

Unsupervised Neural Machine Translation

An Examination of MASS

Mozahid Haque

May 13, 2020

Background

Language to language translation

Background

Language to language translation
Neural Machine Translation (NMT)

Background

Language to language translation
Neural Machine Translation (NMT)

Background

Language to language translation
Neural Machine Translation (NMT)
Google Translate

Background

Language to language translation
Neural Machine Translation (NMT)

Google Translate
Microsoft Translator

Background

Language to language translation
Neural Machine Translation (NMT)

Google Translate
Microsoft Translator
iTranslate

Background

Language to language translation
Neural Machine Translation (NMT)

Google Translate
Microsoft Translator
iTranslate
Babylon

Unsupervised

What about unsupervised NMT?

Unsupervised

What about unsupervised NMT?

Very little...

Unsupervised

What about unsupervised NMT?

Very little...

Recent progress started from 2018—onwards

Proposal

This project proposes an examination of the various techniques employed in state-of-the-art unsupervised NMT, namely Microsoft's Masked Sequence to Sequence learning (MASS).



1.) Similar to BERT

<https://github.com/microsoft/MASS>

1.) Similar to BERT

<https://github.com/microsoft/MASS>

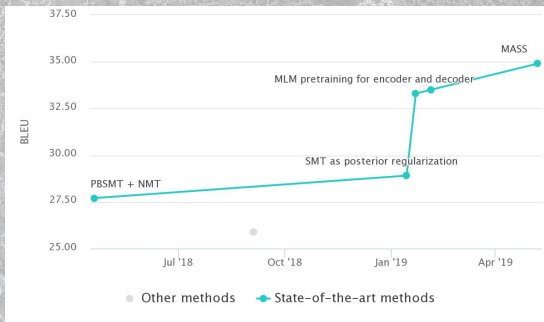
2.) BLEU score higher than XLM

MASS

1.) Similar to BERT

<https://github.com/microsoft/MASS>

2.) BLEU score higher than XLM



Benefits



Learn advanced techniques in NMT

Benefits


Learn advanced techniques in NMT
Fertile research ground

Benefits

Learn advanced techniques in NMT

Fertile research ground

Potential for publication



Thank you!
Questions?

References I



K. Song, X. Tan, T. Qin, J. Lu, T. Liu

MASS: Masked Sequence to Sequence Pre-training for Language Generation, International Conference on Machine Learning, pgs. 5926–5936, 2019.

<https://arxiv.org/pdf/1905.02450.pdf>



P. Linforth

Background Image courtesy of pixabay

<https://pixabay.com/illustrations/neural-network-network-neural-1487112/>



WMT2014 French-English Leaderboard, Papers With Code.

<https://paperswithcode.com/sota/unsupervised-machine-translation-on-wmt2014-1>