Introducción a la Inteligencia Artificial
Facultad de Ingeniería
Universidad de Buenos Aires

Índice

Clase 7

Aprendizaje no supervisado

Machine Learning Supervisado	Machine Learning no Supervisado
Proceso aleatorio $\bar{X}, y$	Proceso aleatorio $\bar{X}$
¿$f_{y/\bar{x}}(y\|\bar{x})$? $\longrightarrow$ Bayes y M.V.	¿$f_{\bar{x}}(\bar{x})$? $\longrightarrow$ Bayes y M.V.
Inferencias, predicciones	Clusterización, Reducción Dimensionalidad

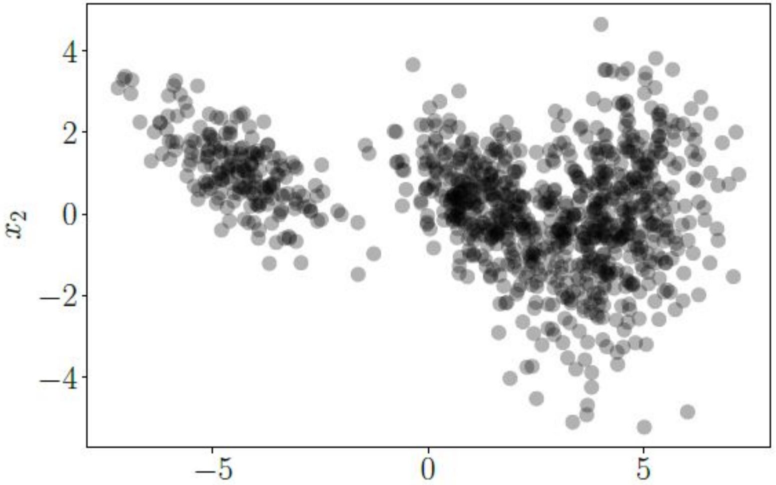
Aprendizaje no supervisado

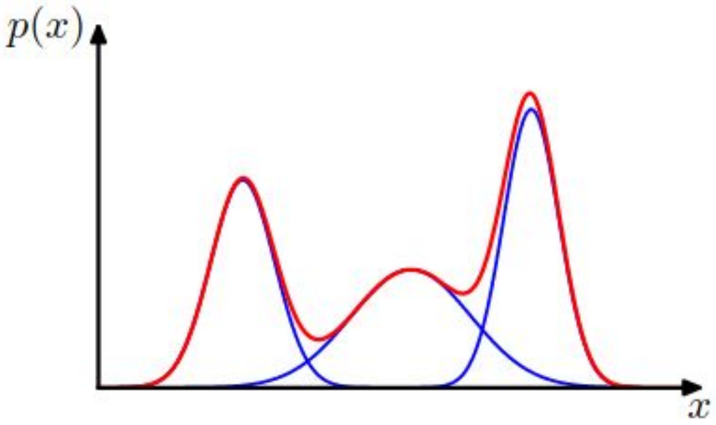
Aplicaciones Generales

- Data Mining
- Pattern Recognition
- Statistical Analysis

Aplicaciones Específicas

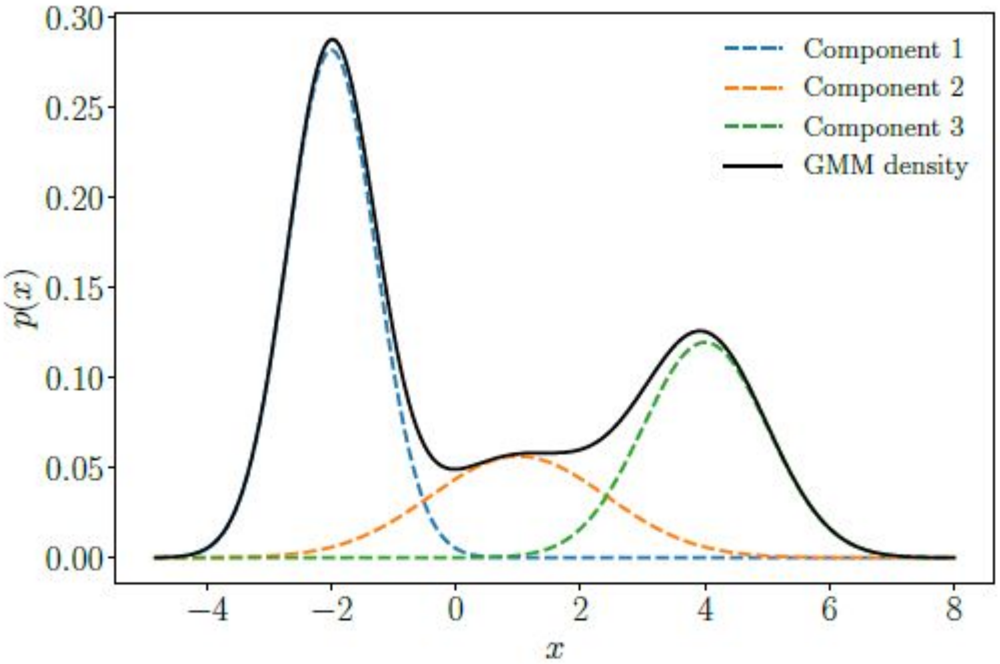
- Density Estimation
- Clustering
- Anomaly Detection
- Object Tracking
- Speech Feature Extraction





Gaussian Mixture Models

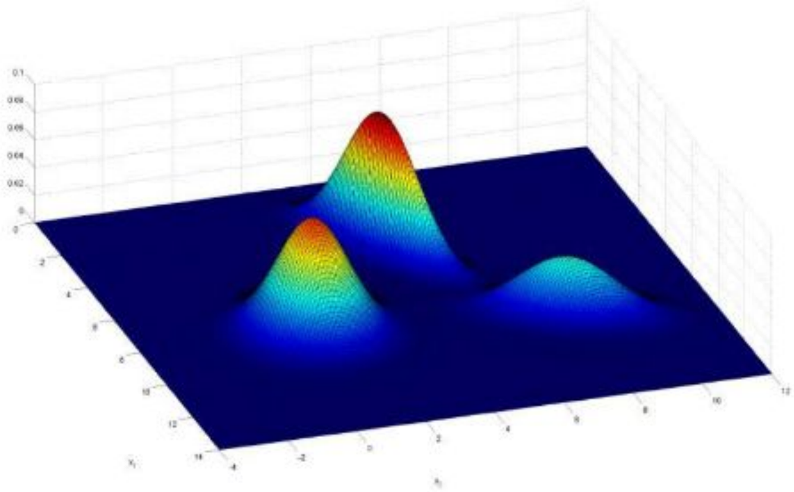
Formulación

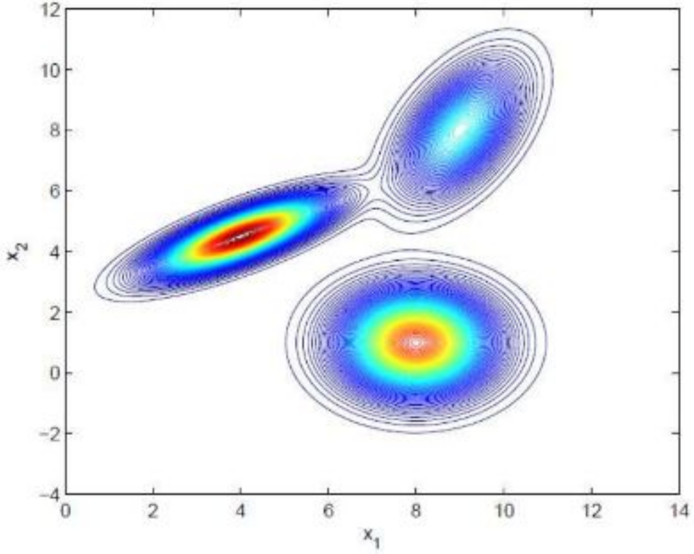


$$p(x \mid \boldsymbol{\theta}) = 0.5\mathcal{N}\left(x \mid -2, \tfrac{1}{2}\right) + 0.2\mathcal{N}\left(x \mid 1, 2\right) + 0.3\mathcal{N}\left(x \mid 4, 1\right)$$

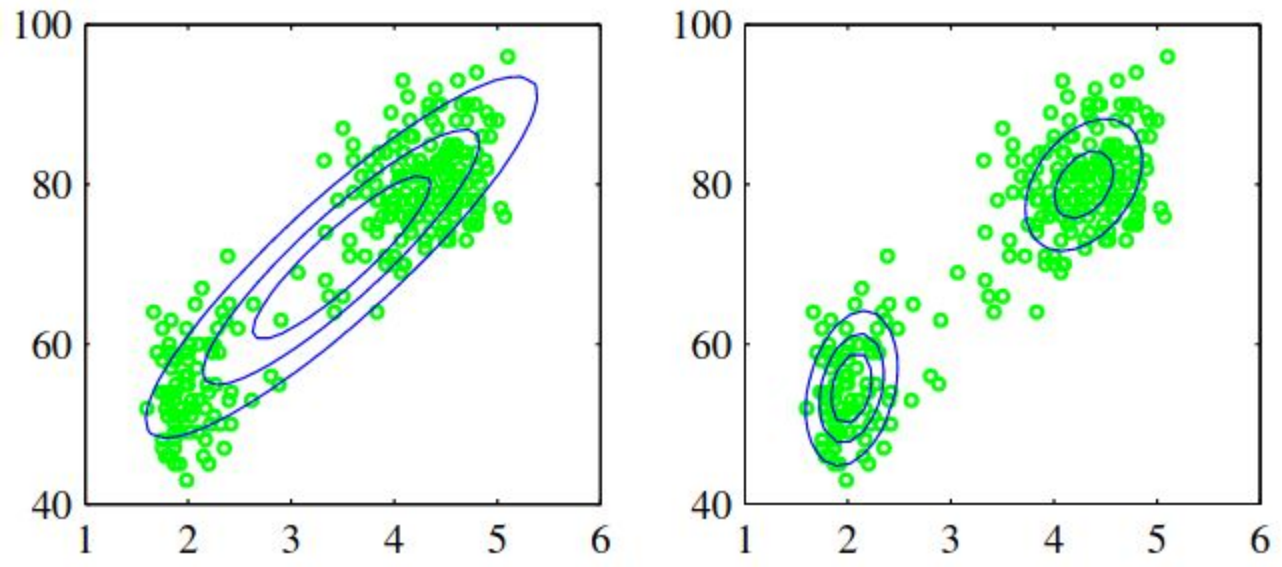
Formulación

$$p(x) = \underbrace{0.3}_{\pi_1}\mathcal{N}\left(x \mid \underbrace{\begin{pmatrix}4\\4.5\end{pmatrix}}_{\mu_1}, \underbrace{\begin{pmatrix}1.2 & 0.6\\0.6 & 0.5\end{pmatrix}}_{\Sigma_1}\right) + \underbrace{0.5}_{\pi_2}\mathcal{N}\left(x \mid \underbrace{\begin{pmatrix}8\\1\end{pmatrix}}_{\mu_2}, \underbrace{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}}_{\Sigma_2}\right) + \underbrace{0.2}_{\pi_3}\mathcal{N}\left(x \mid \underbrace{\begin{pmatrix}9\\8\end{pmatrix}}_{\mu_3}, \underbrace{\begin{pmatrix}0.6 & 0.5\\0.5 & 1.5\end{pmatrix}}_{\Sigma_3}\right)$$



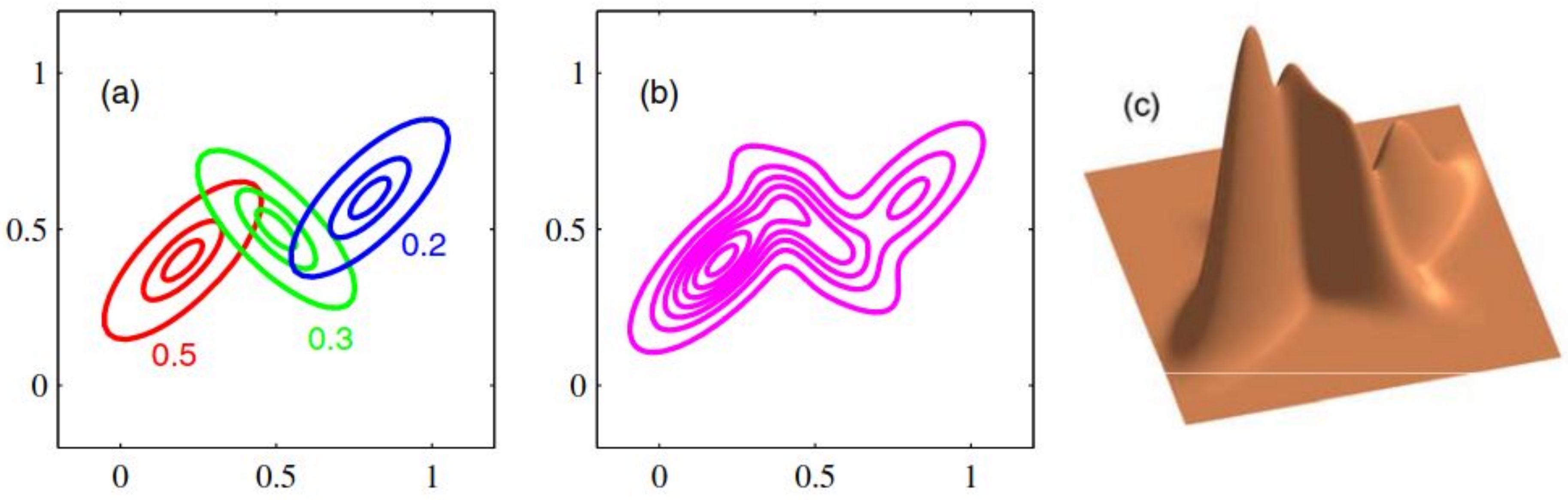


Gaussian Mixture Models: Estudio de fenómenos naturales



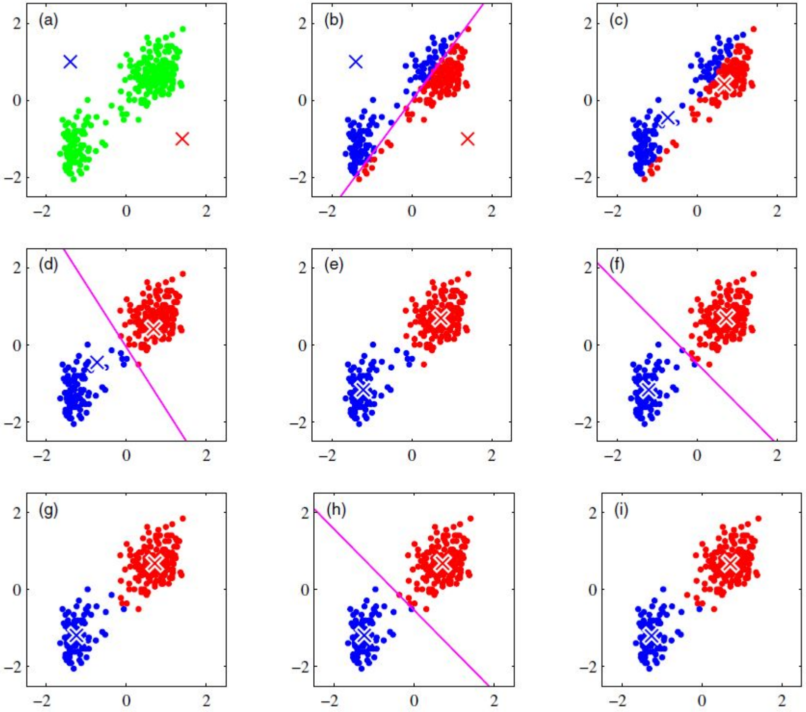
“Old Faithful” dataset. 272 mediciones de erupciones del “Old Faithful” geyser en el Parque Nacional Yellowstone. El eje horizontal representa la duración de una erupción (medida en minutos) y el vertical el tiempo hasta la próxima erupción.

Gaussian Mixture Models: Clustering

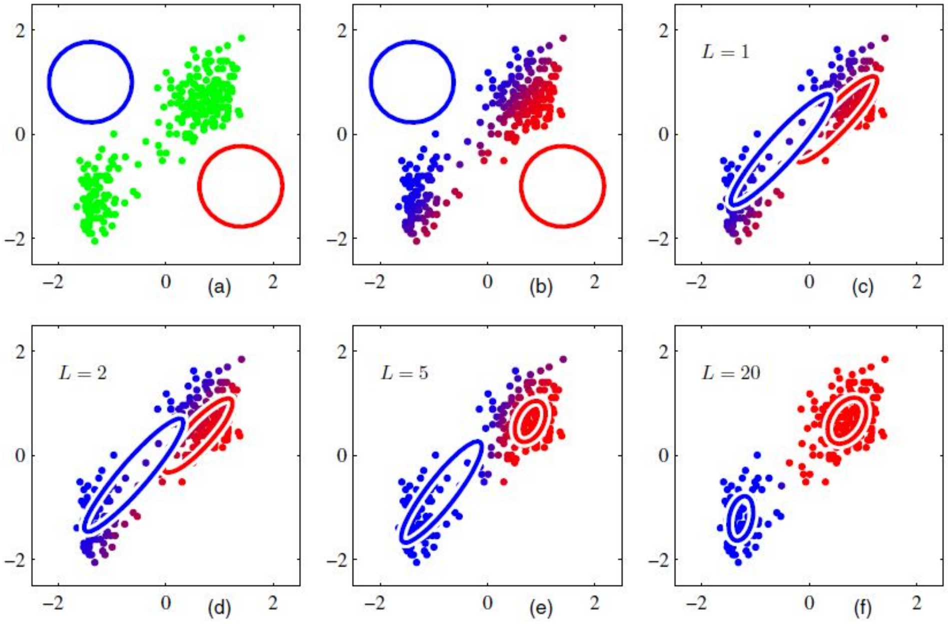


Ejemplo de Gaussian Mixture. En la imagen (a) se muestran las tres distribuciones subyacentes indicando con colores sus variables latentes. En la imagen (b) las curvas de nivel de la distribución conjunta y en la (c) la densidad.

Gaussian Mixture Models: Clustering

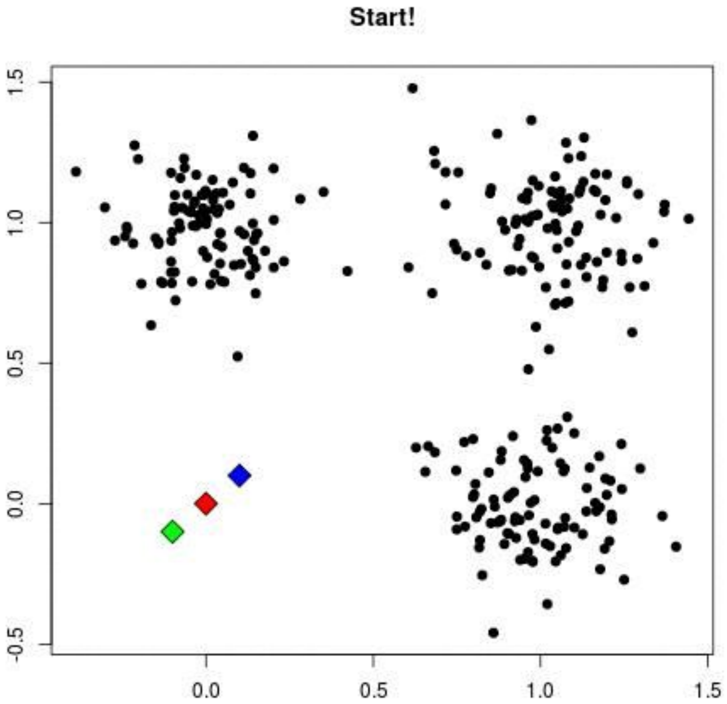


KNN

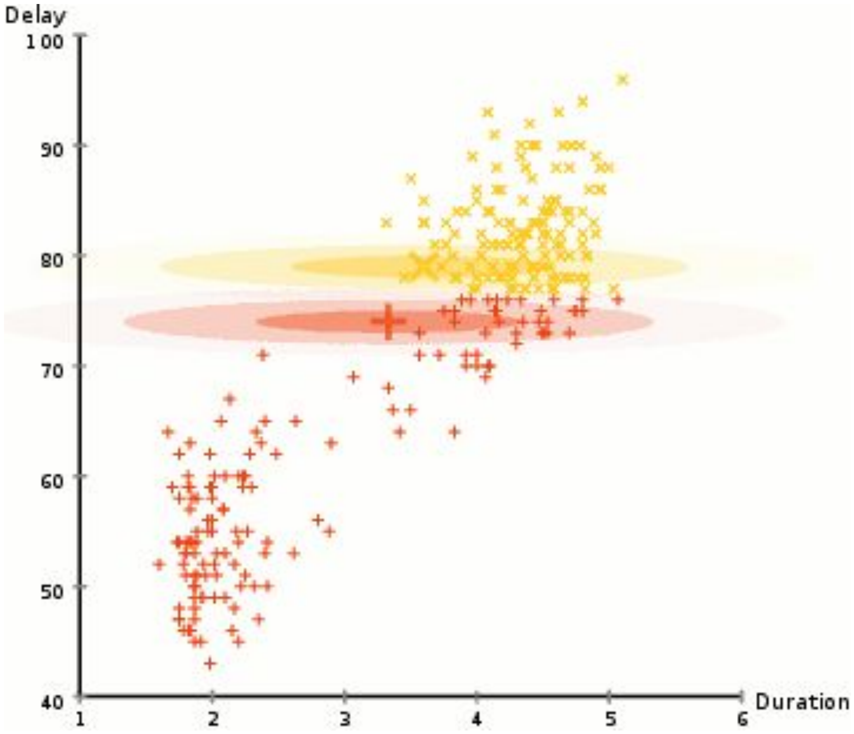


GMM

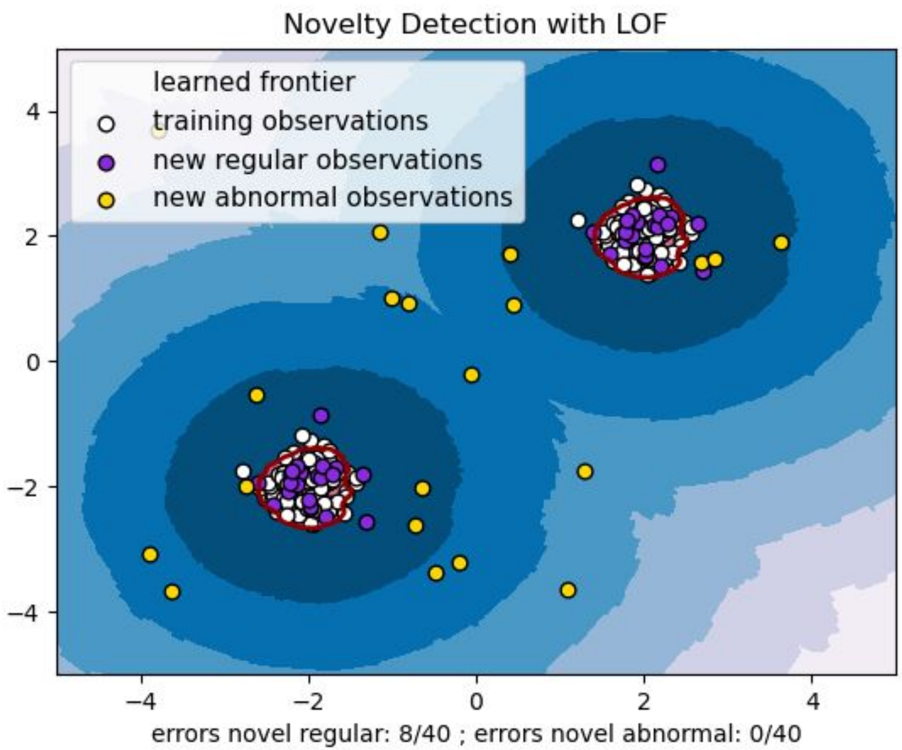
Gaussian Mixture Models - kNN



Gaussian Mixture Models: Clustering

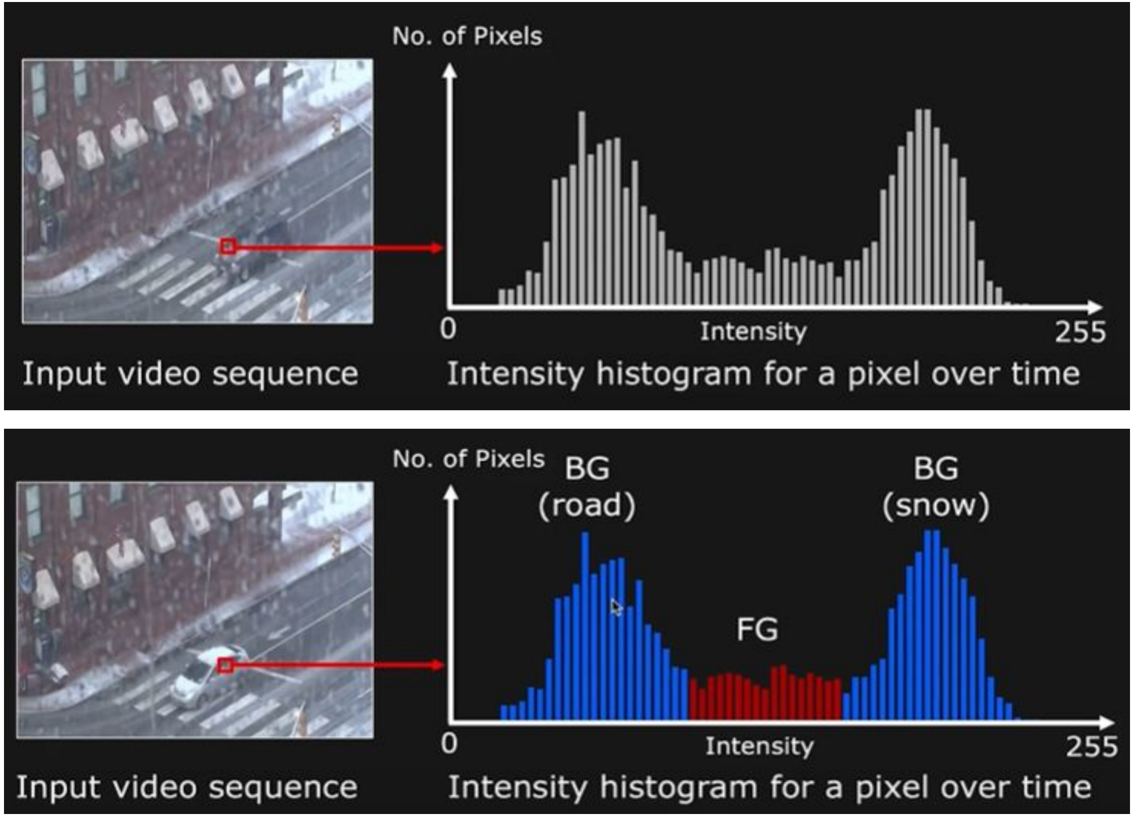


Gaussian Mixture Models: Detección de anomalías



LOF: Local outlier factor

Gaussian Mixture Models - Object Tracking



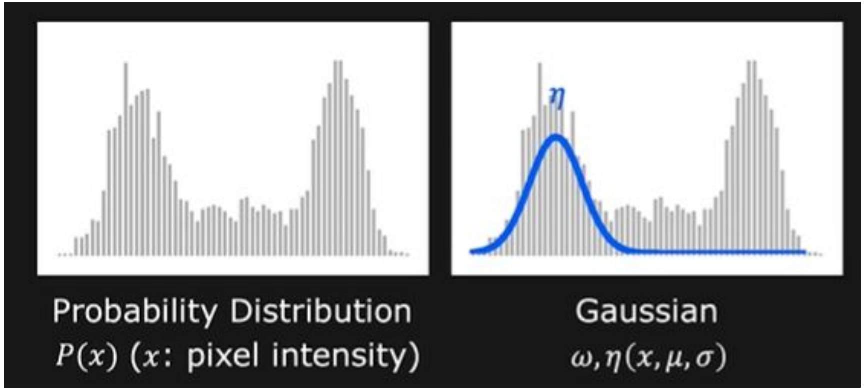
Gaussian Mixture Models - Object Tracking

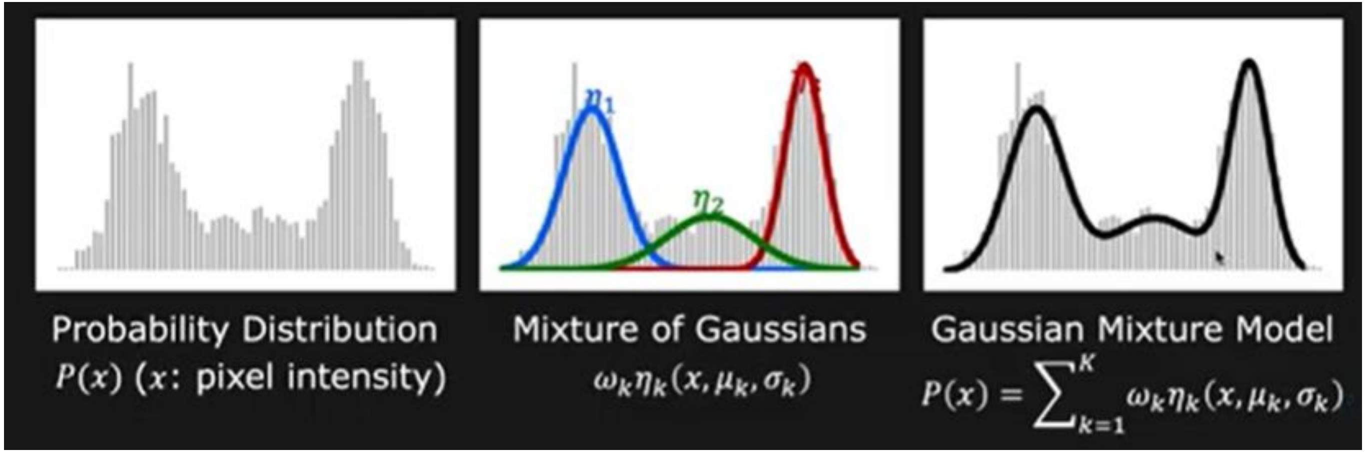
$$\boxed{P(\mathbf{X}) \cong \sum_{k=1}^{K} \omega_k \eta_k(\mathbf{X}, \boldsymbol{\mu}_k, \Sigma_k)} \quad \text{such that} \quad \sum_{k=1}^{K} \omega_k = 1$$

where: $$\boxed{\eta(\mathbf{X}, \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T(\Sigma)^{-1}(\mathbf{X}-\boldsymbol{\mu})}}$$

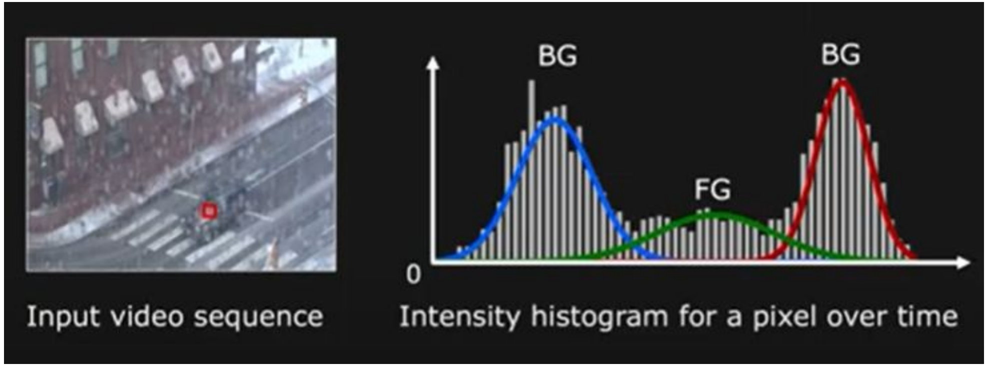
Mean $\boldsymbol{\mu} = \begin{bmatrix} \mu_r \\ \mu_g \\ \mu_b \end{bmatrix}$ Covariance matrix $\Sigma = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$ (can be a full matrix)

Gaussian Mixture Models - Object Tracking





Gaussian Mixture Models - Object Tracking



↑ $\frac{\omega}{\sigma}$ **Background**

↓ $\frac{\omega}{\sigma}$ **Foreground**

Formulación

$$p(\boldsymbol{x}) = \sum_{k=1}^{K} \pi_k p_k(\boldsymbol{x})$$

$$0 \leqslant \pi_k \leqslant 1\,, \quad \sum_{k=1}^{K} \pi_k = 1\,,$$

Mixture Models - General

$$p(\boldsymbol{x} \,|\, \boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_k \mathcal{N}\big(\boldsymbol{x} \,|\, \boldsymbol{\mu}_k,\, \boldsymbol{\Sigma}_k\big)$$

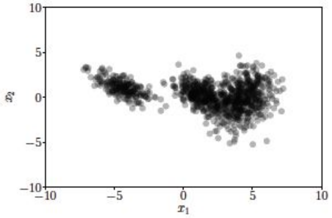
$$0 \leqslant \pi_k \leqslant 1\,, \quad \sum_{k=1}^{K} \pi_k = 1\,,$$

$$\boldsymbol{\theta} := \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \pi_k : \ k = 1, \ldots, K\}$$

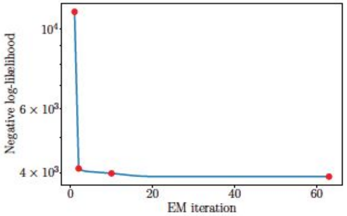
Gaussian Mixture Models

GMM y EM - JAMBOARD

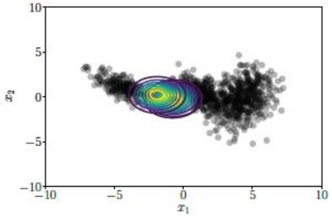
Gaussian Mixture Models - Teoría



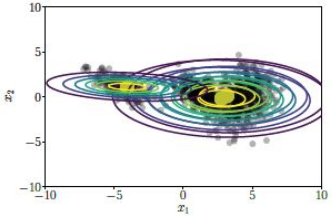
(a) Dataset.



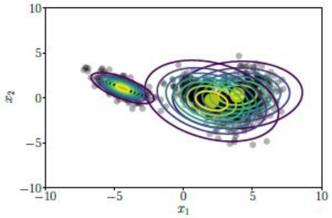
(b) Negative log-likelihood.



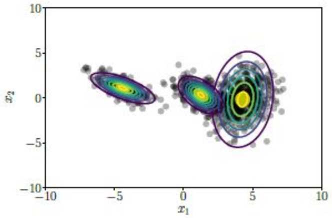
(c) EM initialization.



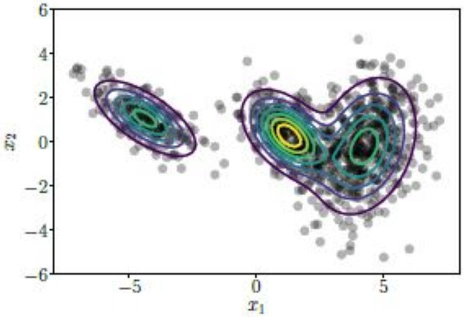
(d) EM after one iteration.



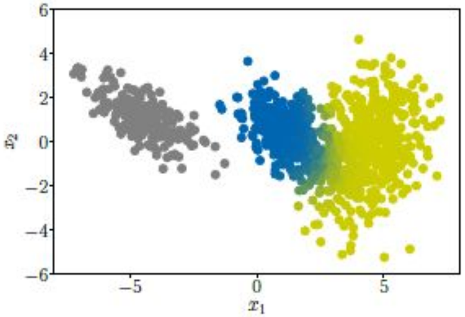
(e) EM after 10 iterations.



(f) EM after 62 iterations.



(a) GMM fit after 62 iterations.



(b) Dataset colored according to the responsibilities of the mixture components.

GMM - NOTEBOOKS

Ejercicio integrador

1. Implementar el algoritmo de Gaussian Mixture Models en NumPy.
2. Aplicar el modelo a un dataset de elección.
3. Comparar los resultados con Scikit-Learn .

Bibliografía

- Mathematics for Machine Learning | Deisenroth, Faisal, Ong
- Pattern Recognition and Machine Learning | Bishop
- Gaussian Mixture Model | John McGonagle, Geoff Pilling, Andrei Dobre
- Expectation-Maximization Algorithms | Stanford CS229: Machine Learning
- First Principles of Computer Vision| Computer Science Department, School of Engineering and Applied Sciences, Columbia University