Aprendizaje Automático

Naïve Bayes

Viviana Cotik

Naïve Bayes (NB)

- Para cualquier problema de **clasificación**
- Uno de ellos: **clasificación de textos**
  - spam
  - en qué carpeta clasificar el e-mail
  - análisis de sentimientos
  - atribución de autoría
  - determinar tema de un artículo

Naïve Bayes

Es un **clasificador probabilístico**

Tipos de clasificadores:

- **“normales”**: devuelven clase más probable.
  - $\hat{y} = f(x)$
- **probabilísticos**: predicen **distribución de probabilidades** sobre un cjto. de clases
  - $P(Y|X)$
    - las probabilidades suman 1
  - Para obtener una sola clase:
    - $\hat{y} = \text{argmax}_y P(Y=y \mid X)$ (clase con mayor probabilidad)

Distribución de probabilidades

Función matemática que proporciona **probabilidad de ocurrencia** de **diferentes resultados posibles de un experimento**.

Naïve Bayes

Clasificador. Aprendizaje supervisado. Supuestos:

- usa regla de Bayes con una suposición Naïve
- con textos: usa una representación particular del documento (**bag of words** o bolsa de palabras)

Naïve Bayes aplicado a clasificación de textos

Entrada:

- documento d
- cant. fija de clases C = {c1, c2,..., ck}
- datos de entrenamiento $(x_1,c(x_1)),....,(x_n,c(x_m))$

Salida:

un clasificador f: d->C

Naïve Bayes

Por ej: Spam-no spam

f(

To: < omitted >
From: Get Rich Click
Subject: Getting better all the time!

Dear Get Rich Click player,

Come play the biggest sweepstakes on the Web. With new ways to win every week, you can't afford to pass this one by! Just click here: http://www.getrichclick.com
****
But First...
****
Get a free $50 gift with a minimum purchase of $50! The Golden Palace offers 28 online casino games. Play for FREE or try your luck for REAL $$$. FREE Software * 24-HR Customer Service * Best Odds * Play to win up to $200,000 INSTANTLY! Click: http://www.goldenpalace.com/indexyy.html

) = c

Ejemplo tomado de: http://web.mit.edu/network/spam/examples/getrich.html

Bag of words (bolsa de palabras)

Texto: **“genial oportunidad. aproveche ya”**

diccionario

a
aprovechar
asa
genial
gol
…
ya

bolsa de palabras*

$$x = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ \dots \\ 1 \end{bmatrix}$$

$$x_i = \begin{cases} \mathbf{1^*} & \text{si } pal_i \text{ está en diccionario*} \\ \mathbf{0} & \text{sino} \end{cases}$$

* podría ser la cant. de palabras en vez de si está o no la palabra

diccionario: diccionario o palabras más frecuentes en mi corpus

Naïve Bayes

documento d, clase c

$$P(c \mid d) = \frac{P(d \mid c)P(c)}{P(d)}$$

$$c_{MAP} = \underset{c \in C}{\operatorname{argmax}}\, P(c \mid d)$$

$$= \underset{c \in C}{\operatorname{argmax}}\, \frac{P(d \mid c)P(c)}{P(d)}$$

$$= \underset{c \in C}{\operatorname{argmax}}\, P(d \mid c)P(c)$$

MAP es "máximo a posteriori" = clase más probable

Regla de Bayes

Eliminamos el denominador

Tomado de curso NLP Stanford (Jurafsky, Manning)

Naïve Bayes

$$c_{MAP} = \underset{c \in C}{\operatorname{argmax}} P(d \mid c) P(c)$$

$$= \underset{c \in C}{\operatorname{argmax}} P(x_1, x_2, \ldots, x_n \mid c) P(c)$$

Documento d representado como atributos x1..xn

Tomado de curso NLP Stanford (Jurafsky, Manning)

Naïve Bayes - Suposiciones de independencia

$$P(x_1, x_2, \ldots, x_n \mid c)$$

- **Bolsa de palabras:** se asume que no importa la posición
- **Independencia condicional:** se asume que las probabilidades de los atributos (features) $P(x_i \mid c_j)$ son independientes dada la clase c

$$P(x_1, \ldots, x_n \mid c) = P(x_1 \mid c) \bullet P(x_2 \mid c) \bullet P(x_3 \mid c) \bullet \ldots \bullet P(x_n \mid c)$$

Tomado de curso NLP Stanford (Jurafsky, Manning)

Naïve Bayes Multinomial

$$c_{MAP} = \underset{c \in C}{\operatorname{argmax}} P(x_1, x_2, \ldots, x_n \mid c) P(c)$$

$$c_{NB} = \underset{c \in C}{\operatorname{argmax}} P(c) \prod_{x \in X} P(x \mid c)$$

Tomado de curso NLP Stanford (Jurafsky, Manning)

Aplicación de NB Multinomial a Clasificación de Textos

positions ← palabras en el documento

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(c_j) \prod_{i \in positions} P(x_i \mid c_j)$$

Tomado de curso NLP Stanford (Jurafsky, Manning)

Corrección para atributos no vistos previamente

Estimadores usando frecuencia de datos

$$\hat{P}(c_j) = \frac{doccount(C = c_j)}{N_{doc}}$$

$$\hat{P}(w_i \mid c_j) = \frac{count(w_i, c_j)}{\sum_{w \in V} count(w, c_j)}$$

Laplace add-1 smoothing

$$c_{MAP} = \text{argmax}_c \hat{P}(c) \prod_i \hat{P}(x_i \mid c)$$

$$\hat{P}(w_i \mid c) = \frac{count(w_i, c) + 1}{\left(\sum_{w \in V} count(w, c) \right) + |V|}$$

Ejemplo en Clasificación de textos

Dan Jurafsky

$$\hat{P}(c) = \frac{N_c}{N}$$

$$\hat{P}(w \mid c) = \frac{count(w,c)+1}{count(c)+|V|}$$

	Doc	Words	Class
Training	1	Chinese Beijing Chinese	c
	2	Chinese Chinese Shanghai	c
	3	Chinese Macao	c
	4	Tokyo Japan Chinese	j
Test	5	Chinese Chinese Chinese Tokyo Japan	?

Ejemplo tomado de curso NLP Stanford (Jurafsky, Manning)

Priors:

$P(c) = \frac{3}{4}$

$P(j) = \frac{1}{4}$

Conditional Probabilities:

P(Chinese|c) = (5+1) / (8+6) = 6/14 = 3/7

P(Tokyo|c) = (0+1) / (8+6) = 1/14

P(Japan|c) = (0+1) / (8+6) = 1/14

P(Chinese|j) = (1+1) / (3+6) = 2/9

P(Tokyo|j) = (1+1) / (3+6) = 2/9

P(Japan|j) = (1+1) / (3+6) = 2/9

Choosing a class:

$P(c|d5) \propto 3/4 * (3/7)^3 * 1/14 * 1/14 \approx 0.0003$

$P(j|d5) \propto 1/4 * (2/9)^3 * 2/9 * 2/9 \approx 0.0001$

Multinomial NB: Aprendizaje

- From training corpus, extract *Vocabulary*

- Calculate $P(c_j)$ terms
  - For each c_j in C do
    $docs_j \leftarrow$ all docs with class $= c_j$

- Calculate $P(w_k \mid c_j)$ terms
  - $Text_j \leftarrow$ single doc containing all $docs_j$
  - For each word w_k in *Vocabulary*
    $n_k \leftarrow$ # of occurrences of w_k in $Text_j$

$$P(w_k \mid c_j) \leftarrow \frac{n_k + \alpha}{n + \alpha \mid Vocabulary \mid}$$

Tomado de curso NLP Stanford (Jurafsky, Manning)

Naïve Bayes con Atributos Categóricos

	Atributos				Clase
Instancia	Cielo	Temperatura	Humedad	Viento	Va a correr?
1	sol	calor	alta	débil	No
2	sol	calor	alta	fuerte	No
3	nublado	calor	alta	débil	Sí
4	lluvia	templado	alta	débil	Sí
5	lluvia	frío	normal	débil	Sí
6	lluvia	frío	normal	fuerte	No
7	nublado	frío	normal	fuerte	Sí
8	sol	templado	alta	débil	No
9	sol	frío	normal	débil	Sí
10	lluvia	templado	normal	débil	Sí
11	sol	templado	normal	fuerte	Sí
12	nublado	templado	alta	fuerte	Sí
13	nublado	calor	normal	débil	Sí
14	lluvia	templado	alta	fuerte	No

P(Sí) = 9/14

P(No) = 5/14

Naïve Bayes con Atributos Categóricos

	Atributos				Clase
Instancia	Cielo	Temperatura	Humedad	Viento	Va a correr?
1	sol	calor	alta	débil	No
2	sol	calor	alta	fuerte	No
3	nublado	calor	alta	débil	Sí
4	lluvia	templado	alta	débil	Sí
5	lluvia	frío	normal	débil	Sí
6	lluvia	frío	normal	fuerte	No
7	nublado	frío	normal	fuerte	Sí
8	sol	templado	alta	débil	No
9	sol	frío	normal	débil	Sí
10	lluvia	templado	normal	débil	Sí
11	sol	templado	normal	fuerte	Sí
12	nublado	templado	alta	fuerte	Sí
13	nublado	calor	normal	débil	Sí
14	lluvia	templado	alta	fuerte	No

P(Sí) = 9/14

P(No) = 5/14

Cielo				
	Sí	No	P(Sí)	P(No)
sol	2	3	2/9	3/5
nublado	4	0	4/9	0/5
lluvia	3	2	3/9	2/5
total	9	5	100%	100%

Temperatura				
	Sí	No	P(Sí)	P(No)
calor	2	2	2/9	2/5
templado	4	2	4/9	2/5
frío	3	1	3/9	1/5
total	9	5	100%	100%

Ídem para:

Humedad

Viento

Naïve Bayes con Atributos Categóricos

Cielo				
	Sí	No	P(Sí)	P(No)
sol	2	3	2/9	3/5
nublado	4	0	4/9	0/5
lluvia	3	2	3/9	2/5
total	9	5	100%	100%

Temperatura				
	Sí	No	P(Sí)	P(No)
calor	2	2	2/9	2/5
templado	4	2	4/9	2/5
frío	3	1	3/9	1/5
total	9	5	100%	100%

P(Sí) = 9/14
P(No) = 5/15

P(Cielo=sol|Juega = Sí) = 2/9
P(Temperatura=templado|Juega = Sí) = 4/9
…

P(Cielo=sol|Juega = No) = 3/5
P(Temperatura=templado|Juega = No) = 2/5
…
Para clasificar (sol, templado, alta, fuerte), se calcula
P(Juega = Sí | x) =
= P(x|Juega = Sí) * P(Sí) = 2/9 * 4/9 * … * … * 9/14
P(Juega = No | x) =
P(x|Juega = No) * P(No) = 3/5 * 2/5 * … * … * 5/14

Resumen

- asume independencia de variables.
- muy usado en **clasificación de textos** (por ej. filtros de spam y en análisis de sentimientos). Sirve como baseline.
- muy rápido
- pocos requerimientos de almacenamiento

Bibliografía

Capítulos de libros:
Mitchell (6.2, 6.9, 6.10)

Paper:
https://ai.stanford.edu/~ang/papers/nips01-discriminativegenerative.pdf