INTRODUCTION

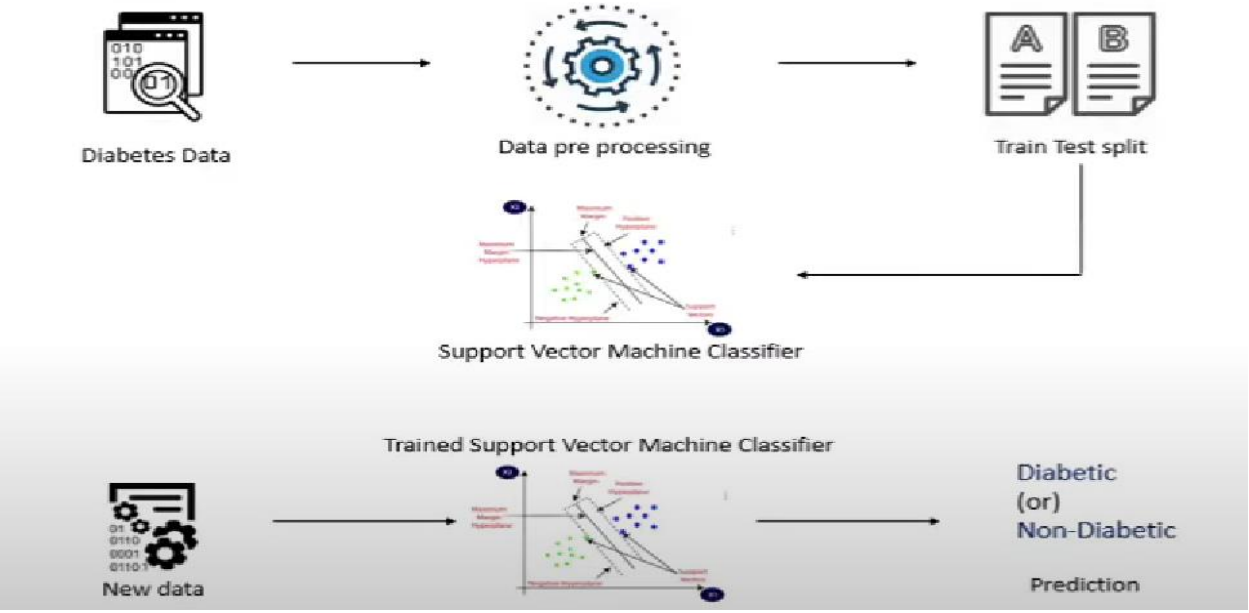
Diabetes is a long-lasting disease that happens when the pancreas fails to create enough insulin, or when the body cannot use the insulin produced efficiently. Insulin is a hormone that controls the level of sugar in the blood.

- According to the International Diabetes Federation (IDF), as of 2021, there were an estimated 303 million people with diabetes worldwide among them 153 million are women.
- Hence a good machine learning model is needed to predict whether a person is diabetic or not subject to his health data.
- Support vector machine is one of the important supervised machine learning classifier algorithm.
- SVM is effective for high-dimensional spaces, making It suitable for data with many feature.
- The SVM algorithm finds the hyperplane that separates the different classes in the data by maximizing the margin between them.

PROBLEM STATEMENT

Diabetes prediction using Support Vector Machine classifier

BLOCK DIAGRAM



METHODOLOGY

- In supervised learning model we feed data to the machine learning model and the model learns from the data and its respective label.
- We train our model with several medical information such as

    1) Glucose level
    2) Insulin
    3) BMI
    4) Skin thickness
    5) Age
    6) Diabetes pedigree function

- The SVM analyzes our model based on the feature as mentioned above and classifies them as diabetic or non-diabetes and the draws the hyperplane.
- Initially we pre-process the data to standardize it.
- We divide the data into training and testing data and we also find the accuracy score of both.
- After we split the data we will feed it to the Support Vector Machine classifier.
- The trained SVM classifier will classify whether the patient Is diabetic or not based on the inputs given.

RESULT

INPUT DATA 1

```
input_data = (5,166,72,19,175,25.8,0.587,51)

# changing the input_data to numpy array
input_data_as_numpy_array = np.asarray(input_data)

# reshape the array as we are predicting for one instance
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

# standardize the input data
std_data = scaler.transform(input_data_reshaped)


prediction = classifier.predict(std_data)
print(prediction)

if (prediction[0] == 0):
  print('The person is not diabetic')
else:
  print('The person is diabetic')
```

```
[1]
The person is diabetic
```

INPUT DATA 2

```
input_data = (1,  85, 66, 29, 0 ,26.6,  0.351,  31)

# changing the input_data to numpy array
input_data_as_numpy_array = np.asarray(input_data)

# reshape the array as we are predicting for one instance
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

# standardize the input data
std_data = scaler.transform(input_data_reshaped)


prediction = classifier.predict(std_data)
print(prediction)

if (prediction[0] == 0):
  print('The person is not diabetic')
else:
  print('The person is diabetic')
```

```
[0]
The person is not diabetic
```

CONCLUSION

- The SVM algorithm is a powerful machine learning technique that has shown excellent performance in classification tasks.
- SVM-based diabetes prediction models have shown high accuracy and specificity in identifying patients at risk of developing diabetes
- Such models can help healthcare professionals to identify at-risk individuals early on and take preventative measures to mitigate the onset of diabetes
- more research and data collection are needed to improve the accuracy and robustness of the model.