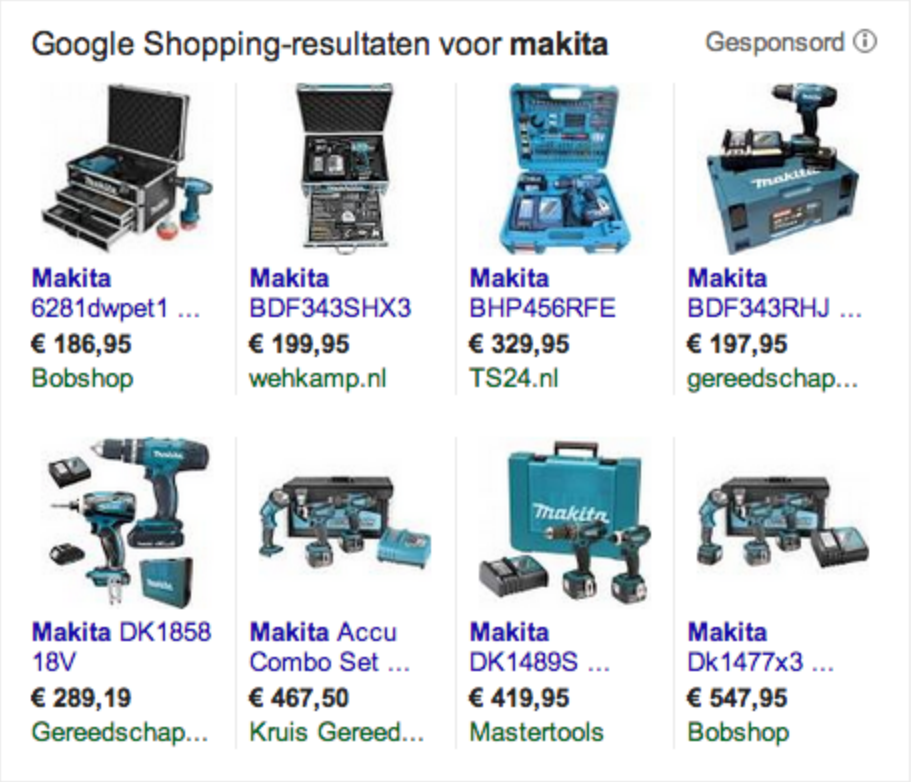
Afstudeeropdracht: vinden van meest rendabele productattributen die de Google AdWords veiling beïnvloeden

Opdracht voor adwordsrobot.com

Situatieschets

Binnen Google is het mogelijk om te zoeken op producten die verkocht worden in webwinkels. Vaak worden deze producten vergezeld door productadvertenties, de zogenaamde Google Shopping advertenties. Bij een zoekopdracht naar `makita` (merk van gereedschap) zijn de volgende Shopping advertenties te zien:



Deze advertenties worden opgezet d.m.v. productgegevens die binnen de webshop(database) beschikbaar zijn. Denk aan merk, productnaam, prijs, producttype, categorie, etc.

Op het moment dat de zoeker een zoekwoord intypt, bepaalt Google welke productadvertenties naar boven moeten komen voor deze zoekopdracht. Als je breed zoekt, bijvoorbeeld op een merk zoals Makita, krijg je een aantal verschillende Makita producten te zien die redelijk uiteen kunnen lopen. Op het moment dat je

iets specifieker zoekt bijv. `makita dk1858` (type boormachine), zul je producten krijgen die meer specifiek aan de zoekopdracht voldoen, zoals hieronder is te zien.

Google verdient geld voor elke klik op hun advertenties. Hoe meer de adverteerder per klik bereid is te betalen, des te groter de kans dat Google zijn producten gaat laten zien, omdat je meer aan zijn klik kan verdienen. Google probeert de zoekopdracht zo goed mogelijk te koppelen aan productaanbod.

Probleemstelling

Binnen Google Shopping advertentieplatform kun je bieden op verschillende attributen van producten. Je kunt besluiten meer te bieden op alle producten van een bepaald merk. Of je kunt besluiten minder te bieden op producten binnen bepaalde prijsklasse. Je kunt ook besluiten om het bod op een specifiek product te verhogen, omdat het goed loopt. Als effect heeft het dat als iemand dan een product zoekt met dat attribuut, dat de kans groter (of kleiner) wordt dat het product getoond wordt, omdat Google ook zoveel mogelijk winst wil maken met een klik.

De moeilijkheid is om de juiste attributen (en hun waarden) te kiezen en voor die groep producten het juiste bod te bepalen. Het kan zijn dat `Nike` (merk) `Voetbalschoenen` (productcategorie) goed verkopen, daarvoor zou je bod hoger kunnen maken om vaker je advertenties naar voren te krijgen.

Het doel is om op productattributen te bieden, zodat de kans op winst het grootst is.

Advertentiestatistieken

Er zijn verschillende statistieken van elk product beschikbaar om mee aan de slag te gaan. Hieronder een aantal:

- `Bid` = Wat de adverteerder bereid is per klik te betalen op dit moment.
- `CostPerClick` = Wat de adverteerder uiteindelijk per klik heeft betaald in een bepaalde periode.
- `Impressions` = Aantal keer dat een advertentie weergegeven is binnen Google binnen een bepaalde periode.

- `Clicks` = Aantal keer dat er geklikt is op een advertentie binnen een bepaalde periode.
- `Conversions` = Aantal keer dat een product gekocht is binnen een bepaalde periode.
- `Cost` = Totaal betaald bedrag aan alle kliks voor desbetreffende product binnen een bepaalde periode.
- Verschillende product attributen per product, zoals voorraad, merk, productnaam, prijs, e.d.

Deze statistieken zijn bijna realtime en kunnen gebruikt worden om biedingen en groepen te sturen.

Te beantwoorden vragen / uitdagingen

Gegeven een voorraad (lijst van producten met product attributen) en advertentiestatistieken (zie hierboven) van deze producten over een periode, dient het volgende te worden onderzocht:

- Welke productattributen zijn betrouwbaar om meer te werken? Het kan bijvoorbeeld zijn dat alle producten van hetzelfde merk zijn of dat bepaalde attributen weggelaten zijn. Voor elke voorraad kan dat dus verschillen.
- Op welke productattributen kun je het beste bieden en hoeveel om de winst zo hoog mogelijk te houden.
- Welke productattributen zijn niet rendabel om op te bieden en kun je beter €0.00 bieden om geen geld aan uit te geven.
- Welke productattributen kies je zodat alle producten uit de voorraad daaronder vallen en je geen dubbele producten binnen een sectie krijgt.
- Welke productattributen hebben nog niet genoeg statistieken om een goed besluit te kunnen maken.
- Prototype van een algoritme dat aan de hand van een voorraad en van advertentiestatistieken een mapping van productattributen/waarden naar bod geeft. Iets in de geest van:

```
merk=nike,   prijs=50-100   => 0.15   # verkoopt veel beter dan gemiddeld
merk=nike,   prijs=0-50     => 0.00   # niet op bieden
merk=nike,   prijs=100-999  => 0.05   # verkoopt gemiddeld
merk=adidas, prijs=0-50     => 0.00   # niet op bieden
merk=adidas, prijs=75-999   => 0.07   # verkoopt beter dan gemiddeld
rest                        => 0.05   # verkoopt gemiddeld
```