UNIVERSITY OF CAMBRIDGE

Language Technology Lab

Normalising Medical Concepts from Personal Health Messages in Social Media

10 December 2016

Nut Limsopatham and Nigel Collier

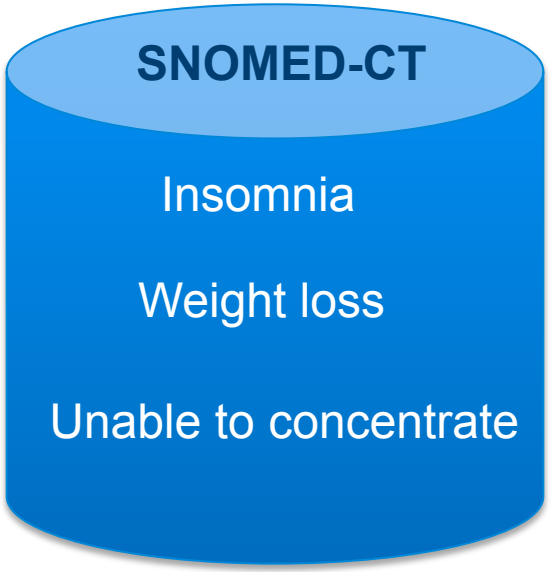
Medical Concept Normalisation

- **Normalise the mentions of medical concepts** in social media text

No way I'm getting any sleep 2nite after taking __DRUG__

__DRUG__ makes me skinny ☺

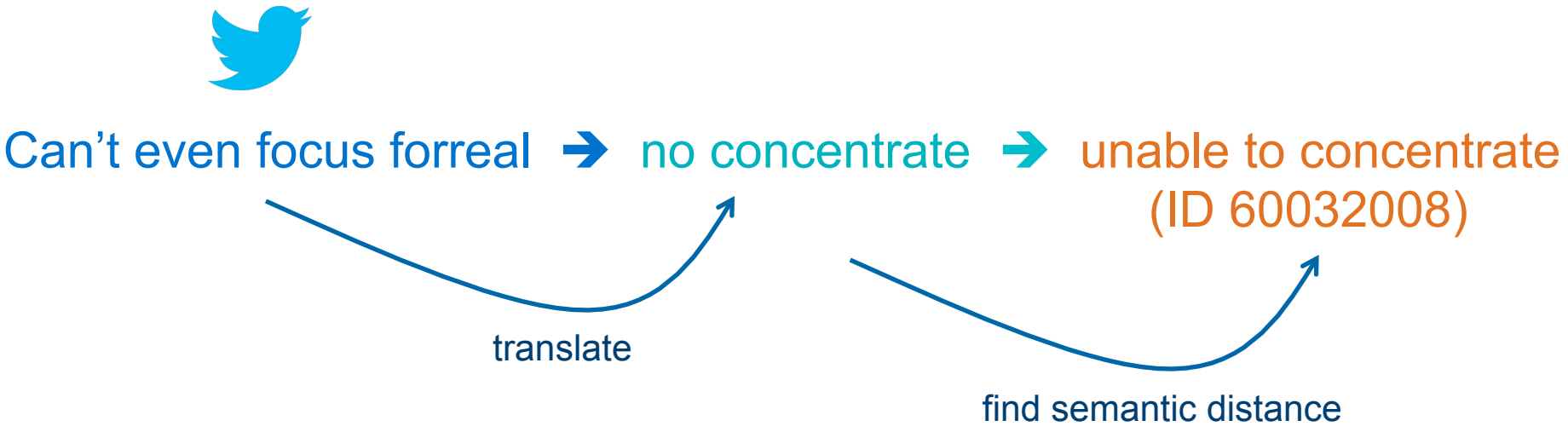
Take __DRUG__ and can't even focus forreal



The datasets are publically available and can be downloaded from https://doi.org/10.5281/zenodo.55013

Adapting Phrase-based MT for Normalisation

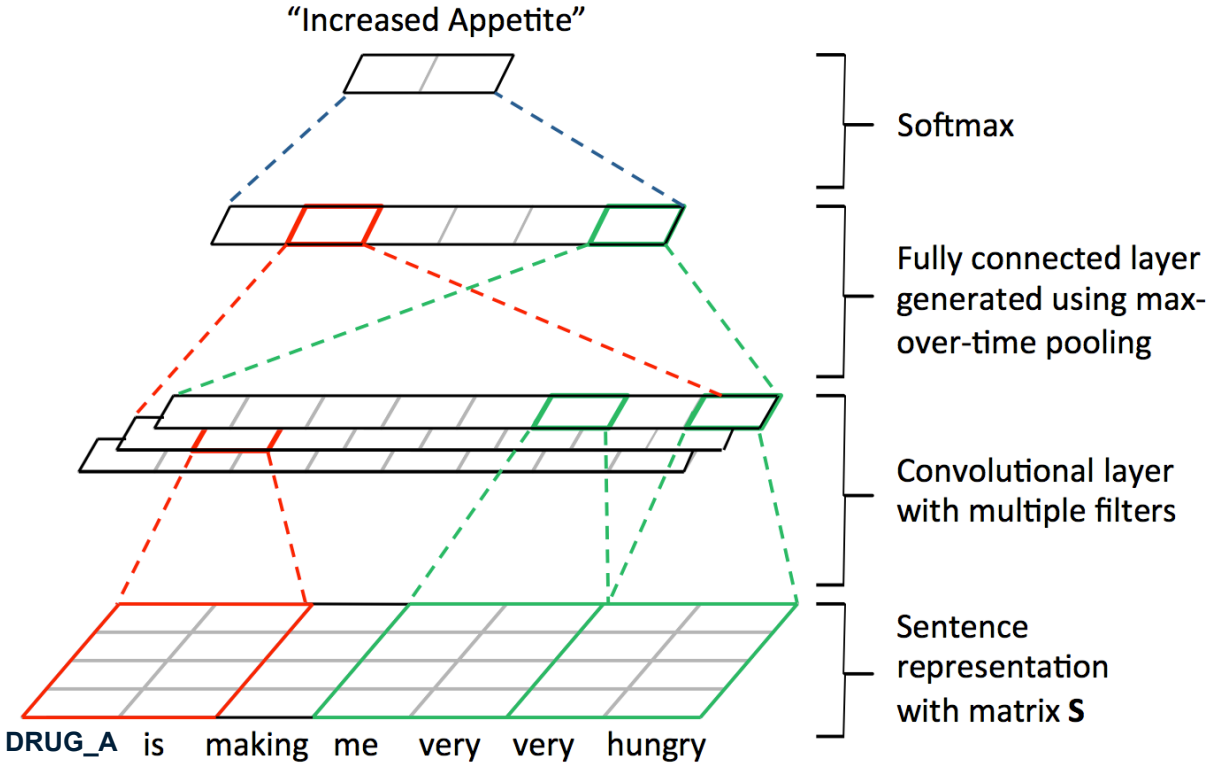
- We use phrase-based MT to ***translate social media text*** to ***formal medical text***, then map the translated text to a medical concept



[4] Limsopatham and Collier (2015), "Adapting phrase-based machine translation to normalise medical terms in social media messages", in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, September, pp. 1675-1680.

A Convolutional Neural Network for Normalisation

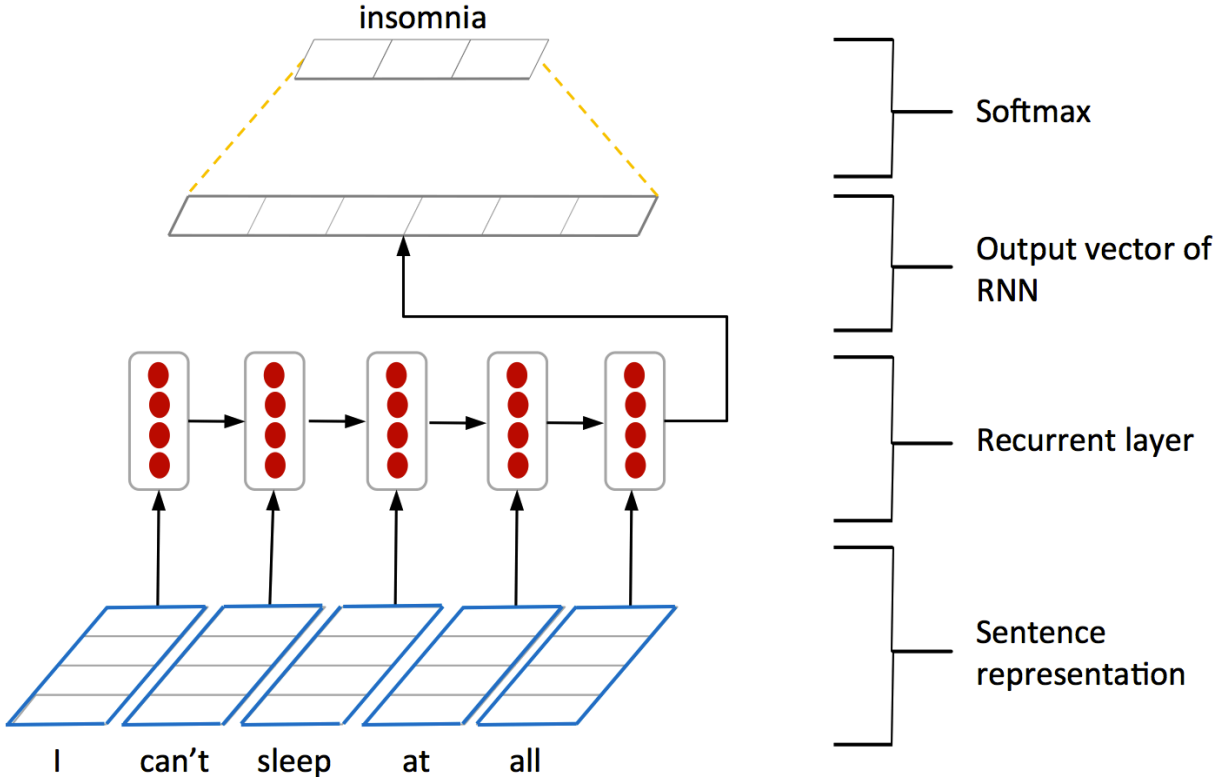
- We use a convolutional neural network to learn the dependence between words in sentences and medical concepts



[4] Limsopatham and Collier (2016), "Normalising Medical Concepts in Social Media Texts by Learning Semantic Representation", in Proceedings of the 54th meeting of the Association for Computational Linguistics (ACL 2016), Berlin, Germany, August 2016.

A RNN Neural Network for Normalisation

- We use a recurrent neural network to learn a representation of a sentence before applying softmax to classify medical concept



[4] Limsopatham and Collier (2016), "Normalising Medical Concepts in Social Media Texts by Learning Semantic Representation", in Proceedings of the 54th meeting of the Association for Computational Linguistics (ACL 2016), Berlin, Germany, August 2016.