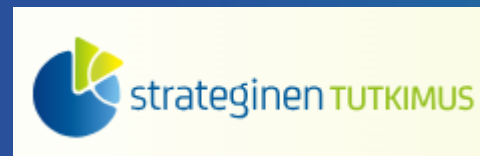


# Luonnollisen kielen käsittely terveydenhuollossa

Ville Vartiainen



# Ville Vartiainen

Dosentti, lääketieteen tohtori

Filosofian maisteri (materiaali- ja nanofysiikka)

Työt tällä hetkellä

- GenAID 2024-2026 (Business Finland)
- GAINS 2025-2028 (-2031) (STN)

Rahoitus

- Suomen lääketieteen säätiö
- Tampereen tuberkuloosisäätiö

Sidonnaisuudet

- Luento- ja konsultointipalkkioita Orionilta

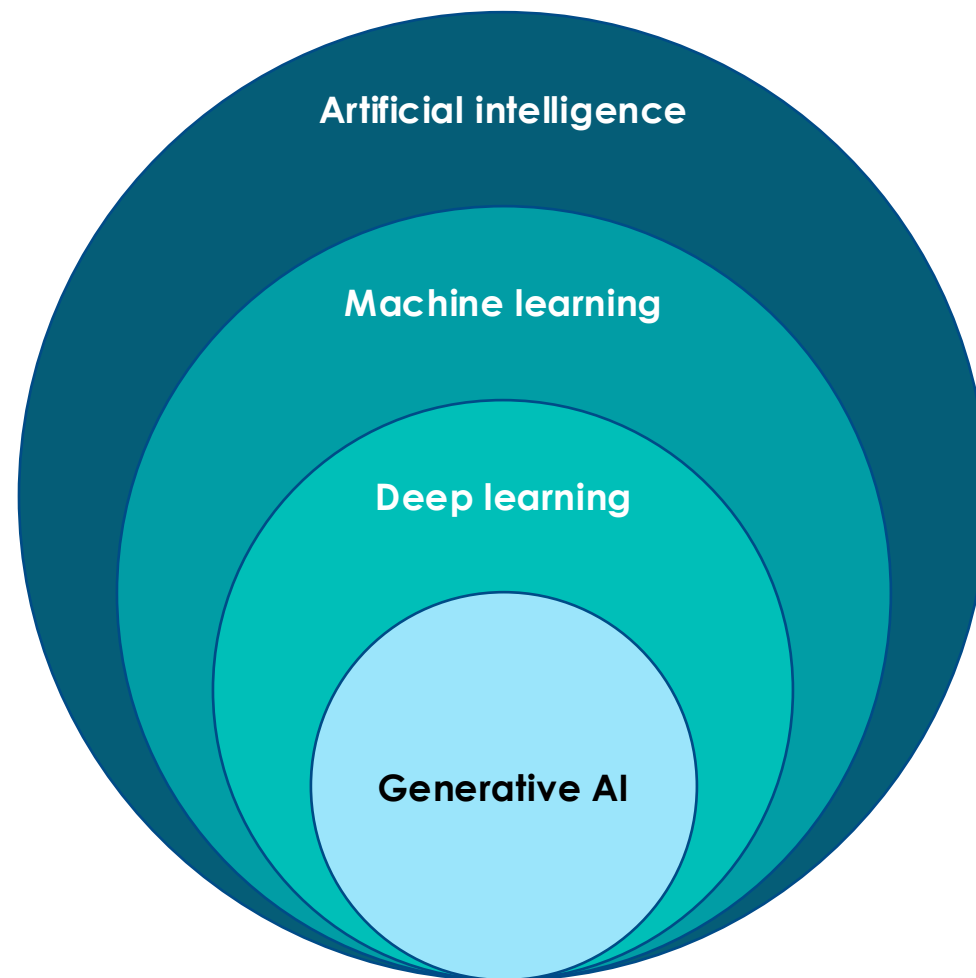


HUS & HELSINGIN YLIOPISTO YHTEISTYÖSSÄ  
HUS & HELSINGFORS UNIVERSITET I SAMARBETE  
A COLLABORATION BETWEEN HUS & UNIVERSITY OF HELSINKI

# SISÄLMYS

- **Neuroverkkojen sielunelämää**
- **Terveydenhuollon konteksti**
- **Esimerkkejä**

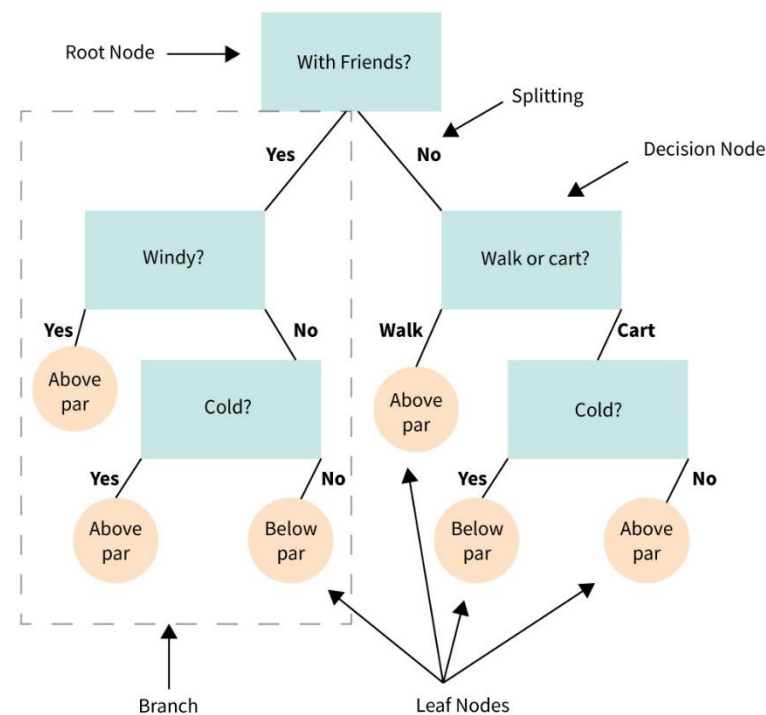
# Määritelmiä



# KONEOPPIMINEN VS KONVENTIONAALINEN STATISTIIKKA

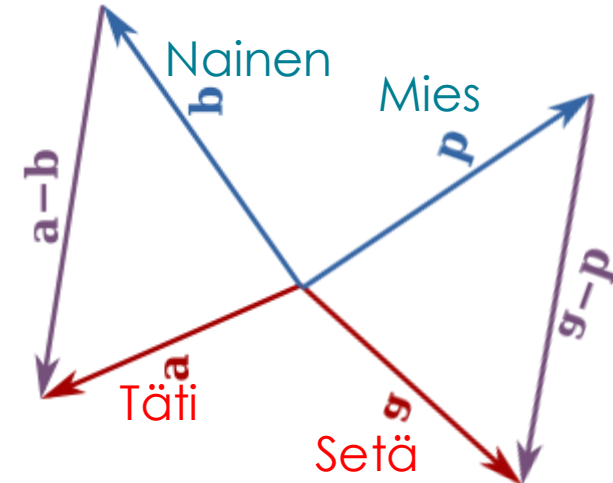
- Koneoppimisella päästään yleensä parempaan ennustavuuteen
- Perinteisellä statistiikalla parametrien tulkinta on helpompaa
- Vrt päätöksentekopuu ja lineaarinen regressio

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i$$



# PIENI SUKELLUS KIELIMALLIEN SIELUNELÄMÄÄN

- Kielimallit muuttavat kielen osat (~sanat) moniulotteisiksi vektoreiksi, jotka kyseisessä vektoriavaruudessa osoittavat merkitykseen
- Vektoreilla voidaan tehdä normaaleja vektoriopin laskutoimituksia
- [But what is a neural network? | Chapter 1, Deep learning \(youtube.com\)](#)
- Esim.
  - $\text{Täti} - \text{Nainen} \approx \text{Setä} - \text{Mies}$
  - $\rightarrow \text{Täti} \approx \text{Setä} - \text{Mies} + \text{Nainen}$
  - $\text{Hitler} - \text{Saksa} + \text{Italia}$   
 $\approx \text{Mussolini}$
  - $\text{Sushi} - \text{Japani} + \text{Saksa}$   
 $\approx \text{Bratwurst}$



# VEKTOREIDEN ILOT

- Matemaattiset operaatiot ovat tietokoneelle yksinkertaisia
  - Vektoireiden pistetulo kertoo saman suuntaisuudesta
  - Voidaan siis katsoa helpolla matemaattisella operaatiolla ovatko merkitykset suurin piirtein samanlaisia
  - Annetaan vektoritietokanta kielimallille kontekstiksi ja pyydetään sitä vastaamaan kontekstin perusteella

# ESIMERKKEJÄ UPOTUSVEKTOREISTA JA NIIDEN VERTAILUSTA - KOSINISIMILARITEETTI

```
get_sentence_embedding("Today is a sunny day")
```

```
tensor([[0.9673]], grad_fn=<MmBackward0>) The weather today is beautiful
```

```
tensor([[0.9029]], grad_fn=<MmBackward0>) It's raining!
```

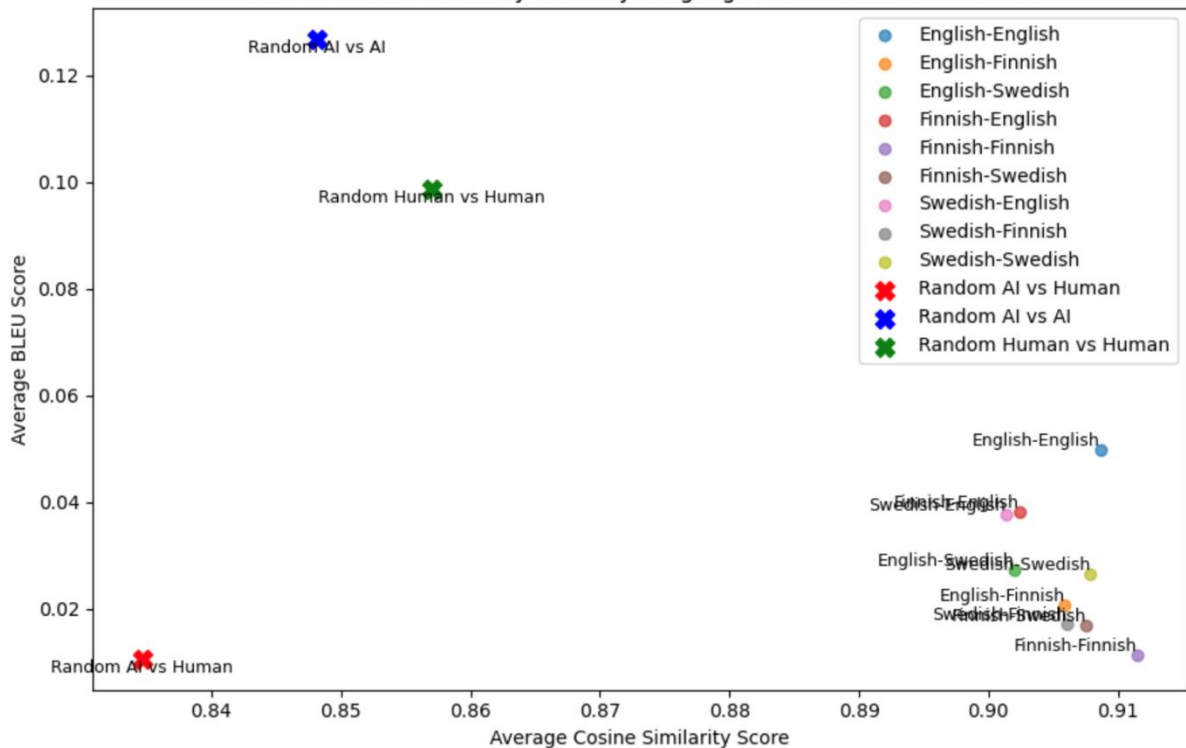
```
tensor([[0.8930]], grad_fn=<MmBackward0>) Dogs are awesome
```

```
util.pytorch_cos_sim( get_word_embedding("The king is angry", "king"), get_word_embedding("The queen is angry", "queen"), )  
tensor([[0.8564]])
```

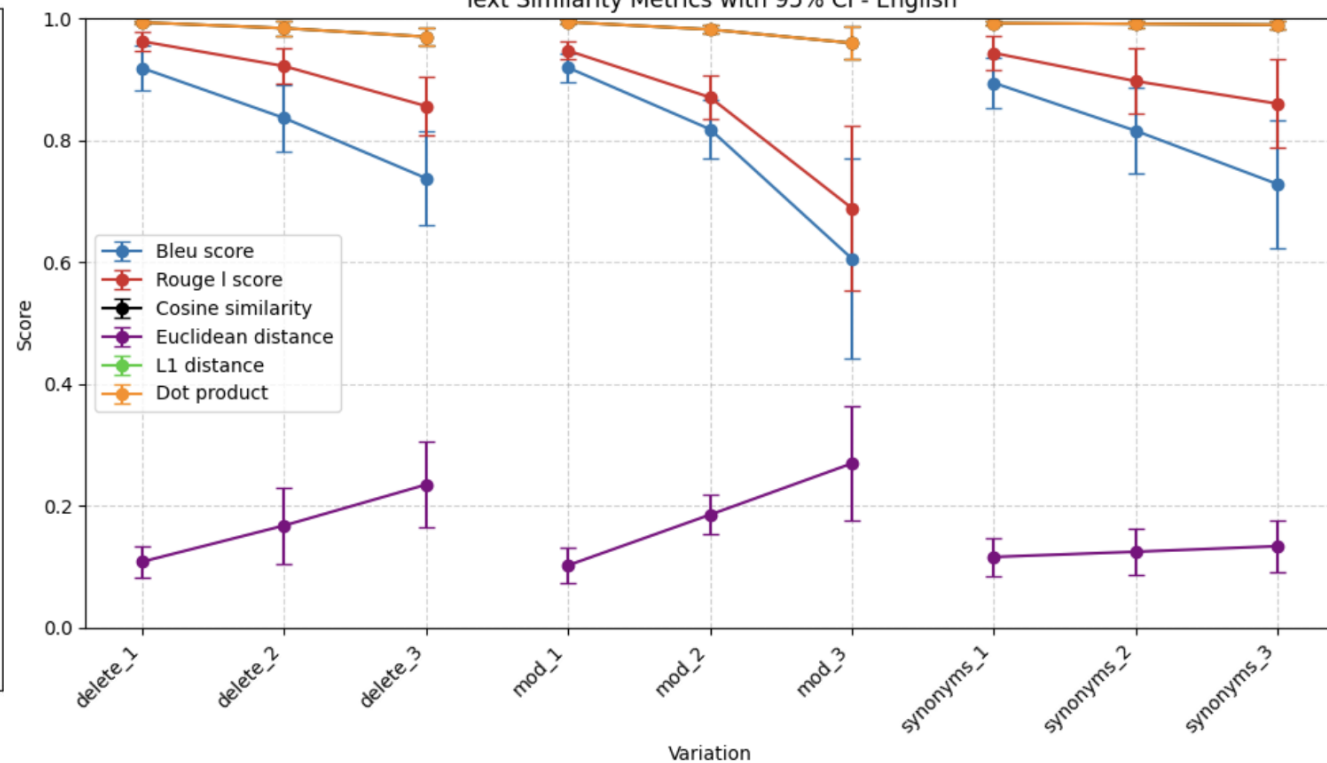
```
util.pytorch_cos_sim( get_word_embedding("The king is happy", "king"), get_word_embedding("The queen is angry", "queen"), )  
tensor([[0.8273]])
```

# MITÄ VÄLIÄ OIKEASTI?

BLEU vs Cosine Similarity Score by Language Pair + Random Baselines



Text Similarity Metrics with 95% CI - English



# ESIMERKKEJÄ

HUS<sup>+</sup>



HUS & HELSINGIN YLIOPISTO YHTEISTYÖSSÄ  
HUS & HELSINGFORS UNIVERSITET I SAMARBETE  
A COLLABORATION BETWEEN HUS & UNIVERSITY OF HELSINKI

Esittäjä

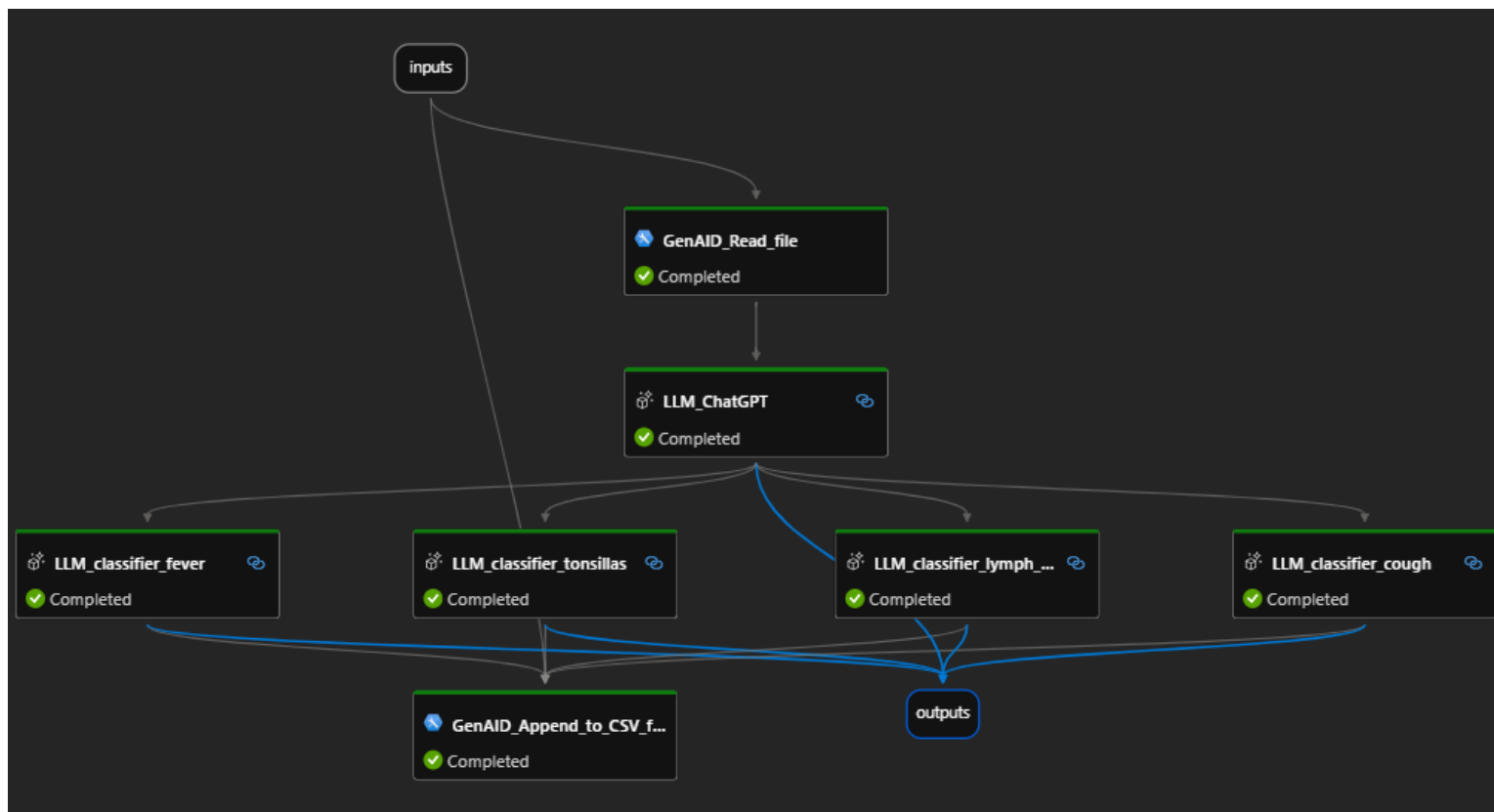
1.4.2026

10

# HOITOSUOSITUSTEN TOTEUTUMINEN



# TIETOJEN RAKENTEISTAMINEN



# Kiitos!



ville.vartiainen@hus.fi



0504306414



Ville Vartiainen



HUS & HELSINGIN YLIOPISTO YHTEISTYÖSSÄ  
HUS & HELSINGFORS UNIVERSITET I SAMARBETE  
A COLLABORATION BETWEEN HUS & UNIVERSITY OF HELSINKI