IIT ROORKEE

NPTEL ONLINE CERTIFICATION COURSE

NAÏVE BAYES PART-4

LECTURE 34

DR. GAURAV DIXIT

DEPARTMENT OF MANAGEMENT STUDIES

NAÏVE BAYES

- Naïve Bayes formula
  - For classification, naïve Bayes formula works quite well
  - Since we don't require probabilities values to be accurate in absolute term, rather just a reasonably accurate rank ordering of these values
  - For the same reason, we should use the numerator only and drop the denominator which is common for all the classes
- Steps when we have a class of interest
  - User specified cut off value for the class of interest

NAÏVE BAYES

- Steps when we have a class of interest
  - Compute the probabilities (P_1, P_2, …, P_p) of belonging to class of interest for each predictor's value (x_1, x_2, …, x_p) taken by the new observation to be classified
  - Compute $P_1 \times P_2 \times \ldots \times P_p \times$ P(Class of interest)
  - Execute previous two steps for all the classes
  - To compute the probability of the new observation belonging to class of interest, divide the value computed in step 2 by the summation of values computed in step 2 for all the classes

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…