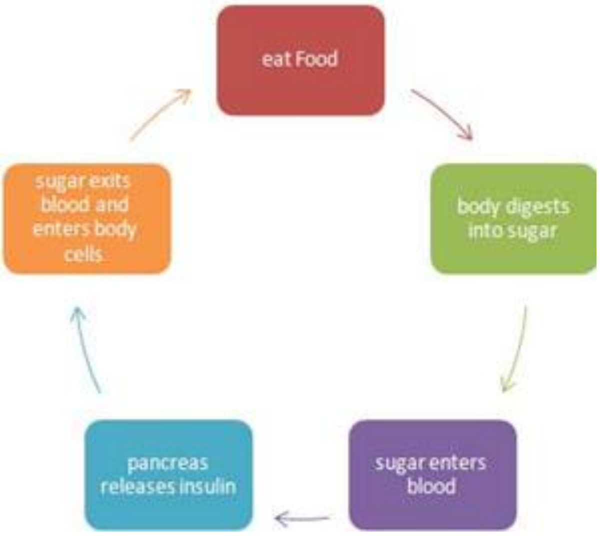

Figure 3: This image indicates how the body without diabetes functions, with Type 2 diabetes, the pancreas is making less insulin, and/or the body experiences issues utilizing that insulin.

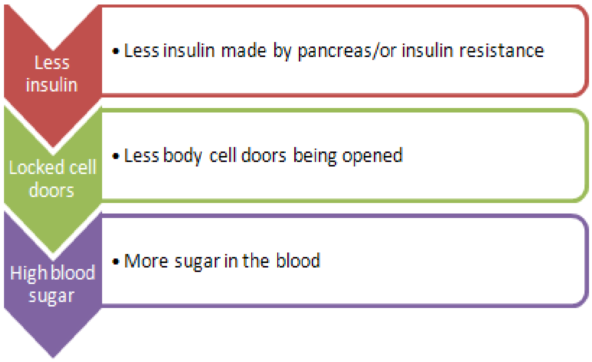

Figure 4: This image demonstrates how what is occurring in the body with Type 2 diabetes and/or pre- diabetes.

What is Pre-Diabetes?

Pre-diabetes implies that the pancreas is making less insulin than it was in the recent past, and/or that your body is getting to be impervious to insulin (not able to utilize the insulin that is accessible). An individual with pre-diabetes has a higher than typical level of sugar in their blood, yet those levels are not yet sufficiently high for a finding of diabetes (hindered fasting glucose, or disabled glucose resilience).

A conclusion of pre-diabetes may be seen as a "reminder" to start to settle on changes in regular choices so that a determination of diabetes may be evaded. The Important thing to know is that Type 2 diabetes can be counteracted by eating healthy, being physically dynamic most days of the week, and losing even a little measure of weight. [2]

Bayesian Analysis Approach

Bayesian Network:

A Bayesian Network consists of the following:

- A set of variables and a set of directed edges between variables.
- Each variable has a finite set of mutually exclusive states.
- The variables together with directed edges form a directed acyclic graph(DAG).
- To each variable D with parents V1,….Vn,there is attached a conditional probability table that is P(D/V1,…,Vn). [3]

D-Separation

Two distinct variables A and B are d-separated,if for all paths between A and B there is an intermediate variable V,(distinct from A and B) such that, the connection is serial or diverging is instantiated and the connection is converging, and neither V nor any of V's descendants have received evidence. If A and B are d-separated, then we call them as d-connected. [3][4]

The Chain Rule or Factoring:
We can always write
P(a, b, c, … z) = P(a | b, c, …. z) P(b, c, … z)
(by definition of joint probability)
Repeatedly applying this idea, we can write
P(a, b, c, … z) = P(a | b, c, …. z) P(b | c,.. z) P(c| .. z)..P(z)
This factorization holds for any ordering of the variables.This is the chain rule for probabilities.

Figure 5: Symptoms and after-effects of Type 2 Diabetes

2. Methodology

The authors selected the Bayesian analysis approach to diagnose type 2 diabetes in patients. For this analysis a dataset of size 50 has been collected from the Abbasi Shaheed Hospital, Karachi ,Pakistan. In this survey, an endeavor is being made to give a review in regards to the applications of Bayes' hypothesis and clinical choice investigation in touching base at a conclusion. With a specific end goal to comprehend demonstrative thinking, it is vital to comprehend the essential numerical dialect of probability and Bayes' hypothesis as connected to clinical medication. [3][4]

Probabilistic reasoning in Clinical Diagnosis:

Probability as connected to clinical indicative thinking may be viewed as a measure of one's quality of conviction that an occasion will happen and range from 0.0 to 1.0. In factual documentation, probability of an occasion A is composed as P[A].[3][6]

Law of Total Probability

The summation rule states that the entirety of probabilities of every single conceivable result of a chance occasion rises to 1.0. On the off chance that there are four conceivable results as K,L,M and N then
P[K] + P[L] + P[M] + P[N] = 1.

Joint Probability

The associative event of any number of occasions will be characterized as joint likelihood of those occasions. In factual documentation, the joint likelihood of two occasions An and B is composed as P[A,B]. [3][5][6]

Contingent Probability

The probability that an occasion A happens, given that the occasion B will be known to happen will be characterized as the contingent likelihood of occasion A given occasion B alternately P[A/B]. The relationship in the middle of joint and contingent probabilities is given by the equation P[A,B] = P[A/B] * P[B]. Authors have used eight parameters/factors/symptoms as part of the diagnosis of diabetes.[3][4]

Figure 6: Main Symptoms of Diabetes

Table1: Factors/symptoms of diabetes

Symptoms (V)	State of facts	Permitted data values
Age(V_1)	Age of the Person: Age is divided into three sub parameters, young (≤ 30),middle-age(30<V1≤50),old age(>50).	Discrete integer values
Family History(V_2)	Either any predecessor or successor is / was suffering type II diabetes.	Yes or No
Alcohol/Smoking (V_3)	Either the person does or does not drinking/Smoking.	Yes or No
Weakness(V_4)	Does a person get weakened in doing a little effort.	Yes or No
Frequency of maturation/Urination(V_5)	Number of times the person passes urine in a day	Discrete integer value.
Pancreatic disease (V6)		Yes or No
Pregnancy (V7)	Either the female subject is pregnant or not	Yes or No
Overweight (V8)		Yes or No

Each symptom or parameter in Table 1 has independent contribution to the prediction of the final result. Mathematically the probability model for a classifier is a conditional model p(D/$V_1, V_2, \ldots, V_n$), over a dependent class variable with a small number of outcomes or classes, conditional on several feature variables V_1 *to* V_n. Baye's Theorem can be demonstrated as:

$$p(D/V_1,V_2,\ldots\ldots,V_n)=$$

$$p(D)p(V_1,V_2,\ldots\ldots,V_n)/\,p(V_1,V_2,\ldots\ldots,V_n)$$

That is

Posterior = (Prior x Likelihood) / Evidence
The numerator can be represented as joint probability model $(D,V_1,V_2,\ldots\ldots,V_n)$. By definition of Conditional probability:

$$=p(D,V_1,V_2,\ldots\ldots,V_n)$$
$$=p(D)\,p(V_1,V_2,\ldots\ldots,V_n/D)$$
$$=p(D)p(V_1/D)p(V_2/D,V_1)p(V_3,\ldots\ldots,V_n/D,V_1,V_2)$$
$$=p(D)p(V_1/D)p(V_2/D,V_1)p(V_3,\ldots\ldots,V_n/D,V_1,V_2)p(V_4,\ldots\ldots,V_n/D,V_1,V_2,V_3)$$
$$=p(D)p(V_1/D)p(V_2/D,V_1)p(V_3,\ldots\ldots,V_n/D,V_1,V_2)\ldots\ldots\ldots p(V_n/D,V_1,V_2,V_3,\ldots\ldots,V_{n-1}).$$

Now the "naive" conditional independence assumptions applied as: presuming that individual symptom for diabetes is conditionally independent of every other symptom for diabetes V_j for i not equal to j. This means that:

$$p(V_i/D,V_j)=p(V_i/D)$$

For i not equal to j and so the joint probability can be expressed as:

$$p(D, V_1, V_2, \ldots, V_n) = p(D)p(V_1/D)p(V_2/D)p(V_3/D) \ldots = p(D)\Pi p(V_i/D)$$

Table 2: (V,D)

	D (Yes)	D (No)
V1(young)	0.0001	0.000001
V1(middle)	0.0015	0.00001
V1(old)	0.002	0.00002
V2 (yes)	0.05	0.00638
V2 (No)	0.025	0.002008
V3 (yes)	0.014	0.0212
V3 (No)	0.1	0.01
V4 (yes)	0.0042	0.006
V4 (No)	0.001	0.00252
V5 (yes)	0.14	0.0065
V5 (No)	0.016	0.03
V6(yes)	0.064	0.001
V6 (No)	0.021	0.02
V7 (yes)	0.106	0.035
V7 (No)	0.102	0.0014
V8 (yes)	0.051	0.008
V8 (No)	0.12	0.03
	0.82	0.18
Sum		1

Formula:

$$P(V/D) = \frac{P(D|V)P(V)}{P(D)} = \frac{P(V,D)}{\sum_D P(V,D)}$$

Using the above table, the authors created the following resultant table of P(V|D).

P (V/D) =

	D(yes)	D(No)
V1	0.0044	0.00017
V2	0.0912	0.0466
V3	0.13902	0.1733
V4	0.0185	0.04736
V5	0.1902	0.20277
V6	0.1036	0.1166
V7	0.2542	0.2022
V8	0.196	0.211
Sum	1	1

3. Conclusion

In the previous decade, there have been colossal advances in the utilization of Bayesian procedure for examination of epidemiologic information, and there are currently numerous useful focal points to the Bayesian approach. Bayesian models can undoubtedly suit in secret variables, for example, a singular's actual illness status in the vicinity of symptomatic slip. The utilization of earlier likelihood circulations speaks to a capable component for fusing data from past studies and for controlling jumbling. Apparatuses are currently accessible that permit disease transmission experts to exploit this capable way to deal with evaluation of presentation sickness relations.

References

[1] Harris M, Zimmet P. Classification of diabetes mellitus and other categories of glucose intolerance. In Alberti K, Zimmet P, Defronzo R, editors. International Textbook of Diabetes Mellitus. Second Edition. Chichester: John Wiley and Sons Ltd; 1997. p9-23.http://www.medicalnewstoday.com/info/diabetes/

[2] Williams textbook of endocrinology (12th ed.). Philadelphia: Elsevier/Saunders. pp. 1371–1435. ISBN 978-1437703245. Bayes. An essay towards solving a problem in the doctrine of chances. Biometrika1958; 45:296-315.

[3] Bayes T. An essay towards solving a problem in the doctrine of chances. Philos Trans R Soc Lond 1763; 53:370–418.Ashby D. Bayesian statistics in medicine: a 25 year review. Stat Med 2006;25:3589-631.

[4] Jilani T. A., Shoaib M., Rehan R., Rehman B. (2014), "A Comparative Study of Data Mining Techniques for HCV Patients' Data", Journal of Applied Environmental and Biological Scinces 12/2014; 4(9S):217-223. (ISI Indexed).

[5] Jilani T. A., H. Yasin and M. Danish (2011), PCA-ANN for Classification of Hepatitis-C Patients, International Journal of Computer Applications 14(7), pp. 1–6, February 2011.

[6] Jilani T. A. , Huda Yasin, Madiha Yasin, Cemal Ardil (2009), Acute Coronary Syndrome Prediction Using Data Mining Techniques- An Application, International Journal of Computational Intelligence, 5(4), pp. 295-299.