ASTEEC Conference Proceeding: Computer Science
3rd International Conference on Information Science and Technology Innovation (ICoSTEC)
July 27, 2024, Yogyakarta, Indonesia

Mining PUBG Winning Rates: Between Chicken Dinner and In Game Traveling Distance

Veronica Wijaya[1], Teddy Surya Gunawan[2], Zakarias Situmorang[3]

[1,2]*Universitas Potensi Utama*
Jl. KL. Yos Sudarso Km. 6,5 No. 3-A, Tanjung Mulia, Medan

[3]*Universitas Katolik Santo Thomas*
Jl. Setia Budi No.479, Tanjung Sari, Medan

[1]veronicawijayawork@gmail.com
[2]tsgunawan@gmail.com
[3]zakarias65@yahoo.com

Abstract—**This research article explores the relationship between travel distance and player performance in PlayerUnknown's Battlegrounds (PUBG), a popular battle royale game. Focusing on walking, riding, and swimming distances, our study employs machine learning algorithms, including Random Forest (RF), Support Vector Machine (SVM), and Multilayer Perceptron (MLP), to predict PUBG player performance, with a primary emphasis on solo matches. Our research encompasses data collection, filtering, model configuration, data sampling, and model evaluation stages to ensure data reliability and model appropriateness. The results highlight the consistent superiority of the Random Forest model, which demonstrates the lowest Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the highest R-squared (R2) values in both training and testing phases. In contrast, the SVM-Sigmoid model consistently delivers poor predictive performance, emphasizing the significance of selecting an appropriate model. In summary, this research unveils the vital role of travel distance in PUBG player performance and underscores the importance of data-driven decision-making and model selection in achieving accurate and reliable predictions. This contribution enhances our understanding of player performance in the dynamic world of PUBG.**

Keywords— **PlayerUnknown's Battlegrounds, Travel Distance, Random Forest, Multilayer Perceptron, Support Vector Machine**

I. INTRODUCTION

PlayerUnknown's Battlegrounds (PUBG) stands as a globally acclaimed battle royale-themed First-Person Shooter game [1]. Originally launched on Personal Computer, it has since proliferated across multiple platforms, including mobile and console gaming [2]. PUBG players' primary objective is to endure and outlast their peers, culminating in the coveted title of 'Winner Winner Chicken Dinner' [3]. To secure survival in PUBG, players traverse the game map, either on foot, by utilizing vehicles scattered haphazardly throughout, while simultaneously scavenging for weaponry, armor, and medical supplies, and dispatching any rivals they encounter [4]. A significant factor motivating players to remain active within the game is the looming threat of the Blue Zone, a restricting area that contracts as the game unfolds, imposing damage on players outside its bounds [5]. Consequently, the distance players cover by walking, using vehicles, or swimming emerges as a critical determinant in their survival prospects.

This research centers on the exploration of a novel dimension within PUBG gameplay: the influence of travel distance on a player's performance. Specifically, this study spotlights three specific features – walking distance, ride distance, and swimming distance – collectively termed as the total travel distance of PUBG players [6].

Prior investigations have indeed employed this dataset to forecast the final rankings of PUBG players, utilizing varying combinations of features. In this study, we draw inspiration from three prior works:

- Dipaktashi Sen et al. conducted research comparing the predictive performance of multiple algorithms, including LightGBM (LGBM), Gradient Boosting Regressor (GBR), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (K-NN), and Linear Regression (LR). They employed 8 and 14 features, which included walking distance, ride distance, and swimming distance. Their findings indicated that the GBR algorithm yielded the best performance, with a Mean Absolute Error (MAE) of 0.062 using 14 features and 0.065 using 8 features [7].
- N.F. Ghazali et al. compared Decision Trees (DT), Linear Regression (LR), and Support Vector Machine (SVM) algorithms to predict the placement of e-sport players, using 26 features that encompassed Total Distance. Their results highlighted the superiority of the SVM algorithm, with the lowest Root Mean Square Error (RMSE) value of 0.09372 [8].
- Madhurya Manjunath Mamulpet's research pitted the LGBM, Multilayer Perceptron (MLP), and Recurrent Neural Network (RNN) algorithms against each other for predicting the winner placement in PUBG, employing the total travel distance as a key feature. Their study underscored LGBM's dominance, with the lowest MAE value of 0.0539 [9].

Building upon the insights from these previous studies, our research endeavors to bring a fresh perspective to the table. We undertake a comparative analysis of three distinctive algorithms - Random Forest (RF), Support Vector Machine

(SVM), and Multilayer Perceptron (MLP) - utilizing three fundamental features: walking distance, ride distance, and swim distance.

RF is one of the ensemble methods known for its high accuracy and low computational burden [10]. This method is popularly used in prediction problems, such as predicting a child's talent based on hobbies, with an average accuracy of 92.75% [11], predicting student performance with an accuracy of 96.88% [12], and predicting student graduation with an average accuracy of 90.45% [13].

SVM is one of the supervised learning algorithms known for its high performance and efficiency, especially on non-linearly separable data [14]. This method shows a high level of accuracy in several prediction studies, such as predicting employee recruitment with 84.59% accuracy [15], predicting stroke disease with 94% accuracy [16], and predicting majors in vocational schools with an accuracy of 89% [17].

MLP is an algorithm that resembles a human neural network, with a feedforward architecture consisting of input, hidden, and output layers [18]. Several studies have demonstrated the high accuracy of this method in various prediction problems, including flood prediction in eastern Iraq (94% accuracy) [19], student performance prediction (87.3% accuracy) [20], and military conscript stress level prediction (90% accuracy) [21].

What sets this research apart is its exclusive focus on solo-player matches within the PUBG ecosystem and a meticulous exploration of the walk distance, ride distance, and swim distance features. By employing Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R-squared (R2) as performance metrics, we aim to discern the most effective model for predicting PUBG player performance. In summary, this research endeavors to unravel the intricate relationship between walk distance, ride distance, swim distance, and the winning rate of PUBG players, specifically in solo matches.

II. Research Methods

This research involves several stages, including data collection, filtering, model configuration, training, testing model predictions, comparing performance results, and analyzing the influence of each feature on predictions. Figure 1 illustrates the stages of this research.

A. *Data Acquisition and Features Filtering*

The foundation of this research lies in data acquisition and features filtering. An extensive dataset of 1,111,742 data points, comprising 26 features and a single target class [22], serves as the starting point. This dataset is refined through a meticulous filtering process, resulting in a reduced dataset of 179,608 data points. The filtration process involves two critical steps:

1. Match type filtering
   Leveraging the 'matchType' feature, the dataset is meticulously filtered to identify PUBG match types characterized as solo matches. This includes the selection of variables such as 'normal-solo,' 'solo,' and 'solo-fpp' [23]–[25]), ultimately yielding the filtered dataset of 179,608 entries.
2. Features Filtering
   Out of the 26 features originally present in the dataset, three features are selected for inclusion in the prediction model: 'walk distance,' 'ride distance,' 'swim distance,' and the 'winPlacePerc' feature. This final selection constitutes the essential components for the predictive model.

B. *Model Configuration*

The model configuration, as depicted in Figure 1, is a pivotal phase of the research. The research employs three distinct algorithms: Multi-Layer Perceptron (MLP), Random Forest (RF), and Support Vector Machine (SVM). Within this framework, a deliberate exploration of model configurations takes place:

1. MLP-Based Models
   These models feature two configuration variations characterized by the utilization of ReLU and Sigmoid activation functions. Each MLP-based model comprises three hidden layers with 100 neurons.
2. SVM-Based Models: SVM models are equipped with two configuration variations, employing the Radial Basis Function (RBF) and Sigmoid kernel functions, each with a gamma value set to 0.01.
3. RF-Based Models: In the Random Forest models, a single variation is employed, incorporating a total of 100 decision trees. These models promise to capture the intricate relationships present within the data.

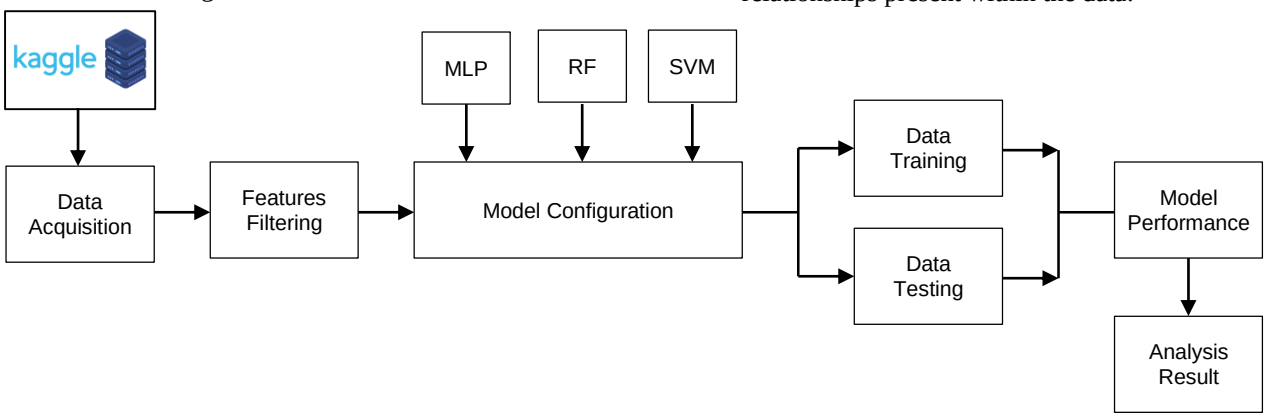

Fig 1. Research Design

The selection of the activation functions for the Multi-Layer Perceptron (MLP) and the choice of kernel functions for Support Vector Machines (SVM) and the utilization of Random Forest in this research are grounded in their distinct advantages. ReLU is preferred for MLP due to its proficiency in capturing non-linear data relationships, computational efficiency through sparse activation, mitigation of the vanishing gradient issue, and ease of implementation and interpretation [26]. Conversely, the Sigmoid activation function is specifically chosen for its suitability in scenarios where interpretability, smooth gradient transitions, and probability outputs are critical, especially in binary classification tasks [27]. In the case of SVM, the deployment of the Radial Basis Function (RBF) kernel is preferred for its prowess in modeling complex non-linear data relationships, making it well-suited for tasks that defy linear assumptions [28]. On the other hand, the Sigmoid kernel in SVM is advantageous for introducing non-linearity into the model, rendering it appropriate for tasks characterized by intricate non-linear patterns [29]. Additionally, Random Forest is harnessed for its capabilities in handling non-linear data relationships, averting overfitting through ensemble learning, robustness to outliers, feature importance insights, and efficacy in high-dimensional datasets [30]. The decision to employ 100 trees in the Random Forest model is founded on its balanced capacity to capture diverse data patterns and prevent overfitting, rendering it an efficient choice for numerous regression tasks within the study [31].

C. Data Sampling

Data sampling is a fundamental element of this study, entailing a deliberate selection of a 70:30 sampling ratio. In this ratio, 70% of the dataset is allocated for training, while the remaining 30% is dedicated to testing. A careful stratified split of the initial dataset, comprising 45,679 data points, yields 31,976 data points for training and 13,703 data points for testing.

Through a statistical examination of both the training and testing datasets, we calculated the means, medians, and measures of dispersion for all features, culminating in the results displayed in Table I and Table II.

TABLE I
TRAINING DATA FEATURES STATISTIC

Feature	Mean	Median	Dispersion
Walk Distance	990.983	616.35	1.068
Ride Distance	651.183	0	2.54
Swim Distance	5.692	0	6.403
Win Percentile	0.474	0.468	0.648

TABLE I
TESTING DATA FEATURES STATISTIC

Feature	Mean	Median	Dispersion
Walk Distance	982.79	607.7	1.072
Ride Distance	641.257	0	2.531
Swim Distance	6.122	0	6.238
Win Percentile	0.471	0.4639	0.652

The analysis of the 70:30 split of the data and the comparison between training and testing data provides insights into the effectiveness of this sampling ratio and the reliability of the training data in representing the testing data.

Firstly, with respect to the sampling result (the effectiveness of 70:30), the 70:30 split is a common practice in dividing datasets for training and testing purposes. In this case, 70% of the data is allocated to the training set, while the remaining 30% is used for testing. This split aims to strike a balance between a sufficiently large training dataset and a suitably sized testing dataset for robust model development and evaluation.

The analysis of feature statistics in the training and testing data reveals noteworthy trends. The similarity in means, medians, and dispersion for most features, such as "Walk Distance" and "Win Percentile," suggests that the 70:30 sampling ratio has been effective in creating training and testing datasets that are reasonably representative of each other. This alignment in statistical characteristics is a positive indication, as it implies that the training and testing data share similar central tendencies and variability.

Regarding the reliability of the training data in representing the testing data, the findings are encouraging. The similar means and medians observed in both datasets for most features indicate that the training data provides a reliable representation of the central tendencies of the testing data. This consistency suggests that the training data is capturing the core characteristics of the testing data.

However, it is important to note that both the training and testing datasets exhibit a substantial presence of zero values in "Ride Distance" and "Swim Distance." This shared characteristic implies that a significant portion of the data in both datasets lacks ride or swim distance information. This aspect should be considered in subsequent analyses and modeling efforts, as it can have an impact on the overall interpretation and predictive modeling.

In summary, the 70:30 sampling ratio appears to have effectively created training and testing datasets that align well in terms of statistical characteristics. This implies that the training data can be considered relatively reliable in representing the testing data, with some specific considerations required for features with numerous zero values. These insights underscore the importance of understanding and accounting for data characteristics in the design and interpretation of analytical and modeling tasks.

D. Model Evaluation

In the pursuit of accurate and robust predictive modeling, the evaluation of model performance is an essential facet of the research process. In this section, we detail the methodology employed for model evaluation using the train-test split approach, and the metrics utilized to assess model efficacy. We rely on well-established metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R2), to provide a comprehensive evaluation of the models' predictive capabilities.

The train-test split method is a widely adopted approach for model evaluation in the realm of machine learning and data science. It involves partitioning the dataset into two distinct

subsets: the training dataset and the testing dataset. The training dataset is used to train the model, allowing it to learn patterns and relationships within the data. Meanwhile, the testing dataset serves as an independent dataset that the model has not seen during training, providing an objective measure of its predictive performance.

To quantify the effectiveness of the models, we employ several key metrics:

1. Mean Squared Error (MSE)
   MSE measures the average squared difference between the actual and predicted values. A lower MSE signifies that the model's predictions are closer to the true values, indicating better predictive performance.
2. Root Mean Squared Error (RMSE)
   RMSE is the square root of the MSE and is expressed in the same units as the target variable. It provides a more interpretable measure of prediction error, and like MSE, a lower RMSE is indicative of better model performance.
3. Mean Absolute Error (MAE)
   MAE calculates the average absolute difference between the actual and predicted values. MAE is less sensitive to outliers compared to MSE and RMSE, making it a valuable metric for assessing model robustness.
4. R-squared (R2)
   R2, also known as the coefficient of determination, quantifies the proportion of variance in the target variable that is explained by the model. R2 ranges from 0 to 1, with higher values indicating a better fit of the model to the data.

These metrics collectively provide a comprehensive view of a model's performance, addressing different aspects of accuracy, error, and explanatory power. By evaluating models using the train-test split method and these metrics, we can make informed decisions about the suitability of our models for the given task, facilitating the selection of the most effective model for predictive purposes.

III. Result

We unveil the significant findings of our investigation, which encapsulate the results achieved through gender classification of college students based on digital image analysis. This classification process leveraged the Inception V3 feature extraction technique in conjunction with the Backpropagation classification approach.

A. Training Performance

In Table III, we present a thorough evaluation of performance metrics for diverse classification models utilized in the analysis of PUBG Winning Rates, with a specific emphasis on the influence of in-game traveling distance. These metrics offer valuable insights into the models' effectiveness in providing accurate predictions for winning rates based on the dataset.

Table III
Performance Metric (Training)

Model	MSE	RMSE	MAE	R2
MLP-ReLU	0.019	0.138	0.101	0.8
MLP-Sig	0.019	0.139	0.101	0.795
RF	0.011	0.105	0.069	0.883
SVM-RBF	0.021	0.144	0.106	0.783
SVM-Sig	6.004	2.45	0.917	-62.32

Table III showcases the results across various metrics, revealing how well each model performed.

Random Forest (RF) emerges as the top performer across most metrics. It achieved the lowest Mean Squared Error (MSE), indicating that its predictions are closest to the actual values. Additionally, RF exhibited the lowest Root Mean Squared Error (RMSE), underlining its accuracy in predicting winning rates. Moreover, RF outperformed other models by having the lowest Mean Absolute Error (MAE), reflecting its robustness in providing accurate predictions. Lastly, RF attained the highest R-squared (R2) value, highlighting its superior ability to explain and predict winning rates effectively.

In contrast, the SVM-Sig model stands out with significantly higher MSE, RMSE, and a negative R2, indicating poor predictive performance. These metrics suggest that the SVM-Sig model's predictions deviate considerably from actual winning rates. This misalignment could be attributed to a mismatch between the model's assumptions and the characteristics of the underlying data.

In summary, the performance metrics in Table III shed light on the relative strengths and weaknesses of the models in the context of PUBG Winning Rates classification, with a particular emphasis on the role of in-game traveling distance as a predictive factor. These metrics are instrumental in evaluating the models' predictive accuracy and reliability, aiding in the selection of the most suitable model for predictive tasks.

B. Testing Performance

Table IV provides a critical evaluation of performance metrics obtained from the testing phase of our study, specifically pertaining to the classification of PUBG Winning Rates with a focus on in-game traveling distance as a pivotal feature.

Table IV
Performance Metric (Testing)

Model	MSE	RMSE	MAE	R2
MLP-ReLU	0.018	0.136	0.101	0.805
MLP-Sig	0.019	0.138	0.101	0.8
RF	0.021	0.146	0.107	0.774
SVM-RBF	0.02	0.142	0.106	0.787
SVM-Sig	6.081	2.466	0.907	-63.293

The metrics presented in Table IV offer a comprehensive insight into the effectiveness of different machine learning models when applied to predict winning rates based on the dataset, and specifically how they perform on unseen test data.

It is evident from the table that Random Forest (RF) outperforms the other models across most metrics in the testing phase. It exhibited the lowest Mean Squared Error (MSE), indicating that its predictions were closest to the actual values, thereby minimizing prediction errors. Moreover,

RF had the lowest Root Mean Squared Error (RMSE), suggesting high accuracy in estimating winning rates and presenting an easily interpretable measure of prediction error. RF also recorded the lowest Mean Absolute Error (MAE), underscoring its ability to provide accurate estimates. Lastly, RF achieved the highest R-squared (R2) value, signifying its exceptional capability to explain and predict winning rates effectively, even on unseen test data.

On the contrary, the SVM-Sig model once again stands out with significantly higher MSE, RMSE, and a negative R2 in the testing phase, indicating its poor predictive performance. These metrics highlight that the SVM-Sig model's predictions deviate substantially from actual winning rates when applied to new data.

In summary, Table IV's performance metrics offer a clear understanding of each model's effectiveness in predicting PUBG Winning Rates based on in-game traveling distance, as observed during the testing phase. These metrics are instrumental in assessing the models' predictive accuracy and reliability when applied to unseen data, aiding in the selection of the most suitable model for predictive tasks.

IV. Discussion

Our study focuses on understanding the impact of travel distance on PUBG player performance, specifically analyzing walking distance, ride distance, and swimming distance in solo-player matches. By utilizing machine learning algorithms and a range of performance metrics, we have gained valuable insights into the predictive accuracy of these models.

A. Training Result

In the training phase, our objective was to evaluate the performance of various machine learning models in predicting PUBG Winning Rates based on in-game traveling distance. The training result, presented in Table III, reveals compelling insights into the effectiveness of these models.

Random Forest (RF) consistently outperforms other models across multiple metrics. It demonstrates the lowest Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the highest R-squared (R2) value. These metrics collectively indicate that RF's predictions are closest to the actual values, underscoring its accuracy, robustness, and explanatory power in predicting winning rates effectively.

In contrast, the SVM-Sig model stands out with significantly higher MSE, RMSE, and a negative R2, suggesting poor predictive performance. This discrepancy can be attributed to a mismatch between the model's assumptions and the underlying data, emphasizing the importance of selecting an appropriate model for the task.

In summary, the performance metrics from the training phase offer a clear perspective on the models' relative strengths and weaknesses in the context of PUBG Winning Rates classification based on in-game traveling distance.

B. Testing Result

In the testing phase, we aimed to assess how well these models perform when applied to new, unseen data while considering in-game traveling distance. The testing result, as presented in Table IV, offers insights into the models' effectiveness in this context.

Random Forest (RF) once again stands out as the top performer in the testing phase, exhibiting the lowest MSE, RMSE, MAE, and the highest R2. These findings highlight RF's consistency in providing accurate predictions, even when confronted with previously unseen data.

On the other hand, the SVM-Sig model maintains its poor predictive performance in the testing phase, as indicated by significantly higher MSE, RMSE, and a negative R2. These results underline the importance of selecting a suitable model, as SVM-Sig's predictions deviate substantially from actual winning rates when applied to new data.

The testing result confirms the predictive accuracy and reliability of these models when extrapolated to unseen data, emphasizing the significance of model selection in PUBG Winning Rates classification based on in-game traveling distance.

C. Comparative Analysis

Comparing the results of the training and testing phases, we observe a consistent trend. Random Forest (RF) outperforms other models, demonstrating its robustness and accuracy in predicting PUBG Winning Rates based on travel distance features. In contrast, the SVM-Sig model consistently lags behind, indicating that its predictions deviate significantly from actual winning rates in both training and testing scenarios.

These findings reiterate the importance of selecting an appropriate model, such as Random Forest, to achieve accurate and reliable predictions in PUBG Winning Rates classification. The choice of model significantly impacts predictive accuracy, especially when considering the role of in-game travel distance as a defining feature.

In summary, our research brings a fresh perspective to understanding the relationship between travel distance features and PUBG player performance. It highlights the pivotal role of model selection in achieving accurate and consistent predictions, underscoring the relevance of data-driven decision-making in the context of gaming performance analysis.

V. Conclusions

In this research article, we delved into the intriguing realm of PlayerUnknown's Battlegrounds (PUBG), a globally acclaimed battle royale game, with a primary focus on the influence of travel distance on player performance. Our investigation specifically scrutinized three vital travel features: walking distance, ride distance, and swim distance, collectively constituting the total travel distance of PUBG players. By harnessing the predictive power of machine learning algorithms, we conducted a comparative analysis of three distinct models: Random Forest (RF), Support Vector Machine (SVM), and Multilayer Perceptron (MLP). Our intent was to unravel the intricate relationship between these travel features and the winning rates of PUBG players,

particularly in solo-player matches. Our journey through this research encompassed several key phases, including data collection, filtering, model configuration, data sampling, and model evaluation. We paid meticulous attention to the methodology employed in each stage, ensuring the reliability and representativeness of the data, as well as the appropriateness of model selection. The results of our investigation unveiled a clear winner among the machine learning models. Random Forest (RF) consistently outperformed the other models in both training and testing phases, exhibiting the lowest Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the highest R-squared (R2) values. These metrics underscore RF's accuracy, robustness, and explanatory power in predicting PUBG player performance based on travel distance. Conversely, the SVM-Sigmoid (SVM-Sig) model consistently delivered poor predictive performance, with significantly higher MSE, RMSE, and a negative R2. This underscored the importance of selecting an appropriate model, as SVM-Sig's predictions deviated substantially from actual winning rates in both training and testing scenarios. In summary, our research elucidates the integral role of travel distance in the world of PUBG, shedding light on the pivotal connection between these features and player performance. Moreover, it reinforces the significance of data-driven decision-making in the context of gaming performance analysis and the critical importance of model selection in ensuring the accuracy and reliability of predictive models. With these insights, our research contributes to the evolving landscape of understanding player performance in the exciting and dynamic world of PUBG.

REFERENCES

[1] S. Putra Perkasa and D. Rahman Nur, "Effectiveness of Player Unknown Battleground (PUBG) Video Game in Improving Vocabulary," *Borneo Educ. J.*, vol. 2, no. 2, pp. 73–77, Aug. 2020, doi: 10.24903/bej.v2i2.629.

[2] M. G. Cobobi, "Antara Individualitas dan Kolektivitas: Studi Hubungan Pemain Gim PlayerUnknown's Battlegrounds (PUBG) Mobile," *J. Animat. Games Stud.*, vol. 6, no. 1, pp. 67–86, Apr. 2020, doi: 10.24821/jags.v6i1.3563.

[3] N. K. Hanafie, B. Bakhtiar, and D. Darmawati, "The Effect of Online Games on Changes in Student Behavior in Middle Schools," *Int. J. Adv. Sci. Educ. Relig.*, vol. 5, no. 2, pp. 48–59, 2022, doi: 10.33648/ijoaser.v5i2.178.

[4] A. Fauzi, "PENGARUH GAME ONLINE PUBG (Player Unknown's Battle Ground) TERHADAP PRESTASI BELAJAR PESERTA DIDIK," *ScienceEdu*, vol. II, no. 1, p. 61, Jul. 2019, doi: 10.19184/se.v2i1.11793.

[5] M. Ilham Rofiqi and H. Hindarto, "Analisis Kecanduan Game Player Unknown's Battlegrounds (PUBG) Mobile dengan Menggunakan Logika Fuzzy," *J. Inform. Polinema*, vol. 7, no. 2, pp. 97–102, Feb. 2021, doi: 10.33795/jip.v7i2.327.

[6] W. Wei, X. Lu, and Y. Li, "PUBG: A Guide to Free Chicken Dinner," pp. 3–8, 2018.

[7] D. Sen, R. K. Roy, R. Majumdar, K. Chatterjee, and D. Ganguly, "Prediction of the Final Rank of the Players in PUBG with the Optimal Number of Features," *Lect. Notes Electr. Eng.*, vol. 888, pp. 441–448, 2022, doi: 10.1007/978-981-19-1520-8_35.

[8] M. A. As'ari, N. F. Ghazali, and N. Sanat, "Esports Analytics on PlayerUnknown's Battlegrounds Player Placement Prediction using Machine Learning Approach," *Int. J. Hum. Technol. Interact.*, vol. 5, no. 1, pp. 17–28, 2021.

[9] M. Manjunath Mamulpet, "PUBG WINNER PLACEMENT PREDICTION USING ARTIFICIAL NEURAL NETWORK," *Int. J. Eng. Appl. Sci. Technol.*, vol. 3, no. 12, pp. 107–118, Apr. 2019, doi: 10.33564/IJEAST.2019.v03i12.020.

[10] K. F. Margolang, M. Zarlis, and Hartono, "Sentiment Classification on Mandalika MotoGP Event Using K-Means Clustering and Random Forest," in *2nd International Conference on Information Science and Technology Innovatin (ICoSTEC)*, 2023, pp. 65–69.

[11] S. Riyadi, Hartono, and Wanayumini, "Predicting Children's Talent Based On Hobby Using C4.5 Algorithm And Random Forest," in *International Conference on Information Science and Technology Innovation (ICoSTEC)*, Mar. 2023, pp. 182–186. doi: 10.35842/icostec.v2i1.54.

[12] S. K. Ghosh and F. Janan, "Prediction of student's performance using random forest classifier," *Proc. Int. Conf. Ind. Eng. Oper. Manag.*, pp. 7089–7100, 2021.

[13] R. Rismayati, I. Ismarmiaty, and S. Hidayat, "Ensemble Implementation for Predicting Student Graduation with Classification Algorithm," *Int. J. Eng. Comput. Sci. Appl.*, vol. 1, no. 1, pp. 35–42, 2022, doi: 10.30812/ijecsa.v1i1.1805.

[14] D. Pardede, Wanayumini, and R. Rosnelly, "A Combination Of Support Vector Machine And Inception-V3 In Face-Based Gender Classification," *Int. Conf. Inf. Sci. Technol. Innov.*, vol. 2, no. 1, pp. 34–39, Mar. 2023, doi: 10.35842/icostec.v2i1.30.

[15] N. A. Sinaga, B. H. Hayadi, and Z. Situmorang, "PERBANDINGAN AKURASI ALGORITMA NAÏVE BAYES, K-NN DAN SVM DALAM MEMPREDIKSI PENERIMAAN PEGAWAI," *J. Tek. Inf. dan Komput.*, vol. 5, no. 1, p. 27, Jul. 2022, doi: 10.37600/tekinkom.v5i1.446.

[16] Y. Pamungkas, A. D. Wibawa, and M. D. Cahya, "Electronic Medical Record Data Analysis and Prediction of Stroke Disease Using Explainable Artificial Intelligence (XAI)," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, no. 4, 2022, doi: 10.22219/kinetik.v7i4.1535.

[17] N. A. Sinaga, R. Ramadani, K. Dalimunthe, M. S. A. A. Lubis, and R. Rosnelly, "Comparison of Decision Tree, KNN, and SVM Methods for Determining Majors in Vocational Schools," *J. Sist. Komput. dan Inform.*, vol. 3, no. 2, p. 94, Dec. 2021, doi: 10.30865/json.v3i2.3598.

[18] M. Abdurohman and A. G. Putrada, "Forecasting Model for Lighting Electricity Load with a Limited Dataset using XGBoost," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 8, no. 2, pp. 571–580, 2023, doi: 10.22219/kinetik.v8i2.1687.

[19] H. M. Idan and K. Q. Hussein, "Comparison study between selected techniques of (ML , SVM and Deep Learning) regarding prediction of Flooding in Eastof Iraq," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 14, pp. 2893–2904, 2021, [Online]. Available: https://turcomat.org/index.php/turkbilmat/article/view/10797

[20] A. S. Hashim, W. A. Awadh, and A. K. Hamoud, "Student Performance Prediction Model based on Supervised Machine Learning Algorithms," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 928, no. 3, 2020, doi: 10.1088/1757-899X/928/3/032019.

[21] S. Bekesiene, R. Smaliukiene, and R. Vaicaitiene, "Using artificial neural networks in predicting the level of stress among military conscripts," *Mathematics*, vol. 9, no. 6, 2021, doi: 10.3390/math9060626.

[22] Kaggle, "PUBG Finish Placement Prediction (Kernels Only)," 2019. [Online]. Available: https://www.kaggle.com/competitions/pubg-finish-placement-prediction

[23] F. Chairani, "ANALISIS PENGARUH E-SERVICE QUALITY DAN HEDONIS TERHADAP E-SATISFACTION PADA GAME ONLINE PUBG MOBILE," STIKOM DINAMIKA BANGSA, 2019. [Online]. Available: http://repository.unama.ac.id/157/

[24] O. Ingkiriwang, V. K. D.D, and Y. Pasoreh, "ANALISIS TENTANG KECANDUAN GAME ONLINE PUBG MOBILE PADA REMAJA DI DESA TOUNELET KECAMATAN KAKAS," *ACTA DIURNA Komun.*, vol. 3, no. 2, 2021, [Online]. Available: https://ejournal.unsrat.ac.id/v3/index.php/actadiurnakomunikasi/article/view/33465

[25] P. M. E-Sport, "PUBG MOBILE Official Competition Rulebook," 2021. [Online]. Available: https://esports.pubgmobile.com/Documents/PUBGMOBILEOfficial

CompetitionRulebookV1.2.0clean_012821.pdf
[26] Z. Zhu, Y. Zhou, Y. Dong, and Z. Zhong, “PWLU: Learning Specialized Activation Functions with the Piecewise Linear Unit,” *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–19, 2023, doi: 10.1109/TPAMI.2023.3286109.
[27] S. Langer, “Approximating smooth functions by deep neural networks with sigmoid activation function,” *J. Multivar. Anal.*, vol. 182, p. 104696, Mar. 2021, doi: 10.1016/j.jmva.2020.104696.
[28] Q. Ai, S. Liu, L. He, and Z. Xu, “Stein Variational Gradient Descent with Multiple Kernels,” *Cognit. Comput.*, vol. 15, no. 2, pp. 672–682, 2023, doi: 10.1007/s12559-022-10069-5.
[29] M.-A. Schulz *et al.*, “Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets,” *Nat. Commun.*, vol. 11, no. 1, p. 4238, Aug. 2020, doi: 10.1038/s41467-020-18037-z.
[30] T. Hanika and J. Hirth, “Conceptual views on tree ensemble classifiers,” *Int. J. Approx. Reason.*, vol. 159, pp. 1–33, 2023, doi: 10.1016/j.ijar.2023.108930.
[31] D. Liu *et al.*, “Optimisation and evaluation of the random forest model in the efficacy prediction of chemoradiotherapy for advanced cervical cancer based on radiomics signature from high-resolution T2 weighted images,” *Arch. Gynecol. Obstet.*, vol. 303, no. 3, pp. 811–820, 2021, doi: 10.1007/s00404-020-05908-5.