Naive Bayes Classifier Example

Standard Naive Bayes Classifier

Class	Document
action	fast car race
action	car race fun race
action	loved fast race
comedy	loved fun couple loved it
comedy	fun couple flies fast very fun

Calculate the priors: $P(action) = \frac{3}{5}$ $P(comedy) = \frac{2}{5}$

Count the words in each class:

counts	car	couple	fast	flies	fun	it	loved	race	very	sum
action	2		2		1		1	4		10
comedy		2	1	1	3	1	2		1	11

Apply Add-1 smoothing:

counts	car	couple	fast	flies	fun	it	loved	race	very	sum
action	3	1	3	1	2	1	2	5	1	$10 + \|V\| = 19$
comedy	1	3	2	2	4	2	3	1	2	$11 + \|V\| = 20$

Calculate the likelihoods for each class:

probs	car	couple	fast	flies	fun	it	loved	race	very	sum
action	$\frac{3}{19}$	$\frac{1}{19}$	$\frac{3}{19}$	$\frac{1}{19}$	$\frac{2}{19}$	$\frac{1}{19}$	$\frac{2}{19}$	$\frac{5}{19}$	$\frac{1}{19}$	1
comedy	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{4}{20}$	$\frac{2}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	1

Exercise

What class does the model predict for the document: **loved plot and fun race**

Naive Bayes Classifier Example

Binary Naive Bayes Classifier

First, **remove duplicates in each document**. That is, create a **set** (rather than a bag) of words for each document:

Class	Document
action	fast car race
action	car race fun ~~race~~
action	loved fast race
comedy	loved fun couple ~~loved~~ it
comedy	fun couple flies fast very ~~fun~~

Calculate the priors: $P(action) = \frac{3}{5}$ $P(comedy) = \frac{2}{5}$

Count the words in each class:

counts	car	couple	fast	flies	fun	it	loved	race	very	sum
action	2		2		1		1	3		9
comedy		2	1	1	2	1	1		1	9

Apply Add-1 smoothing as usual:

counts	car	couple	fast	flies	fun	it	loved	race	very	sum
action	3	1	3	1	2	1	2	4	1	$9 + \|V\| = 18$
comedy	1	3	2	2	3	2	2	1	2	$9 + \|V\| = 18$

Calculate the likelihoods for each class:

probs	car	couple	fast	flies	fun	it	loved	race	very	sum
action	$\frac{3}{18}$	$\frac{1}{18}$	$\frac{3}{18}$	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{4}{18}$	$\frac{1}{18}$	1
comedy	$\frac{1}{18}$	$\frac{3}{18}$	$\frac{2}{18}$	$\frac{2}{18}$	$\frac{3}{18}$	$\frac{2}{18}$	$\frac{2}{18}$	$\frac{1}{18}$	$\frac{2}{18}$	1

Exercise

What class does the model predict for the document: **loved plot and fun race**