

## Tentamen LEREN deel 1

20 oktober 2009

14.00-16.00 uur

### Vraag 1 (2.5 punten)

Een arts wil een adviessysteem alleen gebruiken als het maximaal 5% fouten maakt. Er is een dataset beschikbaar waarmee met Machine Learning een adviessysteem gemaakt kan worden. De "data mining consultant" experimenteert met enkele technieken en bepaalt welke volgens hem het beste werkt. Wat moet de "consultant" doen om rekening te houden met de eis van de arts? Welke experimenten? Wat moet hij doen met de resultaten? Wat kan hij tegen de arts zeggen op grond van de resultaten? (Wees zo specifiek mogelijk.)

### Vraag 2 (2.5 punten)

Mitchell bespreekt neurale netwerken met een sigmoïde drempelfunctie ("sigmoid threshold unit").

1. Een simpel leeralgorithme voor conjunctieve concepten werkt als volgt: als er een voorbeeld van het concept binnenkomt dan vergelijkt het algoritme de huidige hypothese met het voorbeeld. De attributen waarop de hypothese een andere waarde voorspelt dan het voorbeeld heeft worden verwijderd uit de hypothese. Beschrijf dit algoritme in termen van het version space raamwerk, van het backpropagation algoritme voor het leren van neurale netwerken.
2. Hoe werkt in het algemeen backpropagation van de fout (error) door de volgende drempelfunctie: (a) een sigmoid, (b) een harde drempel bij 0 en (c) helemaal geen drempel (geeft de activatie gewoon door)?

### Vraag 3 (2.5 punten)

Stel dat we een leertaak hebben die inhoudt dat we een functie leren van 3 Boolean variabelen (het kan zijn dat niet alle variabelen relevant zijn). Er is enige ruis in het domein.

1. Stel dat we weten dat de te leren functie een conjunctieve functie is. Als we een beslisboom leren, hoeveel "knopen" zal die boom dan maximaal/minimaal hebben, inclusief "wortel" en "bladeren"?
2. Als we een neuraal netwerk leren, hoeveel knopen heeft een neuraal netwerk met 1 "hidden layer" minimaal/maximaal nodig?
3. Stel dat we niks weten over de functie. Hoeveel verschillende beslisbomen zijn er dan mogelijk? (NB: definieer eerst wat "verschillend" betekent!)

### Vraag 4 (2.5 punten)

1. Hoe kan je zien of er overfitting optreedt (bij een beslisbomenleerder of een neuraal netwerk)?
2. Wat kan je ertegen doen bij een beslisbomenleerder?
3. Wat kan je ertegen doen bij een neuraal netwerk?