Genre e-sport gaming tournament classification using machine learning technique based on decision tree, Naïve Bayes, and random forest algorithm

Arif Rinaldi Dikananda, Irfan Ali, Fathurrohman, Ade Rizki Rinaldi, Iin
Software Engineering Department, STMIK IKMI Cirebon, Cirebon, Indonesia

*rinaldi21crb@gmail.com

Abstract. The development of the game industry in this global era no longer presents entertaining games but also provides competitive games. With the recent competitive game, it can be categorized into a sport called e-Sport. Various game developers have also created e-Sport facilities and created a tournament to advance the industry. The increasing number of tournaments that are held in the field of sports from various types of games, it requires a classification for the types of games that are actively holding tournaments from the last few years. The classification used is Naïve Bayes, decision tree and random forest algorithm. Naïve Bayes has become one of the algorithms for data. Naïve Bayes is a classification system based on the theorem of Bayes. Naïve Bayes is a classification system based on the theorem of Bayes. It's also recognized that the Naïve Bayes Classifier is greater than certain other classification methods. As first, the main aspect of Naïve Bayes is a very good (naive) presumption of freedom from any situation or case. It's also recognized that the Naïve Bayes Classifier is greater than certain other classification methods. As first, the main aspect of Naïve Bayes is a very good (naive) presumption of freedom from any situation or case. Decision trees are also well machine learning algorithms used to solve complex classification problems. Decision Tree is a classification method for data mining that aims to predict the behavior of the database. The result from this research is Random forest accuracy 60%.

I. Introduction

In the last few years, the gaming industry has already become a highly profitable industry [1]. E-Sports represents regular sports with player endorsements and sponsorships, audiences packing into tournament venues, lucrative media deals, and college-level recruitment [2]. Playing video games is a leisure activity for most gamers, and can also be a highly earning profession for those who master their match skills and become e-Sports players supported with well corporations [3]. E-Sports can be classified into multiple genres such as Multiplayer Online Battlefield Arena (MOBA), First Person Shooters (FPS), Real Time Strategy (RTS) and Sport Simulation [4]. However, with so many e-Sports game tournaments required classification with machine learning techniques base on However with so many e-sports game tournaments required classification techniques using Naïve Bayes, Decision Tree and Random Forest algorithm

Any of the classification algorithms most have a Naïve Bayes (Naïve Bayes) [5]. The Naïve Bayes Classifier is a classification system based on Bayes' theorem. Naïve Bayes believed the classification for Naïve Bayes was greater than that for many alternative classification methods [6]. Because of its ease in allowing any function to play a role in the final decision, Naïve Bayes is a common approach to engineering learning applications. Simplicity is the same of computational efficiency and this makes the Naïve Bayes algorithm involved in being suitable in a number of areas [7].

The decision tree is a methodology that can reliably identify and/or create regulations and can typically be used for information not identified [8]. Decision trees are most frequently deployed as advanced methods that are widely used and highly important in a number of ways in the area of artificial
intelligence [9]. Back to top A binary decisions branch, splits the information on a selection (parent node) into a subset or subset (child node) by computing the best segregated function set by a scattered selected criterion. There are also no conclusions as to how the particular state of value [10]. The typical decision tree for categorization enforces the basic strategy of choosing between multiple routes by a branch node, however, in short, for instance, is a branching rule with the classification ones[11]. The generic decision tree making is a permanent project, in which the mechanism is reciprocally configured and training data have been partitioned until the state of termination is reached from root point to core node [12].

The Method of a Decision Tree shall constitute models, taken by decision with trumped values, based on attribute information as provided in data. Beyond making a forecast for the document issued, the Decision Tree. They are both fast and precise to ensure that they are both favored in the world of machine learning as well[13]. Decision trees can also manage complex definition, whose definitions can easily be expressed in a feature space type, and referenced technique(s). In several, other words, tree naming may be the feature that maps values (or character values) in the attribute of the decision (or, in other words, in the class labels) of a group of decision classes that reflect an allowable hypothesis [14][15].

2. Methodology
2.1 Dataset
Object classifies of e-sport using gamma algorithm method, which consisted of four genre series of a maximum of 170 of data set. Data set received from a scope of 170 between 1997 and 2020, having an id of Attribute, gender, total income, internet benefit, general pay and all convert table 1 were shown in the figure 1.

| No | Game               | Genre  | Total Earnings | Online Earning | Total Players | Total Tournaments |
|----|--------------------|--------|----------------|----------------|---------------|------------------|
| 1  | Age of Empires     | Strategy | 190075.61   | 33235.72       | 246           | 95               |
| 2  | Age of Empires II  | Strategy | 1480089.01 | 862909.14      | 827           | 549              |
| 3  | Age of Empires III | Strategy | 66463.85    | 24963.85       | 86            | 72               |
| 4  | Age of Empires Online | Strategy | 2668    | 2668           | 16            | 7                |
| 5  | Age of Mythology   | Strategy | 52360     | 360            | 22            | 6                |
| ...|                    |         |              |                |               |                  |
| 170| WWE 2K20           | Fighting Game | 50000      | 50000          | 20            | 2                |

Figure 1. Grouping of e-Sport Gaming Genre Dataset
2.2 E-Sport Gaming Tournament
E-sports is the term used to describe video games that are played competitively and watched by, usually large, audiences. E-sports is an important research field for academia and industry only in terms of size. Goldman Sachs predicted a compound annual growth rate of 22 per cent with a market value of $1.1 billion by 2019 and an estimated 330 million spectators in Super data by 2019 [16]. E-Sports players, such as traditional athletes, have been practicing for long hours and are therefore vulnerable to the negative health effects of prolonged sitting. There is a lack of research on physical activity and the health consequences of prolonged sitting by competitive players.

The goal of this review was to evaluate activity levels, body mass index (BMI) and activity levels in collegiate professional athletes as especially in comparison to age-matched controls. E-Sport, managed video game competitions, is universally acknowledged by the industry as a sport of entertainment. Competitive gaming has rapidly been implemented with the organization of nationally and internationally sports leagues[17].

2.3 Research Method
In the research and decision tree algorithms, Naive Bayes, random forest were used the technique used in a system of mining information and the pre-processing method was achieved by adjusting the representation of data, data preparedness, patterning and assessment. Optimize the use of algorithm choking parameters by choosing only one sub process that selects one. Operator The object of operator is to select one set sub process algorithm with a reliability of another, using some automatic

a. Optimize Parameters
Setting model categories for e-sport gaming using algorithm techniques by optimize parameters as seen in Fig 2.

![Figure 2. Optimize Parameter](image)

Dataset applied a prior information reproprocessing feature consisting of the filter data set by filtering of empty / incomplete data and optimizes the output options of the sub process linear scale minimum at least of maximum rating is stage 3 Sta. The handling of errors that have parameters make it possible to choose a type of execution error when walking set files by mistake that indicates errors, then the executing process is terminated by showing error messages.

b. Select Sub proses
E-sport gaming with algorithm technique is used to optimize the parameter to be introduced in election model categories. Operator is selecting sub method. Operator options a technique to be used in automation electoral, as illustrated in figure 3 of the Table 3 for Processing that is the best algorithm to be used in Engineering Machine Learning

![Figure 3. Cross validation](image)
c. Algorithm

1. Decision Tree

Recursive partitioning was used to create a decision tree. Divide and conquer is a common name for this system, since it uses the characteristic values to break the information into smaller subsets of the same class. Also C4.5 will deal after construction with separate and ongoing details, missing numbers and tree plashing. C4.5 relies on entropy and acquiring ration to pick the best attribute as the root and go on further splitting functions[13].

2. Naïve Bayes

Bayesian theorem provides an equation for calculating posterior probability \( P(c|x) \) from \( P(c) \), \( P(x) \) and \( P(x|c) \):

\[
p(c \mid x) = \frac{p(x \mid c)p(c)}{p(c)}
\]

\( P(c|x) \): the posterior probability of class (c, target) given predictor (x, attributes).

\( P(c) \): the prior probability of class.

\( P(x|c) \): the likelihood, which is the probability of predictor given class.

\( P(x) \): the prior probability of predictor

3. Random Forest

Random Forest has the concept of the actual forest, which says that the more study trees in the forest would be. In addition, if it has a higher number of trees in the forest, Random Forest will provide the best precision. Random Forest does not pay attention to the number of forest trees, can handle missing values, can never over fit, and can handle categorical data as well. Random Forest utilizes the Gini Index measurement to calculate the purity and impurity of the features[13].

3. Result and discussion

3.1 Evaluation

Then its accuracy can be evaluated after information processed to see the output of determining trees, Naïve Bayes and random forest algorithms. The research work include e-sport classifications, using a gaming algorithm system and testing the quality and accuracy of e-sport using gameing technology algorithm used. The accuracy level was checked using the confusion matrix given by accuracy of 59.41% accuracy for Figure 4.

Therefore, to recognize accuracy, applied optimize parameter in the algorithm set by using subprocess operator of a figure 5 acquired the accuracy are Decision Tree 0.571, Random Forest 0.600, and Naïve Bayes 0.471. It is most important for its accuracy by using the random forest algorithms of random forest 0.600 algorithms.

Figure 4. Confusion Matrix
By adding an algorithm for decision tree, Naive Bayes and modification, the operator chooses the algorithms with the best 60% accuracy and value 0.406 is always activated in this approach.

4. Knowledge Presentation
A selection of the study data using a random forest algorithm involving characteristics, total online wages, total revenues of players and total earnings. Tournaments Figure 7 shows that scattering the outcome of a forecast provides visible proof of any future value belonging to the target specification of the e-Sport gaming specification based on practical values predictor. The numbers include total earnings; online and overall player earnings;

Statistic prediction e-sport gaming using algorithm definition based on attribute total benefit, digital salaries, total players and the overall tournament numbers can be viewed in Figure 8 as predicting the first-person shooter genre as much as 48.24%, a predicting fighting game as little as 41.76%, while a predicting online match against 7.06% and strategy of 2.94%.

---

**Figure 5.** Optimize parameters

**Figure 6.** Graph Scatter prediction genre e-Sport Gamming

**Figure 7.** Graph Classification e-Sport Gamming Using Algorithm Technic
Research which has undertaken by others a restriction on the level of training of data as well as on the attribute prediction and which do not go well, affects the value of the basic generated by internal model classifications in every genre to thoroughly assess how accurate and precise the use of gaming technique algorithms from e-ticketing sports are these investigations. More data training more data as well as more variable Attribute predictor is expected to be used in the next study and provided with the trimming of trees method (pruning) to increase the degree of accuracy and not all problems or circumstances needed to be done by an unspecified algorithm for data mining to render data mining algorithms used very accurate. Therefore, in order to make sure the algorithm is very precise, some algorithm comparison tried is needed.

5. Conclusion

Paper explains about this form using gaming system and the algorithms automatic from sport e-ticketing using a gaming ticket procedure. After the operator is selecting sub projections using methods designed to optimize parameters, computer training has gained one optimization at exceptional values precision. Results of procedures parameter for optimizing sub process in the operator selection algorithm decision tree, Naive Bayes, random forest obtained by the random algorithm random forest in that the performer has the greatest precision than Naive Bayes algorithm or tree choice algorithm random.

References

[1] M. Rosell Llorens, “eSport Gaming: The Rise of a New Sports Practice,” *Sport. Ethics Philos.*, vol. 11, no. 4, pp. 464–476, 2017, doi: 10.1080/17511321.2017.1318947.
[2] V. Hodge, S. Devlin, N. Sephton, F. Block, P. Cowling, and A. Drachen, “Win Prediction in Multi-Player Esports: Live Professional Match Prediction,” *IEEE Trans. Games*, vol. PP, pp. 1–1, 2019, doi: 10.1109/tg.2019.2948469.
[3] F. Bányai, Á. Zsila, M. D. Griffiths, Z. Demetrovics, and O. Király, “Career as a Professional Gamer: Gaming Motives as Predictors of Career Plans to Become a Professional Esport Player,” *Front. Psychol.*, vol. 11, no. August, pp. 1–9, 2020, doi: 10.3389/fpsyg.2020.01866.
[4] E. Nagorsky and J. Wiemeyer, *The structure of performance and training in esports*, vol. 15, no. 8 August 2020. 2020.
[5] H. Zhou, *Learn Data Mining Through Excel*. 2020.
[6] N. Salmi and Z. Rustam, “Naïve Bayes Classifier Models for Predicting the Colon Cancer,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 546, no. 5, 2019, doi: 10.1088/1757-899X/546/5/052068.
[7] A. P. Wisbawa et al., “Naïve Bayes Classifier for Journal Quartile Classification,” *Int. J. Recent Contrib. from Eng. Sci. IT*, vol. 7, no. 2, p. 91, 2019, doi: 10.3991/ijes.v7i2.10659.
[8] R. Buaton et al., “Decision Tree Optimization in Data Mining with Support and Confidence,” *J. Phys. Conf. Ser.*, vol. 1255, no. 1, 2019, doi: 10.1088/1742-6596/1255/1/012056.
[9] A. Trabelsi, Z. Elouedi, and E. Lefevre, “Decision tree classifiers for evidential attribute values and class labels,” *Fuzzy Sets Syst.*, vol. 366, pp. 46–62, 2019, doi: 10.1016/j.fss.2018.11.006.
[10] T. Shaikhjina, D. Lowe, S. Daga, D. Briggs, R. Higgins, and N. Khovanova, “Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation,” *Biomed. Signal Process. Control*, vol. 52, pp. 456–462, 2019, doi: 10.1016/j.bspc.2017.01.012.
[11] S. D. Brown and A. J. Myles, *Decision Tree Modeling*, Second Edi., vol. 3. Elsevier, 2020.
[12] J. Yan, Z. Zhang, K. Lin, F. Yang, and X. Luo, “A hybrid scheme-based one-vs-all decision trees for multi-class classification tasks,” *Knowledge-Based Syst.*, vol. 198, p. 105922, 2020, doi: 10.1016/j.knosys.2020.105922.
[13] M. B. Nugroho, “Machine Learning For Beginners Algorithms, Decision Tree & Random Forest Introduction,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.
[14] J. Kozak, *Studies in Computational Intelligence 781: Decision Tree and Ensemble Learning Based on Ant Colony Optimization*. 2019.
[15] O. Nurdianwan, F. A. Pratama, D. A. Kurnia, Kaslani, and N. Rahaningsih, “Optimization of
Traveling Salesman Problem on Scheduling Tour Packages using Genetic Algorithms,” J. Phys. Conf. Ser., vol. 1477, no. 5, pp. 0–6, 2020, doi: 10.1088/1742-6596/1477/5/052037.

[16] D. C. Funk, A. D. Pizzo, and B. J. Baker, “eSport management: Embracing eSport education and research opportunities,” Sport Manag. Rev., vol. 21, no. 1, pp. 7–13, 2018, doi: 10.1016/j.smr.2017.07.008.

[17] J. DiFrancisco-Donoghue, W. G. Werner, P. C. Douris, and H. Zwibel, “Esports players, got muscle? Competitive video game players’ physical activity, body fat, bone mineral content, and muscle mass in comparison to matched controls,” J. Sport Heal. Sci., 2020, doi: 10.1016/j.jshs.2020.07.006.