A Soft Computing Approach to Provide Recommendation on PIMA Diabetes

Syed Muzamil Basha, H. Balaji[1], N. Ch. S. N Iyengar[1] and Ronnie D. Caytiles[2]

SCOPE, VIT University, Vellore-632014
[1]Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad, India
[2]Multimedia Engineering Department, Hannam University, Daejeon, Korea
muza.basha@gmail.com, rdcaytiles@gmail.com,
srimannarayanach@sreenidhi.edu.in

Abstract

The contributions in our research are First, we considered the PIMA dataset with 768 instances and 9 features collected from UCI Repository datasets. Second, we attempt to understand PIMA dataset using Visualization. Third, Dimensionality reduction is performed based on Z Value obtained from Generalized Linear Model (GLM). Four, we applying Machine Learning algorithms and probabilistic modeling Techniques like Linear Discriminant Analysis (LDA), Logistic Regression, Generalized Liner Model (GLMNET), SVM Radial, kNN, Naive Bayes, Regressive Partitioning (rpart), Boosted Tree (C5.0), Bagged CART (treebag), Random Forest (RF), Generalized Boosted Modeling with respect to classification accuracy to find the best model for PIMA diabetes. Five, Evaluating the Diabetes Risk Factor (DRF) from selected features using Generalized Linear Model (GLM). Six, Designing the Membership function and construct Fuzzy rules to obtain different levels of Diabetes.

1. Introduction

Diabetes is one of the major deadliest diseases in the world. It is not only a disease but also have association to other kinds of diseases like heart attack, blindness, kidney diseases, *etc*. Normally identifying process is that patients need to visit a diagnostic center, consult their doctor, and sit tight for a day or more to get their reports. Moreover, every time they want to get their diagnosis report, they have to waste their money in vain. In [1] the ISO/IEEE 11073 family of standards for device communication, specific term codes, formats, and a common core of communication functionality for personal Tele-health are defined in this standard. In our research we would like to develop an novel approach to find a appropriate solution, we have developed a system using Probabilistic and Machine learning models, which has the ability to predict whether the patient has diabetes or not. Furthermore, predicting the disease early leads to treating the patients before it becomes critical. The proposed model have ability to extract hidden knowledge from a huge amount of diabetes-related data collected from UCI data repository, In [12] Deepak Chowdary *et al.*, 2016 conducted survey on stress levels in software professionals, PEGASOS optimization algorithm is taken for classification. Boosting is performed on, each sample to update the error sample proportionally. This model is applied and compared with some of the existing UCI datasets such as Pima-diabetes,

Heart-statlog and Sonar. The aim of our research is to develop a system which can predict the diabetic risk level of a patient with a higher accuracy. This research has focused on developing a system based on **Linear methods**: Linear Discriminant Analysis and Logistic Regression. **Non-Linear methods**: Neural Network, SVM, kNN and Naive Bayes **Trees and Rules**: CART, J48 and PART **Ensembles of Trees**: C5.0, Bagged CART, Random Forest and Stochastic Gradient Boosting. In R Caret provides a good selection of evaluation metrics to use for different problem types include: Classification: **Accuracy**: x correct divided by y total instances. Easy to understand and widely used. **Kappa**: easily understood as accuracy that takes the base distribution of classes into account. Regression: **RMSE**: root mean squared error. Again, easy to understand and widely used. **Rsquared**: the goodness of fit or coefficient of determination. Other popular measures include ROC and LogLoss. The evaluation metric is specified the call to the *train()* function for a given model, so we will define the metric now for use with all of the model training later. In [2] Ibrahim *et al.*, 2016 used Magnetic resonance imaging (MRI) tagging is an technique for evaluating regional heart function. And compared the harmonic-phase (HARP) and sine-wave modeling (SinMod) tagging analysis techniques for evaluating myocardial strain and torsion in healthy controls and type-1 diabetes patients. In [16] Mohamed Hassanein *et al.*, 2017 covered key areas, epidemiology, the physiology of fasting, risk stratification, nutrition advice and medication adjustment. Providing Practical Guidelines to enhance knowledge surrounding the issue of diabetes and Ramadan fasting, thereby empowering healthcare professionals to give updated advice and the best possible support to their patients during Ramadan.

2. Data Analysis on Pima Diabetes Dataset Using R

The dataset considered in the experiment belongs to Pima Indian dataset collected from UCI Machine Learning Repository consisting of 768 observations with 9 variables. In [3] Li *et al.*, 2016 developed an personalized approach that considers the unique social, cultural and demographic characteristic of American Indians. In which real-time diabetes health recommendations are provided by effectively and automatically collecting users' physical and social behavior data.

Table 1. Description of Dataset

Name of the attribute	Type	Range of values (Min - Max)	Explanation
TIMES_PREGNANT	Integer	0 - 17	Frequency of Being Pregnant
PLASMA_GLUCOSE	Integer	0 - 199	Values Obtained from Oral Glucose Tolerance Test
DIASTOLIC_BP	Integer	0 - 122	Diastolic Blood Pressure
TRI_SKIN_FOLD	Integer	0 - 99	Triceps Skin Fold Thickness
INSULIN	Integer	0 - 846	Values Obtained from 2 Hour Serum Insulin
BMI	Number	0 - 67.1	Body Mass Index
D_PED	Number	0.078 - 2.42	Diabetes Pedigree Function
AGE	Integer	21 - 81	Age
DETECTION	Factor	0 - 1	Positive and Negative

A distribution for the variable "PlasmaGC." This distribution is also pretty normal:
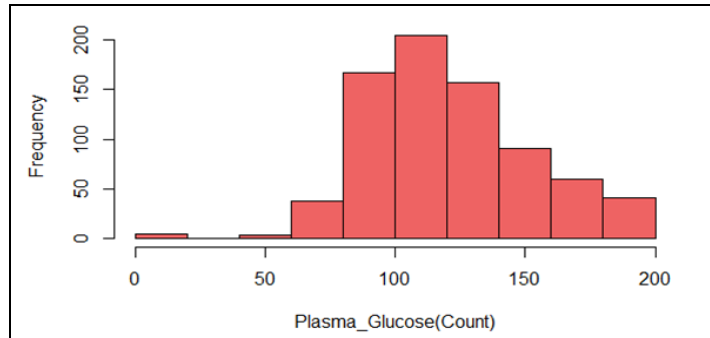

Figure 1. Plasma_Glucose Distribution

Here is a uni-variate distribution plot for the variable "Age." The distribution is skewed right.
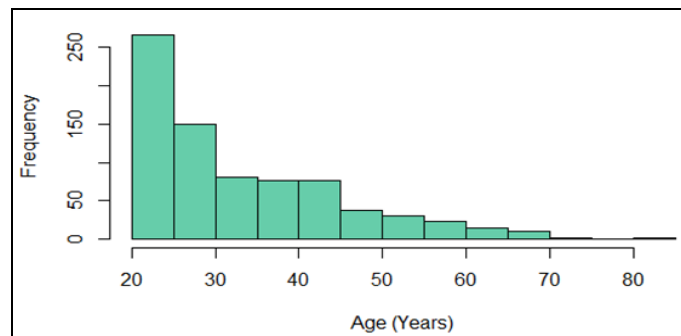

Figure 2. Age Distribution

Here is a distribution plot for the variable "BMI," which appears to be mostly normal, with a few outliers:
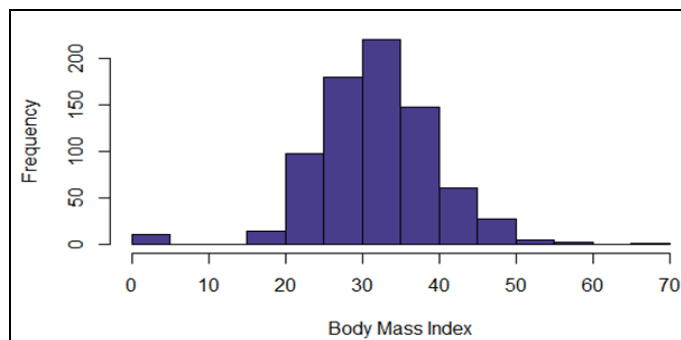

Figure 3. BMI Distribution

Finally, here is a pairs plot for each of the previous variables: In [14] Saud Alzahrani *et al.*, 2017 determined the association between H. pylori infection and risk of incident diabetes in adults at high risk for diabetes by examine the cross-sectional association between presence of H. pylori antibodies and insulin sensitivity, insulin secretion and the disposition index-like measure after a 75-gram oral glucose tolerance test (OGTT). Results in adjusting for matching factors, there was no association between exposure to H. pylori and incident diabetes (odds ratio

[OR] of 1.04 (95% CI, 0.77 to 1.40). In [15] Tsuyoshi Mashitani *et al.*, 2017 assessed the association between dipstick hematuria and estimated glomerular filtration rate (eGFR) decline in Japanese patients with type 2 diabetes, by using Cox proportional hazard model adjusted for potential confounders, this association remained significant even after the exclusion of patients who did not have diabetic retinopathy (hazard ratio: 2.39; 95% CI: 1.13–5.04). And Conclusion that Positive dipstick hematuria was associated with severity of albuminuria and renal function, Suggested that dipstick hematuria is an indicative of more severe diabetic nephropathy.

3. Understand Pima Data with Visualization

The We can see that attributes **age, pedi** and **test** may have an exponential distribution. Also **mass, pres** and **plas** attributes may have a Gaussian or nearly Gaussian distribution. In [4] knoor *et al.*, 2015 proposed a testing method for Android mHealth apps, designed using a threat analysis, considering possible attack scenarios and vulnerabilities specific to the domain.
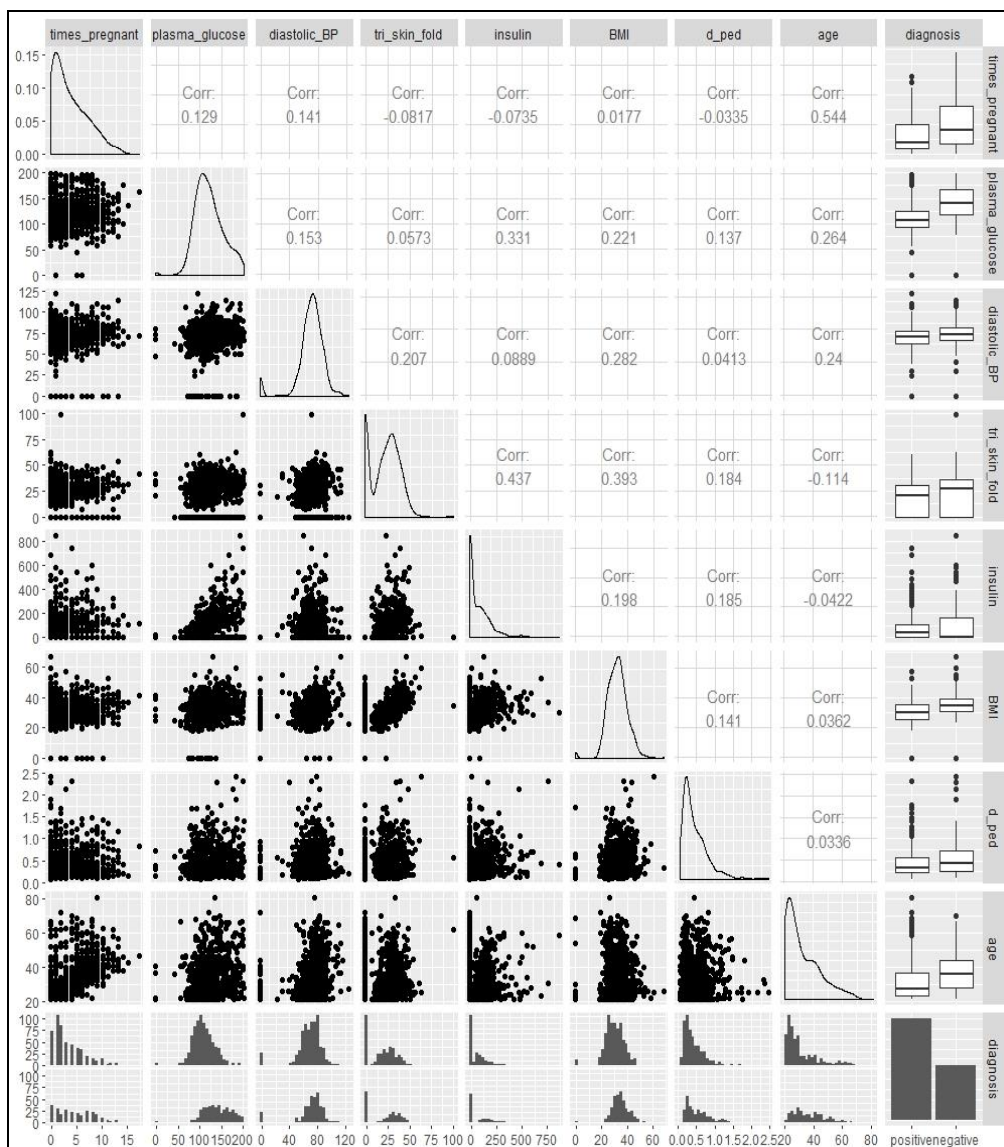

Figure 4. PIMA Visualization using R

4. Methodology

In [5] Zhang *et al.*, 2016 presented a mobile solution for the management of hypertension and diabetes by exchanging the health information between patients and pharmacists, which supports decision making, and assists behavioral interventions for enhanced self-management and healthcare delivery. In [6] Darabi *et al.*, 2015 proposed an architecture of multi agent system which helps type 2 diabetic patients. Each effective parameter of controlling diabetes were considered and providing knowledge for patients' like: suitable diet and monitoring blood glucose level, is assigned to be an agent which works independently at the same time as coordinates and cooperates with other agents.

Figure 5. Methodology
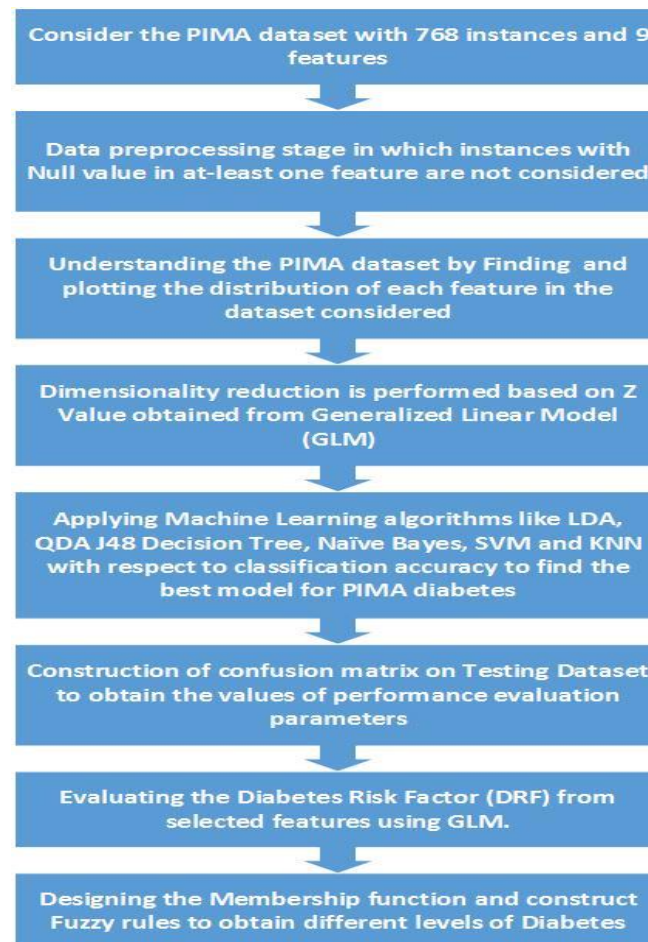
In [7] Hernandez-Contreras *et al.*, 2015 stated that early detection of symptoms related to ulceration risk in patients with diabetes mellitus can prevent future complications. And presented a preliminary classification method to detect temperature patterns which could be related with high-risk patients.

4.1. Simple Linear Regression

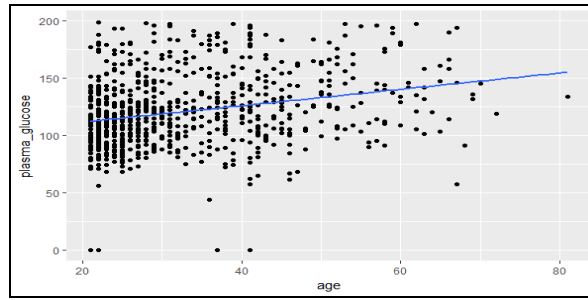PlasmaGC modeled by Age with fitted regression line:

Figure 6. Regression on Plasma_Glucose and Age attributes

The intercept of this linear model indicates that when age is equal to zero, blood glucose concentration has a value of 98 (units for glucose concentration were not specified in the dataset). The Age coefficient of .693 means that for every 1 year increase in age, blood glucose concentration increases by a value of .693. Finally, the adjusted R squared coefficient is .07, meaning that age in years explains 7% of the variation in plasma glucose concentration.
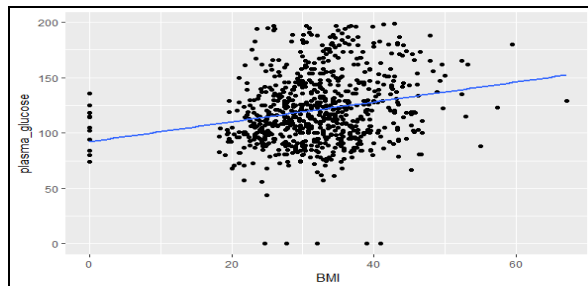

Figure 7. Regression on Plasma_Glucose and BMI Attributes

The intercept of PlasmaGC~BMI is 88.578, which means that when BMI is 0, PlasmaGC is 88.578. However, a BMI of 0 is not possible, so this intercept is not pratically meaningful. The model's BMI coefficient is 1.028, so for every one unit increase in BMI, PlasmaGC is expected to increase by 1.028 units. The adjusted R squared for this model is .0529, so the model only explains 5.29% of the variation in PlasmaGC.
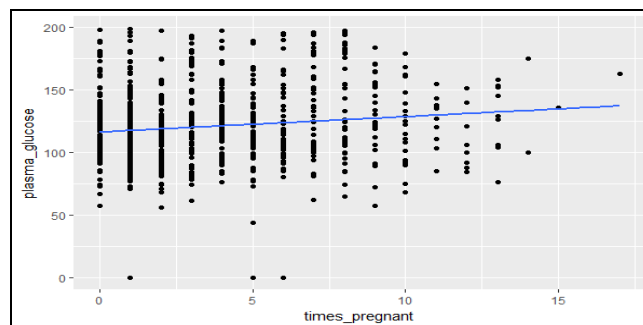

Figure 8. Regression on Plasma_Glucose and times_pregnant Attributes

My final linear regression model, PlasmaGC~Pregnancies, has an intercept coefficient that suggests that when a woman has had 0 pregnancies, her PlasmaGC will have a value of 117.221. The Pregnancies coefficient of 1.159 indicates a 1.159 unit change in PlasmaGC for each additional pregnancy. The adjusted R squared for

this model is .01513– number of pregnancies only accounts for 1.513% of the variation in PlasmaGC.

4.2. Multiple Regression

In the Figure 9, the intercept is 66.715: when Age, Pregnancies, and BMI are equal to zero, PlasmaGC is expected to have a value of 66.715. The explanatory variable Pregnancies has a coefficient of -.2359, so when we hold all other variables constant, each 1 unit change in number of pregnancies yields a -.2359 unit change in PlasmaGC. Data partitioned into training (70%) and testing (30%). Accuracy of 76.82% achieved on the testing dataset. The prediction does not the alternative of consulting the doctor.

```
Call:
lm(formula = plasma_glucose ~ age + times_pregnant + BMI, data = pima)

Residuals:
     Min       1Q   Median       3Q      Max
-132.136  -19.236   -2.392   18.423   83.509

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)     70.0554     5.4282  12.906  < 2e-16 ***
age              0.7240     0.1104   6.559 9.98e-11 ***
times_pregnant  -0.1827     0.3851  -0.474    0.635
BMI              0.8587     0.1382   6.216 8.39e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 30.15 on 764 degrees of freedom
Multiple R-squared:  0.1145,    Adjusted R-squared:  0.111
F-statistic: 32.93 on 3 and 764 DF,  p-value: < 2.2e-16
```

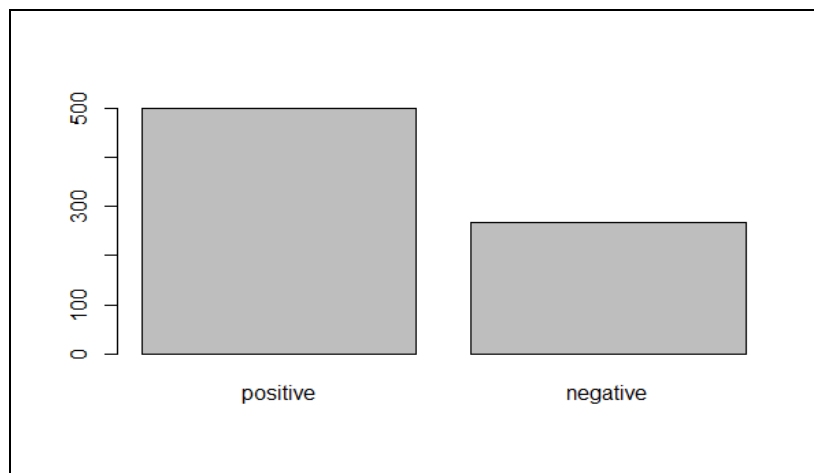**Figure 9. Linear Model Setup in R**

Figure 10. Plot of Estimated Prediction

There are 500 positive and 268 Negative cases considered in our dataset. The data in the dataset made valid.
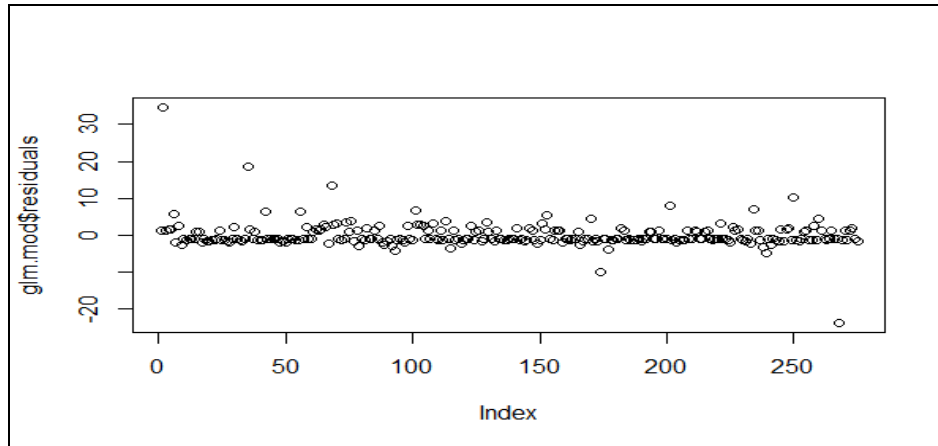
Figure 11. Plot of Generalized Linear Model Residuals
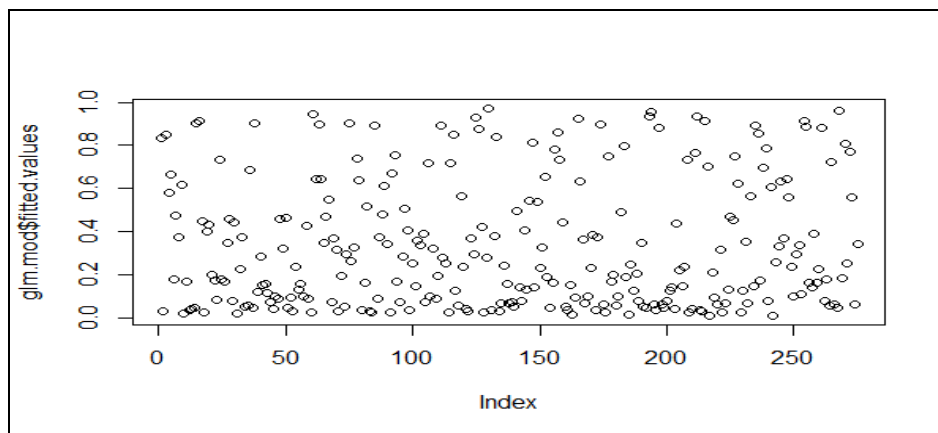
Figure 12. Plot of Generalized Linear Model Fitted Values

Table 2. Summary of Generalized Linear Model

Name of the attribute	Estimated	Std.Error	Z Value	Pr(>|Z|)
TIMES_PREGNANT	0.132891	0.065785	2.020	0.0434 *
PLASMA_GLUCOSE	0.044536	0.007483	5.951	2.66e-09 ***
DIASTOLIC_BP	0.007961	0.015471	0.515	0.6068
TRI_SKIN_FOLD	0.035468	0.020626	1.720	0.0855
INSULIN	-0.00133	0.001652	-0.805	0.4210
BMI	0.034125	0.032801	1.040	0.2982
D_PED	1.235846	0.512210	2.413	0.0158 *
AGE	0.007491	0.020277	0.369	0.7118

From Table 2 we can interpret that the value of Z is higher for PLASMA_GLUCOSE with good significance value and less standard error. In [8] Mathur *et al.*, 2015 provided different ways for improving patient care for amputees suffering from diabetes, by monitoring physiological data such as temperature, gait, heart rate, blood pressure, blood oxygenation (SpO2), respiration and glucose levels. Developed an e-health wearable communication systems delivering improvements in patient care whilst reducing both the demand for resources and the financial burden on healthcare systems.

Table 3. Diabetes Risk Factor

Voa1	Voa2	Voa3	Voa4	Voa5	Voa6	Voa7	Voa8	Diabetes Risk Factor
0	200	2	30	100	28	0.6	35	0.8092543
3	200	2	30	100	28	0.6	35	0.8634029
4	200	2	30	100	28	0.6	35	0.878333
0	200	2	30	100	28	2.42	35	0.9757411
0	200	2	30	100	28	1.42	35	0.9211856
1	140	2	30	100	28	1.42	35	0.4798416
1	100	2	30	100	28	0.42	35	0.04319204

Voa1=Value of attribute 1. From Table 3, it is proved that the change in Voa2 followed by Voa7 will drastically reduce (or) increase the diabetes risk factor.
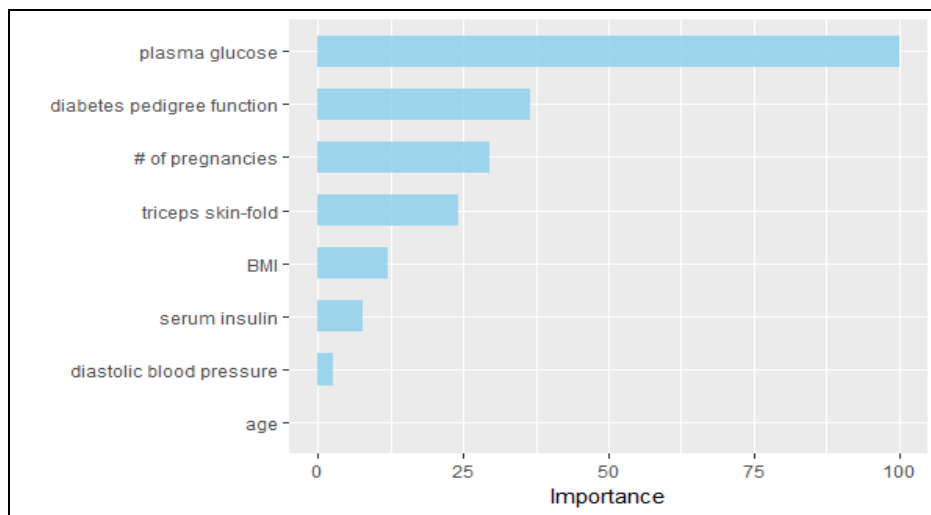
Figure 13. Important Feature Selection based on Z Value

In [9] Ugar *et al.*, 2016 address diabetes mellitus (diabetes) as health problem especially type 2 diabetes mellitus, according to result of the studies of Turkey diabetes apidemiology (TURDEP-I), type 2 diabetes mellitus %7.2 frequency of impaired glucose tolerance (IGT) %6.7 was found. In the studies of TURDEP-II that is published recently has been seen significantly increase of 2 type diabetes mellitus at age of 20 and over. Based on these information, the author aims to form an insulin pen specially designed for visually handicapped people.

Table 4. Confusion Matrix on 0.5 split on Dataset

N=384	Predicted False	Predicted True	
Actual False	212	51	263
Actual True	38	83	121
	250	134	384

In [10] Anggraini *et al.*, 2015 focused on developing mobile encyclopedia application for pregnancy and infant. In which, Forward chaining method was utilized as expert system in giving nutrient recommendation for pregnant mother and infant. Considered several health conditions as inference rules like: Diabetes, hypertension, vegetarian where employed for pregnant women inference rules, while for infant food the rules include baby's age and allergies. Totally, 8 rules for pregnancy nutrient and 52 rules for infant nutrient are employed.

Table 5. Evaluation of Statistical Parameters

Accuracy	0.7682
95% CI	(0.7227, 0.8095)
No Information Rate	0.651
P-Value [Acc > NIR]	4.391e-07
Kappa	0.4782
Mcnemar's Test P-Value	0.2034
Sensitivity	0.8480
Specificity	0.6194
Pos Pred Value	0.8061
Neg Pred Value	0.6860
Prevalence	0.6510
Detection Rate	0.5521
Detection Prevalence	0.6849
Balanced Accuracy	0.7337

5. Results and Discussion

In [13] Alfonso Arranz Martín *et al.*, 2017 reported the clinical characteristics of patients with latent autoimmune diabetes in adults (LADA), and to ascertain their metabolic control and associated chronic complications, data on LADA diagnosis, associated autoimmunity, C-peptide levels, therapeutic regimen, metabolic control, were analyzed. Conducted experiments having, Number of patients (193), In which 56% females. Family history of DM: 62%. Age at DM diagnosis: 49 years. Delay in confirmation of LADA: 3.5 years. Insulin-independence time: 12 months. Baseline serum C-peptide levels: 0.66 ng/ml. Basal-bolus regimen: 76.7%. Total daily dose: 35.1 U/day, corresponding to 0.51 U/kg. With no associated oral antidiabetic drugs: 33.5%. Other autoimmune diseases: 57%. Fasting plasma glucose: 160.5 mg/dL. HbA1c: 7.7%. BMI: 25.4 kg/m2 (overweight, 31.5%; obesity, 8%). Blood pressure: 128/75. HDL cholesterol: 65 mg/dL. LDL cholesterol: 96 mg/dL. Triglycerides: 89 mg/dL. Known chronic complications: 28%. And Concluded that, there is a heterogeneous pancreatic insulin reserve which is negative related to glycemic parameters.

In [11] Litinskaia *et al.*, 2017 aims to improving accuracy of non-invasive blood glucose detection via optical glucometer by more than 2 times by providing a feedback between glucometer and insulin pump and developed mathematical model of blood glucose dynamics on the basis of blood glucose prediction algorithm.
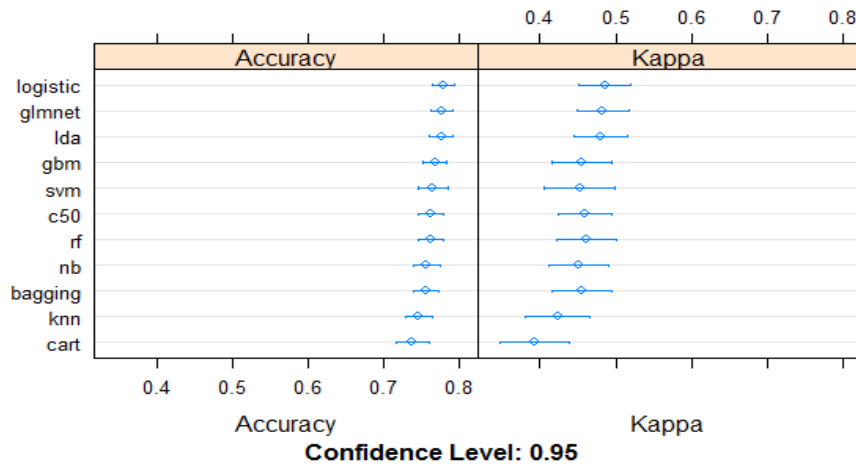
Figure 14. Plot of Kappa and Accuracy Parameters

Table 6. Comparative Study on Predictive Algorithms

lda	logistic	glmnet	svm	knn	nb	cart	c50	bagging	rf	gbm	Resample
0.78	0.77	0.77	0.79	0.70	0.78	0.74	0.82	0.74	0.79	0.73	Fold01.Rep1
0.81	0.79	0.79	0.83	0.79	0.82	0.81	0.75	0.82	0.81	0.82	Fold01.Rep2
0.78	0.78	0.78	0.76	0.71	0.80	0.72	0.66	0.72	0.74	0.75	Fold01.Rep3
0.83	0.82	0.82	0.81	0.79	0.79	0.84	0.74	0.83	0.82	0.84	Fold02.Rep1
0.77	0.78	0.78	0.79	0.71	0.78	0.77	0.75	0.73	0.75	0.79	Fold02.Rep2
0.81	0.81	0.81	0.78	0.81	0.75	0.81	0.74	0.79	0.82	0.77	Fold02.Rep3
0.71	0.72	0.72	0.67	0.68	0.71	0.71	0.82	0.72	0.68	0.71	Fold03.Rep1
0.74	0.74	0.74	0.74	0.77	0.78	0.82	0.78	0.81	0.77	0.77	Fold03.Rep2
0.77	0.78	0.78	0.78	0.71	0.73	0.77	0.82	0.74	0.77	0.79	Fold03.Rep3
0.82	0.82	0.82	0.78	0.71	0.77	0.64	0.83	0.74	0.73	0.78	Fold04.Rep1
0.67	0.68	0.68	0.68	0.71	0.63	0.63	0.79	0.68	0.68	0.67	Fold04.Rep2
0.75	0.74	0.74	0.76	0.74	0.75	0.67	0.75	0.76	0.76	0.79	Fold04.Rep3
0.75	0.75	0.75	0.74	0.77	0.77	0.78	0.77	0.73	0.74	0.81	Fold05.Rep1
0.81	0.79	0.79	0.78	0.73	0.74	0.69	0.75	0.73	0.75	0.78	Fold05.Rep2
0.81	0.81	0.81	0.75	0.68	0.70	0.74	0.73	0.68	0.73	0.73	Fold05.Rep3
0.75	0.76	0.76	0.72	0.78	0.68	0.71	0.71	0.79	0.76	0.75	Fold06.Rep1

|											
0.75	0.77	0.77	0.74	0.70	0.74	0.71	0.73	0.73	0.71	0.73	Fold06.Rep2
0.86	0.86	0.86	0.87	0.79	0.83	0.79	0.73	0.82	0.82	0.84	Fold06.Rep3
0.77	0.77	0.77	0.73	0.82	0.75	0.71	0.77	0.77	0.74	0.77	Fold07.Rep1
0.75	0.77	0.75	0.70	0.71	0.69	0.62	0.70	0.74	0.77	0.75	Fold07.Rep2
0.77	0.78	0.77	0.74	0.70	0.75	0.77	0.82	0.77	0.73	0.77	Fold07.Rep3
0.75	0.74	0.74	0.75	0.75	0.75	0.74	0.77	0.74	0.74	0.75	Fold08.Rep1
0.78	0.79	0.79	0.76	0.78	0.79	0.76	0.70	0.79	0.83	0.79	Fold08.Rep2
0.75	0.74	0.74	0.70	0.69	0.70	0.74	0.71	0.74	0.71	0.71	Fold08.Rep3
0.83	0.83	0.83	0.82	0.71	0.81	0.74	0.75	0.82	0.84	0.83	Fold09.Rep1
0.75	0.77	0.77	0.75	0.74	0.71	0.64	0.79	0.70	0.73	0.73	Fold09.Rep2
0.75	0.77	0.77	0.71	0.74	0.77	0.74	0.86	0.71	0.77	0.75	Fold09.Rep3
0.77	0.77	0.77	0.75	0.78	0.78	0.73	0.79	0.73	0.74	0.73	Fold10.Rep1
0.87	0.87	0.87	0.90	0.90	0.86	0.82	0.76	0.87	0.86	0.84	Fold10.Rep2
0.79	0.81	0.81	0.84	0.79	0.79	0.79	0.80	0.75	0.79	0.78	Fold10.Rep3
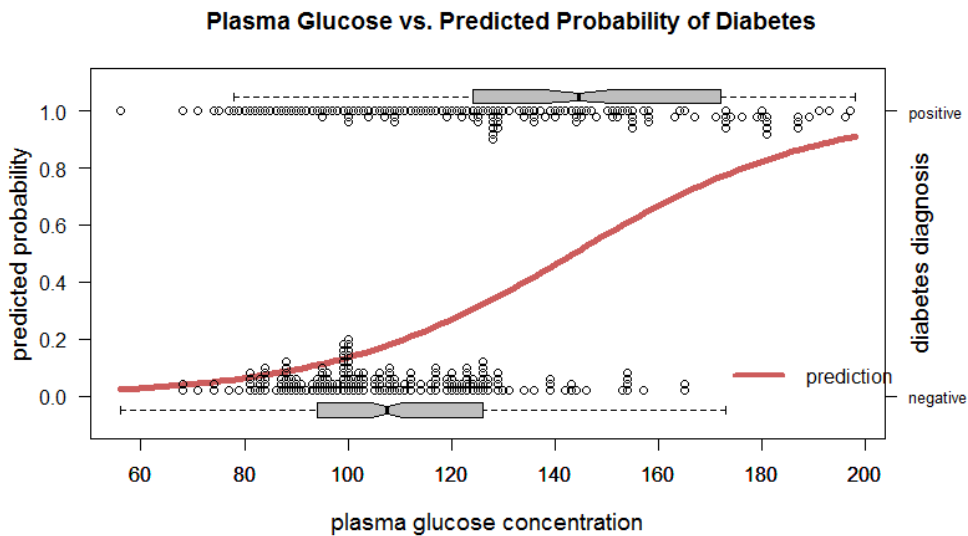
Figure 15. Plot of Generalized Linear Model Prediction

The train function in R programming Language, sets up a grid of tuning parameters for a number of classification and regression routines, fits generalized linear model, and calculates a re-sampling based performance measure. *Diagnosis = Times_Pregnant + Plasma_Glucose + Diastolic_BP + Tri_Skin_Fold + Insulin + BMI + D_Ped + Age*

6. Conclusion

Considering these results, we can infer that, every model has more than 70% of accuracy and less than 50% of kappa value which is high. Likewise, because of the voting process of all the algorithms, it ensures that the final diagnose is very accurate. Besides, we have planned to gather more information from various districts over the nation and grow more exact and general farsighted model.

References

[1] IEEE Draft Health Informatics - Personal Health Device Communication Part 10417: Device Specialization - Glucose Meter," in ISO/IEC/IEEE P11073-10417/D2 January 2015, vol. 1, no. 10, **(2015)** March 11, pp. 100-110.

[2] E. S. Ibrahim, S. Swanson, J. Stojanovska, C. Duvernoy and R. Pop-Busui, "Harmonic phase versus sine-wave modulation for measuring regional heart function from tagged MRI images", 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, **(2016)**, pp. 444-447.

[3] J. Li and J. Kong, "Cell phone-based diabetes self-management and social networking system for American Indians", 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Munich, **(2016)**, pp. 1-6.

[4] K. Knorr and D. Aspinall, "Security testing for Android mHealth apps", 2015 IEEE Eighth International Conference on Software Testing, Verification and Validation Workshops (ICSTW), Graz, **(2015)**, pp. 1-8.

[5] H. Zhang, "A mobile health solution for chronic disease management at retail pharmacy", 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Munich, **(2016)**, pp. 1-5.

[6] Z. Darabi, M. H. F. Zarandi, S. S. Solgi and I. B. Turksen, "An intelligent multi-agent system architecture for enhancing self-management of type 2 diabetic patients", 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Niagara Falls, **(2015)**, pp. 1-8.

[7] D. Hernandez-Contreras, H. Peregrina-Barreto, J. Rangel-Magdaleno, J. Ramirez-Cortes, F. Renero-Carrillo and G. Avina-Cervantes, "Evaluation of thermal patterns and distribution applied to the study of diabetic foot", 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, Pisa, **(2015)**, pp. 482-487.

[8] N. Mathur, I. Glesk and A. Buis, "Issues in wearable mobile sensor platform for lower limb prosthetic users", 2015 17th International Conference on Transparent Optical Networks (ICTON), Budapest, **(2015)**, pp. 1-4.

[9] T. Ugar, O. Kogak and A. Akpek, "New concept design of an insulin pen for visually impaired or blind diabetius mellitus patients", 2016 Medical Technologies National Congress (TIPTEKNO), Antalya, **(2016)**, pp. 1-4.

[10] R. N. E. Anggraini, A. R. Soedjono, F. Y. Sianipar and S. Rochimah, "Infant and pregnancy encyclopedia application", 2015 International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA), Surabaya, **(2015)**, pp. 91-95.

[11] E. L. Litinskaia, N. A. Bazaev, K. V. Pozhar and V. M. Grinvald, "Methods for improving accuracy of non-invasive blood glucose detection via optical glucometer", 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), St. Petersburg, **(2017)**, pp. 47-49.

[12] E. Deepak Chowdary, K. A. Devi, D. Mounika, S. Venkatramaphanikumar and K. V. K. Kishore, "Ensemble classification technique to detect stress in IT-professionals", 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, **(2016)**, pp. 1-6.

[13] A. Arranz Martín, E. Lecumberri Pascual, M. Ángel Brito Sanfiel, V. Andía Melero, L. Nattero Chavez, I. Sánchez López, G. Cánovas Molina, F. Arrieta Blanco and N. González Perez del Villar, "Clinical and metabolic profile of patients with latent autoimmune diabetes in adults in specialized care in Madrid", Endocrinología, Diabetes y Nutrición (English ed.), vol. 64, no. 1, **(2017)** January, pp. 34-39.

[14] S. Alzahrani, J. Nelson, S. F. Moss, J. K. Paulus, W. C. Knowler and A. G. Pittas, "H. pylori seroprevalence and risk of diabetes: an ancillary case–control study nested in the Diabetes Prevention Program", Journal of Diabetes and its Complications, Available online **(2017)** June 3.

[15] T. Mashitani, Y. Hayashino, S. Okamura, M. Kitatani, M. Furuya, T. Iburi, S. Tsujii and H. Ishii, "Association between dipstick hematuria and decline in estimated glomerular filtration rate among Japanese patients with type 2 diabetes: A prospective cohort study [Diabetes Distress and Care Registry at Tenri (DDCRT 14)]", Journal of Diabetes and its Complications, vol. 31, no. 7, **(2017)** July, pp. 1079-1084.

[16] M. Hassanein, M. Al-Arouj, O. Hamdy, W. Mohamad, W. Bebakar, A. Jabbar, A. Al-Madani, W. Hanif, N. Lessan, A. Basit, K. Tayeb, M. Omar, K. Abdallah, A. Al Twaim, M. Akif Buyukbese, A. A. El-Sayed and A. Ben-Nakhi, "Diabetes and Ramadan: Practical guidelines", Diabetes Research and Clinical Practice, vol. 126, **(2017)** April, pp. 303-316.

Authors

Syed Muzamil Basha, he had his Bachelor of Science in Information Technology at SITAMS, MTech in Information Technology (Networking) at VIT University and currently doing his research at VIT University. His research area are Wireless Sensor Networks, Text Mining and Big Data Predictive Analytics.

Dr. H. Bajaj currently working as an Associate Professor in Computer Science and Engineering Department at Sreenidhi Institute of Science and Technology, Hyderabad, Telangana India. His areas of research include Data Warehousing and Mining and Network Security.

N. Ch. S. N. Iyengar (b 1961), he currently Professor at the *Sreenidhi Institute of Science and Technology (SNIST) Yamnapet, Ghatkesr, Hyderabad, Telengana, India.* His research interests include Agent-Based Distributed Computing, Intelligent Computing, Network Security, Secured Cloud Computing and Fluid Mechanics. He had 32+ years of experience in teaching and research, guided many scholars, has authored several textbooks and had nearly 200+ research publications in reputed peer reviewed international journals. He served as PCM/reviewer/keynote speaker/ Invited speaker.

Ronnie D. Caytiles, he had his Bachelor of Science in Computer Engineering- Western Institute of Technology, Iloilo City, Philippines, and Master of Science in Computer Science– Central Philippine University, Iloilo City, Philippines. He finished his Ph.D. in Multimedia Engineering, Hannam University, Daejeon, Korea. Currently, he serves as an Assistant Professor at Multimedia Engineering department, Hannam University, Daejeon, Korea. His research interests include Mobile Computing, Multimedia Communication, Information Technology Security, Ubiquitous Computing, Control and Automation.