Coreference Resolution using Machine Learning Models and Rule-based systems

Sanjukta De · Following
3 min read · Jan 29

The term “Coreference Resolution” refers to a technique used in natural language processing for named entity resolution (NER), information extraction, etc. that locates various mentions of an entity. For example, for the sentence ‘Annabeth is the architect of Olympus. She is an intellectual genius’, ‘Annabeth’ and ‘she’ might be linked to the same entity. To resolve such expressions several approaches have been introduced.

By searching for all mentions of an entity in a dataset, machine learning approaches facilitate coreference resolution, but they are complex and challenging to implement. Inversely, rule-based systems are simple to develop and maintain. [1] outlines a hybrid strategy that blends rule-based

systems and machine learning models. It makes use of a sieve design to ensure that the more significant constraints are prioritized more highly.

There are two steps to the approach. Entity mentions are extracted, sorted, and deleted during the mention processing phase based on syntax, NER, and patterns. By examining all noun phrases and modifying pronouns and categorising them according to various entities, this step promotes strong recall. These mentions act as an input for the sieving stage.

The sieving phase employs a range of constraints, including string matching, head matching, and shared features like gender, number, etc. The modular structure of the method promotes accuracy and provides flexibility for the insertion of new restrictions. In order to facilitate coreference resolution, the sieve architecture additionally integrates shared global entity-level data like gender, number, and animacy (Pass 10). The two sieves that most significantly contribute to the accuracy of this model are sieves 2 (exact match) and 5 (strict head match).

[2] provides a description of an entity coreference resolution method based on automatically recognizable patterns. According to this study, pattern-based systems, as opposed to rule-based systems, allow for the automatic recognition of patterns (rules) based on the input provided by searching across complete clusters of mentions defining an entity.

The sieve architecture described in [1] allows the rules in each stage to infer any information retrieved from a previous phase, which is what makes the rule-based method so competitive even though pattern-based methods are dynamic. Therefore, any coreference choice is dependent on information gathered from previously clustered mentions. So, accurate coreference

resolution begins to shift to the later phases of the approach since high-precision constraints are favored.

Adding more rules for certain exceptions, such as sentences where the noun is defined after the pronoun, could help the sieve architecture function better. For instance, “It was noon and although he had eaten a heavy breakfast, Percy was already feeling hungry.” In such sentence, the model might not be as accurate as it gives more preference to anaphoric candidates than cataphoric ones. [3] outlines the Xrenner model, which focuses primarily on compound modifiers and such phrases (also known as cataphora) like, one-way street.

Utilizing a loss function to rate the quality of a given coreference decision could be another improvement. [4] outlines a method for learning global representations of entity clusters from their mentions by utilising RNN. This performed better than the state of the art and can be used to anticipate pronominal mentions without running extra searches.

References:

[1] Lee, H., Chang, A., Peirsman, Y., Chambers, N., Surdeanu, M. and Jurafsky, D., 2013. Deterministic coreference resolution based on entity-centric, precision-ranked rules. Computational linguistics, 39(4), pp.885–916.

[2] Yang, X. and Su, J., 2007, June. Coreference resolution using semantic relatedness information from automatically discovered patterns. In Proceedings of the 45th annual meeting of the association of computational linguistics (pp. 528–535).

[3] Zeldes, A. and Zhang, S., 2016, June. When annotation schemes change rules help: A configurable approach to coreference resolution beyond OntoNotes. In Proceedings of the Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2016) (pp. 92–101).

[4] Wiseman, S., Rush, A.M. and Shieber, S.M., 2016. Learning global features for coreference resolution. arXiv preprint arXiv:1604.03035.

Coreference Resolution | Machine Learning | Rule Based Systems

Written by Sanjukta De

2 Followers

Following

More from Sanjukta De

Sanjukta De

Monotonicity in NeuralLog

Monotonicity is essential when computing Natural Language Inference (NLI) for system...

3 min read · Jan 21

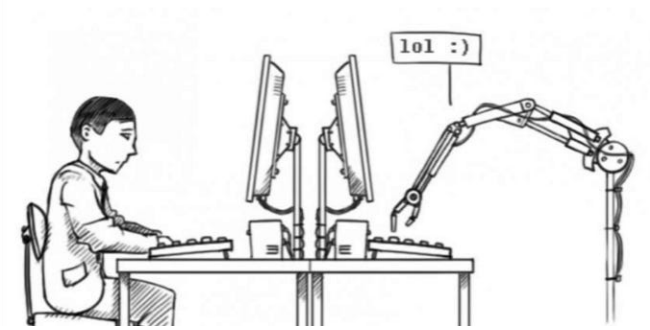
25

Sanjukta De

NP-Completeness Results for Parsing Problems in Language...

Language is ambiguous. The main challenge for linguists nowadays is how to deal with...

3 min read · Jan 9



Sanjukta De

Could a machine ever use and understand language the way...

Recent advances in natural language processing have made it possible for...

3 min read · Jan 8

13

Sanjukta De

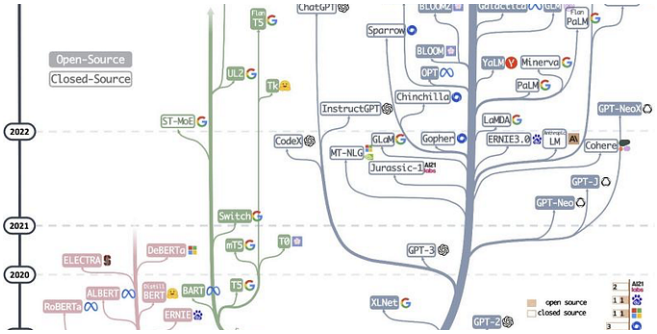
Improve Event Detection using Word Sense Disambiguation

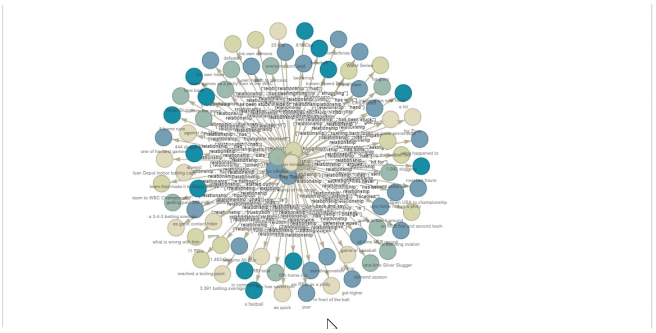
Transfer learning is the process of applying the information obtained by using one...

3 min read · Jan 15

See all from Sanjukta De

Recommended from Medium





Haifeng Li

A Tutorial on LLM

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

15 min read · Sep 14

372

Wenqi Glantz in Better Programming

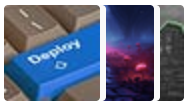
7 Query Strategies for Navigating Knowledge Graphs With...

Exploring NebulaGraph RAG Pipeline with the Philadelphia Phillies

· 17 min read · 4 days ago

501 4

Lists



Predictive Modeling w/ Python

20 stories · 452 saves

Practical Guides to Machine Learning

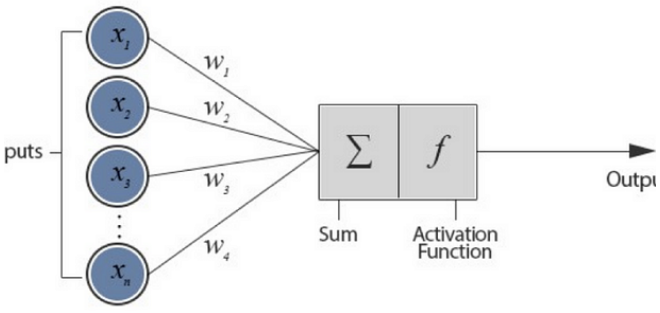
10 stories · 519 saves

Natural Language Processing

669 stories · 283 saves

The New Chatbots: ChatGPT, Bard, and Beyond

13 stories · 133 saves

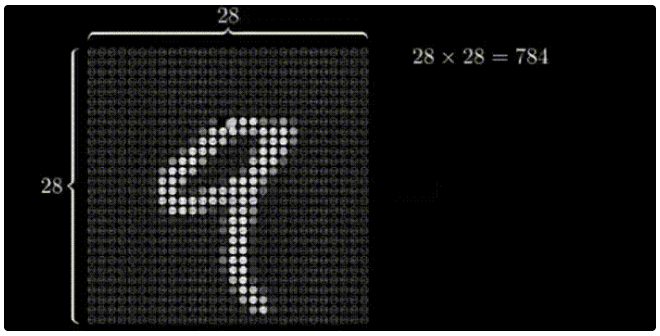


Francesco Franco

Introduction to Neural Networks

With Python implementation

12 min read · Sep 14



Sadaf Saleem

Neural Networks in 10mins. Simply Explained!

What are Neural Networks?

9 min read · May 15

48 252 2

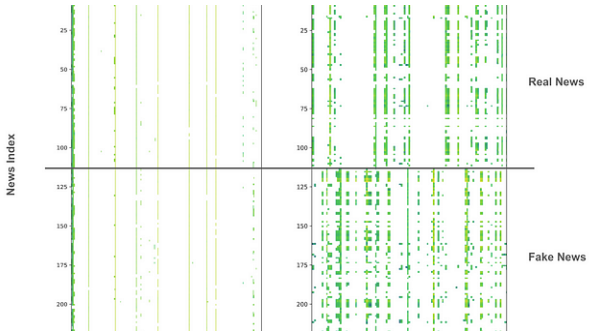
David Shapiro

A Pro's Guide to Finetuning LLMs

Large language models (LLMs) like GPT-3 and Llama have shown immense promise for...

12 min read · Sep 23

283 6



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

1

See more recommendations