PART-2(Email Filtering System Evaluation)

Confusion Matrix

Based on the problem description, we can summarize the classification results in the following confusion matrix:

	Predicted Spam	Predicted Legitimate
Actual Spam	True Positives (TP) = 200	False Negatives (FN) = 50
Actual Legitimate	False Positives (FP) = 20	True Negatives (TN) = 730

1. Accuracy

Accuracy is the proportion of correctly classified emails (both spam and legitimate) out of the total number of emails:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Substituting the values:

$$\text{Accuracy} = \frac{200 + 730}{200 + 730 + 20 + 50} = \frac{930}{1000} = 0.93$$

Thus, the model's accuracy is **93%**.

2. Precision

Precision measures how many of the emails classified as spam were actually spam:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Substituting the values:

$$\text{Precision} = \frac{200}{200 + 20} = \frac{200}{220} \approx 0.909$$

So, the model's precision is **90.9%**.

3. Recall

Recall (also known as Sensitivity or True Positive Rate) measures how many actual spam emails were correctly classified:

$$\text{Recall} = \frac{TP}{TP + FN}$$

Substituting the values:

$$\text{Recall} = \frac{200}{200 + 50} = \frac{200}{250} = 0.8$$

Thus, the model's recall is **80%**.

4. F1-Score

The F1-score is the harmonic mean of precision and recall:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Substituting the values:

$$F1 = 2 \times \frac{0.909 \times 0.8}{0.909 + 0.8} = 2 \times \frac{0.7272}{1.709} \approx 0.851$$

Thus, the F1-score is **85.1%**.

5. False Positive Rate

The false positive rate measures how many legitimate emails were incorrectly classified as spam:

$$\text{FPR} = \frac{FP}{FP + TN}$$

Substituting the values:

$$\text{FPR} = \frac{20}{20 + 730} = \frac{20}{750} \approx 0.0267$$

Thus, the false positive rate is **2.67%**.

Conclusion

The email filtering system has the following performance metrics:

- **Accuracy**: 93%
- **Precision**: 90.9%

- **Recall**: 80%
- **F1-Score**: 85.1%
- **False Positive Rate**: 2.67%

The model performs well overall, but there is a tradeoff between precision and recall. Only a small portion of legitimate emails are mistakenly flagged as spam, but the model could improve its ability to correctly classify all spam emails.