



SINTEF

# Maskinlæringens klimaavtrykk

Hvordan måles det og hva bruker vi  
målingene til?

Erik Johannes Husom

21. september 2023

# Maskinlæringens klimaavtrykk

ARTIFICIAL INTELLIGENCE

## Training a single AI model can emit as much carbon as five cars in their lifetimes

Deep learning has a terrible carbon footprint.

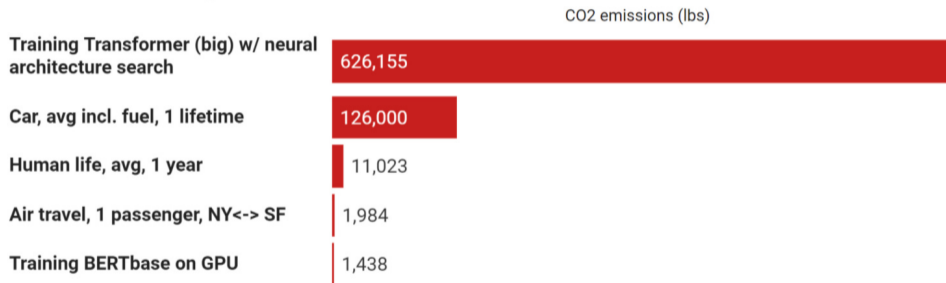
By Karen Hao

June 6, 2019

Overskrift fra MIT Technology Review

## Carbon footprint comparison

Source: Strubell et al, 2019.



Strubell et al. (2019): Energy and Policy Considerations for Deep Learning in NLP



# Agenda



# Agenda

- Maskinl ring og beregningsressurser

# Agenda

- Maskinlæring og beregningsressurser
- Hva slags indikatorer har vi?

# Agenda

- Maskinlæring og beregningsressurser
- Hva slags indikatorer har vi?
- Hvordan utfører vi målinger?

# Agenda

- Maskinlæring og beregningsressurser
- Hva slags indikatorer har vi?
- Hvordan utfører vi målinger?
- Hva kan vi bruke målingene til?





# Livssyklusen til en maskinlæringsmodell



# Livssyklusen til en maskinlæringsmodell

## 1. Datainnsamling



# Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen



## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen
  - Velg modelltype og -konfigurasjon

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen
  - Velg modelltype og -konfigurasjon
  - Læring/trening

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen
  - Velg modelltype og -konfigurasjon
  - Læring/trening
  - Evaluering

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen
  - Velg modelltype og -konfigurasjon
  - Læring/trening
  - Evaluering
6. Eksperimentering: Gjenta steg 4 og 5

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen
  - Velg modelltype og -konfigurasjon
  - Læring/trening
  - Evaluering
6. Eksperimentering: Gjenta steg 4 og 5
7. Utrulling av modellen (*deployment*)

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen
  - Velg modelltype og -konfigurasjon
  - Læring/trening
  - Evaluering
6. Eksperimentering: Gjenta steg 4 og 5
7. Utrulling av modellen (*deployment*)
8. Bruk av modellen (inferens/prediksjon)

## Livssyklusen til en maskinlæringsmodell

1. Datainnsamling
2. Renske data
3. Statistisk analyse, utforsking av data
4. Preprosessering av data
  - Transformere
  - Strukturere
5. Konstruksjon av modellen
  - Velg modelltype og -konfigurasjon
  - Læring/trening
  - Evaluering
6. Eksperimentering: Gjenta steg 4 og 5
7. Utrulling av modellen (*deployment*)
8. Bruk av modellen (inferens/prediksjon)
9. Oppdatering av modellen

# Maskinl ring og beregningsressurser

Tre epoker innen maskinl ring<sup>1</sup>:

---

<sup>1</sup>Sevilla et al. (2022): *Compute Trends Across Three Eras of Machine Learning*



# Maskinlæring og beregningsressurser

Tre epoker innen maskinlæring<sup>1</sup>:

1. *Pre Deep Learning Era* (1952-2010)

---

<sup>1</sup>Sevilla et al. (2022): *Compute Trends Across Three Eras of Machine Learning*

# Maskinlæring og beregningsressurser

Tre epoker innen maskinlæring<sup>1</sup>:

1. *Pre Deep Learning Era* (1952-2010)
2. *Deep Learning Era* (2010-2022)

---

<sup>1</sup>Sevilla et al. (2022): *Compute Trends Across Three Eras of Machine Learning*

Tre epoker innen maskinl ring<sup>1</sup>:

1. *Pre Deep Learning Era* (1952-2010)
2. *Deep Learning Era* (2010-2022)
3. *Large-Scale Era* (2015-2022)

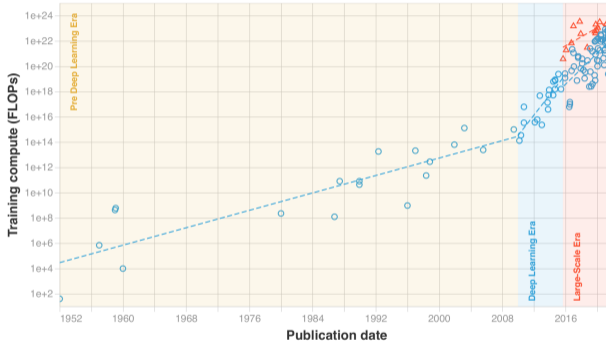
---

<sup>1</sup>Sevilla et al. (2022): *Compute Trends Across Three Eras of Machine Learning*

# Maskinl ring og beregningsressurser

Training compute (FLOPs) of milestone Machine Learning systems over time

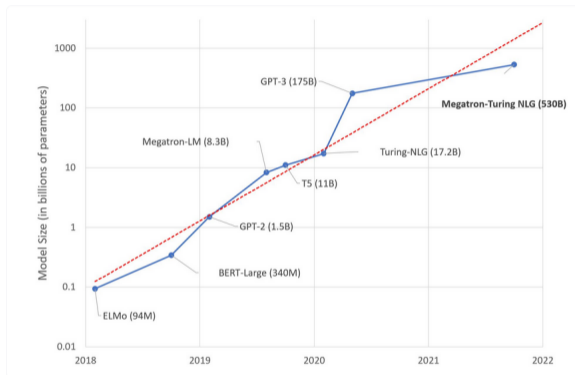
n = 121



Utviklingen av st rrelsen p  maskinl ringsmodeller<sup>2</sup>

<sup>2</sup>Sevilla et al. (2022): *Compute Trends Across Three Eras of Machine Learning*

# Maskinl ring og beregningsressurser



Utviklingen av st rrelsen p  store spr kmodeller

## Maskinl ring og beregningsressurser

- Det siste ti ret med forskning innen maskinl ring<sup>3</sup>:

---

<sup>3</sup>Luccionu et al. (2023): Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning

## Maskinl ring og beregningsressurser

- Det siste ti ret med forskning innen maskinl ring<sup>3</sup>:
  - Klimaavtrykket til ML er  kende

---

<sup>3</sup>Luccionu et al. (2023): Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning

## Maskinlæring og beregningsressurser

- Det siste tiåret med forskning innen maskinlæring<sup>3</sup>:
  - Klimaavtrykket til ML er økende
  - Ca 70% av ML-modellene ble trent ved bruk av høykarbon energikilder (kull, gass, olje)

---

<sup>3</sup>Luccionu et al. (2023): Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning



## Maskinlæring og beregningsressurser

- Det siste tiåret med forskning innen maskinlæring<sup>3</sup>:
  - Klimaavtrykket til ML er økende
  - Ca 70% av ML-modellene ble trent ved bruk av høykarbon energikilder (kull, gass, olje)
  - *Transformer*-modeller er svært populære - og svært karbonintensive

---

<sup>3</sup>Luccionu et al. (2023): Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning

## Maskinlæring og beregningsressurser

- Det siste tiåret med forskning innen maskinlæring<sup>3</sup>:
  - Klimaavtrykket til ML er økende
  - Ca 70% av ML-modellene ble trent ved bruk av høykarbon energikilder (kull, gass, olje)
  - *Transformer*-modeller er svært populære - og svært karbonintensive
  - **Større energiforbruk gir ikke nødvendigvis bedre ytelse**

---

<sup>3</sup>Luccionu et al. (2023): Counting Carbon: A Survey of Factors Influencing the Emissions of Machine Learning

## “Rød” KI vs “grønn” KI

*Red AI stems from the AI community's focus on **accuracy** while paying little attention to efficiency. It leads to a surprisingly large carbon footprint [...]<sup>4</sup>.*

---

<sup>4</sup>Schwartz et al. (2020): Green AI

## “Rød” KI vs “grønn” KI

*Green AI treats **efficiency** as a primary evaluation criterion alongside accuracy. Efficiency entails minimizing data storage, floating point operations, energy consumption, and carbon emissions.<sup>5</sup>*

---

<sup>5</sup>Schwartz et al. (2020): Green AI



## Indikatorer

Hva slags indikatorer har vi for “grønn KI”?

Hva slags indikatorer har vi for “grønn KI”?

- Kategori 1: Beregningsressurser

Hva slags indikatorer har vi for “grønn KI”?

- Kategori 1: Beregningsressurser
- Kategori 2: Klimaavtrykk



## Kategori 1: Beregningsressurser



## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)

## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter

## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter
  - CPU, GPU, TPU

## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter
  - CPU, GPU, TPU
- Modellstørrelse

## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter
  - CPU, GPU, TPU
- Modellstørrelse
  - Antall parametre

## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter
  - CPU, GPU, TPU
- Modellstørrelse
  - Antall parametre
- Tid

## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter
  - CPU, GPU, TPU
- Modellstørrelse
  - Antall parametre
- Tid
  - Trening

## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter
  - CPU, GPU, TPU
- Modellstørrelse
  - Antall parametre
- Tid
  - Trening
  - Inferens/prediksjon



## Kategori 1: Beregningsressurser

- Flyttallsoperasjoner (*floating point operations*)
- Utnyttelse av prosessorenheter
  - CPU, GPU, TPU
- Modellstørrelse
  - Antall parametre
- Tid
  - Trening
  - Inferens/prediksjon
- “Dataeffektivitet” (*data efficiency*)

## Kategori 1: Beregningsressurser

### Fordeler:

- Som regel lette å måle/kalkulere

### Ulemper:

## Kategori 1: Beregningsressurser

### Fordeler:

- Som regel lette å måle/kalkulere
- Muliggjør “benchmarking” og sammenligning

### Ulemper:

## Kategori 1: Beregningsressurser

### Fordeler:

- Som regel lette å måle/kalkulere
- Muliggjør “benchmarking” og sammenligning
- Belyser muligheter for optimalisering

### Ulemper:

## Kategori 1: Beregningsressurser

### Fordeler:

- Som regel lette å måle/kalkulere
- Muliggjør “benchmarking” og sammenligning
- Belyser muligheter for optimalisering

### Ulemper:

- Indirekte mål på klimaavtrykk

## Kategori 1: Beregningsressurser

### Fordeler:

- Som regel lette å måle/kalkulere
- Muliggjør “benchmarking” og sammenligning
- Belyser muligheter for optimalisering

### Ulemper:

- Indirekte mål på klimaavtrykk
- Tar ikke hensyn til variasjon i maskinvare



## Kategori 2: Klimaavtrykk

## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk



## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk
  - Energi (Joules, kWh)

## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk
  - Energi (Joules, kWh)
- Karbonavtrykk (CO<sub>2</sub>eq)

## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk
  - Energi (Joules, kWh)
- Karbonavtrykk (CO<sub>2</sub>eq)
- Karbonintensitet

## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk
  - Energi (Joules, kWh)
- Karbonavtrykk (CO<sub>2</sub>eq)
- Karbonintensitet
  - Utslippsratio relativt til en viss størrelse eller prosess

## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk
  - Energi (Joules, kWh)
- Karbonavtrykk (CO<sub>2</sub>eq)
- Karbonintensitet
  - Utslippsratio relativt til en viss størrelse eller prosess
  - Eksempel: CO<sub>2</sub>eq / kWh

## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk
  - Energi (Joules, kWh)
- Karbonavtrykk (CO<sub>2</sub>eq)
- Karbonintensitet
  - Utslippsratio relativt til en viss størrelse eller prosess
  - Eksempel: CO<sub>2</sub>eq / kWh
- Modellevetid

## Kategori 2: Klimaavtrykk

- Elektrisitetsforbruk
  - Energi (Joules, kWh)
- Karbonavtrykk (CO<sub>2</sub>eq)
- Karbonintensitet
  - Utslippsratio relativt til en viss størrelse eller prosess
  - Eksempel: CO<sub>2</sub>eq / kWh
- Modellevetid
  - Redusere behovet for å måtte trene nye modeller



## Kategori 2: Klimaavtrykk

Fordeler:

- Direkte mål

Ulemper:



## Kategori 2: Klimaavtrykk

Fordeler:

- Direkte mål
- Mer intuitive mål

Ulemper:

## Kategori 2: Klimaavtrykk

### Fordeler:

- Direkte mål
- Mer intuitive mål

### Ulemper:

- Avhengig av hva slags energikilde man bruker

## Kategori 2: Klimaavtrykk

### Fordeler:

- Direkte mål
- Mer intuitive mål

### Ulemper:

- Avhengig av hva slags energikilde man bruker
- Kan være utfordrende å måle nøyaktig

## Ulike indikatorer belyser ulike ting

Model name	Number of parameters	Datacenter PUE	Carbon intensity of grid used	Power consumption	CO <sub>2</sub> eq emissions	CO <sub>2</sub> eq emissions × PUE
GPT-3	175B	1.1	429 gCO <sub>2</sub> eq/kWh	1,287 MWh	502 tonnes	552 tonnes
Gopher	280B	1.08	330 gCO <sub>2</sub> eq/kWh	1,066 MWh	352 tonnes	380 tonnes
OPT	175B	1.09 <sup>2</sup>	231 gCO <sub>2</sub> eq/kWh	324 MWh	70 tonnes	76.3 tonnes <sup>3</sup>
BLOOM	176B	1.2	57 gCO <sub>2</sub> eq/kWh	433 MWh	25 tonnes	30 tonnes

Table 4: Comparison of carbon emissions between BLOOM and similar LLMs. Numbers in *italics* have been inferred based on data provided in the papers describing the models.

Luccioni et al. (2022): Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model

## Hvordan måle?

### Green Algorithms

Towards environmentally sustainable computational science

### Cloud Carbon Footprint

Free and Open Source



Verktøy for måling av indikatorer for grønn KI

# Eksempel: CodeCarbon

```
from codecarbon import EmissionsTracker
tracker = EmissionsTracker()
tracker.start()
# Compute intensive code goes here
tracker.stop()
```

```
from codecarbon import EmissionsTracker

with EmissionsTracker() as tracker:
    # Compute intensive training code goes here
```

```
from codecarbon import track_emissions

@track_emissions
def training_loop():
    # Compute intensive training code goes here
```



CodeCarbon skjermdump



**Hva kan vi bruke målingene til?**

## Hva kan vi bruke målingene til?

- Bevisstgjøring og ansvarliggjøring



## Hva kan vi bruke målingene til?

- Bevisstgjøring og ansvarliggjøring
  - Belyse kostnadene ved bruk av ML-modeller

## Hva kan vi bruke målingene til?

- Bevisstgjøring og ansvarliggjøring
  - Belyse kostnadene ved bruk av ML-modeller
- Sammenligning mellom ulike modeller

## Hva kan vi bruke målingene til?

- Bevisstgjøring og ansvarliggjøring
  - Belyse kostnadene ved bruk av ML-modeller
- Sammenligning mellom ulike modeller
  - Etablere referansemålinger

## Hva kan vi bruke målingene til?

- Bevisstgjøring og ansvarliggjøring
  - Belyse kostnadene ved bruk av ML-modeller
- Sammenligning mellom ulike modeller
  - Etablere referansemålinger
- Balansere ytelse med bærekraft

## Hva kan vi bruke målingene til?

- Bevisstgjøring og ansvarliggjøring
  - Belyse kostnadene ved bruk av ML-modeller
- Sammenligning mellom ulike modeller
  - Etablere referansemålinger
- Balansere ytelse med bærekraft
  - Kombinere indikatorer for ytelse og klimaavtrykk



## Hvordan kombinere indikatorene?

## Eksempel: *Energy-precision ratio*

*Energy-precision ratio* ( $M$ )<sup>6</sup>:

$$M = \text{Error}^{\alpha} \times \text{EPI}$$

- *Error*: Feilrate

---

<sup>6</sup>Lenherr et al. (2021): New universal sustainability metrics to assess edge intelligence

## Eksempel: *Energy-precision ratio*

*Energy-precision ratio (M)*<sup>6</sup>:

$$M = \text{Error}^{\alpha} \times \text{EPI}$$

- *Error*: Feilrate
- *EPI*: Energy consumption per data item classified

---

<sup>6</sup>Lenherr et al. (2021): New universal sustainability metrics to assess edge intelligence



## Eksempel: *Energy-precision ratio*

*Energy-precision ratio* ( $M$ )<sup>6</sup>:

$$M = \text{Error}^{\alpha} \times \text{EPI}$$

- *Error*: Feilrate
- *EPI*: Energy consumption per data item classified
- $\alpha$ : Koeffisient for å vektlegge ytelse foran effektivitet

---

<sup>6</sup>Lenherr et al. (2021): New universal sustainability metrics to assess edge intelligence

## Eksempel: *Green AI Quotient*



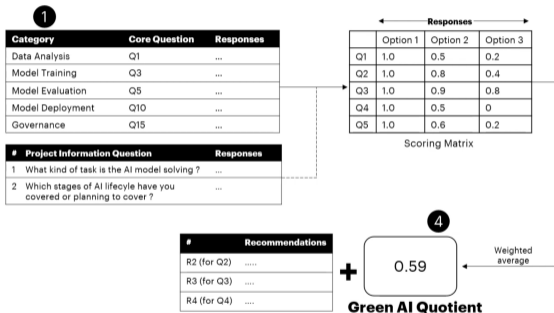
Green AI Quotient: Assessing Greenness of AI-based software and the way forward<sup>7</sup>

---

<sup>7</sup>Sikand et al. (ASE 2023): Green AI Quotient : Assessing Greenness of AI-based software and the way forward

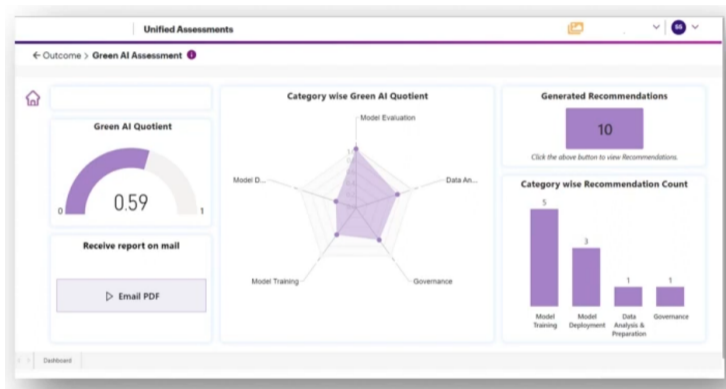
# Eksempel: Green AI Quotient

## GAIQ : Scoring Methodology (Example)



Metode for Green AI Quotient

## Eksempel: Green AI Quotient



Skjermdump fra Green AI Quotient

# Eksempel: d2m – et verktøy for grønn og ansvarlig maskinlæring



d2m

CREATE MODEL

INFERENCE

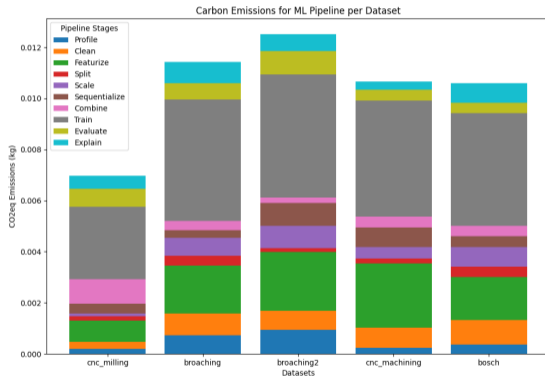
## Overview over existing models

ID	Dataset name	Target variable	R2-score	Training emissions (kg of CO2eq)	Inference emissions (kg of CO2eq)
a31b973b-8c0e-4e59-8362-a355dc97f6d8	cnc_milling	X1_ActualPosition ⓘ	1.00	7.11e-6	2.06e-8
ddc0376f-242b-4bbf-a79f-5719033a8b8e	cnc_milling	X1_ActualPosition ⓘ	1.00	3.75e-6	6.53e-8
315a12f6-0fd1-4d38-9abd-8f65d1eb5f2d	cnc_milling	X1_ActualPosition ⓘ	1.00	4.05e-5	5.23e-8

Create new model

Skjermdump fra SINTEF-verktøyet d2m

# Eksempel: d2m – et verktøy for grønn og ansvarlig maskinlæring



Skjermdump fra SINTEF-verktøyet d2m



## Oppsummering

- Det er viktig å måle klimaavtrykket

## Oppsummering

- Det er viktig å måle klimaavtrykket
- Vi trenger å bruke flere indikatorer



## Oppsummering

- Det er viktig å måle klimaavtrykket
- Vi trenger å bruke flere indikatorer
- Vi har behov for standardiserte indikatorer

## Oppsummering

- Det er viktig å måle klimaavtrykket
- Vi trenger å bruke flere indikatorer
- Vi har behov for standardiserte indikatorer
- Vi må balansere ren ytelse med klimaavtrykket

## Oppsummering

- Det er viktig å måle klimaavtrykket
- Vi trenger å bruke flere indikatorer
- Vi har behov for standardiserte indikatorer
- Vi må balansere ren ytelse med klimaavtrykket
- Hele livssyklusen til et KI-system må tas i betraktning

## Oppsummering

- Det er viktig å måle klimaavtrykket
- Vi trenger å bruke flere indikatorer
- Vi har behov for standardiserte indikatorer
- Vi må balansere ren ytelse med klimaavtrykket
- Hele livssyklusen til et KI-system må tas i betraktning



[Lenke til foiler](#)

## Oppsummering

- Det er viktig å måle klimaavtrykket
- Vi trenger å bruke flere indikatorer
- Vi har behov for standardiserte indikatorer
- Vi må balansere ren ytelse med klimaavtrykket
- Hele livssyklusen til et KI-system må tas i betraktning



Lenke til foiler

- [erik.johannes.husom@sintef.no](mailto:erik.johannes.husom@sintef.no)



SINTEF

Technology for a  
better society