Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 03 October 2024 |
| Team ID | LTVIP2024TMID24947 |
| Project Title | SmartLender - Applicant Credibility Prediction for Loan Approval |
| Maximum Marks | 6 Marks |

**Data Exploration and Preprocessing Report**

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: 4269 rows × 12 columns  Descriptive statistics: |
| Univariate Analysis |  |

|  |  |
| --- | --- |
|  |  |
| Bivariate Analysis |  |
| Multivariate Analysis |  |

|  |  |
| --- | --- |
| Outliers and Anomalies | - |
| **Data Preprocessing Code Screenshots** | |
| Loading Data |  |
| Data Transformation |  |
| Balancing the data |  |
| Feature Engineering | Attached the codes in final submission. |
| Save Processed Data | - |