

# Automaattinen puheentunnistus fysiikan opettajakoulutuksessa

Aleksander Lempinen

# Esittely

- Aleksander Lempinen
  - Aloittanut tietotekniikan laitoksella 2013, koulutusteknologia ja opettajan pedagogiset
  - Kiinnostunut koneoppimisesta ja data-analyysista
  - Tutkimusavustaja opettajankoulutuslaitoksella tammikuu-toukokuu
- Ohjaajat
  - Tommi Kärkkäinen, IT-tiedekunta
  - Daniela Caballero, Chilen Yliopisto
  - Jouni Viiri, OKL

# Tausta

- Opettajankoulutuslaitoksella kerätty paljon dataa
  - Video- tai äänidata
  - Opettajien testitulokset
  - Oppilaiden testitulokset
- Kiinnostusta datan automaattiseen käsittelyyn
  - Puheentunnistus
  - Tekstianalytiikka

# Tutkimus

- Tavoitteena tutkia fysiikan opetuksen laatua automaattisen puheentunnistuksen avulla
- Aikaisemmin tehty:
  - Data kerätty ja analysoitu käsin (Fischer ja muut 2014)
  - Puhedataa muutettu tekstiksi ja analysoitu tekstianalytiikalla ja visualisoitu verkkoanalyysillä (Caballero ja muut 2017)
- Gradun tavoitteet:
  - Kehittää ja verrata eri tapoja esikäsitellä data verkkoanalyysia varten
  - Arvioida eri menetelmien suorituskykyä

# Nykytilanne

- Data
  - Lista fysiikan avainsanoista
  - 25 eri oppitunnin opettajan puheet tekstimuodossa (AaltoASR)
  - Opettajien testit
  - Oppilaiden pre ja post testit
- Työkalut
  - NLTK Snowball stemmer
  - TurkuNLP lemmatizer
  - Finnish Internet Parsebank word2vec
  - NetworkX (verkkoanalyysi)

# Yhteenveto

- Haastava aihe, haastava data
- Vähän tutkittu
- Potentiaalia tutkia lisää (esim väitöskirjaan asti)
- Kiitokset ohjaajille ja OKL:lle mahdollisuudesta tutkia tätä

# Lähteet

- Fischer, Hans E, Peter Labudde, Knut Neumann, and Jouni Viiri. 2014. Quality of instruction in physics: Comparing Finland, Switzerland and Germany. Waxmann Verlag
- Caballero, Daniela, Roberto Araya, Hanna Kronholm, Jouni Viiri, André Mansikkaniemi, Sami Lehesvuori, Tuomas Virtanen, and Mikko Kurimo. 2017. "ASR in classroom today: automatic visualization of conceptual network in science classrooms". In European Conference on Technology Enhanced Learning, 541–544. Springer. doi:10.1007/978-3-319-66610-5\_58.