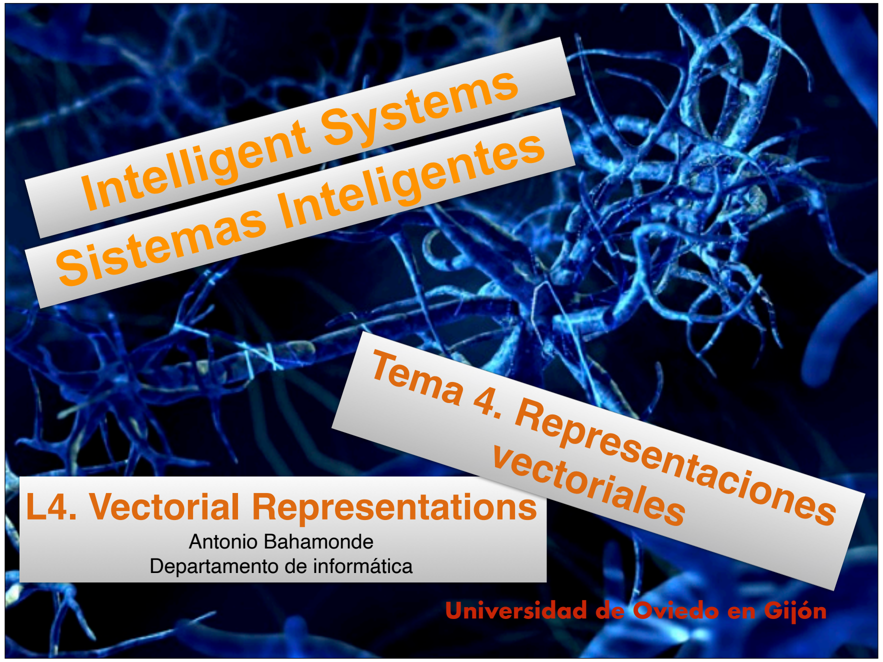


Intelligent Systems
Sistemas Inteligentes

Tema 4. Representaciones vectoriales

L4. Vectorial Representations

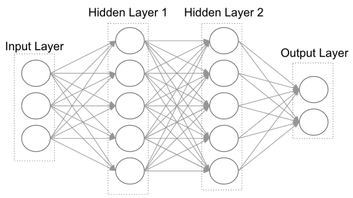
Antonio Bahamonde
Departamento de informática
Universidad de Oviedo en Gijón

Continuous/Discrete

The components of the **input** vector are the characteristics (features) of the examples (cases)

- Continuous
- Discrete
- Mixed

Vectors



$$\boldsymbol{h}_1 = g(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1)$$
$$\boldsymbol{h}_2 = g(\boldsymbol{W}_2\boldsymbol{h}_1 + \boldsymbol{b}_2)$$
$$f(\boldsymbol{x}) = g\Big(\boldsymbol{W}_2\big(g(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1)\big) + \boldsymbol{b}_2\Big)$$

Continuous

Consumer preferences of food products

A consumer likes a product better than another

Expert panels Measurements	Consumer panels
⇩	⇩
Product description	Preference Judgments

Features that entail consumer acceptance of beef meat from seven Spanish breeds

Consumer preferences of food products

- 101 animals
- 7 Spanish breeds
- slaugther weight:
  - 300–350 Kg. (light)
  - 530–560 Kg. (heavy)
- 3 ageing periods: 1, 7, 21 days (*)

(*) Sañudo, C.; Macie, E.S.; Olleta, J.L.; Villarroel, M.; Panea, B.; Albertí, P.. The effects of slaughter weight, breed type and ageing time on beef meat quality using two different texture devices. Meat Science, 66 (2004) 925–932

Consumer preferences of food products

Breed	% fat inter-muscular	Sub-cutaneous	% Bone	%muscl	% fat Intra-muscula
Asturiana Valles	5.01	0.89	15.77	78.32	1.01
Avileña	13.30	3.55	19.30	63.84	2.48
Morucha	13.71	3.93	18.68	63.68	2.59
Parda Alpina	10.27	2.70	20.53	66.50	1.96
Pirenaica	10.11	3.09	17.66	69.14	1.59
Retinta	13.76	4.52	20.40	61.33	2.12
Rubia Gallega	5.90	1.12	16.34	76.64	1.27

Consumer preferences of food products

Each **piece of meat** was described by:

- the weight of the animal,
- ageing time,
- breed,
- 6 physical features describing its texture and
- 12 sensory characteristics rated by 11 different experts (132 ratings).

The dataset has 307 different items.

Discrete

Discrete, n-hot: Text Representation

Dictionary/Corpus

	1	2	3	4	5	6	7	8
	el	un	caballo	gato	blanco	negro	agua	…
“un gato negro”	0	1	0	1	0	1	0	…

Bag of Words (BoW)
Vector Space Model (VSM)

Document classification

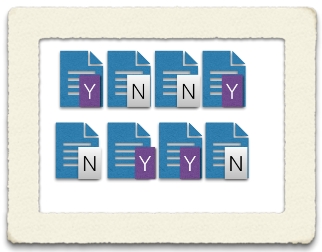
Is this a SPAM email?

Is this a document about economy?

Handmade procedure: keywords

```
If p is in D then return YES
        else return NO
```

Document classification



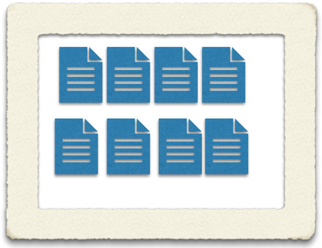
D-i = BoW vector

Inputs	Class
D-1	YES
D-2	No
...	...

Machine Learning algorithm

Learned procedure

Dealing with 2 Documents



Is this document relevant for this query?

Handmade procedure: all words have the same importance

Compute cos and return the documents with the highest value (see next)

Discrete, n-hot: Text Representation

Bag of words

	1	2	3	4	5	6	7	8
	el	un	caballo	gato	blanco	negro	agua	…
"un gato negro"	0	1	0	1	0	1	0	…
"el caballo blanco"	1	0	1	0	1	0	0	
"un caballo negro"	0	1	1	0	0	1	0	

$$
\begin{aligned}
&cos(\text{"un gato negro"}, \text{"un caballo negro"}) \\
&= cos((0,1,0,1,0,1,0),(0,1,1,0,0,1,0)) \\
&= \frac{<(0,1,0,1,0,1,0),(0,1,1,0,0,1,0)>}{\|(0,1,0,1,0,1,0)\| * \|(0,1,1,0,0,1,0)\|} \\
&= \frac{2}{\sqrt{3} * \sqrt{3}} = \frac{2}{3}
\end{aligned}
$$

Dealing with 2 Documents

All words have the same importance? Compute cos …

Some words are more relevant than others, but I do not know which ones

D-i, Q-j arfe BoW vectors

Inputs	Class
D-1, Q-1	YES
D-1, Q-2	No
D-2, Q-43	YES
…	…

Machine Learning algorithm

Learned procedure

Recommender Systems

Recommender Systems

	m1	m2	m3	m4
u1	👎	*	*	👍
u2	*	👍	👎	*
u3	*	*	👎	👍
u4	👍	👍	*	*

Users,

Movies

The goal is to fill this matrix ***according*** to the assessments of users about movies

Recommender Systems: Users and Items representations

Let's assume that we consider 3 features of movies

Action, Romantic, Science fiction

All of them can be quantified (say 0-100)

This can be used both for users and for movies

Recommender Systems: Users and Items representations

Affinity is a typical function to estimate if u is gonna like m-i.

affinity(u,m) = <vec(u), vec(m)>

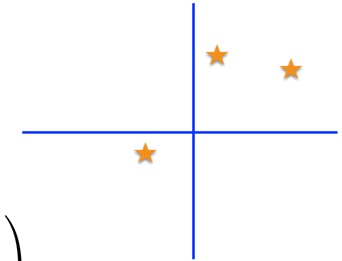
Affinity	m1 = (0, 100, 0)	m2 = (90,1, 100)
u = (100, 1, 80)	0*100+100*1+0*80 = **100**	90*100+1*1+100*80 = **17001**

Recommender Systems: Users and Items representations

How can we find out the features of users and movies?

1-hot embedding

$$\boldsymbol{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$



$$\boldsymbol{Ap} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} p_1 + \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} p_2 + \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} p_3$$

Recommender Systems: Users and Items representations

How can we find out the features of users and movies?

affinity(u,m) = <vec(u), vec(m)> = <Wu,Vm>

W = 3 rows, one column for each user
V = 3 rows, one column for each movie

We must learn the matrices W and V

Recommender Systems: Users and Items representations

affinity(u,m) = <vec(u), vec(m)> = <Wu,Vm>

$$vec(m1) = Vm1 = \begin{pmatrix} 0 & 90 \\ 100 & 1 \\ 0 & 100 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$vec(m2) = Vm2 = \begin{pmatrix} 0 & 90 \\ 100 & 1 \\ 0 & 100 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$vec(u) = Wu = \begin{pmatrix} 100 & \dots \\ 1 & \dots \\ 80 & \dots \end{pmatrix} \begin{pmatrix} 1 \\ \dots \end{pmatrix}$$

Affinity	m1 = (0, 100, 0)	m2 = (90,1, 100)
u = (100, 1, 80)	0*100+100*1+0*80 = **100**	90*100+1*1+100*80 = **17001**

Recommender Systems

	m1	m2	m3	m4
u1	👎	*	*	👍
u2	*	👍	👎	*
u3	*	*	👎	👍
u4	👍	👍	*	*

Users,

Movies

The goal is to fill this matrix ***according*** to the assessments of users about movies

sign (affinity(u, m) + b) = sign (<Wu,Vm>+ b)

Matrix factorization

$$< Wu, Vm > = (Wu)^T(Vm) = u^T(W^TV)m$$

$$W^TV = X$$

$$u^TXm = \sum_{i,j} x_{ij}u_im_j$$

Additional examples

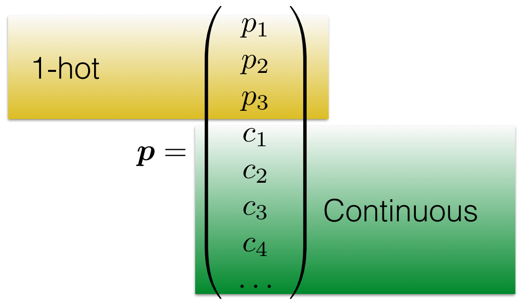
1-hot & n-hot examples

Index	1	2	3	4	5	6	7	8
x1	1	0	0	0	0	0	0	0
x2	0	1	0	0	0	0	0	0
x3	0	0	1	0	0	0	0	0
...								
x8	0	0	0	0	0	0	0	1

```
unknown | Action | Adventure | Animation | Children's | Comedy
| Crime | Documentary | Drama | Fantasy | Film-Noir | Horror |
Musical | Mystery | Romance | Sci-Fi | Thriller | War |
Western |

These are the genres, a 1 indicates the movie is of that
genre, a 0 indicates it is not; movies can be in several
genres at once.
```

Mixed



Consumer preferences of food products

In addition to feature descriptions of consumers and items, we added a **binary identification**

each object (consumer or item) includes in its description a vector with all components 0 but the one (index the ordinal of the object) that has value 1.

Example.

If we have only 3 consumers,

consumer1 =(1,0,0,sex1,age1,job1)^T,

consumer2 =(0,1,0,sex2,age2,job2)^T,

consumer3 =(0,0,1,sex3,age3,job3)^T.

Automatic assessment of open response answers (Peer assessment)

	a1	a2	a3	a4
g1	4	*	*	9
g2	*	7	4	*
g3	*	*	3	8
g4	5	5	*	*

$\mathcal{G}$ graders,

$\mathcal{A}$ answers

$$M(g, a) \in [0, 10]$$

The goal is to fill this matrix ***according*** to the assessments of graders

Automatic assessment of open response answers (Peer assessment)

vectorial representation

- the simple *binary* codification

$$\boldsymbol{g} = (0, \ldots, 1, \ldots, 0)^T \in \mathbb{R}^{|\mathcal{G}|}$$
$$\boldsymbol{a} = (0, \ldots, 1, \ldots, 0)^T \in \mathbb{R}^{|\mathcal{A}|}$$

collaborative filter

Automatic assessment of open response answers (Peer assessment)

Add **content-based** codification in answers

$$\boldsymbol{a} = (\boxed{0, \ldots, 1, \ldots, 0}; \boxed{w_{a,1}, \ldots, w_{a,|Corpus|}})^T \in \mathbb{R}^{|rep(\mathcal{A}))|}$$

where

$$w_{a,i} = \begin{cases} 1 & w_i \in \boldsymbol{a} \\ 0 & \text{otherwise} \end{cases}$$

1-hot

BoW