|  |  |  |  |
| --- | --- | --- | --- |
| **Sr.No.** | **Practicals** | **Date** | **Sign** |
| **1** | **Data Pre-processing and Exploration** |  |  |
| a. | Load a CSV dataset. Handle missing values, nconsistent  formatting, and  outliers. |  |  |
| b. | Load a dataset, calculate descriptive summary statistics, create  visualizations using different graphs, and identify potential  features and target variables **Note**: Explore Univariate and  Bivariate graphs (Matplotlib) and Seaborn for visualization. |  |  |
| c | Create or Explore datasets to use all pre-processing routines like  label encoding, scaling, and binarization.  . |  |  |
| **2** | **Testing Hypothesis** |  |  |
| a | Implement and demonstrate the FIND-S algorithm for finding the  most specific hypothesis based on a given set of training data  samples. Read the training data from a. CSV file and generate the  final specific hypothesis. (Create your dataset) |  |  |
| **3** | **Linear Models** |  |  |
| a | **Simple Linear Regression**  Fit a linear regression model on a dataset. Interpret coefficients,  make  predictions, and evaluate performance using metrics like R squared and MSE. |  |  |
| b | **Multiple Linear Regression**  Extend linear regression to multiple features. Handle feature  selection and potential multicollinearity. |  |  |
| c | **Regularized Linear Models** (Ridge, Lasso, ElasticNet)  Implement regression variants like LASSO and Ridge on any generated  dataset. |  |  |
| **4** | **Discriminative Models** |  |  |
| a | **Logistic Regression**  Perform binary classification using logistic regression. Calculate  accuracy,  precision, recall, and understand the ROC curve**.** |  |  |
| b | Implement and demonstrate k-nearest Neighbor algorithm. Read  the training data from a .CSV file and build the model to classify a  testsample. Print both correct and wrong predictions. |  |  |
| c | Build a decision tree classifier or regressor. Control  hyperparameters like tree  depth to avoid overfitting. Visualize the tree. |  |  |
| d | Implement a Support Vector Machine for any relevant dataset. |  |  |
| e | Implement a gradient boosting machine (e.g., XGBoost). Tune  hyperparameters and explore feature importance. |  |  |
| **5** | **Probabilistic Models** |  |  |
| a | Implement Bayesian Linear Regression to explore prior and  posterior  Distribution. |  |  |
| b | Implement Gaussian Mixture Models for density estimation and  unsupervised clustering. |  |  |
| **6** | **Bayesian Learning** |  |  |
| a | Implement Bayesian Learning using inferences |  |  |