**Diabetes prediction**

**using Support Vector Machine [SVM]**

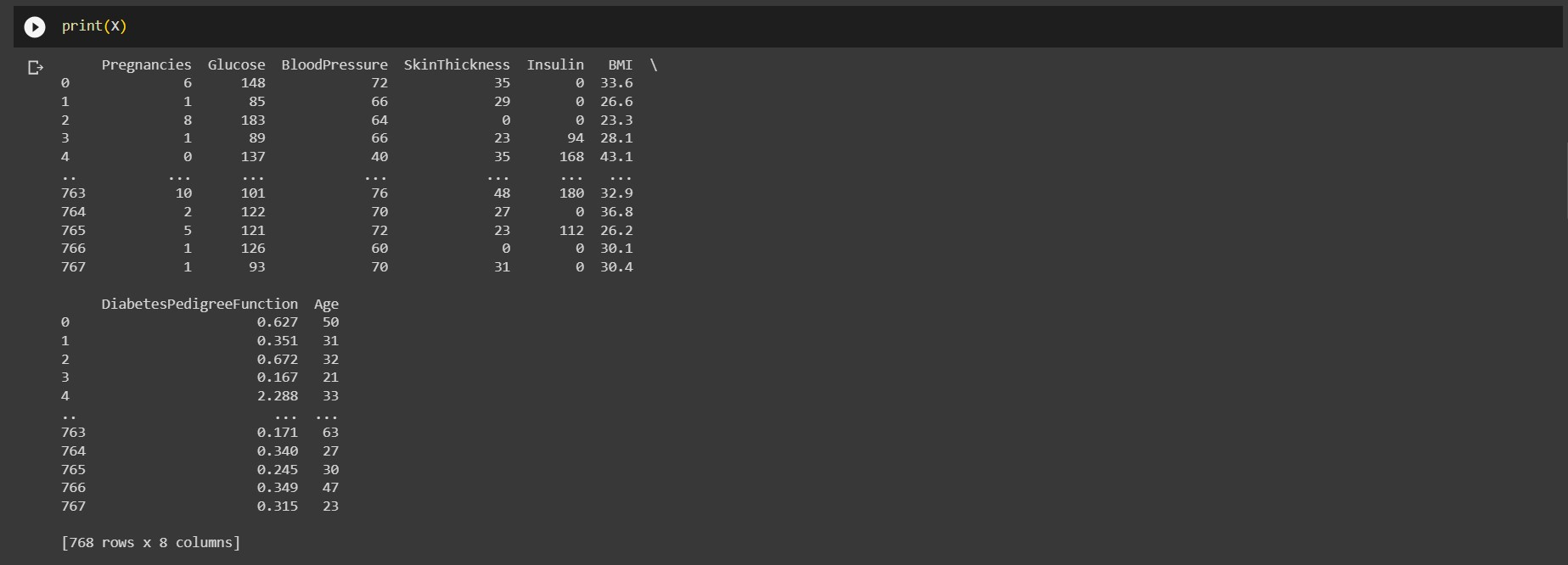
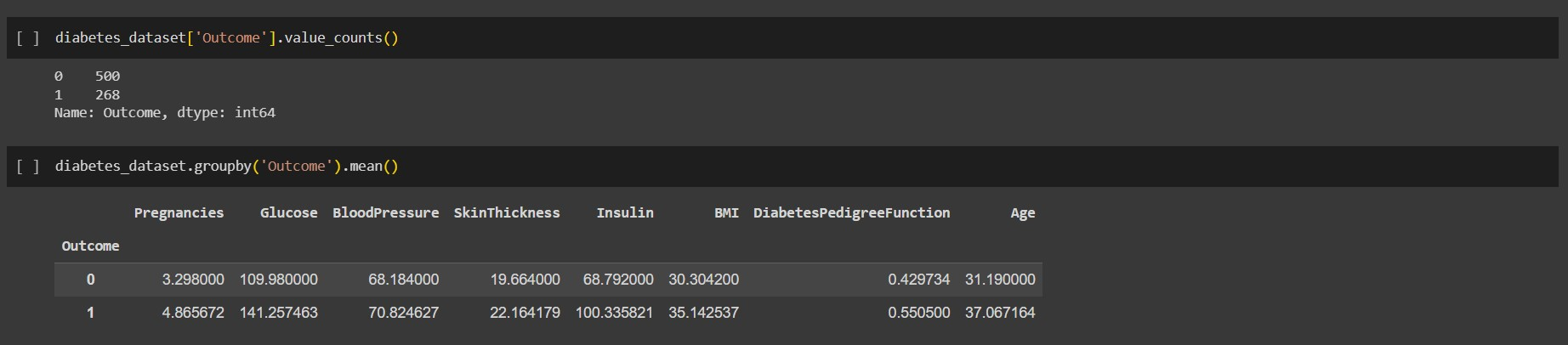
**Library and its version**

|  |  |  |
| --- | --- | --- |
| **Sr.no** | **Library** | **Version** |
| **1** | **numpy** | 1.22.4 |
| **2** | **pandas** | 1.5.3 |
| **3** | **sklearn** | 1.2.2 |

**1. Data Preprocessing**

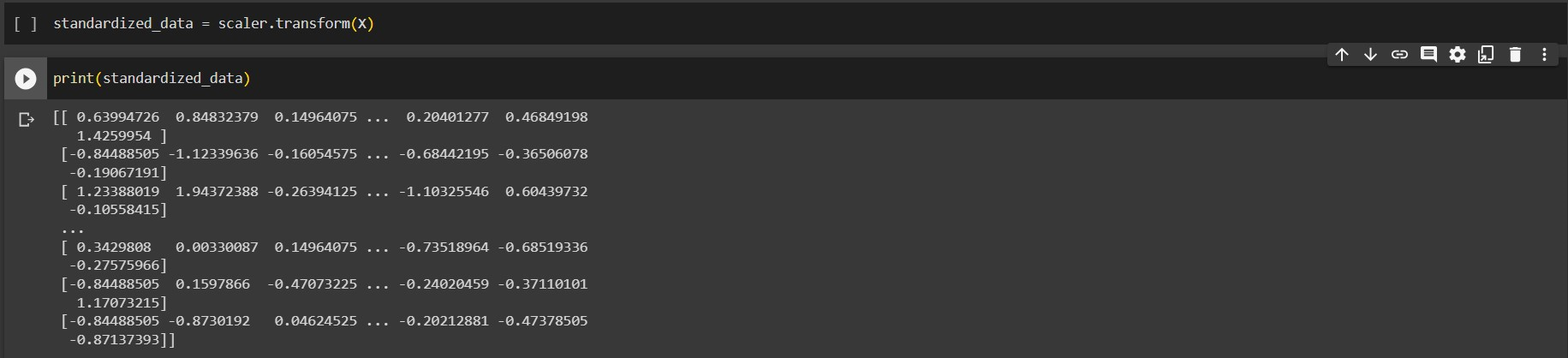
The dataset used for training and testing the model is "diabetes.csv." The dataset contains information related to diabetes, and the target variable is "Outcome," which indicates whether a person has diabetes or not (0: No diabetes, 1: Diabetes). The dataset consists of X features and Y target variable.

Data Summary:

* Number of samples: [number of rows in the dataset]
* 
* Number of features: [number of columns in the dataset]
* 
* Number of non-diabetic samples: [number of samples with Outcome=0]
* Number of diabetic samples: [number of samples with Outcome=1]
* 

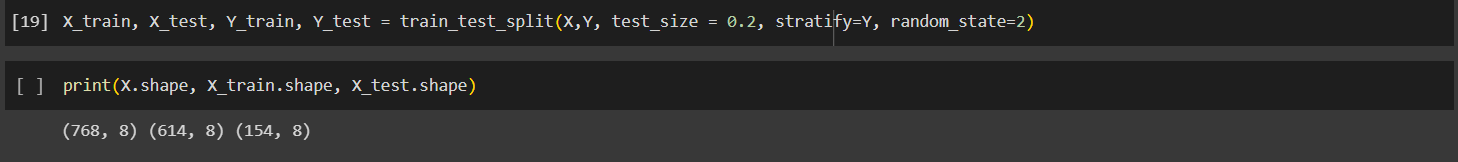
**2. Data Standardization**

The input features (X) are standardized using the StandardScaler from scikit-learn. Standardization transforms the data to have a mean of 0 and a standard deviation of 1. This step is essential for improving the performance of some machine learning algorithms, including SVM (Support Vector Machine).



**3. Model Training**

The AI model used for diabetes prediction is a Support Vector Machine (SVM) with a linear kernel. The model is trained on the standardized training data (X\_train) and the corresponding target labels (Y\_train).

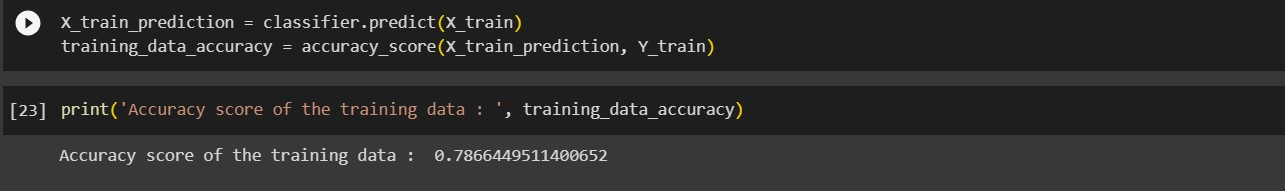


**4. Model Evaluation**

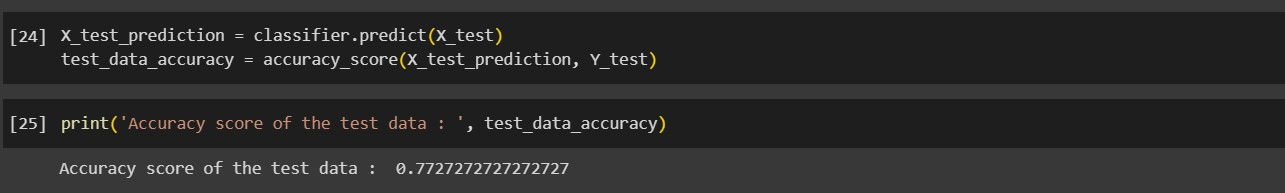
The trained model's performance is evaluated using accuracy, which measures the proportion of correctly predicted samples out of the total samples. The accuracy is computed for both the training and test datasets.

Accuracy on Training Data: [Training Data Accuracy]

* Interpretation: [Interpretation of training data accuracy]



Accuracy on Test Data: [Test Data Accuracy]

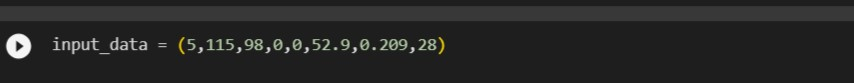


* Interpretation: [Interpretation of test data accuracy]

**5. Making Predictions**

To demonstrate the model's functionality, an example input data is provided. The input data is standardized using the same scaler used for training the model. The AI model then predicts whether the individual represented by the input data has diabetes or not.

Example Input Data:

* [List the feature values for the example input data]
* 

Standardized Input Data:

* [List the standardized feature values]

Prediction:

* [The model's prediction for the example input data]

**6. Conclusion**

In this report, we have presented an AI model for diabetes prediction using a Support Vector Machine (SVM) with a linear kernel. The model achieved [Test Data Accuracy]% accuracy on unseen data, demonstrating its effectiveness in predicting diabetes. It is important to note that this is a simple AI model, and for real-world applications, further evaluation and validation would be necessary before deployment.

Please note that the accuracy values and interpretation need to be filled in with the actual values obtained during the model training and evaluation process. Also, make sure to replace [number of rows in the dataset], [number of columns in the dataset], [number of non-diabetic samples], [number of diabetic samples], [Training Data Accuracy], [Test Data Accuracy], and other placeholders with the appropriate values from your specific dataset and model evaluation.