

BIG DATA

Visa Honkanen, LT, dosentti, CDO

Helsingin Yliopistollinen keskussairaala ja HUS Kuntayhtymä

Big data $\neq$ a lot of data

Velocity
Volume
Variety
Veracity

vauhdilla
valtavasti
vaihtelevaa
virheellistä

Hyvä lääketieteellinen hoito on useimmiten ennakoivaa

TYÖKALUT

- 1) Tiedot hyödynnettävään muotoon
 - 1) Data -> tietoaallas -> kuratointi -> datavarasto
- 2) Keinot jalostaa dataa
 - 1) Analytiikkatyökalut -> yhteistyöpilotit
 - 2) Algoritmikirjastot -> yhteistyöpilotit

HUS Helsingin yliopistollinen sairaala

ANALYTIKKA- JA KEINOÄLYKEHITYSPALVELUT

ANALYTIKKA- JA KEINOÄLYKEHITYSPALVELUT

AI-tiimin toiminta käynnistyi syyskuussa 2017

Analytiikka- ja keinoälykehityspalvelut tammikuussa 2019

AI fokusalueet:

- Tehohoitoennusteet
- kuva-analytiikka
- suomenkielinen tekstianalytiikka

Noin 15 projektia käynnissä

- Tutkimusprojektit 2/3
- Toiminnan kehitysprojektit 1/3

**Analytiikka- ja
tietopalvelut**

1 analyytikko

4 data scientist

2 projektipäällikköä

**Koneoppimissovellukset
ja -järjestelmät**

1 analyytikko

3(+1) data/ML engineer

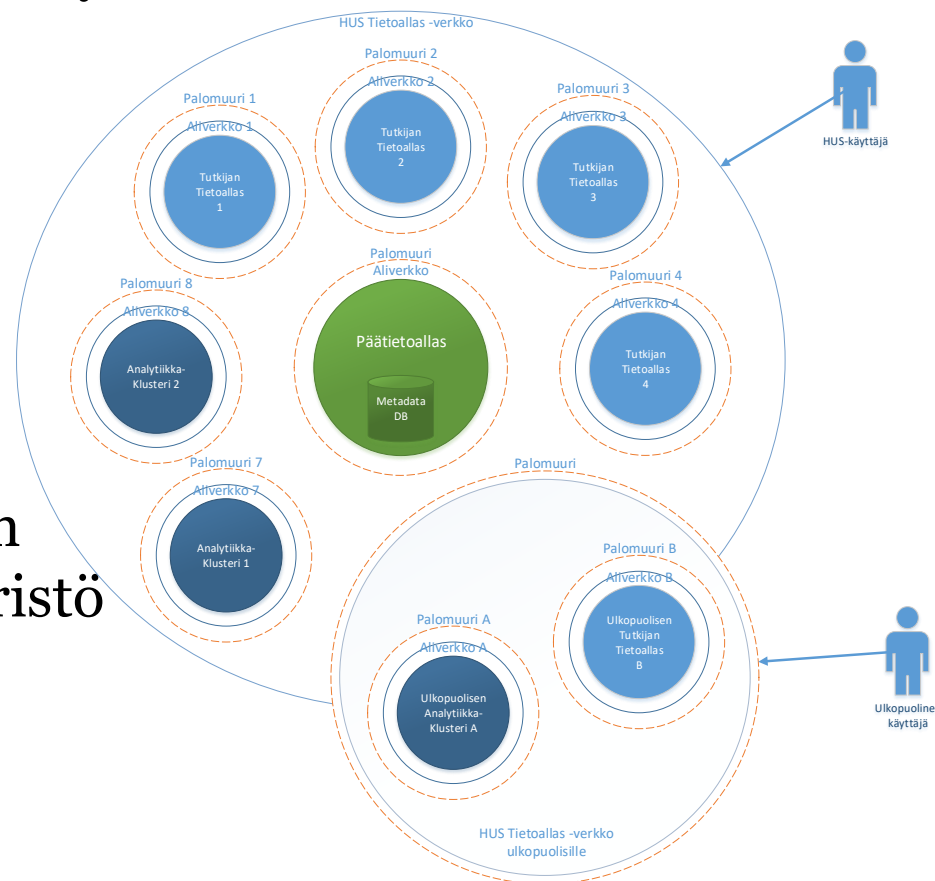
(+1) tietoarkkitehti

2 projektipäällikköä

HUS-TIETOALLAS

Tietoallas tuotantokäytössä 2018

- Spark-klusteri/
Virtuaalikone
- Data ei siirry
HUSin
verkon ulko-
puolelle
- Tietoturvallinen
käsittely-ympäristö



Kymmeniä
potilastietojärjestelmiä
integroituna
Raakadataa

HUSIN TIETOPALVELU

Palvelu käynnistetty syksyllä 2018

- Tietojen siirto tietoaaltaalle
- Tietoluovutukset (tutkimuslupien mukaisesti)
- Biopankkiyhteistyö

Datakatalogi työn alla

Saatavuuspalveluita tulossa kevään 2019 aikana

tietopalvelu@hus.fi

HUS Helsingin yliopistollinen sairaala

DATA

DATA

Tietoaltaassa ”kaikki data” lähdejärjestelmistä

- Raakadataa
- Ei sellaisenaan kovin helposti hyödynnettävissä
- Tarvittu työaika raakadatasta poimintatulokseen (1pv—3kk)

Tietohallinnossa toteutetaan järjestelmäkohtaista tietomallia

- Raportointia varten **jäsennelty** tieto
- Tietopalvelun tarvitsema tieto

Jäsenteleminen vaatii data-osaamista/tuntemusta

HUS Helsingin yliopistollinen sairaala

TEKOÄLYPROJEKTIN KÄYNNISTYMINEN

ENSIARVIOINTI

- Ratkaistava kysymys
 - Selkeä outcome
 - Vaikuttavuuden arviointi
 - Strateginen arvo
- Datan saatavuus, myös tuotantoympäristössä
- Annotoinnin saatavuus
- Kliinikon työaika
- Ohjausryhmän tuki

Lähetelajittelija:

- Outcome: lähetelaatikko
- Vaikuttavuusarvio 2htv
- ~1 vuoden data kertynyt
- Data valmiina tietoaaltaalla
- Annotointi valmiina tietoaaltaalla
- Lähetteitä käsin lajitelleet toimittavat avainsanalistoja

TUOTANNOLLISTAMINEN

- Integrointi järjestelmiin
- Mallin (jatko)kehittäminen
- Tuotannon ympäristön monitorointi
- Tuotannossa olevan mallin monitorointi

Lähetelajittelija:

- RPA integrointi
- Tuotantoympäristö: Flask/REST, jatkossa Azure ML (kehitysympäristössä nyt)
- Tuotantoympäristön monitorointi: Tietoallasympäristön monitoroinnissa (kehitysympäristössä nyt)
- Mallin monitorointi: Azure ML (kehitysympäristössä nyt)

HUS Helsingin yliopistollinen sairaala

HUSIN

ANALYTIKKAPROJEKTEJA

AI-lähetelajittelija

- Vähennetään käsin tehtävää lähetelajittelua, tavoitteena kehittää automaattinen lähetelajittelija Urologialle ✖_{HUS}
- Pilottilajittelija tehty, (robotiikka-)integraatio tehty, tuotannossa

Potilaiden ja päivystysavun ruuhkamittarit ja –ennusteet

- Potilasmäärien ja potilaskäyntien kestoennusteiden näyttäminen potilaille ja päivystysavulle ✖_{HUS}
- Näkymät käytössä HYKS Akuutin toimipisteissä ja päivystysavulla

Päivystyskäyntien arvioidut
kokonaiskestot tällä hetkellä
(keskiarvo)

Aulapotilaat **4 t** 25 min

Vuodepotilaat **6 t** 43 min

Päivystyspisteen
aulanäytöllä

Ennuste
potilaiden
käyntien
kestosta (ennen
siirtymistä kotiin
tai toiselle
osastolle)

Tilannekuvanäytöt

- Potilaspaikka- ja henkilökuntanäkymä jonka avulla voidaan helpommin hahmottaa toimintakokonaisuuden tilanne
- Naistentaudeilla pilotti on ollut käytössä vuodesta 2016, tuotantokäytössä 2019. Käyttöä laajennetaan

Asiakaspalautteen tekoälyanalyysi (POC)



- Vapaamuotoisen asiakaspalautetekstin automaattinen luokittelija pilotoitiin lupaavin tuloksin
- Projekti jäädytetty uuden asiakaspalauttejärjestelmän hankinnan takia

KONEOPPIMISTA

Rannemurtumien tekoälyanalyysi



- Automaattinen leikkaustarpeen arviointi ranteen röntgenkuvista
- Annotointityö käynnissä

Palovammakuvien analysointi (POC)



- Palovammojen vakavuuden automaattinen arviointi kuva-analyysillä.
- Pilottitutkimus tehty lupaavin tuloksin, jatkoprojektia harkitaan

Artificial Intelligence in Prognostication of Traumatic Brain Injuries



- Aivovammapotilaan toipumisennuste
- Hyvät tutkimustulokset on laskettu ja julkaisun lähetetty

Pienten keskosten verenmyrkytyksen ennustaminen

IBM

- Keskosten verenmyrkytyksen ennustaminen muilla potilastiedoilla rikastetusta medianisoidusta potilasmonitoridatasta.
- Artikkelin lähetetty vertaisarviointiin

Nettiterapiadatan louhinta

 HUS

- Ennustetaan potilaiden pysyvyyttä (adherence) ja hoidon lopputulosta
- Tutkimuslupa myönnetty, aineisto kerätty, analysointi aloitettu

Keskosten kehityksen ennustaminen

A? IBM  HUS  NOKIA  UNIVERSITY OF EASTERN FINLAND

- Predict neurodevelopmental outcome of prematurely born using ICU and MRI data
- Tutkimuslupaa ja rahoitusta haetaan, tutkimusryhmä muodostettu

Lukinkalvon alaisen verenvuodon automaattinen havaitseminen

- Tutkimuslupa myönnetty, aineistoa segmentoidaan ja vaihtoehtoisia analysointimenetelmiä testataan



PLANMECA GROUP

NOKIA

Raskausdiabetes

- Mobiilisovellus raskausdiabeetikon omaseurantaan. Ravinto-, aktiivisuus- ja verensokeridatan analyysillä mahdollisuus esimerkiksi potilaan online-ohjaukseen.

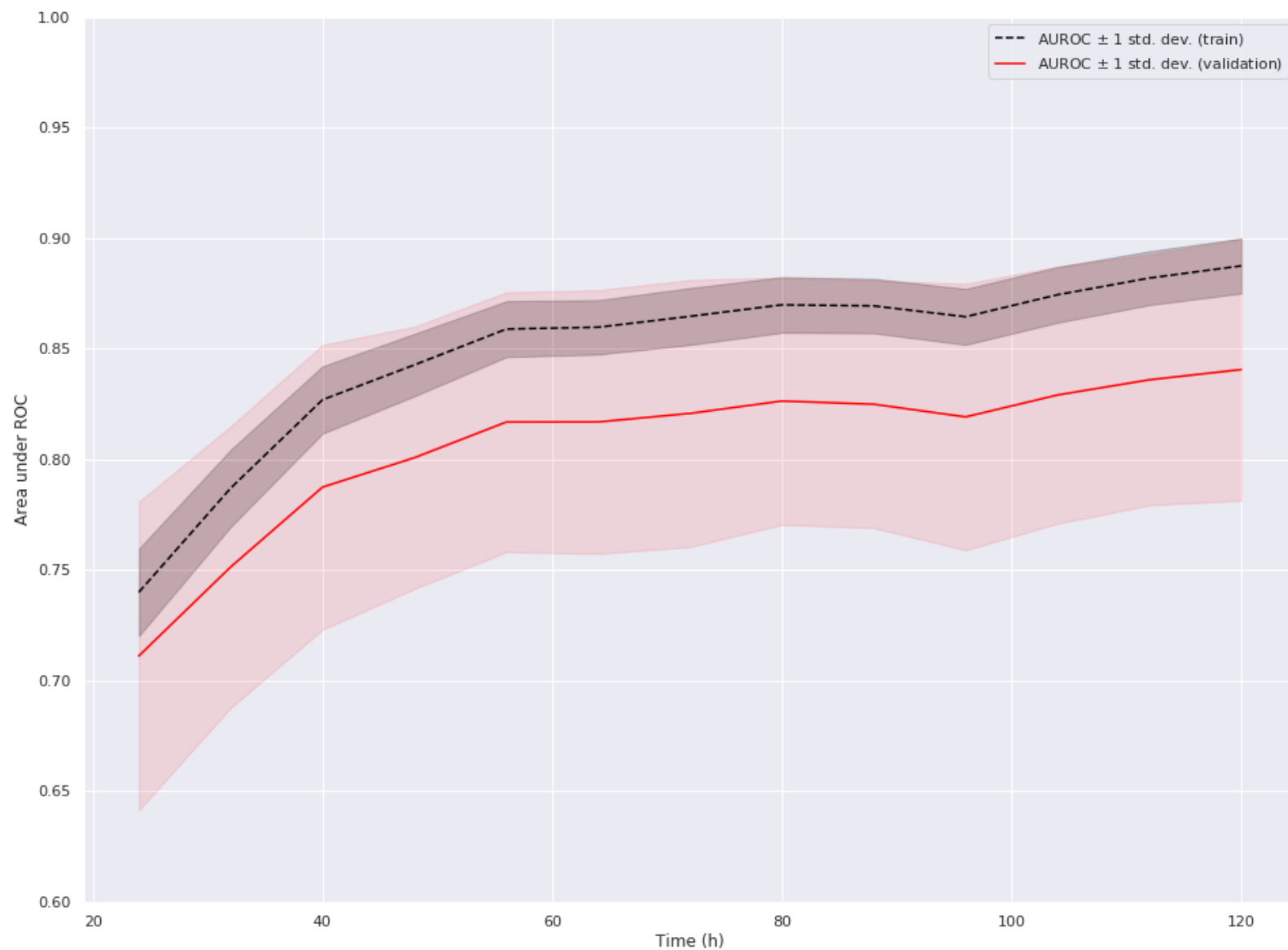


Rare Diseases Early Diagnostics

- Harvinaisten sairauksien löytäminen potilaspolkujen, laboratoriodatan, lausuntojen ja potilaskertomustekstien perusteella



**AI on parhaimillaan dynaamisissa
malleissa haastavissa kliinisissä
tilanteissa**



Luokittelutarkkuus
(AUC) ajan
functiona

**Käyttökelpoinen AI aktiivissa
hoitotilanteessa vaatii vakaa ja
luotettavan datavirran.**

Etiikka

utilitarismi

kategorinen imperatiivi

AIVOVERENVUOTO

TEKOÄLYPROJEKTEISSA OPITTUA

Kaikissa projekteissa tiivis ja toimiva yhteistyö klinisten ja teknisten asiantuntijoiden välillä on ollut välttämätöntä

- Aktiivista työaikaa kuluu yleensä viikkoja tai jopa kuukausia

Menestyksen määrittää datan ja annotoinnin laadukkuus

- Kun objektiivista totuutta ei ole saatavilla, kone voi oppia vain niin hyväksi kuin opettajansa (jos annotoitua dataa on paljon, kone voi kompensoida virheellisiä annotointeja)
- Hyvä lähtöaineisto ja valmis laskentaympäristö mahdollistavat nopean projektin

Tulosten riippumaton arviointi kriittinen osa tekoälyprojekteja

Merkittävä osa projekteista ei tuota käyttökelpoista lopputulosta

MENESTYSTEKIJÄT

Kliininen osaaminen

IT-osaaminen

Analytiikkaosaaminen

Osto-osaaminen

Tieteellinen osaaminen

Paljon hyvälaatuista dataa

Mukaanottava heimokulttuuri

Ihmisiä, joilla on sekä kliinistä että data science ymmärrystä

HUS IT, ANALYTIIKKA- JA KEINOÄLYKEHITYSPALVELUT

Teijo Konttila

Analytiikka- ja tietopalvelut

Miika Leminen

Annu Kaila

Antti Rantala

Walter Grönholm

Henri Laine

Taru Hermens

Koneoppimissovellukset ja –järjestelmät

Eetu Pursiainen

Antti Martikainen

Antonios Thanellas

Kari Laukkanen

Marjo Halonen

NN Tietoarkkitehti

NN Data Engineer