International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-5, March 2020

A Taxonomy for Mental Illness in Healthcare System

Sumitra Mallick, Mrutyunjaya Panda

Abstract: ***The rise in chronic diseases becomes strain in modern healthcare systems, some such disease is the mental illness. However, a solution is required to provide high quality care for the patient that is achieved through precision medicine. The precision medicine for mental illness disease is found to be taken care of because it treats the patient health by considering the individual profiles. The modern healthcare systems are encouraged to develop precision medicine to be used in the future and accessed anywhere in the world. This paper further provides a comprehensive survey of the state of art technologies of precision medicine approach for the mental illness disease. The aim of this research is to develop a taxonomy of mental illness disease by finding their symptoms along with genomes. The survey proposes a Medicare system should be developed for precision medicine considering deep learning techniques are the methodologies to classify the mental illness diseases for better accuracy and efficiency.***

Keywords: Precision Medicine, Mental Illness, Classification, Accuracy, Efficiency.

I. INTRODUCTION

Precision medicine is an emerging approach for disease treatment and prevention that considers individual variability in genes, environment and lifestyle for each person [1]. It takes personal care about the patient and has a very strong potential to consolidate modern e-health systems called as precision medicine. Precision medicine also sometimes referred to as personalized medicine or individual medicine [2]. Precision medicine system can then facilitate disease prediction and treatment, by determining whether an individual runs the risk of developing a disease. Some research has been carried out for genomic interpretation and genetic prediction [3]. sometimes precision medicine can also be defined based on genetic and genomic test where it treats the disease more effectively than traditional treatments. Precision medicine can be viewed as short term treatment due to faster recovery of the patient at whatever level the disease resides. Precision Medicine is very powerful in modern e-health care system that consolidates the personalized drugs to suite to a person. It also takes care of side effects which is very much limited because it applies specific drugs to specific individuals. The author [1] has studied about psychiatric disease that has rich information exists in serial measurement of mental health symptom scores. The basic approach is to plot a trajectory [2, 3] using multiple symptoms to make personalized predictions about future symptoms and some psychiatric events. This approach deals with estimation of a population that is average trajectory for symptom scores of other people and estimates the individual deviations from average trajectory. Then the author try to estimate the risk of experiencing an event using the symptom trajectory. This paper [1] demonstrates a study of anti psychotic therapy for schizophrenia.

It may be one of the biggest challenges globally by 2020 [5, 1] to predict in advance the common mental illness problem of children. Our objective is to facilitate the doctor with right drug at the right time for the right patient [6, 10] one of the most common methodologies is to select the best features that give more accurate result and suitable treatment for the concern patient. Hood et al introduce an "personalized, predictive, preventive, and participatory medicine" model in healthcare which shows the best performance of precision medicine in the patient health [9, 13]. It identifies critical factor with the help of some study which are based on clinicians, clinical researchers, practitioners in pharmaceutical industries, regulatory board members. In the same way the machine learning methodology are also used in the psychiatry field [5] to improve the prediction performance. There is a model called as Gaussian mixture model [10, 12] that is an unsupervised clustering algorithm that fit the data in a given structure.

The remainder of the paper is organized into five sections. In section II the related works of others have discussed in terms of genomic interpretation and genetic prediction, model for prediction, learning metrics and general framework for personalized prediction. In section III the taxonomy of mental illness disease are narrated. In section IV the precision medicine, machine learning and deep learning methodologies adopted by others in specific to classification techniques are elaborated. In section V the accuracy and efficiency of the classification process are viewed. In section VI the concluding remark of the proposed research and the future directions are elicited.

II. RELATED WORK

There are experiments on psychiatric diseases for the interpretation and prediction of genetic data. The author has developed [1] a model for prediction that take repeated symptom scores as input and the future symptoms are predicted. Also a binary event is predicted that is an event like suicidal behavior, hospitalization etc. In many health care sector the health care data which is widely available in Electronic Health Records (EHR's) for analyzing the existing data in order to improve the quality care of patients more efficiently [10].

Revised Manuscript Received on March 30, 2020.
* Correspondence Author
Sumitra Mallick, Department of computer science, KL University, Hyderabad, Telangana, India,
Mrutyunjaya Panda, [2]Department of computer science and applications, Utkal University, odisha, India

More specifically the challenge is to explore the heterogeneity, incompleteness and unbalanced structure. The author [16] has taken a case study for brain tumor diagnosis and the issues are addressed considering a disease subset called Oligodendroglioma which has high rate of responsiveness to chemotherapy.

The other type of disease [21] called Alzheimer's disease where the memory and thinking skill destroy slowly because of irreversible progressive in brain. In the human brain there is loss of connection between nerve cell (neurons) and the brain. The learning metrics derived for the personalized healthcare using two tasks are the disease prediction and clustering [21]. The learning process is also optimized to get improved accuracy by developing a new framework using patient EHR representations and similarities [21]. The similarities of patients are calculated using KNN and triplet loss based distance metric methods which are crucial for the performance of the algorithms. The triplet loss function learns a margin to separate the distance of negative and positive samples where as the KNN classification use the learned representation with positive label for similar pairs and negative label for dissimilar pairs. This paper [21] considers a dataset consists of 100,000 patients over 2 years and predicted for better longitudinal EHR sequence by learning in better way than other models to improve the accuracy.

The three main goals of personalized medicine is to prevent, diagnosis and treat the patients [2]. It is essential that early detection of problem for mental illness patient is to improve the quality. As a result the children may not suffer from complicated problems [7] where the author [9] discusses about mental health problem that basically comes at an early stage. There are eight different machine learning mechanism that are used for analyzing and diagnosis the medical health data. The comparative result shows that the classifier on various machine learning technique i.e. the Multilayer Perception, Multiclass Classifier and LAD Tree produced more accurate results in full attribute set. In this paper, the PANSS (Positive and Negative Syndrome Scale) used to identify two statistically distinct subgroups of patients to investigate whether the two identified subgroup are related to treatment or not.

As the researchers working for the personalized health care, the general framework for personalized prediction is segregated into two novel deep similarity learning [7] i.e. simultaneously learn patient representations and measure pair wise similarity [2]. The first step is to identify the patients suffering from similar kind of disease and the second step is to analyze those patients for diagnosis of the disease. The similarity of disease needs the accuracy and precise distributions of patients in groups are a concern in deep learning [17].

III. MENTAL ILLNESS

Mental illness persists when it is chronic of illness with recurring episode of schizophrenia, anxiety disorder and depression that that effects for longer period. Mental disorder is one type of disease where the patient may do any inactive thing which may lead to some other problem. This mental illness causes unusual and dramatic shifts in mood, energy, hallucination, lack of motivation, loss of interest, sleep disorder, sexuality and clear thinking. Analyzing an individual mental illness data such as behavioral information, motivational behavior, poorly sleeping information, anxiety disorder, eating disorder, poor social adjustment etc. are the point to know how she will respond to a particular therapy. The most significant improvements in performance have been achieved in the EEG datasets with high dimensionality along with changes in sensor placement, activity, and subject [17].

A. Taxonomy

We classify the mental disorder disease [1] into four types 1.schizophrenia 2. Anxiety disorder 3. Depression and 4. Alcoholism disorder that are shown in figure 1.

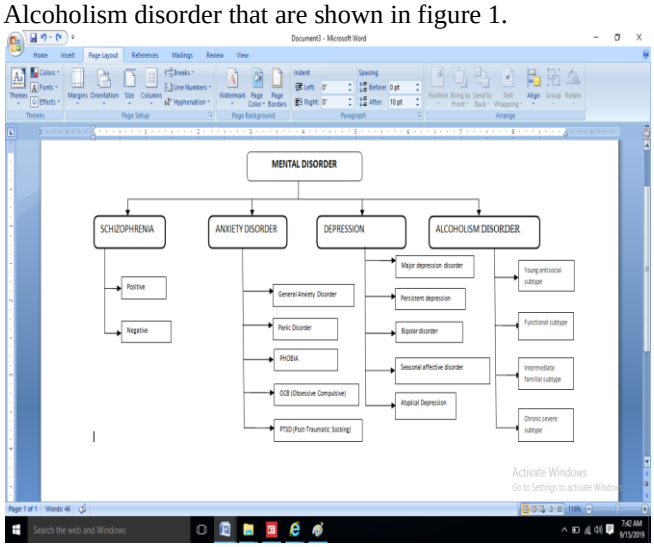

Fig 1. Taxonomy for Mental Disorder

Schizophrenia. Schizophrenia is a type of mental disorder which is also called as chronic disease where it affects on thinking of person, feels and behavior. It seems people have lost their communication with outside world whoever affected with schizophrenia [2]. Mainly it affects to age > 16 to 30. Positive and negative are the two main terms used to characterize the schizophrenia. Positive Schizophrenia tells what is real and presence of symptoms. Negative Schizophrenia which normally refers to the abnormal functionality in the brain like missing of behavior, thinking and perception.

Anxiety disorder. The underlying cause of anxiety disorder involves worry and fear. Some people suffering from this kind of disorder especially have fear in daily activities such as job performance, school work, and relationships. The anxiety disorder is categorized into following types like generalized anxiety disorder, panic disorder, and various phobia-related disorders. General anxiety disorder is the display of excessive anxiety or worry, mostly they feel restless and muscular problem. Its effect continue to remain until six months. Panic Disorder is the fear intensity becomes suddenly high and within a minute it reaches peak like sweating, shaking. A phobia-related disorder is the excessive feeling of actual danger in a particular situation. This causes excessive worry about encountering, heights and flying.

PTSD (Post-traumatic socking stress disorder) develops among people due to the past experience of shocking, scary or dangerous event.

Depression. Depression is the very serious issues amongst women in the country also called as attention-deficit/hyperactivity disorder. Major Depression Disorder is also called as clinical depression having features like depression mood, lack of interest. Persistent depression disorder is the people with depressed mood at least two years. It is also called dysthymia. Bipolar disorder experience the low mood of a person tremendously and it satisfies the testimonial of all major depressions. Seasonal affected disorder is the depression comes generally in winter seasons where the sunlight is low. It occurs during the spring and summer where the patient gains more weight and more sleep. Atypical Depression do not follow the thought behind the actual presentations like excessive eating, strongly reactive mood.

Alcoholism Disorder. Alcoholism Disorder is a dangerous disease because of that 2.5 million death around the world per year. There are four subtypes of disorder i.e. Young antisocial, functional, intermediate familial and chronic severe subtype**.** The young antisocial alcoholism disorder effects to 26 years old on an average. In this disease the effect of drinking starts at 15 which tend to addiction of alcoholics by 18 earlier than other subtypes. Mostly the functional subtype alcoholism disorder affected to middle aged adults those are working peoples which tend to drink in every other day and consume five or more drinks on drinking days. The intermediate familial subtype alcoholism disorder typically states that people start drinking by the age of 17 and they continue to habitat of consuming alcohols by the age of 30. The chronic severe subtype alcoholism disorder is rare among people that mainly effects to men, due to which divorce rates are high and they become frequent users of other kinds of drugs.

IV. METHODOLOGIES ADOPTED

Every day the data that are collected from clinician and pharmacological becomes so much huge which require some statistical analysis to improve the prediction performance. In the same way the machine learning methodology are also used in the psychiatry field to improve the prediction performance. Deep learning [17] also chosen for this research because it overcomes the drawback of machine learning and handles high dimensional data to get accurate performance. There is a model called as Gaussian mixture model [11] is an unsupervised clustering algorithm that fit the data as a number of structures which determine the dimension in which the subgroup are found. The aim is identifying subgroups with in patient rather than classify. Gaussian mixture model with the support vector machine used to segregate the patients where it is explained about the relationship between PANSS (Positive and Negative Syndrome Scale). PANSS used to identify two statistically distinct subgroups of patients to investigate whether the two identified subgroup are related to treatment or not [1].

Classification of medical data and symptoms of the disease which helps in determining the mental disorder based on certain condition [18]. Different classification techniques have been used in order to classify healthcare data such as C4.5 (J48), Random forest (RF) and Random Tree (RT) algorithm. The proposed classification approaches C4.5 (J48) and Random forest (RF) are the efficient techniques which give better classification and disease symptom finding. [20] The machine learning algorithms are classified into different categories like neural network, clustering, regression algorithm, decision tree, bayesian, support vector machine (SVM) and so on [15,7]. [8] After implementing this advanced machine learning techniques, someone can achieve optimized result. There are techniques not limited to AODEsr, Multi Layer Perceptron (MLP), RBF Network, IB1, KStar, Multi-Class Classifier (MCC), FT, LAD Tree for accurate results.

There is a work for novel feature selection called as back propagation neural networks that propose a weight analysis based heuristics to choose only effective features of a disease. The weight analysis based method is called ANNIGMA, where the weights are associated with feature. It is basically used in training stage to remove the noisy data. Deep learning along with machine learning algorithms simulate human learning process where features form different resource are identified, understood and memorized iteratively.

The ability to extract the meaningful information from a large data set is the classification technique. This paper [1] discusses a review of classification techniques and analysis on the clinical dataset [19]. However, Bipolar disorder is one type of psychiatry disease where aim is to identify three subgroup of patient according to their age at onset and based on different phenotypical features. To understand the disease and the symptoms of the disease on medical data according to its condition, an approach for predicting the risk of disease by using polymorphism interaction analysis to explore SNP interactions and colon cancer risk. Gini index and percentage wrong are the two unique scoring functions used to obtained the most likely to predict the risk of disease. The author Sumathi M.R discusses about mental health problem which basically comes at an early stage. Eight different machine learning mechanism are used here for analyzing and diagnosis the medical health data. The comparative result of classifier on various machine learning technique leads to the conclusion that the Multilayer Perceptron, Multiclass Classifier and LAD Tree produced more accurate results in full attribute set [8].

There is a proposed classification mining algorithm for brain tumor disease which provides a simple diagnosis rule. In order to generate the rules the author has used decision tree along with bagging algorithm. The author has claimed the work as a globally optimized hybrid feature selection algorithm derived from artificial neural network [16]. The GANNIGMA (Globally Optimized Artificial Neural Network Input Gain Measurement Approximation) is a hybrid feature selection in the proposed approach finds the significant features which help to generate a simplified rule. GANNIGMA Hybrid feature selection [16, 13] consists of three independent approach called [1]Filter approach,

[2]Wrapper approach and [3]Hybrid approach. [1]filter approach is cost effective but not so accurate. [2]Wrapper approach is not cost effective but accurate. Therefore a [3]Hybrid approach is involved in this feature selection process by combining filter approach with wrapper approach. The proposed hybrid feature selection with ensemble classification technique in this paper [16] combines a maximum relevance and minimum redundancy filter heuristic with globally optimized wrapper heuristic GANNIGMA. Hence the

results indicate that GANNIGMA + MRMR + Bagging + Decision Tree feature selection based proposed ensemble approach provides a more simplified decision rule with higher accuracies.

V. ACCURACY AND EFFICIENCY

In order to improve the accuracy and efficiency of the classification process the filter feature ranking score combined with the wrapper is proved to be good [10]. In the first approach, wrapper process is used for searching better feature for better performance of the model. In this work, the method hybridized a MI (Mutual Information) and MR (Maximum Relevance) with artificial Neural Network (ANN). The unique contribution of this approach is that it amalgamates the potential of wrapper approach to get better features [11]. In order to obtain the most likely interaction that best predict risk of disease uses two unique scoring functions called as Gini index and the percentage wrong. To assess the impact of personalized medicine we addressed three main areas: drugs, genes and consequences that define which genes are linked to responses to current drug treatments for heart stroke [4].

The diagnosis is done with the mental retired children who generally come to the clinician [8]. Initially they have collected 25 attributes data where feature selection algorithm is implemented and then 13 attributes are taken into consideration. Few attributes are relevant to classification only those attributes are considered for prediction of problem and the accuracy level is calculated [20]. The performance of the algorithms can be estimated by using the accuracy and error rate of the classification. The accuracy of the classification can be improved by using weights. In this paper [20] the comparison is done between GWDLM, FDDL (Fisher Discriminant Dictionary Learning), DFDL (discriminant feature oriented dictionary learning) and SVM where it found that GWDLM algorithm have better sensitivity compare to other algorithms. It helps in classifying the patient and having high specificity, it identify the actual disease.

The methodology used in this work [21] is Genetic Algorithm with logistic regression for Alzheimer's disease with high accuracy and built a predictive model. The architecture of the system is the combination of Genetic Algorithm and logistic regression which basically do the searching. The searching is done by Genetic algorithm for selected feature and the output generated by Genetic algorithms will be input for logistic regression. The logistic regression processes those data and produces the result with different variables with best feature set.

VI. CONCLUSION

The study is based on the taxonomy of mental illness disease and their symptoms that are discussed along with the precision medicine and deep learning methodologies adopted by others. The discussion of mental illness disease using precision medicine approach is mostly suggested by the doctors is the need of the hour. It is observed that the deep neural network based classifier with suitable feature selection algorithm has outperformed all the other existing related research in precision medicine in terms of having better accuracy, precision, recall etc. This paper also clarifies the classification algorithms necessitate the ability to extract useful information from large set of data. Our future work will attempt to develop some novel hybrid classifiers with efficient feature selection algorithms that may need to be trained prior to the implementation for real prediction to improve the accuracy and efficiency.

REFERENCES

1. Anthony, T. F., Musliner, K. L., Peter, P. Z., Scott, L. Z.,: A precision medicine approach for psychiatric disease based on repeated symptom scores. In: Journal of Psychiatric Research 95, Published by Elsevier Ltd., pp.147-155. (2013)
2. Katin, A. Mc.L., K.A., Kevin, K.,: Developmental Trajectories of Anxiety and Depression in Early Adolescence. In: J Abnorm Child Psychol. Author manuscript, pp. 311-323. NIH Public Access, (2015)
3. Musliner, K. L., Munk-Olsen, T., Eaton, W. W., E., Zandi, P. Peter. : Heterogeneity in long-term trajectories of depressive symptoms: patterns, predictors and outcomes. In: J. Affect Disord., pp. 199-211. (2016)
4. KenRedekop, W., DeirdreMladsi., Rotterdam., In: The Faces of Personalized Medicine: A Framework for Understanding Its Meaning and Scope: VALUE IN HEALTH 16, pp. S4 – S9. (2013)
5. Fernald, G.H., Capriotti, E., Daneshjou, R., Konrad, J. K., and Altman, R. B., : Bioinformatics challenges for personalized medicine. In: Bioinformatics, vol. 27 no. 13, pp. 1741–1748. (2011)
6. U.S. Food And Drug Administration (2013). Paving the Way for Personalized Medicine: FDA's Role in a New Era of Medical Product Development, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES U.S. FOOD AND DRUG ADMINISTRATION OCTOBER (2013)
7. Qiuling, S., Fenglong, Ma., Ye Yuan., Mengdi, H., Weida, Z., Jing Gao., Aidong, Z., Fellow.: Deep Patient Similarity Learning for Personalized Healthcare. In: IEEE TRANSACTIONS ON NANOBIOSCIENCE, pp. 219-227. IEEE computer society, (2018)
8. Sumathi, M. R., Poorna, B., : Prediction of Mental Health Problems Among Children Using Machine Learning Techniques, In: International Journal of Advanced Computer Science and Applications, Vol. 7, pp.552 – 557. No. 1, (2016)
9. Hood, L., Stephen, H. F., : Predictive, personalized, preventive, participatory (P4) cancer medicine, In: Nature Reviews Clinical Oncology., vol. 8, pp. 184–187. (2011)
10. Azorin., Bellivier, J-M., Frank, K., Adida, A., Belzeaux, M., Fakra, R., Hantouche, E., Lancrenon, E., Sylvie and Golmard JL., Jean-Lousi.: Characteristics and profiles of bipolar I patients according to age-at-onset: findings from an admixture analysis In: Journal of Affective Disorders 150, pp. 993–1000. (2013)
11. Md Shamsul Huda., John Yearwood., Andrew Stranieri.: fv-filter approaches for input feature selection using Maximum relevance-Minimum redundancy and Artificial Neural Network Input Gain Measurement Approximation (ANNIGMA). In: Fourth International Conference on Network and System Security, pp.1-10. Australia (2010)
12. Hall, M. A., Smith, L. A., Feature Subset Selection: A Correlation Based Filter Approach.: In: International Conference on Neural Information Processing and Intelligent Information System, pp. 855-858. Springer. (1997).

13. Subhas C., Misra., Sandip Bisui.: Modelling vital success factors in adopting personalized medicine system in healthcare technology and management. In: 21st Engineering Science and Technology, an International Journal, pp. 532–545. ScienceDirect (2018)
14. Po-Yen, W., Chih- Wen, C., Chanchala, D. Kaddi., Venugopalan, J., Hoffman, R., and May, D. W.: Omic and Electronic Health Record Big Data Analytics for Precision Medicine, In: IEEE Transactions On Biomedical Engineering, Vol. 64, HHS Public Access, NO. 2, pp. 263-273. (2016)
15. Huda, S., John, Y., Herbert, F. J., Mohammad, M. H., Giancarlo, F., and Michael, B.: A Hybrid Feature Selection With Ensemble Classification for Imbalanced Healthcare Data: A Case Study for Brain Tumor Diagnosis. In: Special Section On Healthcare Big Data, IEEE. Translations, Vol. 4, 2016, PP. 9145-9154.(2017)
16. Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., and Yang, GZ.: Deep Learning for Health Informatics. In: IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, Vol. 21, NO. 1, PP. 4-21.(2017).
17. Shivangi, J., Dr. Mohit, G.: A Data Mining Analysis Over Psychiatric Database for Mental healthClassification. In: International Journal on Future Revolution in Computer Science & Communication Engineering ISSN: 2454-4248, Vol. 4, Issue: 1, pp. 241 – 246.(2018)
18. Sanzo, Di. M., Cipolloni, L., Borro, M., Russa, La, R., Santurro, A., Scopetti, M., Simmaco, M., Frati, P., Clinical applications of personalized medicine: a new paradigm and challenge, pp. 194-203. (2017)
19. CHUNXUE, WU., CHONG LUO., NAIXUE, X., ZHANG, W., AND KIM, T-H.: A Greedy DeepLearningMethod for Medical Disease Analysis. In: IEEE Translations and content mining, Vol. 6, pp. 20021-20030. (2018)
20. Johnson, GA. Piers., Luke, Vandewater., William, W., Paul, Maruff., Greg, Savage., Petra, Graham., Lance, S. M., Kathryn, A Ellis., Cassandra, Szoeke., Ralph, N. M., Christopher, C. R., Colin, L. M., David, A., Ping, Z.: Genetic algorithm with logistic regression for prediction of progression to Alzheimer's disease. In: Thirteenth International Conference on Bioinformatics (InCoB2014): Bioinformatics, Johnson et al. BMC Bioinformatics, 15(Suppl 16):S11, Vol. 15. pp. 2-14. (2014)
21. Giampaolo, P., Charles, B. N.: Personalized Medicine in Psychiatry: Back to the Future. In: personalized medicine in psychiatry, 2017, March–April, Elsevier, 2017, Vol. 1-2, DOI: https://doi.org/10.1016/j.pmip.2017.01.001. (2017).

AUTHORS PROFILE

Sumitra Mallick, is currently working as Assistant Professor in CSE in KLU. She is actively engaged in research in the area of machine learning in healthcare sector. Persuing her Ph.D at Utkal University, She has a teaching experience of about 13 years.

Dr. Mrutyunjaya Panda, is currently working as Reader in the Department of computer science and applications at Utkal University, Odisha, India. He has published about 70 papers in International and national journals and conferences. He has also published 7 book chapters to his credit. He has 2 text books and 3 edited books to his credit. He is a program committee member of various international conferences. His active area of research includes Data Mining, Image processing, Intrusion detection and prevention. Social networking, Mobile Communication, wireless sensor networks, Natural language processing, Internet of Things, Text Mining etc.