

# Leren: Programmeeropdracht 5

Week 5 - Semester 1 2014/15

**Individueel maken**

**Deadline 8 december 2014, 12.00 uur**

1. (3 punten) *Gaussian Naive Bayes implementeren*

- (a) Implementeer de Gaussian Naive Bayes Classifier. Pas die toe op de trainingsdata digits123-1 en test het op digits123-2. Bereken de proportie goede classificaties op digits123-2. Een vectorgebaseerde implementatie levert 1 punt meer op dan een implementatie die is gebaseerd op iteratie.
- (b) Kies twee voorbeelden uit die fout geclassificeerd worden in digits123-1 of -2 en geef (minstens) 1 mogelijke oorzaak van de fouten. Probeer dit zo precies mogelijk te doen. Geef een mogelijke manier om deze fout te vermijden.

2. (3 punten) *Gaussian Naive Bayes vergelijken*

Vergelijk de Gaussian Naive Bayes met een andere classifier waarvan je een werkende versie hebt. Train beide hiertoe op digits123-1 en test het resultaat op digits123-2.

- (a) Pas een statistische toets toe om na te gaan of er een verschil is tussen de resultaten. Gebruik hiervoor de accuracy, dus de proportie goed geclassificeerde voorbeelden en zet de grens voor de overschrijdingskans op  $P = 0.05$ .
- (b) Stel dat we omgekeerd te werk zouden gaan, dus trainen op digits123-2 en testen op digits123-1. Verwacht je andere uitkomsten voor accuracies? Is er verschil tussen de toetsingen? Licht je antwoord toe.