# Wat is machine learning?

Machine learning is het proces waarbij computers zich kunnen aanpassen zonder het volgen van vooraf geschreven regels, dit wordt gedaan door het analyseren van data.

## Wat kan machine learning wel?

Machine learning kan ons helpen in het maken vaan beslissingen en het vinden van onregelmatigheden die anders niet gevonden zouden zijn.

## Wat kan machine learning niet?

# Hoe werkt k-means?

K-means clustering is een “unsupervised learning” clustering algorithm, dit betekend dat het gebruikt kan worden om een set van datapoints in groepen de klusteren, zonder labels aan te geven.

Zo werkt het:

De input is een set van punten (x1…. xN), De K in K-means zijn de aantal groepen voor de punten. Daarna plaatsen we K centroids in de ruimte van de punten. Vervolgens herhalen we deze stappen:

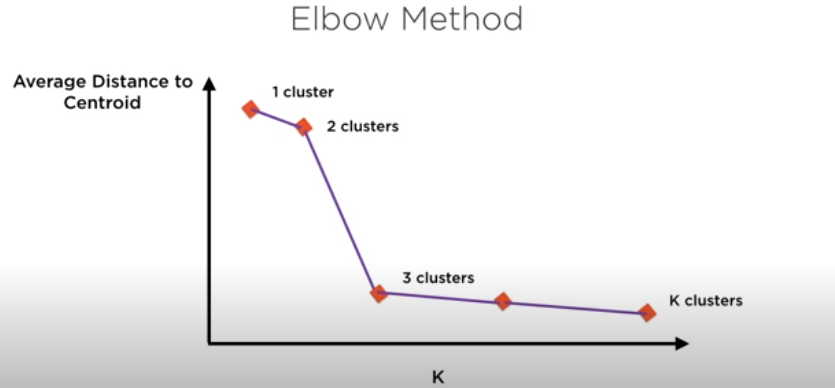
Voor elk punt vind je de dichtstbijzijnde centroid, en geef dat punt aan de gevonden centroid, zo maak je een cluster, vervolgens voeren we de volgende stappen uit op elke cluster:

Voor elke cluster gaan we de positie van de centroid opnieuw berekenen door alle punten te nemen en hun gemiddelde positie uit te rekenen.

Voer dit uit tot dat er geen nieuwe punten aan een centroid worden gegeven.

## Hoe vind je de beste K?’

Er zijn meerdere manieren om dit te doen, je kan de “elbow method” gebruiken. Hierbij gebruiken we een gekozen kost functie met als input K. als je dit op een grafiek zet kan het er zo uit zien:



We zeggen dan dat de plek van de ‘eelboog’ de beste K is (in dit geval 3)

# Hoe werkt logistic regression?

Logistic regression is een classificatie techniek, we hebben meerdere classificatie types: binary en multi classificatie. Binary classificatie is van toepassing wanneer je een dataset in “Ja” en “nee” moet groeperen, multi classificatie is van toepassing wanneer je een dataset moet classificeren in groepen zoals voor welke partij iemand zal stemmen.

# Hoe werken decision trees