Naive Bayes Classifier: Applications and Evaluation Metrics

Dr. Sambit Praharaj, Assistant Professor (II), KIIT

1 Introduction

The Naive Bayes classifier is a probabilistic supervised learning algorithm widely used for text classification tasks due to its simplicity, efficiency, and effectiveness on high-dimensional data. Apart from sentiment analysis, Naive Bayes is applied to several real-world NLP problems. Classifier performance is evaluated using metrics such as Accuracy, Precision, Recall, and F-Measure.

2 Text Classification Tasks Using Naive Bayes (Beyond Sentiment Analysis)

2.1 Spam Detection

Emails or messages are classified as *Spam* or *Not Spam* based on word frequencies. Naive Bayes performs well due to the strong presence of discriminative keywords.

2.2 Topic or Document Classification

Documents are categorized into predefined topics such as Sports, Politics, Technology, or Health. This task is commonly used in news classification and digital libraries.

2.3 Language Identification

Naive Bayes can identify the language of a given text using character or word n-grams. It is effective due to distinct frequency patterns across languages.

2.4 Authorship Attribution

The classifier predicts the author of a document based on writing style, vocabulary usage, and function words. It is useful in plagiarism detection and forensic analysis.

2.5 Fake News Detection

News articles are classified as *Fake* or *Real* based on linguistic cues, sensational terms, and writing patterns.

2.6 Intent Classification

Used in chatbots and virtual assistants to identify user intent such as booking tickets, cancelling orders, or requesting support.

2.7 Question Classification

Questions are classified into types such as Who, What, When, Where, Why, and How. This is an important preprocessing step in question-answering systems.

2.8 Emotion Detection

Instead of simple polarity, text is classified into emotions like Joy, Anger, Sadness, Fear, or Surprise.

2.9 Toxicity and Hate Speech Detection

Text is classified as toxic, abusive, or non-toxic to support content moderation on social media platforms.

2.10 Domain-Specific Text Classification

Naive Bayes is used in medical, legal, and financial domains to classify documents such as medical reports, legal judgments, or financial complaints.

3 Confusion Matrix

A confusion matrix summarizes the prediction results of a classifier in a tabular form for a binary classification problem.

3.1 Definitions

- **True Positive (TP):** Positive instances correctly classified as positive
- **False Positive (FP):** Negative instances incorrectly classified as positive
- **False Negative (FN):** Positive instances incorrectly classified as negative
- **True Negative (TN):** Negative instances correctly classified as negative

3.2 Numerical Example

Consider a spam classification problem where Spam is the positive class. Out of 100 emails, the classifier produces the following results:

Actual / Predicted	**Spam**	**Not Spam**
Spam	40	10
Not Spam	5	45

Thus:

- TP = 40
- FN = 10
- FP = 5
- TN = 45

4 Precision

Precision measures how many of the predicted positive instances are actually positive.

4.1 Formula

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

4.2 Calculation

$$\text{Precision} = \frac{40}{40 + 5} = \frac{40}{45} = 0.8889$$

4.3 Interpretation

Approximately 88.89% of the emails classified as Spam are truly Spam.

5 Recall

Recall measures the ability of the classifier to correctly identify all actual positive instances.

5.1 Formula

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

5.2 Calculation

$$\text{Recall} = \frac{40}{40 + 10} = \frac{40}{50} = 0.8$$

5.3 Interpretation

The classifier successfully detects 80% of all Spam emails.

6 F-Measure (F1-Score)

F-Measure provides a single score that balances Precision and Recall.

6.1 Formula

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

6.2 Calculation

$$F1 = \frac{2 \times 0.8889 \times 0.8}{0.8889 + 0.8} = \frac{1.422}{1.6889} \approx 0.842$$

6.3 Interpretation

The classifier achieves an overall balanced performance of approximately 84.2%.

7 Why Accuracy Alone Is Insufficient

In class-imbalanced datasets, a classifier may achieve high accuracy by predicting the majority class. Precision, Recall, and F-Measure provide a more reliable evaluation, especially in applications such as:

- Spam detection
- Medical diagnosis
- Fraud detection
- Fake news detection

8 Conclusion

Naive Bayes is widely used for multiple text classification tasks beyond sentiment analysis. To evaluate its effectiveness, metrics such as Precision, Recall, and F-Measure are essential, as they provide deeper insight into classifier performance than accuracy alone.