Domain Specific Text Summarization:
Extractive vs Abstractive

Mark Andrew Wallace
MSc in Big Data

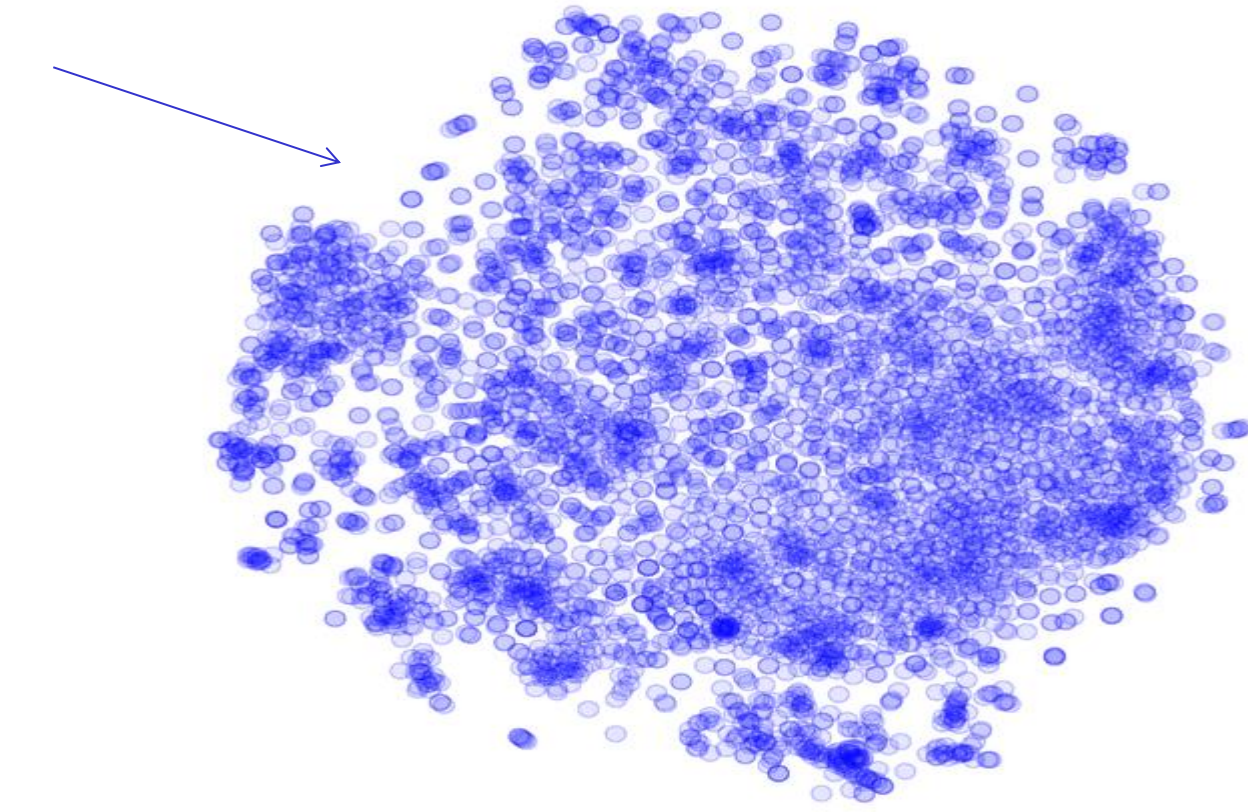**Customisation:**
- Corpus
- POS/NER

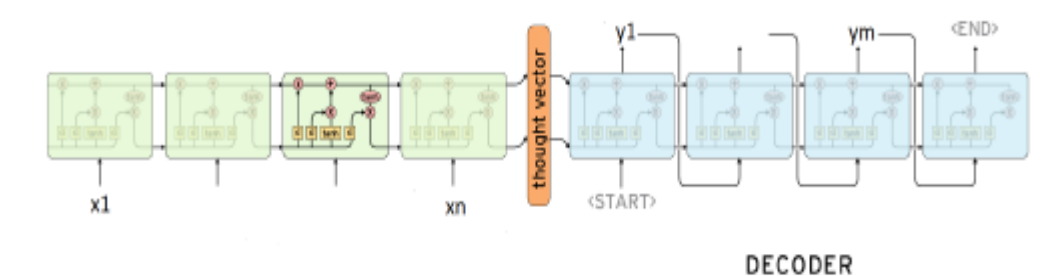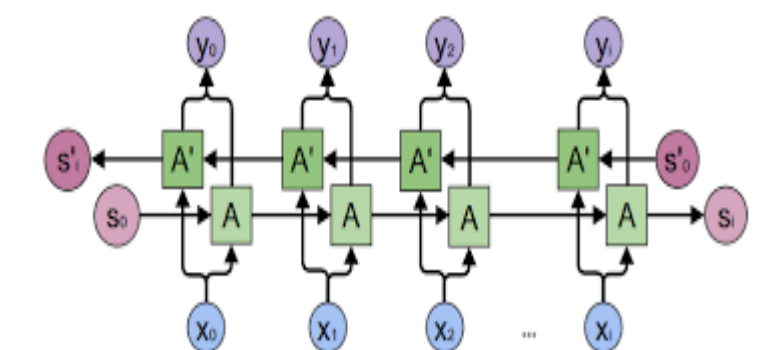**Cleaning:**
- NLTK
- Beautiful Soup

Vectorization:
- Word2Vec
- GloVe

Extractive Summary:

The operator Equinor (formerly known as Statoil) has increased investment levels at Fram in a bid to extend the life of the area.

We assume that a blowdown phase will take place on Fram H Vest, starting in Q4 2018 leading to COP in 2024.

The operator Equinor (formerly known as Statoil) has increased investment levels at Fram in a bid to extend the life of the area. The partners announced in 2017 that they were investing US$122 million (NOK1 billion) in a new processing module at Troll C, where gas for the Fram fields is processed. In May 2018, further investment was announced, which includes three multi-lateral wells being drilled between the Fram F Øst and Fram C Vest fields. Drilling, which will cost US$243 million (NOK1.9 billion) will take place during 2019 and will add 70 mmboe to the field's reserves.

Gas production has been restricted due to capacity constraints on the Troll export pipeline. As such, gas production from areas with high gas to oil ratios (GORs) has been restricted and production is taking place from areas with low GORs. We assume that a blowdown phase will take place on Fram H Vest, starting in Q4 2018 leading to COP in 2024.

These word embeddings are used as inputs into a Bi-RNN Model

ENCODER

DECODER

An extractive summary ranks the sentences in their perceived order of importance and extracts the most important sentences to form a summary

Sentence importance can be calculated by:
- The position of the sentence
- Frequency of the words in the sentence
- Reinforcement learning

Abstractive Summary:

In an attempt to extend the life of the field, Equinor have increased investment in the Fram development.

An abstractive summary learns internal language representations which, in theory, should produce more human-like summaries with original text.

Bi-Directional Recurrent Neural Network Using LSTM Units and Attention