|  |  |
| --- | --- |
| 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 |
| 25-04-2024 | Day-6 | 3.5 Hours |
| Activities done during the day:  Learned about the abstract of the project in detail   * How to gather data for creating dataset? * Decide what data you want to gather * What is Data Preprocessing? * Why is Data Preprocessing important? * 4 Steps in Data Preprocessing   Data Integration  Data Cleaning  Data Transformation  Data Reduction or Dimension Reduction   * Data pre-processing of structured data in machine learning * What are Data and Attributes? * How to analysis data for better understanding     Take available drugs trial dataset from below link after researching many webpages:  <https://archive.ics.uci.edu/dataset/461/drug+review+dataset+druglib+com>    Links:-  <https://www.techtarget.com/searchdatamanagement/definition/data-preprocessing>  <https://www.analyticsvidhya.com/blog/2021/08/data-preprocessing-in-data-mining-a-hands-on-guide/>  <https://towardsdatascience.com/introduction-to-data-preprocessing-in-machine-learning-a9fa83a5dc9d>  <https://www.youtube.com/watch?v=ITy8R4278sk>  <https://www.youtube.com/watch?v=FNLLxYcUnow> | | |
|  | | |