**Poker Rule Induction**

**Introduction to the project:**

We need to predict the best hand you can play based on the cards you’ve been dealt without hand coding rules. The order of cards is also important, which means there are 480 possible Royal Flush hands instead of just four. Identify those, and the other 311,875,200 possible hands correctly. Pretend you never played the game before, are given a history of thousands of games, and are asked to come up with the rules. It is potentially difficult to discover rules that can correctly classify poker hands. Algorithm will need to find rules that are general enough to be broadly useful, without being so broad that they end up being occasionally wrong.

**Data**

Each record in this dataset consists of five playing cards and an attribute representing the poker hand. We are provided with 25,010 poker hands in train.csv and 1,000,000 in test.csv. Each hand consists of five cards with a given suit and rank, drawn from a standard deck of 52. Suits and ranks are represented as ordinal categories:

Each row in the training set has the accompanying class label for the poker hand it comprises. The hands are omitted from the test set and must be predicted by participants. Hands are classified into the following ordinal categories:

**Attribute Information:**

S1 “Suit of card #1” Ordinal (1-4) representing {Hearts, Spades, Diamonds, Clubs}

C1 “Rank of card #1” Numerical (1-13) representing (Ace, 2, 3, … , Queen, King)

S2 “Suit of card #2” Ordinal (1-4) representing {Hearts, Spades, Diamonds, Clubs}

C2 “Rank of card #2” Numerical (1-13) representing (Ace, 2, 3, … , Queen, King)

S3 “Suit of card #3” Ordinal (1-4) representing {Hearts, Spades, Diamonds, Clubs}

C3 “Rank of card #3” Numerical (1-13) representing (Ace, 2, 3, … , Queen, King)

S4 “Suit of card #4” Ordinal (1-4) representing {Hearts, Spades, Diamonds, Clubs}

C4 “Rank of card #4” Numerical (1-13) representing (Ace, 2, 3, … , Queen, King)

S5 “Suit of card #5” Ordinal (1-4) representing {Hearts, Spades, Diamonds, Clubs}

C5 “Rank of card 5” Numerical (1-13) representing (Ace, 2, 3, … , Queen, King)

CLASS “Poker Hand” Ordinal (0-9)

Class Information:

0: Nothing in hand; not a recognized poker hand

1: One pair; one pair of equal ranks within five cards

2: Two pairs; two pairs of equal ranks within five cards

3: Three of a kind; three equal ranks within five cards

4: Straight; five cards, sequentially ranked with no gaps

5: Flush; five cards with the same suit

6: Full house; pair + different rank three of a kind

7: Four of a kind; four equal ranks within five cards

8: Straight flush; straight + flush

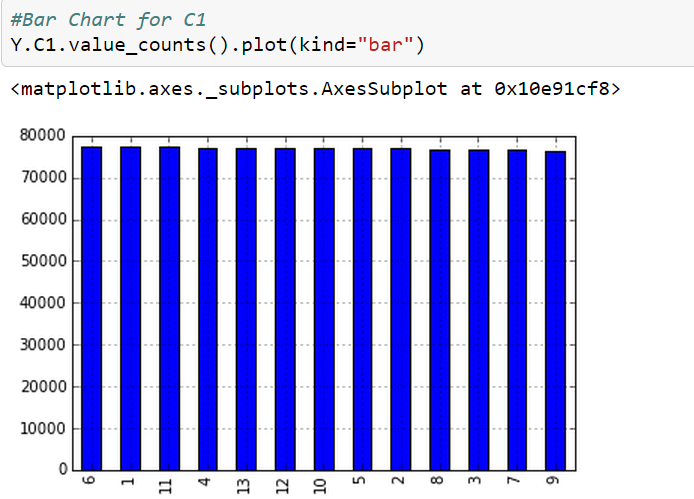
9: Royal flush; {Ace, King, Queen, Jack, Ten} + flush

**Poker Explanation:**

Poker game is purely based on our luck and the type of cards we get. Sometimes bluffing also helps but it is not recommendable. Having good cards in hand will help us raise the amount that we desire and try to win a game. It mainly depends on how confident we are. Even if we feel that we do not have decent cards, we can pass a round for free by just saying "check". We can drop off a game if no decent cards are available with a minimum blind. If someone bets and you are confident to play, you can call the same amount that was raised before.

**EDA:**

I have built the model using gradiant booster analysis and checked missing values for all the attributed individually and found none. I have plot the bar graphs for each and every attribute and a sample bar chart is shown below:



**Machine Learning:**

The machine learning model is built using Gradiant booster method. Using the cross validation method, I have achieved the accuracy of 76% which is quite good. I have performed the analysis on post data analysis as well and shown the bar graph for the hand attribute.

**Post Data Analysis:**

I have imported the result.csv file and checked for the null values and found none. I have shown the results in the EDA file.

**Advantages:**

77% of model accuracy has been achieved and it is quite good.

**Disadvantages of the model:**

Since my model is only 77% accurate, there is 20% more chance that my model is incorrect. I tried my best to improve the accuracy and finally got 77%. Hope this will be considered.