Quantitative Evaluation of Predictive Analytics: A Comparative Study of Machine Learning Models in eSports Outcome Forecasting

Yue Fei

Institute of Health Informatics, School of University College London, Gower Street, London, 7419206790, United Kingdom
rmhiyfe@ucl.ac.uk

Abstract. The popularity of video games such as Pokémon has led to victory prediction receiving increasing attention from researchers and the eSports industry. This study used a dataset containing a variety of Pokémon attributes (including type, attack, defence, speed, and special abilities) and machine learning algorithms such as logistic regression, K-nearest neighbors, neural networks, and decision trees, to predict the outcome of battles based on these attributes and to provide insights into the complex dynamics of Pokémon battles. The results of the study show that the neural network and decision tree outperformed the others, with speed, attack power, and character-type relationships being the most important factors in determining victory or defeat. The models were 95% accurate, highlighting their potential role in shaping strategic decisions in games. In addition to proving the models' effectiveness, this work advances the field of predictive game analysis by emphasizing the crucial strategic components of winning Pokémon battles.

Keywords: Pokémon, eSports games, logistic regression, KNN, neural networks, decision trees.

1 Introduction

1.1 Background

Researchers and the eSports business are becoming more and more interested in victory prediction. Since its introduction in 1996, Pokémon has transcended its original form as a role-playing game (RPG) to become a global cultural phenomenon, including television series, comics, and various video game iterations, including the popular Pokémon Go. The enduring popularity of the series provides a rich field for analytical exploration, particularly through the lens of machine learning, to reveal the strategic elements behind its gameplay.

Y. Wang (ed.), *Proceedings of the 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024)*, Advances in Computer Science Research 115,
https://doi.org/10.2991/978-94-6463-540-9_16

1.2 Related research

An increasing number of researchers are concentrating on the application of sophisticated statistical methods and machine learning to forecast the likelihood that players will win games due to the quick growth of the video game business.

In earlier research, Hao Yi Ong, Sunil Deolalikar, and Mark Peng used logistic regression, support vector machines, and Gaussian discriminant analysis in addition to K-means to cluster player behaviors to predict League of Legends match outcomes [1]. In their work, Johansson and Wikstrom used partial game state data to do real-time prediction to build a model using several parameterized versions of the Random Forest algorithm to forecast the winning Dota2 (MOBA) team [2]. In the meanwhile, Weiqi Wang used hero draft data to forecast Dota2 match results using a multi-layer feed-forward neural network [3].

In several Destiny game styles, Ravari, Spronck, Sifa, and Drachen employed random forests to forecast winning strategies [4]. Additionally, they examined various performance indicators and how each affected each mode. Like this, Ravari, Bakkes, and Spronck also employed Random Forests to forecast the outcomes of games in different Starcraft game modes, and they were able to get the best outcomes [5]. The process of using data analysis and game theory to forecast the results of competitive Pokémon fights was also covered by Dhruva Bhagwat in his paper "Applying Data Analysis and Game Theory to Competitive Pokémon Battle Combinations" [6].

1.3 Objection

The literature review highlights the growing complexity and variety of approaches used for game outcome prediction, especially when machine learning techniques like random forests and neural networks are employed. However, previous research has mainly focused on the application of single models, often ignoring the potential benefits of comparative approaches that juxtapose multiple predictive models. This oversight limits the potential synergies that can be exploited when integrating insights from various model outputs. By offering a thorough comparison of several models, this study seeks to close this gap and provide readers a more nuanced understanding of the predictive capabilities and contributions of each model to the field of e-sports game analytics. Therefore. this paper aims to enhance battle strategy formulation in Pokémon RPGs by employing a suite of statistical and machine learning algorithms to accurately predict the outcomes of one-on-one battles. By pinpointing key determinants of battle success, the study illuminates the strategic underpinnings of Pokémon attributes and offers a new perspective on the application of data analytics in video game strategy development.

2 Introduction

2.1 Data

This study uses a comprehensive dataset published on Kaggle seven years ago by ALOPEZ247 to perform feature-based statistical analyses of Pokémon characters. The dataset contains 800 observations, providing many data points that allow for more detailed analysis and robust validation of the predictive models used. The variables listed in Table 1 were selected because they are all indicative of factors affecting battle outcomes and broadly cover potential strategy combinations and outcomes in Pokémon battles.

Table 1. Description of the data.

Variables	Descriptions	Type
Name	The English name of the Pokémon	character
Attack	Base attack of the Pokémon	numeric
HP	Base health points of the Pokémon	numeric
Legendar	Boolean showing the legendary status of the Pokémon	logic
Defense	Base defense of the Pokémon	numeric
Speed	Base speed of the Pokémon	numeric
Type 1	Primary type, related to the nature (18 possible types)	character
Sp.Def	special defense of the Pokémon	numeric
Sp.Atk	special attack of the Pokémon	numeric

In Pokémon games, certain types of Pokémon may have a natural advantage or disadvantage against other specific types of Pokémon. Boxplots are a standard data visualization tool. As shown in Figure 1, Fire type1 Pokémon usually have higher attack power than Water type1 Pokémon.

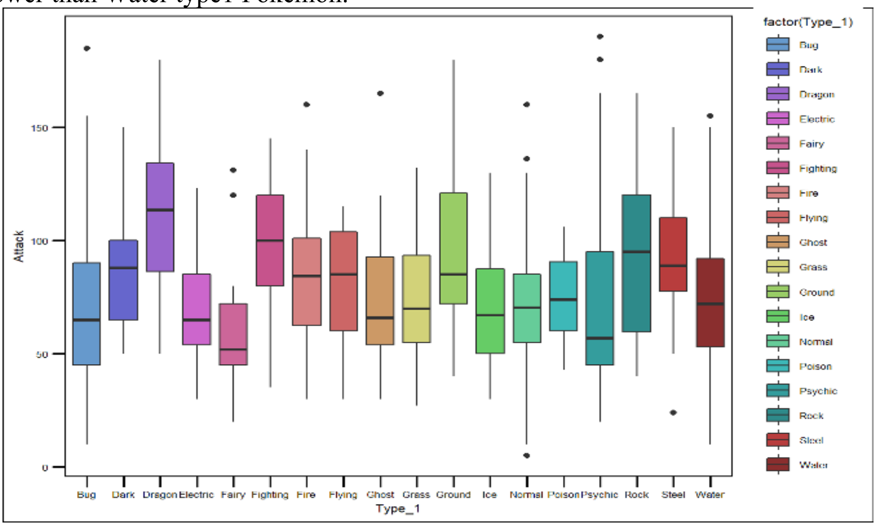

Figure 1. Boxplot of the distribution of Pokémon stats by major type.

2.2 Methodology

Logic Regression. To predict the outcome of a match between two characters, logistic regression is first considered. Since the result of a Pokémon battle can only be binary (i.e., win (y=1) or lose (y=0), logistic regression is the best method for handling this kind of dependent variable. It can predict the likelihood of an event happening, such as the likelihood that the first Pokémon would prevail in the battle. The logistic regression model uses a logistic function to convert linear combinations into probability values. Its mathematical form is as follows:

$$P(Y=1)=\frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n)}} \quad (1)$$

where P(Y=1) is the chance that event, Y occurs; the model predicts a higher probability of event Y occurring, i.e., a larger probability of the first Pokémon winning, if P(Y=1) is greater than or equal to 0.5. X_1, X_2, ..., $X_n$is an input variable indicating various factors affecting the outcome of the matchup, such as attack power, defence power, speed, and so on. β_0, β_1, ..., $\beta_n$are parameters of the model, indicating how much each input variable affects the outcome [7].

Neural Network. Neural networks, which can also predict the outcome of a sparring match, process and learn from data by modelling the way the human brain works, adjusting the weights connecting individual neurons to learn complex patterns and relationships in the data. At the hidden layer, the model transforms the input variables into internal feature representations that can represent complex relationships. This process incorporates nonlinear transformations that allow the network to capture complex patterns and interaction effects in the input data. Finally, in the output layer, the network outputs the probability of the first character winning based on the information processed in the hidden layer [8].

K-Nearest Neighbors (KNN). For regression and classification, non-parametric techniques like the K-Nearest Neighbors algorithm are employed. To determine the most likely winner in Pokémon matchups, this study uses it to locate the 'K' closest training samples in the feature space. The Pokémon then adopts the label that is most popular among its 'k' nearest neighbors, and the prediction is determined by a majority vote among its neighbors. Cross-validation, a method that guarantees the model's performance is not just dependent on the initial train-test split but also generalizes well to fresh data, is used to estimate the ideal number of neighbors, or 'k' [9]. The choice of 'k' profoundly affects the classification's accuracy, with too low values leading to high sensitivity to noise in the training data, and too high values resulting in underfitting. Therefore, it is critical to select an appropriate 'k', which minimizes the prediction error on unseen battles.

Decision Tree. As a predictive modelling technique, decision trees provide an understandable and visual depiction of the decision-making process. Every internal node in the tree represents an attribute, and the branches represent the test results, culminating in leaf nodes that represent the categorization or final judgment. One important parameter in decision tree development is the complexity parameter (cp), which controls the growth of the tree by removing smaller splits. By adjusting cp, overfitting may be controlled, and the model can be made to capture the important patterns in the data without becoming unduly complicated. For a model to be generalizable, depth and simplicity must be balanced [10].

3 Results

3.1 Logic Regression

Since the outcome of a Pokémon matchup depends heavily on the difference in ability between the two participants, rather than just their respective absolute ability values, it is more appropriate to use the difference in ability as the dependent variable. After performing a goodness-of-fit test, this study chose to use a model containing two "lore" variables, with the dependent variables being the difference between attack, defence, sp.attack, sp.defence, speed, and HP, as well as the two lore and type relationships. The Wald-test showed that the difference between Speed, attack, HP, and defence, as well as the type of relationship was the most statistically significant variables in determining the outcome of the battle, with p-values less than 0.001.

The logistic regression model's prediction ability can be evaluated using ROC curves. The model performs better the closer the AUC value is to 1. The model has a significant predictive performance with 93% accuracy for cross validation, according to the results (Figure 2).

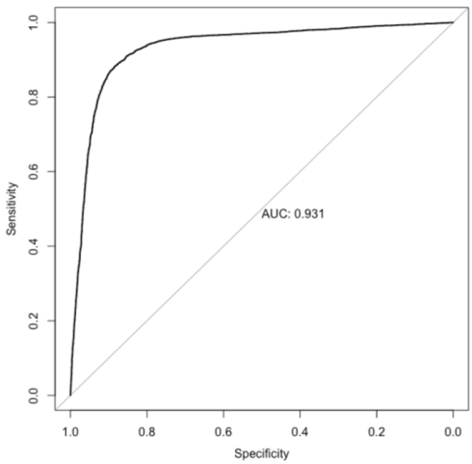

Figure 2. ROC curve.

3.2 Neural Network

To predict the likelihood of the first character winning, a neural network model is constructed, with I1 through I9 standing in for the several variables that influence the win rate. The model nonlinearly converts the input variables into internal feature representations in the hidden layers (H1 to H5), which can represent intricate relationships. The network finally outputs the likelihood that the first character will win at the output layer (O1) using the data analysed in the hidden layer. In this process, the entire network is trained with the training data and the weights of each connection are adjusted to minimize the prediction error. Depending on the number of nodes selected a neural network model can be created (Figure 3).

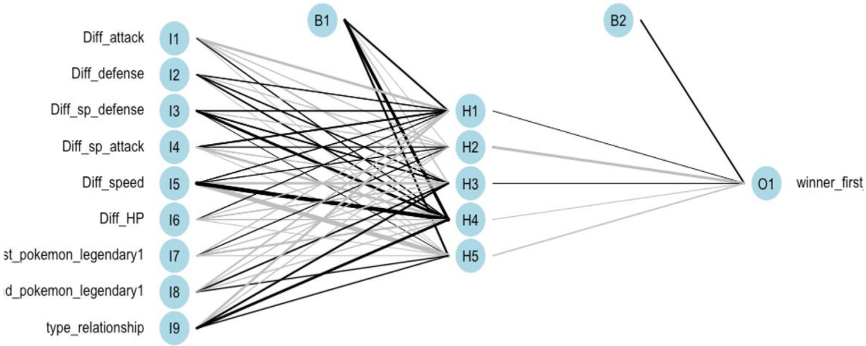

Figure 3. Neural Network model.

Cross-validation was used to assess the predictive ability of the model and the average value of RATE was obtained as 95%, which indicates that the established neural network model has excellent predictive ability.

3.3 KNN

Use cross validation to find out the k value that return us the min error. From the results, k=12 is a good choice (Figure 4). In cross validation it was found that the correctness of model prediction was found to be 0.91 when k=12, which indicates that the neural network model predicts 91% correctly.

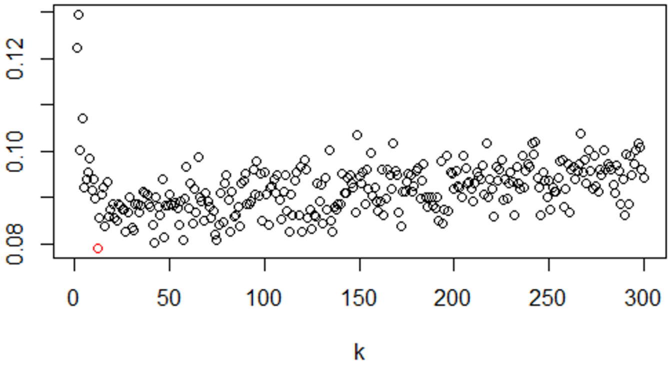

Figure 4. Test misclassification rates.

3.4 Decision Tree

This study compares classification trees and pruned classification trees to assess the impact of complexity on prediction accuracy. Pruning a decision tree involves removing branches that have little to no contribution to the prediction power, which can improve the model's generalizability and prevent overfitting. While a full classification tree uses all available data points and may capture too much noise, resulting in complex models, a pruned tree simplifies the model by considering only the most informative splits.

By comparing a tree with a complexity parameter (cp) of 0.0003 (Figure 5) against one with a cp of 0.0008 (Figure 6), we can observe the trade-offs between model simplicity and predictive performance. Our findings indicate that simplification through pruning does not significantly compromise the model's accuracy, which remains high at 95%. This suggests that pruning, which yields a more interpretable and parsimonious model, does not detract from our ability to accurately predict outcomes of Pokémon battles.

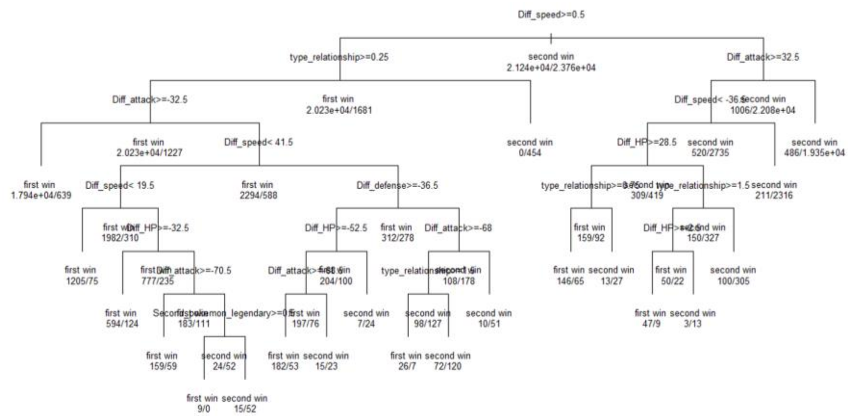

Figure 5. Classification Tree.

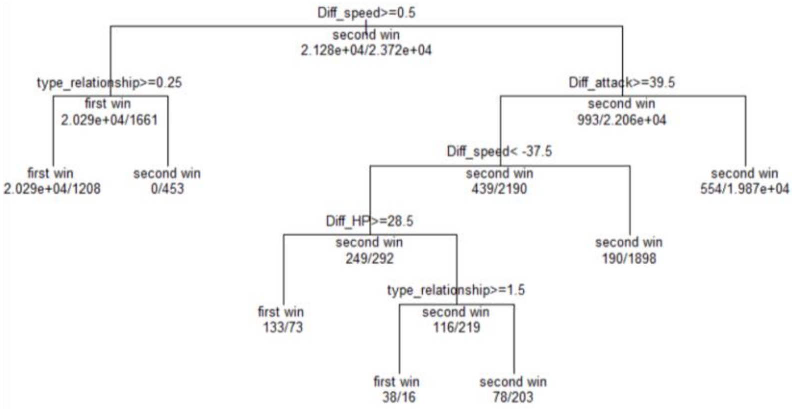

Figure 6. Pruned Classification Tree.

Cross-validation shows that there is not much difference in the accuracy before and after pruning, so the decision tree model after pruning is chosen to better explain the model. The accuracy of the pruned decision tree is 95%, which indicates the excellent performance of using decision trees to predict the outcome of battles. According to the images, differences in speed, attack, HP, and type relationships are the most important statistical variables in determining the outcome of a battle. Combined with logistic regression findings, speed, attack, and type relationships play a decisive role in predicting the outcome of character battles.

4 Conclusion

This study explores the dynamics of Pokémon matchmaking through the lens of machine learning using a dataset containing various Pokémon attributes. The results of the study show that especially speed, attack, and type relationships, play a pivotal role in determining the winner of a Pokémon versus Pokémon battle. Predictive models including neural networks, KNNs and random trees performed well, with neural networks and random trees being particularly accurate.

It's critical to recognize that this study has a lot of limitations even though it offers a thorough investigation of the dynamics of Pokémon battles and shows the great potential of machine learning models in battle outcome prediction. First off, despite being large, the dataset utilized excludes elements that could significantly affect the outcome of battles, like Pokémon movement combinations and special abilities. In addition, the model has not been tested and validated on different datasets, potentially limiting an accurate assessment of its generalization ability. Future research would benefit from the use of a more comprehensive dataset, including more attributes of Pokémon and battle logs, as well as more extensive testing of the model in different contexts. In addition, the model also failed to account for randomness in battles and the complexity of real-time decision-making, both of which may play a key role in actual battles. Therefore, to capture more intricate connections between Pokémon traits and battle situations, future research should investigate deeper learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

References

1. Ong H Y, Deolalikar S and Peng M. (2015). Player Behavior and Optimal Team Composition for Online Multiplayer Games. arXiv:1503.02230
2. Johansson F and Wikström J. (2015). Result Prediction by Mining Replays in Dota 2. Student thesis, Blekinge Institute of Technology, Faculty of Computing, Department of Computer Science and Engineering.
3. Wang W. (2016). Predicting Multiplayer Online Battle Arena (MOBA) Game Outcome Based on Hero Draft Data. (Master's thesis, National College of Ireland)
4. Norouzzadeh Ravari Y, Spronck P, Sifa R and Drachen A. (2021). Predicting Victory in a Hybrid Online Competitive Game: The Case of Destiny. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 13(1), 207-213
5. Norouzzadeh Ravari Y, Bakkes S and Spronck P. (2021). StarCraft Winner Prediction. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 12(2), 2-8
6. Dhruva Bhagwat. Application of data analysis and game theory to competitive Pokémon battle sets. TechRxiv. September 27, 2022
7. Hosmer Jr D W, Lemeshow S and Sturdivant R X. (2013). Applied logistic regression. John Wiley & Sons
8. Goodfellow, I., Bengio, Y. and Courville, A. (2016). Deep Learning. MIT Press
9. James G, Witten D, Hastie T and Tibshirani R. (2013). An Introduction to Statistical Learning. New York: Springer
10. Hastie T, Tibshirani R and Friedman J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer