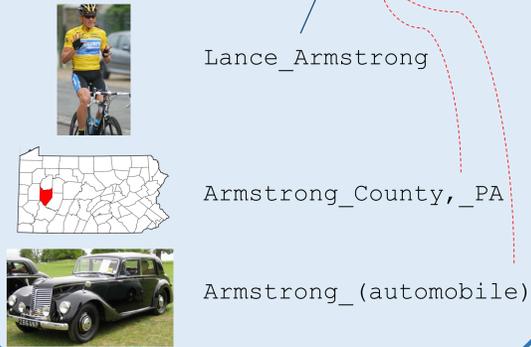


One challenge in entity linking is making use of local and global contextual information to resolve ambiguous mentions. Our models distill multiple granularities of context information into vector representations using convolutional neural networks. Vector representations derived from the source document context and target Wikipedia article are used to compute features that can be incorporated into traditional entity linking systems such as Durrett and Klein (2014).

Entity Linking

Entity linking is a core NLP problem which takes identified spans of text and disambiguates them to the Wikipedia articles they refer to.

...had disqualified Armstrong from ...



What convolutions capture

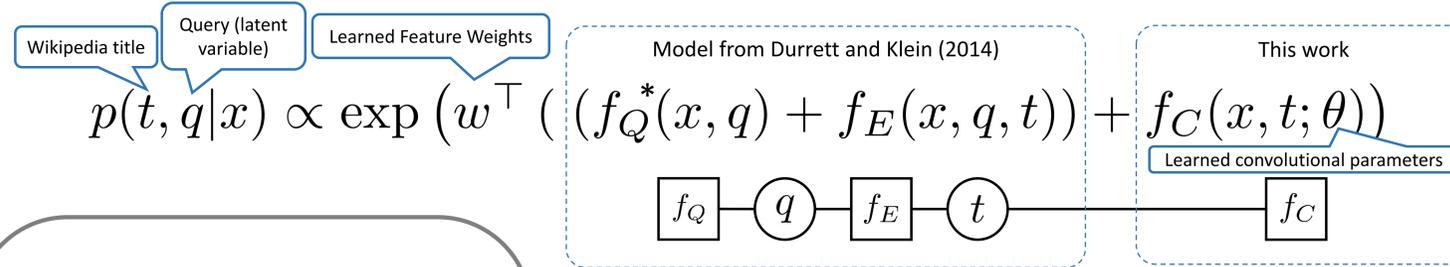
Our model takes a convolution over a rolling 5 word window, which is able to capture semantics similar to that of a bag of 5-grams. A single convolutional filter tends to select phrases which are suggestive of a given topic as shown below. Using 150 different filters allows us to represent documents as mixture of topics.

Example maximal spans from a single convolutional filter

destroying missiles . spy planes
and destroying missiles . spy
by U.N. weapons inspectors .
inspectors are discovering and destroying
an attack using chemical weapons
attack munitions or j-dam weapons
its nuclear weapons and missile

Main contribution

We use convolutional neural networks to extract vector representations from blocks of text. This is similar to a bag-of-words representation in that the convolution acts over n -grams, but learned convolutional filters and low-dimensional vector representations of words give the system additional expressive power. Vectors from the source document context and target Wikipedia article are compared using weighted cosine similarity, giving features that we incorporate into a log-linear model.



Information collected from various granularities of input document

The entire article is fed into a convolutional neural network.

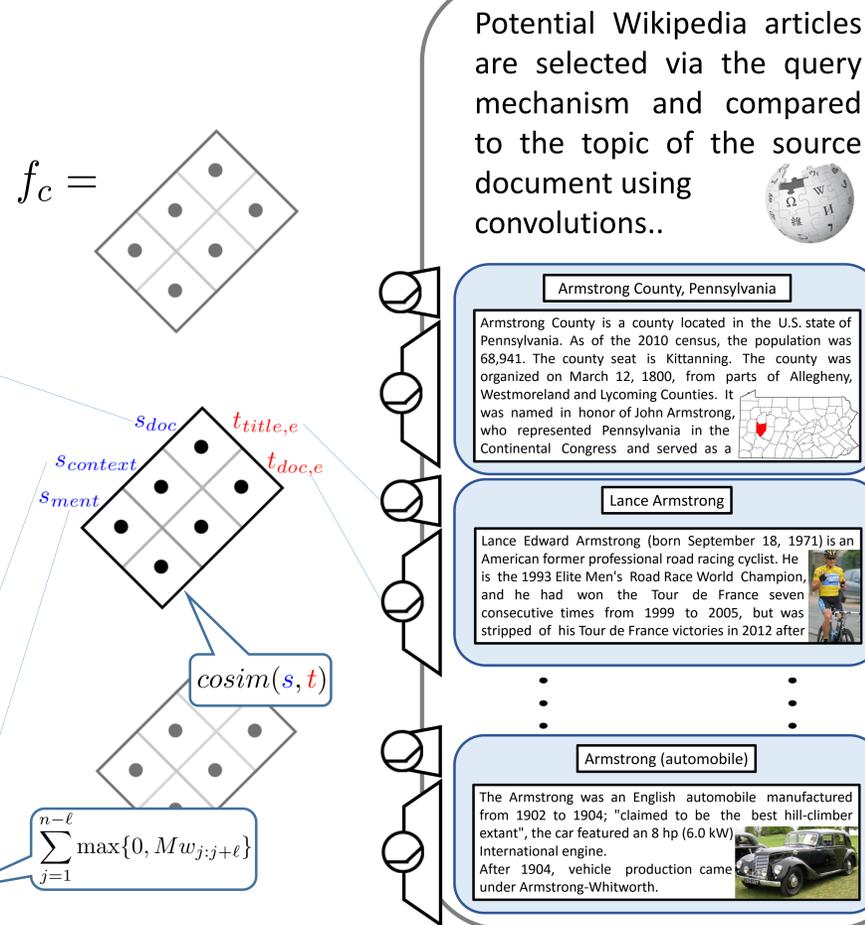
The 1999 Tour de France was a multiple stage bicycle race held from 3 July to 25 July 1999, and the 86th edition of the Tour de France. It has no overall winner—although American cyclist Lance Armstrong originally won the event, the United States Anti-Doping Agency announced in August 2012 that they had disqualified Armstrong from all his results since 1998, including his seven consecutive Tour de France wins from 1999–2005 (which were, originally, the most wins in the event's history); the Union Cycliste Internationale confirmed the result. There were no French stage winners for the first time since the 1926 Tour de France. Additionally, Mario Cipollini won 4 stages in a row, setting the post-World War II record for consecutive stage wins (breaking the record of three, set by Gino Bartali in 1948.)

Immediate context is fed into its own convolution

United States Anti-Doping Agency announced in August 2012 that they had disqualified Armstrong from all his results since 1998

The mention itself is fed into a third convolution

Armstrong



Potential Wikipedia articles are selected via the query mechanism and compared to the topic of the source document using convolutions.

Results

	ACE	CoNLL	WP	Wiki
DK2014	79.6			
AIDA-LIGHT		84.8		
Sparse features	83.6	74.9	81.1	81.5
CNN features	84.5	81.2	87.7	75.7
Full	89.9	85.5	90.7	82.2

Our model outperforms variants that use only sparse features, only dense features, as well as baselines from prior work.

	ACE	CoNLL	WP
$cosim(s_{doc}, t_{doc})$	77.4	79.8	72.9
$cosim(s_{ment}, t_{title})$	80.2	80.9	70.3
All CNN pairs	84.9	86.9	82.0

All convolutions contribute to the performance of the system.

Conclusion

Using multiple granularities in addition to sparse features produces a significant boost in performance on with entity linking systems by efficiently capturing topic information.

* Queries from DK2014, take a surface strings such as "Cycling champion Lance Armstrong" and breaks it down to substrings such as "Lance Armstrong" and "Armstrong" which can match titles of Wikipedia articles.