

A Framework for Efficient Healthcare Resources Utilization using Semi-supervised Machine Learning Algorithm

C. Natarajan^{1*}, J.M. Gnanasekar², N. Janorious Hermia³

¹ Department of Computer Science and Engineering, Saveetha University, Chennai, India

² Department of Computer Science and Engineering, Srivenkateswara College of Engineering & Technology, Chennai, India

³ Department of Computer Science and Engineering, Pallavan college of Engineering, Kanchipuram, India

Corresponding Author: vcnataraj@gmail.com, Tel.: +91-7339590059

DOI: <https://doi.org/10.26438/ijcse/v7i4.894900> | Available online at: www.ijcseonline.org

Accepted: 19/Apr/2019, Published: 30/Apr/2019

Abstract— Electronic Health Records are providing high amount of genetic data and clinical information through the exceptional advances in biotechnology and health sciences. The application of machine learning and data mining methods in biosciences is crucial, more than that very important to transform cleverly all available information into precious knowledge. Diabetes mellitus is defined as a collection of metabolic disorders exerting major pressure on human health worldwide. Large amounts of data generated due to the widespread researches in all areas of diabetes. This study is to present a systematic approach of the applications of machine learning algorithm along with data mining techniques and tools in the field of diabetes research especially in Health Care Resource Utilization (HCRU). There were so many machine learning algorithms used here. Supervised machine learning algorithm, unsupervised machine learning algorithm and Semi-supervised Machine Learning Algorithm (SMLA). This research shows that SMLA such that Transductive Support Vector Machine (TSVM) fits the best for the research in healthcare resource utilization by considering the type of diabetes patient's medical datasets.

Keywords— Diabetes mellitus, Machine Learning, Healthcare resource utilization, Support Vector Machine

I. INTRODUCTION

The big data has the enormous development in the recent years. The most important advances in biotechnology and more exclusively high data rate sequencing result continuously in an easy and economical data production, and guiding the applied biological science into the big data area.

The data can be generated from the surplus amount of digital microscopy, sensors, digital machines, mass spectrometry. Even though the above different technologies producing the huge amount of data, it will not providing the way to extraction of knowledge or interpretation of information in different views by providing different kind of analysis. From this, we can say that the knowledge discovery in Biological data and or biological data mining is more important and significant for providing better care delivery. The main objective is to probe into the rapidly accruing body of biological data and set the fundamental well scrutinized answers to the queries in biological medicine. The skillful and the usefulness of these techniques are derived from the ability of proportionate methods to dig out patterns and generate models from data. In recent days, Diabetes Mellitus is the major problem that maximum human beings facing throughout the world. DM is not a

diseases, it is an imbalance of sugar level in the body condition of all the human beings. To extracting the knowledge for the large volumes of diabetes data sets, Machine learning algorithm and some data mining methods are used. This paper organized in the following sections. The Section II contains the review of the related works based on the objectives of the proposed work. Section III contains the methodologies that are used in the proposed works such as Knowledge discovery, Diabetes Mellitus, Machine Learning Algorithms. The Section IV contains the details about the dataset, Algorithm, Implementation and Results. The Section V explained about the conclusion and Future work plan.

II. RELATED WORKS

The ultimate aim of data mining is to extract useful information from large quantities of data sets. The important part of the knowledge discovery performed by data mining, which is the overall process of tracking down the useful information from the data by conversion of raw data into knowledge [1]. Mining hospital records, such safety issues could be flagged and addressed by hospital management and government regulators study was made to Extracting cardiac dynamics within ECG signal for human

identification and cardiovascular diseases classification using neural networks,[2] cardiac dynamics within ECG signals is extracted accurately by using radial basis function (RBF) neural networks through deterministic learning mechanism. The obtained cardiac system dynamics is represented and stored in constant RBF networks.

The Classification tree algorithm works in a similar fashion like regression tree algorithm. But the Regression tree produces the numeric and as well as continuous variable in case of duration of stay in the health care department [2]. Thus, the regression tree is most appropriate for prognostic type of problems. There are different types of methods and techniques like Bayesian networks, Neural networks ,Multilayer feed-forward networks and perceptron algorithm, Regression trees, SVM, Decision Tree classification[3] and so on can be studied and the end limits of those can be found. Linear regression [4] and polynomial regression falls under this category. The results declared to show that the prediction of presence of coronary artery disease goes on well with top 5 approaches [5]. The performances of various approaches are compared with Hierarchical Cluster Analysis, Multi-Dimensional Scaling (MDS). The logistic regression gives more excellent results in the most of the researches. A study of data mining Approach for prediction and treatments for diabetic disease was conducted at International journal of Science Innovations today. A study of diabetes was conducted in France in 2001[7] to characterize the therapeutic management and control of diabetes and modifiable cardiovascular risk factors in patients with type 2 diabetes receiving specialist care. The study was proposed to 600 dialectologists across France. This is due in part to the severity of diabetes in the patients seen by specialists in diabetes care [8]. However, both attentiveness and application of published recommendations need to be reinforced.

The assessment of Dementia with the help of variables to be predicted, which are selected either by domain experts or by a statistical driven procedure and keep the prediction models which rely on Logistic Regression algorithm[10]. The aim of this study is to enhance the performance of a recent application of Bayesian belief networks using a proxy approach based on Logistic Regression.

The predictive accuracy of regression trees with Logistic Regression models for predicting hospital mortality in heart failure patients. This research shown the conclusion that logistic regression stands ahead than Regression Tress. The recent research inspired by the variety of perspective is employed by Langley that predictive modelling and clustering consume numerous resemblances. The systems TILDE and CLUS [11] use a modified top-down induction of decision trees algorithm to construct clustering trees.

However the distances are used in TILDE and CLUS systems have considered attributes or classes separately, but not both together, even though the idea was presented .Our approach uses a SMLA for the health care resource utilization profiling as such, it is closely related to the approaches for . However, it extends rule induction to the more general task of multi-objective prediction. While some work exists on multi-objective classification with decision trees, the authors are not aware of any work on rule-induction for multi-objective classification. Learning statistical models [12] of phenotypes using noisy labelled training data. J Am Med Inform. Also, little work exists on rule-based regression (some recent examples come from the area of ILP[13], let alone rule-based multi-objective regression Related to rule induction is subgroup discovery which tries to find and describe interesting groups of examples.

III. METHODOLOGY

1. Knowledge Discovery

The most important step in the entire Knowledge discovery in databases (KDD) process is data mining, exemplifying the application of machine learning algorithms in analyzing data. KDD [14] is a field encompassing theories, methods and techniques, trying to make sense of data and extract useful knowledge from them. A complete definition of KDD is the nontrivial process identifying valid, novel, potentially useful, and ultimately understandable patterns in data. It is considered to be a multistep process includes selection, pre-process, transformation, data mining, interpretation and evaluation.

2. Machine Learning Algorithm

The term machine learning is indistinguishable to the term artificial intelligence, given that the prospect of learning is the main distinction of an entity called intellectual in the broadest sense of the word. Machine learning is the systematic field dealing with the ways in which machines learn from experience. The purpose of machine learning is the construction of computer systems that can adapt and learn from their experience. A computer program is said to learn from experience with respect to some class of tasks and performance measure, if its performance is measured by at some class of tasks, and improves with the experience gained. A more complete and proper explanation of machine learning is given in each and every one of the places.

Machine learning algorithms are characteristically classified into three big categories. These are semi-supervised learning, in which the system infers a function from labelled training data, and unsupervised learning, in which the learning system tries to infer the structure of unlabelled

data, and semi-supervised learning system consists of a combination of two algorithms.

3.1. Semi-Supervised Learning

Semi-supervised learning is a very capable approach since it can use readily available unlabelled data to improve supervised learning tasks when the labelled data is insufficient or costly. Semi-supervised learning algorithms use both labelled and unlabelled data for training. The labelled data is typically a small percentage of the training dataset. The goal of semi-supervised learning is to recognize how combining labelled and unlabelled data may change the learning behavior, and design algorithms that take advantage of such a permutation.

Co-training requires two or more views of the data, i.e. disjoint feature sets that provide different complementary information about the instances. Ideally, the two feature sets for each instance are conditionally independent. Self-training can refer to a variety of schemes for using unlabelled data. The implementation of self-training by bagging and majority voting. There are many different semi-supervised learning algorithms. Some of the most commonly used ones including Expectation-Maximization [15] with generative mixture models is an iterative method for finding maximum likelihood estimates of parameters in statistical models, where the models depend on unobserved latent variables. Each iteration of the algorithm consists of an expectation step followed by a maximization step. EM with generative mixture models are suitable for applications where the classes specified by the application produce well clustered data. An ensemble of classifiers is trained on the labelled data instances and then the classifiers are used to classify the unlabelled examples independently.

In inductive learning, the algorithm is trained on explicit training instances but the goal is to learn general rules, which are then applied to the test cases. Transductive SVMs (TSVM) extend general SVMs in that they could also use partially labelled data for semi-supervised learning by following the principles of transduction. In contrast, transductive learning [16] is reasoning from specific training cases to specific testing cases. Graph-based methods are algorithms that exploit the graph structure obtained by capturing pair wise similarities between the labelled and unlabelled instances. These algorithms define a graph structure where the nodes are labelled and unlabelled instances and the edges, which may be weighted, represent the connection of the nodes they connect.

3.2. Supervised Learning

The supervised learning system must learn inductively a function called target function, which is an expression of a model describing the data. The objective function is used to predict the value of a variable, called dependent variable or

output variable, from a set of variables, called independent variables or input variables or characteristics or features. The set of possible input values of the function, called its domain, and also called instances. Each case is described by a set of characteristics such as attributes or features. A subset of all cases, for which the output variable value is called training data or examples. Classification models try to predict distinct classes, such as blood groups, while regression models predict numerical values [17]. Some of the most common techniques are Decision Trees, Rule Learning, and Instance Based Learning, such as k-Nearest Neighbors, Genetic Algorithms, Artificial Neural Networks, and SVM. In order to conclude the best target function, the learning system, given a training set, takes into consideration alternative functions, called hypothesis and denoted by h . In supervised learning, there are two kinds of learning tasks: classification and regression.

3.3 Unsupervised Learning

In the unsupervised learning system tries to discover the unknown structure of data or relations between variables. In that case, training data consists of instances without any equivalent labels.

Association Rule Mining approaches appeared much later than machine learning and is subject to greater influence from the research area of databases. It was proposed in the early 1990s as a market basket analysis, in which the aim was to find correlations in the objects of a database. Based on the shopping cart example, association rules are of the form $X \rightarrow Y$, which means that if you find all of X in a cart it is possible to find Y . The Apriori algorithm [18] is the most well known association rule discovery algorithm. Some of the most well-known applications in biology and bioinformatics include biological sequence analysis, analysis of gene expression data and others. Even though the association rule mining was first introduced as a market basket analysis tool, it has since become one of the most valuable tools for performing unsupervised exploratory data analysis over a wide range of research and commercial areas, including biology and bioinformatics. A meticulous review of discovering frequent patterns and association rules from biological data, including algorithms and applications. Clusters are informative patterns occurring through grouping the similar data items into one group. The separation of a complete dataset into groups of data, so that instances belonging to the same group are as similar as possible and instances belonging to different groups differ as much as possible.

4. Diabetes Mellitus

DM is one of the most common endocrine disorders, affecting more than 1000 million people worldwide. Diabetes Mellitus [19] is defined as a cluster of metabolic disorders mainly caused by abnormal insulin secretion

and/or action. The commencement of diabetes is estimated to rise dramatically in the upcoming years. DM can be divided into quite a lot of different types. However, there are two major clinical types, type 1 diabetes (T1D) and type 2 diabetes (T2D), according to the etiopathology of the disorder. The symptoms of DM include polyurea, polydipsia, and significant weight loss among others. Diagnosis depends on blood glucose levels such as fasting plasma glucose is 7.0 mmol/L. Impaired metabolisms of carbohydrates, fat and proteins caused form the insulin deficiency and it is results in elevated blood glucose levels causes hyperglycemia. T2D appears to be the most common form of diabetes, of all diabetic patients, mainly characterized by insulin confrontation. Other forms of DM, classified on the basis of insulin secretion profile and/or onset, include Gestational Diabetes, endocrinopathies, Maturity Onset Diabetes of the Young, neo-natal, mitochondrial, and pregnancy diabetes. The main causes of T2D include lifestyle, physical activity, dietary habits and heredity, whereas T1D is thought to be due to auto immunological devastation of the Langerhans islets hosting pancreatic- β cells. T1D affects almost 10% of all diabetic patients worldwide, with 10% of them ultimately developing idiopathic diabetes.

IV. IMPLEMENTATION AND RESULTS

4.1 Data Set:

The Medical healthcare datasets are available in more websites and can collect from different network of physicians. Different types of bench marked health care data sets that are available in UCI machine repository and small description about them are collected from that repository. Since these data sets have the values that indicate the significant impact of a particular disease. Hence there is requirement to have proper predictive machine learning algorithms that help the health care resource managers, experts and doctors planning the resource and the treatment accordingly. In this work the accuracy of the approaches like supervised machine learning and SMLA are considered and proved that the SMLA shows the best fir for the healthcare resource management. We select SVM algorithm from the supervised machine learning algorithm, and TSVM for the SMLA.

The test set of dataset is a simple linearly separable two dimensional set. The goal of designing this dataset is to address the uniqueness and desirable quality of SMLA more intuitively. There are overall 600 labelled samples and 200 unlabelled samples in the mixed training data. If i is one of the intermediate levels, we build the set of training data $T.data_i$ by inheriting the support vectors S_{i+1} and adding to them some of their nearest neighbors at level i . If the size of $T.data_i$ is still small enough for applying model selection, and solving TSVM on the whole $T.data_i$, then we use the

parameters C_{i+1} , and γ_{i+1} as initial values for the current, and the retrain levels. If not, the C_{i+1} , and γ_{i+1} are inherited in C_i , and γ_i . The direct application of the TSVM, $T.data_i$ is clustered into K clusters, and pairs of P nearest opposite clusters are retrained, and contribute their solutions to S_i . The number of K is determined in as $T.data_i/Q_t$. The different pairs of clusters are not dependent on the parallel cluster-based retraining.

Algorithm for SMLA

Step 1: Initialize T_i, C_i, S_i, γ_i

Step 2: if i is the intermediate level then
Calculate the best (C_i, γ_i) using UD
 $S_i \leftarrow$ Apply Transductive SVM on X_i

Step 3: Calculate nearest neighbors N_i for support vectors S_{i+1}

Step 4: $T.data_i = S_{i+1} \cup N_i$

Step 5: if $|T.data_i| < Q_t$ then
 $C^O \leftarrow C_{i+1}; \gamma^O \leftarrow \gamma_{i+1}$
Run UD using the initial center C^O, γ^O
Else $C_i \leftarrow C_{i+1}; \gamma_i \leftarrow \gamma_{i+1}$

Step 6: if $|T.data_i| \geq Q_t$ then
 $T.data_i$ into K clusters
 $\forall k \in K$ find P nearest opposite-class clusters
 $S_i \leftarrow$ Apply TSVM on pairs of nearest clusters only
Else
 $S_i \leftarrow$ Apply TSVM directly on $T.data_i$
end if

Step 7: Return S_i, C_i, γ_i

The following table shows that the comparison of supervised learning and semi-supervised Learning algorithm. The figure 1 shown that the graphical view of the comparison of those two data mining algorithms. The training time of the supervised algorithm is greater than that of the semi-supervised one. The accuracy of the semi-supervised for these healthcare data sets especially for the diabetes dataset is the maximum. The expected test errors for the test data also maximum for the supervised learning algorithm.

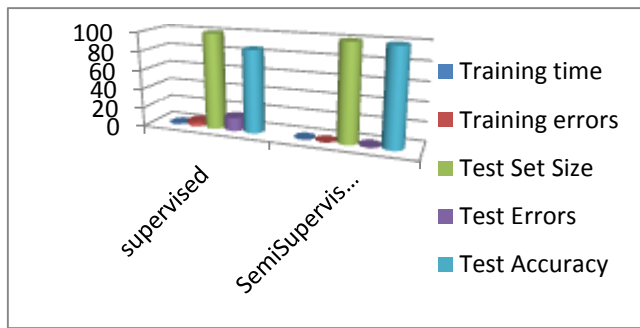


Figure 1: Supervised Vs Semi-supervised

The healthcare data sets are involving to find out the best test accuracy from the supervised and semi-supervised algorithm. The table 2 show that the differences between the supervised and semi-supervised algorithm in the basis of the

Table 1: Comparison of the results of Supervised and semi-supervised on healthcare dataset

Training Algorithm	No.of Vectors	Test Errors	Precision	Recall	Test Accuracy(%)
Supervised	30	125	71.19	98.00	79.17
Semisupervised	30	91	77.43	98.33	84.83
Supervised	39	91	90.98	77.33	84.83
Semisupervised	39	79	89.35	83.00	86.83
Supervised	181	43	89.78	96.67	92.83
Semisupervised	184	46	87.57	98.67	92.33
Supervised	146	109	73.46	99.67	81.83
Semisupervised	140	93	76.61	99.33	84.50

test data errors and the test accuracy. The precision and the recall parameters percentage clearly showing that the semi-supervised algorithm will be the best fit for the healthcare resources utilization. When the numbers of vectors are minimum the test error report for the supervised gives the maximum in the total data set.

Table 2: Accuracy Report for Supervised and Semi-supervised

Algorithm	Time	Errors	Set size	Test Errors	Test Accuracy
SVM	0.77	5	100	14	85
TSVM	0.20	0	100	1	99

The Table 2 concludes that the test errors and the test accuracy depend on the number of vectors we have chosen on the type of algorithm used. The maximum accuracy for the test data shown that the SMLA is having 92.83. The precision and recall percentage also calculated.

The table 3 concludes that the growth rate for the supervised and SMLA. The growth rate calculation depends on the factors of the values of the k and the test accuracy. The growth rate for the supervised learning algorithm will be two points less than the semi-supervised learning algorithm approximately.

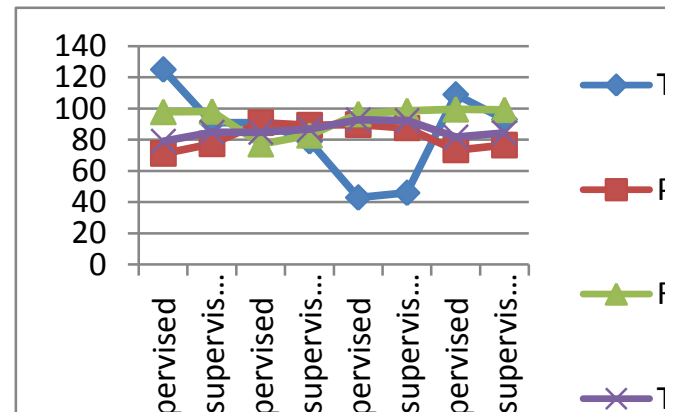


Figure 2: Test Accuracy of Supervised and Semi-supervised

The figure 2 and 3 shows the test accuracy of the supervised and semi-supervised machine learning and growth rates in the graphical methods. Test errors are playing greater level than the test accuracy and the precision and recall percentages.

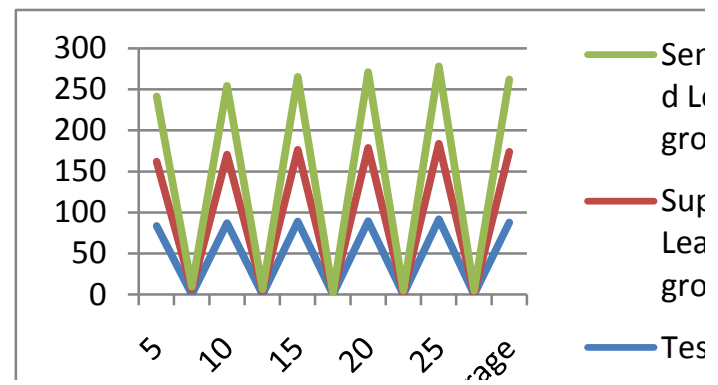


Figure 3: Growth rate of supervised and Semi-supervised

It should include important findings discussed briefly. Wherever necessary, elaborate on the tables and figures without repeating their contents. Interpret the findings in view of the results obtained in this and in past studies on this topic. State the conclusions in a few sentences at the end of the paper. However, valid colored photographs can also be published.

Table 3: Growth rate for supervised and Semi-supervised

K values	Test Accuracy	Supervised Learning growth rate	Semi-supervised Learning growth rate

5	83.1	78.5	80.2
		5.54	3.49
10	86.6	83.6	84.4
		3.46	2.54
15	88.5	87.8	89.3
		0.79	0.9
20	89.1	89.5	92.6
		0.45	3.93
Avg	87.7	86.32	88.26

V. CONCLUSION AND FUTURE SCOPE

The proposed paper is to solve the problems in the comparison between the supervised and SMLA with using TSVM, by a preset number of samples. Predicting the resource allocation correctly is very difficult before training the vector machine, therefore leads to substantial estimation error, especially when the number of the labelled examples is very small. To avoid using more unlabelled examples in a native way, we suggested active learning. Studies have found no correlation between using more unlabelled examples lead to better learning performance, hence more accurate selection of labels is required instead of large number. Semi-supervised learning solves this problem. Compared with vector machine method based on semi-supervised learning, the proposed algorithm added a more various regularization term, which makes full use of the distribution characteristics of unlabelled examples. Compared with supervised, the proposed algorithm has more advantages, as it does not need presetting number of class samples, does not frequently exchange class label, and make use of semi-supervised learning, which selects the best unlabelled data to maximize the performance of classifier. In additional, our algorithm works well on small labelled training dataset, has strong generalization ability, and is applicable to healthcare resource utilization. In future work we have to conduct more astute theoretical analyses about the effectiveness of our proposed method in healthcare resource management. Furthermore, have to design the frame work with semi-supervised learning algorithm for the specific work.

REFERENCES

- [1] Agrawal R, Imielinski T, Swami A. *Mining association rules between sets of items in large databases*. Proceedings of the ACM SIGMOD conference on management of data; 1993. p. 207–16.
- [2] Agrawal R, Srikant R. *Fast algorithms for mining association rules in large data-bases*. Proceedings of the 20th International Conference on Very Large Databases; 1994. p. 478–99.
- [3] Kavakiotis I, Tzanis G, Vlahavas I. *Mining frequent patterns and association rules from biological data*. In: Elloumi M, Zomaya AY, editors. *Biological knowledge discovery handbook: preprocessing, mining and post processing of biological data*. Wiley Book series on bioinformatics: computational techniques and engineering, New Jersey, USA: Wiley-Blackwell, John Wiley & Sons Ltd.; 2014
- [4] Han J, Kamber M, Pei J. *Data mining: concepts and techniques. The Morgan Kaufmann series in data management systems*; 2011.
- [5] Alpaydin E. *Introduction to machine learning*. Cambridge Massachusetts London England: The MIT Press; 2004.
- [6] Guyon I, Elisseeff A. *An introduction to variable and feature selection*. *J Mach Learn Res* 2003; 3:1157–82.
- [7] Witten IH, Frank E, Hall MA. *Data mining: practical machine learning tools and techniques*. 3rd ed. Burlington, MA: Morgan Kaufmann; 2011.
- [8] American Diabetes Association. *Diagnosis and classification of diabetes mellitus*. *Diabetes Care* 2009; 32 (Suppl. 1):S62–7.
- [9] Cox EM, Elelman D. *Test for screening and diagnosis of type 2 diabetes*. *Clinical Diabetes* 2009;4(27):132–8.
- [10] Krentz AJ, Bailey CJ. *Oral antidiabetic agents: current role in type 2 diabetes mellitus*. *Drugs* 2005;65(3):385–411.
- [11] Tsave O, et al. *Structure-specific adipogenic capacity of novel, well-defined ternary Zn(II)-Schiff base materials. Biomolecular correlations in zinc-induced differentiation of 3T3-L1 preadipocytes to adipocytes*. *J Inorg Biochem* Nov 2015; 152:123–37.
- [12] Halevas E, et al. *Design, synthesis and characterization of novel binary V (V)-Schiff base materials linked with insulin-mimetic vanadium-induced differentiation of 3T3-L1 fibro-blasts to adipocytes. Structure-function correlations at the molecular level*. *J Inorg Biochem* Jun 2015; 147:99–115.
- [13] Tsave O, et al. *The adipogenic potential of Cr (III). A molecular approach exemplifying metal-induced enhancement of insulin mimesis in diabetes mellitus II*. *J Inorg Biochem* Oct 2016; 163:323–31.
- [14] "Records in DBLP". Statistics. DBLP. Retrieved 2016–07-16; 2016.
- [15] Agarwal V et al. *Learning statistical models of phenotypes using noisy labeled training data*. *J Am Med Inform*
- [16] Hoyt R, Linnville S, Thaler S, Moore J. *Digital family history data mining with neural networks: a pilot study*. *Perspect Health Information Management* Jan 1 2016.
- [17] Anderson JP et al. *Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: an application of machine learning using electronic health records*. *J Diabetes SciTechnology* Dec 20 2015; 10(1):6–18.
- [18] Anderson AE et al. *Electronic health record phenotyping improves detection and screening of type 2 diabetes in the general United States population: a cross-sectional, unselected, retrospective study*. *J Biomed Inform* Apr 2016; 60:162–8.
- [19] Bashir S, Qamar U, Khan FH. *IntelliHealth: a medical decision support application using a novel weighted multi-layer classifier ensemble framework*. *J Biomed In-form* Feb 2016; 59:185–200.
- [20] *Multilevel Weighted Support Vector Machine for Classification on Healthcare Data with Missing Values*. Talayah Razzaghi, Oleg Roderick, Ilya Safro, and Nicholas Marko Published: May 19, 2016.
- [21] J. H. Friedman. *Multivariate adaptive regression splines*. *Annals of Statistics*, 19(1):1–67, 1991.[9] A. Gawande. *The hot spotters*. New Yorker, January 2011.
- [22] A. K. Jain and Richard C. Dubes. *Algorithms for Clustering Data*. Prentice-Hall, Upper Saddle River, NJ, USA, 1988.
- [23] ChaitraliDangare, S. and Sulaba Apte,S.*Improved study of disease prediction using data mining classification techniques*. *Int.J.Comp.Appl*. 2012, 47(10):75-88.

Authors Profile

Natarajan Chellasamy M.E(Ph.D) was born in Virudhunagar District of Tamil nadu. He had finished his B.E degree in Computer Science and Engineering in Manonmanium Sundaranar Univeristy and he had finished his M.E degree in Computer Science and Engineering in Anna University,Chennai.He is doing his Ph.D in Computer Science and Engineering in Saveetha Univerisy since June 2013. He has published more than 5 research papers in reputed international journals and more than 10 conference papers and it's also available online. His main research work focuses on Big Data Analytics, Data Mining,Neural networks,Data Structures. He has more than 17 years of teaching experience and 5 years of Research Experience.



Dr.J. M. Gnanasekar is working as a Professor in Department of Computer Science and Engineering, Sri Venkateswara College of Engineering (SVCE), Sriperumpudur, Kancheepuram (District), TN, India. He Completed his ME in Computer Science and Engineering and PhD from College of Engineering, Guindy, Anna University, Chennai. His areas of interest include Wireless sensor networks, security, data mining, NLP and cloud computing. He has more than 20 years of experience in teaching and research in the area of computer science and engineering.



N.Janorious hermia M.E (CSE) was born in Kanyakumari District of Tamil nadu. She had finished her B.E degree in Computer Science and Engineering in Manonmanium Sundaranar Univeristy and she had finished her M.E degree in Computer Science and Engineering in sankara SCSVMV University,Kanchipuram.She has published more than 2 research papers in reputed international journals and more than 5 conference papers and it's also available online. Her main interests on Digital priciples and system design, Data Mining,Neural networks,Data Structures. She has more than 10 years of teaching experience.



References

1.Dr.R.Dhanasekaran

*K.S.Rangasamy College of Arts and Science,
Tiruchengode,Tamilnadu, India*

2.Dr.V.Neelanarayanan

*School of Computing Sciences, VIT University, Tamilnadu,
India*

3.Dr.S.Murugavelh

*Faculty of Engineering Sciences, VIT University, Vellore, Tamil
Nadu, I*