

Applied Text Mining in Python

Naïve Bayes Variations

Two Classic Naïve Bayes Variants for Text

- Multinomial Naïve Bayes
 - Data follows a **multinomial distribution**
 - Each feature value is **a count** (word occurrence counts, TF-IDF weighting, ...)
- Bernoulli Naïve Bayes
 - Data follows a **multivariate Bernoulli distribution**
 - Each feature is **binary** (word is present / absent)