

Project Proposal

Introduction

Chess engines such as Stockfish are becoming increasingly more powerful, and are now able to easily outperform even the best humans in much less time. Many players use the evaluations and continuations suggested by these engines to better understand positions and improve their ability. They often achieve this by looking through the various continuations of moves that the engine suggests in order to identify the tactical or strategical advantage that one move has over another.

However, less experienced players will have more difficulty identifying these differences, since they will have encountered the tactical and strategic ideas less frequently in the games they have played. Furthermore, many beginners will not have been explicitly taught about some tactical ideas, and may not be able to understand why a position is favourable for one player even after going deep down a path of suggested moves. Understanding why a certain move was better than another is one of the most important aspects needed to learn from a mistake made in one game, and to make a better move in a similar position in future games.

My project aims to combine approaches to interpretable machine learning models with techniques from more traditional chess engines in order to provide accurate evaluations and move suggestions, as well as giving reasons for them. This would be achieved by training an interpretable model on a dataset of chess positions, with the features being tactical and strategical aspects of the position and the output being the evaluation assigned to the position by a leading chess engine. A chess engine will then be built using the trained model as the evaluation function, and combining the interpretable output of the positions in the search tree to provide reasoning for the evaluation in terms of tactics and strategy.

Work to be Undertaken

1. :

Extensions

1. **GUI to highlight reasoning:**

In order to make it easier for users to visualize the interpretations, I could design a graphical user interface which highlights and explains the tactical and strategic aspects of the position that contribute most to the evaluation.

2. **LLM to provide natural language explanations:**

In order to make the explanations for the evaluation easier to understand

for beginners, I could use a Large Language Model to combine the most important factors with explanations of how they create an advantage for one player to provide a more natural explanation.

Success Criteria

Starting Point

Timetable

Resources