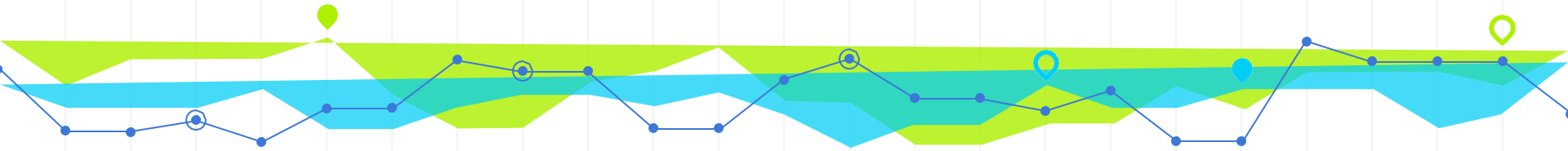
Accuracy of ELO and Betting Odds in Tennis



Tiffany Moi, Manav Sharma, Annie Dai, and Dara Doft

Wharton Moneyball Academy 2017, Sports Analytics Student Research Journal

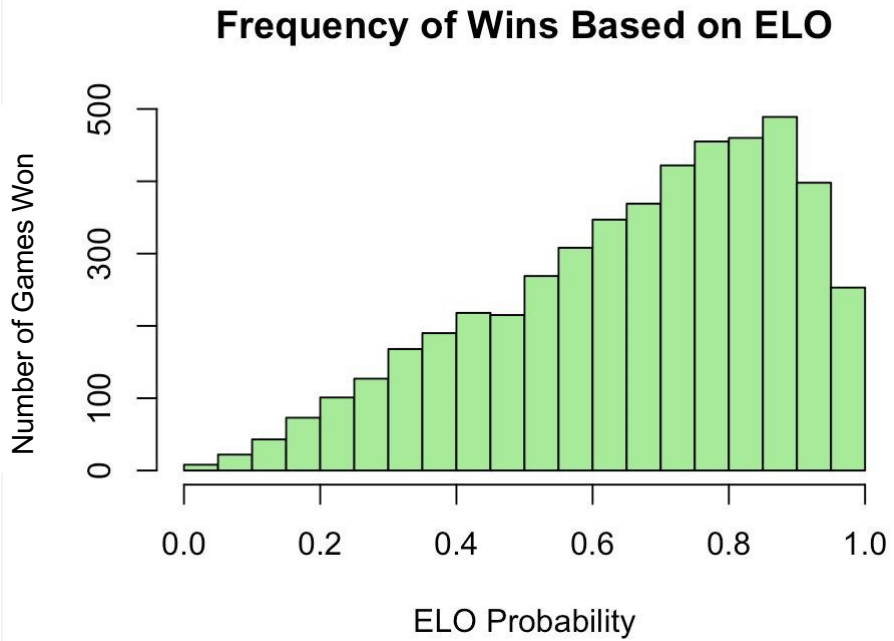
Problem

To see whether or not ELO scores or Betting Odds are accurate in determining the outcomes of matches

Methodology

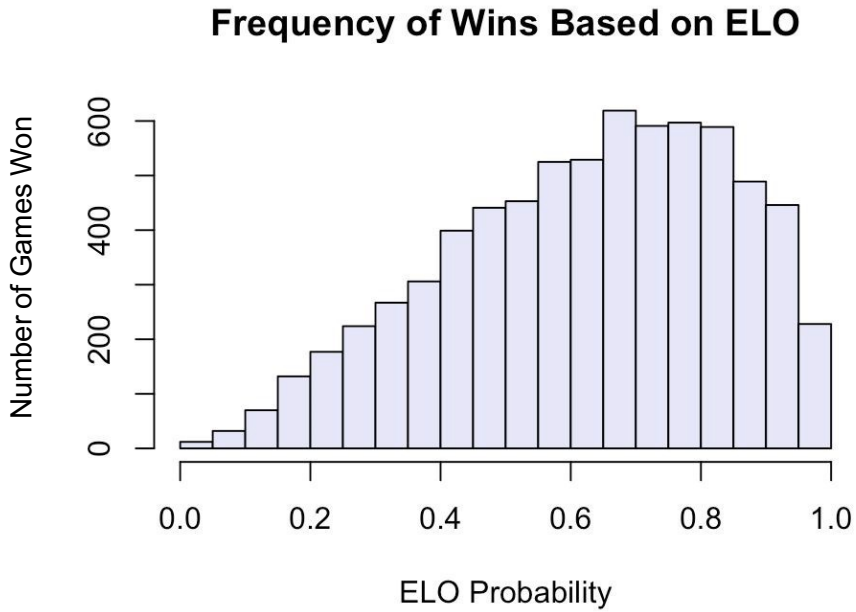
- Collected Data from all Grand Slams dating from 2007 to 2016
- Compared match results with implied probability predictions
- Calculated mean square error and misclassification rate
- Created a Histogram to see frequency
- Created a scatterplot to compare Betting Odd probability to ELO score probability

Results: Men's ELO



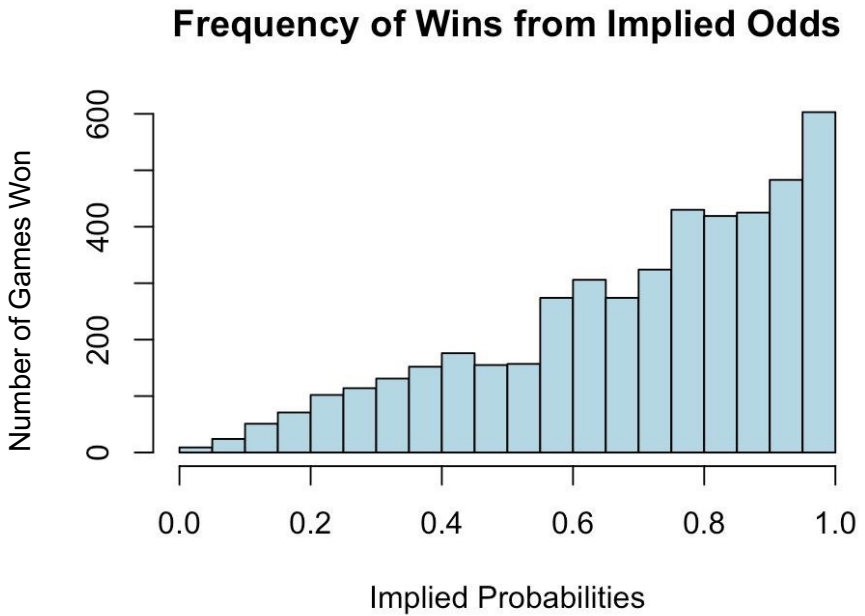
- Misclassification: 0.236
- Mean Squared Error: 0.162
- Analysis: ELO score predictions show that 76.4% of games were predicted correctly

Results: Women's ELO



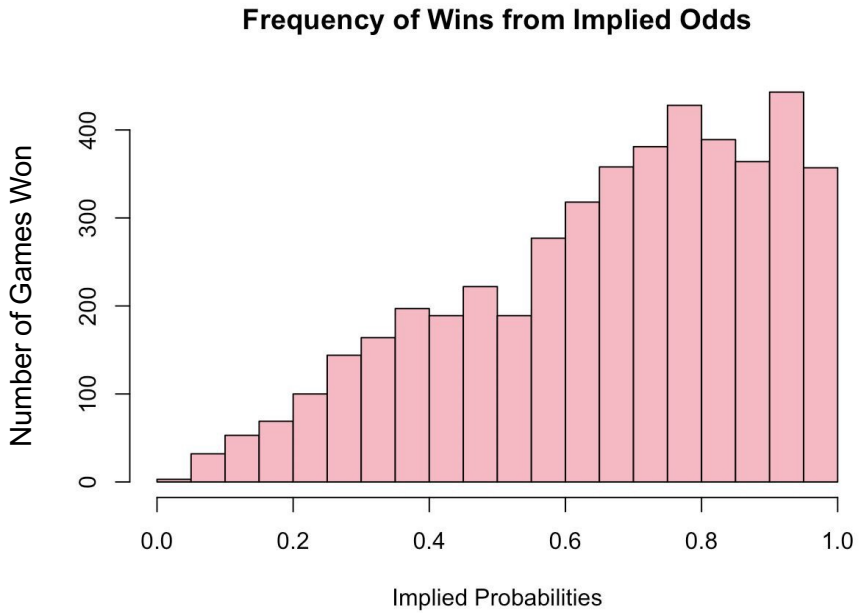
- Misclassification: 0.28
- Mean Squared Error: 0.188
- Analysis: ELO score predictions shows that 71.1% of games were predicted correctly

Results: Men's Betting Odds



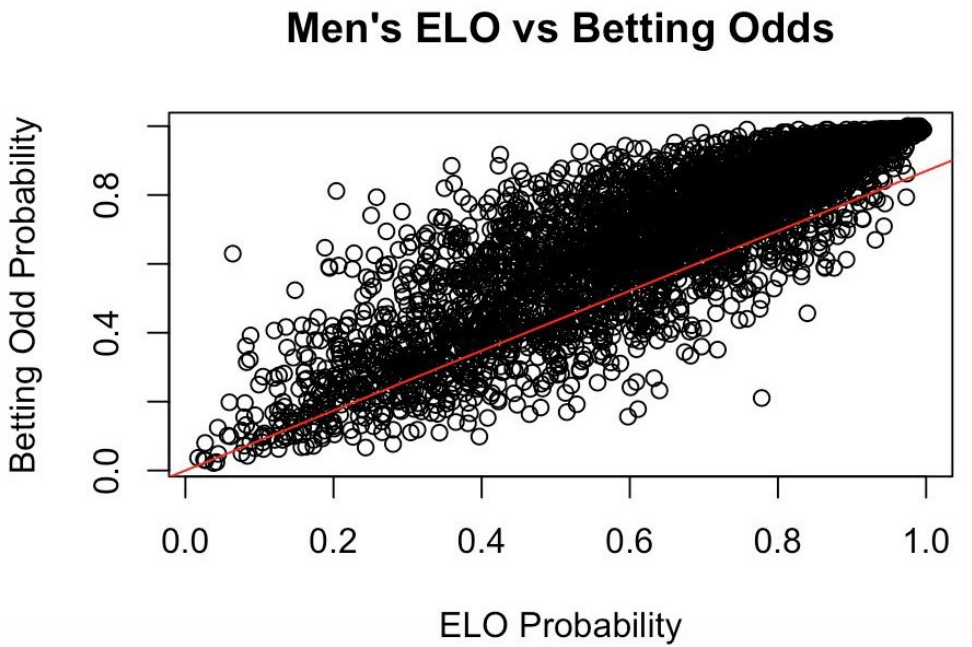
- Misclassification Rate: 0.250
- Mean Squared Error: 0.146
- Analysis: Shows that about 75% of the time the favorite wins going off Bookie odds.

Results: Women's Betting Odds



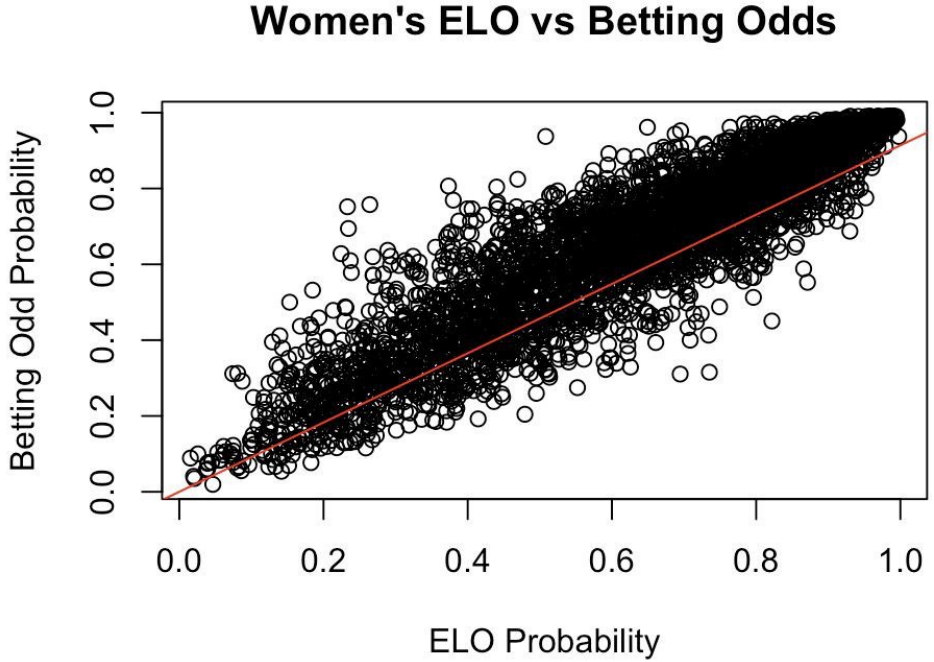
- Misclassification: 0.274
- Mean Square Error: 0.165
- Analysis: The favorite is shown to win about 72.6% based off Bookie odds

Results: Men's Both



- R–Value: 0.8697
- R^2 Value: 0.7564
- Strong Positive Correlation

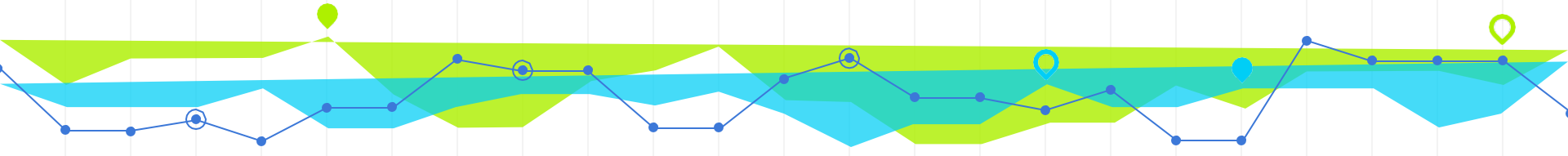
Results: Women's Both



- R–value: 0.9139
- R^2 Value: 0.8353
- Strong Positive Correlation

Conclusion

- Most accurate system is the Men's Betting odds (Accurately predicted 75% matches)
- Comparing both ELO systems the Men's ELO system was accurate with misclassification: 0.236 MSE: 0.162
- When comparing both Betting Odds and ELO for men and women we found strong positive correlation
- But when comparing systems, the Women's ELO vs Betting Odds had a much stronger correlation and higher correlation coefficient than than comparing the Men's Systems against each other.



Faults

- Hard to filter out through the excess amounts of unnecessary data
- Limited betting odds
- Limited variety of tennis stats
- No complete data before 2007