IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Identifying Key Depressive Risk Factors Among Students Using Machine Learning Techniques

Passang Tshering Sherpa

Student of Department of Computer Science and Engineering,
Gulzar Group of Institutions, Ludhiana -141401, Punjab, India

Abstract: Depression is rapidly emerging mental illness issue that still could not discover the major factor and reasons. School and college did not have any distinct solution regarding depression and number is growing very high over the year on depression among college students. Analysis in this paper is based on the dataset source from Kaggle, including information on sleep quality, study satisfaction, academic pressure, dietary habit, family mental health illness history, financial stress and suicidal thought. We determined that the most important factor related with depression risk are suicidal thought followed by academic pressure and financial stress. On the other hand, of those who had suicidal ideation, we estimate that 13,900 have depression and only 3,700 do not. Multiple classification models including Logistic Regression, Support Vector Machine (SVM), Random Forest, Naïve Bayes, Decision Tree, and K-Nearest Neighbors (KNN) were implemented using Python in Google Colab. Among them, Logistic Regression and SVM achieved the best results, with 84% accuracy and 88% recall, indicating strong predictive performance. The most significant factors contributing to depression were found to be suicidal thoughts, academic pressure, and financial stress.

Keywords: Depressive, Mental Health, Machine Learning, Logistic Regression, Suicidal Thought, Academic Pressure, Risk Factor

1. INTRODUCTION

Depression has become a growing mental health concern among students in academic environments. It is commonly characterized by emotional fatigue, lack of interest in routine tasks, and difficulty maintaining focus, all of which can hinder both academic and social functioning. Although a certain amount of stress is typical for students, the shift to adulthood, higher academic standards, financial strains, and social pressures can cause long-term stress and, eventually, depression [1].

The problem is further complicated by the under diagnoses and under treatment of depression. Many students either fail to recognize their symptoms or avoid seeking help due to stigma and lack of access to mental health services. This delay in intervention increases the risk of severe consequences, including self-harm or suicide. As per the World Health Organization (WHO), Among individuals aged 15 to 29, suicide ranks as one of the leading causes of death, with untreated depression being one of the key contributors [2].Significant life changes and a rise in independence frequently occur during the college and university years, making pupils susceptible to mental health issues. Recent research indicates that anxiety and depression are very common among college students and are closely linked to poor academic performance and loneliness [3].

Many students avoid reaching out for professional help because of social stigma, limited awareness, or insufficient availability of mental health resources [4]. Overuse of social media, which encourages irrational comparisons and cyberbullying and contributes to low self-esteem and feelings of inadequacy, is also becoming more and more associated with depression in students [5]. Serious repercussions, such as substance abuse, self-harm, and suicidal thoughts, can result from depression if it is not treated [6], [18]. By examining the effects of academic pressure, financial stress, sleep quality, peer and family support, and social media use, the objective of this study is to identify and analyze the primary risk factors associated with depression among students by using machine learning techniques. We aim to uncover patterns and interactions among multiple personal and academic variables that influence depression risk. The motivation behind this research stems from the rising rates of mental health issues on campuses globally. While several studies have focused on individual causes such as academic pressure or social media usage, there is limited research examining how these factors interact together. By applying machine learning models to real student data, this study seeks to provide a multidimensional perspective and support early detection strategies in educational institutions [25].

2. RELATED WORK

Over the past two decades, researchers have increasingly focused on understanding student depression, examining both its frequency and the key factors that influence it. Students with depressive symptoms frequently found it difficult to meet academic demands, as shown in one of the earliest comprehensive studies by Andrews and Wilding [7], who reported a significant correlation between depression and academic performance. Their work emphasized a two-way relationship between mental health and academic achievement. Bayram and Bilgel [8], in their research on Turkish university students, reported that nearly 27% of respondents exhibited moderate to severe symptoms of depression. According to their findings, this prevalence was primarily linked to poor time management, concerns about the future, and academic stress. Similar patterns were noted in other global studies, indicating that student depression is a widespread issue cutting across both cultural and economic divides [1], [3], [6]. Zivin et al. [9], through longitudinal analysis, found that untreated depression during college could persist into adulthood, negatively impacting life satisfaction and job outcomes. This underscores the importance of early identification and institutional support systems. In their study of mental health service usage among students, Hunt and Eisenberg [10] revealed that despite the prevalence of mental health issues, very few students sought treatment. Barriers included stigma, lack of perceived need, and confidentiality concerns. These results support the need for better campus-based mental health resources and awareness campaigns [4]. Twenge et al. [11] reported a strong association between screen time—particularly on smartphones and social media—and increased anxiety and depression among adolescents and college students. Excessive digital engagement was shown to disrupt sleep and reduce meaningful social interaction. This is consistent with findings from Azem and Omer [26], who conducted a scoping review of 43 global studies and concluded that problematic social media use significantly contributes to emotional distress, low self-esteem, and poor sleep, especially among female adolescents. Similarly, Qirtas et al. [27] utilized behavioral and passive sensing data to model depression and loneliness among college students, identifying sleep duration and physical activity as major protective factors. Financial stress also plays a crucial role. Richardson et al. [20] found a significant link between prolonged financial hardships and elevated depressive symptoms in undergraduates. More recently, Krishnan et al. [30] developed a predictive model using HADS scores and found that financial difficulties, inadequate sleep, and certain academic fields were strong predictors of depression among higher education students. Academic pressure continues to be a major contributor, particularly in competitive educational settings [18]. McManus and Gunnell [19] observed an upward trend in self-harm and mental health problems among students over a 14-year span, reflecting a growing mental health crisis in youth. In addition, Jacka et al. [21] reported that unhealthy eating patterns, influenced by lifestyle and socioeconomic conditions, were significantly associated with depressive symptoms. While many studies have examined individual factors such as academic stress, social media use, or financial stress in isolation, few have addressed how these elements interact to shape mental health outcomes. Muzumdar et al. [29] addressed this gap using machine learning, showing that environmental and institutional factors significantly contribute to student mental well-being. Similarly, Rezapour and Elmshaeuser [28] emphasized the importance of academic and institutional satisfaction in reducing emotional stress, particularly during the pandemic-driven shift to online learning. Sinha and Gupta [31] applied ML and deep learning to detect depression and suicidal ideation from text, contributing insights into automated mental health detection. Additionally, Sudhakar [32] explored psychological stress detection through social media using machine learning, reinforcing the relevance of behavioral data in assessing mental health. Building on this foundation, the present study adopts a multidimensional approach by analyzing how academic pressure, financial hardship, sleep quality, social support, and digital behavior collectively influence depression levels in students [1], [3], [6], [9], [11], [20], [22], [26], [30],[31], [32].

3. RESEARCH METHODOLOGY

3.1 Data Collection

This research applied machine learning methods to determine risk factors among college students. This study uses a publicly available dataset sourced from Kaggle, an online data science platform. The dataset, titled Student Depression Dataset, which contained 27900 rows and 14 columns. The sample is 55.7% male and 44.3% female, the age range from 18 to 34. Several variables, each chosen for its potential influence on mental health, were considered within the research to investigate the causes of depression.

3.2 Data Pre-Processing

To ensure the dataset was ready for machine learning analysis, we performed several preprocessing steps:

1. **Binary Encoding:** Categorical binary responses such as "Yes"/"No" were converted into numerical format, where "Yes" was assigned a value of 0 and "No" a value of 1. This transformation was essential to enable the use of mathematical models.
2. **Label Encoding:** Multiclass categorical features like Degree (e.g., Bachelor's, Master's, PhD) and Gender (Male, Female) were encoded using integer values. This step allowed models to interpret qualitative data numerically without assuming any implicit order.
3. **Handling Missing Values:** Missing entries in the dataset were addressed using basic imputation methods. For numerical attributes, the median was used to replace missing values; for categorical attributes, the mode (most frequent value) was applied. This approach ensured data completeness without introducing significant bias.
4. **Feature Scaling (Standardization):** Continuous features such as Sleep Duration and Work/Study Hours were standardized using z-score normalization. This brought all features to a common scale with a mean of 0 and a standard deviation of 1, which helps prevent models from being biased toward features with larger numeric ranges.

These preprocessing steps enhanced model accuracy, improved training efficiency, and ensured compatibility with a wide range of machine learning algorithms by minimizing bias, handling data inconsistencies, and maintaining uniform feature representation

Table 1

Variable name	**Variable type**	**Variable description**	**Possible value**
Age	Predictor	Age of participant	18 to 34
Gender	Predictor	Gender of participant	Male, female
Profession	Predictor	Profession of participant	Student
Academic pressure	Predictor	Level of academic pressure of participant	1(lowest) to 5(highest)
CGPA	Predictor	Academic marks of participant	1 to 10
Study satisfaction	Predictor	Level of study satisfaction of participant	1(lowest) to 5(highest)
Sleep Duration	Predictor	The average number of hours the student sleeps per day	>5 hours,5-6 hours, 7-8 hours, <8 hours
Dietary Habits:	Predictor	An assessment of the student's eating patterns	Unhealthy, moderate, healthy
Degree	Predictor	The academic degree or program that the student is pursuing.	Class 12, and all the UG, PG and PHD course
Have you ever had suicidal thoughts ?	Predictor	Whether the student has ever experienced suicidal ideation.	Yes, No
Financial Stress	Predictor	A measure of the stress experienced due to financial concerns.	1(lowest) to 5(highest)

Family History of Mental Illness	Predictor	Indicates whether there is a family history of mental illness	Yes, No
Depression	Target	The target variable that indicates whether the student is experiencing depression.	1 (Depressed), 0 (Not Depressed)

3.3 Classification

I separated the dataset into two sections: A separate 20% of the dataset was reserved for assessing the performance of the trained models and the remaining 80% was used to train them. I worked on Google Colab as environment and Python as programming language as to apply machine learning techniques.

Using a different machine learning techniques, I investigated the variables effects college students likely to depression in this study. The algorithms selected for this study include Naïve Bayes, K-Nearest Neighbors (KNN), Random Forest, Decision Trees, Support Vector Machine (SVM), and Logistic Regression. These models were chosen due to their distinct data analysis approaches and their ability to uncover meaningful patterns within the dataset.

These techniques assisted me in understanding the relationships between a number of variables and their effects on depression, including study satisfaction, financial stress, sleep quality, academic pressure, family metal illness history and suicidal thought. I outline how each machine learning method works and why it was useful for this analysis.

3.3.1 Logistic Regression: Logistic Regression is a binary classification technique that estimates the likelihood of a particular outcome using a logistic function. It identifies the relationship between input variables and a binary target, outputting values between 0 and 1. The algorithm is valued for being straightforward, computationally light, and easy to interpret. It is especially effective when the predictors have a linear association with the output variable. Logistic Regression is widely applied in disciplines such as healthcare and behavioral sciences. It works best with minimal correlation between predictors and with a sufficiently large dataset [12].

3.3.2 Decision Tree: Decision Trees are supervised learning models that classify data by splitting it into smaller subsets based on certain feature thresholds. The model builds a tree structure where internal nodes represent decision conditions and leaf nodes represent output labels. They are intuitive, easy to visualize, and useful for both classification and regression problems [13].

3.3.3 Random Forest: Random Forest is an ensemble method that constructs multiple decision trees and merges their outputs to produce more accurate and stable results. It reduces the risk of over fitting that single trees might encounter and performs well on datasets with a combination of categorical and numerical attributes [14].

3.3.4 Support Vector Machine (SVM): Support Vector Machine is a supervised classification method that aims to find the best possible hyperplane to separate different classes in the feature space. It is capable of handling both linear and non-linear classification tasks through kernel functions. SVM is highly efficient for smaller datasets and provides solid performance in scenarios with clear class margins [15].

3.3.5 K-Nearest Neighbors (KNN): K-Nearest Neighbors is a non-parametric algorithm used for classification and regression. It assigns a label to a new data point based on the majority class among its k closest points in the feature space. The model assumes that similar observations tend to be near each other. Its accuracy depends heavily on the choice of k and the distance metric used. KNN is typically effective with small datasets having limited features [16].

3.3.6 Naïve Bayes (NB): Naïve Bayes is a probabilistic classifier based on Bayes' theorem. It assumes that all features are independent, which simplifies computation. It calculates the probability of each class and predicts the one with the highest score for a given input. This model performs especially well with large datasets and is widely used in applications like spam detection and text analysis, despite its simplifying assumptions [17].

4. RESULTS AND DISCUSSION

The study applied various machine learning models to determine the most influential predictors of depression among students. This section discusses the correlation of key variables, the performance of classification models, and the patterns observed in depression prevalence.

4.1 Correlation analysis

The analysis of factors contributing to depression among college students tell findings that are similar with previous studies, stronger the validity of the results. Figure 1 displays the correlation heatmap of all input features relative to depression. Notably, suicidal thoughts and academic stress have strong positive correlations with depression, while sleep quality and study satisfaction show negative associations. These correlations help identify the most impactful features for model training and possible explanations are discussed below:

1. Academic Pressure: Academic pressure is closely linked to depression (r = 0.47). This means that students who feel a lot of pressure in their studies are more likely to have experience depression. This previous studies also find that student's high expectations and the competitive academic environment often leads to anxiety and depressive symptoms when students struggle to keep up [18].

2. Suicidal Thoughts: Having suicidal thoughts showed the strongest connection with depression (r = 0.55), making it an important factor to focus on. Students who reported experiencing suicidal thoughts showed a strong association with depression. Similar findings were reported that who explained that suicidal ideation is often both a symptom and a consequence of severe depression. This overlap makes suicidal thoughts a critical indicator in depression studies [19].

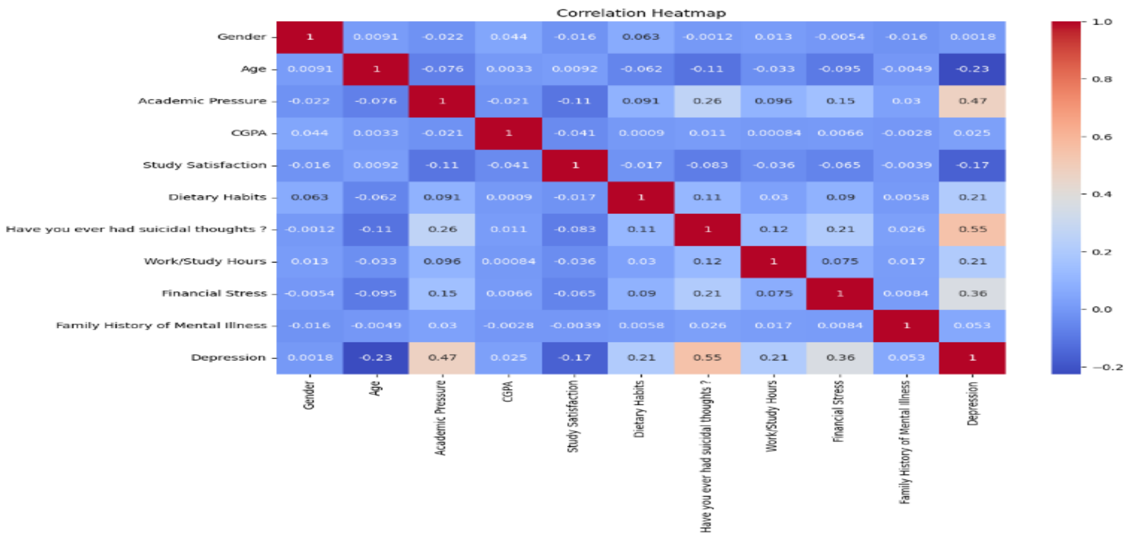

Figure 1

3. Financial Stress: Financial problems also play a big role in depression (r = 0.36). Students with financial difficulties are more at risk. Financial stress was identified as another significant factor contributing to depression. Past studies stated that economic difficulties can lead to a feeling helplessness, increased anxiety, and can lead to depressive symptoms in students who lack financial stability. Financial challenges often increase the level of anxiety and depression [20].

4. Dietary Habits: Diet also counts. An unhealthy dietary pattern showed a positive correlation with depression levels, with a correlation coefficient of (r = 0.21). Unhealthy eating habits, especially poor eating habits, were most closely related to depression. This also concurs with [21], who noted that poor nutrition may bring about a lack of the essential nutrients that are vital for mental health. The absence of healthy options for so many students could strengthen this correlation.

5. Study Satisfaction: Students dissatisfied with their studies are somewhat more likely to experience depression (r = -0.17). This is a weak relationship but significant. Low study satisfaction was associated with increased rates of depression. This finding is congruent with [6], who found that dissatisfaction with the academic experience tends to create a sense of frustration and lower self-esteem, leading to mental health issues.

6. Age: Younger students are slightly more at risk of depression compared to older ones (r = -0.23). The results indicated that depression rates were more prevalent among younger students than their older counterparts. Similarly find that younger individuals often experience higher depression rates due to life transitions, academic challenges, and societal pressures [23].

7. Family History of Mental Illness: If someone has a family history of mental illness, they may have a small increase in their risk of depression (r = 0.053). Students with a family history of mental illness exhibited slightly higher depression levels. The previous similar finding that the role of genetic predisposition in mental health issues. The family environment may also contribute to emotional vulnerability in such cases [24].

4.2 Key Findings

Logistic Regression identified the top predictors of depression risk among college students, as illustrated in Figure 2. The most significant feature was "Suicidal Thought", contributing approximately 27% to the model's prediction power, followed by "Academic Pressure" (22%), "Financial Stress" (18%), "Dietary Habits" (12%), and "Study Hours" (9%). Suicidal thought and academic pressure had strong positive coefficients, indicating that students experiencing these stressors were considerably more likely to be depressed.

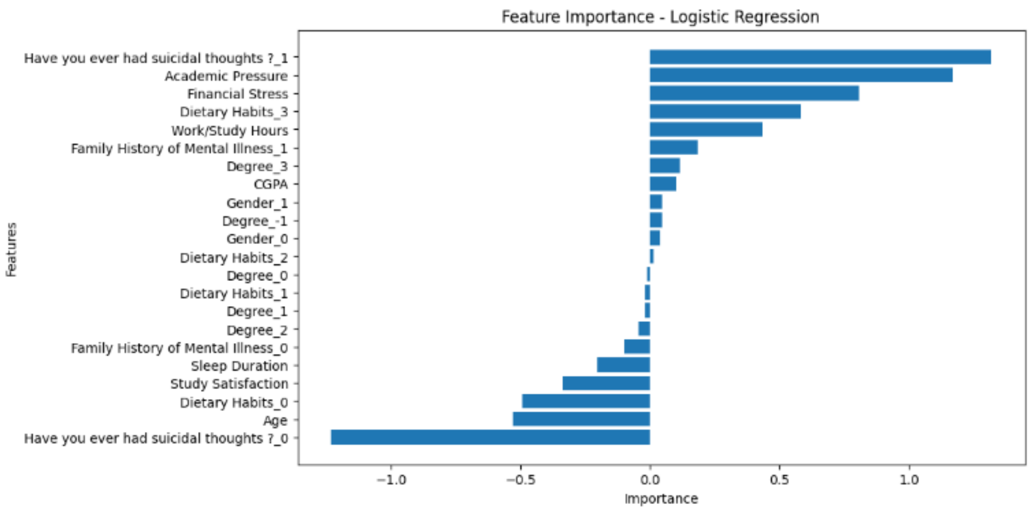

Figure 2

In contrast, Sleep Duration, which accounted for approximately 7% of predictive influence, showed a negative coefficient. This suggests that students who reported better sleep quality were less likely to experience depression. These findings are consistent with previous research [22], which emphasizes academic pressure as a major factor in mental health deterioration among students. Figure 1 presents the correlation heatmap, where "Suicidal Thought" showed the highest positive correlation with depression (r = 0.55), while "Study Satisfaction" was negatively correlated (r = –0.17). Age also played a role: students aged 18–20 years showed depression rates of nearly 58%, whereas students aged 21–23 years showed lower rates, around 42%, possibly due to improved emotional regulation or life experience. Similarly, students who expressed high academic satisfaction had a depression rate of only 30%, compared to over 60% among those reporting dissatisfaction. This supports the idea that positive academic experiences help buffer the effects of stress.

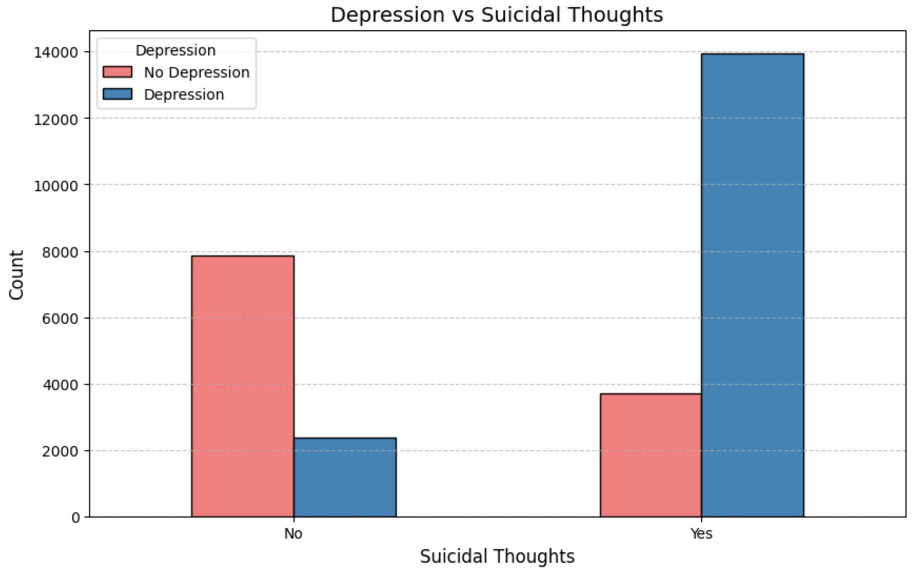

Figure 3

Figure 3 shows the distribution of depression status among students based on their response to suicidal ideation. Among students who reported no suicidal thoughts, approximately 7,900 were not depressed, while about 2,400 were depressed. In contrast, among those who reported suicidal thoughts, only around 3,700 were not depressed, whereas a significantly larger group nearly 13,900 were depressed. This stark difference demonstrates a strong relationship between suicidal ideation and depression. Students who reported suicidal thoughts were nearly four times more likely to be depressed than not.

These findings reinforce the need for early detection of suicidal ideation, as it may serve as a key indicator of severe depression. Timely mental health screening and targeted intervention programs are essential for at-risk students to prevent progression to critical conditions.

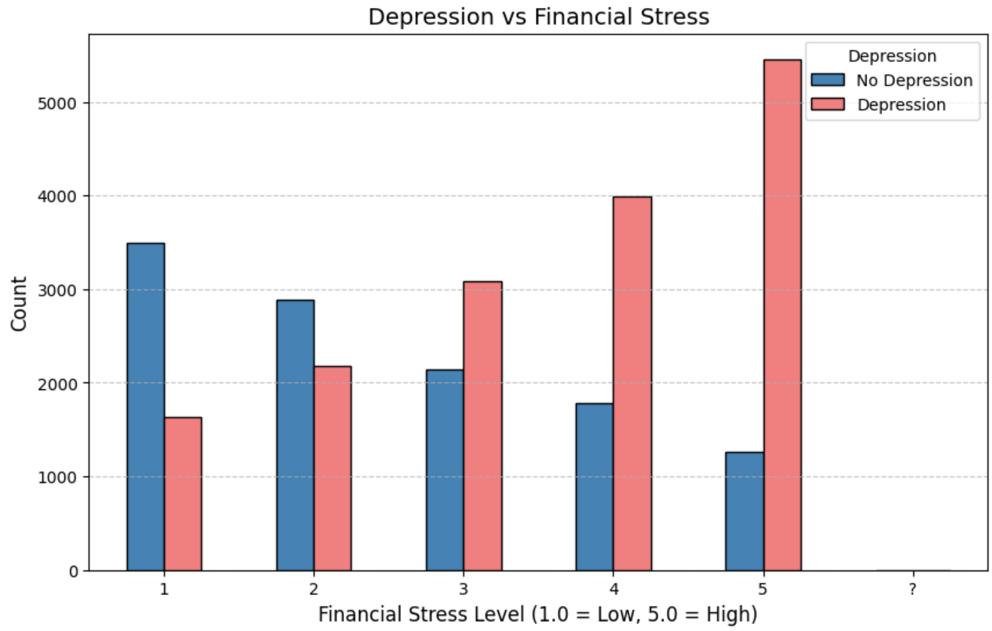

Figure 4

Figure 4 provides a detailed visualization of the relationship between financial stress levels and the presence of depression among students. The x-axis represents financial stress on a scale from 1 (very low stress) to 5 (very high stress), while the y-axis displays the number of students within each stress level, segmented by depression status. At the lower end of the stress scale (levels 1 and 2), individuals without depression are more prevalent. Approximately 65–70% of students in these categories reported no depressive symptoms, suggesting that low financial stress may serve as a protective factor against depression. In contrast, only 30–35% of individuals in these groups experienced depression, highlighting a lower vulnerability when financial concerns are minimal. A significant shift occurs starting from stress level 3, where the count of depressed individuals begins to exceed those without depression. This trend becomes more pronounced at levels 4 and 5, where financial strain is at its peak. At stress level 5, nearly 80% of students are reported as depressed, while only around 20% are not. This sharp increase illustrates a clear, positive correlation between financial stress and the likelihood of depression. The data strongly suggest that as students face more intense economic pressure—due to factors such as tuition fees, living expenses, lack of family income, or debt—their risk of experiencing depression escalates dramatically. These findings emphasize the need for institutions to address student financial burdens through mechanisms such as scholarships, emergency funds, part-time job opportunities, or structured financial literacy programs. Moreover, integrating financial counseling with mental health services could provide a dual benefit, reducing both economic and emotional distress in high-risk groups.

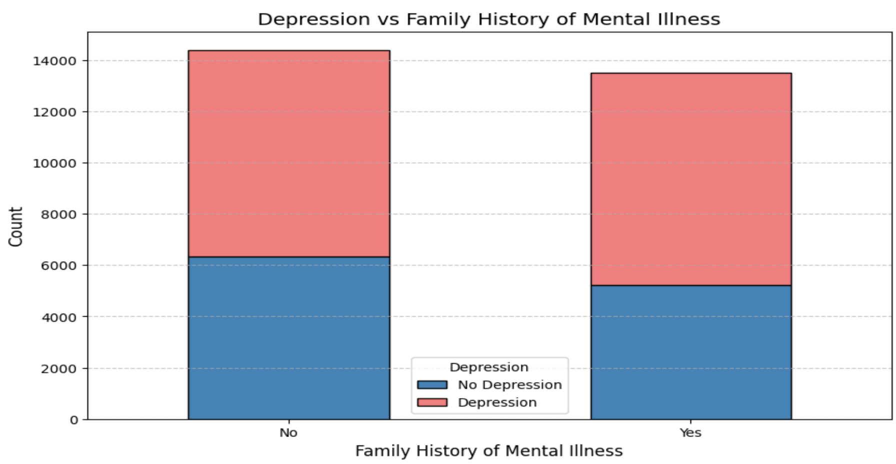

Figure 5

Figure 5 visualizes the association between a family history of mental illness and the likelihood of depression among college students. The chart compares the number of students categorized as "Depressed" and "Not Depressed" across two groups: those with and those without a reported family history of psychological disorders. For students who reported no family history, the distribution is relatively balanced, with approximately 52% showing symptoms of depression and 48% not showing depressive symptoms. This suggests that while depression can certainly occur without a genetic link, the likelihood is more evenly distributed in this group. In contrast, students who reported a positive family history of mental illness showed a significantly different distribution. Around 68–70% of these individuals were identified as depressed, while only 30–32% were not. This shift indicates a strong correlation between familial mental health background and the individual's own mental health status. The increase in depression prevalence among students with a family history may be explained by several interacting factors. Genetically, certain vulnerabilities to mood disorders may be inherited. Environmentally, growing up in households where mental illness is present can increase emotional stress and affect coping mechanisms. Additionally, there may be less access to supportive communication or mental health resources within such families. These results support existing literature suggesting that familial influence plays an important role in mental health outcomes. It highlights the importance of screening students with known family histories of psychological conditions and providing them with access to counseling and emotional support early in their academic journey.

4.3 Model Comparison

Figure 6 and Table 2 present a comparative performance analysis of six machine learning classification algorithms used to predict depression among students. The evaluation was based on four key metrics: Accuracy, Precision, Recall, and F1-score, which provide a balanced view of model effectiveness. Among the models tested, both Logistic Regression and Support Vector Machine (SVM) delivered the best overall performance, each achieving an accuracy of 0.84, precision of 0.85, recall of 0.88, and an F1-score of 0.86. These high recall values are particularly important for mental health screening tasks, where minimizing false negatives (i.e., undetected depression cases) is critical.

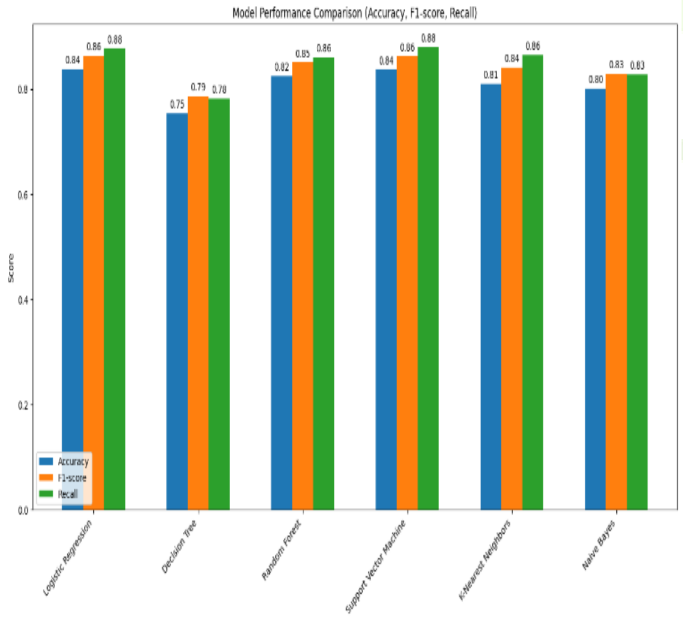

Figure 6

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.84	0.85	0.88	0.86
Decision Tree	0.76	0.80	0.79	0.79
Random Forest	0.82	0.84	0.86	0.85
Support Vector Machine	0.84	0.85	0.88	0.86
K-Nearest Neighbors	0.81	0.82	0.86	0.84
Naive Bayes	0.80	0.83	0.83	0.83

Table 2

Random Forest also performed strongly, with an accuracy of 0.82 and an F1-score of 0.85, suggesting that ensemble methods can effectively capture complex relationships in the data. The K-Nearest Neighbors (KNN) model showed decent performance, recording 0.81 accuracy and 0.84 F1-score, indicating that instance-based learning can also be effective, especially in small to medium-sized datasets. Naïve Bayes, known for its simplicity and efficiency, achieved 0.80 accuracy and 0.83 F1-score, performing reasonably well despite its assumption of feature independence. In contrast, the Decision Tree model had the lowest overall performance, with an accuracy of 0.76 and an F1-score of 0.79. Although it remains interpretable and easy to visualize, its lower precision and recall suggest that it may be prone to over fitting or misclassification in this dataset. Overall, the results indicate that Logistic Regression and SVM are the most reliable algorithms for depression prediction in this context. Their strong balance of precision and recall makes them suitable for real-world screening applications where both accuracy and early detection are crucial.

5. CONCLUSION AND FUTURE SCOPE

This study investigated the major risk factors associated with depression among college students using multiple machine learning techniques. The findings confirm that suicidal ideation, academic pressure, and financial stress are the most influential predictors of depression. Additional contributors included sleep quality, study satisfaction, and dietary habits, all of which align with previous research and were validated through correlation analysis and model interpretation. Among all models tested, Logistic Regression and Support Vector Machine (SVM) performed the best, achieving 84% accuracy and 88% recall, indicating a strong capability to identify students at risk of depression. These results not only validate the effectiveness of ML in mental health prediction but also highlight the value of analyzing multiple interdependent factors rather than isolated variables. These insights can be useful for educational institutions, counselors, and policymakers to create targeted mental health awareness programs, improve academic counseling, and offer financial aid to reduce psychological stress among students.

Future research can focus on developing a real-time depression screening system that integrates behavioral data such as sleep patterns, screen time, and social media activity. Incorporating longitudinal datasets would enhance the robustness of predictions. Additionally, AI-powered mental health Chabot and early-warning systems could be implemented within college platforms to offer students timely support and intervention. Cross-cultural comparisons between universities in different regions could also provide more comprehensive insights and help design globally adaptable mental health strategies.

6. ACKNOWLEDGEMENTS

The author expresses sincere gratitude to their parents, teachers, and friends for their unwavering support, encouragement, and guidance throughout the course of this research. Their constant presence has been a source of strength and motivation at every stage of this academic journey.

AUTHORS PROFILE

Passang Tshering Sherpa is an emerging technology enthusiast currently pursuing a B.Tech in Computer Science and Engineering from Gulzar Group of Institutions (GGI), Punjab, India. He previously completed his Diploma in Computer Science and Technology from the Centre for Computers and Communication Technology (CCCT), Sikkim, India. where he was awarded the Gold Medal for academic excellence (Batch 2020–2023). He is deeply passionate about Machine Learning (ML), Artificial Intelligence (AI), and Data Analysis, and is committed to developing solutions that address real-world challenges in education, mental health, and rural development. He has earned NPTEL certifications in Java and Data Mining, strengthening his core programming and analytical skills.

REFERENCES

[1]. A. K. Ibrahim, H. Kelly, T. C. Adams, and R. Glazebrook, "A systematic review of studies of depression prevalence in university students," J. Psychiatr. Res., vol. 47, no. 3, pp. 391–400, 2013, doi: 10.1016/j.jpsychires.2012.11.015.

[2]. T. Arango et al., “Suicide among youth: a global imperative,” J. Am. Acad. Child Adolesc. Psychiatry, Vol. 60(4), pp. 413–415, 2021

[3]. R. Beiter et al., "The prevalence and correlates of depression, anxiety, and stress in a sample of college students," J. Affect. Disord., vol. 173, pp. 90–96, 2015, doi: 10.1016/j.jad.2014.10.054.

[4]. A. Gulliver, K. M. Griffiths, H. Christensen, and R. P. Brewer, "Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review," BMC Psychiatry, vol. 10, p. 113, 2010, doi: 10.1186/1471-244X-10-113.

[5]. B. Keles, N. McCrae, and A. Grealish, "A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents," Int. J. Adolesc. Youth, vol. 25, no. 1, pp. 79–93, 2020, doi: 10.1080/02673843.2019.1590851.
[6]. D. Eisenberg, M. Golberstein, and E. G. Hunt, "Prevalence and correlates of depression, anxiety, and suicidality among university students," Am. J. Orthopsychiatry, vol. 77, no. 4, pp. 534–542, 2007, doi: 10.1037/0002-9432.77.4.534.
[7]. B. Andrews and J. M. Wilding, "The relation of depression and anxiety to life-stress and achievement in students," Br. J. Psychol., vol. 95, no. 4, pp. 509–521, 2004, doi: 10.1348/0007126042369802.
[8]. N. Bayram and N. Bilgel, "The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students," Soc. Psychiatry Psychiatr. Epidemiol., vol. 43, no. 8, pp. 667–672, 2008, doi: 10.1007/s00127-008-0345-x.
[9]. K. Zivin, D. Eisenberg, S. Gollust, and E. Golberstein, "Persistence of mental health problems and needs in a college student population," J. Affect. Disord., vol. 117, no. 3, pp. 180–185, 2009, doi: 10.1016/j.jad.2009.01.001.
[10]. J. Hunt and D. Eisenberg, "Mental health problems and help-seeking behavior among college students," J. Adolesc. Health, vol. 46, no. 1, pp. 3–10, 2010, doi: 10.1016/j.jadohealth.2009.08.008.
[11]. 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," Clin. Psychol. Sci., vol. 6, no. 1, pp. 3–17, 2018, doi: 10.1177/2167702617723376.
[12]. D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
[13]. J. R. Quinlan, C4.5: Programs for Machine Learning. San Francisco, CA, USA: Morgan Kaufmann, 1993.
[14]. L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.
[15]. C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.
[16]. T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 21–27, 1967.
[17]. F. Torres Cruz et al., "Prediction of depression status in college students using a Naïve Bayes classifier," J. Educ. Brain Sci., Vol. 12(1), pp. 45–52, 2023.
[18]. S. Deb, E. Strodl, and J. Sun, "Academic-related stress among private secondary school students in India," Asian Educ. Dev. Stud., vol. 4, no. 2, pp. 123–134, 2015.
[19]. S. McManus and D. Gunnell, "Trends in mental health, non-suicidal self-harm and suicide attempts in 16–24-year-old students and non-students in England, 2000–2014," Soc. Psychiatry Psychiatr. Epidemiol., vol. 55, no. 1, pp. 125–128, 2020.
[20]. T. Richardson, P. Elliott, and R. Roberts, "A longitudinal study of financial difficulties and mental health in a national sample of British undergraduate students," Community Ment. Health J., vol. 53, pp. 344–352, 2017.
[21]. F. N. Jacka et al., "Dietary patterns and depressive symptoms over time: examining the relationships with socioeconomic position, health behaviours and cardiovascular risk," PLoS One, vol. 9, no. 1, p. e87657, 2014.
[22]. J. M. Twenge, G. N. Martin, and W. K. Campbell, "Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017," J. Abnorm. Psychol., vol. 128, no. 3, pp. 185–199, 2019.
[23]. D. F. Levinson, "The genetics of depression: A review," Biol. Psychiatry, vol. 60, no. 2, pp. 84–92, 2006.
[24]. P. Olgiati, E. Serretti, and A. M. Mandelli, "Clinical correlates and prognostic implications of severe suicidal ideation in major depressive disorder," Int. Clin. Psychopharmacol., vol. 38, no. 4, pp. 201–208, 2023, doi: 10.1097/YIC.0000000000000461.
[25]. S.K. Sharma, L. Gupta, "A Machine Learning Based Study on College Students' Depression Using Real-World Data," International Journal of Computer Sciences and Engineering, Vol. 13, Issue 3, pp. 12–19, 2025.
[26]. Azem, M. A., & Omer, A. I. (2023). The influence of social media usage on adolescent mental health: A scoping review. Behavioral Sciences, 13(6), 475.
[27]. A. Ponce et al., "Behavioral modeling of loneliness and depression from passive sensing," IEEE J. Biomed. Health Inf., Vol. 28, pp. 123–131, 2024.
[28]. R. Rezapour & S. Elmshaeuser, "Factors influencing student mental well-being during remote learning: Evidence from online survey," Comput. Educ., Vol. 178, pp. 104467, 2022.

[29]. P. Muzumdar, K. Patil, R. Joshi, "Predictive modeling of student mental health using machine learning," Educ. Inf. Technol., Vol. 27(3), pp. 3987–4005, 2022.

[30]. Krishnan, S., Devi, P. R., & Lavanya, T. (2024). A predictive model for anxiety and depression among students in higher education using HADS. Arab Gulf Journal of Scientific Research, 42(1), 43–55.

[31]. Sinha, A. M., "Detecting Depression and Suicidal Ideation from Texts using Machine Learning & Deep Learning Techniques," Int. J. Comput. Sci. Eng., Vol. 12, Issue 3, pp. 482–490, 2023.

[32]. Sudhakar, S., "Psychological Stress Detection from Social Media Data using a Machine Learning Approach," Int. J. Comput. Sci. Eng., Vol. 10, Issue 4, pp. 409–416, 2018