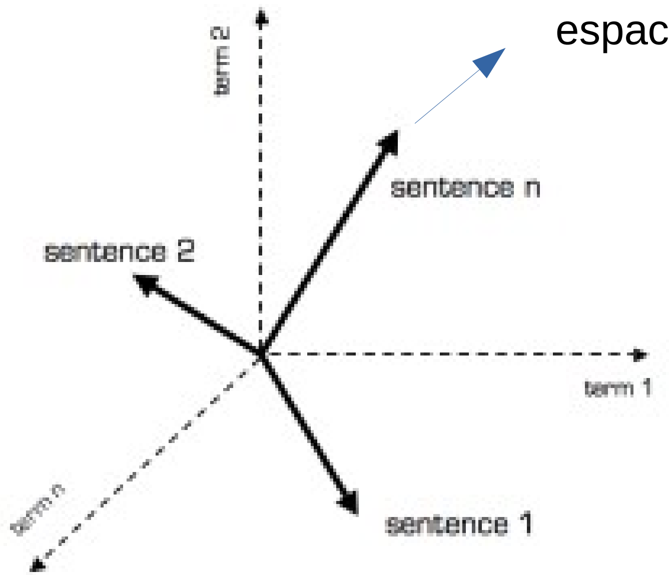
IIC 3800 Tópicos en CC NLP

https://github.com/marcelomendoza/IIC3800

- CLASIFICACIÓN DE DOCUMENTOS -

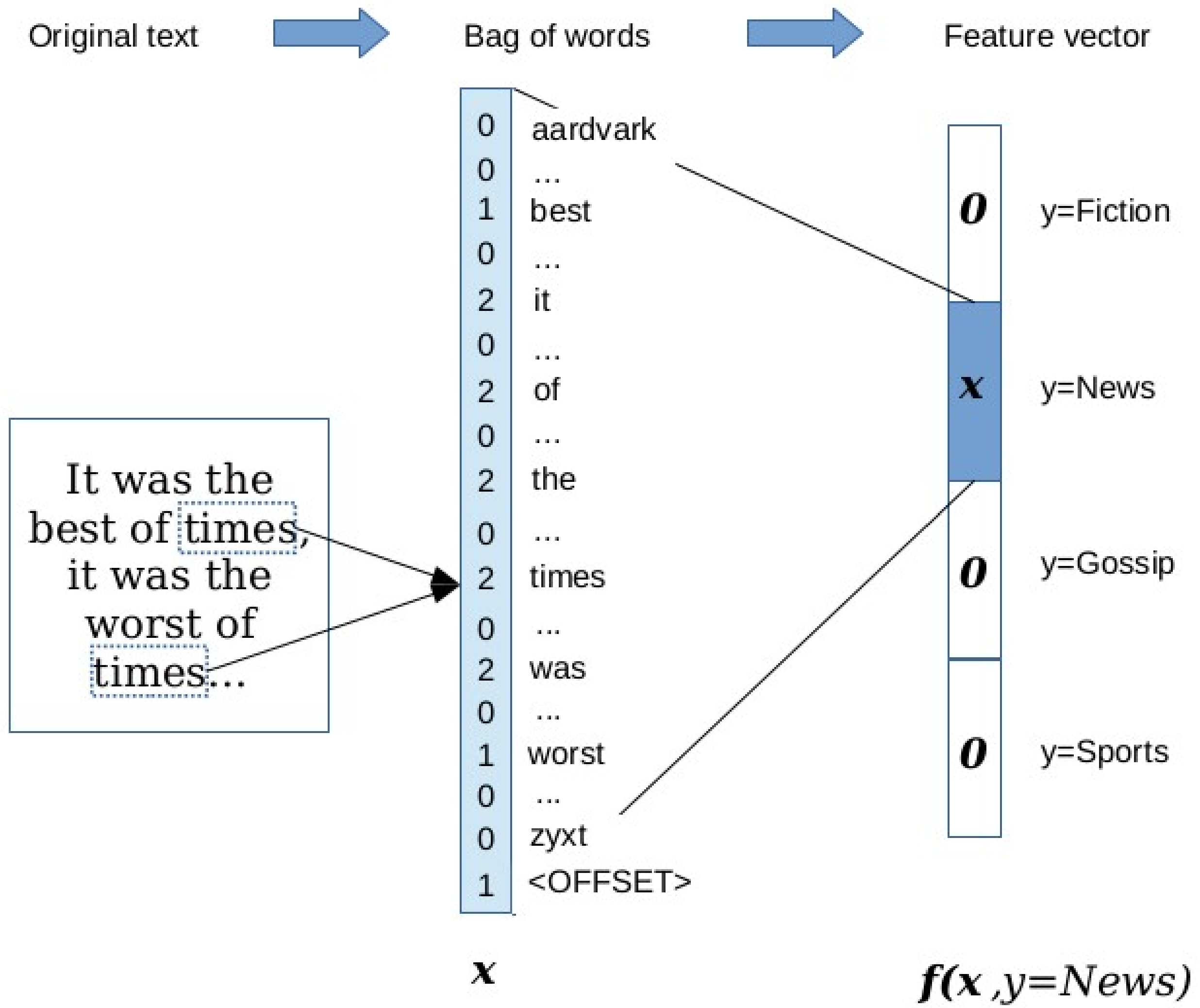
Vector-space model

Cada documento es un vector en el
espacio vectorial de términos



Cada término es una dimensión
en el espacio vectorial de términos

BOW



Original text
Bag of words
Feature vector
It was the best of times, it was the worst of times...
0 aardvark
0 ...
1 best
0 ...
2 it
0 ...
2 of
0 ...
2 the
0 ...
2 times
0 ...
2 was
0 ...
1 worst
0 ...
0 zyxt
1 <OFFSET>
x
0 y=Fiction
x y=News
0 y=Gossip
0 y=Sports
f(x ,y=News)

$f_{i,j}$: # occs. de ti en dj

$\max f_{l,j}$

N : # docs

n_i : # docs donde ti ocurre

Term scoring functions:

- Tf: $$Tf_{i,j} = \frac{f_{i,j}}{\max f_{l,j}}$$

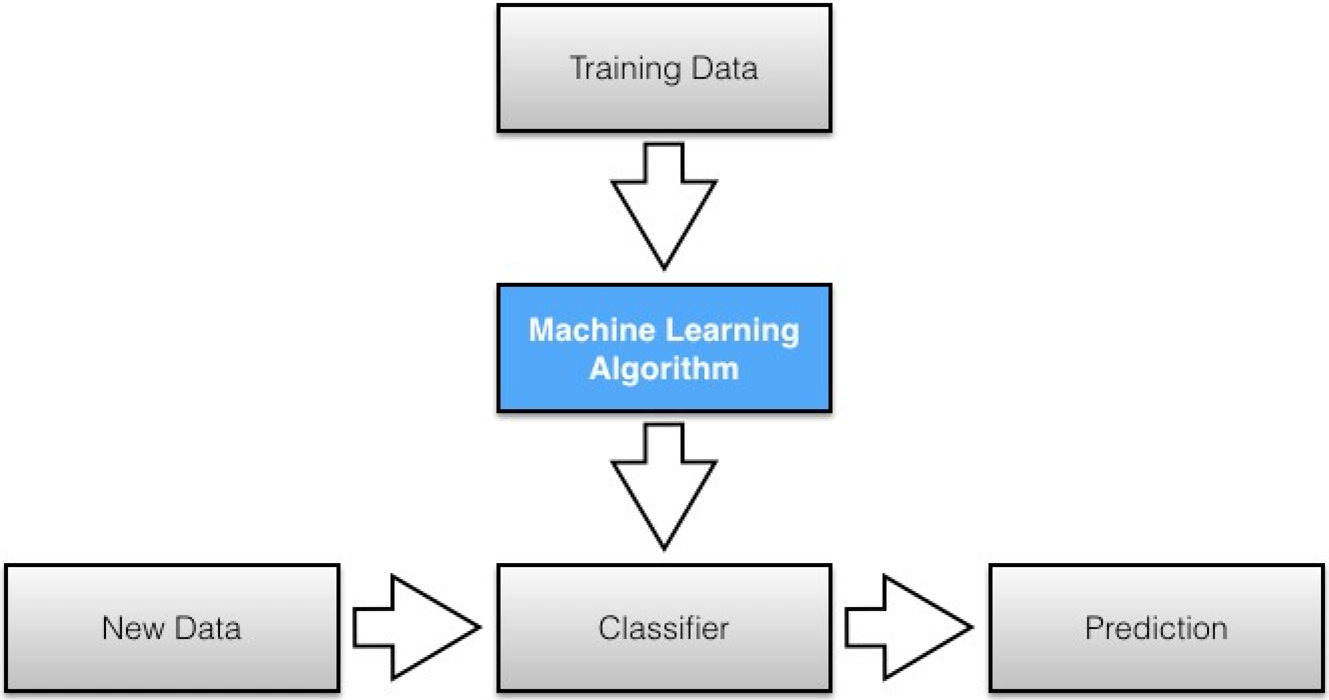
- Tf corregido: $$w_{i,j} = \begin{cases} 1 + \log_{10} f_{i,j} & \text{if } f_{i,j} > 0 \\ 0 & \text{e.t.o.c.} \end{cases}$$

- Idf: $$\mathrm{idf}_{t\,i} = \log_{10} \frac{N}{n_i}$$

- Tf-Idf (Salton): $$w_{i,j} = (1 + \log f_{l,j}) \cdot \log \frac{N}{n_i}$$

- Tf-Idf: $$w_{i,j} = \frac{f_{i,j}}{\max f_{l,j}} \cdot \log \frac{N}{n_i}$$

Síntesis. El enfoque de NLP (clásico)



Naive Bayes

Training (MLE)

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, \mathrm{p}(\boldsymbol{x}^{(1:N)}, y^{(1:N)}; \boldsymbol{\theta})$$

$$= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \prod_{i=1}^{N} \mathrm{p}(\boldsymbol{x}^{(i)}, y^{(i)}; \boldsymbol{\theta})$$

$$= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{i=1}^{N} \log \mathrm{p}(\boldsymbol{x}^{(i)}, y^{(i)}; \boldsymbol{\theta}).$$

Naive Bayes

Training (MLE)

$$\begin{aligned}\hat{\boldsymbol{\theta}} &= \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, \mathrm{p}(\boldsymbol{x}^{(1:N)}, y^{(1:N)}; \boldsymbol{\theta}) \\ &= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \prod_{i=1}^{N} \mathrm{p}(\boldsymbol{x}^{(i)}, y^{(i)}; \boldsymbol{\theta}) \\ &= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{i=1}^{N} \log \mathrm{p}(\boldsymbol{x}^{(i)}, y^{(i)}; \boldsymbol{\theta}).\end{aligned}$$

Generative process:

Algorithm 1 Generative process for the Naïve Bayes classification model

for Instance $i \in \{1, 2, \ldots, N\}$ **do**:
  Draw the label $y^{(i)} \sim \text{Categorical}(\boldsymbol{\mu})$;
  Draw the word counts $\boldsymbol{x}^{(i)} \mid y^{(i)} \sim \text{Multinomial}(\boldsymbol{\phi}_{y^{(i)}})$.

Naive Bayes

Condicionado a y

$$p_{\text{mult}}(\boldsymbol{x};\boldsymbol{\phi}) = B(\boldsymbol{x}) \prod_{j=1}^{V} \phi_j^{x_j}$$

$$B(\boldsymbol{x}) = \frac{\left(\sum_{j=1}^{V} x_j\right)!}{\prod_{j=1}^{V}(x_j!)}.$$

Naive Bayes

Condicionado a y

$$\mathrm{p}_{\text{mult}}(\boldsymbol{x};\boldsymbol{\phi}) = B(\boldsymbol{x}) \prod_{j=1}^{V} \phi_j^{x_j}$$

$$B(\boldsymbol{x}) = \frac{\left(\sum_{j=1}^{V} x_j\right)!}{\prod_{j=1}^{V}(x_j!)}.$$

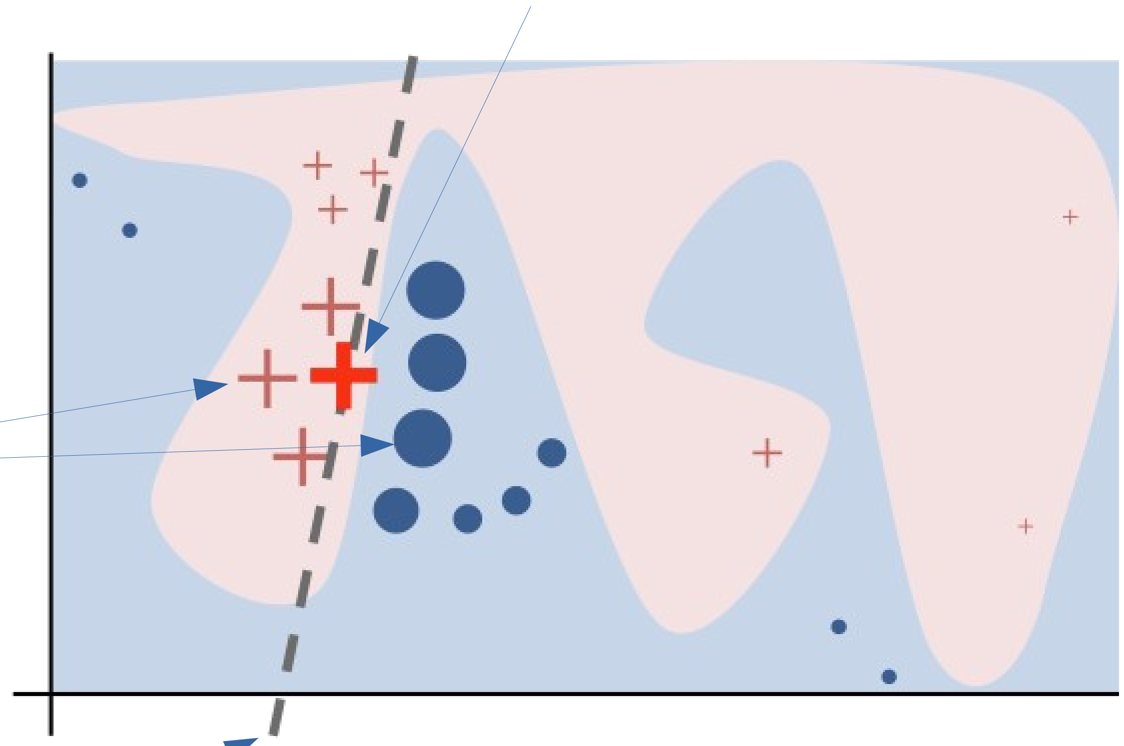
Predicción:

$$\hat{y} = \underset{y}{\operatorname{argmax}} \log \mathrm{p}(\boldsymbol{x}, y; \boldsymbol{\mu}, \boldsymbol{\phi})$$
$$= \underset{y}{\operatorname{argmax}} \log \mathrm{p}(\boldsymbol{x} \mid y; \boldsymbol{\phi}) + \log \mathrm{p}(y; \boldsymbol{\mu})$$

LIME (Local Interpretable Model-Agnostic Explanations)

Instancia a explicar

1. Se perturban (eliminan keywords) de la instancia



2. Se obtienen las predicciones sobre los ejemplos perturbados

3. Se construye un modelo lineal que aproxima al modelo original usando las predicciones (noisy label)

4. Se evalúa el modelo lineal y se obtiene la relevancia de cada keyword

https://github.com/marcelomendoza/IIC3800