Interpretation Analysis of Lab 1:

This code is used to build a logistic regression model that predicts a student's final outcome based on various characteristics.

The analysis of each step is as follows.

1. Loading libraries: - The code loads libraries such as ``caret'' and ``dplyr'' that are required for data manipulation and modeling.

2.Reading data: - Data is read from her CSV file named "oulad-students.csv" and saved in variable "data".

3. Partial data: - Perhaps predict the final result.

4. Variable description: - Summary statistics for selected variables are displayed using summary() to provide an overview of the distribution of the data.

5.Displaying data: - The first few rows of the subset data are displayed using head() to check the structure and values.

6.Missing value removal: - Rows with missing values ​​are removed using na.omit() to ensure the dataset is clean before modeling.

7.Transformation of categorical variables: - The final\_result variable representing the final result is transformed into factors to facilitate modeling and evaluation.

– Other categorical variables must also be converted to factors if they are not already converted.

8.Split the data: - The dataset is split into a training set and a test set (80/20 ratio) using createDataPartition() at caret.

This prepares the data for model training and model evaluation.

9.Training the model: – A logistic regression model ('glm()') is trained using 'final\_result' as the response variable and all other variables as predictors.

10. Making predictions: - Predictions are made on the test data using the trained logistic regression model.

The “type = “response”” argument ensures that predicted probabilities are returned.

11.Predicted class conversion: - To ensure consistency of evaluation, predicted classes are converted to factors at the same level as the final\_result variable in the test data.

12. Model evaluation: - Model performance is evaluated using ``confusionMatrix()'' in ``caret'', which calculates various metrics such as precision, precision, recall, F1 score, etc.

based on the predicted confusion matrix.

will be evaluated using.

Comparison with actual class.

Overall, this code demonstrates a typical workflow for building and evaluating a logistic regression model that predicts student outcomes based on selected characteristics.

Handles data preprocessing, model training, prediction, and evaluation in a structured way.