ML Project Proposal: Sample

1 Introduction

Nowadays Poker has become one of the most popular card games worldwide. Since we are also fans of this game we decided to implement a learning system for Poker hand recognition. We found a suitable dataset online, which already contains a dataset of more than 1 million predefined Poker hands and a corresponding test set. We decided to use the whole training set, because the number of possible combinations of cards that one player can have in his hand is 52!/47! = 311, 875, 200 and so our training set is still only 0.3% of it.

2 Problem

Given a set of five poker cards (each card consists of two values: one ranging from 1-4 corresponding to the suit, the other ranging from 1-13 corresponding to the rank) our system should predict the class in which the five cards fall into. There are ten classes and each corresponds to a possible Poker hand: • 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

3 Input

The input is a vector of ten values corresponding to five Poker cards.

4 Output

The output is the Poker Hand class to which the combination of the five cards Fits.

5 ML Technique

We decided to implement the Naive Bayes classifier, because we think it is the most efficient algorithm in this domain. We load the training set from an .arff file and train our system on it. Our program consists of a GUI, where you can select a test set in .arff format, whose instances are then classified to the appropriate classes. After the classification phase the performance of the algorithm is evaluated according to different measures and displayed in a text area (e.g. it contains MAE and other information). Then it can be stored to file. If we have enough time we also add a visualization in the form of diagrams.

6 Dataset

The detailed description and the dataset itself can be found under the following
URL: http://archive.ics.uci.edu/ml/datasets/Poker+Hand.

Document source: http://www.inf.unibz.it/~zini/ML/slides/project_proposal_example_2.pdf