Demo: Neural Machine Translation

Kyunghyun Cho kyunghyun.cho@umontreal.ca
Bart van Merrienboer bart.vanmerrienboer@umontreal.ca
Caglar Gulcehre caglar.gulcehre@umontreal.ca
Dzmitry Bahdanau d.bahdanau@jacobs-university.de
Jean Pouget-Abadie jean.pougetabadie@gmail.com
Fethi Bougares fethi.bougares@lium.univ-lemans.fr
Holger Schwenk holger.schwenk@lium.univ-lemans.fr
Yoshua Bengio yoshua.bengio@umontreal.ca
and others..
Machine Translation as Nothing, but a *Supervised Learning Task*

La croissance économique a ralenti ces dernières années.

Encode

Decode

\[ [z_1, z_2, \ldots, z_d] \]

Économic growth has slowed down in recent years.

Properties

▶ *Variable-length* input/output
▶ *Many-to-many* mapping
▶ *End-to-end* trainable

Implications

▶ *Order-sensitive*
▶ *Probabilistic*
▶ *End-to-end Smooth*

---

Kalchbrenner&Blunsom (2013), Kalchbrenner et al. (2014), Sutskever et al. (2014), Cho et al. (2014), Grefenstette et al. (2014), Bahdanau et al. (2014)
Learning to Align and Translate Jointly – (1)

Annotation vector $h_j = [\overrightarrow{h}_j; \overleftarrow{h}_j]$

- $\overrightarrow{h}_j$: Forward RNN
- $\overleftarrow{h}_j$: Backward RNN

For each target word $y_t$,

1. Compute $\alpha_{t,j} = f(h_j, s_{t-1})$
   - $\sum_j \alpha_{t,j} = 1$
   - $f$ is a feedforward neural network

2. Get a context vector $c_t = \sum_j \alpha_{t,j} h_j$

And, train the whole model with SGD and backpropagation!
Learning to Align and Translate Jointly – (2)

The agreement on the European Economic Area was signed in August 1992.


"This will change my future with my family," the man said.

"Cela changera mon avenir avec ma famille," dit l'homme.
Demo available now at...

http://lisa.iro.umontreal.ca/mt-demo