Hide menu

Lecture: Naive Bayes

Video: Week Introduction
27 sec

Video: Probability and Bayes' Rule
3 min

Reading: Probability and Bayes' Rule
10 min

Video: Bayes' Rule
4 min

Reading: Bayes' Rule
10 min

Video: Naïve Bayes Introduction
5 min

Reading: Naive Bayes Introduction
10 min

Video: Laplacian Smoothing
2 min

Reading: Laplacian Smoothing
10 min

Video: Log Likelihood, Part 1
6 min

Reading: Log Likelihood, Part 1
10 min

Video: Log Likelihood, Part 2
2 min

Reading: Log Likelihood Part 2
10 min

Week 2 > Laplacian Smoothing

Previous Next

Laplacian Smoothing

We usually compute the probability of a word given a class as follows:

$$P\left(\mathrm{w_i} \mid \text{ class }\right) = \frac{\mathrm{freq}\left(\mathrm{w_i}, \text{ class }\right)}{\mathrm{N_{class}}} \quad \text{class } \in \{ \text{ Positive, Negative } \}$$

However, if a word does not appear in the training, then it automatically gets a probability of 0, to fix this we add smoothing as follows

$$P\left(\mathrm{w_i} \mid \text{class}\right) = \frac{\mathrm{freq}(\mathrm{w_i}, \text{ class }) + 1}{(\mathrm{N_{class}} + \mathrm{V})}$$

Note that we added a 1 in the numerator, and since there are V words to normalize, we add V in the denominator.

N_{class}: frequency of all words in class

V: number of unique words in vocabulary

Mark as completed