Identification of Student Self-Blaming and Academic Burnout Based on Emotional Data Using the K-Nearest Neighbor Algorithm

Siti Mentari, Muhammad Irfan Sarif

Abstract

Academic demands and emotional pressure can lead students to experience self-blaming and academic burnout, which may negatively affect learning performance and psychological well-being. This study aims to identify student self-blaming and academic burnout based on emotional data using the K-Nearest Neighbor (KNN) algorithm. A quantitative research approach was employed using emotional data collected through a structured questionnaire measuring indicators such as feelings of guilt, emotional fatigue, stress, and decreased learning motivation. The collected data were processed through several stages, including data cleaning, normalization, and labeling, before being classified using the KNN algorithm. The results indicate that a considerable proportion of students experience moderate to high levels of self-blaming and academic burnout. The classification model achieved an accuracy of 75.15%, demonstrating satisfactory performance in identifying students' emotional conditions. These findings confirm that emotional data can be effectively utilized as input features for machine learning–based classification in educational contexts. This study provides practical implications as an early detection mechanism to support data-driven academic guidance and psychological intervention programs. Academically, this research contributes to the integration of emotional data analysis, learning analytics, and machine learning techniques for student well-being assessment.

Keywords: *Self-Blaming, Academic Burnout, Emotional Data, K-Nearest Neighbor, Machine Learning*

Siti Mentari[1]
[1]Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: sitimentari618@gmail.com[1]

Muhammad Irfan Sarif[2]
[2]Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: irfanberbagi@gmail.com[2]

Introduction

Emotions play an important role in shaping individuals' behavior and responses to academic experiences. One form of negative emotion that frequently emerges is *self-blaming*, which refers to an individual's tendency to blame oneself for failures or negative events. Tangney et al. explain that self-blaming is closely associated with moral emotions such as guilt and shame, which can influence behavior, decision-making, and psychological well-being [1]. In educational contexts, excessive self-blaming may lead to emotional pressure and decreased self-confidence among students. Ingram further states that negatively focused attention on oneself is correlated with emotional disorders and may exacerbate psychological conditions, particularly when individuals face academic demands [2].

In addition to self-blaming, another psychological issue commonly experienced by students is academic burnout. Academic burnout is defined as a condition of emotional exhaustion accompanied by reduced engagement and motivation toward academic activities as a result of prolonged learning demands [3]. Emotional exhaustion is the core component of burnout and has a direct impact on performance, mental health, and overall well-being. If not properly addressed, this condition may develop into more serious psychological problems and hinder students' academic success [4].

Students' emotions during the learning process can be analyzed as data to gain a deeper understanding of learning dynamics. Pekrun, through the control-value theory, states that academic emotions arise from individuals' perceptions of control and value related to learning activities and directly affect students' motivation, learning strategies, and academic outcomes [5].

The advancement of educational technology has generated large volumes of learning data, creating opportunities for the application of educational data mining and learning analytics to better understand students' learning conditions. Baker and Inventado explain that data mining approaches enable the extraction of hidden patterns from educational data to support data-driven decision-making [6]. One widely used classification algorithm in data analysis is the K-Nearest Neighbor (KNN) algorithm. Cover and Hart introduced KNN as a distance-based classification method that determines data classes based on the majority of nearest neighbors [7]. Machine learning techniques have been widely applied to recognize individuals' emotional states using various data types, including facial expressions, voice signals, and physiological signals, as well as to evaluate emotion classification performance using different learning algorithms [8]. Previous studies have shown that KNN is effectively applied in educational contexts and in the analysis of students' emotional data [9].

Literature Review

Self-Blaming and Negative Emotional Conditions

Self-blaming is a form of negative self-evaluation that occurs when individuals attribute failures or unpleasant events to personal faults. Tangney et al. explain that self-blaming is closely related to moral emotions such as guilt and shame, which play an important role in shaping behavior and individual responses to specific situations. While these emotions may encourage constructive self-reflection, excessive intensity can lead to emotional distress and reduced psychological well-being [1].

In educational settings, a high tendency toward self-blaming may cause students to perceive academic failure as a permanent reflection of personal incompetence. Ingram states that negatively self-focused attention has a significant relationship with emotional disorders such as anxiety and depression. This condition can worsen learning experiences and hinder individuals' ability to manage academic demands adaptively[2].

Academic Burnout

Academic burnout is a psychological condition that arises due to prolonged academic pressure. Schaufeli et al. define burnout among students as a state of emotional exhaustion accompanied by reduced engagement and negative attitudes toward academic activities. This condition is often characterized by emotional fatigue, loss of learning motivation, and decreased interest in academic tasks [3].

Maslach et al. explain that emotional exhaustion is the main component of burnout and has a significant impact on individual performance and well-being [4]. urther research by Maslach and Leiter emphasizes that burnout not only affects academic performance but also has long-term implications for psychological health if not properly managed. Therefore, academic burnout is an important variable to be analyzed in understanding students' psychological conditions within educational environments.

Emotions in Educational Contexts (Emotional Data)

Emotions play a crucial role in the learning process as they influence students' motivation, attention, and cognitive strategies. Through the control-value theory of achievement emotions, Pekrun explains that academic emotions arise from individuals' perceptions of control and value related to learning activities. These emotions directly contribute to the quality of learning engagement and academic outcomes [5]. Furthermore, students' emotions are dynamic and may change throughout the learning process. Affective states can be analyzed computationally to better understand students' learning difficulties in a more objective manner. Therefore, emotional data can be utilized as an important source of information for identifying students' psychological issues, including self-blaming and academic burnout.

Educational Data Mining dan Learning Analytics

The development of educational technology has resulted in large volumes of learning data that can be analyzed to support improvements in educational quality. Baker and Inventado explain that educational data mining and learning analytics aim to extract patterns and knowledge from educational data to understand students' learning behavior and support data-driven decision-making [6].

Romero and Ventura emphasize that learning analytics approaches not only focus on academic outcomes but also include affective and behavioral aspects of students during the learning process. Therefore, the application of data mining in education is a relevant approach for analyzing students' emotional data as a basis for identifying psychological conditions that influence learning processes [10].

K-Nearest Neighbor Algorithm in Educational Data Analysis

The K-Nearest Neighbor (KNN) algorithm is a distance-based classification method widely used in data mining and machine learning. Cover and Hart introduced KNN as a method that determines the class of a data point based on the majority class among its nearest neighbors. This algorithm is known for its simplicity of implementation and its ability to handle multidimensional data [7].

Han et al. explain that KNN is effective in various classification problems because it does not require complex model training. In educational contexts, KNN has been widely applied to analyze learning data and students' emotional data to identify behavioral patterns and support data-driven decision-making. Therefore, KNN is considered an appropriate method for identifying students' self-blaming and academic burnout based on emotional data [11].

Research Methodology

This study employs a quantitative research approach using a machine learning–based classification method to identify student self-blaming and exhaustion based on emotional data. The methodology is designed to systematically process emotional data and evaluate the effectiveness of the K-Nearest Neighbor (KNN) algorithm in classifying students' emotional conditions.

3.1 Research Design

The research design used in this study is a descriptive and experimental approach. Emotional data collected from students are analyzed using a supervised learning technique. The K-Nearest Neighbor algorithm is applied to classify students into predefined categories of self-blaming and exhaustion levels, namely low, moderate, and high. This design allows the evaluation of the algorithm's performance based on classification accuracy and other evaluation metrics.

3.2 Research Instruments

The main instrument used in this study is a structured questionnaire designed to measure students' emotional conditions. The questionnaire consists of several indicators related to self-blaming and exhaustion, such as feelings of guilt, stress, emotional fatigue, and decreased learning motivation. Each indicator is measured using a Likert scale to obtain quantitative emotional data suitable for machine learning analysis.

3.3 Data Collection

Data were collected by distributing questionnaires to students as research respondents. The collected data represent individual emotional conditions during the learning process. Prior to analysis, the data were reviewed to ensure completeness and consistency of responses.

3.4 Data Preprocessing

Data preprocessing was conducted to improve data quality and classification performance. This stage includes data cleaning to remove incomplete or inconsistent responses, data normalization to standardize attribute values, and data labeling based on predefined emotional condition categories. The processed data were then prepared for classification using the KNN algorithm.

3.5 K-Nearest Neighbor Classification

The K-Nearest Neighbor algorithm classifies data based on the similarity between data points, which is calculated using Euclidean distance. The value of k represents the number of nearest neighbors considered in the classification process. In this study, several k values were tested to determine the optimal value that produces the highest classification accuracy.

3.6 Evaluation Techniques

The performance of the classification model was evaluated using several metrics, including accuracy, precision, recall, and confusion matrix analysis. These evaluation techniques were used to measure how well the KNN algorithm identifies student self-blaming and exhaustion based on emotional data.

Results

This section presents the results of the study in accordance with the research objectives, namely identifying student self-blaming and exhaustion based on emotional data and evaluating the performance of the K-Nearest Neighbor (KNN) algorithm. The results include respondent characteristics, distribution of emotional conditions, and classification outcomes.

4.1 Respondent Characteristics

The respondents in this study consisted of 534 students who completed the emotional reflection and psychosocial burden questionnaire. The respondents represented diverse educational backgrounds and learning experiences. Emotional indicators related to self-blaming and exhaustion were measured using a Likert scale ranging from 1 to 5.

Prior to analysis, the collected data were examined for completeness and consistency. Incomplete or inconsistent responses were addressed during the preprocessing stage to ensure data quality. The validated dataset was subsequently used for descriptive analysis and machine learning classification.

4.2 Distribution of Student Self-Blaming Levels

The distribution of student self-blaming levels was obtained by calculating the total self-blaming score for each respondent and categorizing the scores into low, medium, and high levels. The results indicate that a substantial proportion of students experience moderate to high levels of self-blaming, suggesting the presence of internalized academic pressure.

Table 1. Distribution of Student Self-Blaming Levels

No.	Self-Blaming Level	Number of Students	Percentage (%)
1	Low	178	33.3
2	Medium	241	45.1
3	High	115	21.6

These findings indicate that self-blaming is a common emotional response among students and may influence their academic experiences.

4.3 Distribution of Student Burnout Levels

Student burnout levels were determined based on emotional indicators such as emotional fatigue, decreased learning motivation, and academic stress. The results show that burnout is experienced by a considerable number of students, particularly at moderate and high levels.

Table 2. Distribution of Student Burnout Levels

No.	Exhaustion Level	Number of Students	Percentage (%)
1	Low	156	29.2
2	Medium	259	48.5
3	High	119	22.3

The dominance of moderate burnout levels suggests that many students may be at risk of developing academic burnout if preventive measures are not implemented.

4.4 K-Nearest Neighbor Classification Results

The K-Nearest Neighbor (KNN) algorithm was applied to classify student emotional conditions based on emotional data related to self-blaming and burnout. The classification process was conducted after data preprocessing, which included data cleaning, normalization, and emotional level labeling. The dataset was divided into training data and testing data to evaluate the performance of the classification model.

The results of the classification show that the KNN algorithm achieved an accuracy of **75.15%**, indicating that the model is capable of identifying student emotional conditions with satisfactory performance. Emotional indicators such as feelings of guilt, emotional fatigue, and decreased learning motivation contributed significantly to the classification process. This finding confirms that emotional data can be effectively utilized as input features for machine learning–based classification.

Table 3. Confusion Matrix of KNN Classification Results

Actual \ Predicted	Low	Medium	High
Low	41	16	0
Medium	6	39	7
High	0	11	41

Table 3 shows that most classification errors occurred between adjacent emotional categories, particularly between medium and high levels. This indicates that emotional conditions among students tend to overlap, which is consistent with psychological theories stating that emotional states are not strictly separated. Similar findings have been reported in previous studies, where KNN demonstrated reliable performance in classifying emotional and behavioral data in educational contexts.

The application of the KNN algorithm in this study provides a practical contribution as an early detection mechanism for identifying students who may be at risk of experiencing self-blaming and exhaustion. The classification results can support educational institutions in designing data-driven academic guidance and psychological intervention programs.

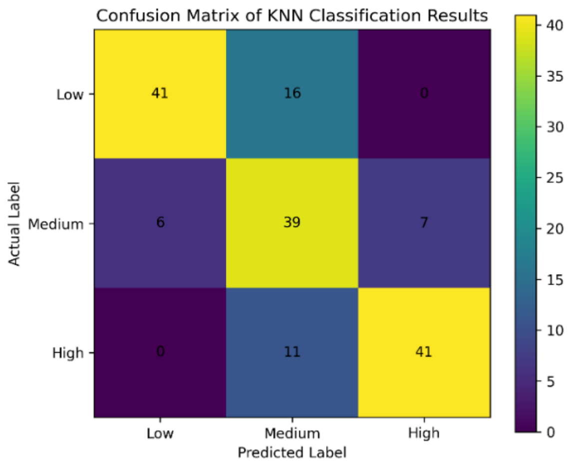

Figure 1. Visualization of KNN Classification Performance Using Confusion Matrix

Conclusion

This study successfully identified student self-blaming and academic burnout (emotional exhaustion) based on emotional data using the K-Nearest Neighbor (KNN) algorithm. The results indicate that a substantial proportion of students experience moderate to high levels of self-blaming and burnout, reflecting emotional challenges that commonly arise in academic learning environments.

The application of the KNN algorithm demonstrated satisfactory classification performance, achieving an accuracy of 75.15%. This result confirms that emotional indicators, including feelings of guilt, emotional fatigue, and decreased learning motivation, can be effectively utilized as input features for machine learning–based classification models in educational settings.

From a practical perspective, the proposed model can function as an early detection mechanism to identify students at risk of academic burnout, enabling educational institutions to implement timely academic guidance and psychological support interventions. From an academic perspective, this study contributes to the integration of emotional data, learning analytics, and machine learning approaches for understanding and monitoring student well-being.

Future studies are encouraged to employ larger and more diverse datasets, compare multiple classification algorithms, and incorporate longitudinal or real-time emotional data to improve model robustness and applicability.

References

[1] J. P. Tangney, J. Stuewig, and D. J. Mashek, "Moral emotions and moral behavior," *Annu. Rev. Psychol.*, vol. 58, pp. 345–372, 2007, doi: 10.1146/annurev.psych.56.091103.070145.

[2] R. E. Ingram, "Self-focused attention in clinical disorders: Review and a conceptual model," *Psychol. Bull.*, vol. 107, no. 2, pp. 156–176, 1990, doi: 10.1037/0033-2909.107.2.156.

[3] W. B. Schaufeli, I. M. Martínez, A. M. Pinto, M. Salanova, and A. B. Barker, "Burnout and engagement in university students a cross-national study," *J. Cross. Cult. Psychol.*, vol. 33, no. 5, pp. 464–481, 2002, doi: 10.1177/0022022102033005003.

[4] C. Maslach, W. B. Schaufeli, and M. P. Leiter, "Utrecht University Repository," vol. 52,

pp. 397–422, 2001.

[5] R. Pekrun, “The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice,” *Educ. Psychol. Rev.*, vol. 18, no. 4, pp. 315–341, 2006, doi: 10.1007/s10648-006-9029-9.

[6] J. A. Larusson and B. White, “Learning analytics: From research to practice,” *Learn. Anal. From Res. to Pract.*, pp. 1–195, 2014, doi: 10.1007/978-1-4614-3305-7.

[7] T. M. Cover and P. E. Hart, “Nearest Neighbor Pattern Classification,” *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.

[8] E. M. G. Younis, S. Mohsen, E. H. Houssein, and O. A. S. Ibrahim, *Machine learning for human emotion recognition: a comprehensive review*, vol. 36, no. 16. Springer London, 2024. doi: 10.1007/s00521-024-09426-2.

[9] S. Mahbub and H. Amiri, “Learning Analytics for Predicting Student Performance in Online Learning Environments,” pp. 1–41, 2020.

[10] C. Romero and S. Ventura, “Educational data mining and learning analytics: An updated survey,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 3, 2020, doi: 10.1002/widm.1355.

[11] J. Ha, M. Kambe, and J. Pe, “Data Mining: Concepts and Techniques,” *Data Min. Concepts Tech.*, pp. 1–703, 2011, doi: 10.1016/C2009-0-61819-5.