International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 10 (2019) pp. 2296-2302

Double Clustering Approach for Predicting Comorbidity Condition in Cardio Vascular Diseases

N.Devi[#], P.Leela Rani[##]

Assistant Professors
Department of Information Technology
Sri Venkateswara College of Engineering
Sriperumbudur, Tamil Nadu 602117, India.

Abstract

Comorbidity is significant during diagnosis and treatment of diseases. It increases the complexity of managing diseases in patients. Comorbid conditions are common among large population. A wide variety of conditions co-occur due to dependencies on one another. Traditional models used Charlson Comorbidity Index(CCI) for assessing the impact of patient demographic and comorbidity burden on patient health outcomes. A process of double clustering approach combining model-based and weighted K-means clustering methods for characterizing and summarizing a patient's comorbid conditions is used. This approach not only helps in data reduction but also preprocess the dataset for better classification. The weighted clustering helps in choosing the right or substantial set for the course. An assessment of performance by comparing two types of support vector machine algorithms is attempted by applying them to patients whose main diagnosis is cardiovascular disease. With this proposed approach, effort is made to improve the classification of patient health records.

1. INTRODUCTION

Chronic coexisting diseases, also called co morbidities[1], is the co-occurrence of one or more disorders in a person. This disorder act together and complicate the course of both the illness. Cardiovascular diseases (CVD) are non-communicable diseases characterized by high impact co morbidities which increase the morbidity and mortality rate. Therefore it is necessary to find the co morbid patterns [2] in the correlated diseases present in a patient who suffers from cardio vascular diseases This necessitates to develop a system that accurately predicts the level of co morbidities present in a CVD patient as well as to predict whether the normal is prune to cardio vascular disease at an early stage so that the person can take necessary steps to prevent it..

2. RELATED WORK

Many researchers[1][2] [4][5] analyzed co morbidities of disorders for prevention and early detection of cardio vascular diseases. For indentifying the co morbidity patterns in CVD, first clinical narratives are mined to extract important entities from medical text. From the extracted features, frequent pattern mining and frequent sequence mining are used to identify co morbidity patterns related to heart failures. Most of the FPM and FSM usually build a prefix tree and ignores the contextual information. Many advancement in data mining approach[3][6][8][9][10] makes it possible to extract entities from free text and deliver additional context attributes beyond the structured information about the patients. Novel hypotheses are generated to discover stable co morbidities and to confirm known ones.

Most medical diagnostic systems use artificial neural networks that perform multivariate analysis to identify the stage of diseases. An ANN consists of many processing elements that are interconnected. These elements include multiple input nodes, weighted intermediate nodes and output nodes. Estimation abilities of ANN make it prevalent to be used in wide area of medicine. Identification of co morbidity patterns of health conditions is critical for evidence-based practice to improve the prevention, treatment and health care of relevant diseases. Sh kay Ng [15] proposed an unified clustering algorithm based on the metric called Somers' D statistic - quantitative measure of co morbidity to identify co morbidity patterns with adjustment for multiple testing in order to curb the false discovery rate.

Cluster analysis, a machine learning approach, has been widely used in medical domain for diagnosing various diseases. Cluster analysis is used when the feature set are characterized by high heterogeneity. Cluster analysis is done by applying multiple clustering algorithms in order to cluster the feature set with high heterogeneity. The clusters are generated such that the intra cluster association among the features is strong where inter cluster association is weak between members of different clusters. In this paper we propose double clustering analysis based approach using LIBSVM and LIBLINEAR SVM to predict the co morbidity pattern in cardio vascular diseases.

3. PROPOSED ARCHITECTURE FOR PREDICTING COMORBIDITY PATTERN IN CVD USING DOUBLE CLUSTERING ANALYSIS

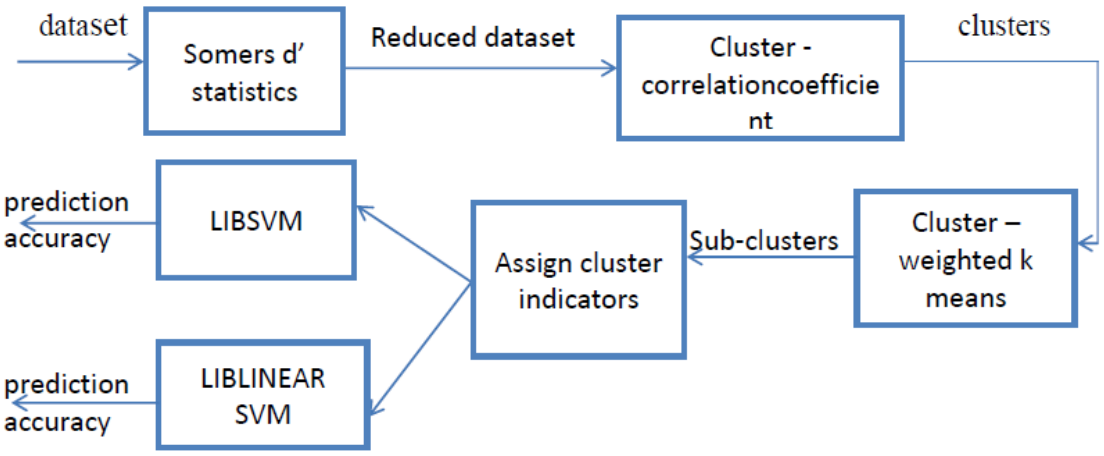

Figure 1. Proposed Architecture for predicting co morbidity pattern

The proposed approach for predicting co morbidity pattern in CVD using Double Clustering Analysis involves the following phases.

Phase 1: Reduce the CVD co morbidity patterns data set by applying Somers d' statistics.

Phase 2: The reduced data set are then grouped into different clusters using model-based clustering algorithm and then with weighted k-means algorithm. Each cluster contains the patient details having similar pattern of co morbid conditions. Each of these clusters is assigned a cluster indicator.

Phase 3: Classification of test data using LIBSVM and LIBLINEARSVM.

3.1 Applying Somers d' statistics

We mimicked the dataset similar to the one provided in heart disease dataset obtained from UCI machine learning repository for training and testing the proposed system. The 29 co morbid conditions considered here are : congestive heart failure, vascular disease, pulmonary circulation disorders, peripheral vascular disease, and hypertension (both uncomplicated and complicated), paralysis, other neurological disorders, chronic pulmonary disease, diabetes without chronic, diabetes with chronic, hypothyroidism, renal failure, liver disease, chronic peptic ulcer disease (includes bleeding only if obstruction is also present), HIV and AIDS (Acquired Immune Deficiency Syndrome), lymphoma, metastatic cancer, solid tumor without metastasis, rheumatoid arthritis/collagen vascular, coagulation deficiency, obesity, weight loss, fluid and electrolyte disorders, blood loss anemia, deficiency anemia, alcohol abuse, drug abuse, psychoses, depression. A sample dataset is shown in table 1.

If there are 29 co morbid conditions then there are 2^29 = 536, 870, 912 maximum number of possible combinations resulting in a very large clustering problem. In order to reduce the problem space, we apply the Somers D' statistics. Suppose if there are only 13 co morbidities in a group, there will be only 8,192 possible combinations and it will be easy to cluster.

Table 3.1 Sample data

Data source	cardiovascular disease	Depression	Congestive heart failure	Psychoses	Diabetes	Neurological disorders
state data organisation	present	extreme rate	major condition	extreme condition	major condition	major
consortia	present	extreme rate	moderate condition	extreme condition	extreme condition	major
state data organisation	absent	minor rate	major condition	extreme condition	extreme condition	extreme
hospital association	present	no class	extreme condition	extreme condition	major condition	major
consortia	absent	no class	no class	moderate condition	extreme condition	major
hospital association	present	mojor rate	extreme condition	extreme condition	major condition	extreme
hospital association	absent	mojor rate	extreme condition	extreme condition	moderate condition	extreme
consortia	absent	no class	moderate condition	major condition	extreme condition	major
state data organisation	present	mojor rate	extreme condition	major condition	extreme condition	no class
consortia	present	no class	extreme condition	extreme condition	extreme condition	major
consortia	absent	extreme rate	extreme condition	extreme condition	no class	moderate
consortia	present	extreme rate	major condition	extreme condition	major condition	extreme
hospital association	absent	extreme rate	extreme condition	major condition	moderate condition	extreme
consortia	absent	extreme rate	minor condition	major condition	major condition	major

Somers D' statistics is a measure a nonparametric measure of the strength and direction of association that exists between an ordinal dependent variable and an ordinal independent variable The ordinal independent variable in our dataset is the disease index i.e., type of cardiovascular disease and the ordinal dependent variable represent the co morbid disease correlated with the type of CVD. As per Somers'D statistics, we assign values between {-1,1}, -1 when all pairs of the variables disagree and 1 when all pairs of the variables agree.

It provides a quantitative measure of co morbidity. It also helps control the false discovery rate [14]. FDR is a method of conceptualizing the rate of type I errors in null hypothesis testing when conducting multiple comparisons. Sometimes, diseases that occur by coincidence or chance are considered to be co morbid. Somers' D Statistics helps to reduce these kinds of false positives.

Concordant and discordant pairs are also used in the computational process of Somers' D Statistics. Two pairs (x_i,y_i) and (x_j,y_j) are concordant if the ranks of both the elements agree (i.e) they are concordant if it satisfies the one of the following constraint:

i) $x_i>x_j$ and $y_i>y_j$ (Eqn.1)

ii) $x_i<x_j$ and $y_i<y_j$. (Eqn.2)

In the proposed case, if both the index and the co morbid disease is present or if both of them are absent, that pair of index and co morbid disease is said to be a concordant pair. Else if, the index disease is present and the co morbid disease is absent or vice versa then that pair of index and co morbid disease is said to be discordant pair.

Somers D' can be defined by the formula,

i) Somers D'= P-Q/min(W_r,W_c) (Eqn.3)

where P is the concordant pair and Q is the discordant pair,

$W_r = P+Q+T_r$,

$W_c = P+Q+T_c$,

T_r - number of tied pairs on row only,

T_c - number of tied pairs on column only.

The value of Somers for the index disease with each co morbid disease is found out to find their dependency in order to ensure if the diseases are co morbid or just a coincidence. If it is just a coincidence, we remove the record from the data set.

3.2 Model based Clustering algorithm

Model based clustering algorithms are unsupervised algorithm. The algorithm starts with optimizing the fit between the data and the mathematical model. Here we use a expectation maximization based model for building the initial clusters. Maximum likelihood is used to estimate model parameters. We denote the model parameters by Θ.More generally, the maximum likelihood criterion is to select the parameters Θ that maximize the log-likelihood of generating the data D:

$$\Theta = \arg(\max L(D|\Theta)) \quad \text{(Eqn.4)}$$

where, $L(D|\Theta)$ is the objective function that measures the goodness of the clustering. Given two clustering with the same number of clusters, we prefer the one with higher $L(D|\Theta)$. Choose Θ that maximizes the likelihood of generating a given set of documents. Once value of $\Theta$is found, we can compute an assignment probability $P(d|\varpi;\Theta)$ for each document-cluster pair. This set of assignment probabilities defines a soft clustering. A commonly used algorithm for model-based clustering is the Expectation-Maximization algorithm or EM algorithm . EM clustering is an iterative algorithm that maximizes $L(D|\Theta)$.

3.3 Weighted K-Means Algorithm

Popular k-means algorithm groups data by firstly assigning all data points to the closest clusters, then determining the cluster means. The algorithm repeats these two steps until it has converged. Weighted k-means [12] is used to improve the clustering scalability and to speed up the clustering process.

Weights are assigned to each data entry in the cluster. The process involves initializing the cluster centers followed by assigning the data to the nearest centroid and re-computing the centroid of the clusters after the addition of each entry into the cluster till all data are assigned.

Algorithm Weighted k-means

Input: a set of *n* data points obtained from the density-biased reservoir sampling,and the number of clusters (*K*)

Output: centroids of the *K* clusters

1) Initialize the *K* cluster centers

2) Repeat

2.1) Assign each data point to its nearest cluster center according to the membership function,

$$m(c_j|x_i)=\frac{||x_i-c_j||^{-p-2}}{\sum_{j=1:k}||x_i-c_j||^{-p-2}} \quad \text{(Eqn.5)}$$

2.2) For each center c_j, recompute the cluster center c_j using the current cluster memberships and weights,

$$c_j=\frac{\sum_{i=1:n} m(c_j|x_i)w(x_i)x_i}{\sum_{i=1:n} m(c_j|x_i)w(x_i)} \quad \text{(Eqn.6)}$$

where $w(x_i)$ is a weight associated with each data point.

3) Until there is no reassignment of data points to new cluster centers.

The assignment of weights is based mainly on two factors namely the frequency of occurrence of the disease and the severity of the disease. For example diseases that occur more

frequently or that with high severity is assigned more weights than those that do not.

3.3 Classification Using SVM

The test data are classified among the cluster groups using Support vector machine[7]. SVM uses decision planes which are used to define the co morbidities conditions. A decision planes are used to separates the co morbidity patterns having different class memberships. SVM constructs a hyper plane that separates the two sets so as to minimize the number of misclassified co morbidity patterns. Many variations of SVM are available, but in this paper LIBSVM[11] and LIBLINEAR SVM[13] are employed. A typical LIBSVM first generates a model from the training data set based on logistic regression and distribution estimation. The next part of the classification is to predict information of a testing dataset using the generated model. In order to make classification faster, we use second order SMO for convergence.

A LIBLINEAR SVM can applied on sparse data set, therefore we applied LIBLINEAR SVM on the original dataset without applying Somers d' statistics. A LIBLINEAR SVM uses a linear regression to generate the kernel model.

4. EXPERIMENTAL ANALYSIS AND RESULT

Initially the procedure begins with applying Somers D' statistics to the given datasets. This module is used mainly for the purpose of reducing the size of the dataset. Somers D' statistic value is generated for all the 29 co morbid conditions. These values are appended at the bottom of the corresponding diseases A new row representing the Somers D' value is inserted at the end of the dataset as shown in the figure 2.

998	extreme condition	extreme loss of functio	moderate condition	extreme condition	extreme condition	extreme loss of functi	one half paralysis	maj
999	moderate conditio	no class	extreme condition	moderate condition	no class	extreme loss of functi	full paralysis	min
1000	moderate conditio	minor loss of function	extreme condition	major condition	extreme condition	extreme loss of functi	full paralysis	extr
1001	major condition	extreme loss of functio	extreme condition	extreme condition	major condition	extreme loss of functi	paralysis in one part	maj
1002	-0.003	-0.001	-0.001	0.001	0.006	0.009	0.011	
1003								

Figure 2. Data set with Somers D' statistic value

present	major con	extreme c	extreme l	full paraly	extreme l	extreme l	extreme c	extreme c	moderate	extreme c	extreme c	extre
absent	extreme c	major con	extreme l	paralysis i	major loss	no class	moderate	major con	moderate	extreme c	major con	mode
absent	extreme c	minor con	major loss	full paraly	extreme l	extreme l	major con	moderate	extreme l	major con	minor con	mode
	0.001	0.006	0.009	0.011	0.012	0.012	0.017	0.022	0.029	0.034	0.035	(

Figure 3. Reduced dataset after filtration based on Somers value

Our objective for applying Somers D' statistic value is to reduce the co morbid diseases that has no significant relation or impact on the index cardiovascular disease, the co morbid conditions with Somers D' values less than zero are filtered as they impose negative effect during prediction. Therefore, only the diseases that have a positive impact or correlation to the index disease are considered. Figure 3 shows the dataset with disease column having only positive Somers D' value.

After obtaining the reduced data set, the data set needs to be clustered. A model based clustering is applied to the dataset to generate the initial set of clusters. Figure 4 shows the distribution of the data points or observations or instances across various clusters. The cluster number to which each datapoint belongs is present in its place.

Then weighted k-means algorithm is applied on the initial set of clusters formed using model based clustering. The figure 5 shows the sub clusters formed after applying weighted k-means. Each sub-cluster is assigned a cluster indicator. Assigning of cluster indicators helps in the process of identifying a sub-cluster uniquely. These cluster indicators are then updated in the data set. The clustered dataset after applying cluster indicators needs to be converted into SVM format which an index:value format.

The last step is to check how accurately the co morbidity condition of Cardio Vascular Diseases are predicted using LIBSVM and LIBLINEAR SVM. The LIBSVM predicts the co morbidity condition of Cardio Vascular Diseases with a accuracy of 48 as shown in figure 6, where as LIBLINEAR SVM produced a accuracy rate of 68% as shown in figure 7.

```
In data(afsom1) : data set 'afsom1' not found
> library(mclust)

    __ _____________    __  _______________
   /  |/  / ____/ /   / / / / ___/_  __/
  / /|_/ / /   / /   / / / /\__ \ / /
 / /  / / /___/ /___/ /_/ /___/ // /
/_/  /_/\____/_____/\____//____//_/    version 5.2.2
Type 'citation("mclust")' for citing this R package in publications.
> data(afsom1)
> m<- Mclust(afsom1)
> m$classification
  [1] 1 1 2 2 2 3 3 3 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2
 [38] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
 [75] 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 1
[112] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[149] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[186] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[223] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[260] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[297] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[334] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[371] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[408] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[445] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[482] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[519] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3
[556] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
[593] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
```

Figure 4. Clusters formed after model based clustering

```
> write(out,file="in1.csv",ncolumns=1,sep=" " ,append=FALSE)
> library(wskm)
> data(Book1)
> dm<-data.matrix(Book1,rownames.force=NA)
> km<-ewkm(dm,3)
> km$classification
NULL
> km$cluster
  [1] 2 1 1 1 1 1 3 3 3 3 3 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [38] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 3
 [75] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
[112] 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[149] 2 2 2 2 2 2 2 2 2 1
> km$size
[1] 54 51 53
>
```

Figure 5. Sub clusters formed after applying weighted k means.

```
C:\Users\Ash\Downloads\pro\libsvm-3.22\libsvm-3.22\windows>svm-train.exe ftr.train
.*
optimization finished, #iter = 1261
nu = 0.982644
obj = -374.278205, rho = -0.034121
nSV = 749, nBSV = 368
Total nSV = 749

C:\Users\Ash\Downloads\pro\libsvm-3.22\libsvm-3.22\windows>svm-predict.exe fts.test ftr.train.model finish.out
Accuracy = 100% (749/749) (classification)

C:\Users\Ash\Downloads\pro\libsvm-3.22\libsvm-3.22\windows>svm-train.exe fintrain.train
.*
optimization finished, #iter = 1261
nu = 0.982644
obj = -374.278205, rho = -0.034121
nSV = 749, nBSV = 368
Total nSV = 749

C:\Users\Ash\Downloads\pro\libsvm-3.22\libsvm-3.22\windows>svm-predict.exe fintest.test fintrain.train.model finish1.out
Accuracy = 47.8088% (120/251) (classification)
```

Figure 7. Prediction using LIBSVM.

```
C:\Users\Ash\Downloads\pro\liblinear-2.1\liblinear-2.1\windows>train.exe fintrain.train
.....................................................................................
optimization finished, #iter = 1000

WARNING: reaching max number of iterations
Using -s 2 may be faster (also see FAQ)

Objective value = -0.047109
nSV = 630

C:\Users\Ash\Downloads\pro\liblinear-2.1\liblinear-2.1\windows>predict.exe fintest.test fintrain.train.model finish2.out
Accuracy = 68.1275% (171/251)
```

Figure 8. Prediction using LIBLINEAR SVM.

5. CONCLUSION AND FUTURE WORKS

In this paper, double clustering approach for finding a pattern and clustering of patient's co morbid conditions is proposed. A comparison between two different types of SVMs is made and it is found that LIBLINEAR SVM outperforms LIBSVM. The main draw back in using LIBSVM and LIBLINEAR SVM is both can classify among two classes i.e. cluster indicator 1 is one class and other cluster indicator are in the second class. In future the approach may be extended by applying multi class classifier.

REFERENCES

[1] Van der Wal HH, Van Deursen VM, Van der Meer P, Voors AA, 2017,."Comorbidities in heart failure. Hand Exp Pharma col; 243:35–66.

[2] Ferruci Luigi, Lash Timothy L., Mor Vincent , Wieland Darryl ,William Satariano William, Silliman Rebecca A., 2007, "Methodology, design, and analytic techniques to address measurement of comorbid disease", The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, no. 3 ,pp. 281-285.

[3] Nasreen S, Azam MA, Shehzad K, Naeem U, Ghazanfar MA., 2014, " Frequent pattern mining algorithms for finding associated frequent patterns for data streams: a survey", Procedia Comput Sci. 2014;37:109–116. doi: 10.1016/j.procs.2014.08.019.

[4] Kudyba S, Gregorio T., 2010," Identifying factors that impact patient length of stay metrics for healthcare providers with advanced analytics.", Health Informatics J;16(4):235–245.

[5] Rani KU,2011," Analysis of heart diseases dataset using neural network approach". Int J Data Min Knowl Manag Process ;1(5):1–8.

[6] Kamath C, 2009, "Scientific data mining: a practical perspective", Philadelphia (PA): Society for Industrial and Applied Mathematics.

[7] Son YJ, Kim HG, Kim EH, Choi S, Lee SK,2010," Application of support vector machine for prediction of medication adherence in heart failure patients". Health c Inform Res ,16(4):253–259.

[8] Kajabadi A, Saraee MH, Asgari S. 2009, "Data mining cardiovascular risk factors", Proceedings of International Conference on Application of Information and Communication Technologies; Oct 14-16; Baku, Azerbaijan. pp. 1-5.

[9] Suryawanshi RD, Thakore DM. 2012, "Classification techniques of data mining to identify class of the text with fuzzy logic", Proceedings of 2012 International Conference on Information and Computer Applications; Feb 17-18; Hong Kong. pp. 263-267.

[10] Sitar-Taut DA, Sitar-Taut AV,2010, "Overview on how data mining tools may support cardiovascular disease prediction" J Appl Comput Sci;4(8):57–62.

[11] Chih-Chung Chang and Chih-Jen Lin, 2011,"LIBSVM : A Library for Support Vector Machines", ACM Transactions on Intelligent Systems and Technology,pp. 2:27:1–27:27.

[12] Kittisak Kerdprasop, Nittaya Kerdprasop, and PairoteSattayatham ,2005, "Weighted K-Means for Density-Biased Clustering", A Min Tjoa and J. Trujillo (Eds.):DaWaK LNCS 3589, pp. 488-497.

[13] Helleputte, T. & Gramme, P., 2015, "Package 'LiblineaR'.Available online https://cran.rproject.org/web/packages/LiblineaR/LiblineaR.pdf

[14] Benjamini, Yoav, and Hochberg yosef 1995, "Controlling the false discovery rate: a practical and powerful approach to multiple testing", Journal of the Royal Statistical Society, Series B (Methodological) pp. 289-300.

[15] S.K. Ng (2015), "A two-way clustering framework to identify disparities in multimorbidity patterns of mental and physical health conditions among Australians", Statistics in Medicine.