International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I4P135

Volume 73 Issue 4, 432-439, April 2025
© 2025 Seventh Sense Research Group®

Original Article

AI for Healthcare: Emotional Data Mining for Problem Risk Analysis in University Students on Social Media Using Text Mining Analytics

Ploykwan Jedeejit[1], Wongpanya S. Nuankaew[2], Patchara Nasa-ngium[3], Pratya Nuankaew[4]

[1]*Department of Marketing, Faculty of Business Administration, Bangkok Thonburi University, Bangkok, Thailand.*
[2]*Department of Computer Science, School of Information and Communication Technology, University of Phayao, Phayao, Thailand.*
[3]*Department of Computer Science, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand.*
[4]*Department, of Digital Business, School of Information and Communication Technology, University of Phayao, Phayao, Thailand.*

[4]*Corresponding Author : pratya.nu@up.ac.th*

Received: 15 November 2024 Revised: 21 March 2025 Accepted: 27 March 2025 Published: 26 April 2025

Abstract *- The domain of adolescent mental health care is a critical preventive measure against issues affecting teenagers. Accordingly, the objectives of this research are to analyze the context in which college-educated adolescents encounter mental health challenges through social media, to establish a predictive model reflecting adolescent sentiment in social media postings employing text mining analytics, and to assess the effectiveness of the developed and selected prototype. This research serves as an application of artificial intelligence and text-mining technologies to foster advancements in the prevention and management of adolescent mental health. The data utilized for analysis consisted of 5,230 social media posts from adolescents enrolled at the University of Phayao, addressing six predominant risk factors: alcohol use, infectious diseases, depression, gaming addiction, pregnancy-related issues, and substance abuse. Five distinct machine learning methodologies have been selected for this study: Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT). The outcomes from the model development and performance testing were notably satisfactory. The models derived using the Support Vector Machine (SVM) and Random Forest (RF) techniques exhibited the highest predictive accuracy, achieving an accuracy rate of 86.81 percent. Consequently, the subsequent research development plan will focus on the practical applications of these findings.*

Keywords *- Adolescent Problems, AI for Healthcare, Emotional Data Mining, Social Media Addiction, Text Mining Analytics.*

1. Introduction

Assessing university students' emotional risk on social media is essential, as these students predominantly use social media as their primary platform to express their feelings, opinions, and personal challenges. However, mining emotional data from social media presents numerous challenges, including the big data challenge, where the vast and unstructured nature of social media data complicates the analytical process. Linguistic variability is another challenge, as the language used on social media often contains slang, emojis, abbreviations, and informal expressions, making it difficult for traditional analytical algorithms. Sentiment analysis accuracy remains a concern since classifying sentiment from textual content may result in inaccuracies, affected by context and sarcasm. Additionally, distinguishing between messages that reflect typical frustrations and those indicating potential risks requiring intervention proves challenging. Lastly, privacy concerns arise as collecting and analyzing student data raises significant issues regarding privacy and ethical considerations in data usage.

Many studies have demonstrated the significance of this research, such as the study titled "Web Emotion Recognition for College Students Based on Microblog Data" [1], which focuses on analyzing the sentiments expressed by college students through social media, especially microblogs. Its main objective is to understand the emotional states of students by examining their online expressions, thus providing valuable insights into their mental well-being. The study entitled "Mining Social Media Data for Understanding Students' Learning Experiences" [2], this research introduces a methodology for analyzing informal social media data to gain insights into students' learning experiences. It illustrates how text mining can reveal patterns related to students' academic and emotional challenges. The research entitled "Analysis Model of College Students' Mental Health Based on Online Social Media Data" [3] undertakes an evaluation of the psychological tendencies manifested by university students. This study employs the processing and analysis of their online social media content over a designated period. The primary aim is to assess the mental health status of these individuals through a thorough multimodal data analysis.

In addition, adolescence is one of the most prone age groups to mental health problems, which manifest as a variety of behavioral problems [4]. Issues that arise are changed and expressed according to the changing times and technology [5], [6], such as being stubborn and disobedient, violating rules and regulations, having a boyfriend and having sex, using drugs, breaking the law, and expressing oneself on social media. Some behavioral problems are so long-standing that they are often challenging to address. Therefore, preventing problems is more necessary and important than solving problems that have already occurred. Such prevention should start with promoting mental health from childhood.

Children with good personality development will have immunity to various mental illnesses and can help prevent mental health problems in adolescents to a great extent. Mental health means a state of mind that is happy, able to have and maintain relationships with others smoothly, able to make oneself useful in an environment that has changed society and the way of life, able to behave appropriately, and free from mental and physical illnesses.

People with poor mental health often have problems adjusting, have psychiatric symptoms such as stress and depression, and have difficulty adjusting even when faced with minor problems. They often have behavioral issues, are quickly ill with mental illnesses, and do not recover well. Common behavioral problems in adolescents include skipping school, gaming addiction, gambling, sexual issues, depression, suicide, and personality disorders.

Another significant issue is social media addiction in teenagers [7], [8], which is using social media for too long, which affects the balance of life, study, and the lack of interaction with people around them. In addition, social media addiction is risky behavior that can cause depression, stress, ADHD, and bipolar disorder. However, current technology has advanced enormously. Examples include artificial intelligence technology in medicine, applied informatics technology for medicine, the development of medical applications, the Internet of Things for tracking patients and healthcare, and many more, enabling society to track and protect teenagers. An innovation that can be used to protect against the potential risks that teens face is text mining technology, which tracks and analyzes social media messages. Text mining is used to analyze text or keywords that can indicate risk issues that the poster faces. Therefore, researchers intend to use text mining techniques to analyze adolescent risk entitled " Emotional Data Mining for Problem Risk Analysis in University Students on Social Media Using Text Mining Analytics".

1.1. Research Objectives

- The first is to study the context of college-educated adolescents' exposure to adolescent mental health problems through social media.
- The second is to develop a predictive model of adolescent voice in social media posts using text mining analytics.
- The last objective is to evaluate the efficacy of the developed and selected prototype.

The data collected for analysis were collected from 5,230 social media posts of adolescents studying at the University of Phayao, consisting of six risk issues of adolescents: alcohol, infectious diseases, depression, gaming addiction, pregnancy problems, and drug problems.

2. Materials & Methods

2.1. Research Scope

The scope of the research is conducted according to the research framework, which is presented in Figure 1; it describes the scope and process of the study. The research consists of four phases: the first phase involves data collection from five social media platforms, including Facebook, Instagram, Twitter, YouTube, and TikTok.

The second phase consists of constructing the model prototype using the CRISP-DM technique, including business understanding, data understanding, data preparation, modelling, evaluation, and deployment. The third phase involves evaluating the prototype model's performance, using the split-test technique and confusion matrix performance, and the last phase involves efficient model selection, using five indicators: accuracy, precision, recall, and f1-score values.

2.2. Data Collection

The data collected and used in this research were social media posts from 2023 – 2024 from students at the University of Phayao, Phayao Province, Thailand. The researchers extracted the data by identifying six keywords: alcohol, infectious diseases, depression, gaming addiction, pregnancy problems, and drug problems. The data obtained is a collection of Thai words published on the link: https://shorturl.at/UJIlP.

The data collection results included label 1: 1,003 depression posts (19.18%), label 2: 1,002 alcohol posts (19.16%), label 3: 1,004 gaming addiction posts (19.20%), label 4: 1,001 infectious diseases posts (19.14%), label 5: 466 drug problems posts (8.91%), and label 6: 754 pregnancy problems posts (14.42%), a total of 5,230 posts.

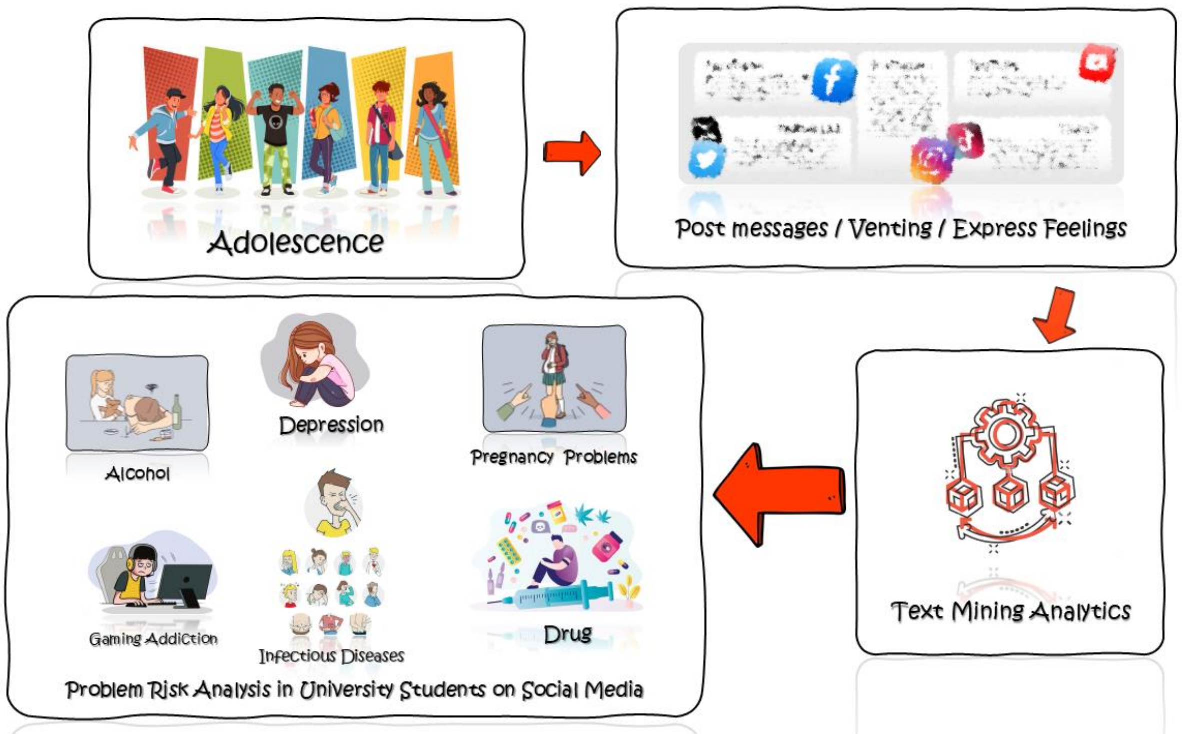

Fig. 1 Research framework

2.3. Modelling

The development process of this research model followed the CRISP-DM data mining technique [9], [10], [11], which consists of six significant stages, as detailed below.

2.3.1. Business Understanding

Understanding the business can be referred to as understanding the research question, which is the origin of the research problem. This research aims to create a risk prediction model for potential impacts and issues among Thai higher education adolescents by applying text mining techniques to analyse the root causes of social media posting problems.

2.3.2. Data Understanding

Understanding data means understanding the context of the data used in the research. At present, data structures are mainly divided into three types: the first type is structured data, which is data in tables and patterns. The second type is unstructured data, which includes images, sounds, conversation sentences, frequency waves, etc., while the third type is semi-structured data, which has special symbols to indicate the meaning of the sentences or data presented.

This research thirdly uses data. While teens often present information in their posts with unstructured data, they also include special symbols such as Hash Tags (#), which indicate their mood and feelings at the time of the post. Therefore, using such data to develop risk prediction models for various adolescent problems is advantageous.

2.3.3. Data Preparation

Data preparation aims to obtain complete, appropriate, and ready data for modeling. Data preparation for this research consists of four parts. 1) Data cleaning: Correct errors and problems identified to create a complete and accurate data set. 2) Data structuring: A model must be created and organized to meet the analysis requirements. 3) Data transformation and enrichment: For data to have structure, it generally needs to be transformed into a unified and usable format. 4) Data validation and release: Automated data validation is performed on the data to check for consistency, completeness, and accuracy.

2.3.4. Modelling

Modelling involves selecting appropriate techniques to develop prototype models and validate and test them simultaneously. This research uses various supervised machine-learning techniques to obtain the most efficient model. Five techniques were selected: Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT). All five techniques can efficiently create classification models and predict desired outcomes [12].

2.3.5. Evaluation

The evaluation stage evaluates the model to obtain an efficient model, which researchers have summarized in the following section. The evaluation technique researchers use is data partitioning for testing, where the data is divided into two parts. The first part is the data used to build the model; it is called the training dataset, while the remaining data is prepared for testing the model, which is called the testing dataset.

2.3.6. Deployment

Deployment is putting the resulting model into practice, where it can be further developed into an application or public policy. This research aims to expand the most efficient model to construct a mobile application that can be connected to an online chat system to caution parents or those involved with teenagers who are at risk of various dangers from the results of this model.

2.4. Evaluation and Model Performance

As mentioned, the evaluation aims to select the most efficient model. The researchers used a technique called the Training and Testing dataset to split the data for testing. The researchers split the data into 80% of the training and 20% of the testing datasets. The technique and indicators used to evaluate the model performance are the Confusion Matrix and five indicators: accuracy, precision, recall, and f1-score values. The Confusion Matrix is an essential tool for evaluating the results of predictions made from a model created using machine learning techniques. The idea is to measure the model's predictive capability for what happens.

The accuracy indicator is the value that shows the overall correctness response rate achieved by using the created model. The precision indicator is the value that indicates the rate at which a given response type or class can predict a response. The recall indicator is the expected accuracy of the model by class, which is a sub-indicator. The f1-score is a harmonic mean of precision and recall created to serve as a single metric to measure the performance of a model. Finally, the support indicator is the frequency of the model predicted correctly. The results of the tests in each technique are shown in Tables 1 to 5 and Figures 2 to 6.

3. Results

3.1. Context of Adolescent Behavior

A study of the context of coping with mental health problems of adolescents through social media of adolescents studying in higher education by collecting data from 2023-2024 found that students at the University of Phayao have a vast six problems, which summarized the data collection as six issues follows: alcohol, infectious diseases, depression, gaming addiction, pregnancy problems, and drug problems. The data obtained is a collection of Thai words published on the link: https://shorturl.at/UJIlP. The data collection results included label 1: 1,003 depression posts (19.18%), label 2: 1,002 alcohol posts (19.16%), label 3: 1,004 gaming addiction posts (19.20%), label 4: 1,001 infectious diseases posts (19.14%), label 5: 466 drug problems posts (8.91%), and label 6: 754 pregnancy problems posts (14.42%), a total of 5,230 posts. From the collected data, it was found that most students at the University of Phayao have problems with game addiction, accounting for 19.20 percent. In addition, other issues follow in a similar proportion, such as depression, which may be a result of excessive addiction to phones and not participating in normal social activities, with a proportion of 19.18 percent. Alcohol problems account for 19.16 percent of the total. These problems need to be prevented.

3.2. The Constructed Model and its Performance

The generated models are reported according to five specified techniques: Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT) techniques. The results are reported by each method, which are shown in Tables 1 to 5 and summarized in Figures 2 to 6.

3.2.1. Naïve Bayes Model Evaluation

Table 1. Naïve bayes model evaluation

Class	Precision	Recall	F1-Score	Support
Label 1	0.6139	0.9652	0.7505	201
Label 2	0.8700	0.8700	0.8700	200
Label 3	0.9530	0.7065	0.8114	201
Label 4	0.8139	0.9400	0.8724	200
Label 5	0.9833	0.6344	0.7712	93
Label 6	0.8667	0.5166	0.6473	151
Macro Avg.	0.8501	0.7721	0.7871	1046
Weighted Avg.	0.8356	0.7983	0.7953	1046
Accuracy	0.7983			1046

Table 1 illustrates the tested results and model performance evaluation using a split-test 80:20 for the Naive Bayes technique. The model had a high accuracy value of 79.83%. In addition, it had high precision, recall, and f1-score values of 83.56%, 79.83%, and 79.53%, respectively. The model performance in the confusion matrix is presented in Figure 2.

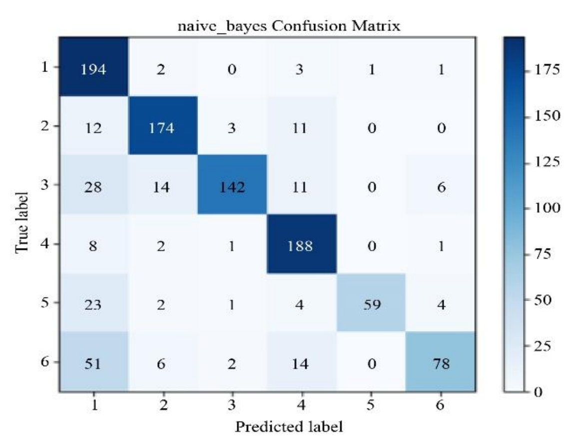

Fig. 2. Naïve bayes confusion matrix

3.2.2. Logistic Regression Model Evaluation.

Table 2. Logistic regression model evaluation

Class	Precision	Recall	F1-Score	Support
Label 1	0.8389	0.8806	0.8592	201
Label 2	0.9275	0.8950	0.9109	200
Label 3	0.8073	0.8756	0.8401	201
Label 4	0.9394	0.9300	0.9347	200
Label 5	0.9861	0.7634	0.8606	93
Label 6	0.7403	0.7550	0.7475	151
Macro Avg.	0.8732	0.8499	0.8588	1046
Weighted Avg.	0.8678	0.8633	0.8639	1046
Accuracy	0.8633			1046

Table 2 demonstrates the tested results and model performance evaluation using a split-test 80:20 for the Logistic Regression technique. The model had a high accuracy value of 86.33%.

In addition, it had high precision, recall, and f1-score values of 86.78%, 86.33%, and 86.39%, respectively. The model performance in the confusion matrix is presented in Figure 3.

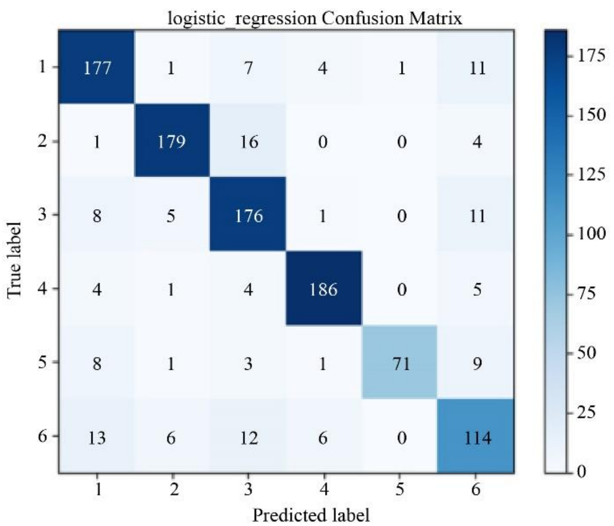

Fig. 3 Logistic regression confusion matrix

3.2.3. Support Vector Machine Model Evaluation.

Table 3. Support vector machine model evaluation

Class	Precision	Recall	F1-Score	Support
Label 1	0.8641	0.8856	0.8747	201
Label 2	0.8927	0.9150	0.9037	200
Label 3	0.8434	0.8308	0.8371	201
Label 4	0.9552	0.9600	0.9576	200
Label 5	0.9500	0.8172	0.8786	93
Label 6	0.7179	0.7417	0.7296	151
Macro Avg.	0.8706	0.8584	0.8636	1046
Weighted Avg.	0.8696	0.8681	0.8683	1046
Accuracy	0.8681			1046

Table 3 shows the tested results and model performance evaluation using a split-test 80:20 for the Support Vector Machine technique. The model had a high accuracy value of 86.81%. In addition, it had high precision, recall, and f1-score values of 86.96%, 86.81%, and 86.83%, respectively. The model performance in the confusion matrix is presented in Figure 4.

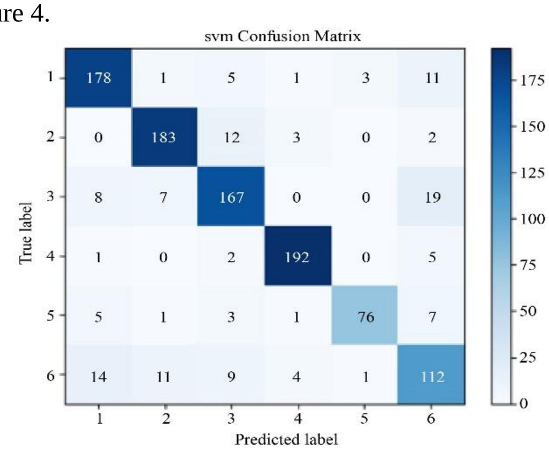

Fig. 4 Support vector machine confusion matrix

3.2.4. Random Forest Model Evaluation.

Table 4. Random forest model evaluation

Class	Precision	Recall	F1-Score	Support
Label 1	0.8564	0.8607	0.8586	201
Label 2	0.9175	0.8900	0.9036	200
Label 3	0.7660	0.8955	0.8257	201
Label 4	0.9550	0.9550	0.9550	200
Label 5	0.9383	0.8172	0.8736	93
Label 6	0.8209	0.7285	0.7719	151
Macro Avg.	0.8757	0.8578	0.8647	1046
Weighted Avg.	0.8717	0.8681	0.8681	1046
Accuracy	0.8681			1046

Table 4 depicts the tested results and model performance evaluation using a split-test 80:20 for the Random Forest technique. The model had a high accuracy value of 86.81%. In addition, it had high precision, recall, and f1-score values of 87.17%, 86.81%, and 86.81%, respectively. The model performance in the confusion matrix is presented in Figure 5.

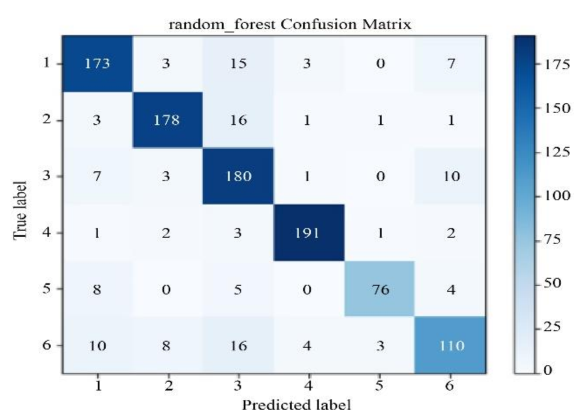

Fig. 5 Random forest confusion matrix

3.2.5. Decision Tree Model Evaluation.

Table 5. Decision tree model evaluation

Class	Precision	Recall	F1-Score	Support
Label 1	0.8586	0.8458	0.8521	201
Label 2	0.8627	0.8800	0.8713	200
Label 3	0.7692	0.7960	0.7824	201
Label 4	0.9337	0.9150	0.9242	200
Label 5	0.7528	0.7204	0.7363	93
Label 6	0.6358	0.6358	0.6358	151
Macro Avg.	0.8021	0.7988	0.8003	1046
Weighted Avg.	0.8150	0.8145	0.8146	1046
Accuracy	0.8145			1046

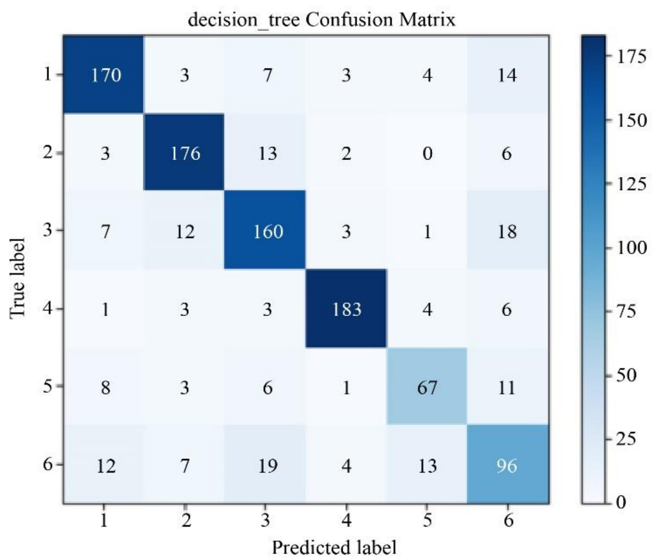

Fig. 6. Decision tree confusion matrix

Table 5 illustrates the tested results and model performance evaluation using a split-test 80:20 for the Decision Tree technique. The model had a high accuracy value of 81.45%. In addition, it had high precision, recall, and f1-score values of 81.50%, 81.45%, and 81.46%, respectively.

The model performance in the confusion matrix is presented in Figure 6. Tables 1 to 5 present the results of the model performance analysis using five techniques, including Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT). The results of the model performance study showed that the Support Vector Machine (SVM) and Random Forest (RF) techniques had the highest accuracy, 86.81%. The second and third-order models with the highest accuracy were the Logistic Regression (LR) and Decision Tree (DT) techniques, with 86.33% and 81.45% accuracy, respectively. In contrast, Naïve Bayes (NB) has the lowest accuracy, with an accuracy of 79.83 percent. However, all models are suitable because they have high accuracy values for all techniques.

4. Discussion

The context in which adolescents may be exposed to social media use has implications and roots in various problems. The leading cause is mobile phone addiction, which affects relaxation and sleep [7]. Lee et al. [7] presented an interesting article about the relationship between phone addiction and sleeplessness, which was conducted in low-income households. Similar to [13] and [14], which describe and summarize the impacts of phone addiction, including depression, social anxiety, and loneliness. Therefore, the context of this research data collection reinforces the importance of concern for caring for adolescents affected by social media.

Due to the problems and importance of this research, a model using artificial intelligence technology has been created as a tool for developing a prototype application for preventing teenage problems. Researchers have used text-mining techniques to analyze initial problem-solving approaches by mining emotional patterns, including post-expressions. This research is consistent with a wide range of other work using text mining to prepare for various problems [15], [16], [17], [18], such as text mining for identifying adolescent substance use and depression, using text mining to extract depressive symptoms, identifying addiction concerns on Twitter, etc. The significance and achievement of this research is that the researchers have developed a prototype model using machine learning techniques to provide a concrete solution for preventing adolescent problems.

5. Conclusion

Preventing and caring for adolescents during this transition period is always better than fixing problems that arise later. Coupled with communication between different generations, it makes it difficult for parents to cope. Therefore, this research aims to develop technology to solve adolescent problems using text-mining techniques with three research objectives: to study the context of college-educated adolescents' exposure to adolescent mental health problems through social media, to develop a predictive model of adolescent voice in social media posts using text mining analytics, and to evaluate the efficacy of the developed and selected prototype.

The data collected and used in the research were from five social media platforms, including Facebook, Instagram, Twitter, YouTube, and TikTok, by University of Phayao students from 2023-2024. The data obtained is a collection of Thai words published on the link: https://shorturl.at/UJIlP. The data collection results included label 1: 1,003 depression posts (19.18%), label 2: 1,002 alcohol posts (19.16%), label 3: 1,004 gaming addiction posts (19.20%), label 4: 1,001 infectious diseases posts (19.14%), label 5: 466 drug problems posts (8.91%), and label 6: 754 pregnancy problems posts (14.42%), a total of 5,230 posts. The technique for developing the predictive model is based on the concept of CRISP-DM, which consists of business understanding, data understanding, data preparation, modelling, evaluation, and deployment. Five techniques are used to develop predictive models: Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT). The development results found that the Support Vector Machine (SVM) and Random Forest (RF) techniques had the highest accuracy, 86.81%. The second and third-order models with the highest accuracy were the Logistic Regression (LR) and Decision Tree (DT) techniques, with 86.33% and 81.45% accuracy, respectively. In contrast, Naïve Bayes (NB) has the lowest accuracy, with an accuracy of 79.83 percent. With the success of the research, the results of this work are a push for researchers to be committed to developing practical programs and mobile applications to be applied to educational organizations and other organizations to be valuable and beneficial in the future.

Acknowledgments

This research project was supported by the Thailand Science Research and Innovation Fund and the University of Phayao. It also received support from many advisors, academics, researchers, students, and staff.

The authors thank everyone for their support and cooperation in completing this research. Moreover, thanks to Claude, ChatGPT, Sci Space, Gemini, and Perplexity Generative AI for their invaluable assistance in gathering information, reading, and providing guidance throughout this research.

References

[1] Yan Wang, “Web Emotion Recognition for College Students Based on Microblog Text Mining,” *Proceedings of the 2023 14th International Conference on E-business, Management and Economics*, New York, NY, USA, pp. 264-269, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Xin Chen, Mihaela Vorvoreanu, and Krishna Madhavan, “Mining Social Media Data for Understanding Students’ Learning Experiences,” *IEEE Transactions on Learning Technologies*, vol. 7, no. 3, pp. 246-259, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[3] Zuqin Lu, “Analysis Model of College Students’ Mental Health Based on Online Community Topic Mining and Emotion Analysis in Novel Coronavirus Epidemic Situation,” *Frontiers in Public Health*, vol. 10, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Robin A. Mcgee, David A. Wolfe, and Susan K. Wilson, “Multiple Maltreatment Experiences and Adolescent Behavior Problems: Adolescents’ Perspectives,” *Development and Psychopathology*, vol. 9, no. 1, pp. 131-149, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[5] Martina Benvenuti et al., “How Technology Use is Changing Adolescents’ Behaviors and their Social, Physical, and Cognitive Development,” *Current Psychology*, vol. 42, no. 19, pp. 16466-16469, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Alison Giovanelli, Elizabeth M. Ozer, and Ronald E. Dahl, “Leveraging Technology to Improve Health in Adolescence: A Developmental Science Perspective,” *Journal of Adolescent Health*, vol. 67, no. 2 Supplement, pp. S7-S13, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Joo Eun Lee et al., “Relationship between Mobile Phone Addiction and the Incidence of Poor and Short Sleep among Korean Adolescents: A Longitudinal Study of the Korean Children and Youth Panel Survey,” *Journal of Korean Medical Science*, vol. 32, no. 7, pp. 1166-1172, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mustafa Ozkan, and Betul Solmaz, “Mobile Addiction of Generation Z and its Effects on their Social Lifes: (An Application among University Students in the 18-23 Age Group),” *Procedia-Social and Behavioral Sciences*, vol. 205, pp. 92-98, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[9] Christoph Schröer, Felix Kruse, and Jorge Marx Gómez, “A Systematic Literature Review on Applying CRISP-DM Process Model,” *Procedia Computer Science*, vol. 181, pp. 526-534, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Rüdiger Wirth, and Jochen Hipp, “CRISP-DM: Towards a Standard Process Model for Data Mining,” *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, Manchester, pp. 29-39, 2000. [Google Scholar] [Publisher Link]

[11] Jacobus Venter, Alta de Waal, and Cornelius Willers, “Specializing CRISP-DM for Evidence Mining,” *Advances in Digital Forensics III, IFIP International Conference on Digital Forensics, National Center for Forensic Science*, Orlando Florida, USA, pp. 303-315, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[12] Devanshi Dhall, Ravinder Kaur, and Mamta Juneja, “Machine Learning: A Review of the Algorithms and Its Applications,” *Proceedings of ICRIC 2019: Recent Innovations in Computing*, pp. 47-63, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] Jin-Liang Wang, Jia-Rong Sheng, and Hai-Zhen Wang, “The Association between Mobile Game Addiction and Depression, Social Anxiety, and Loneliness,” *Frontiers in Public Health*, vol. 7, pp. 1-6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Dalia El-Sayed Desouky, and Hany Abu-Zaid, “Mobile Phone Use Pattern and Addiction in Relation to Depression and Anxiety,” *East Mediterr Health Journal*, vol. 26, no. 6, pp. 692-699, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Elizabeth M. Glowacki, Gary B. Wilcox, and Joseph B. Glowacki, "Identifying #Addiction Concerns on Twitter during the COVID-19 Pandemic: A Text Mining Analysis," *Substance Use and Addiction Journal*, vol. 42, no. 1, pp. 39-46, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Chi-Shin Wu et al., "Using Text Mining to Extract Depressive Symptoms and to Validate the Diagnosis of Major Depressive Disorder from Electronic Health Records," *Journal of Affective Disorders*, vol. 260, pp. 617-623, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Shi-Heng Wang et al., "Text Mining for Identifying Topics in the Literatures About Adolescent Substance Use and Depression," *BMC Public Health*, vol. 16, no. 1, pp. 1-8, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[18] Abdullah J. Sultan, "Addiction to Mobile Text Messaging Applications Is Nothing To 'Lol' About," *The Social Science Journal*, vol. 51, no. 1, pp. 57-69, 2014. [CrossRef] [Google Scholars] [Publisher Link]