

## Factsheet Movie Recommendations

Dit document bevat een samenvatting van het project welke is uitgevoerd door Yoeri Bruchem, Gabriële de la Cruz & Mick van Hulst. Het project is een onderdeel van de Data Science minor aan de Hogeschool Rotterdam. *Het doel voor de groep was om een website op te leveren die films aanbeveelt aan gebruikers van de website.*

### Proces van gebruiker

De bezoeker begint met het aanmaken van een account op de website. Na inloggen krijgt deze een lijst met twintig films te zien. Dit zijn films welke maximaal tien jaar oud zijn en minimaal een gemiddelde beoordeling hebben van drie en een halve ster. Nadat de gebruiker het minimum heeft bereikt (twintig beoordelingen) wordt er een inhoud gebaseerd algoritme gebruikt om films aan te bevelen. Na deze stap krijgt de gebruiker zijn of haar aanbevelingen te zien welke bestaan uit aanbevelingen op basis van een gebruikers gebaseerd algoritme (vergelijking tussen gebruikers) en op inhoud gebaseerde aanbevelingen (vergelijking tussen films).

### Deliverables

Gedurende het project zijn de volgende taken uitgevoerd:

- Opzet infrastructuur: Dit bestond uit het opzetten van de server en het verwerken van de benodigde data. Voor deze data zijn verschillende scrapers ontwikkeld welke de plots van films konden worden verzameld en afbeeldingen van films opgeslagen konden worden.
- Algoritmes voor aanbevelingen: Voor de aanbevelingen zijn verschillende algoritmes gemaakt. Deze worden hier beschreven:
  - Op basis van plots van films zijn er gelijkenissen gezocht tussen films. De films welke van de gebruiker een waardering hebben gekregen worden vergeleken met de films uit de database. Films welke genoeg gelijkenissen tonen, worden aanbeveelt.
  - Op basis van waarderingen van gebruikers zijn er gelijkenissen gezocht tussen gebruikers. De films welke van de gebruiker een waardering hebben gekregen worden vergeleken met andere gebruikers. Films van gebruikers welke gelijkenissen tonen, worden aanbeveelt.
  - Op basis van waarderingen van gebruikers is de afstand tussen gebruikers berekend. Gebruikers die qua waarderingen niet op elkaar lijken krijgen een hoge afstand en gebruikers die veel op elkaar lijken krijgen een kleine afstand. De films van de gebruikers met de kleinste afstanden, worden aanbeveelt.
  - De resultaten zijn geteld en de resultaten van de algoritmes welke het meest voorkomen worden aanbeveelt aan de gebruiker.
- Applicatie: Dit bestond uit het realiseren van een omgeving waar de gebruiker het proces kan uitvoeren welke in de paragraaf hierboven is beschreven.
- Documentatie: Dit bestond uit het verslagleggen van de gebeurtenissen rondom het project.

### Obstakels

Een van de grootste obstakels bij de projectgroep was de ervaring in informatica onder de leden. Sommige hadden ervaring met programmeren, maar niet op het niveau van een informaticastudent. De verwachting was dat de projectgroep in dit project een steile leercurve zou ervaren op het gebied van programmeren. Joep is in het begin van het project uitgevallen. Hij voelde zich niet thuis bij de minor en heeft besloten om vervroegd te beginnen met afstuderen. Dit zorgde voor een verhoging in werkdruk onder de projectleden. Tijdens de tweede periode van het project is Bilal afgevallen. Dit was vrij laat in de periode waardoor het weinig impact had op de taakverdeling.