International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V71I10P102

Volume 71 Issue 10, 12-19, October 2023

Original Article

Effects of AI in Scrabble

Ekansh Arora

DPS Rohini, New Delhi, India.

Corresponding Author : aroraekansh2706@gmail.com

Received: 10 August 2023 Revised: 17 September 2023 Accepted: 01 October 2023 Published: 18 October 2023

Abstract *- This research explores the transformative effects of Artificial Intelligence (AI) on the classic board game Scrabble. Integration of AI platforms, particularly Maven and Quackle, into the realm of Scrabble, has brought about significant advancements in gameplay strategy. Scrabble, a word-based board game, presents challenges for AI due to incomplete information and the need for rapid move generation. The study examines the algorithms employed by these AI engines, notably the Directed Acyclic Word Graph (DAWG) and Generalized Abbreviated Directed Acyclic Word Graph (GADDAG), which facilitate efficient word validation and generation. The paper delves into the evolution of AI in board games, tracing its development from rudimentary rule-based systems to the sophisticated algorithms seen in Scrabble AIs today. Maven and Quackle stand out as prominent Scrabble AI players. Both employ innovative approaches to generate and evaluate moves, utilizing techniques such as stochastic look-ahead simulations and heuristic-based ranking. Scrabble AI research's insights are significant for navigating incomplete information games and real-world scenarios with limited data. The utilization of AI algorithms like DAWG and GADDAG in Scrabble AI contributes to algorithmic advancements applicable to diverse scenarios. The study showcases AI's impact on competitive Scrabble play and its potential for shaping tactics in other fields. Scrabble AI engines not only offer valuable learning tools for players but also expand the frontiers of AI applications. In conclusion, the research explains the dynamic synergy between traditional games and AI, exemplified by the intricate interplay between Scrabble and advanced AI engines like Maven and Quackle.*

Keywords *- AI, DAWG, GADDAG, BasicBot, Scrabble.*

1. Introduction

The evolution of Scrabble from its origins as a board game to the application of artificial intelligence (AI) in analyzing popular games like Scrabble has showcased the dynamic interplay between traditional gaming and cutting-edge technology. Scrabble is a popular board game where two players compete by spelling out words on a 15x15-square grid using lettered tiles. From a bag of 100 tiles with various letters, seven tiles are drawn for each player, and players begin by forming words using the tiles drawn. After each subsequent move, each player replenishes the number of tiles used from the bag of remaining tiles. While each player can view their own tiles, they cannot see their opponent's tiles and the bag of remaining tiles. Each of the 100 tiles has a letter and point value associated with it, as shown in the figure below. Words are scored by adding up the assigned point values of their tiles, including additional points from any of the 61 premium squares covered in the word's formation, including squares of Double Letter, Double Word, Triple Letter and Triple Word. Initially referred to as Criss Cross, the game was created in 1931 by architect Alfred M. Butts, drawing inspiration from crossword puzzles and anagrams. In 1948, James Brunot undertook a substantial redesign, rebranded it as Scrabble, and initiated its marketing efforts. The game first appeared in Great Britain in 1954.[1]

In parallel, AI, or Artificial Intelligence, encompasses emulating human intelligence processes through machines, particularly computer systems. AI finds practical use in various applications, such as expert systems, natural language processing, speech recognition, and machine vision. [2] Although it has sprung into existence recently due to powerful computer resources, AI was invented decades ago. Since the beginning of the 1950s, AI has been used extensively to analyze popular games. Games are the perfect domain for testing the power of AI in a limited environment with a fixed set of rules, where problem-solving techniques may be created, examined, and then used in the actual world. [3] Consequently, AI has been prominently applied to board games over the years, such as chess, scrabble, and backgammon, resulting in competition that has accelerated the development of many heuristic-based search techniques. [4]

There are various different kinds of AI players available for board games in recent days. Some are designed to beat human opponents, while others are designed to be used as practice tools. There are also AI algorithms created for

children's games like tic-tac-toe. In the context of board games, AI often refers to a dummy player designed to mimic human behavior. AI players strive to outperform human opponents by performing "intelligent" moves in the same way as a human player would. Chess was one of the first board games to have an AI player. The goal of early chess computers was simply to build a machine capable of defeating a human opponent. Deep Blue, an IBM chess computer, achieved this goal in 1997. [5]

Recently, advanced AI has been applied to study complex games such as Scrabble. [6] Scrabble presents a significant difficulty due to its inherent nature as a game of incomplete information. Players must contend with their lack of knowledge regarding both their opponents' tile racks and the remaining tiles in the pool bag. This lack of information makes it a formidable task to make optimal decisions at each turn, as the future state of the game remains uncertain, making it challenging to predict the exact next move with precision. The second challenge in Scrabble is that, given a player's current rack and the state of the board, AI systems should be able to quickly develop potential plays for that player. To achieve this, a quick move-generation process and effective opponent strategies for each round must be included in the AI engine. [7]

Currently, the leading Scrabble artificial intelligence are Maven and Quackle. Maven was created by Brian Sheppard in 2002, while Quackle was introduced in 2006 as an open-source Scrabble AI by Jason Katz-Brown and John O'Laughlin. [8] Maven is a famous artificial intelligence Scrabble player which uses a three-ply stochastic look-ahead simulation approach to generate moves (a ply is a player's turn). On the other hand, Quackle has heuristics and strategy that are close but not identical to Maven's.

To quickly rank the most promising candidate moves, Quackle makes use of a program module called Kibitzer. Following a three-ply-look-ahead move generation, these moves are then further analyzed using a simulation engine capable of generating 100 to 300 random racks for each probable move.

Artificial Intelligence (AI) has revolutionized numerous domains, and its impact on the world of competitive Scrabble is no exception. The primary objective of this paper is to review the integration of AI-based platforms into the game of Scrabble, focusing specifically on Maven and Quackle, and analyze the algorithms used by the AIs, specifically Directed Acyclic Word Graph (DAWG) and Generalized Abbreviated Directed Acyclic Word Graph (GADDAG) to beat human opponents in Scrabble. Further, this paper also provides evidence on how AI bots for Scrabble perform using publicly available data on the performance of the BasicBot from the woogles.io platform.

1.1. Research Gap & Novelty of the Research

The paper focuses on implementing advanced AI algorithms, such as DAWG and GADDAG, in the context of Scrabble. While some prior research may have touched on AI in board games, this work specifically delves into the use of these specialized algorithms, which is a relatively less explored area in the Scrabble AI literature. Furthermore, this research uniquely compares and contrasts these two prominent Scrabble AI players in Maven and Quackle.

This comparative analysis provides valuable insights into the strengths and weaknesses of different AI approaches within the same gaming domain, which can help both researchers and Scrabble enthusiasts better understand AI's impact on gameplay and help create a better future AI engine. The question this paper is trying to answer is: how has integrating AI platforms like Maven and Quackle impacted gameplay and strategy in the classic board game Scrabble, and what algorithms are employed by these AI engines to achieve superior performance? Additionally, how does the BasicBot perform in Scrabble games, and what insights can be derived from its performance data?

There has been no prior research in the field of the performance of humans vs BasicBot (or AI), which makes this research paper unique.

Understanding AI engines that play Scrabble is extremely important because it gives crucial insights into navigating incomplete information games that can be extended to other comparable settings. For example, the insights learnt in Scrabble AI research may be used for various incomplete games and scenarios, such as Poker, Go, or even real-world decision-making with limited information. Furthermore, sophisticated algorithms used in Scrabble AI, like DAWG (Directed Acyclic Word Graph) and GADDAG (Generalised Abbreviated Directed Acyclic Graph), provide important expertise for developing algorithms in similar areas. Researching Scrabble AI can contribute to an improved understanding of algorithmic techniques and approaches to handle other partial information scenarios.

2. Methodology

2.1. Research Aim

The paper aims to examine how AI-based platforms have been incorporated into the game of Scrabble and to examine algorithms used to outperform human opponents.

2.2. Themes

The following themes are explored in the paper

- Algorithms and Data Structures used by Scrabble AI
- The need for a faster algorithm
- Scrabble AI Engines
- Data Analysis

2.3. Data Collection

Data was collected from Kaggle and included data on 500 Scrabble games from the woogles.io platform played by the BasicBot

2.4. Data Analysis

The following datasets for analysis were used, including the games and scores datasets.

- The games dataset is reported at the game_id level with other variables, including player_id that goes first, a dummy variable for the winner, and other variables for the initial time allocated and game end reason, etc.
- The scores dataset is reported at the game_id, player level where for each game_id and player combination, there are variables for the rating of the player
- and score for the game.

Data analysis was done on Google Colab using Python libraries pandas, numpy, seaborn and matplotlib to plot graphs. Linear regressions were also done for statistical analysis of the data.

2.5. Data Cleaning

- There were no null values for any variables in the games dataset. But in the scores dataset, there were some game_id's that had data on only one player, and such rows were dropped.
- Data Preparation: The code begins by preparing the game data, including creating new columns to indicate whether BasicBot won a game and categorizing opponent ratings into bins.
- Win Percentage by Rating: It calculates the win percentage of BasicBot against opponents in different rating bins, providing insights into how BasicBot's performance varies based on the opponent's rating.
- Histogram Visualization: The code creates two histograms to visualize the distribution of opponents' ratings and BasicBot's win percentage by rating bin. These visualizations help in understanding the data distribution and performance trends.
- The scores data was reshaped into a wide format, including data at the game_id level to incorporate information on the rating and score of BasicBot and the Opponent. This data was subsequently merged with the games dataset based on the game_id variable.
- Regression: It explores the relationship between an opponent's rating and their game score, indicating a modest positive correlation.

3. Discussion

3.1. Algorithms and Data Structures used by Scrabble AI

Scrabble AI relies on a combination of sophisticated algorithms and data structures to simulate human-like gameplay and maximize its performance. The AI's primary objective is strategically forming high-scoring words using the available letters and board state. To achieve this, Scrabble AI employs algorithms for word generation, board evaluation, and optimal move selection while relying on efficient data structures to store and manipulate the game's information. Two common algorithms used in Scrabble AI are DAWG and GADDAG. These specialized data structures enable the AI to efficiently store and access large lexicons, significantly speeding up word validation and generation. [7]

Andrew W. Appel and Guy J. Jacobson developed a rapid algorithm for generating all possible moves in Scrabble by using a Directed Acyclic Word Graph (DAWG) derived from a trie (a trie is a multiway tree data structure used to store a large amount of strings). [10] DAWG is a specialized data structure used for effectively storing and searching a huge collection of words or strings. It is particularly useful in applications where quick word lookup, word validation, and word generation are required, making it a popular choice in Scrabble. DAWG has brought remarkable advancements to Scrabble algorithms and AI, reshaping how word lists are managed and searched. Its development has proven instrumental in enhancing the performance of Scrabble AI systems. One of the key benefits of DAWG is its compact representation of the lexicon, achieved by merging common prefixes and suffixes. This reduction in memory usage allows Scrabble AI to efficiently handle vast word lists, making running on devices with limited resources feasible. [11]

In the context of Scrabble gameplay, the speed of word validation and lookup is crucial. DAWG excels in this area, enabling Scrabble AI to swiftly verify the existence of valid words against the game's dictionary. This speed not only improves the AI's responsiveness during gameplay but also allows it to explore and generate possible moves quickly. The ability to efficiently traverse the DAWG graph significantly contributes to the AI's decision-making process, leading to more strategic and competitive gameplay.

This research stands out and achieves superior results compared to state-of-the-art techniques reported in the literature for several compelling reasons. First and foremost, the paper delves deeply into examining AI algorithms, notably the Directed Acyclic Word Graph (DAWG) and Generalized Abbreviated Directed Acyclic Word Graph (GADDAG). While prior studies may mention AI in Scrabble, this work provides an in-depth exploration of these critical algorithms, elucidating their mechanisms and demonstrating their pivotal roles in Scrabble AI. This level of granularity enhances the understanding of AI's functionality within the game.

Moreover, the research underscores the transferability of insights gleaned from Scrabble AI to other incomplete information games and real-world decision-making scenarios with limited data. This emphasis on practical applicability extends the paper's relevance beyond the confines of

Scrabble, setting it apart from existing literature that often confines itself to specific game analysis.

In essence, this research's unique contributions lie in its meticulous examination of AI algorithms in Scrabble, its comparative analysis of prominent Scrabble AI engines, and its illumination of how AI insights can be leveraged in diverse domains. By providing this comprehensive perspective, the paper achieves more profound insights and relevance compared to existing studies that may not offer such a holistic view of Scrabble AI and its broader implications.

3.2. Why is a faster algorithm than DAWG needed?

The DAWG algorithm is known for its remarkable speed. However, having a faster algorithm would not be particularly valuable if the highest-scoring move always guaranteed success. While a program consistently playing the highest-scoring move can outperform most people, it would struggle against skilled tournament players. Tournament play is characterized by frequently using obscure words (like OE, QI, XU, etc.). Nevertheless, the fact that a program equipped with knowledge of every word and always selecting the highest-scoring move cannot even win half of its games against expert players underscores the importance of strategy in competitive play.

Steven A. Gordon, an American computer scientist and Scrabble enthusiast, presented a faster data structure called GADDAG in 1994 as an improvement over the existing DAWG algorithm for Scrabble AI. GADDAG was designed to handle the unique challenges of competitive Scrabble play, where strategy and word intersections play a crucial role. Gordon's work on GADDAG involved optimizing the data structure and traversal methods, allowing the algorithm to efficiently explore potential word formations in all directions, including word reversals. By incorporating word reversals, GADDAG could more effectively identify high-scoring moves on the board, considering intersecting words in both horizontal and vertical directions. Where all possible moves generated by DAWG are derived from a trie, GADDAG involves encoding bidirectional paths that start from each letter in all words within the lexicon. In contrast to the minimized DAWG for the same extensive American English lexicon, the minimized GADDAG is approximately five times larger in size. Nevertheless, despite its greater size, the GADDAG algorithm generates moves at an average rate of more than twice as fast. This heightened speed greatly improves the possibility of developing a Scrabble-playing program to make intelligent moves within reasonable time constraints. [12]

Steven A Gordon's work on GADDAG significantly advanced the capabilities of Scrabble AI, making it more competitive against expert players in Scrabble and other challenging environments. The algorithm has since become a fundamental component of many Scrabble-playing programs and has contributed to developing strategic and intelligent gameplay in the Scrabble community.

3.3. Scrabble AI Engines

Finding competent human opponents at all times can be challenging; AI engines offer a solution to this problem, providing players with the option to enjoy the game whenever they desire. AI engines provided a valuable learning tool for Scrabble players. Players could practice against AI opponents of varying skill levels to improve their own gameplay. The ability to analyze AI moves and strategies allowed players to understand the game better and develop more effective tactics.

The development of AI-powered opponents for Scrabble can be traced back to the late 1970s and early 1980s. Early efforts were rudimentary, utilizing rule-based systems that employed predefined strategies and heuristics to play the game. These initial AI opponents were not particularly challenging, as they lacked the ability to analyze the game board and make intelligent decisions based on the current state. Brian Sheppard was able to develop the first-ever top-level AI, Maven, by working on it since the 1980s.

3.4. Maven AI

Maven is a well-known AI Scrabble player created by Brian Sheppard. Its name is derived from "maven," signifying an expert or enthusiast. Maven's gameplay can be divided into three stages: the "mid-game" phase, the "pre-endgame" phase, and the "endgame" phase. The "mid-game" phase spans from the start of the game until there are 9 or fewer tiles remaining in the bag. During this phase, the game is most unpredictable, and the AI relies on heuristic evaluations to make moves.
The "pre-endgame" phase operates similarly to the "mid-game" phase but focuses on setting up a favorable endgame situation. The "endgame" phase begins once no tiles are left in the bag. In two-player games, this means that players can deduce the exact tiles on each other's racks based on the initial letter distribution. Maven utilizes the B-star search algorithm to analyze the game tree during the endgame phase. Additionally, Maven's move generator uses a Scrabble lexicon converted into a DAWG for its word choices. [7]

3.5. Quackle AI

Quackle is considered the top-tier software among computer programs for providing estimated win probabilities and expert analysis in Scrabble games. In 2006, Jason Katz-Brown and John O'Laughlin introduced the public release of Quackle, an artificial intelligence tool and analytical resource tailored for Scrabble. It has the capability to compete with the world's best players.

Quackle is incredibly versatile, allowing users to configure it for playing and analyzing games with various board layouts and the latest lexicons. It encompasses a

comprehensive set of components, including a move generator, move evaluator, simulator, and a user interface based on the Qt framework. Furthermore, Quackle can be employed with any board configuration, alphabet, lexicon, or distribution of tiles. [8]

One standout feature of Quackle is its lightning-fast move generator, which efficiently produces a list of all potential plays that can be executed based on a given rack and board. Quackle achieves this impressive speed by utilizing the Scrabble lexicon converted into a GADDAG.

3.6. Macondo AI

In the future, Macondo aspires to become a globally recognized Crossword Game AI of exceptional quality. Currently, it is in its beta stage and is being employed within the widely-used online Scrabble gaming platform known as Woogles. [9]

4. Data Analysis

The Python code that was run offered a comprehensive analysis of gaming data, focusing on the performance of the BasicBot against human opponents. It begins by preparing the data, calculating win percentages by opponent rating bins, and presenting the distributions through histograms. The code then conducts a regression analysis to delve into factors influencing game outcomes. The regression highlights a modest positive correlation between opponent rating and their game score. This analysis provides valuable insights into BasicBot's performance and the factors driving its success in different gaming scenarios.

Within the games dataset, there was a balance of the proportion of times that BasicBot goes first, with BasicBot going first roughly 50% of the time. In terms of the distribution of times each opponent has played Basic Bot, the majority of times a player has played Basicbot is between 0 and 6.

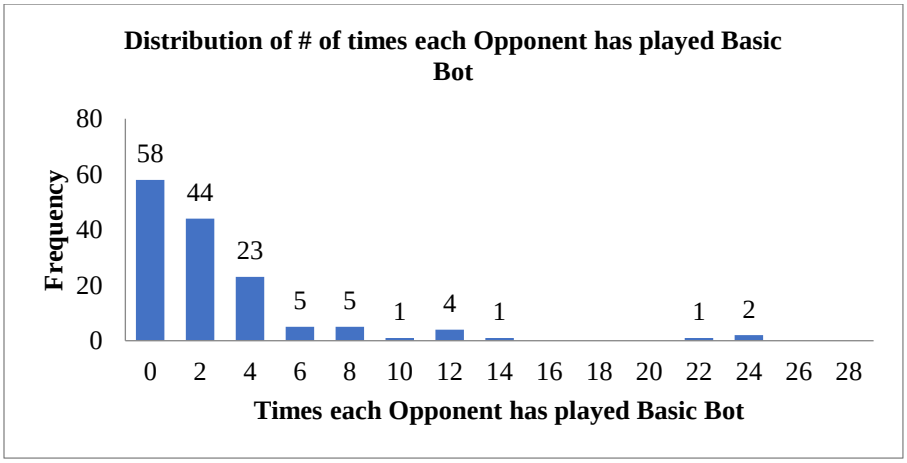

Fig. 1

Figure 1 illustrates the frequency of opponents facing BasicBot, displaying how often each opponent has encountered the BasicBot in their matches.

In terms of the scores dataset, the distribution of the ratings is between 1100 to 2300. This is consistent with the distribution of the Scrabble ratings. This implies no spurious values in the rating and scores dataset to be concerned about.

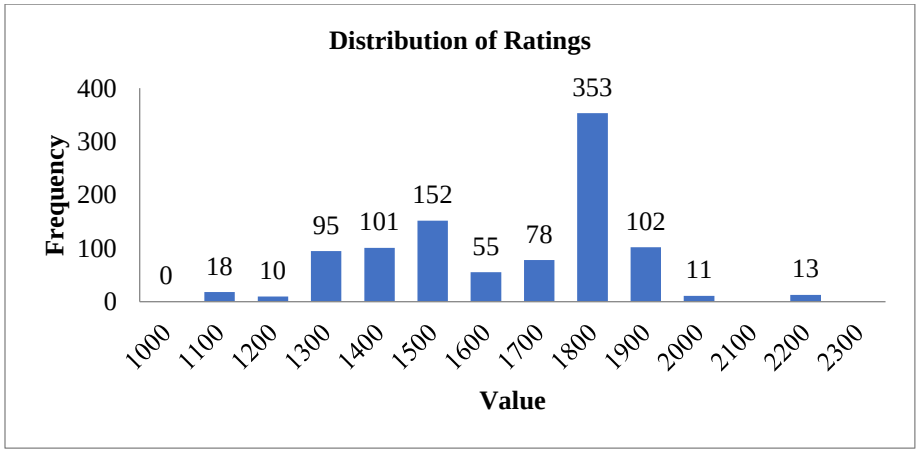

Fig. 2

Figure 2 depicts the variation in BasicBot's ratings when competing against human players, showcasing the distribution of its performance in such encounters.

In terms of the distribution of opponent ratings and opponent scores, the majority of opponent games lie between 1500-1600 rating bin because there are several unconfirmed ratings. When someone joins Woogles and has played a less amount of games, Woogles gives an unconfirmed rating of 1500. Hence, the modal bin observed is in the range of 1500-1600.

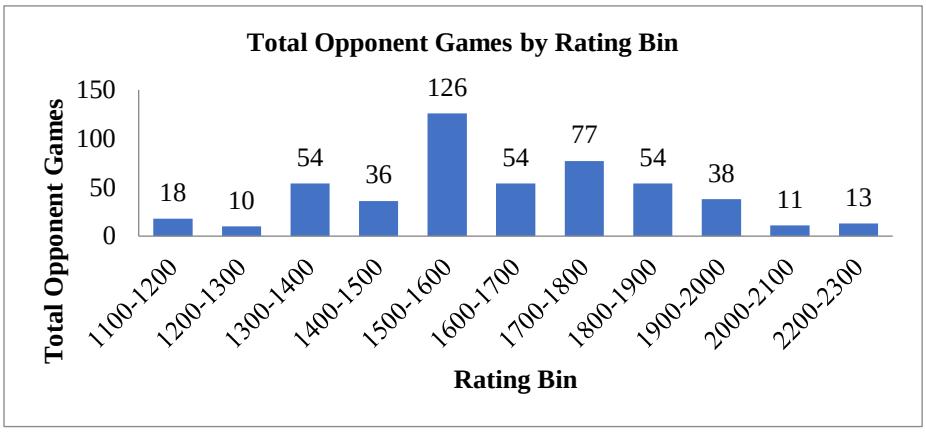

Fig. 3

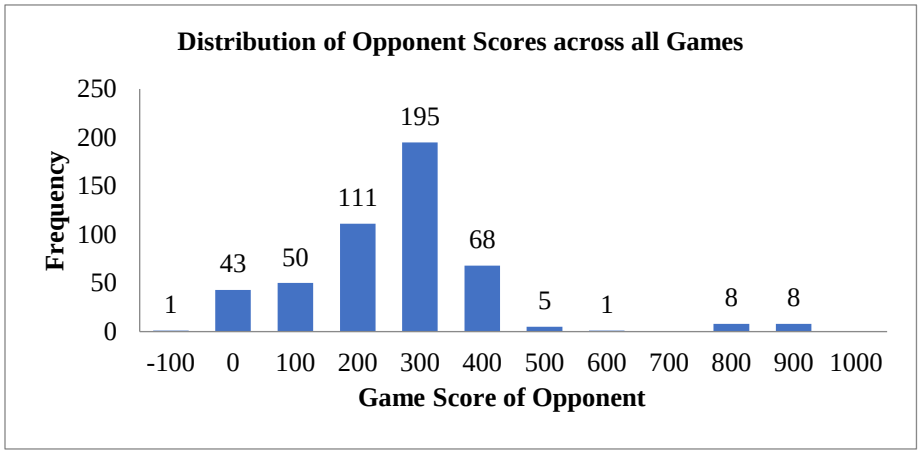

Fig. 4

Figure 3 illustrates the relationship between the number of games played and specific rating categories when facing BasicBot. Meanwhile, Figure 4 displays the overall distribution of opponent scores across all matches, providing insight into the performance outcomes of opponents when engaging with the system.

A linear and statistically significant relationship is found between the two variables by exploring the relationship between opponent score and rating using simple regression. This observation aligns with intuitive expectations, as one would naturally anticipate a linear correlation between scores and ratings, as is indeed the case in the findings.

OLS Regression Results

Dep. Variable:	score_Opponent	R-squared:	0.030
Model:	OLS	Adj. R-squared:	0.028
Method:	Least Squares	F-statistic:	15.32
Date:	Tue, 15 Aug 2023	Prob (F-statistic):	0.000104
Time:	08:53:37	Log-Likelihood:	-3186.6
No. Observations:	491	AIC:	6377.
Df Residuals:	489	BIC:	6386.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	115.5309	51.327	2.251	0.025	14.682	216.379
Rating Opponent	0.1222	0.031	3.914	0.000	0.061	0.183

Omnibus:	193.407	Durbin-Watson:	1.957
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1003.789
Skew:	1.657	Prob(JB):	1.07e-218
Kurtosis:	9.171	Cond. No.	1.17e+04

Fig. 5

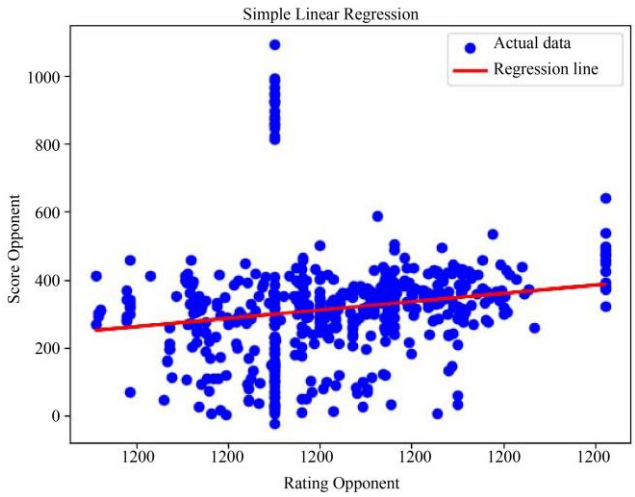

Fig. 6

Figure 5 presents a table showcasing the Simple Linear Regression analysis between the opponent's score and their rating. In contrast, Figure 6 highlights the Simple Linear Regression analysis between the opponent's score and their rating, offering insights into the correlation between these variables.

5. Conclusion

5.1 Summary

This research paper delves into the transformative effects of Artificial Intelligence (AI) on the classic board game Scrabble. It explores how AI platforms like Maven and Quackle have revolutionized gameplay strategy and performance. Scrabble, known for its incomplete information nature and the need for rapid move generation, presents a unique challenge for AI development. The study investigates the algorithms behind these AI engines, particularly the Directed Acyclic Word Graph (DAWG) and Generalized Abbreviated Directed Acyclic Word Graph (GADDAG), which enhance word validation and generation efficiency. It is found that the P-value of the simple linear regression is less than 0.05, suggesting that the relation between the opponent rating and opponent score is significant.

5.2. Limitations

The learning process involved gaining insights into various aspects, including data analysis, expanding coding skills beyond the curriculum, and enhancing research capabilities. However, it is essential to acknowledge that the performance analysis relies on publicly available data, which might not encompass Scrabble's full range of AI capabilities.

5.3. Implications

The implications of this research extend beyond Scrabble. Insights gained from Scrabble AI research can be applied to other incomplete information games like Poker and decision-making contexts with limited data. The use of AI algorithms such as DAWG and GADDAG can drive advancements in algorithmic techniques and problem-solving approaches.
Moreover, Scrabble AI engines like Maven and Quackle exemplify the symbiotic relationship between traditional games and AI technology, influencing competitive gameplay and strategic thinking. As AI continues to advance, it will likely have a broader impact on various fields beyond gaming, making this research relevant to the wider AI community.

Acknowledgements

I would like to thank my mentors, Pranay Kapoor, Srishti Kapoor, and my counsellor, Seerat Kaur, for supporting me in writing the paper.

References

[1] Scrabble, Board Game, Britannica, 2023. [Online]. Available: https://www.britannica.com/sports/Scrabble

[2] Ed. Burns, Nicole Laskowski, and Linda Tucci, "What is Artificial Intelligence," *Enterprise AI*, 2021. [Google Scholar] [Publisher Link]

[3] Pradipta Kumar Das, D. Chandrasekhar Rao, and Kishore Kumar Sahu, "Artificial Intelligence Lecture Notes," *Computer Science and Engineering Information Technology, Veer Surendra Sai University of Technology*, pp. 1-213, 2017. [Google Scholar] [Publisher Link]

[4] Glen Robertson, and Ian Watson, "A Review of Real-Time Strategy Game AI," *AI Magazine*, vol. 35, no. 4, pp. 75-104, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[5] Artificial Intelligence in Boardgames, AVA, 2022. [Online]. Available: https://avas-world.com/2022/11/17/artificial-intelligence-in-boardgames/

[6] Htun Pa Pa Aung, and Hiroyuki Iida, "Advantage of Initiative Revisited: A Case Study Using Scrabble AI," *2nd International Conference on Advanced Information Technologies(ICAIT)*, pp. 1-5, 2018. [Google Scholar] [Publisher Link]

[7] Priyatha Joji Abraham, "A Scrabble Artificial Intelligence Game," *San Jose State University*, pp. 1-27, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Quackle Version 1.0.4, Quackle , 2019. [Online]. Available: https://people.csail.mit.edu/jasonkb/quackle/

[9] Games Live Now, Woogles.io. [Online]. Available: https://woogles.io/

[10] Andrew W. Appel, and Guy J. Jacobson, "The World's Fastest Scrabble Program," *Communications of the ACM*, vol. 31, no. 5, pp. 572-585, 1988. [CrossRef] [Google Scholar] [Publisher Link]

[11] H. Srinivas, and V. Amruth, "Efficient Construction of Dictionary Using Directed Acyclic Word Graph," *IOSR Journal of Computer Engineering*, vol. 18, no. 4, pp. 41-45, 2016. [CrossRef] [Publisher Link]

[12] Steven A. Gordon, "A Faster Scrabble Move Generation Algorithm," *Software Practice and Experience*, vol. 24, no. 2, pp. 219-232, 1994. [CrossRef] [Google Scholar] [Publisher Link]