

Learning Thematic Similarity Metric Using Triplet Networks

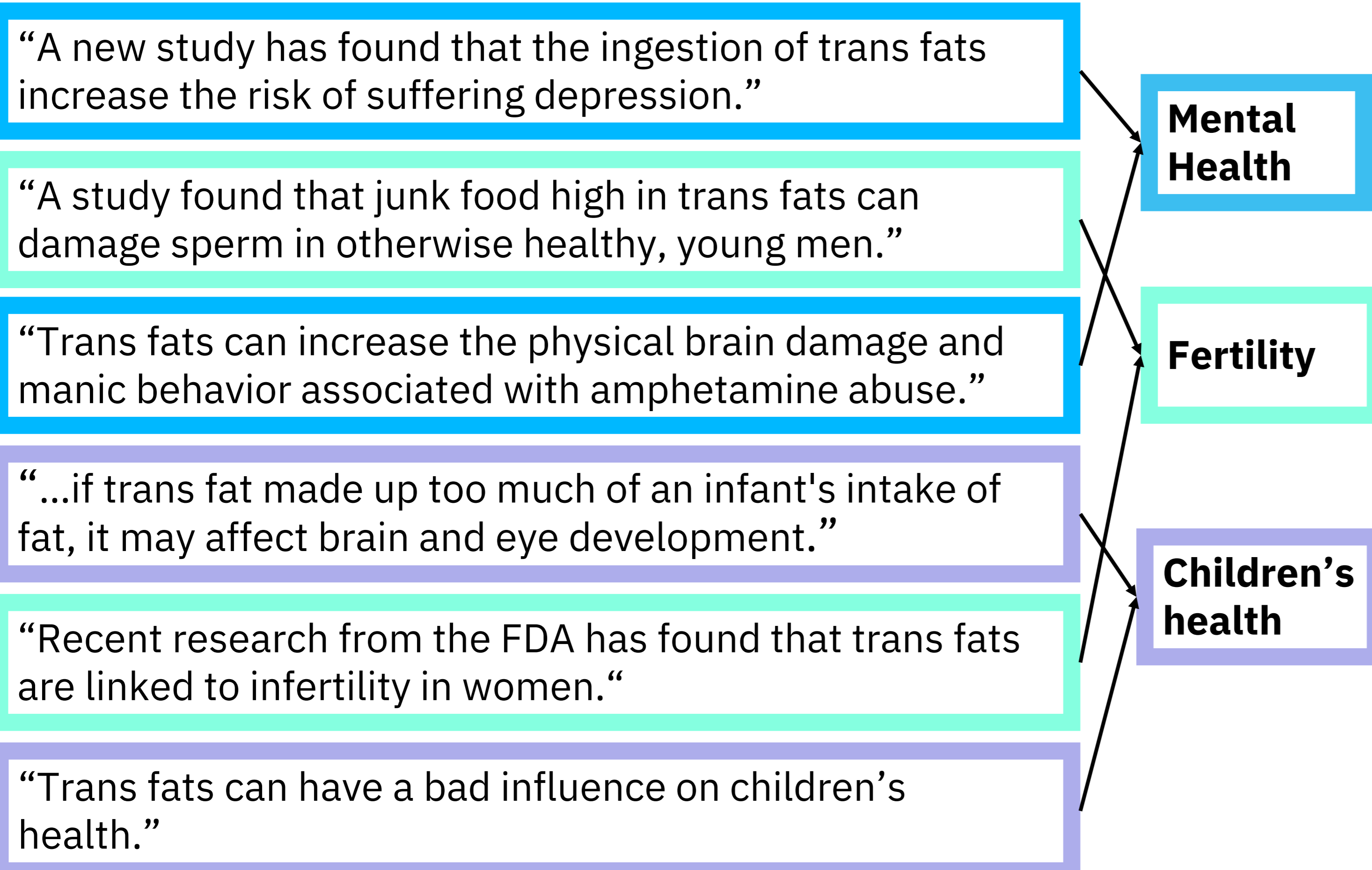
Liat Ein-Dor, Yosi Mass, Alon Halfon, Elad Venezian, Ilya Shnayderman, Ranit Aharonov, Noam Slonim
IBM Research AI

{liate,yosimass,alonal,eladv,ilyashn,ranita,noams}@il.ibm.com

Motivation

- Argumentation mining (Project Debater)
 - Organize arguments into **themes** to generate a compelling argumentative narrative
- Multi-document summarization
 - Organize selected sentences into meaningful sections and paragraphs
- General sentence representations do not capture **thematic** similarity

Example - Trans fats usage in food should be banned



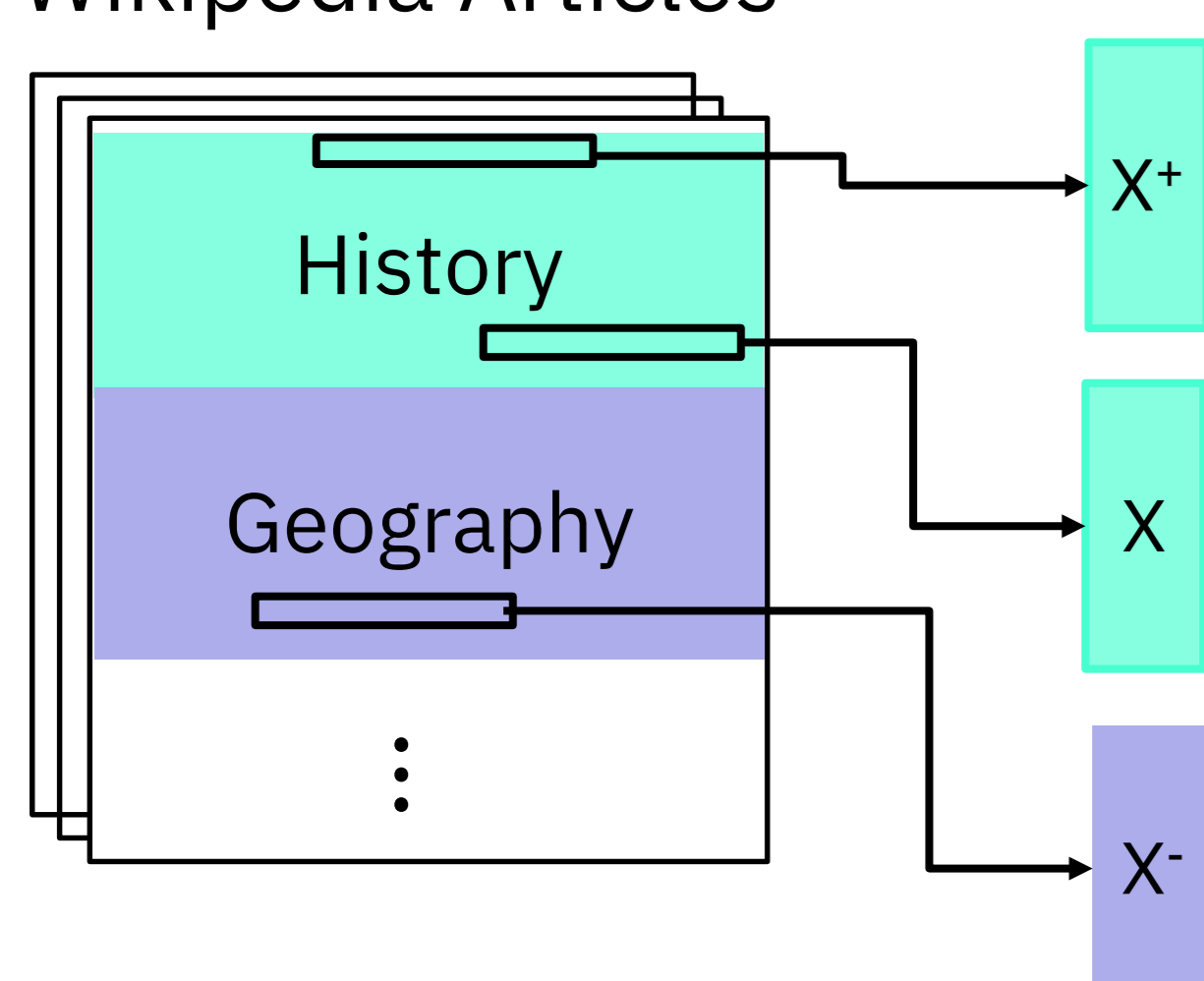
Approach

- Cluster sentences into themes
 - Requires a similarity metric between sentences dedicated to thematic clustering
- Train a Deep Neural Network to learn a thematic similarity through a comparative task.
 - Requires a large set of labeled data
- Create a large weak-supervised labeled data from the partition of Wikipedia articles into sections
- Create a clustering benchmark for evaluation



Generation of Weakly-Labeled Triplets

Wikipedia Articles

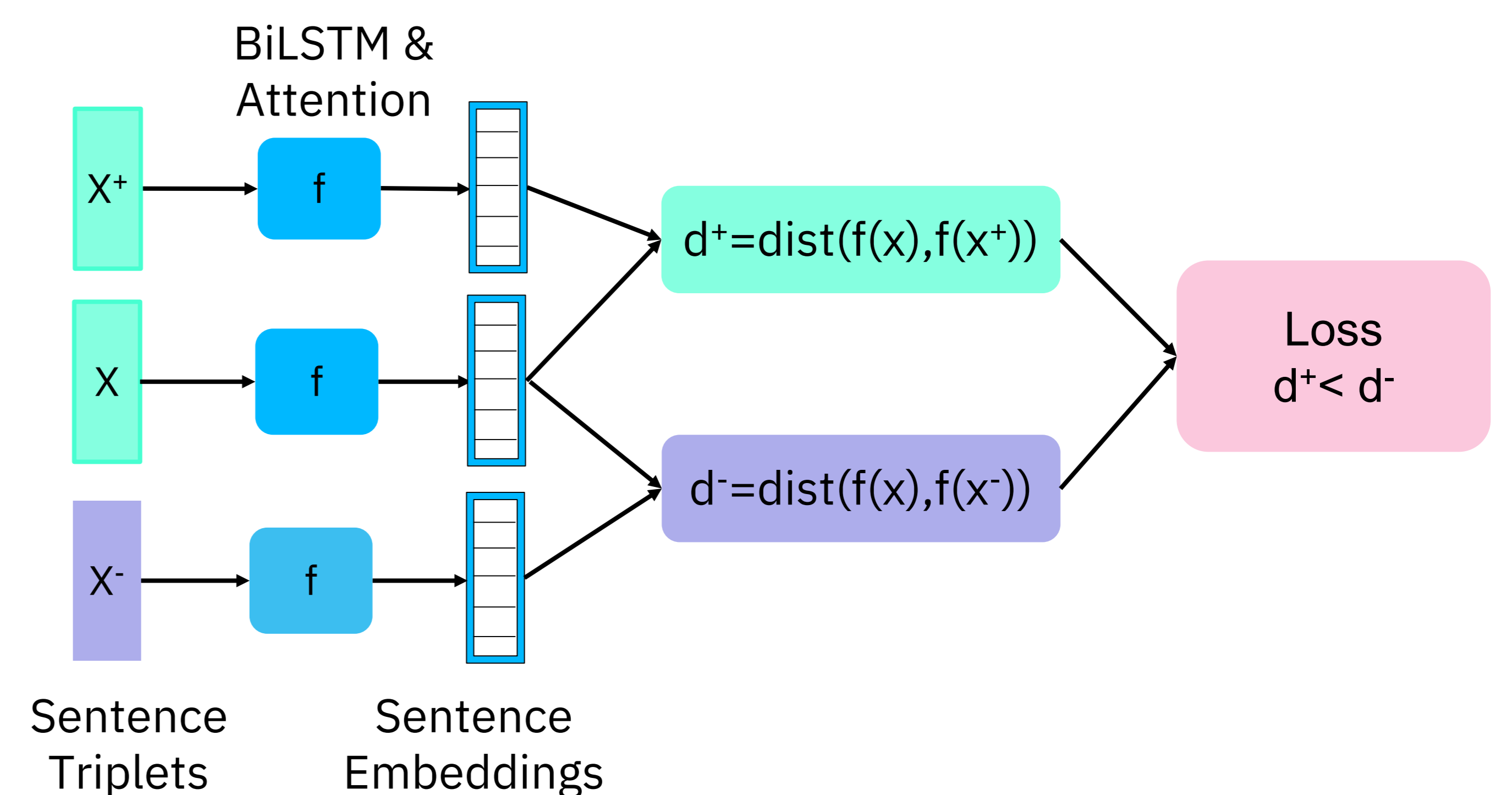


triplets-sen: Two sentences from the same section are closer than a third sentence from a different section

ttriplets-titles: A sentence is closer to its section's title than to another section's title

Dataset statistics	# Triplets - Train/Validation/Test
triplets-sen	1.78M/220K/223K
triplets-titles	1.38M/172K/173K

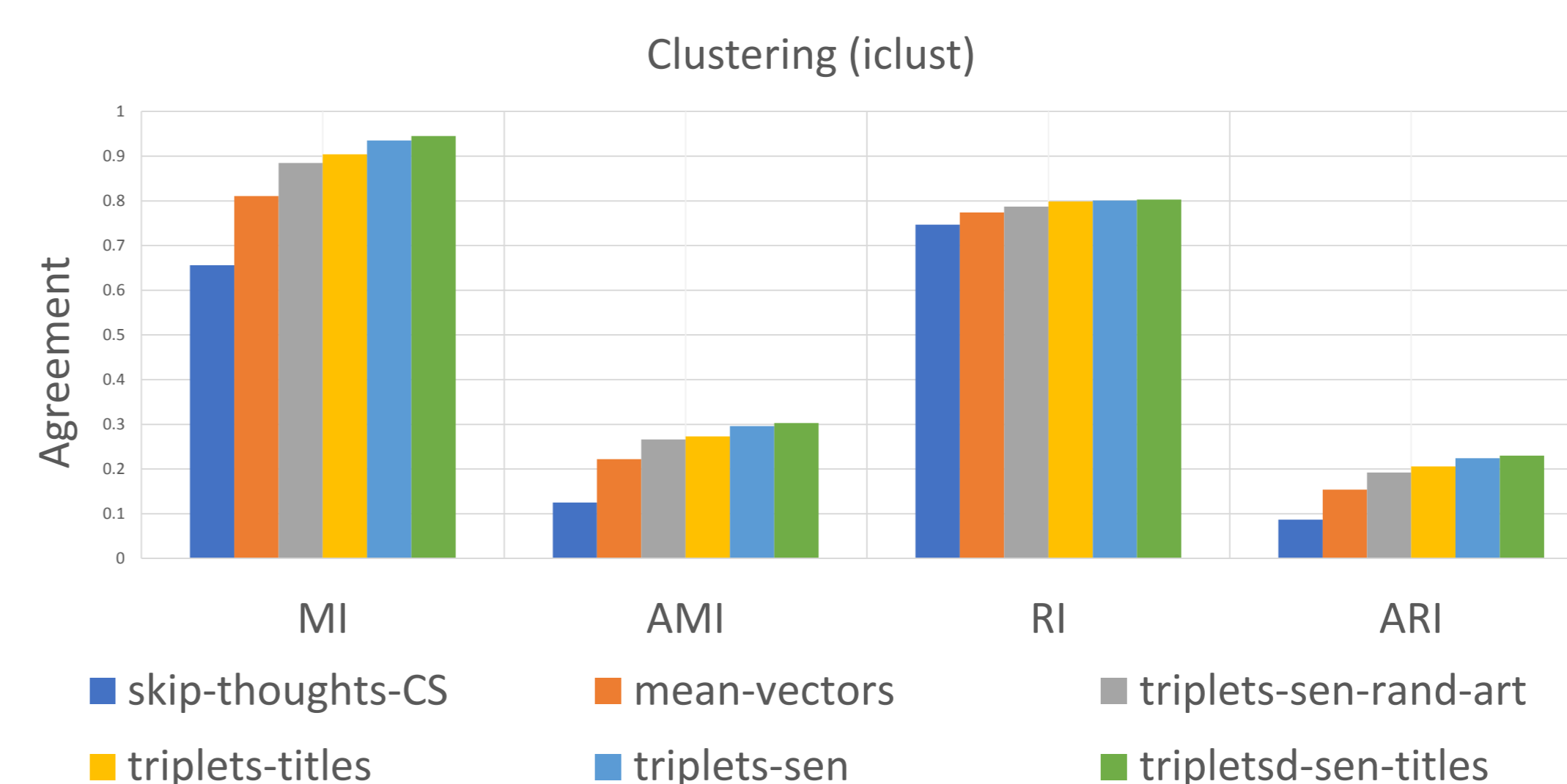
Learn Thematic Similarity (DNN Triplet Network)



Results on the clustering task

Clustering benchmark

692 articles, 5-12 clusters (sections) and 17-1614 sentences per article



Results on the triplets dataset

Method	Accuracy
mean-vectors	0.65
skip-thoughts-CS	0.615
skip-thoughts-SICK	0.547
triplets-sen	0.74

Results on the SICK semantic relatedness task

Method	Pearson	Spearman	MSE
mean-vectors	0.757	0.673	0.455
skip-thoughts-CS	0.858	0.791	0.287
triplets-sen	0.797	0.704	0.372
triplets-titles	0.786	0.685	0.393
triplets-sen-titles	0.818	0.724	0.339

Conclusion:

Triplet networks, trained on weak-supervised data, outperform other methods on thematic clustering tasks and perform on par with them on other semantic relatedness tasks.

Get the clustering benchmark here

www.research.ibm.com/haifa/dept/vst/debating_data.shtml

