

Content filtering

Voor de content filtering heb ik gebruik gemaakt van de prijsklasse van het product. Als de klant naar een product kijkt, zit dat product in een prijsklasse. De klant krijgt producten aanbevolen uit dezelfde prijsklasse. De prijsklassen zijn als volgt:

€0.01 t/m €1.00; 'very cheap'

€1.01 t/m €2.50; 'cheap'

€2.51 t/m €500; 'middle'

€5.01 t/m €7.50; 'expensive'

€7.50 en hoger; 'more expensive'

Dit algoritme is een van de algoritmen die mijn groep en ik gebruikt hebben in het project, in het project hadden we meerdere algoritmen gemaakt die over elkaar heen liepen. Zo hadden we ook een categorie algoritme, die twee algoritmen samen maakte dus samen een nauwkeurigere aanbeveling. Dit algoritme werkt dus beter in samenwerking met een ander algoritme. Alleen werkt hij, alleen haal je er dan niet het meest uit.

Ik heb voor dit algoritme ieder product gepakt een daarvoor 5 random geselecteerde producten gepakt uit dezelfde prijsklasse. Het product en de aanbevelingen heb ik in een database tabel gezet.

	id [PK] character varying	product1 character varying	product2 character varying	product3 character varying	product4 character varying	product5 character varying
1	01-Diversen	25303	46896	23589-blauw	90324	44107
2	01001	9010-rhubarb	45704-zwarteschep	43421	28054	33015
3	01001-chalkwhite	43134-chrysler	41839	16683	34686	5328
4	01001-jetblack	7103	34884	24033	3056	28436
5	01001-lilacchiffon	27163	29901-roze	32116	21795	40127-16
6	01001-motherofpearls	27392	41922	7820	23270	42983
7	01001-pashmina	41333	7813	30628	43038	22982
8	01001-prettyinpurple	39573	48459	22184	36553	33322
9	01001-purplebrust	33701	33300	3763	29150	5142
10	01001-purpleedge	41830	8378	35619	21593	22098
11	01001-rosemauve	4364	35354-zilver	20801	35556-princess	48430
12	01001-tease	47142	32956	9642	36287-oranje	25206
13	01009	34462	33356	36078	31135	27948
14	01019	29233	29217-minniemouzero...	24361	39895-groen	41158
15	01022	39743	7390	24645	20315	36372
16	01035	36350	27990	21370	24062	39925
17	01041	16377	39604	43415	28243	31466
18	01044	26316	28541	26499	20482	45777
19	01080	35163	29364	4523	32209	42913

Collaborative filtering

Voor de collaborative filtering heb ik gebruikt gemaakt van het segment van de klant, een klant wordt in een segment gedaan op basis van hun gedrag in de webwinkel. De klant krijgt een product aanbevolen die een klant uit datzelfde segment heeft bekeken.

Ik heb een database tabel gemaakt die meerdere waardes uit de tabellen 'inner joined'. Zo haal ik de profid en prodid uit profiles_previously_viewed, segment uit profiles en name, category, subcategory, subsubcategory en targetaudience uit products.

Zo ziet het er uiteindelijk uit in de tabel:

	profid character varying	category character varying	targetaudience character varying
1	5ad6681ee208920001...	Kleding & sieraden	Mannen
2	5ad6659b8951860001...	Gezond & verzorging	Jongen
3	5ad652368951860001...	Gezond & verzorging	Kinderen
4	5ad64a3a8c28b10001...	Huishouden	Volwassenen
5	5ad647426cefc100017...	Gezond & verzorging	Vrouwen
6	5ad637cf93916e0001...	Huishouden	Volwassenen
7	5ad6283bbe8d580001...	Baby & kind	Baby's
8	5ad624208c28b10001...	Gezond & verzorging	Volwassenen
9	5ad61db24047770001...	Wonen & vrije tijd	Mannen
10	5ad613024047770001...	Baby & kind	Kinderen
11	5ad5f41a5bb6eb0001...	Gezond & verzorging	Volwassenen
12	5ad5ef0f244a9300017...	Gezond & verzorging	Volwassenen
13	5ad5d8b2e9cdfd0001...	Kleding & sieraden	Vrouwen
14	5ad5bea8e9cdfd00015...	Huishouden	Volwassenen
15	5ad5b82c6c58f20001...	Make-up & geuren	Vrouwen
16	5ad5b1fee9cdfd00015...	Gezond & verzorging	Volwassenen
17	5ad5a680dbc8b80001...	Gezond & verzorging	Mannen
18	5ad59ef6244a930001...	Gezond & verzorging	Mannen
19	5ad575ad6c58f20001f...	Gezond & verzorging	Mannen
20	5ad5225fe9cdfd0001c...	Gezond & verzorging	Mannen