

# Educational Influences on Adolescent Mental Health: Associations and Gender Differences

Veera Venkata Raghunath Indugu<sup>1</sup>, Dedeepya Sai Gondi<sup>2</sup>, Vamsi Krishna Reddy Bandaru<sup>3</sup>, Narendra Lakshmana Gowda<sup>4</sup>

Mutual of Omaha, Exton, PA, USA<sup>1</sup>

Simplyturn Technologies, Princeton, TX, USA<sup>2</sup>

Simplyturn Technologies, The Colony, TX, USA<sup>3</sup>

Ashburn, VA, USA<sup>4</sup>

Corresponding author: Veera Venkata Raghunath Indugu, Email: [veera.indugu@ieee.org](mailto:veera.indugu@ieee.org)

This study explores the connection between schooling factors and teenagers' psychological health, concentrating on individuals aged 11 to 17. Relying on prior studies that emphasize the growing prevalence of mental health challenges among adolescents, the study aims to discover how schooling-related aspects—like academic success, school attendance, and stress tied to academics—influence mental health. By analyzing information from more than 8,000 participants, the study focuses on two major issues: pinpointing associations between education-linked variables and adolescent psychological health and exploring potential gender-related variations in these associations. Several statistical analyses were performed to evaluate the intensity and relevance of these connections. The findings were subsequently utilized to develop a classifier system, which assessed the significance of each variable. The study uncovers critical educational indicators of mental health and provides a foundation for future exploration into less-recognized ways that academic environments affect teenagers' mental well-being.

**Keywords:** Adolescents, Health, Mental, Psychological

## **1. Introduction**

In 2021, the National Health Service (NHS) conducted a survey on the mental health of adolescents aged 5 to 22 years. The findings showed that one in six children aged between 6 and 16 likely suffered from a mental disorder, a considerable increase from the one in nine children reported in a similar assessment in 2017. This trend is supported by the data from various studies conducted over the past two decades, which have documented growing levels of adolescent mental health problems across the globe. The increasing prevalence of mental health disorders has led to comprehensive research into the detection, management, and, most importantly to this study, the potential risk factors and implications of mental health challenges. The common risk factors reported include child maltreatment, overuse of social media, economic hardship, and social isolation. Additionally, poor mental health often correlates with negative outcomes like substance abuse, aggressive tendencies, strained peer interactions, and subpar academic achievement. While extensive research exists on adolescent mental health, relatively little attention has been given to the role of education-related elements in shaping mental health outcomes. This is alarming since children between 5 and 18 spend a significant part of their time in educational settings, which may influence their development of mental health [1–6].

Several patterns emerge between education and mental health. Studies have always shown that problems in mental health negatively impact results in education: reduced grades, lower school attendance, and high dropout rates. Conversely, academic underachievement, increased academic pressures, and unmet educational expectations are associated with worsening mental health outcomes. Therefore, this study will further explore these relationships by basing its exploration on prior literature to reveal new associations. Most existing studies have limited sample sizes, and age ranges, or even focus on certain aspects of the problem. The present study broadens the view by building upon existing evidence and finding other connections between educational factors and mental health. The study looks into data from over 8,000 students aged 11 to 17 in order to determine associations between educational factors and challenges in mental health. The study will be guided by two major questions: (1) What is the nature of the linkages between education-related factors and adolescent mental health for 11 to 17-year-olds? (2) What gender-specific differences are there in these relationships? Even though previous studies have highlighted gender-based patterns, there is no consensus regarding the role that gender plays in such dynamics. Relevant educational variables are identified and assigned a mental health score. Statistical analysis is conducted to determine the significance and strength of these relationships. Features that display meaningful associations will then be applied in a classifier model to judge their relevance and impact on outcomes that are predictive in nature. In doing so, this method intends to give full insight into the interaction between education and mental health and thus lay a foundation for more research in some of these under-explored areas. This approach is intended to provide comprehensive insight into the relationship between education and mental health and provide a basis for further research in relatively under-explored areas. Furthermore, this study incorporates the concept of intersectionality, that is, the interrelation between socio-economic status, cultural contexts, and family environments, in order to have a more holistic understanding of these complex relationships. The analysis is also concerned with exploring how the classroom environment, teacher-student relationship, and availability of academic support services influence adolescent mental health. Findings are expected to inform actionable recommendations for educators, policymakers, and mental health professionals about mitigating the risks associated with educational stressors. Mental health literacy programs, better teacher training, and more inclusive academic policies may emerge as pivotal interventions. This study is therefore not only about deepening understanding of the interplay between education and mental health but also about inspiring data-driven practices that foster resilience and holistic well-being among adolescents.

The study is as follows; the related works will be shown in the next section. The materials and methods are described in Section 3. The experimental analysis is carried out in Section 4, and in Section 5, we wrap up the investigation with some conclusions and recommendations for further research.

## **2. Related Works**

In 2021, NHS Digital reported that an estimated 17.6% of UK children aged 11–16 had a mental health disorder, which represented a significant increase from 12.6% in 2017. In the global arena, the World Health Organisation estimated that 14% of people aged 10–19 years had experienced mental disorders; the most common were anxiety disorders (8.2%), conduct disorders (6%), ADHD (5.5%), and depression (3.9%). The research underscores adolescence as a pivotal phase for mental health, given that 50% of lifelong mental illnesses appear by age 14, with 75% surfacing by the mid-20s. The growing rates of mental health conditions have far-reaching impacts, as 75% of adults facing such challenges report initial symptoms before turning 18. This rising trend in mental health disorders has occurred in concurrence with increasing public awareness over the last few decades. Mental health, which was a largely stigmatized topic in the 1950s, is changing attitude, and society now, in 2022, has become more accommodating, with governments substantially increasing funding for mental health research and treatment. For instance, the UK allocated about £15 billion to its mental health budget in 2022. Recent studies have further examined how adolescent mental health interacts with schooling. There is extensive evidence that has shown associations between mental health challenges and poorer outcomes in education. As an example, [7] described consistent associations of adolescent mental health challenges with lessened educational performance. [8] found that adolescents with mental health problems were twice as likely to fail in achieving five GCSE grades A\* to C and that males were adversely affected more significantly. [9] also found gender-specific differences where depressive symptoms and behavioural problems have been the core issues. However, studies by [10] found that there was no association between the internalizing problem of anxiety and academic performance; therefore, these debates are not yet over. Mental health also significantly impacts school attendance and dropout rates. According to [11] poor mental health at ages 13–14 increased the risks of dropping out, especially among females, although subsequent studies found similar trends among males. [12] found higher dropout rates among male students suffering from mental health problems in higher education. These findings confirm the strong association between mental health and educational outcomes but also point out gender-specific differences. Academic pressures are also known to affect adolescent mental health. The [13] mentioned that 66% of the students aged 15–16 reported experiencing major academic pressure. Poor academic performances, as presented by [14], were also correlated with increased levels of mental distress. High expectations from parents along with personal aspirations and strained familial relationships can heighten these factors. Other contributing factors include being forced to attend school for struggling students, having a negative perception of school, and having low school connectedness. These problems usually lead to emotional distress, loneliness, and even suicidal thoughts. For instance, [15] discovered that 59% of students with mental health issues reported low enjoyment of school. This extensive research opens up the otherwise complex relationship between education and mental health as a two-way bidirectional system. Therefore, this study draws on data from the Millennium Cohort Study to work toward advancing understanding of such critical dynamics and uncovering fresh insights in this field [16–25].

## **3. Materials and Methods**

The primary data source for this study is the Millennium Cohort Study, which covers the relationship of education factors to the mental well-being of teenagers. In this analysis, adolescent mental health is the dependent variable and the education-related factors are considered independent variables. Four leading stages are unfolded in terms of study methodology. The initial phase prepares the dataset and structures it into readiness for analysis, including the distinction between dependent and independent variables. The second stage applies frequency distribution analysis to ascertain the trends prevailing in the dataset. The third stage applies tests of statistical inferences to the educational variables under investigation to observe their association with mental health outcome variables. The last stage constructs a classification model that predicts the mental health statuses of the individuals. The model's variables are checked for relevance and their influence on the outcomes. The MCS is a

longitudinal survey conducted by University College London (UCL). It has been monitoring the lives of nearly 19,000 people born between 2000 and 2002 across the UK. It gives detailed information on physical, socio-emotional, cognitive, and behavioural development, in addition to daily life experiences. This study limits its scope to education and mental health data collected when the participants were 11, 14, and 17 years old. To maintain consistency, the sample size was reduced to just over 8,000 individuals across these three waves. The study utilizes copious longitudinal data and robust statistical techniques to analyze six complex connections between education and adolescent mental health. It aims at enhancing understanding and improving the mental health outcomes of youth by identifying the key relationships and predictors, ultimately contributing to the design of more effective interventions and policies that address adolescent mental well-being and educational support. Moreover, results will help instructors, policymakers, and mental health professionals design a focused approach which deals with all the issues facing adolescents at a particular phase of life; hence, creating a supportive climate for educational growth as well as emotional resilience to young people.

### **3.1 Dataset Analysis**

This study involved collecting and combining information from various sources, including discussions with group members and their guardians, teacher surveys, and academic records, to form a cohesive dataset for analysis. Eight different data files were combined to focus only on educational and psychological health factors, and a Cohort ID was established to ensure consistency across all surveys using anonymized identifiers. Missing values, denoted by negative values, were replaced with NaN entries, and the duplicate records were eliminated. The dataset was further stratified based on gender, which indicated a larger female proportion with a greater percentage of women, at 23.7%, as opposed to men, at 12.7%, identified with 'Abnormal' mental health statuses. The primary focus was mental health, evaluated at age 17, based on questionnaires, diagnoses, and suicide attempts. Mental health was divided into 'Normal' or 'Abnormal' based on SDQ and Kessler 6 scale scores. Independent factors included 92 educational factors, which were a combination of academic performance, attitudes towards education, behavioural problems, homework, academic ambitions, parental involvement, the atmosphere at school, and post-graduation activities. The independent and dependent variables were analyzed with frequency distributions and statistical tests like the Chi-Square Test of Independence and Cramer's V Test to show the significant association between independent and dependent factors. On analyzing, 12 features survived for the integrated dataset, but more differences could be seen among males and females. This study aimed to identify key factors related to mental health disorders in teenagers and their schooling as well as their backgrounds. There are many associations identified, especially on attitudes towards school, expectations from schooling, and behaviour problems, while the last dataset opens further modelling and prediction of mental health outcomes for youths.

## **4. Experimental Analysis**

The prime motivation behind developing the classifier model was to predict the mental health conditions of respondents with the help of several education-based attributes. Post-model training, it was hoped to assess the relative importance of the features involved through SHAP values to understand the effect of different attributes on model prediction. However, there were two main challenges ahead before the final model could be executed. The issue in the first scenario was a large amount of missing data in chosen features. This problem was not faced by earlier statistical tests as they inherently handle missing values. Therefore, classifier models that were to be tested did not have such features and so missing data had to be addressed as shown in Fig. 1. Two imputation strategies were tested: mode imputation, which fills in missing values with the most common value, and KNN imputation, which substitutes missing values with the closest available data point. These methods were applied to five classifiers to evaluate their impact, with KNN imputation yielding slightly better results for the decision tree and random forest classifiers. Consequently, KNN imputation was chosen for handling missing data in the next steps. The second challenge was the significant class imbalance between the 'Normal' and 'Abnormal' categories, with only 18.4% of participants categorized as

'Abnormal'. Without correcting this imbalance, the models would likely favour the majority class, artificially inflating accuracy. To counter this, oversampling was used to equalize the class distribution, and SMOTE was used instead of simple random duplication, creating synthetic samples and thereby avoiding overfitting. Once these problems were sorted out, the next step was to determine which classifier best suited the data. Five commonly used classifiers were tested: logistic regression, naïve bayes, k-nearest neighbour, decision tree, and random forest. Models were trained on selected features and cross-validated with k-fold, measuring accuracy, precision, recall, F1 score, and ROC AUC score, while performance was presented in confusion matrices. The lowest results were seen in logistic regression, followed by naïve bayes, and the best was performed by random forest.

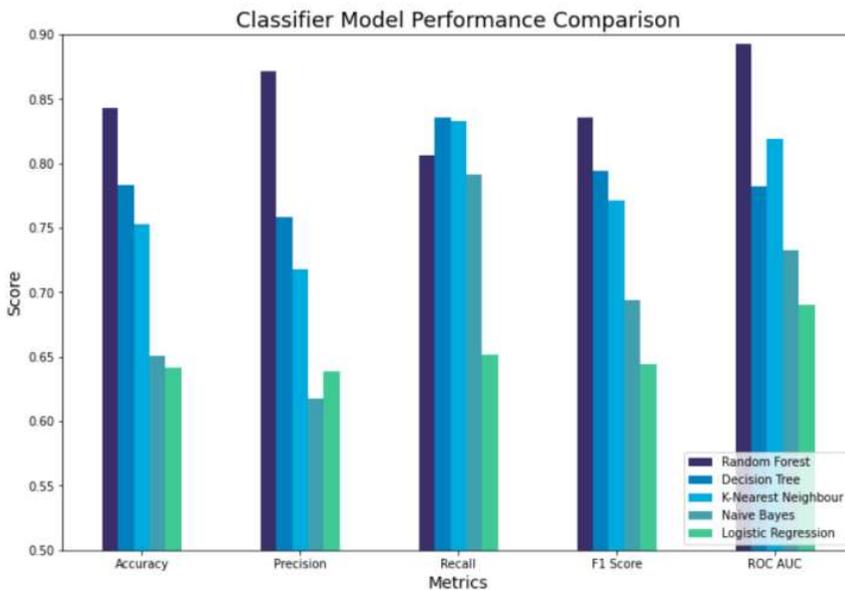


Figure 1. Bar graph comparing the performance of classifier models

After the selection of the final classifier, it was retrained on the data, and feature importance was analyzed using SHAP values as shown in Fig. 2. Utilizing SHAP, a game theory-based method, reduces complexity by providing explanations about the contribution of each feature to predictions. A positive value of SHAP indicates influence towards the 'Abnormal' class, whereas the influence is further placed towards the 'Normal' class in the case of negative values. The SHAP values threw light upon how educational features would contribute to mental health problems. The three models with the final feature set for each group were built on the combined, female, and male groups separately after training the random forest classifier. These models were trained using an oversampled dataset from which SHAP values were later calculated on the test data, keeping the original class proportions. These models were evaluated with k-fold cross-validation, and the performance metrics obtained were compared as shown in Fig. 3. The female group model performed the best on all the metrics, followed by the combined group. The male group model performed impressively with an accuracy of 83%, which is considered high in the industry. The strong results across all models confirmed the validity of the SHAP values used to assess feature importance.

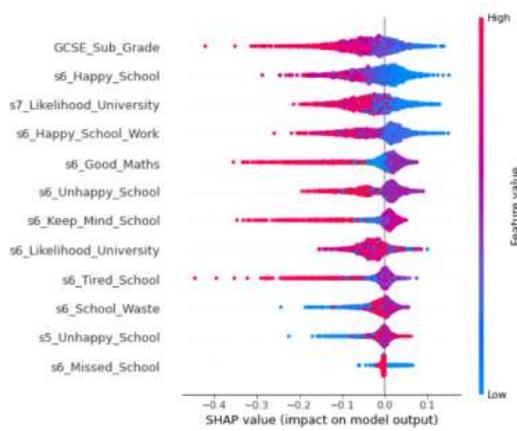


Figure 2. Summary graph of combined model SHAP values

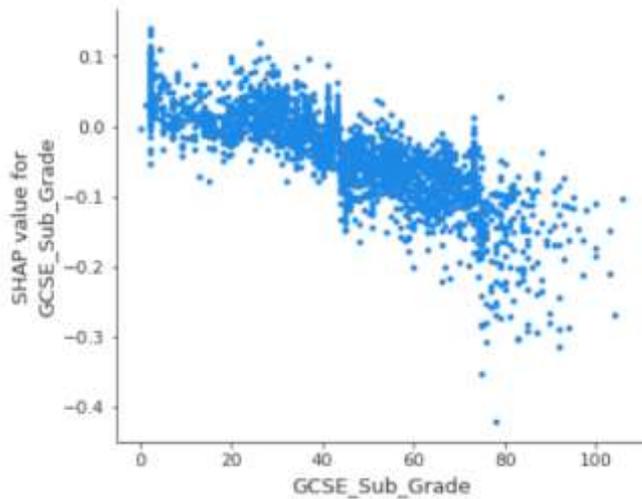


Figure 3. GCSE subject grade for the combined model

Each model ranked the features differently based on the distinct feature sets used as shown in Figs. 4 to 5. The most important features in the combined model were Happy with School (Sweep 6), and Likelihood of University (Sweep 7), whereas Missed School (Sweep 6), Unhappy at School (Sweep 5), and School is a Waste (Sweep 6) were relatively unimportant. The SHAP analysis further brought out how values within each feature influenced the final prediction, such that lower values for the key features Happy with School, Likelihood of University, and Happy with School Work are associated with the 'Abnormal' class, whereas higher values correspond to the 'Normal' class. However, the ordinal nature of some features means the associations do not always clearly reflect positive or negative relationships. For instance, Good at Maths, sweep 6, did not show a definite pattern in relationship to mental health, whereas School is a Waste, sweep 6, was suggestive that perhaps higher values could point to an unexpected positive relationship.

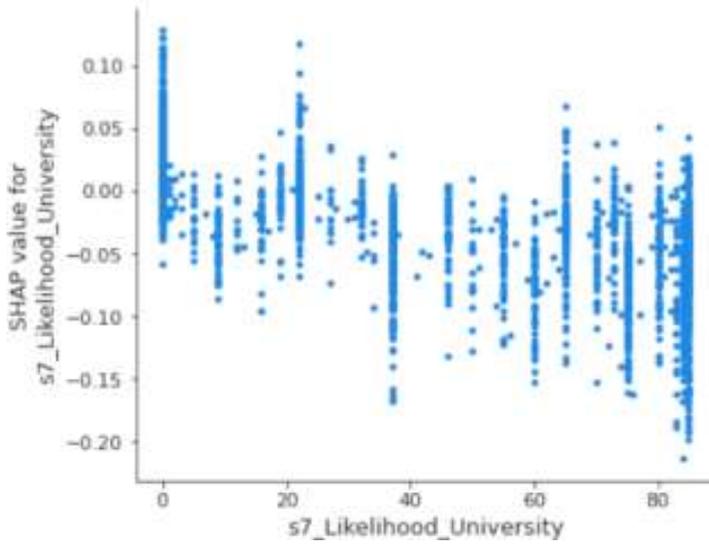


Figure 4. Combined model: university likelihood (sweep 7)

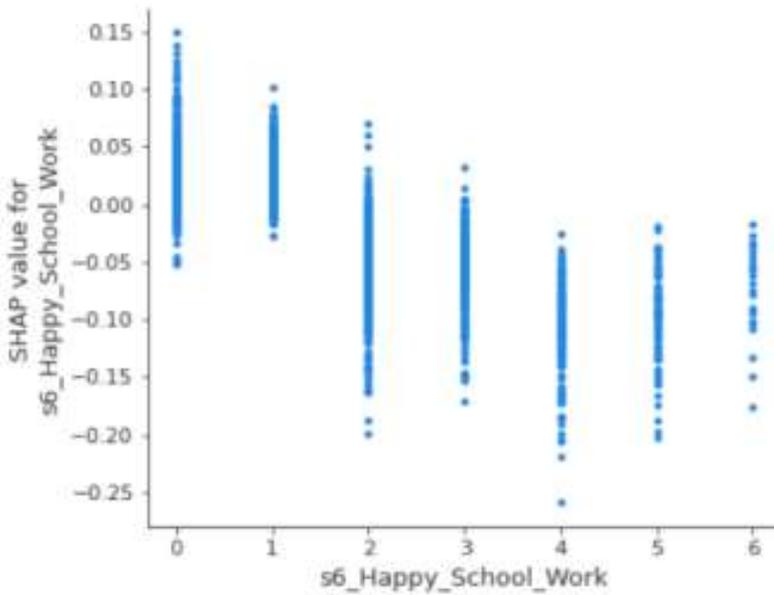


Figure 5. Combined model: content with academic tasks (sweep 6)

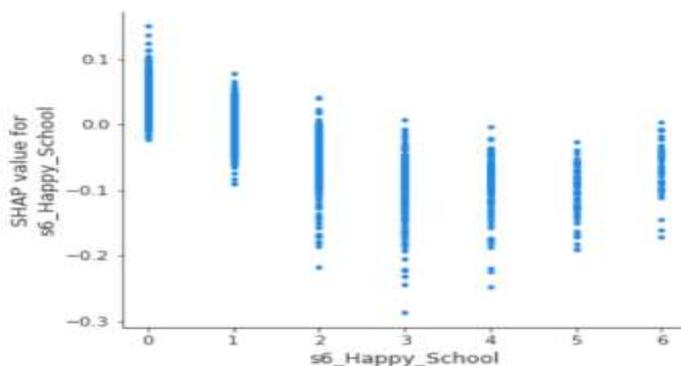


Figure 6. Happy with school: a combined model (sweep 6)

## 5. Conclusion and Future Directions

This study has several constraints that one evaluated when considering the outcomes. Initially, the MCS dataset contains private data, which restricts its accessibility unless thorough training and ethical approval are obtained. Consequently, academic performance information was collected through self-reports, which could lack full accuracy, thus affecting the dependability of the conclusions made regarding academic success. Furthermore, while the study intended to demonstrate a relationship between education and mental health, establishing clear causal links was unattainable due to the inability to satisfy the required conditions for causality. The intricacy of both education and mental health introduces numerous confounding variables, complicating the identification of a direct cause-and-effect relationship. Elements like race, socioeconomic background, and family circumstances were excluded from the study, thus limiting the scope of the study. Despite these limitations, the investigation reveals significant associations between educational factors and mental health, particularly among female subjects. Important findings include the impact of GCSE subject scores on females' mental health, the lack of similar correlations for male participants, and the relationship between school absenteeism and mental health. Moreover, the study underscores the importance of academic expectations, with higher expectations linked to worse mental health in females. Finally, the overrepresentation of females in the sample may account for the more noticeable gender differences identified.

## References

- [1] Kaur, P., Kashyap, G.S., Kumar, A., Nafis, M.T., Kumar, S., Shokeen, V.: From Text to Transformation: A Comprehensive Review of Large Language Models' Versatility. (2024).
- [2] Alharbi, F., Kashyap, G.S., Allehyani, B.A.: Automated Ruleset Generation for "HTTPS Everywhere": Challenges, Implementation, and Insights. *Int. J. Inf. Secur. Priv.* 18, 1–14 (2024). <https://doi.org/10.4018/IJISP.347330>.
- [3] Kashyap, G.S., Sohlot, J., Siddiqui, A., Siddiqui, R., Malik, K., Wazir, S., Brownlee, A.E.I.: Detection of a facemask in real-time using deep learning methods: Prevention of Covid 19. (2024).
- [4] Kanojia, M., Kamani, P., Kashyap, G.S., Naz, S., Wazir, S., Chauhan, A.: Alternative Agriculture Land-Use Transformation Pathways by Partial-Equilibrium Agricultural Sector Model: A Mathematical Approach. (2023).
- [5] Kashyap, G.S., Malik, K., Wazir, S., Khan, R.: Using Machine Learning to Quantify the Multimedia Risk Due to Fuzzing. *Multimed. Tools Appl.* 81, 36685–36698 (2022). <https://doi.org/10.1007/s11042-021-11558-9>.

- [6] Kashyap, G.S., Mahajan, D., Phukan, O.C., Kumar, A., Brownlee, A.E.I., Gao, J.: From Simulations to Reality: Enhancing Multi-Robot Exploration for Urban Search and Rescue. (2023).
- [7] Tielman, M.L., Neerincx, M.A., Bidarra, R., Kybartas, B., Brinkman, W.P.: A Therapy System for Post-Traumatic Stress Disorder Using a Virtual Agent and Virtual Storytelling to Reconstruct Traumatic Memories. *J. Med. Syst.* 41, 1–10 (2017). <https://doi.org/10.1007/s10916-017-0771-y>.
- [8] Large, M., Kaneson, M., Myles, N., Myles, H., Gunaratne, P., Ryan, C.: Meta-analysis of longitudinal cohort studies of suicide risk assessment among psychiatric patients: Heterogeneity in results and lack of improvement over time. *PLoS One.* 11, e0156322 (2016). <https://doi.org/10.1371/journal.pone.0156322>.
- [9] Gradus, J.L., Rosellini, A.J., Horváth-Puhó, E., Street, A.E., Galatzer-Levy, I., Jiang, T., Lash, T.L., Sørensen, H.T.: Prediction of Sex-Specific Suicide Risk Using Machine Learning and Single-Payer Health Care Registry Data from Denmark. *JAMA Psychiatry.* 77, 25–34 (2020). <https://doi.org/10.1001/jamapsychiatry.2019.2905>.
- [10] Arrabales, R.: Perla: A Conversational Agent for Depression Screening in Digital Ecosystems. Design, Implementation and Validation. (2020).
- [11] Almouzni, S., Khemakhem, M., Alageel, A.: Detecting Arabic Depressed Users from Twitter Data. In: *Procedia Computer Science.* pp. 257–265 (2019). <https://doi.org/10.1016/j.procs.2019.12.107>.
- [12] Miner, A.S., Shah, N., Bullock, K.D., Arnow, B.A., Bailenson, J., Hancock, J.: Key Considerations for Incorporating Conversational AI in Psychotherapy. *Front. Psychiatry.* 10, 746 (2019). <https://doi.org/10.3389/fpsy.2019.00746>.
- [13] Rabinowitz, N.C., Perbet, F., Song, H.F., Zhang, C., Botvinick, M.: Machine Theory of mind. In: 35th International Conference on Machine Learning, ICML 2018. pp. 6723–6738. PMLR (2018).
- [14] Billing, D.: Teaching for transfer of core/key skills in higher education: Cognitive skills. *High. Educ.* 53, 483–516 (2007). <https://doi.org/10.1007/s10734-005-5628-5>.
- [15] Duchesne, S., Caroli, A., Geroldi, C., Collins, D.L., Frisoni, G.B.: Relating one-year cognitive change in mild cognitive impairment to baseline MRI features. *Neuroimage.* 47, 1363–1370 (2009). <https://doi.org/10.1016/j.neuroimage.2009.04.023>.
- [16] Marwah, N., Singh, V.K., Kashyap, G.S., Wazir, S.: An analysis of the robustness of UAV agriculture field coverage using multi-agent reinforcement learning. *Int. J. Inf. Technol.* 15, 2317–2327 (2023). <https://doi.org/10.1007/s41870-023-01264-0>.
- [17] Habib, H., Kashyap, G.S., Tabassum, N., Nafis, T.: Stock Price Prediction Using Artificial Intelligence Based on LSTM– Deep Learning Model. In: *Artificial Intelligence & Blockchain in Cyber Physical Systems: Technologies & Applications.* pp. 93–99. CRC Press (2023). <https://doi.org/10.1201/9781003190301-6>.
- [18] Alharbi, F., Kashyap, G.S., Allehyani, B.A.: Automated Ruleset Generation for “HTTPS Everywhere”: Challenges, Implementation, and Insights. *Int. J. Inf. Secur. Priv.* 18, 1–14 (2024). <https://doi.org/10.4018/IJISP.347330>.
- [19] Kashyap, G.S., Siddiqui, A., Siddiqui, R., Malik, K., Wazir, S., Brownlee, A.E.I.: Prediction of Suicidal Risk Using Machine Learning Models, <https://papers.ssrn.com/abstract=4709789>, (2021).
- [20] Alharbi, F., Kashyap, G.S.: Empowering Network Security through Advanced Analysis of Malware Samples: Leveraging System Metrics and Network Log Data for Informed Decision-Making. *Int. J. Networked Distrib. Comput.* 1–15 (2024). <https://doi.org/10.1007/s44227-024-00032-1>.
- [21] Wazir, S., Kashyap, G.S., Malik, K., Brownlee, A.E.I.: Predicting the Infection Level of COVID-19 Virus Using Normal Distribution-Based Approximation Model and PSO. Presented at the (2023). [https://doi.org/10.1007/978-3-031-33183-1\\_5](https://doi.org/10.1007/978-3-031-33183-1_5).
- [22] Kashyap, G.S., Brownlee, A.E.I., Phukan, O.C., Malik, K., Wazir, S.: Roulette-Wheel Selection-Based PSO Algorithm for Solving the Vehicle Routing Problem with Time Windows. (2023).
- [23] Wazir, S., Kashyap, G.S., Saxena, P.: MLOps: A Review. (2023).
- [24] Naz, S., Kashyap, G.S.: Enhancing the predictive capability of a mathematical model for pseudomonas aeruginosa through artificial neural networks. *Int. J. Inf. Technol.* 2024, 1–10 (2024). <https://doi.org/10.1007/S41870-023-01721-W>.
- [25] Kashyap, G.S., Wazir, S., Hamdard, J., Delhi, N., Malik, I.K., Sehgal, V.K., Dhakar, R.: Revolutionizing Agriculture: A Comprehensive Review of Artificial Intelligence Techniques in Farming. (2024). <https://doi.org/10.21203/RS.3-RS-3984385/V1>.