FP- International Journal of Computer Science Research (IJCSR)
Volume 2, Issue 2, Pages 167-175, June 2015

Automatic Text Classification Using Neural Network

Dr. Amit Verma
Professor and Head,
Department of CSE Chandigarh University, Gharuan Punjab
amit.verma@cumail.in

Er.Parminder Kaur
Department of CSE Chandigarh University, Gharuan, Punjab,
parminder.cu@gmail.com

Sandeesh Kaur
Department of CSE Chandigarh University, Gharuan Punjab
sandeesh1691@yahoo.com

ABSTRACT
Text Classification has been growing interest in the research of text mining. Automatically organizing a set of documents into different categories has observed as an active attention. It is of great importance due to enormously increased need to handle large number of electronic documents in real world applications such as web page classification, email processing, spam filtering, language identification etc. Manually organizing documents can be too expensive and takes a lot of time so automatically text classification is attractive because it solves all these problems. There are number of existing approaches, neural network mostly used for text classification due to its greater response time and can better handle noisy data. The aim of this paper is to highlight the performance of neural network in terms of text classification.

Keywords: Text Classification, Clustering, Back Propagation Neural Network.

1. INTRODUCTION

Text Classification is the process of automatically sorting a set of documents into different categories from predefined set [7]. Documents can be classified according to subjects or attributes (such as document type, author, printing year etc). The main goal of text mining is to enable user to extract the relevant information and deals with various operations like retrieval, classification and summarization [7]. This is also important to use data mining for companies and their internet/intranet based applications and information access. With the growth of the text information from the electronic documents and worldwide web, the proper clustering and classification of such enormous amount of information is a crucial step. Information is stored in the unstructured textual format. For this, some of the preprocessing steps are applied to represent the information in clear word format and then important features are selected by feature selection approaches. These approaches follow tokenization, indexes and feature space reduction. Text can be tokenized by using term frequency (TF), inverse document frequency (IDF) and tern frequency inverse document frequency (IFIDF) and by this way features are reduced. Classification phase process the data to make a decision text data belongs to which category. Before doing the actual classification, algorithm can be trained with trained data. So the Computational time can be improved significantly [24]. After that there is need for classification of documents So that documents can be properly classified. If searching is performed by user for particular data then clustering plays a vital role. If cluster is not proper then searching may involve a lot of time and search result may not be accurate. To solve this problem proper cluster and classification is required [1]. For this, back propagation algorithm in neural network helps in calculated the resulted error and it back propagated again and again until the desired outcome comes.

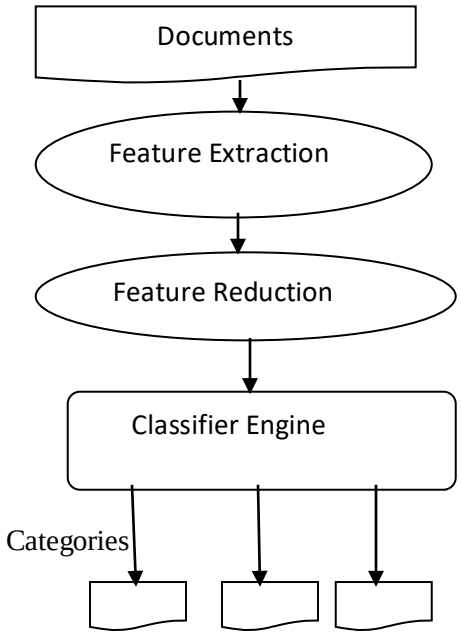

Figure 1. Text Classification Process Diagram

The remaining paper is divided into different sections. Section 2 discussed about detail of clustering techniques and algorithms. Section 3 discussed about different classification techniques. Section 4 gives detail of neural network and its advantages. The Section 5 provides the proposed related work. Section 6 described results and conclusion

2. CLUSTERING TECHNIQUES

All Various clustering techniques are:

2.1 Clustering

It can be considered the most important unsupervised learning problem. It deals with finding a structure in collection of unlabeled data. Clusters are formed on basis of similarity between the objects and dissimilarity between objects belonging to other clusters [6].A good clustering is produced when there is high intra class similarity and low inters class similarity. The cluster similarity can be found by distance function.

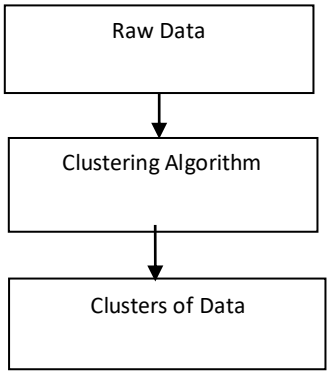

Figure 2[5]. Stages of Clustering

2.2 Clustering Algorithms

These all methods vary in different forms for clustering on basis of procedure used for measuring similarity, threshold variable, Hard and soft clusters like objects that strictly belongs to one cluster and belongs to other cluster to certain degree. These algorithms are classified into following methods.

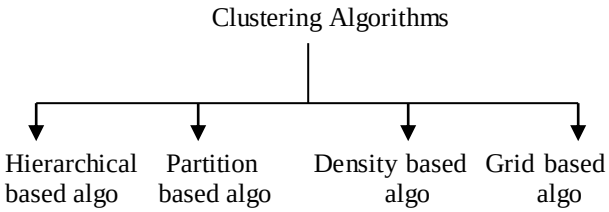

Figure 3. Clustering Algorithms

2.2.1 Hierarchical Methods

In this, cluster hierarchy is building in form of tree. Every cluster node contains a child clusters and child clusters have their common parent. It constructs clusters by step by step procedure. In one step data is not partitioned. It is of two types i.e. Agglomerative and Divisive

Agglomerative that's follow bottom-up approach that starts with single cluster and recursively merge two or more clusters until the objective is achieved in one whole cluster.

Divisive follow top-up approach that starts with whole one cluster and recursively split into two or more cluster until stopping criteria is achieved (requested number of clusters) .instructions.

2.2.2 Partitioning Methods

In this, data is divided into different subsets where each subset represents a cluster. Some greedy heuristic techniques are used because it is infeasible to check all possible subsets. These techniques are used in form of iterative optimization. This means different relocation schemes that iteratively reassign points between k clusters. By this way, in this clusters are revised after being constructed and clusters are gradually improved by relocation algorithm. These results in high quality clusters are formed. It is based on certain criterion functions i.e. minimizing the square errors criterion, which is computed as [5],

$$\text{Error}=\sum |p - m_i|^2 \quad (1)$$

Where p is the Point in cluster

m is the Mean of cluster

Its drawbacks are sometime point is close to centroid of another cluster then overlapping of data may be occurred.

2.2.2.1 Partition algorithms

K-Mean: It is most popular and simple clustering method [21].It is efficient for small data sets and numeric values. It starts with randomly selected some centroid and then assign the object to those cluster that are closest to centroid. Its drawbacks are not able to handle noisy data i.e. sensitive to discover outlier and discover clusters which have convex shapes.

Algo_K-Mean
Initialize k cluster center Determine the distance of each object to centroid **Repeat** Assign each point to nearest cluster center Compute new cluster center **Until** convergence criterion is met

Figure 4[4]. K-Mean algorithm

K-Mode Algorithm: This is also one of important partition method for handling especially categorical data and form of unsupervised learning algorithm. It is extended form of k-mean algorithm that removes the

numeric data limitations and extends the k-mean paradigm to deal with categorical objects by 1) using the simple matching dissimilarity measure 2) use frequency based methods to revise modes of clusters3) replace the means of clusters with modes .By this way, it minimize the cost function [13].

Algo_K-Mode
Select k initial modes one for each of cluster Reassign data object to those cluster whose mode is nearest Recompute new modes until no object has changed cluster membership

Figure 5[13]. K-Mode algorithm

K-Medoid Algorithm: This is partitioning technique of clustering algorithm for efficiently handling noise and outliers as compared to k-mean because it uses the centrally located objects in cluster. It uses the medoid as reference point instead of taking mean value of objects in each cluster. The main strategy of this is to select randomly k initial point as medoid and assign remaining objects to cluster with closest medoid. PAM (Partition around Medoid) are most commonly used algorithm of k-medoid [12].

Algo_ K-Medoid
Select arbitrarily k objects as medoid one for each cluster **Repeat** Assign other remaining objects to those cluster whose medoid is nearest Select non-medoid object O Compute the total points T of swapping object Oi with O If T<O then swap O_i with O to form new set of medoid **Until** no change occur

Figure 6[12]. K-Medoid algorithm

Table 1. Comparative analysis of different partitioning algorithms

Parameters	**K-Mean**	**K-Mode**	**K-Medoid**
Sensitive to Outlier/Noise	Yes	No (Efficient than k-mean)	No (efficient than k-mean)
Capability for efficient handle large data	No	Yes	Yes
Represent reference point	Mean	Mode	Medoid
Complexity	O(NKT)-Time O(N+K)-Space	O(NKT)	$O(k(n-k)^2)$
Type of data	Only numeric	Both Numerical and Categorical	Both numerical and Categorical

Where N is the Number of Objects
K is the Number of Clusters
T is the Number of iterations

2.2.3 Density Based Algorithms

Clusters are formed on basis of regions which have high density [27].It is defined as number of objects in a particular neighborhood. It is capable of finding clusters of arbitrary shapes and can be used to handle noise or outliers. The basic idea behind this is to continue growing the given cluster as long as the number of data points in neighborhood exceeds some threshold. In this, DBSCAN, GDBSCAN and OPTICS algorithms are used.

2.2.4 Grid Based Algorithms

In this, grid structure is formed on basis of quantization of object space into finite number of cells. Its main focus is spatial data i.e. data that model the geometric structure of objects in space, their relationships, properties and operations. Its processing time is fast because it depends on number of cells in each dimension of quantized space instead of number of data objects [3].It is also able to handle noisy data.Its algorithms are STING, Wave cluster and CLIQUE.

3. CLASSIFICATION

It can be classified as supervised learning problem and recognizes patterns that describe the group to which an item belongs [15]. It does this by examining items that already have been classified and inferring a set of rules. It group data according to similarities. It includes different approaches such as decision tree, support machine, K-Nearest Neighbor, Genetic algorithm, Neural Network etc.

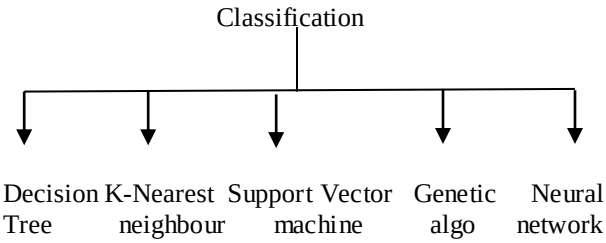

Figure 7. Classification techniques

3.1 Classification techniques

3.1.1 Decision tree

It is a tree like structure that describes the decisions about an item and their possible outcomes. In this, each internal node is represented by input feature and leaf node holds a class label. The basic idea behind this is to choose the best attribute to split the instances and make that attribute a decision node.

3.1.2 K-Nearest Neighbor

It is simplest among all algorithms. It is based on classify objects based on smaller distance. Data points can be classified according to class which has highest confidence. Overlap and distance metric can be used for this [17]. It is powerful, requires no training time, expensive at test time.

3.1.3 Support Vector Machine

It constructs a separating hyper plane between two data sets and maximizes the margin. Support Vectors are those which are present closest to hyper plane. By this way, good separation is achieved by hyper plane that has longest distance between neighboring data points and lower generalization error. It only classifies two classes.

3.1.4 Genetic Algorithm

It was developed by John Holland in 1970 [20].It is adaptive in nature which gives useful solutions to optimization and search problems. It produces adequate solution for particular problem and from large search space. It works effectively on numeric data and on large amount of data. It starts with randomly selected initial population and applies fitness function to evaluate the individuals. Selection is performed according to high fitness individuals. By this way, it manipulates the individuals through genetic operators i.e. selection, mutation and cross over.

Table 2. Benefits and Limitations of various classification techniques

Classification Techniques	Benefits/Advantages	Limitations/Disadvantages
Decision Tree	• Simple to understand and interpret • It handle both numerical and categorical data • Robust in nature because it better handle effects of outlier • Better perform large data in short duration of time • White box in nature because explanation of result is explained by Boolean logic	• Can create over-complex trees that do not generalize data well • Based on greedy algorithms where optimal decisions are made at each node • Don't handle real value parameters as well as Boolean
Support Vector Machine	• Model non-linear class boundaries • Can greatly reduce the entropy of retrieved result • Easy to control frequency of errors • Remove over fitting problem so that it can provide good generalization	• Has only one output • Data is trained one by one • Difficult to understand structure of algorithm
Genetic Algorithm	• Can be used in feature classification and selection • Used for optimization means provide useful solution to optimization and search problems • Efficient method for complex search problem space • Work efficient on large amount of data	• Take too much time and cannot find always exact solution • Not work efficient on small amount of data
	• Ease of maintenance • Better tolerate noisy and missing data • Instances can be classified by more than one output	• Structure of algorithm is complex so difficult to understand • No estimation and prediction errors are calculated

Neural Network	• Makes no assumption about distribution of data	• Needs training to operate • Limited ability to identify possible causal relation due to black box in nature

Table 3. Analysis of different classification techniques with respect to various parameters

PARAMETERS	Decision tree	SVM	Genetic Algorithm	Neural Network
Approach Follows	Self determined and Human intelligence	Non-Parametric	Global search approach	Non-parametric
Speed of training or Training time	Rules and relations between attributes, Selection of attributes, Size of dataset, Splitting of nodes	Choice of kernel parameter, Feature selection, classes separability	Selection of relevant genes and on discard noisy genes	Network topology, Learning rate, Momentum, Epochs, Hidden layer
Accuracy	Selection of attributes	Depends on selection of optimal separating hyper plane	Selection of genes or genetic heritage	Depends upon size of input classes
Complexity	$n^2 * m^3$ where n=no of attributes, m= data samples	Depends upon choice of kernel and kernel parameter	Depends on selection of features	Depends on hidden layers or Network structure

4. NEURAL NETWORK

Neural network is mathematical model based on biological neuron network and represent a brain metaphor for information processing [23]. Animals are able to react adaptively according to changes in their internal and external environment and for perform these behaviour they use nervous system. The nervous system is built by simple units called neurons which are organized in layers. Number of nodes which are interconnected called layers which contain an activation function. Neurons works by processing information by using weighted connection approach. They receive and provide information in form of spikes. Information is presented to network via input layer which then communicate to one or more hidden layers where actual processing is done. Weights are assigned to connections. The hidden layer then communicates to output layer. The main objective is to transfer the inputs to meaningful outputs.

Input layer Hidden layer Output layer layer

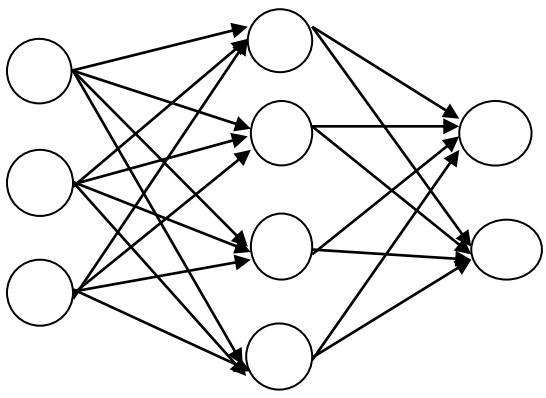

Figure 8. Artificial Neural Network

Tasks to be solved by neural network:

Pattern Recognization- Neural network is trained to recognize or retrieve patterns to solve optimization problems. E.g. a familiar face

Prediction- Model is built to access the class of unlabeled object where a moving object goes, when robot wants to catch it.

Classification- It describe the item belongs to which particular group and for this examine items that are already classified.

Forecasting- It is different from the predictions by estimates the future value of variables within the data.

4.1 Neural Network Topologies

Feed Forward Neuron Network-It is also called the multilayer perception when multiple layers exist. If only one layer exist, then called the perception. In this, information moves in only one direction from input to output. No loops and bidirectional flow of information are presented.

Recurrent Network-It contains the bidirectional flow of information. In this, information can also flow from output layer to previous hidden layers by adjusting weights of connections. It contains the loops and cycles in network.

4.2 Back propagation Algorithm

Rumelhart, Hinton and Williams(1986) presented the back propagation learning algorithm using multilayer

network[26,14].For training a neural network two learning algorithms are used that is supervised learning and unsupervised learning. In supervised learning there is a set of existing inputs and expected outputs. It iteratively assigns a set of weights for prediction of class labels tuples. The predicted output is compared with the resulted value to calculate the error. By this way, during training phase weight of network are adjusted automatically until the desired output comes and used in feed forward neural network to train the network. In unsupervised learning, there is only input values and clustering will performed by network to learn the classes. Its examples are Kohonen network and Hopfield networks.

Algo_Back propagation

Begin
Input: Problem Size, Input Patterns, $iteration_{max}$, $learn_{rate}$
Output: Network
Network ⟵ constructNetworkLayers()
$Network_{weights}$ ⟵ initializeWeights(network,ProblemSize)
For (i= 1 to $iteration_{max}$)
$Pattern_i$ ⟵ SelectInputPattern(InputsPattern)
$Output_i$ ⟵ ForwardPropagate($pattern_i$,Network)
BackwardPropagateError($pattern_i$,$output_i$,Network)
UpdateWeights($pattern_i$,$output_i$,Network, $learn_{rate}$)
End
Return(network)

Figure 9. Pseudocode of back propagation algorithm

4.3 Advantages

a) Neural network perform tasks that a linear network cannot.

b) It can tolerate missing and noisy data.

c) Ease of Maintenance-When element of neural network fails, it can continue without any problems by their parallel nature.

d) It can handle problems in any application and easy to implement.

e) It makes no assumptions about the distribution of data and use middle layer to develop the interval representation of relationship between variables.

4.4 Disadvantages

a) No estimation and prediction errors are calculated.

b) Nature of Neural network is black boxes so in this it is difficult to estimate the relations in hidden layers.

c) It needs training to operate.

5. LITERATURE REVIEW

M.Ruiz et al.[26] described neural network based automatic text categorization is designed for solving problem of recognizing Mesh terms for particular document given the set of important words and test it using MEDLINE dataset. The comparison of counter propagation network against back propagation neural network is done and observed that back propagation takes less training time and error rate is low then counter propagation network.

C.Breen et al. [25] used neural network to examine and accurately classify the images based on visual content. Experiment result show that when images are retrieved based on user's query then system will yield a minimal amount of irrelevant information and maximum amount of relevant information.

A,Selamat et al. [24] used WPCM method for classify web pages. For this, principle component analysis has been used to select the most relevant features and output of this is combined with the feature vector from class profile which contains the most regular words in each class. Then these are used as input to neural network for classification. It gives better classification accuracy with sport news database.

M.Govindarajan et al. [22] proposed back propagation neural network classifier that performs cross validation for original neural network and measure training time for five larger and smaller data sets. So by this way, training time is reduced by more than ten times faster when dataset is larger and accuracy was same for all dataset instead of one smaller data set. This show that bigger the dataset and larger the improvement in speed.

A.Dallali et al. [16] proposed the FCMC-NN for classification of electrocardiography signal and applied to database. First of all Fuzzy c-Mean are used for clustering and diagnose an abnormal heartbeat on

electrocardiogram then output of this is given to neural network for classification. It achieves more optimum cluster centers and reduces training time of neural network classifier and more accurate classification results are obtained.

S.Narwal el al. [15] proposes that data mining tools solve the future trends and behaviors and can answers business questions that traditionally were time consuming to resolve. In this various classification techniques like SVM, decision tree and neural network are applied on XML data. Classification errors and execution time for five datasets are calculated using algorithms.

N.Mathura et al. [14] applied k-mean and back propagation algorithms to recruitment data of an organization after preprocessing the data. Two datasets are used and algorithms were trained with records of one dataset and tested with records of another dataset. It is observed that useful data can be discovered by back propagation algorithm so it has high accuracy than k-mean algorithm.

B.Madasamy et al. [9] described that in medical data the common problem for classification using neural network are complex nature of data, high dimensionality, extract patters etc. Clustering can be done with the help of rule based induction, nearest neighbor, multilayer preceptor and back propagation learning algorithm to reduce the amount of samples and presented to neural network. It improves accuracy and computation time.

A.partra et al. [8] does text classification using BPNN and relevance factor as term weighing method. For this newsgroup20 dataset has been preprocessed by stop words removal, port stemming algorithm. The experiment result shows that the tf.rf performs better than tf in terms of F measure term because of preprocessing time is high.

S.Indu et al. [2] do an extension to k-mean algorithm is done by calculating a weight for each dimension in each cluster and uses this weight value to identify the important dimension of subset. This reduces the sparsity problem of high dimensional data and minimizes the cluster dispersion and entropy is reduced. template.

Table 4. Several Related Studies that deals with Neural Network

Ref.	Title of Paper	Dataset	Algorithms or Method	Parameters	Result and conclusion
[26]	Automatic Text Categorization using Neural Network (1998)	MEDLINE documents	Back propagation and Counter propagation network	Precision, Recall, Error rate(0.2)	Neural network based automatic text categorization is designed for solving problem of recognizing Mesh terms for particular document and test it using MEDLINE dataset. The comparison of counter propagation network against back propagation neural network is done and observed that back propagation takes less training time and error rate is low then counter propagation network.
[24]	Web page Classification Using Neural Networks (2003)	Sport News database	Principle Component analysis(PCA) and Class Profile based Features(CPBF) combined with Neural Network	Accuracy (84.10 %)	Principle component analysis has been used to select the most relevant features and output of this is combined with the feature vector from class profile which contains the most regular words in each class. Then these are used as input to neural network for classification. It gives better classification accuracy as compared to Bayesian and TF-IDF algo's.
[22]	Classifier Based Text Mining for Neural Network (2007)	Five data sets like contact lenses, cpu, weather symbolic, Weather, labor-nega-data	Proposed Back propagation neural network classifier i.e. perform cross validation	Accuracy, Training time, Speed(10 times faster)	Proposed back propagation neural network classifier that performs cross validation for original neural network and measure training time for five larger and smaller data sets. So by this way, training time is reduced by more than ten times faster when dataset is larger and accuracy was same for all dataset instead of one smaller data set. This show that bigger the dataset and larger the improvement in

					speed.
[18]	Naïve Bayes vs. Decision tree vs. Neural Network in classification of training web pages (2009)	Training course web pages	NB classifier, Decision tree, Neural network	Accuracy (95.20%), F-measure (97%)	Automatic analysis and classification of attribute data from training web pages is done and same data sample is run through NB, decision tree. NN classifier. NN is accurate for some domains but in this they need more training and inefficient for handling large number of features. By this way, experiment result show that enhanced NB Perform betters than other techniques.
[14]	Comparison of k-means and Back propagation Data Mining Algorithms (2012)	Organizatio n Recruitment data	K-means algorithm, Back propagation algorithm	Accuracy(80. 40 and 91.42 % with two data sets)	Both k-mean and back propagation algorithms are applied to recruitment data of an organization after preprocessing the data. Two datasets are used and algorithms were trained with records of one dataset and tested with records of another dataset. It is observed that useful data can be discovered by back propagation algorithm so it has high accuracy than k-mean algorithm.
[11]	A comparative study of various methods of classification (2012)	Blood test dataset	Decision tree, Neural Network, Bayesian methods, Rough sets	Accuracy (91.17 %)	Four classification methods are used to identify liver disorder and alcohol by classification of blood test data and evaluate through WEKA tool. The experiment result show that neural network give best results among other method in terms of increasing training size and number of features.
[10]	Automated text classification using dynamic artificial neural network model (2012)	Reuters-21578	DAN2 i.e dynamic architecture for artificial neural network	Precision, Accuracy, Recall	DAN2 approach for automatic text categorization does not require experimentation with parameter setting and network configuration and evaluate its performance using reuters-21578 dataset. Experiment result show that it perform better than KNN and SVM at p=.05 level of significance.

6. RESULTS AND CONCLUSION

In this paper, various techniques of text classification have been discussed. Every technique has its own benefits and limitations. Results from above study concluded that neural network approach in data mining is a promising field of research that works on small as well as large amount of data and reported ability of neural network to detect higher noisy data. In most cases, neural network classification accuracy is better than other classification techniques like SVM, Decision tree, k-nearest neighbor etc because of its characteristics like parallel processing, higher response time, and robustness, less sensitive to noise. It is able to perform very complex pattern matching tasks and its response time is faster.

7. REFERENCES

[1] Shalu Sharma, Sukhvinder Kaur and Ms. Jagdeep Kaur. “Hybrid Clustering and Classification.” International Journal of Advanced Research in Computer Science and Software Engineering, pp.222-225, vol.05, 2015

[2] S.Subha Indu and K.Devika Rani Dhivya,“Implementation of an Entropy Weighted k-mean algorithm for high dimensional sparse data.” International Journal of Advance Research in Computer Science and Management Studies, pp.204-209, vol.02, 2014

[3] Amandeep Kaur Mann and Navneet Kaur. “Grid Density Based Clustering Algorithm.” International Journal of Advanced Research in Computer Engineering and Technology, vol.02, pp.2143-2147, 2013

[4] Patel, Abhishek, and Purnima Singh. "New Approach for K-mean and K-medoids Algorithm." International Journal of Computer Applications Technology and Research 2, no. 1, pp.1-5, 2013

[5] Kaur Mann, Navneet Kaur. "Review Paper on Clustering Techniques." Global Journal of Computer Science and Technology 13, no. 5, 2013

[6] Bansal, Ms Neeraj, and Mr Amit Chugh. "Differentiate Clustering Approaches for Outlier Detection." International Journal of innovative Research in Computer and Communication Engineering, pp.193-196, vol.01, 2013

[7] Bhumika, Prof Sukhjit Singh Sehra and Prof Anand Nayyar. "A Review Paper on Algorithms used for Text Classification", International Journal of Application in Engineering and Management, pp.90-99, vol. 02, 2013

[8] Patra, Anuradha, and Divakar Singh. "Neural Network Approach for Text Classification using Relevance Factor as Term Weighing Method." International Journal of Computer Applications 68, no. 17, pp.37-41, 2013

[9] Madasamy, B., and J. Jebamalar Tamilselvi. "Improving classification Accuracy of Neural Network through Clustering Algorithms." International Journal of Computer Trends and Technology, pp.3242-3246, vol.4, 2013

[10] Ghiassi, M., M. Olschimke, B. Moon, and P. Arnaudo. "Automated text classification using a dynamic artificial neural network model." Expert Systems with Applications 39, no. 12, pp. 10967-10976, 2012

[11] Barnaghi, Peiman Mamani, Vahid Alizadeh Sahzabi, and Azuraliza Abu Bakar. "A comparative study for various methods of classification." In Int. Conference on Information and Computer Networks (ICICN 2012). IPCSIT, vol. 27, 2012.

[12] Velmurugan, Dr T. "Efficiency of k-Means and K-Medoids Algorithms for Clustering Arbitrary Data Points." Int. J. Computer Technology & Applications3, no. 5, pp.1758-1764, 2012

[13] Syal, Rishi, and V. Vijaya Kumar. "Innovative Modified K-Mode Clustering Algorithm." International Journal of Engineering Research and Applications, vol.02, pp.390-398, 2012

[14] Mathuriya, Nitu, and Dr Ashish Bansal. "Comparison of K-Means and Back propagation Data Mining Algorithms." International Journal of Computer Technology and Electronics Engineering (IJCTEE), vol.02, 2012

[15] Sharma, Narendra, Aman Bajpai, and Mr Ratnesh Litoriya. "Comparison the various clustering algorithms of weka tools." International Journal of Advance Research in Computer Science and Software Engineering, pp.886-878, vol. 3, 2012

[16] Dallali, A., A. Kachouri, and M. Samet. "Fuzzy c-means clustering, Neural Network, wt, and Hrv for classification of cardiac arrhythmia." ARPN Journal of Engineering and Applied Sciences 6, no. 10, vol.06, 2011

[17] Gupta, Megha, and Naveen Aggarwal. "Classification techniques analysis." InNCCI2010-National Conference on Computational Instrumentation, CSIO Chandigarh, India, pp.19-20, 2010.

[18] Xhemali, Daniela, Chris J. Hinde, and Roger G. Stone. "Naive Bayes vs. decision trees vs. neural networks in the classification of training web pages." International journal of Computer Science issues, vol.04, pp.16-23, 2009

[19] Gupta, Vishal, and Gurpreet S. Lehal. "A survey of text mining techniques and applications." Journal of emerging technologies in web intelligence 1, no. 1, pp.60-76, 2009

[20] Jimenez, J. F., F. J. Cuevas, J. M. Carpio, and E. Sciagura. "Genetic algorithms applied to clustering problem and data mining." In Proceedings of the 7th WSEAS International Conference on Simulation, Modelling and Optimization, pp. 219-224. World Scientific and Engineering Academy and Society (WSEAS), 2007.

[21] Ahmad, Amir, and Lipika Dey. "A k-mean clustering algorithm for mixed numeric and categorical data." Data & Knowledge Engineering 63, no. 2, pp.503-527, 2007

[22] Govindarajan and R. M. Chandrasekaran. "Classifier Based Text Mining for Neural Network." World Academy of Science, Engineering and Technology 27, pp.200-205, 2007

[23] Singh, Yashpal, and Alok Singh Chauhan. "Neural networks in data mining."Journal of Theoretical and Applied Information Technology 5, no. 6, pp.36-42, 2005

[24] Selamat, Ali, Sigeru Omatu, Hidekazu Yanagimoto, Toru Fujinaka, and Michifumi Yoshioka. "Web page classification method using neural networks."IEEJ Transactions on Electronics, Information and Systems 123, no. 5, pp.1020-1026, 2003

[25] Breen, Casey, Latifur Khan, and Arunkumar Ponnusamy. "Image classification using neural networks and ontologies." In Database and Expert Systems Applications, 2002. Proceedings. 13th International Workshop on, pp. 98-102. IEEE , 2002.

[26] Ruiz, Miguel E., and Padmini Srinivasan. "Automatic text categorization using neural networks." In Proceedings of the 8th ASIS SIG/CR Workshop on Classification Research, pp. 59-72. 1998.

[27] Ester, Martin, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. "A density-based algorithm for discovering clusters in large spatial databases with noise." In Kdd, vol. 96, no. 34, pp.226-231. 1996.