European Journal of Molecular & Clinical Medicine

ISSN 2515-8260 Volume 07, Issue 10, 2020

Machine Learning Approaches For Predicting Patient Severity Levels In T2dm Complications Neuropathy

R. Usharani[1], A. Shanthini[2]

[1]*Research Scholar, Faculty of Engineering & Technology, SRM Institute of Science & Technology, Chennai, Tamil Nadu, India.*
Email: usharanimurugesh@gmail.com
[2]*Associate Professor, Department of Information Technology, SRM Institute of Science & Technology, Chennai, Tamil Nadu, India.*
Email: shanthia@srmist.edu.in

1. INTRODUCTION

The most common complication of both type 1 and 2 diabetes is diabetic neuropathy (DN), a type of nerve damage caused by long-term elevated glucose levels, which occurs in more than half of the individuals affected. Over the span of several decades, the disorder normally progresses progressively and occasionally. Many studies have found that the quality of life of those living with DPN has had a substantial negative effect. In comparison, a significantly improved Mortality for people afflicted with peripheral neuropathy (DPN) with diabetes who have undergone a major amputation. This needs decisive action to tackle this growing public health problem **[1].** The research also uses the algorithms RF, SVM, KNN and NB to search for accuracy and indicates that in classifying the complexity of the T2DM, SVM has the high est accuracy **[2].** Diabetes is a significant metabolic disease that can harmfully annoy patients. It may lead to multiple complications such as heart stroke, diabetic nephropathy and other conditions, if not diagnosed. In keeping a safe life, early diagnosis will help. As cases of diabetes are growing increasingly, these diseases are a cause for global concern. This research also uses the PIMA database and monitors the trend to identify patterns using R-Supervised MLAs-Linear Kernel SVM, RBF-Kernel SVM, KNN, ANN and Reduction of Multifactor Dimensionality (MDR) **[3] [4].** The aim of this analysis was to compare the outputs of SVM, Naïve Net, Decision Stump, and Proposed Ensemble process models of predicting Diabetes using common risk factors. The findings show that the Suggested Ensemble Approach worked well in terms of accuracy of classification **[5].** The number of diabetic patients (characterised by hyperglycemia) in the world is predicted to exceed 642 million in 2024 and requires a great deal of care. So, it is important to use the effective methods to anticipate and recommend to medical practitioners the proper approach.

ISSN 2515-8260 Volume 07, Issue 10, 2020

2. RELATED WORKS

This study finding the correlation among those factors using heatmap visualization techniques in Machine Learning. Many researchers have conducting analysis on the performance of various algorithms and similar approaches to find potential gaps in this area of analysis. This chapter summarises several important works. First of all to make the data correct, relevant and accessible, it is most important to specify the parameters that will assist the study in determining the source of the DPN. Of the 20 features recorded for each patient, 13 were found to impact the progress of the DM patient towards DPN. Age, type of diabetes, degree of education, BMI, blood pressure history, systolic blood pressure, foot ulcer history, drug, weight, history of laser photocoagulation, length, average blood glucose and height were the 13 characteristics identified. **[6, 7]**.

The following Table 1 depicts the various significant factors associated with DPN.

Significant Risk Factors associated with DPN	**References**
Age	[8-11]
BMI (obesity)	[8][13][19]
Blood Glucose	[8][18][19]
HbA1c	[14][18]
Triglycerides	[19-21]
Blood Pressure	[22-24]
Urea Serum Creatinine	[25, 26]
GFR	][27,28]
Hyperglycemia	[29-32]
Microalbuminuria	[33-37]
DiabetesMellitus and its duration of years	[10][38-40]

Table 1: Significant Risk Factors associated with DPN along with references

- **Age and risk of DPN**

The impact of age on diabetic peripheral neuropathy has been described in [8-11] all of the article's reports and they have shown revealed substantial age gaps through the univariate analysis and the multivariate analysis.

- **BMI (Obesity) and risk of DPN**

The effects of BMI on DPN has been analysed in all the studies **[8][13][19].** Obesity has been described as a risk factor in DPN in the Southern German population. Obesity and the existence of at least 2 cardiovascular risk factors (triglycerides or plasma glucose, decreased HDL, increased waist circumference, hypertension) in the general US population aged 40 years or older raise the possibility of peripheral neuropathy (OR: 2.20, 95 percent CI: 1.43-3.39).

ISSN 2515-8260 Volume 07, Issue 10, 2020

- **Blood Glucose and risk of DPN**

Fasting plasma glucose (FPG) **[8][18][19]** High blood sugar (glucose) can injure nerves throughout your body. The results of the studies have been established a syndromic link between neuropathy associated with impaired glucose tolerance and the sensory-predominant neuropathy commonly observed in early diabetes.

- **HbA1c and Risk of DPN**

The results of the Nisar et al and Fargol Booya et al., [16][17] studies confirm reports regarding the association of neuropathy with glycemic control (HbA1C), age, BMI and duration of disease.

- **Triglycerides and Risk of DPN**

The studies **[19][21]** have shown that the presence of triglycerides in diabetics' normal blood functions raises the risk of neurological damage.

- **Blood Pressure and Risk of DPN**

Federica Di Gennaro et al., and Vincenza Spallone assesses the clinical correlates of Morning Blood Pressure Surge (MBPS) in a diabetic population, with particular attention paid to diabetic complications and diabetic neuropathy. Also found that the framework highlights the relevance of assessing BP variability in people with diabetes. These research reports confirm that vascular disease is associated with increased MBPS diabetes in the population, diastolic blood pressure as well as cardiovascular autoimmune disease (CAN) [**22][24]**

- **Urea Serum Creatinine and Risk of DPN**

Diabetic patients with high urea levels, high cholesterol and triglycerides levels should therefore be considered at higher risk for peripheral neuropathy [25]. The researchers found that if the incidence of DPN developed, abnormal lipid profile, elevated urea and reduced RBC levels point to the coexistence of cardiovascular and renal comorbid conditions [26].

- **GFR and Risk of DPN**

The variety of neurologic conditions affecting the central nervous system and the peripheral nervous system are caused by chronic renal failure. The kidneys do not perform well as they should, frequently due to a mismatch of salts and chemicals. The risks of damages to the peripheral nerve may be raised by this mismatch, leading to neuropathy. The research was exploring the relationship between the Glomerular Filtration Rate (GFR) and microvascular complications in type II diabetes mellitus such as diabetic neuropathy and nephropathy Patients of (DM) [27,28]

- **Hyperglycemia and risk of DPN**

The central nervous system (CNS) damage related to hyperglycemia. The obvious hyperglycemia present in diabetes can development of abnormalities and it has been discussed the evidence for insulinopenia in type 1 and insulin resistance in type 2 diabetes as causal factors in the development of DPN. Moreover, data suggest that other factors may also

ISSN 2515-8260 Volume 07, Issue 10, 2020

contribute. The study suggested that these two cases actually represent a single cause of impaired insulin signalling [29-32].

- **Microalbuminuria and risk of DPN**

Diabetic neuropathy observed in this study have been using distal symmetrical sensory motor neuropathy. This design of this study revealed that microalbuminuria can be an essential cause for microvascular complications in type 2 diabetes. The authors concluded that the incidence of microalbuminuria in patients with type-2 diabetes is considered significant. More it also indicates an early stage of diabetic nephropathy. Therefore, strict control of hyperglycemia and hypertension should deal early. It helps to prevent the other host of diabetes related complications occurring in future. Hence, it is clear that microalbuminuria is significantly related to the neurological presence [33-37].

- **Duration of Diabetes and Risk of DPN**

The research framework [10] [38-40] assessed Diabetic neuropathy is strongly associated with duration of diabetes and other factors such as HbA1c and GFR.

- **Heatmap Visualization Technique**

An expansion to the Heat Map has been presented in this work. Algorithm Based Feature Selection. This facilitates the automated collection of threshold parameters that help to enhance the efficiency of generalisation of high-dimensional data such as mass spectrometry. Using several cancer datasets, the authors have conducted a comparative analysis and compared the well-known Recursive Feature Elimination algorithm. The findings showed better efficiency of the grouping, which is very comparable with other related tests. Because of these results, it is possible to create datasets of visual images that are mapped to the original space by generating a heatmap representation of the data, and this will aid in the quest for regions of interest, not just from the point of view of the feature selection, but also for instance selection [41]. The authors illustrated the efficiency of Complex Heatmap to easily expose trends and links between different sources. Four databases of information related to multidimensional genomic evidence in the real world [42]. A new local dimension reduction algorithm suggested by the researchers was used in feature discovery and extraction of features. This article concentrates on extracting the high-dimensional hidden potential structure inMicroarray data for various categories, and explanation and comprehension. The effects of the potential structure information provided with the heatmap graph are also used. The suggested algorithm for the exploration of gene co-expression and coregulation can be efficiently extended to microarray data analysis [43]. A novel research analysis, univariate feature selection, feature importance, and correlation matrix with heat map were applied to find the optimum data subset of erythematosquamous disease in three feature extraction techniques. For calculating model efficiency, four classification techniques are used: Gaussian Naïve Bayesian (NB), Decision Tree (DT), Support Vector Machine (SVM), and Random Forest. The stacking ensemble approach is then introduced to boost the model's predictive efficiency [4]. The heat map image that reveals 11 distinct proteins for HT carcinoma cells that assist in cell survival/apoptosis was considered in this study. Three kinds

of feature selection techniques namely filter, wrapper, and embedded were used and obtain seven marker protiens. But only one of these protiens was used for classification using the KNN and SVM classifiers and yield the results from these protiens [44].

From the related works and the American Diabetes Association recommends aiming for the normal and abnormal range of blood sugar level, Hemoglobin levels, GFR etc which are shown in the following Table 2.

Diabetes Mellitus Factors	Ranges	
	Normal	Abnormal
Blood Glucose	<140 mg/dL	>140 mg/dL
Fasting Plasma Glucose (Pre-breakfast) 2-hours in an oral glucose tolerance test	70-130 mg/dL	>130 mg/dL
Post breakfast Glucose	130-180 mg/dL	>180 mg/dL
HbA1C	5.7 % mg/dL	>6.5 % mg/dL
Age	30-90 (in years)	---
BMI (Body Mass Index)	18-25 kg/m 2	>25 kg/m 2
Triglycerides	<150 mg/dL	150-199 mg/dL
Blood Pressure (BP)	130/80 mm hg	below or higher the normal range at risk
Urea-serum-creatinine	0.6 to 1.4 mg/dL	>1.4 mg/dL
Glomerular Filteration Rate (GFR)	Male: 97 to 137 mL/min Female: 88 to 128 mL/min	Male: >137 mL/min Female: > 128 mL/min
Gender (Male/Female)	---	---
hyperglycemia-Symptoms	No(0)	Yes(1)
hypoglycemic_symptoms	No(0)	-1
microalbuminuria	30 -299 mg	>=300 mg
Diabetes Mellitus (Duration of Years)	<5 years with control	>5 years & elevated level

Table 2: Diabetes Mellitus risk factors ranges according to American Diabetes Association and through Review of various journals mentioned in the Related Works.

These research framework findings confirm that if their diabetes continues for a long time, patients will soon develop a neurological disorder and other diabetic complications.

ISSN 2515-8260 Volume 07, Issue 10, 2020

3. METHODOLOGY

The following diagrammatic representation figure 1 depicts an overview of a general system framework for proposed work

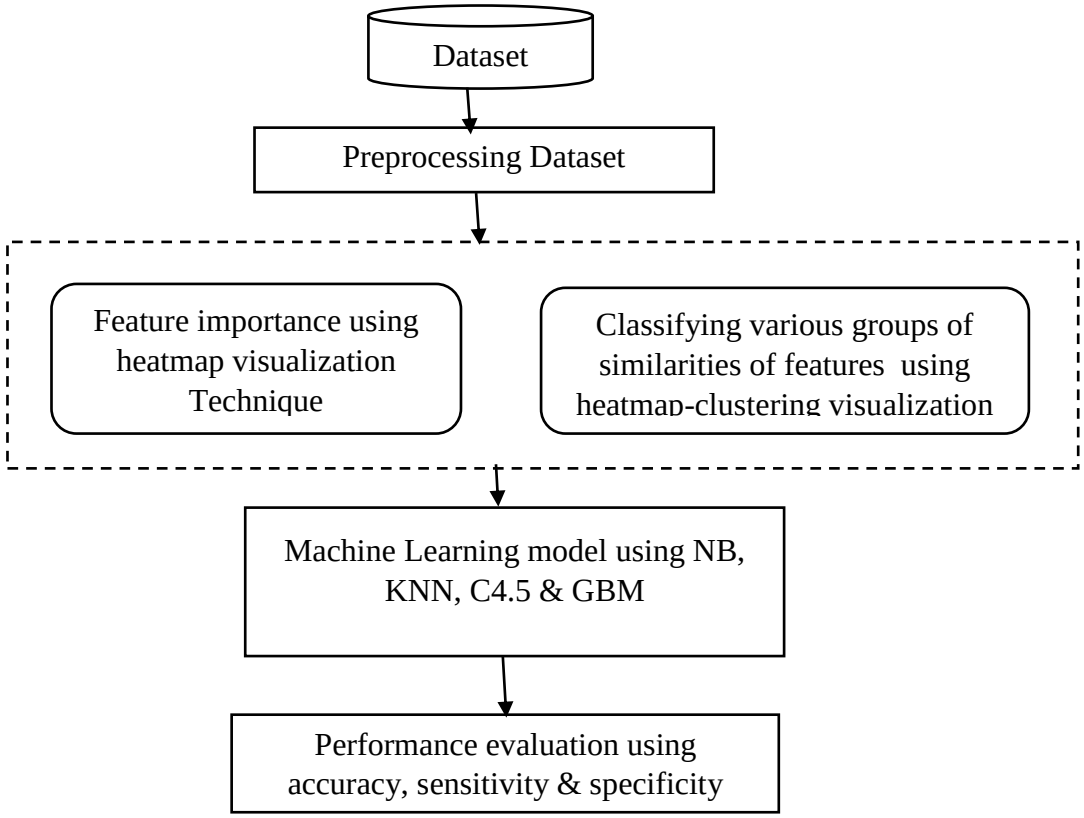

Figure 1. Flowchart of general system framework for proposed work

A. Feature Selection with Heatmap Visualization

Heatmap gives the opportunity to obtain useful information insights from the extensive datasets. It's the perfect method for data exploration strategy. The strong and weak correlations that can distinguish better using this. Simplify the exclusion of obsolete data as well. It may also be possible to do multicollinearity testing. The cell colour in the heatmap relates to the value. It is the best multivariate regression technique. The hue of the gradient allows the values in between to vanish, concentrating on the average [41]. When two variables are correlated, changes in their values have the same cause or an analogous explanation. The growth of one component may be the direct cause of a decrease of another factor in a negative correlation. A weak correlation means that there is a smaller chance of there being an association with the second variable as one variable increases or decreases. Labelling the frequency of the relationship, 0- 0.39 is known to be weak for absolute values of Pearson r, 0.40-0.69 to be mild or moderate, >0.7 to be high and 0.8-1to be a very strong correlation and the background of the findings should be considered. The negative correlated variables were not required for further analysis of the DPN forecast from the heatmap results

ISSN 2515-8260 Volume 07, Issue 10, 2020

of this analysis. Weak, average and high associations between DPN predictor factors were considered in this analysis. For example, from the figure 1, pre_breakfast_bloodGlucose and blood_Glucose, Microalbuminuria and urea-creatinine, GFR and urea-creatinine associations have been the strong chances the patients affected by DPN. Also, the correlations Triglycerides association with BMI, BMI connection with HbA1c were also moderately likely to increase the DPN.

Likewise, the interaction of BP with any BloodGlucose, Age and Triglycerides was rarely likely to raise the DPN. The table 3 summarizes the correlation features and its value that is associated with DPN provided by the heatmap graph.

The following figure 2 depict the Heatmap Visualization Graphs for DPN factors severity

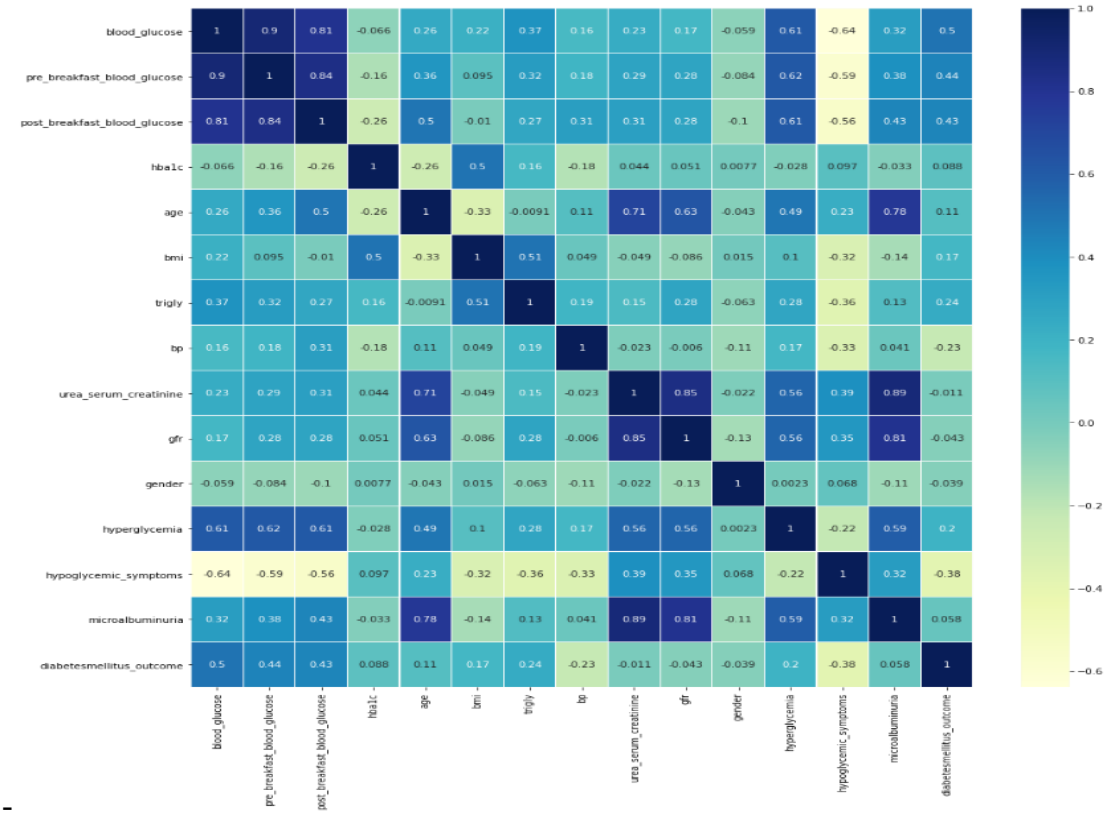

Figure 2. DPN risk factors visualization using Heatmap

Correlation Features of DPN	**Correlation Values**
Pre_breakfast_bloodGlucose & Blood_Glucose	0.91
Micro_albuminuria & Urea_Serum_Creatinine	0.89
GFR & Urea_Serum_Creatinine	0.85
Post_breakfast_bloodGlucose & Pre_breakfast_bloodGlucose	0.84
GFR & Micro_albuminuria	0.81
Post_breakfast_bloodGlucose & BloodGlucose	0.81
Micro_albuminuria & Age	0.78
Urea_Serum_Creatinine & Age	0.71

ISSN 2515-8260 Volume 07, Issue 10, 2020

GFR & Age	0.63
Hyperglycemia & BloodGlucose	0.62
Hyperglycemia & Post-breakfast_bloodGlucose	0.61
Micro_albuminuria & Hyperglycemia	0.59
Hyperglycemia & urea_serum_creatinine	0.56
Hyperglycemia & GFR	0.56
Triglycerides & BMI	0.51
BMI & HbA1C	0.50
Age & Post_breakfast_bloodGlucose	0.50
DM_Duration & bloodGlucose	0.50
Hyperglycemia & Age	0.49
DM_Duration & Pre_breakfast_bloodGlucose	0.44
DM_Duration & Post_breakfast_bloodGlucose	0.43
Microalbuminuria & Post_breakfast_bloodGlucose	0.43
Hypoglycemia & Urea_serum_creatinine	0.39
Micrialbuminuria & pre-breakfast-bloodGlucose	0.38
Triglycerides & BloodGlucose	0.37
Age & Prebreakfast_bloodGlucose	0.36
Hypoglycemia & GFR	0.35
Triglycerides & Pre_breakfast_bloodGlucose	0.32
Microalbuminuria & Hypoglycemia	0.32
Microalbuminuria & Blood Glucose	0.32
BloodPressure & Post-breakfast_bloodGlucose	0.31
Urea_Serum_creatinine & Post-breakfast_bloodGlucose	0.31
Urea_Serum_creatinine & Pre-breakfast_bloodGlucose	0.29
GFR & Pre-breakfast_bloodGlucose	0.28
GFR & Triglycerides	0.28
Triglycerides & Post-breakfastGlucose	0.27
Age & BloodGlucose	0.26
DM_Duration & Triglycerides	0.24
BloodPressure& Triglycerides	0.19
BloodPressure & Pre_breakfast_BloodGlucose	0.18
BloodPressure & Blood Glucose	0.16
BloodPressure & Age	0.10

Table 3. Correlation and Heatmap feature important table

B. Classification of DPN severity levels using Heatmap with Hierarchical Clustering

Heatmaps are related with hierarchical clustering as well. The rows or columns are organised based on similarity by hierarchical clustering. This makes comparisons convenient to see in the data. Heat maps also come with dendrograms. It is a diagrammatic representation the hierarchical association of objects [42, 43]. It is generated most commonly as an output from

ISSN 2515-8260 Volume 07, Issue 10, 2020

hierarchical clustering. This is used to find out the two parameters from the different variations are the most similar and merge them into a cluster.

The figure 3 depicts the pictorial visualization of variety of similarity levels of DPN factors. There are three clusters formed from the given DPN parameters. The clusters with dendrogram and its variety of colour bright-fading with values shown DPN factors that are high severity level group, moderate level group and lower-level affected group as it has been differentiated.

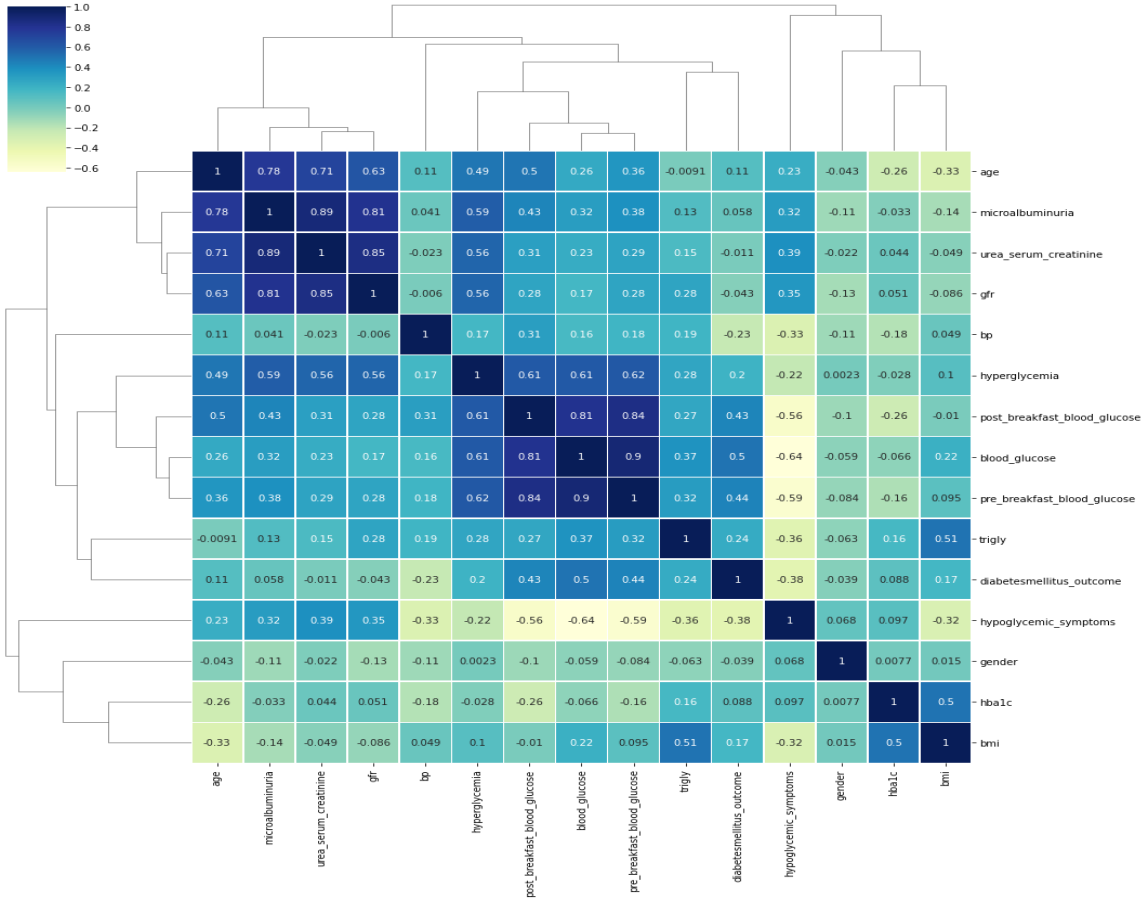

Figure 3. Various group of DPN Criticality level and its risk factors visualization using Cluster-Heatmap

The following table 4 shows a description of the very critical categories of features provided by the cluster of Heatmaps.

Mild -Level (0.1-0.39)	**Moderate-Level (0.4-0.69)**	**Critical-Level (0.7-0.99)**
Hypoglycemia & Urea_serum_creatinine	GFR & Age	Pre_breakfast_bloodGlucose & Blood_Glucose
Micrialbuminuria & pre-breakfast-bloodGlucose	Hyperglycemia & BloodGlucose	Micro_albuminuria & Urea_Serum_Creatinine
Triglycerides & BloodGlucose	Hyperglycemia & Post-breakfast_bloodGlucose	GFR & Urea_Serum_Creatinine
Age &	Micro_albuminuria &	Post_breakfast_bloodGlucose

ISSN 2515-8260 Volume 07, Issue 10, 2020

Prebreakfast_bloodGlucose	Hyperglycemia	& Pre_breakfast_bloodGlucose
Hypoglycemia & GFR	Hyperglycemia & urea_serum_creatinine	GFR & Micro_albuminuria
Triglycerides & Pre_breakfast_bloodGlucose	Hyperglycemia & GFR	Post_breakfast_bloodGlucose & BloodGlucose
Microalbuminuria & Hypoglycemia	Triglycerides & BMI	Micro_albuminuria & Age
Microalbuminuria & Blood Glucose	BMI & HbA1C	Urea_Serum_Creatinine & Age
BloodPressure & Post-breakfast_bloodGlucose	Age & Post_breakfast_bloodGlucose	
Urea_Serum_creatinine & Post-breakfast_bloodGlucose	DM_Duration & bloodGlucose	
Urea_Serum_creatinine & Pre-breakfast_bloodGlucose	Hyperglycemia & Age	
GFR & Pre-breakfast_bloodGlucose	DM_Duration & Pre_breakfast_bloodGlucose	
GFR & Triglycerides	DM_Duration & Post_breakfast_bloodGlucose	
Triglycerides & Post-breakfastGlucose	Microalbuminuria & Post_breakfast_bloodGlucose	
Age & BloodGlucose		
DM_Duration & Triglycerides		
BloodPressure& Triglycerides		
BloodPressure & Pre_breakfast_BloodGlucose		
BloodPressure & Blood Glucose		
BloodPressure & Age		

Table 4. Summary of greatly important categories of features provided by the heatmap cluster.

C. Predictive Analytics Algorithms for DPN:

After selecting the best feature subset, we used various machine learning algorithms to build a predictive model. In this research, we used different algorithms such as Naïve Bayes, C4.5, K-Nearest Neighbor (KNN) and GBM using the programming language Python.

Naïve Bayes:

Naïve Bayes is a fast and eager learning classifier. By using it, forecasts can be made in real time. It helps to capture the complexity of the model in a principled manner by assessing the probabilities of the findings. By doing this, diagnosis and statistical problems can be overcome. It can measure probabilities by measuring the frequency and variations of values in the results [2] [4].

KNN:

It is categorised under the lazy prediction approach. On the basis of comparisons, the simplified methodology encourages new work to be coordinated. The training data was organised into this algorithm. Define k - the number of neighbours close by. Among Distance Away Training and samples. Inaccessibility calculation of the learning sections were grouped and the nearby neighbour was Focused on the minimum - the distance is fixed in the successive phase [44, 45].

GBM:

Gradient boosting is a form of boosting machine learning. It is based on the intuition that when paired with previous ones, the next best possible model minimises the total prediction error. The error is minimised by projections from the current model that are similar to its targets [46].

C4.5:

C4.5 Algorithm extended features of ID3 algorithm suggested by Ross Quinlan et.al al. [46]. The decision tree uses the same knowledge for training as ID3, which contains the learning function. The method of learning should be used for medical data diagnosis in order to predict the importance of the decision. C4.5 in each of the tree divisions nodes, Choose the most efficient data attribute value that separates the data evaluated into class-enriching sub-set data. The forest is created by organised awareness. To make the highest value attribute decision, the normalised data gain is chosen from the C4.5 tree of judgement [4].

4. RESULTS AND DISCUSSION

The proposed study on the Diabetes data collection obtained from different hospitals from 2017 to 2019 has been applied using the programming language of Python. The meaning of the role and correlations gained using the visualisation technique of Heatmap and are described in Figure 1 and Table 3. This feature-relevant heatmap methodology provides a clear picture of data analysis and interpretation of data to enhance the efficiency of the forecast. Larger values were represented by dark colour squares and smaller values by light colour squares. The feature importance determined by the intensity and shades of the colours. Shades of dark blue colour used for strong positive effect, whereas light blue shades are used for average effect and shades of yellow are used for negative effect. The clustered heatmap helpful to identify the various critical levels of DPN patients' conditions. However, the authors sometimes though, with the help of the Clinicians and scientific papers were consulted for the chosen attributes have been medical significance.

ISSN 2515-8260 Volume 07, Issue 10, 2020

After selecting the attributes based on the heatmap visualization technique, applied four ML classifiers namely NB, KNN, C4.5 and GBM. The algorithms are evaluated on the basis of the performance metrics accuracy, sensitivity and specificity and which are presented in Table 3. Table 3 shows that the highest accuracy achieved is 0.82 for NB, 0.72 yields for KNN and almost similar level accuracy achieved for C4.5 and GBM. From this could be concluded, the performance of the KNN, GBM and C4.5 algorithms was almost similar and Naïve Bayes performed better.

The following Table 5 depicts the performance measures accuracy, sensitivity and specificity of DPN prediction using different MLAs.

The DPN prediction performances using various Machine Learning algorithms											
NB			**KNN**			**C4.5**			**GBM**		
Accuracy	**Sensitivity**	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**
0.82	**0.91**	**0.90**	**0.72**	**0.82**	**0.80**	**0.75**	**0.81**	**0.79**	**0.77**	**0.80**	**0.79**

Table 5. The DPN prediction performances using various MLAs

Accuracy, sensitivity, and specificity are used to assess the performance of classifiers which are represented by the following equations (1–3):

$$\text{Accuracy} = \frac{\text{samples correctly classified}}{\text{total samples classified}} = TP + TN / (TP + FP + FN + TN) \quad \text{eqn (1)}$$

,

$$\text{Sensitivity} = \frac{\text{samples correctly classified as positive}}{\text{total positive samples in the dataset}} = TP / (TP + FN) \quad \text{eqn.(2)}$$

,

ISSN 2515-8260 Volume 07, Issue 10, 2020

$$\text{Specificity} = \frac{\text{samples correctly classified as negative}}{\text{total negative samples in the dataset}} = TN / (TN + FP) \quad \text{eqn (3)}$$

,

5. CONCLUSION

A graphic guide to interpreting numeric values is given by heatmaps. They are easy to interpret, so it can be helpful for physicians to reasonably grasp the significance of the features. A better predictive model of health care treatment for people at different risk thresholds, such as low, moderate and high levels of rapid development of DPN, was provided by the clustering heatmap. In a heat map, the outcome of a hierarchical clustering calculation is seen as a dendrogram that can be thought of as a tree. Where they intersect, the lines form nodes that belong to a group. In the given dataset, dendrograms represent the distance or similarity between the parameters [42,43]. Diabetic patients with various classes of DPN severity thresholds and distances between latency and amplitude were seen in this study. In addition, it was possible to create a prediction model using machine learning techniques for the DPN patients. Classifiers accuracy proved the Machine Learning techniques Naïve Bayes, C4.5, KNN and GBM can be preferable to use for this study. The system has achieved an accuracy of 0.82, sensitivity of 0.91 and specificity of 0.90 using NB, an accuracy of 0.72, sensitivity of 0.82 and specificity of 0.80 using KNN, an accuracy of 0.75, sensitivity of 0.81 and specificity of 0.79 using C4.5 and an accuracy of 0.77, sensitivity of 0.80 and specificity of 0.79 using GBM. Other classifiers can also be used to enhance the performance of the proposed framework. In addition, this study encourages clinicians to rapidly identify patients with intensive critical conditions and enables doctors in some cases to reduce the different severity of patient categories.

6. REFERENCES:

[1] Esmaeil S. Nadimi, Tomas Majtner et al, “Facial erythema detects diabetic neuropathy using the fusion of machine learning, random matrix theory and self organized criticality”, Scientific Reports | (2020) | https://doi.org/10.1038/s41598-020-73744-3.

[2] Bharath Sudharsan, Malinda Peeples, Mansur Shomali, “Hypoglycemia Prediction UsingMachine Learning Models for Patients with Type 2 Diabetes”, Journal of Diabetes Science and Technology 2015, Vol. 9(1) 86–90.

[3] Harleen Kaur, Vinita Kumari, “Predictive modelling and analytics for diabetes using a machine learning Approach”, Applied Computing and Informatics, 2018, https://doi.org/10.1016/j.aci.2018.12.004.

[4] Anurag Kumar Verma, Saurabh Pal, “Prediction of Skin Disease with Three Different Feature Selection Techniques Using Stacking Ensemble Method”, Applied Biochemistry and Biotechnology 2019.

[5] K. Nirmaladevi, S. Shanthi, T. Agila, R. T. Dharani, and P. Dhivyapriya, “Analysis and Prediction of Diabetes using Machine Learning,” Test Eng. Manag., vol. 83, no. 9, pp. 14533–14538, 2020.

[6] F. Mao, X. Zhu, S. Liu, X. Qiao, H. Zheng, and B. Lu, "Age as an Independent Risk Factor for Diabetic Peripheral Neuropathy in Chinese Patients with Type 2 Diabetes," vol. 10, no. 3, pp. 592–600, 2019.

[7] Kazemi M, Moghimbeigi A, Kiani J, Mahjub H, Faradmal J, "Diabetic peripheral neuropathy class prediction by multicategory support vector machine model: a cross-sectional study", Epidemiol Health. 2016

[8] Oh, T. J., Kang, S., Lee, J.-E., Moon, J. H., Choi, S. H., Lim, S., & Jang, H. C. , "Association between deterioration in muscle strength and peripheral neuropathy in people with diabetes. Journal of Diabetes and Its Complications" doi:10.1016/j.jdiacomp.2019.04.00

[9] H. Lai, H. Huang, K. Keshavjee, A. Guergachi, and X. Gao, "Predictive models for diabetes mellitus using machine learning techniques," BMC Endocr. Disord., vol. 19, no. 1, pp. 1–9, 2019, doi: 10.1186/s12902-019-0436-6.

[10] Zohaib Iqbal, Shazli Azmi et al, "Diabetic Peripheral Neuropathy: Epidemiology, Diagnosis, and Pharmacotherapy", Clinical Therapeutics/Volume 40, Number 6, 2018

[11] F. Mao, X. Zhu, S. Liu, X. Qiao, H. Zheng, and B. Lu, "Age as an Independent Risk Factor for Diabetic Peripheral Neuropathy in Chinese Patients with Type 2 Diabetes," vol. 10, no. 3, pp. 592–600, 2019.

A. Gordon Smith and J. Robinson Singleton," Obesity and Hyperlipidemia are Risk Factors for Early Diabetic Neuropathy", J Diabetes Complications. 2013 ; 27(5): 436–442. doi:10.1016/j.jdiacomp.2013.04.003

[12] Solomon Tesfaye, Nish Chaturvedi et al, "Vascular Risk Factors and Diabetic Neuropathy", The New England Journal of Medicine,2005.

[13] Jian‑bin Su , Li‑hua Zhao, Xiu‑lin Zhang et al., "HbA1c variability and diabetic peripheral neuropathy in type 2 diabetic patients", Su et al. Cardiovasc Diabetol (2018) 17:47, https://doi.org/10.1186/s12933-018-0693-0

[14] Yun-Ru Lai , Wen-Chan Chiu, Chih-Cheng Huang et al.,"HbA1C Variability Is StronglyvAssociated With the Severity of Peripheral Neuropathy in Patients With Type 2 Diabetes, Frontiers in Neuroscience, 2019.

[15] Fargol Booya Fatemeh Bandarian, et al., "Potential risk factors for diabetic neuropathy: a case control study", BMC Neurology 2005.

[16] Nisar M, Asad A, Waqas A, et al. (August 12, 2015) Association of Diabetic Neuropathy with Duration of Type 2 Diabetes and Glycemic Control. Cureus 7(8): e302. DOI 10.7759/cureus.302.

[17] J. ROBINSON SINGLETON, A. GORDON SMITH, MARK B. BROMBERG "Increased Prevalence of Impaired Glucose Tolerance in Patients With Painful Sensory Neuropathy, DIABETES CARE, VOLUME 24, NUMBER 8, AUGUST 2001.

[18] Natalie C. G. Kwai, William Nigole, Ann M. Poynten et al,"The Relationship between Dyslipidemia and Acute Axonal Function in Type 2 Diabetes Mellitus In Vivo", PLOS

[19] ONE | DOI:10.1371/journal.pone.0153389 April 14, 2016.

[20] Timothy D. Wiggin,1 Kelli A. Sullivan et al, "Elevated Triglycerides Correlate With Progression of Diabetic Neuropathy", DIABETES, VOL. 58, JULY 2009.

[21] B. A. Perkins, A. Dholasania, R. A. Buchanan and V. Bril, "Original Article: Complications Short-term metabolic change is associated with improvement in measures of diabetic neuropathy: a 1-year placebo cohort analysis", Diabetic Medicine 2010

[22] Vincenza Spallone1," Blood Pressure Variability and Autonomic Dysfunction", MICROVASCULAR COMPLICATIONS—NEUROPATHY (R POP-BUSUI, SECTION EDITOR), Current Diabetes Reports, (2018) 18:137.

[23] Joanna Wierzbowska, Robert Wierzbowski, Andrzej Stankiewicz, Brent Siesky, Alon Harris,"Cardiac autonomic dysfunction in patients with normal tension glaucoma: 24-h heart rate and blood pressure variability analysis", Br J Ophthalmol 2012;96:624e628. doi:10.1136/bjophthalmol-2011-300945.

[24] Federica Di Gennaro, Cinzia D'Amato, Roberto Morganti, Carla Greco, Susanna Longo, Diana Corradini, Davide Lauro, Vincenza Spallone, " Morning blood pressure surge is associated with autonomic neuropathy and peripheral vascular disease in patients with diabetes", Journal of Human Hypertension, 2019.

[25] Patrícia Carvalho Machado Aguiar, Marcus Vinícius Della Coletta, Jean Jorge Silva de Souza, "The Association of Dyslipidemia and Peripheral Diabetic Neuropathy: The Influence of Urea", Aguiar et al., Diabetes Case Rep 2016,1:2

[26] Debahuti Mohapatra, Komaladevi Sampath Damodar, " Glycaemia Status, Lipid Profile and Renal Parameters in Progressive Diabetic Neuropathy", Journal of Clinical and Diagnostic Research. 2016 Sep, Vol-10(9): CC14-CC17

[27] Prakash Babaliche, Rishit A. Nadpara, Arif Maldar," Association Between Estimated Glomerular Filtration Rate and Microvascular Complications in Type II Diabetes Mellitus Patients: A 1-Year Cross-Sectional Study, JOURNAL OF THE NATIONAL MEDICAL ASSOCIATION 2018.

[28] Laura Mayeda, Ronit Katz et al, "Glucose time in range and peripheral neuropathy in type 2 diabetes mellitus and chronic kidney disease, BMJ Open Diab Res Care 2020;8:e000991. doi:10.1136/bmjdrc-2019-000991.

[29] John I. Malone ,"Diabetic Central Neuropathy: CNS Damage Related to Hyperglycemia", Diabetes 2016;65:355–357 | DOI: 10.2337/dbi15-0034

[30] Maxim Dobretsov, Dmitry Romanovsky, Joseph R Stimers, "Early diabetic neuropathy: Triggers and mechanisms", World J Gastroenterol 2007 ; 13(2):175-191

[31] BRENDA W.C. BONGAERTS, WOLFGANG RATHMANN et al, "Postchallenge Hyperglycemia Is Positively Associated With Diabetic Polyneuropathy", Diabetes Care 35:1891–1893, 2012

[32] Reddemma Sandireddy, Veera Ganesh Yerra et al, " Neuroinflammation and Oxidative Stress in Diabetic Neuropathy: Futuristic Strategies Based on These Targets", International Journal of Endocrinology Volume 2014

[33] P. K. Bariha, K. M. Tudu, Shiny Thomas Kujur," Correlation of microalbuminuria with neuropathy in type-II diabetes mellitus patients", International Journal of Advances in Medicine (2018)

[34] Ahmad T, Ulhaq I, Mawani M, Islam N. Microalbuminuria in Type-2 Diabetes Mellitus; the tip of iceberg of diabetic complications. Pak J Med Sci. 2017;33(3):519-523. doi: https://doi.org/10.12669/pjms.333.12537

[35] Gunjan Kumar Mandal, Pankaj Kumar Jha, "Study of Microalbuminuria Level in Chronic Complications of type 2 Diabetes",International Journal of Contemporary Medical Research",Volume 4 | Issue 4 | April 2017

[36] ANDREW MORAN, WALTER PALMAS, "Cardiovascular Autonomic Neuropathy Is Associated With Microalbuminuria in Older Patients With Type 2 Diabetes", Diabetes Care 27:972–977, 2004

[37] DAVID S.H. BELL,, CATHERINE H. KETCHUM et al, "Microalbuminuria Associated With Diabetic Neuropathy", DIABETES CARE, VOLUME 15 , APRIL 1992

[38] Prakash Babaliche, Rishit A. Nadpara et al, "Association Between Estimated Glomerular Filtration Rate and Microvascular Complications in Type II Diabetes Mellitus Patients: A 1-Year Cross-Sectional Study", JOURNAL OF THE NATIONAL MEDICAL ASSOCIATION, 2018.

[39] Zohaib Iqbal, MRCP, Shazli Azmi et al, "Diabetic Peripheral Neuropathy: Epidemiology, Diagnosis, and Pharmacotherapy", Clinical Therapeutics/Volume 40, Number 6, 2018

[40] Nisar M, Asad A, Waqas A, et al. (August 12, 2015) Association of Diabetic Neuropathy with Duration of Type 2 Diabetes and Glycemic Control. Cureus 7(8): e302. DOI 10.7759/cureus.302

[41] Carlos Huertas, Reyes Juarez-Ramirez, "Automatic Threshold Search for Heat Map Based Feature Selection: A Cancer Dataset Analysis", International Scholarly and Scientific Research & Innovation 10(7) 2016

[42] Zuguang Gu, Roland Eils and Matthias Schlesner, "Complex heatmaps reveal patterns and correlations in multidimensional genomic data", Bioinformatics, Oxford University Press, 32(18), 2016, 2847–2849

[43] Wenjie You, Zijiang Yang, Mingshun Yuan, and Guoli Ji, "TotalPLS: Local Dimension Reduction for Multicategory Microarray Data", IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS, VOL. 44, NO. 1, FEBRUARY 2014.

[44] Shruti Jain and Ayodeji Olalekan Salau, "An image feature selection approach for

[45] dimensionality reduction based on kNN and SVM for AkT proteins", Cogent Engineering (2019), 6: 1599537.

[46] Minyechil Alehegn, Rahul Raghvendra Joshi, Preeti Mulay, "Diabetes Analysis And Prediction Using Random Forest, KNN, Naïve Bayes, And J48: An Ensemble Approach", INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 09, SEPTEMBER 2019

[47] Ross Quinlan (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA.

[48] Zhongheng Zhang, Yiming Zhao, Aran Canes, Dan Steinberg, Olga Lyashevska, "Predictive analytics with gradient boosting in clinical medicine", Annals of Translational Medicine., Ann Transl Med 2019;7(7):152