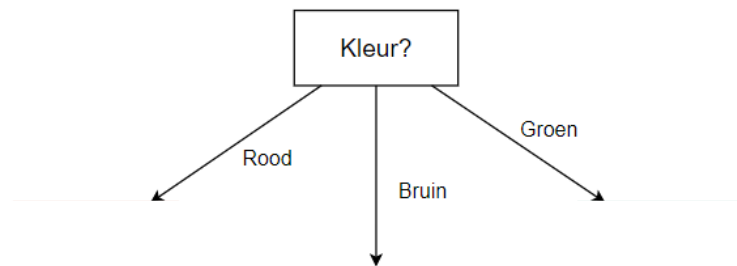


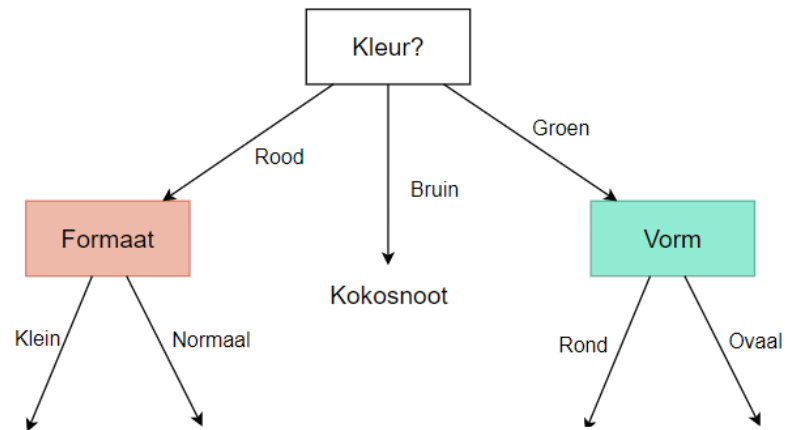
The background features a series of concentric circles in a light gray color, centered on the slide. Overlaid on these circles are stylized circuit board traces in a light blue color. These traces are located in the corners of the slide, with some ending in small circles, resembling nodes or components in a network or tree structure.

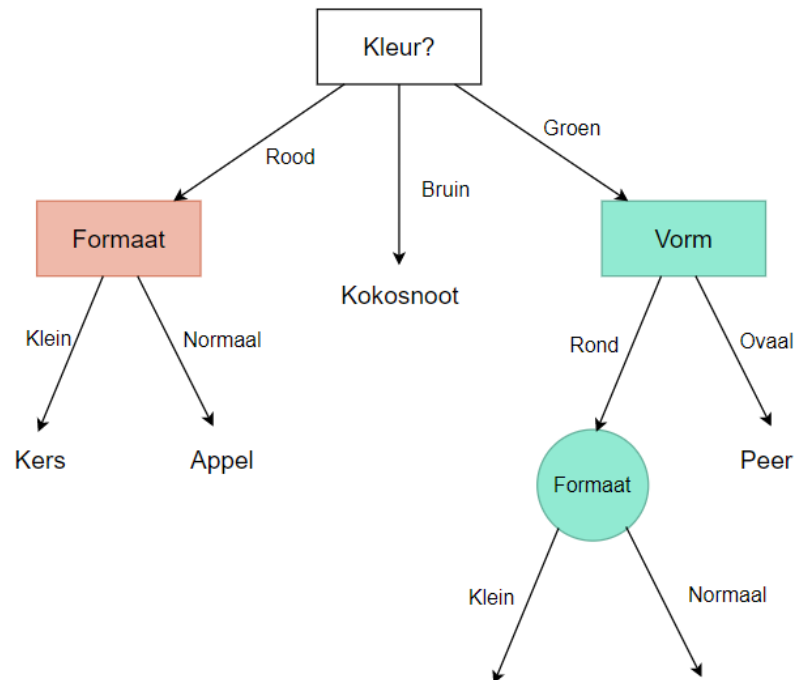
DECISION TREES

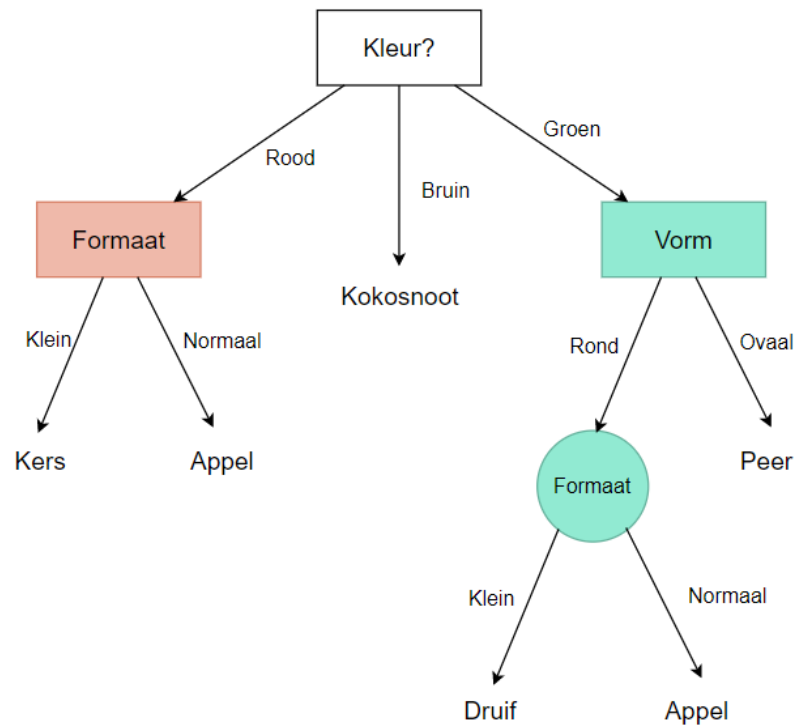
JENS BAETENS

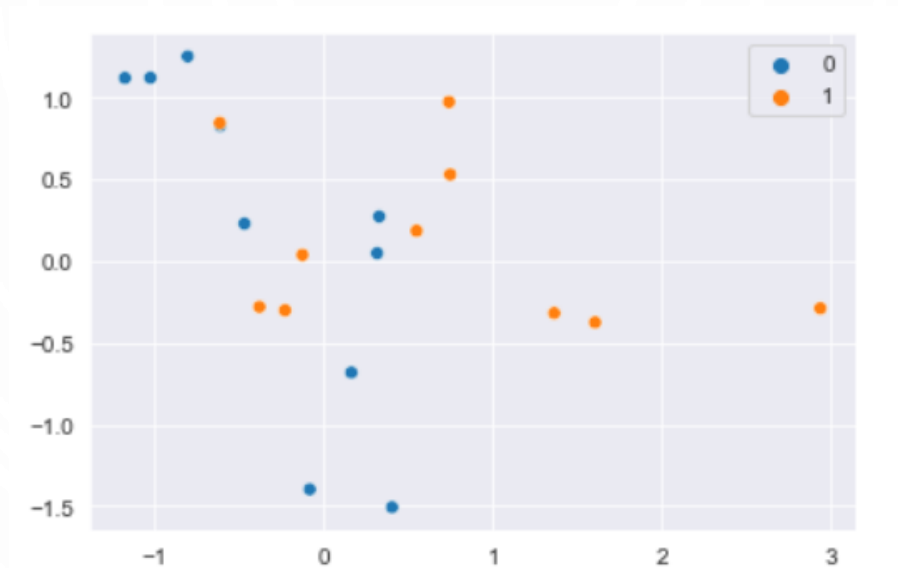


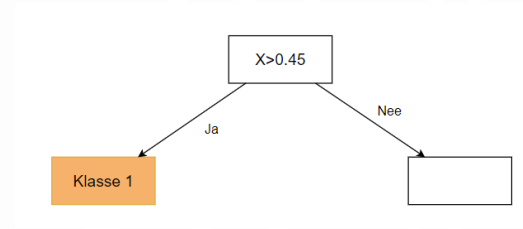
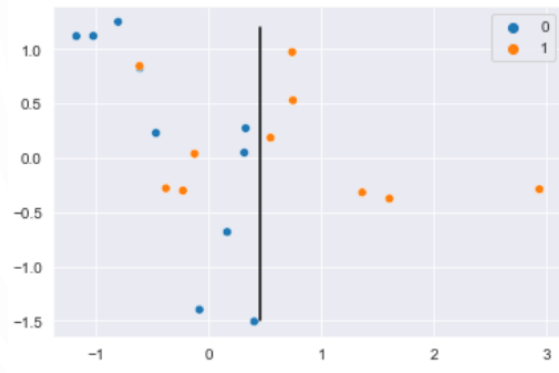


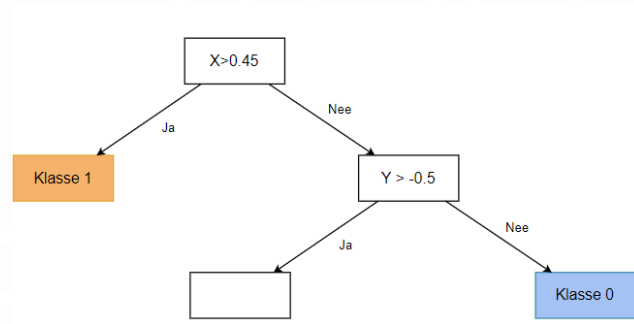
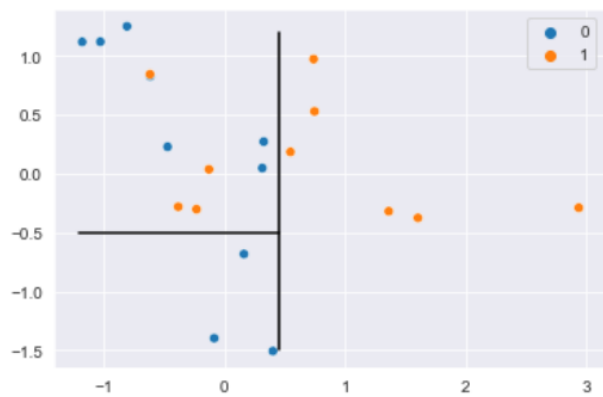


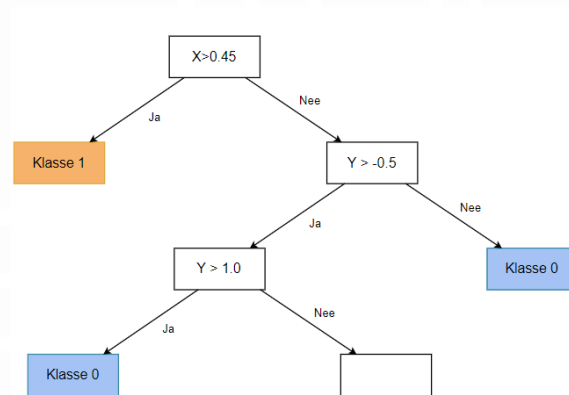
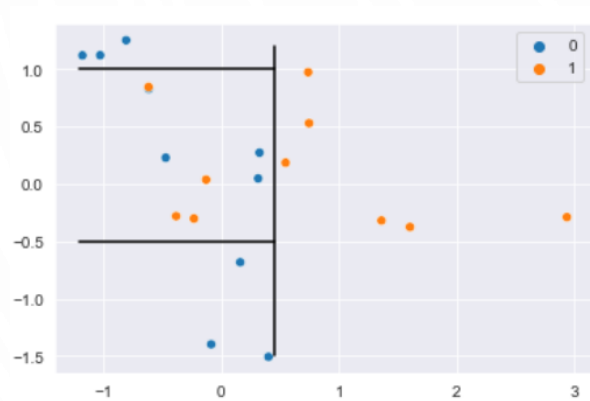


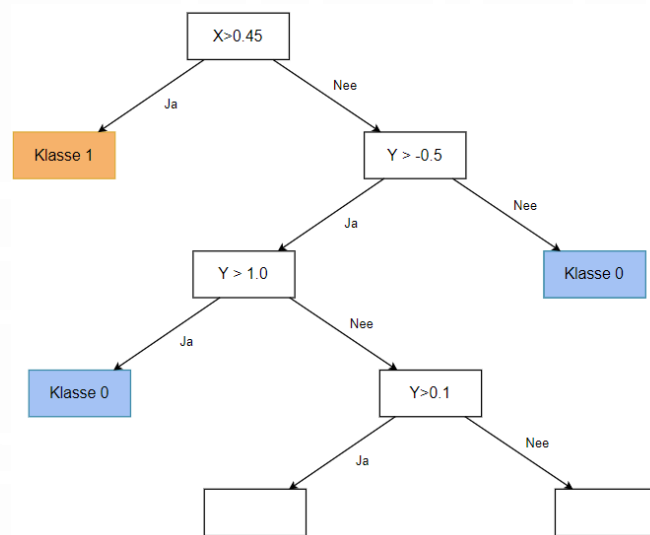
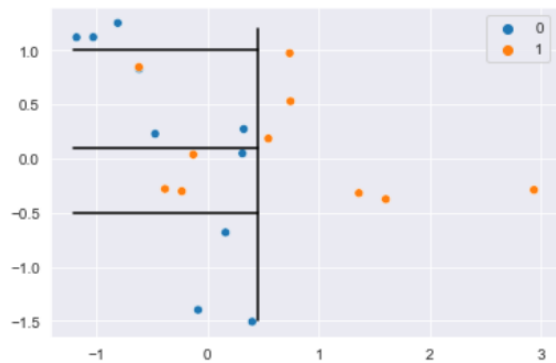


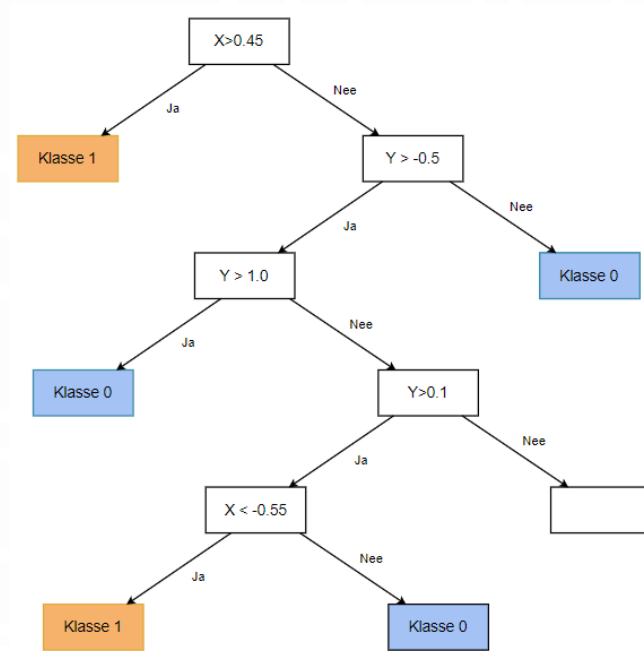
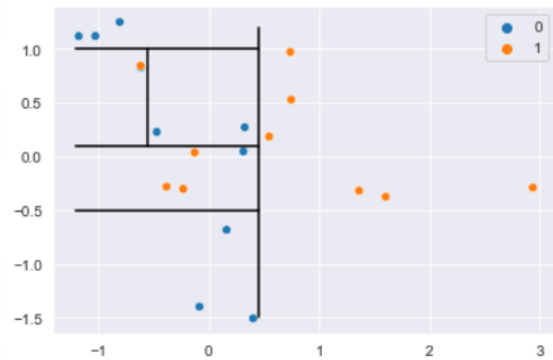


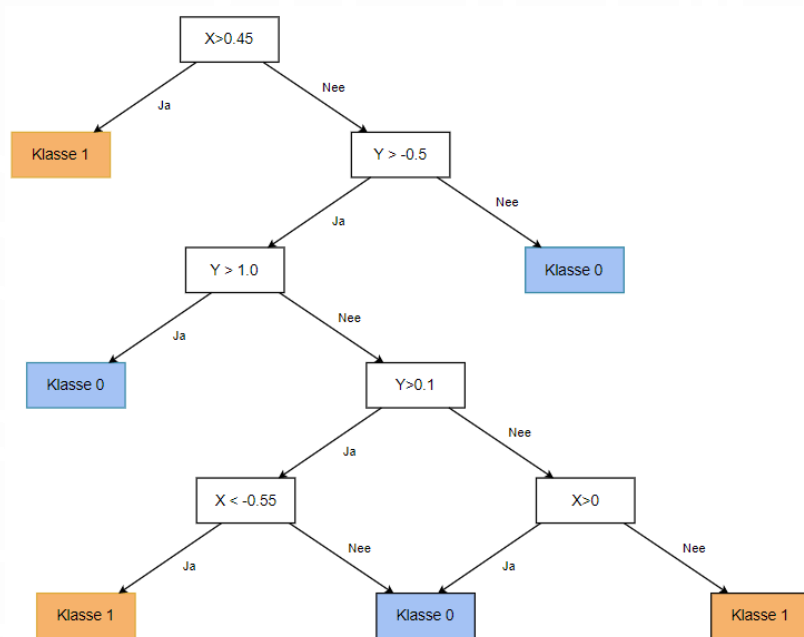
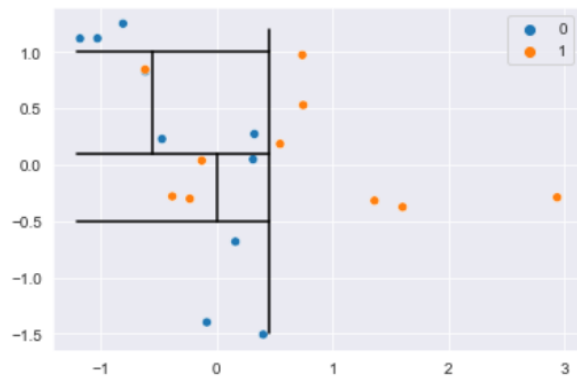












HOE BEPALEN VAN DE SCHEIDINGSLIJN?

Bereken entropie (maatstaf voor de wanorde) voor elke mogelijke verdeling

$$H = \sum_{i=1}^N p_i \log_2\left(\frac{1}{p_i}\right) = - \sum_{i=1}^N p_i \log_2(p_i)$$

P_i is het percentage van elke klasse in een gebied

Bereken het verschil met de entropie voor en na de verdeling

=> Information gained => Hoe groter hoe beter

ALTERNATIEF VOOR ENTROPIE

Logaritmes voor berekenen entropie zijn rekenintensief

Alternatieve manier: Gini impurity

$$G = 1 - \sum_{i=1}^N p_i^2$$

Beste scheidingslijn heeft de kleinste Gini Impurity

RANDOM FORESTS

RANDOM FORESTS

Is Rusland groter dan Afrika?

Ja: ...

Nee: ...





RANDOM FORESTS

Condorcet's jury theorem

Given a jury of voters and assuming independent errors. If the probability of each single person in the jury of being correct is above 50% then the probability of the jury being correct tends to 100% as the number of persons increase.

Nicolas de Condorcet (1743 - 1794)

RANDOM FORESTS

