CS 419
COURSE PROJECT
Poker Hand Classification

- Shubham Agrawal (180040100)
- Bhavini Jeloka (18D100007)
- Mitalee Oza (180100067)
- Latika Patel (180100062)

Flow of the presentation

- Objectives
- Data Set
- Classification
  - Models used
  - Results
  - Comparison Across different models
- Winning Probability Prediction
  - Models used
  - Results
- Challenges Faced

Objectives

- Using machine learning to classify into poker hands given a set of 5 cards
- Generation of probabilities corresponding to training and test data using deterministic approaches (generation of data set)
- Training a regression model on the training data for winning probabilities

Data Set

- Every card is represented as (suit, rank)
- First 10 features of training data represent the 5 selected cards
- Last column of training data represents ranking of the poker hand

Classification Problem

The following models were used for the classification problem:

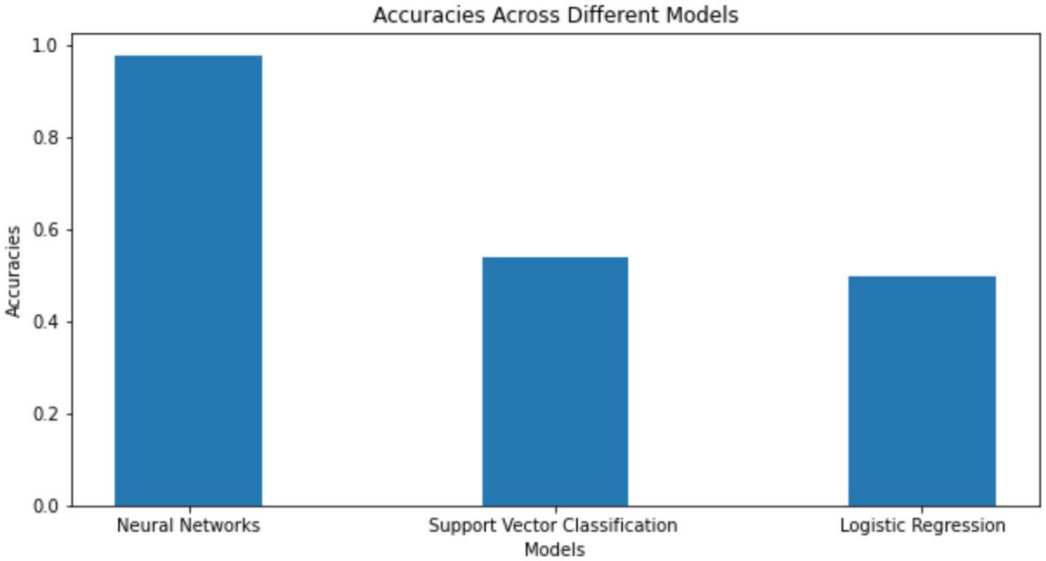
- Neural Networks
- Logistic Regression
- Support Vector Classification

F1 scores have been used to evaluate model performance

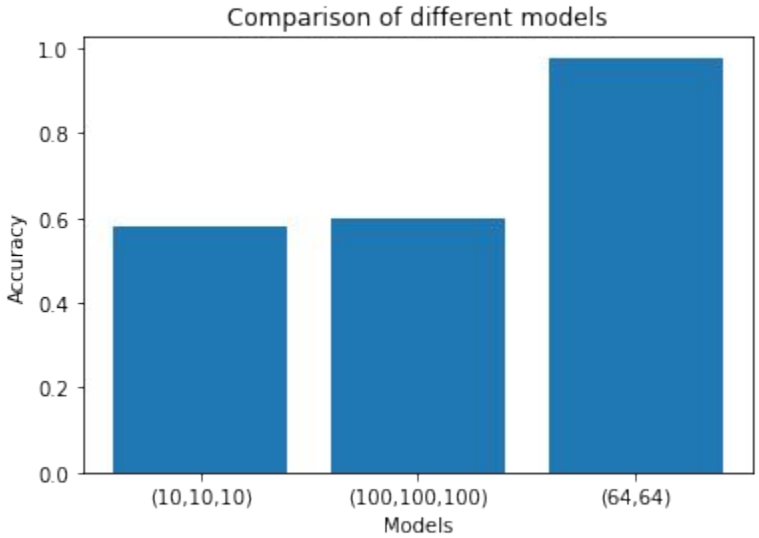
Classification Results

Model	Accuracy
Support Vector Classification Regularization parameter - 1.45, kernel - rbf	0.5409
Logistic Regression Regularization Parameter - 0.032	0.4972
Neural Networks tanh, L2 penalty - 0.0001, Hidden layers and neurons - (64,64)	0.9775

Classification Results



Comparison across different model architectures



Comparison of Neural Networks with different number of hidden layers and neurons

Winning Probability Prediction

The following Regression models have been used to predict winning probabilities:

- Linear Regression
- Ridge Regression

Accuracy is measured evaluating the minimum value of the cost functions

Winning Probability Prediction Results

Model	Error
Linear Regression	0.07
Ridge Regression	0.07

Challenges faced

- Unavailability of data for winning probabilities corresponding to a poker hand
- High computation requirement for generation of winning probability data deterministically
- Unbalanced nature of dataset, difficult to achieve high classification accuracy