International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

A Novel Approach to Predict Kidney Detection Using Support Vector Machine

Natasha Sharma[1], Sahil Dalwal[2]

[1]Department of Computer, Science Bells Institute of Management and Technology Shimla, India
natasha.sharma1003[at]gmail.com
[2]Department of Computer Science Bells Institute of Management and Technology Shimla, India
sahildalwal100[at]gmail.com

Abstract: ***Data Mining is a striking tool for obtaining appreciated information from the huge quantity of available information that can be utilized further for taking the right judgments. Numerous approaches are presented for emerging cost- effective results from the potential data. Mining of data by applying the conditioning rule has been predictably utilized with the objective of the reveling rules in the medical applications. Recognition of different diseases such as kidney, diabetes, and heart attack etc. from huge number of estimates and proofs is an area of great attention for investigators which are not free from false assumptions and unpredictable outcomes. Hence there is terrific requirement to use valuable output resulting from information of patients gathered in our data storehouse. Support Vector Machines, one of the latest approaches for pattern classification, has been widely utilized in large application areas. The objective of this research is to simultaneously optimize parameters and feature subset selection without degrading SVM classification accuracy. We introduce a generic algorithm approach for feature selection named as "H-SVD" and parameters optimization to solve these kinds of problems. In this paper we design an algorithm that enhances the accuracy of diseases prediction system.***

Keywords: SVM, ML

1. Introduction

Data mining is the technique to mine content form huge information series that results in antedating information using process, such as, association, classification etc. In social insurance organization Data mining has exposed wide significance for instance in describing idea treatment strategy foreseeing an alignment hazards considers and searching patient consideration productive cost.

With generating opportunities for reducing or even reversing stride of disease, early identification of persons who are at a greater developing risk a disease is of important clinical value [1].Specifically, 80% of cases of heart and diabetes disease, and 40% of the cancer disease can be preserve defectively at an initial phase [2], checking the requirement for exclusive medical process because of problem arising through normal care of ill patients. In 2012, approximately 76% of Medicare expenditure was on patients, including just 17% of Medicare population, with five or more diseases: Heart disease, Kidney disease, diabetes, renal disease, and lung disease [3].

One kind of a classification algorithm is utilized for prediction and diagnosis of diseases is Support vector machine on the different dataset to diagnose different diseases [10]. To increase the accuracy analysis, it had obtainable to be all more a projecting problem to generate reliable and strong therapeutic choice process. Basically, medicinal analysis is entangled process; consequently the methodology for empathetic this problem is to produce a clever framework, for instance Support Vector Machine.

In this paper we cover the following areas: Section II of the paper represent Literature survey in this field. Section III include proposed methodology and last section illustrate experimental results of our proposed technique.

2. Literature Survey

Review of literature was accompanied in order to attain information of earlier studies. Many articles have been executed utilizing distinct data mining methods, such as, SVM, ANN, NB, DT and Logistic Regression. These methods have been used to produce system in human service research [4].In newest survey controlled; diseases can be managed sufficiently on the off chance that it is examined at an initial phase. For classification, prediction and diagnose of diseases number of data mining algorithms has been projected. For this, data preprocessing should be done. It is a method that includes transforming raw data into comprehensible sign. This supports to fill lost values among data. By examining data utilizing values, it is conceivable for a professional to discover values that are unexpected and inaccurate [6] [7]

Table I: Previous survey on different diseases

Authors	Technique	Diseases	Objectives
Dr. S. Vijayarani et al. [26]	SVM and Naïve bayes	Liver Disease	The main objective of this researchwork is to predict liver diseases using classification algorithms
Mythili T. et al. [27]	Support vector machine	Heart disease	The main objective of this

	(SVM), logistic regression, decision trees, rule based approach		paper is to compare the accuracies of applying rules to the individual results of support vector machine, decision trees, and logistic regression on the Cleveland Heart Disease Database
V. Anuja Kumari et al. [28]	Support Vector Machine	Diabetes disease	The main objective of this paper is to predict the accuracy, sensitivity, and specificity of the SVM.
Mrs. S. Radhimeenaksahi [29]	Support Vector Machine and Artificial Neural Network	Heart disease	The objective of this paper is to break down the use of AI devices for order and expectation of heart illness.
Haiyan Yang et al. [30]	Support vector machine (SVM)	Heart Disease	This objective of the paper is to use chaotic PSO's support vector machine algorithm for the diagnosis of heart disease
Nahla H. Barakat et al.[31]	Support vector machine (SVM)	Diabetes disease	The main motive of the paper is to use SVM for the diagnosis of diabetes. In particular, they use an additional explanation module, which turns the “black box” model of an SVM into an intelligible representation of the SVM’s diagnostic (classification) decision.
Mehmet Fatih Akay [32]	Support vector machines	Breast Cancer disease	The motive of the paper is to combine SVM-based method with feature selection to find breast cancer
Assistant Professor et al. [33]	Support vector machines and Artificial Neural Network	Kidney Disease	The objective of this research work is to predict kidney diseases by using Support Vector Machine (SVM) and Artificial Neural Network (ANN).

3. Proposed Scheme

The overall process of the proposed system contains four sections such as Data preprocessing, feature election, classification and performance evaluation.

A. Data Preprocessing

Data Preprocessing is a technique of data mining to convert raw data into clear design. As we know that original data is always unpredictable, is likely to contain various problems, inadequate, and deficient in certain behaviors. By utilizing some algorithms this data problems can be filled.

B. Feature Selection

To organize and to minimize noise dimension a subset of real variables is exposed which is known as feature selection. The main motive of feature selection is to choose a subset of input variable by extracting features with small analytical information. It changes high-dimensional to low- dimensional data. The input data is changes into a compact representation features set if algorithm input is too large and is supposed to be tremendously redundant. Evading insignificant or redundant features is necessary since that they may have a bad influence on classifier accuracy. Selection of supreme important features subset is an optimization issue. In our research, feature selection is done using H- SVD and Anova which are explained in next section.

C. Classification

To assigns items in a collection to target classes a data mining function is used named as Classification. The motive of classification is to exactly calculate target class for individual instance in data. Testing of Classification models is performed by comparing projected values to recognized target values. For a classification project historical data is naturally categorized into two sets: one for constructing model (training) and other for testing model. For the classification purpose we have SVM algorithms. Classifiers are trained to predict any disease with in no time. In medical diagnosis usage of classifier schemes is growing progressively. In diagnosis there is no uncertainty that data evaluation taken from decisions and patient of professionals are supreme significant factors [5].The results are shown in the below tables. Table 1 shows the performance of different classifiers. It is well recognized in data mining that the following measures provide more informative evaluation of classifier performance when dealing with class-imbalanced data: Accuracy, recall, precision (prec.), and F- measures, [5], which are defined as

Accuracy = (TP+TN)/(TP+FP+TN+FN)
Recall = TP/(TP+FN),
Precision = TP/(TP+FP),
F-measure = (2*recall*precision)/(recall+precision),

In the above definitions, TP is the number of positive cases classified correctly, FN is the number of positive cases classified as negative, FP is the number of negative cases classified as positive, and TN is the number of negative cases classified correctly.

4. Methodology

Many techniques are adapted for predicting diseases. In this proposed work SVM is used to predict the risk of different diseases. The rest of the paper described the dataset and methodology used for diseases prediction.

In computer-based disease diagnosis approaches, in which data is attained from some other sources and is calculated by computer founded applications. Computers have usually been used to build knowledge based clinical decision support systems which used the knowledge from medical experts, and transferring this knowledge into computer algorithms was done manually. This process is time consuming and really depends on the medical expert's opinion, which may be subjective. To handle this problem, machine learning techniques have been developed to gain knowledge automatically from examples or raw data. Medical diagnosis is a significant but complex job that should be accomplished efficiently and its automation would be very useful.

A. Data Set

Data used for proposed system is obtained from UCI (University of California, Irvine C.A) Centre for machine learning and intelligent systems. For the detection of different diseases UCI repository provide a lot of attributes. In our research, we assume different number of the attributes for different diseases. Such as, we take 24 attributes for kidney disease, 6 attributes for Heart and Diabetes and 11 attributes for Brest Cleaning and filtering of the data set is done to remove duplicate records, normalize the values, accounting for missing data and removing irrelevant data items.

For more details of the number of attributes used in research, refer appendix of the paper. B. H-SVD (Singular vector decomposition for Diseases data set) SVD (Singular value decomposition) is the technique that we apply for the extraction of features. To get better results as compare to existing SVD, we have applied improved version of SVD. SVD is a result of linear algebra. SVD shows a significant role in huge applications such as, diseases prediction, face recognition, watermarking, and detection of object and so on. One of the unique features of SVD is that it may be applied on real matrix. For the rectangular matrix M having size p×q, SVD is decomposed in the form of:

M=XDYT D=X
T MY

In above equation, M is a p×q matrix, X is a p× p and Y is a �×q matrix. Both X and Y is orthogonal Matrices. An orthogonal Matrices is the matrix that satisfying the condition

M−1=Mt. With singular value on diagonal D is a p×q diagonal matrix [25]. In above D element can be managed in decreasing order and can be written as

D=diag{d1, d2, ..., dp } p=min (p×q)
d1 ≥d2 ≥d3≥........≥d p>0∧¿
d g+1=dg+2=...=d p=0

In another way we can say that,
X T X=X X T=I, Y T Y=Y Y T=I∧¿

$$D=\begin{bmatrix} d_1 & 0\ldots\ldots & 0 \\ 0 & d_2\ldots\ldots & 0 \\ 0 & 0 & d_p \end{bmatrix}$$

Above decomposition is known as Singular value decomposition. D1, d2, ..., dp

Are Singular values numbers of M. X is the left and Y is the right Singular vector metric of D. X And Y can be described as below:

Where, y_{ij} is rating given by user iforj. We stop training feature when the error rate on a test set increases.

SVD permits us to transform a matrix $Z \in C^{k\times l}$ to diagonal form with the help of unitary matrices i.e. $Z=^{\wedge}O\ \wedge\sum P$¿

In above equation $\wedge_{O\in C}k\ \times l$ has ortho normal columns, $\wedge_{\sum\in C}k\ \times l_{is}$ diagonal, and V $_{\in C}k\ \times l$ unitary. This is a reduced version of SVD also known as practical SVD. On the other hand full version of SVD is of the following form; Z=O∑P¿ O and P are the unitary matrices with$\sum\in C^{k\ \times l}$. A geometric interpretation is used to calculate matrix factors of Z. For classification purpose we use modified SVM. Second method that we use for extraction is Anova.

C. Analysis Of Variance

ANOVA is a statistical technique for comparing means of multiple groups under the null hypothesis that all means are equal. ANOVA produces an F- statistic, defined as:

$F= {}^{\text{variance of the group means}}/\text{mean of the within group variances}$

Essentially, the F value is a comparison of the variance amongst the different groups to the variance amongst all the individuals within those groups. If all means are equal, the F ratio should be approximately one. If the ratio is much larger than one, then the between-groups variability is larger than the within-group variability.

D. Proposed SVM

Support Vector Machine is the favorable process of machine learning centered on the theory of VC dimension and the principle of structural risk minimal, which is component of fine generalization efficiency. The whole concept may also be easily described as follows: searching a superior hyper-plane fulfills the application of classification. Founded on statistical learning theory, the margin scale replicates the generalization ability to a high-quality extent. The larger the margin scale takes, the better the generalization potential of SVMs can have. The brief outline of Modified SVM is as given below. Assume that a given training data set $u_l \in Q^h(l=1,\ 2,,\ n)$with the wanted output $v_l \in \{+1,\ -1\}(l=1,\ 2,,\ n)$ corresponding to the two classes. And assume that there is a different hyper- plane with target function $f\ (u)=\text{¿}k^T.u>+d$, k and d shows the normal vector and the bias respectively. The margin of separation is maximized by$2/\|k\|$, which helps to ensure that all training data to be classified. For non- separable case, slack variable $_{\varepsilon l}\geq 0,\ l=1,\ 2,\ ...,\ n$. A Traditional SVM for optimal separating hyper- plane has the following optimization problem: Minimize $k^{T\,k}2^{+C\Sigma}l=1^{n}$

ε_l

"c" can be calculated using $C=\max(|\acute{}y+3\sigma_{\ y|,\ |\acute{}y-3\sigma\ y|)}$

Such that, $v_{l(k.ul}+d)\geq 1-\varepsilon_l,\ \varepsilon_l>0,\ l=1,\ 2,\ ...,\ n$ In above equation, C is the nonnegative parameter which is chosen by person. Gaussian RBF kernel functions

$\min \alpha,\ \beta\ ^{F(\sigma,\ \alpha(\sigma,\ \beta))=\alpha\ T}(\sigma,\ \beta)e-^{1}2^{\alpha\ T}(\sigma,\ \beta)K_{ext}(\sigma,\ \beta)\alpha(\sigma,\ \beta)$ s.t $\alpha(\sigma,\ \beta)=\text{argmax}\ \alpha^{T}e-^{1}2^{\alpha\ T}K_{ext}(\sigma,\ \beta)\alpha$ s.t $y^{T}\alpha=0,\ _{Ce\geq\alpha\geq 0}$

Step 1: Data is imported and is preprocessed to convert it into a clear format

Step 2: Feature extraction is done using H- SVD to remove redundancy in dataset.
Using a threshold (often it is 80%-90%) for the amount of original data variation explained by the new features leads to the selection of a small number (k) of the new features.
☐ By reducing sum of squared residuals, Parameters are estimated. Before training, from rating from a simple baseline prediction is subtracted – combination of six predictors, with weights chosen with linear regression.

$r_{ij}=y_{ij}-\wedge y_{ij}\ u_{ik}+\text{¿}lrate\star(r_{ij}v_{\ jk}-\lambda u_{ik})$

$u_{\ jk+\text{¿}l}rate_{*(rijuik-\lambda v\ jk)}$ Where y_{ij} is rating given by user $_i$ for $_{j.}$ We stop training feature when the error rate on a test set increases

Step 3: Classification (Training) SVM Algorithm

Tuning support vector regression: Support vector tuning is done with the help of following parameters having Gaussian radial basis function kernel.

Cost/ Penalty factor: It controls tradeoff between the errors of SVM on the training data and margin maximization.

Epsilon: Both complexity and generation capabilities of network depend on its values

Gamma: Biasing and variance is decided by it

After running SVM training and testing for each parameters combination, choose the good values by looking at the least error generated.

Step 4: Performance Evaluation: Testing is done using four parameters such as Accuracy, Precision, Recall and F-measure and comparison is done with KNN, Logistic Regression and Discriminative.

5. Results

The Data Mining techniques are quite helpful in one the stages of Predictive Analysis of Diseases. In this work we mainly focus on four parameters i.e. Accuracy, Precision, Recall, F- measure.

Accuracy: This refers to the ability of the classifiers to correctly measure the intrusions from the training dataset. This is defined as ratio of appropriately classified data to overall classified data.

$\text{Accuracy} = {}^{TP+TN}/TP+FP+TN+FN$

In above equation, TP = True Positive

TN = True Negative
FP = False Positive
FN = False Negative

False Positive Ratio

This is one of the main parameters to find out the effectiveness of various models and also the major concern while network setup. A normal data is considered as abnormal or attack type data. It is defined as:

$_{FPR=}\ FP\ /FP+TN$

False Negative Ratio

This is one of the main parameters used to describe a network intrusion gadget's inability to find out true security events under particular situations. An abnormal data is not detected and considered as normal data. It is defined as:

$FNR = {}^{FP}/FN+TN$

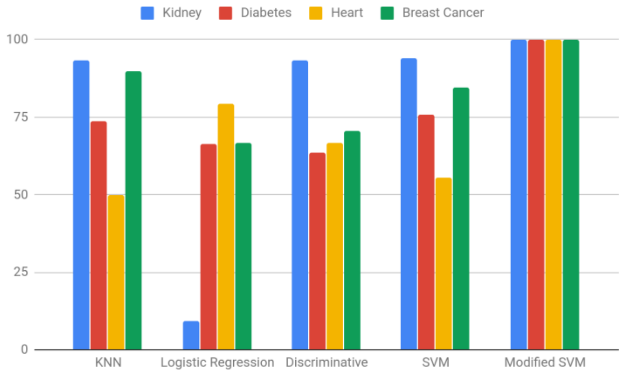

Figure 1: Accuracy Analysis of four Approaches

In this graph we have taken four datasets such as; Kidney, Diabetes, Heart and Brest Cancer. Further four methods are applied on each factor to calculate its accuracy i.e. KNN, Logistic Regression, Discriminative and Modified SVM. When we compared the value of all methods we found that our proposed SVM performs better as compare to all others. In case of Kidney dataset the value of KNN and Discriminative is 90%, value of Logistic Regression is 10, value of our proposed SVM is more i.e. 100%. In the case of Heart dataset value of these methods are 70 for KNN, value of Logistic Regression and Discriminative is almost closer, value of our proposed SVM is 100%.Proposed SVM give 100% in all datasets.

Recall: Recall is how many relevant items are selected. It is a ratio of true positive to the sum of true positive and false negative. In medical diagnosis, test sensitivity (Recall) is the ability of a test to correctly identify those with the disease (true positive rate). If the test is highly Recall and the test result is negative you can be nearly certain that they don't have disease.

Recall=true positives/(true positive+false negative)

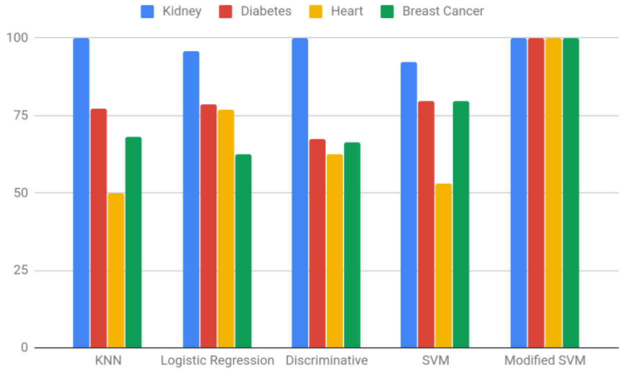

Figure 2: Recall Analysis of four Approaches

In this graph we have taken four datasets such as; Kidney, Diabetes, Heart and Brest Cancer. Further for methods are applied on each factor to calculate its accuracy i.e. KNN, Logistic Regression, Discriminative, SVM and Modified SVM. When we compared the value of all methods we found that our proposed SVM performs better as compare to all others. In case of Kidney dataset the value of KNN, Discriminative and Modified SVM is 100%, value of Logistic Regression and SVM is less than others. In the case of Heart dataset value of these methods are 80 for both for KNN and Logistic Regression, value of proposed SVM is 100%. Proposed SVM give 100% in all datasets.

Precision: Precision is how many selected items are relevant. It is a ratio of true positive to the sum of true positive and false positive. Test specificity (Precision) is the test's ability to correctly recognize those that do not have a disease (true negative rate).If the test output for an extremely precise test is positive user can be nearly certain that they actually have the disease.
Precision=true negatives/(true negative+false positives)

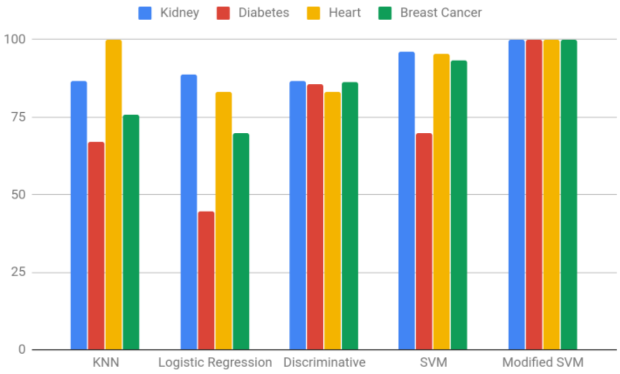

Figure 3: Precision Analysis of four Approaches

F-measure: The F-measure of the system is defined as the weighted harmonic mean of its precision and recall, that is

$$F=\frac{1}{\alpha\frac{1}{P}+(1-\alpha)\frac{1}{R}}$$

Where the weight α∈[0, 1].

In our proposed method we are going to work on above four parameters. For the sake of comparison, result of our proposed classifier i.e. modified SVM is compared with two other classifiers i.e. KNN and Logistic Regression.

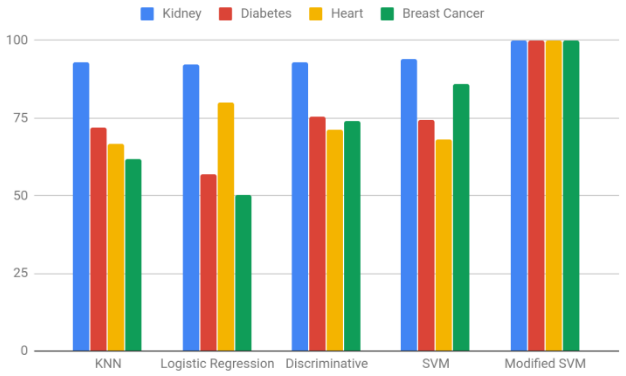

Figure 4: F- measure Analysis of four Approaches

In this graph we have taken four datasets such as; Kidney, Diabetes, Heart and Breast Cancer. Further for methods are applied on each factor to calculate its accuracy i.e. KNN, Logistic Regression, Discriminative and Modified SVM. When we compared the value of all methods we found that our proposed SVM performs better as compare to all others. In case of Kidney dataset the value of KNN, Discriminative and Logistic Regression is 90%, value of Modified SVM is 100%. In the case of Heart dataset value of these methods are lies between 60-80% for KNN, Discriminative and Logistic Regression, value of our proposed SVM is 100%. Proposed SVM give 100% in all datasets.

In above experimentation, the result shows the average performance of different data sets. These models are related on basis of each separate fold or rounds. Experimentation result shows that the Proposed SVM approach is more accurate as compare to other existing System. The accuracy of a proposed model is 100% as shown in the Figure 1. This proposed method performs better existing system. To measure the robustness and effectiveness of any model, comparison of parameters like False Positive Rate and False Negative Rate is computed.

Table 1: Comparison Table

Classifiers	Diseases / Parameters	Kidney	Diabetes	Heart	Breast Cancer
KNN	Accuracy	93.3962	73.6912	50	89.678
	Recall	100	77.3577	50	68.127
	Precision	86.7925	66.9903	100	75.880
	F-measure	92.9293	71.8017	66.6667	61.891
Logistic Regression	Accuracy	9.4165	66.20	79.1667	66.852
	Recall	95.8431	78.4681	76.9231	62.654
	Precision	88.6792	44.6602	83.333	69.789
	F-measure	92.1221	56.9228	80.0	50.369
Discriminative	Accuracy	93.3962	63.651	66.667	70.452
	Recall	100	67.234	62.500	66.400
	Precision	86.7925	85.664	83.333	86.444
	F-measure	92.9293	75.367	71.4236	73.923
SVM	Accuracy	93.96	75.95	55.4196	84.6683
	Recall	92.18	79.52	53.0097	79.5230
	Precision	96.07	69.90	95.4545	93.3824
	F-measure	94.09	74.40	68.1648	85.8972
Modified SVM	Accuracy	100	100	100	100
	Recall	100	100	100	100
	Precision	100	100	100	100
	F-measure	100	100	100	100

6. Conclusion

Data mining strategies had been generously utilized within field of medical diagnosis and prognosis in literature. This work notably places center of attention on feature extraction and classification systems to accurately categorize disease related to kidney. Data mining can be utilized for estimating targeted disease patterns, restorative data extraction, patient backing and administration and discovering of clinical parameters. Here an algorithm is proposed that predicts four different diseases i.e. Kidney, diabetes, heart and Brest cancer which makes use of Proposed Support Vector Machine for classification and H- SVD for feature extraction. From results it is clear that our proposed, algorithm give 100% results in case of Accuracy, Precision, Recall and F- measure and improves performance of system.

Reference

[1] Bates DW et al. “Big Data In Health Care: Using Analytics To Identify And Manage High- Risk And High-Cost Patients” Health Aff. ; 33:1123–1131, 2014. WHO. Chronic diseases and health promotion World Health Organization. 2014. Report available at http://www.who.int/chp/chronic_disease_report/part1/en/index11.html

[2] Abid sarvwar, Vinod Sharma, “Intelligent Bayes Approach to Diagnose Diabetes Type-2” Special Issue of International Journal of Computer Applications on Issues and Challenges in Networking, Intelligence and Computing Technologies, November 2012.

[3] Sonu Kumari, Archana Singh, “A data mining approach for the diagnosis of diabetes mellitus”, IEEE conference on intelligent systems and control, pp-373-375, 2013.

[4] Rakseh Motka, Viral Parmar, “Diabetes Mellitus Forecast Using Different Data mining Techniques IEEE International Conference on Computer and Communication Technology (ICCCT), 2013.

[5] Veena Vijayan V., Aswathy Ravikumar, “Study of Data Mining algorithms for Prediction and Diagnosis of Diabetes Mellitus”, International Journal of Computer Application, Vol 94, pp .12- 16, June 2014.

[6] Harleen Kaur and Siri Krishan Wasan, “Empirical Study on applications of Data Mining Techniques in Healthcare”, Journal of Computer Science, 2006.

[7] C.kalaiselvi, G.M.Nasira, “A New Approach of Diagnosis of Diabetes and Prediction of Cancer using ANFIS”, IEEE Computing and Communicating Technologies, pp 188-190, 2014.

[8] Gunasekar Thangarasu, P.D.D. Dominic, “Prediction of Hidden Knowledge from Clinical Database using Data mining Techniques” International Conference on Computer and Information Sciences(ICCOINS), pp.1-5, 2014.

[9] Velu C.M, K. R. Kashwan, “Visual Data Mining Techniques for Classification of Diabetic Patients”, IEEE International Advance Computing Conference (IACC), pp -1070-1075, 2013.

[10] Asma A. AlJarullah, “Decision discovery for the diagnosis of Type II Diabetes”, IEEE Conference on innovations in information technology, pp-303-307, 2011.

[11] Peter Harrington “Machine Learning in Action”, Manning Publications, 2013

[12] Nirmala Devi M., Appavu alia Balamurugan S., Swathi U.V., “Anamalgam KNN to predict Diabetes Mellitus”, IEEE International Conference on Emerging Trending Computing, Communication and Nano technology (ICECCN), pp. 691 -695, 2013

[13] Jefri Junifer Pangaribuan, Suharjito “Diagnosis of Diabetes Mellitus Using Extreme Learning Machine ”.pp-33-38, 2014

[14] Rojalina Priyadarshini, Nilamadhab Dash, Rachita Mishra “A Novel approach to Predict Diabetes Mellitus using Modified Extreme Learning Machine” International Conference on Electronics and Communication Systems (ICECS), pp.1 5, 2014

[15] Sankaranarayanan.S, Pramananda Perumal T. “Diabetic prognosis through Data Mining Methods and Techniques” International Conference on Intelligent Computing Applications, pp.162- 166, 2014

[16] B.L. Shivakumar, Alby, S. “A Survey on Data-Mining Technologies for Prediction and Diagnosis of Diabetes”, International Conference on Intelligent Computing Applications, pp.167-173, 2014

[17] Poonam Undre, Harjeet Kaur, Prakash Patil “Improvement in Prediction Rate and Accuracy of Diabetic Diagnosis System Using Fuzzy Logic Hybrid Combination”, International Conference on Pervasive Computing (ICPC), pp. 1-4, 2015.

[18] Tom Dent, “Predicting the risk of coronary heart disease”, PHG foundation publisher, 2010.

[19] World Health Organization, “Global status report on no communicable diseases”, 2010.

[20] DSVGK Kaladhar et al., ”Statistical and Data Mining Aspects on Kidney Stones: A Systematic Review and Meta-analysis”, Open Access Scientific Reports, Volume 1, Issue 12, 2012.

[21] Niti Guru, Anil Dahiya, Navin Rajpal, "Decision Support System for Heart Disease Diagnosis Using Neural Network", Delhi Business Review, Vol. 8, No. I (January - June 2007.

[22] Niti Guru, Anil Dahiya, Navin Rajpal, "Decision support System for Heart Disease Diagnosis Using Neural Network", Delhi Business Review, Vol. 8, No. I (January - June 2007.

[23] Chakroborty, Sandipan, and Goutam Saha. "Feature selection using singular value decomposition and QR factorization with column pivoting for text-independent speaker identification." Speech Communication 52.9 (2010): 693-709.

[24] Vijayarani, S., and S. Dhayanand. "Liver disease prediction using SVM and Naïve Bayes algorithms." International Journal of Science, Engineering and Technology Research (IJSETR) 4.4 (2015): 816-820.

[25] Mythili, T., et al. "A heart disease prediction model using SVM-Decision Trees-Logistic Regression (SDL)." International Journal of Computer Applications 68.16 (2013)

[26] Kumari, V. Anuja, and R. Chitra. "Classification of diabetes disease using support vector machine." International Journal of Engineering Research and Applications 3.2 (2013): 1797-1801.

[27] Radhimeenakshi, S. "Classification and prediction of heart disease risk using data mining techniques of Support Vector Machine and Artificial Neural Network." Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on. IEEE, 2016.

[28] Yang, Haiyan, Qifang Luo, and Yongquan Zhou. "Using chaotic adaptive PSO-SVM for heart disease diagnosis." Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on. IEEE, 2009.

[29] Barakat, Nahla, Andrew P. Bradley, and Mohamed Nabil H. Barakat. "Intelligible support vector machines for diagnosis of diabetes mellitus." IEEE transactions on information technology in biomedicine 14.4 (2010): 1114-1120.

[30] Akay, Mehmet Fatih. "Support vector machines combined with feature selection for breast cancer diagnosis." Expert systems with applications 36.2 (2009): 3240-3247.

[31] Vijayarani, S., and S. Dhayanand. "Kidney disease prediction using SVM and ANN algorithms." Int. J. Comput. Business Res 6.2 (2015).