Machine Learning based Detection of Depression and Anxiety

Guna Sekhar Sajja
Information Technology
University of the Cumberlands
Kentucky, USA

ABSTRACT

Anxiety is something that everyone encounters at some point in their lives. Anxiety is a word that is used in everyday situations to represent the discomforts and negative feelings that a person has when they are tense or worried. Machine learning approaches enable computers to create data that may be utilized for factual study to attain a certain range of performances. The use of computer frameworks to automate decision-making based on test data is encouraged throughout the development of the models for the test data. This article presents a model for predicting feelings of anxiety and depression. There is a set of speech data that is used as input into this framework. A preprocessing step was performed on this data set to remove noise from the data and make the original data set more consistent. The input data set is then submitted to a variety of machine learning approaches, including Nave Bayes, Random Forest, and Support Vector Machines (SVM). It is necessary to classify the information. In this section, the categorization results of several approaches are discussed and contrasted.

Keywords

Anxiety, Depression, Prediction, Machine Learning, Classification, Speech data.

1. INTRODUCTION

Anxiety [1] is something that everyone experiences at some point in their lives. Anxiety is a term that is commonly used in everyday situations to describe the uncomfortable and unpleasant feelings that an individual has when presented with stressful or frightening situations. Anxiety may be brought on by several different circumstances. Most of the time, though, it is the result of stress. It is stress-related anxiety that is the most often cited cause for seeking therapy among college students. Every eighth undergraduate student seeks therapy for stress-related issues throughout their time in college, according to the American Counseling Association. Anxiety is defined as an internal fear that is produced by a strong urge to decide. It is a collection of symptoms induced by a lack of ability to cope with the stresses and demands of daily life. Student anxiety is one of the most common problems that they face during their college years. Some people may become overwhelmed by the added demands that come with moving to a new region and being away from family, as well as the stress of attempting to do well in school. Among the most common and common mental illnesses are anxiety disorders, which are also among the most prevalent. Anxiety disorder is characterized by alterations in emotion, thought, behavior, and physiology, among other characteristics. They can manifest itself in a variety of ways among college students, including adjustment disorders, anxiety tests, social phobia, and anxiety disorders induced by drugs and alcohol consumption. [2] [3] [4]

Psychological, physical, and environmental issues are all possible manifestations of anxiety symptoms [4] [5]. Various types of anxiety present themselves in several ways. These include excessive worrying and fear as well as restlessness and overly emotional responses as well as negative thoughts. While some people appear calm when they are frightened, their brains are always thinking, which has a negative impact on their overall quality of life. Individuals who are anxious often complain of chest tightness, a racing or pounding heart, and a pit in their stomach. Some people endure headaches, excessive sweating, and even the need to urinate as a result of their anxiety. In addition to being concerned and terrified, anxiety symptoms include turning blank during a test, feeling powerless when executing duties, and losing interest in a difficult subject.

When a person has depression [6], they are depressed and reluctant to engage in any activity. Depression can have an impact on a person's ideas, behavior, feelings, and physical fitness. Individuals suffering from depression may experience feelings of sadness, anxiety, hesitancy, powerlessness, unworthiness, guilt, irritability, and restlessness. They may lose interest in activities that were formerly delightful, lose appetite or eat excessively, have difficulty concentrating, recalling information, and making decisions, and they may consider or attempt suicide. College is a difficult time for students, and sadness is common throughout this period. Campus life may be stressful, and university students are prone to feeling down in the dumps.

Millions of people suffer with depression [7], which is characterized by a profound sense of despair, gloom, and sadness. It is more than just a "black mood," but rather a persistent sense that a person cannot regulate and that interferes with everyday activities and functioning. According to predictions, depression will overtake heart disease as the leading cause of death by 2023.

Depression [8] may manifest itself in a variety of ways, ranging from mild to severe. Depression is a prevalent mental health problem. The symptoms manifest themselves in four key areas of human work: emotional, cognitive, physical, and behavioral, with the most evident manifestation being mood problems, among other things. Depressive symptoms include changes in nutrition or weight, disturbed sleep and insomnia, decreased energy, a sense of shame, disordered thinking, focusing, and decision-making, persistent death thoughts, and the contemplation of suicide. Depression. Anhedonia, or a lack of interest in things that were formerly enjoyable, is a contributing factor to social disengagement. There are a variety of negative health consequences associated with

depression, including lower immune function and worse recovery from injury or illness.

Machine learning [9] [10] is a branch of information technology that differs from traditional computer methodologies. Algorithms are sets of instructions that are explicitly encoded into a computer and then used by the computer to identify or solve a problem, according to traditional thinking. In a given sector, machine learning technologies allow computers to prepare data inputs for factual inquiry by analyzing patterns in data. It fosters the development of computer frameworks for test data models, with the goal of automating the process of making decisions based on data.

Machine learning is a dynamic field that is always evolving. Uncontrolled learning and supervised training are two of the most common forms of learning. Methods based on machine learning. In the same way, the closest K-algorithm, decision-making, and deep learning are examples of algorithmic strategies used in machine learning applications. Explore the most extensively used programming languages in machine learning by highlighting some of the good and bad qualities of each language. Furthermore, machine learning algorithms that are designed to eliminate these biases throughout the algorithm development process are available.

2. MACHINE LEARNING TECHNIQUES

Using linear regression, real values (housing expenses, number of phone calls, total sales, etc.) are calculated based on the variable being studied (s). In this context, the best possible line is used to create the relationship between the independent and dependent variables. This best-fit line is referred to as a regression line, and it is represented by the linear equation Y= a*X + b. This is known as the Y-regression line on X, and there is another X-regression line, which is represented by X = c*Y + d. This is known as the X-regression line on Y. The simplest way to grasp linear regression is to think back to a time when your child went through it. For example, let us say that we desire a fourth-grade kid to train understudies in his class by increasing weight demands without questioning them about their weights! What can we expect the youngster to do in the future? It is virtually certain that he/she would look (outwardly analyze) people's stature and employment, and then mastermind them utilizing a mixture of these unique aspects. This is, in reality, linear regression in action! When the youngster comprehends that height and construction would be connected to the burden by a connection that is similar to the equation above, he or she has completely grasped the concept. [11] A pattern recognition paradigm, such as the nearest neighbour technique, may be applied to both classification and regression problems. The k in the nearest region is a positive integer that is normally small, and it is sometimes abbreviated as k-NN to indicate that it is a close neighbour. This means that the inspection or regression inside the space will be comprised of k closest training instances. The following is the primary goal of classifying k-NN: Because of this role, you have been accepted into the class. This generates a new object in the class that has the greatest number of members that are closest neighbours to it. If k = 1, the item is classed by the object's nearest neighbor, which is the closest neighbor of the object. [12]

Outwardly, decision trees should speak to choices and appear to advise or guide decision-makers through the process. In machine learning and data mining, decision trees are used as a predictive model to make predictions. In order to make decisions on the data's objective worth, these models summarize the information. With decision tree learning, the goal is to develop a model that can anticipate the estimation of an objective based on the data that is provided. However, the leaves in the prescient model speak to the choices regarding the objective esteem of the facts, however the branches in the prescient model talk to information's traits that are resolved through perception. In the process of "learning" a tree, raw data is divided into subsets based on attribute esteem tests, which are then re-run on each of the inferred subgroups in an iterative fashion. The recursion procedure will be accomplished when the subset at a node has the same proportionate incentive as the intended output at the node in which it is implemented. In order to assess whether or not someone should go fishing, we may take into consideration a number of different factors. This dataset includes information on climate conditions and barometric weight conditions. The decision rules are frequently represented in the form of if-then-else statements. The deeper the tree, the more difficult the set of rules is. [13]

To mimic the processing of light and sound in the human brain, deep learning attempts to replicate the processing of light and sound into vision and hearing. The deep learning engineering system is comprised of several layers of a neural artificial network that was developed on hardware and is driven by biological neural networks, as described above. Deep learning is used to find or modify the highlights of data in a number of nonlinear preparation layers, and it is particularly effective in identifying and changing the highlights of data (or representations). The contribution of the progressive layer is represented by the output of one layer. Deep learning algorithms can be watched and used for data categorization, or they can be left unattended to do pattern analysis on large datasets. Machine learning is now accomplished via the use of algorithms, and deep learning retains the most information and is capable of outperforming humans on some subjective tasks. Deep learning has emerged as the most feasible approach for the artificial intelligence industry as a result of these characteristics. [14]

This is a binary linear classification system that is not probabilistic. It is called the Support Vector Machine (SVM). It produces a training model that categorises the data into one or more target classes based on the information in the data. The data objects are represented as points on a map of space. A noticeable gap separates the

objects belonging to distinct categories, causing the gap's breadth to expand in length. According on whatever side of the gap they originate from, the new instances are assigned to certain target classes. It is also possible to do non-linear classification on input datasets that have not been labelled by employing the support vector machine. Support vector machines classify data via unsupervised learning, which is necessary because there are no objective classes to which the examples may be allocated in the first place. Once a cluster has been built, new instances are added to it based on the functions that have been defined. The author proposes a strong model-based recommendation system based on the non-linear support vector machine, which is described in detail in the paper. In the field of unlabeled data processing, non-linear support vector machine methods are the most often utilized methodology, and they are found in a wide range of industrial applications. [15]

According to author [16], Random Forest is an ensemble learning approach for classification and regression operations

that may be used for classification and regression operations. During the training phase, it generates a huge number of decision trees and uses regression techniques to predict the outcomes of the individual trees. It has a low variance and makes speedy connections between the many parts of the data that is provided for prediction. The reason for the initial lack of interest for this technology is because random forest classification algorithms are difficult to comprehend.

3. RESULT ANALYSIS

There are 600 records in the input data set. 340 records are connected to depression sufferers, while the remaining 260 records are related to normal persons. This data set has been preprocessed to eliminate noise from the data and make the input data set consistent. The input data set is then subjected to different machine learning methods such as Nave Bayes, Random Forest, and SVM. Data classification is carried out. The classification results of several methods are compared. Table 1 displays these findings.

Table1: Accuracy of Machine Learning algorithms

Accuracy of Machine Learning Algorithms	
Algorithm name	**Accuracy (%)**
SVM	**94**
Random Forest	**78**
Naïve Bayes	**76**

4. CONCLUSION

Anxiety is something that everyone encounters at some point in their lives. Anxiety is a term that is commonly used in everyday situations to describe the uncomfortable and unpleasant feelings that an individual has when presented with stressful or frightening situations. When computers are given data inputs for factual enquiry, they can use machine learning approaches to arrange the data such that the answers are within a specified range. It encourages the use of computer frameworks to develop models for test data, with the goal of automating decision-making based on data inputs in order to save time. This article proposes a model for predicting feelings of anxiety and depression. There is a set of speech data that is used as input into this framework. A preprocessing step was performed on this data set in order to remove noise from the data and make the original data set more consistent. The input data set is then submitted to a variety of machine learning approaches, including Nave Bayes, Random Forest, and Support Vector Machines (SVM). It is necessary to classify the information. In this section, the categorization results of several approaches are discussed and contrasted.

5. REFERENCES

[1] A. Ahmed, R. Sultana, M. T. R. Ullas, M. Begom, M. M. I. Rahi and M. A. Alam, "A Machine Learning Approach to detect Depression and Anxiety using Supervised Learning," 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), 2020, pp. 1-6, doi: 10.1109/CSDE50874.2020.9411642.

[2] Priya, A., Garg, S., &Tigga, N. (2020). Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. Procedia Computer Science, 167, 1258-1267. doi: 10.1016/j.procs.2020.03.442

[3] Richter, T., Fishbain, B., Markus, A. et al. Using machine learning-based analysis for behavioral differentiation between anxiety and depression. Sci Rep 10, 16381 (2020). https://doi.org/10.1038/s41598-020-72289-9

[4] Su, C., Xu, Z., Pathak, J. et al. Deep learning in mental health outcome research: a scoping review. Transl Psychiatry 10, 116 (2020). https://doi.org/10.1038/s41398-020-0780-3

[5] Sharma A and Verbeke WJMI (2020) Improving Diagnosis of Depression With XGBOOST Machine Learning Model and a Large Biomarkers Dutch Dataset (n = 11,081). Front. Big Data 3:15. doi: 10.3389/fdata.2020.00015

[6] Cacheda F, Fernandez D, Novoa FJ, Carneiro V Early Detection of Depression: Social Network Analysis and Random Forest Techniques J Med Internet Res 2019;21(6):e12554 doi: 10.2196/12554

[7] Nemesure, M.D., Heinz, M.V., Huang, R. et al. Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. Sci Rep 11, 1980 (2021). https://doi.org/10.1038/s41598-021-81368-4

[8] Kumar, P., Garg, S., & Garg, A. (2020). Assessment of Anxiety, Depression and Stress using Machine Learning Models. Procedia Computer Science, 171, 1989-1998. doi: 10.1016/j.procs.2020.04.213

[9] Ihmig FR, H. AG, Neurohr-Parakenings F, Schäfer SK, Lass-Hennemann J, Michael T (2020) On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals. PLoS ONE 15(6): e0231517. https://doi.org/10.1371/journal.pone.0231517

[10] Moshe I, Terhorst Y, Opoku Asare K, Sander LB, Ferreira D, Baumeister H, Mohr DC and Pulkki-Råback L (2021) Predicting Symptoms of Depression and Anxiety Using Smartphone and Wearable Data. Front. Psychiatry 12:625247. doi: 10.3389/fpsyt.2021.625247

[11] M. Huang, "Theory and Implementation of linear regression," 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2020, pp. 210-217, doi: 10.1109/CVIDL51233.2020.00-99.

[12] K. Taunk, S. De, S. Verma and A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), 2019, pp. 1255-1260, doi: 10.1109/ICCS45141.2019.9065747.

[13] A. Rajeshkanna and K. Arunesh, "ID3 Decision Tree Classification: An Algorithmic Perspective based on Error rate," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 787-790, doi: 10.1109/ICESC48915.2020.9155578.

[14] Y. Yadav, V. Kumar, V. Ranga and R. M. Rawat, "Analysis of Facial Sentiments: A deep-learning Way," 2020 International Conference on Electronics and

Sustainable Communication Systems (ICESC), 2020, pp. 541-545, doi: 10.1109/ICESC48915.2020.9155622.

[15] J. Huang, J. Zhou and L. Zheng, "Support Vector Machine Classification Algorithm Based on Relief-F Feature Weighting," 2020 International Conference on Computer Engineering and Application (ICCEA), 2020, pp. 547-553, doi: 10.1109/ICCEA50009.2020.00121.

[16] Z. Bingzhen, Q. Xiaoming, Y. Hemeng and Z. Zhubo, "A Random Forest Classification Model for Transmission Line Image Processing," 2020 15th International Conference on Computer Science & Education (ICCSE), 2020, pp. 613-617, doi: 10.1109/ICCSE49874.2020.920190.