|  |  |
| --- | --- |
| Name Of The Student | Himanshu |
| Internship Project Topic | TCS iON RIO-210: Build a Classification Model for Drug Trials Dataset |
| Name of the Organization | TCS iON |
| Name of the Industry Mentor | Himdweep Walia |
| Name of the Institute | Amity University |

|  |  |  |
| --- | --- | --- |
| Date | Day # | Hours Spent |
| 30-05-2024 | Day-37 | 5 Hours |
| Activities done during the day:  **Project Hands-on - Import related modules and dataset file**  **Link of the google drive google Colab file :-**  <https://colab.research.google.com/drive/1VQRq0l6oc9Uj4cOOqiuhkfS1JmpKr3fU?usp=sharing>  Upload the dataset in google colab.    **Import required libraries in python:**   |  | | --- | | **import pandas as pd** |   pandas (all lowercase) is a popular Python-based data analysis toolkit which can be imported using import pandas as pd.  It presents a diverse range of utilities, ranging from parsing multiple file formats to converting an entire data table into a NumPy matrix array.  This makes pandas a trusted ally in data science and machine learning.  **Import** = “Bring this functionality or library to my python script”  **Pandas** = The library you want to import, in this case, it’s pandas  **As** = The python nomenclature for creating as alias.  **pd** = The standard short name for referencing pandas   |  | | --- | | **import numpy as np** |   NumPy, which stands for Numerical Python, is a scientific computing library built on top of the Python programming language.  The import numpy portion of the code tells Python to bring the NumPy library into your current environment.   |  | | --- | | **import seaborn as sns**  **import matplotlib.pyplot as plt** |   Seaborn & Matplotlib the both python libraries is used for data visualization and viewing the graphs   |  | | --- | | **from sklearn.svm import SVC** |   Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outliers detection.  It is C-support vector classification whose implementation is based on svm. The module used by scikit-learn is sklearn.svm.SVC. This class handles the multiclass support according to one-vs-one scheme.   |  | | --- | | **from sklearn.ensemble import RandomForestClassifier** |   A random forest is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and use averaging to improve the predictive accuracy and control over-fitting.   |  | | --- | | **from sklearn.model\_selection import train\_test\_split** |   The train\_test\_split function of the sklearn.model\_selection package in Python splits arrays or matrices into random subsets for train and test data, respectively.  To use the train\_test\_split function, we’ll import it into our program   |  | | --- | | **from sklearn.metrics import classification\_report** |   A classification report is a performance evaluation metric in machine learning. It is used to show the precision, recall, F1 Score, and support of your trained classification model.  **Code:-**    **Output:-** | | |
|  | | |