ICIC Express Letters
Volume **17**, Number **5**, May **2023**

ICIC International ©2023 ISSN 1881-803X
pp. 531–540

A NEW PROBABILISTIC WEIGHTED VOTING MODEL FOR DEPRESSIVE DISORDER CLASSIFICATION FROM CAPTIONS AND COLORS OF IMAGES

Wongpanya Nuankaew[1], Damrongdet Doenribram[2], Chatklaw Jareanpon[2]
Pratya Nuankaew[3] and Preut Thanarat[4,*]

[1]Polar Lab, Department of Information Technology
[2]Polar Lab, Department of Computer Science
[4]Polar Lab, Department of Creative Media
Faculty of Informatics
Mahasarakham University
Khamriang Sub-District, Kantarawichai District, Maha Sarakham 44150, Thailand
{ 62011262001; chatklaw.j }@msu.ac.th; 60011252001@alumni.msu.ac.th
*Corresponding author: preut.t@msu.ac.th

[3]School of Information and Communication Technology
University of Phayao
Maeka Sub-District, Muang District, Phayao 56000, Thailand
pratya.nu@up.ac.th

Received August 2022; accepted October 2022

Abstract. *Depression disorder is a significant issue leading to suicide. Previously published research on depression has found many associations with posted images, personalities, and emotions of social media users. Identifying patients early at a primary stage will help reduce the graveness levels and consequently the morality rate of attempted suicide. We use a sample taken from Twitter and Instagram. This research aims 1) to find an optimal number of features, and the proper number of classifiers, and 2) to propose a new probabilistic weighted voting model for depressive disorder classification from captions and colors of images. This method uses the single classification combinations of being support vector machine, K-nearest neighbors, decision trees, Naïve Bayes, gradient boosting tree, and generalized linear models. Giving probabilistic weight to single classifiers, up to 6 probabilistic weighted voting ensembles were created. The proposed model achieved an accuracy of 87.23% and was more effective than other models. It theoretically and experimentally performed significantly better than single classifiers and majority vote ensemble models alone. Finally, the model effectively classified patients with depression disorder using the captions and colors of images that they have posted on social media.*
Keywords: Probabilistic weighted voting, Weighted voting ensemble, Depression classification, Text classification, Social media

1. **Introduction.** Major depression disorder (MDD) is a mental health disorder that is the leading cause to suicide. MDD patients encounter emotional and physical difficulties as it affects how patients feel, think, and behave. Identifying individuals and diagnosing patients is needed for a timely treatment plan that can help to reduce the suicide [1]. In recent decades, social media posting data has been extensively studied to analyze the depression disorders [2]; posting such as text, images, emoticons, retweets, shares, and likes can help to diagnose the symptoms of depression. The posting behavior of users reflects their feelings at that moment. Through their post, they may express their feelings with text or image, or both. Therefore, the analysis of text information and image content together may help classify and identify depression better than only text

DOI: 10.24507/icicel.17.05.531

or images. This research analyzes the textual features and image features for depressive disorder classification from captions and colors of images from social media.

The weighted voting ensemble algorithm (WVEA) is applied to improving the performance of the classification model by selecting the highest weights vote given to the single classifiers [3]. It can solve the instability problem of majority vote; when the number of base classifiers is even, the majority vote may decrease performance. If the number of the efficacious classifiers is less than the weak classifiers that lose voting, it wrongly predicts and affects the performance of the model. Researchers have proved the performance of weighted voting ensemble preferable to majority vote because it averts errors generated by a single classifier [3-6]. Probabilistic weighted voting ensemble learning is used for depressive classification [5,6] and classification model [7,8]. The WVEA result was shown to be higher if an appropriate number of classifiers was chosen [7], and this algorithm could be applied to the majority voting problem in many domains.

Rojarath and Songpan [7] proposed WVEA models to find the appropriate number of classifiers using the probability weighted-voting and cost-sensitive probability [8] for weighted voting in ensemble models with the UCI data set. Their results were better than all the other basic classification and ensemble models. Osamor and Okezie [3] developed an improved performance model for the prediction of tuberculosis diagnosis for binary classes, and their results showed the highest accuracy of 95% which was superior to the accuracies of the single classifiers.

The related work issues in the context of depressive disorder have been published using ML and data mining (DM). Researchers have combined a single base classification with the ensemble method to give better prediction results. Doenribram et al. [9] presented a classification of depression on Twitter using the Naïve Bayes (NB) and voting ensemble method. Their targets were grouped into a depressed group and a non-depressed group. They compared different models based on the number of the attributes by top-k with the information gain (IG) method, and the highest accuracies given the training set and testing set were 95.85% and 80.00%, respectively. Nuankaew et al. [5] presented the WVEA framework using probability weighted voting that solved the unweighted vote and optimal features problem in Doenribram's work. A good performance was achieved as the number of features and the processing time was reduced. Zulfiker et al. [6] implemented a depression classification model using single classifiers, ensemble and weighted voting with feature selection algorithms and the synthetic minority oversampling technique was used to adjust imbalanced data. Yazdavar et al. [10] proposed the multimodal prediction framework for predicting depressive behaviors based on content-based, image-based, and network-based models using ML algorithms and ensemble feature selection with statistical techniques. This framework was effective showing an above-average F1 score of 5%. Liu and Shi [11] proposed a hybrid feature selection and ensemble method for identifying depressive users with the synthetic minority oversampling technique by analyzing text and posting behaviors. The stacking model achieved an accuracy greater than the single model and other ensemble models. All researchers presented the models for analyzing depressive disorder to suggest corresponding treatment.

The main contribution of the research described here is a proposal for a new probabilistic weighted voting method that models depression by classifying different feature data. The objectives are to 1) find an optimal number of features and the proper number of classifiers, and 2) to propose a new probabilistic weighted voting model for depressive disorder classification from captions and colors of images. The working principle of the proposed model depends on weight selection and class possibilities corresponding to predictions of classifiers for binary classes. This work may help better classify people with mental illness through their social media posts.

This research is structured as follows. Section 2 describes the research methodology used followed by a description of the data and experiments. Section 3 presents the results and discussion. Section 4 provides information regarding the conclusion and future work.

2. **Methodology.** The Introduction analyzed modeling approaches that can be applied to achieving a better classification result. The research used the posting text to classify depressive disorders that may not be comprehensive in expressing feelings. Using a variety of social media data to analyze depression is to increase the efficiency of identifying the symptoms manifested by user's behavior [10]. Images can be used to communicate meanings and expressions of emotion. Previous studies have shown that depressed patients are likely to expose their moods through their preference for darker shades in daily life [12] and their images posted were more likely to be blue, gray, and dark [13]. Using image features to analyze depressive disorder has been an important developing field of research. This research uses ML and DM for the "new probabilistic weighted voting model for depressive disorder classification from captions and colors of images" (NPWVM). The following sections describe the five phases of modeling: data collection, data pre-processing, training phase, testing phase, and evaluation.

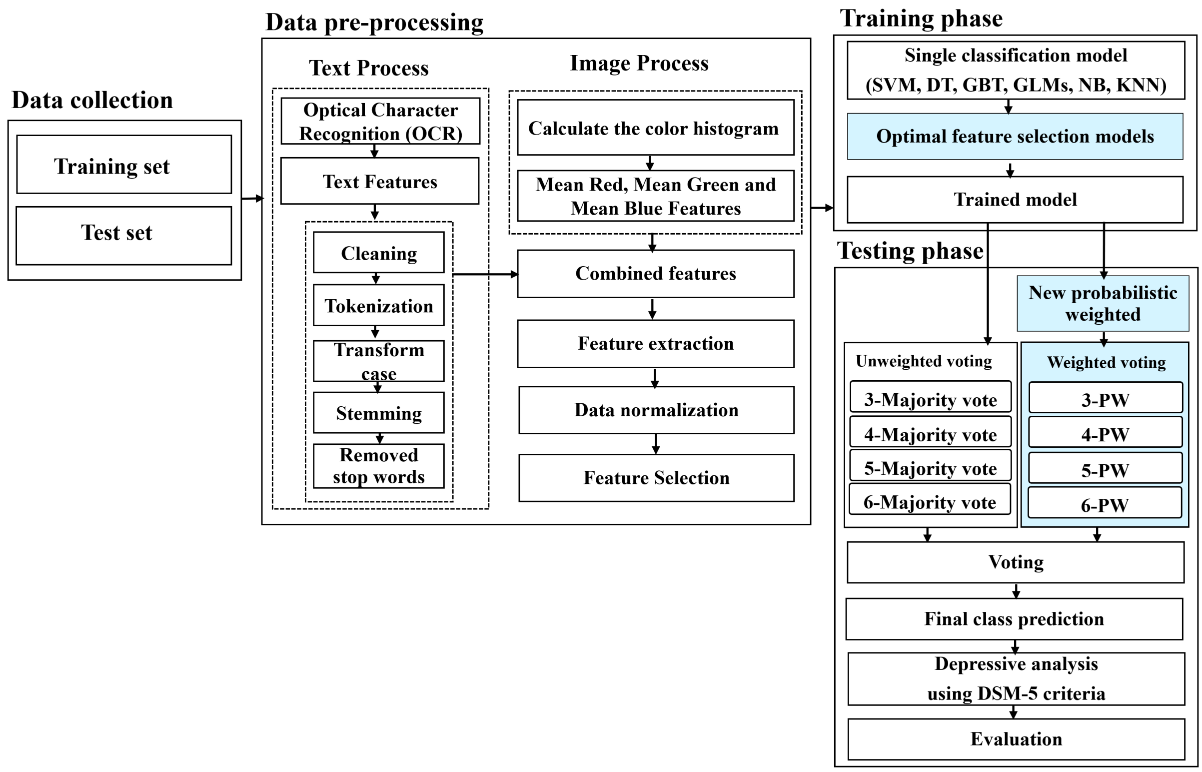


FIGURE 1. Illustration of the proposed new probabilistic weighted voting model for depressive disorder classification

2.1. **Data collection.** The primary dataset is captions and images compiled from Twitter and Instagram posts in 2009 until 2022. It was partitioned into a training set and a test set. The training set used images from social media hashtags that pointed out the nine depressive symptoms and intrapersonal normality of 5,300 samples including 3,200 (60.38%) depression (hashtags from Doenribram et al. [9]) and 2,100 (39.62%) non-depression samples (hashtags: #happy #enjoy #smile). Captions, colors of images, and other visual properties are extracted from each post. The mean RGB values (red, green, and blue) and height and width of an image are used to calculate the RGB histogram features. This training set is used in the model training phase. The test set has 62,257 captions and images of Twitter posted in English by famous 47 vocalists, actors, actresses,

TABLE 1. Summary of the test set data

Type	**Number of people**	**Number of images**	**Avg. number of images**	**Number of captions**	**Avg. number of captions**	**Total**	**Avg. number of total**
Depression	27 (57.45%)	20,830 (68.65%)	771.48	17,415 (54.57%)	645.00	38,245 (61.43%)	1,416.48
Non-depression	20 (42.55%)	9,511 (31.35%)	475.55	14,501 (45.43%)	725.05	24,012 (38.57%)	1,200.60
Total	47	30,341	–	31,916	–	62,257	–

and celebrities. 27 of them are known to be diagnosed with depression and 20 are not, as detailed in Table 1. The test set is used to test the new probabilistic weighted voting model on final classification.

2.2. **Data pre-processing.** The data pre-processing modifies and removes noise, and unnecessary data to prevent misleading feature data in modeling. Several types of information are retrieved from social media posts such as captions, images, retweets, shares, and public user information. Captions and colors are extracted from images. This process consists of two parts: the processing of image features, and the processing of text features.

Processing image features. Image elements include colors, textures, shapes, sizes, spaces, points, and other characteristics. Colors in an image are an essential feature for image representation [14]. Histogram regeneration is the method of extracting color features of an image for usage. An image is composed of pixels and each pixel is defined by an RGB (red, green, and blue) triplet.

The color histogram of an image also indicates the distribution of light intensity. It shows the number of pixel intensities and varies from 0 to 255. It is generated with three-channel RGB values, separating the components of each color R [0, 255] – G [0, 255] – B [0, 255] for each pixel. A color histogram considered by the statistical method reduces the color feature dimensions like mean, standard deviation, and skewness [15]. In this research, the color histograms of the image are calculated using the mean color of the image [16]. The formula of the mean of the RGB histogram is defined by Equation (1).

$$\bar{x} = \frac{\sum_{j=1}^{M \cdot N} x_{ij}}{M \cdot N} \tag{1}$$

when $M \cdot N$ is the size of $M \times N$ of image pixels and x_{ij} is the value of image pixel j of color channel i. We create the histogram of the three features with the mean value of red, green, and blue colors that are used to analyze the model.

Processing text features. Captions and texts are extracted from images using an optical character recognition (OCR) program. To prevent a decrease in model performance, noises like special characters, links, numbers, retweets, and users are removed. Subsequently, many preparatory actions are performed including separating strings into tokens or words, spell checking, stemming for root word, transforming to lowercase, and removing stop words by following the standard stop words list of English language.

Feature extraction process. We extract the means of RGB color features through color histogram calculation and combine them with the text features that are recognized in the images for the representation. All weighted features are normalized to the standardized terms ranging from 0 to 1. This extracting process uses the bag-of-words (BoW) method with the binary term occurrence for the weighted vectors, which can serve as a feature representation for our models. The counting of frequency in a sentence of the binary term is weighed as 1 when the words are detected and as 0 when not [9]. The feature selection is applied to finding an appropriate k constant using IG and ranking the features by the top k. IG is used to reduce unnecessary features in building models [5].

2.3. **Training phase.** The process is designed to train the model with the base classifiers and the ensemble technique to obtain the optimal feature selection models and the adjustment of weights. This research compares the classifiers that are frequently applied to outcomes prediction using decision trees (DT) [7], NB, KNN, SVM [4], generalized linear models (GLMs), and gradient boosting trees (GBT) [18] for building the trained model. All 6 algorithms are the trained models whose individual performance is used to calculate the weighted adjusting for the final make of the decision model.

Weighted adjusting. This research chooses the trained model for the weighted adjusting from the optimal feature selection models. The defining adjusted weight of each model is calculated by the probabilistic of each class and each classifier based on binary classes. The weight of each class is obtained by the correct prediction rate of the classes from the true positive values [8].

$$tb\text{-}rate = (True\ positive\ class)/(Total\ member\ of\ class) \tag{2}$$

Here, *tb-rate* denotes the true positive rate of binary class. The formula that calculates the new probability weighted (y) [8] is obtained by

$$new\ probabilistic\ weighted\ (y) = weightmax \frac{\sum_{n=c}^{class} (Probability_j \times tb\text{-}rate(d))}{c} \tag{3}$$

Here, *new probabilistic weighted* (y) is the maximum weight value of the class. $Probability_j$ is the probability value of the number j of samples in a class, and c is the number of single models used to create learning models.

The highest weighted class is chosen for the final decision in each instance. This research applies the theory of the IG by top k for feature selection, and the weighted voting ensemble algorithm depending on Equations (2) and (3). We use SVM, KNN, DT, NB, GBT, and GLMs to design the proposed weighted voting ensemble models. We then compare the performance of the majority vote and our proposed model. The classifiers are implemented by obtaining SVM, KNN, DT, NB, GBT, and GLMs as the single classifiers which are chosen for the training model of the NPWVM. The NPWVM modeling calculates the new probabilistic weights by Equations (2) and (3). The highest average weighted of each model is taken into account in order to decide the new probabilistic weighting using Algorithm 1. The results of Algorithm 1 obtain the new probabilistic weights of the model that are also used for the final prediction in the test set, as described in Algorithm 2. We present each model's workflows that use a different number of classifiers and the different classifiers for the most efficient model. The probabilistic weighted voting model using three classifiers is called 3-PW model. The four-classifier one is called 4-PW model. The five-classifier one is called 5-PW model. The six-classifier one is called 6-PW model. The majority vote models are called 3-Majority vote, 4-Majority vote, 5-Majority vote, and 6-Majority vote, respectively.

2.4. **Testing phase.** The trained models and the weighted value of each model are applied for performance testing. We compare the classifications between the single classification, unweighted vote, and the NPWVM models.

Depressive analysis. MDD patients have five depressive symptoms or more of the nine symptoms which last for at least 10-14 consecutive days, according to DSM-5 criteria. This process explains the calculations used for analyzing the depression as follows: 1) count the frequency symptom (0-3 points) accumulative in the duration of 14 days, 2) calculate the new frequency until finished, and 3) sum the points and compare to the severity level of depression being 0-7 for non-depression, 8-12 for mild, 13-18 for moderate,

Table 2. The calculation of the new probabilistic weighted voting model

Algorithm 1. Calculation of the new probabilistic weighted voting model

Input: Training set

Output: The new probabilistic weighted (y_j)

Given a set of the classifiers $c = (c_1, c_2, \ldots, c_k)$, a set of the new weights $w = (w_1, w_2, \ldots, w_m)$, a set of probability $P = (P_1, P_2, \ldots, P_j)$, and a set of the binary class $d = (0, 1)$
  for $i \leftarrow 1$ to k **do**
    for $j \leftarrow 1$ to m **do**
      $tb\_rate(d) \leftarrow (true\ positive\ class\ d)/(total\ member\ of\ class\ d)$ *in* Equation (2)
      $w_{(m,d)} \leftarrow sum(P_j \bullet tb\_rate(d))/(the\ number\ of\ classifier\ c)$ *in* Equation (3)
      $y_j \leftarrow weightmax\ w_{(m,d)}$
      $new\_w(y_j) \leftarrow y_j$
    return $weightmax_{j=1,\ldots} w_{(m,d)}$
    end for
  end for

Table 3. The calculation of the final prediction and depressive analysis

Algorithm 2. Calculation of the final prediction and depressive analysis

Input: Test set

Output: The final prediction of the member (z_l)

Given a set of the member of the test set $z = (z_1, z_2, \ldots, z_l)$,
a set of the prediction class of each instance $j = (j_1, j_2, \ldots, j_m)$
for $z \leftarrow 1$ to l **do**
  for $j \leftarrow 1$ to m **do**
    $prediction\_class_j \leftarrow P_j \bullet y_j$ # use y_j from Algorithm 1
    $score \leftarrow 0;$
   for $date\_time \leftarrow 1$ to 14 **do**
    if $prediction\_class_j == depressive\_symptoms$ **then**
     $total\_score(z_l) \leftarrow sum(score(j_m))$
    end if
   end for
  end for
  if $total\_score(z_l) <= 27$ **and** $total\_score(z_l) >= 13$ **then**
    $z_l \leftarrow true$
  else $z_l \leftarrow false$
  end if
end for

and 19-27 for severe. These are criteria determined and used by the American Psychiatric Association [17]. The testing results of the models on severity level of depression were analyzed using Matlab R2020a program (in Algorithm 2). The results of the final predictions are evaluated to determine the performance of the models.

2.5. **Evaluation.** The quality of the model is evaluated using 10-fold cross-validations that compose of 90% training data and 10% testing data. The efficiency metrics used are accuracy, precision, recall, and F1 [8].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

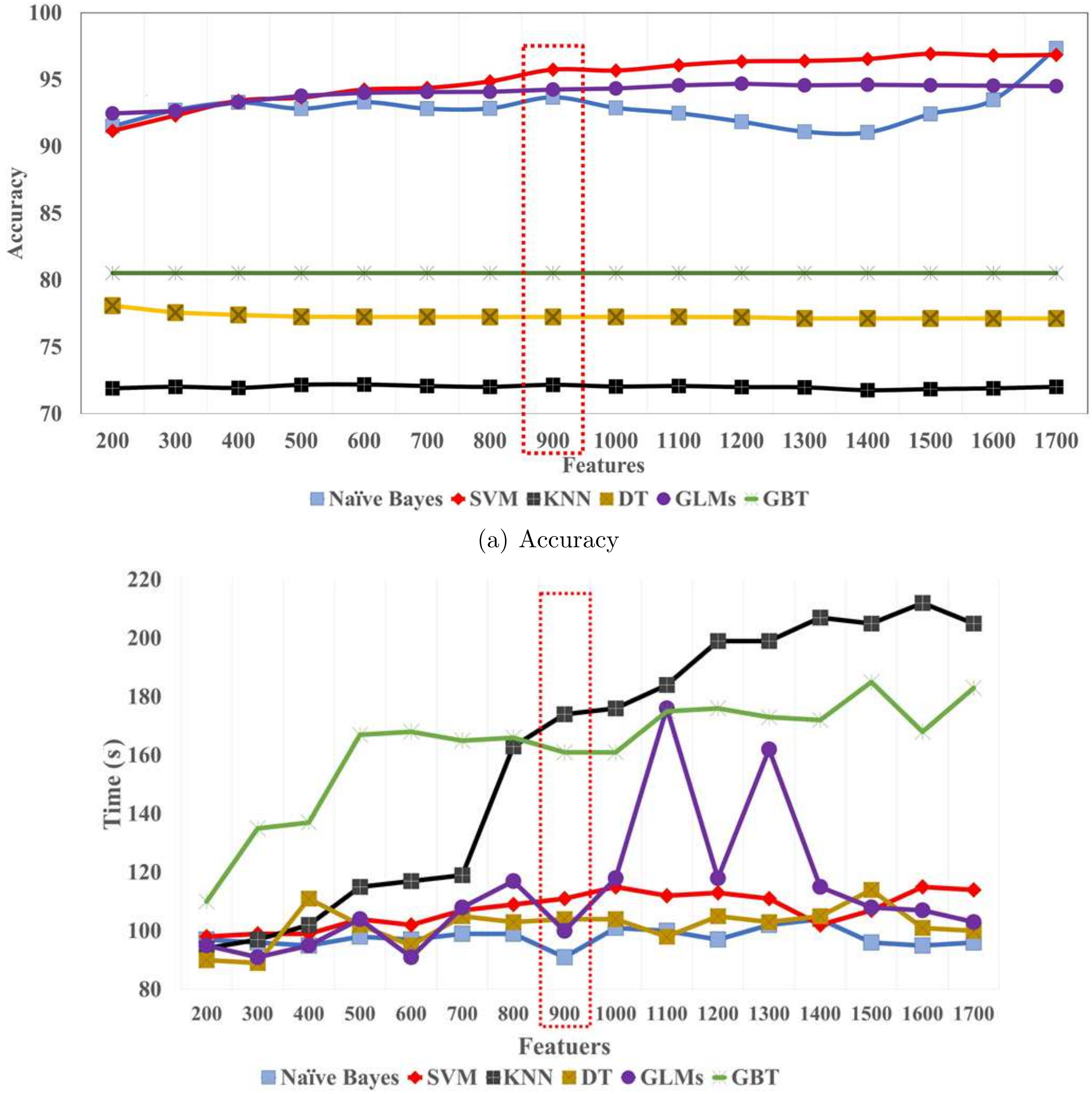
$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

$$\text{F1} = \frac{2 * (\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{7}$$

Here, TP, FP, TN, and FN denote true positive, false positive, true negative and false negative counts, respectively.

3. **Results and Discussion.** To find the effective model, six classifiers, the NB, SVM, KNN, DT, GLMs, and GBT with IG are used to find the number of features and top k values are ranged between 200-1700 features to find the weighting of each feature. Then, we optimize the features by increasing the accuracy in balance with the processing time. As the results, the accuracy and the time efficacy are shown in Figure 2.



(a) Accuracy

(b) Time

Figure 2. The accuracy (a) and time (b) consumption shown to different numbers of feature selections in the six classifications

As shown in Figure 2(a), the accuracies of NB and KNN increase and decrease slightly with the number of features between 200 and 900. After that, both values start to decrease until 1,400 features at which they rise again. The accuracy of SVM model increases as the number of features increases. At 800 to 900 features, the accuracy clearly increases when compared to NB and KNN. For GLMs classifier, the accuracy increases a bit up until 1,200 features, whereas DT and GBT are relatively stable. In Figure 2(b), the processing time increases as the number of features increases for SVM and KNN classifiers, however, that of SVM decreases when the number of features ranges between 1,000-1,400. The other models show decreases in processing time between 800 and 900 features. Considering the balance between accuracy and processing time, 900 features is the optimal number of divisions to use for the classification training model.

With 900 features, the training model is experimented, and the performance is shown in Table 4. The DT classifier model has the highest recall of 98.00% in the depression class. The precision of the NB classifier model is as high as 97.89%. The F1 of the SVM classifier model is as high as 96.44% in the depression class. Lastly, the SVM classifier model with the accuracy of 95.75% outperforms all the other models.

Table 4. The performance of the training model with 900 features

Classifiers	Precision (%)		Recall (%)		F1 (%)		Accuracy (%)
	Yes	No	Yes	No	Yes	No	
SVM	97.75	92.91	95.16	92.91	**96.44**	**94.75**	**95.75**
DT	74.00	**93.97**	**98.00**	47.52	94.33	63.12	78.00
KNN	75.60	66.18	79.59	60.86	77.54	63.41	72.04
NB	**97.89**	88.22	91.50	**97.00**	94.59	92.40	93.68
GBT	86.51	72.84	80.19	80.95	83.23	76.68	80.49
GLMs	95.71	92.12	94.75	93.52	95.71	92.12	94.26

Note. Yes = depression class and No = non-depression class

As the performance of the test data set shown in Table 5, the 4-PW and 5-PW models perform highest in most metrics. The 5-PW model has the highest recall rate at 92.57% in the depression class. The 4-PW scores highest on precision and F1 with the rates of 88.89% in the depression class and 85.00% in the non-depression class. With an accuracy of 87.23%, the 4-PW model also outperforms the other models.

In Table 4, the KNN and DT single model are obviously weak learners. They do not deem viable for this specific decision condition setting as they are statistically not applicable to this dataset. In order to be used for a majority voting model, the performance of training set must be carefully monitored. The 3-Majority vote, 5-Majority vote, 3-PW and 6-PW models are equally accurate due to the limited amount of test set data, which causes similar predictions. The training and the single models are also nearly effective.

The performance of the training set in the non-depression class of some models is not as well as in the depression class as the number of training data may have been too few. The 4-Majority vote model does not predict correctly as the voted scores are tie for 7 persons from the test set. The 4-PW model and the 6-PW model calculate the probabilistic weighted voting and the results are better than the 4-Majority vote and 6-Majority vote, because they obtained the new weights from the trained model that increase effectiveness for the final prediction. The processing times of the models are not significantly different. The proposed model can solve the instability problem of the majority vote.

4. **Conclusion and Future Work.** This research presents the new weighted voting model to classify depressive disorder classification from captions and colors of images. Additionally, a corpus consisting of 5,300 captions and colors of images is developed to perform the depression classification and measure the performance of the NPWVM

Table 5. The performance of the test set

Classifiers	Precision (%)		Recall (%)		F1 (%)		Accuracy (%)	Time (m)
	Yes	No	Yes	No	Yes	No		
Single model								
SVM	88.46	80.95	85.19	85.00	86.79	82.93	85.11	5.19
DT	88.46	80.95	85.19	85.00	86.79	82.93	85.11	**4.18**
KNN	81.25	54.84	48.15	85.00	60.47	66.67	63.83	37.45
NB	88.24	60.00	55.56	**90.00**	68.18	72.00	70.21	5.45
GBT	88.46	80.95	85.19	85.00	86.79	82.93	85.11	5.32
GLMs	85.19	80.00	85.19	80.00	85.19	80.00	82.98	5.40
Ensemble model								
3-Majority vote	85.71	84.21	88.89	80.00	87.27	82.05	85.11	32.39
4-Majority vote	83.33	69.57	74.07	80.00	78.43	74.42	76.60	31.54
5-Majority vote	85.71	84.21	88.89	80.00	87.27	82.05	85.11	31.36
6-Majority vote	85.19	80.00	85.19	80.00	85.19	80.00	82.98	33.03
3-PW model	85.71	84.21	88.89	80.00	87.27	82.05	85.11	31.29
4-PW model	**88.89**	**85.00**	88.89	85.00	**88.89**	**85.00**	**87.23**	31.54
5-PW model	83.33	88.24	**92.57**	75.00	87.72	81.08	85.11	32.39
6-PW model	85.71	84.21	88.89	80.00	87.27	82.05	85.11	31.33

Note. Yes = depression class and No = non-depression class

modeling using the test set of 47 people. As optimized, 900 features are used in the training phase of the model. Although the single classification and majority voting ensemble models (combinations of SVM, KNN, DT, NB, GBT, and GLMs) from captions and image colors are able to classify and detect depression, the NPWVM models obtain better classification results. The 5-PW model has the highest 92.57% recall. Meanwhile, the 4-PW model has the best accuracy of 87.23% and scores the highest precision and F1 at 88.89%. Our model is effective in the resolving depression class. The recall of the NB model (N) outperformes in the non-depression class though with a high overfitting rate in the depression class. With the results of the 4-Majority vote model and 6-Majority vote models, even classifiers are more likely to affect the performance. As performed, the proposed model can solve the problem of weak classification, the instability of the majority voting, and the probability values that are very close to each other and affect the effectiveness. We find analyzing MDD with caption and color features of images on social media is of high effectiveness and allows inferring the mental health states.

In the future, it will be intriguing to examine further into the performance of models in deeper detailed classification such as severity levels of depression or other pre-processing methods for feature vector like term frequency – inverse dense frequency (TF-IDF) [19].

Acknowledgment. Infrastructure, hardware, and software used in this study are partially supported by Mahasarakham University, Rajabhat Maha Sarakham University, and University of Phayao. The data sets were made publicly available by Twitter and Instagram. The authors would like to thank technicians and respondents for their generous support.

REFERENCES

[1] X. Li et al., Predictors of suicidal ideation, suicide attempt and suicide death among people with major depressive disorder: A systematic review and meta-analysis of cohort studies, *Journal of Affective Disorders*, vol.302, no.1, pp.332-351, 2022.

[2] D. William and D. Suhartono, Text-based depression detection on social media posts: A systematic literature review, *Procedia Computer Science*, vol.179, no.2019, pp.582-589, 2021.

[3] V. C. Osamor and A. F. Okezie, Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis, *Scientific Reports*, vol.11, no.1, pp.1-11, 2021.

[4] A. Onan, S. Korukoğlu and H. Bulut, A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification, *Expert Systems with Applications*, vol.62, pp.1-16, 2016.

[5] W. Nuankaew, P. Nuankaew, D. Doenribram and C. Jareanpon, Weighted voting ensemble for depressive disorder analysis with multi-objective optimization (draft), *Current Applied Science and Technology*, vol.23, no.1, https://li01.tci-thaijo.org/index.php/cast/article/view/253085, 2023.

[6] M. S. Zulfiker et al., An in-depth analysis of machine learning approaches to predict depression, *Current Research in Behavioral Sciences*, vol.2, pp.1-12, 100044, 2021.

[7] A. A. Rojarath and W. B. Songpan, Probability-weighted voting ensemble learning for classification model, *Journal of Advances in Information Technology*, vol.11, no.4, pp.217-227, 2020.

[8] A. Rojarath and W. Songpan, Cost-sensitive probability for weighted voting in an ensemble model for multi-class classification problems, *Applied Intelligence*, vol.51, no.7, pp.4908-4932, 2021.

[9] D. Doenribram, C. Jareanpon, J. Jiranukool and S. Sonbua, Major depressive disorder classification from user behaviors from Twitter, *The 24th International Symposium on Artificial Life and Robotics (AROB2019)*, Beppu, Japan, pp.241-246, 2019.

[10] A. H. Yazdavar et al., Multimodal mental health analysis in social media, *PLoS ONE*, vol.15, no.4, pp.1-27, 2020.

[11] J. Liu and M. Shi, A hybrid feature selection and ensemble approach to identify depressed users in online social media, *Frontiers in Psychology*, vol.12, no.1, pp.1-13, 2022.

[12] H. R. Carruthers, J. Morris, N. Tarrier and P. J. Whorwell, The manchester color wheel: Development of a novel way of identifying color choice and its validation in healthy, anxious and depressed individuals, *BMC Medical Research Methodology*, vol.10, pp.1-13, 2010.

[13] A. G. Reece and C. M. Danforth, Instagram photos reveal predictive markers of depression, *EPJ Data Science*, vol.6, no.1, pp.1-12, 2017.

[14] A. Shahbahrami, D. Borodin and B. Juurlink, Comparison between color and texture features for image retrieval, *Proc. of the 19th Annual Workshop on Circuits, Systems and Signal Processing*, 2008.

[15] S. M. Singh and K. Hemachandran, Content based image retrieval based on the integration of color histogram, color moment and gabor texture, *International Journal of Computer Applications*, vol.59, no.17, pp.13-22, 2012.

[16] N. El-Bendary, E. El Hariri, A. E. Hassanien and A. Badr, Using machine learning techniques for evaluating tomato ripeness, *Expert Systems with Applications*, vol.42, no.4, pp.1892-1905, 2015.

[17] American Psychiatric Association, *Diagnostic and Statistical Manual Psychiatric, Mental Disorder Edition "DSM-5"*, 5th Edition, American Psychiatric Association, 2013.

[18] A. Chowanda, D. Suhartono, E. W. Andangsari and K. Z. Zamli, Machine learning algorithms exploration for predicting personality from text, *ICIC Express Letters*, vol.16, no.2, pp.117-125, 2022.

[19] M. H. Nguyen, A label-oriented approach for text classification, *International Journal of Innovative Computing, Information and Control*, vol.16, no.5, pp.1593-1609, 2020.