the Landlord (Dou Di Zhu)”

Dou Di Zhu

- Basic rules of Dou Di Zhu
  - Dou Di Zhu is a three players card game with a 54-card deck
  - Two sides: one player will be the landlord, and two other players will be the peasants
- The game has three stages:
  - Dealing cards
  - Bidding landlord
  - Playing cards



Dealing Cards

- At the beginning of a game, 17 cards are dealt to each of the three players as their hand cards.
  - There are totally 51 cards are dealt to the players.
- The three remaining cards are given to the landlord after the landlord is selected.

Bidding landlord

- After dealing 51 cards to players, players bid to become the landlord.
  - ✓ In normal game, people like rolling dice to decide the first player to select landlord.
  - ✓This player can give up to become the landlord, and the player to his right can choose whether or not to become the landlord. Keep going until one player select to become the landlord.
- The three left cards belong to the landlord's hand cards.
  - ✓Landlord has 20 cards, and every peasant has 17 cards.

Playing Cards

- Playing Order
  - ✓The landlord will play cards first at each round of the game. The next player always will be the current player's right side person.
- When the first player played all hand cards, game over.
  - ✓If the player is peasant, he and his peasant teammate win this round together.
  - ✓If the player is landlord, only the landlord win the game, and two peasants lose.

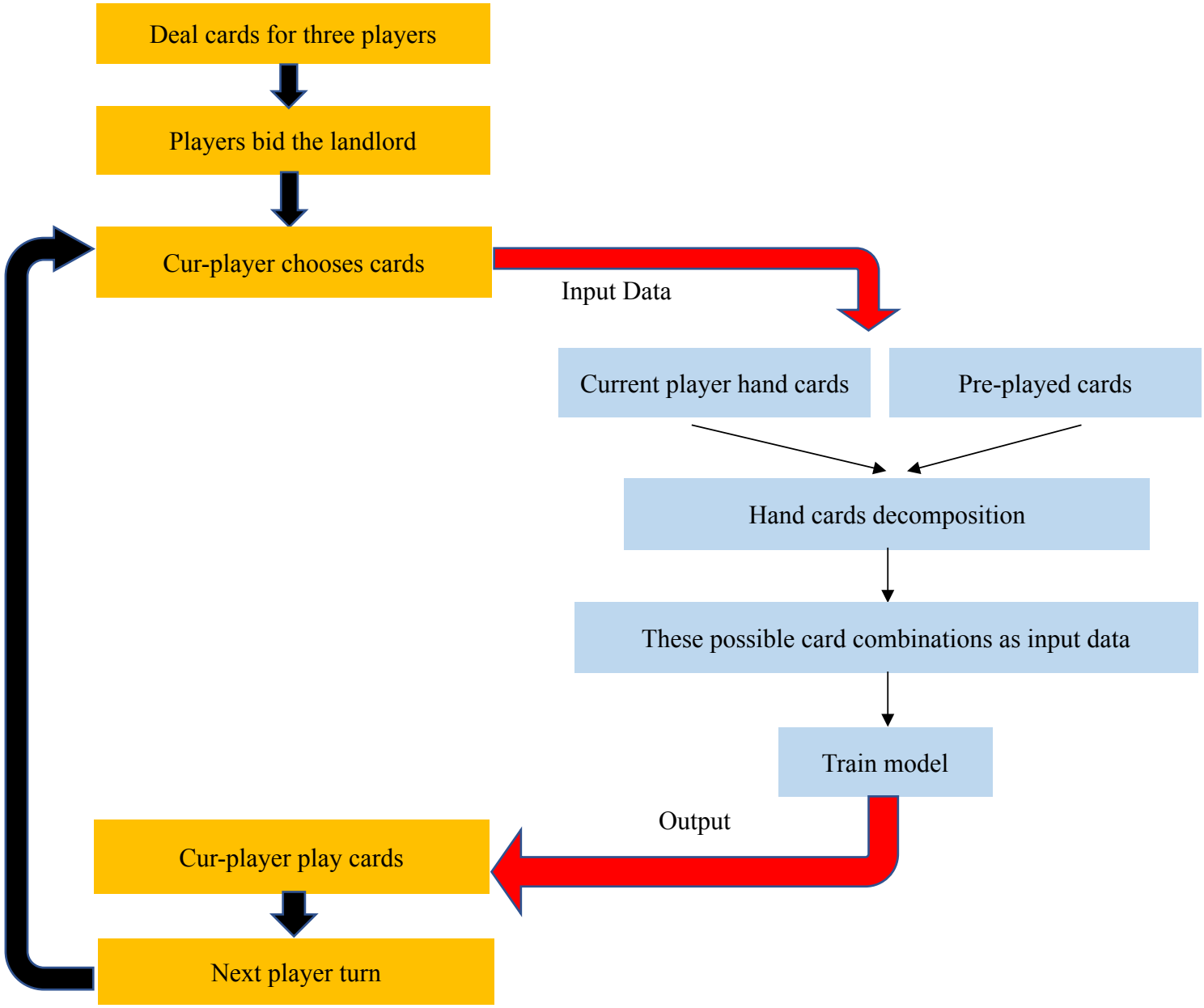
Card Combination

Card combination	Description
Rocket	Same as the Joker Bomb, both jokers (Red and Black), is the highest Bomb.
Bomb	Four cards with the same points. (e.g. AAAA)
Single	One single card. (e.g. A)
Pair	Two cards with the same points. (e.g. AA)
Triplet	Three cards with the same points. (e.g. AAA)
Triplet with an attached card/pair	Triplet with an attached card/pair. (e.g. AAA+B or AAA+BB)
Single Sequence	Five of more Singles in sequence excluding 2 and Jokers. (E.g. ABCDE or ABCDE...)
Double Sequence	Three of more pairs in sequence excluding 2 and Jokers. (E.g. AABBCC or AABBCC...)
Pass	Choose not to play a card this turn. It is also called as a trivial pattern.

Tools & Environments

- Python 3.6
  - Numpy package
- Tensorflow 2.2
  - Tensorflow: Created by the Google Brain team, is an open source library for numerical computation and large-scale machine learning.

Game Flow


Deal cards for three players
Players bid the landlord
Cur-player chooses cards
Input Data
Current player hand cards
Pre-played cards
Hand cards decomposition
These possible card combinations as input data
Train model
Output
Cur-player play cards
Next player turn

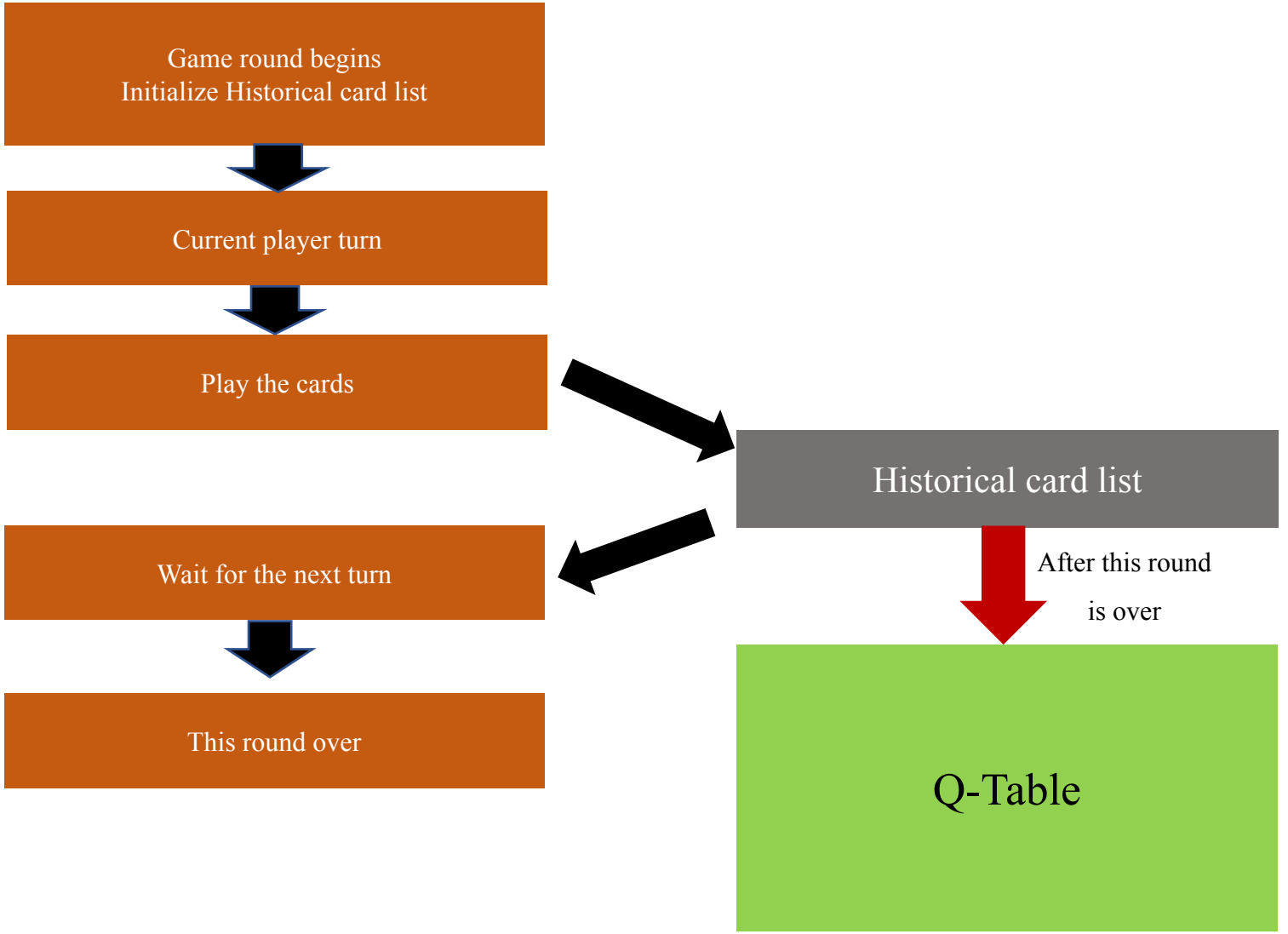
Dataset

- One deck has 54 cards with four suits.
- Use the order number 0-53 to present the 54 cards.
- Like order number 0, 1, 2, 3 corresponds to the card: 3-Heart, 3-Tile, 3-Clover, and 3-Pike.
  - card = n // 4
  - n is the card order number, to calculate the card number from 0-14
  - Result 0 is card 3, 1 is card 4, …, 9 is card J, 10 is card Q, …, 13 is card Black Jack, 14 is card Red Jack.

Q-Learning model

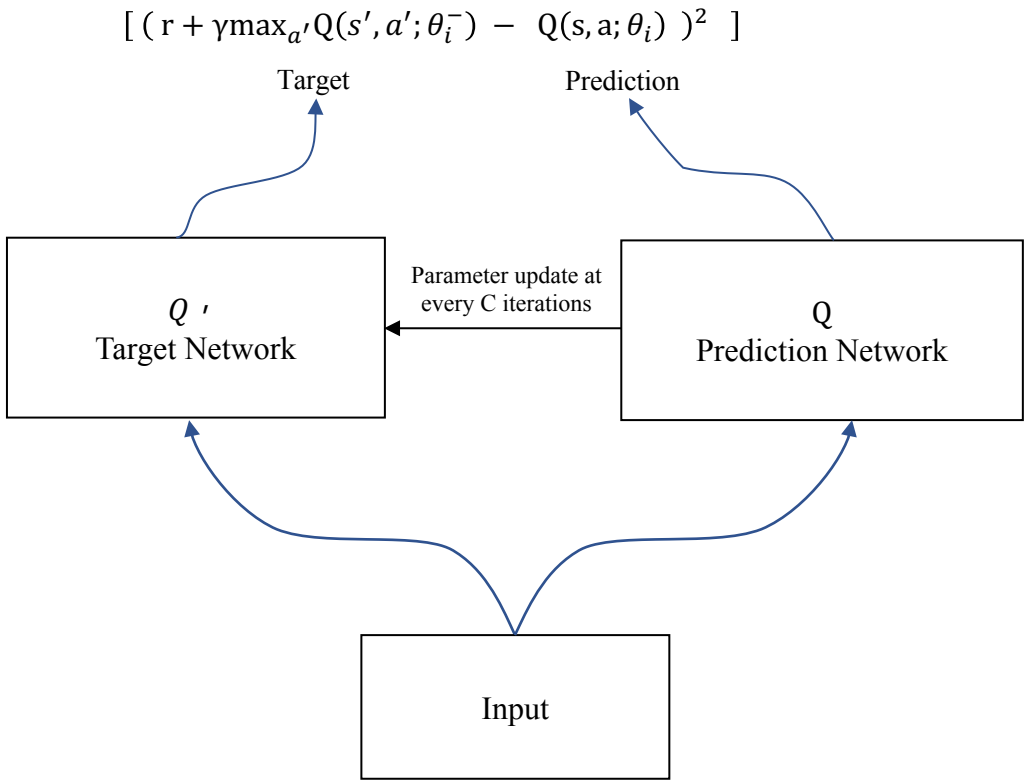
- Design a Q-learning model for playing the Dou Di Zhu.
- Utilized the Q-learning strategy that each player has an independent Q-Table to store the different playing action and corresponded reward.
- Every game round, the players' played cards keep saving in a temporary list separately for each turn.
- When the round is over, these card combinations will transfer into Q-Table, and update their rewards based on the win or loss of the game round.

Update the Q-learning strategy



DQN model

- The DQN model has two network: Target Network & Prediction Network.
- Except input layer and out layer, There are two hidden layers.
- One size of 500 memory pool
- The “e-greedy” strategy
- Every 300 iterations update the Target Network



Zhou rule-based model

- Zhou rule-based model is a kind of Rule-based model.
- This model based on a priority:
  - Triplet cards better than Sequence cards
  - Sequence cards better than Pair cards
  - Pair cards better than Single card
  - Single card better than Bomb cards
  - Bomb cards better than Rocket

Experiments and Observations

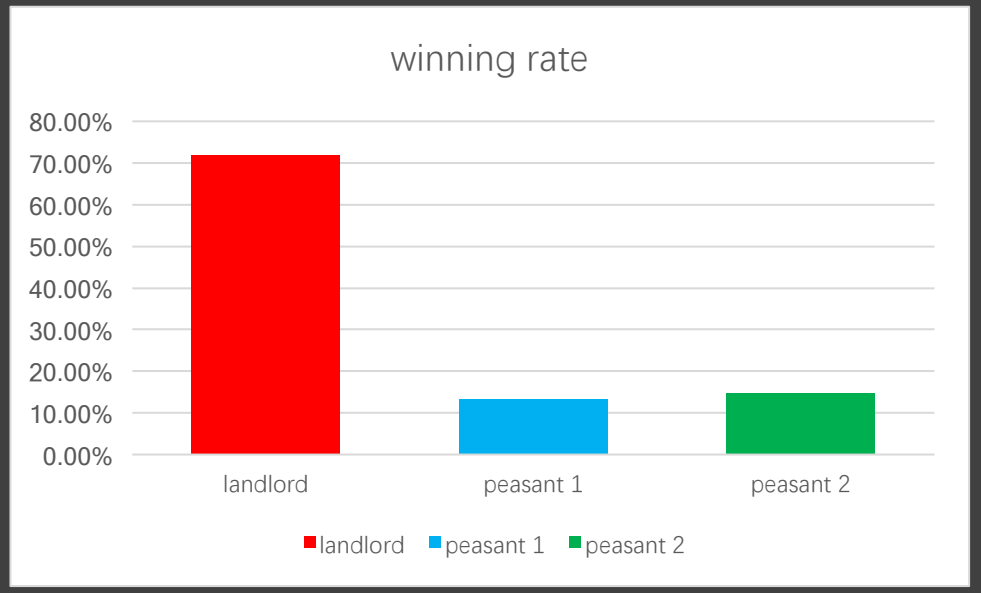
- Project is executed on a MacBook Pro with 8 GB memory; or desktop PC of Windows 10, with 24GB memory.
- I train each model for 100,000 games and test it for 10,000 games (Except Zhou rule-based model, we can directly test it 10,000 games).

- **Experiment 1:** DQN model Versus Random method
- **Experiment 2:** DQN model Versus Q-learning model
- **Experiment 3:** Q-learning model Versus Random method
- **Experiment 4:** Zhou rule-based model Versus Random method

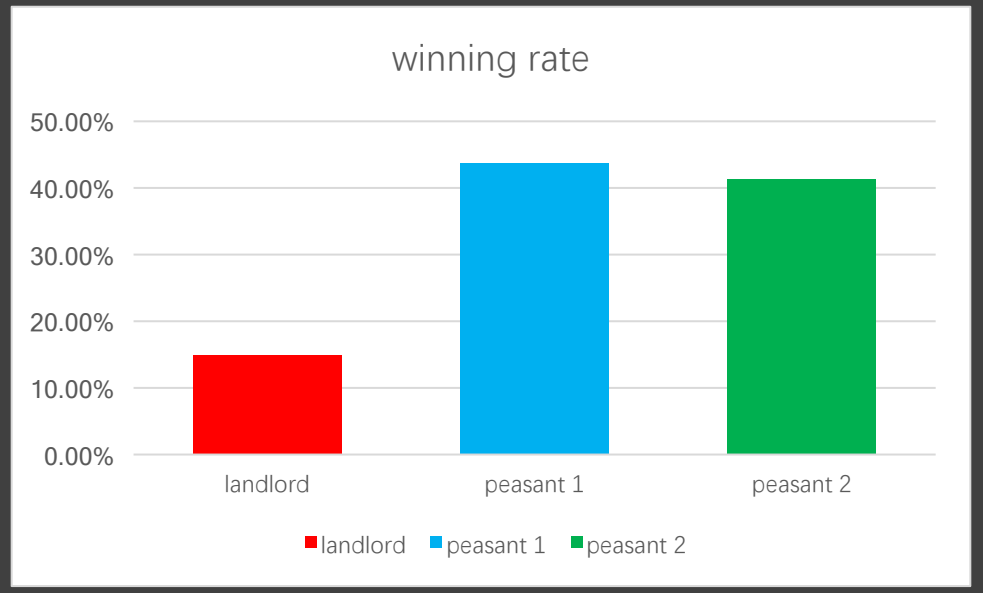
Experiments and Observations

- **Experiment 1:** DQN model Versus Random method

DQN landlord VS Random peasants



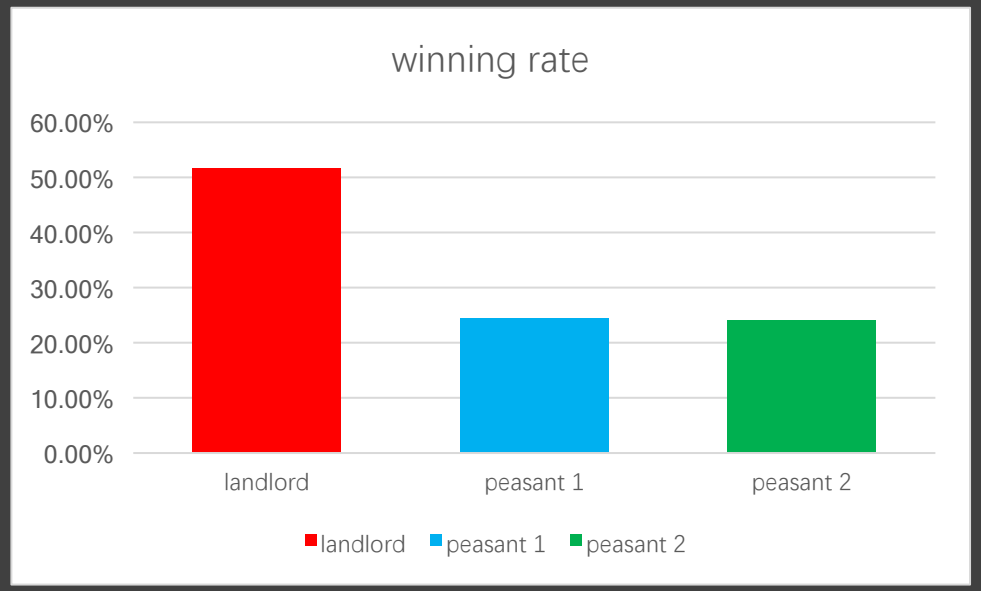
Random landlord VS DQN peasants



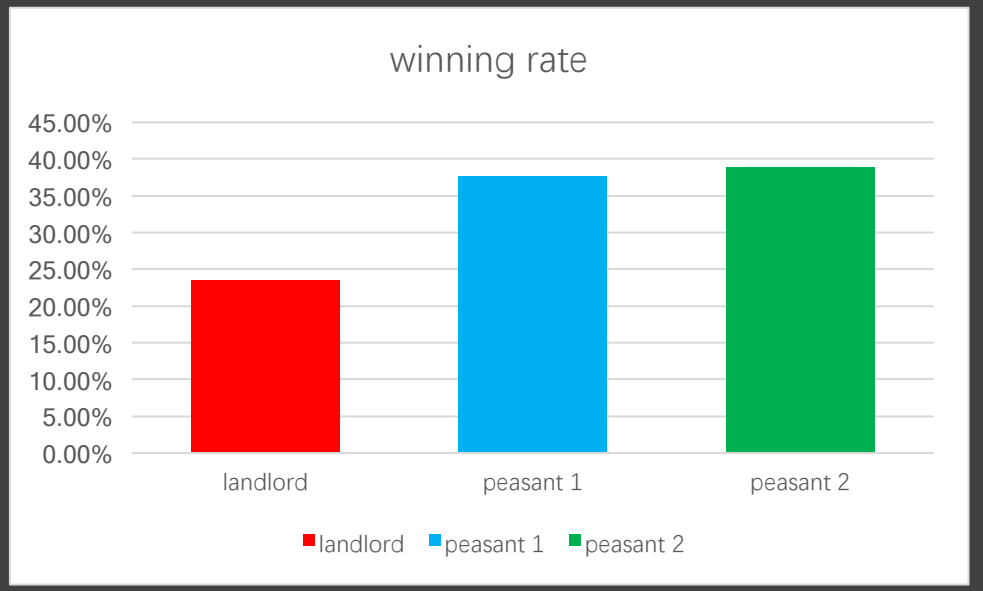
Experiments and Observations

- **Experiment 2:** DQN model Versus Q-learning model

DQN landlord VS Q-learning peasants



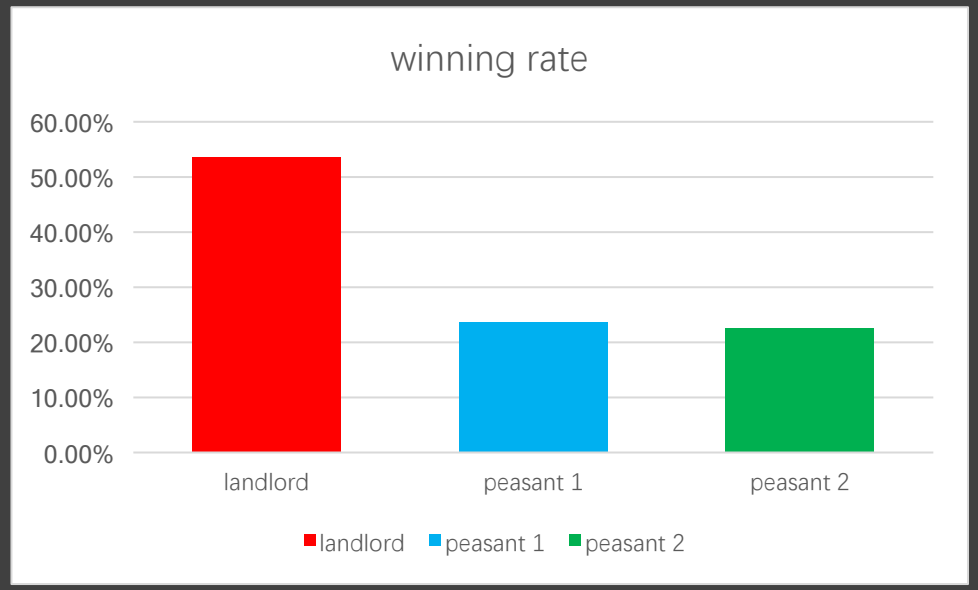
Q-learning landlord VS DQN peasants



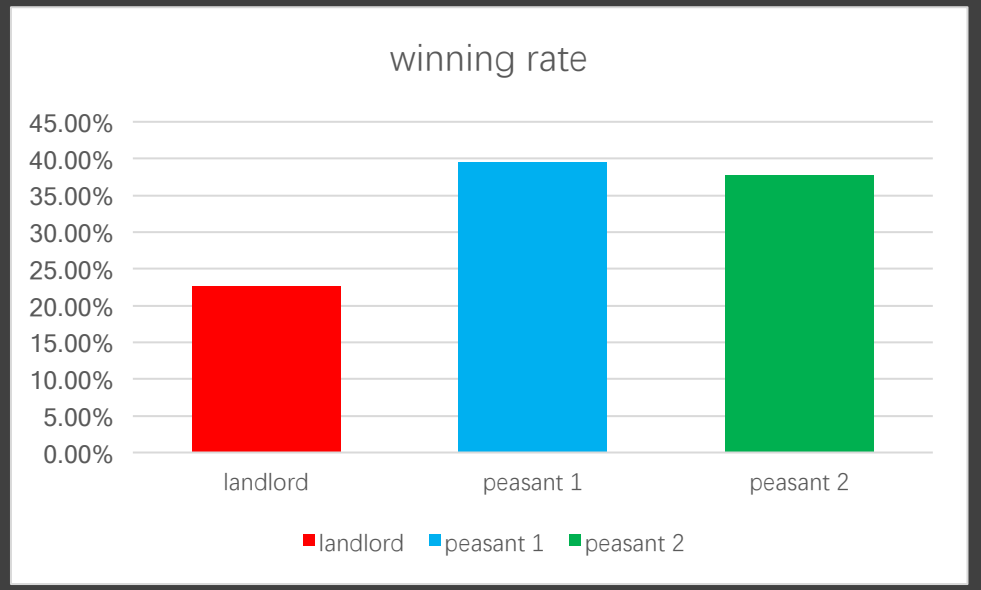
Experiments and Observations

- **Experiment 3:** Q-learning model Versus Random method

Q-learning landlord VS Random peasants



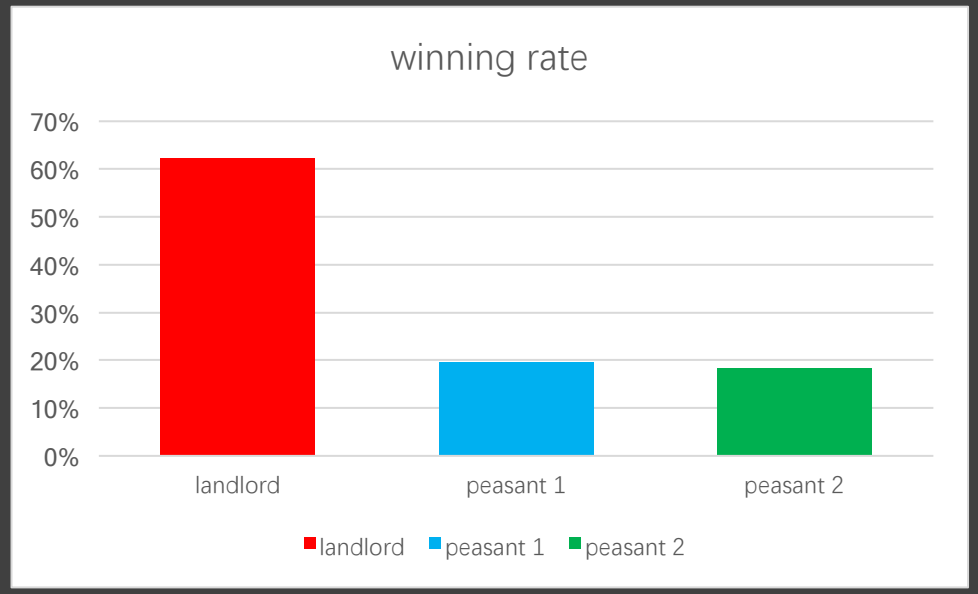
Random landlord VS Q-learning peasants



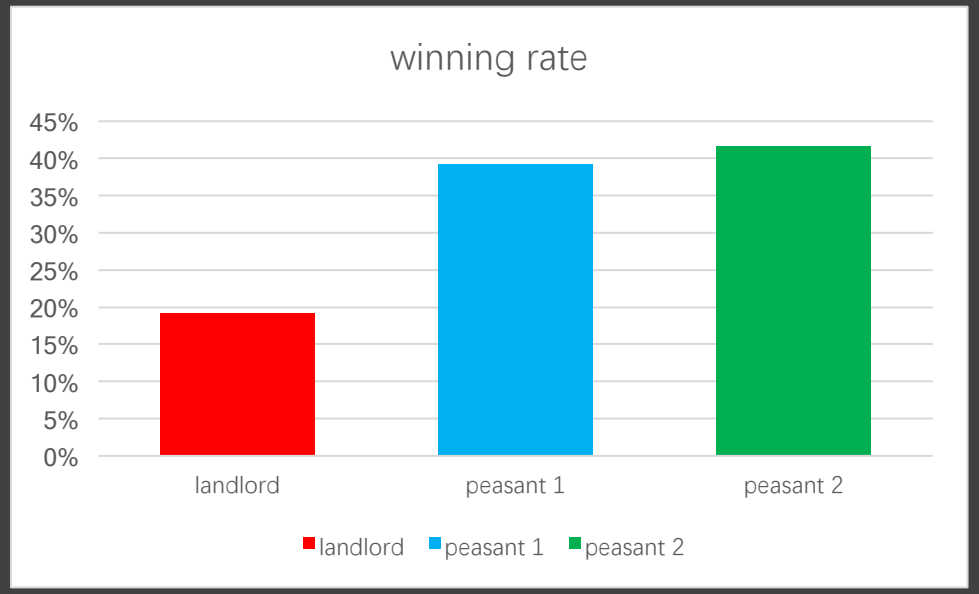
Experiments and Observations

- **Experiment 4:** Zhou rule-based model Versus Random method

Rule-based landlord VS Random peasants

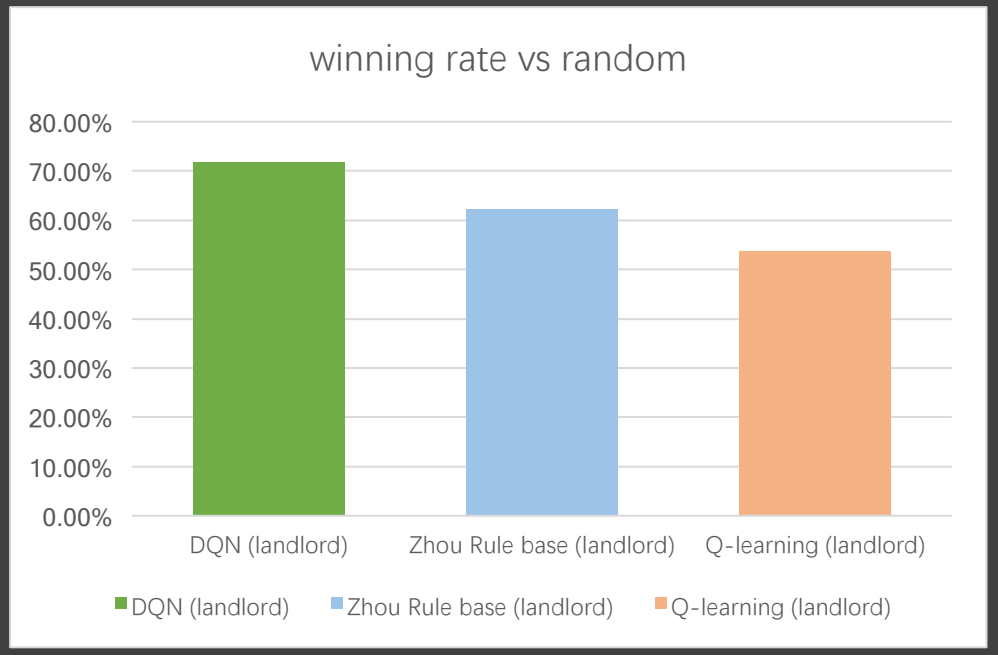


Random landlord VS Rule-based peasants



Experiments and Observations

- Observation
- we compare the results of the DQN model VS the random, the Q-learning model VS the random, and the Zhou rule-based model VS the random based on the previous test result.
- The DQN model has more than a 10% higher winning rate than the other two models.



Conclusion and Future Work

- The DQN model has a 10% higher win rate than the Q-learning model and Zhou rule-based model when playing as the landlord, and a 5% higher win rate than the other models when playing as a peasant.

- In the future work, we will make more research on bidding part of Dou Di Zhu.
- Try more different models based on DQN, like DDQN, Prioritized Replay DQN, and Dueling DQN on the Dou Di Zhu game.

Questions

Thank You!