**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | LTVIP2024TMID24892 |
| Project Title | Liver Patient Identification – prediction of liver patient |
| Maximum Marks | 6 Marks |

**Data Exploration and Preprocessing :**

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python

employed for preprocessing tasks like normalization and feature engineering. Data cleaning will

address missing values and outliers, ensuring quality for subsequent analysis and modeling, and

forming a strong foundation for insights and predictions.

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | Dimension:  584rows X 12columns  Descriptive statistics: |
| Univariate Analysis |  |
| Bivariate Analysis |  |
| Multivariate Analysis |  |
| Outliers and Anomalies |  |
| **Data Preprocessing Code Screenshots** | |
| Loading Data |  |
| Handling Missing Data |  |
| Data Transformation | There is no need of Standardization and Normalization of our dataset, as we using Ensemble Technique. |
| Feature Engineering |  |
| Save Processed Data | - |