CLUSTERING DICHOTOMOUS DATA FOR HEALTH CARE

D. Ashok Kumar[1] and M.C. Loraine Charlet Annie[2]

[1]Department of Computer Science, Government Arts College, Trichy-620022
akudaiyar@yahoo.com
[2]Department of Computer Science, Government Arts College, Trichy-620022
lorainecharlet@gmail.com

ABSTRACT

Dichotomous data is a type of categorical data, which is binary with categories zero and one. Health care data is one of the heavily used categorical data. Binary data are the simplest form of data used for heath care databases in which close ended questions can be used; it is very efficient based on computational efficiency and memory capacity to represent categorical type data. Clustering health care or medical data is very tedious due to its complex data representation models, high dimensionality and data sparsity. In this paper, clustering is performed after transforming the dichotomous data into real by wiener transformation. The proposed algorithm can be usable for determining the correlation of the health disorders and symptoms observed in large medical and health binary databases. Computational results show that the clustering based on Wiener transformation is very efficient in terms of objectivity and subjectivity.

KEYWORDS

Dichotomous Data, Clustering

1. INTRODUCTION

Technological advancements in the form of computer-based patient records software and personal computer hardware are making the collection of and access to health care data more manageable which demands in-depth analysis. Business, financial, and scientific data can be easily modeled, transformed and applied formulas in contrast to medical data, whose underlying structure is poorly classified in mathematical terms. The health care or medical datasets usually are very large, complex, and heterogeneous and vary in quality. The characteristics of the data may not be optimal for mining or analytic processing. The challenge here is to convert the data into appropriate form for clustering. Medical data is fragmented and distributed so that the confidence that can be placed on the data mining results is a great challenge. Hence appropriate data mining techniques are required here for processing the raw and sparse categorical medical data; clustering is an important data mining technique. Clustering is the process of grouping similar objects in such a way that two objects from the same cluster are more similar than two objects from different clusters. In medical field, clustering technique can be used to group the patients into different categories like normal, abnormal, intensively cared.

Categorical variables are characterized by values which are categories. Two main types of these variables can be distinguished: dichotomous, for which there only are two categories, and multi-categorical. For dichotomous data, both categories have the same importance. A common

DOI : 10.5121/ijist.2012.2203

example for dichotomous data is (male, female). For multi-categorical, more than two categories can be represented. Dichotomous variables are often coded by the values zero and one. Binary data are the simplest form of data used in information systems for databases; it is very efficient based on computational efficiency and memory capacity to represent categorical type data.

For better understanding of the patient's case, close ended questions i.e yes/no questions can be used for Patients' profiles. Initially patients will be examined and medical experts' makes the patients' profile by answering the yes or no questions. Relevant attributes can be used for these types of profiles. The binary database used here has value as 1 for yes and 0 for no answers. For segregating the patients, obtained binary data from the patents' profiles need to be grouped it means clustered. Hence clustering is the best data mining technique for this application.

The characteristics of clinical data, including issues of data availability and complex representation models, can make data mining applications challenging. Data mining is a step in the Knowledge Discovery in Databases where a discovery-driven data analysis technique is used for identifying patterns and relationships in datasets. The question becomes how to bridge the two fields, data mining and medical science, for an efficient and successful mining of medical data. The eventual goal of this data mining effort is to identify factors that will improve the quality and cost effectiveness of patient care. Evaluation of stored clinical data may lead to discovery of trends and patterns hidden within the data that could significantly enhance our understanding of disease progression and management. Due to the high volume of the medical databases, current data mining tools requires grouping of data from the medical database [13], [14]. The right mining technique, for the right dataset is to select the best data partitioning technique. Two general steps of the data mining techniques include data preparation and knowledge discovery. The data preparation stage involved two stages: data cleaning and transformation. The data cleaning step eliminates inconsistent data like missing values data, incomplete data. The transformation aims at converting the data fields to numeric fields like 0 for "no" and 1 for "yes" answers. Knowledge discovery in databases (KDD) is defined as the nontrivial extraction of implicit, previously unknown, and potentially useful information from data. Here binary data clustering is used to group the patients using the profiles created with the close-ended questions.
Clustering is the problem of identifying the distribution of patterns and intrinsic correlations in large binary data sets by partitioning the data points into similarity classes. A common goal of the medical data mining is the detection of some kind of correlation. Similarity measures are usually used for clustering the Dichotomous variables. Extensive research [1, 2 and 17] is going on clustering large binary data sets due to its wide assorted applications. The K-means clustering algorithm remains one of the most popular and Standard clustering algorithms used in practice. Variants of K-means algorithm are On-line K-means, Scalable K-means and Incremental K-means [3]. The main reasons for its usage are that it is simple to implement and it is fairly efficient. The popular k-means algorithm is given in the following section.

2. K-MEANS ALGORITHM

The K-means is a centroid based Partitional clustering algorithm [4], which is detailed in Algorithm1. This K-means algorithm has been discovered by several researchers across different disciplines, most notably [5, 6, 7, 8 and 9]. The popular heuristics for solving the k-means problem is based on a simple iterative scheme for finding a locally minimal solution. K-means algorithm works optimally with categorical and numeric data so that this is the best for binary data clustering. It is simple and fairly fast [10], results are easy to interpret and it can work under a variety of conditions hence it stand as the standard algorithm for clustering. The fundamental idea is to find K average or mean values, about which the data can be clustered. The k-means algorithm is a simple iterative method to partition a given dataset into a user specified number of clusters K. K-means is initialized from some random or approximate solution. The step one randomly picks up K feature vectors as centers for K clusters. The step 2 finds similar feature vectors for each cluster by Euclidean distance. Updating the cluster centroid helps to make more similar and clear cluster. The K-means algorithm divides the feature vectors into K clusters by minimizing the total within the class sum of squares at step 3 and then it halts the process.

Algorithm 1: K-means Algorithm for clustering binary data

The K-means algorithm is build upon the following operations:

Step 1: Choose initial cluster centroids $Z_1, Z_2, \ldots, Z_K$ randomly from the p points

$X_1, X_2, \ldots X_p$, $X_i \in R^q$

where q is the number of features/attributes

Step 2: Assign point X_i, $i = 1, 2, \ldots, p$ to cluster C_j, $j = 1,2,\ldots,K$

if and only if $\| X_i - Z_j \| < \| X_i - Z_t \|$, $t = 1, 2,\ldots,K$. and $j \neq t$.

Ties are resolved arbitrarily.

Step 3: Compute the new cluster centroids $Z_1^*, Z_2^*, \ldots, Z_K^*$ as follows:

$Z_i^* = \frac{1}{l_j}\sum_{X_j \in C_j} X_i$ where $i = 1, 2,\ldots,K$, and

l_j =Number of points in C_j.

Step 4: If $Z_i^* = Z_i$, $i = 1, 2,\ldots, K$ then terminate. Otherwise $Z_i \leftarrow Z_i^*$ and go to step 2.

Except the first step, the other three steps are repeatedly performed in the algorithm until the algorithm converges. Note that in case the process does not terminate normally at Step 4, then it is executed for a maximum number of iterations. The k-means algorithm converges when the assignments no longer change. The number of iterations required for convergence varies and may depend on p, this algorithm is linear based on the dataset size.

3. CLUSTERING METRICS

Clustering metrics plays an important role in obtaining good clusters. Clustering metrics assess similarity between the components of a vector, which in general can be formalized as

$$d_{ij} = \alpha \sum_{l=1}^{p} f(x_l, y_l) \quad (1)$$

where α is a coefficient depending on vectors x_l and y_l, and varying across similarity measures. The function f may represent the sum, difference, probability, or some other function applied to its arguments. Manhattan distance, Euclidean distance, Hamming are common functions. The selection of a distance measure plays an important role in partitioning the data set.

Depending on the type of attributes like numeric, continuous, categorical used similarity measures varies. Consider a m-dimensional Euclidean space, the distance between any two points, $x = [x_1, x_2,\ldots,x_p]$ and $y = [y_1, y_2,\ldots,y_p]$ is given as Common distance/ Squared Euclidean Distance in equation (1) i.e.L_2 norm which is represented as Sq.Eucli.

$$\sqrt{\sum_{i=1}^{p}(x_i - y_i)^2} \qquad (2)$$

and the Manhattan distance /Cityblock Distance is given in equation (3) i.e L_1 norm which is the sum of absolute differences, represented as CB

$$\sum_{i=1}^{p}|x_i - y_i| \qquad (3)$$

3.1. Hamming Distance

The most heavily used measure for binary data clustering is the Hamming distance. Usually, Binary data are represented as vectors and Vectors are represented as bit sequences. Hamming distance counts only exact matches between bit sequences. The Hamming distance between two vectors is the number of coefficients in which they differ. It is simply defined as the number of bits that are different between two bit vectors. Hamming distance is an easy-to-define metric. For binary strings a and b the Hamming distance is equal to the number of ones in a XOR b. Hamming distance between x and y objects is given in equation (4).

$$d_{xy} = q+r \qquad (4)$$

where q is the number of variables with value 1 for the x^{th} object and 0 for the y^{th} object. r is the number of variables with value 0 for the x^{th} object and 1 for the y^{th} object.

Brief introduction is obtained up to now about K-means algorithm & its performance measures. The new proposed idea to improve the Binary data clustering is done through Wiener Transformation which is explained in the following section.

4. WIENER TRANSFORMATION APPROACH

A new binary data clustering technique based on Wiener Transformation is proposed here. Wiener Transform is efficient on large linear spaces. Usually Wiener Transformation is used in Image Restoration for noise-removal filtering [12]. In this paper, the binary data is transformed into real data using the Wiener Transformation, which is a statistical transformation. The approach is based on a stochastic framework. The transformed data is clustered using the K-means algorithm. Its main advantage is the short computational time it takes to find a solution so that the clustered data is very efficient compared to normal binary data clustering.

The input for wiener transformation is stationary with known autocorrelation. It is a causal transformation. It is based upon linear estimation of statistics [12]. The Wiener transformation is optimal in terms of the mean square error. The Wiener filter is a filter proposed by Norbert Wiener. The syntax for Wiener filter is Y = wiener2 (X, [p q], noise) for two-dimensional image which is normally used for image restoration. Wiener2 function is used because input is a 2-dimensional matrix. This equation is used here for data mining task.

The input X is a two-dimensional matrix and the output matrix Y is of the same size. Wiener2 uses an element-wise adaptive Wiener method based on statistics estimated from a local neighbourhood of each element. Usually binary data are represented as vectors. Wiener estimates the local mean μ in equation (5) and variance σ^2 in equation (6) for each element on every vector of the binary input matrix using the equations given below.

$$\mu = \frac{1}{pq}\sum_{n_1,n_2 \in \eta} X(n_1,n_2) \tag{5}$$

$$\sigma^2 = \frac{1}{pq}\sum_{n_1,n_2 \in \eta} (X^2(\ n_1,n_2) - \mu) \tag{6}$$

where η is the p-by-q local neighbourhood of each element in the input matrix X. Here input is considered on vector basis and hence p, q is considered as default value p=q=3. wiener2 then creates a element-wise Wiener filter Y for every vector using the above said mean and variance are given in equation (7).

$$Y(n_{1,},n_{2)} = \mu + \frac{\sigma^2 - \upsilon^2}{\sigma^2}(X(n_{1,},n_2) - \mu) \tag{7}$$

where υ^2 is the average of all the local estimated variances. Here the binary data is pre-processed by transforming into real data using the Wiener Transformation for a vector. The transformed data is clustered using the above mentioned K-means algorithm.

5. RESULTS AND DISCUSSIONS

Dichotomous data is transformed to real domain by the linear Wiener transformation. Then K-means clustering algorithm is executed on the transformed data. Wiener transformation is used in this clustering approach because it based on neighbourhood elements since clustering means finding similarity. The qualitative evaluation of quantitative results are done using the best metrics for binary data clustering like Inter-cluster distance, Intra-cluster distance, Sensitivity and Specificity.

Lens dataset [18] is a database for fitting contact lenses. There are 3 Classes; class1: the patient should be fitted with hard contact lenses, class2 : the patient should be fitted with soft contact lenses, class 1 : the patient should not be fitted with contact lenses. Number of Instances are 24 and Number of Attributes are 4, all are nominal values.

5.1. Inter-cluster distance

The Inter-cluster distance μ is calculated for K-means clustering with distances Squared Euclidean Distance, City Block Distance, Euclidean distance and Hamming Distance for the Actual Dataset (AD) and the Wiener Transformed Dataset (WTD).

Inter-cluster distance means the distances between different clusters, and it should be maximized i.e distance between their centroids. The inter-cluster distance μ for K clusters $C_1, C_2, \ldots C_K$ with centroids Z_i, i=1...K is given in equation (8).

$$\mu(C_1, C_2, \ldots, C_K) = \sum_{i=1}^{K}\sum_{j=i+1}^{K} |Z_i - Z_j\ | \tag{8}$$

Table 1 : Inter - Cluster distance for Lens binary dataset

Distance Measure	μ_1		μ_2		μ_3		μ_4		μ_5		Average μ	
	AD	WTD	AD	WTD	AD	WTD	AD	WTD	AD	WTD	AD	WTD
Sq.Eucl	21.72	40.07	21.72	28.85	34.25	30.85	34.25	24.73	15.56	14.77	25.5	27.85
CB	29.3	43.37	32.07	33.29	17.27	17.64	18.52	18.72	17.27	25.81	22.89	27.77
Eucli	27.96	28.84	36.25	70.87	29.10	39.52	21.72	21.73	27.11	40.36	28.426	40.26
HD	30.27	55.11	17.27	21.35	21.72	21.73	38.65	26.24	19.08	38.65	25.4	32.62

From Table 1, it is observed that the Inter-cluster distance μ_i i= 1, 2, .., 5 for various distance measures in WTD outperforms the AD. Five executions are done here because K-means clustering varies based on the initialization of centroids. Table 1 shows that the Inter-distance (μ) for Squared-Euclidean distance is 26.25074 for WTD but for AD is 26.29906. For City Block distance μ is 27.76638 for WTD but for AD is 22.88762 For Euclidean distance μ is 40.2646 for WTD but for AD is 28.42598. For Hamming distance μ is 32.61574 for WTD but for AD is 25.39932. Total average inter-distance for AD is 25.553 and 32.124 for WTD that means 6.57 improvements for μ on a average.On an average Wiener transformed clusters are best compared to normal clusters.

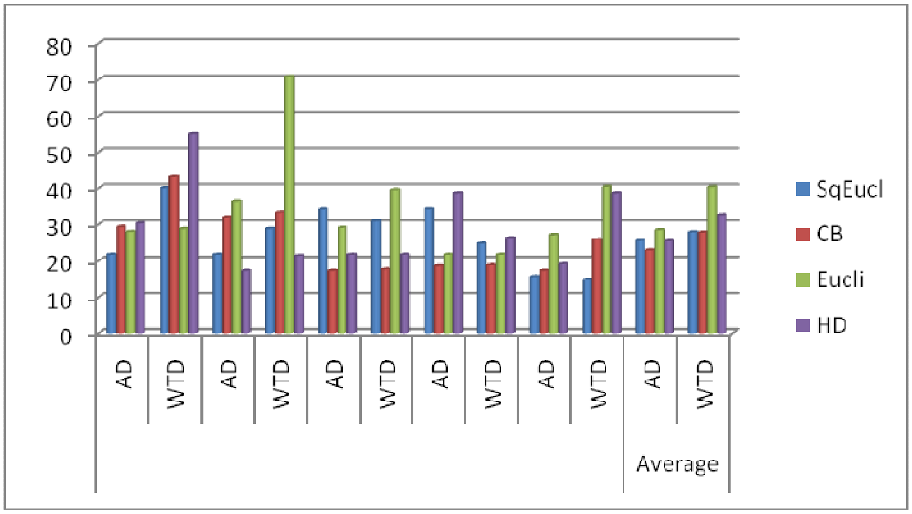

Figure1. Inter – cluster distance analysis for all distance measures

Fig 1 the inter-cluster distance (IED) measures for the clusters of K-means algorithm with actual binary data and wiener transformed data as input using Squared Euclidean Distance, City Block Distance and Hamming Distance during clustering. Five different executions are done because clusters vary based on centroid initialization. Final column depicts the average improvement.

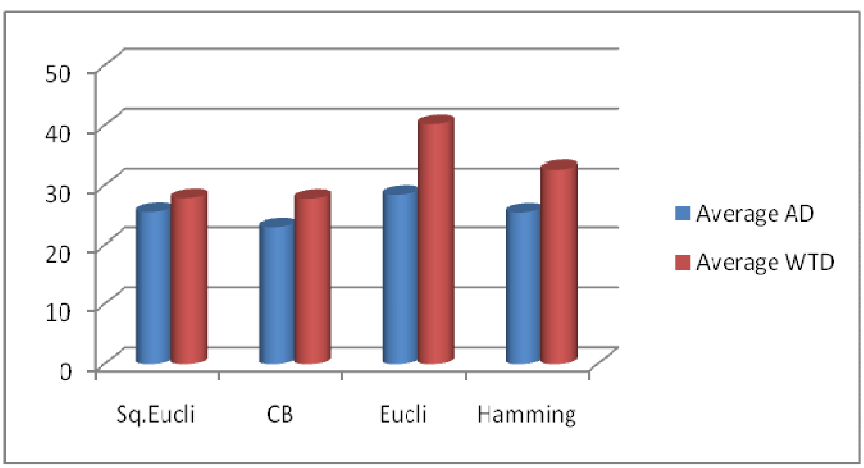

Figure 2. Inter-distance Average performance analysis of various distance measures

Fig 2 depicts the overall average performance measures for K-means clustering for AD and WTD with Squared Euclidean Distance, City Block Distance and Hamming Distance for clustering. City Block Distance has higher efficiency.

5.2. Intra-cluster distance

Intra-cluster distance is the sum of distances between objects in the same cluster, and it should be minimized. The intra-cluster distance υ for K clusters C_1, C_2, .. C_K centroid Z_i, i=1..K is given in equation (9).

$$\upsilon(C_1, C_2, \dots, C_K) = \sum_{i=1}^{K} \sum_{X_j \in C_i} |X_j - Z_i| \qquad (9)$$

The Intra-cluster distance υ is calculated for the Squared Euclidean Distance, City Block Distance and Hamming Distance for the Actual binary dataset and the Wiener transformed dataset. The results are tabulated in Table 2 for lens dataset.

Table 2. Intra-Cluster Distance for Lens binary dataset

Distance Measure	υ_1		υ_2		υ_3		υ_4		υ_5		Average υ	
	AD	**WTD**	**AD**	**WTD**	**AD**	**WTD**	**AD**	**WTD**	**AD**	**WTD**	**AD**	**WTD**
SqEucl	0.97	0.96	1.1	1.01	1.1	1.01	1.10	0.91	1.10	0.91	1.08	0.96
CB	0.91	1.01	1.01	1.05	1.11	1.05	1.25	1.01	1.1	0.97	1.08	1.01
Eucli	1.14	1.05	1.10	1.01	0.93	0.91	1.10	1.09	0.9	1.01	1.03	1.01
HD	0.99	0.96	1.1	1.07	1.1	1.01	1.08	0.97	1.02	1.05	1.06	1.01

The Table 2 shows that the Intra-distance (υ) for Squared-Euclidean distance is 0.9611 for Wiener Transformed Dataset (WTD) but for Actual Dataset (AD) is 1.075. For City Block distance υ is 1.01704 for WTD but for AD is 1.07706. For Euclidean distance υ is 1.01434 for WTD but for AD is 1.03284. For Hamming distance υ is 1.00996 for WTD but for AD is 1.05974. Total average inter-distance for AD is 1.06116 and 1.00061 for WTD that means 0.061 improvements for υ on a average.

Fig 3 depicts the intra-cluster distance (IAD) measures for the clusters of K-means algorithm with actual binary data and wiener transformed data as input using Squared Euclidean Distance, City Block Distance, Euclidean distance and Hamming Distance during clustering. Wiener transformation clustering is best compared to normal binary data clustering using K-means algorithm.

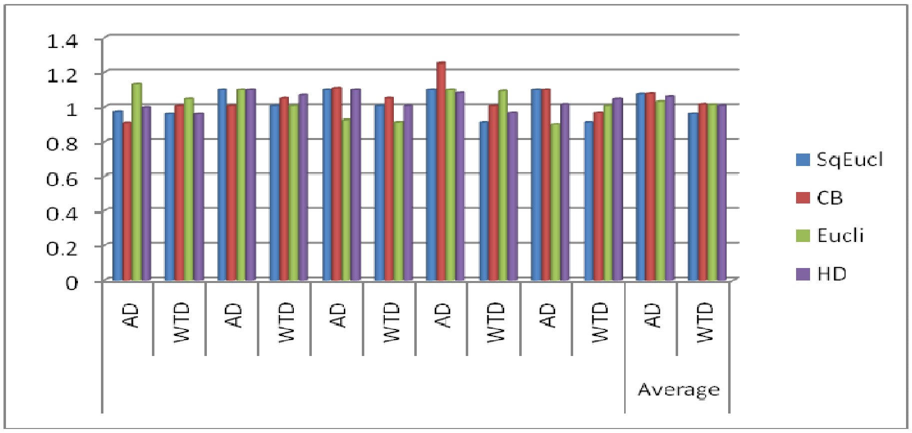

Figure 3. Intra cluster distance analysis for various distance Measures

Fig 4 depicts the average intra-cluster distance measures of Squared Euclidean Distance, City Block Distance, Euclidean Distance and Hamming Distance. For Squared Euclidean Distance and City Block Distance WTD has better performance.

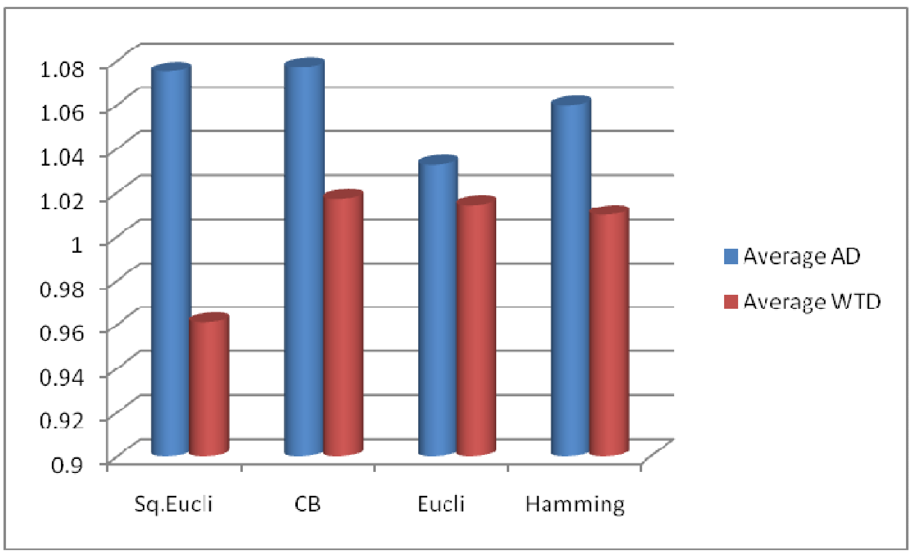

Figure 4. Intra distance average performance of various distance measures

5.3. Statistical Measures

Sensitivity and Specificity are statistical measures used in medical field; same measures are used here for evaluating wiener transformed binary data clustering. Sensitivity measures the ability of test to be positive when the condition is actually present, or how many of the positive test examples are recognized. A sensitivity of 100% means that the test recognizes all actual positives.

$$Sensitivity\ (SE) = \left(\frac{TP}{TP+FN}\right) * 100 \qquad (10)$$

Specificity measures the ability of a test to be negative when the condition is actually present, or how many of the negative test examples are excluded. A specificity of 100% means that the test recognizes all actual negatives.

$$Specificity\ (SP) = \left(\frac{TN}{FP+TN}\right) * 100 \tag{11}$$

Predictive accuracy gives an overall evaluation.

$$Predictive\ Accuracy(PA) = \left(\frac{TP+TN}{TP+TN+FP+FN}\right) * 100 \tag{12}$$

Where,

TP - Number of True Positives
TN - Number of True Negatives
FP - Number of False Positives
FN - Number of False Negatives

Table 3. Sensitivity, Specificity and Positive Predictive Value

Distance Measure	**Sensitivity (%)**		**Specificity (%)**		**Positive (%) Predictive Value**	
	AD	**WTD**	**AD**	**WTD**	**AD**	**WTD**
Sq.Eucli	96	97	93	94	93	93
CB	92	94	97	97	97	98
Eucli	92	94	97	97	97	98
Hamming	92	93	99	98	99	99

Sensitivity and Specificity for WTD are always higher than the AD which is observed from Table 3. From the lens dataset 8 persons are in each group for hard lens, soft lens and no lens category. Consider Sq.Eucli distance, K-means clustering with AD finds 5 persons, 4 persons and 4 persons correctly for hard lens, soft lens and no lens category respectively; but WTD clustering groups 6 persons, 5 persons and 6 persons correctly for hard lens, soft lens and no lens category respectively. Hence Sensitivity is increased 97% for Sq.Eucli distance. Consider CB distance, K-means clustering with AD finds 4 persons, 6 persons and 5 persons correctly for hard lens, soft lens and no lens category respectively; but WTD clustering groups 6 persons, 6 persons and 6 persons correctly for hard lens, soft lens and no lens category respectively. Hence Sensitivity is increased as 94% for CB distance. Consider Eucli distance, K-means clustering with AD finds 5 persons, 6 persons and 4 persons correctly for hard lens, soft lens and no lens category respectively; but WTD clustering groups 6 persons, 6 persons and 6 persons correctly for hard lens, soft lens and no lens category respectively. Hence PA value is increased by 1% as 98% for Eucli distance. Consider Hamming distance, K-means clustering with AD finds 5 persons, 7 persons and 6 persons correctly for hard lens, soft lens and no lens category respectively; but WTD clustering groups 6 persons, 7 persons and 6 persons correctly for hard lens, soft lens and no lens category respectively. Hence Sensitivity is increased by 1% as 93% for Eucli distance. It means Wiener Transformation clustering for binary data is very effective compared to AD K-means clustering.

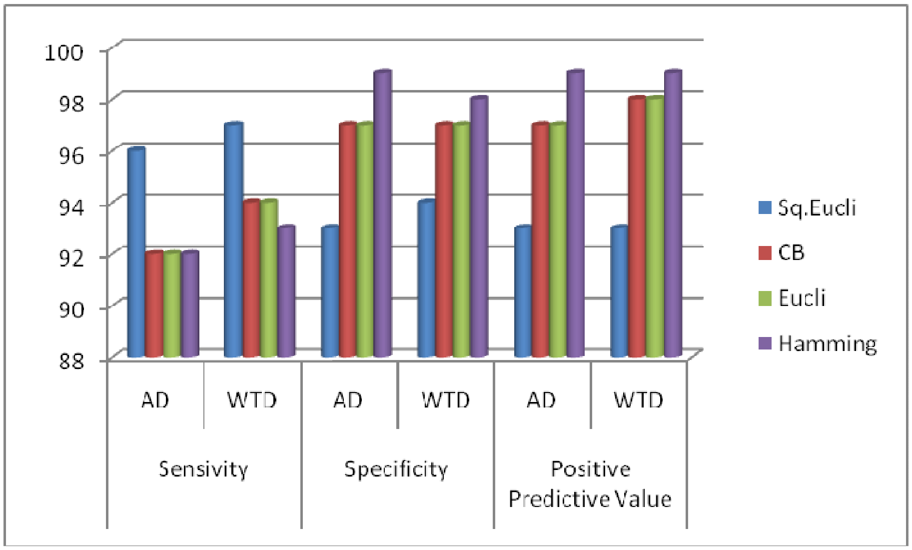

Figure 5. Average Performance of Sensitivity and Specificity with various distance measures

Fig 5 depicts the average performance of Sensitivity, Specificity and Predictive Accuracy Value (PA) Value for the Actual binary data and the wiener transformed binary data clustering with Squared Euclidean Distance, City Block Distance and Hamming Distance measures.

3. CONCLUSIONS

A novel clustering approach has been proposed for dichotomous health care data. Here binary health care data is transformed into real domain using the linear Wiener Transformation. Then the wiener transformed data is clustered using the standard K-means algorithm. The proposed binary data clustering technique empirically works well for finding good clusters since K-means algorithm works fine in real domain and also the wiener transformation clustering is based on neighbourhood elements. Statistical measures such as Sensitivity, Specificity and positive predictive value which are similar to healthcare domain are optimized in its average performances. From the clustering optimality measures it is observed that clustering dichotomous data using the Wiener transformation is more efficient than normal clustering with K-means algorithm.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their significant and constructive critiques and suggestions, which improved the paper very much.

REFERENCES

[1] C. Aggarwal and P. Yu , "Finding generalized projected clusters in dimensional spaces", In ACM SIGMOD Conference, pp. 70-81, 2000.

[2] Tao Li, "A general model for clustering binary data", In Proc. of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pp.188, 2005.

[3] Carlos Ordonez, "Clustering Binary Data Streams with K-means", In 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 12-19, 2003.

[4] J. McQueen, "Some methods for classification and analysis of multivariate observations", In Proc. 5th Berkeley Symp Mathematics, statistics and probability, pp. 281-296, 1967.

[5] S.P. Lloyd, "Least squares quantization in PCM", IEEE Transaction Information Theory, Special issue on Quantization, pp. 129-137, 1982.

[6] E. Forgy, "Cluster Analysis of multivariate data: Efficiency vs. interpretability of classification", Biometrics, pp. 768, 1965.

[7] H.P. Friedman and J. Rubin, "On some invariant criteria for grouping data", Journal of American Statistical Association, 62, pp.1159-1178, 1967.

[8] J. McQueen, "Some methods for classification and analysis of multivariate observations", In Proc. of 5th Berkeley Symp Mathematics, statistics and probability, pp. 281-296, 1967.

[9] T. Kanungo, D.M. Mount, N. Netanyahu, C. Piatko, R. Silverman and A.Y. Wu A, "A local search approximation algorithm for k-means clustering", Computational Geometry: Theory and Applications, pp.89-112, 2004.

[10] P. Bradley, U. Fayyad and C. Reina, "Scaling Clustering Algorithms to Large Databases", In Proc. of the 4th ACM SIGKDD, pp. 9-15, 1998.

[11] M. Sultan and D.A. Wigle, "Binary tree-structures vector quantization approach to clustering and visualizing microarray data", Bioinformatics, Vol.18, 2002.

[12] Frederic Garcia Becerro, "Report on Wiener filtering", Image Analysis, Vibot, 2007. [Online]http://eia.udg.edu/~fgarciab/

[13] Krzysztof J. Cios , William Moore, Uniqueness of Medical Data Mining, pp.109-143, 2002.

[14] Han J and Kamber M., "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, San Francisco, CA, pp.32, 2001.

[15] Kurgan, L.A., Cios, K.J., Tadeusiewicz, R., Ogiela, M. and Goodenday, L.S. "Knowledge Discovery Approach to Automated Cardiac SPECT Diagnosis" Artificial Intelligence in Medicine, vol. 23:2, pp. 149-169, 2001.

[16] Cios K. J. and Kurgan L. "Hybrid Inductive Machine Learning: An Overview of CLIP Algorithms", In: Jain L.C., and Kacprzyk J. (Eds). New Learning Paradigms in Soft Computing, Physica-Verlag Springer, 2001.

[17] Gower, J.C. and Legendre, "P. Metric and Euclidean properties of dissimilarity coefficients". Journal of Classification, 3, pp.5-48, 1986.

[18] Witten, I. H. and MacDonald, B. A. "Using concept learning for knowledge acquisition. International Journal of Man-Machine Studies," pp. 349-370, 1988.