



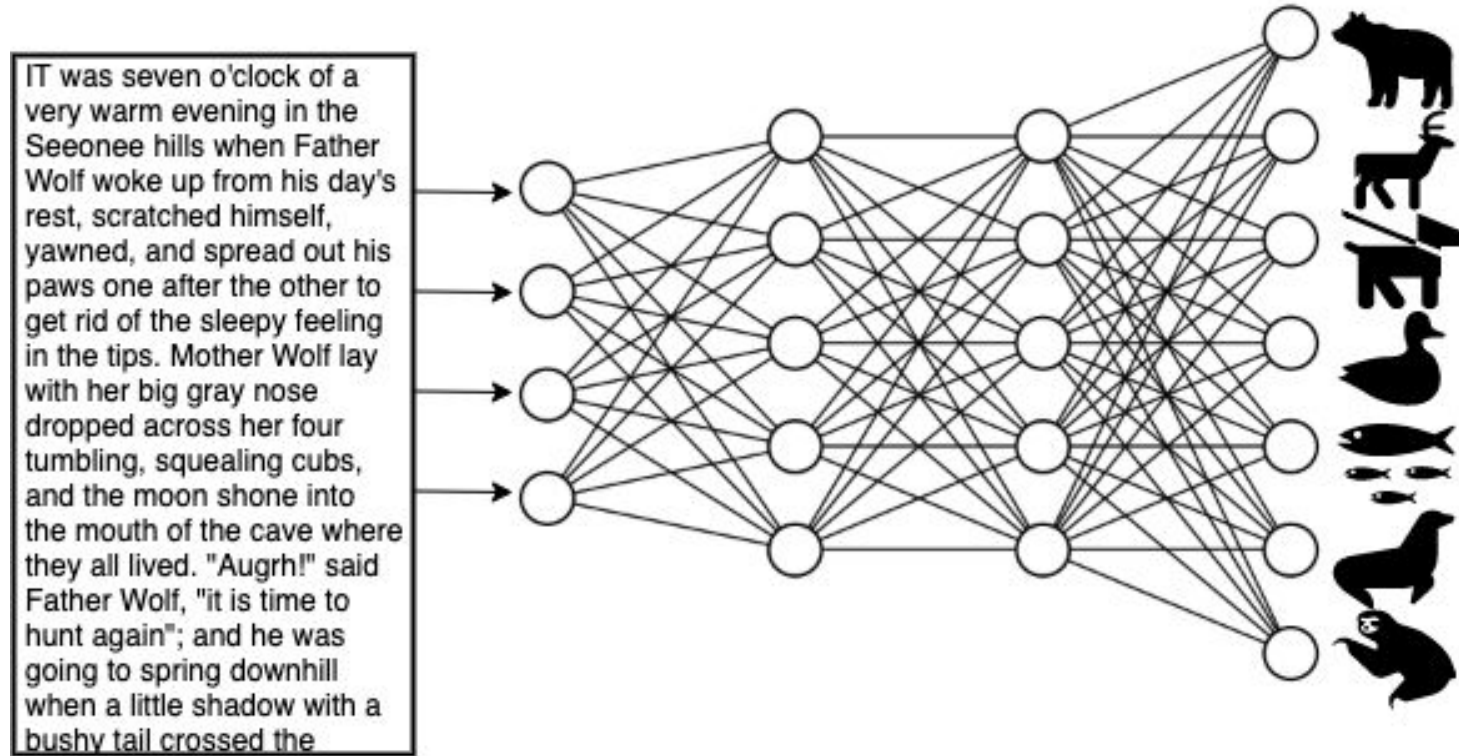
Training for Brilliant Minds

Neural networks for XMTC

Markus Koskela

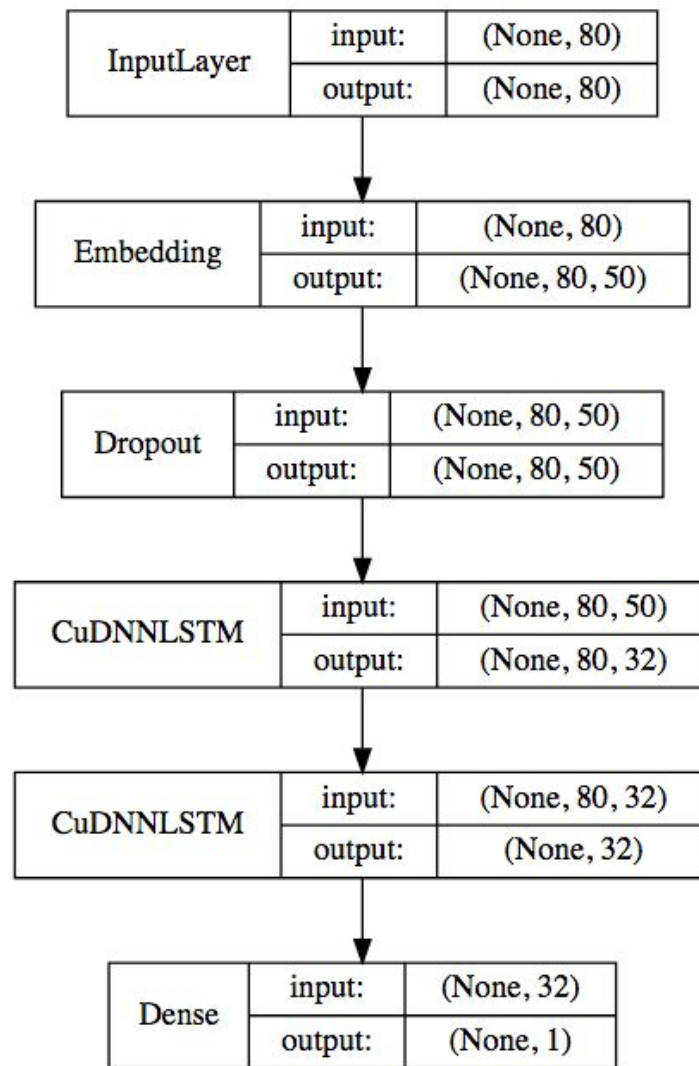


Neural networks for text



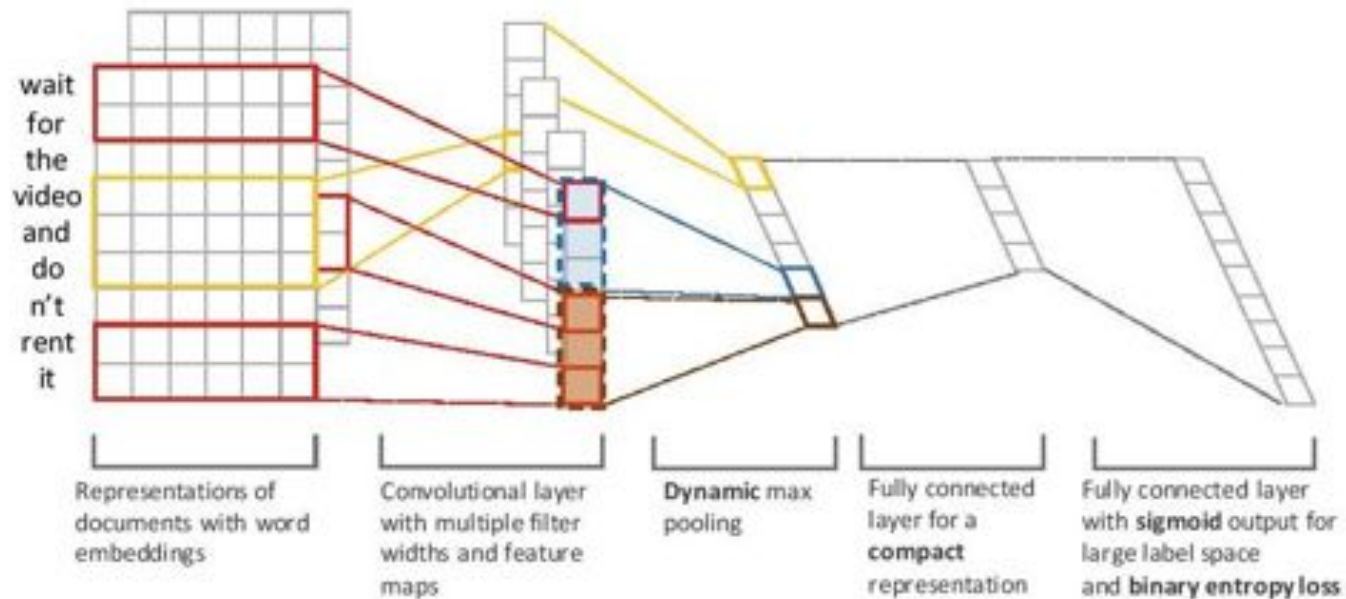
Neural networks for text

- first layer is usually an embedding
- then there are three main approaches (that can also be combined):
 - convolutional layers
 - recurrent layers
 - attention & dense layers
- last layers are often dense



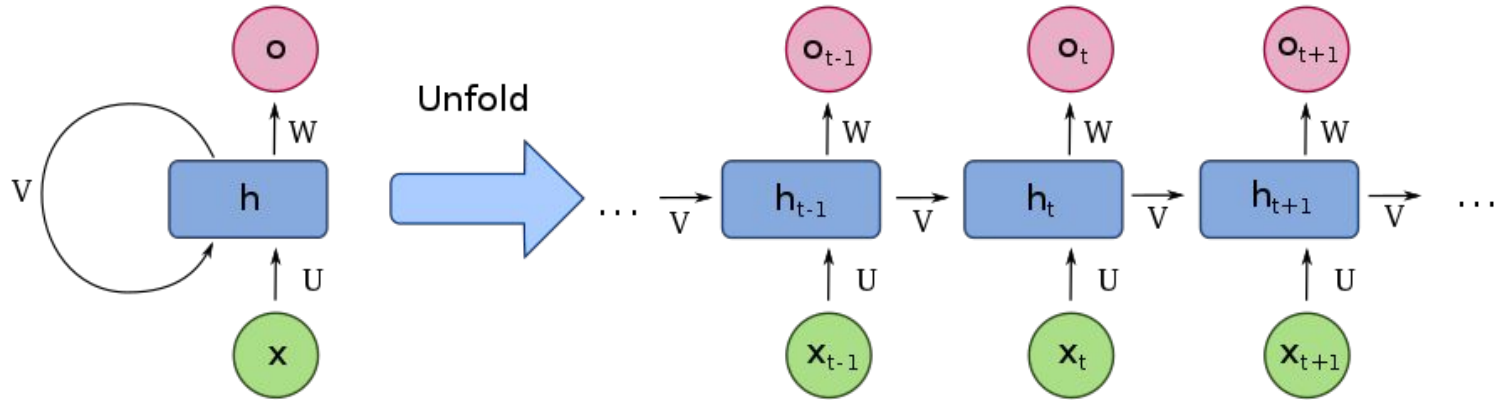
Convolutional neural networks

- Neural networks with convolutional kernels of fixed size
- XML-CNN, Bow-CNN



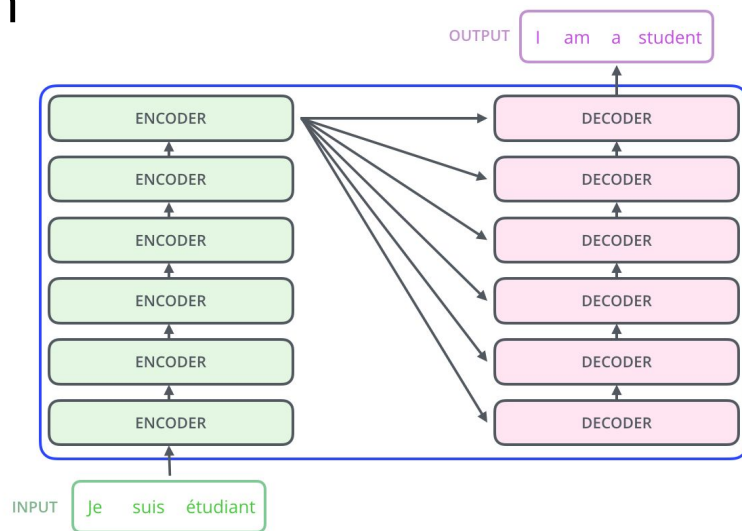
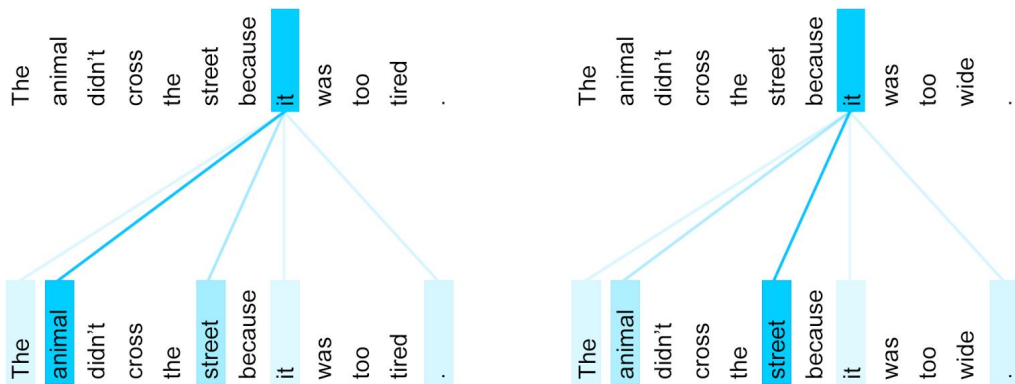
Recurrent neural networks

- Neural networks with memory and recurrent connections, i.e. loops
- AttentionXML



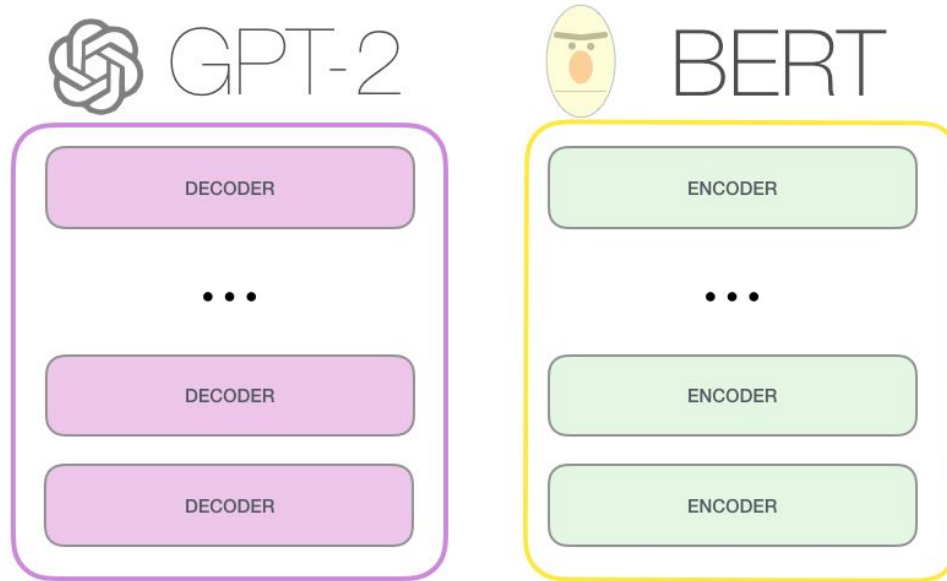
Attention & dense layers

- Transformer models consist of layers that transform a sequence to another sequence
- Self-attention: relating different positions of a sequence in forming its representation



Models based on Transformer

ALL WORK
AND NO PLAY
MAKES JACK
A DULL [???

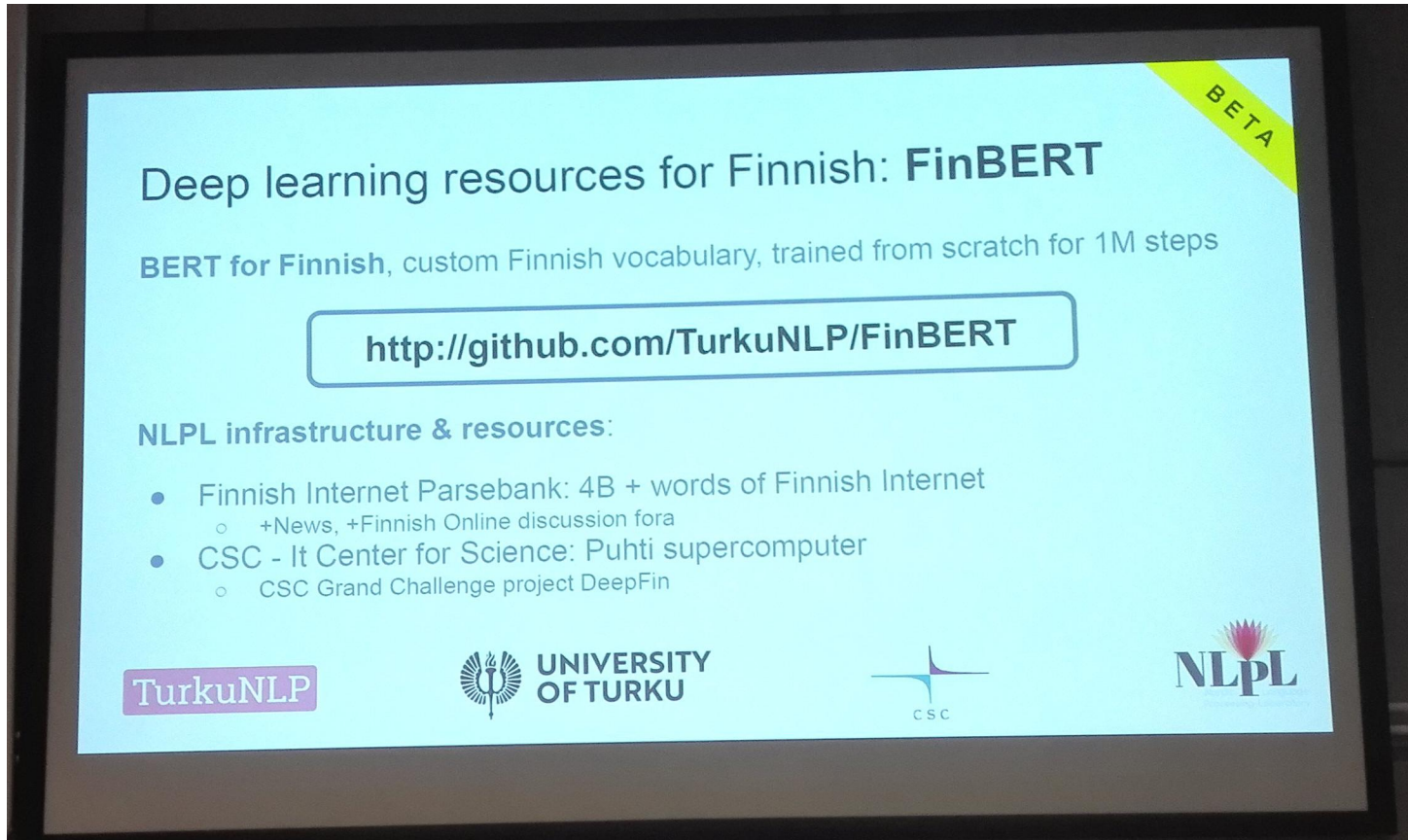


ALL WORK AND
NO [???] MAKES
JACK A DULL
BOY

Image from: <https://jalammar.github.io/illustrated-gpt2/>

- BERT-like: masked LM: *predict the masked words*
- GPT-like: left-to-right LM: *predict the next word*

Models based on Transformer



Deep learning resources for Finnish: **FinBERT**

BERT for Finnish, custom Finnish vocabulary, trained from scratch for 1M steps

<http://github.com/TurkuNLP/FinBERT>

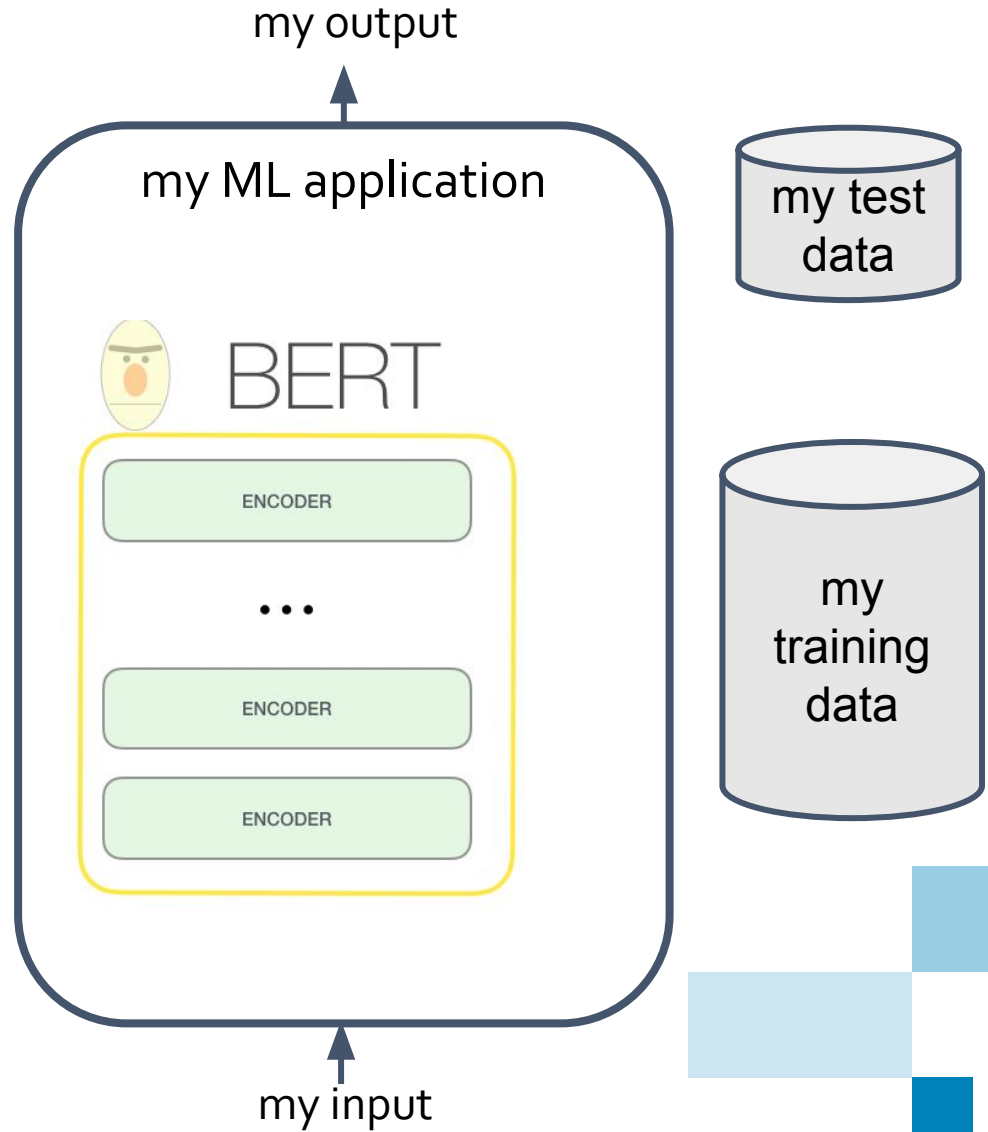
NLPL infrastructure & resources:

- Finnish Internet Parsebank: 4B + words of Finnish Internet
 - +News, +Finnish Online discussion fora
- CSC - It Center for Science: Puhti supercomputer
 - CSC Grand Challenge project DeepFin

TurkuNLP UNIVERSITY OF TURKU CSC NLPL

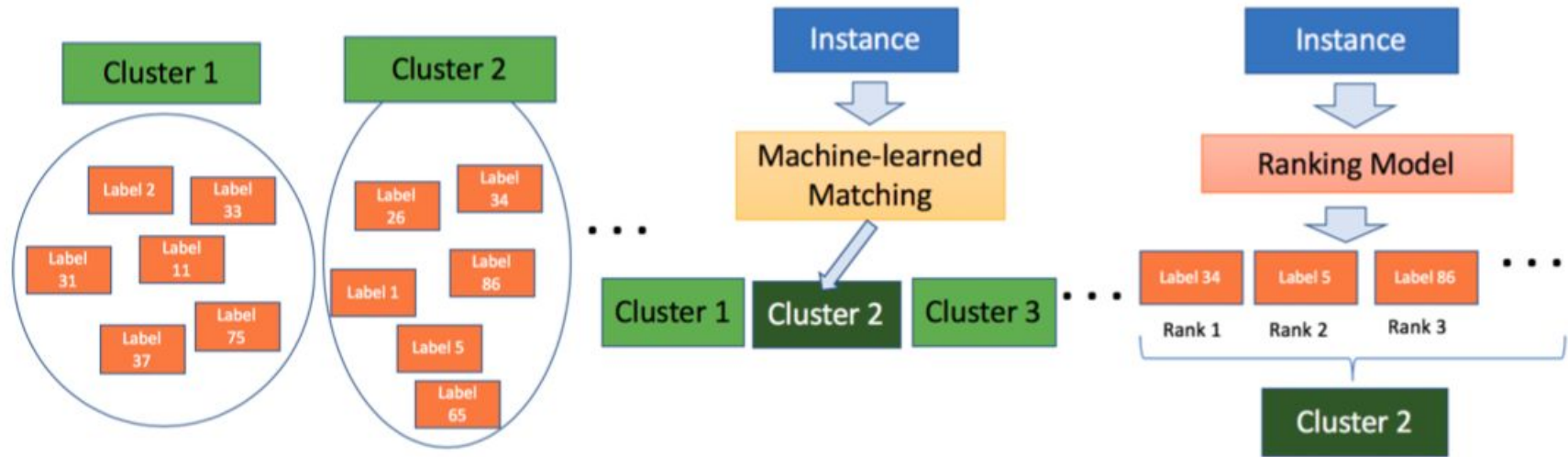
Fine-tuning

- Pre-trained general model can be fine-tuned to a specific task
- Some amount of further training with task-specific data is needed
- Lots of different kinds of applications exist

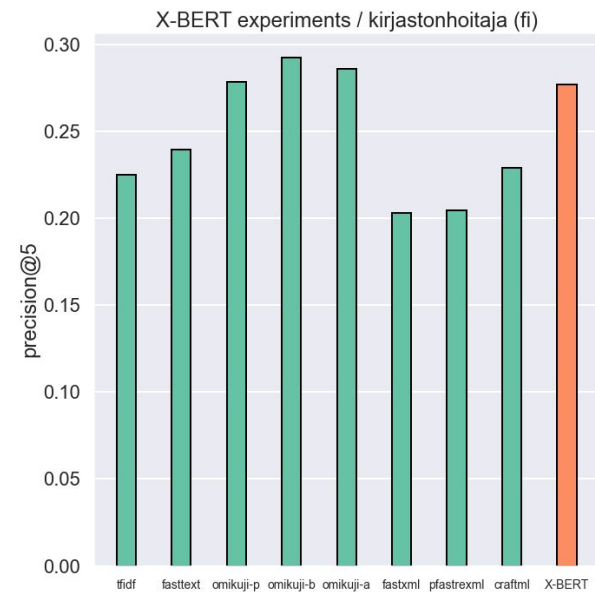
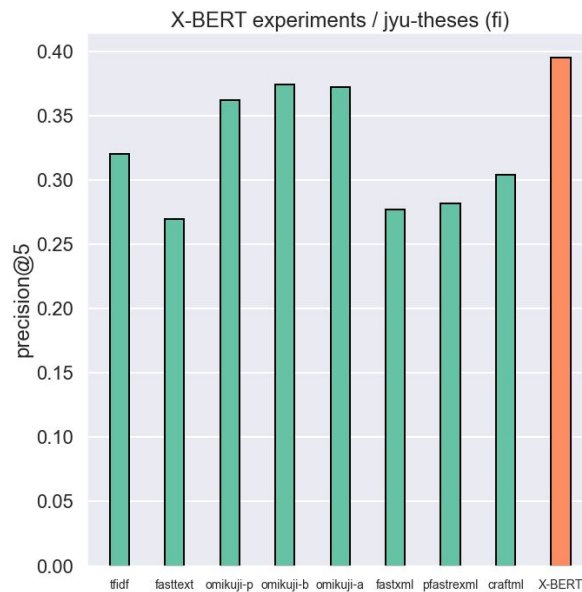
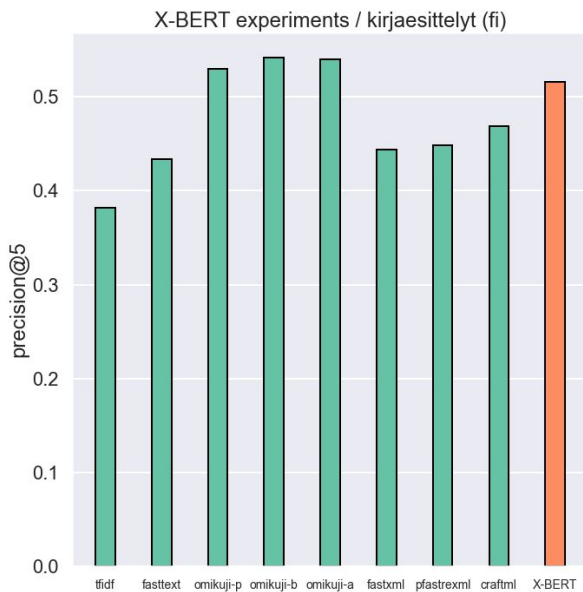


X-Transformer

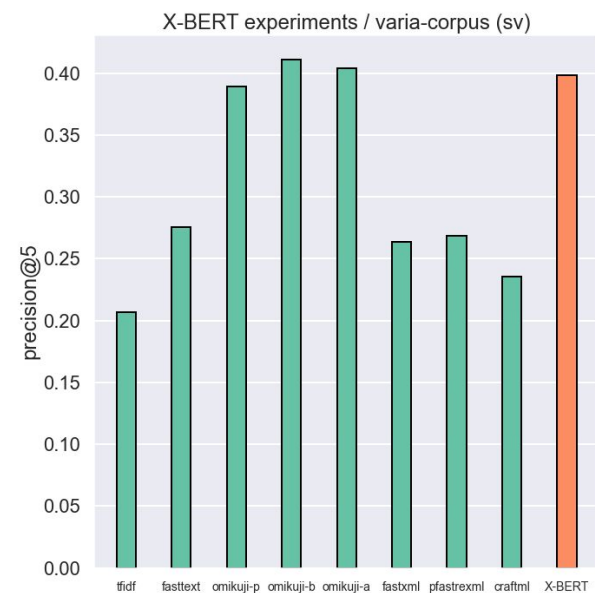
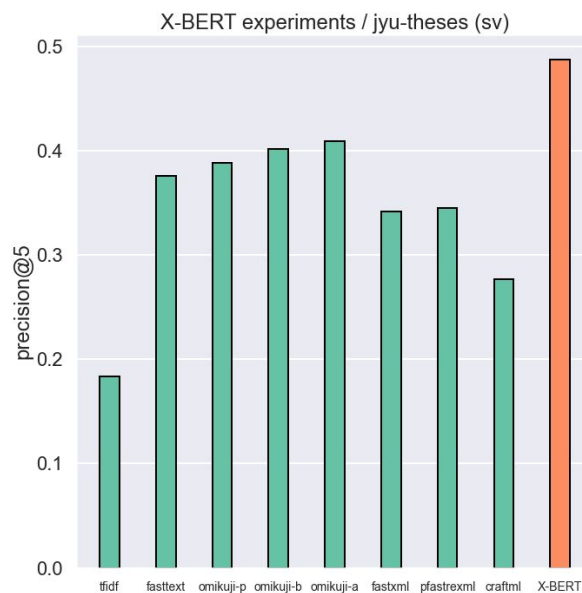
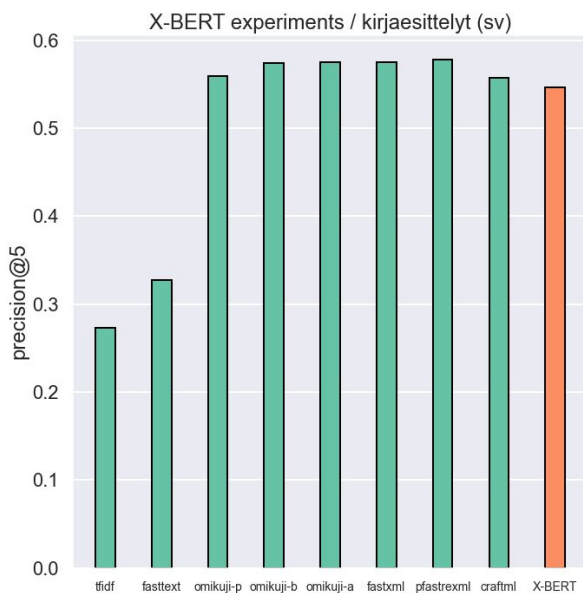
- First “working” application of Transformers to XMTC



Some results: Finnish



Some results: Swedish



Some results: English

