

Mining Behavioral Patterns from Web Activity

Meno: Veronika Gondová

Vedúci práce: Mária Bieliková

Dátum: 23.03.2017



PeWe@FIIT
personalized web group

Čo môže zahŕňať segmentácia zákazníkov?



Kto to je?

Doménové charakteristiky

(lojálny, cestovateľ)

Demografia

(vek, bydlisko, plat, pohlavie)

Osobnosť

(osobnosť Big5, impulzívnosť, zvedavosť)

Zručnosti

(navigácia, webová gramotnosť)

Ako sa správa?

Aktivita

(čo nakupuje spolu, sekvencie)

Rozhodovanie - dôležitosť

(cena, lokalita, kategória, vyhľadávanie)

Záujmové oblasti

(texty, fotografie, prezentácia ponuky)

Kontext nákupu

(darček, cielený nákup, nuda)

Emócie

krátkodobé

Ako viem využiť vzory v aktivite na klasifikáciu?

Klasický postup



+ Vzory v aktivite na webe

Hypotéza



Aké typy vzorov môžeme využiť?

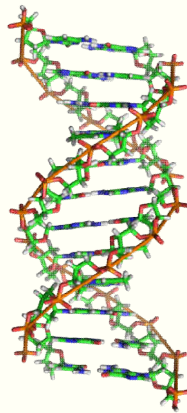


Asociačné pravidlá

Napr. aké eventy sa spolu nachádzajú v session

Sekvenčné vzory

Napr. aké sekvencie eventov predchádzajú kúpe



Event pri hľadaní vzorov
= dvojica (typ eventu + typ dealu)

Asociačné pravidlo
{view služby , view jedlo, search služby} -> {basket služby}
Podpora = s, Dôvera = c

Sekvenčný vzor
(view jedlo, search služby, view služby)

Aké dáta máme k dispozícii a aké sú problémy?

22 miliónov eventov

- View
 - List
 - Search
 - Basket
 - Purchase_processed
- + referrer, filtre

2 660 000 používateľov

ID vs. Cookies

Krátkodobé ponuky

Abstrakcia cez main category
alebo LDA



Pri 33% mám iba ID

Neviem namapovať ak už
pred tým nenamapované

Počet eventov na usera

32000 – 1 event
Počet ľudí s menej ako 10
eventami: 1 600 000

**Prihlásený vs. neprihlásený
používateľ**

Párovanie: iba 150 000

Aké sú možné scenáre?

Sekvenčné vzory v session končiacimi kúpov

Eventy vyskytujúce sa spolu v session končiacich kúpov

Sekvenčné vzory v správaní sa používateľa do prvej kúpy (bez session)

Pri 33% mám iba ID

Neviem namapovať ak už pred tým nenamapované

Počet eventov na usera

32000 – 1 event

Počet ľudí s menej ako 10 eventami: 1 600 000

Prihlásený vs. neprihlásený používateľ

Párovanie: iba 150 000