CONDITIONAL GENERATORS OF WORDS DEFINITIONS

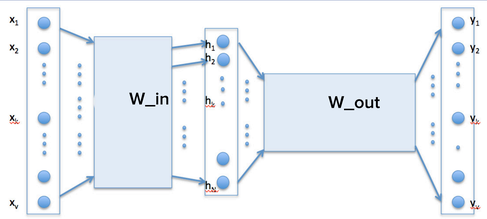
Artyom Gadetsky[1], Ilya Yakubovsky[3], Dmitry Vetrov[1,2]

[1] National Research University Higher School of Economics; [2] Samsung — HSE Laboratory; [3] Joom

0. Introduction

1. By combining word embeddings with RNNs we can train definitions generator
2. Additional way to explore word embeddings
3. Possible way to interpret neural network's predictions

1. Word2Vec Skip-gram



$$p(Y|X) \rightarrow \max_{\{In,Out\}}$$

2. Word ambiguity

$$In(\texttt{'Waterloo'}) = ?$$

$In(\texttt{'Euston'})$ $In(\texttt{'Station'})$ — $In(\texttt{'Napoleon'})$ $In(\texttt{'Austerlitz'})$ $In(\texttt{'Battle'})$ — $In(\texttt{'ABBA'})$ $In(\texttt{'Voulez-vous'})$ $In(\texttt{'Mamma mia'})$

- Suppose we want to answer the question `When was the Battle of Waterloo?`
- Well... It depends on whether the following holds true:

$$In(\texttt{'Waterloo'}) - In(\texttt{'Battle'}) + In(\texttt{'Date'}) \approx In(\texttt{'1815'})$$

3. Adaptive Skip-gram

- Introduce latent variable z_i that indicates meaning of the particular word occurrence x_i
- Search vector representations for meaning rather than word $\mathbf{In(x_i, z_i)}$

`Who won the Battle of Waterloo?`

Probabilities of meanings
0.0000098
0.997716
0.0000309
0.00207717
0.00016605

Closest words:
"sheriffmuir"
"agincourt"
"austerlitz"
"jena-auerstedt"
"malplaquet"
"königgrätz"
"mollwitz"
"albuera"
"toba-fushimi"
"hastenbeck"

`Our train has departed from Waterloo at 1100pm`

Probabilities of meanings
0.948032
0.00427984
0.000470485
0.0422029
0.0050148

Closest words:
"paddington"
"euston"
"victoria"
"liverpool"
"moorgate"
"via"
"london"
"street"
"central"
"bridge"

4. Definition Modeling

$$p(D|w^*, C) = \prod_{t=1}^{T} p(w_t|w_{i<t}, w^*, C)$$

$w^* -$ word being defined
$D = \{w_1, \ldots, w_T\} -$ definition
$C = \{c_1, \ldots, c_m\} -$ context

5. Data

cat[1]

NOUN

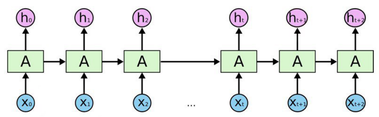
1 A small domesticated carnivorous mammal with soft fur, a short snout, and retractable claws. It is widely kept as a pet or for catching mice, and many breeds have been developed.

> Felis catus, family Felidae (the cat family); it was probably domesticated in ancient Egypt from the local race of wildcat. The cat family also includes the ocelot, serval, margay, lynx, and the big cats

– Example sentences

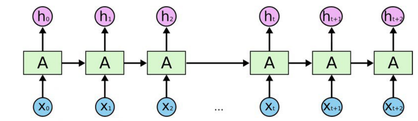
'Various species have been used as models of human asthma, including guinea pigs, mice, rats, cats, and dogs.'

6. Adaptive Skip-gram based



$$v* = disambiguation(word|context)$$
$$x_i = [emb(w_i); v^*]$$

7. Soft-attention based



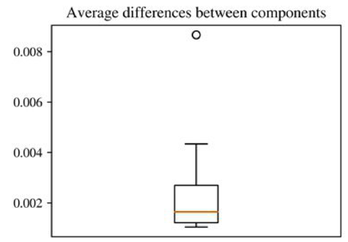
$$v^* = \begin{bmatrix}1.1\\-0.3\\4.2\end{bmatrix} \odot \begin{bmatrix}0.1\\0.9\\0.8\end{bmatrix}$$

$emb(word)$ $mask(context)$

8. Generated definitions

Word	Context	Definition
star	she got star treatment	a person who is very important
star	bright star in the sky	a small circle of a celestial object or planet that is seen in a circle
sentence	sentence in prison	an act of restraining someone or something
sentence	write up the sentence	a piece of text written to be printed
head	the head of a man	the upper part of a human body
head	he will be the head of the office	the chief part of an organization, institution, etc
reprint	they never reprinted the famous treatise	a written or printed version of a book or other publication
rape	the woman was raped on her way home at night	the act of killing
invisible	he pushed the string through an inconspicuous hole	not able to be seen
shake	my faith has been shaken	cause to be unable to think clearly
nickname	the nickname for the u.s. constitution is 'old ironsides '	a name for a person or thing that is not genuine

9. Mask exploration

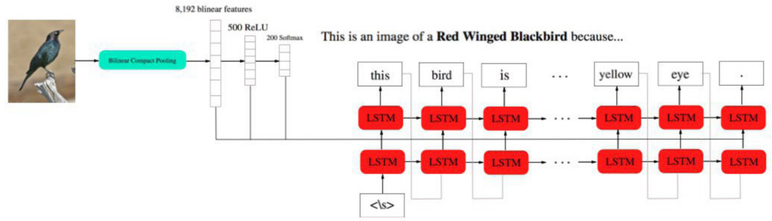


Word	Noun Definition
judge	a person who is responsible for the legal proceedings of a jury
answer	a statement that solves a problem or explains how to solve the problem

Noun Context	Verb Definition
he is due to appear before a judge and jury on monday	act as a judge
he knocked and entered without waiting for an answer	express the opinion of

10. Interpretability

$$(x, y, E(x)) - \text{object, label and explanation}$$



11. Acknowledgments

This work was partly supported by Samsung Research, Samsung Electronics, Sberbank AI Lab and the Russian Science Foundation grant 17-71-20072.

12. References

- Thanapon Noraset, Chen Liang, Larry Birnbaum, and Doug Downey. 2017. Definition modeling: Learning to define word embeddings in natural language. In 31st AAAI Conference on Artificial Intelligence, AAAI 2017, pages 3259–3266. AAAI press.
- Sergey Bartunov, Dmitry Kondrashkin, Anton Osokin, and Dmitry Vetrov. 2016. Breaking sticks and ambiguities with adaptive skip-gram. In Artificial Intelligence and Statistics, pages 130–138.
- Shane Barrat. 2017. InterpNET: Neural Introspection for Interpretable Deep Learning. NIPS