

Predictive Modelling of a Chess Player's Style using Machine Learning

Aniketh Kalagara

Received December 18, 2023

Accepted March 29, 2024

Electronic access April 15, 2024

For many novice and professional chess players, the emergence of 'unbeatable' artificial intelligence (AI) chess models has caused an increase reliance on this technology for memorization, losing the creative and analysis provoking aspect of the game. In this paper, I propose the use of AI in chess to enhance the creative aspect of chess rather than promoting memorization of the best moves. Specifically, I built AI models that assist chess players in categorizing their own personal human playing styles, based on the moves they make during games. The AI model was trained on games of five of the world's most renowned chess legends: Magnus Carlsen, Hikaru Nakamura, Robert James Fischer, Vishwanathan Anand, and Garry Kasparov. A set number of 900 games were taken per player in which a variety of features were extracted in regards to their importance to identifying playing style ranging from the game length to the piece advancement. Logistic regression and neural network models were trained using these 4500 games before finally testing each model to measure the success of the model based on the accuracy of each model on correctly identifying these players based on the features extracted. When tasked with associating these features with one of the 5 famous players, the best model (a neural network) performed 33.5% better than random guessing.

Introduction

Over the last few decades, chess has been hit with an increased influence of technology to assist players in the sport. The world championship is a prime example, months of strenuous memorization of every single possible line provided by computers, the two participants memorize hundreds of games provided. Existing research places emphasis on the use of chess as a tool for growth in the field of artificial intelligence as the complexity of the game provides researchers an opportunity to create models to "solve" the game. In this computer-assisted environment, many assume that Artificial Intelligence (AI) models will soon completely take over the sport. However, it must be noted that there is no set way to play a chess game. After the first move is played, 400 possible board setups exist. After the second move, there are 197,472 possible setups. After the third: 121 million. AI can cut down those possibilities to a few 100 of the "best" moves (in terms of equally high winning probably). Although at the moment chess is too complicated for AI to solve, the avenues of research involve enabling AI to generate results for specific positions, limiting the thought-provoking aspect which in itself brought the game to popularity.

*Chess AI: Competing Paradigms for Machine Intelligence*¹ is an example of said research. Within the paper itself, the study takes a look at comparison of two chess engines, Stockfish and Leela to solve an extremely well-known endgame problem within the chess community, Plaskett's Puzzle. The study involves understanding algorithmic differences to decide the

best chess engine. In the discussion section of the study, the authors mention, "Stockfish and LCZero represent two competing paradigms in the race to build the best chess engine." This represents the influence of interpretive chess AI to enable chess endgames to be solved easier, limiting the creativity of individuals as they try to weave their games towards reaching said positions. Although this research drives progress towards making chess an easier game to play, this will cause limitations, especially within younger players as their reliance on these artificial intelligence systems and engines like Stockfish and Leela to produce gameplay like a machine, due to their ease in solving complicated chess puzzles like Plaskett's Puzzle, which inhibits growth within players.

*Aligning Superhuman AI with Human Behavior: Chess as a Model System*² is another such example of AI being aligned with humans through chess. The purpose of this study was to enable connection between humans and AI to predict human moves at a higher accuracy than existing chess engines such that they can be used as a training partner. In this study, although there is a use of chess AI, the use of chess is not interpretive, rather it is using chess AI for the playing aspect of the game. The authors develop "Maia", a version of the strongest chess engine in the world, Alpha Zero, to predict human moves which will help chess players in identifying their mistakes based on Maia's prediction of the possibility of a mistake on each move. The abstract says, "our results suggest that there is substantial promise in designing artificial intelligence systems with human collaboration in mind by first accurately modeling granular human decision-making."

This research defines the use of chess AI I sought for, the use of chess AI for the purpose of playing, rather than interpreting moves and providing an easy solution out, rather enabling AI to assist players in the process of developing their ability in the game itself.

Taking inspiration to research like the above, I want to place my own spin on using playing chess AI to enable the growth of chess players. Specific to my personal experience and through research of existing papers and studies, I identified a field which was not explored but is extremely fundamental in a chess player's growth, the playing style of a chess player. The influence of a playing style on the moves and ideas one creates in a game is what defines why chess is so competitive. Ranging from an extremely aggressive playing style involving sacrifices and dubious decisions to throw the opponent off, versus a calm positional playing style which involves spreading a game over a long period of time, grinding out the opponent until they tire out and make a mistake. The variety in playing styles are endless and each individual differs ever so slightly. To enable new chess players to connect to a set playing style can enable immense and quicker growth as they are able to study games, ideas and concepts based on players of a similar playing style, such that one develops their own in the process. In this project, I propose this new use of chess AI: to help players identify their own playing styles. Then, when they're faced with the decision of their next move, they'll know which move is most aligned with their personal style of play.

The availability of top-level chess games is vast, with thousands of games available in PGN (Portable Game Notation) formats throughout chess databases. Chess players of the recent past have different playing styles, yet, they have achieved greatness in the sport. An example can be of the following players: Magnus Carlsen, Robert James Fischer, Garry Kasparov, Vishwanathan Anand, and Hikaru Nakamura. Within this elite list contains four world champions and a renowned world number 2. The success these players have obtained are in different eras. Hence, the development of their playing styles will obviously differ, showcased by the vast amount of analysis which top level analysts and chess authors have highlighted in many books of these five greats. To put it simply, these players provide an insight to a difference in playing styles. Magnus Carlsen was an endgame mastermind, his ability to make even the most boring and equal of positions into a win is what makes him arguably the best player to ever play over the chess board, a resilience and positional master. Robert James Fischer was dynamic, not only in his personality but as a chess player, his understanding of the game was beyond advanced, dismantling what was the Soviet machine during his time, creating games on the basis of a fast-paced and creative playing style, creating unique ideas in often dull positions. Garry Kasparov in comparison was an all-out aggressor, creating powerful combinations and attacking sequences in his games without showing mercy was the

highlight of the most dominant world champion in the history of chess. In comparison, Vishwanathan Anand, India's greatest ever chess player, is a flexible player, shifting his playing style according to the needs of the position with a pure focus on identifying the best continuation regardless of if he'd have to play defensively or aggressively. Finally, Hikaru Nakamura, is a tactical and calculative genius, a player who thrives on deep but quick calculations in positions to perform complicated sacrifices and ideas, a playing style which has contributed to being arguably the greatest online chess player due to his quick calculative ability.

Using these clear, unique differences in playing styles of these greats as a baseline to explore playing styles can also be used in the effectiveness of AI in identifying the unique qualities these five contain in their approach in a chess game. Through accurate identification of playing styles, I can also assist novice players in their own identification of playing style through connection with these legends of the game. Hence, I defined my research question as: "How effectively can AI assist novice players in identifying their playing styles by matching their moves to those of top grandmasters?"

A successful model will accurately match chess games – based solely on their moves – to the players who played them. Then, this model can be deployed to analyze moves of novice chess players, such that they can identify which of the five "greats" their own playing style matches the most.

A notable study worth highlighting, conducting similar research is Classification of Chess Games and Players by Styles Using Game Data by M. G. P. B. Jayasekara in 2018³. This research heavily inspired me due to the similarity it had to the research I intended to produce. Existing literature apart from this did not have specificities towards connecting games to playing style, rather involved direct prediction of the raw moves of the games. Jayasekara's work inspired me to use feature extraction to map games according to their features to enhance the idea of creating my models according to the playing style, rather than the sole game in itself, which does not highlight a playing style, rather just enables model to produce an output regarding how close a new game is to another game in the dataset. The size of his dataset and model creation were also different, as he mapped the features according to clustering and classification methods. Jayasekara assumed the playing style for the top-level player he studied and concluded that each playing style is differentiable from the other based on the clusters of each playing style created and how each cluster has games which are accurately mapped to the playing style identified, based on classification methods such as Logistic Regression, Naïve Bayes, Random Forest, etc. In comparison, I intend to map the features to the player itself and rather than proving a difference in playing style, due to this research proving said difference based on his conclusions, I intend to map these different playing styles using Logistic Regression and deep learning Neural Network models to enable

the AI to predict the game and map it to a player based on the features extracted, which map it to the idea of having a specific playing style. A strongly performing model will then enable novice chess players to input their own games and based on the features extracted, map them to a playing style of either of these five top level grand-masters.

Methodology

Dataset Used and Pre-Processing

This study's original dataset contains games of twelve extremely strong players of different eras⁴. The dataset contained the player name, color they played as in the game, the opponent and opponent FIDE classical rating, the result for the player, among other game data, but most importantly, the game given in standard chess notation, PGN.

To boost the interpretability of the results for novice chess players and to lower computational barriers, I decided to cut the dataset down from 12 players to five. This decision was chosen by the clear differentiability between the playing styles, as highlighted in the introduction.

Prior to performing any further changes to the dataset and beginning the data manipulation process, all missing or NA values were taken out of the dataset such that during the process of cutting down the number of games, no games without any applicable information would be present, hence, no cause for syntax errors within the code produced.

Since the dataset had a varying amount of games per player, the dataset was cut down such that each player had only 900 games, such that the models can fairly guess the player without the bias of an increased amount of games per player causing errors in results.

Feature Extraction

Due to the massive number of possibilities a chess game has, machine learning models cannot use the raw moves of a game as its features. Otherwise, every game is like a new game, and there are no generalizable patterns to learn from. Hence, I had to extract common features from games, such that we could use them for prediction on future games.

There are eight distinct features that I extracted in which each feature had a specific reason for its presence in the feature extraction, highlighted below.

Game Length

Players with more conservative playing styles might get into longer games or a player who prefers grinding out opponents, such as Magnus Carlsen.

Number of Trades

Trades represent pieces getting taken off the board in succession. When considering the number of trades, equal and unequal trades were considered as trades. Even if a bishop was traded for a rook (a bishop being of lower value than the rook), we considered it as a trade.

This was done because trades represent game simplification, indicating that the larger the trades, the more the player likes playing with less pieces on the board. A player who is much more capable of playing endgames and prefers the board having less pieces is also an indicative of playing style as games completed with less pieces on the board would mean a clear preference, such as Magnus Carlsen.

Queen Lifetime

The queen lifetime as the name says represents how long the queen has stayed on the board. The queen, being the strongest piece on the board, has a strong presence. Hence, some players tend to want to remove it quickly to ensure stretched out and positional games in comparison to players who want to keep it on the board for longer representing an attacking playing style. The use of the queen on the board for a longer period of time would mean an aggressive or attacking playing style as it provides avenue for combinations, such as Hikaru Nakamura.

Number of Central Pawns

A chess board is divided into eight columns or in chess terms, files, represented by letters "a" to "h". The pawns located on the "c", "d", "e" and "f" files are considered to be pawns which can occupy the center. Along with the eight files, there are eight rows or in chess terms, ranks, represented by numbers "1" to "8". The central ranks are considered to be ranks "4" and "5".

The number of central pawns represent how closed or open the game is, highlighting a player's style depending on the positions which arise out of these pawns and their positions, or otherwise known as pawn structure. A player with different kinds of pawn structures in games would represent an adaptable and flexible playing style as there is no preference noticed due to the different number of central pawns in the games seen, such as Vishwanathan Anand.

Piece Advancement

This metric measures how many times each player goes on to the "opponent's side of the board". A chess board as mentioned has 8 ranks. The "white side of the board" is considered to be ranks "1" to "4" and of course black has ranks "5" to "8".

The more advanced one's position is, the more attacking it is considered to be, with less advancements correlating to a more positional style of play. A player like Robert James

Fischer would have some pieces advanced and some pieces not as advanced to enable the movement of some pieces but maintain control and defense towards his own position, such that there is opportunity to create a unique pattern when a position is dull with his piece positions being varied.

Queen Entry

Again, using the queen as a metric is important due to the value which the piece provides in the game. In this scenario, the queen entry is simply defined by the amount of moves each side's queen takes to make its first move.

The faster the queen enters the game, the more aggressive that player intends to play, such as Garry Kasparov.

Castling

Castling is a special move which involves the subsequent movement of the king and rook. Castling can only occur once in the game given that the king or rook involved in the castling has not moved prior. The process of castling is an extremely common practice associated with king safety. There are two types of castling as shown in figure 1, the king moving to the right for white or left for black is called "king-side castling" and the other is known as "queen-side castling".

A player who switches up the castling they do, whether it's not castling at all throughout the game or between the two types, often correlates to a dynamic, creative style of play, such as Robert James Fischer.



Fig. 1 Castling

Piece Moves

This metric simply checks the number of piece moves done by white's and black's knights, bishops, rooks, queens and kings. Again, using the first letter of PGN notation, every instance a specific letter is present, the counter for the specific white or black piece counter is appended.

The number of piece moves will differentiate between player to player depending on the openings and strategies one uses in the game. When the number of piece moves is mapped to each player, the difference in playing style can be seen with the priority placed on certain pieces over other pieces, hence, providing insight to the playing style according to which piece a player prefers.

Cleanup

After extracting the features, I finalized the dataset with the columns correlating to the name of the grandmaster, the metrics and the color. The color they played influences their playing style, because black is considered to be more defensive compared to white as they play second. Also, the color the player of focus played will come with any PGN notation inputted, hence, we decided to take it directly from the dataset and use it as a ninth metric.

To prepare for modeling, I one-hot encoded the categorical variables (the castling metric and color metric). This means that the categorical variables were re-classified in terms of numerical, preferably binary outputs. For example, if the player is white, there can be a white metric in which the value is 1, representing true, and the black metric where black is 0, representing false. This can be achieved for any categorical variable type through "one-hot encoding".

Figure 2 shows one-hot encoding, we use one-hot encoding to make data into these binary values such that it can be used in AI model creation. Once the one-hot encoded data is placed back into the dataset, it is ready to be placed into a model.

Logistic Regression Model

Prior to feature extraction, the model wasn't split and the twelve players still remained. Hence, I decided prior to splitting to the five chosen grandmasters, I wanted to test the model using just two players to see how it performed on a logistic regression model to see if it is ready for genuine model creation.

Firstly, a logistic regression model is my baseline model of choice to see how much better the model can perform compared to random guessing. Using a baseline model helps give an idea on how the main, ideal model would perform in comparison as it ideally should be better and given it is not, we know that there are bugs in the code. Logistic regression is a model used to predict binary outputs of 0 and 1, or true and false. In this scenario, the logistic regression model will predict the occurrence of whether

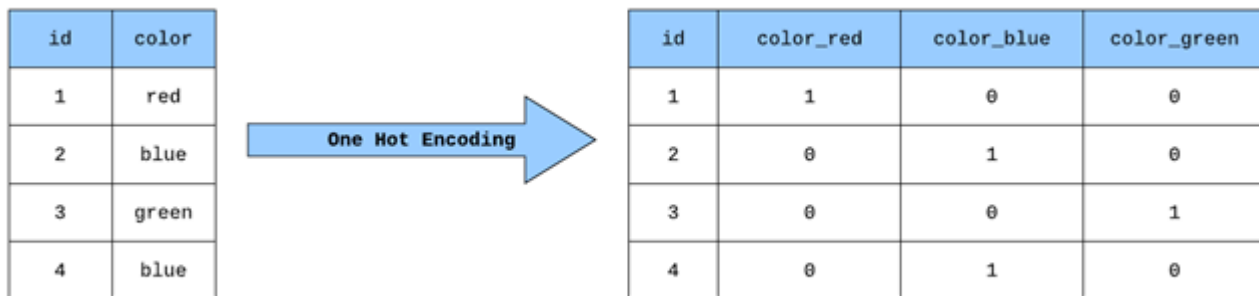


Fig. 2 One-Hot Encoding

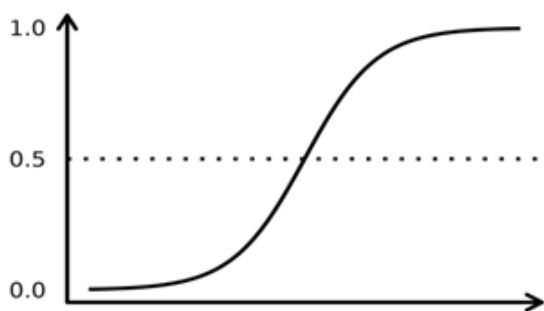


Fig. 3 Logistic Regression Diagram

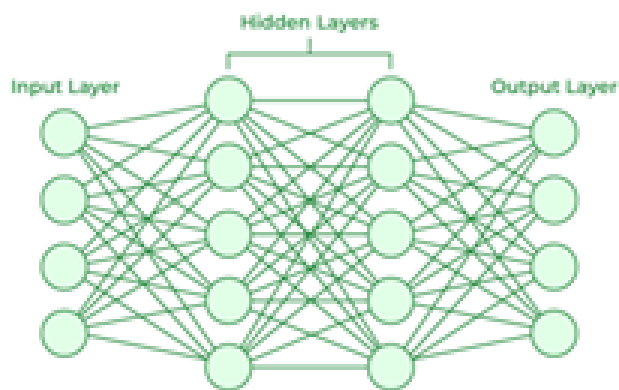


Fig. 4 Neural Network Diagram

the game is of one player or not. The model will predict on a scale from 0 to 1 and whichever the model predicts is closer to the true value, it will predict that as the player.

A diagrammatic representation of logistic regression is displayed with a sigmoid or S-shaped curve structure. As said earlier, the model will base its result on such a scale and accordingly predict between the two players. If the model successfully predicts above random guessing, it can be applied to the chosen five players as well. The training to testing split, which involves the dataset being split randomly into data which the model will be trained by and a test set for the model to predict results on, is a 70 to 30 train to test split. Every time any model is run, there will be a random state factor which means every time, there will be a random split occurring, which means there won't be the same data in a train set and a test set ever.

Prior to creation of any model, due to the dataset being of a random amount of games per each player, it is important that the amount of games per player remains equal and the players chosen are known to be of different styles such that the best accuracy can be obtained and the model can be accurately measured to see if the feature extraction did not have any bugs. Hence, for all models created, the baseline amount of games will be 900. Along with that, for the two-player logistic regression,

we took two players of fairly different playing styles. A sacrificial, extremely dubious and aggressive playing style of the 8th world champion, Mikhail Tal and the grinding, positional and extremely accurate endgame playing style of the 16th world champion and one of our five players as well, Magnus Carlsen.

Once the model has been used on two players, we can move on to providing a baseline model for five players as we now know the model functions accurately. Using the logistic regression model will enable us to check whether the neural network model will perform better or not and see how much better it performs in comparison to random guessing.

Neural Network Model

Here is the actual model which will be used to check for five players. This model will be intended to answer the research question to the best of its ability. Prior to creation of the model as well, 900 games were the baseline to measure the results of the neural network.

A neural network is a method of AI which is inspired from the human brain, using layers connected by neurons to help computers identify and predict results. Neural networks use different methods such as the number of times a model should be trained known as epochs, the amount of layers and the amount of weights or nodes which connect to each neuron, etc. The output of this neural network will be a five-node layer for the five players which the neural network should be tested on. Throughout the process, the neural network will be tuned in the form of the variety of metrics which can be provided by a neural network. The process of tuning will undergo multiple trial and error processes to identify the best performing model with the highest accuracy.

Different model architectures were tested on the basis of training and validation datasets derived from the 70% split. The trial-and-error process involved looping through different model architectures, changing the number of hidden layers and the number of nodes per layer while keeping epochs constant. Model tuning is extremely important as identifying the best model to finally test the dataset on will enable for the best results to be obtained. The model used the same activation functions for the hidden layers and output layer, which are validation functions which decide the number of weights on each node in the neural network, manage the learning rate through each node of the neural network and more to help validate and perform at the highest accuracy. The optimizer used was Adam, the most popular neural network optimizer which helps provide a positive correlation between the accuracy and loss of a model by reducing overfitting and increase computational efficiency by reducing computation time. As mentioned, the validation metric used to measure the success of the model will be the accuracy and the model will be tuned to produce the highest accuracy. The accuracy is the percentage of the number of correct predictions the neural network produces. The validation accuracy will be measured of each model during the tuning process to decide on the best model architecture. The validation accuracy comes from the training of the data from the 70

The model will be tuned by changing the number of hidden layers and nodes per layer as mentioned. There were a variety of combinations created, ranging from 1-10 hidden layers and 10-100 neurons per layer where each model was trialed through creation of training models using the different combinations. These choices were made as if there were any more hidden layers or neurons per layer, the CPU used for producing the results would not be able to produce accurate results either due to over-fitting or lack of processing ability.

Finally, as per a multitude of trials, the final model architecture had 3 hidden layers with 40 neurons per layer due to it having the highest validation accuracy of 27.5

Calculations

Once the best accuracy is obtained from the neural network model, a simple calculation should be made to see how well it performed above random guessing. Often when AI models provide an accuracy, it is seemingly quite low, but this low accuracy is not a true showcase of how well the model has performed.

First, let's consider the two-player model. If the model is randomly guessing every player for every game, there will be a 50% accuracy. For the five-player model, the random guess accuracy would be 20%. Taking the actual accuracy of either model trained, we can identify how well the model is performing above random guessing. For example, the accuracy of the logistic regression model on two players is 55%. To identify how well it has performed above random guessing, we subtract 55 from 50, giving 5% and divide that according to the random guessing accuracy, being 50%. This results in a value of 0.1 or a 10% better performance than random guessing. Although a 55% accuracy seems really bad, the model is actually predicting something, which means it can be applied given further testing. Hence, using this metric as a measure of accuracy enables a better understanding on the performance of an AI model which will be used for all three models and their best performing accuracy which will be shown in the following results table.

Results

The following table presents the results of all three models:

Model Type	Accuracy	Percent Improvement Over Random Guessing
Logistic Regression (2 Players)	61.6%	23.2%
Logistic Regression (5 Players)	24.5%	22.5%
Neural Network (5 Players)	26.7%	33.5%

Table 1 Comparison of Model Performances

The results clearly highlight that, initially, the model did function given two players and post that, had a good improvement from the logistic regression model to the neural network by a 11% percent improvement over random guessing.

While looking through existing studies in regards to chess game to player prediction similar to my study, I identified a study where researchers create a model which predicts a player's game based on only the game notation rather than feature extraction. Detecting Individual Decision-Making Style: Exploring Behavioral Stylometry in Chess⁵ which used game notation to identify the player to their game at a staggering 98% accuracy. This as a benchmark would be a lot higher than my presented study, but an observation made was that the high accuracy was attributed to the first 15 moves of a player's game, given a large dataset, there would be patterns with the openings played by a player

which can then connect them to a grandmaster as seen in the study, versus my identification of games according to features extracted. The research conducted was also more in depth and the models produced were much more complicated, attributing to the high accuracy.

In a complex domain like chess and correlating games to playing styles, measuring a player's playing style according to features extracted will enable a deeper understanding towards the playing style of a player. To have a model perform at 33.5% better than random guessing is powerful as this means that the model is able to produce results noteworthy enough for a player to have their playing style decided upon. For programmers, the accuracy is a validation metric to ensure a model is performing real calculations and predicting rather than guessing. Now that there is an assurance that the model is not randomly guessing, this can be applied to assist novice chess players in identifying their own playing style as the model will always provide an insight towards how close one's games and game features are towards one of these five greats of the game.

Discussion

The research in comparison to existing research enables an improved growth on what already exists. The study of course as mentioned, does not reach as high of an accuracy in comparison to other studies presented. The limitations of the study could be a factor to play, whether it was the CPU strength which limited the different type of model combinations which could be made and the use of only 900 games per each of the five players to perform the calculations. Given a larger dataset to train, the model is bound to perform better with an improved amount of information to create predictions on. Varying the type of optimizer also used such as Stochastic Gradient Descent to measure how it performs on the classification done. Use of different classification models as well could be a factor to play as only two were used to measure the accuracy of the model produced, such as Random Forest Classifier.

With more experimentation and even larger neural network models, along with stronger processing power, the model can be maximized given the features extracted. A larger dataset can help assist in creating an even better trained model producing higher accuracies. Along with that, we can see that the number of features extracted are only nine, with so many factors affecting a player's playing style, ranging from piece sacrifices to openings played, there are factors which can affect a player's style not being fed into the model. As mentioned above, the opening plays a key role in identifying the style of a player as each player has a pattern in the openings played as those result in positions which a player is comfortable in. This will also heavily assist novice players as studying games of a specific grandmaster which the model identifies to be closest to the player's playing style will enable them to learn the best open-

ings, helping them find positions they will also feel comfortable in. Openings are an example of one of many key features which can be extracted given more in-depth feature extraction on the model, such as number of sacrifices, number of checks, and center control of pieces (apart from pawns). Larger experimentation in model architectures, optimizers used and overall changes in classification methods used are also improvements which can be added to enhance the study further.

Conclusion

The application of this research can be applied to help newer chess players identify their playing styles. What I wish to do with this research would be creating an application in which players input their own games and according to the features extracted, help lead themselves towards these top grandmasters, whether it is these five or anyone else. This proposed application can enable novice players to learn and grow in their chess journey through being able to know their playing style. This discovery can enable them to study games of a specific player, understand ideas and openings played by these players along with developing their understanding of the game based on the information provided by the application and the prediction made by the model. An individual although, will be restricted to these five players as the model only predicts according to their playing styles, but as a novice chess player, this basic understanding just provides them a general direction to go in, as when a player becomes more experienced, they start to develop their own playing style, which is the sole purpose of this research. Although, this research is just a stepping stone to what is a vast opportunity for growth in terms of expanding on understanding a chess player's playing style. Improving accuracy, obtaining a larger dataset, experimentation on the models used, larger feature extraction, etc, can help provide the best information to novice chess players. My goal is to enable chess to be a game supported by AI, rather than being interpreted and solved by it, a goal I hope this research begins to reach.

Acknowledgements

Thank you for the guidance of Dashiell Young-Saver from Harvard University in the development of this research paper.

References

- 1 S. Maharaj, N. Polson and A. Turk, *Chess AI: Competing Paradigms for Machine Intelligence*.
- 2 R. McIlroy-Young, S. Sen, J. Kleinberg and A. Anderson, *Aligning Superhuman AI with Human Behavior: Chess as a Model System*.
- 3 M. Jayasekara, *Classification of Chess Games and Players by Styles Using Game Data*.

-
- 4 L. Renyu, *Do Chess Masters Have Different Playing Styles?* Kaggle, <https://www.kaggle.com/code/liury123/do-chess-masters-have-different-playing-styles>.
 - 5 R. McIlroy, R. Wang, S. Sen, J. Kleinberg and A. Anderson, *Detecting Individual Decision-Making Style: Exploring Behavioral Stylometry in Chess*, <https://papers.nips.cc/paper/2021/file/ccf8111910291ba472b385e9c5f59099-Paper.pdf>.