

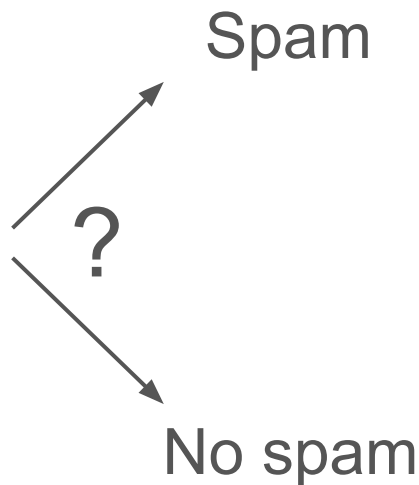
Clasificación de Textos

Clasificación de Textos

Detección de spam

From: selenac.sespa.es
Reply-to: sgtclark0654@hotmail.com
Subject: Oportunidad Única!

Tengo una propuesta comercial que podría interesarte. Notifícame si estás interesado en más detalles.

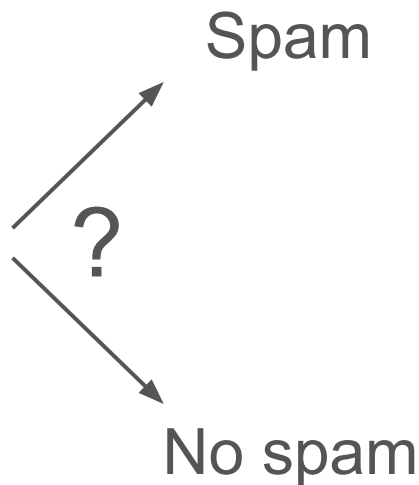


Clasificación de Textos

Detección de spam

From: selenac.casas@sespa.es
Reply-to: sgtclark0654@hotmail.com
Subject: Oportunidad Única!

Tengo una propuesta comercial que podría interesarte. **Notifícame** si estás interesado en **más detalles**.



Clasificación de Textos

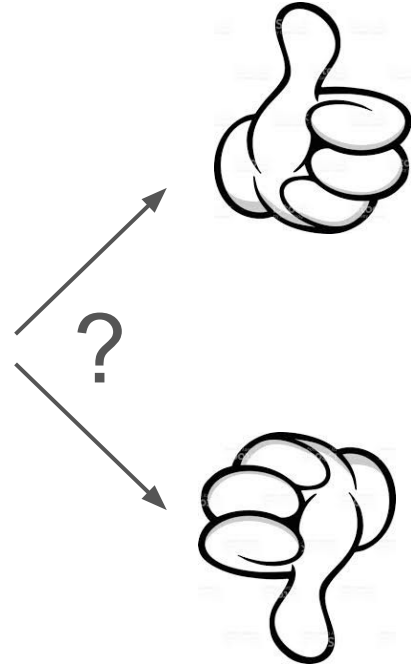
Detección de sentimiento

@juanma25



Que **bien** que estoy!

Empanadas + Netflix , infalible!



Clasificación de Textos

De que país es este tweet?

@amiguis90

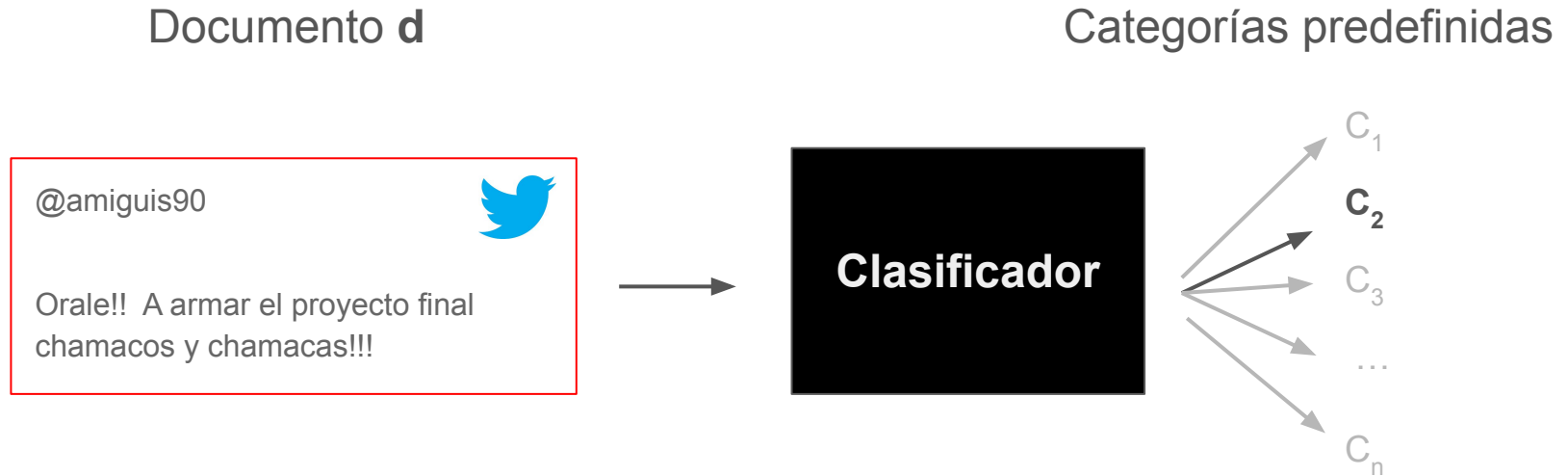


Orale!! A armar el proyecto final **chamacos**
y **chamacas!!!**



Clasificación de Textos

Idea general



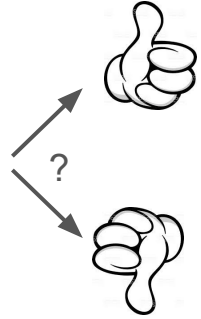
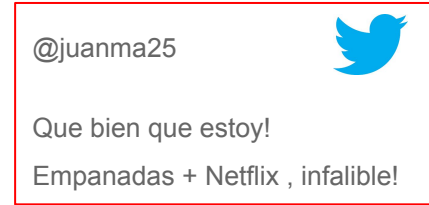
Métodos:

Opción 1

- Uso de reglas y expresiones regulares

Opción 2

- Supervised Machine Learning

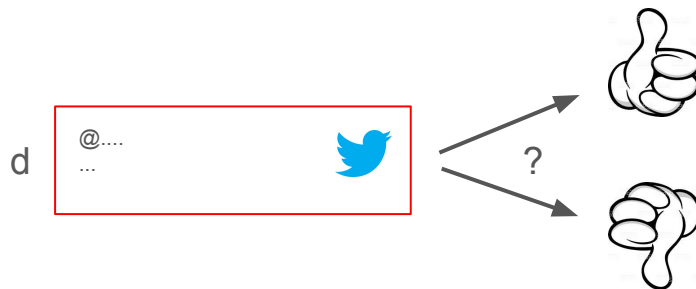


Supervised Machine Learning

Input $\longrightarrow \{ (d_1, c_1), (d_2, c_2), (d_3, c_3), \dots, (d_n, c_n) \}$

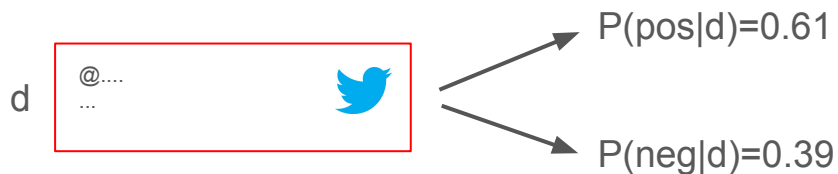


Clasificador



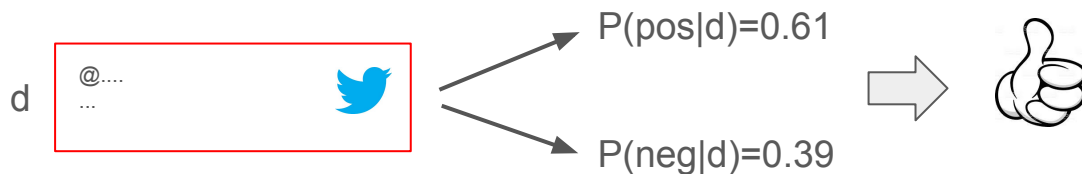
Naïve Bayes

- Input $\longrightarrow \{ (d_1, c_1), (d_2, c_2), (d_3, c_3), \dots, (d_n, c_n) \}$
- Dado un nuevo $(d, ?)$, estima $P(\text{pos}|d)$ y $P(\text{neg}|d)$



Naïve Bayes

- Input $\longrightarrow \{ (d_1, c_1), (d_2, c_2), (d_3, c_3), \dots, (d_n, c_n) \}$
- Dado un nuevo $(d, ?)$, estima $P(\text{pos}|d)$ y $P(\text{neg}|d)$
- Selecciona la clase de mayor probabilidad (Maximum A Posteriori)



$$c_{MAP} = \underset{c \in C = \{\text{pos}, \text{neg}\}}{\operatorname{argmax}} P(c|d)$$

Naïve Bayes

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

Teorema de Bayes

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

- $P(c|d)$: Probabilidad de que el documento d sea de la clase c (posterior)
- $P(d|c)$: Probabilidad de obtener el documento “ d ” dado que es de la clase c (likelihood)
- $P(c)$: Probabilidad de la clase c (prior)
- $P(d)$: Probabilidad de de obtener un documento d

Naïve Bayes

$$\begin{aligned}c_{MAP} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(c|d) \\&= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} \frac{P(d|c)P(c)}{P(d)}\end{aligned}$$

Teorema de Bayes

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

- $P(c|d)$: Probabilidad de que el documento d sea de la clase c (posterior)
- $P(d|c)$: Probabilidad de obtener el documento “ d ” dado que es de la clase c (likelihood)
- $P(c)$: Probabilidad de la clase c (prior)
- $P(d)$: Probabilidad de de obtener un documento d

Naïve Bayes

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

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

No depende de c

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

Teorema de Bayes




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

- $P(c|d)$: Probabilidad de que el documento d sea de la clase c (posterior)
- $P(d|c)$: Probabilidad de obtener el documento “ d ” dado que es de la clase c (likelihood)
- $P(c)$: Probabilidad de la clase c (prior)
- $P(d)$: Probabilidad de de obtener un documento d

Naïve Bayes

$$\begin{aligned}c_{MAP} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(d|c)P(c) \\ &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1, x_2, x_3, \dots, x_m | c)P(c)\end{aligned}$$

Set de entrenamiento

| | | |
|-------|---|---|
| d_1 | @nos_paso_a_todos Odio al heladero! Me puso casi todo de americana! | $c_1 =$  |
| d_2 | @maria_1991 Muy buena peliiii, Las quiero amigas! | $c_2 =$  |
| ... | | |
| d_n | @pedro_1990 Que feo que me salio este mate | $c_n =$  |

Quiero predecir

| | | |
|-----|----------------------------------|---------------|
| d | @juanma25 Que bien que estoy! | $c_{MAP} = ?$ |
|-----|----------------------------------|---------------|

Naïve Bayes

$$\begin{aligned} c_{MAP} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(d|c)P(c) \\ &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1, x_2, x_3, \dots, x_m | c)P(c) \end{aligned}$$

estimación

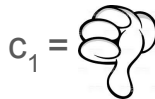
$$\frac{\#c}{\#D}$$



Set de entrenamiento

d_1

@nos_paso_a_todos
Odio al heladero! Me puso
casi todo de americana!



d_2

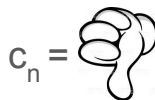
@maria_1991
Muy **buena** peliiii,
Las **quiero** amigas!



...

d_n

@pedro_1990
Que **feo** que me salio
este mate



Quiero predecir

d

@juanma25
Que bien que estoy!

$c_{MAP} = ?$

Naïve Bayes

$$\begin{aligned}
 c_{MAP} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(d|c)P(c) \\
 &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1, x_2, x_3, \dots, x_m | c)P(c)
 \end{aligned}$$

Difícil de
estimar


estimación

$$\frac{\#c}{\#D}$$

Set de entrenamiento

d_1

@nos_paso_a_todos
Odio al heladero! Me puso
casi todo de americana!

$c_1 =$ 

d_2


@maria_1991
Muy **buena** peliiii,
Las **quiero** amigas!

$c_2 =$ 

...

d_n

@pedro_1990
Que **feo** que me salio
este mate

$c_n =$ 

Quiero predecir

d

@juanma25
Que bien que estoy!

$c_{MAP} = ?$

Naïve Bayes

$$\begin{aligned}
 c_{MAP} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(d|c)P(c) \\
 &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1, x_2, x_3, \dots, x_m | c)P(c)
 \end{aligned}$$

Difícil de
estimar

estimación

$$\frac{\#c}{\#D}$$

Multinomial Naïve Bayes approximation

- **Bag of words:** La posición de las palabras no importa
- **Independencia Condicional:** $P(x_i|c)$ son independientes


$$P(x_1, x_2, \dots, x_m | c) = P(x_1 | c) \cdot P(x_2 | c) \dots P(x_m | c)$$

$$P(x_1, x_2, \dots, x_m | c) = P(Que|c)P(bien|c)P(que|c)P(estoy|c)P(!|c)$$

Set de entrenamiento

d_1

@nos_paso_a_todos
Odio al heladero! Me puso
casi todo de americana!

$c_1 =$ 

d_2


@maria_1991
Muy **buena** peliiii,
Las **quiero** amigas!

$c_2 =$ 

...

d_n

@pedro_1990
Que **feo** que me salio
este mate

$c_n =$ 

Quiero predecir

d

@juanma25
Que bien que estoy!

$c_{MAP} = ?$

Multinomial Naïve Bayes

$$\begin{aligned}c_{MNB} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1|c)P(x_2|c)\dots P(x_m|c)P(c) \\ &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(c) \prod_i P(x_i|c)\end{aligned}$$

Estimations

$$\begin{aligned}c_{MNB} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1|c)P(x_2|c)\dots P(x_m|c)P(c) \\ &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(c) \prod_i P(x_i|c)\end{aligned}$$

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

Fracción del training set con clase c

$$\hat{P}(w_i|c) = \frac{\operatorname{count}(w_i, c)}{\sum_{w \in V} \operatorname{count}(w, c)}$$

Fracción de apariciones de la palabra w_i entre todas las palabras de los documentos de clase c

Bolsa de palabras de todos los documentos de clase c

| Palabras | Frec. |
|----------|-------|
| a | 1921 |
| arriba | 121 |
| | |
| zebra | 0 |


Ejemplo


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


$$\hat{P}(w_i|c) = \frac{\text{count}(w_i, c)}{\sum_{w \in V} \text{count}(w, c)}$$

$$c_{MNB} = \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1|c)P(x_2|c)\dots P(x_m|c)P(c)$$

$$= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(c) \prod_i P(x_i|c)$$

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$


Ejemplo


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


$$\hat{P}(w_i|c) = \frac{\text{count}(w_i, c)}{\sum_{w \in V} \text{count}(w, c)}$$

$$c_{MNB} = \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1|c)P(x_2|c)\dots P(x_m|c)P(c)$$

$$= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(c) \prod_i P(x_i|c)$$

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$

$$\propto \frac{2}{3} \left[\begin{array}{c} \\ \end{array} \right]$$


Ejemplo


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


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

BoW de todos los documentos de clase *neg*

| Palabras | Frec. |
|----------|-------|
| estoy | 2 |
| triste | 2 |
| mojado | 1 |
| hambre | 1 |

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?


$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$


$$\propto \frac{2}{3} \begin{bmatrix} \quad \end{bmatrix}$$


Ejemplo

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

$$\hat{P}(w_i|c) = \frac{\text{count}(w_i, c)}{\sum_{w \in V} \text{count}(w, c)}$$

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

BoW de todos los documentos de clase *neg*

| Palabras | Frec. |
|----------|-------|
| estoy | 2 |
| triste | 2 |
| mojado | 1 |
| hambre | 1 |


$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$


$$\propto \frac{2}{3} \left[\frac{2}{6} \cdot \frac{1}{6} \right] = \frac{1}{27}$$


Ejemplo

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

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

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

BoW de todos los documentos de clase *neg*

| Palabras | Frec. |
|----------|-------|
| estoy | 2 |
| triste | 2 |
| mojado | 1 |
| hambre | 1 |

BoW de todos los documentos de clase *pos*

| Palabras | Frec. |
|----------|-------|
| estoy | 1 |
| feliz | 1 |

$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$


$$\propto \frac{2}{3} \left[\frac{2}{6} \cdot \frac{1}{6} \right] = \frac{1}{27}$$


$$P(pos|d_{test}) \propto P(pos) [P(estoy|pos)P(hambre|pos)]$$


Ejemplo

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

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

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

BoW de todos los documentos de clase *neg*

| Palabras | Frec. |
|----------|-------|
| estoy | 2 |
| triste | 2 |
| mojado | 1 |
| hambre | 1 |

BoW de todos los documentos de clase *pos*

| Palabras | Frec. |
|----------|-------|
| estoy | 1 |
| feliz | 1 |

$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$

$$\propto \frac{2}{3} \left[\frac{2}{6} \cdot \frac{1}{6} \right] = \frac{1}{27}$$

$$P(pos|d_{test}) \propto P(pos) [P(estoy|pos)P(hambre|pos)]$$

$$\propto \frac{1}{3} \left[\frac{1}{2} \cdot 0 \right] = 0$$

Laplace (add-1) smoothing

$$\begin{aligned}\hat{P}(w_i|c) &= \frac{\text{count}(w_i,c)+1}{\sum_{w \in V} \text{count}(w,c)+1} \\ &= \frac{\text{count}(w_i,c)+1}{\left[\sum_{w \in V} \text{count}(w,c) \right] + |V|}\end{aligned}$$

add- α smoothing

$$\begin{aligned}\hat{P}(w_i|c) &= \frac{\text{count}(w_i,c)+\alpha}{\sum_{w \in V} \text{count}(w,c)+\alpha} \\ &= \frac{\text{count}(w_i,c)+\alpha}{\left[\sum_{w \in V} \text{count}(w,c) \right] + \alpha|V|}\end{aligned}$$

Bolsa de palabras de todos los documentos de clase c

| Palabras | Frec. |
|----------|--------|
| a | 1921+1 |
| arriba | 121+1 |
| | |
| zebra | 0+1 |

| Palabras | Frec. |
|----------|-----------------|
| a | 1921 + α |
| arriba | 121 + α |
| | |
| zebra | 0 + α |

Ejercicio: ahora con smoothing ($\alpha=0.1$)

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


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


neg


| Palabras | Frec. |
|----------|-------|
| estoy | 2 |
| triste | 2 |
| mojado | 1 |
| hambre | 1 |

pos

| Palabras | Frec. |
|----------|-------|
| estoy | 1 |
| feliz | 1 |

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$

\propto

$$P(pos|d_{test}) \propto P(pos) [P(estoy|pos)P(hambre|pos)]$$

\propto

Ahora con smoothing ($\alpha=0.1$)

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


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


neg


| Palabras | Frec. |
|----------|-------|
| estoy | 2 |
| triste | 2 |
| mojado | 1 |
| hambre | 1 |

pos

| Palabras | Frec. |
|----------|-------|
| estoy | 1 |
| feliz | 1 |

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$

$$\propto \frac{2}{3} \left[\frac{2.1}{6.5} \cdot \frac{1.1}{6.5} \right] \approx 0,036$$

$$P(pos|d_{test}) \propto P(pos) [P(estoy|pos)P(hambre|pos)]$$

$$\propto$$

Ahora con smoothing ($\alpha=0.1$)

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


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


neg


| Palabras | Frec. |
|----------|-------|
| estoy | 2 |
| triste | 2 |
| mojado | 1 |
| hambre | 1 |

pos

| Palabras | Frec. |
|----------|-------|
| estoy | 1 |
| feliz | 1 |

d_1 Estoy feliz $c_1 =$ 

d_2 Estoy triste y mojado $c_2 =$ 

d_3 Estoy triste y con hambre $c_3 =$ 

d_{test} estoy con hambre $c_{test} =$?

$$P(neg|d_{test}) \propto P(neg) [P(estoy|neg)P(hambre|neg)]$$

$$\propto \frac{2}{3} \left[\frac{2.1}{6.5} \cdot \frac{1.1}{6.5} \right] \approx 0,036$$

$$P(pos|d_{test}) \propto P(pos) [P(estoy|pos)P(hambre|pos)]$$

$$\propto \frac{1}{3} \left[\frac{1.1}{2.5} \cdot \frac{0.1}{2.5} \right] \approx 0,0059$$

Problemas numéricos

$$\begin{aligned}c_{MNB} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1|c)P(x_2|c) \dots P(x_m|c)P(c) \\ &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(c) \prod_i P(x_i|c)\end{aligned}$$

La productoria de muchos números muy chicos da un número muy muy muy chico

Problemas numéricos

$$\begin{aligned}c_{MNB} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(x_1|c)P(x_2|c) \dots P(x_m|c)P(c) \\ &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} P(c) \prod_i P(x_i|c)\end{aligned}$$

La productoria de muchos números muy chicos da un número muy muy muy chico

Solución:

La clase con mayor probabilidad, también tendrá mayor $\log(\text{probabilidad})$

$$\begin{aligned}c_{MNB} &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} \log[P(c) \prod_i P(x_i|c)] \\ &= \underset{c \in C = \{pos, neg\}}{\operatorname{argmax}} \log P(c) + \sum_i \log P(x_i|c)\end{aligned}$$

Tips

Cómo mejorar (o quizás no) un sistema de clasificación de texto:

➤ **Normalización:**

- Colapso de tokens (números, fechas, nombres)

Tips

Cómo mejorar (o quizás no) un sistema de clasificación de texto:

➤ **Normalización:**

- Colapso de tokens (números, fechas, nombres)
- **Lematizar** o hacer **stemming**

➤ **Aumentar** el peso de ciertos tokens (duplicar tokens):

- Tokens del título (Cohen & Singer 1996)
- Primera oración de cada párrafo (Murata, 1999)
- En oraciones que contienen palabras del título (Ko, 2002)

Tips

- Se pueden descartar las palabras muy **poco frecuentes** o las **muy frecuentes**

,
,
,
,

Tips

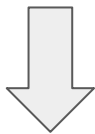
- Se pueden descartar las palabras muy **poco frecuentes** o las **muy frecuentes**
- Se pueden descartar **stopwords** (“de”, “la”, “los”, “que”,...)

Tips

- Se pueden descartar las palabras muy **poco frecuentes** o las **muy frecuentes**
- Se pueden descartar **stopwords** (“de”, “la”, “los”, “que”,...)
- Para el *sentiment analysis* se pueden **propagar negaciones**

Por ejemplo:

Esto no está bueno ni rico. En la casa de...



Esto no no_está no_bueno no_ni no_rico En la casa de...

Tips

- Se pueden extraer n-gramas

“Esto no está bueno ni rico.”

- Unigramas:

[“Esto”, “no”, “está”, “bueno”, “ni”, “rico”, “.”]

- Bigramas:

[“Esto no”, “no está”, “está bueno”, “bueno ni”, “ni rico”, “rico .”]

- Trigramas:

[“Esto no está”, “no está bueno”, “está bueno ni”, ...]

Clasificación de Textos

Naïve Bayes:

- No tiende a over-fitear, high-bias, (Ng and Jordan, 2002)
- No tiene muchos parámetros a ajustar
- Es intuitivo
- Es rapido

Clasificación de Textos

Naïve Bayes:

- No tiende a over-fitear, high-bias, (Ng and Jordan, 2002)
- No tiene muchos parámetros a ajustar
- Es intuitivo
- Es rapido

Pero:

- Es un primer algoritmo a probar, es probable que otros métodos de clasificación supervisada funcionen mejor, como: SVM, NNs

Detección de hate speech en twitter

Train set



80% del dataset

Test set



20% del dataset

Evaluación de los métodos

Identificación de mensajes con hate speech:
Sobre el test set

| | Eran <i>hate</i> | Eran <i>no-hate</i> |
|-----------------------------------|------------------|---------------------|
| Identificados como <i>hate</i> | 8 (tp) | 8 (fp) |
| Identificados como <i>no-hate</i> | 2 (fn) | 92 (tn) |

Matriz de confusión

Notar que estoy evaluando la capacidad de identificar los posts de **hate**, y no la capacidad de identificar los de **no-hate**

Precision y Recall

Precision: fracción de los identificados como *hate* que fueron correctamente clasificados

Recall: fracción de los que eran *hate*, que efectivamente fueron identificados como *hate*

| | Eran <i>hate</i> | Eran no- <i>hate</i> |
|------------------------------------|------------------|----------------------|
| Identificados como <i>hate</i> | 8 (tp) | 8 (fp) |
| Identificados como no- <i>hate</i> | 2 (fn) | 92 (tn) |

$$Precision = \frac{tp}{tp+fp} = \frac{8}{16} \approx 0.5$$

$$Recall = \frac{tp}{tp+fn} = \frac{8}{10} \approx 0.8$$

Precision y Recall

Casos límites:

Clasifico siempre como *hate*

| | Eran <i>hate</i> | Eran <i>no-hate</i> |
|-----------------------------------|------------------|---------------------|
| Identificados como <i>hate</i> | 10 (tp) | 100 (fp) |
| Identificados como <i>no-hate</i> | 0 (fn) | 0 (tn) |

$$Precision = \frac{tp}{tp+fp} = \frac{10}{110} \approx 0.09$$

$$Recall = \frac{tp}{tp+fn} = \frac{10}{10} \approx 1$$

Clasifico como *hate* solo si estoy muuuy seguro

| | Eran <i>hate</i> | Eran <i>no-hate</i> |
|-----------------------------------|------------------|---------------------|
| Identificados como <i>hate</i> | 3 (tp) | 1 (fp) |
| Identificados como <i>no-hate</i> | 7 (fn) | 99 (tn) |

$$Precision = \frac{tp}{tp+fp} = \frac{3}{4} \approx 0.75$$

$$Recall = \frac{tp}{tp+fn} = \frac{3}{10} \approx 0.3$$

F-measure

El F-measure (F1-score) es un trade off entre el Precision y el Recall y se calcula como el promedio armónico entre ambos

$$F = \frac{2}{\frac{1}{P} + \frac{1}{R}} = \frac{2PR}{P+R}$$

$$Precision = \frac{tp}{tp+fp} = \frac{8}{16} \approx 0.5$$

$$Recall = \frac{tp}{tp+fn} = \frac{8}{10} \approx 0.8$$

$$\longrightarrow F = 0.616$$

Macro vs Micro averaging

Cuando no hay una clase privilegiada

| | Eran bot | Eran troll | Eran normal |
|----------------|----------|------------|-------------|
| Clasif: bot | 10 | 20 | 0 |
| Clasif: troll | 10 | 40 | 0 |
| Clasif: normal | 0 | 0 | 1000 |

| bot | eran pos | eran neg |
|-------------|----------|-----------|
| Clasif: pos | 10 (tp) | 20 (fp) |
| Clasif: neg | 10 (fn) | 1040 (tn) |

| troll | eran pos | eran neg |
|-------------|----------|-----------|
| Clasif: pos | 40 (tp) | 10 (fp) |
| Clasif: neg | 20 (fn) | 1010 (tn) |

| normal | eran: pos | eran: neg |
|-------------|-----------|-----------|
| Clasif: pos | 1000 (tp) | 0 (fp) |
| Clasif: neg | 0 (fn) | 80 (tn) |

Macro vs Micro averaging

| bot | eran pos | eran neg |
|-------------|----------|-----------|
| Clasif: pos | 10 (tp) | 20 (fp) |
| Clasif: neg | 10 (fn) | 1040 (tn) |

precision = 0.5

| troll | eran pos | eran neg |
|-------------|----------|-----------|
| Clasif: pos | 40 (tp) | 10 (fp) |
| Clasif: neg | 20 (fn) | 1010 (tn) |

precision = 0.67

| normal | eran: pos | eran: neg |
|-------------|-----------|-----------|
| Clasif: pos | 1000 (tp) | 0 (fp) |
| Clasif: neg | 0 (fn) | 80 (tn) |

precision = 1

Macro-averaging precision = $(0.5 + 0.67 + 1)/3 = 0.72$

Macro vs Micro averaging

| | | |
|-------------|----------|-----------|
| bot | eran pos | eran neg |
| Clasif: pos | 10 (tp) | 20 (fp) |
| Clasif: neg | 10 (fn) | 1040 (tn) |

precision = 0.5

| | | |
|-------------|----------|-----------|
| troll | eran pos | eran neg |
| Clasif: pos | 40 (tp) | 10 (fp) |
| Clasif: neg | 20 (fn) | 1010 (tn) |

precision = 0.67

| | | |
|-------------|-----------|-----------|
| normal | eran: pos | eran: neg |
| Clasif: pos | 1000 (tp) | 0 (fp) |
| Clasif: neg | 0 (fn) | 80 (tn) |

precision = 1

Macro-averaging precision = $(0.5 + 0.67 + 1)/3 = 0.72$

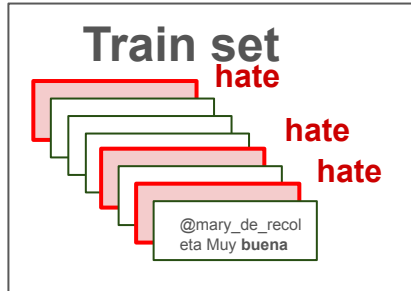
Micro-averaging table

| | | |
|-------------|-----------|-----------|
| Total | eran pos | eran neg |
| Clasif: pos | 1050 (tp) | 30 (fp) |
| Clasif: neg | 30 (fn) | 2130 (tn) |

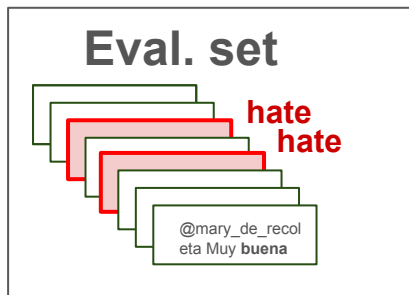
Micro-averaging precision = $1050/1080=0.97$

Training - Evaluation sets

Training set
80%



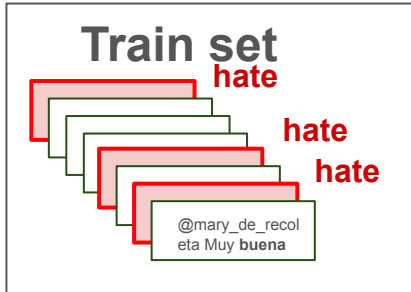
Eval. set
20%



Opción 1

Cross Validation (5-fold CV)

Training set
80%

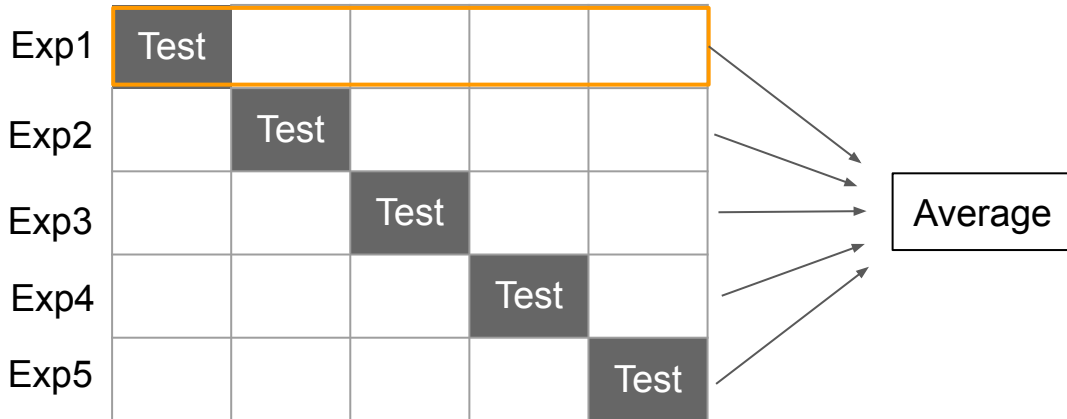


Eval. set
20%



Opción 1

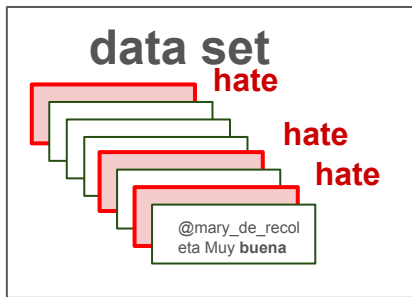
5-Fold Cross-Validation



Opción 2

Y si quiero probar muchos modelos?

Evalúo muchos modelos y elijo el mejor



Modelos

1) f-score=0.56

2) f-score=0.61

...

351) f-score=0.83

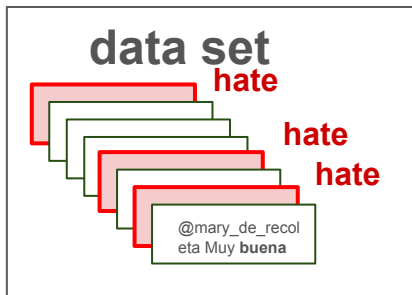
...

9581) f-score=0.59

9582) f-score=0.66

Y si quiero probar muchos modelos?

Evalúo muchos modelos y elijo el mejor



Modelos

1) f-score=0.56

2) f-score=0.61

...

351) f-score=0.83

...

9581) f-score=0.59

9582) f-score=0.66

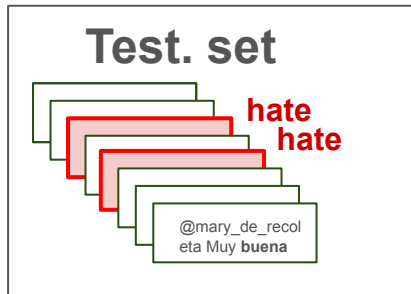


Y si quiero probar muchos modelos?

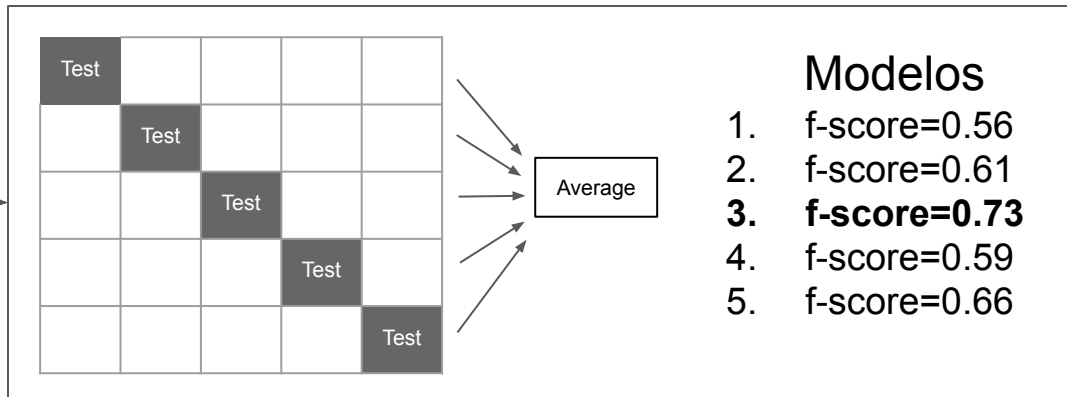
Training set
80%



Test. set
20%



Evalúo muchos modelos y elijo el mejor

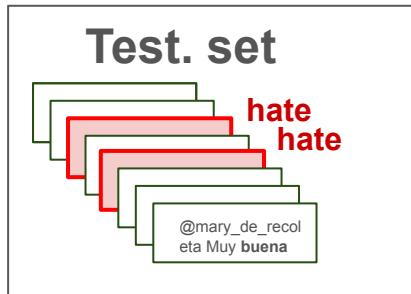


Cross Validation + Test set

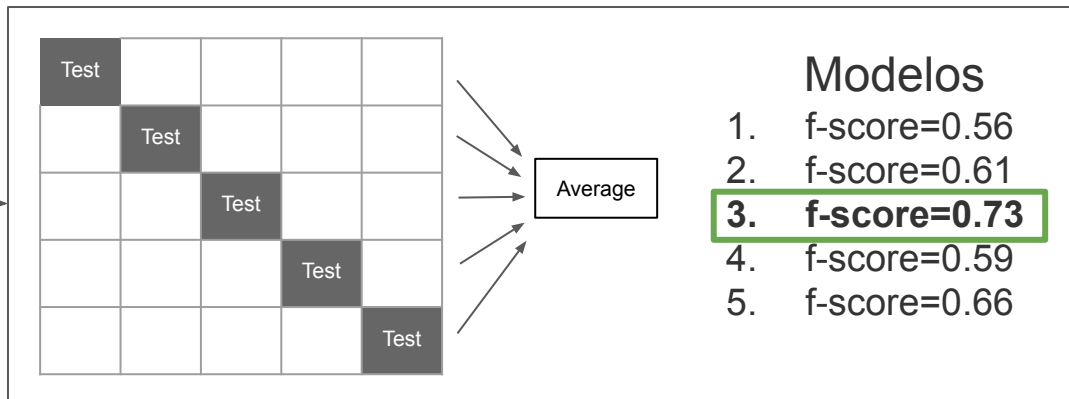
Training set
80%



Test. set
20%



Evalúo muchos modelos y elijo el mejor



Modelo 3 entrenado
con el Training set

f-score=0.70

FIN