



CODING BLOCKS

Code Your Way To Success

Multinomial Event Model Naive Bayes

Steps to Compute Posterior Probabilities for Text Classification

Given an

- an input document $x = \langle w_1, w_2, \dots, w_n \rangle$
- the category that this document belongs to c

Training Time

- increment the count of total documents we have learned from N .
- increment the count of documents that have been mapped to this category N_c .
- if we encounter new words in this document, add them to our vocabulary, and update our vocabulary size $|V|$.
- update $\text{count}(w, c) \Rightarrow$ the frequency with which each word in the document has been mapped to this category.
- update $\text{count}(c) \Rightarrow$ the total count of all words that have been mapped to this class.

Testing Time

Compute Prior Prob of each class c -

$$P(c) = N_c / N$$

- N_c : No of documents having class c
- N : Total no of documents during training time

Likelihood ($P(X|Y=c)$)

We need to iterate through each word in the document and calculate:

$$P(w | c) = [\text{count}(w, c) + 1] / [\text{count}(c) + |V|]$$

This is the count of how many times this word has appeared in class c , plus 1, divided by the total count of all words that have ever been mapped to class c , plus the vocabulary size. This uses the Laplace-Smoothing, so we don't get tripped up by words we've never seen before. This equation is used both for words we have seen, as well as words we haven't seen.

$$x = \langle w_1, w_2, w_3, \dots, w_n \rangle$$

$$P(x|c) = \text{Product}(P(w_i|c))$$

- We multiply each $P(w | c)$ for each word w in the new document, then multiply by $P(c)$

), and the result is the probability that this document belongs to this class.

Final Prediction

Predict the class which has highest $P(Y=C|x)$ posterior probability.