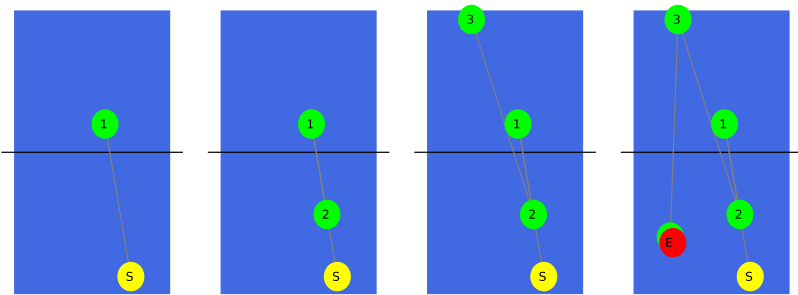


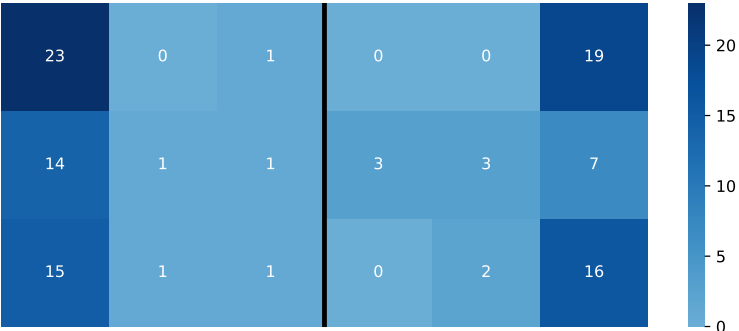
# Supervised Learning for Table Tennis Match Prediction

Sophie Chiang

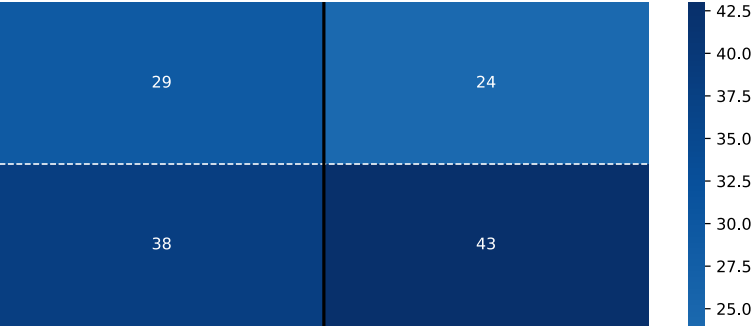
## Features



Match progression can be plotted for each set, recording the location of each ball bounce. The primary source of data are automatic captures from TTNNet.



To reduce the dimensionality of the problem, a rally can be represented as the location of the winning shot. Furthermore, each half of the table can be split into nine equal sections, and the location of winning shots can be grouped e.g., forehand and backhand.

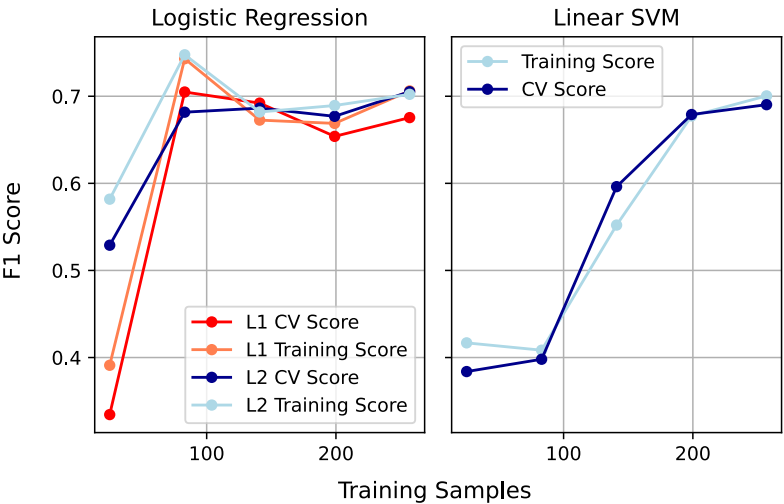


## Abstract

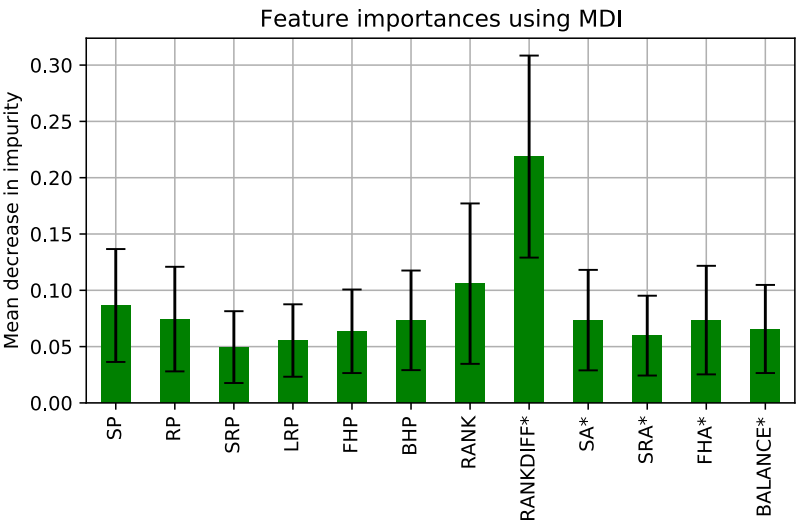
Machine learning has shown to be useful in many different fields of classification and prediction. I apply these techniques specifically to predicting the outcome of a table tennis match, and a similar idea has been applied to different sports for arranging effective training for players to improve performance for future matches. Player and match statistics from data on historical matches were used to predict the outcome, and 12 features were extracted and derived to capture the quality of players and their opponents. The utility of these features are also demonstrated in an ablation study, and the results can serve as a baseline for future table tennis prediction models and can feed back to prediction research in similar ball sports.

## Method

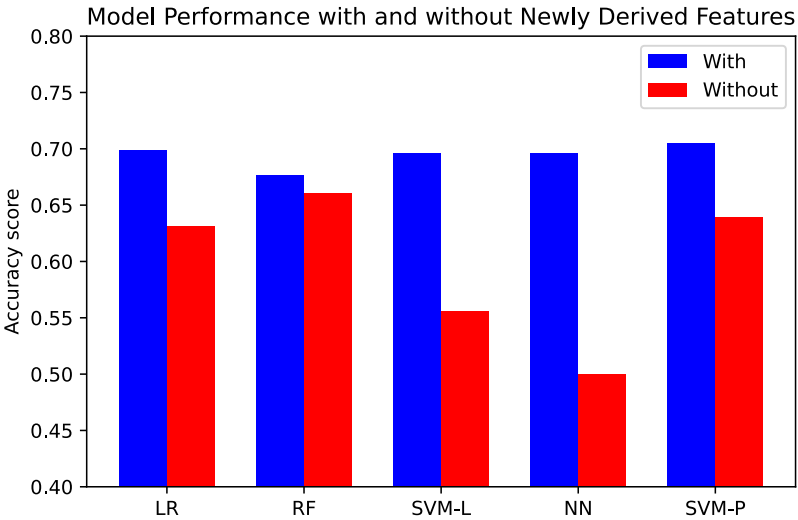
State-of-the-art supervised machine learning models used include logistic regression, random forest, support vector machines and neural networks.



## Results and Conclusions



The importance of features can also be extracted and visualized from a random forest classifier.



To evaluate model performance, I calculate the accuracy of how well each model performs against a test set of matches i.e., the proportion of how many matches are correctly predicted. The set of newly derived additional features can be concluded to increase model accuracy, as validated in the figure shown above.