



Haittapahtumien tunnistaminen HAIPRO aineistosta

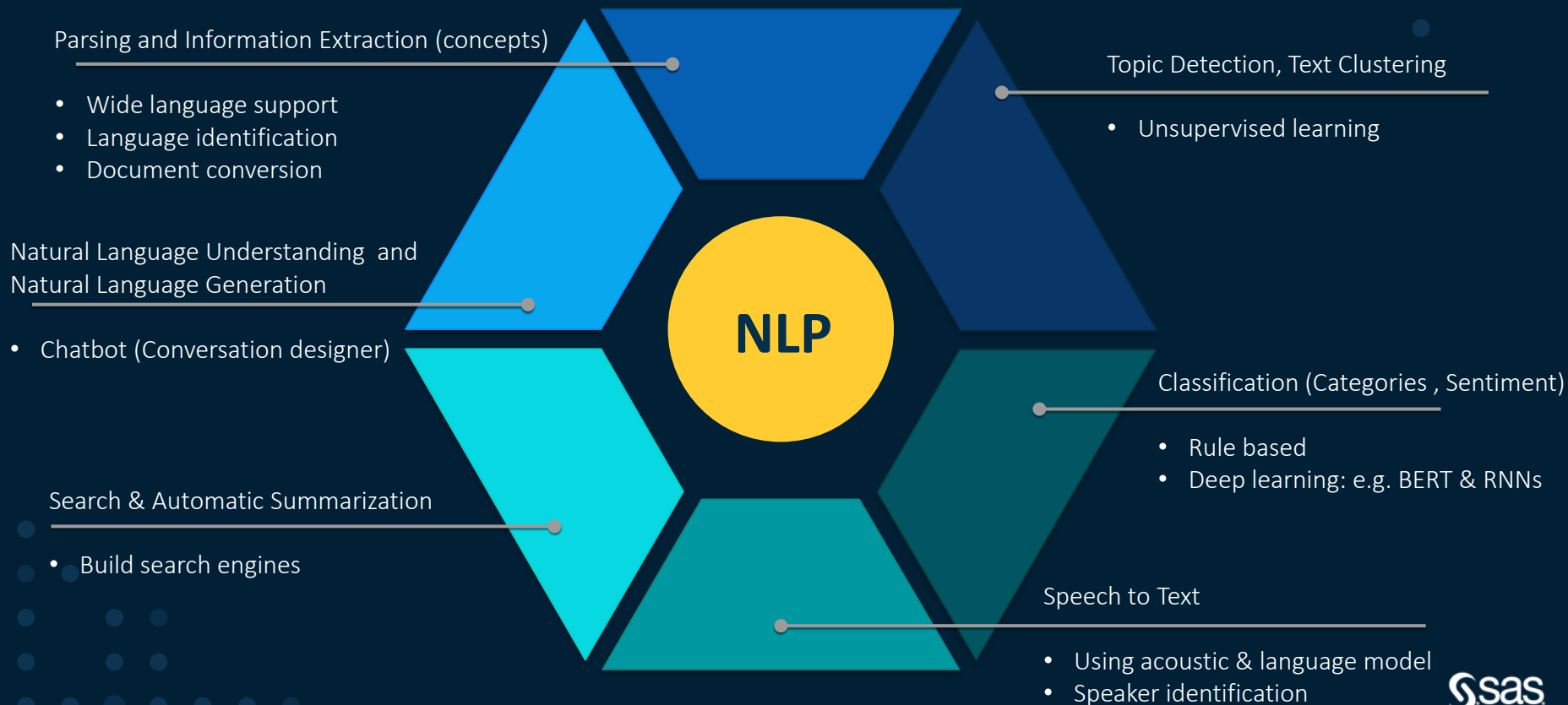
SAS Visual Text Analytics

Antti Heino, AI & Cloud Advisor, SAS Institute

8.6.2021



Text Analytics / NLP Capabilities



Global language support

Arabic

Chinese

Croatian

Czech

Danish

Dutch

English

Farsi

Finnish

French

German

Greek

Hebrew

Hungarian

Hindi

Indonesian

Italian

Japanese

Kazakh

Korean

Norwegian

Polish

Portuguese

Romanian

Russian

Slovak

Slovene

Spanish

Swedish

Tagalog

Thai

Turkish

Vietnamese

Statistics & Machine Learning



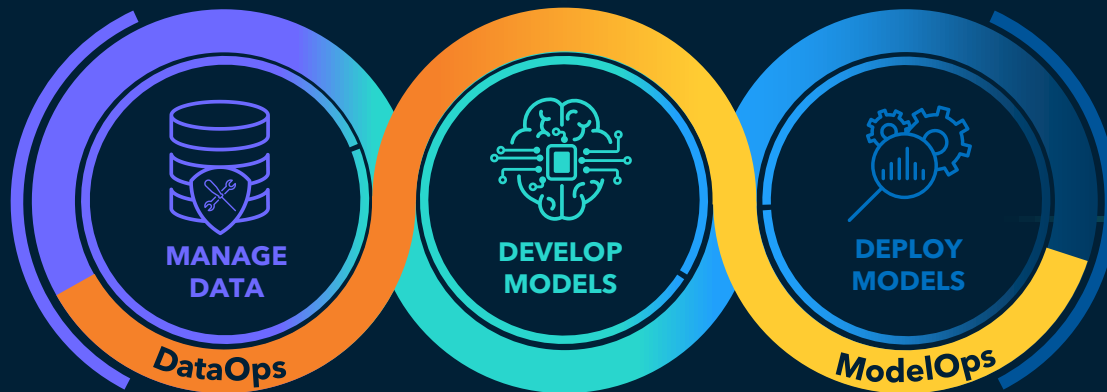
Forecasting, Optimization



Visualization



Deployment



Data Management



Decision Management



Natural Language Processing



Computer & Machine Vision

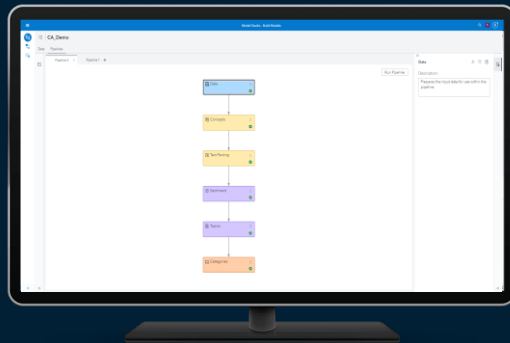
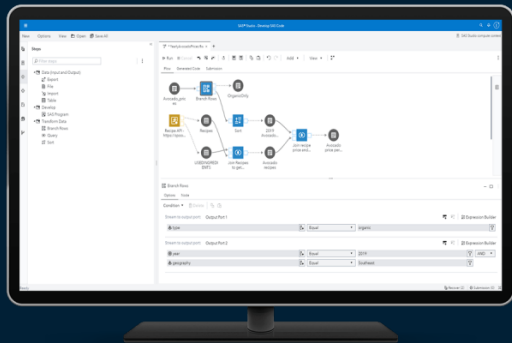


Koko analyysiketju samalla alustalla

DATA

TEKSTIANALYTIikka

TULOSTEN TARKASTELU



Projektin vaiheet

DATA



Projektissa käytettiin 15 tuhatta OYS:n haittatapahtuma-kuvausta

TEKSTIANALYTIikka



- Tekstiä analysoitiin SAS:in Viya analytiikka-alustalla
- Toteutettiin tekstin anonymisointi, ATC luokittelu ja haittatapahtumaan johtaneen syyn luokittelu

TULOSTEN TARKASTELU



- Tekstianalytiikan tuloksista muodostettiin visualisointeja, joilla pystytään analysoimaan uusia yhteyksiä haittatapahtumaan johtaneiden syiden ja ATC luokkien välillä
- Tuloksia tutkimalla voidaan löytää tekijöitä, joita muokkaamalla voidaan parantaa potilasturvallisuutta

Anonymisointi

HAIPRO ilmoitusten tapahtumakuvaukset saattavat sisältää henkilötietoa. Datan hyödyntämisen kannalta se olisi pyrittävä anonymisoimaan

TOTEUTETUT TOIMENPITEET

Datasta etsitään henkilötietoa vastaavia kohtia ja ne korvataan merkkijonolla, joka kertoo millainen tieto on poistettu

- Nimet, osoitteet, puhelinnumerot, henkilötunnukset, sähköpostiosoitteet



Anonymisointi

- Tapahtuman kuvauksesta etsitään merkkijonoja, jotka voivat olla henkilötietoa ja ne korvataan prxchange funktiolla
 - Korvaava tekstipätkä (esim. XXXXA, XXXXB) auttaa tunnistamaan mitä on pyritty anonymisoimaan
- Nimien (varsinkin taivutettujen) anonymisointi johtaa joidenkin sanojen poistamiseen, jotka muistuttavat nimiä
- Nimien osalta piti pystyä jättämään tekstiin järjestelmien nimet
- Anonymisointia on testattu niin ettei se vaikuta tunnistettujen ATC luokkien lukumäärään

```
>> /
36
37 ⊖ data work.anon_step1;
38     set public.WORK_LÄÄKEHAIPROTSASPOC2020_2; /*REPLACE inputdata and caslib*/
39
40 /* Social security numbers and phone numbers */
41 Tapahtuman_kuvaus_23 = prxchange('s/\d\d\d\d\d\d\d?\S\d\d\d\d\d/w/XXXXA/', -1, Tapahtuman_kuvaus_23);
42 Tapahtuman_kuvaus_23 = prxchange('s/(358)\ ?\d+\ ?\d+\ ?\d+/XXXXB/', -1, Tapahtuman_kuvaus_23);
43
```


ATC luokittelu

ATC luokittelu tapahtuu lääkenimien ja vaikuttavien aineiden perusteella. Lisäksi huomioidaan apteekkinimikkeet

Noin 11 tuhatta luokittelutermiä ajetaan HAIPROissa olevaa tapahtumakuvausta vasten.

Jokainen tapahtuma luokituu niin moneen eri ATC-luokkaan kuinka monta eri termiä sieltä tunnistetaan.



ATC luokittelu

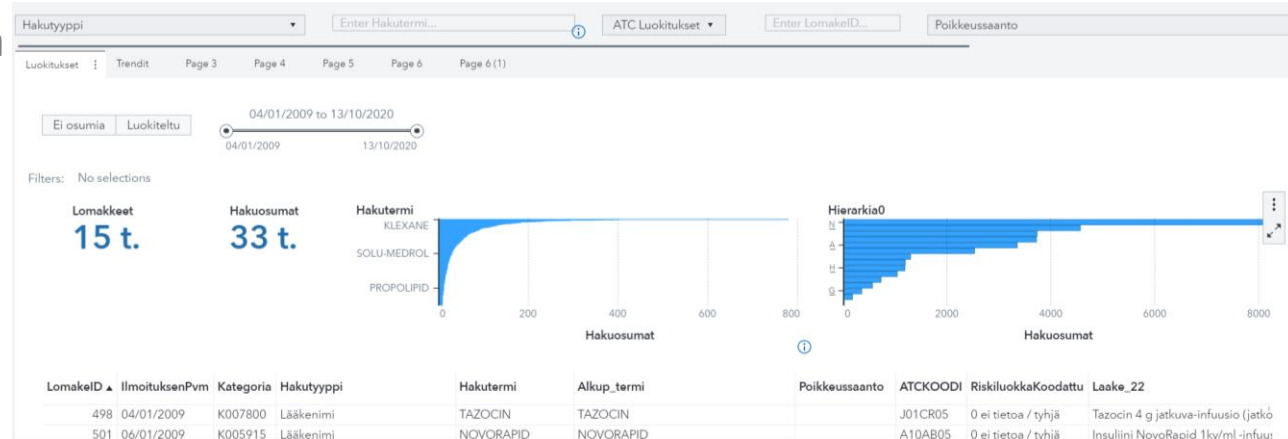
- Luokittelutermejä oli niin paljon (11 tuhatta), että luokittelun ajaminen dataa vasten päätettiin tehdä ohjelmallisesti
- Vaikuttavat aineet ja Lääkenimet tauluista kerätään luokittelutermit ja niistä muodostetaan luokittelusäännöt
- Jokainen haipro-ilmoitus saa 1/0 tuloksen jokaisesta luokasta sen mukaan löytyvätkö hakutermit siitä

	Alkup_termi	Kategoria	K_saanto
1	kolekalsiferoli, yhdistelmävalmisteet	K000001	1:Top/K000001;(OR, "kolekalsiferoli**")
2	Tiamiini (B1-vitamiini)	K000002	2:Top/K000002;(OR, "tiamiini**")
3	Sulbutiamiini	K000003	3:Top/K000003;(OR, "sulbutiamiini**")
4	Benfotiamiini	K000004	4:Top/K000004;(OR, "benfotiamiini**")
5	b1-vitamiinin ja b6- ja/tai b12-vitamiinin yhdistelmävalmisteet	K000005	5:Top/K000005;(SENT, "b1**", "b6**", "b12**", "vitamiini**")

```
11 /******  
12 /* VAIHE 1: */  
13 /* Data jonka pohjalta saadaan luokittelusäännöt */  
14 /******  
15  
16  
17 /* Vaikuttaviin aineisiin perustuvat hakutermit */  
18 data va; set &lahdedata_vaikuttavat_aineet;  
19 format Hakutyyppe $50;  
20 format Poikkeussaanto $500. ;  
21 Hakutyyppe = "Vaikuttava aine" ;  
22 Hakutermit = VaikuttavaAine 1;  
23 Alkup_termi = VaikuttavaAine;  
24 Poikkeussaanto = PoikkeavaHakuehtoAtc;  
25 KEEP Hakutyyppe Hakutermit Alkup_termi ATCKOODI Poikkeussaanto;  
26 WHERE HuonoHakusana ne -1;  
27 run;  
28  
29 /* Lääkenimiin perustuvat hakutermit */  
30 data la; set &lahdedata_laakenimet;  
31 format Hakutyyppe $50;  
32 format Poikkeussaanto $500. ;  
33 Hakutyyppe = "Lääkenimi" ;  
34 Hakutermit = LAAKENIMI 1;  
35 Alkup_termi = LAAKENIMI;  
36 Poikkeussaanto = PoikkeavaHakuehtoLääkenimi;  
37 KEEP Hakutyyppe Hakutermit Alkup_termi ATCKOODI Poikkeussaanto;  
38 WHERE HuonoLääkenimiHakusana ne 1;  
39 run;  
40  
41 /* Manuaaliset hakutermit - Vaikuttava aine */  
42 data va_man; set &lahdedata_vaikuttavat_aineet;  
43 format Hakutyyppe $50. ;  
44 format Poikkeussaanto $500. ;  
45 Hakutyyppe = "Vaikuttava aine (manuaalinen syöttö)";  
46 Alkup_termi = VaikuttavaAine;  
47 Poikkeussaanto = PoikkeavaHakuehtoAtc;  
48  
49 do i=1 by 1 while(scan(HakusanaManuaalisestiAtc,i,') ^=' ');
```

ATC luokittelu

- Tuloksena saadaan 15 tuhannelle haipro- ilmoitukselle n. 30 tuhatta ATC osumaa
- Saadaan arvokasta tietoa mitkä ATC luokat ja hakutermit esiintyvät eniten ilmoituksissa
- ATC hierarkiassa päästään näkymässä porautumaan alemmille tasoille ja tuloksia voi rajata esim. päivämäärän mukaan



Ristiinanalyysi haittatapahtumien aiheuttajien kanssa

ATC-luokkien lisäksi HAIPRO-ilmoituksia luokiteltiin seuraaviin kiinnostaviin luokituksiin:

LASA ("Look alike – sound alike")

Vireystila (Viittaus henkilökunnan vireyteen)

Potilastietojärjestelmä

Antoreitti

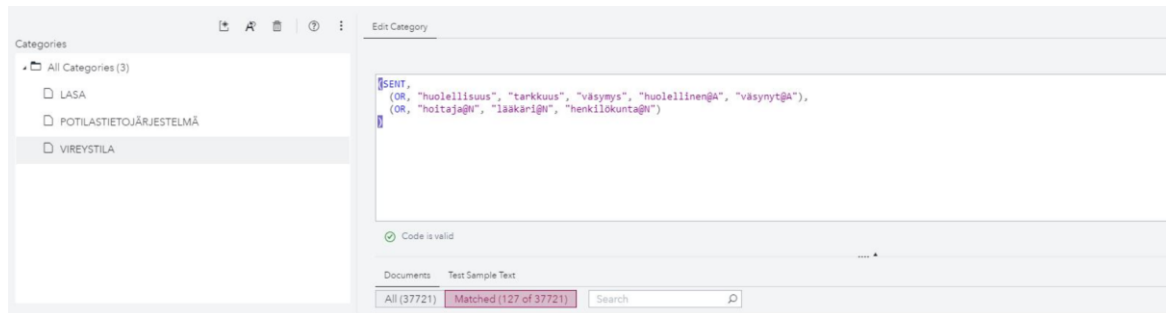
Potilas Outcome

Saatuja tuloksia vertailtiin raportilla ilmoituksista löydettyihin ATC-luokkiin



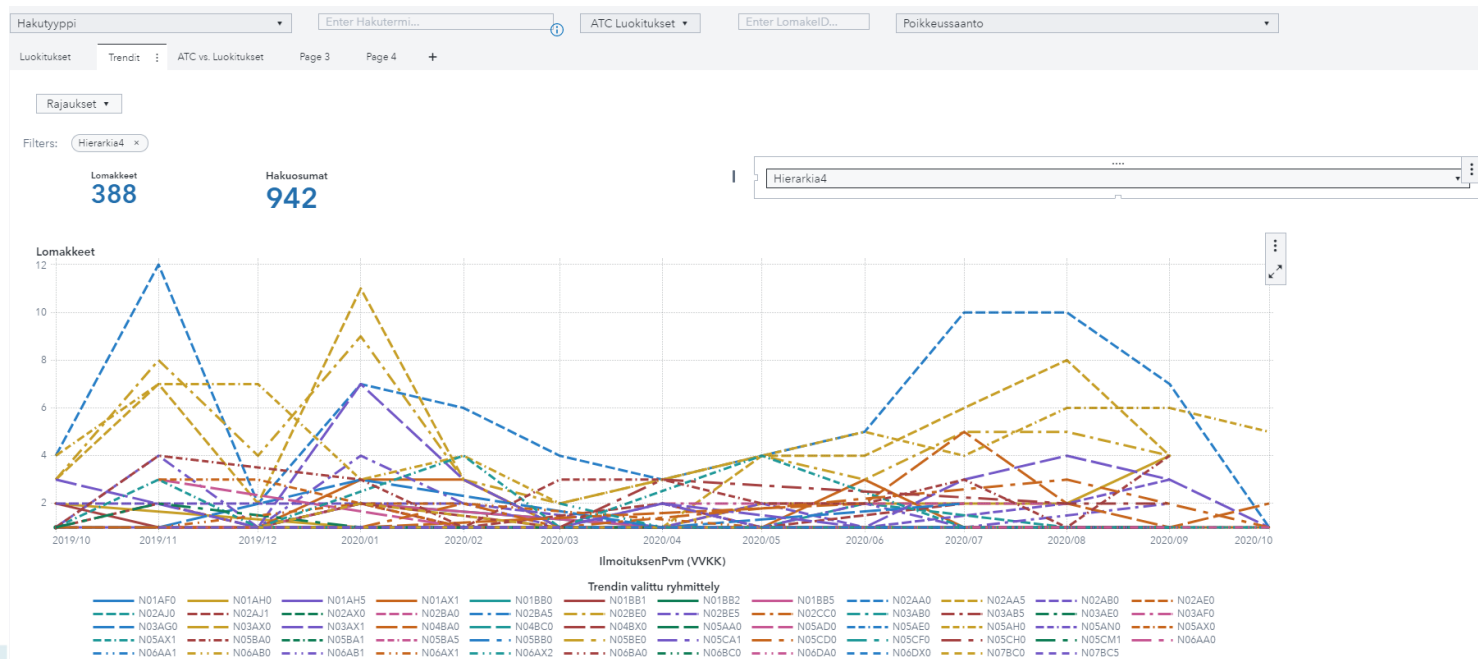
Ristiin analyysi muiden luokkien kanssa

- Luokittelusäännöt ovat samankaltaisia kuin ATC-luokissakin
- Näitä luokkia on kuitenkin vähemmän, joten ne voidaan tehdä visuaalisessa käyttöliittymässä
- Suomen kielen tuki helpottaa sääntöjen tehokasta muodostamista



Esimerkkejä tuloksista

N-lääkkeiden ilmoitukset vuoden ajalta



OYS, Sami Sneck, sami.sneck@ppshp.fi
OYS & OY Outi Laatikainen, outi.laatikainen@oulu.fi
SAS Institute, Antti Heino, antti.heino@sas.com
Aureolis, Ville Huikari, ville.huikari@aureolis.com



ppshp



More information

- Visual Text Analytics demo (identifying need for car recalls)
 - <https://www.youtube.com/watch?v=NrkwlJcDI6M>
- Deep learning for text data demo (BERT)
 - <https://www.youtube.com/watch?v=Wrcx3ZwqaD4>
- Product page
 - https://www.sas.com/en_us/software/visual-text-analytics.html
- Documentation & support
 - <https://support.sas.com/en/software/visual-text-analytics-support.html>