

Leren: Schrijfopdracht 4

Week 5 - Semester 1 2014/15

Mag in teams van 2 gemaakt worden
Deadline 8 december 2014 12.00 uur

1. 2 punten) *Beslisbomen*

Bereken met de hand 1 iteratie van k-means clustering en 1 iteratie van Expectation Maximization (EM) voor de onderstaande 1-dimensionele data. Neem aan dat er drie clusters zijn, initialiseer de gemiddelden met 1, 3 en 8 en ga voor EM uit van een normale/Gaussian verdeling. Neem aan dat de standaardafwijkingen in het domein voor alle clusters gelijk zijn aan 2. Bereken de "cost" (voor k-means) en de likelihood (voor EM) voor en na deze stap.

Data: 1, 2, 3, 3, 4, 5, 5, 7, 10, 11, 13, 14, 15, 17, 20, 21

2. (1 punt) *Overzicht*

We hebben in de cursus Leren een aantal leeralgorithmen gezien. Wat is een leertaak (of een variant) waarvoor geen van de besproken algoritmen echt geschikt is, maar die wel aangepakt kan worden met 1 van de benaderingen die we gezien hebben, zoals minimaliseren van kosten met gradient decent, Bayesiaans, informatie/eenvoud (zoals beslisbomen)? Beschrijf kort (a) de leertaak, (b) waarom de methoden die we hebben bepreken hiervoor niet geschikt zijn en (c) wat een mogelijke benadering is en (d) hoe deze leertaak aangepakt zou kunnen worden.