International Journal of Research in Advent Technology, Vol.6, No.7, July 2018
E-ISSN: 2321-9637
Available online at www.ijrat.org

A Survey on Data Mining Algorithm for Brain Disorder

Gayathri S[1], Dr. M. Punithavalli[2]

[1]Ph.D Research Scholar Department of Computer Applications, Bharathiar University
2Associate Professor, Department of Computer Applications, Bharathiar University, Tamil Nadu, India
Email: selvamgayathri06@gmail.com[1], mpunithasrcw@yahoo.co.in2

Abstract - Nowadays, most of the people addicted to alcohol resulting in they have an Alcohol Use Disorder (AUD). Alcoholism can increase the risk of specific cancers and also damage to the brain, liver and the other organs. It can easily affect the brain cells namely neurons and cause Wernicke-Korsakoff syndrome which is known as brain damage characterized by severe amnesia, confabulation and dementia. In addition, hepatic encephalopathy is also a brain damage due to liver failure which may cause by alcoholism, cirrhosis, etc. Such severe injuries are predicted and diagnosed by using data mining techniques. Data mining is the most dominant method to mine useful patterns or data from any image or textual data. In recent years, data mining is the most essential in the medical industry since it has significant utility in the healthcare domain in the real world. The most popular data mining techniques to understand the different features of the health dataset are clustering and classification algorithms. It offers a support for identifying a reliable relationship between the patients profiling or psychoanalysis and outcome. In this paper, data mining techniques that have been developed for predicting brain injuries are studied and analyzed based on their merits and demerits. The main objective of this paper is to review and analyze the data mining algorithms developed for brain injury detection in a medical field in order to improve the diagnosis to the patients. Moreover, different performance metrics are also compared to those algorithms to understand their efficiency in brain damage detection.

Keywords— Alcoholism, Alcohol use disorder, Wernicke-Korsakoff syndrome, Brain injury, Data mining algorithms, TBI, GLCM, MRI.

1. INTRODUCTION

Generally, alcohol is considered to be a depressant. It has several effects on the brain's neural activities such as increasing dopamine discharge and stimulating endorphin production due to its major ingredient named ethanol. Such modified activities cause normal functions of brain neurons. Drinking characters can be categorized as recreational, heavy and addictive drinking. Normally, recreational drinking involves fitness risks and has no consequences for ourselves or others. Heavy drinking has a clinically essential impairment or distress. It is considered to be a serious health issue that can have major consequences in both personal and professional life. Addictive drinking is a maladaptive pattern of drinking in which the person suffering from Alcoholism due to issues at work or personal life. Alcohol addiction or alcoholism refers to a disease that affects the human [1]. In other words, it is a chronic disease characterized by an inability to control drinking or absorption level of alcohol. The symptoms of Alcoholism are increased quantity or frequent use, High tolerance for alcohol or lack of hangover symptoms, Drinking at unsuitable times or places , Hiding alcohol or hiding when drinking Increased weariness, hopelessness or the other emotional problems and Wanting to be where alcohol is present and avoiding circumstances where there is none. Physical symptoms may happen when the person is unable to drink [2]. This is called withdrawal. Such symptoms are signs of an alcohol addiction. The body feels it's unable to act and function since it must without the alcohol. Withdrawal symptoms include Nausea, Sweating, Shaking, Hallucinations, Convulsions and Vomiting. Alcoholism can cause several short-term and long-term effects. Short-term effects of alcoholism may risky as long-term effects.

1.1. Short-term effect of Alcohol

Many short-term effects of alcoholism may create the following problems

- Slow reaction time
- Poor reflexes
- Reduce brain activity
- Lowered inhibitions
- Blurry vision
- Breathing problem
- Impatience

1.2. Long-term effect of Alcohol

- Brain defects including Wernicke-Korsakoff syndrome
- Liver syndrome

- Diabetes complications
- Heart problems
- Increased risk of cancer
- Vision problems
- Bone loss
- Birth defects during pregnancies
- Risk of death from drunk driving
- Increased risk of homicide and suicide

The alcohol addiction are confirmed by a doctor based on the person's activities and health issues or using questionnaires to evaluate a dependence on alcohol including behavior such as slurred speech or shivering when it is consumed. There are reliable tests available for the actual use of alcohol whereas the general test is Blood Alcohol Content (BAC). BAC is most helpful to verify the alcohol tolerance.

1.3 The other tests for recognizing the Alcoholism:

- Macrocytosis
- Elevated GGT (Gamma-Glutamyl Transferase)
- Moderate elevation of AST (Aspartate Transaminase) and ALT (Alanine Transaminase). The ratio of AST and ALT is 2:1.
- High CDT (Carbohydrate Deficient Transferrin)

In contrast, an individual who drinks heavily over a long period of time may have brain deficits that persevere well after he/she achieves abstemiousness. Several factors impact how and to what extent alcohol affects the brain including:

- How much and how often a person drinks
- The age at which he/she first began drinking and how long he/she has been drinking
- The person's age, qualification level, gender, genetic background and history of family alcoholism
- Whether he/she is at risk as a result of prenatal alcohol exposure
- His/her general health information

The negative side effects of alcohol i.e., severe injuries in the human brain are detected efficiently by using data mining algorithms in order to provide a better diagnosis. Data mining can address the problem of how best to utilize this data for discovering new knowledge and improving the decision-making process. Nowadays, data mining is mostly used in medical science for the accurate diagnosis of the patient disease by considering the most significant features from a huge amount of data [3]. Such features can be extracted through several learning processes such as classification and regression, clustering, association rule learning and feature selection. Data mining tools are used for controlling human restrictions such as subjectivity or errors due to the weariness and providing equipped suggestions for the decision processes [4]. Several methodologies and implementation are available for each data mining process. However, the critical issue of using data mining process is the reliability analysis of an extracted knowledge of data. Such concern is more significant in the medical field while predicting criteria for new patients. Also, data mining technologies may suffer from circularity and require separate datasets for validation. Therefore, this paper studies different methods used for detecting brain injuries in order to identify the best accuracy model. Moreover, the advantages and limitations of those methods are also investigated to further improve the prediction of brain injuries precisely. The rest of the paper is organized as follows: Section 2 describes the different data mining techniques proposed for predicting the brain injuries. Section 3 presents the performance comparison of those techniques and Section 4 concludes the survey, thus summarizes the entire discussion on brain injury prediction using data mining.

2 LITERATURE SURVEY

A hybrid method [5] was proposed for Magnetic Resonance Imaging (MRI) image classification. In this method, Neural Network (NN) based method was proposed for classifying a given MR brain image as normal or abnormal. Initially, wavelet transform was employed for extracting the features from images and then Principle Component Analysis (PCA) was applied for reducing the dimensions of features. Then, the reduced features were transmitted to Back Propagation (BP) NN with which Scaled Conjugate Gradient (SCG) was adopted for finding the optimal weights of the NN. This method was applied to 66 images and the classification accuracies of training and testing were computed.
A new approach [6] was proposed for automated diagnosis according to the MRI brain images classification. This approach has two processes such as feature extraction and classification. Initially, the features from MRI images were obtained by using Discrete Wavelet Transformation (DWT). The extracted features were reduced by using PCA for obtaining more significant features. Then, Feed Forward back propagation Artificial Neural Network (FP-ANN) and KNN classifiers were applied for the classification process.
MRI mammogram image classification [7] was proposed by using an ID3 algorithm. The objective was to segment the mammogram images automatically and classify them as benign, malignant or normal according to the ID3 algorithm.

In this system. A hybrid method of data mining algorithm was used for predicting the texture features which are used in the classification process. The size and stages of a tumor were detected by using the ellipsoid volume formula which is computed over the segmented region. This algorithm provides a high level of accuracy within less amount of time period. Automatic 3D segmentation of human brain images [8] was proposed in which the brain scans were processed in
2D and 3D. It consists of different processes such as preprocessing, segmentation, etc., and employed in 3D image processing extension for RapidMiner platform.
An excessive alcohol effect [9] was analyzed which causes the brain damage. In this analysis, a significant group of teenagers was identified along AUD without other psychiatric disorders. In addition, brain morphology in teenagers was analyzed and compared to age and gender-matched fit controls. MRI data were examined by using FSL's FIRST software for subcortical volumes and also cortical Gray Matter (GM) was analyzed by using Voxel-Based Morphometry (VBM) and Regions-Of-Interest (ROI) analysis. Moreover, a large amount of decreased GM density in AUD was compared to the control located in the left lateral frontal, temporal and parietal lobes.
Brain Tumor prediction and classification [10] was proposed by using a decision tree. The aim was to focus on the image mining for classifying the brain tumor in the brain MRI images. This system has four phases for classifying the medical images with proper diagnosis. Pre-processing was achieved by using the median filtering process and features were extracted by using texture feature extraction method. Then, the extracted features from the CT scan images were used for mining the association rules. Moreover, a decision tree was used for the classification process.
A hybrid approach [11] was proposed for MRI brain image classification by combining Weighted Firefly (WFF) and K-means algorithm called WFF-K-means and Modified Cuckoo Search (MCS) and K-means algorithm called MCS-K-means. In this approach, a novel Computer-Aided Diagnosis (CAD) technique was proposed for the classification of the MR brain images. Initially, different features such as color, texture and shape features were derived from the segmented image. Such features were given to the Multi-Class Support Vector Machine (MC-SVM) classifier with a hybrid feature selection algorithm which is trained with data labeled by experts and the brain images were detected at high accuracies. Performance analysis of data mining algorithms [12] was presented for classification of the medical image. The objective of this study was analyzing the performance of image mining algorithms according to the classification accuracy, processing time, error rate, sensitivity and specificity. In medical image analysis, image classification was the most important technique which has two level processes. Initially, the typical model was constructed based on the determined collection of concept or data classes. Then, the model was used for the classification process and better method of image mining was identified according to the performance analysis. Different algorithms used in this study were Regression Tree (CART), K-means, Naive Bayes (NB), Decision Tree (DT), K-Nearest Neighbor (KNN) and Support Vector Machine (SVM).

Brain image segmentation [13] was proposed by using semi-supervised clustering technique. The aim of this technique was partitioning the brain images into different non-overlapping homogeneous regions by using the image pixel intensity values as features. The objective functions were three cluster validity indices which are concurrently optimized by using AMOSA i.e., a modern Multiobjective Optimization technique according to the Simulated Annealing concepts. The initial two cluster validity indices were symmetry distance based Sym-index and Euclidean distance based I-index according to the unsupervised properties. The last cluster validity index was a supervised data based cluster validity index namely Minkowski index.
Traumatic Brain Injury (TBI) [14] was predicted and diagnosed which is associated with Substance-Related Disorder (SRD). In this analysis, the relationship between TBI and SRD was estimated with Cox proportional hazard regression models. The data were collected from the Taiwan National Health Insurance Research Database for several patients with or without TBI during 2000–2010.Identification and classification of brain tumor MRI images [15] were proposed by using feature extraction based on the DWT and probabilistic neural network. Initially, MRI images were segmented into different regions by using region growing technique. Here, morphological filtering was used for removing the noise from the images. Then, the Gray-Level Co-Occurrence Matrix (GLCM) was applied for extracting the statistical features and DWT method was applied for extracting the wavelet coefficients. After that, a probabilistic neural network classifier was introduced for training and testing the detection accuracy of tumor locality in brain MRI images.

3 DISCUSSIONS

This section illustrates an overview of merits and demerits of different MRI brain image pre-processing, segmentation and classification techniques whose functional scenarios are discussed in the above section. Through the literature survey on pre-processing, segmentation and classification techniques, the following limitations are observed. Most of the techniques have high computation time.

KNN algorithm has the limitation such as the determination of values of parameter K i.e., nearest neighbor values. An ID3 algorithm has the drawbacks such as only one attribute was tested at a time to make a decision. Based on data mining techniques are observed and an ideal solution is identified to overcome those issues in brain image classification.

Table.1 COMPARISON OF PERFORMANCE EFFECTIVENESS OF DIFFERENT MRI BRAIN IMAGE SEGMENTATION AND CLASSIFICATION ALGORITHMS

Ref. No.	Methods Used	Merits	Demerits	Performance Metrics
[5]	BPNN with SCG, PCA	High classification accuracy.	Computation time for feature extraction was high.	Classification accuracy=100%, **Mean Computation time:** Feature extraction=0.023s, Feature reduction=0.0187s, NN classification =0.0035sec
[6]	DWT, PCA, FP-ANN, KNN	Better efficiency, inexpensive and robust.	Computation cost was moderately high and also require for determining the value of parameter K i.e., number of nearest neighbors.	**DWT-FP-ANN based classifier:** Classification accuracy=90%, Sensitivity=94%, Specificity=90% **DWT-KNN based classifier:** Classification accuracy=99%, Sensitivity=100%, Specificity=90%
[7]	ID3 algorithm based classification	High level of accuracy.	Simultaneous classification of image leads to high computational cost and only one attribute was tested at a time to make a decision.	Classification accuracy=99.9%,
[8]	Automatic 3D segmentation	Better accuracy.	Requires more efficient classification algorithms.	**Training Dataset:** Sensitivity=99.13% Specificity=96.15% Accuracy=98.38% **Testing Dataset:** Sensitivity=92.5%, Specificity=100%, Accuracy=95.08%
[10]	Decision tree based brain tumor prediction and classification	Better accuracy and sensitivity.	Computation time was not evaluated.	Precision=100%, Sensitivity=93%, Specificity=100%, Accuracy=96%
[11]	CAD technique, IGSFFS, MC-SVM, WFF-K-means,	Higher performance.	Computationally expensive.	WFF-K-means: Classification accuracy=98.93%, Sensitivity=97.87% Specificity=99.52% MCS-K-means: Classification accuracy=99.31%,

	MCS-K-means			Sensitivity=98.23% Specificity=99.73%
[12]	CART, NB, DT, KNN, SVM based medical image classification	Better classification accuracy.	DT has less efficiency since its credibility and sensitivity were not accurate.	Classification accuracy: SVM=87%, CART=91%, K-means=96%, NB=90%, DT=59%
[13]	Brain image segmentation using semi-supervised clustering, AMOSA	Automatic segmentation of the MR brain images.	Requires further improvement on classification accuracy since it uses only intensity values as the features.	Nil.
[14]	TBI prediction	High sensitivity.	Some important data were not presented such as family history, marital status, etc.	Nil
[15]	DWT, Probabilistic neural network	High accuracy.	Requires more efficient segmentation and feature extraction techniques.	PSNR=14.011dB, MSE=6.121, Accuracy=95% (Testing dataset)

4 CONCLUSION

In this article, a detailed study on MRI brain image classification using data mining techniques was presented. It is obvious all researchers have tried in different techniques for preprocessing, extracting the features, segmentation and classification of MRI brain images to achieve better results than the other classification techniques. According to this analysis, some significant results for the detection of a brain tumor or injuries due to different factors such as alcohol, etc., are addressed. In future work, the classification of MRI brain images will be required to improve the accuracy and reduce the computation time efficiently. Based on this classification, the brain damage due to alcoholism will be reduced significantly.

REFERENCES

[1] Enoch, M. A; & Goldman, D. (2002):Problem drinking and alcoholism: diagnosis and treatment. American family physician, 65(3), 441-454.

[2] Smith, C. M; & Barnes, G. M. (1983). Signs and symptoms of hangover: prevalence and relationship to alcohol use in a general adult population. Drug and alcohol dependence, 11(3-4), 249-269.

[3] Mathew, S. T;Nachamai, M. (2015). Clustering of brain MRI image using data mining algorithm. International Journal of Advanced Computational Engineering and Networking, 3(4), 37-40.

[4] Megalooikonomou, V; Ford, J;Shen, L; Makedon, F;Saykin, A. (2000). Data mining in brain imaging. Statistical Methods in Medical Research, 9(4), 359-394.

[5] Zhang, Y; Dong, Z;Wu, L; Wang, S. (2011). A hybrid method for MRI brain image classification. Expert Systems with Applications, 38(8), 10049-10053.

[6] Rajini, N. H; Bhavani, R.(2011, June). Classification of MRI brain images using k-nearest neighbor and artificial neural network. In Recent Trends in Information Technology

(ICRTIT), 2011 International Conference on (pp. 563-568). IEEE.

[7] Angayarkanni, A. P; Kamal, B. D. N. B. (2012). MRI mammogram image classification using ID3 algorithm. In 2012 IET Conference on Image Processing, pp. 1-5.

[8] Uher, V; Burget, R (2012, July). Automatic 3D segmentation of human brain images using data-mining techniques. In Telecommunications and Signal Processing (TSP), 2012 35th International Conference on (pp. 578-580). IEEE.

[9] Fein G;Greenstein D;Cardenas, V. A; Cuzen, N. L; Fouche, J. P;Ferrett, H;Stein, D. J. (2013). Cortical and subcortical volumes in adolescents with alcohol dependence but without substance or psychiatric comorbidities. PsychiatryResearch:Neuroimaging, 214(1), 1-8.

[10] Naik, J; Patel, S. (2014). Tumor detection and classification using decision tree in brain MRI. International Journal of Computer Science and Network Security (IJCSNS), 14(6), 87

[11] Rajalakshmi N; Lakshmi Prabha V. (2015). MRI brain image classification—a hybrid approach. International Journal of Imaging Systems and Technology, 25(3), 226-244.

[12] Keerthana, P; Thamilselvan, P; Sathiaseelan, J. G. R. (2016). Performance Analysis of Data Mining Algorithms for Medical Image Classification. International Journal of Computer Science and Mobile Computing, 5(3), 604-609.

[13] Saha S;Alok, A. K; Ekbal, A. (2016). Brain image segmentation using semi-supervised clustering. Expert Systems with Applications, 52, 50-63.

[14] Wu, C. H; Tsai, T. H; Su, Y. F; Zhang, Z. H; Liu, W; Wu, M. K; Lin, C. L. (2016). Traumatic Brain Injury and Substance Related Disorder: A 10-Year Nationwide Cohort Study in Taiwan. Neural plasticity, 2016.

[15] Shree, N. V; Kumar, T. N. R. (2018). Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. Brain informatics, 5(1), 23-30.