MET CS 677 - Final Project Discussion

By: Saurav Banerjee

This document discusses the observation as seen by running multiple Regressor Algorithms on the KRK Dataset.

About the Data Set

The King-Rook-King Dataset focuses on the chess end game positions of White King, White Rook and Black King. The dataset contains the optimal depth-of-win for White in 0 to 16 moves, otherwise drawn {draw, zero, one, two, ..., sixteen}.

More details on this dataset is available at - http://archive.ics.uci.edu/ml/datasets/Chess+%28King-Rook+vs.+King%29

Aim of the Project

Out of many possible prediction heuristics possible with this dataset, the goal of this project is to predict the next move of the black king, which will increase winning chances for white in the game.
The dataset provides us with the depth-of-win field. This field has been used to derive the heuristic. The idea here is to assume black's next move by looking at the ("depth-of-win" - 1) positions, and finding the position on ("depth-of-win" - 1), which has minimum difference in area between the current position and ("depth-of-win" – 1)'s position. Below diagram explains the heuristic –

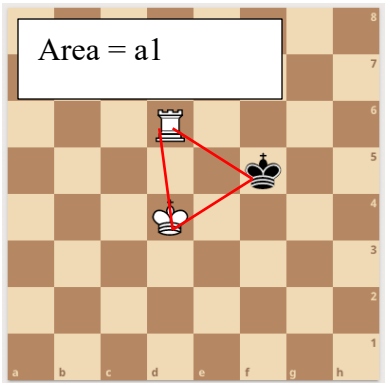
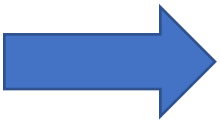


Fig.1. depth-of-win = x



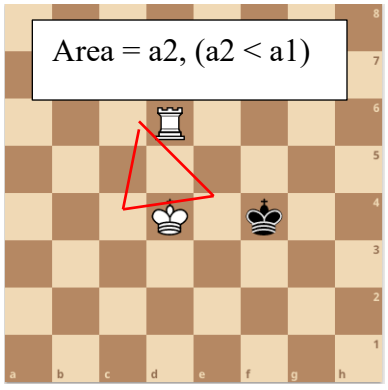
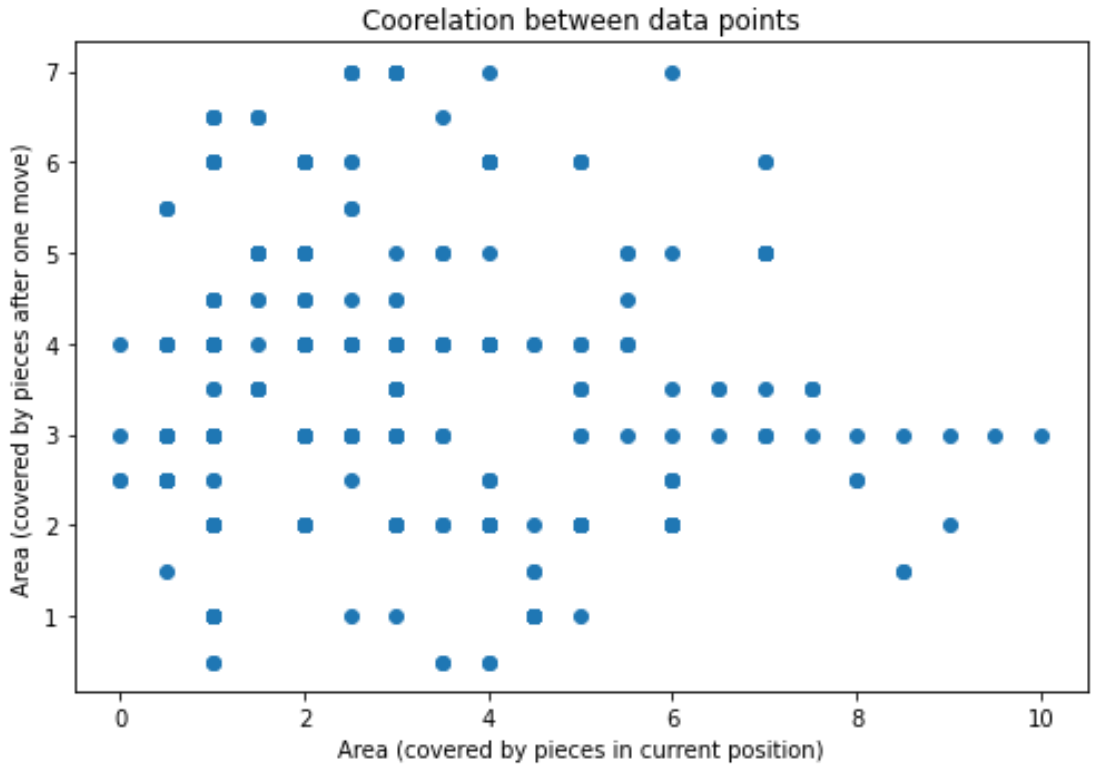


Fig.2. depth-of-win = x -1, with area closest to a1.

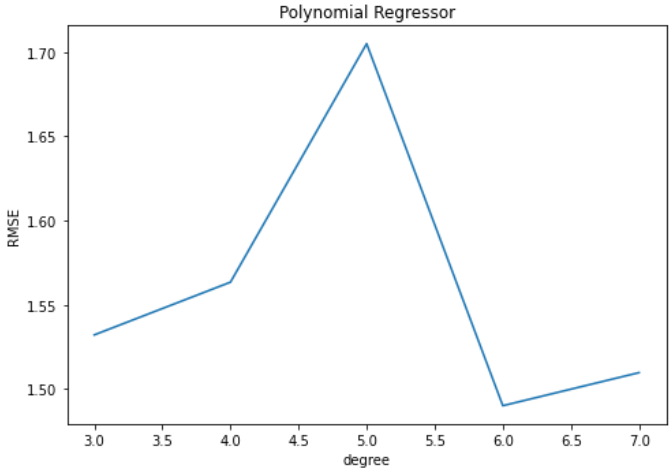
Predictions

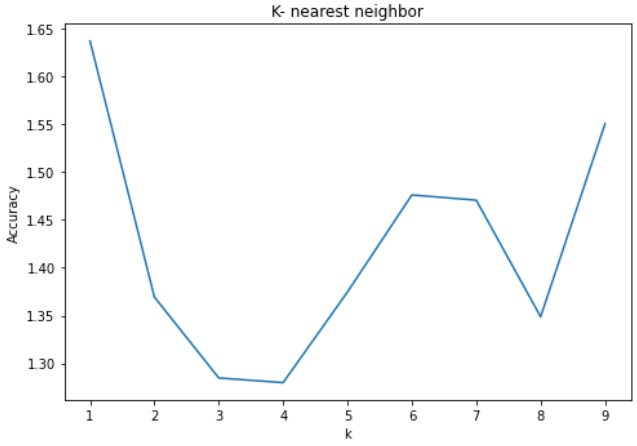
To predict the heuristics, Polynomial, K- nearest neighbor, Decision Tree, Random Forest, SVN regressors were used to evaluate the possibility of such a area heuristic to predict the next best move in a chess game.

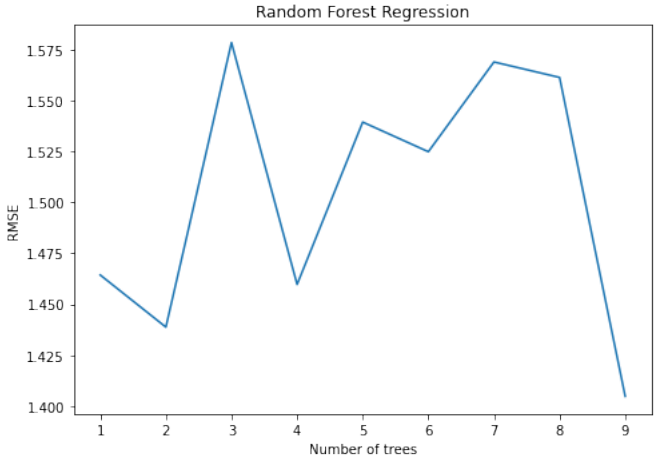
There was no specific correlation found between the area covered by the pieces in the current position, with those of the area as predicted by the heuristic. This ruled out any chances of using a Linear Regression for prediction. (Pearson's coefficient = -0.1096)



Below are the Root Mean Squared Errors plots for the regressors –







All the regressors reported high error rates suggesting, that the heuristic is not a good measure to predict the next best move.

RMSE Summary

Regressor	Avg. RSME
Polynomial	1.560087598
K-Neibhour	1.421306604
Decision Tree	1.51894771
Random Forest	1.504520112
Support Vector	1.587345443

Conclusion

Based on the above prediction analysis, we find the area predictor is not a good prediction model to predict the next best move to improve chances for a white win.