

Sequence-to-sequence learning (Seq2Seq)

- LSTM/RNN neuroverkko
- Malli palauttaa yksittäisen arvon sijasta esim lauseen
- Input myös sarja/jakso esim. lause

- Chatbot(esim Siri)
- Kääntäjät(Google Translate)
- Text summarizing

- Encode decode malli
- Ensin luodaan sanasto -> jokainen sana/merkki muutetaan vektoreiksi
- Word_2_id ja id_2_word funktiot
- Teksti vektoroidaan käsitellään ja muutetaan lopuksi takaisin tekstiksi

- Tekstin vektorointiin kaksi esi treenattua menetelmää
- Google Word2vec
- Stanfordin yliopiston GloVe(count-pohjainen)

- Outputilla max pituus
- Jos output lyhyempi sarjan perään lisätään padding-dataa