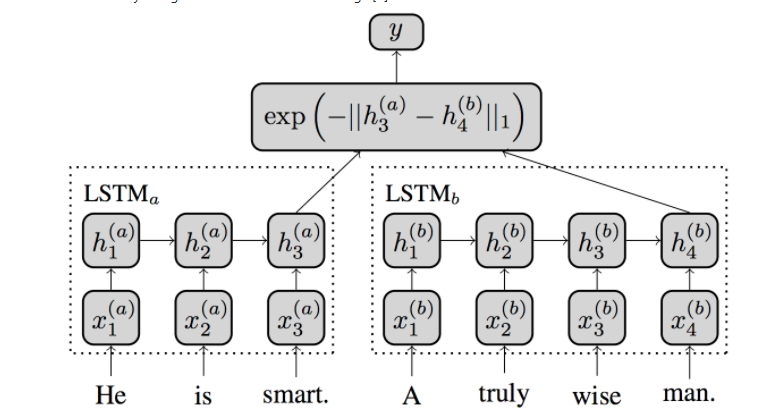
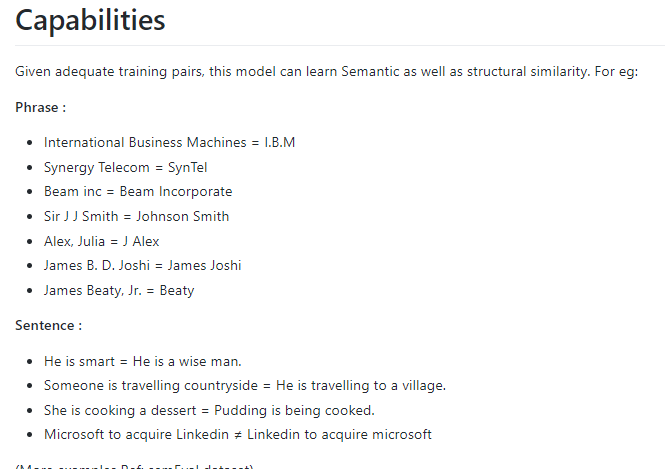
# **Semantic Similarity Approaches**

# **siamese network for text similarity (TensorFlow – LSTM & RNN)**

****

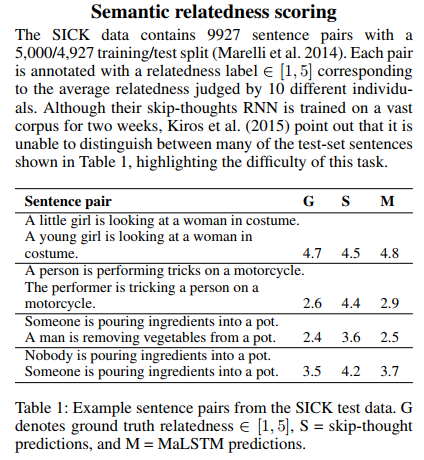
****

<https://github.com/dhwajraj/deep-siamese-text-similarity>

1. **Sentence Similarity with Keras**

**Siamese Recurrent Architectures for Learning Sentence Similarity**

<http://www.mit.edu/~jonasm/info/MuellerThyagarajan_AAAI16.pdf>

****

**Implementation**

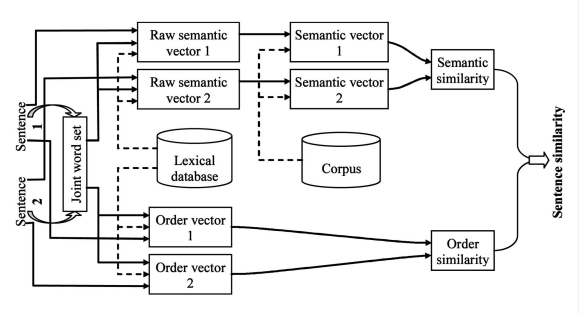
<https://stackoverflow.com/questions/39289050/sentence-similarity-using-keras>

1. **Getting word similarity by using Synset of wordNet Database**

Python/NLTK implementation of the algorithm described in the paper –

## **Sentence Similarity Based on Semantic Nets and Corpus Statistics by Li, et al.**

* This paper focuses directly on computing the similarity between very short texts of sentence length.
* It presents an algorithm that takes account of semantic information and word order information implied in the sentences.
* The semantic similarity of two sentences is calculated using information from a structured lexical database and from corpus statistics

****

**Implementation**

<https://github.com/sujitpal/nltk-examples/blob/master/src/semantic/short_sentence_similarity.py>

1. **Sematch**

**Library** : <https://github.com/gsi-upm/sematch-demo>

**Demo** : <http://sematch.cluster.gsi.dit.upm.es/>

1. **A Grammar-Based Semantic Similarity Algorithm for Natural Language Sentences**

* This paper presents a grammar and semantic corpus based similarity algorithm for natural language sentences
* Unlike existing approaches use fixed term set of vocabulary, cooccurrence terms [[1](https://www.hindawi.com/journals/tswj/2014/437162/#B49)–[3](https://www.hindawi.com/journals/tswj/2014/437162/#B51)], or even word orders [[8](https://www.hindawi.com/journals/tswj/2014/437162/#B33)], the proposed approach directly extracts the latent semantics from the same or similar links.

<https://www.hindawi.com/journals/tswj/2014/437162/>

# [**Doc2Vec Sentence Clustering**](https://stackoverflow.com/questions/43476869/doc2vec-sentence-clustering) **- by similarity (**sklearn)

Disadvantages – manual intervention is required to decide no of clusters.