

# **DIABETES MELLITUS DIAGNOSTIC EXPERT SYSTEM**

A Dissertation submitted in fulfillment of the requirements for the Degree  
of

## **MASTER OF ENGINEERING** *in* **Electronic Instrumentation & Control Engineering**

*Submitted by*

Jaskaran Bir Singh  
801351005

*Under the Guidance of*  
Dr. Sunil Kumar Singla  
Assistant Professor, EIED



**2015**

**Electrical and Instrumentation Engineering Department**

**Thapar University, Patiala**

*(Declared as Deemed-to-be-University u/s 3 of the UGC Act., 1956)*

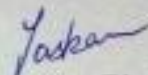
**Post Bag No. 32, Patiala – 147004**

**Punjab (India)**

## DECLARATION


I hereby certify that the work which is presented in dissertation entitled, "**DIABETES MELLITUS DIAGNOSTIC EXPERT SYSTEM**", in partial fulfillment of the requirements for the award of the degree of Master of Engineering in Electronic Instrumentation and control, submitted to Electrical & Instrumentation Engineering Department of Thapar University, Patiala is as authentic record of my own work carried under the supervision of **Dr. Sunil Kumar Singla**. It refers others researcher's work which are duly listed in the reference section. The matter contained in this dissertation has not been submitted, neither in part nor in full to any other degree to any other university or institute except as reported in text and references.

Place: Patiala  
Date: 14/07/2015

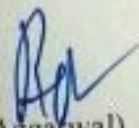
  
(Jaskaran Bir Singh)  
Roll No.: 801351005

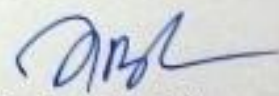
It is certified that the above statement made by the student is correct to the best of my knowledge and belief.

Date: 14/07/2015

  
(Dr. Sunil Kumar Singla)  
Assistant Professor  
Electrical & Instrumentation Engineering Department  
Thapar University, Patiala

*Countersigned by:*

  
(Dr. Ravinder Aggarwal)  
Head  
Electrical & Instrumentation Engineering Department  
Thapar University, Patiala

  
(Dr. S. S Bhatia)  
Dean (Academic Affairs)  
Thapar University, Patiala

## CLAIM OF AUTHENTICITY

This is to certify that the work presented in this dissertation entitled "DIABETES MELLITUS DIAGNOSTIC EXPERT SYSTEM" submitted by Mr. Jaskaran Bir Singh in the partial fulfillment of the requirement for the award of degree of Master of Engineering in electronics instrumentation and control engineering at Thapar university, Patiala, is an authentic record of the candidate's own work carried by him under the supervision and guidance of Dr. Sunil Kumar Singla .

The matter embodied in this report has practical significance and can be implemented by doctors to diagnose Diabetes Mellitus and its types.

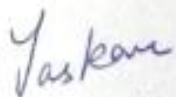
  
D.M. (Dr. Rohit Aggarwal)  
(Endocrinology)  
Aggarwal Health Care Centre D. M.  
(Endocrinology and Metabolism)

## ACKNOWLEDGEMENT

First of all, I would like to express my Gratitude to Dr. Sunil Kumar Singla Assistant Professor, Electrical and Instrumentation Engineering Department (EIED), Thapar University Patiala for his patient guidance and support throughout this dissertation. I am truly very fortunate to have the opportunity to work with him. I found this guidance to be extremely valuable.

I am also thankful to our Head of Department, Dr. Ravinder Aggarwal as well as PG Coordinator, Nirbhow Jap Singh, Assistant Professor, Electrical and Instrumentation Engineering Department (EIED) Thapar University, Patiala. I would also like to thank Dr. Rohit Aggarwal, Physician (Diabetes Specialist) and my friends who devoted their valuable time and help me in all possible ways towards successful completion of this work. I thank all those who have contributed directly or indirectly to this work.

Lastly, I would like to thank my parents for their years of unyielding love and encourage they have always wanted the best for me and I admire their determination and sacrifice.

  
**Jaskaran Bir Singh**  
801351005

# TABLE OF CONTENTS

	<b>Page</b>
<b>CERTIFICATE</b>	i
<b>CLAIM OF AUTHENTICITY</b>	ii
<b>ACKNOWLEDGEMENT</b>	iii
<b>LIST OF TABLES</b>	vi
<b>LIST OF FIGURES</b>	vii
<b>NOMENCLATURE</b>	viii
<b>ABSTRACT</b>	ix
<b>CHAPTER 1 INTRODUCTION</b>	<b>1-8</b>
1.1 Diabetes	2
1.2 Types of Diabetes	3
1.3 Major Symptoms used for Diagnosis of Diabetes	3
1.4 Causes of Diabetes	4
1.4.1 Causes of Type 1 Diabetes	4
1.4.2 Causes of Type 2 Diabetes	5
1.4.3 Causes of Gestational Diabetes	5
1.5 Blood Glucose Regulation Hormones	5
1.5.1 Role of Insulin	5
1.5.2 Role of Glucagon	6
1.6 Expert System	6
1.7 Diabetes Mellitus Diagnosis Expert System	7
1.8 Objectives of the dissertation	8
1.9 Organization of dissertation	8
<b>CHAPTER 2 LITERATURE SURVEY</b>	<b>9-15</b>
2.1 Artificial Neural Network and k- Nearest Neighbor	11
2.1.1 Artificial Neural Network (ANN)	11
2.1.2 Human Biological Neuron and Artificial Neuron	12
2.1.3 Learning in ANN	13
2.1.4 <i>k</i> - Nearest Neighbor ( <i>k</i> -NN)	14
2.1.5 <i>k</i> -NN Classification	14
<b>CHAPTER 3 METHODOLOGY USED</b>	<b>16-32</b>
3.1 Overview of Methodology	16
3.1.1 Flow Chart	17
3.2 Database used	18
3.2.1 Normalization of data	19
3.3 Symptoms Based Diagnosis Algorithm	20
3.4 ANN and <i>k</i> -NN Based Diagnosis	22

3.5	Artificial Neural Network (ANN) Training Technique	23
3.5.1	Method used for Training	24
3.5.2	Weight Updation Delta Rule	25
3.6	Training of ANN Network Using NPR Tool of MATLAB R2009a	26
3.7	$k$ -NN ( $k$ -Nearest Neighbor)	27
3.7.1	Method used	27
3.7.2	Selection of Value of $k$	28
3.8	Type of Diabetes Diagnosis	29
3.8.1	For Type 1 Diabetes Rule used	29
3.8.2	For Type 2 Diabetes Rule used	29
3.9	Diet Recommendation	30
<b>CHAPTER 4</b>	<b>RESULTS</b>	<b>33-41</b>
4.1	Test Confusion Matrix with 65% Training data using ANN	33
4.2	Gradient and Validation Curve	33
4.3	Receiver Operating Characteristics Plot	34
4.4	Mean Square Error	34
4.5	Test Confusion Matrix with 85 % Training data using ANN	35
4.6	ROC Plot	35
4.7	Gradient and Validation Curve	36
4.8	Mean Square Error	36
4.9	ANN Forecast and Target Data for first 20 Samples	37
4.10	$k$ -NN Test Accuracy Results	38
4.11	Overall Comparison Table	40
4.12	User Interface Standalone Application .exe	41
<b>CHAPTER 5</b>	<b>CONCLUSIONS AND FUTURE SCOPE</b>	<b>42-43</b>
5.1	Conclusion	42
5.2	Future Work	43
<b>REFERENCES</b>		<b>44-46</b>
<b>APPENDIX A</b>		<b>48</b>
<b>APPENDIX B</b>		<b>51</b>

# LIST OF TABLES

Table No.	Caption	Page
3.1	Class Repetition Table	18
3.2	Statistics Of Pima Indians Diabetes Database	19
3.3	NPR Tool Training Properties Used	27
3.4	$k$ -NN Properties Table	28
4.1	Test Confusion Matrix	33
4.2	Means Square Error (MSE) Table	34
4.3	Test Confusion Matrix	35
4.4	Mean Square Error	36
4.5	ANN Forecast Table	37
4.6	$k$ -NN Test Accuracy Table	38
4.7	Overall Accuracy Comparison	40

# LIST OF FIGURES

Figure No.	Caption	Page
1.1	Worldwide Diabetics Distribution	1
1.2	Beta Cells	2
1.3	Worldwide Symbol for Diabetes	2
1.4	Insulin and Glucagon Released by Pancreas	6
1.5	Expert system Block Diagram	7
2.1	Neuron Cell in Human Brain	12
2.2	ANN Model of Biological Neuron	12
2.3	Nearest Neighbor based on Euclidean Distance	14
3.1	Four Layer Model of Diabetes Diagnosis Expert System	17
3.2	Process of Normalization	19
3.3	exe Input Symptoms Questions	21
3.4	Input Parameters, ANN and $k$ -NN based Diagnosis of Diabetes	23
3.5	Back propagation of Errors in Back Propagation Network	24
3.6	NN train Tool in MATLAB R2009a	26
3.7	Type Diagnosis and Diet Chart Recommendation on UI Window	31
3.8	Files generated from MATLAB Runtime Compiler output.	32
4.1	Gradient and Validation Plot	33
4.2	ROC Test Curve	34
4.3	ROC Test Curve	35
4.4	Gradient and Validation Plot	36
4.5	.exe of User Interface	41

# NOMENCLATURE

<b>ANN</b>	Artificial Neural Network
<b>BMI</b>	Body Mass Index
<b>DI</b>	Diabetes Insipidus
<b>DM</b>	Diabetes Mellitus
<b>ESTA</b>	Expert System Shell for Text Animation
<b>GDA</b>	Generalized Discriminant Analysis
<b>GDM</b>	Gestational Diabetes
<b>GSVD</b>	Generalized Singular Value Decomposition
<b>IRDA</b>	Intelligent Diet-Recommendation Agent
<i>k</i> -NN	<i>k</i> Nearest Neighbor
<b>LS SVM</b>	Least Square Support Vector Machine
<b>MCR</b>	MATLAB Compiler Runtime
<b>MSE</b>	Mean Square Error
<b>NN</b>	Neural Network Tool
<b>NPR</b>	Neural Pattern Recognition Tool
<b>ROC</b>	Receiver Operating Characteristics
<b>SVM</b>	Support Vector Machine
<b>SVR</b>	Support Vector Regression

# ABSTRACT

The most challenging problem troubling the world today is spread of Diseases and the consequent sufferings of the innocent inhabitants of this planet. The doctors and researchers all over the world work, jointly and separately, day and night to combat this grave problem to the maximum extent they can. Inspire of this conscious and collective effort mistakes are due to wrong diagnostics, the starting step toward implementing a cure.

Engineer can contribute in this filed by developing a software or computer program, called Expert System, which can Aid the Doctor in diagnosing a Disease. This expert system will surely bring down the mistakes which in turn will be boon for humanity. One such system to diagnose various kinds of diabetes, using the concept of symptoms, ANN and  $k$ -NN on MATLAB R2009a platform, has been developed in this work.

In the present work a prototype expert system has been developed for clinical diagnostics by training the Artificial Neural Network and  $k$  Nearest Neighbor for Diabetes Mellitus diagnosis. The design aligns the structure and work of diagnostics mechanism similarly to the decision making process of a physician by dividing the diagnostics in three parts. The first part of the diagnostics is to find the risk of patient suffering from diabetes using weighted symptoms of the patient. Second part is to find output class value of diabetes of a patient using ANN (Artificial Neural Network) and  $k$ -NN (closer the output value to 1 more the patient is diabetic sufferer). Third part of our Project classifies whether the patient is Type 1 or Type 2 Diabetic along with respective Diet charts and preventive prescriptions. Standalone .exe application of the above project has been made to serve as potable project that can run on any Windows 95 or later platform without requirement of MATLAB R2009a software installed on that PC.

Diabetes disease has affected almost all countries in the world. This Disease causes 2.7% of the total casualties in the world each year [1]. Today, according to report of International Diabetes Federation (IDF), there are nearly 230 million patients suffering from Diabetes worldwide, 80% of the total patients are from America developing countries [1, 2]. Estimated Diabetic sufferers by 2025 are 380 million across the world. In 2010 Diabetes was identified as seventh most leading cause of death based on 69,701 death certificates in U.S [2]. The worldwide Diabetic distribution is shown in Fig 1.1.

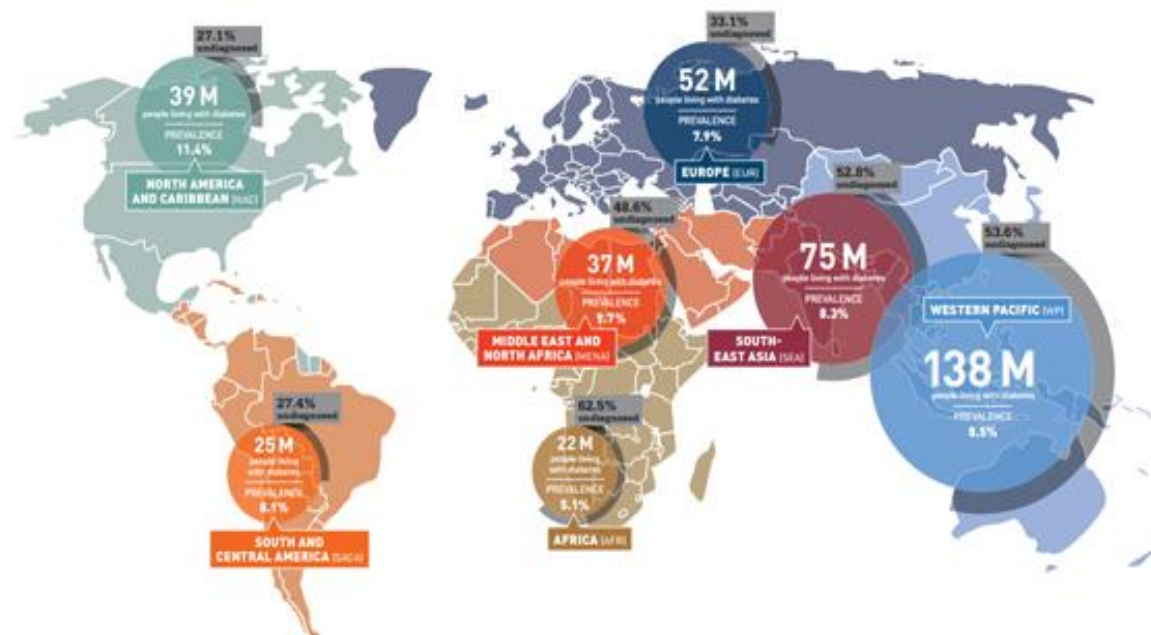


Fig 1.1 Worldwide Diabetic's Distribution [1]

Hence Diagnosis of Diabetes is very important to save millions of sufferers worldwide. Use of Expert system has been greatly done in medical diseases diagnosis in recent decade in developed countries. Even by now Expert Systems aren't used for all medical disease diagnosis purposes which have caused wastage of human resource, money and causing suffering with many humans lives.

**1.1 Diabetes** It is a disease related to metabolism of the body, the way we use digested food for energy. The digestive system turns the foods in the form of glucose, a form of sugar that enters the blood. It's the responsibility of Insulin that helps the absorption of glucose in the body and releases energy for body. This disease occurs when insulin is either not produced in the body or body does not use made insulin properly [2].

Pancreas is responsible for production of insulin. Pancreas contains cluster of cells called islets, insulin is made in beta cells of islets as shown in Fig 1.2.

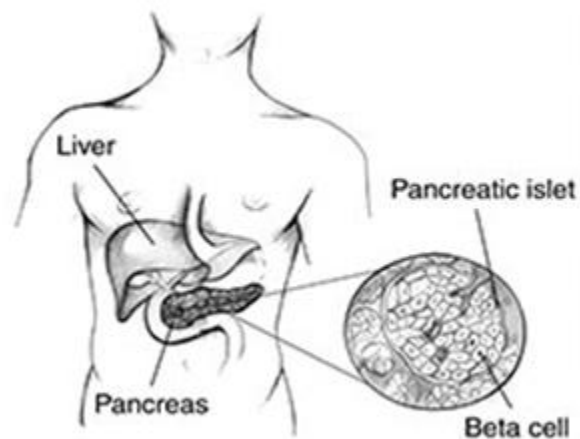


Fig 1.2 Beta Cells [4]

If insulin is not made in beta cells or insulin doesn't initiate proper action then level of glucose rises in the body that causes pre Diabetes or Diabetes. Thus, during Diabetes body cells are deprived of energy in spite of having high blood glucose concentration in the body. By this moment increased blood glucose level proves to be harmful, leading to complications for instances hear problem, stroke, kidney problem, blindness, dental disease, and amputations. This disease may lead to other problems at later stage of life.

Blue circle shown in Fig 1.3 is worldwide symbol used for Diabetes.

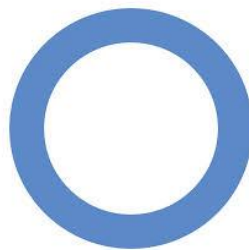


Fig 1.3 Worldwide Symbol for Diabetes [2]

## 1.2 Types of Diabetes

Diabetes is mainly classified into following Types

- a) **Type 1 Diabetes:** Also known as Juvenile-Onset Diabetes. World's 15% of total Diabetic population is Type 1 Diabetic. In this body's defense mechanism attacks cells which produce insulin. This category of Diabetes usually found in children's and adults but can affect people of any age. This type of Diabetics need regular dosage of injections of insulin to counter elevated levels of glucose in blood, failing so they may die [4, 5].
- b) **Type 2 Diabetes:** This type of Diabetes is called noninsulin dependent Diabetes or adult onset Diabetes. World's 85% of total Diabetic population is Type 2 Diabetic This can occur to any age group people. 90% of the world's population has Type 2 Diabetes cases. Overweight is not always the cause of Type 2. During initial stage people with this Type of Diabetes can manage through routine exercise but with time they have to turn on to oral drugs [4, 5].
- c) **Gestational Diabetes (GDM):** This category of Diabetes is associated with elevated blood glucose level for the period of pregnancy. Known for this category of Diabetes that can occur to both mother and baby and develops 1 in 25 pregnancies. This type of Diabetes disappears after pregnancy but 50% women with GDM turns into Type 2 Diabetes patients mostly after decade period [4, 5].

**1.3 Major Symptoms used for Diagnosis of Diabetes** There are many Symptoms of Diabetes but most common signs and symptoms of Diabetes are [7]

- a) Frequent urination
- b) Excessive thirst
- c) Increased hunger
- d) Weight loss
- e) Tiredness
- f) Lack of interest and concentration
- g) Numbness of hands or feet
- h) Blurred vision
- i) Frequent infections
- j) Slow-healing wounds

k) Vomiting and stomach ache

Type 2 Diabetes symptoms can be mild or slow while in Type 1 symptoms occurs suddenly making Type 2 Diabetes hard to detect [5].

## **1.4 Causes of Diabetes**

Diabetes disease may be result of many causes that can directly or indirectly prove to be reason behind occurrence of this disease. Some of the most important reasons that can trigger this disease are explained below

**1.4.1 Causes of Type 1 Diabetes** Type 1 Diabetes occurs when body does not have adequate amount of insulin or either beta cells that produce insulin are died off due to autoimmune action of defense system of the body where body's defense systems destroys beta cells in the body that leads to this type of Diabetes [6,7]. The causes of Type 1 Diabetes are

- (i) **Genetic Susceptibility** Genes are carried from parents to children, specific type of genes has been considered to be one of prominent cause for Type 1 Diabetes. Certain genes Human Leukocytes Antigen (HLA) have risk of developing Type 1 Diabetes [5].
- (ii) **Autoimmune Destruction** of Beta Cells, T cells damage beta cells that produce insulin. T cells are a type of white cells. Insulin itself in some cases is the key that causes destruction of beta cells; this type of response of body defense system that destroys its beta is called autoimmune action. This process leads to Type 1 Diabetes [5].
- (iii) **Environmental Factors** Certain factors like junk food, adulterated food, toxins, viruses and infections may indirectly trigger Type 1 Diabetes or may become cause behind autoimmune action [1, 5].
- (iv) **Viruses and Infections** Virus is not direct cause behind Diabetes but researcher have shown a link between them , people with viral infections are diagnosed with Type 1 after later stage, that occurs mostly in winters when viral infection is at peak [5].
- (v) **Infant Feeding practices** Children's who get breastfed milk have reduced risk of budding Type 1 in comparison to who get cow's milk in small age. As breastfed milk has more vitamin D in comparison to cow's milk [5,7].

### 1.4.2 Causes of Type 2 Diabetes

This category of Diabetes is caused by many factors including insulin resistance, a condition in which the body's muscle, fat, and liver cells do not use insulin effectively. This may cause Type 2 Diabetes [7]. Some of important causes of Type 2 Diabetes are

- (i) **Genetic Susceptibility** People carrying TCF7L2 gene have increased risk of developing Type 2 Diabetes and with copies of same gene have more risk of developing the same disease [4, 5].
- (ii) **Obesity and Physical Inactivity** High BMI index can lead to obesity; obsessive people with low or no physical activity have more chances of developing Type 2 Diabetes than non obsessive people [4, 5].
- (iii) **Insulin Resistance** Obsessive people have common problem of insulin resistance, muscle fat, liver cells do not respond to insulin, for that pancreas has to produce more insulin in order for compensation [4,8].

### 1.4.3 Causes of Gestational Diabetes:

Major causes of GSD are hormonal change and metabolic burden during pregnancy.

- (i) **Insulin Resistance** Effect of insulin resistance increases in women due to pregnancy and other factors that occur in women's life. If the pancreas can't create sufficient insulin payable to beta cell dysfunction, gestational Diabetes occurs. Increased weight gain during pregnancy further boosts risk [2, 5].
- (ii) **Family History** Genes play significant role in development of risk of this category of Diabetes genetics is the reason why this type occurs more frequently in African Americans, American Indians [2, 7].

## 1.5 Blood Glucose Regulation Hormones

Insulin and Glucagon control glucose regulation level in the body. These are very important hormones related to metabolism of the body.

### 1.5.1 Role of Insulin

The function of the Insulin is to lower the blood glucose level in the body. Insulin is made in beta cells, therefore, when the glucose level in the blood raises pancreas secrete hormone called Insulin to lower the glucose level in the body. This hormone works in opposite manner to Glucagon hormone [1].

### 1.5.2 Role of Glucagon

The function of the glucagon is to raise the low blood glucose level in the body. Glucagon is produced in alpha cells, thus function of Glucagon is opposite to of insulin. Both these hormones are produced in pancreas organ of the body [1].

Thus, Diabetes Mellitus disease is caused to due to lack of insulin producing beta cells or their inability to produce required Insulin for proper metabolism of body. Therefore both these hormones play a vital role in proper metabolic functions of the body [1, 2]. Fig 1.4 shows the release of Insulin and Glucagon by pancreas.

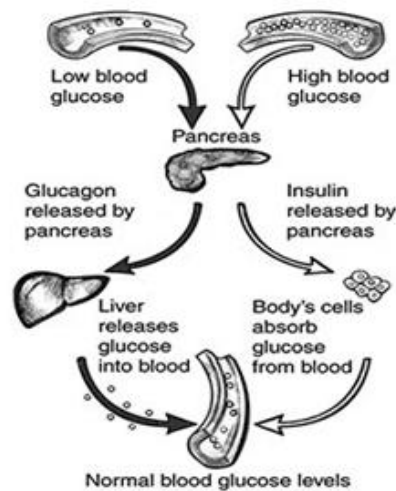


Fig 1.4 Insulin and Glucagon released by Pancreas [2]

### 1.6 Expert System

Expert System is a software code that imitate or tries to imitate human expert brain. Its function is to provide expert advice (decision, recommendations) like a domain expert is consulted. Expert System gathers all the skills of domains experts or from other knowledge sources by training the network and use this gained expertise to provide recommendation or decision support to persons with lesser skills [6]. It consists of two subsystems i.e. the inference engine and the knowledge base. Expert system solves complicated problems by using most developed artificial intelligence models like Fuzzy Logic, Neural Network, Decision Trees etc.[5, 9]. Practically, medical field was the first field that made use of Expert System (ES) technology. Few initial most successful expert systems were Mycin, Nursexpert, Centaur, Diagnoser and Medics [10]. It is a software code with the aim to imitate the way of thinking process of experts in dealing with real world situations that don't have exact solutions and

gathers knowledge and expertise from experienced domain experts and from other knowledge sources and then helps in providing decision support to the patients with lesser domain skills or expertise of diseases diagnosis [6].

Since the existence of artificial intelligence many Experts Systems have been developed to aid doctors and patients in diseases diagnosis. Few of the Expert Systems are listed below:

- a) PUFF: Pulmonary disease diagnosis
- b) AI/COAG: Blood disease diagnosis
- c) AI/RHEUM: Rheumatic disease diagnosis
- d) CADUCEUS: Internal medicine disease diagnosis
- e) ANNA: Monitoring and treatment analysis
- f) BLUEBOX: Depression diagnosis and treatment
- g) MYCIN: Microbial disease diagnosis and treatment
- h) ONCOCIN: Treatment and management of patient's chemotherapy
- i) ATTENDING: Anesthesia management education
- j) GUIDON: Microbial disease education [10].

### 1.7 Diabetes Mellitus Diagnosis Expert System

A computer program that can diagnose various types of Diabetes is known as Diabetes Expert system. Input to this type of system can be various patient symptoms like age, tests reports and other symptoms. Output of the system is particular type of Diabetes detected and a preventive measure from that type of Diabetes. It is made to assist the human or physician to take their decision unless high emergency is not there.

The block diagram of Diabetic Expert System is shown in Fig. 1.5.

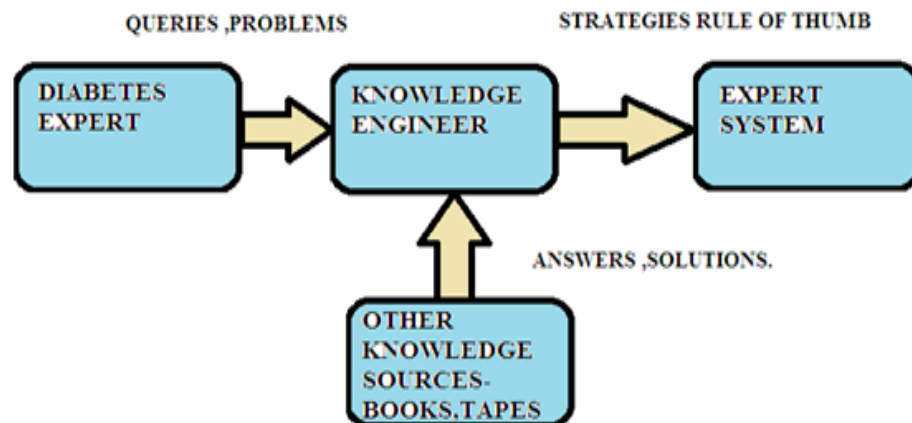


Fig 1.5 Expert System Block Diagram [6]

This Expert System can be used to aid human expert or patients to diagnose Diabetes and provides necessary prescription to patients with Diabetes. Knowledge source can be human expert (physician), books, research papers or internet. The role of engineer is to make software that can behave like a human Diabetes Expert with the combined knowledge from physician, books, internet or other knowledge sources [6].

Diabetes expert system makes use of artificial intelligence for coming to conclusions just like human expert do. Various techniques which can be used are

- (i) Fuzzy Logic
- (ii) Artificial Neural Network
- (iii)  $k$ - Nearest Neighbor
- (iv)  $k$  means
- (v) Support Vector Machine.

## **1.8 Objectives of the Dissertation**

The various objectives of this work are given below

- (i) Study of various existing Expert Systems used for Diabetic Diagnosis.
- (ii) Collection of database or selection of existing database for Diabetes Mellitus.
- (iii) Development of an Expert System for the Diagnosis of Diabetes and its analysis.

## **1.9 Organization of dissertation**

Dissertation structure is as below

**Chapter 2** Includes literature survey about the algorithms used in past, ANN and  $k$ -NN for Diabetes Diagnosis Expert System.

**Chapter 3** Consists of Methodology used for Diabetes Diagnosis Expert System.

**Chapter 4** Includes the Results, confusion matrix as a Figure of Merit for showing the effectiveness of this system.

**Chapter 5** is of Conclusion and Future Recommendations.

A lot of work has been done in the past in the field of Diabetes Diagnosis using Expert Systems and various improvements have been done from past continuously to improve diagnostic results review of few such relevant literature is given below.

**Duarte Spring [11]** implemented an Expert system in sun's java language made in java shell, these shells are specifically designed for programming Expert Systems and one such shell used here is jess which is quite user friendly and advanced version of CLIPS. The input to JESS is information of patient necessary to make diagnostics e.g. age, symptoms, choice of tests etc. The output from this system is diagnostic of disease of particular category of Diabetes. This approach uses java developer kit (variously known as Jdk and Sdk version 1.2).

**Garcia et al. [12]** implemented an Expert System named Esdiabetes that is made to diagnose and control blood glucose level. System is represented in WxCLIPS than CLIPS 6.0 as this version is comparatively more user friendly and easy to use than earlier used version. Users are asked questions until enough information is there to produce decision.

**Bhimrao et al. [13]** used natural therapy and integrated almost every natural therapy for Diabetes using Expert System Shell for Text Animation (ESTA). ESTA can write rules and has an inference engine. This system asks user questions i.e. select the disease and their symptoms for which natural treatment is required.

**Wiley et al. [14]** describes use of Support Vector Regression (SVR) and 3 upcoming application of AI are made available automatic problems detection in BG control (b) offering solutions to the detected problems (c) remembering the effective and/or ineffective solutions for individual Patients type1 diabetes (T1D).

**Beulah et al. [15]** have developed a system that can be accessed from any part of the world .This was designed in dot net framework instead of java. This system uses cloud storage i.e. software is available online, so it can be accessed from any part of the world. It also provides early diagnosis of all three types of Diabetes that is Type1, Type 2 and GSD, by noticing symptoms of the patient.

**Akter et al. [16]** provides Preventive measures and routine health checks for diagnosis and management of Diabetes Mellitus. Authors have developed an automated system with good interface. This performs diagnosis by rules given by medical doctors and experts.

**Kalpana et al. [17]** have presented the Fuzzy logic approach on 5 parameters of Pima Indians Diabetes database. They have used triangular membership function in MATLAB and DM parameter as 1 for positive result and 0 for negative result. Five fuzzy rules are used for fuzzification and centroid method of defuzzification is used. OR (maximum) method of connection is used for interconnection between rules. This paper focuses Study mainly on young aged between 26-30 years.

**Jaafar et al. [18]** have done diabetes forecasting using artificial neural network and Back propagation algorithms for which learning of weights are used. Test is done on 768 data, 268 were found Diabetic out of them. Neural network tool of MATLAB is used. Author used 790 epochs till the gradient is reached with appropriate accuracy.

**Lee et al. [19]** have implanted fuzzy based 3 ontology diet recommendation model for type 2 Diabetes and also intelligent diet-recommendation agent (IRDA) is made for recommendation on diet to patient with various Type 2 Diabetes. Type 2 fuzzy ontology model is developed using Type 2 fuzzy sets. Diet recommendation charts are developed especially in consideration of Taiwanese diet.

**Chang lee et al. [20]** have produced five layer model that include a fuzzy knowledge layer, fuzzy group relation layer, fuzzy group domain layer, fuzzy personal relation layer and fuzzy personal domain layer, this model is made to diagnose the patients with Diabetes disease. They implemented the code in C++, schematic decision support agent (SDSA) is made to assist the decision making. This model is based on fuzzy ontology.

**Rahman et al. [21]** have studied Relative discussion of classification techniques and their comparison using 3 mining tools: WEKA, TANGARA and MATLAB. Pima Indians Database is used with 768 in all 3 tool, error and accuracy are figure of merit used for comparison. Highest accuracy of 81.33% has been reported of J48graft in Weka.

**Park et al. [22]** have shown that the limitations of LDA are removed by generalized singular value decomposition (GSVD). They have removed problem of nonsingular scatter matrix. GSVD is made to solve the generalized Eigen value problems by kernel function. This GSVD is then

used for classification of Diabetes. Authors have shown advantages of using GSVD over LDA for Diabetes database classification.

**Baudat et al. [23]** have shown a new technique called Generalized Discriminant analysis (GDA) to work with nonlinear discriminate analysis. GDA map's the input to high dimensional feature space, this makes classification easy. Different kernel function operator has been used to accomplish wide range of nonlinearities here. Data classification is done for proving validation of results.

**Ahmed et al. [24]** have studied current state of research in the area of Diabetes diagnostics Expert System and in this also key features and problems of existing systems are highlighted, only analysis of past systems is done and indicated that with Pace diseases diagnosis technology is improves.

**Polat et al. [25]** have implemented Generalized Discriminant Analysis (GDA) and LS SVM (least square support vector machine based Expert System, in the proposed system both techniques are used in cascaded form i.e. both are used one after other, first stage used GDA and second stage uses LS SVM, both methods are used with 10 cross validation.

**Chikh et al. [26]** have shown model that uses Artificial Immune Reorganization System (AIRS2) and Fuzzy  $k$ -NN. The authors proposed a cascade system that uses both algorithms. Pima Indians Diabetes Database is used and 10 fold cross validation is used and highest accuracy of 82.69% is achieved. Authors have shown the advantage of using two cascade systems to improve misclassification error over single layer system.

## **2.1 Artificial Neural Network and $k$ - Nearest Neighbor**

ANN and  $k$ -NN both are machine learning algorithm, both are part of soft computing techniques used to create artificial intelligence in machines to perform logical decision ability like humans have naturally.

### **2.1.1 Artificial Neural Network (ANN)**

ANN is artificial machine model that exists in soft computing that tries to copy the human brain and work similar as human biological neuron does. Human brains works entirely differently from how artificial neuron used in computers work. In ANN training or learning of particular task is achieved through updation of weights unlike biological neuron, the way the updation occurs depends upon the learning algorithm used.

Specific properties of Neural Networks are

- (i) Adaptive weights that gets updated as per learning occurs
- (ii) It can approximate nonlinear functions [9]

### 2.1.2 Human Biological Neuron and Artificial Neuron

Human brain has around  $10^{11}$  neurons. Artificial neural is made in similar way as human neuron functions, thus artificial neuron is made by looking at human brain neuron. Human brain neuron consists of cell body, dendrites and axon. The role of the dendrites is to input electrochemical signal into the neuron and whereas axon brings the signal out from neuron to outside the cell to communicate with other neurons [9]. The human neuron and artificial neuron is shown in Fig 2.1 and 2.2 below

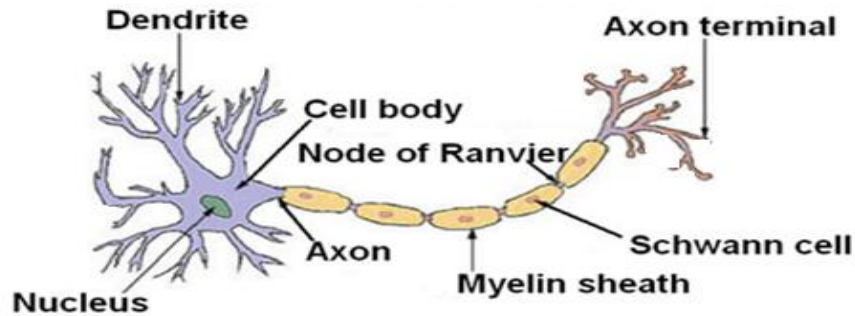


Fig 2.1 Neuron Cell in Human Brain [9]

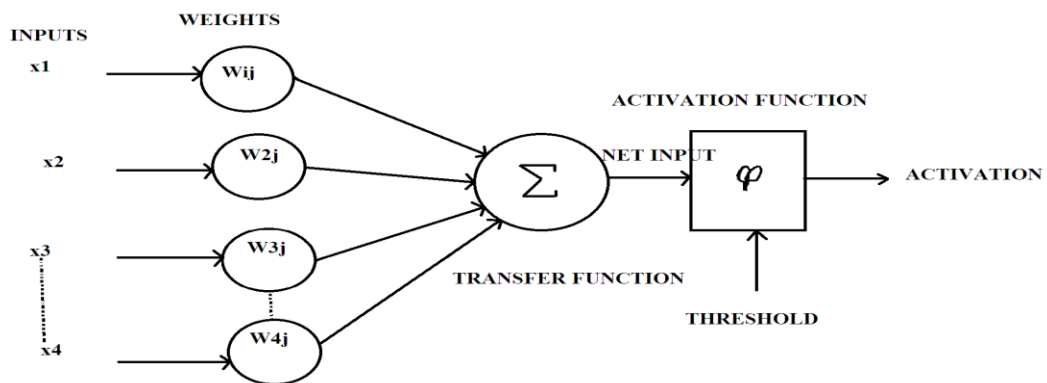


Fig 2.2 ANN Model of Biological Neuron [9]

Neuron from its axon produces output, but output from axon occurs only when collective sum of inputs from dendrites is above a certain threshold level, output from one axon excite the dendrites of other neurons through a gap called synaptic gap called synapse. Some synapse

encourage neurons to shoot and other don't, while it's a still a mystery how exactly our brain works but it is somehow formulated as linked with interconnection between neurons.

Artificial Neural Network try to model this low level functionality of the brain, whereas, in human brain many complex electrochemical process are going on even when simple task of thinking is performed. The argument for the neural net approach to AI is that, if we can model the low level activities correctly, the high level functionality may be produced as an emergent property [9].

Software model ANN has processing element having many inputs and associated weight, activation function that decides the threshold. ANN is trained by updating the associated weights change and weights are updated till minimum error occurs. Another connection to neurons can be made with value as input -1 and threshold. Neuron will give output only if sum is greater than zero.

Thus, we conclude that human brain neuron is not exactly similar to artificial neuron. Biological neuron has much more complex compared to artificial neuron. Pulse arrangements in biological are quite complex as compare to artificial one [9].

### **2.1.3 Learning in ANN**

One of most popular property of ANN is to learn from past experience same is the case of human brain that too can learn from past experience. In reality till now we have learned just few about how actually human brain functions, during learning of our brain synapse gap changes. This change in Artificial Neural Network is represented by weight updation as learning progresses weights gets updated till minimum error is reached. In human brain more relevant information makes synaptic gaps stronger thus becomes easy to recall and vice versa [9].

Thus ANN model is made to imitate and behave like human brain does by changes weights between layers in the network, which represents strengthening and weakening of synaptic gaps of humans. By this ANN network learns or gets trained just like humans have ability to do.

Learning is extremely helpful for solving real life problems where exact mathematical equation does not applies in other words in these cases programmer cannot apply mathematical formulae, so, learning algorithm solves the above problem, ANN uses one such learning algorithm.

Two types of learning's are used in ANN

- a) Supervised
- b) Unsupervised

Supervised learning is done in presence of targets, i.e. targets are provided along with inputs to converge as per targets i.e. output should be = targets. In this learning is done in presence of instructor. This type of learning requires algorithm to generalize the training data to forecast unseen situations [9]. Back propagation is form of supervised learning i.e. targets are provided in this learning.

Unsupervised learning is done in absence of targets i.e. no targets are provided with inputs, the most common method of this type of learning is cluster analysis, k-means. This type of learning does not occur in presence of instructor [9], but commonly used form of learning is supervised learning.

#### 2.1.4 *k*- Nearest Neighbor (*k*-NN)

*k*-NN is also known as lazy learning or instant based learning is another machine learning technique based on Euclidean distance between the neighbors. In lazy learning little processing is done during training, major part of processing done only at instance last minute when test sample is given, hence, named instance learning.

#### 2.1.5 *k*-NN Classification

In *k*-NN classification, the test sample is compared with most similar training sample already stored during training; similarity is computed by distance metric called as Euclidean distance. Based on the nearest distance the test sample is classified to that output class training samples. If  $k = 1$ , then the object is simply assigned to the class of that single nearest neighbor [27].

In *k*-NN regression, the output is the belongings value for the object. This value is the average of the values of its *k*-NNs [27]. The nearest neighbor based on Euclidean distance is shown in Fig 2.3

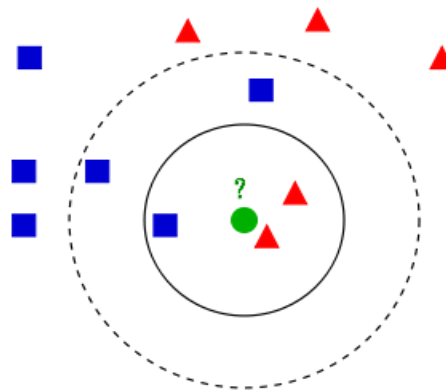


Fig 2.3 Nearest Neighbor based on Euclidean Distance [27]

The test sample (green circle) should be classified either to the first class of blue squares or to the second class of red triangles. If  $k$  is equal to 3 (solid line circle) it is designated to the 2nd class because there are two triangles and only 1 square inside the inner circle. If  $k = 5$  (dashed line circle) it is assigned to the first class (3 squares vs. 2 triangles inside the outer circle) [27].

In learning phase of  $k$ -NN only storing of samples and target class value is done, rest of all processing is done in testing phase of samples.  $k$  is user defined value, Euclidean distance is commonly used measure for calculating distance between neighbors but in some case hamming distance may be used for text classification.  $k$  value is set by hit and trial method starting from  $k=1$ , then increasing the value of  $k$ , bigger the database of training samples larger should be the value of  $k$ , if  $k=$  infinite then  $k$ -NN prediction error will be equal to bays error rate.

### **3.1 Overview of Methodology**

Diabetes Diagnostic Expert System is made to help human expert i.e. doctor, patients in diagnosis of Diabetes Mellitus and its Types. Mainly there are two Types of Diabetes

- a) Type 1 or Juvenile Onset Diabetes
- b) Type 2 or Non-Insulin dependent Diabetes

The Diabetes Expert System also provides necessary prescription to cure the type of Diabetes patient has by providing a Diet chart for Type 1 and Type 2 patients. The work done in this dissertation has following three stages:

- i. Diagnosis Based on 12 Symptoms of Diabetes
- ii. Diagnosis Based on Test parameters using ANN and  $k$ -NN
- iii. Types 1 or Type 2 Classification and Diet recommendation

First stage of this application asks for basic 12 questions to the user and based on the weighted average of the input the system decided whether the user is possibly Diabetic sufferer or not. The second part of the work asks the user to have eight different parameters and user is asked to enter the numeric value of all the input parameters. Based on these eight inputs, application using ANN and  $k$ -NN performs computations using MATLAB Runtime Compiler (MCR) [28] and result is displayed on user interface window. ANN and  $k$ -NN is trained using Neural Pattern Recognition (NPR) tool of Neural Network. Normalized data is then trained using Back Propagation Scaled Gradient method of ANN; errors are back propagated till validation stops. Means Square Error (MSE) and accuracy are performance factors that determine the accuracy of the expert system. Similarly another classifier  $k$ -NN is used, with highest accuracy of 78.51%. We have used two classifiers at a time with same inputs at a time to avoid misclassification error that may occur in results of either classifier decisions. Decision of second stage is supported by numerical value of the class between 0 and 1. The closer the value to 1 class the more likely the user is Diabetic. In the third stage of this dissertation Diabetic Mellitus is further classified.

Diabetes Mellitus is mainly of Two Types

- a) Type 1
- b) Type 2

Once the patient is confirmed whether he/she is Diabetic based on the above symptoms and 8 numeric input values application predicts whether the user is Type 1 Diabetic or Type 2. The Diet charts for Type 1 and Type 2 Diabetic patients have also been provided.

**3.1.1 Flow Chart** consists of four stages where three stages are for diagnosis using classification technique and last fourth stage provides Diet recommendation based on Type of Diabetes patient is diagnosed. Flow chart is shown in Fig 3.1 below

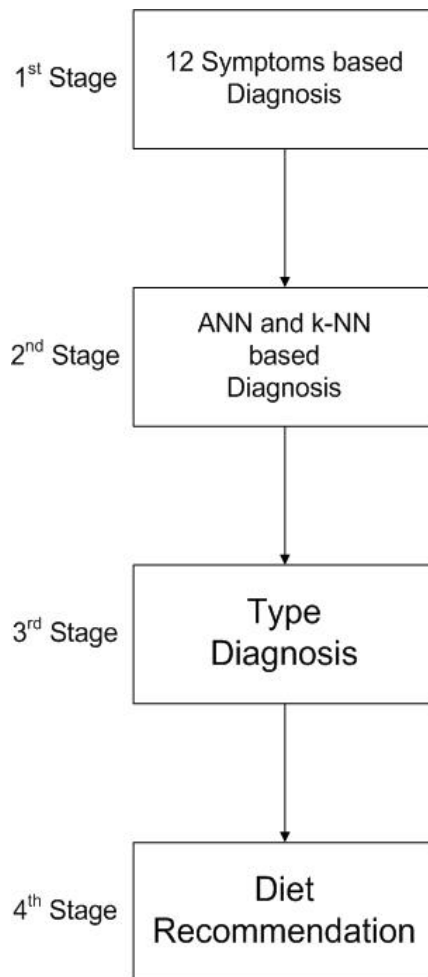


Fig 3.1 Four Layer Model of Diabetes Diagnosis Expert System

**3.2 Database used** Pima Indians Diabetes database from University of California Irvine (UCI) repository has been used for training of ANN and *k*-NN. This database has 8 input parameters and one output parameter DM. Database is 768 patients on 8 parameters and 9<sup>th</sup> output parameter [1, 3].

Parameters of this database are as follows

1. Number of times pregnant
2. Plasma glucose concentration (glucose tolerance test)
3. Diastolic blood pressure (mm Hg)
4. Triceps skin fold thickness (mm)
5. 2-Hour serum insulin (mu U/ml)
6. Body mass index (weight in kg/(height in m)<sup>2</sup>)
7. Diabetes pedigree function
8. Age (years)
9. Class value (0 or 1)

9<sup>th</sup> parameter is output parameter DM to indicate whether the patient is Diabetic or not,

- (i) DM= 0 Indicates Test Result Negative
- (ii) DM=1 Indicates Test Result Positive

Class 0 has 500 repetitions and class 1 has 268 repetitions out of total 768 samples. Classes and repetition of classes in database is shown in Table 3.1

Table 3.1 Class Repetition Table [3]

Classes	Number of repetitions
0	500
1	268

Database contains 500 samples with negative test results (Non-Diabetic patients) and 268 samples with positive test results (Diabetic patients). Brief statistical analysis of eight parameters (input) of Pima Indians Diabetes Database is shown in Table 3.2.

Table 3.2 Statistics of Pima Indians Diabetes Database [3]

Attribute Number	Mean	Standard Deviation
1	3.8	3.4
2	120.9	32.0
3	69.1	19.4
4	20.5	16.0
5	79.8	115.2
6	32.0	7.9
7	0.5	0.3
8	33.2	11.8

### 3.2.1 Normalization of data

Normalization of data of 768 samples is done using following algorithm in MATLAB editor, process of Normalization of data is shown in Fig 3.2 below

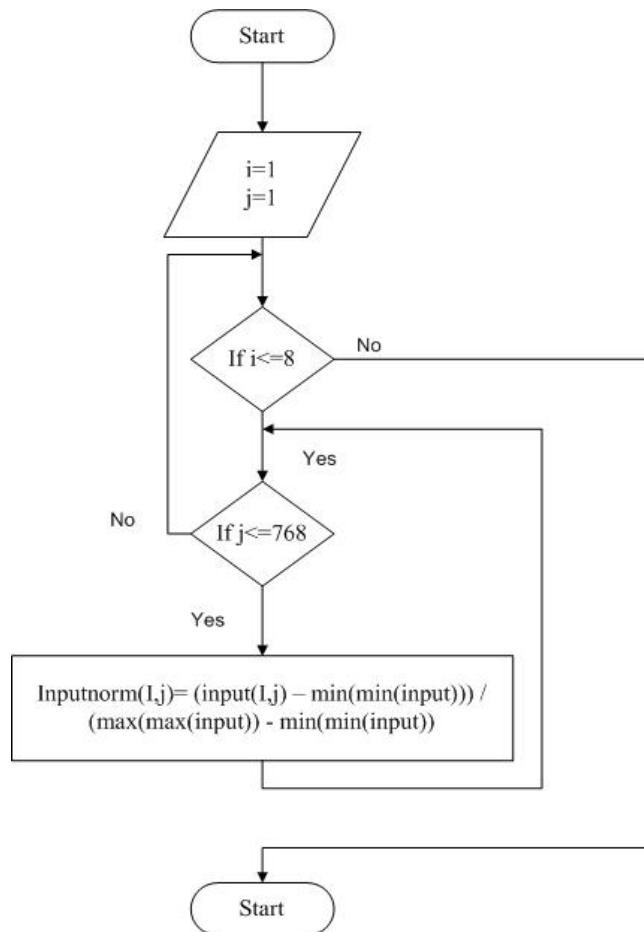


Fig 3.2 Process of Normalization

By normalization of input samples all input samples are converted in the value range of 0 to 1. The main advantage of normalization is that it helps to reduce redundancy. This reduces the disk space for samples.

### 3.3 Symptoms Based Diagnosis Algorithm

This is first stage of this work, symptoms are given weights each symptom is given a different weight as per its accountability to confirm a patient Diabetic. Here total sum of all symptoms are calculated by adding all the symptoms whose inputs are Yes given by user on user interface window.

Following questions are asked in the first stage:

- 1) Do you feel blurred vision?
- 2) Do you often feel thirst?
- 3) Do you often feel tiredness?
- 4) Are you having weight loss?
- 5) Do you have frequent urination?
- 6) Are you having slow healing of wounds?
- 7) Do you have genital itching?
- 8) Do you feel vomiting and stomach pain?
- 9) Do you feel lack of interest and concentration?
- 10) Do you feel Extreme hunger?
- 11) Do you feel Tingling or Numbness in Hands, Legs or Feet?
- 12) Do you feel very dry or itchy skin [7]?

The patient response is recorded in the form of y or n.

y -----→Means Yes

n -----→Means No

The user interface window is shown in Fig 3.3.

```
Please Enter the symptoms(specify the answer with "Y" for Yes or "N"
Do you feel blurred vision?:Y
Do you often feel thirst?:N
Do you often feel tiredness?:Y
Are you having weight loss?:N
Do you have frequent urination?:Y
Are you having slow healing of wounds?:N
Do you have genetical itching?:N
Do you feel vomiting and stomach pain?:N
Do you feel lack of intrest and concentration?:Y
Do you feel Extreme hunger?:N
Do you feel Tingling or Numbness in Hands, Legs or Feet?:Y
Do you feel very dry or itchy skin.?:Y

SYMPTOMS BASED DIAGNOSIS RESULT=====>>>>
=====>>JASKARAN IS POSSIBLY NOT DIABETIC <<=====
```

Fig 3.3 .exe Input Symptoms Questions

Based on the input to these responses the result is declared with percentage of possibly of Diabetes. Symptoms are allocated weight as per their importance in deciding factor in diabetes possibility. Total weighted sum is calculated as shown below

$$\text{Total}=\text{sym}(1)*2+\text{sym}(2)*2+\text{sym}(3)*2+\text{sym}(4)+\text{sym}(5)*2+\text{sym}(6)*2+\text{sym}(7)+\text{sym}(8)+\text{sym}(9)+\text{sym}(10)+\text{sym}(11)*2+\text{sym}(12)$$

Here,

- Sym (1) ==→ Signifies Symptom of Blurred Vision
- Sym (2) ==→ Signifies Symptom of Thirst
- Sym (3) ==→ Signifies Symptom of Tiredness
- Sym (4) ==→ Signifies Symptom of Weight Loss
- Sym (5) ==→ Signifies Symptom of Frequent Urination
- Sym (6) ==→ Signifies Symptom of Slow Healing Of Wounds
- Sym (7) ==→ Signifies Symptom of Genital Itching
- Sym (8) ==→ Signifies Symptom of Vomiting and Stomach Pain
- Sym (9) ==→ Signifies Symptom of Lack of Interest and Concentration
- Sym (10) ==→ Signifies Symptom of Feel Extreme Hunger
- Sym (11) ==→ Signifies Symptom of Feel Tingling or Numbness in Hands, Legs or Feet
- Sym (12) ==→ Signifies Symptom of Very Dry or Itchy Skin

In the end, total sum is calculated and compared with a predefined cutoff weight value, if

1) Total >cut off = Then Possibly Diabetic

2) Total < cut off = Then Possibly Not Diabetic

We have used 12 most important and confirmative symptoms of Diabetic patients. Percentage of Diabetic is provided as:

If Total  $\geq 11$

Perc =  $(\text{Total}/18) * 100$ .

Here,

Perc ==  $\rightarrow$  Signifies percentage.

### 3.4 ANN and $k$ -NN Based Diagnosis

Second stage of this work is trained using Pima Indian Diabetes Database. Input parameters used for training are numeric values that user has to feed from user interface window, following are eight input parameters

1. Number of times pregnant
2. Plasma glucose concentration a 2 hours in an oral glucose tolerance test
3. Diastolic blood pressure (mm Hg)
4. Triceps skin fold thickness (mm)
5. 2-Hour serum insulin ( $\mu$  U/ml)
6. Body mass index (weight in kg/(height in m)<sup>2</sup>)
7. Diabetes pedigree function
8. Age (years) [1, 3]

Output (9<sup>th</sup> parameter)

9. Class variable (0 or 1)

0----- $\rightarrow$  Represents patient is Non Diabetic

1----- $\rightarrow$  Represents patient is Diabetic.

Input parameters, ANN and *k*-NN based Diagnosis of Diabetes is shown to user as shown in Fig 3.4.

```

Now Enter The Test Details
Number Of Times Pregnant<0-17>:0
Plasma Glucose Concentration a 2 Hours In An Oral Glucose Tolerance Test:<0-199>
155
Diastolic Blood Pressure <mm Hg><0-122>:120
Triceps Skin Fold Thickness <mm><0-99>:90
2-Hour Serum Insulin <mu U/ml><0-846>:800
Body Mass Index <weight in kg/<height in m>^2><0-67>:60
Diabetes Pedigree Function<0.078-465>:350
Age <years><21-81>:25

ANN BASED DIAGNOSIS RESULT=====>>>>
=====>>JASKARAN IS DIABETIC<<===== 0.8091

KNN BASED DIAGNOSIS RESULT=====>>>>
=====>>JASKARAN IS DIABETIC <<=====

```

Fig 3.4 Input Parameters, ANN and *k*-NN based Diagnosis of Diabetes

Before training data is normalize between range 0 and 1. In this network are trained using training samples from database. Then forecasting of test samples is done from trained network. ANN network with 86.1% is used for forecasting and *k*-NN network is trained each time we give test samples.

### 3.5 Artificial Neural Network (ANN) Training Technique

Back Propagation Scaled Gradient technique is used for training of network in ANN, Back propagation (supervised learning) means “Backward Propagation of Errors”, a method used for training in ANN, back propagation is used along with optimization method scaled gradient. Gradient is given to optimization method which is used to change the weights as per the sign of gradient. Back propagation consists of two sub parts propagation of errors and weight updation. Output performance is Mean Square Error (MSE). Diagram of ANN using back propagation is shown below

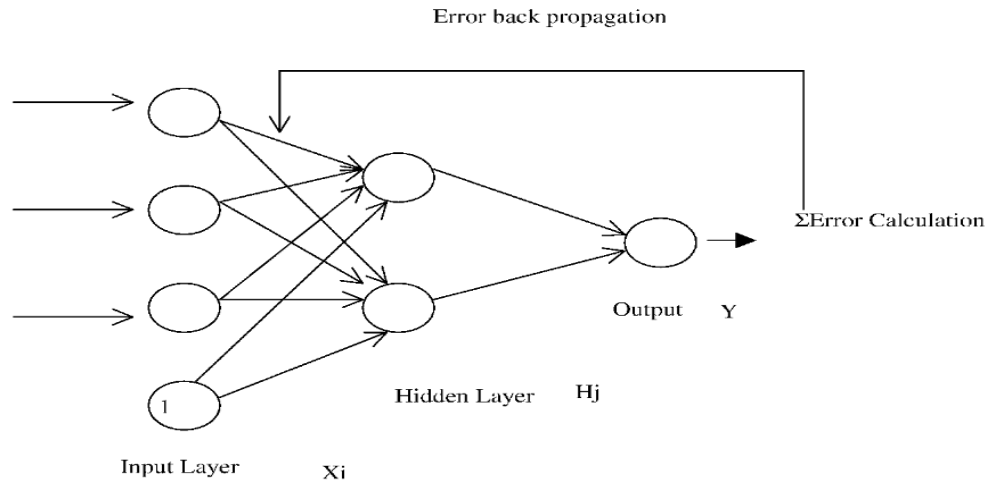


Fig 3.5 Back Propagation of Errors in Back Propagation Network [9]

The Back Propagation consists of two steps: propagation of errors and weight update.

### Step 1: Propagation of Errors

Propagation of errors can be divided into

1. Forward propagation of input from side to side the neural network in order to generate the propagation's production activations.
2. Backward propagation of the output activations all the way through the neural network using the training pattern target in order to produce the deltas of all output and hidden neurons.

### Step 2: Weight Updation

For each weight synapse follow the following steps:

1. Multiply its output delta and input activation to get the slope of the weight.
2. Subtract a ratio of the gradient from the weight.

The learning rate decides the speed of learning which is expressed as above ratio. The higher the fraction, the quicker the neuron trains and the lesser the fraction, training accuracy is good. Whether the error is increasing or decreasing is decided by sign of gradient of weights.

#### 3.5.1 Method used for Training

Algorithm for a 3 layer network

Initialize weights

do  
 for Each training  
 Prediction = neural-net- output  
 Actual=teacher –output (ex)  
 Compute error (prediction –actual)  
 Compute  $\Delta W_h$  for all weights from hidden layer to output layer  
 Compute all weights from input to hidden layer  
 Update network weights  
 Until all examples classified correctly or another stopping criteria satisfied  
 Return the network

In this technique error propagates backwards from output to input .this algorithm calculates the Gradient of errors of the network that is used to adjust to network weights. This algorithm allows rapid convergence on acceptable local minima for errors.

These networks are multilayer Perceptron model and hidden layer used should have nonlinear activation function for more than one layer which is equivalent to many single layer models, linear network.

### 3.5.2 Weight Updation Delta Rule

Delta rule is also known as gradient descent rule in machine learning. For a neuron  $j$  with activation function  $g(x)$ , the delta rule for  $j_s$   $i$ th weight  $w_{ji}$  is given by:

$$\Delta w_{ji} = \alpha(t_j - y_j)g'(h_j).x_i \dots \dots \dots (i)$$

Where,

$\alpha$  is called learning rate

$g(x)$  is transfer function

$t_{ij}$  is the output targeted

$h_j$  is weighted sum of the inputs.

$y_i$  is the actual output

$x_i$  is the  $i$ th input

It can be written as

$$h_j = \sum x_i \cdot w_{j,i} \text{ and } y_j = g(h_j) \dots \dots \dots \text{(ii)}$$

The Delta rule commonly with linear transfer function is

$$\Delta w_{ji} = \alpha (t_j - y_j) x_i \dots \dots \dots \text{(iii)}$$

### 3.6 Training of ANN Network Using NPR Tool of MATLAB R2009a

In MATLAB R2009a Neural Pattern Recognition tool (NPR) [28] is used for training data and testing data, test with highest accuracy is used as network for application. Test accuracy with 115 samples for testing is 86.1 is used in the network. NPR tool is available in MATLAB toolboxes as: Start << toolboxes << more << neural network << neural pattern recognition tool [28]. NN train Window used for training of Network is shown in Fig 3.6 below

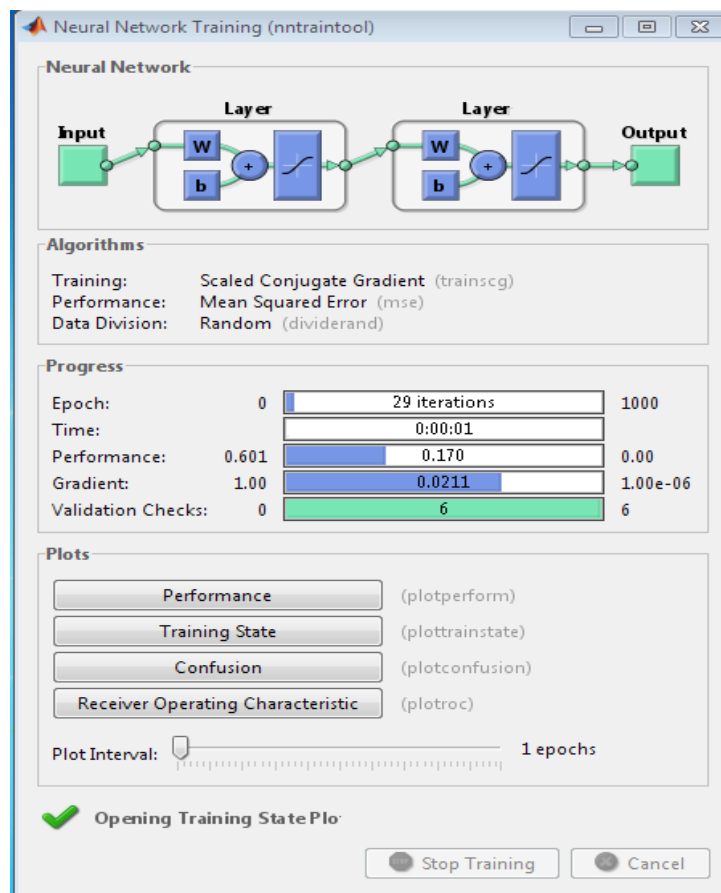


Fig 3.6 NN train Tool in MATLAB R2009a [28]

In NPR Tool following properties have been used as shown in Table 3.3 below

Table 3.3 NPR Tool Training Properties used

Training method	Scaled conjugate gradient
Figure of Merit	Mean square error and accuracy
Data division	Random
Total Epochs	29 iterations
Time required	1 sec
Gradient reached	0.0211
Validation checks	6
No of hidden layers used	64
Training samples	615

### 3.7 *k*-NN (*k*-Nearest Neighbor)

*k*-NN is also known as lazy learning or instant based learning is another machine learning technique based on Euclidean distance between the neighbors. This technique uses Euclidean distance for classification as a parameter. In this type of learning little processing is done during training, major part of processing done only at instance last minute when test sample is given hence, named instance learning.

#### 3.7.1 Method used

When *k*-NN is used as classifier, *k* is user defined constant. In this Euclidean distance is used as a distance metric.

*k*-NN is trained every time the program is executed NN is a simple algorithm that stores all available classes and classifies new cases based on similarity measure. Also called memory based or example based learning [10]

- (i) If  $k=1$ , select the nearest neighbour

(ii) If  $k > 1$ ,

(a) for classification select the most frequent neighbour

(b) for regression calculate the average of  $k$  neighbours,

It calculates the Euclidean distance as

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \dots \dots \dots (iv)$$

### 3.7.2 Selection of Value of $k$

Determined experimentally Start with  $k=1$  and use a test set to validate the error rate of the classifier. That value of  $k$  is chosen for which minimum is error rate.

Table 3.4  $k$ -NN Properties Table

Value of $k$ used	10
Data division	Random
Classification Method	Instance based
Scaling factor for weights $m=$	2
No of training samples	726
Test labels	Infinite
Info =	1

Normalization of database data is done before providing it to NPR tool of MATLAB for training purposes.

We have trained the ANN using Neural Pattern Recognition Tool (NPR) of MATLAB 7.8.0 (R2009a) and used the trained network using .mat files in the code of MATLAB. For  $k$ -NN network is trained every time the code is executed [28].

### 3.8 Type of Diabetes Diagnosis

This is third stage of this work, in this Diabetic Mellitus is further classified. Diabetes Mellitus is mainly of two Types

- a) Type 1
- b) Type 2

After the patient is confirmed to be Diabetic by second stage i.e. ANN and k-NN, then this stage further determines which category of Diabetes the patient or user has. BMI (Body Mass Index), Age and Race play an important role in determining the Type of diabetes, these parameters in combination with each other helps in determining Type of Diabetes [7]. Rules for determining the Type are stated below

#### 3.8.1 For Type 1 Diabetes Rule used

Necessary Condition for Type 1 Diabetes:

- 1) Sample (6)  $< 30/846$  and
- 2) Sample (8)  $< 40/846$

Here, sample 6 is BMI index (  $\text{kg/ m}^2$  )

Sample 8 is Age (years).

Thus, if a patient or user, Body Mass Index (BMI) is less than 30 and his/her Age is less than 40 then patient is Type 1 Diabetic [29].

#### 3.8.2 For Type 2 Diabetes Rule used

Necessary Condition for Type 2 Diabetes:

- 1) Sample (6)  $> 30/846$  and
- 2) Sample (8)  $> 40/846$  and
- 3) Race == 1.

or

- 1) Race not equal to 1 and
- 2) Sample (6)  $\geq 30/846$  and
- 3) Sample (8)  $\geq 25/846$  and

Here, Sample 6 is BMI index (  $\text{kg/ m}^2$  )

Sample 8 is Age (years)

Race number:

Race 1 White population

Race 2 Black population

Race 3 South Asian Gray population

Thus, from above rule it concludes that if patient or users, BMI is greater than equal to 30 and his/her age is greater than 40 along with user belongs to white community then user/patient is Type 2 Diabetic or if patient/user is other than white community i.e. from Black or South Asian Gray Community and his/her BMI is greater than equal to 30 along with his/her age is greater than or equal to 25 then patient is confirmed Type 2 Diabetic. Indian community belongs to South Asian Gray [29].

### **3.9 Diet Recommendation**

This is fourth and final stage of this work, after the patient/user is shown which Type of Diabetes he/she has, the finally user is asked few question about whether he/she want to know about Diabetes, Diet chart for Diabetes he/she has, detailed prescription for lifestyle planning is also provided in this. Questions are asked in the following ways

1) Do you want Diet chart?

Y or y -----→ provides diet chart to Diabetic respective of type of Diabetes one has

N or n -----→ carries the user to next stage of the program

2) Do you need detailed prescriptions about Diabetes lifestyle plan?

Y or y -----→ provides detailed knowledge about all the prescription needed to follow to cure Diabetes.

N or n -----→ carries the user to next stage of the program.

Fig 3.7 shows how system asks for Diet chart and prescription to the user on user interface window

```
CATEGORY:====>>>>
TYPE 2 DIABETES

Do you want to know diet chart for type 2 diabetes?
    1. Yes          2. No
Please Select an Option:2

Do you want detailed prescription of diabetes management?
    1. Yes          2. No
Please Select an Option:1
Warning: Image is too big to fit on screen; displaying at 17%
> In imuitools\private\initSize at 73
  In imshow at 262
  In final at 254
  0.7000

PRESS ANY KEY TO QUIT
```

Fig 3.7 Type Diagnosis and Diet Chart Recommendation on User Interface Window

PDF is provided for detail knowledge of the patient, Diet chart along with detailed prescription PDF presentation is provided to the patient in this system. User is asked whether he/she wants detailed prescription or not, as per requirement of patient system automatically provides the required information to the patient.

Standalone application of the above Expert System is made in deploy tool of MATLAB R2009a with inbuilt all the MATLAB Compiler Runtime (MCR) package, this facilitates the user to make portable program of MATLAB that can be run on any system without the need of MATLAB being installed on that PC or laptop [28]. The following four files are generated from MCR compilation process output as shown in Fig 3.8.

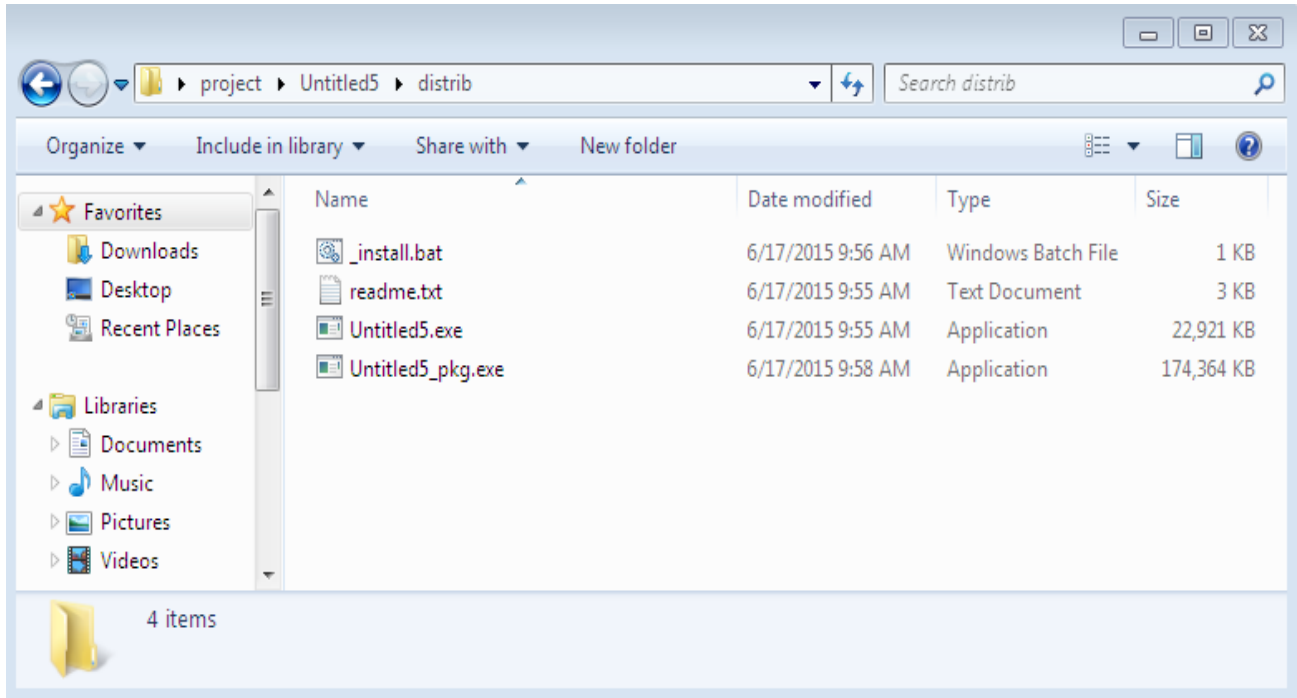


Fig 3.8 Files generated from MATLAB Runtime Compiler output.

These four files can be carried in folder, to make Expert System portable. These files contain package.exe file, when clicked it automatically installs MCR on new PC where MATLAB is not installed before.

**4.1 Test Confusion Matrix with 65% Training data using ANN**

Test Confusion matrix consists of two classes Class 1 and Class 2, percentage shown in green box indicates percentage of truly classified samples in Class 1 and Class 2 respectively. Percentage shown in white box indicates misclassified percentage of samples in both Class1 and Class 2 respectively. Overall test accuracy of 269 samples for 65% training is 81.0%. Test confusion matrix is shown in Table 4.1.

Table 4.1 Test Confusion Matrix

CLASS	1	2
1	53 73.6%	19 26.4%
2	32 16.2%	165 83.8%
OVERALL ACCURACY	81.0%	

**4.2 Gradient and Validation Curve**

Gradient curve shows the value of gradient reached after 24 is 0.040256 and maximum number of validation checks used is 6 for 24 epoch (iterations). This plot is shown in Fig 4.1.

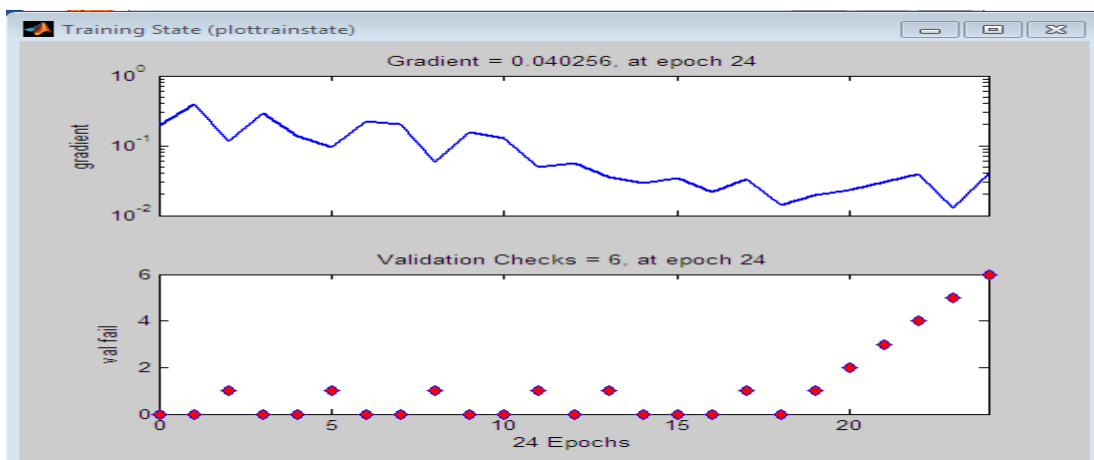


Fig 4.1 Gradient and Validation Plot

### 4.3 Receiver Operating Characteristics Plot

Receiver Operating Characteristics (ROC) curve is a performance index of a binary classifier system as discrimination threshold is varied. Curve is true positive (recall) vs. false positive (fall-out). Two colors curve represent two binary classes Class1 and Class2 here. This plot is shown in Fig 4.2.

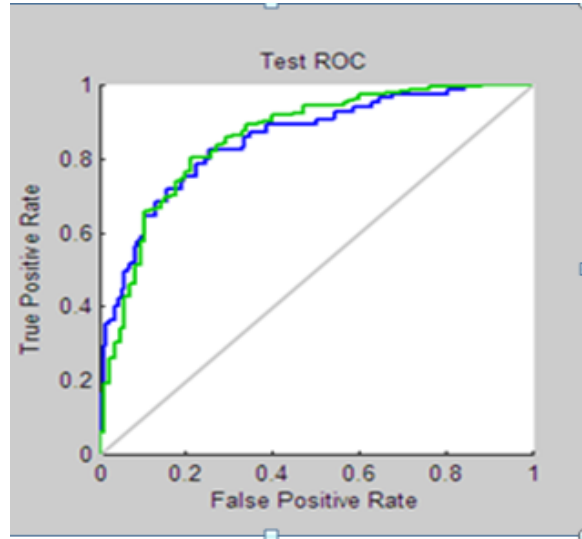


Fig 4.2 Roc Test Curve

### 4.4 Mean Square Error

This is Figure of Merit of classifier used to determine how effectively classifier is predicting. MSE depicts the average of squares of the errors i.e. difference between estimator and what is estimated. Lesser the % MSE for classifier more accurate the predictions will be. In this work we have achieved 18.95 % mean square error with 269 test samples. This is shown in Table 4.2.

Table 4.2 Means Square Error (MSE) Table

<b>RESULTS</b>	<b>NO. OF SAMPLES</b>	<b>MSE</b>	<b>%MSE</b>
<b>TRAINING</b>	<b>461</b>	<b>1.63955</b>	<b>24.51193</b>
<b>VALIDATION</b>	<b>38</b>	<b>1.46905</b>	<b>21.05263</b>
<b>TESTING</b>	<b>269</b>	<b>1.37987</b>	<b>18.95910</b>

#### 4.5 Test Confusion Matrix with 85 % Training data using ANN

Test Confusion matrix consists of two classes Class 1 and Class 2, percentage shown in green box indicates percentage of truly classified samples in Class1 and Class 2 respectively. Percentage shown in white box indicates misclassified percentage of samples in both Class1 and Class 2 respectively. Overall test accuracy of 115 samples for 85% training is 86.1%. This is shown in Table 4.3 below

Table 4.3 Test Confusion Matrix

CLASS	1	2
1	27 93.1%	2 6.9%
2	14 16.3%	72 83.7%
OVERALL ACCURACY	86.1%	

#### 4.6 ROC Plot

ROC curve for 85% Training the two colors are closer towards 1 true positive compared to with 65% training ROC test curve is in shown in Fig 4.3.

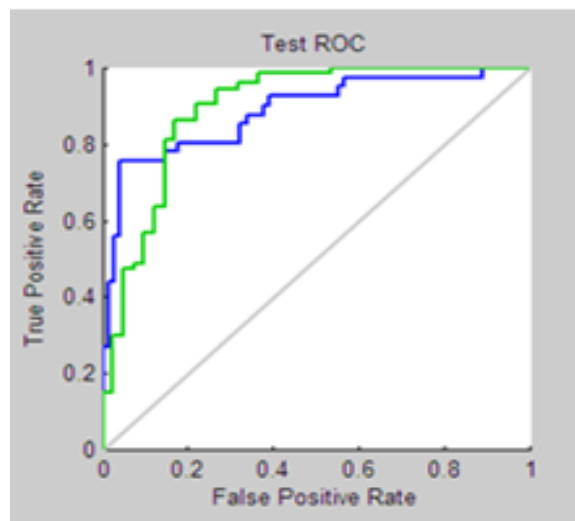


Fig 4.3 ROC Test Curve

#### 4.7 Gradient and Validation Curve

Gradient curve shows the value of gradient reached after 29 are 0.021116 and maximum number of validation checks used is 6 for 29 epochs (iterations). This is shown in Fig 4.4.

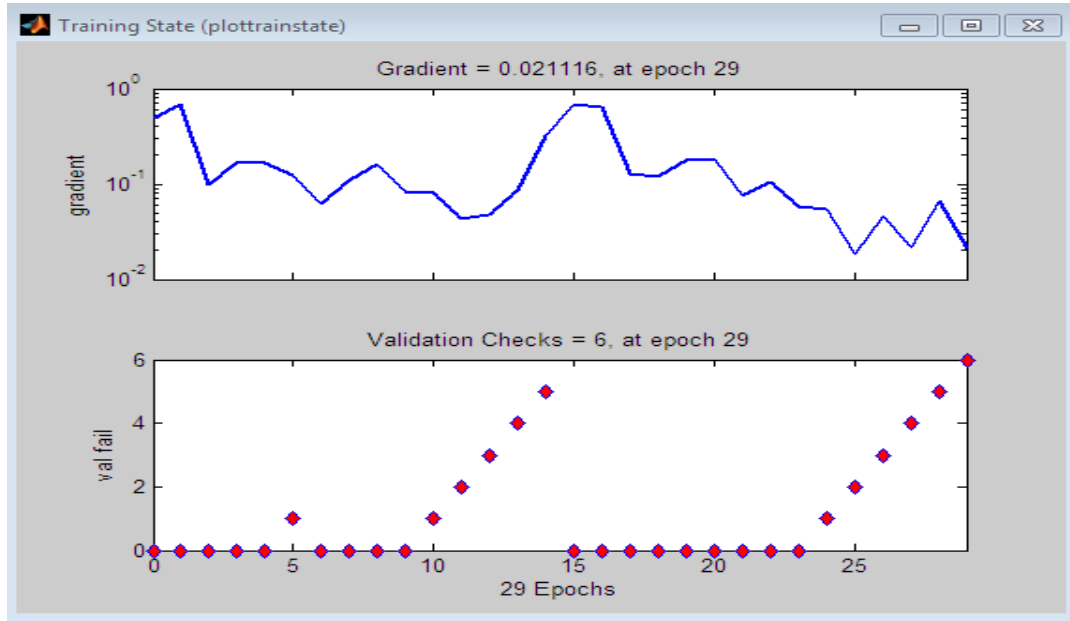


Fig 4.4 Gradient and Validation Plot

#### 4.8 Mean Square Error

Mean square of 85% training is comparatively less compared to 65% training. The MSE for 85% training is 13.913% compared to 18.9% for 65% training. Lesser the % MSE betters the accuracy of the classifier. This is shown in Table 4.4.

Table 4.4 Mean Square Error

RESULTS	NO. OF SAMPLES	MSE	%MSE
TRAINING	615	1.69805	24.39024
VALIDATION	38	1.78700	26.31578
TESTING	115	1.30121	13.913

#### 4.9 ANN Forecast and Target Data for first 20 Samples

ANN network after training is tested for samples from the database with known targets, but for test targets are not provided, ANN can forecast the targets of test samples. Table 4.5 shows the forecast value of targets of test samples.

Table 4.5 ANN Forecast Table

<b>SNO.</b>	<b>FORECAST</b>	<b>ACTUAL</b>	<b>ERROR</b>
1	0.6761	1	0.3239
2	0.0742	0	-0.0742
3	0.7319	1	0.2681
4	0.0692	0	.0692
5	0.6808	1	0.3192
6	0.2393	0	-0.2393
7	0.08552	1	0.91448
8	0.5539	0	-0.5539
9	0.8648	1	0.1352
10	0.1073	1	0.8927
11	0.3779	0	-0.3779
12	0.7982	1	0.2018
13	0.4318	0	-0.4318
14	0.8605	1	0.1395

15	0.5677	1	0.4323
16	0.3242	1	0.6758
17	0.4501	1	0.5499
18	0.2790	1	0.721
19	0.4876	0	-0.4876
20	0.2427	1	0.7573

#### 4.10 *k*-NN Test Accuracy Results

*k*-NN is trained using samples from database, Accuracy is the Figure of Merit use to define the effectiveness of the classifier, it can be clearly seen from the Table below as we increase the no of training samples the accuracy increases, *k*-NN accuracy with different training samples is shown in Table 4.6 below

Table 4.6 *k*-NN Test Accuracy Table

<b>NO. OF TRAINING SAMPLES</b>	<b>TEST ACCURACY FOR 115 SAMPLES</b>
100	70.8333
120	69.5313
140	69.5313
150	69.7917
160	70.0521
170	69.5313
180	68.6198
190	68.8802

200	70.7031
220	70.8333
250	72.9167
300	74.2188
350	73.8281
320	73.4375
400	75
450	74.0885
420	74.4792
500	74.4792
520	74.8698
538	75.3906
550	75.5208
600	75.6510
615	76.9531
700	77.9948
765	78.2552
750	78.3854
767	78.1250

755	78.2552
720	78.2552
725	78.3854
<b>726</b>	<b>78.5156</b>
727	78.3854
768	78.1250

Highest accuracy of 78.5156 % for 726 samples as Training data has been achieved.

#### 4.11 Overall Comparison Table

Overall accuracy Table shows highest accuracy with given same database of 768 samples is 86.1% for 115 test samples. Highest accuracy of  $k$ -NN achievable with given database is 78.5 % for 115 test samples. Overall accuracy comparison Table of all three techniques used is shown in Table 4.7.

Table 4.7 Overall Accuracy Comparisons

Method used	Total no of sample	Training samples%	Test samples%	No of true positives	No of false positives	Overall accuracy
ANN	768	60%+5%	35%	218	51	81.0%
ANN	768	80%+5%	15%	99	16	86.1%
$k$ -NN	768	85%	15%	88	27	76.9%

#### 4.12 User Interface Standalone Application .exe

Standalone application .exe is made using deploy tool of MATLAB R2009a [28], this facilitates to make portable exe programs that are directly executable on windows platform without being of MATLAB R2009a installed on the computer. It carries MATLAB Compiler Runtime (MCR) [28] that can performs all the computations necessary for using ANN and *k*-NN that actual MATLAB software can do. MCR is installed on remote computer automatically in just 10 minutes as compared to MATLAB requires approximate 45 minutes to get installed.

```
C:\Users\jas\Desktop\project\Untitled5\distrib\Untitled5.exe
*** TEST FOR DIABETES ***

Please Enter Your Name:TTTT
1. White
2. Black
3. South Asian Gray
Please Specify your race number:3

Please Enter the symptoms(specify the answer with "Y" for Yes or "N"
Do you feel blurred vision?:Y
Do you often feel thirst?:N
Do you often feel tiredness?:Y
Are you having weight loss?:N
Do you have frequent urination?:Y
Are you having slow healing of wounds?:N
Do you have genietal itching?:Y
Do you feel vomiting and stomach pain?:N
Do you feel lack of intrest and concentration?:Y
Do you feel Extreme hunger?:Y
Do you feel Tingling or Numbness in Hands, Legs or Feet?:Y
Do you feel very dry or itchy skin.?:N

SYMPTOMS BASED DIAGNOSIS RESULT=====>>>>
=====>>TTTT IS 61.111111 PERCENT POSSIBLY DIABETIC <<=====

Now Enter The Test Details
Number Of Times Pregnant(0-17):0
Plasma Glucose Concentration a 2 Hours In An Oral Glucose Tolerance Test:(0-199)
179
Diastolic Blood Pressure (mm Hg)(0-122):50
Triceps Skin Fold Thickness (mm)(0-99):36
2-Hour Serum Insulin (mu U/ml)(0-846):159
Body Mass Index (weight in kg/(height in m)^2)(0-67):37.8
Diabetes Pedigree Function(0.078-465):.455
Age (years)(21-81):22

ANN BASED DIAGNOSIS RESULT=====>>>>
=====>>TTTT IS DIABETIC<<===== 0.7783

KNN BASED DIAGNOSIS RESULT=====>>>>
=====>>TTTT IS DIABETIC <<=====
```

Fig 4.5 .exe of User Interface

#### 5.1 Conclusion

Since there is vast number of people worldwide who suffers from Diabetes, the use of Expert System for predicting Diabetes is highly helpful. Most of the times the patients are not aware of the disease, its symptoms and don't know how to manage it. Also during most of cases patients don't have access of physicians during high time in order to tackle the above problem so the need of Expert System for Diabetes is highly desirable. One such system has been developed by us using ANN and *k*-NN. The developed system can present needed information regarding diagnosis and recommendations for patients with Diabetes, as system is made by gathering expertise from many medical specialists and experts, so it has wider practical importance for patients.

The proposed expert system first of all finds the possibility of the patient being Diabetic. Further, classification of Type 1 and Type 2 type of Diabetes after the patient is confirmed Diabetic makes the system more informative and useful. Further treatment, suggestions are provided for giving prescriptions to patient with particular type of Diabetes whether Type1 or Type 2. Diet chart is presented for patients with Type 1 and Type 2 after patient is confirmed Diabetic which was not provided in earlier models.

A good prediction accuracy of 86.1 % using ANN and 76.9 % using *k*-NN has been achieved, this is considerably better as far previous model using same database were made using ADAP with accuracy of 72.3%. Overall accuracy of our system is far better than similar models. Use of two classifiers at the same time reduces the misclassification by one classifier thus, final result is cross checked from both results this is major advantage of using two classifiers at the same time in decision support system.

#### 5.2 Future Work

Diabetes diagnosis Expert System is a new and efficient program. However, chances of improvement are always there, therefore future work of this dissertation can be carried as

- 1) Diagnostics of Diabetes Insipidus using ANN and  $k$ -NN can be done.
- 2) More input Parameters can be added.
- 3) LS- SVM can be incorporated which may improve performance.
- 4) Fuzzy  $k$ -NN can also be used.
- 5) GUI can be implemented to make it more users friendly.

## REFERENCES

---

- [1] American Diabetes Association, “American Diabetes Association”, 2015. [Online]. Available: <http://www.diabetes.org/?referrer=https://www.google.co.in/>. [Accessed on 14- Jun- 2015].
- [2] Wikipedia Diabetes mellitus 2015. [Online]. Available: [https://en.wikipedia.org/wiki/Diabetes\\_mellitus](https://en.wikipedia.org/wiki/Diabetes_mellitus). [Accessed: 14- Jun- 2015].
- [3] Archive.ics.uci.edu, “Index of /ml/machine-learning-databases/pima-Indians-diabetes”, 2015. [Online]. Available: <https://archive.ics.uci.edu/ml/machine-learning-databases/pima-indians-diabetes/>. [Accessed: 14- Jun- 2015].
- [4] [Online]. Available: [http://diabetes.niddk.nih.gov/dm/pubs/riskfortype2/risk\\_508.pdf](http://diabetes.niddk.nih.gov/dm/pubs/riskfortype2/risk_508.pdf). [Accessed: 15- Jun- 2015].
- [5] Diabetes.co.uk, “Diabetes Symptoms - Diabetes Mellitus Signs”, 2015. [Online]. Available: <http://www.diabetes.co.uk/diabetes-symptoms.html>. [Accessed: 14- Jun- 2015].
- [6] Wikipedia, “Expert system”, 2015.[Online]. Available: [https://en.wikipedia.org/wiki/Expert\\_system](https://en.wikipedia.org/wiki/Expert_system). [Accessed: 14- Jun- 2015].
- [7] American Diabetes Association, “Standards of medical care in diabetes -2007”, diabetes care, vol.30, no.1, pp. S4-S41, 2007.
- [8] D. Campos, M. Hernandez, R. Femat and A. Gorrillo, “Fuzzy-based controller for glucose regulation in type 1 diabetic patients by subcutaneous route”, IEEE Trans, Biomed .Engineering, vol. 53, no. 11, pp.2201-2210, November 2006.
- [9] Wikipedia, “Artificial neural network”, 2015.[Online]. Available: [https://en.wikipedia.org/wiki/Artificial\\_neural\\_network](https://en.wikipedia.org/wiki/Artificial_neural_network). [Accessed: 14- Jun- 2015].
- [10] I.M. Ahmed, A.M. Ahmoud, M. Aref and A. B. M. Salem, “A study on Expert Systems for Diabetic Diagnosis and Treatment,” Recent Advances in Information Science, pp. 363-367.
- [11] F. Hill and J. Ernest, “Jess, the java expert system shell”, Distributed Computing Systems, Sandia National Laboratories, USA, 1997.

- [12] M. Garcia, A. Gandhi and T. Singh, "Diabetes (An expert system in diabetes)", JCSC 16, 2001.
- [13] S. Kumar and B. Bhimrao, "Development of knowledge Base Expert System for Natural treatment of Diabetes disease", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 3, No. 3, 2012.
- [14] M. Wiley and R. Bunesco, "Emerging Applications for Intelligent Diabetes Management Cindy Marling", Association for the Advancement of Artificial Intelligence, 2011.
- [15] P. Devamalar, M. Bai and T. Srivatsa, "An Architecture for a Fully Automated Real-Time Web-Centric Expert System", World Academy of Science, Engineering and Technology, 2007.
- [16] I.M. Ahmed, A.M. Ahmoud, M. Aref and A. B. M. Salem, "A study on Expert Systems for Diabetic Diagnosis and Treatment," Recent Advances in Information Science, pp. 363-367.
- [17] M. Kalpana and S. Kumar, "Fuzzy Expert System For Diagnosis of Diabetes Using Fuzzy Determination Mechanism", International Journal For Computer Science and Emerging Technology.
- [18] B. Jaafar and M. Ali, "Diabetes mellitus forecast using artificial neural network (ANN)", Asian Conference on Sensors and the International Conference on new Techniques in Pharmaceutical and Biomedical Research, pp. 135-139, Sept 2005. Doi: 10.1109/ASENSE.2005.1564523.
- [19] C. Lee, M. Wang and H. Hagra, "A Type-2 Fuzzy Ontology and Its Application to Personal Diabetic-Diet Recommendation", IEEE Transactions on Fuzzy Systems, vol.18,no.2, pp.374,395, April 2010. Doi: 10.1109/TFUZZ.2010.2042454.
- [20] C. Lee and M. Wang, "A Fuzzy Expert System for Diabetes Decision Support Application," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol.41, no.1, pp.139-153, Feb. 2011 doi: 10.1109/TSMCB.2010.2048899.
- [21] R. Rahman and F. Afroz, "Comparison of Various Classification Techniques Using Different Data Mining Tools for Diabetes Diagnosis", JSEA, vol. 06, no. 03, pp. 85-97, 2013.

- [22] C. Park and H. Park, "Nonlinear Discriminant Analysis Using Kernel Functions and the Generalized Singular Value Decomposition", *SIAM. J. Matrix Anal. & Appl.*, vol. 27, no. 1, pp. 87-102, 2005.
- [23] G. Baudat and F. Anouar, "Generalized Discriminant Analysis Using a Kernel Approach", *Neural Computation*, vol. 12, no. 10, pp. 2385-2404, 2000.
- [24] I.M. Ahmed, A.M. Ahmoud, M. Aref & A. B. M. Salem, "A study on Expert Systems for Diabetic Diagnosis and Treatment," *Recent Advances in Information Science*, pp. 363-367.
- [25] K. Polat, S. GÃneÅÿ and A.Arslan, "A cascade learning system for classification of diabetes disease: Generalized Discriminant Analysis and Least Square Support Vector Machine", *Expert Systems with Applications*, vol. 34, no. 1, pp. 482-487, 2008.
- [26] T. Zeki, M. Malakooti, Y. Ataeipoor and S. Tabibi, "An Expert System for Diabetes Diagnosis", *American Academic & Scholarly Research Journal*, vol. 4, no. 5, 2015.
- [27] Wikipedia, "K-nearest neighbors algorithm", 2015. [Online]. Available: [https://en.wikipedia.org/wiki/K-nearest\\_neighbors\\_algorithm](https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm). [Accessed: 14- Jun- 2015].
- [28] MATLAB documentation, functions and data sheet [Online]. Available: <http://www.mathworks.in/products/matlab/> [Accessed: 14- Jun- 2015].
- [29] Ucsfhealth.org, "Diabetes Mellitus Signs and Symptoms | Conditions & Treatments | UCSF Medical Center", 2015. [Online]. Available: [http://www.ucsfhealth.org/conditions/diabetes\\_mellitus/signs\\_and\\_symptoms.html](http://www.ucsfhealth.org/conditions/diabetes_mellitus/signs_and_symptoms.html). [Accessed: 11- Jul- 2015].
- [30] J. Cantais, D. Dominguez, V. Gigante, L. Laera1 and V. Tamma, "An example of food ontology for diabetes control", *International Semantic Web Conference*, 2005.
- [31] S. Lee, W. Jian and K. Huang, "A fuzzy ontology and its application to news summarization", *IEEE Trans Syst. Man Cybern B Cybern*, vol. 35, no. 5, pp. 859-880, Oct. 2005
- [32] W. Szajnar and G. Setlak, "A concept of building an intelligence system to support diabetes diagnostics", *Studia Informatica*, 2011.
- [33] K. Polat and S. Gunes, "An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease", *Digital Signal Process.*, vol. 17, no. 4, pp. 702 -710 2007

- [34] G. Underwood, "Diabetes Aid: A System for the diagnosis and management of diabetes using a palm pilot", Master of Science, University of Florida, 2015.
- [35] J. W. Smith, J. E. Everhart, W. C. Dickson, W. C. Knowler and R. S. Johanne, "Using the ADAP learning algorithm to forecast the onset of diabetes mellitus," In Proceedings of the Annual Symposium on Computer Application in Medical Care , American Medical Informatics Association, November, 1994, pp.261.

**A.1 Diet Chart for Type 1 Diabetes**

The secret to a Type 1 Diabetes Diet is focusing on nutritious foods that are high in vitamins and minerals, while cutting down on those that are high in fat, sugar, or empty carbohydrates. Actually, it’s a good Diet for anyone, Diabetic or not. Food categories with food names is shown below in Table A.1

Table A.1 Food Categories and names

<b>Food Categories</b>	<b>Food Names</b>				
Whole grains	Brown rice	Bran Cereal	Whole Grain Breads		
Carbohydrates	Beans	Starchy Vegetables	Fruit juices	Pasta	Bread
Proteins	Meat	Proteins	Beans	Eggs	

1) Fruits

Fruits are natural sources of sugar and should be counted as carbohydrates if you’re using a diet plan. The best ways to eat them are fresh or frozen. Citrus fruits, like oranges and grapefruit, are best.

2) Vegetables

Starch is a form of sugar that naturally occurs in many common vegetables, such as potatoes, corn, and peas. These vegetables aren’t bad. However, they contain more carbohydrates than other vegetables and should be used sparingly. Focus instead on vegetables that have few carbohydrates but are rich in vitamins, minerals, fiber, and phytochemicals. These include:

- (a) Most green leafy vegetables
- (b) Asparagus
- (c) Beets
- (d) Carrots
- (e) Celery

- (f) Cucumber
- (g) Onions
- (h) Peppers
- (i) Sprouts
- (j) Tomatoes

Always choose fresh or frozen vegetables without added salt or sauces.

### 3) Whole grains

Your body will convert whole grains into sugar, but at least they'll be packing nutrition and extra fiber. Brown rice, bran cereal, and whole grain breads are great sources of whole grains.

### 4) Carbohydrates

Carbohydrates turn to sugar in your blood. This raises your glucose level about an hour after you eat. There are three types of carbohydrates: Starches, Sugars.

### 5) Proteins

Proteins are extremely important in maintaining muscle and repairing wounds. Besides meat, proteins are found in beans and eggs. Proteins won't raise your blood sugar. However, processed or fatty meats also contain fat, sodium, and cholesterol. While these substances have no effect on Diabetes, too much of them can have other adverse health effects.

## **A.1.1 When to Eat**

Breakfast really is the most important meal of the day. A healthy breakfast can get your blood sugar back up after a night's rest. Exercise and physical activity lowers blood sugar, so if you're going to do intense exercise, you'll want to measure your blood sugar before and after you exercise.

Eating smaller meals and progressively snacking throughout the day can make your blood sugar easier to monitor and prevent levels from peaking. Fruits, vegetables, nuts, and other foods travel easily and are great to have on hand when you need them. Eating big meals filled can spike your blood sugar. While it's not a good idea, you still have to live your life. In times like the holidays, big, delicious meals are not only inevitable, they're tradition. Be sure to keep an eye on your blood sugar and use your insulin accordingly.

Food pyramid for Type 1 Diabetes is shown in Fig A.1 below

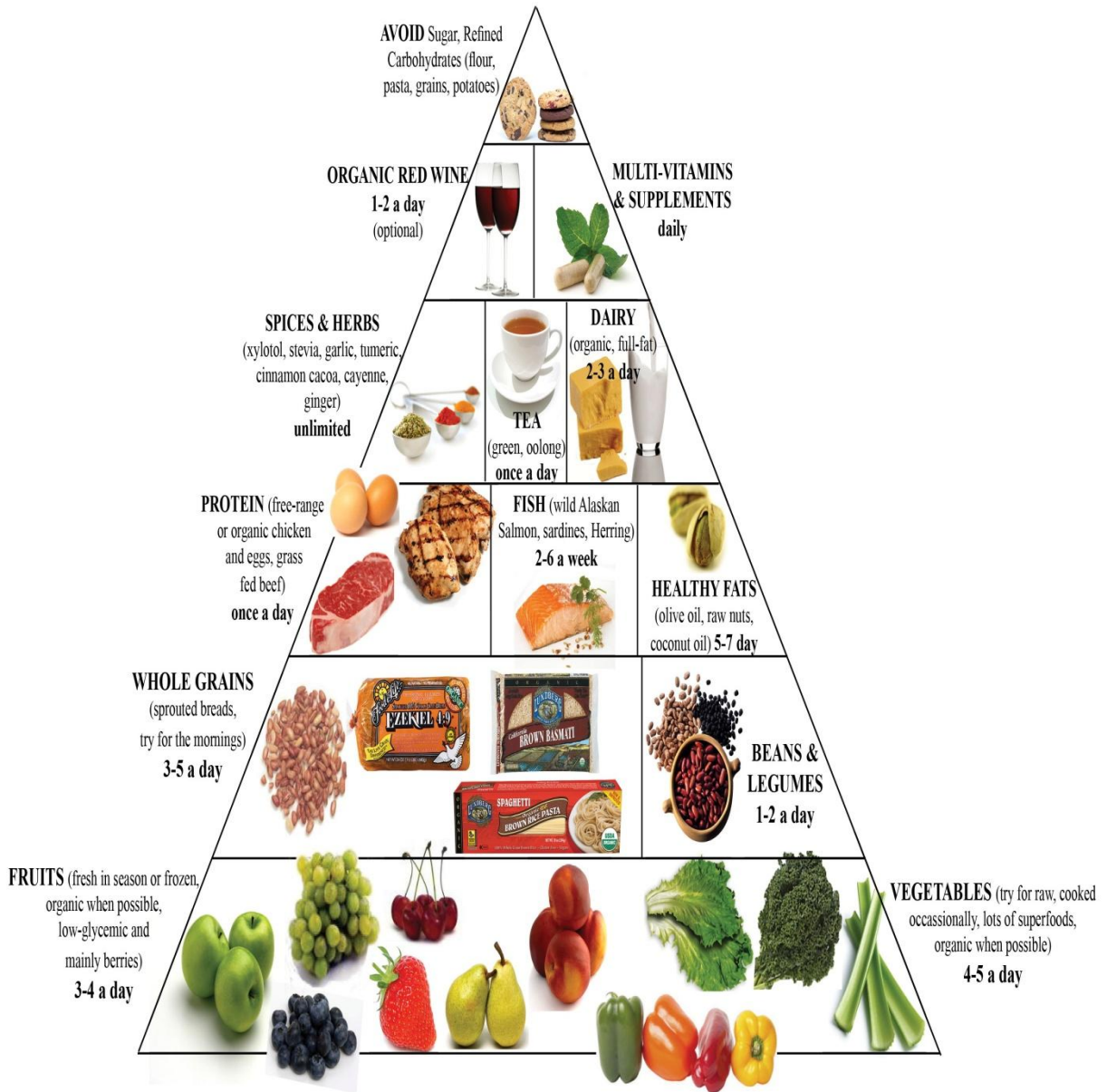


Fig A.1 Food pyramid for Type 1 Diabetes

## **B.1 Diet Chart for Type 2 Diabetes**

This eating plan is low in refined grains and sugar, low in saturated and trans fat and high in fibre. It focuses on eating regularly timed meals and selecting low glycemic index foods.

Focus on Carbohydrate, Fibre and Glycemic Index. By choosing appropriate portions of carbohydrate containing foods and selecting ones that have more fibre and a lower glycemic index, you help improve your blood glucose control.

Low GI Foods and Their Glycemic Index Glycemic index is a scale (0-100) ranking how quickly carbohydrate containing food will digest into glucose in our blood. High GI foods break down quickly where as low GI foods break down slowly. With low GI foods you feel full longer and you body's insulin has more time to perform its job and remove glucose from the blood. Low GI foods are generally close to the farm. " low gi foods (55 or less) choose most often medium gi foods Go Low GI Foods and Their Glycemic Index Glycemic index is a scale (0-100) ranking how quickly a carbohydrate containing food will digest into glucose in our blood. High GI foods break down quickly where as low GI foods break down slowly. With low GI foods you feel full longer and you body's insulin has more time to perform its job and remove glucose from the blood.

### **B.1.1 Tips for Using the Glycemic Index**

Think low and slow choose from the low and medium GI foods most often. Balance your choices with lean protein choices and healthy fats. Protein and fat also slow the digestion, keep you feeling full and slow the release of glucose into your blood. Balancing Carbohydrates (Carbs") Along with the glycemic index approach, it is important to understand and control the amount of carbohydrates you eat and drink to better manage your blood glucose. Here are some tips:

1. Carbohydrate is the nutrition term used for starch, sugar and fibre.
2. Spacing your carbohydrates over the day helps your body to keep a stable glucose level

3. Starch choices that are higher in fibre contribute less to raising your blood glucose. Use the Nutrition Facts panel to choose cereal, bread, crackers, rice and other grain or starch choices that have at least 2 grams of fibre per serving
4. Breakfast should be 1/3 starch or grain, 1/3 fruit and 1/3 protein
5. Lunch and supper should be 1/2 vegetables, 1/4 starch and 1/4 protein
6. Snack choices can help prevent dips in blood glucose that can occur if your meals are more than 4 hours apart or if you are very active
7. Choose snacks with about 20 grams of carbohydrate. Fruit, vegetables, high fiber granola bars, whole grain crackers with cheese or peanut butter are just few options. Sugar and Refined Starches Table sugar, brown sugar, honey, maple syrup and molasses are all simple sugars and they all have 5 grams of carbohydrates or “carbs” in a teaspoon. Simple sugars such as these and „white“ or refined starches break down quickly and cause blood glucose to rise quickly.
  - a) A quick rise in blood glucose can often be followed by an equally quick drop
  - b) This can lead to cravings for more sugary foods and feelings of hunger and fatigue
  - c) This can lead to cravings for more sugary foods and feelings of hunger and fatigue
  - d) This can lead to cravings for more sugary foods and feelings of hunger and fatigue
  - e) Avoid adding sugar, reduce the amount or substitute a non-sugar sweetener for baking such as Splenda (sucralose).

Food names based on GI are given below in Fig B.1

<b>LOW GI FOODS (55 OR LESS)</b> CHOOSE MOST OFTEN	<b>MEDIUM GI FOODS (56-69)</b>	<b>HIGH GI FOODS (70+)</b> CHOOSE LESS OFTEN
Whole grain bread	Couscous	White bread
Pumpkin bread	Rye bread	Instant mashed potatoes
Oatmeal	Instant Oatmeal	Corn Flakes, Rice Krispies
All-Bran cereal	Shredded Wheat	Refined, sweetened cereals
Converted rice	Cream of Wheat	Instant rice
Brown & Basmati rice	Whole grain crackers	Bagels
Bulgur, Barley, Quinoa	Pita bread	Waffles/pancakes – made with white flour
Firm cooked pasta	Long grain white rice	Soda crackers
Beans, peas, lentils	Apricot, banana	French fries
Apples, peaches, pears	Cantaloupe	Dried dates/figs
Grapefruit, oranges	Pineapple, raisins	Sweetened fruit juice
Berries, cherries, grapes	Canned fruit in juice	Parsnips, pumpkin
Kiwi, Mango, Plum	Cranberry juice	Rutabaga, turnip
Avocado	New potatoes	Broad beans
Sweet Potato	Beets	Refried beans
Carrots, broccoli	Sweetened condensed milk	Ice cream
Cauliflower, corn		Soft drinks
Leafy vegetables		Glucose
Low fat milk, soymilk, yogurt and cottage cheese		
	*Adapted from "The GI Diet" Rick Gallop	

Fig B.1 Food names based on GI