

# Business rules voor Recommendation Engine

Door Lucas van der Horst

## Rule 1:

Is content filtering, het geeft tien product id's van producten met dezelfde:

- Category
- Subcategory
- Subsubcategory
- Targetaudience

Want mensen die naar producten kijken zijn sneller ook geïnteresseerd in producten in dezelfde categorie of voor dezelfde doelgroep.

## Rule 2:

Is collaborative filtering, geeft tien id's van profielen van hetzelfde segment.

Ik wilde nog dat die dan id's zou geven van producten die bekeken zijn door het meest matchende profiel uit dat segment. En dat had ik ook voor 60% af maar kon het toen tijd gedwongen niet afmaken.

## Distance rules:

Ik heb ook twee regels geschreven om te kijken hoe dichtbij producten/profielen bij elkaar zitten. Des der grotere kleinere distance, des der meer de producten/profielen op elkaar lijken.

Ik wilde hierop twee aparte tables maken die dan bij het preprocessen gevuld werden met informatie over welke producten/profiles dichter bij elkaar liggen. Later kwam ik erachter dat de processing capaciteit daarvoor gigantisch hoog ligt.

Maar het kan wel gebruikt worden om bijvoorbeeld een sample aan producten te nemen en die te sorteren. Of zoals bij Rule 2 vergelijken met het meest matchende profiel.

## Products distance:

Basis afstand = 0

Als het merk niet hetzelfde is, wordt de afstand groter met 1.

Als het type niet hetzelfde is, wordt de afstand groter met 1.

Als de categorie niet hetzelfde is, wordt de afstand groter met 1.

Als de sub categorie niet hetzelfde is, wordt de afstand groter met 1.

Als de sub sub categorie niet hetzelfde is, wordt de afstand groter met 1.

Als de doelgroepen niet hetzelfde is, wordt de afstand groter met 5.

Als de één een deal heeft maar de ander niet, wordt de afstand groter met 1.

Het prijsverschil in centen gedeeld door 500 wordt ook toegevoegd aan de afstand.

De meeste van deze zijn vanzelfsprekend behalve misschien het prijsverschil. Maar ik denk wel dat er een overeenkomst zit dat mensen die meer geld over hebben / hogere kwaliteit willen / grote dingen kopen ten opzichte van mensen die meer willen besparen / kleinere dingen kopen.

## Profiles distance:

Basis afstand = 0

Als het segment niet hetzelfde is, wordt de afstand groter met 5.

Het datums verschil in jaren wordt er ook aan toegevoegd.

Ik heb ook het datums verschil meegenomen zodat als bijvoorbeeld de webshop z'n focus veranderd meer op een andere markt, dat recentelijke klanten minder effect hebben van klanten die zijn weggegaan voor de shift.