Outcome Prediction Using Naïve Bayes Algorithm in The Selection of Role Hero Mobile Legend

A S Chan1*, F Fachrizal1* and A R Lubis1*
1Department of Information and Computing, Politeknik Negeri Medan, Medan 20155, Indonesia.

*Email: chakadoang@gmail.com, ferrypolmed@gmail.com, arifridho@polmed.ac.id

Abstract. According to the e-Marketer market research institute, the net population of the country reached 83.7 million people in 2014. With this amount, Indonesia has been included in the 20 most internet user countries in the world. MOBA is currently in great demand. The high level of competition of this type of MOBA game attracts a lot of game players' attention to tournaments or competitions officially both regionally and internationally. In 2013 there were around 71.5 million e-Sports enthusiasts around the world. Every year the number of Indonesian gamers is estimated to increase by around 33 percent. With a growth of 33 percent, this certainly makes a great opportunity for game developers. Currently, the games circulating in Indonesia is mostly from foreign countries such as America, Japan, and South Korea. Each hero has a different role (ability). The foresight of players and the ability of individual skills are determined by selecting the composition of hero roles on a team to achieve victory in the legendary mobile gameplay. Naïve Bayes Classifier is the simple Statistical Bayesian Classifier. Naïve Bayes is able to classify role heroes well enough to predict the victory based on the role hero chosen by the player. Naïve Bayes implementation can explore the characteristics of the attribute and dataset of the chosen role hero. Naïve Bayes algorithm method can predict the Victory in mobile legend gameplay from existing hero roles based on datasets and training data.

1. Introduction
According to the e-Marketer market research institute, the net population of the country reached 83.7 million people in 2014 [1]. With this amount, Indonesia has been included in the 20 largest internet user countries in the world. It can be ascertained that the number of internet users will continue to grow along with the increasingly cheap and easy internet connection, the spread of networks, and the ease of obtaining hardware to access the internet such as computers, laptops, and smartphones.

The type of game that is currently in great demand is MOBA [2]. MOBA (Massive Online Battle Arena) is a type of online game that combines two types of games, namely RTS (Real-Time Strategy) and RPG (Role Playing Game) [3] where players run one character from two teams opposite in order to destroy the opponent's base. Each character played has a role with the strengths and weaknesses of each so it is required to cooperate with team members to win the game [4]. Examples of games of this type include World of Warcraft, DotA, League of Legends, Mobile Legends, Vainglory, etc [5].

The high level of competition from this type of MOBA game attracts a lot of attention from game players until it is often held in official tournaments or competitions both regionally and internationally. This competition is called the e-Sports tournament [6]. in this tournament players can get a variety of attractive prizes from game merchandise to cash. In 2013 there were around 71.5 million e-Sports enthusiasm in worldwide. Every year, the number of Indonesian gamers is estimated to increase by around 33 percent [7]. With a growth of 33 percent, this certainly makes a great opportunity for game
developers. Actually, in Indonesia, there are very many game developers, there are around 80 game developers in Indonesia.

Currently, the most popular online game for kids, teens to adults is an Android-based mobile legend game. Mobile Legend: Bang Bang is a game (online Multiplayer Battle Arena) MOBA on Android that is similar to DOTA in PC games [9]. For gamers who have played DOTA 2 in Indonesia Mobile Legend: Bang-bang seems to be increasingly popular.

The Naive Bayes classifier is well studied [10]. Naïve Bayes’ is a conditional probability model. Despite its simplicity and strong assumptions, the naïve Bayes classifier has been proven to work satisfactorily in many domains [11]. Bayes’ rule provides a framework for incorporating judgment (prior beliefs) with observational data [12].

2. Method
2.1. Classification
One of the tasks that can be done with data mining is classification [13]. The classification was first applied to plants that classify a particular species, such as those carried out by Carolus von Linne (also known as Carolus Linnaeus) who first classified species based on physical characteristics. Furthermore, he is known as the father of classification [14].

In the classification, there are target variable categories. The methods/models that have been developed by researchers to resolve classification cases

2.2. Algorithm Naïve Bayes
The Naive Bayes Classifier technique is based on Bayesian Theorem and it is used when the dimensionality of the inputs is high. The Naive Bayes method is a method that can be used to classify documents statistically [15]. This method can predict the probability of membership of a class from data. Naive Bayes is one of the most effective and efficient inductive learning algorithms for machine learning and data mining. The independence of this attribute in actual data is rare, but even though the independence assumption of these attributes is violated the performance of classifying naïve Bayes is quite high [16], this is evidenced in various empirical studies

Opportunity $p (C = ci | X = xj)$ [17] shows the probability of the Xi attribute with the value xi given class c, wherein Naive Bayes, class C is of the qualitative type while attribute Xi can be of a qualitative or quantitative type.

When the Xi attribute is of quantitative type, the probability of $p (X = xi | C = cj)$ will be very small so that the opportunity equation cannot be relied upon for quantitative type attribute problems. So to deal with quantitative attributes, there are several approaches that can be used such as normal distribution [18].

$$ F = N(Xi; \mu_c, \sigma_e) = \frac{1-(X1-\mu_c)^2}{\sqrt{2\mu_c\sigma_e}} \tag{1} $$

Numerical values will be mapped to values in the form of intervals, Naïve Bayes' calculations depicted as follows [19].

$$ \rho(I = ij | C = cj) = \frac{p(I = Ij) p(C = ci | I = Ij)}{P(C = Ci)} \tag{2} $$

Explanation: $p(I=ij|C=ci)$ : interval opportunities i-j for ci class
$p(C=ci|I=ij)$ : ci class opportunities for i-j interval
$p(I=ij)$ : probability interval of j on all formed interval
$p(C=ci)$ : chance of an -i class for all classes in the dataset
The advantage of naïve Bayes is it requires short computational time for training. And improves the classification performance by removing the irrelevant features. And then it has good performance. And then less accurate as compared to other classifiers on some datasets [20].

2.3. Game

The game comes from English words that have the basic meaning of the game. The game, in this case, refers to the notion of intellectual playability. Games can also be interpreted as the arena of the decisions and actions of their players, there are targets that the players want to achieve. Intellectual agility at a certain level is a measure of the extent to which the game is interesting to play optimally [21]. The game also significantly sharpens the power of analysis of its users to process information and make quick decisions that are accurate.

3. Research Result

3.1. Training Data

Naïve Bayes calculations are done by counting 36 hero data, the criteria specified in this victory prediction include:
- Role, The role is a type of hero who uses melee weapons or long-range weapons
- Type, The type of hero is based on the type of power possessed such as the type of fighter, witch, defense and others
- Hero Statistics, The default ability level of each hero that has different values according to the type of hero used
- HP, The amount of health power of each hero varies according to the type of hero used as tanks have more HP than wizards

Facts for probability:

| There is data | = 36 |
|--------------|------|
| P(Y = WIN)   | = 19/36 = 0,527 |
| P(Y = LOSE)  | = 17/36 = 0,472 |

Fact:
- P(FIGHTER|Y = WIN) = 7/36 = 0,194
- P(FIGHTER|Y = LOSE) = 7/36 = 0,194

Do all calculations in the same way with formulas. After displaying statistical facts data from the 36 data sets then the next step is to create a dataset for training from the 36 data above as in the following table: data

3.2. Training Data

The next prediction data is based on the data set above using the naïve Bayes algorithm

\[ \rho(Y | X) = \frac{p(X | Y)p(Y)}{P(X)} \]  \hspace{1cm} (3)

Then specify what attributes the data set has as shown in the following table:

| id Attribute | name Attribute | stat Attribute |
|--------------|----------------|----------------|
| 1            | TYPE           | Discovered     |
| 2            | ROLE           | Discovered     |
| 3            | STAT           | Discovered     |
| 4            | HP             | Discovered     |
| 5            | RESULT         | To Find        |
After the attributes on the mobile legend game are set, the next step is to give an id rating on each attribute so as to make it easier to do calculations using the naïve Bayes algorithm. The next step is to make predictions based on existing training data. Prediction is done 3 times randomly.

**Table 2. Prediction.**

| TYPE   | ROLE     | Stat | HP  | Result |
|--------|----------|------|-----|--------|
| MELEE  | FIGHTER  | LOW  | FAT | ??     |
| RANGE  | MAGE     | HIGH | THIN| ??     |
| MELEE  | TANK     | HIGH | THIN| ??     |
| RANGE  | MARKSMAN | LOW  | THIN| ??     |

Calculating the probability uses MELEE type with the FIGHTER role, having LOW statistics with the hp FAT.

**Table 3. Prediction hero 1.**

| Attribute | Point | Result |
|-----------|-------|--------|
| TYPE      | MELEE | WIN    | 6 11 | 0.5454 |
| ROLE      | FIGHTER | WIN | 3 11 | 0.2727 |
| STAT      | Low   | WIN | 5 11 | 0.4545 |
| HP        | FAT   | WIN | 6 11 | 0.545 |
| Result    | WIN   | 0.03688 | 11 | 0.4057 |
| Result    | LOSE  | 0.00439 | 8  | 0.0351 |

Calculating the probability of WINNER uses type RANGE with MAGE role, having HIGH statistics with hp THIN.

Based on the results of the above calculation for the prediction of the selection of hero types with RANGE with the MAGE role, having HIGH statistics with a THIN cellphone is to LOSE with a percentage of 0.292 for LOSE and 0.112 for WIN.

**Table 4. Prediction hero 2.**

| Attribute Discovered | Point Attribute | Result |
|----------------------|-----------------|--------|
| TYPE                 | RANGE           | WIN    | 5 11 | 0.45454 |
| ROLE                 | RANGE           | LOSE   | 5 8  | 0.625   |
| ROLE                 | MAGE            | WIN    | 2 11 | 0.18181 |
| ROLE                 | MAGE            | LOSE   | 3 8  | 0.375   |
| STAT                 | High            | WIN    | 3 11 | 0.27272 |
| STAT                 | High            | LOSE   | 2 8  | 0.25    |
| HP                   | THIN            | WIN    | 5 11 | 0.45454 |
| HP                   | THIN            | LOSE   | 5 8  | 0.625   |
| Result               | WIN             | 0.0102452 | 11 | 0.112697 |
| Result               | LOSE            | 0.0366210 | 8  | 0.292968 |
Calculates the probability of WIN uses MELEE type with TANK role, has LOW statistics with hp FAT. Based on the results of the above calculations for the prediction of the selection of the hero with the MELEE type with the TANK role, having a LOW statistic with the FAT is a victory with a percentage of 0.405 for WIN and 0.070 for LOSE.

Table 5. Prediction hero 3.

| Attribute | Point Attributes | Result |
|-----------|------------------|--------|
| TYPE      | MELEE WIN        | 6 11   |
|           | MELEE LOSE       | 3 8 375|
| ROLE      | TANK WIN         | 3 11 2722|
|           | TANK LOSE        | 2 8 25|
| STAT      | Low WIN          | 5 11 4544|
|           | Low LOSE         | 2 8 25|
| HP        | FAT WIN          | 6 11 4545|
|           | FAT LOSE         | 3 8 375|
| Result    | WIN              | 0.0368827 11 0.4057099|
|           | LOSE             | 0.0087890 8 0.0703125|

Calculating the probability of winning with the type RANGE with the role of MARKSMAN, having LOW statistics with THIN.

Based on the results above for the prediction of the results of the selection of hero with type RANGE with the role of MARKSMAN, having LOW statistics with a THIN cellphone is a victory with percentage 0.281 for WIN and 0.195 for LOSE.

Table 6. Prediction hero 4.

| Attribute | Point Attributes | Result |
|-----------|------------------|--------|
| TYPE      | RANGE WIN        | 5 11 4545|
|           | RANGE LOSE       | 5 8 625|
|           | MARKSMAN WIN     | 3 11 2722|
| ROLE      | MARKSMAN LOSE    | 2 8 25|
| STAT      | Low WIN          | 5 11 4544|
|           | Low LOSE         | 2 8 25|
| HP        | THIN WIN         | 5 11 4545|
|           | THIN LOSE        | 5 8 625|
| Result    | WIN              | 0.0256130 11 0.281743|
|           | LOSE             | 0.0244140 8 0.19531|

The results obtained that predicted calculations for wins and losses based on the chosen role hero are for the hero with the role fighter having the highest prediction rate of victory, while the role hero mage, the player is predicted to experience defeat, then for the predicted role hero tank and marksman will win even with smaller presentations.

4. Conclusion

The conclusion of the prediction of choosing a hero role on a mobile legend game using the naïve Bayes algorithm is quite good and in accordance with predictions, and can determine the possibility of the hero's WIN or LOSE. And the application of the Naïve Bayes algorithm method can predict the Victory in mobile legend gameplay from existing hero roles based on datasets and training data so that they can predict systematically. so that players can practice weaker heroes and have a low winning presentation, so opponents will have difficulty predicting the weakness of the legendary mobile player.
References

[1] A. Drachen et al., “Skill-Based Differences in Spatio-Temporal Team Behaviour in Defence of The Ancients 2 (DotA 2),” vol. 2, no. DotA 2.

[2] T. E. Batsford, “Calculating Optimal Jungling Routes in DOTA2 Using Neural Networks and Genetic Algorithms,” 2013.

[3] A. Tlili, F. Essalmi, L. Jemni, B. Ayed, and M. Jemni, “A smart educational game to model personality using Learning Analytics,” 2017.

[4] K. Song, T. Zhang, and C. Ma, “Predicting the winning side of DotA2,” pp. 2–5.

[5] K. Wang and W. Shang, “Outcome prediction of DOTA2 based on Naïve Bayes classifier,” Proc. - 16th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2017, no. 1994, pp. 591–593, 2017.

[6] C. Eggert, M. Herrlich, J. Smeddinck, and R. Malaka, “Classification of Player Roles in the Team-Based Multi-player Game Dota 2,” pp. 112–113, 2015.

[7] Asosiasi Penyelenggara Jasa Internet Indonesia - APJII, Profil Pengguna Internet Indonesia 2014, 2014.

[8] Muhammad Abdul Basith A., “1.”

[9] P. Yang and D. L. Roberts, “Identifying Patterns in Combat that are Predictive of Success in MOBA Games.”

[10] J. D. M. Rennie, L. Shih, J. Tjeevan, and D. R. Karger, “Tackling the Poor Assumptions of Naive Bayes Text Classifiers,” no. 1973, 2003.

[11] L. Dey, S. Chakraborty, A. Biswas, B. Bose, and S. Tiwari, “Sentiment Analysis of Review Datasets Using Naive Bayes’ and K-NN Classifier,” Int. J. Inf. Eng. Electron. Bus., vol. 8, no. 4, pp. 54–62, 2016.

[12] J. Karandikar, T. Mcleay, and S. Turner, “Tool wear monitoring using naive Bayes classifiers,” pp. 1613–1614, 2015.

[13] A. S. Chan, “Jurnal Ilmiah Informatika (JIF) Prediksi Kedatangan Wisatawan Pada Pariwisata Kota Batam Dengan Menggunakan Teknik Knowledge Data Discovery,” pp. 1–6, 2018.

[14] N. R. Fatahillah, “Implementation Of Naive Bayes Classifier Algorithm On Social Media (Twitter) To The Teaching Of Indonesian Hate Speech,” pp. 128–131, 2017.

[15] Ö. F. Arar and K. Ayan, “A Feature Dependent Naive Bayes Approach and Its Application to the Software Defect Prediction Problem,” Appl. Soft Comput. J., 2017.

[16] A. Choi, “Structured Features in Naive Bayes Classification,” pp. 3233–3240.

[17] H. X. Goulart, M. D. L. Tosi, D. Soares-gonc, and R. F. Maia, “Hybrid Model For Word Prediction Using Naive Bayes and Latent Information,” pp. 1–16.

[18] S. G. M. R. Meybodi and M. D. M. M. Ebadzadeh, “Bayesian network structure training based on a game of learning automata,” Int. J. Mach. Learn. Cybern., 2016.

[19] M. Fatmawati and I. Mubarog, “Naive bayes algorithm performance for smartphone sentiment analysis in social media,” vol. 1, no. 2, pp. 76–81, 2020.

[20] S. D. Jadhav and H. P. Channe, “Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques,” Int. J. Sci. Res., vol. 5, no. 1, pp. 1842–1845, 2016.

[21] C. Botella et al., “Computers in Human Behavior Treating cockroach phobia using a serious game on a mobile phone and augmented reality exposure: A single case study,” Comput. Human Behav., vol. 27, no. 1, pp. 217–227, 2011.