**NML 502 - Learning How to Optimally Play Blackjack**

Jeremy David, Ken Groszman

3/20/2017

Statement of Problem:

Blackjack, sometimes called twenty-one, is the world’s most popular casino game. It is also commonly cited as the casino game where the player has the best odds of winning, with the most widely accepted “optimal strategy” giving the player 49-51 odds of winning in the long-run. In this project, we will teach an artificial neural network how to play blackjack and see what strategy it suggests to maximize profitability. As an extension, we will see if we can simulate teaching our network how to “count cards”, a strategy used by professional gamblers to increase their probability of winning. We hope that through this project, we will either confirm or debunk the common blackjack “optimal strategy” and gauge the effect of a variety of factors on a player’s profitability in a blackjack game.

*Generating Data*:

In summary, we will generate testing and training data by simulating entire hands of blackjack and selecting what the optimal strategy would be for the player in order to maximize profits.

We will use MATLAB to create a vector of cards corresponding to the number of decks which will then be randomly permuted to simulate shuffling. Cards will be “drawn” from this deck as needed in order to simulate a round of the game. The dealer’s strategy will be simulated, according to common casino rules (hit anything less than 16 or a soft 17). We will then simulate all of the player’s different possible strategies (hit, stand, split, double down) to reveal which

To evaluate the network, we will measure the net “payout” after a certain number of simulated hands. This makes the results of our network consequential and transferrable to a real life situation.

*Data Description Technical Details:*

|  |  |
| --- | --- |
| **Origin of Data** | Generated using blackjack simulation script, as explained above |
| **Description of Data Meaning** | ANN INPUT:  A vector representing the cards that a player can see during a blackjack game (his own two cards, the dealer’s face-up card, and the two cards of all the other players at the table)  ANN DESIRED OUTPUT:  A one-in-n output vector representing the player’s optimal strategy for that simulated hand. An optimal strategy is defined as the least risky logical strategy that beats the dealer. |
| **# of Data Points** | We can generate an arbitrary number of training and testing data points. Given the large number of permutations of possible draws from a deck of cards, we hope to start with 1,000,000 training points and increase it as needed. |
| **Input Features and Encoding** | The number of input dimensions will vary based on the situation:  3 FIXED INPUTS:   * Two features for the player’s two cards * One feature for the dealer’s face-up card   VARIABLE INPUTS:   * A variable number of features for the other cards visible to the player, based on the number of players at the table and cards played in past hands (for “counting cards”   Encoding:  Each card {A, 2, ..., 10, J, Q, K} encoded as a corresponding integer {1, …, 13} |
| **Output Classes and Encoding** | Optimal strategies for a particular hand:   * Stand (hit zero times) * Hit once * Hit twice * Hit three times   In more advanced versions, we hope to include:   * Split * Double Down   Encoding:  Each of these strategies will be encoded as a corresponding unit vector |
| **Output Class Distribution** | Based on 10,000 hands simulated for one player and four decks:   * Stand: 65.4% * Hit once: 25.8% * Hit twice: 7.2% * Hit three times: 1.5% |

**Objectives:**

Our objective is to create a neural network that maximizes profits playing blackjack. What we would like to get out of the network is as follows (in order of importance)

1. Given the dealer’s and player’s cards, make optimal hit/stay decisions
2. When given the dealer’s card, our own cards, and several previous cards, to be able to make even better hit/stay decisions (i.e. “Count Cards”) (an extension of the neural net trained in objective 1)
3. When given the dealers card, our own cards, and several pervious cards, be able to make optimal hit/stay decisions, as well as split/double down decisions. (this would include adding a second neural net that is trained to make split/double down decisions)

Ultimately we would like to be able to come close to recreating, or surpassing, the current “optimal strategy”, and be able to make a profit after simulating many games.

**Technical Approach:**

Our approach will be a two-layer network using back propagation as our learning rule. The “Desired output” will come from simulating a blackjack deck and determining what the optimal hit/stay decision would be. Output layer dimensionality will be determined by the number of possible strategies (stand, hit once, hit twice, etc.) and will be encoded as unit vectors.

In order to add the “counting cards” feature (objective #2), we would take one of two approaches. We would either include the “card count” as an input PE based on which cards have been removed from the deck, or the N previous cards as N more input PEs.

In order to determine success, we will see what percentage of games the network wins, and its net payout. After some number of learning steps, a recall will be done, simulating multiple blackjack games using the decision that is outputted from the net. Profit will then be calculated based on consistent bets of $100.

Learning parameters such as the learning rate, momentum and hidden PEs will be tweaked and optimized.

I think we say all of this aboveThis process will have two separate steps. The first step will be generating the training data by simulating many games of blackjack and the second will be running this data through the neural net.

As a second phase of the project, we will add training a second neural net to make split/double down decisions. This neural net will be used first, and then after the betting decision is made, the hit/stay neural net will be used to play the hand.

