

The Impact of Technology on Mental Health: Cyberbullying, Generative AI, And Sentiment Analysis Emotion-Aware Cyberbullying Detection and Intervention Using Transformer Models

Sayed Mahak¹, Aditya Saxena², Geetanjali Rautela³, Ritik Saxena⁴
^{1,2,3,4}Department of Computer Applications, Invertis University, Bareilly
doi.org/10.64643/IJIRTV12I11-207265-459

Abstract—The rapid growth of digital communication has significantly increased instances of cyberbullying, adversely affecting users' mental well-being through anxiety, depression, and social withdrawal. Existing systems largely rely on reactive moderation techniques, often failing to detect early signs of distress. This research proposes a proactive framework leveraging transformer-based models, such as BERT and GPT, implemented via the Hugging Face library, for real-time cyberbullying detection. The proposed system utilizes advanced Natural Language Processing techniques and sentiment analysis to capture contextual meaning and emotional nuances in text, enabling the identification of subtle indicators of harmful interactions. In addition, a sentiment-aware chatbot is integrated to provide adaptive and empathetic responses, suggest coping strategies, and offer immediate support during emotionally critical situations. The system also incorporates an alert mechanism to notify human moderators when high-risk scenarios are detected. The methodology includes data preprocessing, model fine-tuning, and performance evaluation using standard classification metrics. Key challenges such as sarcasm detection, bias mitigation, and data privacy are also addressed. The proposed approach aims to shift cyberbullying management from reactive filtering to proactive intervention, enhancing digital well-being through intelligent and timely support systems.

Index Terms—Cyberbullying, Mental Health, Natural Language Processing (NLP), Sentiment Analysis, Transformer Models, Hugging Face, Chatbots, Deep Learning, Text Classification, Digital Well-being.

I. INTRODUCTION

Day by day, living with screens shapes fresh ways people share feelings, build bonds, leave traces of

themselves. Platforms that ping back, chat apps lighting up pockets, forums buzzing late at night - each part of the background noise now. Though voices cross continents fast, though quiet users find crowds who listen, weight grows behind eyes staring at glow. Emotion bends under pressure from endless scrolls, quick judgments, lives half-shared. Distance blurs even when thumbs move swiftly across glass. Trouble hides in design meant to hold attention too tight.

One decade ago, signs began pointing toward rising mental health struggles worldwide, especially within younger populations. At the same time digital tools became part of daily life, scientists started questioning how screen habits might influence emotional states. Though apps and online platforms now offer new ways to access support - including virtual counseling, automated messaging services, and continuous tracking - these advances come alongside downsides like online harassment, pressure from idealized images, and constant connectivity stress.

What adds more layers here is tech such as generative AI systems that craft text nearly indistinguishable from human writing, opening paths for fresh ways to communicate while simultaneously fueling unease over fake content and robotic abuse.

Tools designed to spot feelings in written words stand apart; they give insight into mood patterns, sometimes flagging those who may struggle with psychological well-being.

These advances stretch both opportunity and worry across unseen terrain.

This study sets out to critically explore such patterns, tackling three central issues:

- How does technology influence mental health, both positively and negatively?
- What is the role of cyberbullying in shaping psychological outcomes in digital environments?
- How can emerging technologies such as generative AI and sentiment analysis be leveraged to improve mental health while minimizing risks?

II. PROBLEM STATEMENT

Digital tech's swift rise reshapes mental well-being, bringing chances alongside hurdles. Though apps like social networks and smart systems offer easier access to help and knowledge, they simultaneously fuel growing problems including stress, low mood, and loneliness. One pressing issue emerges through frequent online abuse, where constant digital attacks target users - particularly young people - and trigger deep emotional harm.

While tools such as generative AI and sentiment analysis begin shaping responses to mental health needs - offering ways to spot issues sooner and deliver assistance - they come with complications. These include producing damaging material, missing subtle human emotions, raising data privacy questions, and opening doors to abuse.

Even with more studies emerging, few models bring together ways to weigh technology's mental health trade-offs. Because of this gap, looking closely at digital effects - like online harassment, artificial intelligence that creates content, or emotion detection tools - matters greatly. Solutions must center people, hold up under scrutiny, avoid danger where possible, yet still deliver value when used.

III. THE DUAL NATURE OF TECHNOLOGY IN MENTAL HEALTH

A. Technology as a Positive Force

Now available to many more people, mental health support reaches farther due to digital tools. Because of technological advances, care once limited by location or expense opens up widely. Those held back by social judgment, high fees, or remote living find help through virtual therapy platforms. Web-based guidance materials also offer practical steps without needing

face-to-face visits. Peer networks thrive online, creating connection where isolation once dominated. Studies show digital platforms help teens express emotions, find peer backing, build self-image [11]. Because online groups act as protected zones, people open up about personal stories while gaining guidance this cuts loneliness.

Right now, artificial intelligence tools are playing a bigger role in helping people manage mental well-being. Because they operate instantly, chatbots give on-the-spot guidance when someone needs it most. Where access to therapists is limited, such digital helpers become especially useful. Their availability fills gaps that traditional services often leave open.

Studies indicate social media may encourage people to express themselves while also helping manage emotions [11] notes this alongside its role in fostering connections among peers.

B. Negative Effects of Technology

Even with its advantages, reliance on tech carries serious downsides for psychological well-being. Too much time online often links to higher rates of anxiety, low mood, and emotional strain. Platforms meant for connection may instead trigger unfavorable comparisons - prompting feelings of insufficiency or discontent among users. Looking at recent findings, more time spent online links closely to growing psychological issues along with dangers tied to cyberspace [14]. Harmful actions like bullying, spreading false information, or compulsive usage often grow stronger once they enter digital settings.

Connected devices mix private moments with shared ones, so mental downtime grows rare. Because people stay reachable nearly every hour, fatigue often builds without warning.

It magnifies joy just as easily as it spreads anxiety. A tool that reflects mood, sometimes distorts it. Through screens, feelings grow louder. Relief appears alongside distress. Not neutral, never passive. Its influence bends with how it's used. Moments of comfort emerge next to waves of pressure.

C. The Amplification Effect

What stands out when examining tech's impact on mental well-being is how it magnifies what's already there. Instead of introducing fresh emotional patterns, digital tools intensify habits people already have. Support among individuals may grow stronger,

spreading faster through networks. On the flip side, cruelty like harassment often escalates, lasting longer due to constant access.

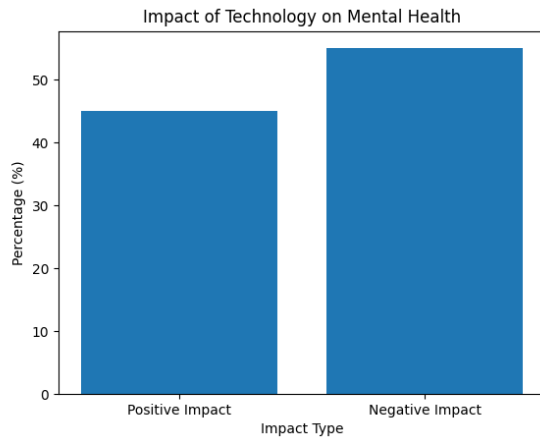


Fig. 1(a): Comparative representation of positive and negative impacts of technology on mental health

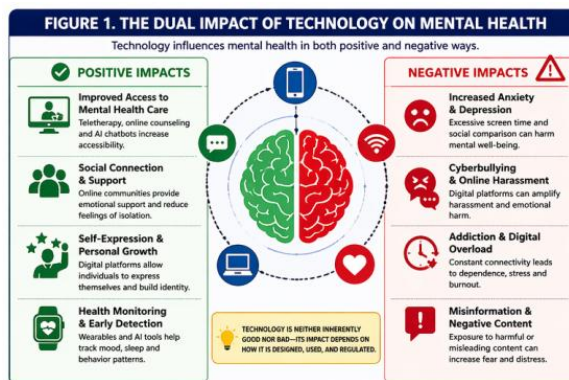


Fig. 1(b): Prevalence of mental health issues associated with high technology use (Compiled from [1]; [15]; [14]; [2])

IV. CYBERBULLYING AND MENTAL HEALTH

A. Concept and Distinctive Features

What happens when screens become weapons? Cyberbullying involves hurting people through digital tools - messages, posts, or images sent online. Not face-to-face, but still damaging. Distance doesn't soften impact; it sometimes makes it worse. The behavior spreads fast, often beyond control. Victims may feel trapped because escape isn't easy - even at home, the harassment follows.

Another twist: bystanders see everything yet rarely step in. Because visibility is high, shame can deepen. While physical bullying might stop after school ends,

this kind continues nonstop. Anonymity gives attackers cover, encouraging bolder actions. For these reasons, emotional wounds run deep - and recovery takes time

Anonymity, allowing perpetrators to avoid accountability Persistence, since harmful material stays available on the web Across vast regions, people gain access to observe or join. Accessibility extends beyond cities into remote areas. Through diverse channels, engagement becomes possible for many. Participation opens up through multiple pathways. Exposure grows as more individuals connect remotely. Inclusion reaches farther due to expanded networks What sets cyberbullying apart is how its traits amplify harm while resisting oversight.

B. Psychological Impact

A substantial body of research demonstrates the severe mental health consequences of cyberbullying: Besides feeling down, those affected often struggle with constant worry - self-worth tends to erode over time [1]. A link stands clear between thoughts of suicide and deep emotional pain [15]. Victims may develop long-term trauma and PTSD-like symptoms [13]. Emotional pain from online harassment often runs deep, with research showing teens particularly at risk of developing suicidal thinking [5]. Though invisible, the psychological weight emerges clearly through personal accounts collected in qualitative work.

C. Mechanisms of Psychological Harm

Cyberbullying affects mental health through multiple mechanisms:

i. Chronic Exposure

Stuck in a cycle of abuse, those targeted find no relief from ongoing pressure.

ii. Cognitive Overload and Rumination

Stuck in loops of negativity, the mind tightens its grip on pain. Rather than easing up, repeated rumination deepens discomfort. Thoughts that linger too long shift balance toward strain. Dwelling amplifies unease instead of quieting it.

iii. Social Withdrawal

Isolation often grows when people pull back from others. Social contact fades, step by step.

Studies show cyberbullying exerts ongoing mental strain because it follows people everywhere they go online [4].

iv. Links to Suicide and Serious Harm

Feelings of depression can stand between online bullying and thoughts about suicide [16].

Among those affected, major declines in psychological well-being appear clearly in broad research [12].

D. Interpretation:

Cyberbullying is not just harmful- it can be life-threatening, making it a critical public health issue requiring urgent intervention.

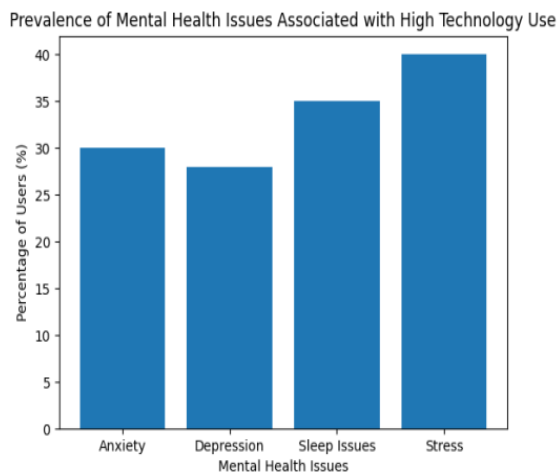


Fig. 2(a): Psychological effects experienced by victims of cyberbullying.

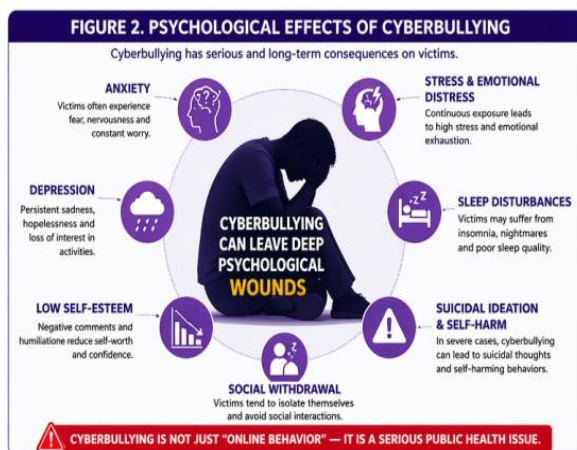


Fig. 2(b): Psychological effects experienced by victims of cyberbullying (Compiled from [1]; [15]; [14]; [2])

V. SENTIMENT ANALYSIS IN MENTAL HEALTH

A. Concept and Methodology

Beginning with raw text, sentiment analysis applies natural language processing alongside algorithms trained on data patterns. Through examination of vocabulary, sentence structure, or shifts in voice, it detects emotional signals hidden beneath surface words. Sometimes subtle cues reveal more than explicit statements this method captures those traces.

Classification emerges by linking textual traits to affective states: joy, anger, indifference appear in code-readable form. Rather than relying on human judgment, systems derive meaning from statistical regularities across vast examples.

B. Mental Health Applications

Precise identification of initial symptoms linked to depression and anxiety emerges through recent findings [3].

Tracking shifts in mental well-being across communities [17]. Monitoring mental health becomes possible without interruption or physical intrusion through these setups.

C. Role in Detecting Cyberbullying

Patterns of harassment get flagged by AI systems, spotting toxicity in speech [7]. Systems operating in real time help manage damaging material automatically, [6].

Interpretation: Sentiment analysis transforms mental health care from reactive to proactive.

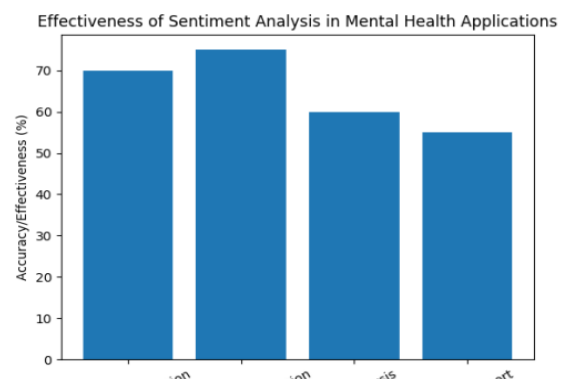


Fig. 3(a): Applications of sentiment analysis in mental health detection and monitoring

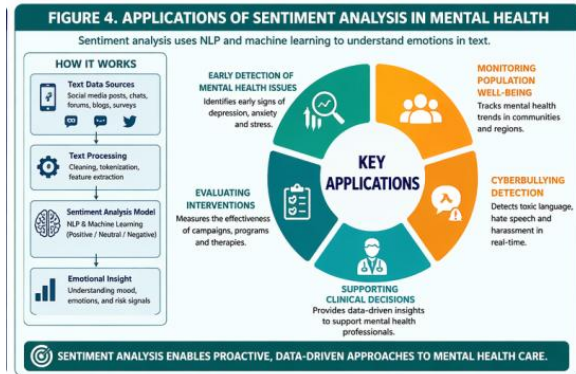


Fig. 3(b): Applications of sentiment analysis in mental health detection and monitoring (Compiled from [1]; [15]; [14]; [2])

VI. GENERATIVE AI AND MENTAL HEALTH

A. Opportunities

AI chatbots provide 24/7 mental health support. Besides cutting costs, these approaches ease social judgment while opening doors to care. Not everyone finds help easily - yet practical solutions emerge when access widens unexpectedly. Stigma fades somewhat where services become routine parts of daily life. Early detection becomes possible through their support, while also enabling ongoing tracking of emotional states. Studies suggest artificial intelligence could reshape access to psychological care [2].

B. Risks and Limitations

AI may generate incorrect or harmful advice. Emotional insight remains limited when set beside human capabilities. Potential for bias and misinformation. Still, artificial intelligence risks enabling damaging material - fueling online harassment, for instance [9]. Though useful, it might also spread verbal attacks more widely. With little oversight, such tools could intensify harm instead of reducing it.

C. Human-AI Interaction Challenges

Users may over-rely on AI for emotional support. Still, issues tied to honesty, openness, and responsibility carry weight today [8], [10]. Think of generative AI as something that backs up human effort, rather than taking its place. Interpretation: Though useful, it cannot mirror the depth of personal attention people bring. Its role fits

alongside caregivers, not instead of them. Human connection remains central when technology steps in. Care involves nuance machines do not grasp. So, balance matters - leaning on AI without letting go of what only humans provide. And, Generative AI must be viewed as a support tool, not a replacement for human care.

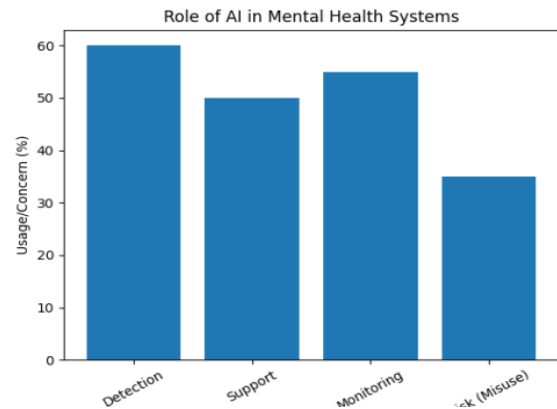


Fig. 4(a): Key roles and risks of AI in mental health systems

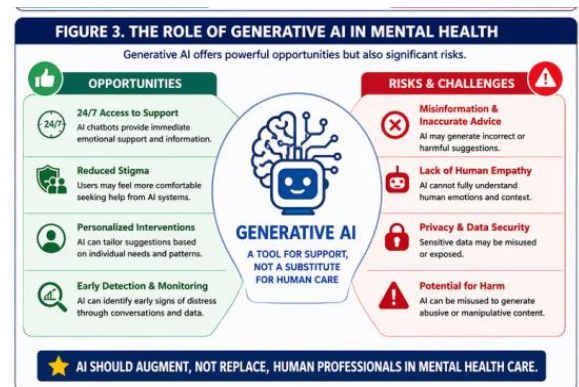


Fig. 4(b): Key roles and risks of AI in mental health systems. (Compiled from [1]; [15]; [14]; [2])

VII. PROPOSED SYSTEM: AI-BASED SYSTEM FOR CYBERBULLYING DETECTION AND MENTAL HEALTH SUPPORT

A. Background

The increasing use of social media and digital communication platforms has led to a rise in cyberbullying, which significantly impacts mental health. Traditional monitoring systems are often insufficient to detect harmful content in real time or provide timely psychological support. This case study presents a multi-layer AI-based system that integrates

cyberbullying detection, sentiment analysis, and a mental health chatbot to address these challenges.

B. Problem Context

Users on platforms such as social media, forums, and chat applications frequently encounter harmful interactions. These interactions can lead to emotional distress, anxiety, and depression.

There is a need for an intelligent system that can:

- Detect cyberbullying content
- Analyze emotional states
- Provide immediate support through automated intervention

C. Proposed Solution

A multi-layer AI system is designed to combine:

- Cyberbullying detection
- Sentiment and emotion analysis
- A mental health chatbot for intervention

D. System Architecture

i. Input Layer

The system collects data from multiple sources:

- Social media posts
- Chat messages
- Online forums

ii. Preprocessing Layer

The input data is cleaned and standardized through:

- Tokenization
- Stopword removal
- Lemmatization
- Handling emojis and slang expressions

iii. Cyberbullying Detection Layer

This layer uses advanced transformer-based models such as:

- BERT
- RoBERTa
- HateBERT

Tasks performed:

- Binary classification (bullying vs. non-bullying)
- Multi-class classification (types of bullying)

Rationale: Transformer models capture contextual meaning more effectively than traditional machine learning approaches.

iv. Sentiment and Emotion Analysis Layer

This layer evaluates:

- Sentiment polarity (positive/negative)
- Emotional states (anger, sadness, fear)

Transformer-based approaches provide higher accuracy compared to classical methods.

v. Decision Engine

Based on model outputs:

- Bullying content is flagged
- Emotional distress triggers intervention

vi. Mental Health Chatbot Layer

The chatbot is built using transformer models (e.g., DialoGPT, GPT variants).

Functions include:

- Providing empathetic responses
- Suggesting coping strategies
- Detecting crisis situations and escalating when necessary

E. Methodology

i. Data Collection

- Cyberbullying datasets from platforms like Twitter and Reddit
- Mental health datasets including conversational logs

ii. Model Training

- Fine-tuning transformer-based models
- Feature extraction using embeddings

iii. Evaluation Metrics

System performance is measured using:

- Accuracy
- Precision
- Recall
- F1-score

Hybrid AI models have demonstrated accuracy levels of up to 94–95%.

F. Implementation

i. Tools and Technologies

- Python
- Hugging Face Transformers
- PyTorch / TensorFlow
- Flask / FastAPI

ii. System Pipeline

1. Input text is received
2. Processed through a BERT-based classifier
3. Sentiment and emotion analysis is performed
4. Chatbot generates a response if needed

G. Results and Discussion

The system demonstrates several key outcomes:

- Improved cyberbullying detection using NLP techniques
- Superior performance of transformer models over traditional methods
- Effective scalability of chatbot-based mental health support

Limitations:

- Difficulty in detecting sarcasm and slang
- Ethical concerns such as bias and data privacy

VIII. ETHICAL ISSUES AND CHALLENGES

Key concerns include:

- Privacy violations from analyzing personal data
- Algorithmic bias affecting vulnerable populations
- Misuse of AI for harassment or manipulation
- Lack of regulation and accountability

Ethical frameworks and responsible AI design are essential for safe deployment.

IX. FUTURE DIRECTIONS

Future research and development should focus on:

- Digital resilience training for users [14].
- Improved AI accuracy and fairness.
- Integration of human oversight in AI systems
- Development of global policies and regulations

X. CONCLUSION

What shapes technology's effect on mental health isn't the tech itself, but human choices behind it. Access to support has widened because apps and platforms reach people who once had few options. A shift toward earlier intervention emerged as algorithms began flagging emotional shifts through behavior patterns. Still, darker outcomes trail close behind - harassment online grows harder to escape each year. Dependence on screens alters attention spans, often without users

noticing at first. Decisions coded into artificial intelligence can deepen bias if left unchecked.

What makes cyberbullying so harmful? It runs constantly, ignoring clocks and borders, often feeding feelings of sadness, fear, or even thoughts of ending one's life. Responses must come fast - schools, lawmakers, online spaces - all needing to move together, yet few solutions feel big enough just yet.

Even so, progress in understanding emotions through machines brings real potential. By moving beyond wait-and-see approaches, such tools might detect warning signs long before breakdowns occur - though success depends heavily on how clearly, they work. Accuracy matters just as much as fairness in handling personal information.

Looking ahead, mental health in a digital era hinge on equilibrium - where progress meets accountability, where machines complement emotional understanding. These approaches need to embrace diversity, respect cultural contexts, yet remain grounded in strong moral principles.

Surprisingly, tools meant to connect us might instead widen inner struggles. Whether they lift people up depends on nothing more than intentional choices in their application.

REFERENCES

- [1] Afrianto, "Combating cyberbullying and promoting adolescent mental health: A focus on evidence-based interventions," *J. Health Literacy Qual. Res.*, vol. 3, no. 1, 2023, doi: 10.61194/jhlqr.v3i1.539.
- [2] L. Balcombe, "AI chatbots in digital mental health," *Informatics*, vol. 10, no. 4, Art. no. 82, 2023, doi: 10.3390/informatics10040082.
- [3] A. Biró, A. Cuesta-Vargas, and L. Szilágyi, "Precognition of mental health and neurogenerative disorders using AI-parsed text and sentiment analysis," *Acta Univ. Sapientiae Informatica*, vol. 15, pp. 359–403, 2023, doi: 10.2478/ausi-2023-0022.
- [4] R. Dennehy, S. Meaney, M. Cronin, and E. Arensman, "The psychosocial impacts of cybervictimisation and barriers to seeking social support: Young people's perspectives," *Child. Youth Serv. Rev.*, vol. 111, Art. no. 104872, 2020, doi: 10.1016/j.childyouth.2020.104872.

- [5] S. Jabeen and N. Ehsan, "Horrors of cyberbullying: A qualitative study of cyberbullied adolescents," *J. Prof. Appl. Psychol.*, 2023, doi: 10.52053/jpap.v4i2.154.
- [6] R. K., "Spotting of cyberbullying on social media using machine learning," *Int. J. Sci. Res. Eng. Manage.*, 2024, doi: 10.55041/ijrem31868.
- [7] S. Mahbub, E. Pardede, and A. S. M. Kayes, "Detection of harassment type of cyberbullying: A dictionary approach and its impact," *Secur. Commun. Netw.*, vol. 2021, pp. 1–12, 2021, doi: 10.1155/2021/5594175.
- [8] K. J. Mukta and S. Akter, "Effect of social support on mental health: The study of social networks in times of crisis," *J. Posthumanism*, 2025, doi: 10.63332/joph.v5i7.2899.
- [9] C. Noval, "Implementation of artificial intelligence in the legal protection of cyber bullying victims on social media," *J. Social Sci.*, 2024, doi: 10.46799/jss.v5i6.945.
- [10] A. Polillo, K. Cleverley, D. Wiljer, F. Mishna, and A. Voineskos, "Digital disconnection: A qualitative study of youth and young adult perspectives on cyberbullying and adoption of detection tools," *J. Adolesc. Health*, 2024, doi: 10.1016/j.jadohealth.2023.11.395.
- [11] Y. Qian, "The effect of social media on mental health among adolescents," *Interdiscip. Humanit. Commun. Stud.*, 2024, doi: 10.61173/225rnr90.
- [12] S. K. Rajappan, A. A. Ali, A. F. Hamad, Z. Abbas, M. Alotaibi, F. A. Riyadh, S. Marzouq, and M. Suchithra, "Relationship between cyberbullying and mental health of teenage victims," *J. Family Med. Prim. Care*, vol. 14, pp. 2237–2241, 2025, doi: 10.4103/jfmpc.jfmpc_1592_24.
- [13] W. R. W. Rosli, S. N. Ya'cob, M. H. Abu Bakar, and M. S. M. Bajury, "Governing the risks of cyber bullying in the workplace during the COVID-19 era," *Malays. J. Soc. Sci. Humanit.*, vol. 6, no. 10, 2021, doi: 10.47405/mjssh.v6i10.1079.
- [14] I. N. F. Saputri, N. Solichah, S. A. Wahyu, A. F. Z. Ani, B. A. F. Aprasanjani, and U. F. Masykuroh, "Research development about digital resilience: A scoping review," *J. Indones. Psychol. Sci.*, 2024, doi: 10.18860/jips.v4i2.28048.
- [15] A. M. Schenk and W. Fremouw, "Prevalence, psychological impact, and coping of cyberbully victims among college students," *J. School Violence*, vol. 11, no. 1, pp. 21–37, 2012, doi: 10.1080/15388220.2011.630310.
- [16] R. Y. S. Subardjo, D. Mokhtar, M. R. Kamaluddin, N. S. M. Aun, Z. Khan, and Y. Jie, "Understanding the pathway from cyberbullying to suicidal ideation: Depression as a mediating factor," *Psikohumaniora: J. Penelit. Psikol.*, vol. 9, no. 2, 2024, doi: 10.21580/pjpp.v9i2.23357.
- [17] J. P. Varghese, A. Paul, J. Cherian, J. Joseph, A. P. Joseph, and A. Sebastian, "Mapping and visualizing the intersection of sentiment analysis and mental health," *Salud, Ciencia y Tecnología – Serie de Conferencias*, 2025, doi: 10.56294/sctconf20251543.