

Reconocimiento de emociones en señales de voz

Alumnos: Leonardo Arato

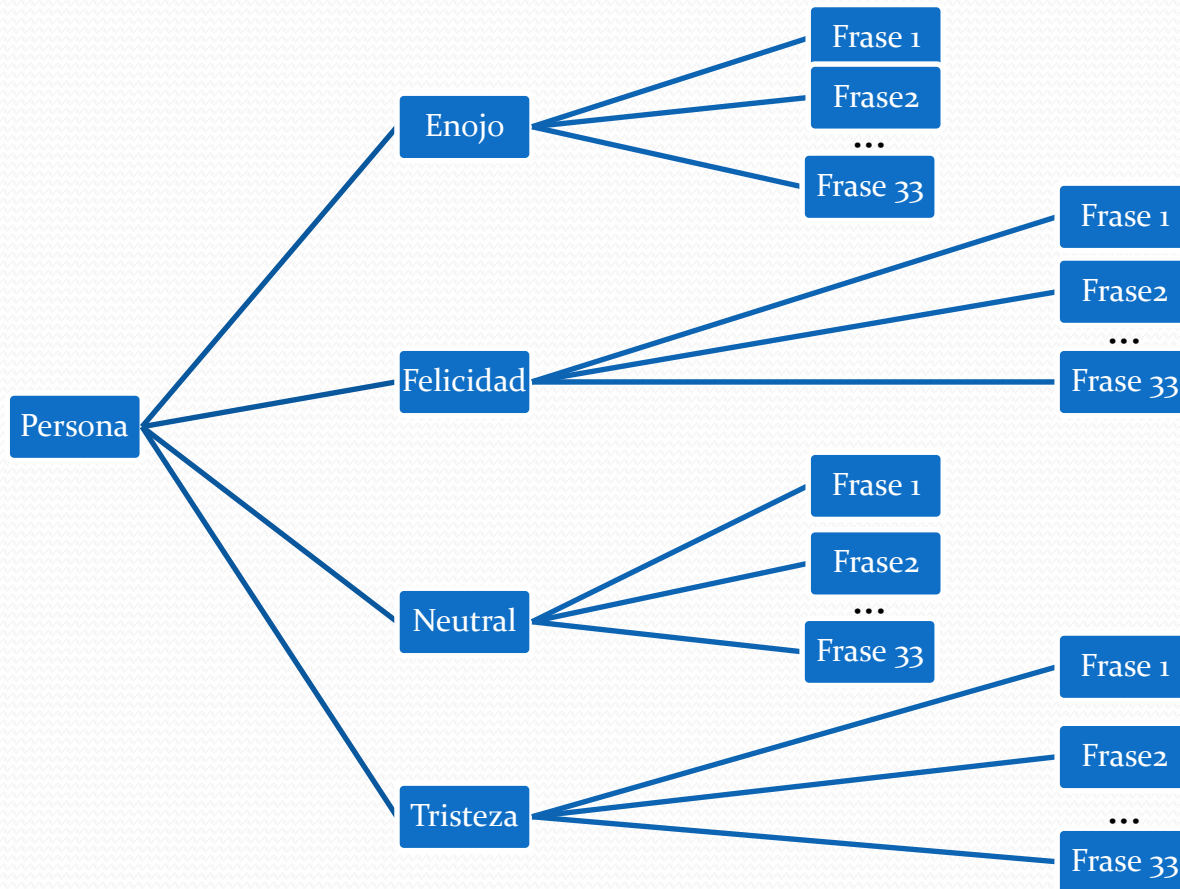
Nahuel Boutet

Tutor: Dr. H. Leonardo Rufiner

Motivaciones

- Humano-Máquina
- Reconocimiento de llamadas de emergencia.
- Call Centers.
- Detección de enfermedades psicológicas

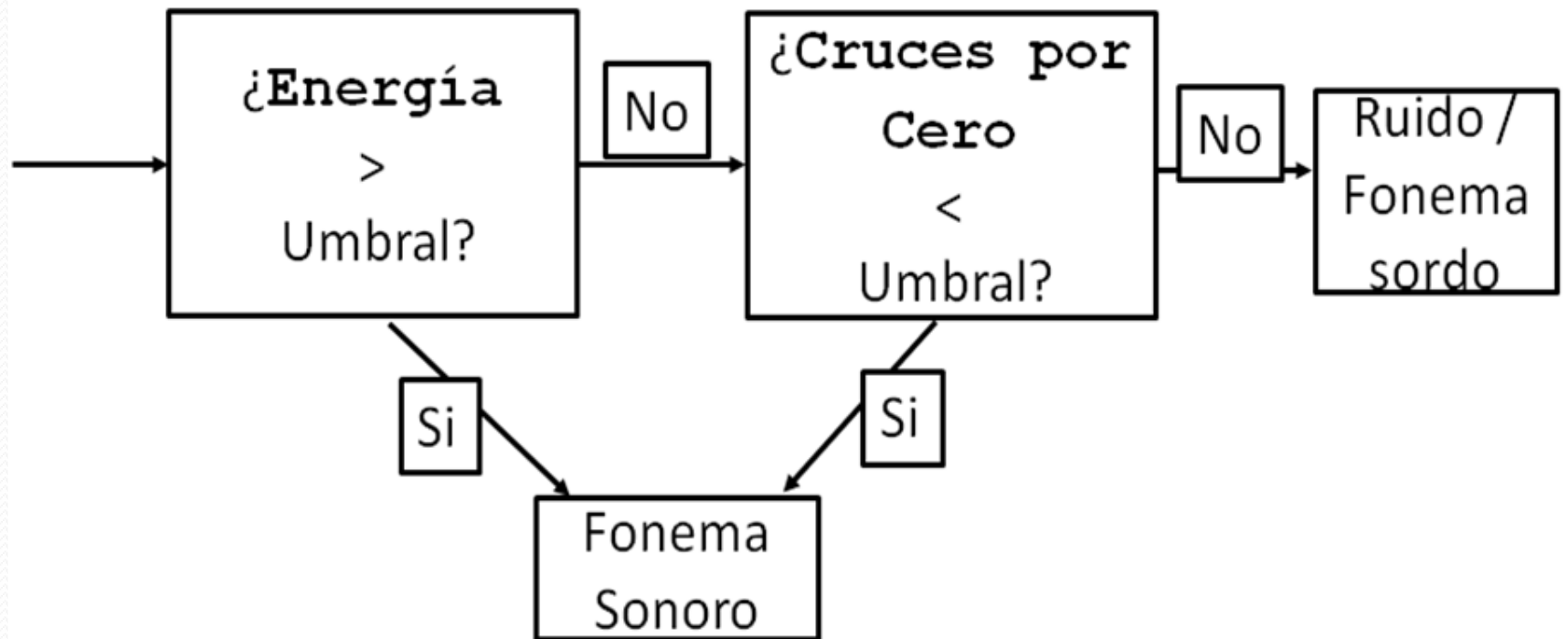
Base de datos



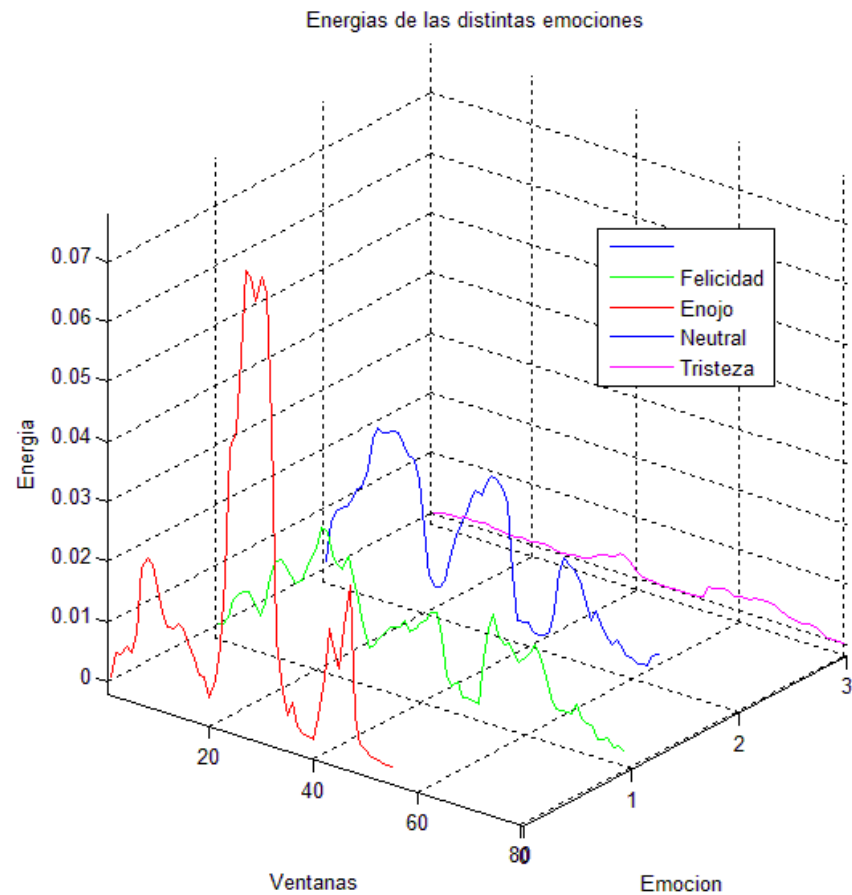
Tratamiento de señales

- Ventana de Hamming
- Fonemas Sonoros
- Energía
- Frecuencia Fundamental
- Frecuencias Formantes

Tratamiento de señales

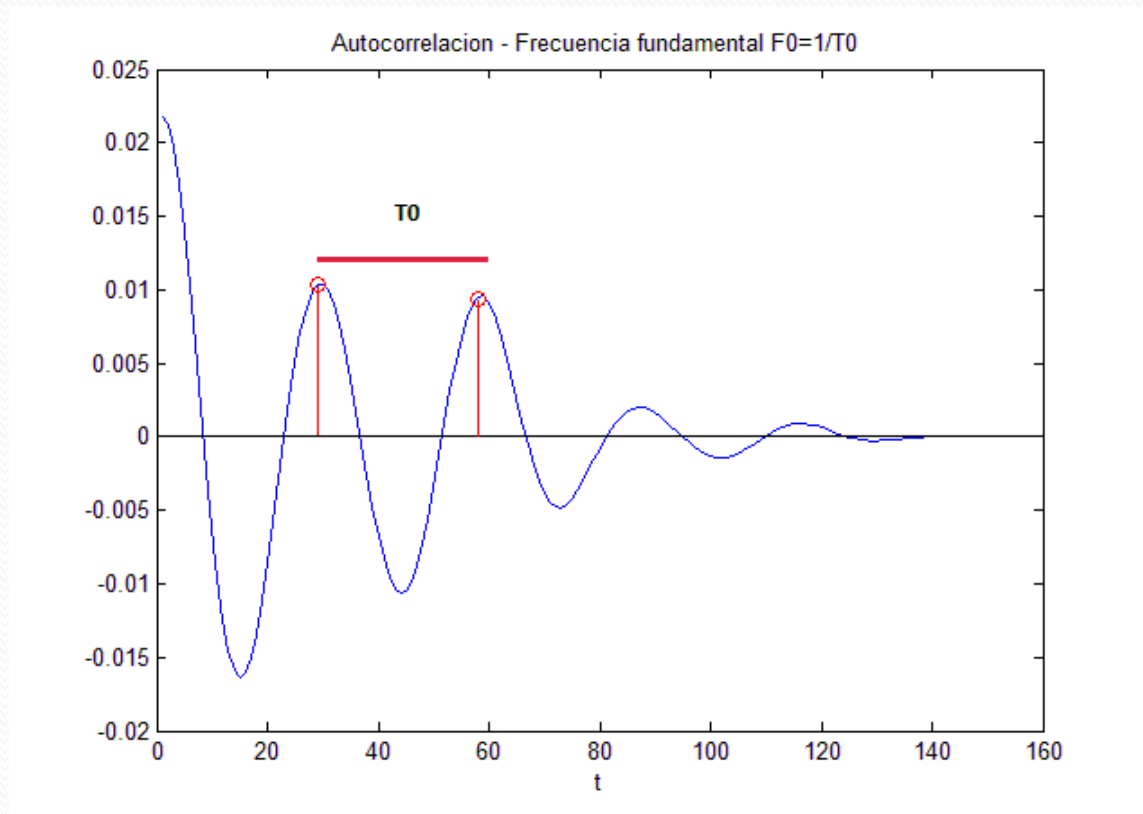


Energías



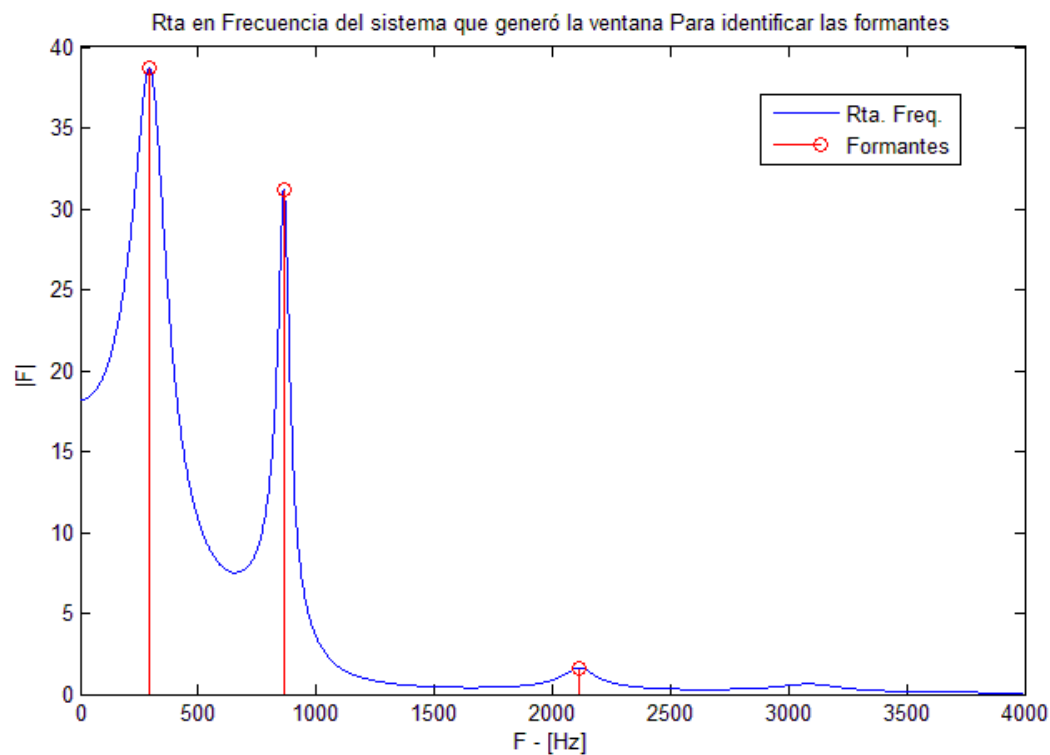
Frecuencia Fundamental

- Auto-Correlación.



Frecuencias Formantes

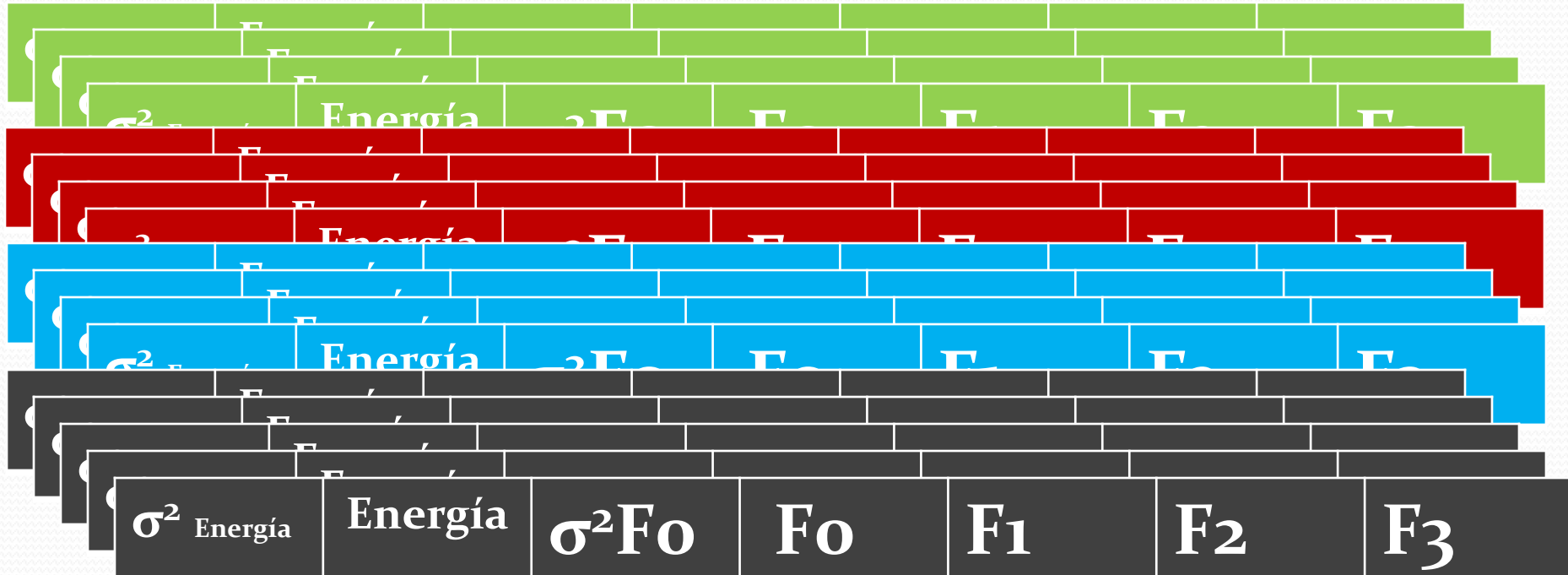
- LPC.
- Respuesta en Frecuencia del sistema.



Vector Característico

σ^2 Energía	Energía	$\sigma^2 F_0$	F_0	F_1	F_2	F_3
--------------------	---------	----------------	-------	-------	-------	-------

Conjunto de vectores característicos por emoción



Distancia Euclídea

Vc=	σ^2 Energía	Energía	$\sigma^2 F_0$	F_0	F_1	F_2	F_3
p=	0.0953	0.0938	17.3258	152.5	488.9	996.4	2156
q=	0.0469	0.0493	18.27	156.8	487.4	1050.6	2300.2
p-q=	0.0484	0.0445	-0.9442	-4.3	1.5	-54.2	-144.2

$$d_E(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}.$$

Distancia
Euclídea

154.11

Distancia de Mahalanobis

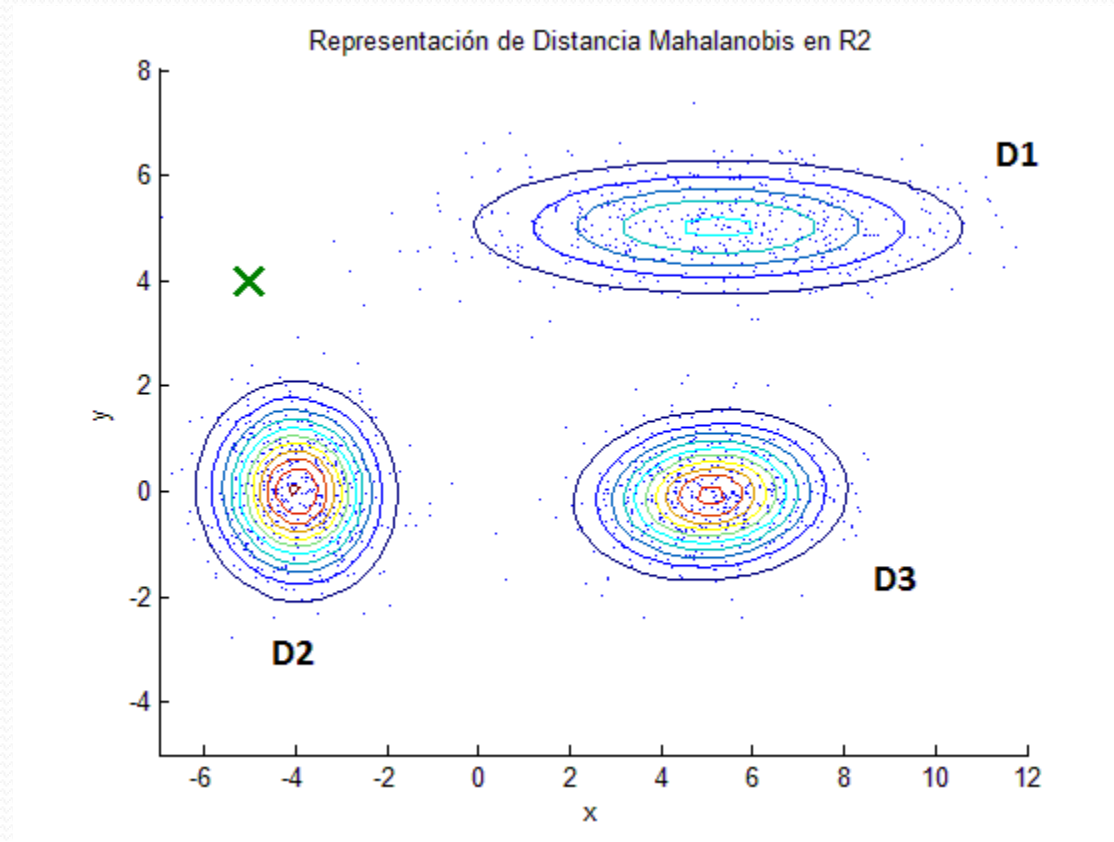
Pondera según la varianza: las variables con menos varianza tendrán más importancia que las de mayor varianza. La expresión quedaría:

$$d_2(\vec{x}_1, \vec{x}_2) = \sqrt{\left(\frac{(x_{11} - x_{12})}{\sigma_1}\right)^2 + \left(\frac{(x_{21} - x_{22})}{\sigma_2}\right)^2}$$

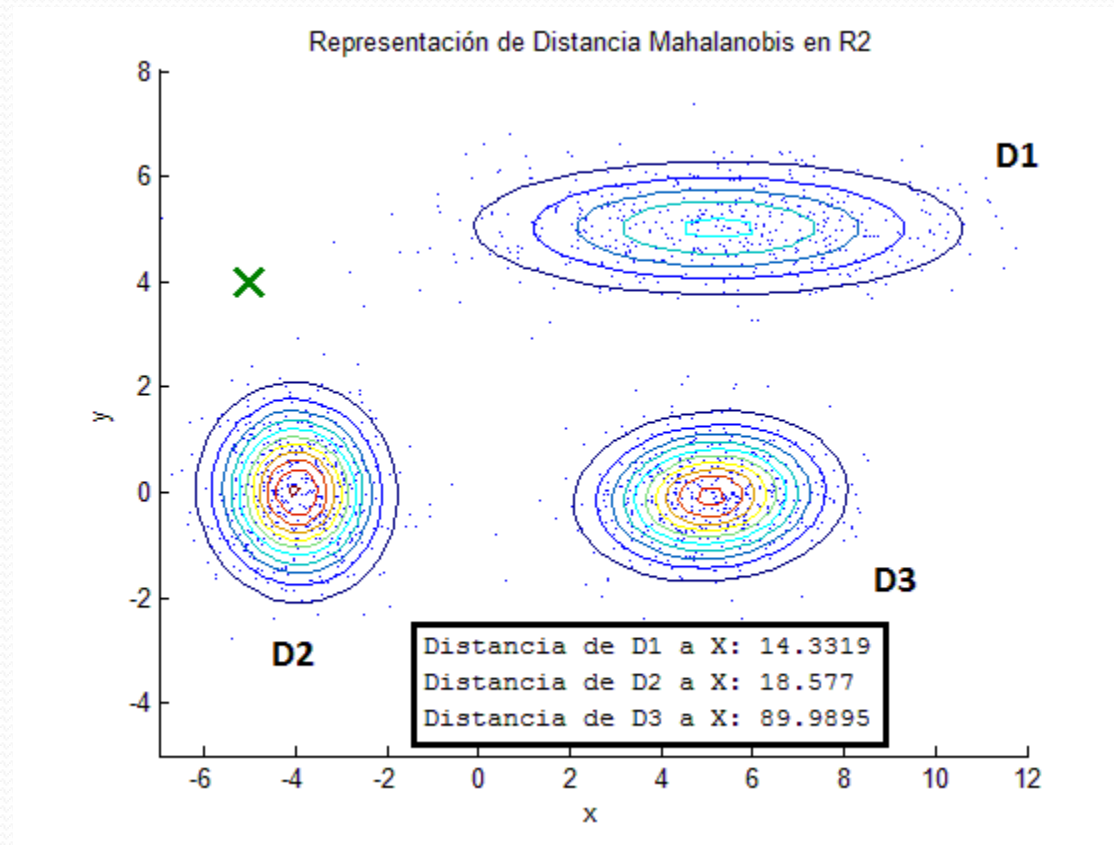
Fórmula general:

$$d_m(\vec{x}_1, \vec{x}_2) = \sqrt{(\vec{x}_1 - \vec{x}_2)^T \Sigma^{-1} (\vec{x}_1 - \vec{x}_2)}$$

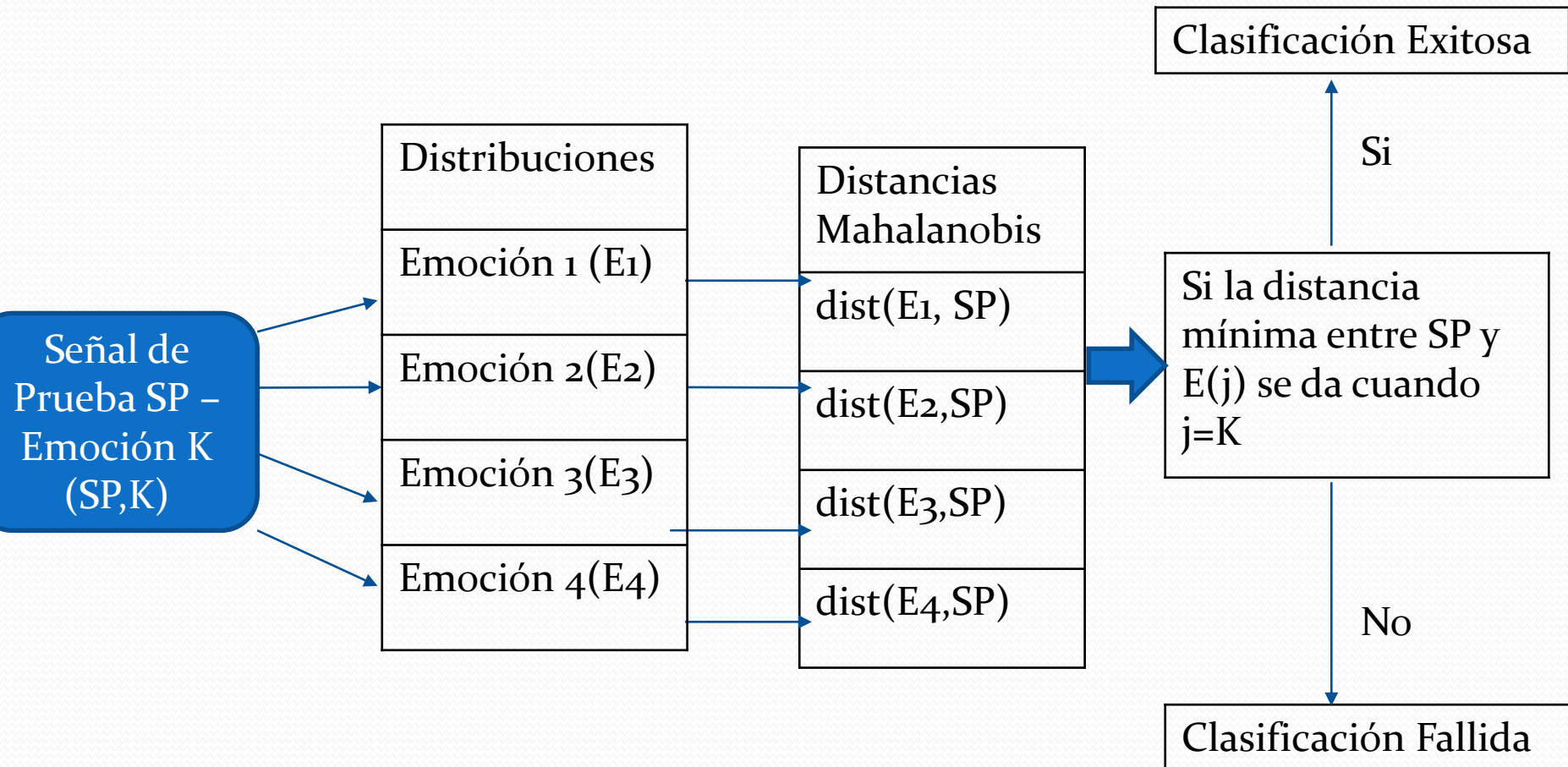
Distancia de Mahalanobis



Distancia de Mahalanobis



Clasificación



Resultados-

Ventana de 160 muestras	Hombres	Mujeres
<i>Emoción</i>	<i>Porcentaje de Acierto</i>	<i>Porcentaje de Acierto</i>
Enojo	98.21 %	84.37 %
Felicidad	73.21 %	93.75 %
Neutral	76.78 %	84.37 %
Tristeza	78.57 %	75.00 %

Ventana de 256 muestras	Hombres	Mujeres
<i>Emoción</i>	<i>Porcentaje de Acierto</i>	<i>Porcentaje de Acierto</i>
Enojo	100 %	84.37 %
Felicidad	73.21 %	93.75 %
Neutral	76.57 %	84.37 %
Tristeza	80.35 %	84.37 %

Señal + Ruido

SNR = 20dB	Hombres	Mujeres
<i>Emoción</i>	<i>Porcentaje de Acierto</i>	<i>Porcentaje de Acierto</i>
Enojo	98.21 %	81.25 %
Felicidad	53.57 %	87.5 %
Neutral	73.21 %	50.00 %
Tristeza	71.42 %	78.12 %

SNR = 0dB	Hombres	Mujeres
<i>Emoción</i>	<i>Porcentaje de Acierto</i>	<i>Porcentaje de Acierto</i>
Enojo	77.5 %	74.37 %
Felicidad	26.78 %	18.75 %
Neutral	0.0 %	0.0 %
Tristeza	0.0 %	0.0 %



Fin