

THOMAS
MORE

JESSA
ZIEKENHUIS



OK logistiek Jessa

Proof of concept

Met steun van

**AGENTSCHAP
INNOVEREN &
ONDERNEMEN**



Vlaanderen
is ondernemen

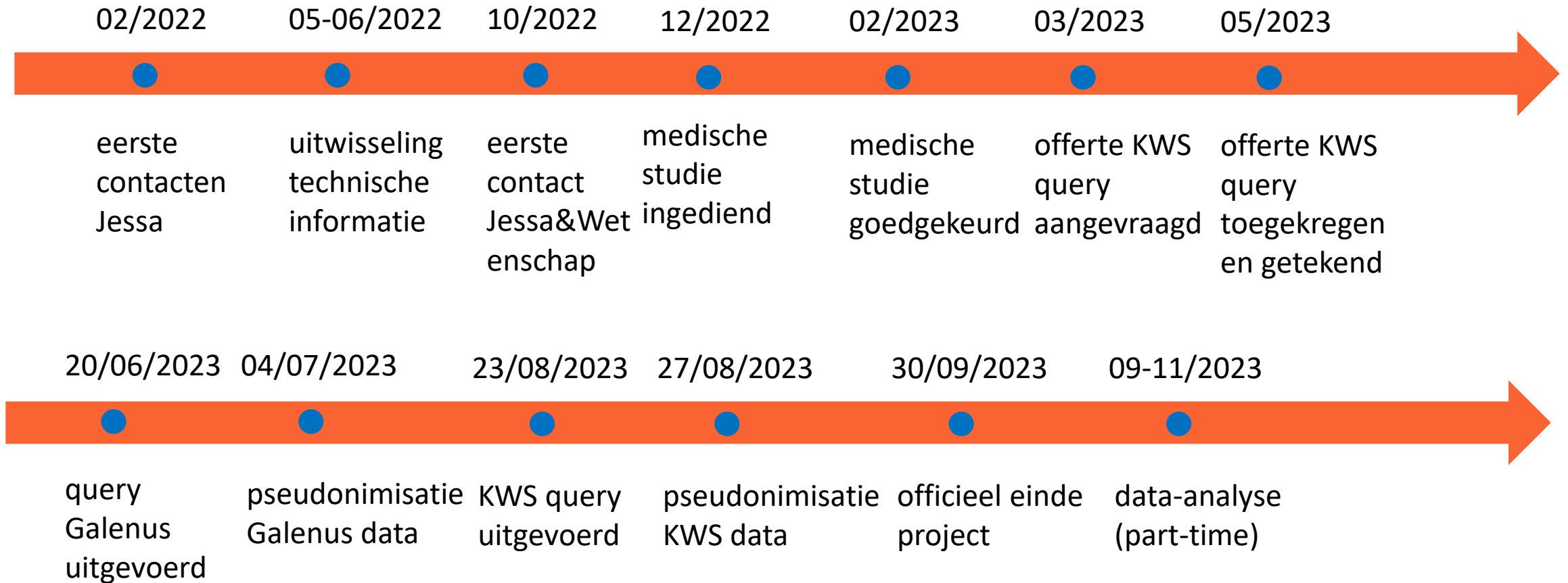
Opmerking vooraf

- Doelstelling TETRA projecten
 - technologische opportuniteiten (use cases) aandragen
 - kennisoverdracht
 - inspiratie
- Gericht op
 - begeleidingsgroep
 - ≠ 1 specifieke “klant”
- niet marktverstrend
 - ≠ onderzoek & ontwikkeling (start to finish)
- verdere uitwerking : O&O project

Uitdagingen Logicare (Algemeen)

- Overbevraging zorgmedewerkers
 - nasleep COVID
 - prioriteit bij zorgtaken (uiteraard!)
- Format TETRA
 - impliciete verwachting = minimale werklast begeleidingspartners
 - ↔ toegang interne (medische!) data
- Aanvraag klinische studies
- Data readiness – Data maturiteit – Datasilo's
- Wisselende contactpersonen

Tijdslijn



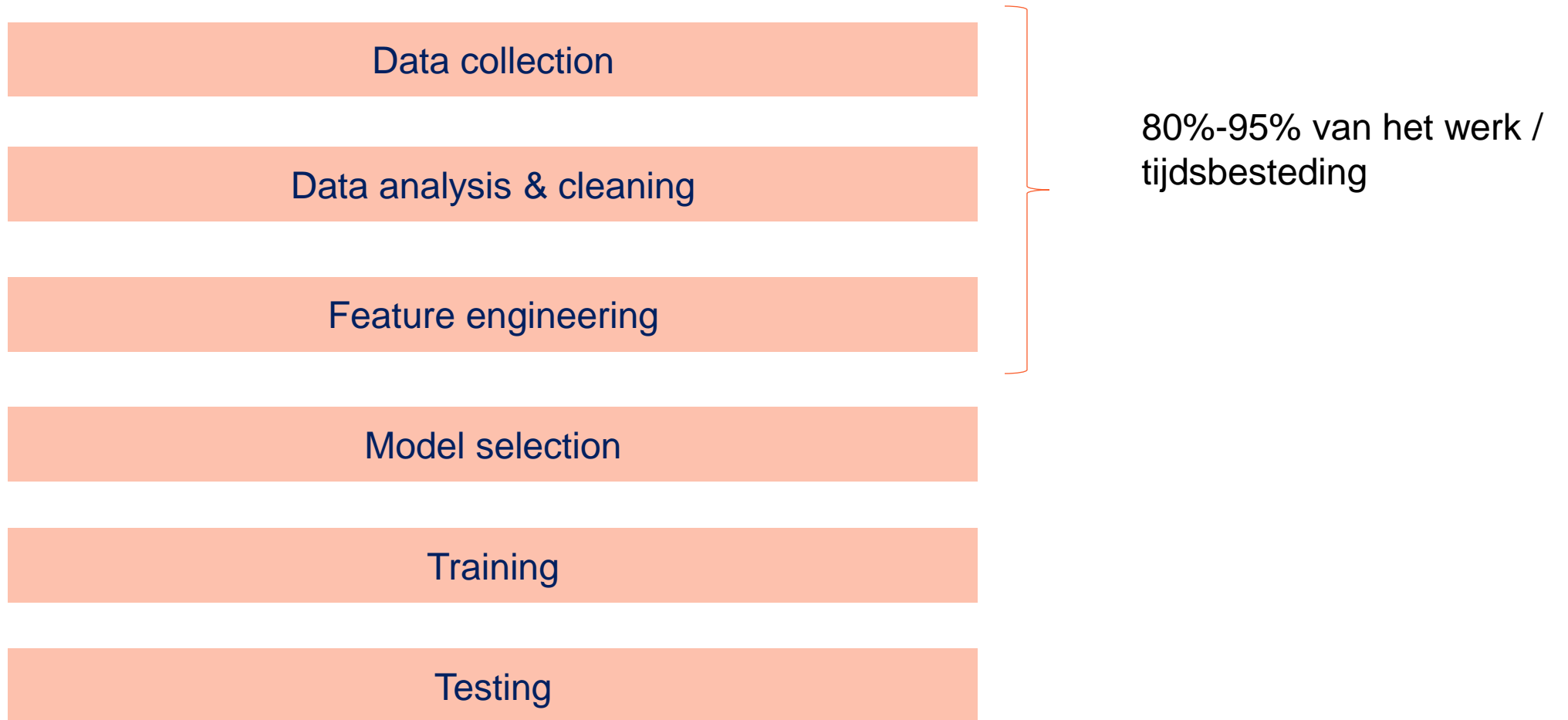
Onderzoeksvraag - initieel

- Startsituatie
 - lijsten (excell) van klaar te zetten items op basis van ingreep-arts-OK ruimte
 - verbruiksdata wijkt substantieel af (niet gebruikte items, niet-klaargelegde items)
 - voorraden niet volledig onder controle
- Verbetering mogelijk met data analytics en AI ?
- Betere match mogelijk mbv patientdata ?
- focus : Invasief Medische Hulpmiddelen (IMH's)

Medische studie

- Retrospectieve medische studie bij Jessa & Wetenschap
 - zeer goede begeleiding
 - duidelijke procedures, stappen en doorlooptijden
- Inhoudelijk
 - Indieningsbrief
 - Studieprotocol
 - Datamanagement plan
 - GCP certificaten
 - Studieovereenkomst
 - Contactinformatie
 - Goedkeuring diensthoofden van betrokken afdelingen

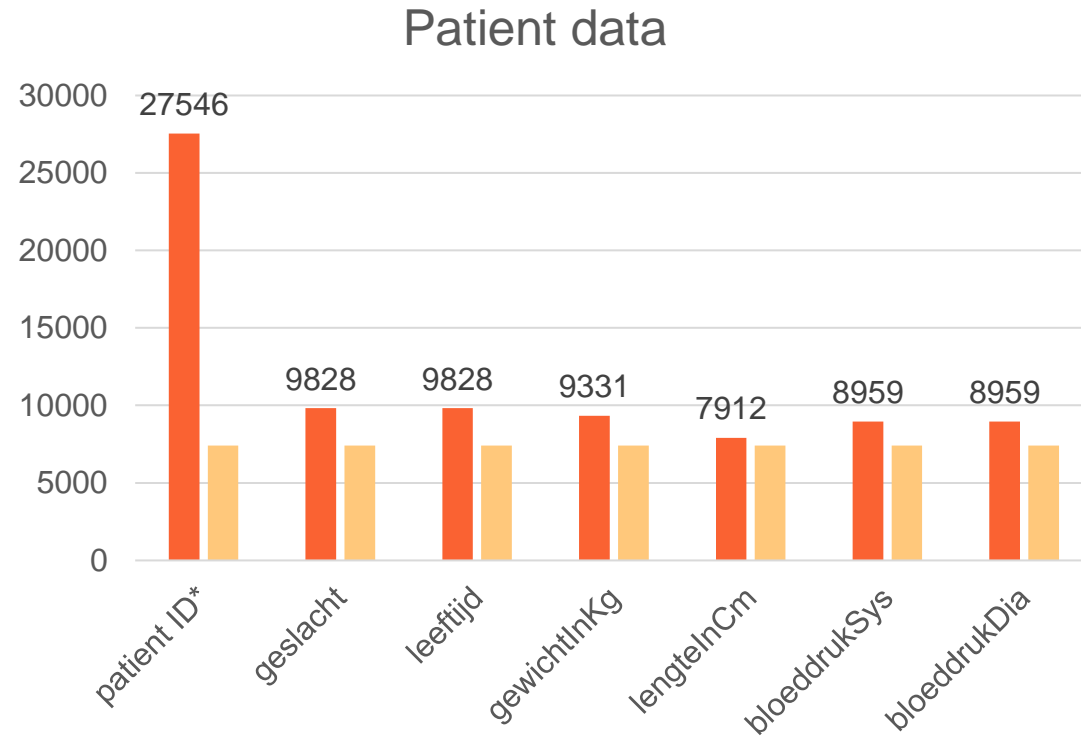
AI werkwijze



Data collectie

- OK data
 - Juliette (Galenus)
 - korte communicatielijn, snelle respons
- Patientdata :
 - KWS (Nexuzhealth)
 - indirecte communicatie, tragere respons
- Gevoeligheid
 - patiëntgegevens : beide
 - gegevens medewerkers (arts): OK data
- Toegang : enkel mogelijk via medische studie
 - Data wordt gepseudonimiseerd
 - Data enkel toegankelijk voor onderzoekers geregistreerd in het datamanageplan
→ afgeschermdde omgeving

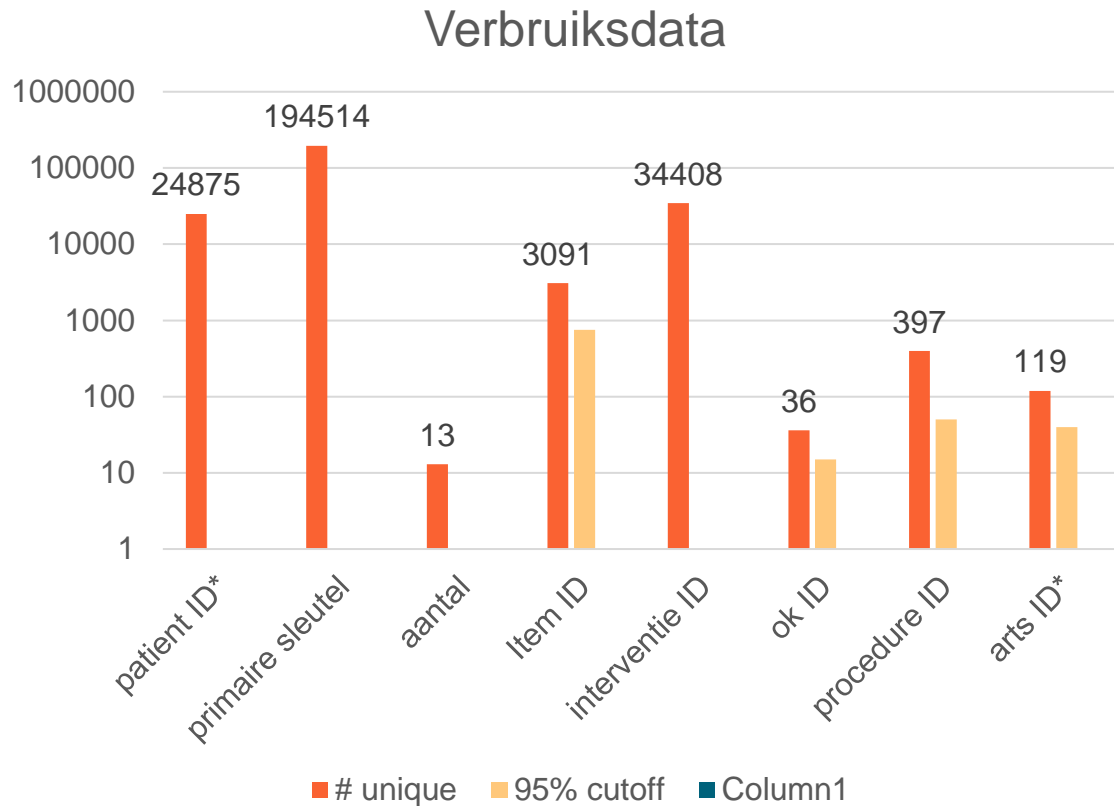
Data Elementen – analyse vooraf



- slecht 30% ingevuld
- slechts 25% ingevuld en gelinked aan verbruiksdata

(* = gespeudonimiseerd) ■ # ingevuld ■ # volledig en gelinked

Data Elementen – analyse vooraf



Groot aantal items
⇔ 10% van aantal ingrepen
Relatief groot aantal procedures

(* = gespeudonimiseerd)

Data analyse - vaststelling

- Veel verschillende lijsten van items voor 1 type ingreep-arts-OK
- → IMH's variëren sterk

Procedure ID	OK	Arts	# ingrepen	# verschillende lijsten	# verschillende items	#ingrepen/ #verschillende lijsten	#ingrepen/ #verschillende items
OFTA251	SA-OK-12	52	2191	475	102	4,6	21,5
OFTA251	SA-OK-12	86	2052	517	135	4,0	15,2
CARD657	VJCATH01	54	1948	1086	320	1,8	6,1
OFTA251	SA-OK-11	117	1344	401	122	3,4	11,0
CARD657	VJCATH02	51	1259	658	338	1,9	3,7
CARD657	VJCATH02	54	980	562	280	1,7	3,5
CARD657	VJCATH02	47	875	452	260	1,9	3,4
CARD657	VJCATH02	91	751	501	416	1,5	1,8
CARD657	VJCATH01	91	690	452	380	1,5	1,8
OFTA251	SA-OK-11	63	648	288	120	2,3	5,4
CARD044	VJCATH05	112	569	434	122	1,3	4,7
CARD104	VJCATH04	102	508	191	77	2,7	6,6
OFTA251	SA-OK-12	110	482	167	54	2,9	8,9
CARD657	VJCATH01	47	451	239	186	1,9	2,4
OFTA251	SA-OK-12	1	443	213	69	2,1	6,4
CARD044	VJCATH05	108	390	281	89	1,4	4,4
CARD044	VJCATH05	3	382	305	91	1,3	

Recommender systeem

Per ingreep bekijken we volgende rijen

#items om uit te kiezen (0 = item niet gekozen, 1 = item gekozen)

werkelijk	0	0	0	1	0	0	1	0	0	0	1	0	0	0	...	0	0	1	0	0	1	0	0	0	0	1	0	0	0		
voorspeld	0	0	0	1	0	0	0	0	0	0	1	0	1	0	...	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0
score	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	...	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1

0 = fout voorspeld

1 = correct voorspeld

Eerste resultaat

Ruwe eerste test voor 1 combinatie (procedure/OK/arts)

- aantal items om uit te kiezen : 151
- aantal voorbeelden om op te trainen : 231

→ gemiddeld 3% van de items fout voorspeld !

- Dit lijkt in eerste instantie goed
- ... maar is misschien nog niet klinisch bruikbaar omwille van het grote aantal items van waaruit gekozen wordt (151) → 3% ~ 5 items
- In meer detail, het aantal juist voorspelde items per ingreep:
 - min : 135
 - gemiddeld : 145
 - max : 148

verwacht	0	0	0	1	0	0	1	0	0	0	1	0	0	0	...	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0
voorspeld	0	0	0	1	0	0	0	0	0	0	1	0	1	0	...	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0
score	1	1	1	1	1	1	0	1	1	1	1	1	0	1	...	1	1	1	1	1	1	1	1	1	1	1	0	1	1	

Eerste resultaat

Vertaling naar meer klinische bruikbare cijfers

Ruwe eerste test voor 1 combinatie (procedure/OK/arts):

- gevoeligheid om nodige items te suggereren
 - min: 0.0, median: 0.43, mean: 0.52, max: 1.0
- performantie in het niet-selecteren van niet-benodigde items
 - min: 0.96, median: 0.96, mean: 0.96, max: 0.96
- Gemiddeld wordt ongeveer de helft van de benodigde items voorspeld
- Er worden maar 4% foute items voorspeld

verwacht	0	0	0	1	0	0	1	0	0	0	1	0	0	0	...	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0
voorspeld	0	0	0	1	0	0	0	0	0	0	1	0	1	0	...	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0
score	1	1	1	1	1	1	0	1	1	1	1	1	0	1	...	1	1	1	1	1	1	1	1	1	1	1	0	1	1	

Eerste resultaat - bis

- De data aangeleverd vanuit het galenus systeem bevatten naast de invasief medische hulpmiddelen ook implantanten.
- Tijdens de nabespreking van het VOKA event werd duidelijk dat deze het minst afhangen van de gegevens van de patient.
- De volgende test (bis) is uitgevoerd voor de dezelfde selectie van (procedure/OK/arts)
- aantal items om uit te kiezen : 52
- aantal voorbeelden om op te trainen : 184

Eerste resultaat - bis

Vertaling naar meer klinische bruikbare cijfers

Ruwe eerste test voor 1 combinatie (procedure/OK/arts) – geen implantaten

- gevoeligheid om nodige items te suggereren
 - min: 0.0, median: 0.5, mean: 0.54, max: 1.0
- performantie in het niet-selecteren van niet-benodigde items
 - min: 0.96, median: 0.96, mean: 0.96, max: 0.96
- Gemiddeld wordt ongeveer de helft van de benodigde items voorspeld
- Er worden maar 4% foute items voorspeld
- ➔ Dit is iets beter maar niet spectaculair.
- ➔ Het aantal items om uit te kiezen is kleiner, maar dit veroorzaakt ook problemen : het kleine aantal samples (231) laat geen volledige stratificatie toe tussen training en test

Eerste resultaat - tris

- De oorspronkelijke bedoeling was om te vergelijken met de klaarzetlijsten. Deze zijn specifiek voor een combinatie van
 - (type ingreep, ok, arts)
- Om meer statistiek te krijgen in de gegevens (en een betere stratificatie van de labels tussen train en test) hebben na het VOKA event besloten om een test uit te voeren door verder aggregatie van de gegevens
- Deze test agereert de vorige ingreep over alle artsen, maar nog steeds voor dezelfde OK ruimte
- aantal items om uit te kiezen : 115
- aantal voorbeelden om op te trainen : 360

Eerste resultaat - tris

Vertaling naar meer klinische bruikbare cijfers

Ruwe eerste test voor 1 combinatie (procedure/OK) – geen implantaten

- gevoeligheid om nodige items te suggereren
 - min: 0.0, median: 0.67, mean: 0.61, max: 1.0
 - performantie in het niet-selecteren van niet-benodigde items
 - min: 0.96, median: 0.96, mean: 0.96, max: 0.96
- Dit is duidelijk beter
- Gemiddeld wordt twee derde van de benodigde items voorspeld.

Eerste resultaat - quater

- Deze test is een uitbreiding van de vorige
- De gegevens worden geaggregeerd over alle OK's en artsen m.a.w. alle gegevens voor 1 type ingreep worden gebruikt
- aantal items om uit te kiezen : 156
- aantal voorbeelden om op te trainen : 796

Eerste resultaat - quater

Vertaling naar meer klinische bruikbare cijfers

Ruwe eerste test voor 1 combinatie (procedure) – geen implantaten

- gevoeligheid om nodige items te suggereren
 - min: 0.0, median: 0.5, mean: 0.54, max: 1.0
 - performantie in het niet-selecteren van niet-benodigde items
 - min: 0.96, median: 0.96, mean: 0.96, max: 0.96
- Dit is duidelijk slechter
- Gemiddeld wordt de helft van de benodigde items voorspeld
- De vermoedelijke oorzaak is het grote aantal labels dat in test beland maar dat niet in training aanwezig omwille van items die weinig gebruikt worden.

Eerste resultaat - quint

- Deze test is een uitbreiding van de vorige
- Alle items die slechts 1 voorkomen, worden verwijderd uit de dataset
- aantal items om uit te kiezen : 37
- aantal voorbeelden om op te trainen : 458

Eerste resultaat - quint

Vertaling naar meer klinische bruikbare cijfers

Ruwe eerste test voor 1 combinatie (procedure) – geen implantaten

- gevoeligheid om nodige items te suggereren
 - min: 0.17, median: 0.75, mean: 0.75, max: 1.0
 - performantie in het niet-selecteren van niet-benodigde items
 - min: 0.96, median: 0.96, mean: 0.96, max: 0.96 Dit is duidelijk slechter
- Gemiddeld wordt de 75% van de benodigde items voorspeld

Conclusies

- Al lijkt zeker een toegevoegde waarde te kunnen leveren
 - De eerste resultaten werden becomen zonder veel modellen te tunen te doen. Ze zijn gericht op het verkennen van de effecten van de data op de performantie.
 - De resultaten van het 'recommender' algoritme kunnen nog verfijnd worden in een tweede stap door een ander ML algoritme
 - De resultaten tonen dat het mogelijk is om de benodigde items voor een ingreep te voorspellen.
 - Via extra overleg met de domein experts (bv in verband met aggregatie van gegevens, of samenvoegen van categoriën) kan dit tot een bruikbaar tool worden ontwikkeld.

Conclusies

- Centraal datawarehouse (alle data) is een belangrijke stap
 - Data van andere diensten/delen leveren een meerwaarde bij het optimaliseren/beheersen van dienstgebonden processen
 - Data wordt best beschikbaar gemaakt aan de mensen op de werkvloer die de meest gedetailleerde domeinkennis hebben
 - niet alleen voor directie en hoger kader
 - Verhogen van locale “datageletterdheid” is een belangrijk piste om nuttig gebruik te maken van de data
 - toegankelijkheid niet enkel via IT of externe leveranciers
 - Eigen onderzoeksvraag : kunnen software systemen meer menselijk geletterd en meer domein geletterd gemaakt worden ?
 - bv via conversatie interfaces