# Tutorial 4: Naïve Bayes and Multiclass Precision, Recall

## Naïve Bayes

We want to build a naive bayes sentiment classifier using **add-1 smoothing**, as described in the lecture (not binary naive bayes, regular naive bayes). We denote the two classes positive and negative as + or – respectively. Here is our training corpus:

**Training Set**

- the movie has no plot

- honestly pretty boring

+ pretty interesting movie

**Test Set**

pretty enjoyable plot

Answer the questions below given the sets above.

1. Compute the prior for the two classes + and –

1 mark

1. Compute the following likelihoods:
   1. P(boring | - )
   2. P (boring | + )
   3. P (movie | + )

2 marks

1. Then compute whether the sentence in the test set is of class positive or negative (you may need a calculator for this final computation).

3 marks

## Multiclass Precision, Recall and F1

We want to measure the accuracy of a Twitter moderation bot. The classifier has the following classes: Safe content (S), Potentially Offensive Content (O), Prohibited Content (P). The classifier’s performance is given in the table below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Gold S** | **Gold O** | **Gold P** |
| **Predicted S** | 30 | 20 | 10 |
| **Predicted O** | 50 | 60 | 10 |
| **Predicted P** | 20 | 20 | 80 |

1. Compute the micro average precision and recall.

2 marks

1. Compute the macro average precision and recall.

2 marks

# Answer Key

## Naïve Bayes

Answer 1:

**P(-) = 2/3  
P(+) = 1/3**

Answer 2:

|V| = 9  
n- = 8  
n+ = 3

1. **P('boring' | - ) = (1 + 1) / (n- + |V|) = (1 + 1) / (8 + 9) = 2/17 = 0.1176**
2. **P('boring' | + ) = (0 + 1) / (n+ + |V|) = (0 + 1) / (3 + 9) = 1/12 = 0.0833**
3. **P('movie' | + ) = (1 + 1) / (n+ + |V|) = (1 + 1) / (3 + 9) = 2/12 = 0.1667**

Answer 3:

P (c | "pretty enjoyable plot") ∝ P(c) \* P ("pretty enjoyable plot" | c)

= P(c) \* P(pretty | c) \* P(enjoyable | c) \* P(plot | c)

≈ P(c) \* P(pretty | c) \* P(plot | c), 'enjoyable' is unknown

P( - | "pretty enjoyable plot") ∝ (2/3) \* (2/17) \* (2/17) = 0.009227

P( + | "pretty enjoyable plot") ∝ (1/3) \* (2/12) \* (1/12) = 0.004630

P( - | "pretty enjoyable plot") is greater, so the **test set sentence is classified as class negative**.

## Multiclass Precision, Recall and F1

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Gold S** | **Gold O** | **Gold P** |
| **Predicted S** | 30 | 20 | 10 |
| **Predicted O** | 50 | 60 | 10 |
| **Predicted P** | 20 | 20 | 80 |

Confusion Matrix

### Precision Micro-Average:

Precision for Class S = True Positives\_S / (True Positives\_S + False Positives\_S) = 30 / (30 + 20 + 10) = 30 / 60 = 0.5

Precision for Class O = True Positives\_O / (True Positives\_O + False Positives\_O) = 60 / (50 + 60 + 10) = 60 / 120 = 0.5

Precision for Class P = True Positives\_P / (True Positives\_P + False Positives\_P) = 80 / (20 + 20 + 80) = 80 / 120 = 0.6667 (approximately)

Micro-Average Precision = (Total True Positives) / (Total True Positives + Total False Positives) = (30 + 60 + 80) / (60 + 120 + 120) = 170 / 300 = 0.5667 (approximately)

### Recall Micro-Average:

Recall for Class S = True Positives\_S / (True Positives\_S + False Negatives\_S) = 30 / (30 + 50 + 20) = 30 / 100 = 0.3

Recall for Class O = True Positives\_O / (True Positives\_O + False Negatives\_O) = 60 / (20 + 60 + 20) = 60 / 100 = 0.6

Recall for Class P = True Positives\_P / (True Positives\_P + False Negatives\_P) = 80 / (10 + 10 + 80) = 80 / 100 = 0.8

Micro-Average Recall = (Total True Positives) / (Total True Positives + Total False Negatives) = (30 + 60 + 80) / (100 + 100 + 100) = 170 / 300 = 0.5667 (approximately)

### Precision Macro-Average:

Macro-Average Precision = (Precision for Class S + Precision for Class O + Precision for Class P) / Number of Classes = (0.5 + 0.5 + 0.6667) / 3 ≈ 0.5556 (approximately)

### Recall Macro-Average:

Macro-Average Recall = (Recall for Class S + Recall for Class O + Recall for Class P) / Number of Classes = (0.3 + 0.6 + 0.8) / 3 ≈ 0.5556 (approximately)