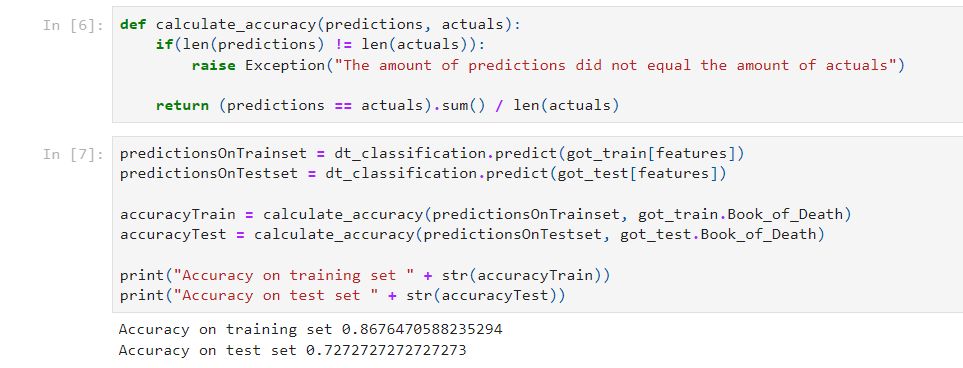
## Assignment 101

Data Science kan gebruikt worden om op basis van datasets inzichten te krijgen en data te voorspellen. Met een gegeven dataset kan data verwerkt worden aan de hand van wiskundige formules(figuur 1). De resultaten kunnen gevisualiseerd woorden aan de hand van grafieken(boxplots, histogrammen, scatterplots etc.). Deze inzichten zijn nodig om vragen te kunnen beantwoorden zoals, hoe en waarom. Het is de bedoeling dat de data begrepen wordt en op basis daarvan met berekeningen een voorspelling gemaakt kan worden voor oningevulde data.

Chart, bar chart

Description automatically generated

Figuur 2

Figuur

Business Intelligence en Data Science zijn te vergelijken. Ze gebruiken beide data en verwerken dit om hier informatie en inzichten uit te halen(figuur 2). Bij Artificial Intelligence wordt er veel voorspeld. Een AI moet kunnen leren van informatie en de informatie in de toekomst kunnen gebruiken. Een decision tree is hier een goed voorbeeld van. AI kan een decision tree gebruiken om successen en falen vast te stellen. (figuur 3)

Diagram

Description automatically generated

Figuur 3

## Assignment 102

Een data scientist kijkt welke vragen beantwoord moeten worden en waar de data voor de antwoorden te vinden is. Dit kan door verschillende analyses. Zo kan er één variabele(categorisch of numeriek) bekeken worden, twee, of meerdere. Bij het bekijken van één variabele krijgt een data scientist visualisatie voor categorische waardes en analytische inzichten voor numerieke waardes. Hieruit kan gezien worden hoe vaak een waarde voorkomt of welke waarde het meest voorkomt.Chart, histogram

Description automatically generated Chart, histogram

Description automatically generated  
Bij het bekijken van twee variabele wordt er gekeken of er verbanden zijn tussen de twee variabele. Dit kunnen alle combinaties zijn. Dit zal meer antwoordt geven op waarom de data is zoals het is en of combinaties veel voorkomend zijn.

Chart, bar chart

Description automatically generatedChart, scatter chart

Description automatically generated

Bij meer variabele kan een data scientist op basis van bestaande data, een voorspelling doen. Dit is handig voor een bedrijf om te kijken wat bijvoorbeeld klanten gaan doen op het moment dat Ook is het mogelijk om de data te clusteren om te zien welke data allemaal bij elkaar