

# Practica 5

## SVM y medidas de calidad



Clasificación de Patrones

## ÍNDICE

- 5.1 Clasificadores SVM
- 5.2 Validación de parámetros
- 5.3 Base de datos SPAM
- 5.4 Medidas de calidad



Clasificación de Patrones

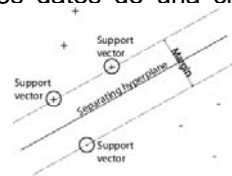
2

## 5.1 SVM

Clasificador SVM lineal:

**Datos separables linealmente:** Un clasificador SVM busca el mejor hiperplano separador que separa los datos de una clase respecto a la otra. Función a minimizar:

$$L = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^N \alpha_i \left( y_i (\mathbf{w}^T \mathbf{x}_i + w_0) - 1 \right)$$



Se obtiene un problema convexo en las variables  $\alpha_i$  que puede resolverse usando software estándar de optimización:

$$L = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^N \alpha_i \alpha_k y_i y_k \mathbf{x}_i^T \mathbf{x}_k \quad \text{subject to} \quad \begin{cases} \alpha_i \geq 0 \\ \sum_{i=1}^N \alpha_i y_i = 0 \end{cases}$$

La clasificación del vector  $\mathbf{x}$  se reduce a:

$$\hat{y} = \text{sign}(g(\mathbf{x})) = \text{sign}(\mathbf{w}^T \mathbf{x} + w_0) = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i \mathbf{x}_i^T \mathbf{x} + w_0\right)$$



Clasificación de Patrones

3

## 5.1 SVM

Clasificador SVM lineal:

**Datos NO separables linealmente:** No existe un hiperplano que pueda separar las dos clases.

$$y_i (\mathbf{w}^T \mathbf{x}_i + w_0) \geq 1 - \xi_i \quad i = 1, \dots, N$$

se introduce una penalización para los valores no-nulos de  $\xi$ :

$$L = \frac{1}{2} \|\mathbf{w}\|^2 + P \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i \left( y_i (\mathbf{w}^T \mathbf{x}_i + w_0) - (1 - \xi_i) \right) - \sum_{i=1}^N \beta_i \xi_i$$

**El parámetro de penalización P es un parámetro a validar**

Un valor alto de P conduce a soluciones sobre entrenadas (Overfitting)



Clasificación de Patrones

4

## 5.1 SVM

**Clasificador SVM NO lineal:** Uso de funciones de Kernel

$$L = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^N \alpha_i \alpha_k y_i y_k K(\mathbf{x}_i, \mathbf{x}_k) \quad \text{subject to} \quad \begin{cases} 0 \leq \alpha_i \leq P \\ \sum_{i=1}^N \alpha_i y_i = 0 \end{cases}$$

Ejemplo kernel gaussiano:

$$K(\mathbf{x}_i, \mathbf{x}_k) = \exp\left(-\frac{1}{\sigma^2} \|\mathbf{x}_i - \mathbf{x}_k\|^2\right)$$

La clasificación del vector  $\mathbf{x}$  se reduce a:

$$\hat{y} = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + w_0\right)$$

**El parámetro de penalización  $\sigma^2$  es un parámetro a validar**



Clasificación de Patrones

5

## 5.2 Validación de Parámetros $P$ y $\sigma^2$

División de la BD en 3 subconjuntos:

Train:  $X_{\text{train}}$ ,  $Labels_{\text{train}}$ , 60% 70%, etc,...

Validación:  $X_{\text{val}}$ ,  $Labels_{\text{val}}$ , 20% 15%

Test:  $X_{\text{test}}$ ,  $Labels_{\text{test}}$ , 20% 20% 15%

**Validación de los parámetros (fuerza bruta):**

Para cada par de valores a probar ( $P, \sigma^2$ )

- Se diseña el clasificador SVM con la BD de train
- Se mide el error de validación.

Posteriormente:

- Se selecciona el clasificador correspondiente al menor error de validación.
- Se mide el error de test con el clasificador seleccionado



Clasificación de Patrones

6

### 5.3 Base de datos SPAM

Cada vector corresponde a un e-mail recibido y es de  $d=57$  características.

Objetivo: Construcción de un filtro de spam personalizado

#### Parámetros de la base de datos:

- Total de vectores de la base de datos : 4601 (1813 spam y 2788 mail).
- Clases: 2
- Las características corresponden al número de veces que aparece una determinada palabra o carácter en e-mail
- Objetivos: Construcción de un filtro de spam personalizado.

**Pre-procesado de la base de datos** (Agiliza convergencia algoritmo de programación con restricciones, previene overfitting):

- Cuantificación binaria de las características 1:54
- Supresión de las características 55, 56 y 57



Clasificación de Patrones

7

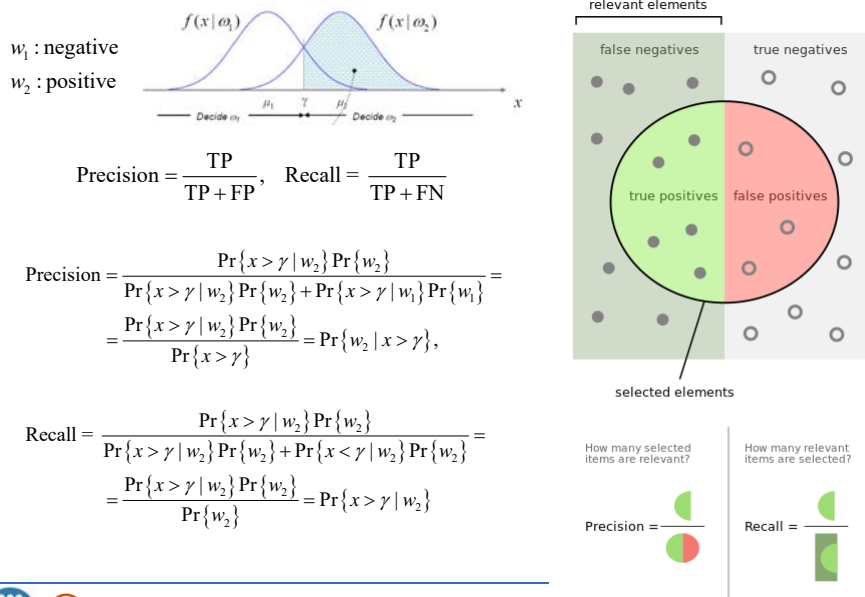
### 5.3 BD SPAM: Content of a feature vector

Number	Feature	Number	Feature
1	word_freq_make: continuous.	30	word_freq_labs: continuous.
2	word_freq_address: continuous.	31	word_freq_telnet: continuous.
3	word_freq_all: continuous.	32	word_freq_857: continuous.
4	word_freq_3d: continuous.	33	word_freq_data: continuous.
5	word_freq_our: continuous.	34	word_freq_415: continuous.
6	word_freq_over: continuous.	35	word_freq_85: continuous.
7	word_freq_remove: continuous.	36	word_freq_technology: continuous.
8	word_freq_internet: continuous.	37	word_freq_1999: continuous.
9	word_freq_order: continuous.	38	word_freq_parts: continuous.
10	word_freq_mail: continuous.	39	word_freq_pm: continuous.
11	word_freq_receive: continuous.	40	word_freq_direct: continuous.
12	word_freq_will: continuous.	41	word_freq_cs: continuous.
13	word_freq_people: continuous.	42	word_freq_meeting: continuous.
14	word_freq_report: continuous.	43	word_freq_original: continuous.
15	word_freq_addresses: continuous.	44	word_freq_project: continuous.
16	word_freq_free: continuous.	45	word_freq_re: continuous.
17	word_freq_business: continuous.	46	word_freq_edu: continuous.
18	word_freq_email: continuous.	47	word_freq_table: continuous.
19	word_freq_you: continuous.	48	word_freq_conference: continuous.
20	word_freq_credit: continuous.	49	char_freq_:: continuous.
21	word_freq_your: continuous.	50	char_freq (: continuous.
22	word_freq_font: continuous.	51	char_freq [ continuous.
23	word_freq_000: continuous.	52	char_freq !: continuous.
24	word_freq_money: continuous.	53	char_freq \$: continuous.
25	word_freq_hp: continuous.	54	char_freq #: continuous.
26	word_freq_hpl: continuous.	55	capital_run_length_average: continu
27	word_freq_george: continuous.	56	capital_run_length_longest: contin
28	word_freq_650: continuous.	57	capital_run_length_total: continu
29	word_freq_lab: continuous.		



Clasificación de Patrones

## 5.4 Medidas de calidad: Precision and recall



Clasificación de Patrones

9

## 5.4 Medidas de calidad:

Precision, Recall=Sensitivity, Specificity,

$$P = \frac{\text{Clasificado SPAM correctamente}}{\text{Clasificado SPAM}} = \frac{TP}{TP + FP}$$

$$S = \frac{\text{Clasificado SPAM correctamente}}{\text{\# total de SPAM}} = \frac{TP}{TP + FN}$$

$$Es = \frac{\text{Clasificado MAIL correctamente}}{\text{\# total de MAIL}} = \frac{TN}{TN + FP}$$

**F score:** A measure that combines precision and recall is the harmonic mean of precision and recall, the traditional F-measure or balanced F-score:

$$F\_score = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Prior measures:

$$\Pr\{\text{clase=SPAM}\} = \frac{\text{\# elementos SPAM de la BD de test}}{\text{\# elementos de la BD de test}}$$

$$\Pr\{\text{clase=MAIL}\} = \frac{\text{\# elementos MAIL de la BD de test}}{\text{\# elementos de la BD de test}}$$



Clasificación de Patrones

10