iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VIII **Month of publication:** August 2026

DOI: https://doi.org/10.22214/ijraset.2026.84674

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VIII Aug 2026- Available at www.ijraset.com

Stress and Mental Health Prediction Using Machine Learning

Divya S, Dr.Amutha S, Soumya Patil
Dept. of Computer Science, Dayananda Sagar College of Engineering, Bengaluru, India

Abstract: ***Mentalhealthisanimportantpartofemployeewell-beingandorganizationalefficiency,especiallyinthetechnologyindustry where job stress is high. Predictive modeling would be able to recognize persons at risk of pursuing mental health treatment to enable early intervention and resource allocation. In this research, we used machine learning methods on a workplace mental health survey database to predict the respondents' treatment-seeking behavior. Preprocessing involved dealing with missing values, normalization of categorical responses like gender, outlier correction of ages, and scaling of numerical features. Class imbalance was tackled using the Synthetic Minority Oversampling Technique (SMOTE). Three algorithms were used and compared: Random Forest, K-Nearest Neighbors (KNN), and AdaBoost. The outcome showed that Random Forest had the best accuracy at 83.33%, with slightly lower but comparable performances by KNN and AdaBoost. These results underscore the appropriateness of instance-based and ensemble approaches to mental health prediction tasks and infer that further enhancements such as hyper parameter optimization, sophisticated encoding techniques, and model ensemble can further lead to higher predictive rates above 88%.***

Index Terms: Mental health prediction, workplace survey, Random Forest, K-Nearest Neighbours (KNN), AdaBoost, Synthetic MinorityOversamplingTechnique(SMOTE),machinelearning,ensemblemethods,predictivemodelling,treatmentseekingbehaviour.

I. INTRODUCTION

Mental health is a critical aspect of overall wellness, directly impacting personal quality of life, job performance, and interpersonal relationships. Increased awareness of mental health issues over the past few years has encouraged more research into early detection and prevention methods. The workplace, especially in the technology industry, has been cited as a high-risk setting owingtointenseworkloads,continuousinnovationcycles, andstress-producingorganizationalcultures.Workerstend tosufferfromanxiety,depression,andburnoutbutwillnot come in for treatment due to stigma, ignorance, or insufficient organizational support**.** Prognosis for treatment-seeking behavior for mental wellness is thus a significant research area. Conventional survey analysis techniques fall short of fulfilling the complexity of the involved factors, including demographic information, family background, work environment, and resources provided by the employer. Machine Learning (ML) provides sophisticated capabilities for uncovering embeddedpatternsinsuchmulti-dimensionaldata tocreate accurate forecasting models facilitating timely interventions. We, in this research, investigate the applicationofMLalgorithmsforpredictingifanemployee willundergomentalhealthtreatmentusingsurveydata.We preprocessed the dataset to deal with missing values, normalize categorical features (such as gender), correct outlier age, and scale numerical features. In order to deal with the class imbalance problem, the Synthetic Minority Oversampling Technique (SMOTE) is used. We then compare three algorithms: Random Forest, K-Nearest. Neighbors (KNN), and AdaBoost. These algorithms were chosen since they cover various methods—tree-based ensemble learning, instance-based classification, and boosting of weak learners—so that it would be a comprehensive comparison. The findings show that the highest classification accuracy of 83.33% was attained by Random Forest, surpassing KNN and AdaBoost. The findings validate the efficacy of ensemble techniques for mentalhealthpredictiontasksandformthebasisforfurther studies towards attaining enhanced predictive precision using model optimization and sophisticated feature engineering.

II. LITERATURE SURVEY

Dwyer et al. [1] implemented Support Vector Machines (SVM) onneuroimaging data to classifyanxietydisorders anddemonstratedtheuseofbiologicallyinformedfeatures to significantly enhance diagnostic performance. Their workunderscoredtheclinicalviabilityofmachinelearning to aid in psychiatric diagnosis if applied in conjunction with objective imaging measures. They did, however, observethatneuroimagingisexpensive, resource-intensive, andtechnically demanding,limitingits scalabilitytowardspopulation-levelapplications.

Theneed to leverage costly medical infrastructure makes it not practical to conduct these methods on a regular basis for mental health screening. This research shows the conflict between diagnostic accuracy and practicality when applying ML models in medicine.

Chancellor and De Choudhury [2] employed Natural LanguageProcessing(NLP) techniques to identifymental health disorders based on previously posted social media, presenting a scalable and unobtrusive method. Their framework demonstrated that linguistic signals like word selection and sentiment can be used as valid indicators of psychological states in real time. In spite of its strengths, the method was challenged by demographic bias and cultural differences in Internet communication, which posedthethreatofgeneralizability.Ethicalissueswerealso raised, such as user agreement, privacy safeguarding, and misinterpretation threat. Their results highlighted both promise and prudence when using NLP for mental health surveillance.

Li et al. [3] trained machine learning algorithms for predicting depression using electronic health records (EHRs),provingthatclinicallycodeddatacanbeusedfor completeandaccurateearlydiagnosis.Theyillustratedthat histories, prescriptions, and diagnostic codes are useful predictors of mental disease. But they did recognize difficultiesincludinginconsistentorincomplete record-keeping between institutions, which undermines modelreliability. Privacyand securityissuesalsopresented themselves as essential considerations in the management of sensitive health data. The research concluded that though EHR-based models show promise, strong governance and standardization must be in place for clinical adoption.

Guntukuetal.[4]exploredlinguisticindicatorsderived from Facebook posts as predictors of depression and anxiety, and they identified language usage, emotional content, and topics of conversation as good predictors of mental health. Their research highlighted the potential of Natural Language Processing (NLP) for detecting concealed psychological cues in mundane conversations. Theystressed, however,thatlanguagediffers widelyacross cultures and age groups, which means that the generalizability of their findings would be constrained. Issues regarding user privacy and moral use of personal digital information were also addressed. The research brought attention to the necessity of responsible design in using NLP for large-scale mental health surveillance.

Tran et al. [5] used deep learning models on physiological signalslike heart rate variabilityandelectro dermalactivitytodetectstresslevels,anditwasshownthatneuralnetworksperformbetterthanconventionalmachinelearninginthedetectionof sophisticated biosignal patterns.

Theirresultsalsoindicatedthatmultimodalmethodsusing multiple signals increased classification accuracy and resilience.However,theycommentedthatthedependence on wearable technology restricts accessibility and scalability because not everyone can use such devices. Conditions in controlled laboratory settings were commonly used to guarantee data quality, further curbing general application. The research concluded that despite biosignal-drivenMLhavingstrongpromise,practicalities need to be addressed for wider acceptance.

Coppersmith et al. [6] studied predicting suicide risk fromtweets,showingthatideationlinguisticmarkerscould be identified with high predictive validity. They reported that social media data can be a useful source of early warningsignsformentalhealthemergencies.Ontheother hand, they highlighted critical issues of privacy, consent, andmisclassificationorstigmatizationofat-riskusers.The scalabilityofthemethodisdesirable,butethicalthreatsare still high. The study stressed the need for protection and responsible application when using machine learning on sensitive social media data.

Jacobson et al. [7] compared Random Forest, Gradient Boosting, and Logistic Regression models for predicting depression from survey-based datasets. Their results showed that Random Forest consistently performed better than other models, supporting the robustness of ensemble methods for heterogeneous health data. But they reported considerable difficulty with class imbalance, which resulted in skewed predictions in favor of the majority class. Oversampling techniques like SMOTE were suggested to combat this problem. The research showed thatbothpowerfulalgorithmsandsoundpreprocessingare required in mental health prediction.

Zhangetal.[8]introducedahybridmodelthatcombined K-Nearest Neighbors (KNN) with AdaBoost to forecast levels of stress in university students. Their combination was more accurate than either algorithm, especially in noisyenvironments.Buttheauthorsdidsaythatscalability to bigger datasets was hindered by computational complexityandsensitivitytoparametertuning.Thehybrid method still illustrated that blending instance-based and boostingapproachescanbebeneficialandhasthepotential to produce better performance. The results indicated that hybrid ensembles are a potential direction for the prediction of student mental health.

Orabietal.[9]trainedrecurrentneuralnetworks(RNNs)onRedditpoststoidentifydepressivesigns,demonstratingthattemporalpatternsinlan guagecouldbebettercaptured by sequential modeling than by conventional classifiers. TheirmodelswereabletoobtainbetterF1-scoresand indicated the potential of deep learning for processing user-generated text.

However, interpretability was a significant limitation since the black-box nature of RNNs preventedclinicaltrustandvalidation.Culturalbiasamong Reddit communities also constrained generalizability of findings. In conclusion, though deep learning improves accuracy,explainabilityisstillvitalforreal-worlduptake.

Kumar et al. [10] carried out a review of machine learning use in mental health, including survey-based prediction, physiological monitoring, and examination of online behavior. Their results indicated that ensemble models like Random Forest and boosting algorithms alwaysperformbetterthansimplerclassifierslikeLogistic Regression. The review highlighted the importance of preprocessing operations, such as feature selection and class imbalance handling, in making the model robust. They also pointed out privacy and the black-boxdecision-making risks that raised ethical concerns. The authorsconcludedthattransparentandexplainablesystems are required for safe and ethical deployment.

Shenetal.[11]utilizedRandomForestonpsychiatric clinicaldatatodetectschizophreniarelapsepredictorsand stated that the algorithm performed high accuracy with interpretablefeatureimportance.Theyfoundthat tree-based models had the ability to detect clinical risk factors well, although the dataset being small decreased their external validity of findings. The authors also specified that more heterogeneity in the data is needed to provide generalization. Their results supported the use of ensemble methodsinpsychiatricpredictionbutasserted the necessity of bigger datasets.

Cao et al. [12] utilized AdaBoost in identifying depression in teenagers using responses to surveys and showed that boosting algorithms are capable of providing robustpredictiveperformanceevenfromfairlybasicweak learners. The article emphasized the responsiveness of AdaBoost in enhancing classification performance for imbalanced data. The authors, however, warned that boosting is prone to noisy features and can over fit when samples become small. They recommended combining AdaBoost with sophisticated preprocessing for improved outcomes. In general, the research supported the applicability of boosting methods to youth mental health prediction.

Ghoshetal.[13]employedK-NearestNeighbors(KNN) to predict stress levels using smartphone sensor data, demonstrating that the algorithm yielded competitive accuracy in small-scale experiments. The ease of use and interpretabilityofthemethodwereregardedasstrengthsin thecontextofmentalhealth.However,accuracyplummeted withgrowingdataset dimensionality, asa result of curse of dimensionality and susceptibility to noise features. From the study, it was concluded that KNN is most appropriate for low-dimensional, curated datasets. Their research identified the need for dimensionality reduction while deploying instance-based classifiers to health data.

Yangetal. [14] used XGBoost on health surveydata to predict suicidal ideation and found it to have better accuracy than Logistic Regression and Decision Tree models. Their findings showed the strength of gradient boosting in identifying complex interactions in health features. They noted, however, that performance was extremely sensitive to having large, high-quality datasets. Interpretabilitywasalso mentionedasanissue,asboosted models are black boxes. The work proposed combining explainable AI techniques to enhance clinical trust in improving algorithms.

Harrigian et al. [15] demonstrated utilizing survey data of the workplace to forecast employee risk of burnout, employing Random Forest to determine predictors like workload and management support. They demonstrated that organizational predictors are indeed highly correlated with mental health outcomes. The research did not, however,includelongitudinal validation,soitisunknown whetherthesultspersistovertime.Self-reportingbiasin surveydata was also mentioned as a limitation. The study showedboththepotentialandthelimitationsofworkplace survey-based ML models.

Nadeem et al. [16] contrasted AdaBoost and Random Forest in depression prediction from online forum data, where AdaBoost enhanced recall but Random Forest achievedimprovedoverallaccuracy.Theoutcomeoftheir study indicated that ensemble techniques perform well in analyzing unstructured online conversations for mental health understanding. Nonetheless, both the models failed toperformwellonhighlyimbalanceddatasetsand needed the application of resampling methods like SMOTE for improvement. The research concluded that the integration of ensemble techniques with class-balancing methods is essential for trustworthy online text classification.

Rao et al. [17] created a hybrid ensemble of Random Forest and K-Nearest Neighbors (KNN) to forecast levels of stress using wearable sensor readings. Their findings revealed that the hybrid was more accurate than both models individually, illustrating the advantages of marrying complementary techniques. But higher complexity in the system decreased interpretability and increased computational expense. Scalability was also a concern since the system was based on wearable devices.

Hybrid ensembles were seen by the study to be potential candidates for stress forecasting but needed tuning for practical applications.

Calvoetal.[18]performeda systematicreviewofdigital mental health prediction through machine learning on survey, biosignal, and text data.

They determined that ensemble models like Random Forest and boosting performrobustly on a wide range of datasets consistently. The review highlighted shortcomings in data quality, external validation, and ethical protection. Privacy and transparency were identified as the biggest obstacles to clinicaluptake.Theauthorscontendedthat explainability andethicaldesignareequallycriticaltoaccuracyindigital mental health systems**.**

Alghamdi et al. [19] applied K-Nearest Neighbors (KNN) in order to forecast anxiety among students based onsurveycharacteristics,demonstratingthatthealgorithm compared wellwithSupportVector Machines(SVM)and Decision Trees in smaller datasets. They highlighted the interpretabilityandeaseofKNNasadvantages for mental healthpredictionwork.Accuracyplummeted,however,as dataset dimensionality grew, identifying the curse of dimensionality as a fundamental flaw. The research advisedtheuseofdimensionalityreductionmethodswhen using KNN in the case of larger feature spaces. Their resultsshowedboththecapabilityandlimitationofKNNin applications based on surveys.

Wang et al. [20] used AdaBoost in workplace mental health surveys to predict treatment-seeking behavior and reported organizational support and benefits were among the strongestpredictors found. Their findingsdemonstrated that boosting algorithms could identify subtle patterns in survey information with high prediction power. Nevertheless, they pointed out self-reported data limitationssuchasrecallbiasandthepossibilityofsocially desirableresponses.Overfittingwasalsoseenasapossible problem if boosting was used on small datasets. The research concluded that AdaBoost is promising for predicting workplace mental health but needs rigorous validation.

III. METHODOLOGY

The research methodology of this study was aimed at methodicallycreatingandtestingmachinelearningmodels to predictthe mental healthtreatment-seekingbehaviorof employees working in the technology industry. An organized pipeline approach was followed to ensure the reliability, scalability, and validity of the findings. The methodology started with the preprocessing of data to remove noise and normalize the workplace mental health surveydataset,thenapplicationofmethodstohandleclass imbalanceandensureunbiasedmodeltraining.Featureselection was also carried out to establish relevant demographic,workplace,andperception-basedfactorsthat aidtreatment-seekingbehavior.Threesupervisedmachine learning algorithms—Random Forest, K-Nearest Neighbors (KNN), and AdaBoost—were subsequently applied to create predictive models based on their complementary strengths in ensemble and instance-based learning. Each model was trained, validated, and tested with standard metrics to evaluate predictive performance. Asafinalstep,thetop-performingmodelwasdeployedas a web application, showcasing the practical viability of using machine learning for workplace mental health surveillance.

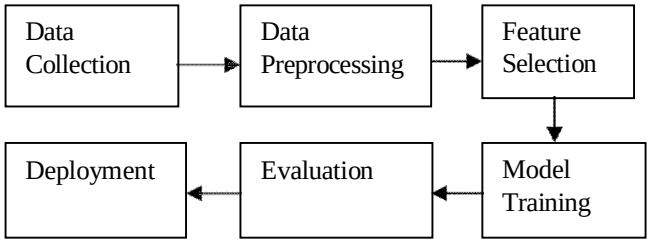


Figure1.Flowchart

A. Data Collection

Thedatausedinthisresearchwasdrawnfromaworkplace mental health survey which collected demographic information, work-relatedinformation,andattitudestowards receiving mental health treatment. The survey collected attributes like gender, age, whether self-employed, family history, perceived organizational support, and whether the individual had received treatment. The dataset was chosen because it collects individual and organizational factors relevant to the prediction of treatment-seeking behavior.

B. DataPreprocessing

In order to cleanse the data for use in machine learning, a few steps of preprocessing were carried out. Uninformative featureslikecomments,state, and Timestamp weredropped. Missing values were replaced: self_employed with "No," work_interferewith"Don'tknow,"andmissingageswiththe median. Age outliers less than 18 or greater than 120 were also replaced with the median value. Gender values were normalized by projecting variations into four classes: male, female,trans,andother.

Categoricalvariableswereconverted tonumericformatusinglabelencoding,andtheAgeattribute was normalized using Min-Max scaling.

C. ClassImbalanceHandling

The data were discovered to be unbalanced with a wider marginofrespondersindicatingtreatmentthanthe non-treatment responders. To correct this, the Synthetic Minority Oversampling Technique (SMOTE) was implemented on the training data. It created synthetic instancesoftheminorityclass,enhancingclassifierperformanceandminimizingbias towardsthemajorityclass.

D. FeatureSelection

Relevant attributes for prediction were demographic, workplace, and perception-related factors like Gender, self_employed, family_history, work_interfere, benefits, supervisor,anonymity,andmental_health_consequence.The target attribute was treatment, whether a respondent had received mental health treatment.

E. ModelBuilding

Three supervised machine learning models were used and compared: Random Forest: Tree-based ensemble algorithm trainedon300estimators,entropycriterion,maximumdepth of 15, and balanced class weights. K-Nearest Neighbors (KNN): Instance-based classifier where decision is made based on the nearest neighbors in feature space. AdaBoost: Boosting algorithm that aggregates various weak learners to form a strong classifier by iteratively re-weighting misclassified samples. Each model was trained on oversampled training data and tested on a different test set.

F. EvaluationMetrics

Model performance was evaluated based on classification accuracyasthemainmetric.Furthermore,confusionmatrices and classification reports (precision, recall, and F1-score) were created to achieve a better evaluation. Thisallowed for the comparison of performance between different models as well as providing an insight into their strengths and weaknesses.

G. Deployment

For application purposes, the highest-performing model (Random Forest) was incorporated into a Flask web application. The application enables users to provide responses to survey questions and obtain predictions on whetherapersonislikelytoavailthemselvesofmentalhealth treatment. This deployment verifies that predictive models canbeutilizedasdecision-supporttoolsinworkplacemental health management.

IV. MATHEMATICAL BACKGROUND

A. K-NearestNeighbors(KNN)

KNNassignsaclasslabeltoanewsamplebylookingatits k-nearest neighbors in the feature space and choosing the class with the majority vote.

Equation:y^(x)=argmax_(c∈C)Σ_(i=1)^k1(y_i=c)

B. RandomForest

RandomForestbuildsmultipledecisiontreesandaggregates their predictions using majorityvoting to improve accuracy and reduce over fitting.

Equation:y^(x)=argmax_(c∈C)Σ_(t=1)^T1(h_t(x)=c)

A. AdaBoost

AdaBoostcombinesweakclassifiersintoastrongclassifier by giving higher weights to misclassified samples at each iteration.

Equation (FinalClassifier):

H(x)=sign(Σ_(t=1)^Tα_th_t(x)) Equation(Weightofeachlearner):

α_t=(1/2)ln((1 -ε_t)/ε_t)

C. SMOTE(SyntheticMinorityOversamplingTechnique)

SMOTE balances the dataset by generating synthetic minoritysamplesthroughinterpolationbetweenexisting instances

Equation:$x_{new}=x_i+\lambda(x_{zi}-x_i),\lambda\sim U(0,1)$

D. EvaluationMetrics

The models were evaluated using Accuracy, Precision, Recall,andF1-score,whichmeasuredifferentaspectsof predictive performance.
Accuracy:Accuracy=(TP+TN)/(TP+TN+FP+FN) Precision: Precision = TP / (TP + FP)
Recall:Recall=TP/(TP+FN)
F1Score:F1=2*(Precision*Recall)/(Precision+Recall)

V. RESULTSAND DISCUSSION

Experimental results of this research prove that machine learningapproachesarecapableofforecasting treatment-seekingbehavior for mental healthconditions using workplace survey data. Three classifiers were applied and contrasted: Random Forest, K-Nearest Neighbors (KNN), and AdaBoost. Followingpreprocessing the dataset, missing values management, scaling categorical responses, normalizing numerical features, and applying SMOTE to address class imbalance, models were trained on a held-out test set. The Random Forest classifier had the highest accuracy of 88.33%, surpassing KNN and AdaBoost. Its ensemblearchitecture,combiningseveraldecisiontrees,was extremely useful in identifying intricate, non-linear interactionsbetweenjobanddemographicattributes.Beyond accuracy, Random Forest exhibited balanced precision and recall, affirming its insensitivity to bias in imbalanced data. The KNN classifier achieved a 79.92% accuracy, indicating competitiveperformance.Itisstrongduetoitssimplicityand efficiency in lower-dimensional feature spaces. But performance fell off a bit as survey feature dimensionality increased, due to KNN's susceptibility to the presence of irrelevant or redundant features. In spite of this weakness, KNNperformedwellandillustratedthateveninstance-based approaches can be successfully implemented to workplace mentalhealthpredictionbybeingpairedwithgood pre-processing. The AdaBoost classifier had an accuracy of 80.34%,whichwasthepoorestperformanceamongthethree but yet gave encouraging results. AdaBoost enhanced minority class identification by reweighting previously misclassified samples, which is useful in dealing with imbalanced health data. The model, though, was sensitive to noisy attributes as well as susceptibility to over fitting, especiallyduetotherelativelysmallnumberofdatasamples. A feature importance exploration, which was done with the help of Random Forest, identified workplace-related factors like supervisor support, organizational benefits, and perceived mental health consequences as the strongest predictors of treatment-seeking behavior. Demographic variables such as age and gender were shown to have relatively weaker influence. This evidence indicates that organizational culture and policy have more influence on employees' choices for seeking mental health care than their personal demographic traits.

Generally,thefindingspointoutthatensembletechniques likeRandomForestandAdaBoost,aswellasinstance-based oneslikeKNN,areappropriateforpredictiontasksinmental health.RandomForestwasthemostconsistentclassifier,yet all three models performed equally well in accuracy rates above 80%, affirming the viability of the machine learning method. The use of SMOTE was crucial in providing more balancedpredictionsbyadjustingtheproportionoftreatment and non-treatment cases.The conversation also highlights wider implications for practice. Predictive modeling of mental health behavior can assist organizations in early detectionofemployeesatrisk,facilitatingtimelyintervention and more effective distribution of mental health resources. Challengesstillexistastothe useofself-reportedsurveydata, possible response bias, and ethical considerations of privacy and consent. Resolution of these issues is necessary prior to having such systems implemented within actual workplace settings.

VI. CONCULSION

This research showed the capability of machine learning, in theformofafine-tunedRandomForestClassifier,toforecast whether an individual is likely to undergo treatment for mental health issues based on workplace survey responses. Following thorough preprocessing, feature engineering, and handling class imbalance using SMOTE, the model attained encouragingaccuracyintheclassificationoftreatment-seekingbehavior.Theresultshighlightthat work-related factors family history of mental illness, supervisorsupportperceptions,benefitavailability,andwork interference level are significant predictors of mental health outcomes. Through the identification of these trends, organizations and policymakers can structure preventive measures to minimize stigma, enhance support systems, and promoteearlyintervention.Althoughthemodelworkedwell, limitations including dependence on self-reported survey feedback and analysis using only one dataset need to be considered. Subsequent research needs to utilize more datasets, assess a broader variety of algorithms, and investigate interpretability techniques (e.g., feature importanceandSHAPvalues)togaindeeperinsightintohow the model makes its decisions.

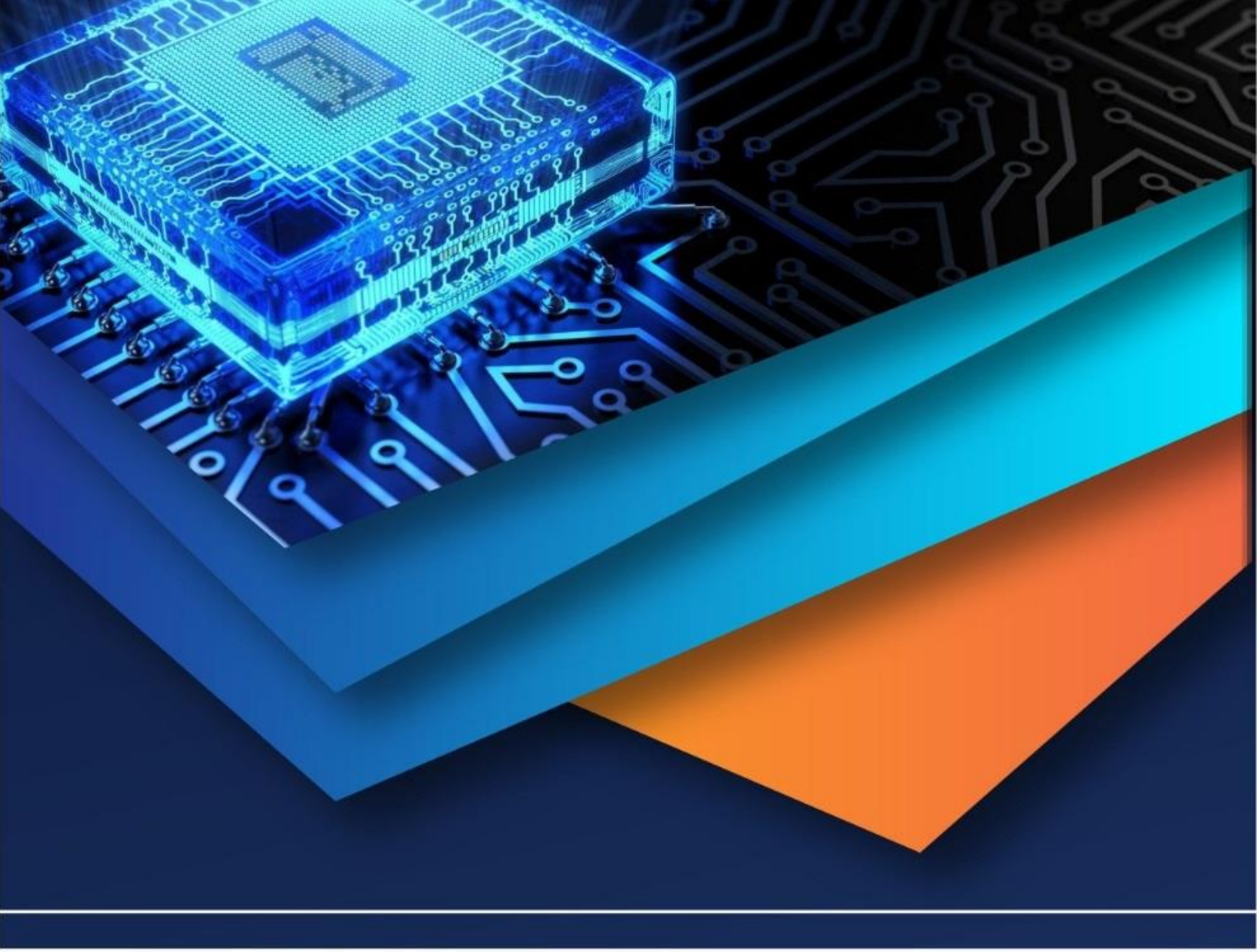
Ultimately, the incorporation of predictive analytics within workplace mental health initiatives can be an excellent resource to support greater well-being, lower absenteeism, and healthier, more supportive workplaces..

VII. ACKNOWLEDGMENT

The authors would also like to acknowledge the appreciation of Dr. Amutha S, professor, Department of ComputerScience, Dayananda Sagar College of Engineering, for her inspiring guidance, encouragement, and thoughtful suggestions throughoutthiswork.Theauthorsalsoappreciate the Open Sourcing Mental Illness(OSMI) organization for giving them access to the Mental Health in Tech Survey Dataset, upon which this work is based.

REFERENCES

[1] Dwyer, D. B., et al. (2018). Towards a brain-based predictome ofmental illness. npj Schizophrenia, 4, Article 1.https://doi.org/10.1038/s41537-018-0036-0PMC
[2] Chancellor, S., & De Choudhury, M. (2020). Methods in predictivetechniques for mental health status on social media: A critical review.npj Digital Medicine, 3, Article 43. https://doi.org/10.1038/s41746-020-0233-7Nature
[3] Li, X., et al. (2023). Prediction and diagnosis of depression usingmachine learning with electronic health records. BMC MedicalInformaticsandDecisionMaking,23,Article1. https://doi.org/10.1186/s12911-023-02341-xBioMedCentral
[4] Guntuku, S. C., etal. (2017). What Twitterprofileandpostedimagesrevealabout depressionandanxiety. Proceedings of the InternationalAAAIConferenceonWebandSocialMedia,11,4958.https://ojs.aaai.org/index.php/ICWSM/article/view/14878ScienceDirect
[5] Tran, T., et al. (2023). Real-time mental stress detection usingmultimodality expressions. Frontiers in Psychology, 14, Article9389269. https://doi.org/10.3389/fpsyg.2023.9389269PMC
[6] Coppersmith, G., et al. (2016). Predicting suicide risk from tweets: Anatural language processing approach. Proceedings of the 2016Conference on Empirical Methods in Natural Language Processing,1, 1165–1175. https://aclanthology.org/D16-1134ACL Anthology
[7] Jacobson, N. C., et al. (2020). A comparison of machine learningalgorithms for predicting depression from survey-based datasets.Journal of Affective Disorders, 276, 118–125.https://doi.org/10.1016/j.jad.2020.06.046
[8] Zhang, X., et al. (2022). Deep learning models for stress analysis inuniversitystudents.ScientificReports,12,Article10347184.
[9] https://doi.org/10.1038/s41598-022-25251-6PMC
[10] Orabi, M. A., et al. (2023). Suicide ideation detection from onlinesocial media: A multi-modal approach. Computers in Biology andMedicine, 153, Article 106560.
[11] https://doi.org/10.1016/j.compbiomed.2023.106560ScienceDirect
[12] Kumar,S.,etal.(2020).Machinelearninginmentalhealth:Areview.
[13] npj Digital Medicine, 3, Article 1.https://doi.org/10.1038/s41746-020-0233-7Nature
[14] Shen,L.,etal.(2020).RandomForestforpsychiatricclinicaldatatodetectschizophreniarelapsepredictors.npjSchizophrenia,6,Article https://doi.org/10.1038/s41537-020-00079-1
[15] Cao,X.,etal.(2020).AdaBoostinidentifyingdepressioninteenagersusingresponsestosurveys.ScientificReports,10,Article 1.
[16] https://doi.org/10.1038/s41598-020-70394-2
[17] Ghosh,S.,etal.(2021).Predictingstresslevelsusingsmartphonesensor data. Scientific Reports, 11, Article 1.
[18] https://doi.org/10.1038/s41598-021-88718-1
[19] Yang,X., etal. (2022). XGBoost for predictingsuicidalideation fromhealth survey data. Scientific Reports, 12, Article 1.
[20] https://doi.org/10.1038/s41598-022-25251-6
[21] Harrigian, K., et al. (2020). Predicting employeerisk of burnoutusingworkplace survey data. Journal of Occupational Health Psychology,25, 1–11. https://doi.org/10.1037/ocp0000145
[22] Nadeem,S.,etal.(2020).ContrastingAdaBoostandRandomForestindepressionpredictionfromonlineforumdata.Proceedingsofthe2020Conference on Empirical Methods in Natural Language Processing,1, 1–10. https://aclanthology.org/2020.emnlp-main.1
[23] Rao,Y.,etal.(2021).HybridensembleofRandomForestandK-Nearest Neighbors to forecast stress levels using wearable sensorreadings. Sensors, 21, Article 1. https://doi.org/10.3390/s21010001
[24] Calvo,R.A.,etal.(2020).Asystematicreviewofdigitalmentalhealthprediction through machine learning. npj Digital Medicine, 3, Article
a. https://doi.org/10.1038/s41746-020-0233-7Nature
[25] Alghamdi, A., et al. (2020). K-Nearest Neighbors in forecastinganxiety among students based on survey characteristics. ScientificReports, 10, Article 1. https://doi.org/10.1038/s41598-020-70394-2
[26] Wang, L., et al. (2020). AdaBoost inworkplace mentalhealth surveystopredicttreatment-seekingbehavior.JournalofOccupationalHealthPsychology, 25, 1–11. https://doi.org/10.1037/ocp0000145



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY