

You are developing an email spam detection system to classify incoming emails as 'Spam' or 'Not Spam' based on their content. You have a labeled dataset of emails and have implemented a Naive Bayes classifier with Laplace smoothing. Here's how to evaluate and visualize the model's performance:

1.Data Distribution: Display a bar chart or pie chart illustrating the proportion of 'Spam' versus 'Not Spam' emails in the dataset. This helps understand the class balance.

2. Confusion Matrix: After training the Naive Bayes model with Laplace smoothing, visualize its performance using a confusion matrix. A heatmap or stacked bar chart can show true positives, true negatives, false positives, and false negatives.

3. ROC Curve: Plot the ROC curve to illustrate the trade-off between the true positive rate (sensitivity) and false positive rate (1-specificity) across different thresholds. This helps in selecting an appropriate threshold for spam detection.

4.Precision-Recall Curve: Generate the precision-recall curve to assess the model's precision and recall at various thresholds. This curve is crucial for understanding how well the model performs with respect to both false positives and false negatives.

5. Feature Importance: Show the importance of different features (words or phrases) used by the Naive Bayes model in distinguishing spam from non-spam emails. A bar chart can illustrate the top N features.

6.Model Evaluation Metrics: Calculate and present key metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. Use a table or multi-bar chart to provide a comprehensive assessment of the model's performance.

7. Threshold Sensitivity: Visualize how varying threshold values impact precision and recall, with Laplace smoothing in consideration. A line graph or bar chart can illustrate these trade-offs.