Tekno-Pedagogi 15 (2) (2025)
ISSN: 2088-205X | E-ISSN: 2715-7415
Tekno-Pedagogi : Jurnal Teknologi Pendidikan
https://online-journal.unja.ac.id/pedagogi
https://doi.org/10.22437/teknopedagogi.v15i2.46957

Classification of Anxiety Levels in Vocational Students Through Life Story Analysis Using Multi-class SVM

Ni Putu Dea Sillviari[1], I Made Candiasa[2], Gede Indrawan[3]
[1,2,3]Universitas Pendidikan Ganesha, Singaraja, Indonesia
Corresponding author email: dea.sillviari@student.undiksha.ac.id

Info Article
Received: 12 July 2025
Revised: 14 July 2025
Accepted: 25 August 2025
Online Version: 24 Sept 2025

Abstract
Anxiety is a psychological condition frequently experienced by vocational high school students due to academic pressure, practical training demands, and uncertainty about future careers. This study aims to (1) classify the anxiety levels of vocational students based on their personal narratives shared via WhatsApp conversations, and (2) compare the performance of two kernel types in the Multi-class SVM classification model. This quantitative study used a computational experimental design involving 670 Grade X students from a vocational school in Gianyar, Bali. A total of 1,476 narrative texts were collected and labelled into five anxiety levels using the DASS-42 scale: normal, mild, moderate, severe, and very severe. The classification process applied TF-IDF vectorization and compared the RBF and Sigmoid kernels. Evaluation results showed that the Sigmoid kernel achieved the highest accuracy (81.42%) and macro-average F1-score (0.7914), demonstrating superior ability to recognize minority classes. The novelty of this study lies in the integration of life story analysis from digital narratives with Multi-class SVM, highlighting its effectiveness in detecting hidden psychological conditions through everyday student communication. These findings provide empirical evidence that computational text analysis can support school-based mental health monitoring. In conclusion, the proposed model not only offers a reliable tool for early psychological screening but also opens opportunities for the development of technology-assisted interventions in educational settings.

Keywords: Anxiety; Classification; Multi-class Support Vector Machine; Vocational High School

INTRODUCTION

Anxiety is a psychological condition commonly experienced by adolescents, characterized by subjective discomfort such as worry, tension, and emotional distress (Soen et al., 2022). It may also refer to temporary feelings of nervousness or fear that arise in challenging life situations (Muhammad et al., 2021), such as during job interviews or medical examinations (Adhara et al., 2023). Anxiety is a natural and frequent occurrence among teenagers (Oktamarin et al., 2022), particularly those enrolled in vocational high schools. It often disrupts concentration (Damayanti et al., 2021), leading students to lose focus during learning activities, feel out of control, daydream frequently, and develop an intense fear of the future (Angelin et al., 2021). This may have a negative impact on physical health and hinder the learning process (Handayuni & Ifdil, 2020). Once students experience mental health issues, their motivation to continue studying tends to decrease significantly (Lumban Gaol, 2016).

Vocational high schools are formal educational institutions that provide mid-level skills training as a continuation of junior secondary education (Rahmadhani et al., 2022). In line with its objectives, vocational high school aims to equip students with the competencies needed to enter the workforce, develop professional attitudes, make career choices (Kurnia & Hakim, 2023), and continue to higher education (Widarto et al., 2024). Vocational education responds to current and future demands of business and industry, preparing graduates to become productive, adaptive, and creative members of society (Prasetyowati et al., 2021). As a distinct form of education, vocational high schools emphasize not only academic knowledge but also practical skills that prepare students for immediate employment (Putri et al., 2022). The challenge of balancing academic responsibilities with job-readiness often causes emotional strain among students (Setiawan et al., 2023), one manifestation of which is anxiety. Such anxiety may affect academic performance, social relationships, and psychological development (Purnama et al., 2024). Many vocational students worry about their ability to secure employment that matches their skills and training. Heavy workloads, frequent evaluations, and high expectations from parents and teachers also contribute to heightened anxiety (Susatya et al., 2023). In general, anxiety disorders can be classified into two types: trait anxiety, which is a personality tendency to perceive non-threatening situations as threatening, and state anxiety, a temporary emotional reaction characterized by subjective worry and physiological arousal (Manuaba et al., 2022). For vocational students who are in their adolescent years social media plays an important role as a platform to express feelings and cope with anxiety through sharing and interaction (Rachmawati & Kusumah, 2023). In education, social media also supports learning and communication, especially in online environments. Among various platforms, WhatsApp is one of the most widely used communication applications. It is not only utilized for personal communication but also serves as a medium for discussions without face-to-face interaction (Sasmita et al., 2022). For students, writing on paper is often perceived as time consuming, energy draining, and mentally demanding, which results in a reluctance to write (Wibawa, 2020). Many students also lack confidence in their writing abilities, although writing can stimulate creativity and idea generation. Consequently, writing on paper is frequently replaced by digital alternatives, particularly via social media.

Life story narratives for vocational students generally refer to documents that describe personal data, education history, competencies, and school experiences (Demartoto et al., 2020). Today, posts and conversations on social media especially WhatsApp can also be considered forms of life stories, representing both ongoing and past personal experiences (Hatta & Ulhaq, 2022). These narratives typically contain reflections on academic performance, technical skills, and emotional responses to various life events (Isnaini et al., 2023). Teachers may use these types of media as informal communication and reflection tools through WhatsApp group chats (Bustomi & Yuliana, 2023), which in turn can be analyzed to help classify students' anxiety levels (Arif et al., 2020). To assess anxiety effectively, it requires a special instrument, in particular when using life story analysis as a basis (Muqorobin & Triana, 2022). However, tools that are specifically designed for this context remain scarce. Most existing instruments, such as the Hamilton Anxiety Rating Scale (HARS), Beck Anxiety Inventory (BAI), and Generalized Anxiety Disorder Scale (GAD-7), focus on general clinical symptoms (Beka-Dede et al., 2022). These tools often emphasize physical and cognitive signs of anxiety and may overlook contextual factors unique to vocational education (Widhiyanti et al., 2021), such as performance pressure, career uncertainty, and workplace adaptation challenges (Widhiyanti et al., 2021). In addition, traditional questionnaire-based approaches can produce biased or inaccurate results due to discomfort in self-reporting, thus hampering proper classification (Yolanda & Mulya, 2024).

This study introduces a new approach to analyse and classify vocational students' anxiety levels through life story narratives collected from WhatsApp conversations. Specifically, this study aims to (1) classify students' anxiety levels into five categories using a machine learning model based on Multi-class SVM, and (2) compare the performance of two kernel types Radial Basis Function (RBF) and Sigmoid in detecting anxiety patterns in narrative texts. Through dual focus, the research is expected to support early detection efforts and enhance the effectiveness of psychological interventions in vocational school settings (Pusparani et al., 2025).

Compared to previous studies, such as of that (Paramartha et al., 2022), which applied SVM to classify public sentiment (positive, negative, neutral) in Twitter data related to public policy, this study offers a novel contribution by implementing Multi-class SVM to classify five psychological anxiety levels in unstructured personal narratives obtained from WhatsApp an informal and private communication platform. In addition to the difference in classification granularity, this approach also departs from Paramartha et al.'s use of public sentiment data by leveraging psychologically grounded indicators (based on the DASS-42 scale) and focusing on authentic, context-specific expressions from students. While the prior study focused on general emotional tone, this research targets clinically relevant anxiety levels for mental health intervention in an educational context. Machine learning provides a powerful framework for converting unstructured narrative data into meaningful information that supports data-driven decision-making (Hendrayana et al., 2023). This method demonstrates how machine learning can classify critical insights from complex data through feature extraction and classification (Handayani, 2021).

Multi-class SVM is particularly well-suited for handling non-linear and high-dimensional data (Indrawan et al., 2022). In this study, the algorithm is used to classify students' anxiety into five levels normal, mild, moderate, severe, and very severe based on their WhatsApp life stories. The classification is enabled by SVM's ability to construct hyperplanes that separate data into more than two categories (Altintaş et al., 2021). Data preprocessing includes normalization, tokenization, stopword removal, and transformation using TF-IDF (Aditya et al., 2025). These steps enable objective and precise analysis of student narratives. By using Multi-class SVM, anxiety patterns in student narratives can be classified more accurately and objectively. Thus, this study is expected to generate relevant findings that support effective psychological intervention and mental health management in vocational education settings. The approach contributes not only to academic knowledge but also offers practical implications for schools seeking to implement early detection systems for student anxiety using digital, data-driven methods.

RESEARCH METHODS

Research Design

This study uses a quantitative approach with a computational experimental design. Unlike traditional experimental designs involving treatment and control groups, this research focuses on evaluating the performance of a machine learning algorithm (Multi-class SVM) in classifying students' anxiety levels based on narrative data. The anxiety-related texts collected from WhatsApp conversations were divided into training (80%) and testing (20%) sets. The experiment involved tuning hyperparameters using GridSearchCV and comparing two kernel types: Radial Basis Function (RBF) and Sigmoid. Both were chosen due to their effectiveness in handling non-linear patterns and high-dimensional feature spaces, which are common characteristics of text data represented through TF-IDF (Syahira & Kurniawan, 2024). The RBF kernel is well-suited for capturing complex, non-linear relationships within the data, while the Sigmoid kernel is known for its ability to model sparse distributions (Rabbani et al., 2023), making it potentially more sensitive to subtle variations in short or informal narrative texts such as WhatsApp conversations. The design allows the researcher to assess the classification accuracy and generalization performance of the model, representing a form of experimental evaluation typical in computational studies.

Research Subject

The subjects of this research were vocational high school students in Class X at one of Vocational High School in Gianyar, Bali. The total number of students involved in this study was 670, who voluntarily participated in WhatsApp chat-based discussions. The sampling technique used was purposive sampling, selecting students who actively engaged in chat discussions and were willing to share their life stories through WhatsApp messages. The research was conducted from January to February 2025. The data collection and interaction with research subjects took place at one of Vocational High School in Gianyar Regency, Bali.

Research Procedure

The figure below illustrates the systematic flow of this study, starting from data collection, keyword determination, text pre-processing, anxiety word weighting, classification using Multi-class SVM, to visualization and recommendation. The research was conducted in the following systematic steps:

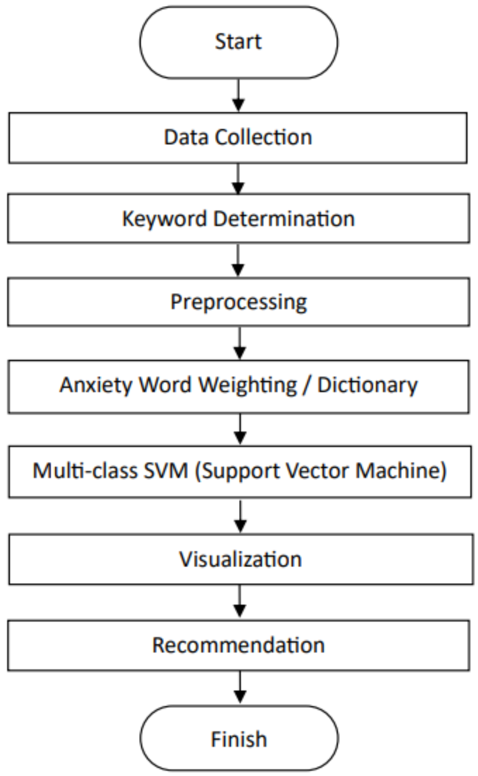


Figure 1. Research Stages

Instruments, and Data Collection Techniques

The primary research instrument is an anxiety dictionary developed based on the DASS-42 scale and related psychological literature. This dictionary includes categorized keywords associated with five anxiety levels: Normal, Mild, Moderate, Severe, and Very Severe.

Table 1. Example of word weighting for anxiety

Word	Word Weight	Anxiety Category
Senang, baik, damai Happy, well, peaceful	1	Normal
Gelisah, cemas, bingung Restless, anxious, confused	2	Mild
Kecewa, malas, stress Disappointed, unmotivated, stressed	3	Moderate
Panik, kewalahan, tertekan Panicked, overwhelmed, pressured	4	Severe
Lelah, hancur, histeris Exhausted, devastated, hysterical	5	Very Severe

To ensure content validity, the dictionary was validated through expert judgment involving a psychology expert, who assessed the contextual relevance and accuracy of each keyword. This table serves as the foundation for the initial annotation process, where each word is assigned a weight to automatically determine the student's anxiety level using a dictionary-based approach.

WhatsApp conversation data was collected from January to February 2025. Each day, students were given one open-ended question related to their personal experiences, a question they responded to freely in WhatsApp group chats or private chats. The process involved the use of Google Colab with Python programming to automate data extraction and convert it into a structured CSV format. All conversations were filtered to remove irrelevant or non-personal messages. Ethical considerations and privacy were fully observed throughout the process. All student data were anonymized prior to analysis to ensure confidentiality and comply with ethical research standards. Written informed consent was obtained from all participants, and parental permission was secured for minors.

Data Analysis Technique

To ensure accurate classification of students' anxiety levels, a structured data analysis technique was employed, consisting of sequential preprocessing, labelling, and classification stages. The analysis process was designed to transform unstructured narrative texts into measurable features that could be processed using machine learning algorithms. This study applied a computational approach that integrates text preprocessing, feature extraction, and model optimization to achieve robust classification results. Each stage of the analysis was carefully constructed to minimize noise, enhance feature representation, and improve the model's predictive performance. The following steps outline the complete data analysis procedure, from text normalization to hyperparameter tuning in the classification model. The text analysis process included the following preprocessing stages:

1. Normalization
2. Stopword removal
3. Stemming
4. Tokenization

After preprocessing, the data were labeled using a combination of dictionary-based labeling and classification using Multi-class SVM with the One-vs-Rest strategy (Jayadi et al., 2023). The analysis began with preprocessing stages including text normalization, stopword removal, stemming, and tokenization (Idris et al., 2023). Then, keywords were weighted and labeled according to anxiety categories. Narrative data were converted into vectorized form using TF-IDF. Classification was performed using Multi-class SVM with One-vs-Rest strategy, comparing the performance of two kernels. To optimize performance, GridSearchCV with 5-fold cross-validation was applied for hyperparameter tuning. The model was trained and tested using an 80:20 data split.

Evaluation Metrics

Several standard evaluation metrics were used to comprehensively measure the performance of the classification model. Evaluation metrics are essential in determining how well the model predicts anxiety levels and in identifying its strengths and limitations. This study adopts widely used metrics in machine learning classification tasks to ensure reliability and comparability of the results. The use of multiple metrics allows for a more balanced evaluation, as each provides a different perspective on model accuracy and robustness. The following section outlines the evaluation metrics applied in this study, including accuracy, precision, recall, and F1-score, along with their respective formulas. The model performance was measured using the following metrics:

1. Accuracy
2. Precision
3. Recall
4. F1-score

Accuracy describes how accurately a model correctly classifies data. The formula is (Putra & Wardijono, 2020):

$$Accuracy = \frac{(\mathrm{TP} + \mathrm{TN})}{(TP + TN + +FP + FN)}$$

Precission indicates the accuracy between the requested data and the predictions made by the model. The formula is:

$$Precision = \frac{\mathrm{TP}}{(TP + FP)}$$

Recall measures how successfully the model retrieves relevant information. The formula is:

$$Recall = \frac{\mathrm{TP}}{(TP + FN)}$$

F1-Score describes a model's success in retrieving information. The formula is:

$$F1 - Score = 2\mathrm{x}\frac{\mathrm{Precision\ x\ Recall}}{Precision + Recall}$$

Description:

a. TP = True Positive
b. FP = False Positive
c. TN = True Negative
d. FN = False Negative

Visualization and Implementation

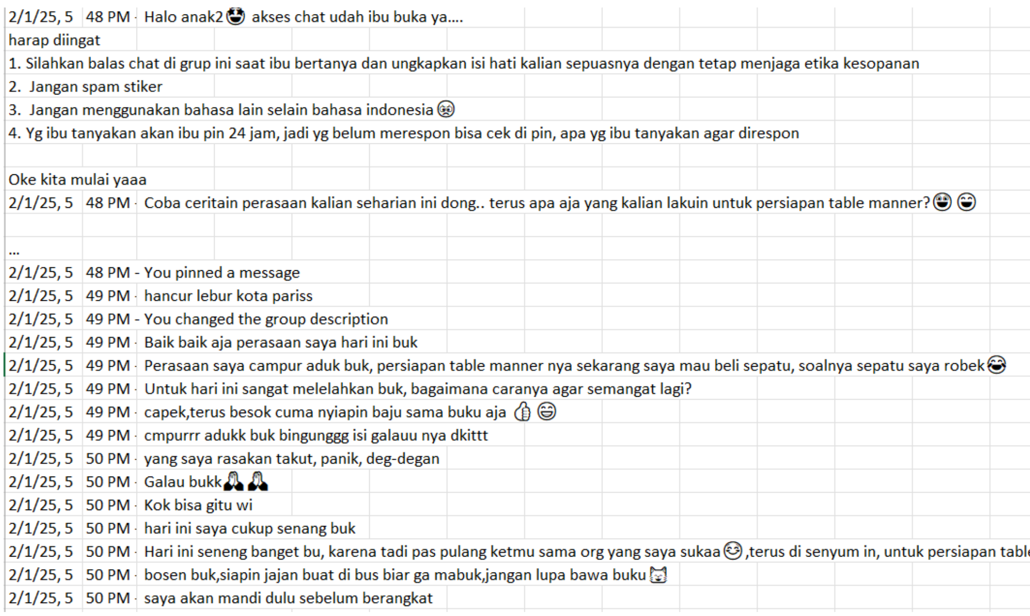
The classification results were visualized using confusion matrices and tables to interpret the model’s accuracy and class distribution. The model’s ability to classify anxiety patterns in student narratives provides a foundation for psychological intervention planning, especially for students classified with Severe or Very Severe anxiety. This approach is considered effective as it combines the strengths of text analysis and machine learning to produce an objective, measurable, and scalable solution for mental health detection in vocational education.

RESULTS AND DISCUSSION

Results

The results of this study clearly address the research objectives: (1) classifying students' anxiety levels into five categories, and (2) comparing the performance of the RBF and Sigmoid kernels in the Multi-class SVM model. A total of 1,476 narrative texts were collected and labelled into five anxiety categories: Normal, Mild, Moderate, Severe, and Very Severe based on the DASS-42 scale. The dataset was pre-processed using normalization, tokenization, stemming, stopword removal, and vectorized with TF-IDF. After applying Multi-class SVM with both RBF and Sigmoid kernels, the classification results were evaluated using accuracy, precision, recall, and F1-score.

Participant recruitment and data collection were carried out from January to February 2025, involving all Grade X students. Each day, students were given one question related to their personal information or experiences, to which they responded freely through WhatsApp groups or private chats. This process maintained ethical standards and participant privacy. A total of 4,257 data entries were collected and subsequently extracted for further analysis.



2/1/25, 5 48 PM · Halo anak2 akses chat udah ibu buka ya....
harap diingat
1. Silahkan balas chat di grup ini saat ibu bertanya dan ungkapkan isi hati kalian sepuasnya dengan tetap menjaga etika kesopanan
2. Jangan spam stiker
3. Jangan menggunakan bahasa lain selain bahasa indonesia
4. Yg ibu tanyakan akan ibu pin 24 jam, jadi yg belum merespon bisa cek di pin, apa yg ibu tanyakan agar direspon

Oke kita mulai yaaa
2/1/25, 5 48 PM · Coba ceritain perasaan kalian seharian ini dong.. terus apa aja yang kalian lakuin untuk persiapan table manner?

...
2/1/25, 5 48 PM - You pinned a message
2/1/25, 5 49 PM · hancur lebur kota pariss
2/1/25, 5 49 PM - You changed the group description
2/1/25, 5 49 PM · Baik baik aja perasaan saya hari ini buk
2/1/25, 5 49 PM · Perasaan saya campur aduk buk, persiapan table manner nya sekarang saya mau beli sepatu, soalnya sepatu saya robek
2/1/25, 5 49 PM · Untuk hari ini sangat melelahkan buk, bagaimana caranya agar semangat lagi?
2/1/25, 5 49 PM · capek,terus besok cuma nyiapin baju sama buku aja
2/1/25, 5 49 PM · cmpurrr adukk buk bingunggg isi galauu nya dkittt
2/1/25, 5 50 PM · yang saya rasakan takut, panik, deg-degan
2/1/25, 5 50 PM · Galau bukk
2/1/25, 5 50 PM · Kok bisa gitu wi
2/1/25, 5 50 PM · hari ini saya cukup senang buk
2/1/25, 5 50 PM · Hari ini seneng banget bu, karena tadi pas pulang ketmu sama org yang saya sukaa ,terus di senyum in, untuk persiapan table
2/1/25, 5 50 PM · bosen buk,siapin jajan buat di bus biar ga mabuk,jangan lupa bawa buku
2/1/25, 5 50 PM · saya akan mandi dulu sebelum berangkat

Figure 2. Raw Data

Pre-processing

In this anxiety analysis study, data cleaning was aligned with pre-processing procedures to ensure optimal results. The pre-processing steps included temporary normalization, custom stopword removal, stemming, tokenization, and data splitting.

Temporary Normalization

Temporary normalization process, a crucial step in preparing the text for further preprocessing. This stage involved standardizing the textual data to reduce inconsistencies and ensure uniformity across the collected WhatsApp conversations. Table 2 below provides a result of the temporary normalization, showcasing how the raw data was transformed into a standardized format suitable for further processing.

Table 2. Temporary normalization result

Data Chat	Normalized Chat
Hancur lebur kota pariss	*Hancur lebur kota paris*
Paris is completely destroyed	Paris is completely destroyed
Baik baik aja perasaan saya hari ini buk	*Baik baik aja perasan saya hari ini bu*
I’m feeling just fine today, Ma’am	I’m feeling just fine today, Ma’am
Campurrr adukk buk bingunggg isi galauu nya dikittt	*Campur aduk bu bingung isi galau nya dikit*
All mixed up, Ma’am, a bit confused and anxious	All mixed up, Ma’am a bit confused and anxious
Galau bukk	*Galau bu*
Feeling anxious, Ma’am	Feeling anxious, Ma’am
Kok bisa gitu wi	*Kok bisa gitu wi*
How could that happen, Wi?	How could that happen, Wi?
Hari ini saya cukup senang buk	*Hari ini saya cukup senang bu*
I’m quite happy today, Ma’am	I’m quite happy today, Ma’am

This initial normalization process serves as a means to clean and simplify the text to make it more uniform, thereby improving accuracy in further analysis and processing. At the preliminary stage, case folding, removal of repeated characters, and the use of a slang dictionary have been carried out to make the data more consistent for the next normalization phase.

Custom Stopword Removal

The next stage is custom stopword removal, which aims to remove words that do not significantly contribute to the anxiety analysis. Words such as conjunctions, prepositions, and other connecting words that frequently appear in the text but are irrelevant to classifying anxiety levels are removed at this stage. This process is crucial to ensure that only words that have emotional meaning or relevance to the ongoing anxiety context are retained, thereby improving the model's accuracy in detecting important patterns in the narrative. Table 3 below shows the results of the stopword removal process, tailored to the student's anxiety context.

Table 3. Custom stopword removal result

Normalized Chat	Cleaned Text
hancur lebur kota paris	*hancur lebur kota paris*
Paris is completely destroyed	Paris is completely destroyed
baik baik aja perasan saya hari ini bu	*baik baik aja perasan saya hari ini bu*
I’m feeling just fine today, Ma’am	I’m feeling just fine today, Ma’am
campur aduk bu bingung isi galau nya dikit	*campur aduk bu bingung isi galau nya dikit*
All mixed up, Ma’am — a bit confused and anxious	All mixed up, Ma’am — a bit confused and anxious
galau bu	*galau bu*
Feeling anxious, Ma’am	Feeling anxious, Ma’am
kok bisa gitu wi	*kok bisa gitu*
How could that happen, Wi?	How could that happen?
hari ini saya cukup senang bu	*hari ini saya cukup senang bu*
I’m quite happy today, Ma’am	I’m quite happy today, Ma’am

In the implementation within research related to anxiety, stopword removal was specifically conducted by eliminating 14 unnecessary words while still maintaining relevance to the context of anxiety. In this stage, irrelevant words were removed, focusing on conjunctions that do not contribute to anxiety analysis, such as “yang,” “dan,” “juga,” “karna,” “tetapi,” “kalau,” “tapi,” “wi,” “le,” “lalu,” “agar,” “sedangkan,” “padahal,” and “karena.”

Stemming

Following the stopword removal stage, the data then undergoes stemming, which aims to simplify words into their basic form. At this stage, various affixes attached to words are removed, so that each word is converted into a simpler and more consistent root form. Table 4 below shows the results of the stemming process, where words have been simplified to facilitate further analysis.

Table 4. Stemming result

Cleaned Text	Teks Bersih
hancur lebur kota paris	*hancur lebur kota paris*
Paris is completely destroyed	Paris is completely destroyed
baik baik aja perasan saya hari ini bu	*baik baik aja rasa saya hari ini bu*
I’m feeling just fine today, Ma’am	I’m feeling just fine today, Ma’am
campur aduk bu bingung isi galau nya dikit	*campur aduk bu bingung isi galau dikit*
All mixed up, Ma’am — a bit confused and anxious	All mixed up, Ma’am — a bit confused and anxious
galau bu	*galau bu*
Feeling anxious, Ma’am	Feeling anxious, Ma’am
kok bisa gitu	*kok bisa gitu*
How could that happen?	How could that happen?
hari ini saya cukup senang bu	*hari ini saya cukup senang bu*
I’m quite happy today, Ma’am	I’m quite happy today, Ma’am

The stemming process was carried out by converting words into their root forms. This was done by removing prefixes, suffixes, or affixes. Additionally, the stemming process included the initialization of a manually compiled root word dictionary based on the collected chat data.

Tokenization

The tokenization process, transforms long narrative text into chunks that can be more easily analyzed by the model. Tokenization allows each word or group of words to be treated as a separate element, allowing for further processing of the data in numerical form that can be used by the classification algorithm. Table 5 below shows the results of the tokenization process, showing how the text has been divided into individual tokens that are easier to process.

Table 5. Tokenization result

Clear Texts	Tokenized Text
hancur lebur kota paris	*['hancur', 'lebur', 'kota', 'paris']*
Paris is completely destroyed	['destroyed', 'completely', 'city', 'paris']
baik baik aja rasa saya hari ini bu	*['baik', 'baik', 'aja', 'rasa', 'saya', 'hari', 'ini', 'bu']*
I’m feeling just fine today, Ma’am	['just', 'fine', 'feeling', 'I', 'today', 'Ma’am']
campur aduk bu bingung isi galau dikit	*['campur', 'aduk', 'bu', 'bingung', 'isi', 'galau', 'dikit']*
All mixed up, Ma’am — a bit confused and anxious	['mixed', 'up', 'Ma’am', 'confused', 'bit', 'anxious']
galau bu	*['galau', 'bu']*
Feeling anxious, Ma’am	['anxious', 'Ma’am']
kok bisa gitu	*['kok', 'bisa', 'gitu']*
How could that happen?	['how', 'could', 'that', 'happen']
hari ini saya cukup senang bu	*['hari', 'ini', 'saya', 'cukup', 'senang', 'bu']*
I’m quite happy today, Ma’am	['today', 'I', 'am', 'quite', 'happy', 'Ma’am']

In research related to anxiety analysis, the data was adjusted through a tokenization process before proceeding to the next stages. This aimed to optimize the results obtained. The tokenization process is intended to break the text into word or phrase units.

Data Splitting

The data splitting stage involves dividing the dataset into two parts: training data and testing data. This splitting is important to ensure that the model can be trained on some of the data and tested on data that was not involved in the training process, thereby producing a more objective performance evaluation. Table 6 below shows the results of the data separation process, which will be used to train and test the Multi-class SVM model in the next stage.

Table 6. Data splitting result

Tokenized Text
Hancur
Destroyed
Lebur
Shattered
Kota
City
Paris
Paris
Baik
Good
Baik
Good
Aja
Just
Rasa
Feeling

To simplify the word weighting process for the development of an anxiety dictionary, the data was split by placing each separated word into a single line. Through this process, a total of 9,118 words were obtained, each of which would later be assigned an anxiety weight based on the values determined by counselors or psychological experts. The output shown in this table facilitates the word weighting and TF-IDF vectorization process, which is essential for transforming narratives into numerical format that can be processed by the SVM algorithm.

Anxiety Words Weighting

The word weighting process was conducted manually under the guidance of counselors. Anxiety levels were scored from 1 to 5, excluding zero, as anxiety is considered a common experience among vocational school students.

Table 7. Anxiety words weighting

Words	Anxiety Score	Anxiety Cathegory
baik, senang, biasa, tenang good, happy, normal, calm	1	Normal
risih, gelisah, galau, ragu uncomfortable, restless, anxious, doubtful	2	Mild
stres, pusing, tertekan, kecewa stressed, dizzy, pressured, disappointed	3	Moderate
krisis, darurat, berantakan, kacau crisis, emergency, messed up, chaotic	4	Severe
mampus, mati, hancur, keras doomed, dead, devastated, intense	5	Very Severe

Sentence Labeling

Labeling process was conducted using a combination of keyword-based and multi-class SVM approaches. Cleaned text was separated into dictionary entries and student responses. Identified keywords were labeled according to anxiety levels, while unidentified text was classified using SVM with TF-IDF and GridSearchCV for optimal accuracy. The results of both methods were combined to form a final dataset containing text, extracted ideas, and anxiety labels for further analysis.

Table 8. Labeling result

Id	Text	label
1	*hancur lebur kota paris* Paris is completely destroyed	Very Severe
2	*baik baik aja rasa saya hari ini bu* I’m feeling just fine today, Ma’am	Normal
3	*campur aduk bu bingung isi galau dikit* All mixed up, Ma’am — a bit confused and anxious	Mild
4	*galau bu* Feeling anxious, Ma’am	Mild
5	*kok bisa gitu* How could that happen?	Normal
6	*hari ini saya cukup senang bu* I’m quite happy today, Ma’am	Normal

Training & Testing

The training and testing of the SVM model for classifying anxiety levels began with loading and inspecting the dataset, followed by class distribution visualization. Data were split into 80% training and 20% testing sets and transformed into TF-IDF vectors. Optimal hyperparameters were determined using GridSearchCV with cross-validation. The best-performing model was used for prediction on the test data and evaluated using accuracy, a classification report, and a confusion matrix for 2D visualization. The results were saved in a csv file. The used kernels included the Radial Basis Function (RBF) kernel and the Sigmoid kernel for comparison of the resulting accuracies. Comparison of different codes and kernels is carried out in this classification process by adjusting the needs of each kernel and not reducing the fairness of the code in the comparison of the two, so that the accuracy results can differ according to the code needs of each kernel. The comparison can be seen from the description of the differences and similarities of the code as follows:

Table 9. Code comparison on each kernel

Step	RBF Version	Sigmoid Version	Notes
1. Library & Data Load	Identical	Identical	Only script repetition
2. Preprocessing (TF-IDF)	TF-IDF vectorization	TF-IDF vectorization + Normalization	Normalizer() used only in sigmoid version due to kernel's sensitivity to data scale
3. Oversampling	Same (RandomOverSampler)	Same (RandomOverSampler)	No difference
4. SVM Kernel	kernel='rbf'	kernel='sigmoid'	Main difference: the kernel used
5. Parameter Tuning (GridSearchCV)	C, gamma, max_iter	C, gamma, max_iter, coef0	coef0 added for sigmoid as it affects kernel output
6. Cross Validation Scheme	Same (StratifiedKFold)	Same (StratifiedKFold)	Identical
7. Evaluation & Output	classification_report.csv, confusion_matrix.csv, Confusion matrix plot	classification_report_sigmoid_tuned.csv, confusion_matrix_sigmoid_tuned.csv, Sigmoid confusion matrix plot	Evaluation concept is the same, only filenames differ
8. Additional Info	Accuracy and best parameters	Accuracy and best parameters	Same

Classification Results Using Radial Basis Function (RBF) Kernel

Throughout this section, we present the evaluation results for the RBF kernel, which demonstrate how this model classifies students' anxiety levels based on the processed data, as well as the performance metrics such as precision, recall, and F1-score for each anxiety category.

	precision	recall	f1-score	support
Normal	0.7530120481927711	0.946969696969697	0.8389261744966443	132.0
Ringan	0.8636363636363636	0.5	0.6333333333333333	38.0
Sedang	0.8507462686567164	0.8382352941176471	0.8444444444444444	68.0
Parah	0.9523809523809523	0.6451612903225806	0.7692307692307693	31.0
Sangat Parah	0.95	0.7037037037037037	0.8085106382978723	27.0
accuracy	0.8108108108108109	0.8108108108108109	0.8108108108108109	0.8108108108108109
macro avg	0.8739551265733608	0.7268139970227256	0.7788890719606127	296.0
weighted avg	0.8285146215273441	0.8108108108108109	0.8037266385224403	296.0

Figure 3. RBF Kernel Result

Based on precision, the highest values were found in the "Severe" class (0.9524) and "Very Severe" class (0.95), indicating strong performance in classifying predictions for these categories. In contrast, the "Normal" class showed the lowest precision (0.7530), suggesting a relatively high number of false positives. In terms of recall, the "Normal" class had the highest value (0.9469), meaning the model successfully captured nearly all actual cases. The lowest recall appeared in the "Mild" class (0.5), revealing that half of the true mild anxiety cases were missed. The F1-score, which balances precision and recall, was highest in the "Normal" class (0.8389), while the "Mild" class again had the lowest (0.6333), indicating difficulty in classifying this category accurately. The distribution of test data was: Normal (132), Mild (38), Moderate (68), Severe (31), and Very Severe (27), with "Normal" being the most dominant class and "Very Severe" the least.

Overall, the RBF kernel achieved an accuracy of 81.08%, which is a solid result for multi-class classification. The macro-average F1-score was 0.7789, indicating performance inconsistency across classes, while the weighted average F1-score of 0.8037 reflected better overall balance considering class distribution. The optimal hyperparameter configuration for the RBF kernel, obtained via GridSearch and One-vs-Rest strategy, was: C=100, gamma='scale', and max_iter=1000. A high C value emphasizes correct classification on training data, while gamma='scale' adapts to feature variance. The iteration limit ensured convergence without excessive training time.

Classification Results Using Sigmoid Kernel

The classification results were also analyzed using the Sigmoid Kernel, which shows an evaluation of accuracy, precision, recall, and F1-score for each anxiety category, as well as how this kernel identifies anxiety patterns in student narratives.

	precision	recall	f1-score	support
Normal	0.7828947368421053	0.9015151515151515	0.8380281690140845	132.0
Ringan	0.7857142857142857	0.5789473684210527	0.6666666666666666	38.0
Sedang	0.8194444444444444	0.8676470588235294	0.8428571428571429	68.0
Parah	0.9523809523809523	0.6451612903225806	0.7692307692307693	31.0
Sangat Parah	0.9130434782608695	0.7777777777777778	0.84	27.0
accuracy	0.8141891891891891	0.8141891891891891	0.8141891891891891	0.8141891891891891
macro avg	0.8506955795285315	0.7542097293720184	0.7913565495537326	296.0
weighted avg	0.821275181687081	0.8141891891891891	0.8101131459582164	296.0

Figure 4. Sigmoid Kernel Result

The Multi-class SVM model using the Sigmoid kernel showed strong performance, particularly in the "Severe" (0.9524) and "Very Severe" (0.9130) classes, which recorded the highest precision values. However, the "Normal" class had the lowest precision (0.7829), indicating a relatively high number of false positives. In terms of recall, the "Normal" class performed best (0.9015), capturing most actual cases in that category. The "Mild" class had the lowest recall (0.5789), suggesting that a significant number of severe anxiety cases were not correctly classied. The highest F1-score was achieved in the "Moderate" class (0.8429), while the "Mild" class remained the weakest (0.6667), although slightly better than in the RBF kernel. Class distribution remained unchanged: Normal (132), Mild (38), Moderate (68), Severe (31), and Very Severe (27).

Overall, the Sigmoid kernel model achieved a slightly higher accuracy of 81.42%, outperforming the RBF kernel. This indicates better generalization to unseen data and more balanced performance across categories. The macro-average F1-score was 0.7914, and the weighted average F1-score was 0.8101, both showing improvement over the RBF results. The optimal hyperparameter configuration for the Sigmoid kernel was: C=10, coef0=-1.0, gamma='scale', and max_iter=1000. The moderate C value contributed to better generalization, while the negative coef0 parameter specific to sigmoid kernels improved model behavior for this dataset. The gamma setting adapted to feature variance, and the iteration limit ensured efficient convergence.

Performance Comparison of RBF and Sigmoid Kernels

The performance of the two tested kernels, was compared to assess their effectiveness in classifying students' anxiety levels. Table 10 below presents a comparison of the accuracy results, average F1-score, and performance on specific anxiety categories between the two kernels, providing an overview of the advantages and disadvantages of each method in the context of student narrative analysis.

Table 10. Comparison kernels

Kernel Type	Accuracy	Macro F1-Score	Highest Precision (Class)	Lowest Recall (Class)
RBF	81.08%	0.7789	Very Severe (0.95)	Mild (0.50)
Sigmoid	81.42%	0.7914	Severe (0.95)	Mild (0.5789)

The Sigmoid kernel slightly outperformed the RBF kernel in overall accuracy (81.42% compared to 81.08%) and macro-average F1-score (0.7914 vs. 0.7789), demonstrating a more balanced and consistent classification performance across all five anxiety levels. Notably, both kernels achieved high precision in identifying critical categories RBF in Very Severe (0.95) and Sigmoid in Severe (0.95). However, the Sigmoid kernel showed a marked improvement in recall for the Mild category, increasing from 0.50 to 0.5789, which indicates better detection of this underrepresented and often ambiguous class. This improvement is significant, as misclassifying lower levels of anxiety can lead to missed opportunities for early intervention.

The results suggest that the Sigmoid kernel is more sensitive to subtle, context-dependent variations in short narrative text, especially given the informal nature of WhatsApp communication. Its capacity to handle sparsity and capture non-linear relationships within the TF-IDF-transformed feature space makes it particularly effective for analyzing emotionally nuanced student responses. Furthermore, the ability to classify not only majority categories like Normal and Moderate, but also more complex and less frequent categories such as Mild, Severe, and Very Severe, highlights the Sigmoid kernel's robustness and adaptability. These findings underscore the importance of selecting an appropriate kernel in machine learning models for psychological text analysis. In practical terms, the superior performance of the Sigmoid kernel enhances the reliability of anxiety level classification, which is crucial in educational settings where early detection can inform timely counseling and support strategies. Thus, the Sigmoid kernel is not only technically advantageous but also pedagogically and psychologically impactful, paving the way for more effective mental health monitoring through automated narrative analysis.

Data Visualization

The data visualization results are presented in the form of confusion matrix tables based on the outcomes produced by each kernel, allowing for a comparison of classification results. In this confusion matrix, it can be seen that the "Normal" class is classified very well by the model, with 125 out of 132 data points correctly classified. Only 7 samples were misclassified, distributed as "Mild" (2) and "Moderate" (5). For the "Mild" class, the model correctly classified only 19 out of 38 samples, while the rest were misclassified as "Normal" (16), "Moderate" (2), and "Severe" (1). This indicates that the model has difficulty distinguishing mild anxiety from other categories, especially from the "Normal" class. In the "Moderate" class, the model demonstrates fairly good performance, with 57 out of 68 samples classified correctly. Most misclassifications fall into the "Normal" (10) and "Mild" (1) categories. Meanwhile, the "Severe" class shows more varied results, with only 20 out of 31 data points correctly classified. The rest were misclassified as "Normal" (8), "Moderate" (2), and "Very Severe" (1).

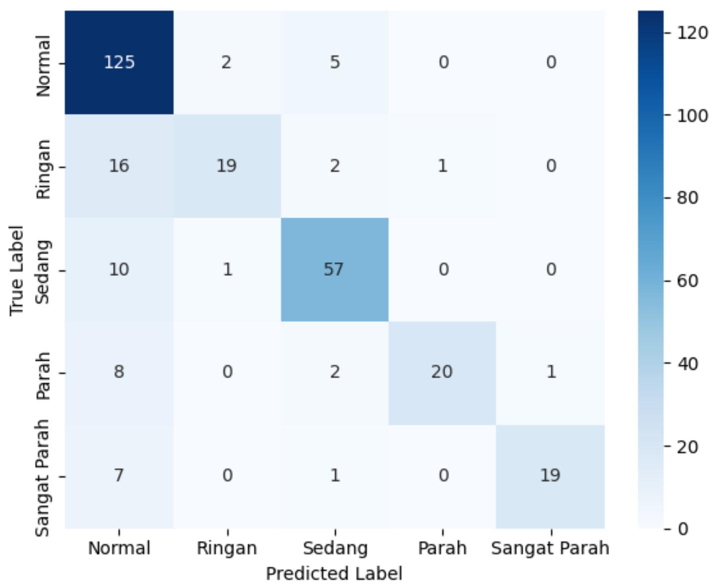


Figure 5. RBF Confusion Matrix

Lastly, in the "Very Severe" class, the model correctly classified 19 out of 27 samples. However, 8 data points were misclassified, mostly as "Normal" (7) and one as "Moderate." In general, the model is capable of classifying majority classes such as "Normal" and "Moderate" quite well. However, for minority classes like "Mild," "Severe," and "Very Severe," there are still notable misclassifications, most likely due to class imbalance and the similarity in characteristics among categories. This suggests that although the overall accuracy is relatively high, the model still needs improvement in recognizing low-frequency classes to ensure more balanced classification performance.

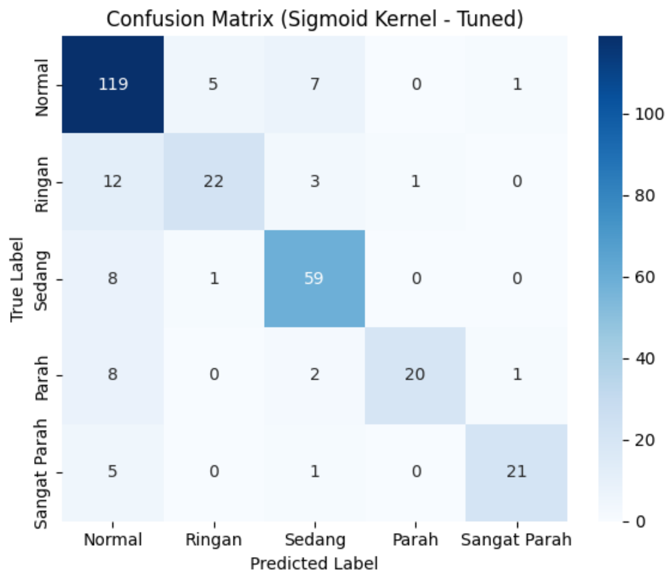


Figure 6. Sigmoid Confusion Matrix

The confusion matrix above represents the classification results using the tuned Sigmoid kernel, illustrating the model's performance in categorizing data into five anxiety levels: Normal, Mild, Moderate, Severe, and Very Severe. The model shows solid performance, particularly in the "Normal" and "Moderate" classes. In the "Normal" category, out of 132 samples, 119 were correctly classified, while the remaining were misclassified as Mild (5), Moderate (7), and Very Severe (1). This demonstrates the model's strong ability to recognize characteristics of normal anxiety. For the "Mild" class, 22 out of 38 samples were correctly classified, with the remaining 16 misclassified into "Normal" (12), "Moderate" (3), and "Severe" (1). Although misclassifications still exist, the model shows improvement over the previous result, especially in reducing confusion between the "Mild" and "Normal" classes.

The model also performs very well in the "Moderate" class, with 59 out of 68 data points correctly classified. Only a few samples were misclassified into "Normal" (8) and "Mild" (1), indicating the model's strong capability in recognizing moderate anxiety. For the "Severe" class, the model correctly classified 20 out of 31 samples, with the rest misclassified into "Normal" (8), "Moderate" (2), and "Very Severe" (1). This shows that although the performance is decent, the model still struggles to distinguish the "Severe" class from neighbouring levels. Finally, in the "Very Severe" category, 21 out of 27 samples were correctly classified, while the remaining were misclassified into "Normal" (5) and "Moderate" (1). This reflects an improvement in the model's accuracy in detecting the most critical anxiety level.

Discussions

Overall, the tuned Sigmoid kernel demonstrates improved accuracy, particularly in classifying minority classes such as Mild and Very Severe. Although some misclassifications persist especially between adjacent severity levels the Sigmoid kernel exhibits greater sensitivity to emotional subtleties embedded in informal, short-form narratives. These results are in line with findings from international studies such as (Altintaş et al., 2021) and (Arif et al., 2020) which emphasize the effectiveness of non-linear SVM kernels for multi-class classification tasks involving emotional or psychological data from social media sources. Expanding this perspective, (Sheykhmousa et al., 2020) conducted a systematic review comparing SVM with other machine learning methods in domains requiring the analysis of high-dimensional and non-linear data, consistently highlighting the superiority of SVM under such conditions. Similarly, (Paramartha et al., 2022) reported that while SVM may underperform in general sentiment classification compared to rule-based approaches like SentiStrength, it excels in capturing deeper, layered emotional features when combined with appropriate feature engineering. These insights reinforce the findings of the present study and support the use of the Sigmoid kernel in processing sparse, context-specific narrative data for early psychological screening.

In this study, the Multi-class Support Vector Machine (SVM) approach especially with the Sigmoid kernel proved effective in classifying vocational students' anxiety levels based on WhatsApp life story narratives. The Sigmoid kernel achieved higher accuracy (81.42%) and macro-average F1-score (0.7914) compared to the Radial Basis Function (RBF) kernel, indicating better generalization performance and heightened sensitivity in detecting less frequent but critical anxiety classes such as Severe and Very Severe. This performance advantage is attributed to the Sigmoid kernel's suitability for handling sparse, non-linear data, which mirrors the characteristics of short, informal WhatsApp messages. The study's methodology supports these outcomes, particularly the integration of TF-IDF vectorization with a DASS-42-based anxiety dictionary. The preprocessing pipeline including normalization, stopword removal, stemming, and tokenization enabled meaningful and consistent feature extraction from the narrative data. Moreover, the use of hybrid labeling, combining dictionary-based and supervised learning approaches, allowed the model to capture both lexically defined and statistically inferred indicators of anxiety.

These results are also consistent with those of Paramartha et al. (2022), who evaluated SVM for sentiment analysis on Twitter and found a lower accuracy (73.33%) compared to SentiStrength, largely due to challenges in detecting emotional nuance in general, unstructured public content. In contrast, the present study benefits from structured, context-specific narratives focused explicitly on anxiety, allowing SVM to perform more reliably. Nonetheless, the model still encountered classification errors between closely related classes such as Normal and Mild, revealing a limitation of TF-IDF in representing semantic nuances. To address this, future research could explore more advanced text representations, such as word embeddings (e.g., Word2Vec, GloVe) or transformer-based architectures (e.g., BERT), which are better suited to capturing contextual meaning and psychological subtleties in language. Despite these limitations, the approach presented here offers a practical, low-cost solution for early mental health screening that leverages widely used digital platforms like WhatsApp making it particularly relevant for resource-constrained educational settings. Ultimately, this method provides school counselors with a scalable, data-driven tool to proactively identify students at risk of elevated anxiety. The successful application of Multi-class SVM in this context further affirms its relevance and adaptability in educational research, especially in domains involving non-linear, high-dimensional, and emotionally complex datasets. By bridging machine learning with real-world school counseling practices, this study lays the groundwork for future advancements in automated, narrative-based mental health assessment.

Recommendations

Based on the results of the study, it was found that several students exhibited signs of anxiety categorized as Severe and Very Severe. However, the number of detected cases may not fully reflect students' actual psychological conditions due to the limitations of written expression, especially in informal platforms like WhatsApp. This is consistent with insights from school counselors, who observed that students with higher levels of anxiety often struggle to articulate their feelings, especially in written formats such as chat messages. This limitation highlights the need for multi-modal data and ongoing support systems. To respond to these findings, several recommendations are proposed.

1. Schools should develop structured guidance and counseling programs informed by machine learning-based anxiety assessments. The classification results can assist counselors in identifying high-risk students and tailoring intervention strategies that are both timely and specific to individual needs.
2. The integration of machine learning models particularly Multi-class SVM into school systems can be achieved through the development of technology-based platforms, such as mobile apps or Learning Management Systems (LMS). These platforms can be designed to automatically monitor students' digital narratives (e.g., reflective journals, chat logs, or weekly submissions) and flag patterns indicative of psychological distress. Practically, schools can implement this system as part

of a broader digital counseling infrastructure. For example, student inputs can be anonymized and analyzed in real time using cloud-based services, with alerts sent directly to counselors or mental health staff when high levels of anxiety are detected.

3. Training programs should be provided to teachers and staff to interpret model outputs and understand how to intervene appropriately.
4. Collaboration with psychologists, mental health professionals, and IT developers is essential to ensure the system is clinically sound, ethically compliant, and technically reliable.
5. Schools should actively promote mental health literacy by organizing seminars, workshops, or peer-support initiatives that normalize discussions around anxiety and encourage help-seeking behavior.
6. Pedagogical adjustments are also important. Shifting toward project-based, collaborative, or inquiry-based learning can reduce academic pressure while fostering emotional resilience and peer engagement.

In summary, integrating this machine learning model into a technology-enhanced school environment offers both preventive and responsive benefits. It allows institutions to move beyond reactive mental health strategies toward a more proactive, data-informed framework that supports student well-being holistically and sustainably.

CONCLUSION

This research employed a quantitative approach with a computational experimental design to classify the anxiety levels of vocational students based on their WhatsApp life story narratives. A total of 1,476 sentences were labeled into five anxiety categories using the DASS-42 scale and processed using a Multi-class SVM model with TF-IDF vectorization. The model's performance was evaluated by comparing the effectiveness of Radial Basis Function (RBF) and Sigmoid kernels. The results showed that the Sigmoid kernel achieved the highest accuracy (81.42%) and macro-average F1-score (0.7914), and was particularly effective in detecting Severe (10.5%) and Very Severe (9.1%) levels of anxiety. These findings confirm the potential of Multi-class SVM to objectively classify psychological conditions from unstructured narrative data and support the implementation of early detection systems in vocational education settings. This method enables data-driven interventions and opens new possibilities for the development of digital counseling tools that help educators identify and support at-risk students more efficiently.

For future development, this approach could be enhanced by incorporating multimodal data, such as audio recordings of student reflections, video-based facial expression analysis, or biometric signals, to capture a richer spectrum of emotional cues beyond text. Integrating these modalities with textual analysis could improve model robustness, reduce misclassification in subtle emotional states, and offer a more holistic assessment of students' psychological well-being. Such innovations would significantly strengthen the practical application of machine learning in school-based mental health monitoring systems and expand its relevance across diverse educational and cultural contexts.

ACKNOWLEDGMENTS

We would like to thank all parties who have supported this research, especially the students and teachers at one of Vocational School in Gianyar, Bali, who participated, as well as colleagues who provided valuable input in the research process.

REFERENCES

Adhara, H., Erika, E., & Wahyuni, S. (2023). Social anxiety levels among adolescents at social child welfare institutions in Pekanbaru City. *International Journal of Nursing and Health Services (IJNHS)*, *6*(4), 246–256. https://doi.org/10.35654/ijnhs.v6i4.691

Aditya, E., Astawa, I. G. K., Limbong, K. G., Indrawan, G., Indrawan, G., & Gunawan, M. A. O. (2025). Analisis sentimen pengguna sistem e-kinerja desa Kabupaten Jembrana menggunakan metode naive bayes. *Jurnal Teknologi dan Sistem Informasi Bisnis*, *7*(1), 8–14. https://doi.org/10.47233/jteksis.v7i1.1693

Altintaş, E., Uylaş Aksu, Z., & Gümüş Demir, Z. (2021). Machine learning techniques for anxiety disorder. *European Journal of Science and Technology*, *31*(31), 365–374. https://doi.org/10.31590/ejosat.999914

Angelin, L., Wetik, S. V., & Ferlan, P. A. (2021). Tingkat stres berhubungan dengan kejadian Generalized Anxiety Disorder (GAD) pada mahasiswa tingkat akhir. *Jurnal Keperawatan Jiwa*, *9*(2), 399. https://doi.org/10.26714/jkj.9.2.2021.399-408

Arif, Ashjan, B., Ghufran, M., Taghreed, S., Nada, A., Nada, A., & Maryam, A. (2020). Classification of anxiety disorders using machine learning methods: A literature review. *Insights of Biomedical Research*, *4*(1). https://doi.org/10.36959/584/455

Beka-Dede, E. V., Mola, S. A. S., & Nabuasa, Y. Y. (2022). Implementasi hamilton anxiety rating scale untuk mendiagnosis tingkat kecemasan pada mahasiswa Dalam Penyusunan Skripsi. *Jurnal Komputer dan Informatika*, *10*(1), 55–64. https://doi.org/10.35508/jicon.v10i1.6353

Bustomi, A. R., & Yuliana, N. (2023). Peran aplikasi whatsapp dalam dinamika ilmu komunikasi. *Triwikrama: Jurnal Multidisiplin Ilmu Sosial*, *2*(4), 2023–2054

Damayanti, P. S., Putra, A., & Srirahmawati, I. (2021). Pengembangan kecerdasan emosional melalui pendidikan karakter pada peserta didik di Sekolah Dasar. *Equilibrium: Jurnal Pendidikan*, *9*(3), 348–356. https://doi.org/10.26618/equilibrium.v9i3.5992

Demartoto, A., Sunesti, Y., Haryono, B., & Mundayat, A. A. (2020). Life story of patient with supervision's fighting against Covid-19 in Surakarta Indonesia. *Jurnal Analisa Sosiologi*, *9*(2), 423–435. https://doi.org/10.20961/jas.v9i2.44436

Handayani, R. N. (2021). Optimasi algoritma support vector machine untuk analisis sentimen pada ulasan produk tokopedia menggunakan PSO. *Media Informatika*, *20*(2), 97–108. https://doi.org/10.37595/mediainfo.v20i2.59

Handayuni, T. S., & Ifdil, I. (2020). The concept of anxiety in practice exam among vocational school students. *Education and Social Sciences Review*, *1*(1), 23. https://doi.org/10.29210/07essr48000

Hatta, H., & Zia Ulhaq, M. (2022). Penggunaan media sosial whatsapp di kalangan mahasiswa program studi desain komunikasi visual Universitas Negeri Makassar. *Jurnal Kependidikan Media*, *11*(3), 153–161. https://doi.org/10.26618/jkm.v11i3.9492

Hendrayana, I. G., Divayana, D. G. H., & Kesiman, M. W. A. (2023). Komparasi Metode Svm, K-Nn dan Nbc pada analisis sentimen. *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, *4*(1), 191–198. https://doi.org/10.35870/jimik.v4i1.157

Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis sentimen terhadap penggunaan aplikasi shopee mengunakan Algoritma Support Vector Machine (SVM). *Jambura Journal of Electrical and Electronics Engineering*, *5*(1), 32–35. https://doi.org/10.37905/jjeee.v5i1.16830

Indrawan, G., Setiawan, H., & Gunadi, A. (2022). Multi-class SVM classification comparison for health service satisfaction survey data in bahasa. *HighTech and Innovation Journal*, *3*(4), 425–442. https://doi.org/10.28991/HIJ-2022-03-04-05

Isnaini, Z., Supriyono, A., & Rachma, S. N. (2023). Efektifitas penggunaan media sosial whatsapp dalam komunikasi kelompok. *MASSIVE: Jurnal Ilmu Komunikasi*, *3*(1), 15. https://doi.org/10.35842/massive.v3i1.84

Jayadi, B. V., Handhayani, T., & Lauro, M. D. (2023). Perbandingan KNN dan SVM untuk klasifikasi kualitas udara di Jakarta. *Jurnal Ilmu Komputer dan Sistem Informasi*, *11*(2). https://doi.org/10.24912/jiksi.v11i2.26006

Kurnia, K., & Hakim, D. L. (2023). Identification of career choice for vocational high school students: A review. *Jiptek*, *16*(2), 150. https://doi.org/10.20961/jiptek.v16i2.67732

Lumban-Gaol, N. T. (2016). Teori stres: Stimulus, respons, dan transaksional. *Buletin Psikologi*, *24*(1), 1. https://doi.org/10.22146/bpsi.11224

Manuaba, I. B. N. W., Dantes, G. R., & Indrawan, G. (2022). Analisis sentimen data provider layanan internet pada twitter menggunakan support vector machine dengan penambahan algoritma levenshtein distance. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, *5*(2), 9–17. https://doi.org/10.47970/siskom-kb.v5i2.261

Muhammad, A. D., Rosyidawati, N. H., Sudrajat, A. A., Khairunnisa, N. H., Rahmawati, B. D. Z., Khatimah, W. H., Apriyani, P., Andriani, A. P. D., Widyastuti, P. A., Suryani, D. S., Nur Azizah, P. F. S., & Yuniasih, D. (2021). Anxiety of final semester students: Mini review. *Ahmad Dahlan Medical Journal*, *2*(2), 85–92. https://doi.org/10.12928/admj.v2i2.5432

Muqorobin, M., & Triana, H. (2022). Validitas instrumen tingkat kecemasan siswa terhadap pembelajaran matematika dalam penelitian eksperimen. *JISIP (Jurnal Ilmu Sosial Dan Pendidikan)*, *6*(2), 4173–4181. https://doi.org/10.58258/jisip.v6i2.3121

Oktamarin, L., Kurniati, F., Sholekhah, M., Nurjanah, S., Oktaria, S. W., Sukmawati, S., & Apriyani, T. (2022). Gangguan kecemasan (axiety disorder) pada anak usia dini. *Jurnal Multidisipliner Bharasumba*, *1*(02), 119–134. https://doi.org/10.62668/bharasumba.v1i02.192

Paramartha, A., Mudjihartono, P., & Emanuel, A. W. R. (2022). Analysis of emotional aspect of twitter postings about large scale social limitations in suppressing the spread of Covid-19. *International Journal of Information Science and Management*, *20*(2), 1–14. https://doi.org/20.1001.1.20088302.2022.20.2.1.6

Prasetyowati, D., Indiati, I., & Nayla, A. (2021). Analisis keterlaksanaan perencanaan dan proses kegiatan pembelajaran praktik di SMK selama pandemic Covid 19. *Jurnal Riptek*, *15*(2), 69–74. https://doi.org/10.35475/riptek.v15i2.121

Purnama, G. W., Permana, A. A. J., Ananda, K. N., Purnami, N. L. I., Nugraha, G. N. A., & Yogi, I. B. S. M. (2024). Implementasi sistem pakar untuk klasifikasi tanaman padi (*Oryza sativa* L.). *Jurnal Pendidikan Teknik Elektro*, *13(2)*(11), 171–185. https://doi.org/10.23887

Pusparani, A. D., Kesiman, M. W. A., & Aryanto, Y. E. K. (2025). Identification of little tuna species using Convolutional Neural Networks (CNN) Method and ResNet-50 Architecture. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)*, *8*(1), 86–93. https://doi.org/10.24014/ijaidm.v8i1.31620

Putra, D. P., & Wardijono, B. A. (2020). Analisis akurasi penerapan algoritma support vector machine menggunakan kernel radial basis function pada penentuan kelayakan kredit (studi kasus german kredit data). *Jurnal Ilmiliah KOMPUTASI*, *19*(2), 175–180. http://dx.doi.org/10.32409/jikstik.19.2.2786

Putri, I. R., Setiawan, A., & Nasrudin, D. (2022). Energy literacy profile of vocational high school teacher candidates for renewable energy engineering expertise program. *Jurnal Ilmiah Pendidikan Teknik Dan Kejuruan*, *15*(2), 99. https://doi.org/10.20961/jiptek.v15i2.67623

Rabbani, S., Safitri, D., Rahmadhani, N., Sani, A. A. F., & Anam, M. K. (2023). Perbandingan evaluasi kernel SVM untuk klasifikasi sentimen dalam analisis kenaikan harga BBM. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *3*(2), 153–160. https://doi.org/10.57152/malcom.v3i2.897

Rachmawati, A. D., & Kusumah, Y. S. (2023). Evaluation of industrial work practice program in Vocational High School 1 Rejotangan. *Jurnal Pendidikan Vokasi*, *13*(1), 60–68.

Rahmadhani, S., Ahyanuardi, & Suryati, L. (2022). Vocational high school students' competency needs to the world of work. *Mimbar Ilmu*, *27*(2), 349–355. https://doi.org/10.23887/mi.v27i1.42161

Sasmita, A., Pradnyana, G. A., & Divayana, D. G. H. (2022). Pengembangan sistem analisis sentimen untuk evaluasi kinerja Dosen Universitas Pendidikan Ganesha dengan Metode Naïve Bayes. *JST (Jurnal Sains Dan Teknologi)*, *11*(2), 451–462. https://doi.org/10.23887/jstundiksha.v11i2.44384

Setiawan, H., Gunadi, A., & Indrawan. (2023). Klasifikasi pelayanan kesehatan berdasarkan data sentimen pelayanan kesehatan menggunakan Multiclass Support Vector Machine. *Jurnal Sistem dan Informatika (JSI)*, *17*(1), 47–54. https://doi.org/10.30864/jsi.v17i1.512

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 6308–6325. https://doi.org/10.1109/JSTARS.2020.3026724

Soen, C. C., Hardjasasmita, I. M., & Ulitua, A. E. (2022). Generalized anxiety disorder: Diagnosis and treatment. *Jurnal Muara Medika Dan Psikologi Klinis*, *1*(2), 133. https://doi.org/10.24912/jmmpk.v1i2.14865

Susatya, E., Triatmaja, A. K., Suratno, E., & Faisyal Fachri, F. (2023). Impact of the center of excellence vocational High School Program on the interest of new learners. *Journal of Vocational Education Studies*, *6*(1), 62–75. https://doi.org/10.12928/joves.v6i1.7256

Syahira, M. A., & Kurniawan, R. (2024). Analisis sentimen cyberbullying pada media sosial X menggunakan Metode Support Vector Machine. *Jurnal Media Informatika Budidarma*, *8*(3), 1724. https://doi.org/10.30865/mib.v8i3.7926

Wibawa, I. P. G. S. T., Sariyasa, & Suarni, N. K. (2020). Pengaruh model pembelajaran inkuiri berbasis STEM dengan penugasan berbantuan Aplikasi Whatsaap terhadap partisipasi orang tua dan hasil belajar IPA siswa kelas IV SD se-gugus VIII mengwi. *Jurnal Pendidikan Dasar Indonesia*, *4*(2), 79–90. https://ejournal-pasca.undiksha.ac.id/index.php/jurnal_pendas/article/view/3319/1786

Widarto, W., Nawanksari, S., & Utami, P. M. (2024). The urgency of interpersonal skills, leadership and self-discipline for vocational high school students majoring in aesthetics and hairdressing. *Jurnal Pendidikan Dan Pengajaran*, *57*(3), 520–531. https://doi.org/10.23887/jpp.v57i3.77440

Widhiyanti, A. A. S., Candiasa, I. M., & Aryanto, K. Y. E. (2021). Implementasi AHP-Topsis dan naïve bayes dalam sistem pendukung keputusan pemberian bimbingan konseling siswa. *SINTECH (Science and Information Technology) Journal*, *4*(2), 129–138. https://doi.org/10.31598/sintechjournal.v4i2.731

Yolanda, A. M., & Mulya, R. T. (2024). Implementasi metode support vector machine untuk analisis sentimen pada ulasan Aplikasi Sayurbox di Google Play Store. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, *6*(2), 76–83. https://doi.org/10.35580/variansiunm258