

# **Diabetics Prediction Using Machine Learning**

December 7, 2022

C V Rahul Bharadwaj-CB.EN.U4CSE19416

D. Bargav Srinivas-CB.EN.U4CSE196612

K. D Saran-CB.EN.U4CSE19125

J. Venkat Rithvik-CB.EN.U4CSE19326

Project Mentor: Vijay Kumar Sundar, Asst.Prof, Dept. of Computer Science & Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham.  
Coimbatore

## **Abstract**

Diabetes is a disease caused due to high glucose level in a human body. Diabetes ought not be overlooked in the event that it is untreated, Diabetes might cause a few significant issues in an individual like: heart related issues, kidney issue, circulatory strain, eye harm and it can likewise influences different organs of human body.

Diabetes can be controlled on the off chance that it is anticipated before. To accomplish this objective this task work we will do early expectation of Diabetes in a human body or a patient for a higher exactness through applying, Various Machine Learning Techniques. AI procedures Provide improved outcome for expectation by con-structing models from datasets gathered from patients.

## **Acknowledgement**

We would like to thank our supervisor Dr.VIJAY KUMAR SUNDAR for their invaluable guidance and support throughout our research. We would also like to thank Dr.Anantha Narayanan V and Dr. Govindarajan J for their help in reviewing our project to take further steps in project

## **Introduction**

Diabetes is a major health problem worldwide, affecting millions of individuals every year. It is estimated that over 422 million people worldwide have diabetes and the number continues to rise. Early detection and prevention are critical in controlling the spread of diabetes, and machine learning can be used to predict the likelihood of developing diabetes. In this thesis, we propose a novel machine learning approach to diabetes prediction.

## **Motivation**

Existing machine learning models for diabetes prediction have been limited in their accuracy and scope. Existing models typically rely on a small set of features such as age, gender, and body-mass index, which may not accurately reflect the complex interplay of factors that contribute to diabetes. Additionally, existing models often neglect to consider genetic factors, which play a significant role in diabetes prediction. Our goal is to develop a more accurate model for predicting diabetes based on a combination of demographic, clinical, and genetic data.

## **Major Contribution**

The major contribution of a machine learning project for diabetes prediction would be to develop an automated decision-support system. This system would be capable of using predictive analytics to identify individuals who are at high risk of developing diabetes, and predicting the onset times of diabetes for these individuals. Additionally, the system would be able to provide tailored advice to individuals who are at risk of developing diabetes, in order to reduce their chances of developing the disease. This system could be used to provide early interventions, potentially saving lives and reducing healthcare costs.

### **Algorithm Used**

In this project we used Random forest Algorithm, Decision tree, Xg boost classifier and Support vector machine in this we are getting high accuracy in Xg boost classifier(0.779527)

### **Accuracy For our ML Models**

Random forest-0.775

Decision Tree-0.70866

Xg boost Classifier-0.779527

SVM-0.74803

### **Observation**

Machine learning can be used to make predictions about the likelihood of a person developing diabetes, based on a range of features such as age, gender, BMI, and family history. By analyzing a large dataset of patient records, a machine learning model can be trained to accurately identify the risk factors associated with the condition. This can be used to identify individuals who might benefit from further monitoring or early intervention, allowing for early diagnosis and management of diabetes.

### **Proposed System**

The proposed system in diabetics prediction using machine learning will involve gathering and analyzing data from various sources to identify patterns and trends. The data will include patient information such as age, gender, BMI, blood pressure, glucose levels, etc.

The machine learning algorithm will then be used to build a predictive model from the data. The model will be trained using historical data and then used to make predictions about future cases. The system will then be used to provide an early warning for at-risk patients and provide doctors with the information needed to make appropriate decisions.

The proposed system will also involve monitoring and tracking patient data in real-time, so that changes in health can be detected and alerted to health care providers. This will enable early intervention and more effective treatment of diabetes. Finally, the system will be used to provide personalized treatment plans and advice tailored to each patient's individual needs. This will help to ensure that the most effective treatments are being used and improve the overall prognosis for those with diabetes.

## Random forest

```
from sklearn import metrics

predictions = rfc.predict(X_test)
print("Accuracy_Score =", format(metrics.accuracy_score(y_test, predictions)))
```

```
Accuracy_Score = 0.7637795275590551
```

## Decision Tree

```
from sklearn import metrics

predictions = dtree.predict(X_test)
print("Accuracy Score =", format(metrics.accuracy_score(y_test,predictions)))
```

```
Accuracy Score = 0.7047244094488189
```

## Xg Boost classifier

```
from sklearn import metrics

xgb_pred = xgb_model.predict(X_test)
print("Accuracy Score =", format(metrics.accuracy_score(y_test, xgb_pred)))
```

```
Accuracy Score = 0.7795275590551181
```

## SVM

```
from sklearn import metrics  
  
print("Accuracy Score =", format(metrics.accuracy_score(y_test, svc_pred)))
```

Accuracy Score = 0.7480314960629921

## Conclusion

The conclusion of the machine learning approach to predicting diabetes is that it can be a very effective tool for identifying and predicting diabetes. By using the right algorithms and the right data, it is possible to create models that can accurately predict the likelihood of an individual having diabetes. However, it is important to remember that machine learning does not replace traditional medical diagnosis and should not be used as a replacement for medical advice.

## References

1. <https://ieeexplore.ieee.org/document/9588925/figures#figures>
2. <https://ieeexplore.ieee.org/document/8819841/citations#citations>
3. <https://ieeexplore.ieee.org/document/8679365>
4. <https://ieeexplore.ieee.org/document/8679365>
5. <https://ieeexplore.ieee.org/document/8819841>

Thank You