


SCC0602 - Algoritmos e Estruturas de Dados I

K-Nearest Neighbors (k-NN)



Professor: André C. P. L. F. de Carvalho, ICMC-USP
PAE: Rafael Martins D'Addio
Monitor: Joao Pedro Rodrigues Mattos

K-Nearest Neighbors (k-NN)

- One of the simplest machine learning algorithms
- Instance-based learning
 - Data points or vectors = instances
 - No previous training required
 - No parameter tuning!

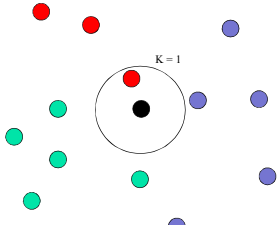
© André de Carvalho - ICMC/USP 2

K-Nearest Neighbors (k-NN)

- Learning is just a matter of finding the k most similar (closest) known data points and pick the most popular class among them
 - A distance metric is necessary to define the "closeness" among data points

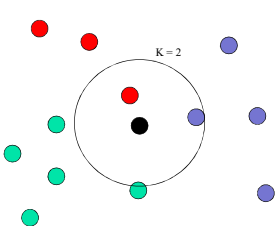
© André de Carvalho - ICMC/USP 3

K-Nearest Neighbors (k-NN)



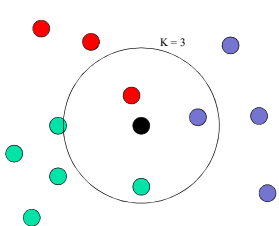
© André de Carvalho - ICMC/USP 4

K-Nearest Neighbors (k-NN)

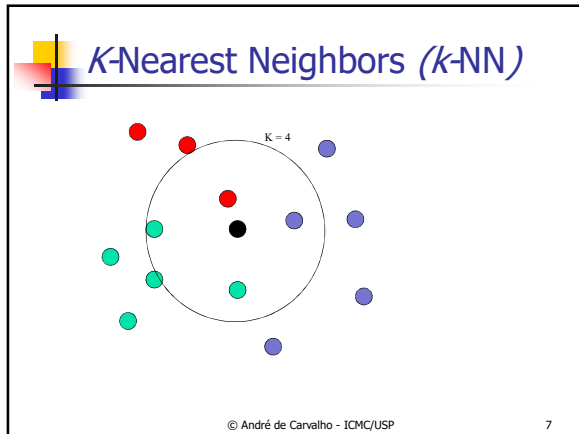


© André de Carvalho - ICMC/USP 5

K-Nearest Neighbors (k-NN)



© André de Carvalho - ICMC/USP 6



- ### K-Nearest Neighbors (k -NN)
- Design issues
 - Distance metric
 - Euclidean? Manhattan? Other?
 - Size of k
 - Voting method
 - What should we do when there are ties?
- © André de Carvalho - ICMC/USP 8