International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 3, June 2024, pp. 3277~3286
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i3.pp3277-3286

Fine tuning attribute weighted naïve Bayes model for detecting anxiety disorder levels of online gamers

Anastasya Latubessy[1,4], Retantyo Wardoyo[2], Aina Musdholifah[2], Sri Kusrohmaniah[3]
[1]Doctoral Program Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta, Indonesia
[2]Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta, Indonesia
[3]Department of Psychology, Faculty of Psychology, Universitas Gadjah Mada, Yogyakarta, Indonesia
[4]Department of Informatics Engineering, Faculty of Engineering, Universitas Muria Kudus, Kudus, Indonesia

Article Info

Article history:

Received Feb 7, 2024
Revised Feb 23, 2024
Accepted Feb 25, 2024

Keywords:

Anxiety disorder
Attribute weighted
Fine tuning
Naïve Bayes
Online gamer

ABSTRACT

This research applies the fine tuning attribute weighted naïve Bayes (FTAWNB) model using ordinal data. It is known that in previous research, the FTAWNB model outperformed its competitors on the dataset used. However, the FTAWNB model has not been applied in the mental health domain that uses ordinal data. Therefore, this research used the anxiety gamers dataset to test the fine-tuning attribute weighted naïve Bayes (FTAWNB) model. Anxiety disorders are mental health disorders that can indicate the emergence of a gaming disorder. Gamers can experience anxiety disorders classified into four classes, namely minimal, mild, moderate, and severe anxiety. Then compare the results by FTAWNB obtained with three other naïve Bayes algorithms, namely Gaussian naïve Bayes, multinomial naïve Bayes, and categorical naïve Bayes, using the same dataset. Model performance is measured based on accuracy, precision, recall, and processing time. The test results show that the FTAWNB outperforms the other three models' accuracy, precision, and recall, with an accuracy value of 99.22%. While the accuracy of Gaussian NB is 91.132%, Categorical is 91.592%, and multinomial naïve Bayes is 61.104%. However, the FTAWNB takes slightly longer than the other three models' processing time. The FTAWNB takes 0.07 seconds to build the model and 0.05 seconds to test the model on training data.

Corresponding Author:

Retantyo Wardoyo
Department of Computer Science and Electronics, Universitas Gadjah Mada
Building C, 4th Floor, Sekip Utara, Bulaksumur, Senolowo, Sinduadi, Mlati District, Sleman Regency, Yogyakarta Special Region 55281, Indonesia
Email: rw@ugm.ac.id

1. INTRODUCTION

Everyone loves games because they are fun. However, in 2018 the World Health Organization (WHO) determined gaming disorder as a psychiatric disorder. Gaming disorder occurs when there are uncontrolled game patterns, marked by: i) control disorders for the time used to play games, ii) cannot control the priority of life, and iii) no matter the adverse effects [1]. Playing games can also be an escape because it is more fun in cyberspace, where many players exist [2]. Therefore, experiencing difficulties expressing emotions can also be associated with internet gaming disorder [3]. Some studies assume that the high prevalence of anxiety disorders is related to Internet gaming disorder [4]–[6]. The inability to control excessive worry indicates general anxiety disorder (GAD) [7].

Several studies in computer science raise cases of anxiety disorders to be solved using computer science-based models [8]–[11]. A study used questionnaire data on four machine learning classifiers for the prediction of anxiety and depression. The four machine learning classifiers used are decision tree, random forest, naive Bayes and linear regression. In this research, decision tree provides the best accuracy on the dataset used [8]. Meanwhile, a study used a dataset of anxiety disorders in online gamers from Kaggle on nine machine learning classifiers. In this research, multi-layer perceptron provided the best accuracy while gaussian naive Bayes produced the lowest accuracy compared to the other nine classifiers on the dataset used [9]. A different study used deep learning for the classification of anxiety, depression and comorbidities. The data used in this research is a text dataset from Reddit [10]. Another study compared auto-sklearn, naive Bayes and logistic regression models to predict mood and anxiety disorders. This research concludes that machine learning methods are more suitable for complex datasets [11]. One reliable classification model often used to analyze and identify psychological problems is naïve Bayes [8], [9], [11]–[15]. However, naïve Bayes does not provide the best accuracy for classifying anxiety disorders in gamers compared to the other eight classification algorithms. It is well known that naïve Bayes has a weakness in conditional independence [16]. Many researchers have attempted to fix these weaknesses. Some of the naïve Bayes model developments only emphasize attribute weighting to improve performance [17]–[20]. On the other hand, some only use the fine-tuning process to enhance the performance of naïve Bayes [21], [22].

Meanwhile, the fine tuning attribute weighted naïve Bayes (FTAWNB) model is one of the developments of the naïve Bayes model, which improves the performance of naïve Bayes by combining the concepts of attribute weighting and fine-tuning [23]. These two things are considered equally important in improving the performance of naïve Bayes. It is known that naïve Bayes (NB) uses (1) and (2) to estimate the probability of its class membership and predict its class label. Where *C* is the set of all possible c class labels, m is the number of attributes; A_j is the value of the jth attribute A_j of x, $P(c)$ is the prior probability of class c, and $P(a_j|c)$ is the conditional probability of $Ajaj$ which is in class c, which can be estimated by (3) and (4). Where q is the number of classes, n is the number of training instances, n_j is the total value of the j_{th} attribute A_j, ci is the class label of the th training instance i, $a_{i,j}$ is the value of the j_{th} attribute of the I_{th} training instance, and the indicator function $\delta(x, y)$ is one if $x = y$ and zero otherwise.

$$P(c|x)_{NB} = \frac{P(c)\prod_{j=1}^{m} P(a_j|c)}{\sum_{c\in C} P(c)\prod_{j=1}^{m} P(a_j|c)} \quad (1)$$

$$C(x)_{NB} = argmax_{c\in C} P(c|x) \quad (2)$$

$$P(c) = \frac{\sum_{i=1}^{n} \delta(ci,c) + \frac{1}{q}}{n+1} \quad (3)$$

$$P(a_j|c) = \frac{\sum_{i=1}^{n} \delta(aij,aj)\delta(ci,c) + \frac{1}{nj}}{\sum_{i=1}^{n} \delta(ci,c)+1} \quad (4)$$

The difference between the NB standard and the FTAWNB lies in the formula for calculating the conditional probability $P'(a_j\,|c)$ shown in (5) and (6), where the FTAWNB will be given attribute weights and fine tuning.

$$P(c|x)_{FTAWNB} = \frac{P(c)\prod_{j=1}^{m} P\prime(a_j|c)}{\sum_{c\in C} P(c)\prod_{j=1}^{m} P\prime(a_j|c)} \quad (5)$$

$$C(x)_{FTAWNB} = argmax_{c\in C} P(c|x) \quad (6)$$

The FTAWNB model has been compared with four competitors, such as a correlation-based featured weighting filter for NB (CAWNB) [18], fine tuned naïve Bayes (FTNB) [21], boosted naïve Bayes [24], and standard naïve Bayes (NB) [25]. Tested using 60 datasets from the UCI repository and obtained an outstanding accuracy value compared to its competitors on the dataset used of 98.04% and followed successively by CAWNB with an accuracy of 97.95%, FTNB with 97.87%, and NB with 97.84.

This study investigated the effects of the FTAWNB model on ordinal data. While earlier studies have explored the impact of FTAWNB on various data sets, they have not explicitly addressed its influence on ordinal and text data. Thus, using the online gamers anxiety disorder dataset from Kaggle as the dataset, we applied the FTAWNB model to ordinal data in the current study. Then compare the results obtained with three other naïve Bayes algorithms, namely Gaussian NB, multinomial NB, and categorical

NB, using the same dataset. Model performance is compared based on accuracy, precision, recall, and processing time.

2. RELATED WORK

Since 2001, the WHO has reported that each region has obstacles in diagnosing health problems [26]. WHO also discussed the shortage of treatment facilities and the long, time-consuming diagnosis process [27]. Therefore, many countries use technological sophistication and knowledge development to address mental health problems. Several studies use naive Bayes to identify mental health problems. Even in 2023 and above, research on anxiety disorders, depression, and other mental health cases will often use machine learning models for predictions [28]–[32].

In research on anxiety disorder, a study conducted a comparison of nine classification algorithms, including Gaussian naïve Bayes (GNB), k-nearest neighbor (KNN), decision tree, random forest AdaBoost, support vector machine (SVM), gradient boosting, multi-layer perceptron (MLP) and XGBoost. MLP provides the best accuracy of 99.96% of the nine algorithms, and GNB provides the lowest accuracy of 79.46% [9]. In research on internet addiction by Ioannidis *et al.* [33] naïve Bayes outperformed in cross-validation settings, but its performance varies more in PR-AUC. Furthermore, a study evaluated internet addiction using 100 student data samples. Of the 100 data samples used, the naive Bayes model can classify 88 data correctly [34]. While a study compared the naive Bayes model with two other machine learning algorithms: multinomial logistic regression and auto-sklearn for predicting mood and anxiety disorders. All three models managed to bode well, but the auto-sklearn model outperformed them both [11].

One other mental disorder is depression. Several studies conduct an analysis of depression using the naive Bayes model resulting in an accuracy of 74.35% using a dataset from Reddit [13], 94% using a dataset from the Australian data archive [12], and 86.364% using survey data [35]. In addition, a study uses questionnaire data to classify depression and internet addiction. The study produced the accuracy of the naive Bayes model of 88.9% for depression and 84.1% for the accuracy of Internet addiction [15]. Based on some of this literature, the naive Bayes model has yet to provide the best accuracy compared to other machine learning algorithms used in those studies. This article uses the development of the naive bayes model called the fine tuning attribute weight naive Bayes model for detecting anxiety disorder levels of online gamers, which is one of the mental health problems.

3. METHOD

It starts with the data collecting process, where the data used is public data on anxiety disorders in online gamers taken from Kaggle [36]. Then continue with the data preprocessing process, where there are 13,464 instances in the dataset. This study uses eight attributes: attributes one to seven are statements of general anxiety disorder symptoms taken from the GAD7 scale, and the eight attribute is the class label of the GAD level. There are four class levels used, namely minimal, mild, moderate, and severe anxiety. The weight of each statement is as follows: Not at all0, Several days1, more than half the days2, and nearly every day3. Meanwhile, the final score is obtained by adding all the weights obtained from the GAD7 statements based on the answers from the participants with the anxiety scale between 0 to 21. The scale of each class is shown in Table 1.

Table 1. Classification GAD7 level

Level Anxiety	Score	Class
Minimal	0-4	0
Mild	5-9	1
Moderate	10-14	2
Severe	15-21	3

The next step is followed by the model validation process using 10-fold cross-validation. The data is divided into ten parts with the same amount of each. For each fold (1 to 10 fold), one part becomes testing data, and the remaining nine become training data. There are two algorithms in FTAWNB, namely, the training algorithm and the classification algorithm. Therefore, it will continue building the FTAWNB model after training, which will generate class predictions. Figure 1 shows a flowchart of the classification process in this study. Meanwhile, the FTAWNB model consists of two phases, namely the initializing phase and the fine-tuning conditional probabilities phase, as shown in Figure 2.

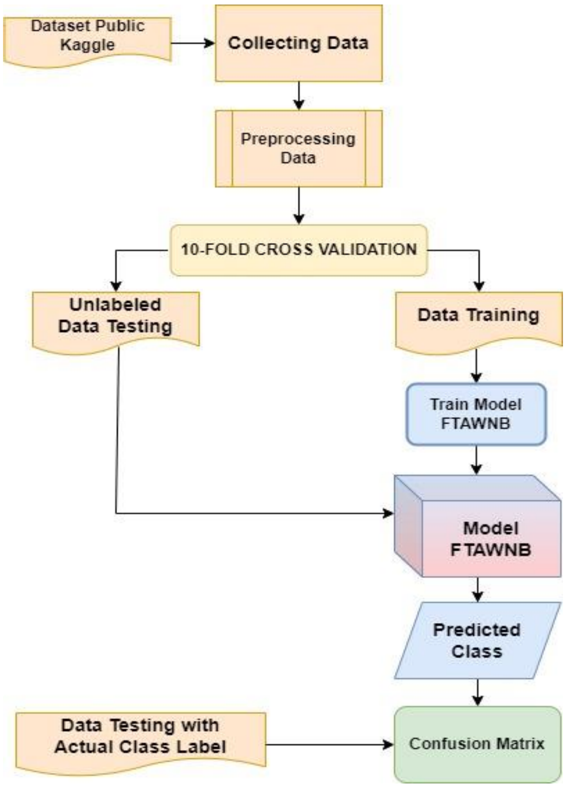


Figure 1. Workflow diagram of the proposed classification process

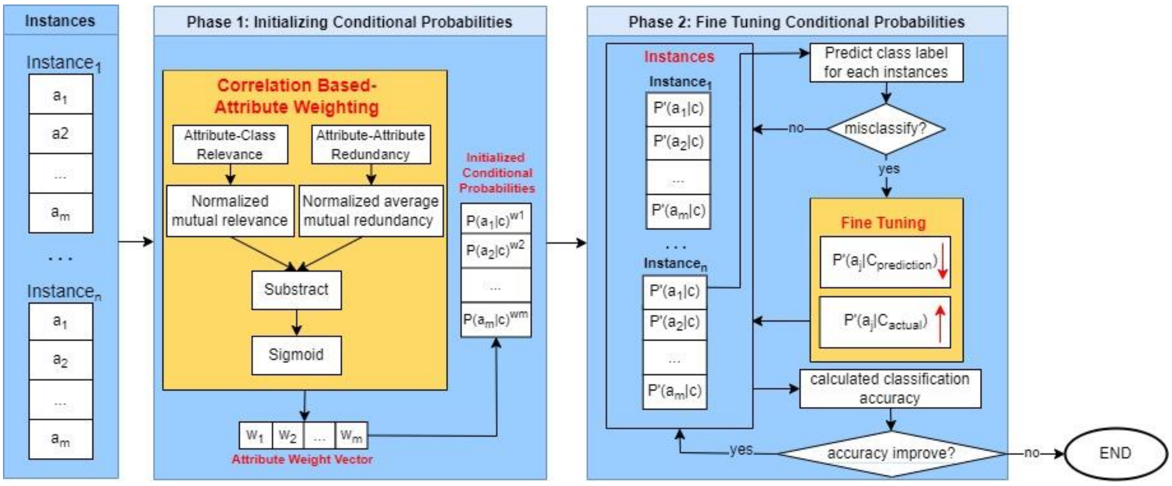


Figure 2. Framework FTAWNB

3.1. Initializing conditional probabilities phase

In the first phase, the weights of the attributes are calculated by taking into account the classes' relevance and the attributes' redundancy. This stage is called initializing conditional probabilities. conditional probabilities are initialized and construct the same information as an attempt to measure the correlation between each pair of discrete random variables. Calculations of attribute-class relevance and attribute-attribute inter-correlation, respectively, are defined in (7) and (8). $I(A_j;C)$ represents attribute-class relevance, while $I(A_j;A_k)$ represents attribute inter-correlation.

$$I(A_j;C) = \sum_{aj}\sum_{c} P(a_j,c) log \frac{P(a_j,c)}{P(a_j)P(c)}, \quad (7)$$

$$I(A_j; A_k) = \sum_{aj} \sum_{ak} P(a_j, a_k) log \frac{P(a_j, a_k)}{P(a_j)P(a_k)}, \quad (8)$$

To keep $I(A_j; C)$ and $(A_j; A_k)$ in the range [0,1], normalization is performed into $NI(A_j; C)$ and $NI(A_j; A_k)$ using 9 and 10.

$$NI(A_j; C) = \frac{I(A_j; C)}{\frac{1}{m} \sum_{j=1}^{m} I(A_j; C)}, \quad (9)$$

$$NI(A_j; A_k) = \frac{I(A_j; A_k)}{\frac{1}{m(m-1)} \sum_{j=1}^{m} \sum_{k=1 \wedge k \neq j}^{m} I(A_j; A_k)}, \quad (10)$$

Then, the subtraction process is carried out using (11) to obtain the weight of the j_{th} attribute D_j. Based on (11), it is shown that the results of the proportional reduction between the normalized mutual relevance and the normalized average mutual redundancy produce the weight of each attribute. Because the D_j value defined by (11) may have a negative value, the standard sigmoid logistic function converts D_j to [0, 1] with (12). Where w_j is the discriminatory weight of the j_{th} attribute.

$$D_j = NI(A_j; C) - \frac{1}{m-1} \sum_{k=1 \wedge k \neq j}^{m} NI(A_j; A_k) \quad (11)$$

$$w_j = \frac{1}{1 + e^{-Dj}} \quad (12)$$

Because the D_j value defined by (11) may have a negative value, the standard sigmoid logistic function converts D_j to [0, 1] with (12). Where w_j is the discriminatory weight of the j_{th} attribute.

3.2. Fine tuning conditional probabilities phase

In the second phase, fine-tuning is carried out on the conditional probabilities of the training instances. First, alternately predict the class label ($C_{prediction}$) from each training instance $T_i (i = 1,2, \ldots, n)$. If a training instance is incorrectly classified ($C_{prediction} \neq C_{actual}$), fine-tune the appropriate conditional probabilities. The fine-tuning formula is more clearly shown by (13) and (14), where c_{actu} and c_{pred} are the actual class and class prediction of each misclassification of the training instance.

$$P'(a_j | c_{actu}) = P'(a_j | c_{actu}) + \delta(a_{j,} c_{actu}) \quad (13)$$

$$P'(a_j | c_{pred}) = P'(a_j | c_{pred}) - \delta(a_{j,} c_{pred}) \quad (14)$$

Then parameter η∈ [0,1] controls the learning rate. Likewise, $\delta(a_j, c_{pred})$ must be reduced in proportion to the error, the difference between β. $P'(a_j|c_{pred})$ and $P'min(a_j|c_{pred})$, and learning rate η. Based on this analysis, the formulas for changing the step size $\delta(a_j, cactu)$ and $\delta(aj, c_{pred})$ are shown in (15) to (17).

$$\delta(a_j, c_{actu}) = \eta. (\alpha \, . \, P'_{max}(a_j | c_{actu}) - P'(a_j | c_{actu})) \, . error \quad (15)$$

$$\delta(a_j, c_{pred}) = \eta. (\beta \, . \, P'(a_j | c_{pred}) - P'_{min}(a_j | c_{pred})) \, . error \quad (16)$$

$$error = P(c_{pred} | T_i) - P(c_{actu} | T_i) \quad (17)$$

4. RESULTS AND DISCUSSION

This study uses the FTAWNB model to classify the level of anxiety disorder among online gamers. The original FTAWNB source code uses the Java programming language, which is included in the Weka classifier. Meanwhile, in this study, the FTAWNB code was run using the Python interpreter by calling system functions in the OS module. We found that the FTAWNB model can be applied to ordinal data. The proposed method in this study tended to have an inordinately higher proportion of performance. Our study suggests that higher accuracy is not associated with the poor performance of the naive Bayes model in previous research using the same dataset [9]. The proposed method may benefit from initializing probabilities and fine-tuning phases without adversely impacting accuracy.

Classification of online gamer anxiety disorder levels using the FTAWNB model provides an accuracy of 99.22%, with a total of 13,359 correctly classified instances out of 13,646 existing instances. This model cannot classify only 105 instances correctly according to their actual class. Table 2 shows the results of the cross-validation. Model performance is measured based on precision and recall. Table 3 shows the ratio of true positive predictions compared to each class's overall positive predicted outcomes. The minimum class gets the highest precision score of 99.7%, while the moderate class gets the lowest precision score of 97.5%.

The success rate of the FTAWNB model in finding information for each class is also shown in Table 3. The success rate of detecting anxiety disorder at the minimum level is the highest compared to others. The acquisition can see a class minimal recall value of 99.8%. Then sequentially, the recall value is 98.8% for mild, 98.6% for moderate, and 97.0% for severe anxiety disorder. The amount of data that is positive and correctly predicted as positive (TP Rate) and the amount of data that is negative but expected as positive (FP Rate) is shown in Table 3. The results show that in each class, there are still instances that should have negative values but are predicted as positive, although the percentage is tiny.

The FTAWNB model correctly predicted 7,429 instances in the minimal class, 3,622 in the mild class, 1,591 in the moderate class, and 717 correct instances in the severe class. For the number of incorrectly classified instances, shown in Table 4 with the following explanation, 17 instances that should be in the minimum class are classified as mild classes. While the instances should be in the mild class, 25 instances are classified as a minimal class and 19 as a moderate class. Furthermore, 22 instances that should be in the class severe are classified as class moderate.

Table 2. Stratified cross-validation

Parameter	Result
Accuracy	99.22%
Correctly classified instances	13,359
Incorrectly classified instances	105
Kappa statistic	0.9871
Mean absolute error	0.15
Root mean squared error	0.2217
Relative absolute error	49.7762%
Total number of instances	13,464

Table 3. Detailed measure by class

Class	Precision (%)	Recall (%)	TP Rate (%)	FP Rate (%)
0	99.7	99.8	99.8	0.4
1	99.2	98.8	98.8	0.3
2	97.5	98.6	98.6	0.3
3	98.9	97.0	97.0	0.1

Table 4. Confusion matrix

Class	Classified as			
	0	1	2	3
0	7,429	17	0	0
1	25	3,622	19	0
2	0	14	1,591	8
3	0	0	22	717

This research also carried out tests using the same dataset, namely anxiety gamers data from Kaggle on several naïve Bayes models such as gaussian naïve Bayes (GNB), categorical naïve Bayes (CNB), and multinomial naïve Bayes (MNB). The performance comparison results of the four models are shown in Table 5. FTAWNB outperforms the other three models by achieving 99.22% accuracy on the dataset used.

Figures 3 and 4(a)-(c) present two different graph types (bar graphs and pie graphs) that compare the performance of the four classifier models. The MNB model had the lowest accuracy, at 61.10%; the GNB had 91.13%; the CNB had 91.59%; and the FTAWNB had the highest accuracy, at 99.22%. In terms of precision, FTAWNB maintained its lead with a precision value of 92.2%, while MNB produced the fewest results. This time around, the GNB model performed better than the CNB model in terms of precision, scoring 89.5% compared to 89.2% for the CNB model. In terms of recall, the FTAWNB model continues to perform better, with a result of 99.2%. GNB, CNB, and MNB have the lowest results, at 88.7%, 87.5%, and 31.4%, respectively.

Table 5. Performance comparison

Model	Performance by		
	Accuracy (%)	Precision (%)	Recall (%)
FTAWNB	99.22	99.2	99.2
GNB	91.13	89.5	88.7
CNB	91.59	89.2	87.5
MNB	61.10	39.4	31.4

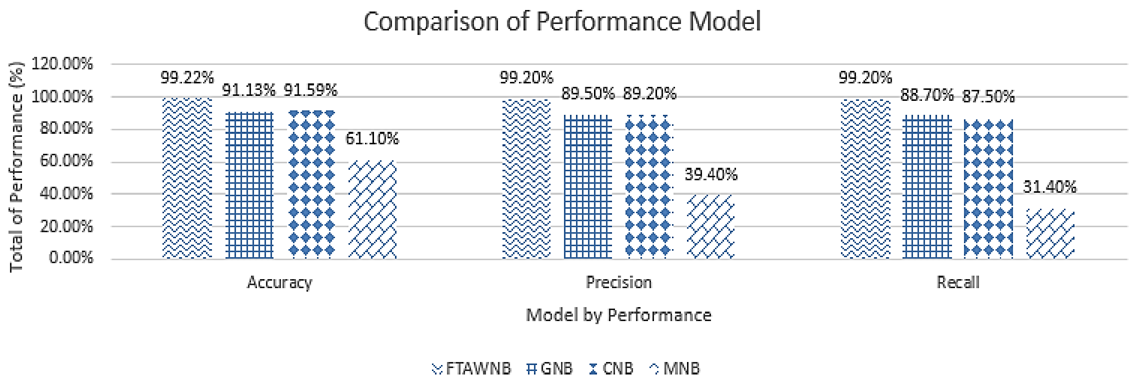
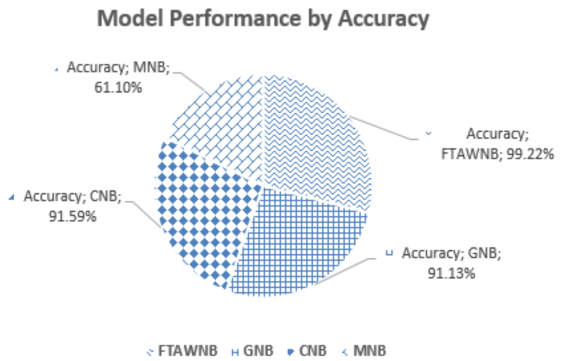
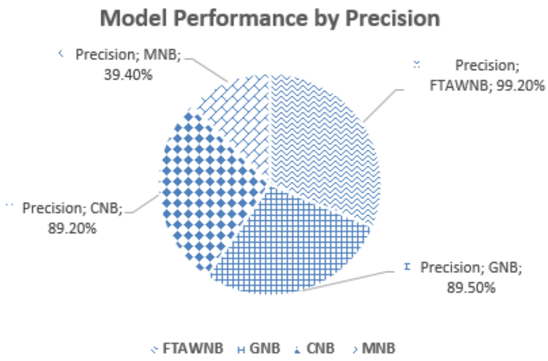


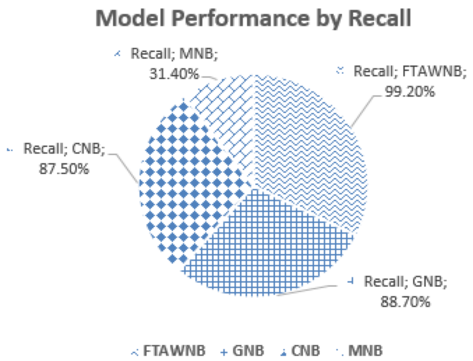
Figure 3. Performance comparison of FTAWNB model with the other naïve Bayes models



(a)



(b)



(c)

Figure 4. Comparison of model performance (a) accuracy, (b) precision, and (c) recall

With a processing time of 0.00468 seconds, the MNB model requires less time to process than other models. Meanwhile, it takes 0.00601 seconds to process the GNB model and 0.00897 seconds to process the CNB model. In terms of accuracy, precision, and recall, the FTAWNB model excels. However, it takes longer for processing time compared to the other three models. According to Table 6, which compares the processing times of the four models, FTAWNB takes 0.07 seconds to build models.

Table 6. Comparison of processing time

Model	Processing time (s)
FTAWNB	0.07
GNB	0.00601
CNB	0.00897
MNB	0.00468

This study explored a comprehensive FTAWNB model with ordinal data. However, further and more in-depth studies may be needed to confirm its performance, especially regarding processing time. Our analysis shows that compared to the GNB, CNB, and MNB models, the FTAWNB model is more robust in terms of accuracy, precision, and recall. Future studies may explore the FTAWNB model with feasible ways of producing faster processing times. The presence of two phases in the FTAWNB makes this model take longer to build a model. As shown in Figure 5, the FTAWNB takes 0.07 seconds, which is much different from the other three naïve Bayes models.

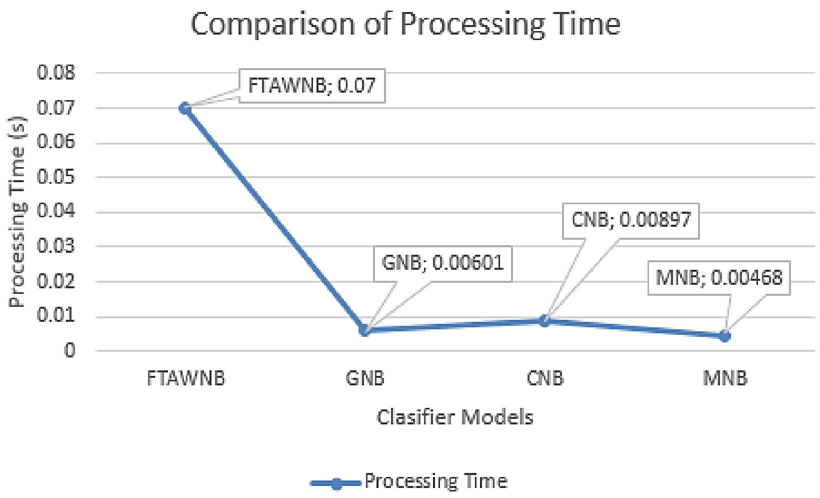


Figure 5. Comparison graph of processing time between the FTAWNB model and three other naïve Bayes models

5. CONCLUSION

This study compares the FTAWNB model with three other naïve Bayes models, namely, Gaussian NB, Categorical NB, and Multinomial NB using the same dataset. We proved that the FTAWNB model still works well on ordinal data and can be used for classification in the mental health domain. The FTAWNB model can provide the best accuracy, precision, and recall based on the test results using online gamers' anxiety disorder ordinal data. However, this model cannot outperform the other three models regarding processing time. FTAWNB requires a longer processing time because it has to go through two phases: the conditional probability initialization phase and the fine-tuning phase.

ACKNOWLEDGEMENTS

The authors would like to thank Universitas Muria Kudus, Universitas Gadjah Mada, the Center for Educational Financial Services (Puslapdik), and the Education Endowment Fund (LPDP) for supporting the authors by providing investing facilities so this research can run smoothly.

REFERENCES

[1] World Health Organization, "ICD-11 chapter 06 mental, behavioural or neurodevelopmental disorders," in *ICD-11 International Classification of Diseases 11th Revision*, 2018, pp. 1–195.
[2] P. D. McDonald, "The principle of division in roger caillois's man, play and games," *Games and Culture*, vol. 15, no. 8, pp. 855–873, 2020, doi: 10.1177/1555412019853080.
[3] C. Bonnaire and D. Baptista, "Internet gaming disorder in male and female young adults: The role of alexithymia, depression, anxiety and gaming type," *Psychiatry Research*, vol. 272, pp. 521–530, 2019, doi: 10.1016/j.psychres.2018.12.158.

[4] J. Burkauskas *et al.*, "Association of Internet gaming disorder symptoms with anxiety and depressive symptoms and substance use: an international cross-sectional study," *Middle East Current Psychiatry*, vol. 29, no. 1, 2022, doi: 10.1186/s43045-022-00180-6.
[5] D. A. Gentile *et al.*, "Pathological video game use among youths: a two-year longitudinal study," *Pediatrics*, vol. 127, no. 2, pp. 319–329, 2011, doi: 10.1542/peds.2010-1353.
[6] C. Y. Wang, Y. C. Wu, C. H. Su, P. C. Lin, C. H. Ko, and J. Y. Yen, "Association between internet gaming disorder and generalized anxiety disorder," *Journal of Behavioral Addictions*, vol. 6, no. 4, pp. 564–571, 2017, doi: 10.1556/2006.6.2017.088.
[7] V. del Barrio, *Diagnostic and statistical manual of mental disorders*. American Psychiatric Publishing, 2004.
[8] R. Aggarwal and A. Goyal, "Anxiety and depression detection using machine learning," in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, May 2022, no. May, pp. 141–149, doi: 10.1109/COM-IT-CON54601.2022.9850532.
[9] A. A. Rahman, L. I. Khalid, M. I. Siraji, M. M. Nishat, F. Faisal, and A. Ahmed, "Enhancing the performance of machine learning classifiers by hyperparameter optimization in detecting anxiety levels of online gamers," in *24th International Conference on Computer and Information Technology, ICCIT 2021*, 2021, pp. 18–20, doi: 10.1109/ICCIT54785.2021.9689911.
[10] V. Borba de Souza, J. Campos Nobre, and K. Becker, "DAC stacking: a deep learning ensemble to classify anxiety, depression, and their comorbidity from Reddit texts," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 7, pp. 3303–3311, Jul. 2022, doi: 10.1109/JBHI.2022.3151589.
[11] W. A. van Eeden *et al.*, "Predicting the 9-year course of mood and anxiety disorders with automated machine learning: a comparison between auto-sklearn, naïve Bayes classifier, and traditional logistic regression," *Psychiatry Research*, vol. 299, 2021, doi: 10.1016/j.psychres.2021.113823.
[12] U. M. Haque, E. Kabir, and R. Khanam, "Detection of child depression using machine learning methods," *PLoS ONE*, vol. 16, pp. 1–13, 2021, doi: 10.1371/journal.pone.0261131.
[13] P. Jain, K. R. Srinivas, and A. Vichare, "Depression and suicide analysis using machine learning and NLP," *Journal of Physics: Conference Series*, vol. 2161, no. 1, 2022, doi: 10.1088/1742-6596/2161/1/012034.
[14] S. Joshi, A. Ghosh, and S. Shukla, "Machine learning model for depression prediction during COVID-19 pandemic," in *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, Apr. 2022, pp. 1–5, doi: 10.1109/ICDCECE53908.2022.9792792.
[15] N. Rubaiyat *et al.*, "Classification of depression, internet addiction and prediction of self-esteem among university students," in *2019 22nd International Conference on Computer and Information Technology*, 2019, pp. 18–20, doi: 10.1109/ICCIT48885.2019.9038211.
[16] S. Chakraborty, S. H. Islam, and D. Samanta, "Data classification and incremental clustering using unsupervised learning," in *EAI/Springer Innovations in Communication and Computing*, 2022, pp. 73–99.
[17] L. Jiang, C. Li, S. Wang, and L. Zhang, "Deep feature weighting for naive Bayes and its application to text classification," *Engineering Applications of Artificial Intelligence*, vol. 52, pp. 26–39, 2016, doi: 10.1016/j.engappai.2016.02.002.
[18] L. Jiang, L. Zhang, C. Li, and J. Wu, "A correlation-based feature weighting filter for naive Bayes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 2, pp. 201–213, 2019, doi: 10.1109/TKDE.2018.2836440.
[19] L. Yu, S. Gan, Y. Chen, and M. He, "Correlation-based weight adjusted naive Bayes," *IEEE Access*, vol. 8, pp. 51377–51387, 2020, doi: 10.1109/ACCESS.2020.2973331.
[20] H. Zhang, L. Jiang, and L. Yu, "Attribute and instance weighted naive Bayes," *Pattern Recognition*, vol. 111, 2021, doi: 10.1016/j.patcog.2020.107674.
[21] K. El Hindi, "Fine tuning the naïve Bayesian learning algorithm," *AI Communications*, vol. 27, no. 2, pp. 133–141, 2014, doi: 10.3233/AIC-130588.
[22] K. M. El Hindi, R. R. Aljulaidan, and H. AlSalman, "Lazy fine-tuning algorithms for naïve Bayesian text classification," *Applied Soft Computing Journal*, vol. 96, 2020, doi: 10.1016/j.asoc.2020.106652.
[23] H. Zhang and L. Jiang, "Fine tuning attribute weighted naive Bayes," *Neurocomputing*, vol. 488, pp. 402–411, 2022, doi: 10.1016/j.neucom.2022.03.020.
[24] C. Elkan, "Boosting and naive Bayesian learning," in *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, Newport Beach, CA, USA, 14–17 August 1997.
[25] P. Langley, W. Iba, and K. Thompson, "Analysis of Bayesian classifiers," in *Proceedings Tenth National Conference on Artificial Intelligence*, 1992, pp. 223–228.
[26] B. Morris, "54th world health assembly," *International Journal of Health Care Quality Assurance*, vol. 14, no. 6, 2001, doi: 10.1108/ijhcqa.2001.06214fab.005.
[27] WHO, "Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level," Report, Sixty-Fifth World Health Assembly, *World Health Organization*, 2012.
[28] A. Rahmadani *et al.*, "Depression, anxiety, and stress disorders detection in students during the CoVID-19 pandemic using naïve Bayes algorithm," *BIO Web of Conferences*, vol. 75, 2023, doi: 10.1051/bioconf/20237501003.
[29] E. Zhou *et al.*, "Prediction of anxious depression using multimodal neuroimaging and machine learning," *NeuroImage*, vol. 285, 2024, doi: 10.1016/j.neuroimage.2023.120499.
[30] Y. Jin *et al.*, "Discovery of depression-associated factors among childhood trauma victims from a large sample size: Using machine learning and network analysis," *Journal of Affective Disorders*, vol. 345, pp. 300–310, 2024, doi: 10.1016/j.jad.2023.10.101.
[31] Y. Zhong *et al.*, "A machine learning algorithm-based model for predicting the risk of non-suicidal self-injury among adolescents in western China: a multicentre cross-sectional study," *Journal of Affective Disorders*, vol. 345, pp. 369–377, 2024, doi: 10.1016/j.jad.2023.10.110.
[32] Q. Liu *et al.*, "Abnormal multi-layered dynamic cortico-subcortical functional connectivity in major depressive disorder and generalized anxiety disorder," *Journal of Psychiatric Research*, vol. 167, pp. 23–31, 2023, doi: 10.1016/j.jpsychires.2023.10.004.
[33] K. Ioannidis *et al.*, "Problematic internet use (PIU): Associations with the impulsive-compulsive spectrum. An application of machine learning in psychiatry," *Journal of Psychiatric Research*, vol. 83, pp. 94–102, 2016, doi: 10.1016/j.jpsychires.2016.08.010.
[34] C. Nandhini and K. Krishnaveni, "Evaluation of internet addiction disorder among students," *Indian Journal of Science and Technology*, vol. 9, no. 19, 2016, doi: 10.17485/ijst/2016/v9i19/93864.
[35] M. L. Joshi and N. Kanoongo, "Depression detection using emotional artificial intelligence and machine learning: a closer review," *Materials Today: Proceedings*, vol. 58, pp. 217–226, 2022, doi: 10.1016/j.matpr.2022.01.467.
[36] D. Agrawal, "Online gaming anxiety data," Kaggle, https://www.kaggle.com/datasets/divyansh22/online-gaming-anxiety-data (accessed Feb 7, 2024).

BIOGRAPHIES OF AUTHORS

Anastasya Latubessy is currently pursuing her doctoral program in the Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta, Indonesia. She took her undergraduate (S.Kom) in Informatics Engineering Department, Faculty of Information Technology, Universitas Kristen Satya Wacana, Salatiga, and Master (M.Cs) in Computer Science and Electronics Department, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Indonesia. Her research areas include machine learning, artificial intelligence, and psychoinformatics. She can be contacted at email: anastasyalatubessy1987@mail.ugm.ac.id.

Retantyo Wardoyo is a lecturer and a researcher at the Department of Computer Science, Universitas Gadjah Mada. He obtained his bachelor's degree from Mathematics in Universitas Gadjah Mada, Indonesia. He obtained his master's degree in computer science at the University of Manchester, UK, and his doctoral degree from Computation at the University of Manchester Institute of Sciences and Technology, UK. His research interests include intelligent systems, reasoning systems, expert systems, fuzzy systems, vision systems, group DSS and clinical DSS, medical computing and computational intelligence. He can be contacted at email: rw@ugm.ac.id.

Aina Musdholifah is a lecturer and a researcher at the Department of Computer Science, Universitas Gadjah Mada. She obtained her bachelor's and master's degree from Computer Science and Electronics Department, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Indonesia. She obtained her doctoral degree from computer science at the Universiti Teknologi Malaysia. Her research interests include genetic algorithm, fuzzy logic and bioinformatics. She can be contacted at email: aina_m@ugm.ac.id.

Sri Kusrohmaniah is a lecture and researcher at the Faculty of Psychology, Universitas Gadjah Mada. She obtained her bachelor's and master's degree from Faculty of Psychology, Universitas Gadjah Mada, Indonesia. She obtained her doctoral degree from Universiti Kebangsaan Malaysia. Her research interests include psychology and cognitive sciences, computer perception, memory and attention, and cognitive science. She can be contacted at email: koes_psi@ugm.ac.id.