

Introduction

Sentiment Analysis aims to detect if a text is conveying a positive or negative opinion, or sentiment. While Document Level Sentiment Analysis aims to detect whether an entire text is positive or negative, like a restaurant review, Targeted Sentiment Analysis (TSA) is more fine-grained and looks inside each sentence to determine: What, if anything, in this sentence is spoken positively or negatively about?

This is a difficult task that is best approached with a large, manually annotated dataset. Before the NoReC_{fine} dataset was released in 2020, TSA was out of reach for Norwegian, but is now possible. With limited data, though, comes the need for transfer learning and other methods for enhancing the dataset. We have created a state-of-the-art model for TSA in Norwegian based on the NoReC_{fine} dataset, and we have explored mitigating the lack of training data in the target language and domain, through machine-translation from Norwegian into English, and through Cross-domain and Cross-lingual training data.

NoReC_{fine} - a fine-grained dataset for Norwegian Sentiment Analysis

- Each sentence is annotated for polar expressions and their relation to a target and a holder. The relation to the target contains polarity and intensity.
- There are 8634 sentences from 327 reviews in the training set. Two thirds of the targets receive positive sentiment, while one third receive negative sentiment.
- The reviews are concerning a variety of products and productions. This wide domain variety adds to the complexity of the task, but also allows for cross-domain experiments.

The following sentences show examples where one section is annotated as a positive sentiment target (green), while another section is annotated as a negative sentiment target (red).

Maggios myke soulvokal og tidvis fengende låter redder den uinspirerte opptredenen fra å bli en fadese .

The soft soul vocal and partly catchy songs of Maggio save the un-inspired performance from totally flopping .

Denne sesongen av Merlin er mer actionfylt enn de første to .

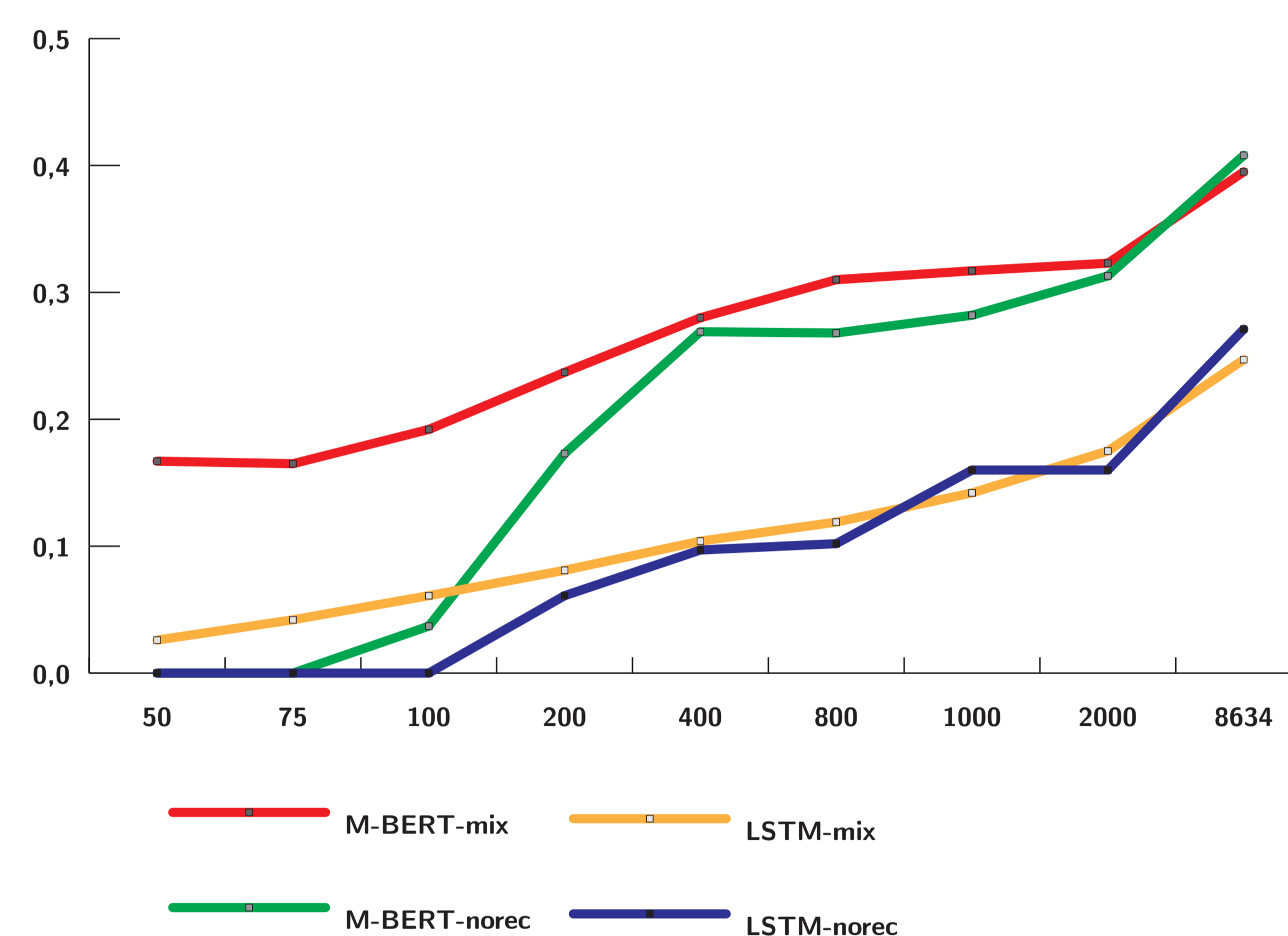
This season of Merlin is more action-packed than the previous two .

Uten **Susanne Sundfør** hadde ikke dette vært all verdens **låt** .

Without Susanne Sundfør , this would not have been much of a song .

Transformers over LSTM

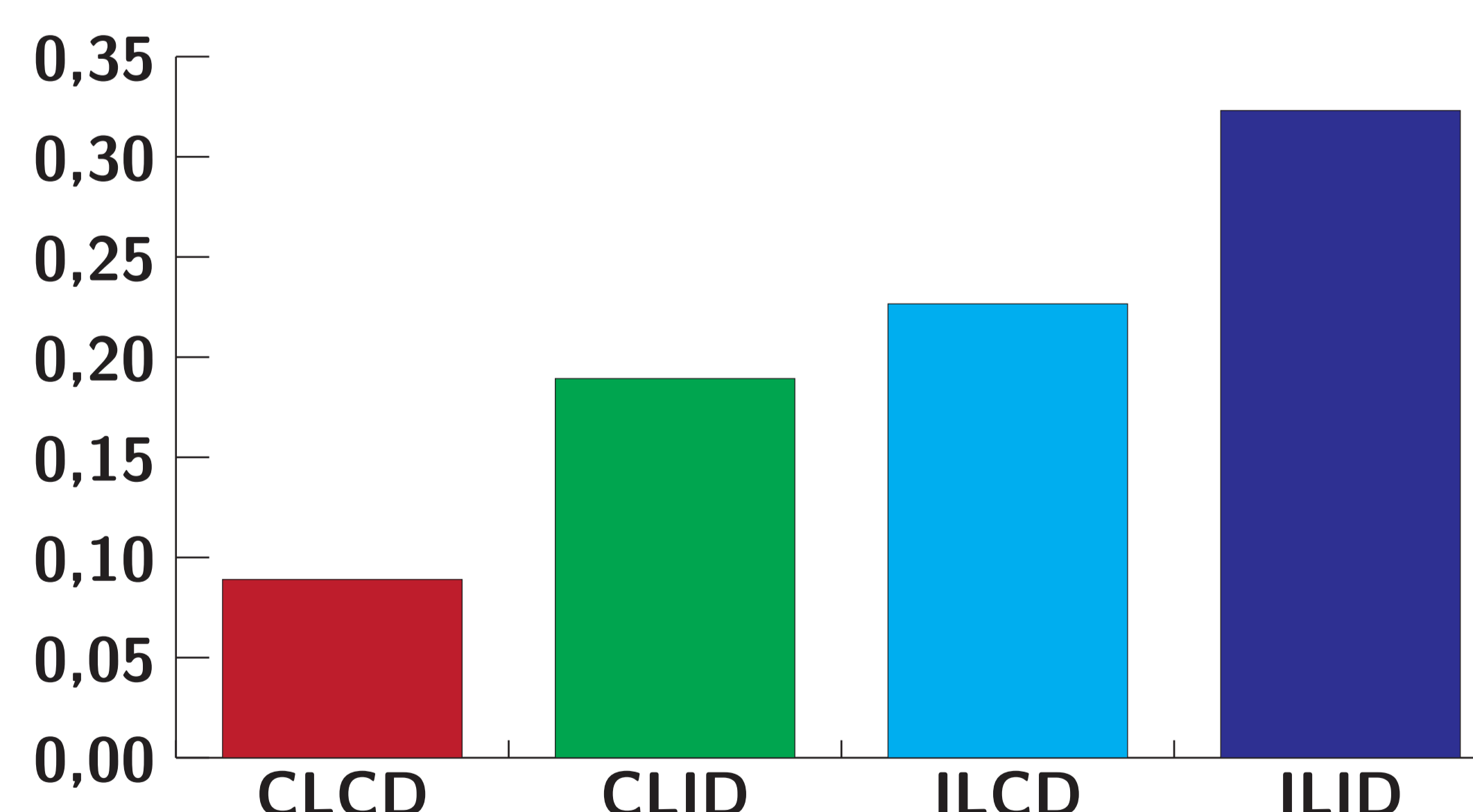
Multilingual Bert (M-BERT) is a pretrained language model providing contextual word representations for more than a hundred languages, including Norwegian and English. Building our experiments on this Transformer-based architecture provided substantial benefits over an LSTM-based architecture with bilingual word embeddings. The Transformer-based solution is able to utilize both Norwegian and English data for fine-tuning, and benefits from mixed training data especially when Norwegian training sentences are only a few hundred.



The line graph shows how F-score improves as number of Norwegian training sentences increase. Architectures M-BERT and LSTM. With "mix" experiments, the Norwegian training sentences are mixed with a full set of English training data. All experiments are evaluated strictly so that the target beginning and end needs to be exactly as annotated, and the polarity needs to be correct.

There is no place like home

Although the M-BERT pretrained language model can utilize TSA data from another domain, and even another language, we find that fine-tuning on training data from the same language and domain (In-Language In-Domain), certainly outperforms data from the same language but different domain (In-Language Cross-Domain) which outperforms data from the same domain but another language (Cross-Lingual In-Domain) which outperforms TSA data from both a different language and different domain (Cross-Lingual Cross-Domain). Most noteworthy in our findings may be that Cross-Lingual In-Domain data almost match In-Language Cross-Domain data.



The bar graph shows how F-score improves as distance between training data and test data is reduced from Cross-Lingual Cross-Domain, to In-Language In-Domain.