



# Recommender Systems en de Meerwaarde van een Gepersonaliseerde Ervaring bij Streaming Services

Auteur: **Jens Van Gelder**

Interne promotor: **Louis-Philippe Kerkhove**

Externe promotor: **Steven Schellens**

**Academiejaar: 2017-2018**

Doel: tot het bekomen van het diploma Hoger Onderwijs, één cyclus, volledig leerplan,  
| studiegebied Handelswetenschappen, Bachelor in het Informatiemanagement en de multimedia |  
| Optie Informatie Management & Security |





# Recommender Systems en de Meerwaarde van een Gepersonaliseerde Ervaring bij Streaming Services

Auteur: **Jens Van Gelder**

Interne promotor: **Louis-Philippe Kerkhove**

Externe promotor: **Steven Schellens**

**Academiejaar: 2017-2018**

Doel: tot het bekomen van het diploma Hoger Onderwijs, één cyclus, volledig leerplan,  
| studiegebied Handelswetenschappen, Bachelor in het Informatiemanagement en de multimedia |  
| Optie Informatie Management & Security |

"Ik, [Jens Van Gelder](#), verklaar dat, voor zover ik er weet van heb, deze scriptie geen materiaal bevat dat ooit in eender welke instelling is gebruikt om een diploma, van welke aard ook, te behalen of dat eerder werd gepubliceerd of geschreven door een ander persoon, behalve daar waar deze scriptie referenties bevat naar andere werken."

# Inhoudsopgave

<b>2 Samenvatting</b>	7
<b>3 Inleiding</b>	8
3.1 Probleemstelling	8
3.2 Hoofdvraag	8
<b>4 Theoretisch kader/Literatuuronderzoek</b>	9
4.1 Wat is een recommender system?	9
4.2 Content-Based Recommender Systems	10
4.2.1 Voor- en nadelen	11
4.3 Collaborative Recommender Systems	11
4.3.1 Memory-based Collaborative Filtering	13
4.3.1.1 User-based	13
4.3.1.2 Item-based	13
4.3.2 Model-based Collaborative Filtering	14
4.3.3 Voor- en nadelen	15
4.4 Hybrid Systemen	16
<b>5 Creëren van een gepersonaliseerde ervaring</b>	17
<b>6 Use Cases</b>	18
6.1 Netflix	18
6.1.1 Business model	18
6.1.2 Recommender systems	20
6.1.2.1 Personalized Video Ranker (PVR)	20
6.1.2.2 Top-N Video Ranker	21
6.1.2.3 Trending Now	21
6.1.2.4 Because You Watched & Video-Video Similarity	21
6.1.3 Survey	22
6.1.4 Besluit	23
6.2 Youtube	23
6.2.1 Business model	23
6.2.2 Recommender systems	24
6.2.3 Survey	26
6.2.4 Besluit	27
6.3 Spotify	27
6.3.1 Business model	27
6.3.2 Recommender systems	28
6.3.3 Survey	30
6.3.4 Besluit	31
<b>7 Inzet van de Eindgebruiker</b>	32
<b>8 Besluit</b>	32
<b>9 Bibliografie</b>	33

# 1 Woord vooraf

Tijdens mijn opleiding Information Management & Security heb ik veel verschillende dingen bijgeleerd over verschillende deelvakken in de wereld van Data. Deze vakken bevatten zaken zoals Business Analytics, Business Intelligence maar ook werden ons dingen als Entrepreneurship en Business Management aangeleerd. Al de expertises die ik heb opgedaan de laatste 3 schooljaren heb ik nu gecombineerd in deze scriptie.

Mijn interesses lagen vooral bij het technische aspect van deze opleiding. Daarom dat ik gekozen heb om een analyse te maken op de werking achter een interessante en steeds meer voorkomende techniek, namelijk Recommender systems. Naast de werking van deze systemen wou ik ook graag nagaan hoe de grote Streaming services, waar ik dagelijks ook gebruik van maak, deze technieken implementeren om een gepersonaliseerde ervaring aan te bieden. Ook leek het me interessant om te weten te komen hoe effectief deze systemen zijn en of deze wel degelijk een meerwaarde bieden.

Graag wil ik Michelle Lenaerts bedanken voor het helpen zoeken naar een goed onderwerp voor dit eindwerk. Ook wil ik mijn promotors bedanken. Louis-Philippe Kerkhove voor het geven van een kickstart en tips over de mogelijk inhoud van dit eindwerk en Steven Schellens voor mij te introduceren in de wereld van Machine Learning.

## 2 Samenvatting

Het aanbieden van een gepersonaliseerde ervaring begint steeds meer en meer de norm te worden voor online diensten zoals Youtube, Amazon en Netflix. Deze diensten maken hiervoor gebruik van technieken genaamd Recommender systems. In deze scriptie zullen we onderzoeken wat een Recommender system is, wat de verschillende soorten systemen zijn en hoe deze werken. Eens we een goed begrip hebben van wat een Recommender system is en hoe deze werken gaan we onderzoeken hoe deze gebruikt worden door grote bedrijven aan de hand van 3 use cases: Youtube, Netflix en Spotify.

Per use case gaan we deze bedrijven analyseren op hun werking en implementaties van Recommender systems. We gaan kijken wat het business model is van elk deze bedrijven en hoe deze mogelijk een invloed hebben op het type Recommender system dat wordt toegepast.

We gaan de methodologie van de verschillende bedrijven vergelijken met elkaar en aan de hand van een survey proberen te achterhalen wat de gebruikers van deze diensten denken over de al dan niet gepersonaliseerde ervaring die wordt aangeboden.

Uiteindelijk is het doel om tot een conclusie te komen of het aanbieden van een gepersonaliseerde ervaring door streaming services al dan niet meerwaarde biedt aan deze service.

## 3 Inleiding

### 3.1 Probleemstelling

Met de stijgende populariteit van media streaming services zoals Spotify en Netflix is het belangrijk voor deze bedrijven om maandelijks zoveel mogelijk abonnementen van klanten te behouden en nieuwe klanten te overtuigen om een nieuw abonnement te starten. De grote uitdaging hier zal zijn om hun gebruikers geïnteresseerd te houden in hun services. Hoe kunnen deze bedrijven een leuke, interessante en gepersonaliseerde ervaring aanbieden aan elk van hun klanten?

### 3.2 Hoofdvraag

Bedrijven zoals Netflix en Spotify beschikken over een groot aanbod aan films, muziek en series dat op aanvraag beschikbaar staat voor hun klanten. Via een maand- of jaarabonnement kunnen klanten gebruik maken van deze dienst waar en wanneer ze willen. Dit type bedrijven of diensten worden ook wel streaming services genoemd. Er zijn ondenkbaar veel verschillende genres en sub-genres in de muziek-, film- en tv-wereld met iedereen dat wel zijn voorkeuren heeft. Het is noodzakelijk voor streaming services om de voorkeur van hun klanten op hun aanbod te voorspellen om zo een gepersonaliseerde ervaring te creëren en nieuwe media voor te stellen aan de klant zodat deze geboeid blijft gebruik maken van hun diensten. Deze noodzakelijkheid in het creëren van een gepersonaliseerde ervaring werd onderzocht door Netflix:

Consumer research suggests that a typical Netflix member loses interest after perhaps 60 to 90 seconds of choosing, having reviewed 10 to 20 titles (perhaps 3 in detail) on one or two screens. The user either finds something of interest or the risk of the user abandoning our service increases substantially. The recommender problem is to make sure that on those two screens each member in our diverse pool will find something compelling to view, and will understand why it might be of interest. (Gomez-Urbe & Hunt, 2015)

Uit dit onderzoek blijkt dus dat klanten geïnteresseerd houden een groot probleem kan zijn voor streaming services. Hoe de grote streaming services dit probleem trachten op te lossen wordt onderzocht in deze scriptie via volgende onderzoeksvraag:

“Hoe worden Recommender System toegepast voor het creëren van een gepersonaliseerde ervaring bij de verschillende populaire streaming services en hoe effectief zijn deze in het geëngageerd houden van gebruikers?”



## 4 Theoretisch kader/Literatuuronderzoek

Volgende situatie komt je misschien wel bekend voor. Je bent online aan het zoeken naar een nieuw paar schoenen op bijvoorbeeld Zalando. Je denkt het juiste paar schoenen gevonden te hebben en wanneer je ze probeert te bestellen krijg je eerst nog een aantal andere kledingstukken voorgesteld waarvan de site denkt dat deze goed samen zouden gaan bij je nieuwe schoenen en stijl. Achter de schermen werkt hier een software tool die gebruik maakt van machine learning algoritmes die op basis van jouw keuzes en smaak nieuwe aanbevelingen kan geven. Deze tools worden ook wel Recommender Systems genoemd. In het *Recommender System Handbook (2010)* wordt er een zeer goede definitie gegeven van wat een Recommender system is:

Recommender Systems are software tools and techniques providing suggestions for items to be of use to a user. The suggestions provided are aimed at supporting their users in various decision-making processes, such as what items to buy, what music to listen, or what news to read. (Rokach, Ricci, & Shapira, 2010, p. 1)

Zoals bovenstaande definitie al aanhaalt worden Recommender Systems gebruikt in meerdere plaatsen waarvan je nog niet wist dat deze technologie achter de schermen werkte. Facebook gebruikt het om pagina's voor te stellen die je zou liken, Amazon gebruikt Recommender Systems om koopjes voor te stellen gebaseerd op eerdere aankopen en LinkedIn gebruikt deze technologie om professionals over heel de wereld samen te brengen.

In dit eindwerk gaan we de focus leggen op hoe de populaire streaming services als Netflix en Spotify gebruik maken van deze tools om een gepersonaliseerde ervaring te bieden aan elk van hun klanten. Maar alvorens we hier aan beginnen moeten we eerst duidelijk begrijpen wat een Recommender System juist is en hoe deze technologie werkt.

### 4.1 Wat is een recommender system?

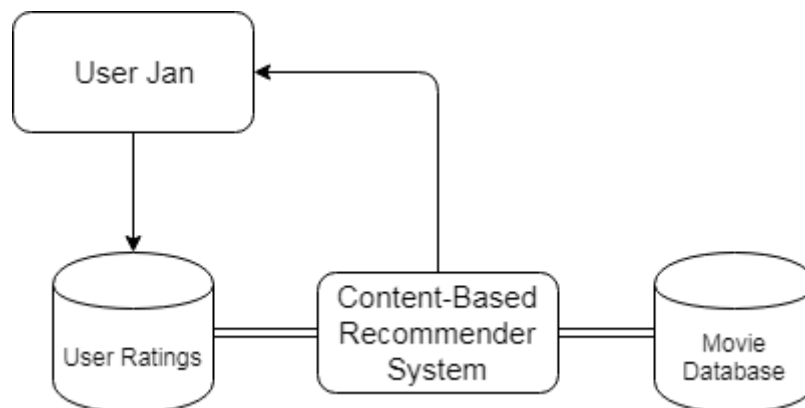
Eerder gaf het grote Recommender System Handbook (2010) ons al een simpele definitie: Een techniek of software dat aanbevelingen maakt voor users. Rokach, Ricci & Shapira melden verder ook dat het ontwerpen van een Recommender System gebruik maakt van multidisciplinaire inspanningen op verschillende gebieden van de informatica, met name Machine Learning (ML) en datamining, information retrieval, en human-computer interactie (Rokach, Ricci, & Shapira, 2010, p. 21). ML is het in staat stellen van computers tot het herkennen van patronen in data. Hiervoor worden wiskundige algoritmes gebruikt om computers zelfstandig te laten leren uit data. ML is een breed veld binnen Artificial Intelligence (AI) dat staat voor het creëren van intelligentie in computers. Via het gebruik van deze ML algoritmes zal het dus mogelijk zijn om patronen te ontdekken in de kijkgeschiedenis van een klant om zo uit deze data nieuwe gepersonaliseerde aanbevelingen te geven.

Je kan een Recommender System vergelijken met een filter waar al je mogelijke uitkomsten door gaan maar enkel de juiste informatie en aanbevelingen voor jouw interesses doorlaat. Momenteel kan je alle Recommender Systems opdelen onder drie

verschillende modellen. Content-based, Collaborative en de Hybrid Recommender Systems.

## 4.2 Content-Based Recommender Systems

Een content-based systeem maakt aanbevelingen lijkend op media dat de gebruiker in het verleden leuk vond (Adomavicius & Tuzhilin, 2005). De gelijkenissen tussen deze media worden opgenomen door dit systeem. Hierdoor kan het bijvoorbeeld in staat zijn te detecteren wanneer het genre “action” vaak voorkomt in de lijst van films dat jij een hoge score hebt gegeven. Wanneer het systeem een bepaald attribuut vaak opnieuw ziet opduiken zal deze instinctief items met dit attribuut beginnen voorstellen aan de gebruiker. Een artikel over content-based recommendation & user privacy in social-tagging systems beschrijft dit systeem als volgt: *“a proper technique for representing the items and users’ profiles, and a strategy to compare items and users, and produce a recommendation”* (Puglisi, Parra-Arnau, Forné, & Rebollo-Monendero, 2015, p. 3.4). Om dit meer context te geven in de vorm van data management kan je voorstellen dat dit systeem twee databanken naast elkaar gaat zetten om de gelijkenissen tussen de twee te kunnen vinden. Enerzijds de databank gevuld met de informatie van de gebruiker met alle attributen (naam van een film, acteurs, genre) dat hij een hoge of lage rating heeft gegeven met dan als tweede de volledige databank aan media van de streaming services zoals aangegeven in onderstaande [Figuur 1](#).



*Figuur 1: Vereenvoudigde voorstelling van een gebruiker dat films een score geeft en nieuwe aanbevelingen krijgt gebaseerd op deze ratings via een Content-Based systeem.*

Om een Content-Based recommender systeem te laten werken gaat onze “User Jan” uit [Figuur 1](#) eerst en vooral een goed profiel moeten vormen dat zijn smaak en voorkeur duidelijk weergeeft. Zoals eerder is aangehaald zou het systeem kunnen opmerken dat Jan meerdere actiefilms met Arnold Schwarzenegger een hoge score heeft gegeven. Content-Based systemen zullen deze trend opmerken en gaan hun grote databank van alle beschikbare films doorzoeken naar de tags “Actie” en “Arnold Schwarzenegger”. Eens het systeem nieuwe films vindt met deze tags kan het dit gaan voorstellen aan Jan als iets waar hij mogelijk interesse kan hebben.

### 4.2.1 Voor- en nadelen

Isinkaye, Folajimi, & Ojokoh halen aan als voordelen bij Content-Based recommender systems dat deze de mogelijkheid hebben om nieuwe items aan te bevelen, zelfs als er geen algemene beoordelingen zijn van gebruikers. Dus zelfs als de database geen algemene gebruikersvoorkeuren bevat, wordt de nauwkeurigheid van aanbevelingen niet beïnvloed. Ook als de gebruikersvoorkeuren veranderen, heeft deze de capaciteit om zijn aanbevelingen in een korte tijd aan te passen (Isinkaye, Folajimi, & Ojokoh, 2015).

Een groot probleem met dit systeem is dat het geen rekening houdt met hoe andere gebruikers iets scoren. Stel nu dat het recommender system heeft besloten dat een gebruiker grote fan is van Johnny Depp. Het systeem zal de film Dark Shadows (2012) voorstellen aan de user omdat Depp acteert in deze film, niet wetende dat de algemene ratings voor deze film zeer laag liggen. De kans dat de gebruiker toch geen fan is van de film zal zeer groot zijn waardoor met andere woorden het recommender system zijn job niet correct heeft verricht. Adomavicius & Tuzhilin gebruiken hiervoor de term “*Overspecialization*”. Zoals net voorgesteld in het voorbeeld met Johnny Depp is dit wanneer het systeem alleen items kan aanbevelen die scoren in hoge mate tegen het profiel van een gebruiker, is de gebruiker beperkt tot zijn aanbevolen items die vergelijkbaar zijn met de items die al zijn beoordeeld. Idealiter moet de gebruiker worden gepresenteerd met een diverse reeks opties en niet met een homogene set van alternatieven (Adomavicius & Tuzhilin, 2005).

Nog een mogelijk probleem dat kan opduiken tijdens het ontwerpen van een Content-Based recommender system volgens Adomavicius & Tuzhilin is dat deze systemen afhankelijk zijn van goed gelabelde data. Om goede aanbevelingen te kunnen maken moeten items in bezit zijn van een rijke verzameling aan beschrijvingen of tags. Hoe dieper de gebruiker kan ingaan op zijn smaak hoe makkelijker het wordt voor onze systemen om een goede aanbeveling te maken (Adomavicius & Tuzhilin, 2005). Bijvoorbeeld niet enkel het muziekgenre en de artiest weergeven als beschrijving bij een album maar ook de types van instrumenten om keywords over de inhoud van de songteksten. Dit zorgt ervoor dat je nog steeds aanbevelingen zal krijgen die gelijken van genre op je reeds beoordeelde items maar ook kunnen passen in het thema van liedjes waar je momenteel naar op zoek bent.

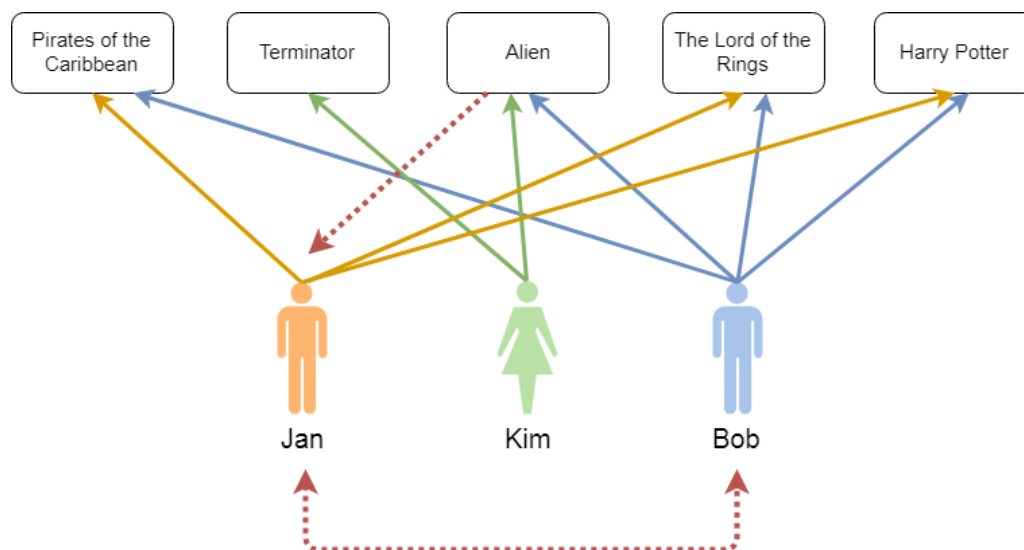
## 4.3 Collaborative Recommender Systems

In tegenstelling tot een content-based recommender system gaat een Collaborative systeem aanbevelingen maken gebaseerd op ratings van andere gebruikers met dezelfde smaak (Adomavicius & Tuzhilin, 2005). De GroupLens Research Group (2001) beschrijft de werken van een Collaborative Recommender System als volgt:

Recommender systems apply data analysis techniques to the problem of helping users find the items they would like to purchase at E-Commerce sites by producing a predicted likeliness score or a list of top-N recommended items for a given user. Item recommendations can be made using different methods. Recommendations can be based on demographics of the users, overall top

selling items, or past buying habit of users as a predictor of future items.  
(Sarwar, Karypis, Konstan, & Riedl, 2001)

In de context van een streaming service zal dit soort systeem dus aanbevelingen maken niet enkel gebaseerd op jouw smaak maar ook op de ratings van gelijkaardige gebruikers of algemeen hoog beoordeelde films of muziek. In plaats van gebruik te maken van labels of beschrijvende informatie bij items, zoals het genre van een film, die nodig waren bij hiervoor behandelde Content-Based technieken. Hebben we hier de algemene activiteit nodig van onze klanten om gelijkenissen te gaan ontdekken in kijk/luister patronen bij users. Eens er een link is gevonden in smaak tussen bijvoorbeeld twee gebruikers “Jan” & “Bob” kunnen er items worden aangeraden die onze gebruiker “Jan” bijvoorbeeld wel al heeft bekeken maar “Bob” nog niet kent. Dit wordt verduidelijkt via onderstaande [Figuur 2](#).



*Figuur 2: Vereenvoudigde voorstelling van een gebruiker dat films kijkt en nieuwe aanbevelingen krijgt gebaseerd op gelijkaardige gebruikers via een User-Based Collaborative systeem.*

[Figuur 2](#) toont ons 3 gebruikers; Jan, Kim en Bob. Er is een vergelijking gemaakt tussen Jan en Bob omdat ze beide Pirates of the Caribbean, The Lord of the Rings en Harry Potter gezien hebben. Bovenop deze 3 films heeft Bob ook nog eens Alien gezien. Aangezien het systeem een duidelijke overeenkomst heeft gevonden tussen de smaak en het kijkpatroon van deze twee gebruikers zal er via Collaborative filtering een aanbeveling worden gemaakt naar Jan om ook Alien te kijken. In tegenstelling tot Content-Based filtering(hoofdstuk 4.2) wordt er hier dus geen rekening gehouden met de specifieke beschrijving of tags van een item (film,album). Het enige wat we nodig hebben is een grote verzameling van gebruikers hun geschiedenis om goede aanbevelingen te kunnen maken op andere gelijkaardige gebruikers. Last.fm is bijvoorbeeld een streaming service die gebruik maakt van gelijkenissen tussen verschillende gebruikers. Enkele andere grote bedrijven die ook gebruik maken van Collaborative filtering zijn social media sites als Facebook en LinkedIn. Deze sites maken gebruik van je lijst vrienden en connecties om zo nieuwe aanbevelingen te kunnen maken.

Onderzoekers hebben een aantal Collaborative algoritmes ontwikkeld die kunnen worden opgedeeld in twee hoofdcategorieën: Memory-based en Model-based algoritmes (Sarwar, Karypis, Konstan, & Riedl, 2001). Deze worden hieronder verder besproken.

### 4.3.1 Memory-based Collaborative Filtering

Memory-based algoritmes zijn gebaseerd op voorspellingen van beoordelingen over de gehele verzameling van eerder beoordeelde items van de gebruikers (Sarwar, Karypis, Konstan, & Riedl, 2001).

De items die al door de gebruiker zijn beoordeeld, spelen een relevante rol bij het zoeken naar een gelijkaardige gebruiker. Zodra een overeenkomst tussen gebruikers is gevonden, kunnen verschillende algoritmen worden gebruikt om de voorkeuren van gelijkaardige gebruikers te combineren om aanbevelingen te genereren. Vanwege de effectiviteit van deze technieken hebben ze een groot succes bereikt in toepassingen in het echte leven. Memory-based Collaborative filtering kan op twee manieren worden bereikt door middel van user-based en op item-based technieken (Isinkaye, Folajimi, & Ojokoh, 2015).

#### 4.3.1.1 User-based

In vorig hoofdstuk, 4.3 Collaborative Recommender Systems, zijn we al in contact gekomen met User-based algoritmes in de vorm van [Figuur 2](#). Zoals de naam suggereert maken User-based algoritmes gebruik van gelijkenissen in de activiteit van gebruikers.

We hebben al uitgebreid de werking van deze technieken besproken in hoofdstuk 4.3. Hier is nog eens een kort en bondige samenvatting in de woorden van Isinkaye, Folajimi, & Ojokoh:

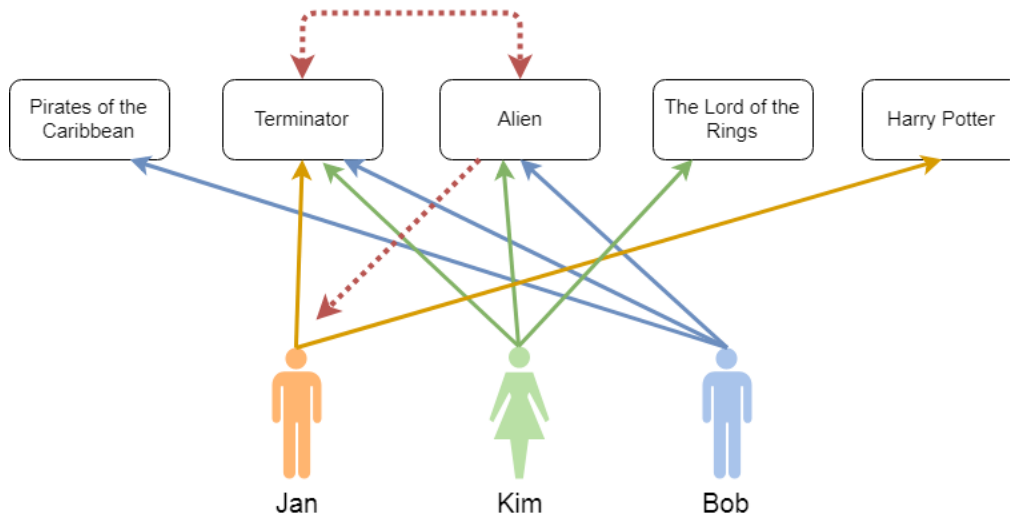
User based collaborative filtering technique calculates similarity between users by comparing their ratings on the same item, and it then computes the predicted rating for an item by the active user as a weighted average of the ratings of the item by users similar to the active user where weights are the similarities of these users with the target item. (Isinkaye, Folajimi, & Ojokoh, 2015)

Naast User-based filtering hebben we ook nog Item-based filtering.

#### 4.3.1.2 Item-based

In het vorige deel gebruikten we User-based Collaborative filtering om gelijkenissen te vinden tussen gebruikers. Voor Item-based Collaborative filtering gaan we het omgekeerde doen. Voorspellingen maken gebaseerd op de gelijkenissen tussen items in plaats van de gebruikers. Het bouwt een lijst van item overeenkomsten door alle items op te halen die door een actieve gebruiker worden beoordeeld uit de User-Item matrix, het bepaalt hoe vergelijkbaar de opgehaalde items zijn met het item waar we van willen te weten komen of we dit item moeten aanraden aan de gebruiker, waarna het de meest vergelijkbare items selecteert en de bijbehorende overeenkomsten ook in kaart brengt (Isinkaye, Folajimi, & Ojokoh, 2015).

De exacte wiskundige berekening achter deze techniek ligt buiten de scope van deze scriptie maar via onderstaande [Figuur 3](#) zullen we via een vereenvoudigd diagram de werking proberen verduidelijken.



*Figuur 3: Vereenvoudigde voorstelling van een gebruiker dat films kijkt en nieuwe aanbevelingen krijgt gebaseerd op gelijkaardige films via een Item-Based Collaborative systeem.*

Ons voorbeeld maakt weer gebruik van 3 users en hun beoordelingen op 5 verschillende items. Alle 3 de users hebben de film Terminator gezien en deze een goede score gegeven. Onze vraag wordt nu of we de film Alien moeten aanraden aan onze gebruiker Jan. Hiervoor gaan we gelijkenissen moeten zoeken in onze films. We kunnen zien dat Kim en Bob ook beide fan waren van Terminator maar daarnaast kunnen we ook zien dat ze beide fan waren van Aliens, ons doelwit. In deze kleine dataset van 5 items zal het systeem dit detecteren als grootste overeenkomst tussen de voorgeschotelde films waardoor we de film Alien kunnen voorstellen aan Jan met een grote zekerheid dat hij hier ook fan van zal zijn. Met andere woorden heeft het systeem een verband kunnen leggen tussen de twee films Terminator en Alien dankzij het gebruik en de ratings van andere users.

### 4.3.2 Model-based Collaborative Filtering

Volgens Bergner, Yoav, et al (2012) is het idee achter een Collaborative filtering systeem dat wanneer meerdere gebruikers communiceren met overlappende subsets van items, informatie uit de interacties kan worden geëxtraheerd en gebruikt om probabilistische voorspellingen te maken over mogelijke toekomstige interacties. Memory-based systemen proberen dit te doen door gebruik te maken van gelijkenis tussen gebruikers op basis van hun eerdere interacties zoals hierboven al is aangehaald. In onze voorbeelden [Figuur 2 & 3](#) zien we hoe een algoritme kan voorspellen dan onze user Jan identieke interacties zal hebben in de toekomst als de meest vergelijkbare user Kim (of Bob). Deze beschrijvende aanpak zal niet proberen om zin te maken van deze interacties zelf. Daarentegen, maakt Model-based Collaborative filtering gebruik van deze informatie verborgen in de interacties om ze een reeks parameters te modelleren voor de users en items die, wanneer ze samen worden gebracht, een probabilistische voorspelling kunnen maken over de



ontbrekende interacties (Bergner, Yoav, et al, 2012). Het bouwen van dit model kan gedaan worden via machine learning of data mining technieken. Voorbeelden van deze technieken zijn onder meer: Singular Value Decomposition (SVD), Matrix Completion Technieken, Latent Semantic methodes, en Regression & Clustering (Isinkaye, Folajimi, & Ojokoh, 2015). Deze technieken vallen buiten de scope van deze scriptie maar het is belangrijk om te weten dat al deze mogelijke oplossingen bestaan.

Het idee achter een Model-based systeem is dus om vooraf een model, een soort template, te voorzien waar je data op zal kunnen passen om gelijkenissen te vinden en voorspellingen te maken. De beschikbaarheid van dit model zorgt ervoor dat onze processen veel sneller en real-time zullen werken. Het gebruik van een model zorgt ook tot de mogelijkheid van het gebruiken van een grotere schaal aan data (Sarwar, Karypis, Konstan, & Riedl, 2001).

### 4.3.3 Voor- en nadelen

In tegenstelling tot Content-based systemen heeft Collaborative filtering veel minder tags en gelabelde data nodig om accurate voorspellingen te maken. Collaborative systemen hebben enkel user ratings nodig om goed te presteren (Lee, Sun, & Lebanon, 2012). Mede hierdoor zijn Collaborative technieken het meest voorkomend momenteel in de wereld van Recommender systems.

Collaborative Recommender systems zijn zeer effectief in het genereren van een gepersonaliseerde ervaring maar deze techniek lijdt volgens Wu, Liu, Xie, Ester, & Yang (2016) onder het moeilijk schalen met een grote dataset (p. 85). Scalability is een probleem dat overal in de data wereld vaak opduikt en ook hier kan het voor problemen zorgen. Als de dataset blijft groeien kan niet enkel de performance van onze systemen achteruitgaan maar moeten de aanbevelingen worden geüpdatet mochten er nieuwe user-groepen ontdekt worden.

Naast scalability halen Isinkaye, Folajimi, & Ojokoh (2015) ook nog 3 andere potentiële problemen die kunnen opduiken bij Collaborative Recommender systems: Het Cold-start probleem, het Data sparsity probleem en Synonymy.

Het Cold-start probleem valt voor wanneer het niet mogelijk is om voorspellingen te maken door een gebrek aan ratings. Dit kan voorvallen door het aanmaken van een nieuwe gebruiker dat nog niet de kans heeft gehad om items een rating te geven maar ook zien we dit soms gebeuren bij het toevoegen van een nieuw item of wanneer de community rond het systeem nog relatief nieuw is en dus nog niet genoeg input heeft kunnen vergaren (Bobadilla, Ortega., Hernando, & Bernal, 2012). Dit sluit aan bij het Data sparsity probleem wat inhoudt dat er te weinig data aanwezig is om accurate voorspellingen te maken. Dit kan leiden tot bepaalde items waar het niet voor mogelijk zal zijn om aanbevelingen te maken.

Het laatste probleem Synonymy wordt door Isinkaye, Folajimi, & Ojokoh gedefinieerd als volgt:

Synonymy is the tendency of very similar items to have different names or entries. Most recommender systems find it difficult to make distinction between closely related items such as the difference between e.g. baby wear and baby

cloth. Collaborative Filtering systems usually find no match between the two terms to be able to compute their similarity. (Isinkaye, Folajimi, & Ojokoh, 2015)

Deze problemen kunnen een grote negatieve impact hebben op onze voorspellingen en meestal is er niet meteen een concrete oplossing voor deze problemen. Er zijn verscheidene fixes en workarounds gevonden over de jaren door ontwerpers van Recommender systems maar in deze scriptie gaan we hier niet verder op ingaan. Zelfs al met deze mogelijke nadelen zijn Collaborative Recommender systems nog steeds onder de meest gebruikte en succesvolle systemen tot op heden.

## 4.4 Hybrid Systemen

Zoals hierboven net is aangehaald kunnen Recommender systems vaak problemen met zich mee brengen. Hybrid Recommender systems zijn één van de mogelijke oplossingen voor het verbeteren van deze issues. Hybrid Recommender systems zijn systemen die twee of meer van de hierboven beschreven technieken combineren om de prestaties van de aanbevelingen te verbeteren. Zo is het mogelijk om twee verschillende Content-based systemen te combineren of een combinatie te gebruiken van Content-based en Collaborative Filtering om elkaars zwaktes tegen te gaan.



## 5 Creëren van een gepersonaliseerde ervaring

Met de spontane opkomst en groei van online webshops, streaming services en andere online diensten over de laatste jaren is het nodig voor deze bedrijven om klanten een goede reden te geven om voor hen dienst te kiezen. Een van de manieren waarop bedrijven meerwaarde proberen te geven aan hun diensten is door het personaliseren van hun service gericht op te wensen van de klant. Met andere woorden willen ze proberen een dienst te voorzien die specifiek is afgestemd om elke individuele klant of groep van klanten.

Columnist Nick Iyengar (2017) stelt het voordeel van personalisatie voor als volgt: Als we onze bezoekers / gebruikers / klanten meer relevante, nuttige ervaringen kunnen aanbieden, zouden ze hierop moeten reageren door dieper in te gaan op deze dienst en vaker te kiezen om te abonneren.

De Recommender systems die we daarnet uitgebreid hebben besproken in ons Theoretisch kader zijn hiervoor perfect geschikt. Als we bijvoorbeeld in onze webshop enkel items kunnen laten zien aan onze klanten waarvan we weten dat deze hier interesse in heeft. Dit soort voorspellingen kunnen gemaakt worden aan de hand van Recommender system maar enkel wanneer een bedrijf beschikt over genoeg data. Zowel data over hun eigen producten als data over de gebruikers en algemene trends zijn nodig voor een Recommender system om goede voorspelling te maken.

Het opstellen van een Recommender system kan moeilijker zijn dan gedacht. Vooraleer een bedrijf deze technieken begint te overwegen is het nodig om te weten wat men er specifiek mee wilt bereiken. Het business model van een bedrijf en het uiteindelijke doel van wat ze willen dat klanten gaan doen moet de basis zijn van het Recommender Systeem. Een dit vast staat kan er verder gekeken worden naar de verschillende werkmethodes en types van systemen die we hebben aangehaald in Hoofdstuk 4 van deze scriptie.

In het volgend hoofdstuk gaan we analyseren hoe 3 van de grootste Online Streaming Services, namelijk Netflix, Youtube en Spotify, gepersonaliseerde ervaringen proberen te creëren en tot in welke mate deze wel degelijk effectief zijn in het behouden van de interesse van klanten.

## 6 Use Cases

In de volgende secties zullen we gebruik maken van bovenstaande literatuuronderzoek om een analyse uit te voeren op 3 populaire Streaming Services: Netflix, Youtube en Spotify. Eerst en vooral moeten we een goed beeld hebben van de business models van deze bedrijven en wat zij verwachten of nodig hebben van hun klanten om succesvol te zijn. Welke meerwaarde heeft een Recommender system voor deze bedrijven en waarop worden deze gebaseerd?

We gaan de verschillende Recommender system methodologieën vergelijken met hun Business model met een nadruk op het creëren van een gepersonaliseerde ervaring. Aan de hand van een zelf georganiseerde survey gaan we ook trachten na te gaan hoe succesvol deze bedrijven hierin zijn.

### 6.1 Netflix

#### 6.1.1 Business model

Netflix is een online streaming service en content provider dat gebruikers in de vorm van subscriptions in staat stelt om hun grote verzameling van films en TV series te bekijken op verschillende platformen zoals Tablets, TVs, PCs en game consoles. Netflix biedt deze dienst aan in de vorm van een maandelijks abonnement waarin gebruikers gedurende deze tijd on-demand, zonder reclame, kunnen kijken naar hun favoriete series en films. Onderstaande [figuur 4](#) is een voorbeeld van hoe Netflix hun abonnementen werken.

STAP 1 VAN 3

### Kies een abonnement dat bij je past.

Downgrade of upgrade wanneer je wilt

	Basic	Standaard	Premium
Prijs per maand nadat de gratis maand is afgelopen op 9/07/18	7,99 €	10,99 €	13,99 €
HD beschikbaar	×	✓	✓
Ultra HD beschikbaar	×	×	✓
Aantal schermen dat je tegelijkertijd kunt gebruiken	1	2	4
Kijk op je laptop, tv, telefoon en tablet	✓	✓	✓
Eindeloos veel series en films	✓	✓	✓
Altijd opzegbaar	✓	✓	✓
Eerste maand gratis	✓	✓	✓

Beschikbaarheid van HD en Ultra HD is onderworpen aan je internetprovider en de mogelijkheden van het apparaat. Niet alle content is beschikbaar in HD of Ultra HD. Ga naar [Gebruiksvoorwaarden](#) voor meer details.

DOORGAAN

source: netflix

**Figuur 4:** De verschillende abonnementen dat Netflix aanbiedt inclusief specifieke abonnement bonussen.

Het algemene doel van Netflix is dus om zoveel mogelijk huidige abonnees te verkrijgen. Zowel het behouden van huidige klanten als het verkrijgen van nieuwe gebruikers is van groot belang voor het bedrijf om geld te maken. Zoals hierboven is aangehaald staat Netflix standvast dat ze geen pay-per-view of ad-supported content business models willen gebruiken. Netflix beschrijft hun focus als volgt:

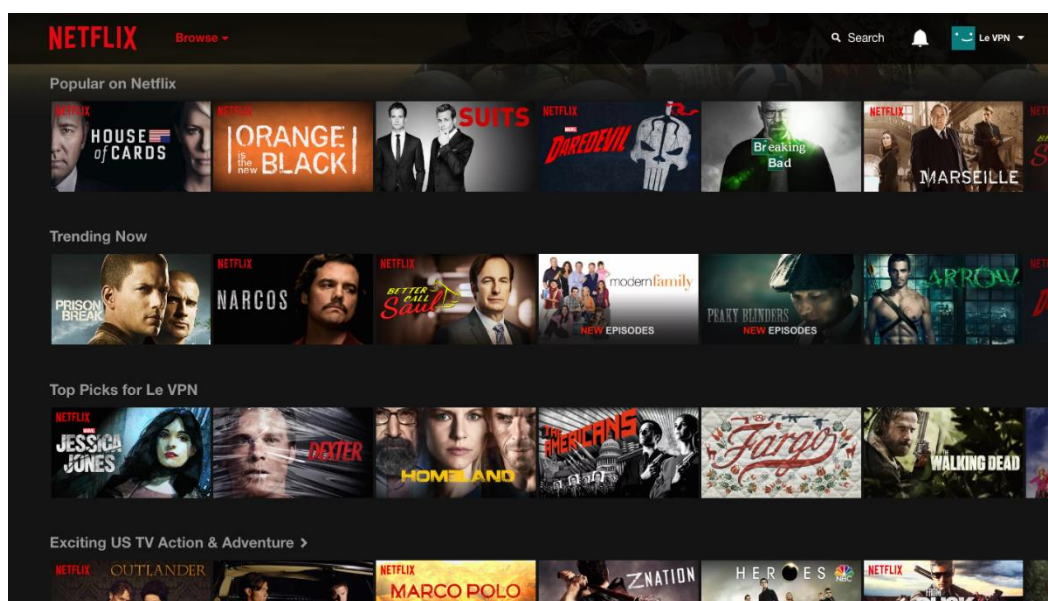
We are about the freedom of on-demand and the fun of binge viewing. We are about the flexibility of any screen at any time. We are about a personal experience that finds for each person the most pleasing titles from around the world. To deliver this experience to our members, we expect to spend around \$1.3 billion on technology & development in 2018. (Netflix, 2018)

Het doelpubliek van Netflix zijn dus mensen die graag hun series en films on-demand willen kijken op hun platform naar keuze in plaats van te moeten wachten op reclames en wekelijkse releases. Naast hun on-demand service is hun immens grote verzameling aan content ook zeker een meerwaarde voor klanten om te kiezen voor hun diensten. Netflix produceert ondertussen ook hun eigen series die enkel verkrijgbaar zijn op hun platform wat nogmaals klanten een goede reden geeft om te investeren in Netflix als Streaming service.

Naast deze Value propositions ziet Netflix hun Recommender system ook aan als een van de grote redenen waarom klanten voor hun dienst moeten kiezen en mogelijks langer geïnteresseerd en klant zullen blijven. We gaan nu analyseren op welke manier ze dit ondernemen en of dit wel degelijk zoveel meerwaarde geeft als Netflix zegt.

### 6.1.2 Recommender systems

De internet-tv-ruimte is jong en de concurrentie is rijp, dus innovatie is van cruciaal belang. Een belangrijke pijler van Netflix volgens Gomez-Urbe & Hunt (2015) is het aanbevelingssysteem dat leden helpt om nieuwe video's te vinden in elke kijk sessie. Het Netflix Recommender systeem is niet één algoritme, maar eerder een verzameling van algoritmes die verschillende use cases uitwerken die dan samen komen om de complete Netflix-ervaring te creëren.



source: Netflix, [whatisvpnandhowdoesitworks.blogspot.com](http://whatisvpnandhowdoesitworks.blogspot.com)

*Figuur 5: Een lijst van gepersonaliseerde aanbevelingen op de Netflix homepage.*

Bovenstaande [Figuur 5](#) toont ons 4 verschillende soorten aanbevelingen die Netflix maakt voor zijn gebruikers. Gomez-Urbe & Hunt delen de gebruikte types Recommender systemen in als 4 algemene technieken in hun artikel *The Netflix Recommender System: Algorithms, Business Value, and Innovation* (2015): Personalized Video Ranker (PVR), Top-N Video Ranker, Trending Now, Because You Watched en Video-Video Similarity (Gomez-Urbe & Hunt, 2015). We gaan in de volgende secties deze 4 systemen verder analyseren en vergelijken met de Recommender Systems dat we hebben besproken in Hoofdstuk 4 van deze scriptie.

#### 6.1.2.1 Personalized Video Ranker (PVR)

Het PVR algoritme maakt gebruik van de tags en beschrijvingen die samenhangen met de verschillende films om en gepersonaliseerde aanbeveling te maken gebaseerd op bijvoorbeeld een specifiek genre. Een voorbeeld hiervan in [Figuur 5](#) is de rij "Exciting US TC Action & Adventure". In de context van het theoretisch kader van deze scriptie

(Hoofdstuk 4) kunnen we deze techniek vergelijken met Content-based Recommender systems (4.2). Over tijd zal PVR zowel kijkgeschiedenis en persoonlijke ratings gebruiken om deze dan te vergelijken met films die passen in de gekozen zoekterm, zoals het genre. Eens deze vergelijkingen zijn gemaakt kunnen we deze gebruiken om nieuwe films te gaan voorspellen aan users.

### 6.1.2.2 Top-N Video Ranker

Top-N ranker combineert persoonlijke ratings met de meest populaire films in de Netflix catalogus om de rij “Top Picks” voor te stellen in [Figuur 5](#).

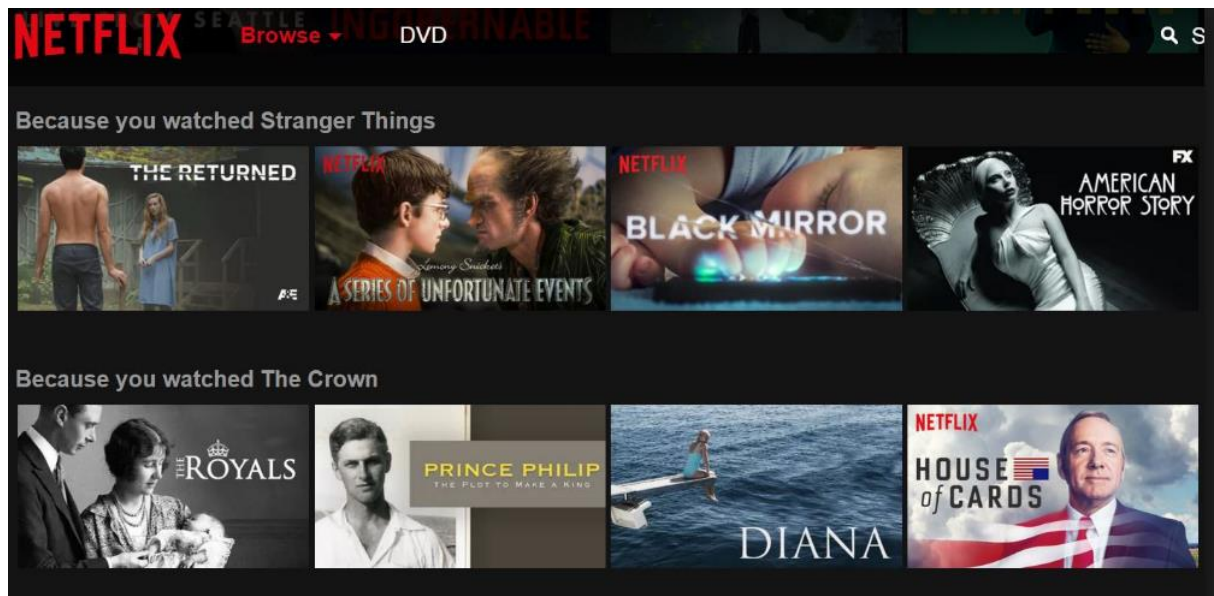
Met andere woorden, het richt zich op het hoofd van de ranglijst en selecteert films die het beste passen bij voorkeuren van de gebruiker (Goltsman, 2017).

Net zoals PVR maakt dit gebruik van Content-based Recommender technieken om een gepersonaliseerde voorspelling te maken maar nu met de focus op de meest populaire films.

### 6.1.2.3 Trending Now

Trending Now is in essentie gelijk aan de Top-N ranker maar de focus ligt hier op seizoensgebonden aanbevelingen. Hier zullen gebruikers bijvoorbeeld meer horror films te zien krijgen tijdens de periode rond halloween. Bepaalde genres op tags die tijdsgebonden zijn worden vergeleken met de voorkeuren van gebruikers om tijdens deze periodes interessante nieuwe aanbevelingen te maken.

### 6.1.2.4 Because You Watched & Video-Video Similarity



source: Netflix

Figuur 6: Een weergave van Because You Watched Recommendations op Netflix.

**Figuur 6** toont ons het “Because You Watched” systeem dat Netflix gebruikt om aanbevelingen te maken gebaseerd op 1 specifieke video die een gebruiker in het verleden heeft gezien. Bij Netflix refereren ze naar deze techniek als het Video-Video Similarity algoritme, of makkelijker gezegd “sims” (Gomez-Urbe & Hunt, 2015). Dit systeem zoekt naar films die lijken op elkaar en houdt in eerste instantie geen rekening met de voorkeur van de gebruiker. Pas later wanneer deze aanbevelingen worden gemaakt aan de gebruiker zal zijn voorkeur ook worden toegepast op wat hij uiteindelijk te zien krijgt. Deze techniek hebben we ook al aangehaald in het Theoretisch kader van deze scriptie, namelijk Item-based Collaborative systems (hoofdstuk 4.3.1.2). Het maakt een vergelijking puur op de films zonder rekening te houden van de algemene ratings.

### 6.1.3 Survey

Volgende surveys zijn opgesteld om de algemene ervaring van Netflix gebruikers in kaart te brengen in de context van hun gepersonaliseerde ervaringen dat Netflix aanbiedt zoals aangehaald in de vorige sectie. Deze surveys hebben gedurende anderhalve maand lang data vergaard over de mening van Netflix gebruikers.

In **Figuur 7** gaan we na hoe vaak gebruikers een rating achter laten op films of series die ze bekeken hebben.



**Figuur 7:** Een survey op het geven van ratings op Netflix.

Uit deze survey blijkt nu dat 38% van gebruikers uit onze test data een rating achterlaten.

Een meer interessante opmerking is dat 12% van de gebruikers ervan overtuigd is dat ze nooit ratings achterlaten op films. Dit betekent dat Netflix mogelijk 12% aan data verliest om betere voorspellingen te maken. Deze 12% verliest mogelijk ook veel kansen om zelf een accurate gepersonaliseerde ervaring te beleven op Netflix.



**Figuur 8:** Beoordeling van de gepersonaliseerde aanbevelingen.

Maar als we dan verder gaan kijken naar **Figuur 8** zien we dat dit toch niet het geval blijkt te zijn. Maar 4% beweerd nog nooit een goede aanbeveling gekregen te hebben terwijl 18% nog niet volledig verkocht blijkt te zijn om de systemen maar al wel eens een goede aanbeveling heeft gekregen. De algemene perceptie over de accuraatheid van Netflix hun Recommender systems blijkt zeer hoog te liggen uit deze data.



#### 6.1.4 Besluit

Uit deze analyse op de gebruikte Recommender systems van Netflix en de algemene indruk op de gebruikers aan de hand van bovenstaande surveys hebben we een beter beeld gekregen op waarom Netflix deze technieken toepast en of deze wel degelijk effectief zijn.

Netflix neemt het aanbevelen van nieuwe films en series zeer slim aan. Ze maken gebruik van een Hybrid systeem waarin ze zowel gebruiken maken van Collaborative filtering als Content-based filtering. Via dit Hybride systeem zijn ze in staat op meerdere niveaus aanbevelingen te maken. Dit gaat van films die een grote gelijkenis hebben met elkaar tot het aanbevelen van romantische films tijdens de periode van valentijn. In [Figuur 5](#) zagen we hoe Netflix al meteen via de hoofdpagina in staat is om op eerste zicht meer dan 20 aanbevelingen te maken, dit zonder zelfs nog maar te hoeven verder navigeren door een menu of lijst.

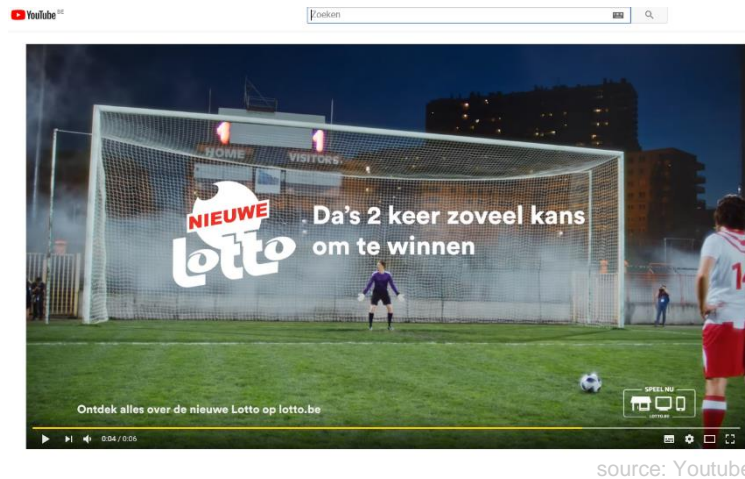
Aangezien Netflix een stabiele lijst aan abonnees nodig heeft om winst te kunnen maken zagen zij Recommender systems en een gepersonaliseerde dienst aan als een grote meerwaarde om klanten te lokken maar ook te behouden. Uit onze data vergaard uit de survey merken we nu ook wel degelijk mensen algemeen tevreden zijn over hun gepersonaliseerde streaming service.

## 6.2 Youtube

### 6.2.1 Business model

We hebben net gezien bij de Netflix use case dat zij er alles aan proberen te doen om een reclameloze streaming service aan te bieden. Youtube werkt net tegenovergesteld tot deze methodologie. YouTube is nu 's werelds grootste website voor het delen van video's en de op één na grootste zoekmachine op het internet. Elke dag worden er meer dan een miljard uren besteed aan het bekijken van YouTube-video's, er wordt elke minuut meer dan 400 uur aan inhoud geüpload, het platform is beschikbaar in 76 verschillende talen en wordt niet alleen gebruikt door het gewone volk maar ook door grote merken, beroemdheden en veel andere YouTube-exclusieve beroemdheden om zichzelf te promoten en om geld te verdienen via YouTube (Dutta, 2017).

Het grootste deel van alle content op Youtube staat gratis beschikbaar voor iedereen die wil. Om toch nog winst te kunnen maken zorgt dit ervoor dat Youtube genoodzaakt is om advertenties te plaatsen voor- en tijdens videos. Het doel hier zal zijn gebruikers zo lang mogelijk te laten kijken naar video's en zo vaak mogelijk ervoor te laten zorgen dat gebruikers zullen blijven verder klikken op nieuwe video's.



source: Youtube

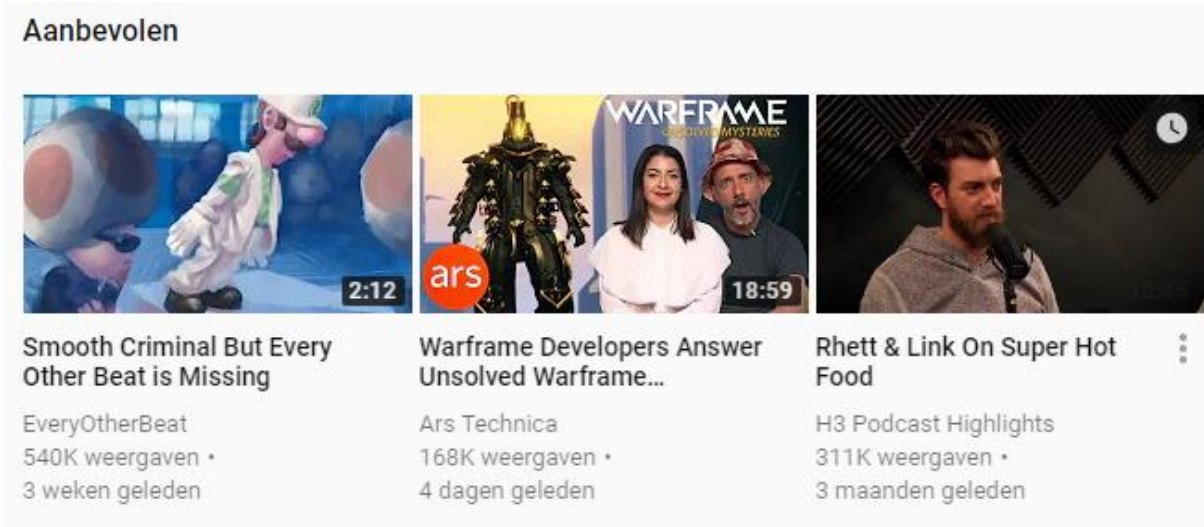
*Figuur 9: Voorbeeld van een advertentie voor een Youtube video. De gele balk onderaan geeft weg dat er een advertentie bezig is.*

Net zoals Netflix biedt Youtube ook een maandelijks abonnement aan genaamd Youtube Red. De Value propositions voor deze maandelijks subscription is het bekijken van video's zonder ooit advertenties te krijgen tijdens of voor de video's. Ook geeft Youtube Red membership toegang tot bepaalde exclusieve content enkel verkrijgbaar via dit model.

## 6.2.2 Recommender systems

Omdat Youtube beschikt over zoveel data in de vorm van verschillende video's duikt er al meteen een groot probleem op voor het toepassen van Recommender systems op al deze hopen data. In hoofdstuk 4.3.3: Voor- en nadelen van Collaborative filtering systemen hadden we het al reeds over hoe Scalability een groot probleem kan vormen. Youtube lost dit op door eerst en vooral al deze video's te beginnen opdelen in verschillende kleine clusters van video's die op elkaar lijken of over hetzelfde onderwerp gaan. Daarna gaat Youtube beginnen met deze clusters te vergelijken met jouw profiel gebaseerd op kijkgeschiedenis maar ook zal er worden rekening gehouden met welke video's je in het verleden hebt geliked. Ook zal het rekening houden met persoonlijke informatie als geslacht, leeftijd en land wanneer deze beschikbaar is om mensen met een gelijkaardig profiel te kunnen vinden (Rodriguez, 2017). Via al deze informatie en het verkleinen van hun dataset zal Youtube al snel in staat zijn om gelijkaardige video's voor te stellen gericht op jouw voorkeuren.



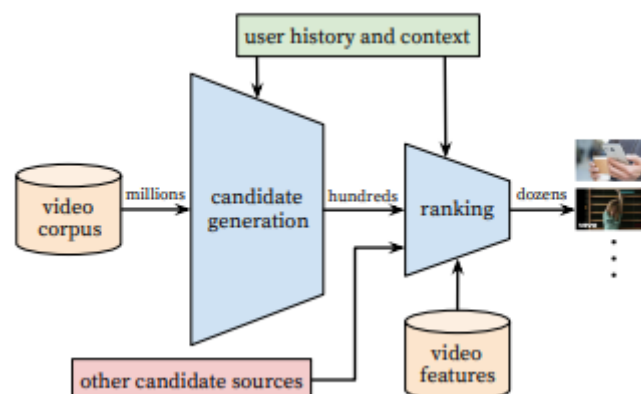


source: Youtube

*Figuur 10: Aanbevolen video's van Youtube op de home-page.*

Als we even teruggaan naar het Theoretisch kader voor deze scriptie kunnen we de observatie maken dat Youtube gebruik maakt van User-based Collaborative filtering om een gelijkenis te vinden tussen gebruikers. Maar ook Item-based Collaborative filtering gebruikt om deze clusters aan gelijkaardige video's te vinden in hun data.

Google engineers stellen de Youtube Recommender systems voor als een Deep Collaborative filtering model dat effectief in staat is tot het vinden van verborgen structuren en linken in data om zo goede voorspellingen te kunnen maken (Covington, Adams, & Sargin, 2016). Ze geven ons ook een vereenvoudigde representatie, in de vorm van onderstaande [Figuur 11](#), van de data "funnel" waar mogelijke video's worden opgehaald en geranked vooraleer ze worden voorgesteld aan de gebruiker.

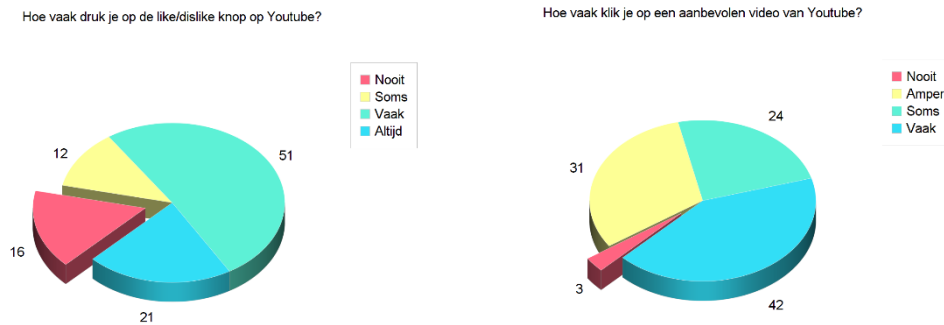


source: (Covington, Adams, & Sargin, 2016)

*Figuur 11: De Recommendation system architectuur voorgesteld door google engineers Covington, Adams, & Sargin.*

### 6.2.3 Survey

Volgende surveys zijn opgesteld om de algemene ervaring van Youtube gebruikers in kaart te brengen in de context van hun gepersonaliseerde ervaringen dat Youtube aanbiedt zoals aangehaald in de vorige sectie. Deze surveys hebben gedurende anderhalve maand lang data vergaard over de mening van Youtube gebruikers.



*Figuur 12 (links): Het aantal mensen dat gebruik maakt van de like knop.  
Figuur 13 (rechts): Hoe vaak mensen verder klikken op aanbevolen video's.*

Uit deze surveys blijkt dat net zoals bij Netflix het grootste aantal users een rating achter laat na ze een video hebben bekeken, dit zien we in [Figuur 12](#). Dit is een goed teken voor Youtube aangezien dit uiteindelijk zal zorgen voor een dieper en breder Recommender system. Uit [Figuur 13](#) leren we dat het grootste aantal gebruikers ook wel degelijk vaak gebruik gaan maken van de aanbevelingen van Youtube. Maar liefst 42% van de test-groep klikt vaak op een aanbevolen video uit nieuwsgierigheid of interesse met maar 3% dat nooit gebruik maakt van deze dienst.

Ten laatste willen we ook nog te weten komen hoe accuraat de Recommender systemen van Youtube zijn in het aanbevelen van video's. In [Figuur 14](#) zien we dat de algemene ervaring met aanbevolen video's goed tot zeer aangenaam is. Met maar 1% van de test-groep waar Youtube nog niet in geslaagd is om een gepersonaliseerde ervaring aan te bieden. Een reden hiervoor kan zijn dat deze gebruikers pas net een account hebben aangemaakt en dat het systeem dus lijdt aan het Cold-start problem (Hoofdstuk 4.3.3) of dat deze gebruikers nog geen account hebben aangemaakt waardoor het moeilijk wordt om gebruiker gericht aan de slag te gaan.



*Figuur 14: De tevredenheid van gebruikers en hun aanbevolen video's.*

## 6.2.4 Besluit

In tegenstelling tot Netflix is het doel van Youtube niet gericht op het behouden en verkrijgen van abonnees. Het doel van Youtube is om zoveel mogelijk mensen te laten klikken op video's om geëngageerd actief te blijven op hun platform dat hen uiteindelijk revenue oplevert in de vorm van advertenties voor en tijdens sommige video's.

De manier waarop Youtube Recommender systems gebruikt is door het aanmaken van een lijst vol aanbevolen video's die gepersonaliseerd zijn op jouw voorkeur, kijkgeschiedenis en jouw demografie. Hiervoor maken ze gebruik van een Hybrid Collaborative filtering systeem dat zowel gebruikers met elkaar vergelijkt als video-video gelijkenissen probeert te vinden in hun catalogus om zo de correcte video's te kunnen aanbevelen gepersonaliseerd voor de gebruikers.

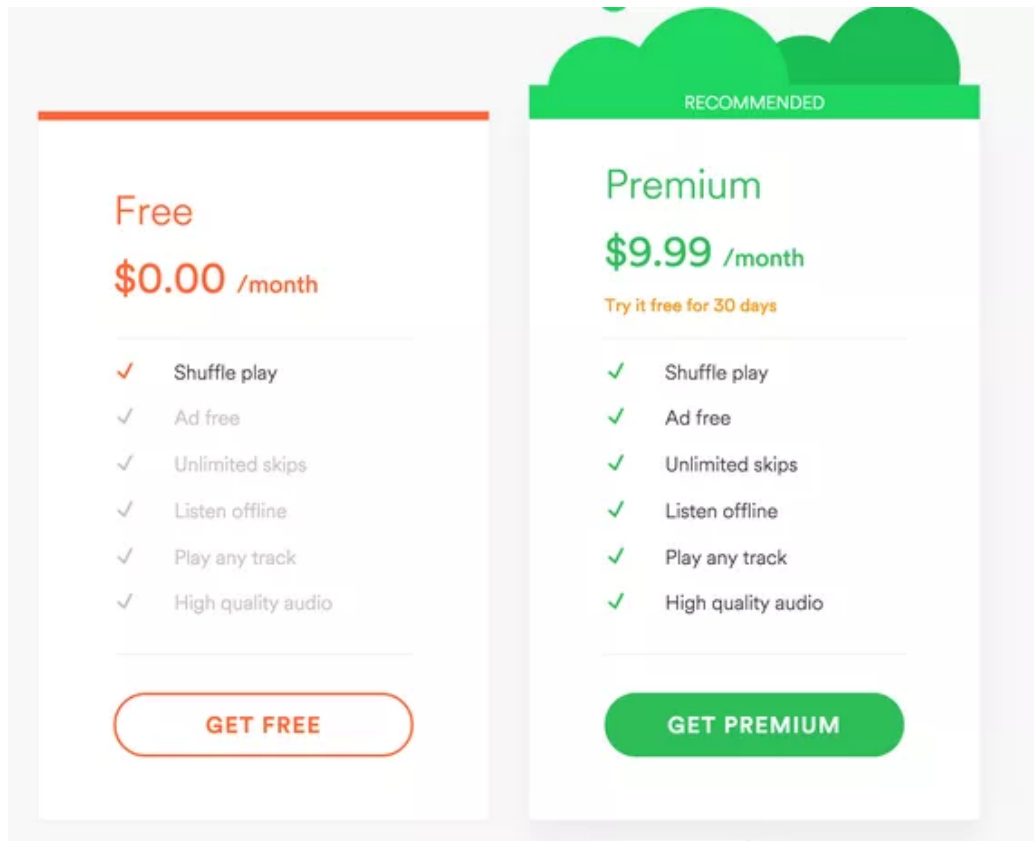
Het aanbod van video's dat Youtube ons aanbiedt is onwaarschijnlijk groot en blijft groeien met de dag. Dit betekent dat Scalability een groot probleem is voor Youtube en hun Recommender system. We hebben dit veel voorkomend probleem al eerder besproken in hoofdstuk 4.3.3. De oplossing die Youtube hiervoor gevonden heeft is het gebruik maken van Deep Learning technieken om data samen te clusteren en deze cluster steeds meer te verkleinen na rekening te houden met specifieke informatie die ze over tijd hebben samengesteld over de gebruiker. Dit is hierboven te zien in [Figuur 11](#).

Uit onze surveys blijkt dan ook dat Youtube zeer competent is in het creëren van een gepersonaliseerde ervaring. Met het grootste aantal van de gebruikers dat al goede ervaringen heeft gehad met de aanbevolen video's dat dus uiteindelijk meer clicks en views oplevert voor Youtube.

## 6.3 Spotify

### 6.3.1 Business model

Spotify gebruikt een mix van de vorige Business models die we gezien hebben bij Netflix en Youtube. Net zoals Youtube biedt het hun grote muziekcatalogus gratis aan met hun revenue stream hier zijnde dat deze gebruikers advertenties zullen te horen krijgen. De grootste omzet bij Spotify is eerder wel hun Premium abonnement dat vergelijkbaar is met hoe Netflix werkt. Gebruikers met een Premium abonnement krijgen geen advertenties te horen samen met nog enkele voordelen zoals de mogelijkheid om te luisteren naar je muziek offline en nog andere voordelen die hieronder te zien zijn in [Figuur 15](#).



source: Feedough, Spotify

*Figuur 15: De voordelen van een Premium Spotify abonnement.*

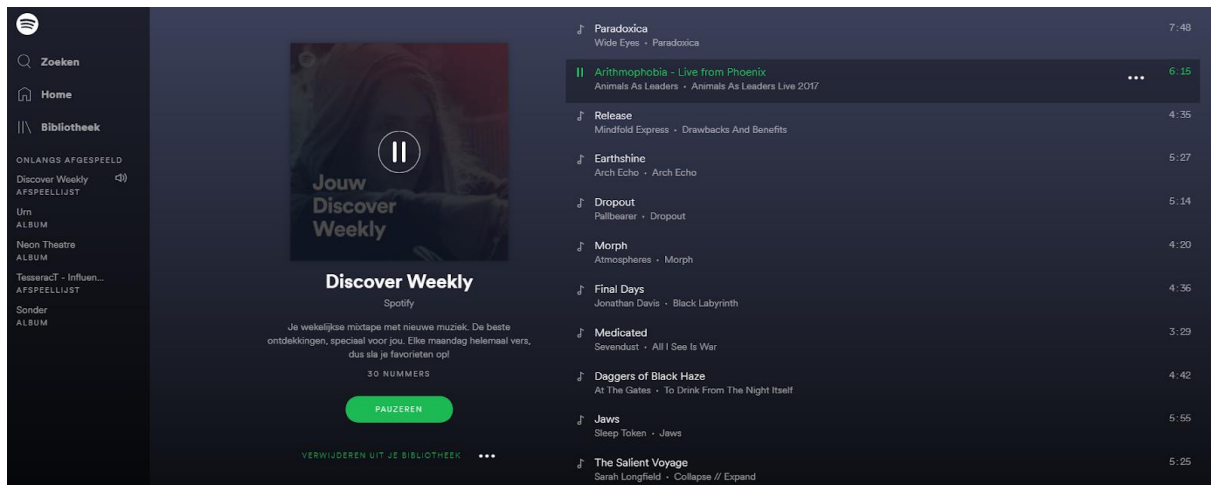
De manier waarop de gratis gebruikers advertenties te zien of te horen krijgen op Spotify gebeurt op een aantal verschillende manieren die zijn onderzocht in een Brand Study van Rameez M Sydeek (2018) op het entrepreneur platform Feedough. Deze advertenties maken gebruik van technieken als Recommender systems om gepersonaliseerde reclame voor te stellen aan hun gratis gebruikers. Hij haalt volgende manieren van reclame aan gebruikt door Spotify:

Spotify supports different kinds of advertisements: Sponsored Playlists, Branded Moments, Sponsored Sessions, Audio, Video Takeovers, Display, Homepage Takeovers, Overlay, Advertiser Pages and Branded Playlists. These ads vary in type, size and user engagement. (Sydeek, 2018)

We gaan niet verder ingaan op deze specifieke manieren van advertenties maar in de context van deze scriptie is het interessant om te weten dat de besproken technieken hier niet enkel kunnen gebruikt worden voor het aanbevelen van video's of muziek maar ook voor het aanbieden van gepersonaliseerde reclame.

### 6.3.2 Recommender systems

Spotify biedt een gepersonaliseerde ervaring aan in de vorm van het Discover Weekly systeem dat te zien is in onderstaande [Figuur 16](#).



source: Spotify

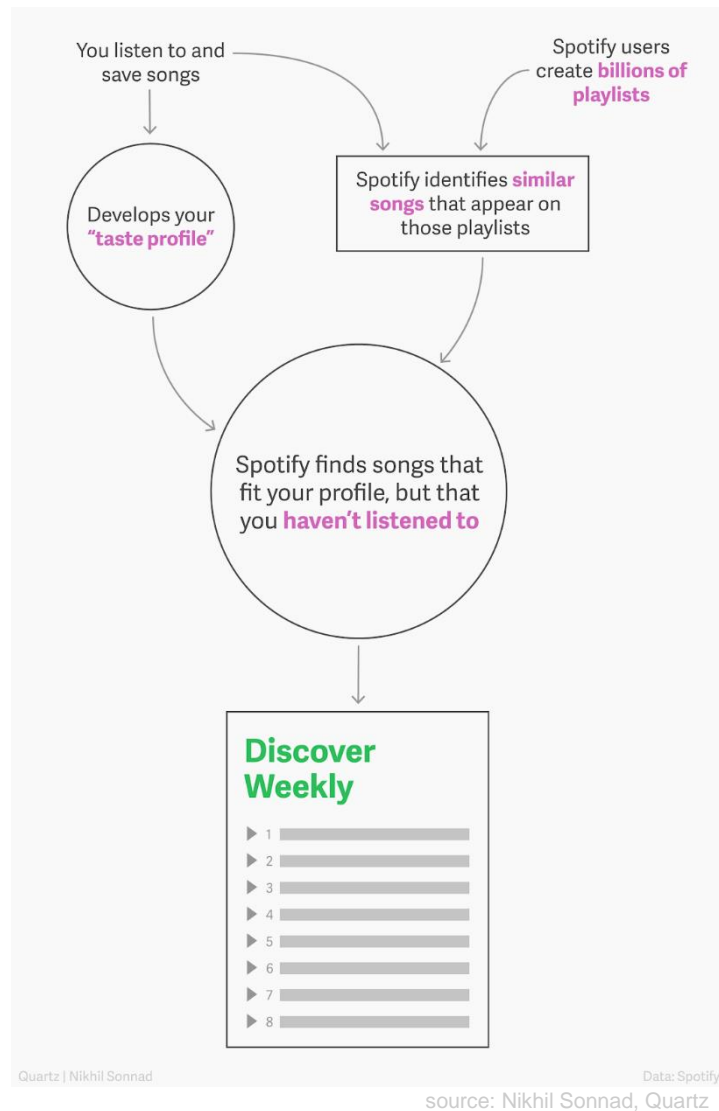
*Figuur 16: Spotify's Discover Weekly playlist.*

Dit systeem stelt wekelijks een aantal artiesten en liedjes voor die je nog nooit eerder hebt beluisterd op hun site gebaseerd op de specifieke voorkeur van een gebruiker en de muziek waar iemand naar heeft geluisterd in het verleden.

“We now have more technology than ever before to ensure that if you’re the smallest, strangest musician in the world, doing something that only 20 people in the world will dig, we can now find those 20 people and connect the dots between the artist and listeners”, is een quote van Matthew Ogle, in een interview met Quartz (Pasick, 2015).

Zoals de quote van Ogle aanhaald zijn hun systemen in staat om gelijkenissen te vinden in zowel gebruikers als muzikanten om tot in zeer specifieke sub-genres en niet bekende artiesten nog steeds goede aanbevelingen te maken. Met andere woorden maakt Spotify dus net zoals Youtube gebruik van een Hybrid Collaborative Recommender System om hun voorspellingen te maken.

Naast informatie van artiesten en gebruikers maakt Spotify ook gebruik van informatie als muziekblogs en playlists aangemaakt door andere gebruikers om patronen te vinden in data. Dit gebeurt via Natural Language Processing (Pasick, 2015). Dit is het vinden van betekenissen in geschreven en gesproken tekst wat dan ook weer gebruikt kan worden om gelijkenissen te vinden in data. Onderstaande flowchart toont aan hoe Spotify in staat is goede aanbevelingen te maken (Figuur 17).

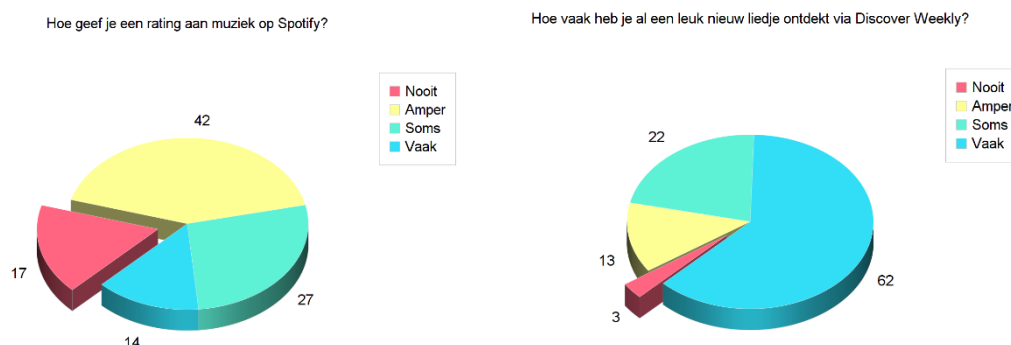


Figuur 17: Spotify's flowchart van het aanbevelen van nieuwe muziek.

### 6.3.3 Survey

Volgende surveys zijn opgesteld om de algemene ervaring van Spotify gebruikers in kaart te brengen in de context van hun gepersonaliseerde ervaringen dat Spotify aanbiedt zoals aangehaald in de vorige sectie. Deze surveys hebben gedurende anderhalve maand lang data vergaard over de mening van Spotify gebruikers.

In [Figuur 18](#) zien we dat net zoals bij onze vorige twee Streaming services mensen nog steeds niet gewend zijn aan het achterlaten van een rating op muziek die ze hebben beluisterd. Gelukkig voor Spotify blijken hun systemen voorbereid te zijn op een tekortkoming van ratings aangezien 62% van de test-groep in [Figuur 19](#) al vaak nieuwe leuke muziek hebben ontdekt via het Discover Weekly systeem. Een reden waarom Spotify nog steeds in staat is om goede voorspelling te maken is omdat ze meer focussen op hoe lang en hoe vaak gebruikers naar een bepaald nummer luisteren in plaats van de ratings dat een bepaalde artiest krijgt. Bijvoorbeeld als een gebruiker liedjes van een specifieke artiest steeds binnen de paar seconden overslaan zal dit een rode vlag oproepen in het Discover Weekly systeem om deze artiest minder te beginnen aanraden.



*Figuur 18 (links): Het aantal mensen dat ratings geeft aan muziek op Spotify.*

*Figuur 19 (rechts): Hoe vaak mensen nieuwe interessante muziek hebben ontdekt via Spotify's Discover Weekly systeem.*

### 6.3.4 Besluit

Spotify maakt gebruik van verschillende bronnen om een goede gepersonaliseerde ervaring te kunnen bouwen voor hun gebruikers. Enkele voorbeelden van deze bronnen die we net hebben aangehaald zijn de luistergeschiedenis van gebruikers, verschillende genres, en algemene ratings. Deze bronnen zijn algemene zaken die door bijna elke Recommender system gebruikt worden. Waar Spotify zich hier van diversifieert is dat zij ook gaan kijken in zaken als blogs en reviews gemaakt muziekliefhebbers om mogelijk nieuwe gelijkenissen te kunnen vinden tussen gebruikers en aanbevelingen te maken.

Uit onze surveys blijkt ook dat van de 3 besproken streaming services Spotify degene blijkt te zijn met de beste aanbevelingen. Maar liefst 62% van de test-groep heeft al vaak een goede ervaring gehad met het Discover Weekly systeem van Spotify.

## 7 Inzet van de Eindgebruiker

Uiteindelijk is het de verantwoordelijkheid van de bedrijven om accurate aanbevelingen te genereren voor hun klanten maar wat veel eindgebruikers vaak vergeten of niet beseffen is dat zij zelf ook een grote rol spelen in het krijgen van goede gepersonaliseerde ervaring.

Uit voorgaande surveys in Hoofdstuk 6 bleek steeds dat een groot aantal van de test-groep nooit een rating achterlaat wat ervoor zorgt dat bedrijven beschikken over minder data om voorspellingen mee te gaan maken. Des te meer data er is om mee te werken hoe accurater de Recommender systems zullen werken.

Het is ook nodig voor een eindgebruiker om te weten hoe ze zorgvuldig moeten omgaan met Recommender systems om de best mogelijke ervaring eruit te krijgen.

## 8 Besluit

Recommender Systems zijn een geavanceerde technologie dat bedrijven in staat stelt om een gepersonaliseerde ervaring aan te bieden aan gebruikers. Door gebruikers en items te vergelijken met elkaar kunnen deze systemen efficiënt nieuwe trends en gelijkenissen vinden voor bepaalde gebruikersgroepen. Uit ons onderzoek blijkt dat er veel verschillende soorten technieken bestaan om aanbevelingen te maken voor klanten en dat al deze verschillende systemen in samenwerking met elkaar kunnen werken voor nog snellere en betere voorspellingen. Om te weten welke technieken het best werken voor elk bedrijf moeten we eerst gaan kijken naar het business model en het exacte doel dat men wilt bereiken.

Youtube zal bijvoorbeeld gebruikers zo vaak mogelijk willen laten doorklikken op nieuwe video's terwijl Netflix een zo aangenaam en comfortabele gebruikservaring wil aanbieden om gebruikers te overtuigen om steeds hun abonnement te verlengen. Er is dus eerst en vooral een goed begrip nodig op het doelpubliek voor het Recommendation system.

Via onze analyse op 3 grote Streaming services en een survey over de algemene bevindingen van klanten over de gepersonaliseerde ervaringen dat deze Streaming services aanbieden hebben we geleerd dat deze systemen wel degelijk in staat zijn accurate aanbevelingen te maken. Een interessante bevinding is dat veel gebruikers nooit een rating achterlaten desondanks dit probleem zijn de 3 besproken streaming services nog steeds in staat om goede aanbevelingen te maken.

Uit de afgenomen surveys kunnen we steeds afleiden dat voor de 3 besproken use cases gebruikers steeds goede ervaringen hebben opgelopen met de gepersonaliseerde aanbevelingen en kunnen we concluderen dat deze systemen wel degelijk een meerwaarde bieden voor Streaming services.



## 9 Bibliografie

- Adomavicius, G., & Tuzhilin, A. (2005). *Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions*. IEEE.
- Bobadilla, J. O. (2012). *A collaborative filtering approach to mitigate the new user cold start problem*. *Knowledge-Based Systems*, 26, 225-238.
- Burke, R. (2007). *Hybrid web recommender systems*. In *The adaptive web* (pp. 377-408). Springer, Berlin, Heidelberg.
- Covington, P. A. (sd). *Deep neural networks for youtube recommendations*. In *Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 191-198). ACM.
- Dutta, A. (2017). *Youtube Business Model | How does Youtube make money?* Feedough.com.
- Goltsman, K. (2017). *Under the Hood of Netflix Recommender System*. Datascience Foundation.
- Gomez-Uribe, C. A., & Hunt, N. (2015). *The Netflix recommender system: Algorithms, business value, and innovation*. ACM Trans.
- Lee, J. S. (2012). Lee, J., Sun, M., & Lebanon, G. arXiv preprint arXiv:1205.3193.
- Netflix. (2018). *Netflix's View: Internet entertainment is replacing linear TV*. Opgehaald van Netflix.com: <https://ir.netflix.com/netflixs-view-internet-tv-replacing-linear-tv>
- Pasick, A. (2015). *The magic that makes Spotify's Discover Weekly playlists so damn good*. Quartz.
- Popescul, A. P. (2001). *Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments*. In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence* (pp. 437-444). Morgan Kaufmann Publishers Inc.
- Puglisi, S., Parra-Arnau, J., Forné, J., & Rebollo-Monendero, D. (2015). *On content-based recommendation and user privacy in social-tagging systems*. Computer Standards & Interfaces.
- Rodriguez, A. (2018). *Youtube's recommendation drive 70% of what we watch*. Quartz.
- Rokach, L., Ricci, F., & Shapira, B. (2010). *Recommender System Handbook*. Springer US.
- Sarwar, B. K. (2001). *tem-based collaborative filtering recommendation algorithms*.
- Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). *Item-based collaborative filtering recommendation algorithms*. In *Proceedings of the 10th international conference on World Wide Web*: ACM.

Sydeek, R. M. (2018). *Spotify Business Model | How Does Spotify Make Money?*  
Feedough.com.

Wu, Y., Liu, X., Xie, M., Ester, M., & Yang, Q. (2016). *CCCF: Improving collaborative filtering via scalable user-item co-clustering*. ACM.