

VELKOMMEN TIL HON2200

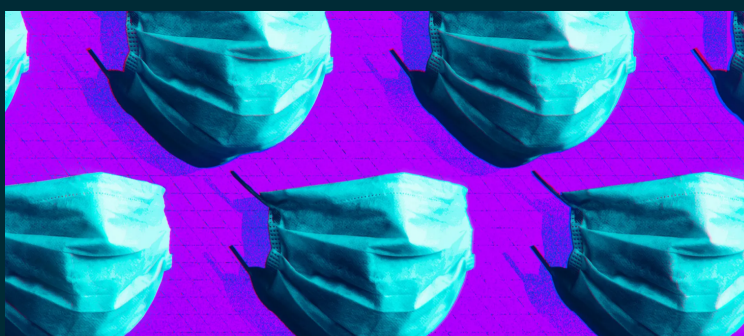
MÅL FOR DAGENS ØKT

- Hva er data-drevne prosjekter?
- Hvorfor har maskinlæring blitt så populært?
- Kunne lese inn og utforske datasett med pandas

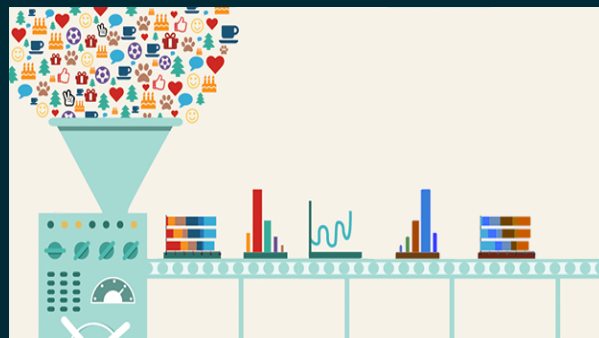
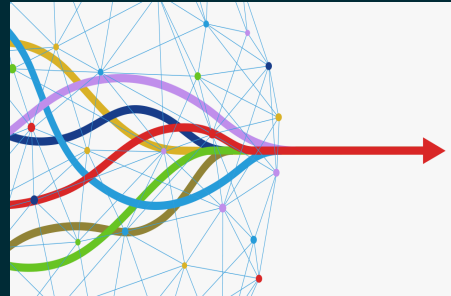
HVA ER SKUMLEST?

English-Norwegian translation in context

That's a stupid question in Norwegian



DATA-DREVEN?



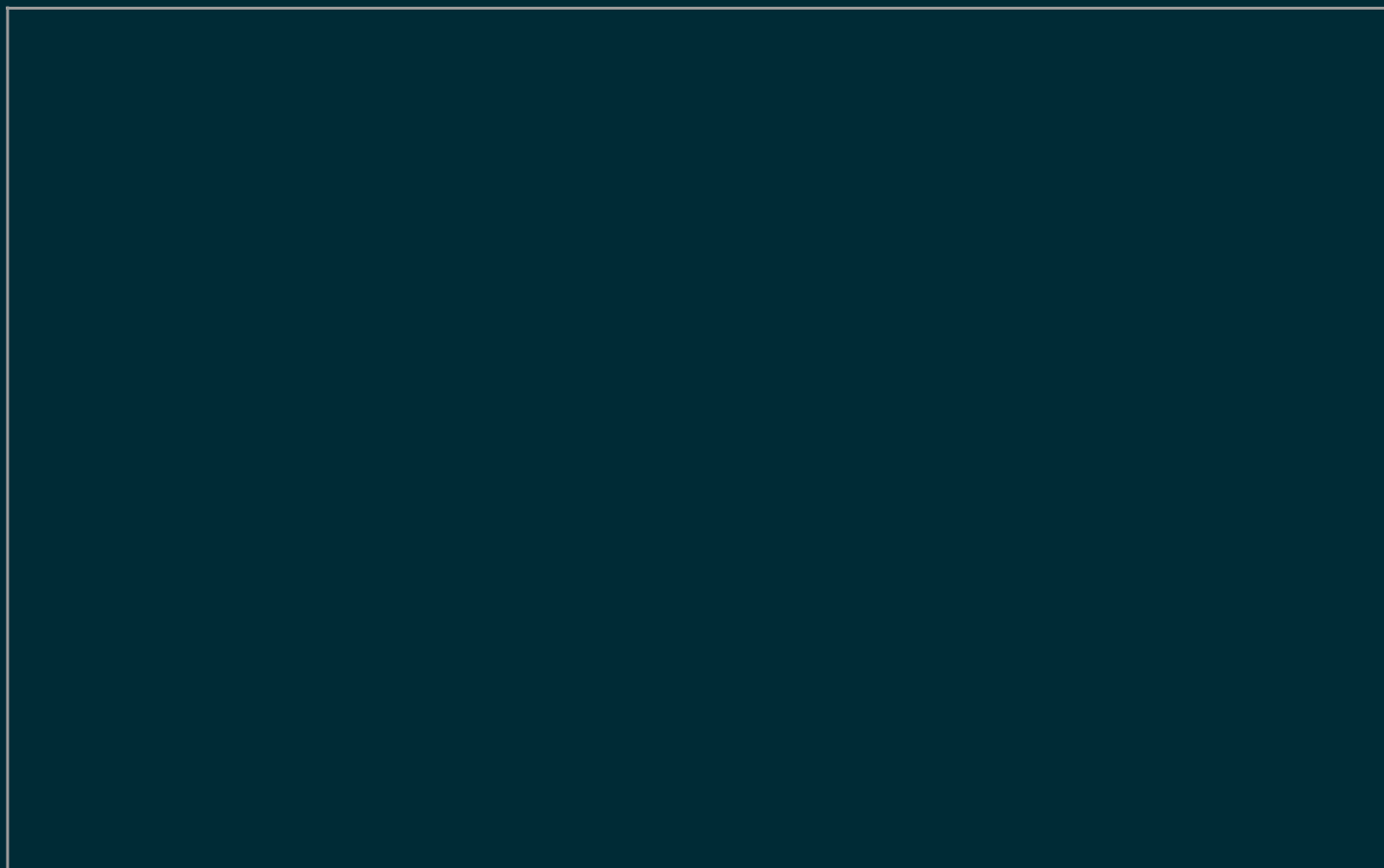
PRAKTISK INFORMASJON

- Undervisning på tirsdager 16:15-18:00
- 1 oblig + 2 prosjekter + muntlig eksamen
- Nettressurser lenket fra emnesiden. Disse endres gjennom semesteret. Lærebok på nett (ISL)
- Fysisk undervisning. Noe streaming for utveksling.
- Undervisere

KOMMUNIKASJONSKANALER

- Emnesiden (og mine studier). Særlig timeplanen
- Discourse hon2200-discourse.uio.no
- Github Classroom for innleveringer
- Epost / telefon (finnes på uio.no)

KURSNETTSIDEN



HON2200 – DATA-DREVNE PROSJEKTER

- Statistisk modellering
- Maskinlæringsmodeller
- Etske vurderinger

STATISTIKK I SAMFUNN OG VITENSKAP

UNDERVEISOPPGAVE

Kom opp med et sett med forutsetninger som gjør følgende utsagn

- Sant
- Usant

Kostnadene ved tiltakene mot koronaviruset var større enn kostnadene ved sykdommen selv

BETINGET SANNSYNLIGHET

- $C(\text{corona}|\text{ingen tiltak})$
- $C(\text{corona}|\text{tiltak})$
- $C(\text{tiltak})$

FORSLAG TIL FORUTSETNINGER

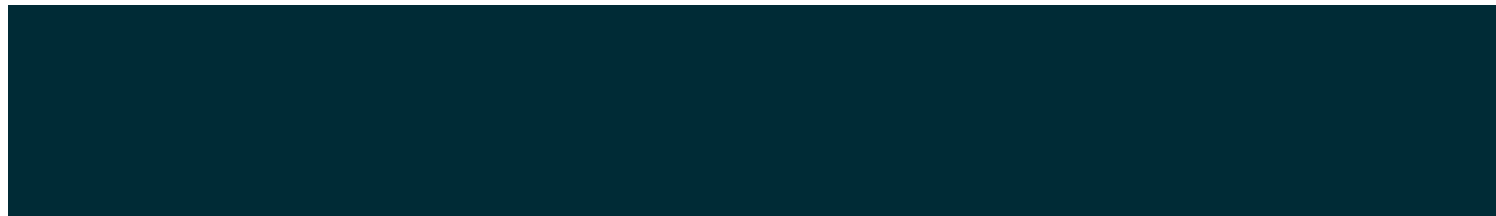
- $C(\text{corona}|\text{ingen tiltak}) < C(\text{tiltak})?$
- $C(\text{corona}|\text{tiltak}) < C(\text{tiltak})?$
- Hva legger vi i begrepet "kostnad"?

BURDE KORONATILTAKENE VÆRT TJENNET TIDLIGERE?

- Europe must come together to confront omicron
- COVID-19 will continue but the end of the pandemic is near
- Covid-19 vaccines and treatments: we must have raw data, now
- Effectiveness of COVID-19 vaccines against the Omicron variant of concern, medArxiv preprint (Ikke les hele denne)

Rådata fra UK:

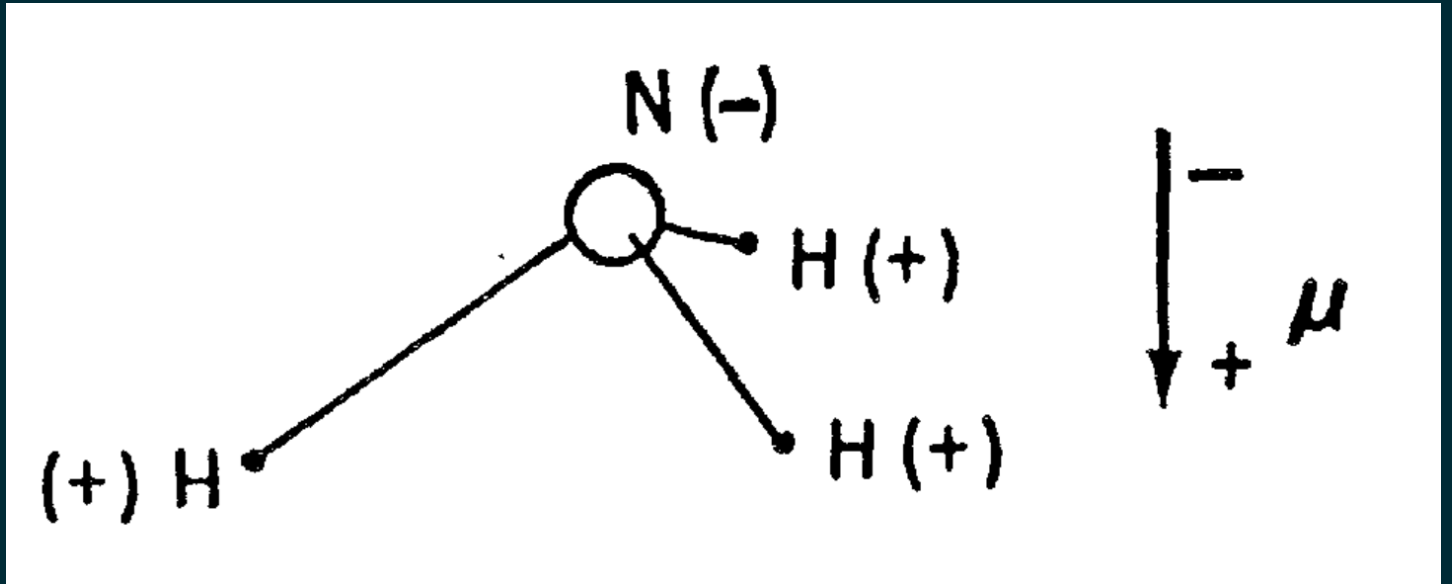
<https://www.gov.uk/government/news/covid-19-omicron-variant-raw-data-released>



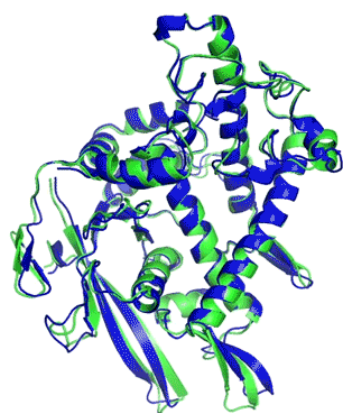
UNDERVEISOPPGAVE

1. Skriv ned på et ark om du har landet på ja eller nei og hvorfor
2. Les opp valget og begrunnelsen for gruppa (alle gjør dette)
3. Identifiser kriteriene de forskjellige personene i gruppa har brukt for å ta et valg
4. Er det noe informasjon dere mangler for å ta et godt valg basert på kriteriene?
5. Oppsummering i plenum. Én i gruppa presenterer

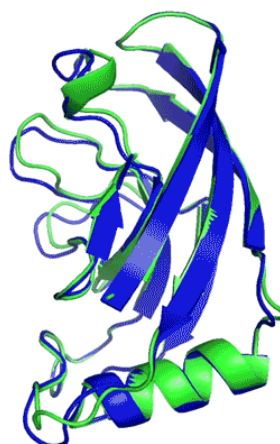
STORDATA – MORE IS DIFFERENT



ALPHAFOLD



T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)



T1049 / 6y4f
93.3 GDT
(adhesin tip)

- Experimental result
- Computational prediction

()

DATBEHANDLING I JUPYTER LAB

(Live-demo)

Studentrepresentanter: Emelie (SV) Mia (MN)
Sveinung (HF)