Member-only story

Rule-based technique for coreference resolution

Nikhil Verma · Following
3 min read · Jun 22, 2022

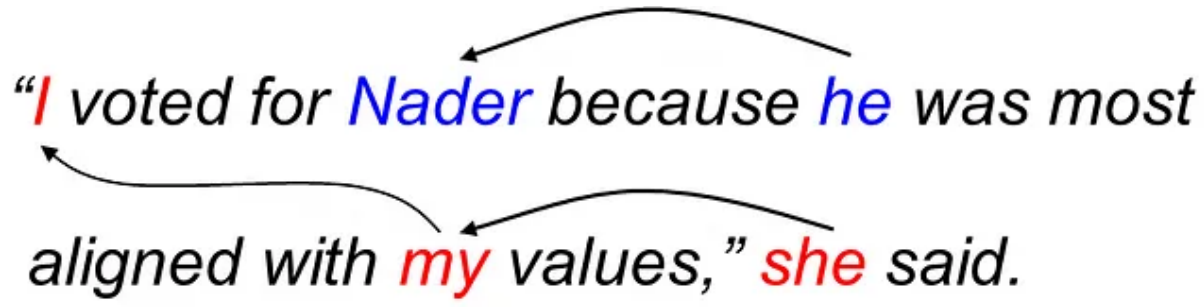


Photo from google search

Set of sentences that convey an understanding of text, when read together are referred to as discourse. In these sentences, the words that can be grouped by identity of reference are called as coreference. Using intuition from the past that coreference resolution mechanism using syntactic/semantic rules could outperform ML based techniques, the authors present a two layer model which uses hand-written rules called as sieves. It uses linguistic features from text and is entity based. It uses an entity centric approach i.e. coreference decision is globally informed. It helps in sharing features across each sieve in the framework.

The Mention Detection stage, which is a high recall stage, stands at the top of the stack because we do not want to miss mentions that are guaranteed to affect the final score. Following it the featurised text is passed through coreference resolution slabs. These slabs focus on increasing precision of the system. Here text is processed through 10 sieves. The system's precision ordering allows it to first link high-confidence mention-pairs, and only later consider lower confidence sources of information [1].

Notable sieves worth mentioning that contribute around 97% model improvement in paper are:-

- Exact string match: Two mentions are linked if they have the exact same string. Notable is that this sieve may not always improve precision as in the field of biomedical [2].
- Strict head match: Linking a mention to an antecedent based on the matching of their head words which are not just naive but follow some constraints.
- Pronominal coreference resolution: All sieves focus on nominal coreference resolution so in the end it's important to focus on pronominal coreference by considering features such as number, gender, person, animacy, NER label and pronoun distance.

Why rule-based method is competitive against pattern-recognition

No doubt that rule-based systems are only based on a certain set of rules and are static in nature whereas pattern recognition algorithms define their own rule-set that are based on input and outputs it receives.

The sieve based architecture incorporates lexical, syntactic and semantic information, and has access to document-level information. Pattern Recognition techniques mostly use a mention detection for entity recognition and then mention clustering (or graph partitioning) to split them into few different groups which hinders global information sharing.

Unlike sieve based approaches which focus first on recall and then on precision, mention clustering can lead to cascading of errors and depleting model's performance. ML techniques (such as Mention-pair, Entity-Mention and Mention-Ranking etc.) performed poorly due to insufficient knowledge or features for hard common mention cases.

It's only lately in 2017 after attention and transformer architecture introduction that the community started using LSTM and attention based techniques for end-to-end coreference resolution [3].

How to improve it?

To improve ML models, one idea is to consider all possible mentions simultaneously and assign probability for being a coreferrant w.r.t coreferee. Another thought is to create a neural net having fed a pair of coreference clusters and then try to merge clusters if required. This approach might outperform 2013 ML models but not [3].

References

1. Lee, Heeyoung, Mihai Surdeanu, and Dan Jurafsky. "A scaffolding approach to coreference resolution integrating statistical and rule-based models." Natural Language Engineering 23.5 (2017): 733–762.

2. Bell, Dane, et al. “Sieve-based coreference resolution in the biomedical domain.” arXiv preprint arXiv:1603.03758 (2016).

3. Lee, Kenton, et al. “End-to-end neural coreference resolution.” arXiv preprint arXiv:1707.07045 (2017).

NLP | Coreference Resolution | Deep Learning | Literature Review

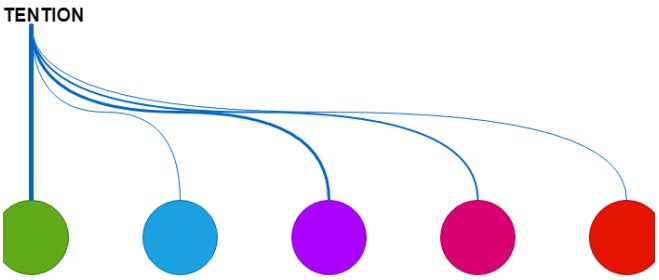
Written by Nikhil Verma

Following

911 Followers

Knowledge shared is knowledge squared | My Portfolio https://lihkinverma.github.io/portfolio/ | My blogs are living document, updated as I receive comments

More from Nikhil Verma



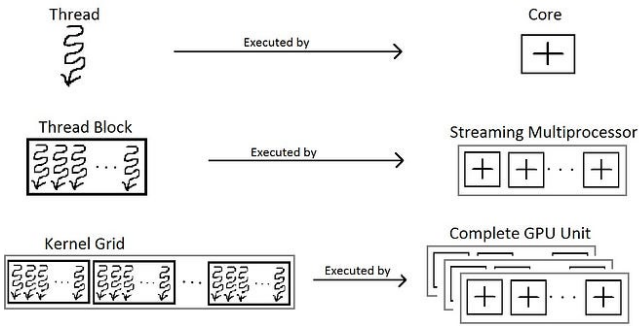
Nikhil Verma

Query, Key and Value in Attention mechanism

Transformers are like bread and butter of any new research methodology and business ide...

5 min read · Mar 26, 2022

94 1



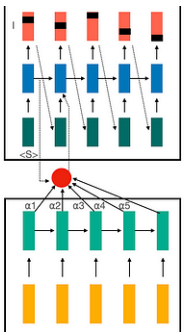
Nikhil Verma

The Symphony of Numbers: Inside the GPU's Power of 2 Architecture

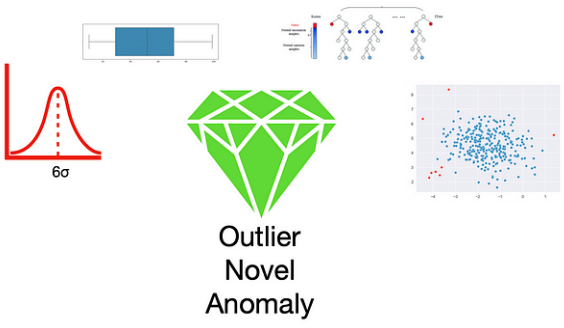
In the heart of the data center, where the pulse of artificial intelligence reverberates,...

4 min read · Sep 14

3



Nikhil Verma



Nikhil Verma

Understanding the Attention mechanism

Language modelling is the task of generating next word given previous words of a sequen...

3 min read · Jan 1

16

Outlier detection techniques

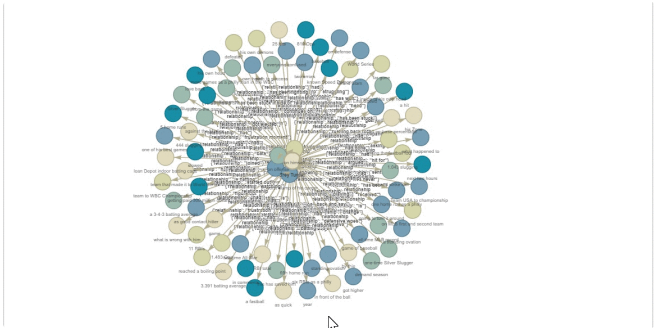
Pattern! Patterns! Patterns!

5 min read · Jul 15, 2021

97

See all from Nikhil Verma

Recommended from Medium

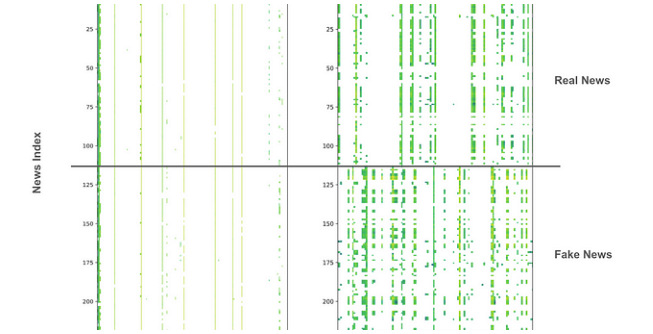


Wenqi Glantz in Better Programming

7 Query Strategies for Navigating Knowledge Graphs With...

Exploring NebulaGraph RAG Pipeline with the Philadelphia Phillies

17 min read · 4 days ago



Sherry Wu in Stanford CS224W GraphML Tutorials

Spread No More: Twitter Fake News Detection with GNN

By Li Tian, Sherry Wu, Yifei Zheng as part of the Stanford CS224W course project.

11 min read · May 16

501 4 1

Lists

Natural Language Processing

669 stories · 283 saves

The New Chatbots: ChatGPT, Bard, and Beyond

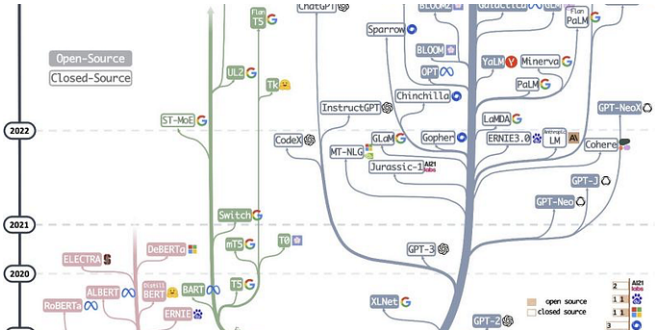
13 stories · 133 saves

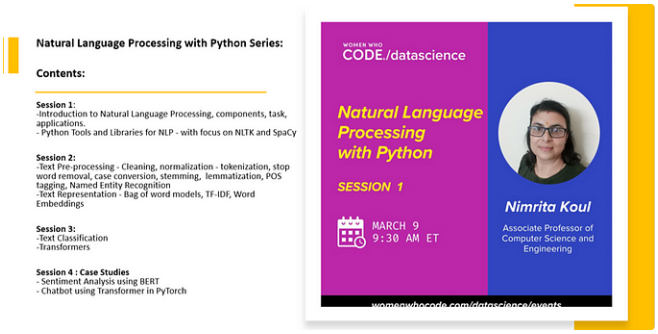
Practical Guides to Machine Learning

10 stories · 519 saves

New_Reading_List

174 stories · 133 saves





Haifeng Li

A Tutorial on LLM

Generative artificial intelligence (GenAI), especially ChatGPT, captures everyone's...

15 min read · Sep 14

Nimrita Koul

NLP with Python Part 2 NLTK

This is the second article in the series of articles on Natural Language Processing...

5 min read · Apr 5

372 2

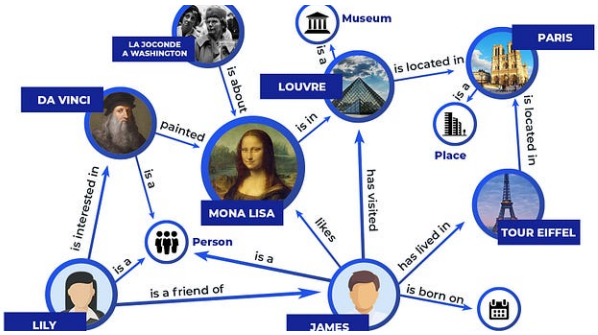
Alex Reed

Writing Your First NLP Python Script: From Text Preprocessing t...

Welcome to the world of Natural Language Processing (NLP)! NLP is a fascinating field a...

5 min read · Sep 22

56



Alla Chepurova in DeepPavlov

Improving Knowledge Graph Completion with Generative LM...

Combining both internal LM knowledge and external data from KG

13 min read · Sep 5

36

See more recommendations