AI BASED DIABETES PREDICTION SYSYEM

(Understand the requirements and plan the design of system)

Submitted by

J.Gophisha III ECE A

Gce Tirunelveli(9508)

Phase 1 document submission

IBM NAAN MUDHALVAN

WHAT IS DIABETES?

- Diabetes is a disease that occurs when your blood glucose, also called blood sugar, is too high.
- Glucose is your body’s main source of energy. Your body can make glucose, but glucose also comes from the food you eat
- Insulin is a hormone made by the pancreas that helps glucose get into your cells to be used for energy.
- If you have diabetes, Then your body doesn’t make enough insulin, or doesn’t use insulin properly. Glucose then stays in your blood and doesn’t reach your cells

TYPES OF DIABETES

.

Two types:

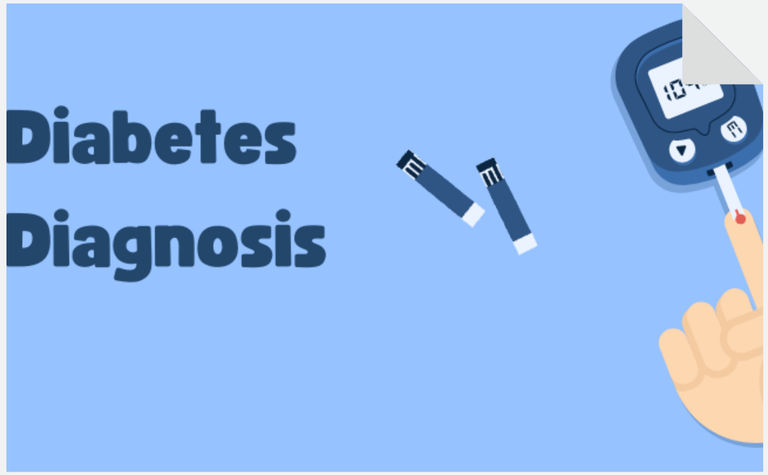
- Type1
- Type2

•If you have type 1 diabetes, your body makes little or no insulin. Your immune system attacks and destroys the cells in your pancreas that make insulin

•If you have type 2 diabetes, the cells in your body don’t use insulin properly. The pancreas may be making insulin but is not making enough insulin to keep your blood glucose level in the normal range

You are more likely to develop type 2 diabetes if you have risk factors, such as overweight or obesity, and a family history

PROBLEM STATEMENT



The system aims to determine whether a patient is at risk of diabetes based on various characteristics and features and provide personalized preventive measures, allowing individuals to take proactive actions to manage their health.

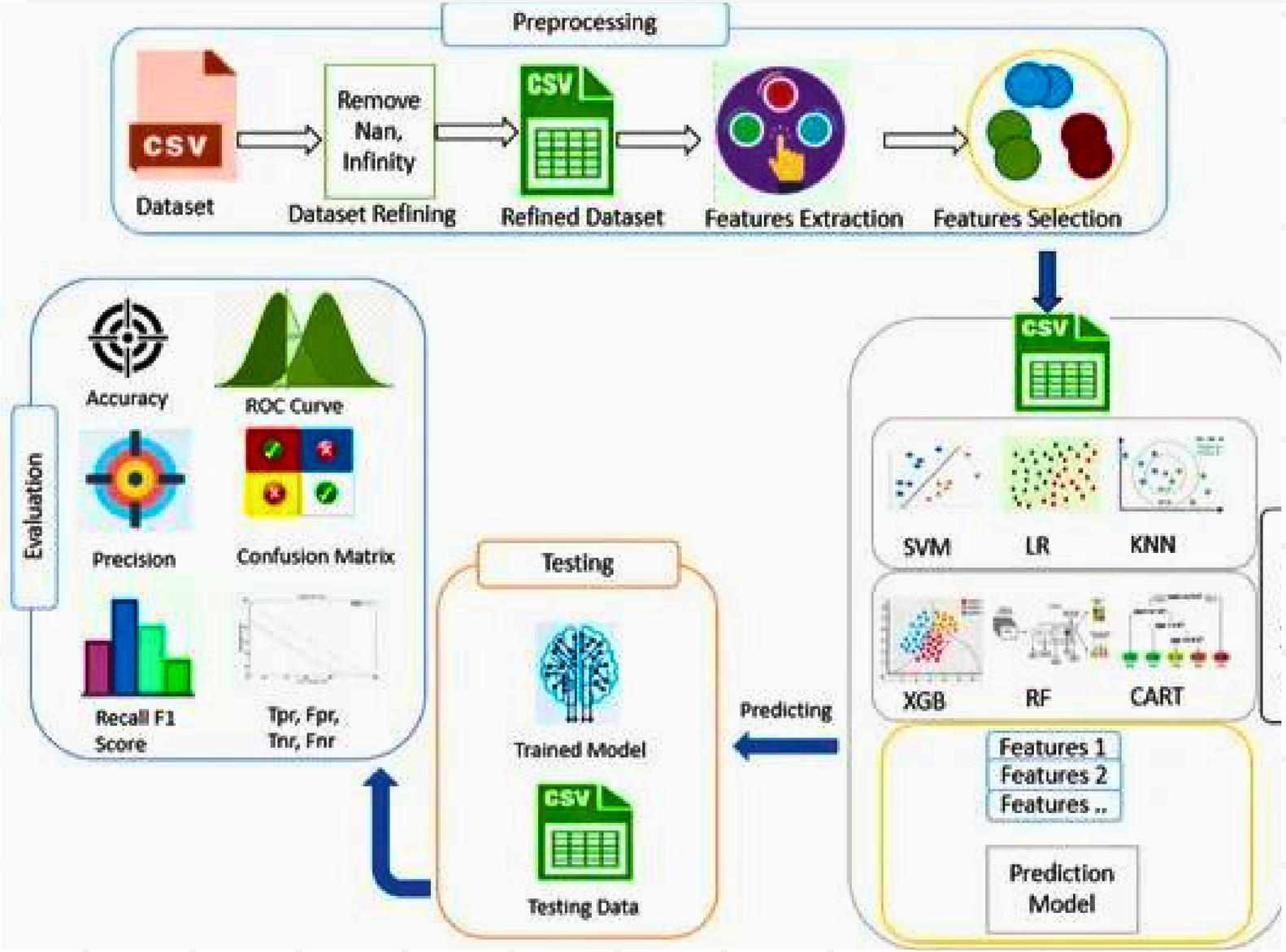
WHAT IS THE PLAN FOR DEVELOPING THIS PROJECT

COLLECT DATASET → PREPROCESS THE DATA → SELECT FEATURES

↓

TRAIN A MACHINE LEARNING MODEL → EVALUATE THE MODEL → DEPLOY THE MODEL

PROPOSED MODEL



DESIGN OF THE PROBLEM

1.data collection

	A	B	C	D	E	F	G	H	I
1	ıcies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigre	Age	Outcome
2	6	148	72	35	0	33.6	0.627	50	
3	1	85	66	29	0	26.6	0.351	31	
4	8	183	64	0	0	23.3	0.672	32	
5	1	89	66	23	94	28.1	0.167	21	0
6	0	137	40	35	168	43.1	2.288	33	1
7	5	116	74	0	0	25.6	0.201	30	0
8	3	78	50	32	88	31	0.248	26	1
9	10	115	0	0	0	35.3	0.134	29	0
10	2	197	70	45	543	30.5	0.158	53	1
11	8	125	96	0	0	0	0.232	54	1
12	4	110	92	0	0	37.6	0.191	30	0
13	10	168	74	0	0	38	0.537	34	1
14	10	139	80	0	0	27.1	1.441	57	0
15	1	189	60	23	846	30.1	0.398	59	1
16	5	166	72	19	175	25.8	0.587	51	1
17	7	100	0	0	0	30	0.484	32	1
18	0	118	84	47	230	45.8	0.551	31	1
19	7	107	74	0	0	29.6	0.254	31	1
20	1	103	30	38	83	43.3	0.183	33	0
21	1	115	70	30	96	34.6	0.529	32	1
22	3	126	88	41	235	39.3	0.704	27	0
23	8	99	84	0	0	35.4	0.388	50	0
24	7	196	90	0	0	39.8	0.451	41	1
25	9	119	80	35	0	29	0.263	29	1
26	11	143	94	33	146	36.6	0.254	51	1
27	10	125	70	26	115	31.1	0.205	41	1

The initial step is to gather a large dataset that includes patient demographics, medical history, lab results, and other related information. This data can be obtained from electronic health records (EHRs) or other sources

DATASET ATTRIBUTES

- Pregnancis: Number of times pregnant
- Glucose: Plasma glucose concentration a 2 hours in an oral glucose tolerance test
- Blood Pressure: Diastolic blood pressurure(mm Hg)
- SkinThickness: Triceps skin foldt hickness (kg mm)
- Insulin 2-Hour serum insulin (mu U/ml)
- BMI: Body mass index (weight in kg/(height in m)^2)
- DiabetesPedigreeFunction: Diabetes pedigree function
- Age: Age (years)
- Outcome: Class variable (0 or 1) 268 of 768 are 1, the others are 0

Dataset source: https://www.kaggle.com/datasets/mathchi/diabetes-data-set

STEPS

2.DATE PREPROCESSING

Clean and preprocess the data. Handle missing values, outliers, and normalize features for consistency.

3.FEATURES ENGINEERING

New features may be created from the existing data to improve the predictive performance of the system.

STEPS

4.MODEL TRAINING

machine learning model is trained on the prepared data. The model learns to identify patterns in the data that are associated with diabetes. Choose machine learning algorithms suitable for classification tasks. Common choices include logistic regression, decision trees, random forests,

5.EVALUATION

Select appropriate evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to measure the model's performance.

STEPS

6.CONTINUOUS MONITORING

Implement a system for monitoring the model's performance over time. Retrain the model periodically with new data to keep it up to date.

7.DEPLOYMENT

Integrate the trained model into a user-friendly interface. This could be a web or mobile app for easy accessibility.

BENIFITS

EARLY DETECTION

AI systems can detect diabetes and its associated complications at an earlier stage than traditional methods. This allows for earlier intervention and improved patient outcomes.

PERSONALIZED RISK ASSESSMENT:

AI systems can provide personalized risk assessments for individuals based on their individual risk factors. This can help individuals to make informed decisions about their health and lifestyle.

SCALABILITY

AI systems can be scaled to analyze large amounts of data from large populations. This makes them ideal for use in public health initiatives to identify and target individuals at high risk of developing diabetes.

CONCLUSION

AI-based diabetes prediction systems have the potential to revolutionize the way we detect, prevent, and manage diabetes. These systems have the potential to improve the lives of millions of people worldwide.

THANK YOU