

An Epidemic of Neurodegenerative Disease Analysis Using Machine Learning Techniques



M. Menagadevi, V. Vivekitha, D. Thiyagarajan, and G. Dhivyasri

Introduction

Neurodegenerative diseases are a collection of neurological disorders in which neurons in the central nervous system die or are harmed, causing significant disabilities and, in the worst-case scenario, death. Figure 1 shows the major neurological disorder, which affects all age groups. Typically, they are stumbled upon in elderly age. However, the development of the disease could occur sooner. Their prevalence has risen dramatically in recent years, and this trend is projected to continue. As the world's population ages, the trend will continue. Neurodegenerative illnesses are difficult to manage and can be costly, since the cause is unknown and there is no recognized cure. Currently, treatments are focused on reducing symptoms.

The use of machine learning algorithms in medical and scientific research has gotten a lot of attention in recent years. New technologies have made it possible to rapidly accumulate patient data, such as ultrasonography and MRI readouts; omics profiles of biological samples; electronically collected clinical, behavioral, and activity data; and social media-derived information in the last decade. The number of characteristics (or variables) recorded in each observation can occasionally surpass the total number of observations in these large health datasets, making them high-dimensional. Machine learning advancements can be of great assistance to diagnosis and monitoring, such as illness onset detection, disease characterization,

M. Menagadevi (✉) · V. Vivekitha · G. Dhivyasri
Department of Biomedical Engineering, Dr. N.G.P Institute of Technology,
Coimbatore, TN, India

D. Thiyagarajan
Department of Artificial Intelligence and Machine Learning, School of Engineering, Malla
Reddy University, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

G. R. Kanagachidambaresan et al. (eds.), *System Design for Epidemics Using
Machine Learning and Deep Learning*, Signals and Communication
Technology, https://doi.org/10.1007/978-3-031-19752-9_17

285

Fig. 1 Neurological disorders

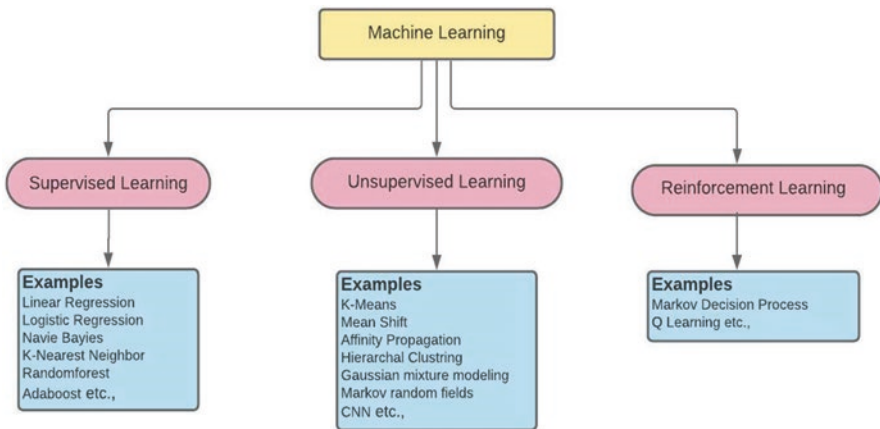
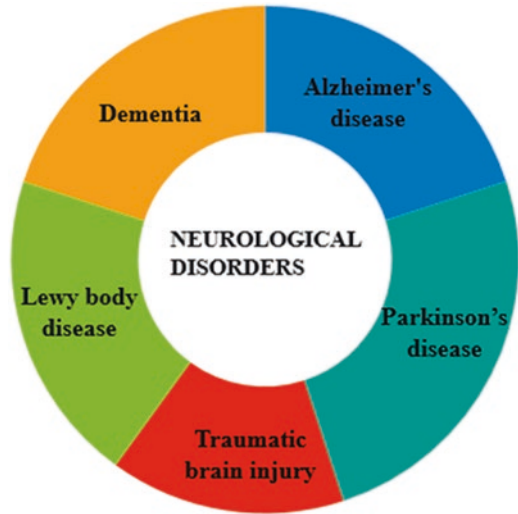


Fig. 2 Divisions of machine learning algorithm

and disease improvement quantification of differential diagnosis. ML approaches have been used to assist physicians with CAD to diagnose the diseases [1, 2].

Classes of machine learning algorithm:

- Supervised learning
- Unsupervised learning
- Reinforcement learning (Fig. 2)

Supervised Learning

These approaches are most typically used to analyze data linked to neurodegenerative diseases. Dataset is required to train the model. Often, these labels necessitate human curation or professional evaluation; for example, a radiologist is required to label a series of radiographs, the size of a brain area on an MRI scan, and the labeling of MRI scan pictures, and a neuropathologist is necessary. The computer may then use this model to predict the label for fresh, unlabeled datasets using the new input characteristics. It can be difficult to collect big enough volumes of correct labels for supervised machine learning. Figure 3 shows the supervised learning model.

Classification is a process that includes finding rules in which new objects are assigned to a predefined category, and it requires two steps. The first step is the building of classification with the help of training dataset that has class attributes. The second step is to analyze the performance of the classification using test dataset. The supervised process is in which prediction is iteratively performed using training data [24, 25]. In supervised learning, some of the commonly used algorithms are decision tree, linear classification, and Naive Bayes.

The dataset must first be preprocessed, which involves a task that transforms the raw data into a finished dataset. It can be fed into the data model. The data cleanup,

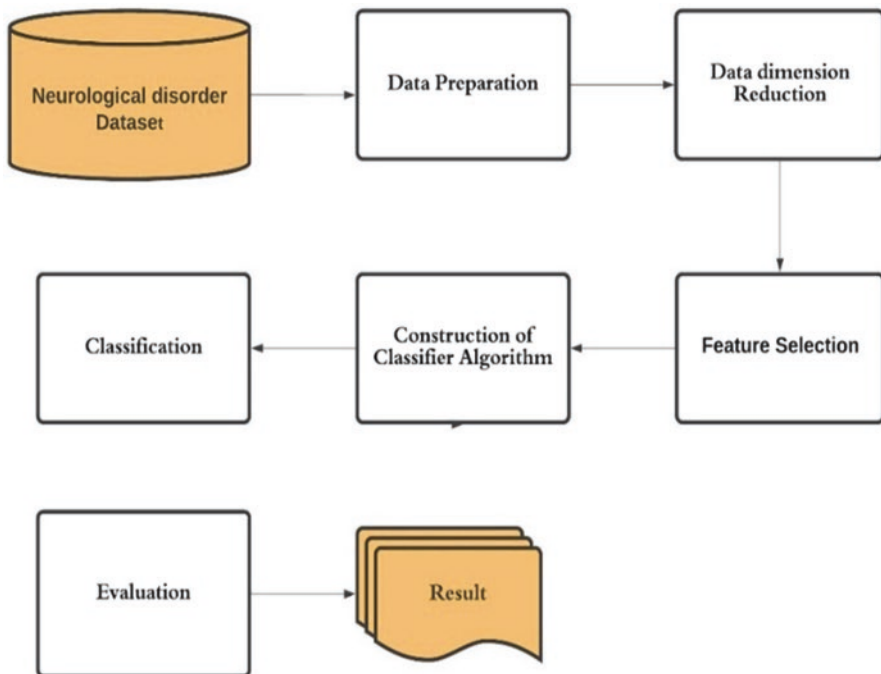


Fig. 3 General block of supervised learning model

data transformation, feature selection, and feature mapping stages in the preprocessing phase all contribute to the classifier’s accuracy. Data cleaning is a step that looks upon any missing value in the dataset, and data transformation ensures that the data values fall in a small range. The feature selection is one of the significant steps that limit finite features, which can provide a higher detection rate, and the final step is the feature mapping, which helps in finding a decision region. This can maximize the class separately between the given classes in the newly mapped space. In the training phase, it practically supplies at least 70% of the dataset to create a model that can categorize and predict the right label of the testing set. Using data evaluation is we can analyze the performance of the algorithm in the neurological disorder dataset. Data evaluation includes accuracy sensitivity specificity and error rate. Many literatures suggest that using a supervised algorithm for neurological disorder classification, it was found that it has low specificity, about 0–54%, and the accuracy of the decision was predicted to be nearly 80% only.

Decision Tree

The decision tree (DT) is one of the most well-known and widely used machine learning methods. It is a mimic of the human thinking ability to make a decision. It’s a tree-structured decision-making technique. The root, leaf nodes, branches, and internal nodes are the four basic components. The root of a tree connects its classes, with leaf nodes representing classes, branches representing outcomes, and internal leaves representing processes. The classification rules are the paths from the root to the leaves, and the steps involved in predicting the class of the given dataset are given below. Figure 4 shows the basic flow of the decision tree algorithm [31].

- Step 1: It starts from the root nodes of the tree which contains the complete dataset.
- Step 2: Select the attribute/feature value for the root node. Create a branch for each possible attribute/feature value.

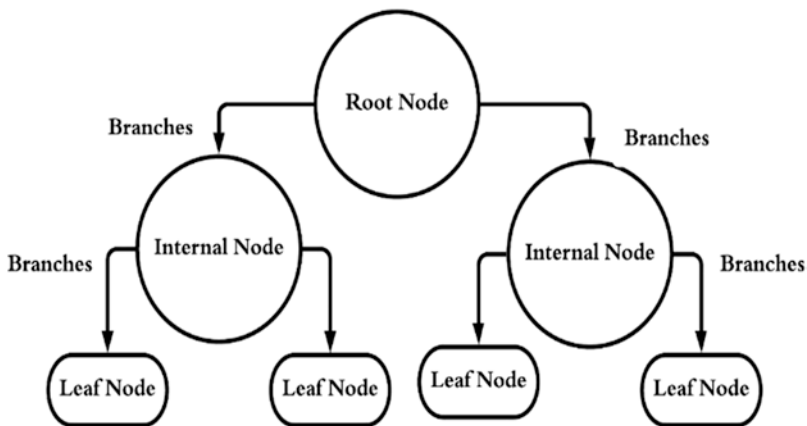


Fig. 4 Basic flow of decision tree algorithm

Step 3: Generate the decision tree which contains the best attribute.

Step 4: Repeat recursively for each branch, and make a new decision tree using the subset of the dataset created in step 2.

The process is continued until a stage where you cannot further classify the nodes and call them as final node/leaf.

Figure 5 explains the use of the decision tree algorithm for the image dataset for analyzing the brain image. Neurological disorder image datasets can be collected from various datasets (e.g., OASIS, TOF-MRA, Allen Brain Atlas, Alzheimer's Disease Neuroimaging Initiative (ADNI), the fMRI Data Center, etc.). The images were then divided into two groups: testing and training. After collecting the results, the features of the image were extracted using a suitable technique, and the testing was carried out using a decision tree algorithm (C4.5, ID3, CHAID, etc.). Figure 6 shows the output of lesion detection in the brain MRI image using decision tree algorithm.

Linear Regression

Linear regression was first developed in the discipline of statics to aid in the understanding of the relationship between numerical input and output. A linear model is a linear regression. It assumes that the input variable (x) and output variable (y) have

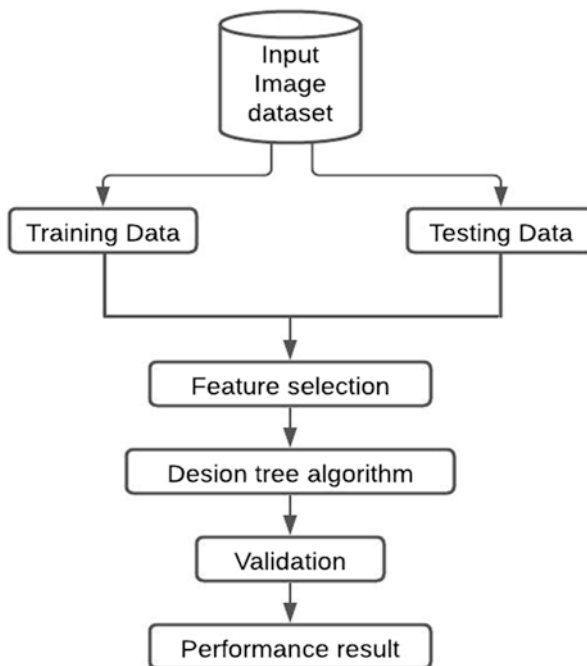


Fig. 5 Decision tree algorithm for image analysis

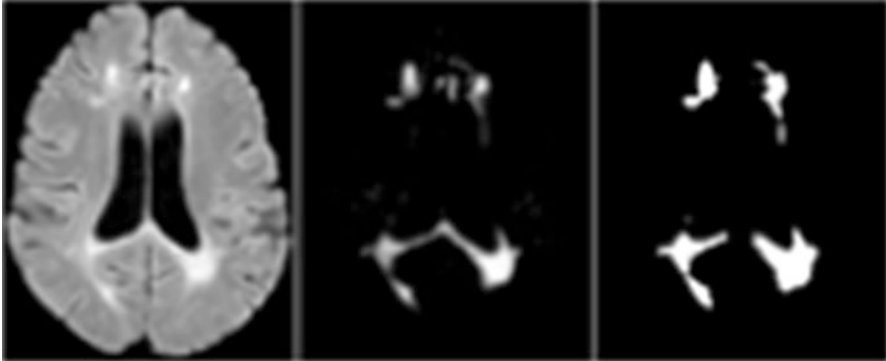


Fig. 6 Decision tree algorithm for lesion detection

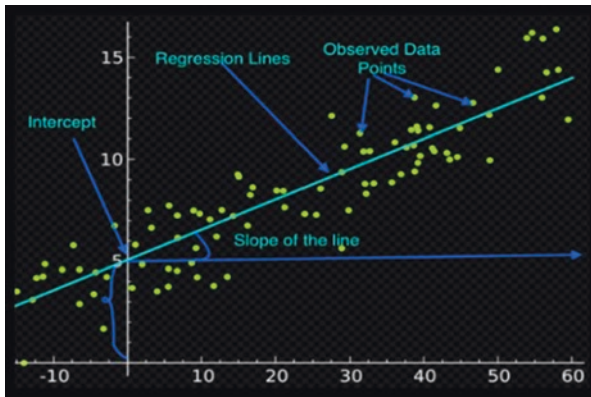


Fig. 7 Plot of linear regression

a linear relationship (y). Simple linear regression and multivariate linear regressions are two separate forms. Simple regression is an approach in which the response prediction is made using a simple feature. The linear regression algorithm’s purpose is to discover the optimal values for b_0 and b_1 , such that the best fit line can be found. Figure 7 shows the general plot of the linear regression model [32].

$$y = b_0 + b_1x_1 \tag{1}$$

When multiple features are used, it is called as multiple linear regression

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \tag{2}$$

By filtering a linear equation to the observed data, multiple linear regression aims to model the relationship between one or two features and response. Regression analysis formula

$$y = mx + b \tag{3}$$

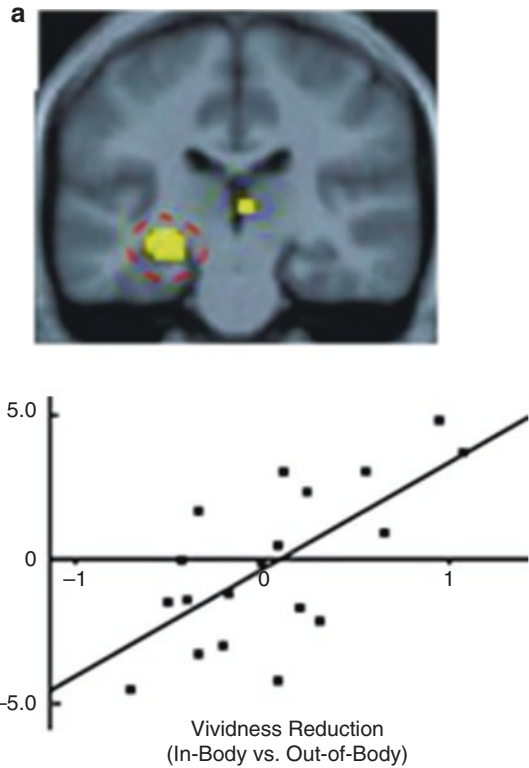
Regression lines are used to predict the values of “y” (dependent variable) for a given value of “x” (independent variable). The best-fit regression line times to reduce the sum of squared distance between observed (actual) and forecasted data points. The intercept of regression lines in estimating the value of “y” (dependent variable) when “x” (independent variable) has no influence.

Figure 8 [34] shows the output of the linear regression model for analyzing the left hippocampus of the functional MRI dataset; Figure 8b shows the linear model analysis of the BOLD signal in the image.

Logistic Regression

One of the supervised machine learning algorithms that may be used to model the probability of an event or class is the logistic regression model. It is used when the data is separable linearly and the outcome is binary. In binary classification, there are primarily two types: one is logistic regression, which is a single independent

Fig. 8 (a) Linear regression analysis of the fMRI data (b) Linear analysis of BOLD signal



variable to predict the output, and the other one is multiple logistic regression, where multiple independent variables are used to predict the output [33]. Logistic regression models the data using the sigmoid function for single logistic regression which is given as

$$e = 1 / (1 + e^{-(b_0 + b_1x)}) \tag{4}$$

Logistic regression models the data using the sigmoid function for multiple logistic regression which is given as

$$e = 1 / (1 + e^{-(b_0 + b_1x + b_2x^2 + \dots + b_nx^n)}) \tag{5}$$

The probability of occurrence of a neurological disorder can be obtained by using this algorithm with high accuracy. The input image is preprocessed, the feature is extracted using a suitable algorithm, and the image is divided into training and testing data. Then, the logistic regression predicts the disorder based on the obtained features. Figure 9 shows the flow diagram of logistic regression [34].

Naive Bayes

The Naive Bayes algorithm is one of the simple and most effective supervised clas-

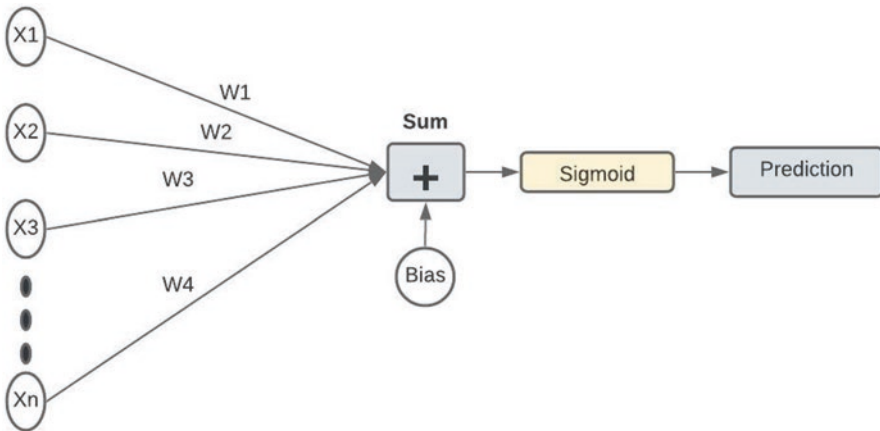


Fig. 9 Flow diagram of logistic regression

sification algorithms. It helps in building fast machine learning that can predict rapidly. It is an object probability-based probabilistic classifier, where naïve assumes

the occurrence of a certain feature independent of the occurrence of the other feature and Bayes, as it depends on the Bayes principle: $P(y|X) = \frac{P(X|y)P(y)}{P(X)}$, where y is a class variable and x is a dependent feature vector (of size n), which is given as $X = \{x_1, x_2, x_3, \dots, x_n\}$. Apply Bayes' theorem to the given dataset in the following way: Now, if any two events A and B are independent, then $P(A, B) = P(A)P(B)$. Then; the final result will be given as

$$P(y|x_1, \dots, x_n) = \frac{P(x_1|y)P(x_2|y)\dots P(x_n|y)}{P(x_1)P(x_2)\dots P(x_n)} \quad (6)$$

It can also be rewritten as

$$P(y|x_1, \dots, x_n) = \frac{P(y) : \prod (I = i \dots n) P(x_i|y)}{P(x_1)P(x_2)\dots P(x_n)} \quad (7)$$

We can now eliminate that term because the denominator remains constant for every given input:

$$(y|x_1, \dots, x_n) \propto P(y) : \prod (i = 1 \dots n) P(x_i|y) \quad (8)$$

It was found the best classification results were obtained by mostly using a Naive Bayes classifier for analyzing MRI images from Parkinson's and Alzheimer disease [5].

Support Vector Machine

SVM is abbreviated as support vector machine. It is the most widely used supervised learning method for classification and regression; however, it is more commonly used for classification. The goal of the SVM method is to determine the best line or decision boundary for categorizing n -dimensional space into classes, so that subsequent data points can be easily placed in the relevant category. This ideal decision boundary is referred to as a hyperplane. SVM is used to choose the extreme points/vectors that help build the hyperplane. These extreme points are known as support vectors, and hence, algorithm is called support vector machine. The margin is the difference between two lines on the closest data points of different classes [26]. The perpendicular distance between the line and the support vectors can be used to determine the margin. A wide margin is considered a good margin, whereas a tiny margin is considered a poor margin. The training culminated in the creation of a decision surface that splits the space into two subspaces. Each subspace represents a different training data class. Once the training is complete, the test data is mapped to the feature space. These data are then assigned a class based on which

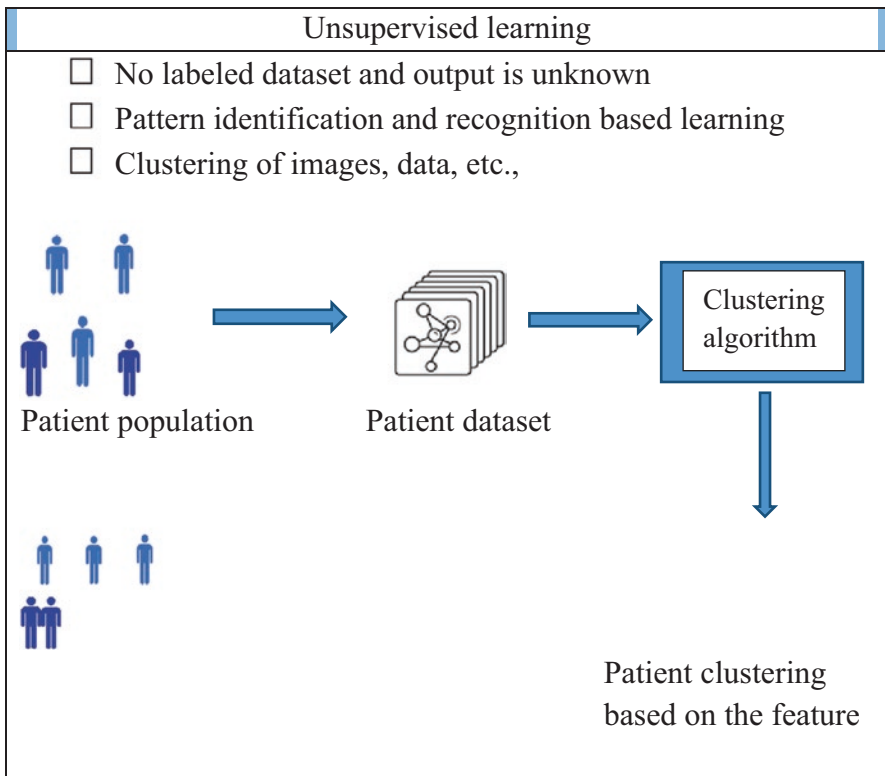
subspace they are mapped to. The fundamental goal of SVM is to divide datasets into classes in order to find the maximum marginal hyperplane (MMH), which can be done in two steps [29]:

1. To begin, SVM will construct hyperplanes that best divide the classes iteratively.
2. The hyperplane that correctly separates the classes will then be chosen.

By applying the above steps, a linear discrimination SVM model, which is a binary classifier, divides the space into two classes of MRI images by predicting the hyperplane.

Unsupervised Learning

Unsupervised machine learning techniques are effective for applications like clustering, since they don't require tagged data. For example, image classification may be studied using clustering methods. Unsupervised clustering, in addition to analyzing existing data, predictions can also be made using algorithms. A model can be trained using the dataset from various data sources.



K-Means Algorithm

A popular unsupervised learning approach is the k-means clustering algorithm. This approach uses an integer k and n observations. The result is a k -set partition of the n observations, with each observation belonging to the cluster with the closest mean. The operations of k-means are summarized in the stages below.

1. Set up k cluster centers. This can be achieved in practice by selecting k center at random.
2. Points are derived from n observations or k randomly generated center points.
3. Calculate how far each observation is from the cluster centers.
4. Each point should be assigned to the cluster with the smallest distance between its center and the other cluster centers.
5. As the cluster mean, recalculate the placements of the k centers.
6. Once more, compute the distance between each data point and the newly estimated centers.

Steps 3 and 4 should be repeated until all data points have been allocated to the same cluster (data points do not move).

Dataset of individuals with Alzheimer's (AZ) disease having longitudinal evaluations. The National Alzheimer's Coordinating Center database, supplying neuropathological data. The AZ disease in the image was clustered using the means clustering technique. The researchers wanted to look at subgroups of patients with different extrapyramidal sign progression trajectories and their clinical and neuropathological correlates [4]. K-means algorithm to cluster the neurological dataset is based on the disease condition using the above steps [3].

Mean-Shift Clustering Algorithm

The mean-shift clustering method is a centroid-based approach that aids unsupervised learning in a variety of situations. It is one of the most effective image processing and computer vision algorithms. It works by relocating data points to centroids, which then become the mean of other points in the region. The mode searching algorithm is another name for it. The benefit of the method is that it distributes clusters to data without automatically specifying the number of clusters depending on bandwidth.

Unlike the K-means cluster technique, mean-shift does not need a priori cluster number specification. In relation to the data, the algorithm determines the number of clusters [5]. For fMRI (functional magnetic resonance imaging) analysis of neurological disorders the temporal properties of the image is taken. Simulated and actual fMRI data were utilized to compare cluster analysis with mean-shift clustering, which uses a feature space that includes temporal and spatial features [6]. Figure 10 shows the flow of mean-shift algorithm used for clustering medical images.

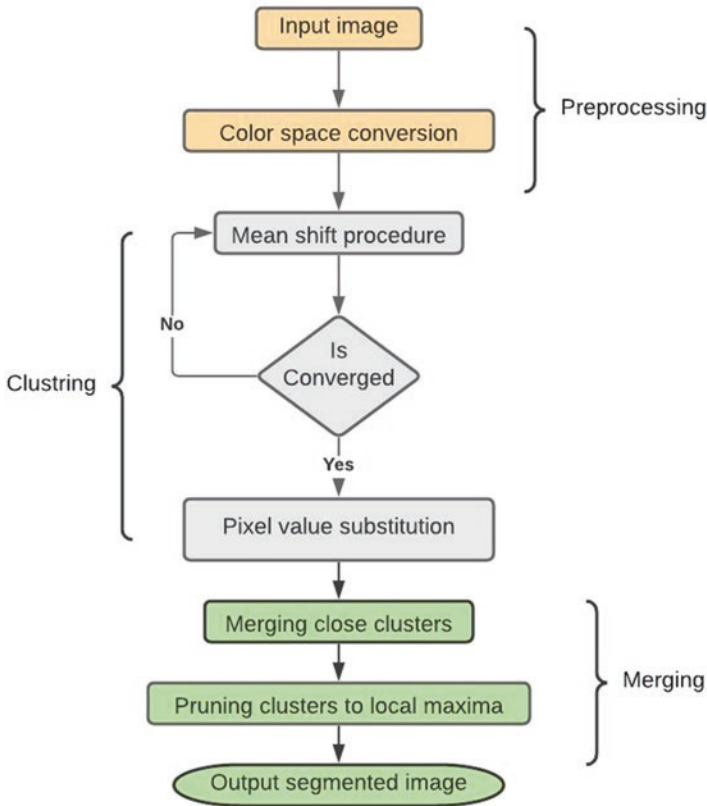


Fig. 10 Flowchart of mean-shift algorithm used for medical image classification

Affinity Propagation and Hierarchical Clustering

To split the brain MRI images into multiple clusters, an affinity propagation approach was used. The affinity propagation method uses tissue-segmented and anatomically parcellated pictures to characterize the similarity between brain images [7]. After clustering, a representative exemplar image is found as the single topic atlas for each clustered, and the MRI images fitting to the same subgroup are recognized. Figure 11 depicts the brain MRI image's tissue segmentation. The algorithm of affinity propagation is:

- Algorithm works based on resemblances among data points.
- All data points are considered as the potential cluster center.
- Cluster centers are identified that represents the dataset.

Affinity propagation requires two data:

- Similarities between data points $s(i,k)$
- Preferences $s(k,k)$

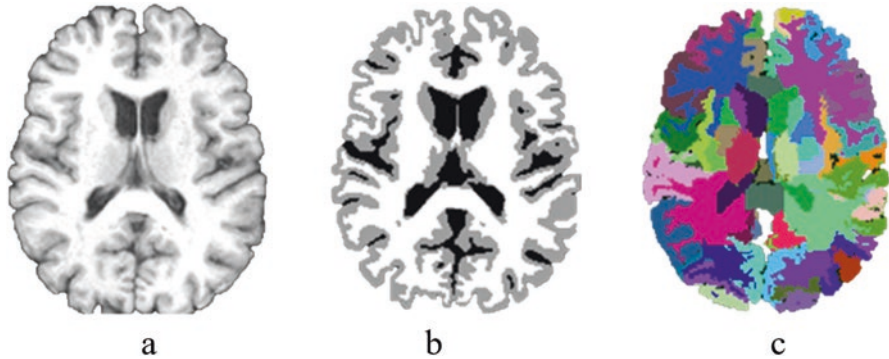


Fig. 11 (a) Original image. (b) Tissue segmented image using affinity propagation. (c) Structure parcellation using affinity propagation

i represents a data point.

k represents the exemplar.

The hierarchical clustering-based segmentation (HCS) is an unsupervised method for separating various sections in image data. The HCS method segments the images by splitting an image into its regions at hierarchical degrees of permitted dissimilarity across the different regions [36]. As the allowable threshold value is increased, the hierarchy represents the uninterrupted merging of similar, adjacent, and disjoint regions [8, 9].

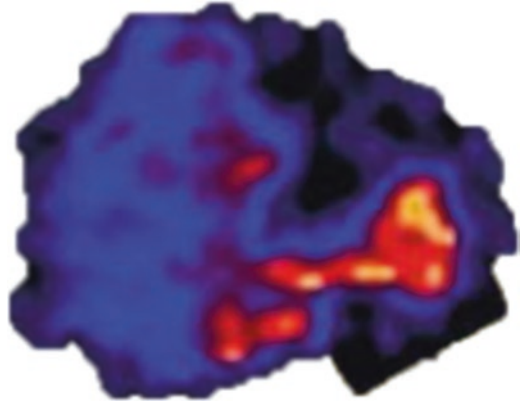
Hierarchical clustering-based segmentation algorithm includes the following.

1. Labeling of each pixel in the image:
 - Label every pixel according to the pre-segmentation if the image has already been segmented.
 - If initial segmentation isn't possible, label every pixel as a separate section.
 - The amount of distinction between areas that can be allowed should be set to zero.
2. Dissimilarity value is calculated among in the given image.
 - The least dissimilarity value should be used as the threshold value.
3. If the threshold and dissimilarity value are equal, then merge all the values close to the dissimilarity value or else move to step 6.
4. If the regions merged in the above step is greater than 0, reclassify the pixels on the boundary of the combined areas with the remaining regions until no further classification is done.

Save the region data for this iteration as an intermediate segmentation among the combined regions after all accessible border pixels have been classed, and move on to step 2.

Otherwise, move to step 5, if the areas merged in step 3 is equal to 0.

Fig. 12 Hierarchical clustering for brain image



5. Move to step 7 if the number of regions in the image is less than the predetermined value, otherwise go to step 6.
6. If the current dissimilarity permitted value is well below the maximum permissible value, go to step 2 and gradually increase the dissimilarity allowed value. If not, go to step 7.
7. Save the data from the current iteration.

Figure 12 [35] shows the output of the hierarchical clustering algorithm of brain image by applying the abovementioned procedure.

Density-Based Spatial Clustering (DBSC)

Density-based clustering is a term used to describe unsupervised learning approaches for identifying unique groups/clusters in input images. DBSCAN is a density-based clustering base technique. It can detect clusters of various shapes and sizes in a large amount of noisy data including outliers [10].

Two parameters are used in the DBSC algorithm:

minPts: The minimum number of clustered points required for a region to be considered dense (a threshold)

eps (): A distance metric that may be used to find points in the vicinity of a given location

- The algorithm selects a point in the image dataset at random.
- If there are at least “minPoint” points within a radius, we consider it to be part of the same cluster.
- The clusters are then extended, by repeating the neighborhood computation for each nearby point.

Static and dynamic functional connectivity of the fMRI images can be analyzed using the DBSC algorithm by applying the above procedure [11, 12].

Gaussian Mixture Modeling

Histogram thresholding is one of the most often used methods for segmenting pictures, and Gaussian-mixture-based segmentation is based on it. The assumption in histogram thresholding is that an image has two areas or classes: target and background, each with a unimodal gray-level distribution. As a result, the segmentation problem entails selecting an appropriate threshold for partitioning the picture into target and background regions. The probability density function (PDF) of the gray levels in the picture is a combination of two Gaussian density functions with specific means, standard deviations, and proportions in Gaussian-mixing segmentation techniques [12, 13].

$$P(x) = \sum_{a=1}^b p_a N(x | \alpha_a, \sigma_a^2)$$

where b is the number of regions

$p_a > 0$ -weights, i.e., is $\sum_{a=1}^b p_a$.

$$N(\alpha_a, \sigma_a^2) = \frac{1}{\sigma_a \sqrt{2\pi}} \exp \frac{-(x - \alpha_a)^2}{2\sigma_a^2}$$

where

α_a – mean of class a .

σ_a^2 – standard deviation of class a .

Figure 13 [35] shows the input MRI image and Gaussian output image [14]. White and gray matter in the brain MRI image can be classified to analyze neurological disorders, like attention deficit hyperactivity disorder (ADHD).

Convolutional Neural Network (CNN)

CNN stands for convolutional neural network and is a sort of deep learning neural network. In short, consider CNN to be a machine learning system that can take an input image, assign relevance (learnable weights and biases) to various aspects/objects in the image, and distinguish between them. Figure 14 shows the block diagram of CNN [15, 30, 35].

CNN extracts data from images by extracting features.

The following are the components of any CNN:

- A grayscale image serves as the input layer.
- The output is a multi-class labeling system.
- Convolution layers, rectified linear unit layers, pooling layers, and a fully linked neural network are all hidden layers in the CNN architecture.

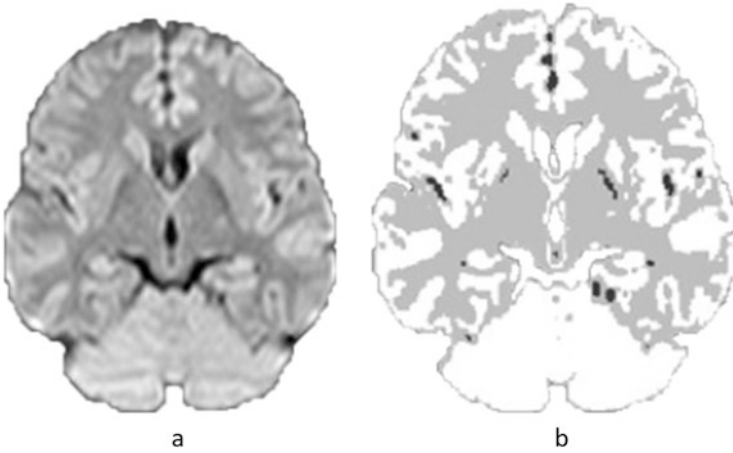


Fig. 13 (a) Input MRI image, (b) Gaussian image

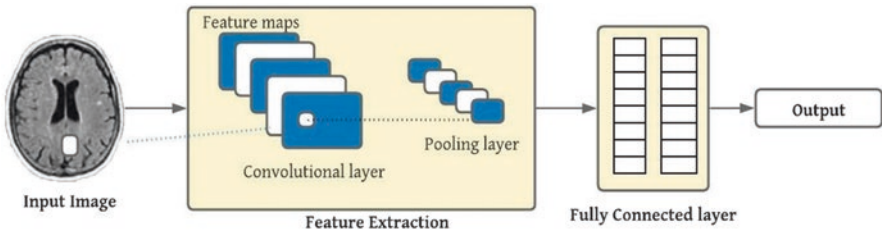


Fig. 14 Block diagram of CNN

Convolutional Layer

It has a convolutional filter of size $N \times M \times D$, where D is the image depth. N and M are the image pixel. Kernels are convolved over the image region, and the dot product among the filter entries is computed during the forward pass [16].

Activation Functions

For nonlinear transformation of medical images, the following activation functions are used:

- Sigmoid
- Tan hyperbolic
- Rectified linear unit

Pooling Layer

Convolved features are down sampled in the pooling layer. By dimensionality reduction, it reduces the amount of computing power necessary to process the data. It avoids overfitting and minimizes translation and rotational variance of images by aggregating data over space. It also minimizes spatial size by aggregating data across space or feature types. The input to a pooling process is partitioned into a collection of rectangular patches. Depending on the pooling, each value is replaced with a single value. Maximum pooling and average pooling are the two forms [27, 28].

Fully Connected Layer

The fully connected layer is comparable to a deep neural network in that each node has interconnections to all of the inputs and that each link has weights associated with it. The total of all inputs multiplied by the weights yields the final output. The classifier work is performed by the fully connected layer, which is preceded by the sigmoid activation function. In Table 1, various types of CNN are used for the classification of neurological disorders. The table highlights some widely used CNN networks used for the analysis of disorders.

S. no.	CNN	Developed by	Neurological disorders
1.	Modified LeNet [17]	R.A. Hazarika, A. Abraham, D. Kandar and A.K. Maji	Mild cognitive impairment (MCI), cognitively normal (CN), Alzheimer’s disease (AD)
2.	VGG16 [18]	Jain, R., Jain, N., Aggarwal, A. and Hemanth, D.J.	AD, MCI
3.	AlexNet, GoogLeNet, ResNet50 [19]	Khagi, B., Lee, B., Pyun, J.-Y. and Kwon, G.-R.	AD, healthy control (HC)
4.	ResNet-18 [20]	M. Raza, M. Awais, W. Ellahi, N. Aslam, H.X. Nguyen and H. Le-Minh	CN, significant memory concern (SMC), early mild cognitive impairment (EMCI), mild cognitive impairment (MCI), late mild cognitive impairment (LMCI), and AD
5.	VGG 19 [21]	Bhatele, K.R. and Bhadauria, S.S.	Alzheimer’s disease (AD) and Parkinson’s disease (PD)
6.	U-Net [22]	Fan, Zhonghao, Li, Johann, Zhang, Liang, Zhu, Guangming, Li, Ping, Lu, Xiaoyuan, Shen, Peiyi, Shah, Syed, Bennamoun, Mohammed, Hua, Tao and Wei, Wei	AD, late MCI, early MCI
7.	3D CNN [23]	Yee, Evangeline, Ma, Da, Popuri, Karteek, Wang, Lei and Beg, Mirza Faisal	Stable dementia, stable normal control

Conclusion

Advanced CNN approaches and models in supervised and unsupervised algorithms, as well as advances in high-speed computing techniques, provide a unique opportunity to predict and manage a variety of neurological illnesses, such as Alzheimer's disease, Parkinson's disease, and schizophrenia, among others. The most popular CNN models were investigated in this article for diagnosing neurological disorders using MRI, fMRI, and CT scan dataset. The use of CNN techniques to classify neurological illnesses identified in the literature has been described.

References

1. A.-M. Tautan et al., Artificial intelligence in neurodegenerative diseases: a review of available tools with a focus on machine learning techniques. *Artif. Intell. Med.* **117**, 102081 (2021)
2. A. Tagaris, D. Kollias, A. Stafylopatis, G. Tagaris, S. Kollias, Machine learning for neurodegenerative disorder diagnosis—survey of practices and launch of benchmark dataset. *Int. J. Artif. Intell. Tools* **27**(03), 1850011 (2018)
3. M.B.T. Noor, N.Z. Zenia, M.S. Kaiser, et al., Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain Inf.* **7**, 11 (2020)
4. G. Tosto, S.E. Monsell, S.E. Hawes, G. Bruno, R. Mayeux, Progression of extrapyramidal signs in Alzheimer's disease: clinical and neuropathological correlates. *J. Alzheimers Dis.* **49**, 1085–1093 (2016)
5. L. Ai, J. Xiong, Temporal–spatial mean-shift clustering analysis to improve functional MRI activation detection. *Magn. Reson. Imaging* **34**, 1283 (2016)
6. H. Cho, S. Kang, S.I. Cho, Y.H. Kim, Image segmentation using linked mean-shift vectors and its implementation on GPU. *IEEE Trans. Consum. Electron.* **60**, 719–727 (2014)
7. G. Li et al., Grouping of brain MR images via affinity propagation. The Midwest symposium on circuits and systems conference proceedings: MWSCAS. Midwest symposium on circuits and systems (2009)
8. C. Hejjun, J. Zheng, Y. Guo, Y. Shen, A robust image segmentation method using hierarchical color clustering, pp. 1–6. <https://doi.org/10.1145/3028842.3028843> (2016).
9. A. Selvan, L. Cole, L. Spackman, C. Wright, Hierarchical cluster analysis to aid diagnostic image data visualization of MS and Other medical imaging modalities, in *Imaging Mass Spectrometry Methods and Protocols*, Methods in Molecular Biology, ed. by L. Cole, vol. 1, (Humana Press, Totowa, 2017), pp. 95–123
10. <https://www.kdnuggets.com/2020/04/dbscan-clustering-algorithm-machine-learning.html>
11. D. Rangaprakash, T. Odemuyiwa, D. Narayana Dutt, et al., Density-based clustering of static and dynamic functional MRI connectivity features obtained from subjects with cognitive impairment. *Brain Inf.* **7**, 19 (2020)
12. H. Bi, H. Tang, G. Yang, et al., Accurate image segmentation using Gaussian mixture model with saliency map. *Pattern. Anal. Applic.* **21**, 869–878 (2018)
13. J.B. Ashly, S.N. Kumar, F.A. Lenin, K.H. Ajay, V. Suresh, An improved Gaussian mixture model based on prior probability factor for MR brain image segmentation, in *Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing*, ed. by K. Das, J. Bansal, K. Deep, A. Nagar, P. Pathipooranam, R. Naidu, vol. 1057, (Springer, Singapore, 2020)

14. L. Moraru, S. Moldovanu, L.T. Dimitrievici, N. Dey, A.S. Ashour, F. Shi, S.J. Fong, S. Khan, A. Biswas, Gaussian mixture model for texture characterization with application to brain DTI images. *J. Adv. Res.* **16**, 15–23., issn 2090-1232 (2019)
15. D.R. Sarvamangala, R.V. Kulkarni, Convolutional neural networks in medical image understanding: a survey. *Evol. Intel.* **15**, 1–22 (2021)
16. R.A. Hazarika, A. Abraham, D. Kandar, A.K. Maji, An improved LeNet-Deep Neural Network Model for Alzheimer's disease classification using brain magnetic resonance images. *IEEE Access* **9**, 161194–161207 (2021)
17. E. Yagis et al., Effect of data leakage in brain MRI classification using 2D convolutional neural networks. *Sci. Rep.* **11**(1), 22544 (19 Nov 2021)
18. R. Jain, N. Jain, A. Aggarwal, D.J. Hemanth, Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cogn. Syst. Res.* **57**, 147 (2019)
19. F. Ramzan et al., A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural Networks. *J. Med. Syst.* **44**(2), 37 (2019)
20. B. Khagi, B. Lee, J.-Y. Pyun, G.-R. Kwon, CNN models performance analysis on MRI images of OASIS dataset for distinction between Healthy and Alzheimer's patient, in *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, (IEEE, Piscataway), pp. 1–4
21. K.R. Bhatele, S.S. Bhadauria, Classification of neurodegenerative diseases based on VGG 19 deep transfer learning architecture: a deep learning approach. *Biosci. Biotechnol. Res. Commun.* **13**(4), 1972–1980 (2020)
22. Z. Fan, J. Li, L. Zhang, G. Zhu, P. Li, X. Lu, P. Shen, S. Shah, M. Bennamoun, T. Hua, W. Wei, U-net based analysis of MRI for Alzheimer's disease diagnosis. *Neural Comput. Applic.* **33**, 13587 (2021). <https://doi.org/10.1007/s00521-021-05983-y>
23. E. Yee, D. Ma, K. Popuri, L. Wang, M.F. Beg, Construction of MRI-based Alzheimer's disease score based on efficient 3D convolutional neural network: comprehensive validation on 7,902 images from a multi-center dataset. *J. Alzheimers Dis.* **79**, 1–12 (2020). <https://doi.org/10.3233/JAD-200830>
24. S.A. Mostafa, A. Mustapha, S.H. Khaleefah, M.S. Ahmad, M.A. Mohammed, Evaluating the performance of three classification methods in diagnosis of Parkinson's disease, in *SCDM 2018. AISC*, ed. by R. Ghazali, M. Deris, N. Nawi, J. Abawajy, vol. 700, (Springer, Cham, 2018), pp. 43–52. https://doi.org/10.1007/978-3-319-72550-5_5
25. G.R. Ramani, G. Sivagami, Parkinson disease classification using data mining algorithms. *Int. J. Comput. Appl.* **32**(9), 17–22 (Oct 2011)
26. D. Gil, M. Johnson, Diagnosing Parkinson by using artificial neural networks and support vector machines. *Glob. J. Comput. Sci. Technol.* **9**, 63–71 (2009)
27. S. Joshi, D. Shenoy, G.G. Vibhudendra Simha, P.L. Rrashmi, K.R. Venugopal, L.M. Patnaik, Classification of Alzheimer's disease and Parkinson's disease by using machine learning and neural network methods, in *2010 Second International Conference on Machine Learning and Computing*, (2010), pp. 218–222. <https://doi.org/10.1109/ICMLC.2010.45>
28. S. Sivaranjini, C.M. Sujatha, Deep learning based diagnosis of Parkinson's disease using convolutional neural network. *Multimed. Tools Appl.* **79**, 15467–15479 (2019). <https://doi.org/10.1007/s11042-019-7469-8>
29. L. Khedher, I.A. Illan, J.M. Gorriz, J. Ramirez, A. Brahim, A. Meyer-Baese, Independent component analysis-support vector machine-based computer-aided diagnosis system for Alzheimer's with visual support. *Int. J. Neural Syst.* **27**(03), 1650050 (2017)
30. A. Segato, A. Marzullo, F. Calimeri, E. De Momi, Artificial intelligence for brain diseases: a systematic review. *APL Bioeng.* **4**(4), Article ID 041503 (2020)
31. J. Maroco et al., Data mining methods in the prediction of dementia: a real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC. Res. Notes* **4**(1), 299 (2011)

32. A. Mozhdehfarahbakhsh, S. Chitsazian, P. Chakrabarti, T. Chakrabarti, B. Kateb, M. Nami, An MRI-based deep learning model to predict Parkinson's disease stages. *J. Critic. Rev.* Preprint from medRxiv (Feb 2021). <https://doi.org/10.1101/2021.02.19.2125208>
33. L. Billeci, A. Badolato, L. Bachi, A. Tonacci, Machine learning for the classification of Alzheimer's disease and its prodromal stage using brain diffusion tensor imaging data: A systematic review. *PRO* **8**(9), 1071 (2020)
34. L. Bergouignan, L. Nyberg, H.H. Ehrsson, Out-of-body-induced hippocampal amnesia. *Proc. Natl. Acad. Sci. U. S. A.* **111** (2014). <https://doi.org/10.1073/pnas.1318801111>
35. G. Li, L. Guo, T. Liu, Grouping of brain MR images via affinity propagation, in *Conference Proceedings (Midwest Symposium Circuits Systems)*, (2009 May 24), pp. 2425–2428. <https://doi.org/10.1109/ISCAS.2009.5118290>
36. A. Selvan, S. Pettitt, C. Wright, Hierarchical clustering-based segmentation (HCS) aided interpretation of the DCE MR images of the prostate. (2015). <https://doi.org/10.13140/RG.2.1.1496.1123>