|  |  |
| --- | --- |
| Name Of The Student | Muhammed Nabeel CP |
| Internship Project Topic | RIO – 25 Classification Model – Build a Model that Classifies the Side Effects of Drug |
| Name of the Organization | TCS- iON |
| Name of the Industry Mentor | Himalaya Aashish |
| Name of the Institute | ICTAK |

|  |  |  |
| --- | --- | --- |
| Date | Day # | Hours Spent |
| 01/03/2021 | 10 | 4.5 |
| Activities done during the day:   * Checked for outliers in the dataset by plotting box plots. * An outlier is a data point that lies outside the overall pattern in a distribution. * Outliers are detected in the column ‘Ease of Use’ and ‘Useful Count’. * Indexed the outliers with IQR Method. * IQR is the range between the first and the third quartiles namely Q1 and Q3: IQR = Q3 – Q1. The data points which fall below Q1 – 1.5 IQR or above Q3 + 1.5 IQR are outliers. * Thus, the indexed outliers are removed from the dataset. * Changed the index after removing outliers. | | |