Poker Hands

Jeswanth Yadagani(jy3012),
Kavita Anant(ka2744),
Prutha Parmar(prp2126)

About Data

Dataset Used: Poker Hand Data Set - UCI Machine Learning Repository

- Each record is a hand consisting of five playing cards drawn from a standard deck of 52.
- Two attributes: suit and class, describe each card.
- Rank attribute describes the “Poker Hand”.
- 25,010 train samples.

Score for Poker Hand

```
> head(train_data)
  S1 C1 S2 C2 S3 C3 S4 C4 S5 C5 Class
1  1 10  1 11  1 13  1 12  1  1     9
2  2 11  2 13  2 10  2 12  2  1     9
3  3 12  3 11  3 13  3 10  3  1     9
4  4 10  4 11  4  1  4 13  4 12     9
5  4  1  4 13  4 12  4 11  4 10     9
6  1  2  1  4  1  5  1  3  1  6     8
```

```
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
```

See-5 vs RAGA

See-5 Algorithm

```
If (R3 = R5)∧(R5 = R4)∧(R5 != R1)∧(R2
!= R4) then (SCORE = 9)
If (R2 != R1)∧(R2 = R5)∧(R2 != R3)∧(R2
= R4) then (SCORE = 9)
```

RAGA Algorithm

```
If (NumEqualValues(class = rank) =
3) then (SCORE = 9)
```

Feature Engineering

Necessary to make model ignore positional component!

- Sort by Class
- Sort by Suite
- Maximum number of cards for the same class
- Maximum number of cards for the same suite

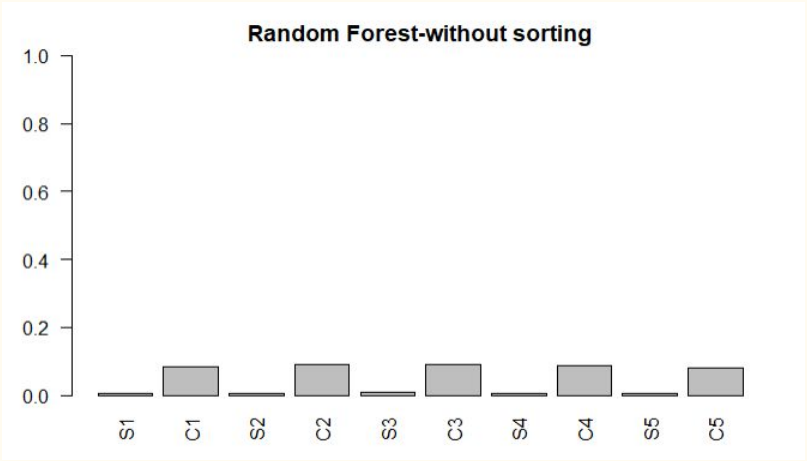
Training-

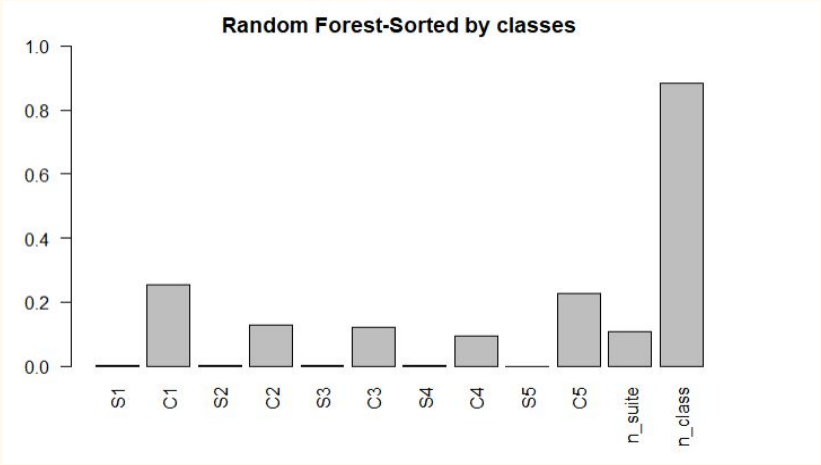
	Card 1		Card 2		Card 3		Card 4		Card 5		
Instance#	S1	C1	S2	C2	S3	C3	S4	C4	S5	C5	Class
1	Heart	10	Heart	K	Heart	Q	Heart	A	Heart	J	Royal Flush(9)
2	Heart	A	Spade	Q	Club	K	Heart	J	Diamond	10	Straight (4)

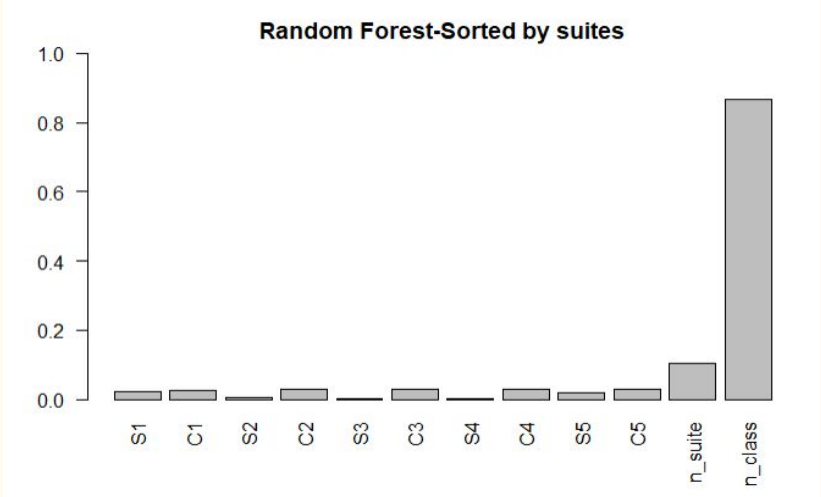
Test-

Heart	A	Heart	Q	Heart	K	Heart	J	Heart	10

Feature Importance







Baseline Accuracies

Model	Train classification Accuracy	Test Classification Accuracy
Multi Class Linear Regression	42.38	42.83
Decision Trees	42.38	42.83
Random Forest	58.20	58.39
SVM	60.22	55.67

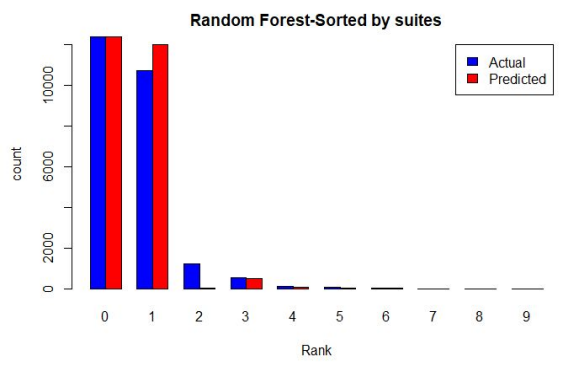
Improved Accuracies(class sorted)

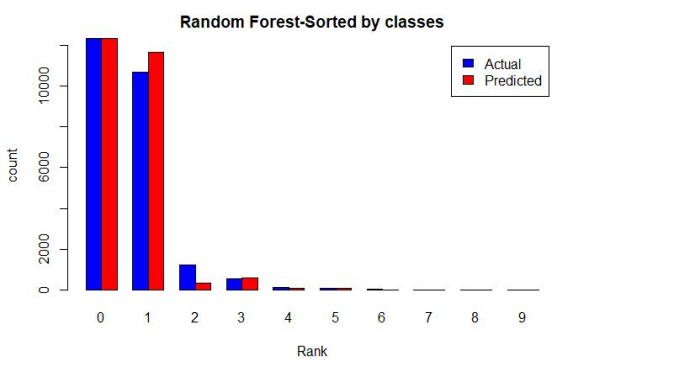
Model	Train classification accuracy	Test Classification accuracy
Multi Class Linear Regression	92.44	92.36
Tree Regression	94.45	94.38
Random Forest	**95.56**	**95.41**
SVM	94.50	94.42

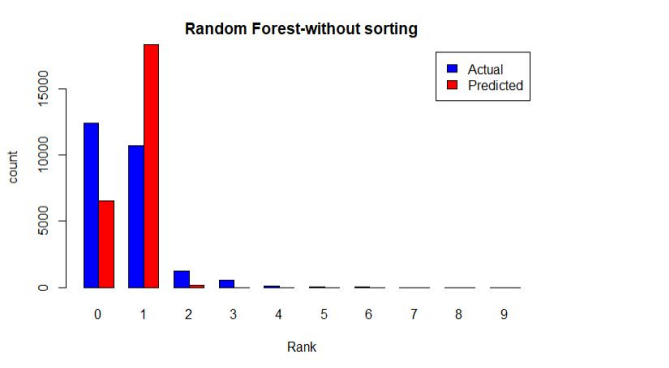
Improved Accuracies(suite sorted)

Model	Train classification accuracy	Test Classification accuracy
Multi Class Linear Regression	92.33	92.26
Tree Regression	94.38	94.36
Random Forest	94.17	94.19
SVM	**94.5**	**94.42**

Random Forest actual vs predicted ranks







Comparison of Results

	Original Paper results		Our project results	
	Evolutionary search(See-5)	RAGA	Baseline	After feature engineering
Training Accuracy	64.25%	90.39%	60.22%	95.56%
Test Accuracy	36.16%	57.6%	55.67%	95.41%

References

- http://www.wseas.us/e-library/conferences/crete2002/papers/444-494.pdf

THANK YOU!