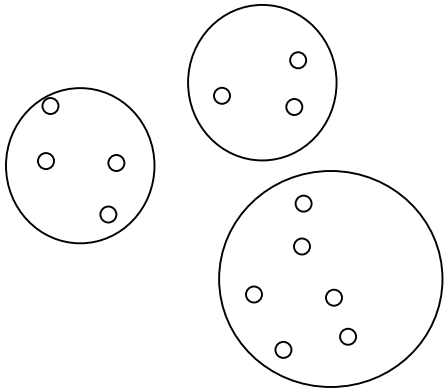
APPLIED MACHINE LEARNING

Classification

Clustering, semi-supervised clustering and classification

Clustering

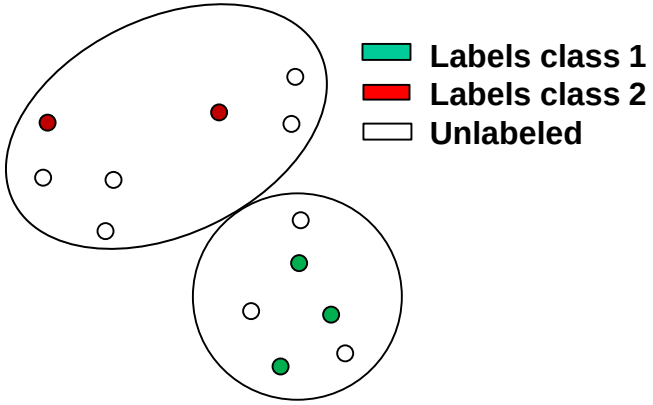
No labels for the points!



Group points according to the geometrical distribution of points

Semi-supervised clustering

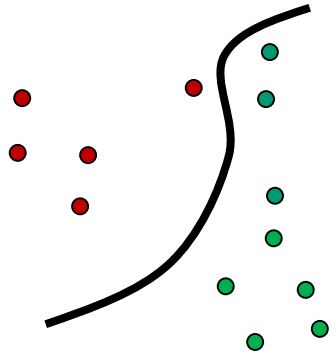
Labels a fraction of the points



Use the labels to choose hyperparameters of clustering using F1-measure.

Classification

All points are labelled



Use the labels to determine the boundary between the two classes

Binary Classification: Formalization

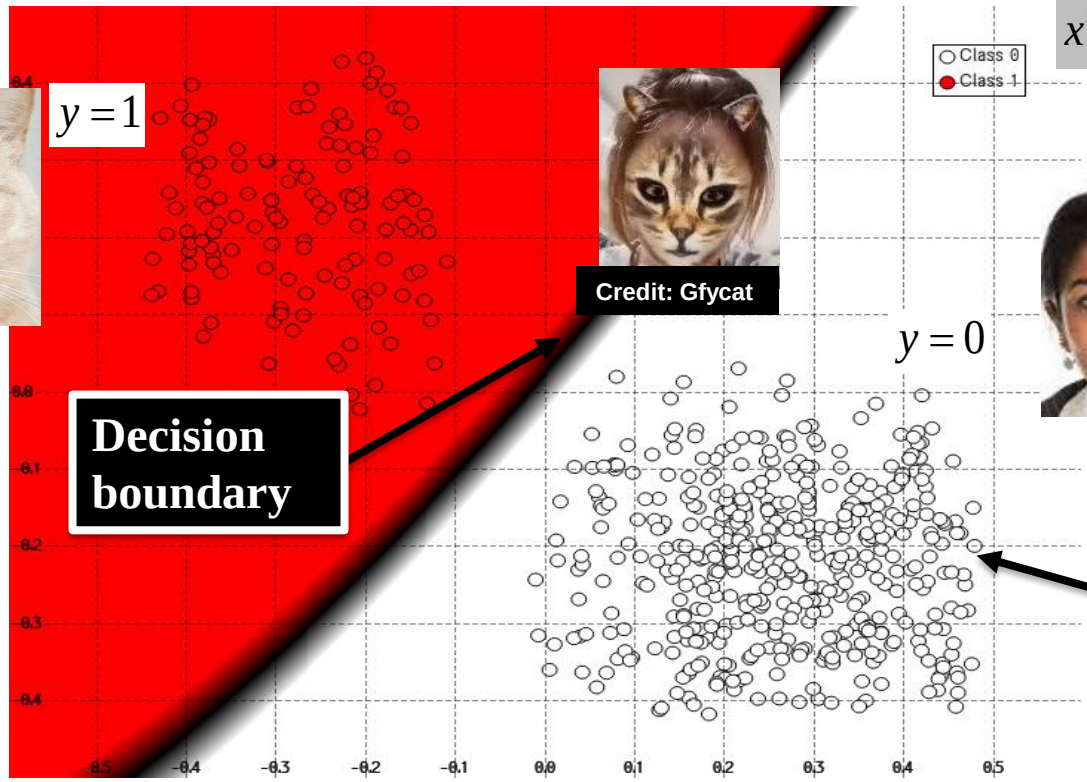
Class label $y \in \mathbb{N}$

$x^i \rightarrow y$? and $y = \{0,1\}$

$y = 1$

Credit: Gfycat

$y = 0$

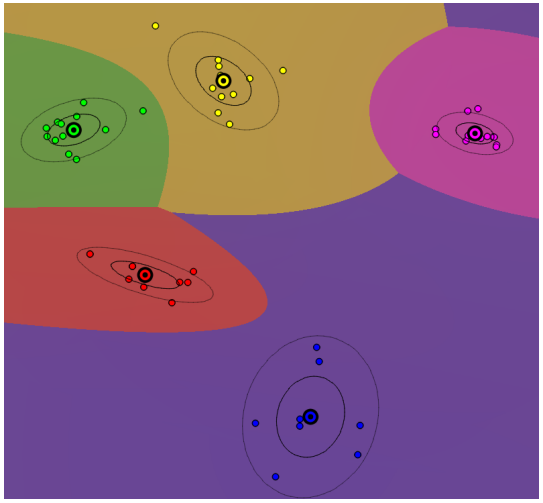


Datapoint

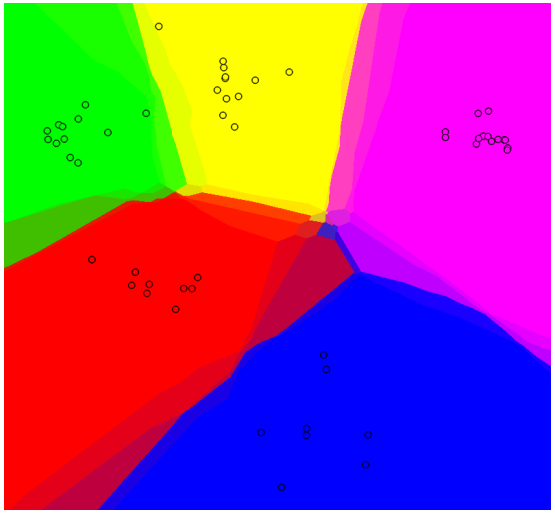
$x \in \mathbb{R}^N$

- Classification is a supervised learning technique
  - → Number and class types known, class label given for all training points.
- Class label is discrete.
- Goal: find a decision boundary which separates at best the classes.

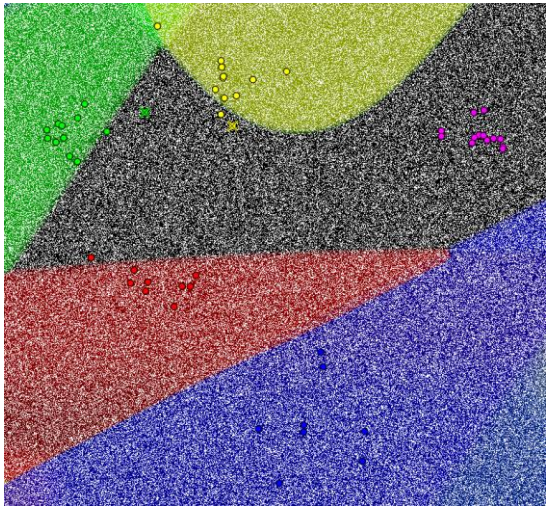
Classification techniques covered in applied machine learning course



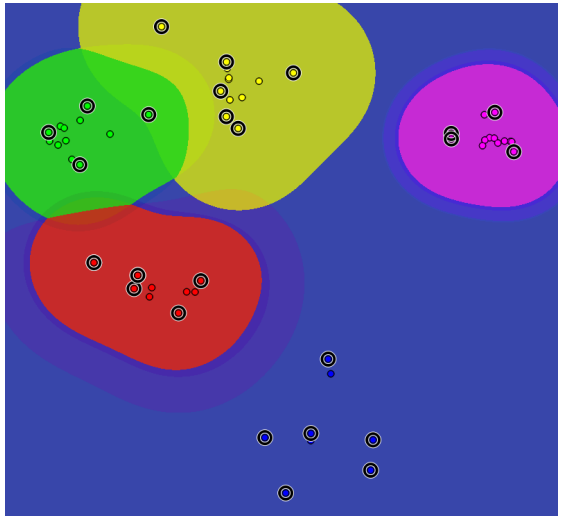
Gaussian Mixture Models (GMM)



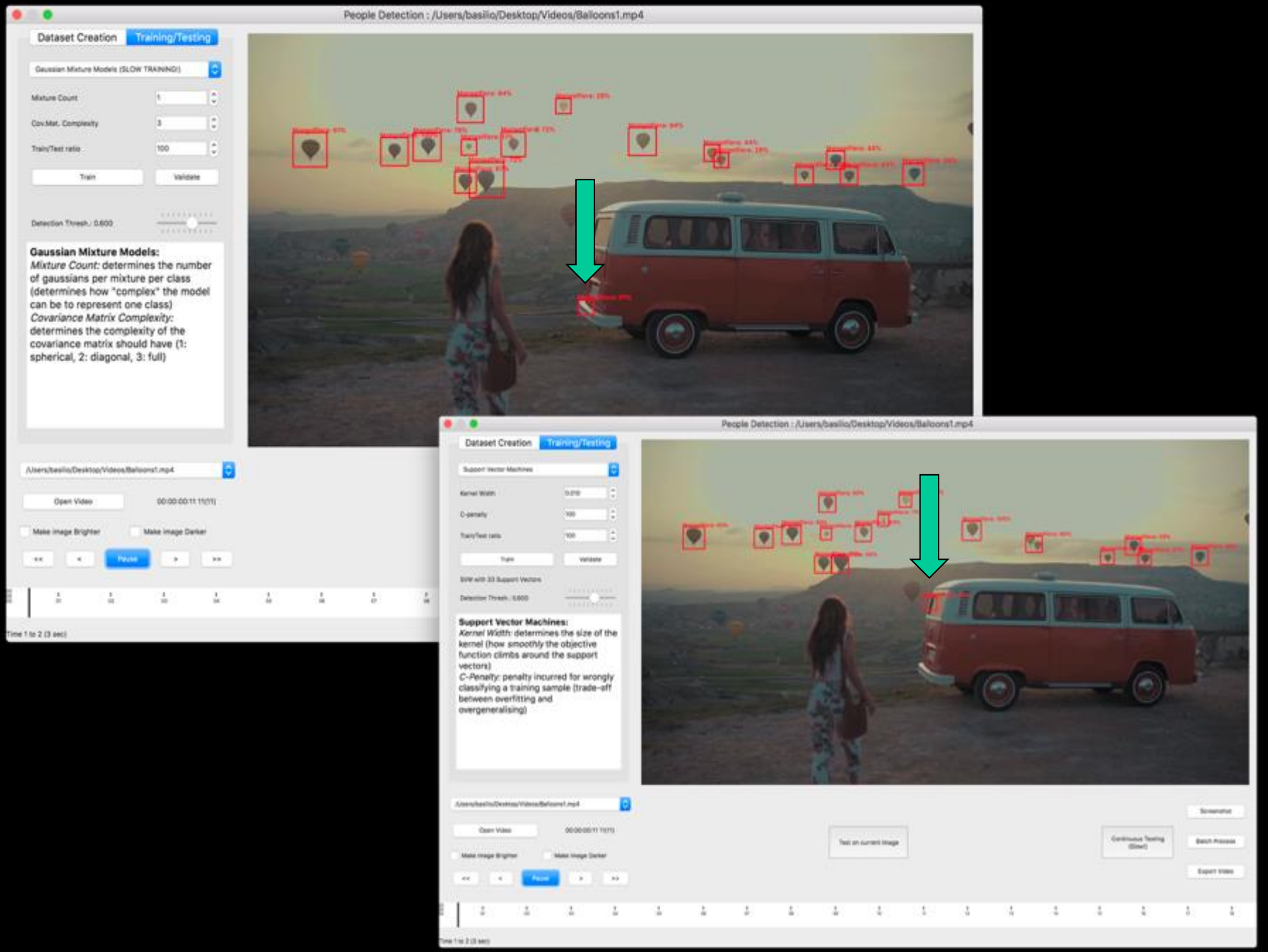
K-nearest neighbors (KNN)



Neural Networks (NN)



Support Vector Machine (SVM)


People Detection : /Users/basilio/Desktop/Videos/Balloons1.mp4
Dataset Creation
Training/Testing
Gaussian Mixture Models (SLOW TRAINING!)
Mixture Count
Cov.Mat. Complexity
Train/Test ratio
Train
Validate
Detection Thresh.: 0.600
Gaussian Mixture Models:
Mixture Count: determines the number of gaussians per mixture per class (determines how "complex" the model can be to represent one class)
Covariance Matrix Complexity: determines the complexity of the covariance matrix should have (1: spherical, 2: diagonal, 3: full)
/Users/basilio/Desktop/Videos/Balloons1.mp4
Open Video
Make Image Brighter
Make Image Darker
Pause
People Detection : /Users/basilio/Desktop/Videos/Balloons1.mp4
Dataset Creation
Training/Testing
Support Vector Machines
Kernel Width
C-penalty
Train/Test ratio
Train
Validate
Detection Thresh.: 0.600
Support Vector Machines:
Kernel Width: determines the size of the kernel (how smoothly the objective function climbs around the support vectors)
C-Penalty: penalty incurred for wrongly classifying a training sample (trade-off between overfitting and overgeneralising)
Open Video
Make Image Brighter
Make Image Darker
Pause