**TABLE OF CONTENTS**

|  |  |
| --- | --- |
|  | Page No. |
| TITLE PAGE .................................................................................................................... | 1 |
| DECLARATION .............................................................................................................. | 2 |
| CERTIFICATE …........................................................................................................... | 3 |
| ACKNOWLEDGEMENT.................................................................................................. | 4 |
| ABSTRACT...................................................................................................................... | 5 |
|  |  |
| CHAPTER 1 INTRODUCTION | 6-9 |
| 1.1.          Introduction ……………………................................................... | 6 |
| 1.2 Problem Statement.……………………....................................... | 7 |
| 1.2.          Objective………………………………………………………… | 8 |
| 1.3.          Scope……………………………………………………………... | 9 |
| CHAPTER 2 LITERATURE REVIEW…………………………………………….... | 10-13 |
| CHAPTER 3 PROPOSED METHODOLOGY …………………………………........ | 14-15 |
| 3.1 Methodology…………………………………………………………... | 14 |
| 3.2 Flowchart………………………………………………………………. | 15 |
| CHAPTER 5 CONCLUSION …....................................................................................... | 16 |
| REFERENCES….............................................................................................................. | 17 |

Project Synopsis

on

**Mental Health Analysis Using Machine Learning**

Submitted as a part of course curriculum for

**Bachelor of Technology**

in

**Computer Science**

A picture containing text, clipart

Description automatically generated

**Submitted by**

Vanshika Namdev

1900290310184

Vedant Panday

1900290120125

Shivansh Kumar Yadav

1900290120109

**Under the Supervision of**

Mrs. Neha Shukla

Assistant Professor

**KIET Group of Institutions, Ghaziabad**

**Department of Computer Science**

**Dr. A.P.J. Abdul Kalam Technical University**

**2022-2023**

**DECLARATION**

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Signature of Students:

Name: Shivansh Kumar Yadav

Roll No.: 1900290120109

Name: Vanshika Namdev

Roll No.: 1900290310184

Name: Vedant Panday

Roll No.: 1900290120125

**CERTIFICATE**

This is to certify that Project Report entitled “**Mental Health Analysis Using ML**” which is submitted by **Shivansh Kumar Yadav, Vedant Panday, Vanshika Namdev** in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science of Dr A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

**Guide Name:** Mrs. Neha Shukla

(Assistant Professor)

**ACKNOWLEDGEMENT**

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to Ms. Neha Shukla, Assistant Professor, Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his/her constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of Dr. P. K Singh, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Name: Vanshika Namdev

Roll No: 1900290310184

Name: Vedant Panday

Roll No.: 1900290120125

Name: Shivansh Kumar Yadav

Roll No.: 1900290120109

**ABSTRACT**

Increases in the occurrence and global burden of mental illness have made the prevention and treatment of mental disorders a public health priority. A 2017 US report showed that an estimated 46.6 million adults have been affected by a mental illness. This equates to nearly 20% of the US population alone.

Responding to the need for more effective mental health services, the role of digital technology for improving access, engagement, and outcomes of therapeutic treatment is increasing in importance and has led to a wide range of health technologies and applications. These include mobile apps and wearable devices to assist symptom monitoring and health risk assessments, computerized treatments, and mental health peer or community support**.**

Our objective is to use various machine learning techniques to analyse mental health. Machine learning approaches have emerged as a novice way of constructing meaningful representations from unstructured data.

There are very less cases of early mental illness detection, and even when diagnosed a lot of patients do not come forward to seek help from the professional counselors because of insecurity, inferiority and social pressures that are associated with these diseases and our approach will help to overcome these barriers

**Chapter 1: INTRODUCTION**

* 1. **INTRODUCTION**
* Increases in the occurrence and global burden of mental illness have made the prevention and treatment of mental disorders a public health priority.
* A 2017 US report showed that an estimated 46.6 million adults have been affected by a mental illness. This equates to nearly 20% of the US population alone.
* Responding to the need for more effective mental health services, the role of digital technology for improving access, engagement, and outcomes of therapeutic treatment is increasing in importance and has led to a wide range of health technologies and applications.
* These include mobile apps and wearable devices to assist symptom monitoring and health risk assessments, computerized treatments, and mental health peer or community support**.** 
  1. **PROBLEM STATEMENT**
* Around 1 in 5 of the world's children and adolescents have a mental disorder.
* Depression is a common mental disorder. Globally, it is estimated that 5.0% of adults suffer from depression (1).
* Almost 800 000 people die by suicide every year; 1 person dies from suicide every 40 seconds. Suicide is the second leading cause of death in individuals aged 15-29 years.
* People with severe mental disorders die 10 to 20 years earlier than the general population.
* The global economy loses about US$ 1 trillion per year in productivity due to depression and anxiety.
  1. **RESEARCH OBJECTIVES**
* Our objective is to use various machine learning techniques to do mental health analysis on the public data available on the social media. Machine learning approaches have emerged as a novice way of constructing meaningful representations from unstructured data.
* There are very less cases of early mental illness detection, and even when diagnosed a lot of patients do not come forward to seek help from the professional counselors because of insecurity, inferiority and social pressures that are associated with these diseases and our approach will help to overcome these barriers.
  1. **SCOPE**
* Right now, we are only working on analysis of mental health using social media data to perform prediction, but we can further extend this to provide solutions for the predicted disorder.
* We can also apply deep learning techniques to evaluate and perform our analysis on mental health.

**Chapter 2**

**LITERATURE REVIEW**

**2.1 Application of deep and machine learning techniques for multi-label classification performance on psychotic disorder diseases**

This study presents Psychotic Disorder Diseases (PDD) dataset with five labels: bipolar disorder, vascular dementia, attention-deficit/hyperactivity disorder (ADHD), insomnia, and schizophrenia as a multi-label classification problem. The study also investigates the use of deep neural network and machine learning techniques such as multilayer perceptron (MLP), support vector machine (SVM), random forest (RF) and Decision tree (DT), for identifying hidden patterns in patients’ data. The study furthermore investigates the symptoms associated with certain types of psychotic diseases and addresses class imbalance from a multi-label classification perspective. The performances of these models were assessed and compared based on an accuracy metric. The result obtained revealed that deep neural network gave a superior performance of 75.17% with class imbalance accuracy, while the MLP model accuracy is 58.44%. Conversely, the best performance in the machine learning techniques was exhibited by the random forest model, using the dataset without class imbalance and its result, compared with deep learning techniques, is 64.1% and 55.87%, respectively. It was also observed that patient’s age is the most contributing feature to the performance of the model while divorce is the least. Likewise, the study reveals that there is a high tendency for a patient with bipolar disorder to have insomnia; these diseases are strongly correlated with an R-value of 0.98.

**2.2 A Dataset for Research on Depression in social media**

Language provides a unique window into thoughts, enabling direct assessment of mental-state alterations. Due to their increasing popularity, online social media platforms have become promising means to study different mental disorders. However, the lack of available datasets can hinder the development of innovative diagnostic methods. Tools to assist health practitioners in screening and monitoring individuals under potential risk are essential. In this paper, the have presented a new a dataset to foster the research on automatic detection of depression. To this end, we present a methodology for automatically collecting large samples of depression and non-depression posts from online social media.

**2.3 Mental health analysis using deep learning for feature extraction**

There is an immense need to analyze and monitor a person‟s mental health as justified in our previous work. Feature extraction is a decisive part of all the data mining related tasks. We utilize deep learning feature extraction algorithm like sentence embedding to analyze mental health of persons from their social media posting and behavioral features and combine it with the traditional machine learning algorithms to enhance their performance. Newly, deep learning approaches have emerged as a novice way of constructing meaningful representations from unstructured data. Not only they are good in the data encoding but also carry the semantic meaning with them which help in modeling better. We find that deep learning feature extraction helps in classifying the normal users from the non-normal users as compared to their traditional counterparts. Also, the newer models attain a very low false positive rate. The best accuracy received by our model is 89%.

**2.4 A Chatbot-based Mobile Application to Predict and Early-prevent Human Mental Illness**

The paper outlines the design of a chatbot mobile application for human mental health, which can determine whether an individual has a mental illness and suggest prevention methods by using machine learning algorithms. There are some popular mental counseling applications in the markets, most of which are based on one mental survey or examination. We noticed that this can be done in a more natural/intelligent way - texting, talking or even videoing. Our application can interact with users by texting, talking, and videoing, which are based on natural language processing (NLP) and integrated continuous emotion dialogue analysis and sentence/audio generation

**2.5 Deep learning for prediction of depressive symptoms in a large textual dataset**

This work proposes an efficient approach using Long Short-Term Memory (LSTM)-based Recurrent Neural Network (RNN) to identify texts describing self-perceived symptoms of depression. The approach is applied on a large dataset from a public online information channel for young people in Norway. The dataset consists of youth’s own text-based questions on this information channel. Features are then provided from a one-hot process on robust features extracted from the reflection of possible symptoms of depression pre-defined by medical and psychological experts. The features are better than conventional approaches, which are mostly based on the word frequencies (i.e., some topmost frequent words are chosen as features from the whole text dataset and applied to model the u nderlying events in any text message) rather than symptoms.

**2.6 COVID-19 pandemic and mental health consequences: Systematic review of the current evidence**

During the COVID-19 pandemic general medical complications have received the most attention, whereas only few studies address the potential direct effect on mental health of SARS-CoV-2 and the neurotropic potential. Furthermore, the indirect effects of the pandemic on general mental health are of increasing concern, particularly since the SARS-CoV-1 epidemic (2002–2003) was associated with psychiatric complications.

**2.7 Application of Machine Learning Methods in Mental Health Detection: A Systematic Review**

This study reviewed articles published in major databases between 2007 and 2018 through keyword searches. The articles were screened based on their titles and abstracts before the full texts were reviewed. The articles were coded in accordance with data set (e.g., data sources, keywords, and geographical locations), method of data analysis, machine learning or deep learning technique, classifier performance, and feature extraction method. 22 articles were selected for review from the total of 2770. As OSNs exhibit high potential as a data source in early detection of mental health problems, most researchers used text analysis on a new data set extracted from different OSNs sources. The extracted data were examined using a statistical analysis or machine learning techniques. Several studies also applied multimethod techniques, which included distributing questionnaires while requesting for the respondents' consent to later access and extract information from his/her OSNs account.

**2.8 Behavioral Modeling for Mental Health using Machine Learning Algorithms**

This research work proposes to apply various machine learning algorithms such as support vector machines, decision trees, naïve bayes classifier, K-nearest neighbor classifier and logistic regression to identify state of mental health in a target group. The responses obtained from the target group for the designed questionnaire were first subject to unsupervised learning techniques. The labels obtained because of clustering were validated by computing the Mean Opinion Score. These cluster labels were then used to build classifiers to predict the mental health of an individual. Population from various groups like high school students, college students and working professionals were considered as target groups. The research presents an analysis of applying the aforementioned machine learning algorithms on the target groups and also suggests directions for future work.

**2.9 Machine Learning for Mental Health in social media: Bibliometric Study**

Social media platforms provide an easily accessible and time-saving communication approach for individuals with mental disorders compared to face-to-face meetings with medical providers. Recently, machine learning (ML)-based mental health exploration using large-scale social media data has attracted significant attention. Publications addressing social media and ML in the field of mental health were retrieved from the Scopus and Web of Science databases. We analyzed the publication distribution to measure productivity on sources, countries, institutions, authors, and research subjects, and visualized the trends in this field using a keyword co-occurrence network. The research methodologies of previous studies with high citations are also thoroughly described.

**2.10 Predicting mental health problems in adolescence using machine learning techniques**

Predicting which children will go on to develop mental health symptoms as adolescents is critical for early intervention and preventing future, severe negative outcomes. Although many aspects of a child’s life, personality, and symptoms have been flagged as indicators, there is currently no model created to screen the general population for the risk of developing mental health problems. Additionally, the advent of machine learning techniques represents an exciting way to potentially improve upon the standard prediction modelling technique, logistic regression. Therefore, we aimed to I.) develop a model that can predict mental health problems in mid-adolescence II.) investigate if machine learning techniques (random forest, support vector machines, neural network, and XGBoost) will outperform logistic regression.

**Chapter 3 PROPOSED METHODOLOGY**

**3.1 Methodology**

* First, we are going to extract data. For this we have used Twitter API for extracting tweets by searching keywords like depression, suicidal, anxiety, trauma and many more.
* Then we are going to clean and visualize this data.
* After preprocessing, we will apply machine learning multi-class algorithms to predict the disorder.
* Then we are going to evaluate performance of our model and will use various metrics. Further, we can deploy this using Flask.

**3.2 Flowchart**

Diagram

Description automatically generatedFig 1. This figure represents the flowchart for mental health analysis using machine learning.

**CONCLUSION**

* We have already filed the **patent** for this project and hopefully it will be published soon.
* We are going to publish a **research paper** on our research so that we can share our project’s result with everyone else and to further improve it even more.
* We hope that our outcome will contribute to address this serious issue in our society and help in mental health awareness too.

**References**

[1]Elujide, I., Fashoto, S., Fashoto, B., Mbunge, E., Folorunso, S. and Olamijuwon, J., 2021. *Application of deep and machine learning techniques for multi-label classification performance on psychotic disorder diseases*.

[2] Rissola, E., Bahrainian, S. and Crestani, F., 2020. *A Dataset for Research on Depression in Social Media*.

[3] J.Joshi, D., Makhija, M., Nabar, Y., Nehete, N. and S. Patwardhan, D., 2021. *Mental health analysis using deep learning for feature extraction | Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*.

[4] Podrazhansky, A., Zhang, H., Han, M. and He, S., 2020. *A Chatbot-based Mobile Application to Predict and Early-prevent Human Mental Illness | Proceedings of the 2020 ACM Southeast Conference*.

[5] Uddin, M., Dysthe, K., Folstad, A. and Brandtzaeg, P., 2021. *Deep learning for prediction of depressive symptoms in a large textual dataset*.

[6]Vindegaard, N. and Benros, M., 2021. *COVID-19 pandemic and mental health consequences: Systematic review of the current evidence*.

[7] RAHMAN, R., OMAR, K., NOAH, S., DANURI, M. and AL-GARADI, M., 2020. *Application of Machine Learning Methods in Mental Health Detection: A Systematic Review*.

[8] Srividya, M., Mohanavalli, S. and Bhalaji, N., 2018. *Behavioral Modeling for Mental Health using Machine Learning Algorithms*.

[9] Kim, J., Lee, D. and Park, E., 2021. *Machine Learning for Mental Health in Social Media: Bibliometric Study*.

[10] Tate, A., McCabe, R., Larsson, H., Lundström, S., Lichtenstein, P. and Kuja-Halkola, R., 2020. *Predicting mental health problems in adolescence using machine learning techniques*.