A High-Precision SVM Model for Depression Detection Among Farmers

S. Vidya
School of Computer Science and Applications, REVA University, Bangalore, India | Bangalore Institute of Technology, Bangalore, India
vidyachandru.cs@gmail.com (corresponding author)

G. T. Raju
Department of Computer Science and Engineering, SJC Institute of Technology, VTU, India
gtraju1990@yahoo.com

M. Vinayaka Murthy
School of Computer Science and Applications, REVA University, Bangalore, India
mvinayaka.murthy@reva.edu.in

Received: 29 June 2025 | Revised: 13 July 2025, 31 July 2025, and 4 August 2025 | Accepted: 14 August 2025

 DOI: https://doi.org/10.48084/etasr.13046

ABSTRACT

Depression among farmers is an increasing concern, driven by a complex interplay of agricultural, socioeconomic, and behavioral stressors. This study employs machine learning to identify depression in farmers by integrating diverse factors such as economic conditions, social behavior, and farming-related challenges. This study offers a novel data-driven framework for the detection of depression among Indian farmers by merging socioeconomic and psychological indicators—an area with limited prior investigation. The integration of Patient Health Questionnaire 9 (PHQ-9) assessments with an SVM-RFE model establishes a robust and accurate method for large-scale mental health screening in agrarian populations. Data were gathered from 1,069 farmers in Mandya, India, using structured surveys and the validated PHQ-9. The SVM model demonstrated high performance, achieving 96.59% accuracy, 96.47% precision, 96.59% recall, and a 96.00% F1-score, surpassing several other advanced classification algorithms. The findings underscore the significant impact of economic instability, social isolation, and limited access to mental healthcare on farmers' psychological well-being. Notably, gender-based disparities emerged, with approximately 40% of female farmers found to be more susceptible to depression. The study highlights the urgent need for integrated mental health support systems and agricultural policy reforms, advocating for scalable, AI-powered early detection tools tailored to rural farming communities.

Keywords-farmers' mental health; depression detection; SVM; agriculture

I. INTRODUCTION

Agriculture remains vital to the Indian economy, engaging 42.3% of the population and contributing 18.2% to GDP. Despite hurdles, the sector averaged 4.18% annual growth over five years [1]. Farmer suicides have slightly decreased to 5,579 in 2020, down from 5,957 in 2019, with Karnataka accounting for 1,072 of these cases [2]. Farmers primarily seek guidance on seed selection, pricing, fertilizers, marketing, pests, and crop diseases. They also favor traditional sources of information, likely due to limited access to modern technologies, which many found difficult to use in farming [3]. Approximately 42 farmer suicides were reported in Mandya Taluk between 2016 and 2019 [4]. In the farming community, mental illness appears to be particularly stigmatizing, and farmers seem to hesitate to contact the healthcare system for help [5,6]. Manic depression, also known as bipolar disorder or Major Depressive Disorder (MDD), is a common type of depression often observed among farmers [7]. Figure 1 shows a rise in Karnataka’s agricultural GSDP, growing from ₹533.95 billion in 2012 to ₹2,011.62 billion by 2023. This growth underscores agriculture's continued importance in the state's economy, supporting rural livelihoods, food security, and job creation. Advancements in farming methods, increased government backing, and technology adoption have all contributed to this upward trend.

Machine learning has become valuable in the detection of depression by identifying complex emotional and behavioral patterns often overlooked by traditional methods. By analyzing vast datasets, such as speech, text, facial expressions, and physiological responses, these models enhance diagnostic accuracy and reduce subjectivity. They also enable early detection and tailored treatment plans, significantly advancing the quality and effectiveness of mental healthcare.

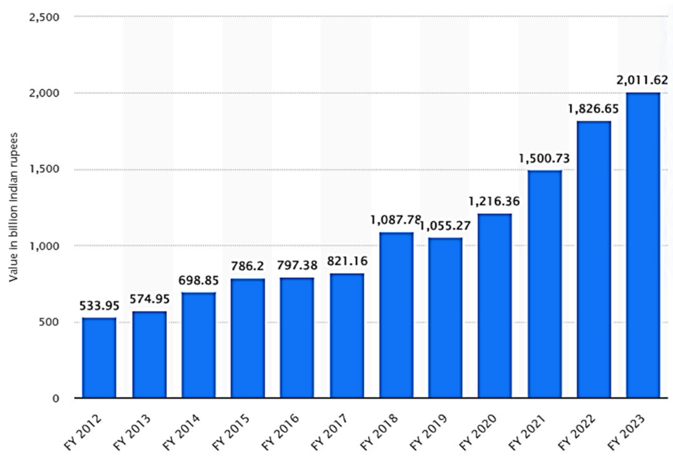

Fig. 1. Gross state domestic product for agriculture for Karnataka, India, from the financial year 2012 to 2023. Courtesy: https://www.statista.com

This study aimed to utilize machine learning techniques for the early detection of anxiety and depression among farmers in Mandya. Early identification of these mental health conditions has the potential to prevent severe outcomes, including loss of life, while also contributing to the broader well-being of the agricultural community. Indirectly, such advances could promote greater interest in agriculture among younger generations, generate employment opportunities, and positively affect the country's GDP by strengthening the rural economy.

A. *Key Contributions of the Proposed System*

This study presents a novel framework for identifying key determinants of depression among farmers, a population that remains underrepresented in mental health studies. Data for this study were collected using structured surveys as the principal instrument for collecting responses. The analysis incorporated the PHQ-9 questionnaire in conjunction with a supplementary tool that captures socio-demographic, economic, personal, behavioral, agricultural information, and technological variables. This comprehensive approach ensures a multifaceted understanding of the contributing factors to depression.

The study in [8] combined EEG signals with PHQ assessments to detect depression among farmers, but it lacked clinical validation and was not widely accepted in psychiatric practice. This study also suffered from drawbacks, such as small sample sizes, risk of misdiagnosis, and complex technical requirements. This study addresses these issues using non-invasive, scalable methods based on socio-demographic and PHQ-9 data. By leveraging a robust dataset and applying the Support Vector Machine (SVM) algorithm, this study proposes an accurate and interpretable model to identify the root causes of depression. The insights derived from this work aim to support evidence-based decision-making and enable targeted mental health interventions.

II. RELATED WORK

The study in [9] explored depression detection through social media, focusing on Twitter data. This study introduced a word-embedding optimization approach to classify user text, comparing different deep learning models. This study faced challenges such as short, informal posts, misspellings, and unstructured language due to the format of tweets. In [10], machine learning, deep learning, and NLP techniques were applied to detect depression among Bangladeshi university students, achieving high accuracy and recall through models such as Random Forest (RF) and RoBERTa. Resampling methods and explainable AI tools, such as SHAP and LIME, improve prediction quality and transparency, supporting broader applicability across populations.

A questionnaire-based study in Thrace, Greece, explored the links between depression and factors such as demographics, lifestyle, and health using four machine learning models—LR, SVM, XGBoost, and NNs [11]. A genetic algorithm aided in feature selection, while SHAP explained each feature's impact on depression prediction. In [12], NHANES 2013-2014 data were used with various machine learning models to predict depression based on PHQ-9 scores, with XGBoost delivering the best results across all metrics. SHAP analysis revealed key risk factors, including income-to-poverty ratio, gender, BMI, hypertension, and renal function. In [13], a video-based system was developed to detect depression in college students through facial feature analysis. Using Gabor filters for feature extraction and SVM for classification, the system was trained on both positive and negative emotional expressions to identify depressive symptoms. In [14], current machine learning approaches for early depression detection were reviewed, using data sources such as audio, video, images, PHQ responses, social networks, and EEG signals. In [15], machine learning techniques were investigated to detect depression in college students using PHQ-9 data, finding RF to be the most effective with 98% accuracy. XGBoost and SVM also performed strongly, highlighting the value of such models in enhancing campus mental health support.

In the context of agricultural risk, in [16], price fluctuations and production uncertainties were identified as primary challenges, with diversification emerging as an effective risk management approach. In [17], an interdisciplinary effort combined machine learning techniques with social media data and traditional mental health assessments to detect depression and suicidal ideation. Algorithms such as SVM, RF, and neural networks enhanced prediction capabilities, although concerns about language diversity and privacy remain. In [18], behavioral indicators of depression were investigated using smartphone data from 629 users over 22 days. This analysis focused on 22 behavioral features, including routine regularity, entropy, and standard deviation-linked with PHQ-8 scores, applying correlation and bivariate linear mixed models to understand their relationship with depressive symptoms.

In [19], a hybrid machine learning model (EKLS) was proposed to detect emotions through facial expressions using data from the COVID era, achieving 99.82% accuracy. However, this method may face skepticism from psychiatrists due to its reliance on non-clinical, visual cues. In [20], sentiment analysis was applied to IMDB and Spotify reviews using models such as GRU, LSTM, Linear SVC, and AdaBoost, combined with TF-IDF and hyperparameter tuning. GRU with TF-IDF achieved top accuracy at 97%, outperforming others, while TF-IDF proved more effective than ANOVA for text classification.

III. PROPOSED METHODOLOGY

Primary data was collected from rural, urban, and semi-urban areas across the seven taluks of the Mandya region: Mandya, Maddur, Malavalli, Pandavapura, Srirangapatna, Nagamangala, and Krishnarajpet. The sampling was stratified based on land types and the predominant crops cultivated in each area. Figure 2 illustrates the key factors that influence farmers' lives, categorized into economic, personal, behavioral, and technological domains. Economic stressors, such as income volatility, mounting debt, and fluctuating market prices, contribute significantly to financial instability. Behavioral patterns, along with the use of technology, play a vital role in determining both productivity and mental well-being. Figure 3 emphasizes how personal factors, such as health status, educational background, and social relationships, impact farmers' decision-making and productivity. Poor access to healthcare and low educational levels increase economic and psychological vulnerabilities, while strong family support can foster resilience. In contrast, social isolation is associated with increased mental health risks. Personal data were organized into four categories for structured analysis. Technological influences, including the application of precision farming tools, are acknowledged as productivity enhancers. However, their adoption remains limited due to high costs and limited accessibility in certain regions.

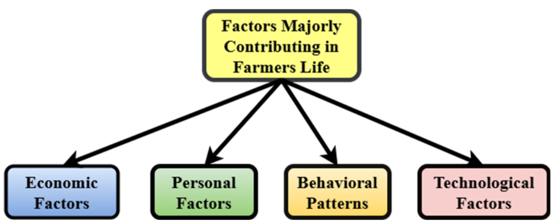

Fig. 2. Factors majorly influencing a farmer's life.

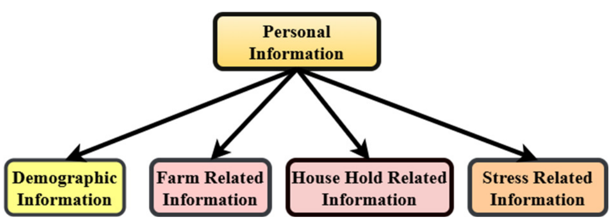

Fig. 3. Personal information of a farmer, divided into four categories while collecting the dataset.

A. Population

The data collection process incorporated variables such as residence type, gender, age group, and population category. From 1,453 initial responses, 1,069 were finalized after data cleaning using stratified sampling for demographic balance. The final sample included 641 male and 428 female participants, with age groups evenly distributed between 25–55 years, and 107 respondents above 55. Geographically, respondents were spread across rural (535), urban (214), and semi-urban (321) areas. Depression levels were evaluated using PHQ-9, a reliable nine-item tool aligned with DSM-IV criteria, utilizing a three-point Likert scale to measure symptom frequency over a two-week period [21] as detailed in Table I.

TABLE I. OPTIONS AND ENCODED VALUES

Options	Value
NA	0
SD	1
MD	2
ND	3

*NA=Not at all
SD= Several Days
MD=More than half the days
ND= Nearly everyday

B. Dataset

The questionnaire comprised 52 questions, each representing a distinct variable collected from participants. These features spanned demographic details (e.g., age, gender, marital status, education, household size), agricultural practices (e.g., farm type and size, crop variety, market access, farming experience), and economic conditions (e.g., subsidies, loans, crop loss). Social and environmental factors, such as satisfaction with family, job, and living conditions, were also included. Psychological health was assessed using the PHQ-9, capturing key depression indicators such as anhedonia, low mood, fatigue, sleep/appetite disturbances, self-worth, focus, psychomotor activity, and suicidal thoughts, alongside the perceived impact on daily functioning. Together, these variables supported a holistic analysis of mental health and stressors affecting farming communities.

C. Data Preprocessing

The dataset was preprocessed using Python's scikit-learn library. The standard scaler was applied to normalize the numerical features, ensuring consistent scaling across variables, while the OneHotEncoder was used to convert categorical variables into binary feature representations. Following preprocessing, the dataset was divided into training and testing subsets using an 80:20 data split ratio using a stratified sampling strategy to preserve the class distribution. These splits were kept fixed for all models. This allowed for effective model training on the larger portion of the data and independent validation on the remaining subset to evaluate model performance and generalizability.

```
Algorithm 1: Preprocessing
    Input: Dataset
    for each column x in data:
      if x is a categorical column:
        Encode x using Label Encoding:
      end if
    end for
    Output: Preprocessed dataset data
```

```
Algorithm 2: Farmer Depression Detection
             Model
Input:
  F: Set of farmer-specific features
  A: Set of agricultural factors
  S: Set of socioeconomic variables
  E: Set of environmental stressors
Output:
  D: Depression risk score (0-100)
```

```
1. Initialize D = 0
2. For each feature f in F:
2.1. Calculate weight wf based on feature
     importance
2.2. D=D+(wf*f)
3. For each factor a in A:
3.1. Calculate weight wa based on factor
     importance
3.2. D=D+(wa*a)
4. For each variable s in S:
4.1. Calculate weight ws based on variable
     importance
4.2. D=D+(ws*s)
5. For each stressor e in E:
5.1. Calculate weight we based on
     stressor importance
5.2. D=D+(we*e)
6. Normalize D to scale 0-100:
   D=(D-Dmin)/(Dmax-Dmin)*100
7. Apply machine learning classifier C to
   D:
   Risk_category = C(D)
8. Return D and Risk_category
```

D. Feature Extraction

Data preprocessing was carried out to prepare the dataset, derived from an extended version of PHQ-9, for machine learning model development. Table II outlines the classification of depression severity among farmers, mapping specific score ranges to categories such as minimal, mild, moderate, and severe.

TABLE II. LABELS BASED ON SCORES FOR THE FARMERS DATASET

Total score	Severity of depression	Stage/label
0-4	Minimal or none	0
5-9	Mild	1
10-14	Moderate	2
15-19	Moderately severe	3
20-27	Severe	4

Figure 4 presents the machine learning pipeline used for depression detection. The pipeline begins with data preprocessing, which involves handling missing values and encoding categorical variables. Relevant features are then selected using Recursive Feature Elimination (RFE) to improve model accuracy and reduce complexity. A comprehensive hyperparameter tuning and validation strategy was adopted to ensure the robustness of the SVM-RFE model and to reduce the risk of overfitting. A grid search procedure was used to optimize key SVM parameters, including the regularization parameter and kernel-specific parameters, depending on the selected kernel (e.g., RBF or linear). The tuning process was integrated within a nested stratified k-fold cross-validation framework to maintain class balance and prevent biased performance estimation. Feature selection via Recursive Feature Elimination (RFE) was conducted strictly within each training fold, ensuring that the test data remained completely unseen during both feature ranking and model training stages. This prevents data leakage and protects against overestimating model performance, a common risk in clinical and psychological datasets where sample sizes can be limited and class imbalance is frequent. By employing this nested cross-validation structure alongside systematic hyperparameter optimization, the evaluation of the model's accuracy becomes more reliable and generalizable. This methodological rigor is especially important in mental health research, where predictive modeling must be interpretable and clinically applicable.

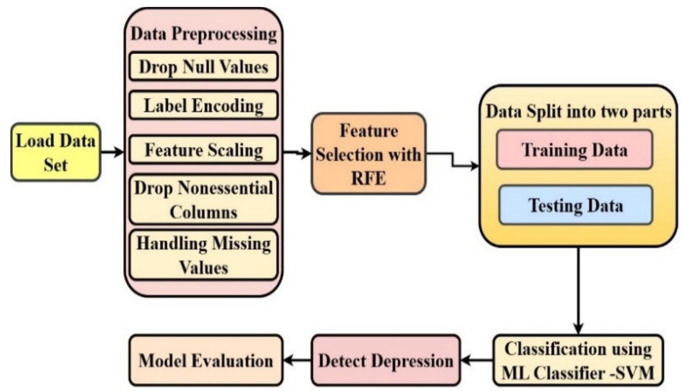

Fig. 4. Methodology.

E. Support Vector Machine (SVM)

SVMs are highly effective in depression detection, particularly due to their ability to handle high-dimensional data and identify optimal hyperplanes that distinguish between individuals with and without depressive symptoms. SVMs are especially well-suited for psychological research, where datasets are often limited in size. Their use of kernel functions allows for the modeling of complex, non-linear relationships, ideal for interpreting nuanced questionnaire responses and behavioral indicators. SVMs also exhibit strong robustness to noise and are capable of efficiently managing multi-dimensional inputs, such as aggregated questionnaire scores and various psychological variables. Moreover, their ability to generalize well to unseen data makes them valuable tools for both research and clinical diagnostic applications. These characteristics collectively position SVMs as a powerful method for accurately identifying depression across diverse populations. The decision function for an SVM is expressed as:

$$f(x) = sign(\sum_{i=1}^{n} a_i\, y_i K(x_i, x) + b) \qquad (1)$$

where x is the input feature vector, x_i is the support vector (training examples), y_i are labels of the training examples, $K(x_i, x)$ is the kernel function (linear, polynomial, RBF, etc.), a_i are Lagrange multipliers determined during training, and b is the bias term. The $sign$ function determines the class label.

```
Algorithm 2. Steps to identify the key
reason for Depression
1. Input: PHQ and standard questionnaire
   responses.
2. Calculate PHQ Score
   Sum responses to PHQ questions.
```

```
3. Depression Assessment
   If PHQ score < threshold,
     output "Not Depressed" and stop.
4. Analyse Key Reasons
   Evaluate economic,
   personal/agricultural factors,
   technological factors, and
   psychological factors for depression
   triggers.
5. Output Results: Depression status and
   identified key reasons.
```

IV. RESULTS AND DISCUSSIONS

A. Performance Evaluation Metrics

The SVM model showed strong performance. As shown in Figure 5, the confusion matrix reports the correct classification of 7 Extremely Depressed (ED), 75 Normal (ND), 83 Slightly Depressed (SD), and 48 Very Depressed (VD) instances. Although these results are promising, further validation using larger and more diverse datasets is recommended to enhance generalizability.

Figure 6 presents the precision-recall curves, which evaluate the model's effectiveness in distinguishing between depression levels. Curves approaching the top-right corner indicate higher classification performance, demonstrating the model's reliability in identifying depressive symptoms.

B. Results

Out of the 1,069 farmers, 625 (58.4%) were identified as experiencing depression, while 444 (41.6%) were classified as non-depressed. These findings indicate a notably high prevalence of depression within the surveyed farming population.

TABLE III. RESULTS USING MACHINE LEARNING ALGORITHMS

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
SVM	96.59	96.47	96.59	96.00
CNN	85.3	71	70	70
LSTM	85.3	87	89	87
RF	95.64	96.00	96.00	96.00
CatBoost	96	95.5	95.5	95.5

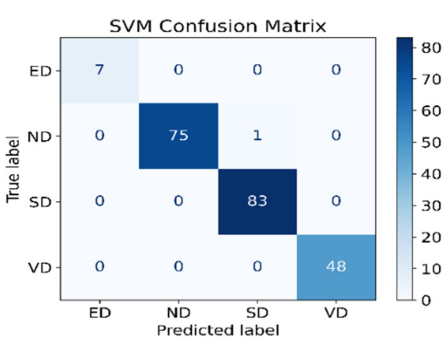

Fig. 5. Confusion matrix for the performance of SVM.

(a)

(b)

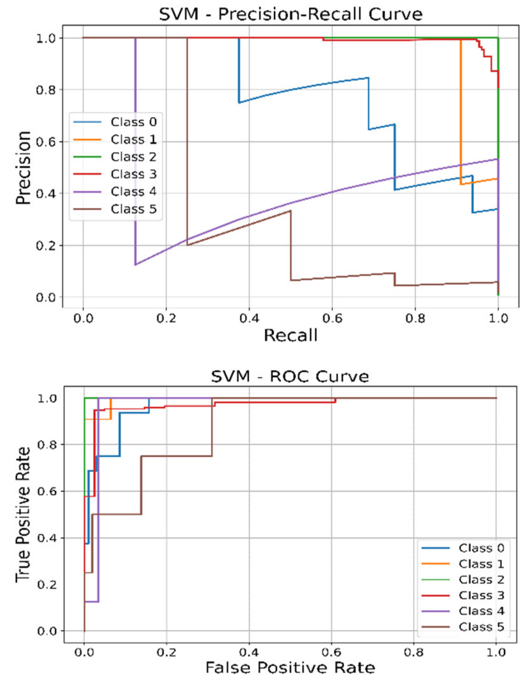

Fig. 6. SVM model evaluation: (a) Precision-Recall trade-off for SVM multi-class classification, (b) ROC curves for SVM multi-class classification.

C. Gender Disparity

The results revealed a significant gender disparity in mental health. Approximately 40% of female farmers were found to be experiencing depression, substantially higher than their male counterparts. Contributing factors include increased household responsibilities, work-related stress, crop failures, and social isolation, all of which disproportionately impact women in agricultural settings.

The research also highlights an increased vulnerability among farmers aged 35 to 45 years, a demographic particularly susceptible to depression. Individuals in this age group often face a combination of challenges, including managing family and work obligations, coping with financial stress, and enduring the physically demanding nature of farming, all contributing to their elevated risk of mental health issues.

D. Impact of Economic Factors

Significant challenges faced by 80% of farmers involve high input costs, fluctuating commodity prices, and labor shortages, factors that collectively threaten the economic viability of their agricultural operations. This persistent financial stress is a major contributor to the onset of depression among farmers.

The SVM model achieved high performance. As shown in Table III, the model achieved an accuracy of 96.59%, a precision of 96.47%, a recall of 96.59%, and an F1-score of 96.00%. Here, accuracy denotes the proportion of correct predictions out of the total predictions made, indicating the model's reliability in identifying depression in the target population.

E. Comparative Study and Findings

As presented in Table IV, the SVM model achieved a prediction accuracy of 96.59%, outperforming both the RF and CatBoost models, which recorded accuracies ranging from 77.45% to 92.36%. This superior performance can be attributed to comprehensive data collection, effective feature engineering, and the inherent robustness of the SVM algorithm. By incorporating a diverse range of socioeconomic and agricultural variables, the model demonstrates high precision in detecting depression within the farming community.

TABLE IV. COMPARATIVE PERFORMANCE ASSESSMENT BETWEEN THE PROPOSED AND EXISTING APPROACHES

Ref.	Number of participants	Data collected	Model	Accuracy (%)
[22]	355	Socio-demographic	RF	89.3%
[23]	520	PHQ-9, Job-Seeking stress-related Questions, Socio-demographic	Stacking classifier	77.45%
[24]	348	Occupational information, Socio-demographic	RF	85.5%
[25]	2121	Socio-demographic	RF	89%
[26]	111	Twitter posts, comments	SVM	82.5%
[27]	6588	Socio-demographic, economic, and clinical information	RF	86.20%
[28]	470	Socio-demographic, occupational information	CatBoost	89.30%
[29]	604	Socio-demographic and psychosocial information	AdaBoost	92.56%
[30]	tweets_ combined_12 .csv	Twitter social media posts	Gradient Boosting	90%
Proposed model	**1069**	**Socio-demographic, personal information, farm-related, technological factors, behavioral factors, psychosocial information, and PHQ-9**	**SVM**	**96.59%**

V. LIMITATIONS AND FUTURE WORK

Although this study is based solely on data from the Mandya district in Karnataka, which may introduce regional specificity, it provides a focused foundation for understanding local agrarian mental health patterns. Recognizing that cultural and agricultural differences between regions may affect the findings. Minor irregularities in the ROC curves are likely due to limited data and class imbalance, with such effects common in clinical datasets. Future research will extend validation efforts using a large dataset from diverse farming communities to enhance the model's generalizability. Addressing these challenges with larger and more balanced datasets in future work will likely lead to smoother classifier performance and more robust insights.

VI. CONCLUSION

This study highlights the prevalence of depression among farmers in the Mandya region, emphasizing the complex interplay of economic hardship, social isolation, and limited access to mental health services. Although government subsidies and loan schemes provide some financial relief, they are often insufficient to address the deeper systemic challenges faced by the farming community. To alleviate these issues, the promotion of modern agricultural machinery can help reduce labor dependency and improve operational efficiency. Increasing awareness and adoption of agricultural insurance schemes can offer critical financial protection against crop failure and market volatility. In addition, integrating mental health support services, stress management programs, and community-based support networks is essential to improve the psychological well-being of farmers. These interventions not only support individual health but also contribute to increased agricultural productivity and, in turn, foster broader national economic growth.

ACKNOWLEDGMENT

The authors acknowledge the help of Dr. B. R. Harish, Professor and Head of Community Medicine, Mandya Medical College, Dr. J. Bindiya, LMO, MD Psychiatry, Mandya Medical College, and Dr. Vinod, MD Psychiatry, Sankalpa Clinic, Mandya. The authors also acknowledge REVA University, SJC Institute of Technology, and Bangalore Institute of Technology for the extended support in conducting this research.

REFERENCES

[1] "Provisional Estimates of Annual GDP for 2023-24 And Quarterly Estimates of GDP for Q4 of 2023-24." https://www.pib.gov.in/www.pib.gov.in/Pressreleaseshare.aspx?PRID=2022323.

[2] "India: agriculture gross domestic product of Karnataka 2024| Statista," *Statista*. https://www.statista.com/statistics/974112/india-agriculture-gsdp-of-karnataka/.

[3] K. S. Parvathamma and B. N. Hemavathi, "Agricultural Information Needs of Small Holder Farmers of Karnataka: A Pilot Study at Ramanagara District," *International Journal of Information Studies*, vol. 17, no. 2, pp. 51–63, Apr. 2025, https://doi.org/10.6025/ijis/2025/17/2/51-63.

[4] I. P. Irvan and N. N. Yuliarmi, "Analysis of impact factors on farmers income," *International Research Journal of Management, IT and Social Sciences*, vol. 6, no. 5, pp. 218–225, Sep. 2019, https://doi.org/10.21744/irjmis.v6n5.731.

[5] C. Coughenour and L. Swanson, "Determinants of Farmers' Satisfactions With Farming and With Life: A Replication and Extension," *Journal of Rural Social Sciences*, vol. 09, no. 1, Dec. 1992.

[6] A. Ramesh and C. Madhavi, "Occupational stress among farming people," *Journal of Agricultural Sciences – Sri Lanka*, vol. 4, no. 3, pp. 115–125, Mar. 2010, https://doi.org/10.4038/jas.v4i3.1650.

[7] S. Daghagh Yazd, S. A. Wheeler, and A. Zuo, "Key Risk Factors Affecting Farmers' Mental Health: A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 16, no. 23, Dec. 2019, Art. no. 4849, https://doi.org/10.3390/ijerph16234849.

[8] H. M. Mallikarjun, A. Chhetri, G. S. Apoorva, G. Jadhav, and B. V. Sheetal, "Depression Predictor Model for Farmers using Machine Learning Techniques," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 2S, pp. 540–543, Dec. 2019, https://doi.org/10.35940/ijitee.B1047.1292S19.

[9] A. H. Orabi, P. Buddhitha, M. H. Orabi, and D. Inkpen, "Deep Learning for Depression Detection of Twitter Users," in *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, New Orleans, LA, USA, Mar. 2018, pp. 88–97, https://doi.org/10.18653/v1/W18-0609.

[10] G. H. A. Masud, R. I. Shanto, I. Sakin, and M. R. Kabir, "Effective depression detection and interpretation: Integrating machine learning, deep learning, language models, and explainable AI," *Array*, vol. 25, Mar. 2025, Art. no. 100375, https://doi.org/10.1016/j.array.2025.100375.

[11] C. Mimikou *et al.*, "Explainable Machine Learning in the Prediction of Depression," *Diagnostics*, vol. 15, no. 11, Jun. 2025, Art. no. 1412, https://doi.org/10.3390/diagnostics15111412.

[12] T. Vu *et al.*, "Prediction of depressive disorder using machine learning approaches: findings from the NHANES," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, Feb. 2025, Art. no. 83, https://doi.org/10.1186/s12911-025-02903-1.

[13] N. S. Parameswaran and D. Venkataraman, "A computer vision based image processing system for depression detection among students for counseling," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 1, pp. 503–512, 2019, https://doi.org/10.11591/ijeecs.v14.i1.

[14] S. Vidya, G. T. Raju, and M. Vinayaka Murthy, "Review on Depression Detection Using Machine Learning Techniques," in *Proceedings of 4th International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications*, vol. 873, V. K. Gunjan and J. M. Zurada, Eds. Springer Nature Singapore, 2024, pp. 715–728.

[15] S. Vidya, N. Swamy, and H. Lalitha, "Leveraging Machine Learning Algorithms for Depression Prediction in College Students Using PHQ-9 Scores," in *2024 IEEE 6th PhD Colloquium on Emerging Domain Innovation and Technology for Society (PhD EDITS)*, Bangalore, India, Nov. 2024, pp. 1–2, https://doi.org/10.1109/PhDEDITS64207.2024.10838637.

[16] N. Jankelova, D. Masar, and S. Moricova, "Risk factors in the agriculture sector," *Agricultural Economics (Zemědělská ekonomika)*, vol. 63, no. 6, pp. 247–258, Jun. 2017, https://doi.org/10.17221/212/2016-AGRICECON.

[17] S. Jain, S. P. Narayan, R. K. Dewang, U. Bhartiya, N. Meena, and V. Kumar, "A Machine Learning based Depression Analysis and Suicidal Ideation Detection System using Questionnaires and Twitter," in *2019 IEEE Students Conference on Engineering and Systems (SCES)*, Allahabad, India, May 2019, pp. 1–6, https://doi.org/10.1109/SCES46477.2019.8977211.

[18] K. O. Asare, Y. Terhorst, J. Vega, E. Peltonen, E. Lagerspetz, and D. Ferreira, "Predicting Depression From Smartphone Behavioral Markers Using Machine Learning Methods, Hyperparameter Optimization, and Feature Importance Analysis: Exploratory Study," *JMIR mHealth and uHealth*, vol. 9, no. 7, Jul. 2021, Art. no. e26540, https://doi.org/10.2196/26540.

[19] B. Sahoo and A. Gupta, "An Ensemble Kernelized-based Approach for Precise Emotion Recognition in Depressed People," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18873–18882, Dec. 2024, https://doi.org/10.48084/etasr.8785.

[20] M. A. Soeleman, D. Trisatya, A. Marjuni, P. N. Andono, and K. Kondo, "Sentiment Analysis on Movie and Product Reviews Using Supervised Learning and Ensemble Learning," *International Journal of Intelligent Engineering and Systems*, vol. 18, no. 4, pp. 169–198, May 2025, https://doi.org/10.22266/ijies2025.0531.11.

[21] K. Kroenke, R. L. Spitzer, and J. B. W. Williams, "The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener," *Medical Care*, vol. 41, no. 11, pp. 1284–1292, Nov. 2003, https://doi.org/10.1097/01.MLR.0000093487.78664.3C.

[22] R. Rois, M. Ray, A. Rahman, and S. K. Roy, "Prevalence and predicting factors of perceived stress among Bangladeshi university students using machine learning algorithms," *Journal of Health, Population and Nutrition*, vol. 40, no. 1, Dec. 2021, Art. no. 50, https://doi.org/10.1186/s41043-021-00276-5.

[23] U. B. Munir, M. S. Kaiser, U. I. Islam, and F. H. Siddiqui, "Machine Learning Classification Algorithms for Predicting Depression Among University Students in Bangladesh," in *Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering*, 2022, pp. 69–80, https://doi.org/10.1007/978-981-16-7597-3_6.

[24] A. Priya, S. Garg, and N. P. Tigga, "Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms," *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020, https://doi.org/10.1016/j.procs.2020.03.442.

[25] Md. I. H. Nayan *et al.*, "Comparison of the Performance of Machine Learning-based Algorithms for Predicting Depression and Anxiety among University Students in Bangladesh: A Result of the First Wave of the COVID-19 Pandemic," *Asian Journal of Social Health and Behavior*, vol. 5, no. 2, pp. 75–84, Apr. 2022, https://doi.org/10.4103/shb.shb_38_22.

[26] H. S. Alsagri and M. Ykhlef, "Machine Learning-Based Approach for Depression Detection in Twitter Using Content and Activity Features," *IEICE Transactions on Information and Systems*, vol. E103.D, no. 8, pp. 1825–1832, Aug. 2020, https://doi.org/10.1587/transinf.2020EDP7023.

[27] K. S. Na, S. E. Cho, Z. W. Geem, and Y. K. Kim, "Predicting future onset of depression among community dwelling adults in the Republic of Korea using a machine learning algorithm," *Neuroscience Letters*, vol. 721, Mar. 2020, Art. no. 134804, https://doi.org/10.1016/j.neulet.2020.134804.

[28] A. Sau and I. Bhakta, "Screening of anxiety and depression among the seafarers using machine learning technology," *Informatics in Medicine Unlocked*, vol. 16, 2019, Art. no. 100149, https://doi.org/10.1016/j.imu.2018.12.004.

[29] Md. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, and M. S. Uddin, "An in-depth analysis of machine learning approaches to predict depression," *Current Research in Behavioral Sciences*, vol. 2, Nov. 2021, Art. no. 100044, https://doi.org/10.1016/j.crbeha.2021.100044.

[30] M. K. Myee, R. D. C. Rebekah, T. Deepa, G. D. Zion, and K. Lokesh, "Detection of Depression in Social Media Posts using Emotional Intensity Analysis," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16207–16211, Oct. 2024, https://doi.org/10.48084/etasr.7461.