

Practical-6

Practical Problem (Statement)

1. Import a dataset from a web storage location (CSV). Name the dataset df. The dataset should contain at least the following variables:
 - o admit — binary outcome (1 = admitted, 0 = not admitted)
 - o gre — GRE score (numeric)
 - o gpa — undergraduate GPA (numeric)
 - o rank — prestige/rank of undergraduate institution (categorical: 1,2,3,4)
2. Perform logistic regression to model the probability of admission (admit) using gre, gpa, and rank as predictors.
3. Report and interpret:
 - o Model coefficients and p-values (significance)
 - o Odds ratios and 95% confidence intervals
 - o Multicollinearity check (VIF)
4. Check model fit / predictive performance:
 - o Train/test split (70/30) and predict on test set
 - o Confusion matrix, accuracy, precision, recall, F1-score
 - o ROC curve and AUC
 - o Calibration plot and Hosmer–Lemeshow style test (grouped deciles) — comment on calibration
5. Produce clear comments/interpretation for each output.

Dataset:

admit,gre,gpa,rank

0,380,3.61,3

1,660,3.67,3

1,800,4.00,1

1,640,3.19,4

0,520,2.93,4

1,760,3.00,2

1,560,2.98,1

0,400,3.08,2

1,540,3.39,3

0,700,3.92,2

0,800,4.00,4

0,440,3.22,1

1,760,4.00,1

0,700,3.08,2

1,700,4.00,1

0,480,3.44,3

0,780,3.87,4

0,360,2.56,3

0,800,3.75,2

1,540,3.81,1

0,500,3.17,3

1,660,3.63,2

0,600,2.82,4

0,680,3.19,4

1,700,3.63,2

0,380,3.33,4

1,540,3.44,3

0,460,2.91,3

0,700,3.45,3

0,800,3.80,2

1,580,3.46,2

0,500,2.97,4

0,660,3.43,1

1,600,3.77,2

0,660,3.40,3

0,640,3.22,3

0,480,3.81,4

0,700,3.15,2

0,560,2.98,1

1,800,3.94,1

0,420,3.35,2

0,680,3.43,4

0,600,3.36,3

1,800,4.00,2

0,660,3.44,4

1,700,3.45,3

0,480,2.78,2

1,780,3.86,4

0,360,2.56,4

0,800,3.75,2

1,540,3.81,1

0,500,3.17,3

1,660,3.63,2

0,600,2.82,4

0,680,3.19,4

1,700,3.63,2

0,380,3.33,4

1,540,3.44,3

0,460,2.91,3

0,700,3.45,3

0,800,3.80,2

1,580,3.46,2

0,500,2.97,4

0,660,3.43,1

1,600,3.77,2

0,660,3.40,3

0,640,3.22,3

0,480,3.81,4

0,700,3.15,2

0,560,2.98,1

1,800,3.94,1

0,420,3.35,2

0,680,3.43,4

0,600,3.36,3

1,800,4.00,2