Kyle Sherman

CS-2610

3/28/2022

Research Project: Introduction

There are 3 phases of a standard chess game with white getting the first move: The opening, the middlegame, and the endgame. While the middlegame and endgame do play an important role, the opening sets the state for the flow and development of the rest of the game. A good opening will give a player an advantage when entering the middle game, prepare pieces to launch an attack, and set the defense all at the same time. They play such an important role that many professional players will spend years studying and perfecting different openings to see which ones will impact their game.

Like many others around the world, I am an avid fan of the game and play it often – though I am far from the definition of a good player. Over the years, and hundreds of losses, I have found myself curious of the following questions which will serve as my primary research for this project:

1. Is there an implicit bias in favor of the white piece as a result of getting the first move?
2. Which opening’s have the highest win rate, and which have the lowest win rate?
3. Is there a correlation between the above and the player’s ELO (skill rating)?

I have acquired a dataset from Kaggle consisting of more than 20,000 chess games played on the 2nd largest online chess platform ‘Lichess.com’. The uploader of the data collected it using Lichess.com’ public access API. The observations contain all the necessary data from over 20,000 chess games which should be enough to find any potential correlations relating to my research questions.

The following are the included variables:

* Rated (T/F)
* Start Time
* End Time
* Number of Turns
* Game Status
* Winner
* Time Increment
* White Player Rating
* Black Player Rating
* All Moves in Standard Chess Notation
* Opening Eco (Standardized Code for any given opening)
* Opening Name
* Opening Ply (Number of moves in the opening phase)