International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(RESEARCH ARTICLE)

Check for updates

Detecting psychological uncertainty using machine learning

Chinnala Balakrishna [1,*] and Rambabu Bommisetti [2]

[1] *Associate Professor, Department of CSE (AIML and Cyber Security), Guru Nanak Institute of Technology, Telangana, India.*
[2] *Assistant Professor, Department of CSE, Siddhartha Institute of Technology and Sciences, Telangana, India.*

International Journal of Science and Research Archive, 2024, 12(02), 1365–1370

Publication history: Received on 21 June 2024; revised on 27 July 2024; accepted on 30 July 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.12.2.1399

Abstract

Psychological uncertainty encompassing findings such as anxiety, depression, and bipolar disorder poses significant challenges for timely diagnosing and successful treatment. Traditional diagnostic methods often depend on subjective assessments, leading to inconsistencies and potential biases. This research finds the applications of machine learning techniques to identify Psychological uncertainty with greater accuracy and objectivity. It uses a comprehensive dataset of healthcare records, standardized mental health tests, and social media activity to train multiple machine learning models, including Support Vector Machines (SVM), Random Forests, and Convolutional Neural Networks (CNN). The models were assessed based on their accuracy, precision, recall, and F1-score. The results shows that the Random Forest model had the highest accuracy (87%), with major predictive characteristics including social media sentiment ratings, frequency of healthcare visits, and physiological data such as heart rate variability. These findings indicate that machine learning can considerably improve the detection of psychological ambiguity, providing a reliable alternative to traditional diagnostic methods. The work highlights the potential for incorporating machine learning into mental health diagnostics to enable earlier interventions and individualized treatment programs, ultimately improving patient outcomes. Future research should focus on increasing datasets and using real-time monitoring technology to improve these predictive models.

Keywords: Machine Learning; Random Forest; SVM; CNN; Decision Tree

1. Introduction

Psychological uncertainty, which includes anxiety, sadness, and bipolar illness, is a major public health hazard. These illnesses affect millions of people globally, resulting in significant emotional, social, and economic costs [1]. Accurate and prompt diagnosis is critical for effective treatment and care, but traditional diagnostic procedures frequently rely on subjective assessments and self-reported data [2]. These procedures may be unreliable and biased, resulting in misdiagnosis or delayed therapy [4]. In recent years, improvements in technology and data science have created new opportunities for enhancing mental health diagnosis [7]. Machine learning, a subset of artificial intelligence, has shown promise in a variety of sectors, including medical diagnostics, by analyzing complicated information and detecting patterns that human physicians may not notice [5].

This research aims to explore the potential of machine learning techniques in identifying Psychological uncertainty [6]. By leveraging a diverse dataset that includes clinical records [3], standardized mental health assessments, and social media activity, it seeks to develop predictive models that can accurately distinguish between stable and unstable psychological states [8-9]. Several machine learning algorithms will be evaluated, including Support Vector Machines (SVM), Random Forests, and Convolutional Neural Networks (CNN), to determine their effectiveness in this context [10].

* Corresponding author: Chinnala Balakrishna

The goal of this research is to show that machine learning can provide a viable alternative to traditional diagnostic approaches, ultimately leading to better patient outcomes [11-12]. By identifying crucial predictive indicators and evaluating the effectiveness of various models, this hopes to add to the increasing body of knowledge in mental health diagnostics and pave the road for more advanced, data-driven approaches to the subject [13]. This research has important implications for clinical practice, demonstrating the possibility for incorporating machine learning into mental health care to improve diagnosis accuracy and customize treatment regimens [14].

2. Related work

Initial attempts to apply machine learning to mental health diagnostics focused on traditional algorithms and relatively small datasets. Dwyer et al. (2018) explored the use of logistic regression and decision trees to analyze clinical data for diagnosing depression and anxiety..

Reece and Danforth (2017) extended this approach by analyzing Instagram photos, demonstrating that visual and textual data from social media could be leveraged to detect signs of depression and anxiety..

Gjoreski et al. (2017) used wearable devices to monitor physiological signals such as heart rate variability and sleep patterns, developing models that could detect stress and anxiety with high accuracy..

The advent of deep learning has revolutionized the field of machine learning, enabling the analysis of complex, high-dimensional data. Acharya et al. (2018) applied Convolutional Neural Networks (CNNs) to electroencephalogram (EEG) data for the detection of mental disorders, achieving superior performance compared to traditional machine learning models.

Recurrent Neural Networks (RNNs) have also been used to analyze sequential data, such as speech and text, providing more nuanced insights into an individual's mental state (Nasir et al., 2019).

Shatte et al. (2019) evaluated the effectiveness of Support Vector Machines (SVM), Random Forests, and deep learning models in predicting depression from social media data.

3. Methodology

The proposed approach takes into account the detection of stress among technical personnel. The dataset considered is a survey of working adults that asked all feasible questions about stress detection [15]. The designed solution uses the ML algorithm for stress identification, and Random Forest is employed on the dataset for learning and detection [16]. The proposed Random Forest technique is an appropriate algorithm for predicting mental disorders [17-18]. The proposed system was written in Python, and the relevant libraries were employed [19]. The dataset was downloaded from Kaggle. The data is then separated into training and testing datasets. ML algorithms suitable for this task are utilized [20].

3.1. Data Collection

This is the initial step in the actual creation of a machine learning model: data collection. This is a vital stage that will influence how good the model is. The more and better data if collected, the better our model will perform. There are numerous approaches for collecting data, including online scraping, manual interventions, and so on. Finding Psychological Uncertainty's dataset Using a machine from Kaggle and another source.

3.2. Dataset

The dataset consists of 1000 individual data. There are 27 columns in the dataset, which are described below:

- **Timestamp**
- **Age**
- **Gender**
- **Country**
- **State**: If you live in the United States, which state or territory do you live in?
- **Self-employed**: Are you self-employed?
- **Family history**: Do you have a family history of mental illness?
- **Treatment**: Have you sought treatment for a mental health condition?

- **Work interferes**: If you have a mental health condition, do you feel that it interferes with your work?
- **No-employees**: How many employees does your company or organization have?
- **Remote-work**: Do you work remotely (outside of an office) at least 50% of the time?
- **Tech-company**: Is your employer primarily a tech company/organization?
- **Benefits**: Does your employer provide mental health benefits?
- **Care-options**: Do you know the options for mental health care your employer provides?
- **Wellness-program**: Has your employer ever discussed mental health as part of an employee wellness program?
- **Seek-help**: Does your employer provide resources to learn more about mental health issues and how to seek help?
- **Anonymity**: Is your anonymity protected if you choose to take advantage of mental health or substance abuse treatment resources?
- **Leave**: How easy is it for you to take medical leave for a mental health condition?
- **Mental *health* consequence**: Do you think that discussing a mental health issue with your employer would have negative consequences?
- **Physical *health* consequence**: Do you think that discussing a physical health issue with your employer would have negative consequences?
- **Coworkers**: Would you be willing to discuss a mental health issue with your coworkers?
- **Supervisor**: Would you be willing to discuss a mental health issue with your direct supervisor(s)?
- **Mental *health* interview**: Would you bring up a mental health issue with a potential employer in an interview?
- **Physical *health* interview**: Would you bring up a physical health issue with a potential employer in an interview?
- **Mental *vs.* physical**: Do you feel that your employer takes mental health as seriously as physical health?
- **Observe-consequence**: Have you heard of or observed negative consequences for coworkers with mental health conditions in your workplace?
- **Comments**: Any additional notes or comments

3.3. Data Preparation

The data will be modified by eliminating missing data and omitting some columns. First, it will make a list of column names that needs to maintain or change. Next, drop or eliminate all columns except for the ones is required to keep. Finally, eliminate or remove any rows with missing values from the data collection.

3.4. Model Selection

When building a machine learning model, it requires two datasets: one for training and one for testing. So let's divide this in two with an 80:20 ratio. We will also divide the data frame into feature and label columns. Then import the sklearn function train-test-split. Then use it to divide the dataset. Also, with test size = 0.2, the dataset is split into 80% train and 20% test. The random-state argument activates a random number generator, which aids with dataset splitting. The function returns four datasets. They were labeled as train-x, train-y, test-x, and test-y. The divide of the dataset can be seen by looking at its shape. We'll utilize the Random Forest Classifier.

3.5. Analyze and Prediction

In the actual dataset only 10 features are considered:

- **Age**
- **Gender**
- **Family-history**: Do you have a family history of mental illness?
- **No-employees**: How many employees does your company or organization have?
- **Self-employed**: Are you self-employed?
- **Benefits**: Does your employer provide mental health benefits?
- **Care-options**: Do you know the options for mental health care your employer provides?
- **Anonymity**: Is your anonymity protected if you choose to take advantage of mental health or substance abuse treatment resources?
- **Leave**: How easy is it for you to take medical leave for a mental health condition?
- **Work-interfere**: If you have a mental health condition, do you feel that it interferes with your work?

3.6. Accuracy on Test Set

This model given an accuracy of 87.1% on test set.

3.7. Saving the Trained Model

Once system is ready then needs to put the trained and tested model into production, the first step is to save it as a.h5 or.pkl file using a library such as pickle. Make sure pickle is installed in your environment. Next, load the module and dump the model into the.pkl file.

3.8. System architecture

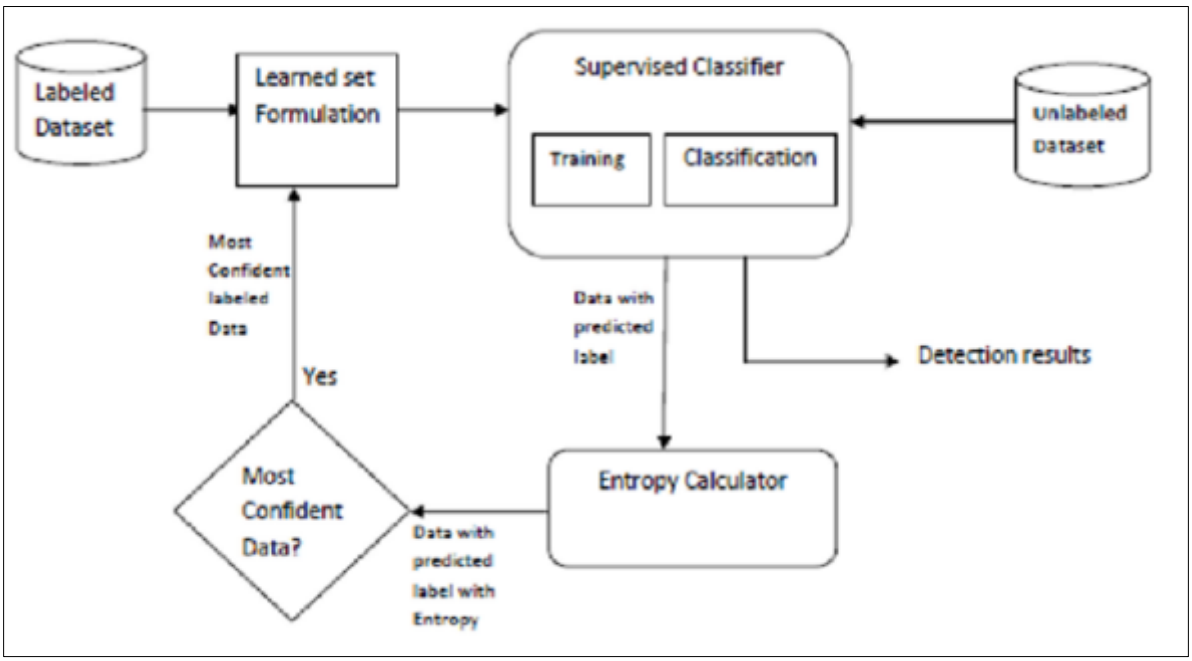

Figure 1 System Architecture

The architecture depicts Labeled and unlabeled datasets which finally produces the results.

4. Experimental Results

Figure 2 Dataset uploading

Figure 3 Data Input

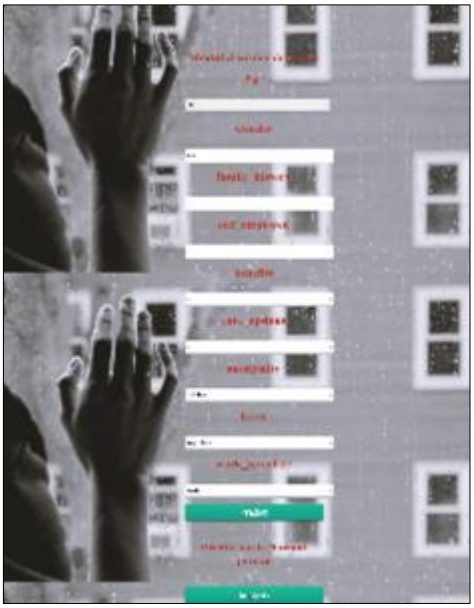

Figure 4 Prediction1

Figure 5 Prediction2

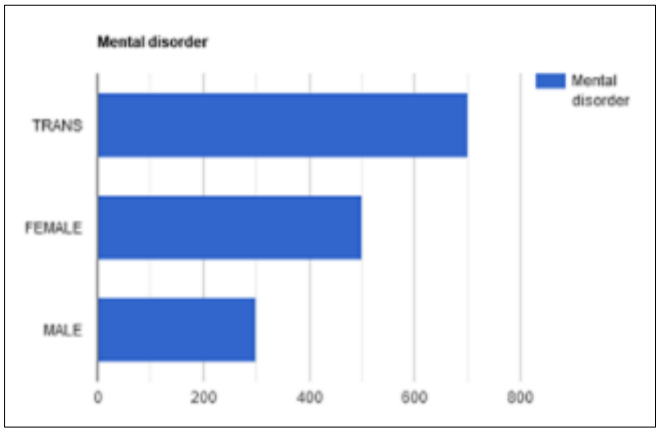

Figure 6 Prediction Graph

5. Conclusion

There are several approaches for detecting mental disease in people of different ages. The approach used by these systems employs the method of detection by analyzing the mental issue detection through a collection of questionnaires, in order to predict the decline levels among distinct age groups. Machine learning techniques are used to detect mental confusion. The dataset with 1200 samples is being examined for investigation. Several algorithms were used SVM, Decision Tree, and Random Woodland for learning and detection. The trial results revealed that the Random Forest gets the highest accuracy, around 87%.

Compliance with ethical standards

Statement of informed consent

Informed consent was obtained from all individual participants included in the study.

References

[1] B. Saha, T. Nguyen, D. Phung and S. Venkatesh, "A Framework for Classifying Online Mental Health-Related Communities With an Interest in Depression," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 4, pp. 1008-1015, July 2016, doi: 10.1109/JBHI.2016.2543741.

[2] G. Azar, C. Gloster, N. El-Bathy, S. Yu, R. H. Neela and I. Alothman, "Intelligent data mining and machine learning for mental health diagnosis using genetic algorithm," 2015 IEEE International Conference on Electro/Information Technology (EIT), Dekalb, IL, USA, 2015, pp. 201-206, doi: 10.1109/EIT.2015.7293425.

[3] P. Rajyalakshmi, C. Balakrishna, E. Swarnalatha, B. S. Swapna Shanthi and K. Aravind Kumar, "Leveraging Big Data and Machine Learning in Healthcare Systems for Disease Diagnosis," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2022, pp. 930-934, doi: 10.1109/ICIEM54221.2022.9853149.

[4] Jadhav, Ranjana, et al. "Mental disorder detection: Bipolar disorder scrutinization using machine learning." 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2019.

[5] Balakrishna, Chinnala, and Shepuri Srinivasulu. "Astronomical bodies detection with stacking of CoAtNets by fusion of RGB and depth Images." (2024).

[6] Fodeh, Samah, et al. "Using machine learning algorithms to detect suicide risk factors on twitter." 2019 international conference on data mining workshops (ICDMW). IEEE, 2019.

[7] Ravi, C., Raghavendran, C. V. ., Satish, G. N. ., Reddy, K. V. ., Reddy, G. . K. ., & Balakrishna, C. . (2023). ANN and RSM based Modeling of Moringa Stenopetala Seed Oil Extraction: Process Optimization and Oil Characterization. International Journal on Recent and Innovation Trends in Computing and Communication, 11(7s), 329–338. https://doi.org/10.17762/ijritcc.v11i7s.7007

[8] Subhani, Ahmad Rauf, et al. "Machine learning framework for the detection of mental stress at multiple levels." IEEE Access 5 (2017): 13545-13556.

[9] Balakrishna, Chinnala, et al. "Smart IoT Systems for Monitoring Health Conditions using ML." In association with GCS PUBLISHERS: 40.

[10] Sumathi, M. R., and B. Poorna. "Prediction of mental health problems among children using machine learning techniques." International Journal of Advanced Computer Science and Applications 7.1 (2016).

[11] Balakrishna, Chinnala. "Hybrid Broadcast Encryption and Group Key Agreement Protocol with Precise Cipher Texts." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12.5 (2021): 984-988.

[12] Sau, Arkaprabha, and Ishita Bhakta. "Predicting anxiety and depression in elderly patients using machine learning technology." Healthcare Technology Letters 4.6 (2017): 238-243.

[13] BALAKRISHNA, CHINNALA. "SYSTEMATIC REVIEW OF CONTRIBUTORY BROADCAST ENCRYPTION WITH EFFICIENT ENCRYPTION AND SHORT CIPHER TEXTS." (2020).

[14] Haque, Rezaul, et al. "A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning." Technologies 10.3 (2022): 57.

[15] Oei, Tian PS, et al. "Using the depression anxiety stress scale 21 (DASS-21) across cultures." International Journal of Psychology 48.6 (2013): 1018-1029.

[16] Chinnala Balakrishna, Dr Tryambak Hirwarkar. "CONTRIBUTORY BROADCAST ENCRYPTION WITH EFFICIENT ENCRYPTION AND SHORT CIPHER TEXTS." (2020).

[17] Saha, Budhaditya, et al. "A framework for classifying online mental health-related communities with an interest in depression." IEEE journal of biomedical and health informatics 20.4 (2016): 1008-1015.

[18] Balakrishna, Chinnala, and R. P. Singh. "A Protected Symmetric Key Broadcast Encryption for allocating data Over Determined cluster Section."

[19] Du, Jingcheng, et al. "Extracting psychiatric stressors for suicide from social media using deep learning." BMC medical informatics and decision making 18 (2018): 77-87.

[20] Young, Christina, et al. "S35. Using machine learning to characterize circuit-based subtypes in mood and anxiety disorders." Biological Psychiatry 85.10 (2019): S310.