**Charotar University of Science and Technology [CHARUSAT]**

**Chandubhai S. Patel Institute of Technology [CSPIT]**

**U & P U. Patel Department of Computer Engineering**

**Lab Manual**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Subject code | : | CE-350 | Semester | : | 6 | Academic Year | : | 2021 |
| Subject name | : | Data Warehouse and Data Mining | | | | | | |

**Practical – 7**

**Aim:**

**Performance Evaluation Matrix [Confusion Matrix] Understanding and Practical Implementation in Google colab.**

* **Prerequisite**

S/W: - Anaconda, Jupyter Notebook

Little bit knowledge about the Jupyter Notebook and how it works on Data Mining.

* **Step to install Anaconada Tool:**

1: [Download the Anaconda installer](https://www.anaconda.com/download/#windows).



2: RECOMMENDED: [Verify data integrity with SHA-256](https://docs.anaconda.com/anaconda/install/hashes/). For more information on hashes, see [what about cryptographic hash verification?](https://conda.io/projects/conda/en/latest/user-guide/install/download.html#cryptographic-hash-verification)

3: Double click the installer to launch.

**Note**

To prevent permission errors, do not launch the installer from the [Favorites folder](https://docs.anaconda.com/anaconda/user-guide/troubleshooting/#distro-troubleshooting-favorites-folder).

**Note**

If you encounter issues during installation, temporarily disable your anti-virus software during install, then re-enable it after the installation concludes. If you installed for all users, uninstall Anaconda and re-install it for your user only and try again.

4: Click Next.

5: Read the licensing terms and click “I Agree”.

6: Select an install for “Just Me” unless you’re installing for all users (which requires Windows Administrator privileges) and click Next.

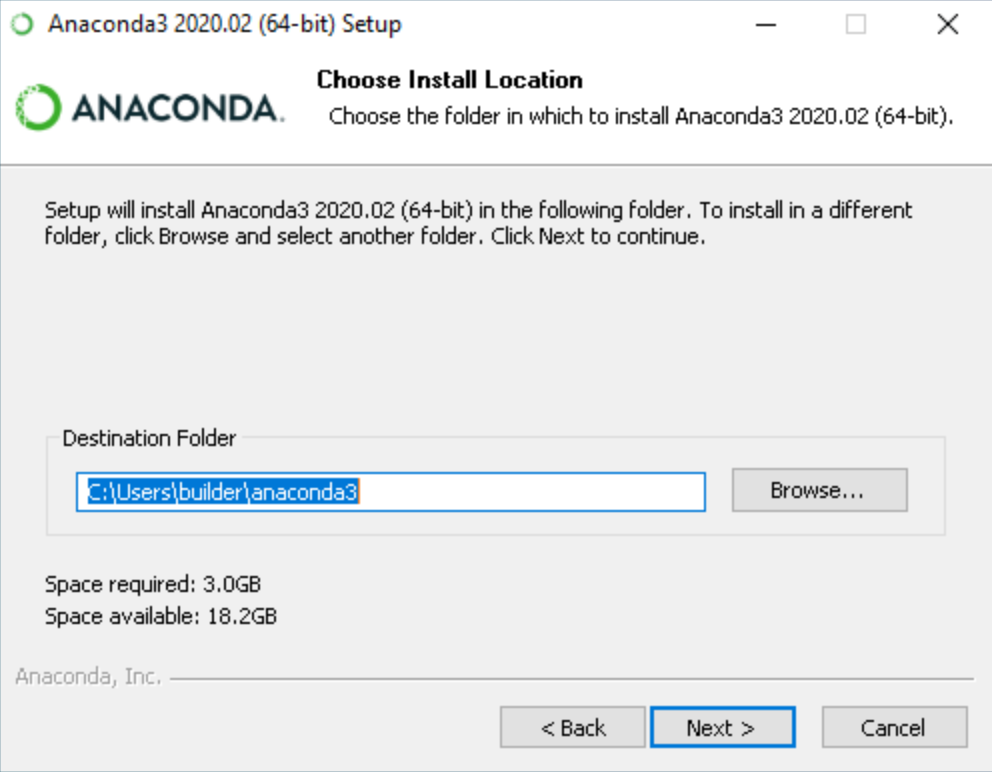
7: Select a destination folder to install Anaconda and click the Next button. See [FAQ](https://docs.anaconda.com/anaconda/user-guide/faq/#distribution-faq-windows-folder).

**Note**

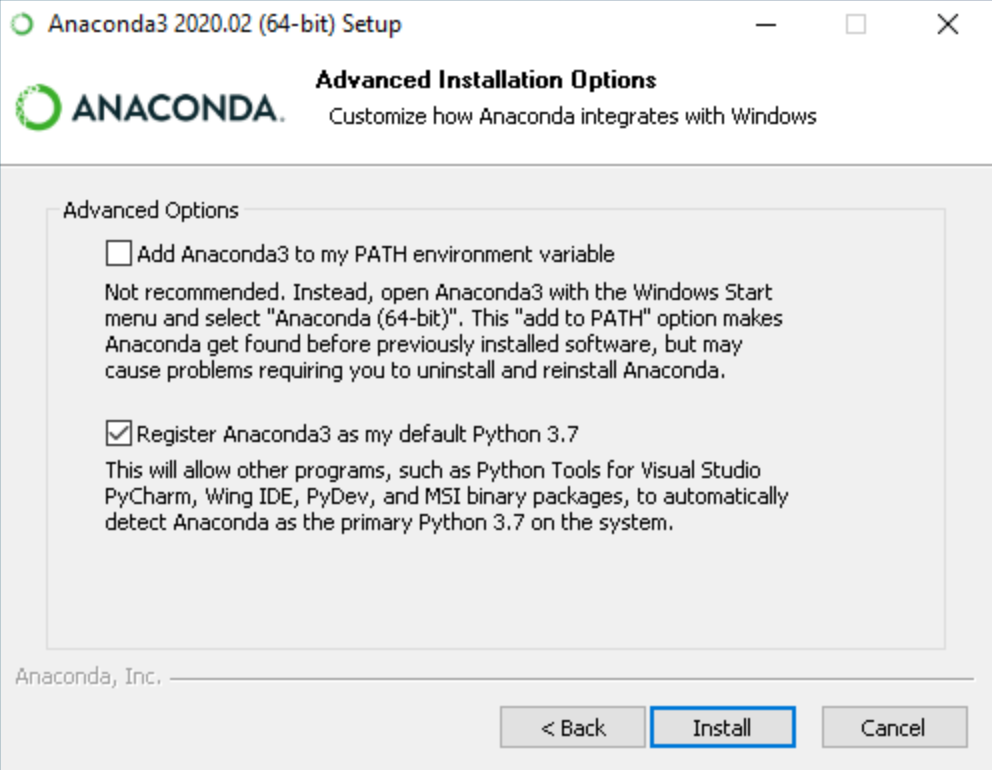
Install Anaconda to a directory path that does not contain spaces or unicode characters.

**Note**

Do not install as Administrator unless admin privileges are required.

[](https://docs.anaconda.com/_images/win-install-destination.png)

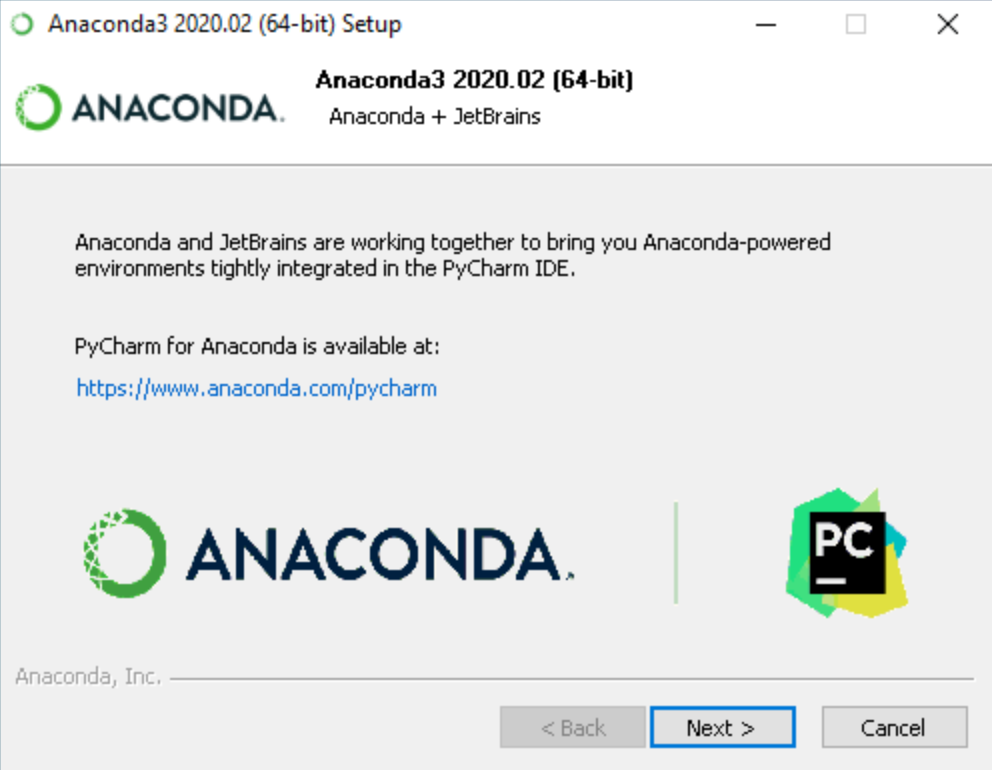
8: Choose whether to add Anaconda to your PATH environment variable. We recommend not adding Anaconda to the PATH environment variable, since this can interfere with other software. Instead, use Anaconda software by opening Anaconda Navigator or the Anaconda Prompt from the Start Menu.

[](https://docs.anaconda.com/_images/win-install-options.png)

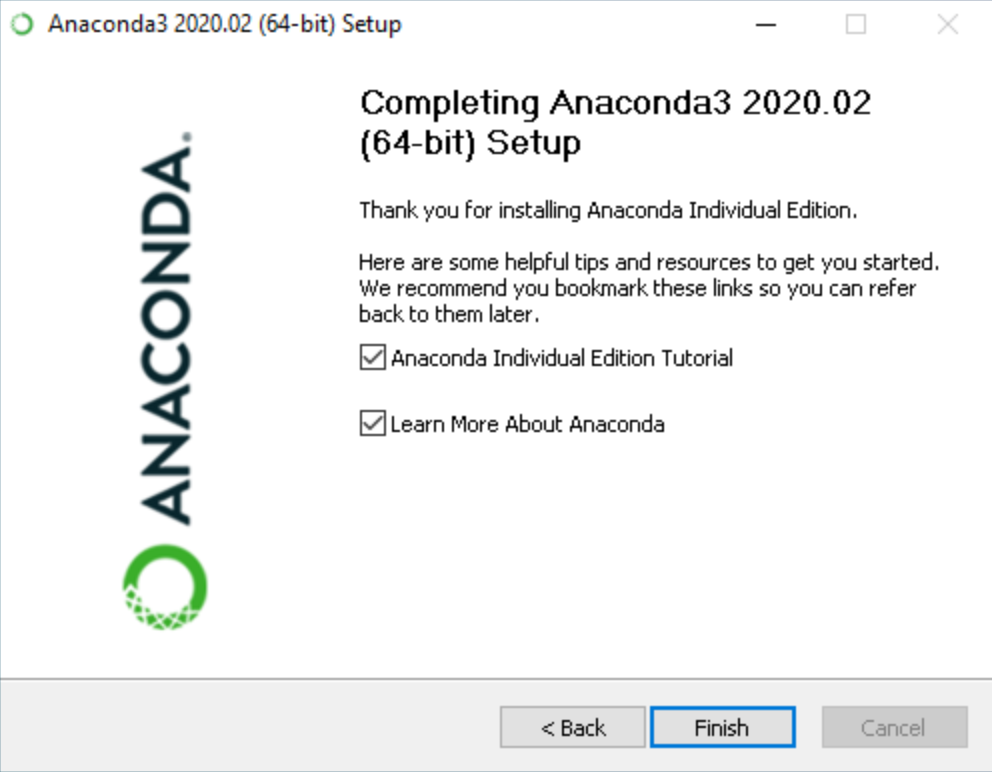
9: Choose whether to register Anaconda as your default Python. Unless you plan on installing and running multiple versions of Anaconda or multiple versions of Python, accept the default and leave this box checked.

10: Click the Install button. If you want to watch the packages Anaconda is installing, click Show Details.

11: Click the Next button.

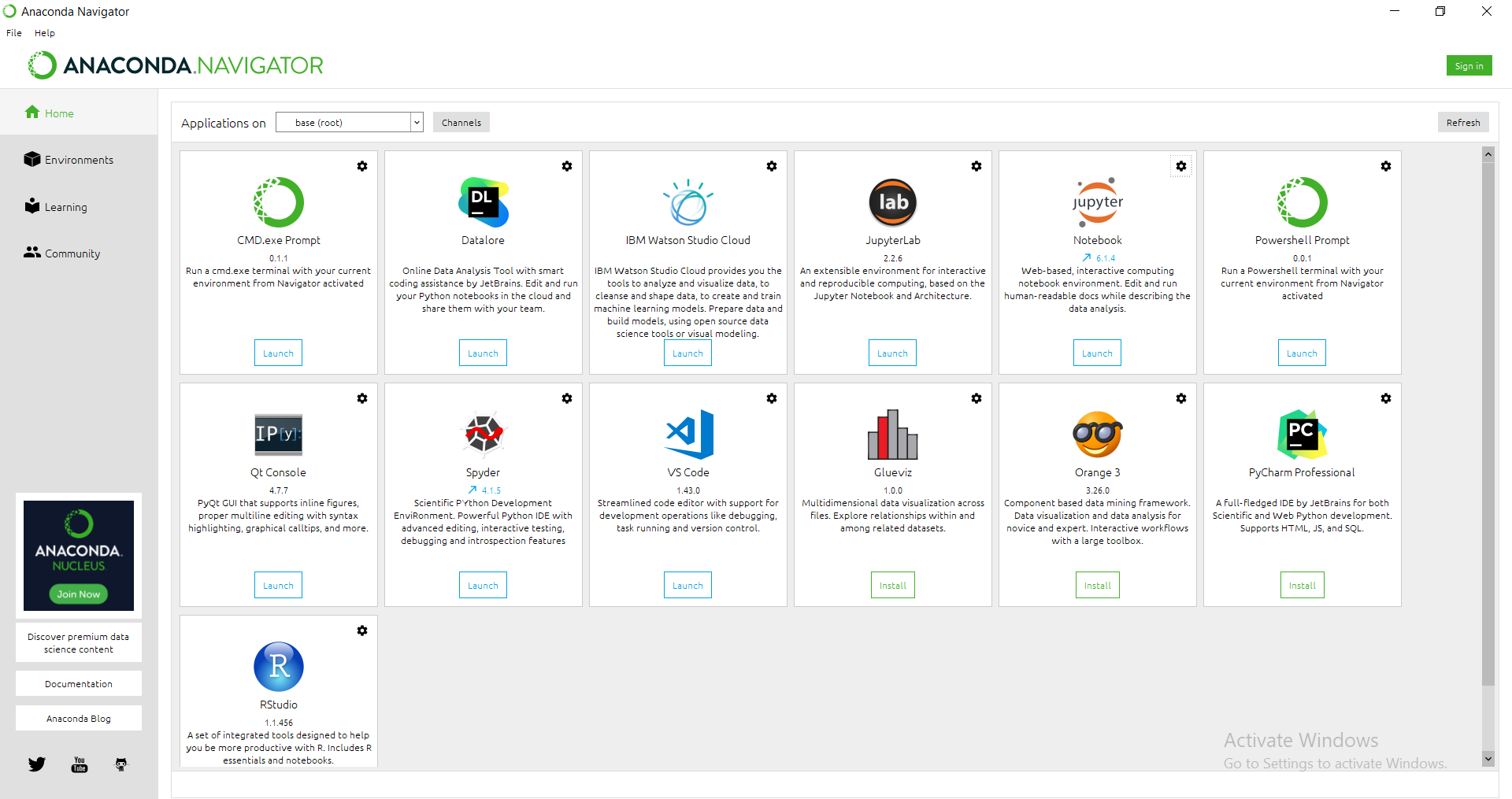
[](https://docs.anaconda.com/_images/win-install-pycharm.png)

12: After a successful installation you will see the “Thanks for installing Anaconda” dialog box:

[](https://docs.anaconda.com/_images/win-install-complete.png)

13: Click the Finish button.

14: [Verify your installation](https://docs.anaconda.com/anaconda/install/verify-install/).



* **Algorithm / process flow:**

1. Import pandas and linear\_model from sklearn

2. Convert csv file into dataframe

3. Handling Missing Values

a. Identify NaN in columns

b. Replace missing values with fillna()

4. Drop index column (Serial No.)

5. Dimensional Reduction

a. Prepare test\_dataset

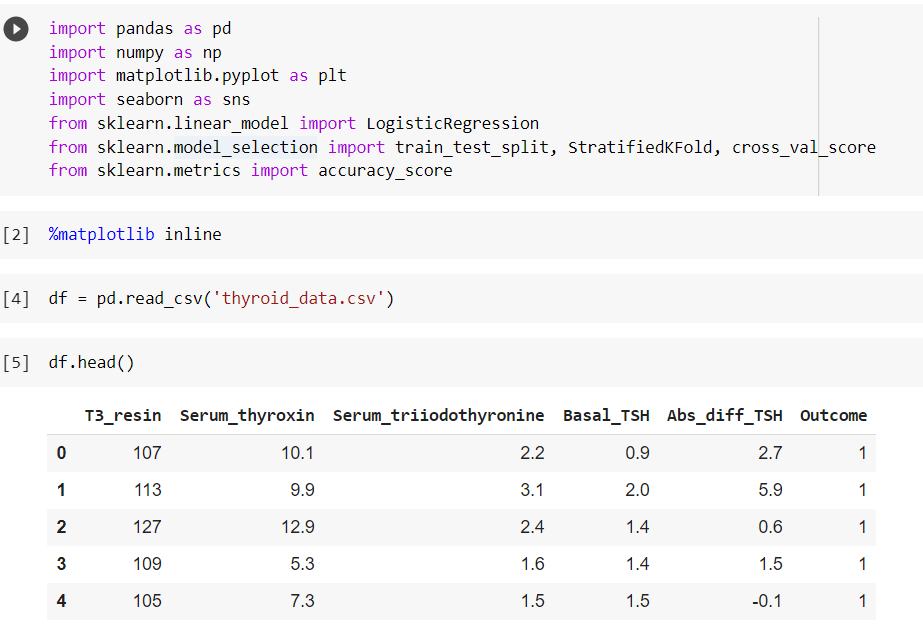
b. Select train\_dataset

6. Apply linear regression and prepare model by finding best fit line

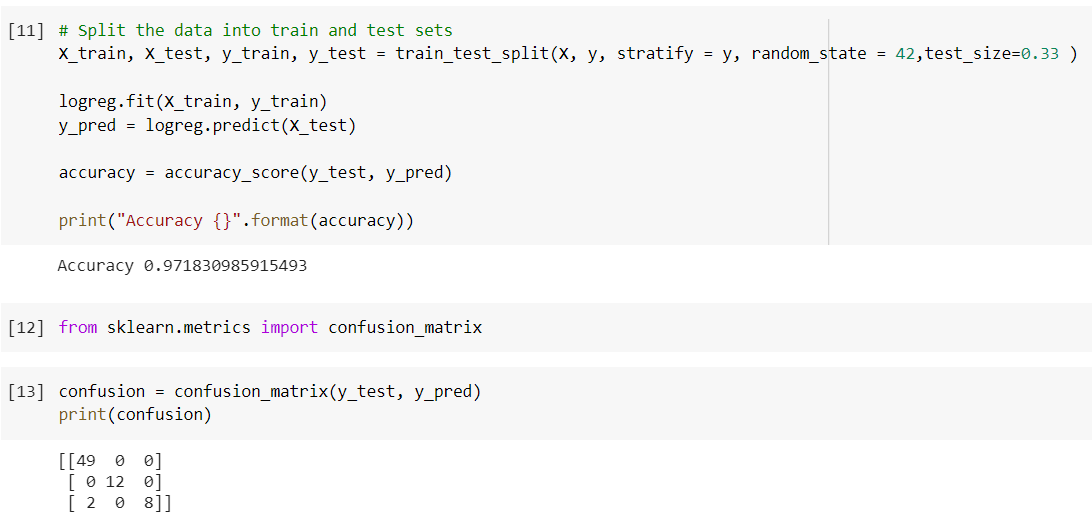
7. Predict ‘Chance of Admit’ on the basis of model and given input values

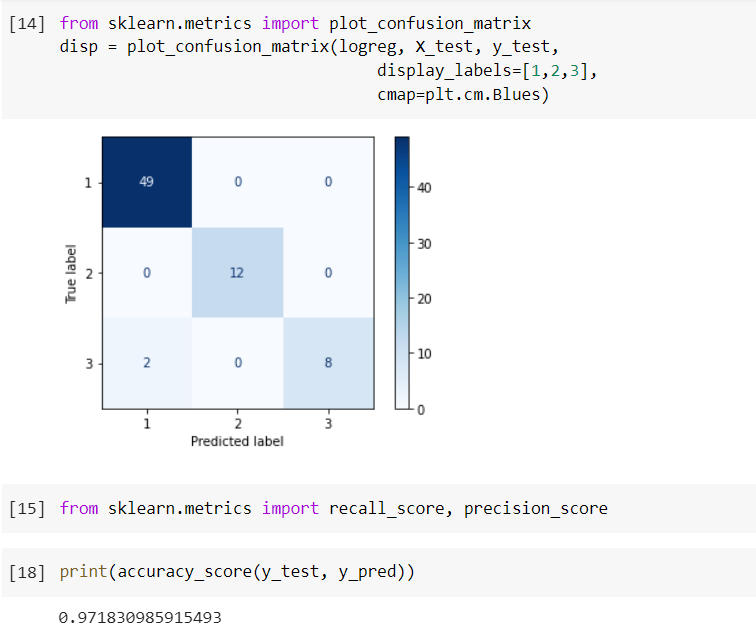
8. Print the predicted value

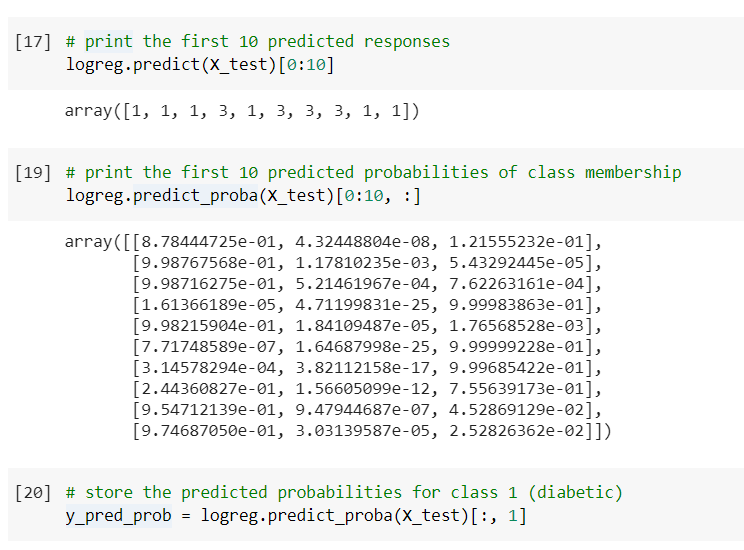
**Screenshot/Output**



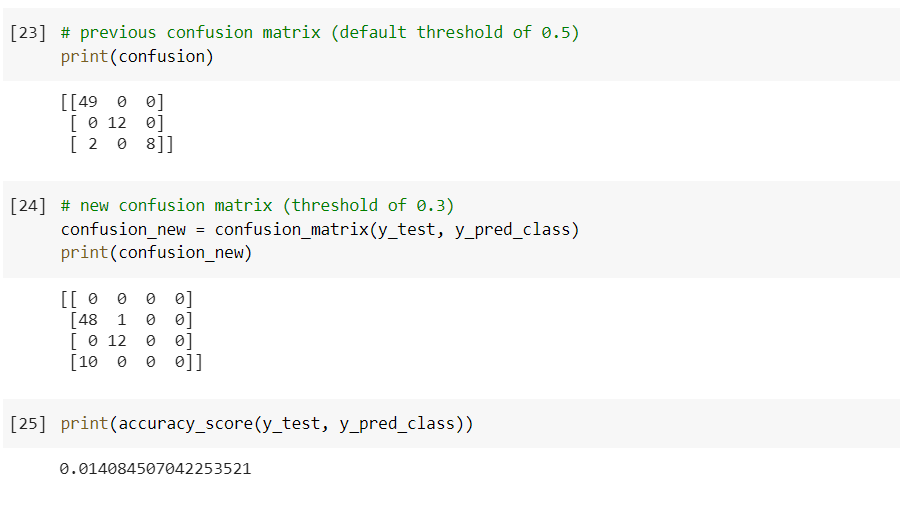












* **References:**
  + [http://education.abcom.com/admit-prediction/](http://education.abcom.com/admit-prediction/%20)
  + <https://www.youtube.com/watch?v=JrIo06yYXyg&t=1s>
* **Question & Answers:**
* **What is confusion matrix?**

A **confusion matrix** is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.

* **How do you solve a confusion matrix?**

Step 1) First, you need to test dataset with its expected outcome values.

Step 2) Predict all the rows in the test dataset.

Step 3) Calculate the expected predictions and outcomes:

* **Why do we need confusion matrix?**

**Confusion matrices** are used to visualize important predictive analytics like recall, specificity, accuracy, and precision. **Confusion matrices** are useful because they give direct comparisons of values like True Positives, False Positives, True Negatives and False Negatives.

* **How do I import confusion matrix?**
* # Importing the dependancies
* from sklearn import metrics
* # Predicted values
* y\_pred = ["a", "b", "c", "a", "b"]
* # Actual values
* y\_act = ["a", "b", "c", "c", "a"]
* # Printing the confusion matrix
* # The columns will show the instances predicted for each label,
* # and the rows will show the actual number of instances for each label.
* print(metrics.confusion\_matrix(y\_act, y\_pred, labels=["a", "b", "c"]))
* # Printing the precision and recall, among other metrics
* print(metrics.classification\_report(y\_act, y\_pred, labels=["a",
* "b","c"]))
* **What is F1 score in confusion matrix?**

**F1 Score** becomes 1 only when precision and recall are both 1. **F1 score** becomes high only when both precision and recall are high. **F1 score** is the harmonic mean of precision and recall and is a better measure than accuracy.

* **Conclusion :**

In this Practical, we learned about what is evaluation matrix (confusion matrix) and how to implement it in google colab.

|  |  |  |  |
| --- | --- | --- | --- |
| **Prepared By:** | Rajiv Kumar Gupta (18CE137) | **Date:** | 04-14-2021 |