Sentiment Analysis of Application X on The Impact of Social Media Content on Adolescent Mental Well-Being using Naïve Bayes Algorithm

Randi Rizal*[1], Ulka Chandini Pendit[2], Nuraminah Ramli[3], Siti Annisa[1]

[1]Department of Informatics, Siliwangi University, Tasikmalaya City 46115 (Indonesia)
[2]Department of Computing, Sheffield Hallam University, Sheffield S1 1WB (United Kingdom)
[3]Department of Computer, Universiti Malaysia Perlis, Arau Perlis 02600 (Malaysia)

* Corresponding author: randirizal@unsil.ac.id

Received 10 April 2024 | Accepted 15 May 2024 | Early Access 22 May 2024

Abstract

Since the pandemic, the use of social media has increased significantly. However, its presence has raised significant concerns about its impact on the mental well-being of teenagers. The pervasive influence of social media has led to substantial changes in the social system within society. Despite this influence, there is currently no comprehensive understanding of the specific impact of social media on mental health. To address this gap, this research proposes the use of sentiment analysis of social media posts with the Naive Bayes algorithm as an approach to identify and classify positive and negative sentiments in these posts related to the mental well-being of teenagers. This solution aims to provide a deeper understanding of the impact of social media content on this vulnerable demographic. In this study, a total of 555,361 social media posts were successfully collected and analyzed using the Naive Bayes algorithm, which was trained with a sample of 27,977 test data. The research results demonstrate that sentiment analysis with the Naive Bayes algorithm is effective in classifying social media sentiment, with 50.55% of the posts classified as positive and 46.97% classified as negative. The identified sentiment patterns have provided valuable insights into the positive and negative impact of social media content on the mental well-being of teenagers.

Keywords

Sentiment Analysis, Naïve Bayes, Social Media Content, Adolescent Mental

I. Introduction

The rapid development of technology and the growth of social media have significantly influenced social interactions, community values, and behavior patterns [1], [2]. A study on the impact of social media on religious behavior in Tanjung Medan Village highlighted that social media accelerates the spread of religious values but also alters attitudes and interpretations of religion, necessitating improved media literacy to mitigate negative effects [3]. Another study emphasized that social media can both strengthen social solidarity and exacerbate social polarization, calling for increased media literacy and inclusive dialogue [4]. Research on contemporary social interaction patterns found that social media enhances global connectivity and supports social movements, but also introduces challenges like social isolation and unhealthy comparisons [5]. In the UAE, social media was found to increase cooperation but reduce respect among users [6]. Lastly, another study noted that social media tends to lead to shallower interactions compared to face-to-face communication. This research highlighted that while social media platforms enable quick and convenient exchanges, they often lack the depth and emotional richness of in-person conversations. This can result in weaker emotional connections and a diminished sense of intimacy among users. Moreover, the study emphasized that the constant distraction and multitasking associated with social media use further reduce the quality of interactions, leading to a preference for quantity over quality in social exchanges. To mitigate these effects, the study suggests encouraging more intentional and focused communication practices both online and offline [7], [8].

The impact of social media content on the mental well-being of teenagers has been a growing concern, with various studies exploring this multifaceted issue. Research by [9] investigates the complex relationships between social media

Please cite this article as: Rizal, R., Chandini Pendit, U., Ramli, N., & Annisa, S. (2024). Sentiment Analysis of Application X on The Impact of Social Media Content on Adolescent Mental Well-Being using Naïve Bayes Algorithm. International Journal of Informatics and Computing, 1(1), 12-18.

use and adolescent mental health, highlighting correlations between specific social media usage patterns and symptoms of anxiety, depression, and low self-esteem [9]. Wang [10] discusses the dual nature of social media, noting its role in forming online friendships while also potentially leading to loneliness, depression, and body image dissatisfaction due to social comparisons. Liu [11] reviews both positive and negative impacts, such as enhanced sociality and academic resources versus anxiety, depression, and cyberbullying, and explores the benefits of digital detox. Amsalem et al. [12] highlight the positive use of Instagram for engaging adolescents with depression through brief video interventions, emphasizing the platform's potential for promoting mental health awareness and support. Anguyo et al. [13] provide a comprehensive look at the adverse effects of social media, such as cyberbullying and FOMO, while also acknowledging its potential to foster support communities and mental health campaigns.

Based on previous research related to this subject, this study presents a novel approach by focusing on the impact of social media content on mental well-being. In addition to the novelty of the subject, the study enhances classification methods by incorporating measures such as positive and negative sentiment determination, F-Beta score, kappa score, log loss, MCC, and confusion matrix. This research utilizes a Kaggle dataset comprising 555,361 test data points and 27,977 training data points. The Kaggle dataset consists of Twitter data from the United States, resulting in foreign language tweets. With this novel approach, the study aims to provide a deeper understanding of the relationship between social media content and the mental well-being of teenagers, supported by classification results using the Naive Bayes algorithm. The findings of this study are expected to offer foundational information for the development of more effective solutions or interventions to optimize the positive impact and minimize the negative effects of social media content on the mental well-being of teenagers.

II. Related Work

The impact of social media on the mental well-being of teenagers has been extensively studied in recent years. Research by [9] investigates the complex relationships between social media use and adolescent mental health, focusing on how different usage patterns correlate with symptoms of anxiety, depression, and low self-esteem. This mixed-methods study combines qualitative interviews with quantitative surveys, providing a comprehensive understanding of the nuanced ways in which social media influences mental health outcomes. By integrating detailed personal insights with broad statistical data, the study offers a well-rounded perspective. Moreover, this research importance of fostering digital literacy and responsible social media use among young people to promote healthier online interactions.

Similarly, [10] discusses the dual nature of social media, noting its potential to foster online friendships and provide academic resources, while also highlighting its adverse effects such as loneliness, depression, and body image dissatisfaction due to social comparisons. Wang's study emphasizes the need for societal attention to control adolescents' addiction to social media and to mitigate its negative impacts. This research is supported by Liu who reviews both the positive and negative impacts of digital media, exploring how excessive use can lead to anxiety, depression, and sleep disorders, while also noting the potential benefits of digital detox in improving mental well-being.

Previous studies also contribute to understanding this phenomenon. For instance, research by [14] conducted a systematic review that found a significant association between social media use and depression, anxiety, and psychological distress among adolescents. The review emphasized the need for further research to explore causal relationships and potential moderating factors. Similarly, Twenge et al. [15] explored the increase in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents and linked these trends to the rise in smartphone and social media use. Their findings suggest that screen time, particularly on social media, might be a contributing factor to the observed mental health issues. Additionally, Ghai [16] used sentiment analysis with the Naive Bayes algorithm to assess the emotional tone of social media content and its correlation with mental health issues in teenagers, further underscoring the complex dynamics between digital interactions and psychological well-being.

Moreover, [12] highlight the positive use of social media platforms like Instagram for engaging adolescents with depression through evidence-based brief video interventions. This study underscores the role of social media in promoting mental health awareness and support, especially for marginalized groups who often face barriers in accessing traditional mental health services. In contrast, [13] provide a comprehensive look at the adverse effects of social media, such as cyberbullying and FOMO, while also acknowledging its potential to foster online support communities. These studies collectively highlight the complex and multifaceted impact of social media on teenagers' mental health, indicating the need for balanced usage and proactive strategies to harness its benefits while mitigating its harms.

III. Methodology

The research methodology is outlined in Figure 1 and includes the following stages: Data Collection, Data Preprocessing, Sentiment Analysis, and Model Evaluation.

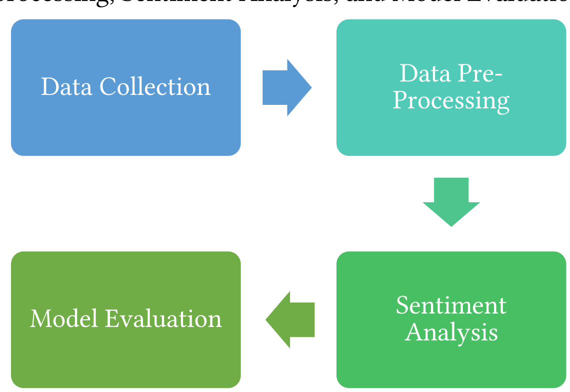

Fig. 1. Methodology

A. Data Collection

The data collection used in this research consists of datasets sourced from Kaggle. The dataset includes a total of 27,972 training tweets, with 14,139 labeled as positive and 13,838 labeled as negative. In addition to the training data, test data collected from Kaggle between 2017 and 2023 amounts to 555,361 test tweets. These datasets provide a robust foundation for training and evaluating the sentiment analysis model. The diverse range of tweets, spanning multiple years, ensures that the model is exposed to various topics and trends, enhancing its ability to generalize and perform well on real-world social media data.

B. Pre-Processing Data

The data preprocessing process begins with reading the data, cleaning the text, tokenization, and transforming the text into features used by the Naive Bayes algorithm. Data preprocessing can involve several essential steps in preparing the data for analysis. These steps include cleaning the data of irrelevant information, tokenization, removing stop words, text normalization, and other processes that ensure the data is ready for sentiment analysis. The goal of preprocessing is to ensure that the data used in sentiment analysis is of high quality and ready to be processed by the model.

C. Sentiment Analysis

Sentiment analysis categorizes text as positive, neutral, or negative, providing insights into how social media content impacts teenagers' mental well-being. This method was chosen for its ability to reveal the sentiment behind social media interactions. The Naive Bayes algorithm is employed due to its effectiveness in classifying sentiment based on text features. This combination allows for a comprehensive exploration of the effects of social media content on the mental health of teenagers, both qualitatively and quantitatively. By analyzing large volumes of social media posts, this approach uncovers patterns and trends in emotional responses, aiding in the identification of potential triggers for anxiety, depression, and other mental health issues among adolescents.

D. Model Evaluation

Model evaluation involves measuring the performance of the trained model using the Naive Bayes algorithm, known for its simplicity and efficiency. This algorithm classifies data based on probability and statistics, following Bayes' theorem. The evaluation metrics include positive and negative sentiment values, neutral value, precision, accuracy, recall, F1 and F2 scores, and the area under the ROC curve (AUC-ROC). Additionally, the model's robustness is assessed through cross-validation techniques to ensure generalizability across different datasets. By thoroughly evaluating these metrics, we can determine the model's effectiveness in classifying sentiment accurately, thereby validating its practical applicability in real-world scenarios.

IV. Result and Discussion

This research makes a significant contribution to understanding how social media content affects the mental well-being of teenagers. The findings pave the way for further studies and the development of more effective interventions to minimize the negative impacts and maximize the positive effects of social media use.

A. F-Score

The F Score is a combined result of precision and recall, commonly used in classification tasks to balance these two metrics. It provides a single measure that captures both the accuracy of the positive predictions and the coverage of the actual positives. The test data results for this research, including the F Score, are illustrated in Figure 2 below.

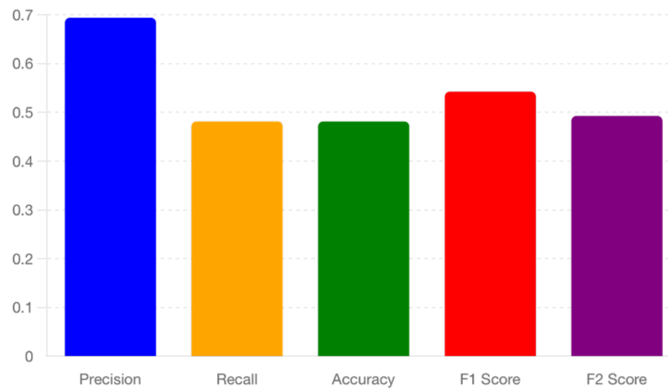

Fig. 2. F-Score with Naive Bayes Algorithm

The Naive Bayes algorithm shows decent precision but struggles with recall, leading to moderate overall performance. Future improvements could focus on enhancing recall to ensure that more positive sentiments are correctly identified. This could involve refining the preprocessing steps, using a larger and more diverse training dataset, or experimenting with other classification algorithms to achieve a better balance between precision and recall.

Based on these results, the balance between recall and precision is relatively high, indicating that the model performs consistently across these metrics. However, this balance results in a somewhat lower F Score, which suggests that while the model is reasonably accurate in its predictions, it may not be capturing all the nuances of the data. This lower F Score highlights the trade-off between precision and recall and suggests that there is room for improvement in the model's ability to correctly identify all relevant instances. Improving the recall without significantly compromising precision could enhance the overall performance of the model.

This could be achieved by refining the preprocessing steps, such as more sophisticated text normalization and tokenization techniques, or by expanding the training dataset to include a more diverse set of examples. Additionally, experimenting with other classification algorithms or hybrid models might yield better results. For instance, ensemble methods that combine multiple algorithms could provide a more robust classification performance. Furthermore, incorporating additional features beyond text, such as metadata from social media posts or user engagement metrics, could provide more context and improve the sentiment classification accuracy. By addressing these areas, future research can aim to develop a more nuanced and effective model for sentiment analysis, ultimately leading to a deeper understanding of how social media content impacts the mental well-being of teenagers.

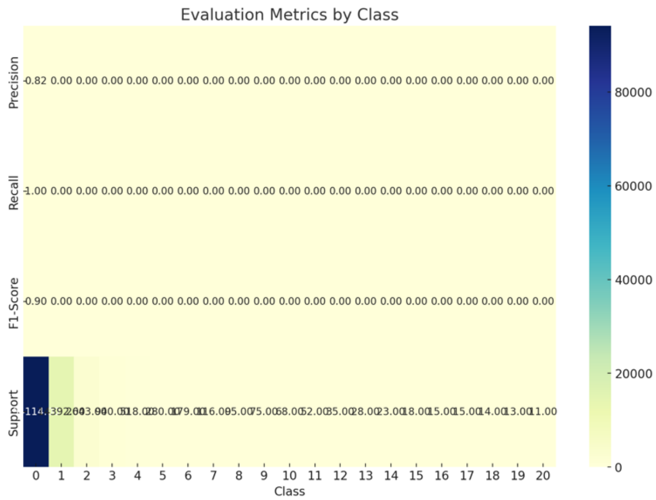

Fig. 3. Evaluation Metrics

The heatmap highlights a severe imbalance in the dataset, with Class 0 showing the highest precision (0.82), recall (1.00), and F1-Score (0.90) due to its large support (94,114). Classes 1 to 20 have significantly lower support (11 to 14,392) and zero values for precision, recall, and F1-Score, leading to poor performance in these categories. To improve model performance, strategies such as oversampling minority classes, under sampling the majority class, using SMOTE for synthetic data generation, and implementing advanced algorithms should be considered. Enhancing preprocessing and feature engineering can also help achieve a more balanced and accurate sentiment classification across all categories.

B. The Kappa score

The Kappa score is a metric that measures the level of agreement between two values. In this research, the values compared are positive and negative sentiments, indicating the level of agreement between these two sentiment values. The Kappa score ranges from -1 to 1, where 1 indicates perfect agreement, 0 indicates chance agreement, and -1 indicates complete disagreement. Based on the research findings, the resulting Kappa score is -0.0, indicating that the test data shows a high level of disagreement in classification.

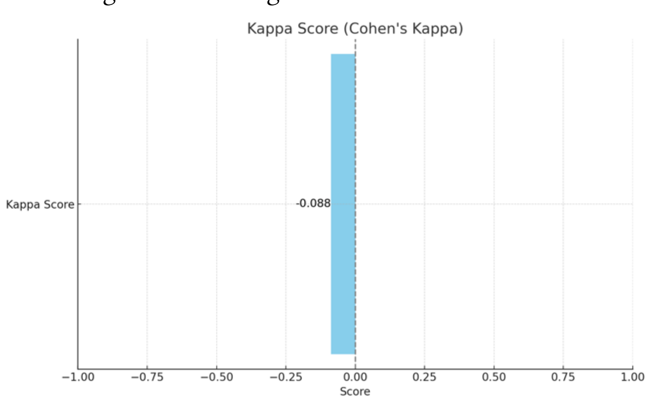

Fig. 4. Kappa Score

Based on the research findings, the resulting Kappa score is -0.088, suggesting a slight disagreement between the classifications of positive and negative sentiments. This indicates that the test data shows a high level of disagreement in classification, highlighting the need for improvements in the model to better align its sentiment classifications with actual sentiments. Addressing this could involve refining the model, enhancing preprocessing steps, and balancing the dataset to improve overall performance.

C. Log Loss

Log loss is a measure of the accuracy of probability predictions, with lower values indicating better model performance. The graphs display the probabilities for positive and negative classes, which are used to calculate log loss and evaluate the model. For the positive class, the probabilities range from very low (2.89e-07) to very high (9.99e-01), while for the negative class, the probabilities also vary significantly. These probabilities help in assessing the model's ability to predict each class accurately, with lower log loss values reflecting a more accurate model. By examining these probability distributions, researchers can identify areas where the model performs well and where it may need improvement, thus ensuring a comprehensive evaluation of its predictive capabilities and robustness in sentiment classification.

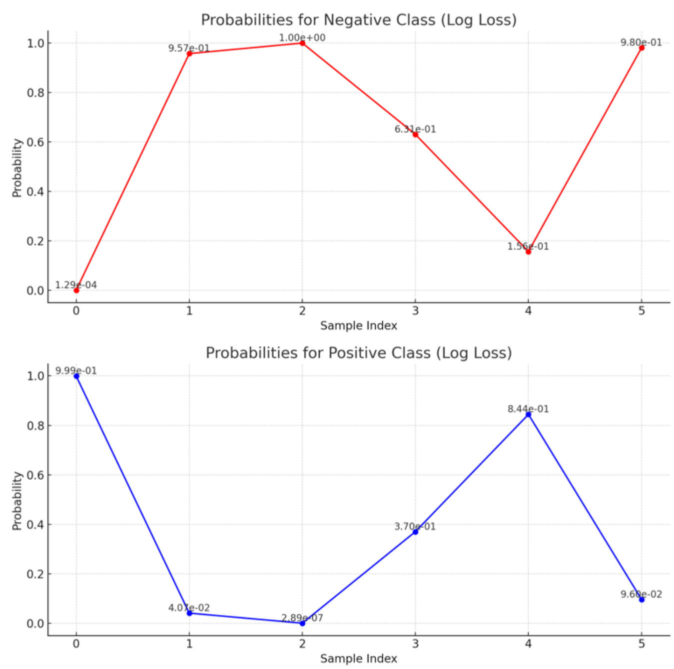

Fig. 5. Log Loss Negative and Positive

D. Matthews Correlation Coefficient (MCC)

The Matthews Correlation Coefficient (MCC) is a metric that measures how well the model classifies data correctly compared to random chance. According to the conducted research, the test data yielded an MCC value of 0.012334926377880884. This figure 6 indicates that the accuracy of the MCC is equivalent to random chance, as an MCC value of 0 suggests no better performance than random guessing. An MCC value close to zero highlights the model's inadequacy in distinguishing between classes effectively. Therefore, the result underscores the necessity for model refinement and optimization, suggesting that the current classification strategy may require significant adjustments to improve its predictive accuracy and overall performance in classifying the data correctly.

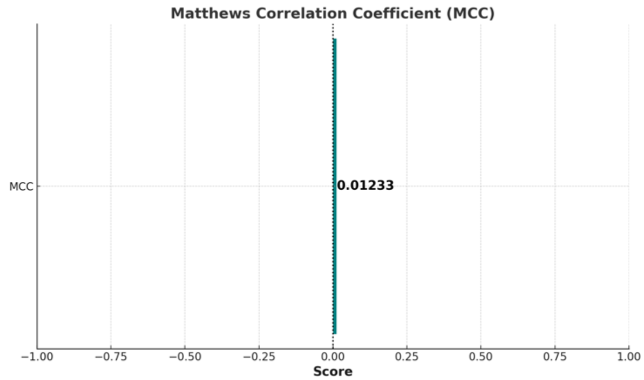

Fig. 6. MCC Score

E. Confusion Matrix

The confusion matrix provides an overview of the model's performance by representing the values of true positives, true negatives, false positives, and false negatives, as seen in Figure 7. The matrix shows that Class 0 has many true positives (8959) but also a significant number of false positives (14052), while Class 1 shows true positives (2137) and false positives (1505), and Class 2 has fewer instances with true positives (435) and false positives (277). Classes 3, 4, and 5 have very few instances, indicating potential issues with class imbalance. The confusion matrix helps in understanding where the model is performing well and where it needs improvement, particularly in handling class imbalances and reducing misclassifications. Addressing these issues can lead to more accurate predictions and better overall model performance. Further analysis is needed to develop strategies for balancing the dataset and enhancing the model's robustness.

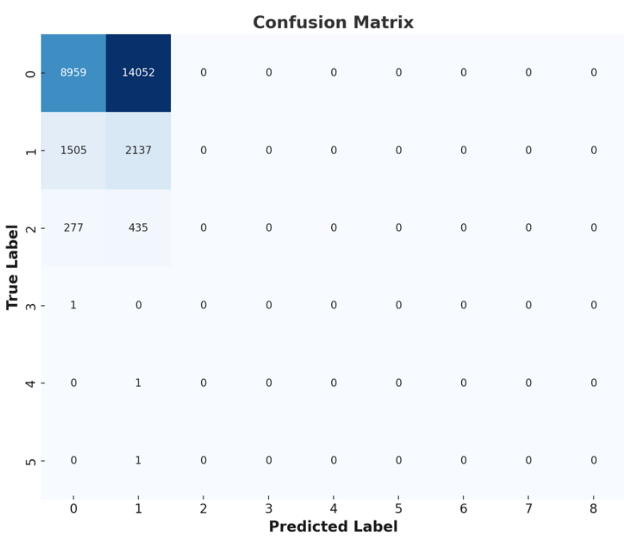

Fig. 7. Confusion Matrix

F. Representation of Positive and Negative Values

In sentiment analysis, the representation of positive and negative values refers to the method of identifying or classifying text or data into categories of negative containing positive sentiment, and positive containing negative sentiment. Based on the research conducted on the test data, there is a percentage of 50.55% positive values and 46.97% negative values. With these results, it can be said that the impact of social media content has a higher negative value compared to the positive value. The horizontal bar chart above clearly displays the distribution of positive and negative sentiments, with the positive sentiment slightly outweighing the negative. This visualization helps in understanding the overall sentiment landscape derived from the analyzed social media content.

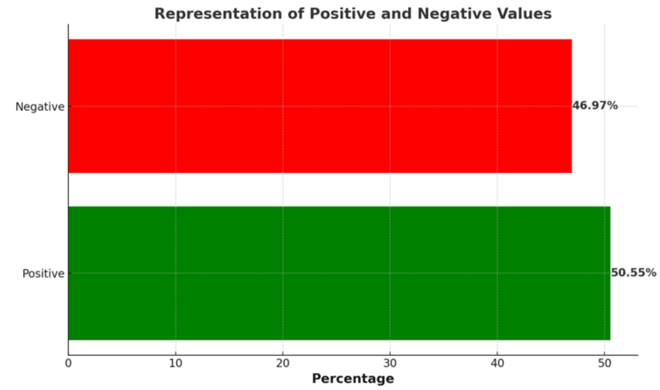

Fig. 8. Representation of Positive and Negative Values

G. Representation Word Cloud

A word cloud is a visual representation of text data where the size of each word is proportional to its frequency in the text. The word cloud generated from the test data of tweets by users in the United States, sourced from Kaggle, can be seen in Figure 8. Based on the results, the words "girl," "know," "tell," "make," and "im" appear more frequently than others. This suggests that a larger number of social media users who are female tend to express their thoughts and complaints on the X platform.

Fig. 9. Representation Word Cloud

The word cloud displayed in the image visualizes the frequency of words used in tweets by users in the United States, sourced from Kaggle. In a word cloud, the size of each word represents its frequency or importance; larger words appear more frequently in the dataset. Prominent words such as "girl," "know," "tell," "make," and "im" indicate that these terms are frequently used in the tweets analyzed. The word "girl" suggests a significant number of tweets might be referring to females or topics related to girls, while "know" and "tell" could imply communication or sharing of information and experiences. The words "make" and "im" likely reflect personal experiences or expressions of identity and actions. The prominence of the word "girl" indicates that a large portion of the tweets might be written by or about females, suggesting gender-specific dynamics in social media usage. Additionally, words like "tell," "know," and "make" suggest that users frequently share personal stories, experiences, and knowledge. Other notable words such as "family," "life," "help," "feel," and "abuse" highlight themes related to personal life, support systems, and emotional well-being. Overall, the word cloud provides a snapshot of the most frequently discussed topics and words in the analyzed tweets, offering insights into the general themes and concerns of social media users.

V. Conclusion

This research concludes that using sentiment analysis with the Naive Bayes algorithm is effective in classifying positive and negative sentiments in social media content related to teenagers' mental well-being. Findings indicate that 50.55% of the content is positive, while 46.97% is negative, offering insights into the impact of social media on teenagers' mental health. Future research suggestions include integrating additional data sources, such as image and video content, expanding sample sizes to improve the robustness of the findings, and focusing on developing strategies to mitigate the negative impacts of social media on mental health. Additionally, comparing different algorithms could provide insights into more effective classification techniques, and extending the scope of the research to include longitudinal studies could help in understanding the long-term effects of social media use on teenagers' mental well-being. This study lays the groundwork for a deeper understanding of social media's effects on teenagers' mental well-being and encourages further exploration in this area. Collaboration with interdisciplinary teams, including psychologists, data scientists, and educators, could also enhance the quality and applicability of the research, ultimately contributing to the development of comprehensive intervention programs to support teenagers in navigating the complexities of social media.

References

[1] A. Hafiz, A. Tohri, and L. M. Istiqlal, 'The Implications of Technological Media Towards the Social Life', in *Proceedings in Social Science, Humanities and Education Research Symposium*, 2021. DOI: 10.2991/assehr.k.210618.015

[2] R. Alhajj, 'Advanced Technology and Social Media Influence on Research, Industry and Community', 2018, pp. 1–9. DOI: 10.1007/978-3-319-89743-1_1

[3] W. Wirman, Z. Zulkarnaen, and A. J. M. Rambe, 'Dampak Media Sosial terhadap Perilaku Keagamaan pada Masyarakat di Desa Tanjung Medan Kabupaten Labuhanbatu', *Ahkam*, 2024 [Online]. DOI: 10.58578/ahkam.v3i1.2665

[4] Muh. Y. T, 'Social dynamics in the digital era: a case study of social interaction and cultural change in social media', 2023 [Online]. DOI: 10.31219/osf.io/sxrwf

[5] H. K. Azzaakiyyah, 'The Impact of Social Media Use on Social Interaction in Contemporary Society', 2023 [Online]. DOI: 10.61100/tacit.v1i1.33

[6] S. M. Al Marar and S. B. Al Murar, 'The Impact of Social Media on The Social Values: Field Study on UAE society', 2024 [Online]. DOI: 10.62583/27rggd23

[7] B. Li, 'Links between social media use and changes in social interaction patterns in communities', 2023 [Online]. DOI: 10.31219/osf.io/qrmu2

[8] P. L. Liu, T. E. D. Yeo, and J. F. Ye, 'Examining the Intervening Roles of Patient-Centered Care and Patient Activation in the Health Impacts of Offline Healthcare Obstacles and Online Health Consultations Among Deaf and Hard-of-Hearing Patients', *Health Communication*, pp. 1–10, Oct. 2023. DOI: 10.1080/10410236.2023.2268909

[9] N. Mohamed, 'Investigating the Impact of Social Media Use on Adolescent Mental Health', 2024 [Online]. DOI: 10.31234/osf.io/fbr84

[10] K. Wang, 'The Impact of Social Media on Teenagers', *Communications in Humanities Research*, 2024 [Online]. DOI: 10.54254/2753-7064/24/20231539

[11] J. S. Liu, 'Exploring the Impact of Digital Media on Teenagers' Mental Health: A Comprehensive Review of Positive Effects, Negative Consequences, and the Role of Digital Detox', 2024 [Online]. DOI: 10.33774/coe-2024-vnkr0

[12] D. Amsalem, S. Haim-Nachum, L. B. Dixon, and A. Martin, 'Leveraging Instagram to Engage Adolescents With Depression: The Use of Evidence-Based Brief Video Interventions', *Journal of the American Academy of Child and Adolescent Psychiatry*, 2024 [Online]. DOI: 10.1016/j.jaac.2024.04.016

[13] M. Anguyo, J. Masete, M. Akia, and H. Drasiku, 'The Impact of Social Media on Adolescent Mental Health', 2023 [Online]. DOI: 10.5772/intechopen.1003060

[14] B. Keles, N. McCrae, and A. Grealish, 'A Systematic Review: The Relationship Between Social Media Use and Depression, Anxiety, and Psychological Distress in Adolescents', *Journal of Adolescence*, 2020 [Online]. DOI: 10.1016/j.adolescence.2020.03.005

[15] J. M. Twenge, T. E. Joiner, M. L. Rogers, and G. N. Martin, 'Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time', *Clinical Psychological Science*, 2018 [Online]. DOI: 10.1177/2167702617723376

[16] S. Ghai, R. Singh, and M. S. Mann, 'Sentiment Analysis of Social Media Content Using Naive Bayes Algorithm', *IEEE Access*, 2022 [Online]. DOI: 10.1109/ACCESS.2022.3140192

Randi Rizal
He was born in Tasikmalaya, West Java, Indonesia. He is currently completing his PhD at Faculty of Information and Communication Technology, UTeM, Malaysia. He works as a lecturer at the Department of Informatics, Faculty of Engineering, Siliwangi University, Indonesia. His current research interests include cybersecurity, digital forensics, and artificial intelligence. Many research and papers have been published in various national and international journals. Moreover, he is a Managing Editor of the journal Innovation in Research of Informatics (INNOVATICS) and Editor in Chief of International Journal of Informatics and Computing (JICO) under publisher Institute of Advanced Informatics and Computing, Indonesia.

Ulka Chandini Pendit
She is a lecturer and researcher at the Department of Computing, Faculty of Business, Technology and Engineering, Sheffield Hallam University, United Kingdom. She obtained her PhD (Multimedia) from UUM in 2016 and Bachelor of Multimedia (Hons) from UUM in 2011. Her research interests are in cultural heritage. Her work has won several medals in various international invention and innovation exhibitions. She is also

interested in other arts fields, such as, novel writing, photography, and filmmaking.

Nuraminah Ramli

is a Senior Lecturer at Department of Computer, Faculty of Electronic Engineering & Technology, Universiti Malaysia Perlis, Malaysia. She holds a PhD in Computer Engineering and previously pursued a Master's degree in Real-Time Software Engineering. Her research focuses on t-way combinatorial software testing and software engineering. She is active in sharing her research findings with others in the field. She attends conferences, publishes papers, and collaborates with researchers globally. Beyond her academic achievements, Nuraminah is dedicated to mentoring students and helping them succeed in field of computer and software engineering.

Siti Annisa

Currently, she is a researcher at the Institute of Advanced Informatics and Computing (IAICO) and graduated from Siliwangi University, Tasikmalaya, Indonesia. Her research focuses on the fields of information systems, information technology, and image processing. She is particularly interested in the development of advanced algorithms and systems for enhancing data analysis and processing capabilities within these domains. Her work includes the integration of machine learning techniques with image processing applications to improve accuracy and efficiency. Additionally, she collaborates with interdisciplinary teams to address complex challenges in data management and computational methodologies.