Mellom data og innsikt bor metodene

10. januar 2019

Ingrid K. Glad, Matematisk institutt, UiO

BigInsight


sfi
Senter for forskningsdrevet innovasjon
Norges forskningsråd
BigInsight
STATISTICS FOR THE KNOWLEDGE ECONOMY
UiO : University of Oslo
NR Norsk Regnesentral
NORWEGIAN COMPUTING CENTER


Statistisk sentralbyrå
Statistics Norway
nav
ABB
Skatteetaten
DNV·GL
Oslo
University Hospital
BigInsight
UiO : University of Oslo
HYDRO
CANCER
Registry of Norway
NR
Norsk
Regnesentral
NORWEGIAN COMPUTING CENTER
DNB
UNIVERSITETET I BERGEN
folkehelseinstituttet
telenor
Gjensidige

BigInsight

Personalised solutions

Transient predicition

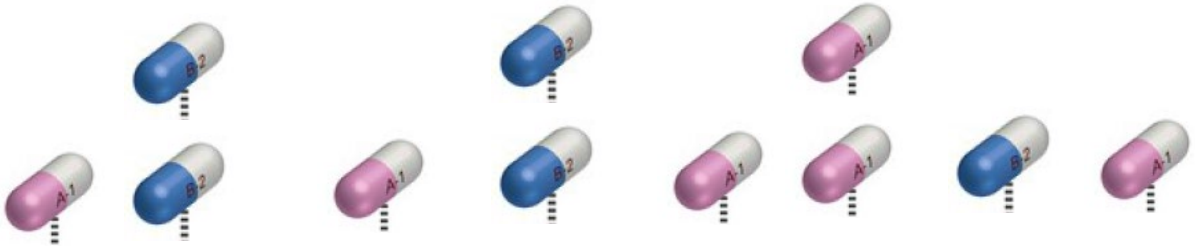
Helse Markedsføring Svindel Sensorer Energi Forklarbar AI

nyskapende – nøyaktig – ansvarlig – forklarlig – åpen – rettferdig


PERSONTILPASSET CANCER-TERAPI

Persontilpasset cancer-terapi:
Gi hver enkelt pasient skreddersydd behandling

- mange ulike medisiner
- mange kombinasjoner
- ulike doser
- i ulik rekkefølge



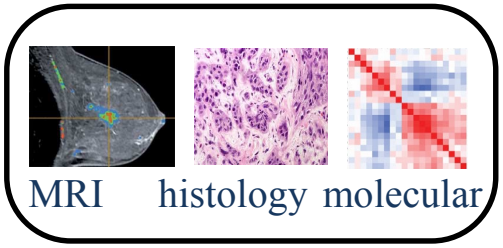
tid

Mange virtuelle kopier av samme pasient

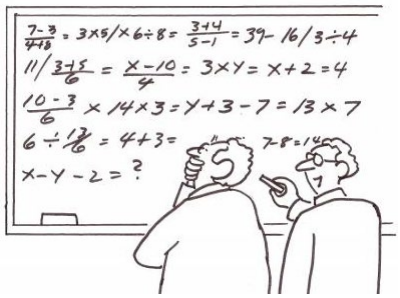
- Simulerer terapi-strategier
- Hvilken virket best?

kreft-pasient

personlige kliniske data



multi-skala matematisk modell for cancer-vekst



$$\frac{d[\mathrm{TP53}]_{\boldsymbol{x}}}{dt} = k_7 - k_7' \frac{K(\boldsymbol{x},t)}{K_{\mathrm{TP53}} + K(\boldsymbol{x},t)}[\mathrm{TP53}]_{\boldsymbol{x}},$$

$$\frac{d[\mathrm{sVEGF}]_{\boldsymbol{x}}}{dt} = k_8 - k_8' \frac{K(\boldsymbol{x},t)}{K_{\mathrm{VEGF}} + K(\boldsymbol{x},t)}[\mathrm{sVEGF}]_{\boldsymbol{x}} + k_8'' \frac{[\mathrm{TP53}]_{\boldsymbol{x}}[\mathrm{sVEGF}]_{\boldsymbol{x}}}{J_5 + [\mathrm{sVEGF}]_{\boldsymbol{x}}}$$

$$s_K \frac{\partial K}{dt} - D_k \nabla^2 K = r_k(K_0 - K)\mathcal{G}(\boldsymbol{x},t) - \frac{\phi_k K}{K_1 + K}\delta(\boldsymbol{x},t)$$

$$s_V \cdot \frac{\partial V}{dt} - D_v \nabla^2 V = r_v([\mathrm{sVEGF}]_{\boldsymbol{x}})\delta(\boldsymbol{x},t) + k_d C - k_a A_e V - \psi_v V.$$

$$s_A \cdot \frac{\partial A}{dt} - D_A \nabla^2 A = r_A(A_1(t) - A)\mathcal{G}(\boldsymbol{x},t) + k_d C - k_a A V$$

$$\frac{\partial G^j}{dt} - D_{G^j} \nabla^2 G^j = r_{G^j}(G_1^j(t) - G^j)\mathcal{G}(\boldsymbol{x},t) - \psi_{G^j} G^j$$

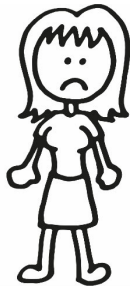
$$\frac{dG_1^2(t)}{dt} = -\frac{q_2}{w_1}G_2^2(t) - \frac{q_3}{w_1}G_3^2(t) - \frac{cl_2}{w_1}G_1^2(t) + \frac{q_3}{w_3}G_3^2(t) + \frac{q_2}{w_2}G_2^2(t)$$

$$\frac{dG_2^2(t)}{dt} = -\frac{q_2}{w_2}G_2^2(t) + \frac{q_2}{w_1}G_1^2(t) \qquad \frac{dA_1(t)}{dt} = -\frac{q}{v_1}A_1(t) - \frac{cl}{v_1}A_1(t) + \frac{q}{v_2}A_2(t)$$

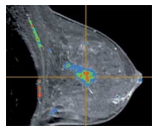
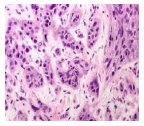
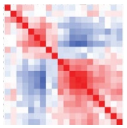
$$\frac{dG_3^2(t)}{dt} = -\frac{q_3}{w_3}G_3^2(t) + \frac{q_3}{w_1}G_1^2(t) \qquad \frac{dA_2(t)}{dt} = -\frac{q}{v_2}A_2(t) + \frac{q}{v_1}A_1(t),$$

Modellere biologiske prosesser
- cellevekst og respons på medisin

kreft-pasient

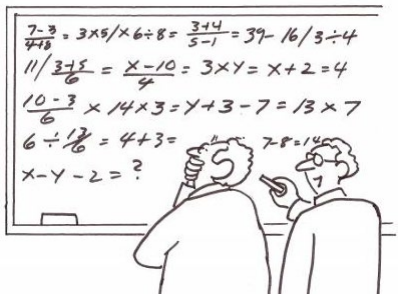


personlige kliniske data

MRI histology molecular

multi-skala matematisk modell for cancer-vekst



$$\frac{d[\mathrm{TP53}]_{\boldsymbol{x}}}{dt} = k_7 + k_7' \frac{K(\boldsymbol{x},t)}{K_{\mathrm{TP53}} + K(\boldsymbol{x},t)}[\mathrm{TP53}]_{\boldsymbol{x}},$$

$$\frac{d[\mathrm{sVEGF}]_{\boldsymbol{x}}}{dt} = k_8 + k_8' \frac{K(\boldsymbol{x},t)}{K_{\mathrm{VEGF}} + K(\boldsymbol{x},t)}[\mathrm{sVEGF}]_{\boldsymbol{x}} + k_8'' \frac{[\mathrm{TP53}]_{\boldsymbol{x}}[\mathrm{sVEGF}]_{\boldsymbol{x}}}{J_5 + [\mathrm{sVEGF}]_{\boldsymbol{x}}}$$

$$s_K \frac{\partial K}{dt} - D_k \nabla^2 K + r_k(K_0 - K)\mathcal{G}(\boldsymbol{x},t) - \frac{\phi_k K}{K_1 + K}\delta(\boldsymbol{x},t)$$

$$s_V \cdot \frac{\partial V}{dt} - D_v \nabla^2 V + r_v([\mathrm{sVEGF}]_{\boldsymbol{x}})\delta(\boldsymbol{x},t) + k_d C - k_a A_c V - \psi_v V.$$

$$s_A \cdot \frac{\partial A}{dt} - D_A \nabla^2 A + r_A(A_1(t) - A)\mathcal{G}(\boldsymbol{x},t) + k_d C - k_a A V$$

$$\frac{\partial G^j}{dt} - D_{G^j}\nabla^2 G^j = r_{G^j}(G_1^j(t) - G^j)\mathcal{G}(\boldsymbol{x},t) - \psi_{G^j} G^j$$

$$\frac{dG_1^2(t)}{dt} = -\frac{q_2}{w_1}G_2^2(t) - \frac{q_3}{w_1}G_3^2(t) - \frac{cl_2}{w_1}G_1^2(t) + \frac{q_3}{w_3}G_3^2(t) + \frac{q_2}{w_2}G_2^2(t)$$

$$\frac{dG_2^2(t)}{dt} = -\frac{q_2}{w_2}G_2^2(t) + \frac{q_2}{w_1}G_1^2(t)$$

$$\frac{dG_3^2(t)}{dt} = -\frac{q_3}{w_3}G_3^2(t) + \frac{q_3}{w_1}G_1^2(t)$$

$$\frac{dA_1(t)}{dt} = -\frac{q}{v_1}A_1(t) - \frac{cl}{v_1}A_1(t) + \frac{q}{v_2}A_2(t)$$

$$\frac{dA_2(t)}{dt} = \frac{q}{v_2}A_2(t) + \frac{q}{v_1}A_1(t),$$

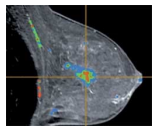
Estimere parametre fra pasient-data

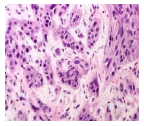
Persontilpassede simuleringer av behandlings-strategier

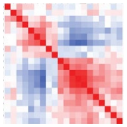
kreft-pasient

onkolog

personlige kliniske data

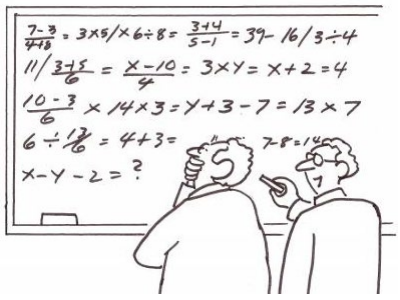




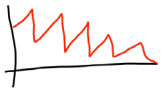


MRI histology molecular

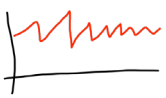
multi-skala matematisk modell for cancer-vekst



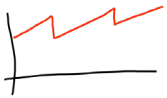
persontilpassede simuleringer



strategi 3

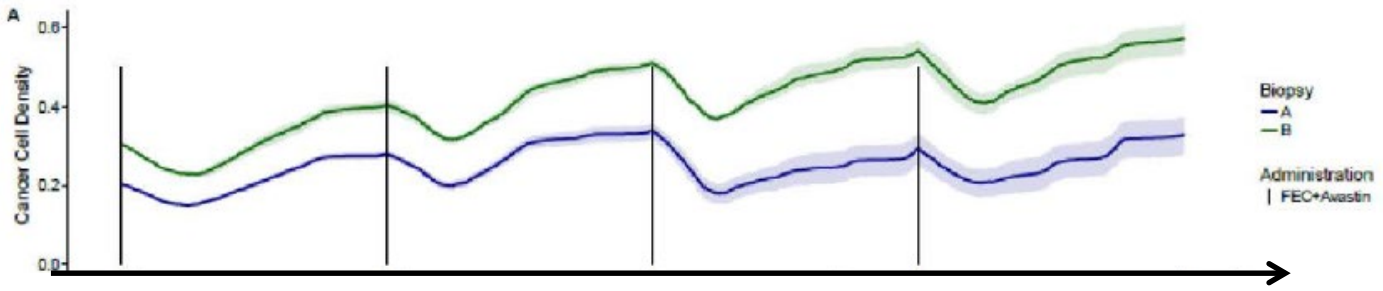


strategi 2



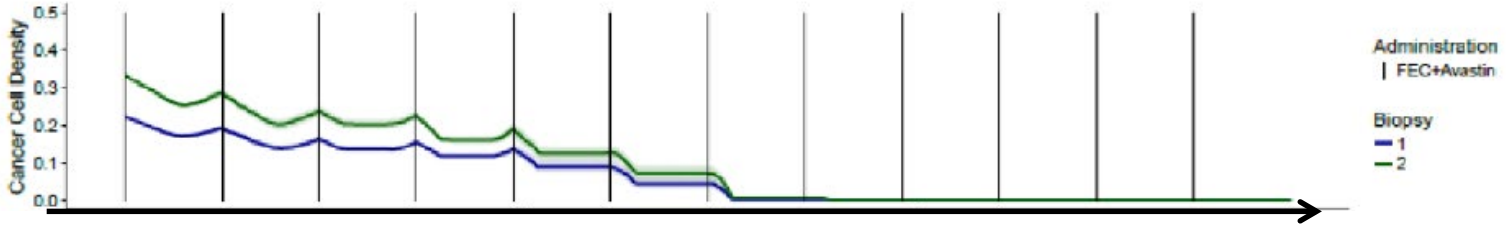
strategi 1

Uncertainty



12 uker

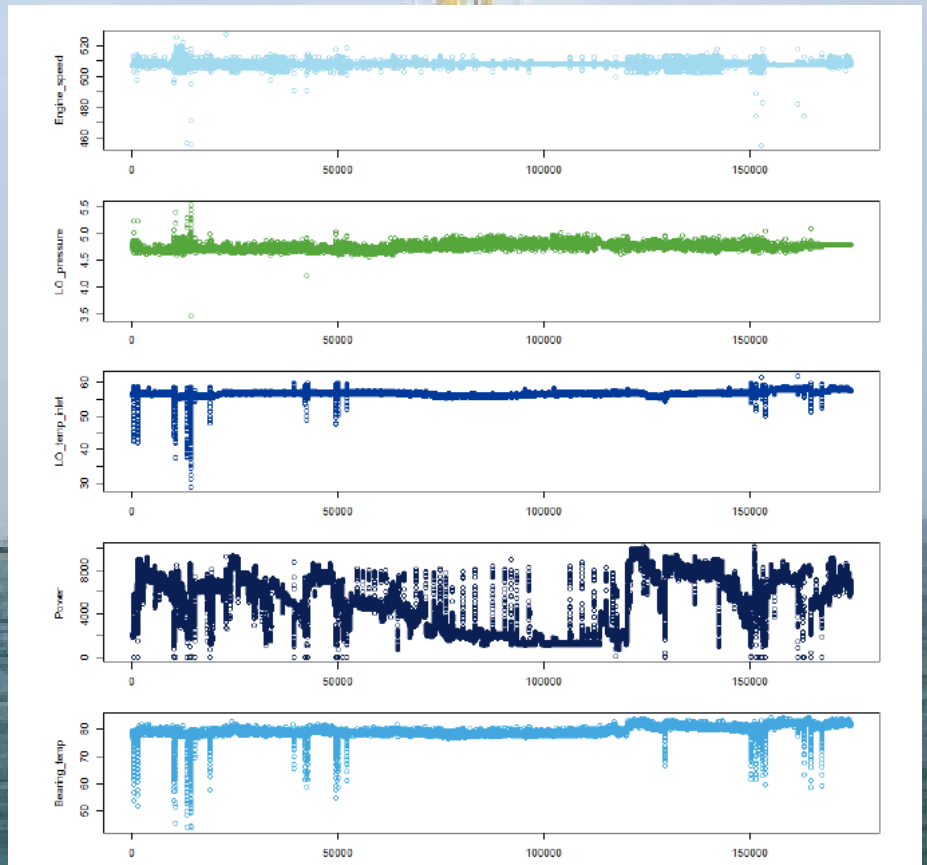
Kjemoterapi hver 3. uke



12 uker

Kjemoterapi hver uke i stedet

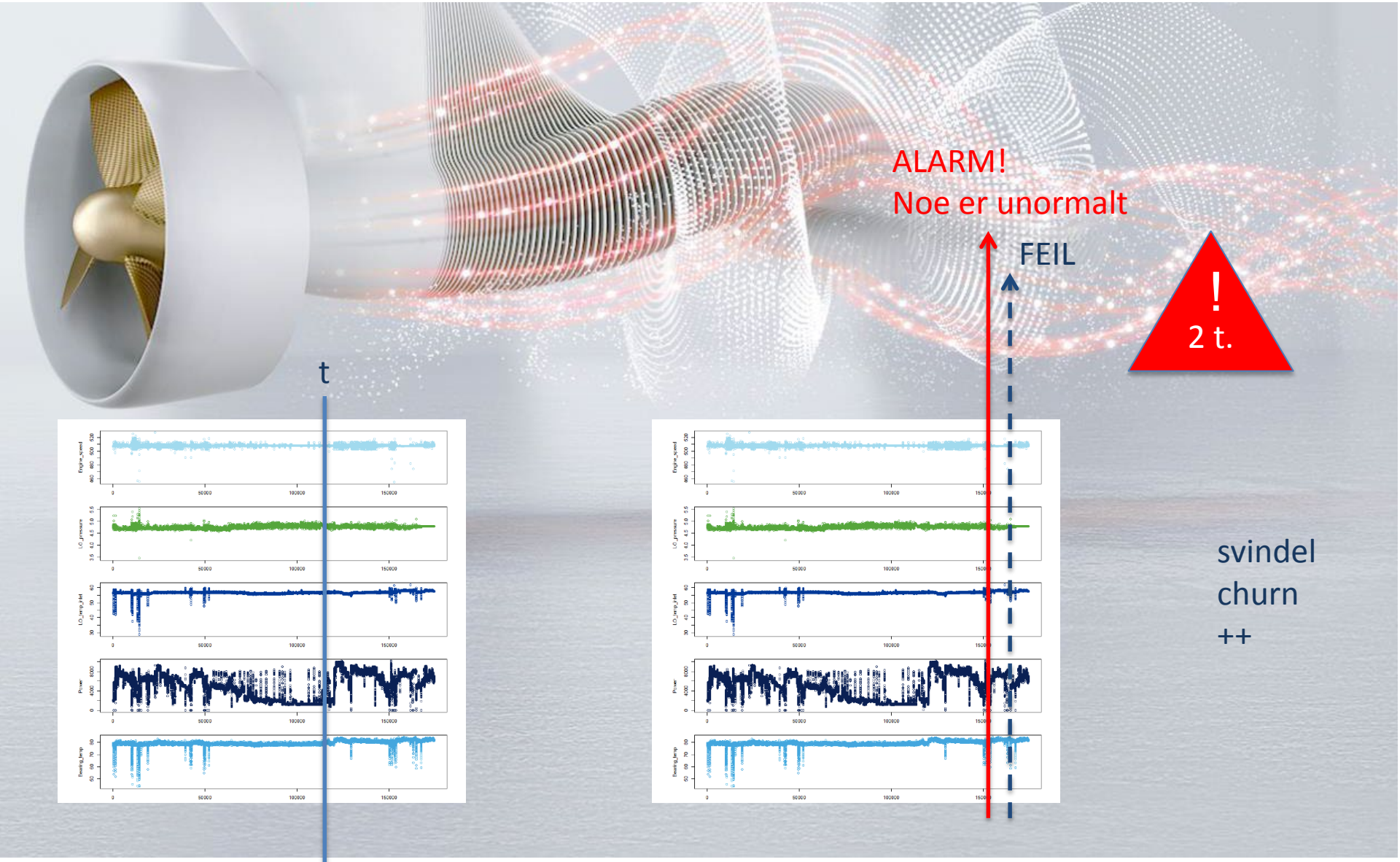

SENSOR-DATA FOR SIKKERHET TIL SJØS


Engine_speed
460
480
500
520
0
50000
100000
150000
LO_pressure
3.5
4.0
4.5
5.0
5.5
LO_temp_inlet
30
40
50
60
Power
0
4000
8000
Bearing_temp
50
60
70
80

+

+

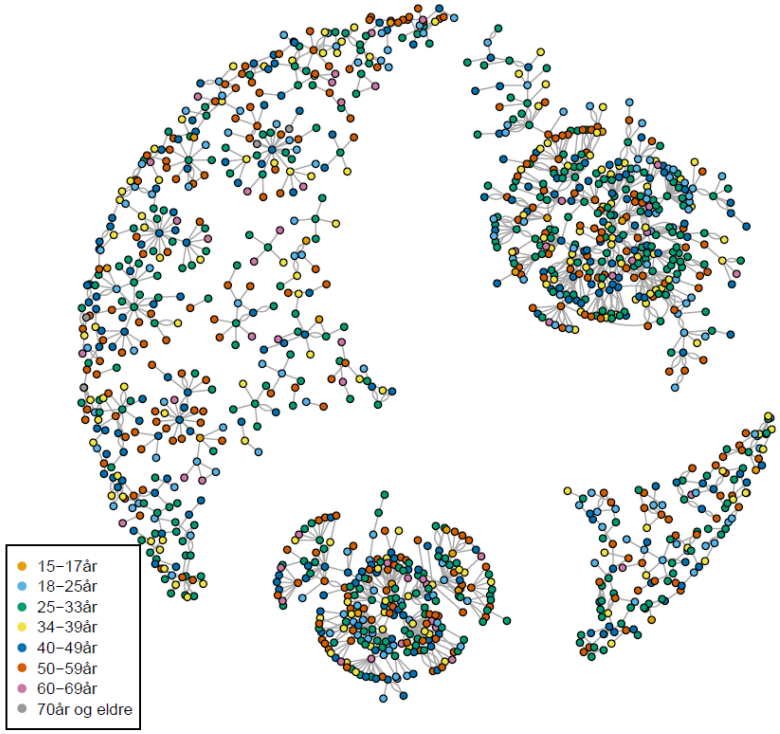
+


ALARM!
Noe er unormalt
FEIL
! 2 t.
t
svindel
churn
++
Lære fra normale tilstander
For ethvert tidspunkt t: observer og rekonstruer
residual$_t$
test$_t$
ALARM
trenger mer obs
NORMALT

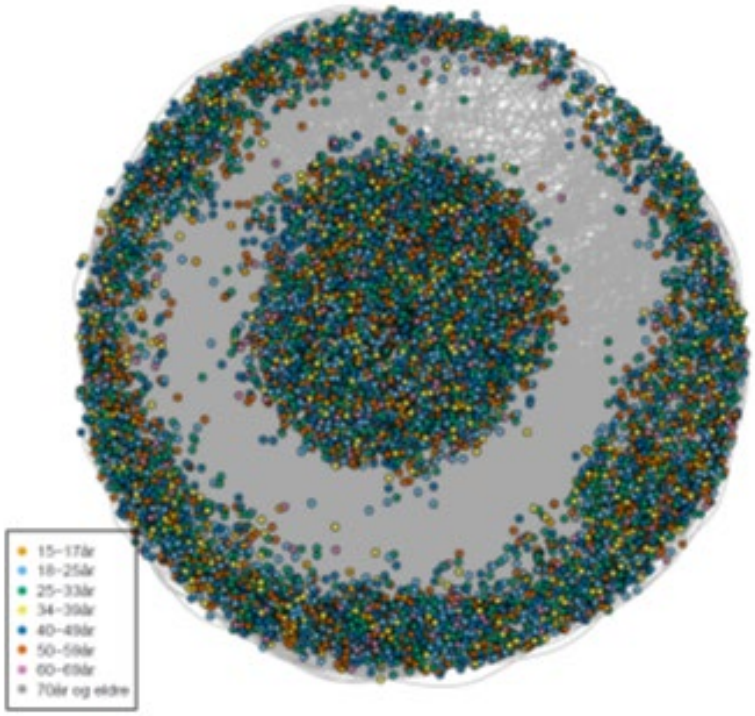

VIPPS! – SOM ILD I GRESS

Vipps-nettverket: Den første uken

30. mai 2015



5. juni 2015



Datasett første 1.5 år:

1.8 millioner brukere

28.9 million transaksjoner

Hva er underliggende vekst-modell for Vipps-nettverket?

Kombinasjon av

- preferanse-modell (jo flere kanter, jo mer sannsynlig å opprette nye)
- fitness-modell (hver node har egen innebygget egenskap («fitness»))

Likelihood-estimering fra 28.9 millioner transaksjoner gir en tilpasset vekstmodell

- preferanse-delen av modellen ≈ Facebook og Twitter
- svært ulik fitness for ulike individer

Kan brukes til:
- Å få innsikt i struktur og prosess
- Identifisere nøkkelpersoner i nettverket
- Predikere fremtidig utvikling

BigInsight

BigInsight Day 2018 at DNB