

AI-BASED ON DIABETES PREDICTION SYSTEM

ABSTRACT:

- Diabetes is one of the most deadly and chronic diseases which cause an increase in blood sugar.
- Diabetes remains untreated and unidentified many difficulties may arise due to that.
- The tedious work is in identifying the process which results in visiting the clinic and consulting the doctor.
- This tedious work has been solved with the rise in the approaches used by machine learning.
- Over recent years a plenty of growth has been seen doing over health informatics by focusing on the technology of the presentation, generation and application of clinical information in health care.
- Its accessibility and understanding have become easier with language technologies.
- This paper gives a comprehensive prospect of work accomplished to develop a model that can predict

the possibility of diabetes in patients with extreme accuracy

PROGRAM:

```
# Import necessary libraries
```

```
import numpy as np
```

```
import pandas as pd
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LogisticRegression
```

```
from sklearn.metrics import accuracy_score,  
confusion_matrix,classification_report
```

```
#load the dataset (replace 'diabetes.csv' with your dataset)
```

```
data=pd.read_csv('diabetes.csv')
```

```
#split the dataset into features(X) and target (Y)
```

```
X= data.drop('outcome',axis=1)
```

```
Y= data['outcome']
```

```
#split data into training and testing sets
```

```
X_train,X_test,Y_train,Y_test = train_test_split(X,Y,test_size =  
0.2,random_state=42)
```

```
#create a random forest classifier model
```

```
Model = RandomForestClassifier(n_estimators =  
100,random_state =42)
```

```
#Train the model
```

```
Model.fit (X_train,Y_TRAIN)
```

```
#Make predictions on the test set
```

```
Y_pred = model.predict(X_test)
```

```
#Evaluate the model
```

```
accuracy = accuracy_score(y_test,y_pred)
```

```
confusion = confusion_matrix(y_test,y_pred)
```

```
classification_report_str =  
classification_report(y_test,y_pred)
```

```
#print the evaluation results
```

```
Print(f'Accuracy:{accuracy:.2f}')
```

```
Print('Confusion Matrix:')
```

```
Print(confusion)
```

```
Print(Classification Report:')
```

```
Print(classification_report_str)
```

INPUT:

The patient is likely does not have diabetes

OUTPUT:

Pregnancies:5

Glucose:120

BloodPressure:70

SkinThichness:30

Insulin:80

BMI:25

DiabetePedigreeFuction:0.4

Age:35

CONCLUSION :

The main objective of this paper is to improve the correctness of predictive models. The accuracy can be achieved by either refining the performance of the data or with the help of an algorithm. For achieving best results the data can be improved at the earlier phase. PIMA dataset has been taken for performing the accuracy checks on each classifier. Among all, the genetic algorithm leads over

others. In this research, the author has concluded one more important factor that the accuracy of a model is highly dependent on the database