PERFORMANCE METRICS Part-2

LECTURE 17

DR. GAURAV DIXIT

DEPARTMENT OF MANAGEMENT STUDIES

PERFORMANCE METRICS

- Performance Metrics in presence of a class of interest

$$\text{sensitivity} = \frac{n_{1\,1}}{n_{1\,0}+n_{1\,1}} = \text{true positive fraction}$$

$$\text{specificity} = \frac{n_{0\,0}}{n_{0\,0}+n_{0\,1}} = \text{true negative fraction}$$

- ROC (receiver operating characteristic) curve
  - Used to plot {sensitivity, 1-specificity} points as the cutoff value increases
  - Top left corner points reflect wanted performance

PERFORMANCE METRICS

- Open Excel and RStudio

- Rank Ordering of records for class of interest
  - Based on estimated probabilities of class membership

- Lift curve is used to display the effectiveness of the model in rank ordering of cases
  - Constructed using validation partition scores

Key References

- Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data by EMC Education Services (2015)

- Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner by Shmueli, G., Patel, N. R., & Bruce, P. C. (2010)

Thanks…