

SEMEVAL-2019

TASKS

Yadong Liu

INTRODUCTION

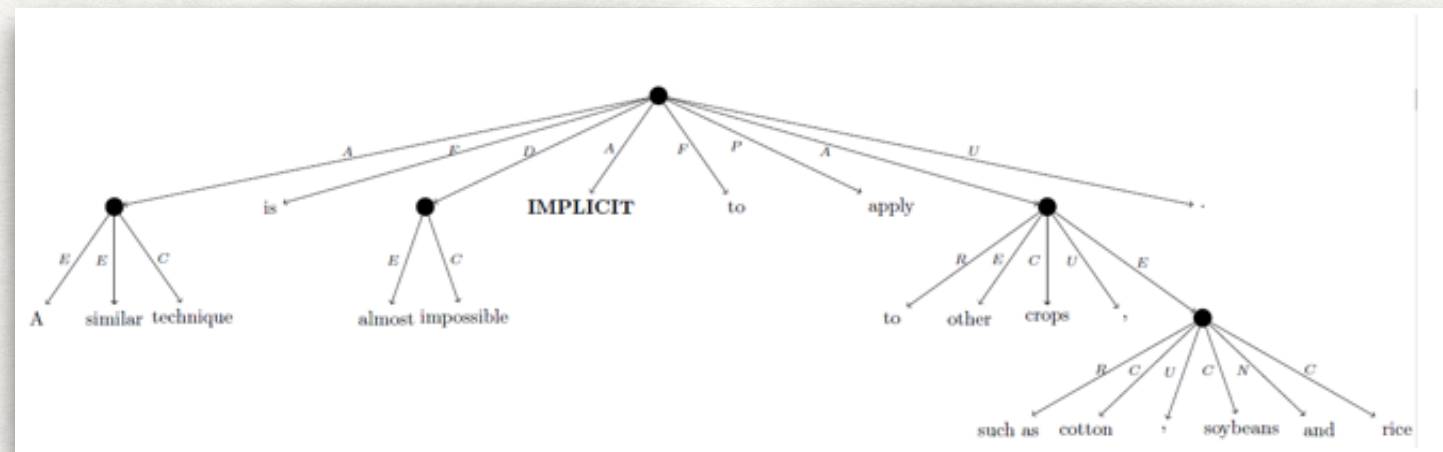
- What is SemEval?
- SemEval, aka Semantic Evaluation, is a series of evaluations of computational semantic analysis systems.
- The evaluations aim to provide mechanisms to help with problems and solutions identification for meaning computations.
- <https://en.wikipedia.org/wiki/SemEval>
- Participants need to submit a solution for a/some task(s) with the given dataset.

SEMEVAL-2019 TASKS

- Overview
- 12 tasks divided into 5 different categories have been released so far.
- <http://alt.qcri.org/semEval2019/index.php?id=tasks>
- Frame semantics and semantic parsing
- Opinion, emotion and abusive language detection
- Fact vs fiction
- Information extraction and question answering
- NLP of scientific applications

TASK 1 CROSS-LINGUAL SEMANTIC PARSING WITH UCCA

- https://competitions.codalab.org/competitions/19160#learn_the_details
- This task is about parsing text using the UCCA semantic annotation.
- UCCA is a semantic representation scheme that is applicable cross-linguistically. It represents semantics of human speech utterances as directed acyclic graphs.



- Public data and Starting kit can be obtained from the link above.

TASK 2 UNSUPERVISED LEXICAL SEMANTIC FRAME INDUCTION TASK

- <https://competitions.codalab.org/competitions/19159>
- Lexical frame induction refers to the process of verbs and their dependent words grouping in type-feature structure (i.e. frames) in a fully unsupervised way.
- Task 1. Grouping Verbs to Frame Type Clusters. (Group verbs in clusters based on their meanings)
- Task 2.1. Clustering arguments of verbs to frame-specific slots
- Task 2.2. Clustering arguments of verbs to generic roles (semantic role labelling)
- You must register with LDC to get access to the data. https://competitions.codalab.org/competitions/19159#learn_the_details-datasets

TASK 3 EMOCONTEXT

- <https://www.humanizing-ai.com/emocontext.html>
- This task asks you to classify the emotion of the utterance as one of the classes: happy, sad, angry or others with a given textual dialogue.
- Training data sets include 15K records for emotion class, and 15K records for others.
- Dataset can be obtained by joining the EmoContext LinkedIn group indicated on the webpage above.

TASK 4 HYPERPARTISAN NEWS DETECTION

- <https://pan.webis.de/semeval19/semeval19-web/>
- This task requires participants to determine whether a news article follows a hyperpartisan argumentation. For instance, whether it shows blind, unreasoning or prejudiced allegiance to one party or person.
- A training dataset that contains 1 million articles labeled by the overall tendency will be released in September. A trial dataset will be provided upon registration.
- Registration link: https://docs.google.com/forms/d/e/1FAIpQLScqDhTNqCHiA_XZoKLq0QdEOmcIjcuF9RJoPwdekuy1GuoESrw/viewform

TASK 5 SHARED TASK ON MULTILINGUAL DETECTION OF HATE

- <https://competitions.codalab.org/competitions/19935>
- This task is about the detection of Hate Speech towards two target groups, women and immigrants in Spanish and English.
- Task A. Hate Speech Detection against Immigrants and Women. It is a classification task to predict whether a tweet is hateful or not.
- Task B. Aggressive behaviour and Target Classification. It asks system to classify whether a hateful tweet is aggressive or not.
- Join the google group for more details: [semeval2019-task5-hateval\[at\]googlegroups.com](mailto:semeval2019-task5-hateval[at]googlegroups.com)
- Dataset is available TODAY!

TASK 6 OFFENSEVAL

- <https://competitions.codalab.org/competitions/20011>
- OffensEval: Identifying and Categorizing Offensive Language in Social Media.
- Task A. Offensive language identification
- Task B. Automatic categorization of offence types
- Task C. Offence target identification.
- The trial data is available upon registration. <https://competitions.codalab.org/competitions/20011#participate>
- The training data will be available on Oct 10, 2018.

TASK 7 RUMOUREVAL

- <https://competitions.codalab.org/competitions/19938>
- This task asks to detect the veracity of rumours.
- Task A classifies responses of a rumour according to stance. i.e. support, deny, query, and comment.
- Task B classifies rumour for veracity.
- Data. Join the google group for details
- <https://groups.google.com/forum/#!forum/rumoureval>

TASK 8 FACT CHECKING IN COMMUNITY QUESTION ANSWERING FORUMS

- <https://competitions.codalab.org/competitions/20022>
- Task A classifies whether a question is asking for an opinion, a factual information or just socializing.
- Task B classifies whether an answer to a factual information is true, false, or not considered as a proper answer.
- Data will be available tomorrow!
- Join the google group for details. [semeval-2019-task-8](#)
- Email organizer at semeval-2019-task-8-organizers@googlegroups.com

TASK 9 SUGGESTION MINING FROM ONLINE REVIEWS AND FORUMS

- <https://competitions.codalab.org/competitions/19955>
- Suggestion mining refers to tips/advice/recommendations extraction from unstructured text.
- Task A. Participants are required to perform suggestion mining within a specific domain.
- Task B. Participants are required to perform suggestion mining cross different domains.
- Dataset can be obtained here. <https://github.com/Semeval2019Task9?tab=repositories>

TASK 10 MATH QUESTION ANSWERING

- <https://competitions.codalab.org/competitions/20013>
- This task asks participant to train a math QA system to answer SAT questions. Most of the questions are 5-way multiple choice, and some have a numeric answer.
- You can get started at <https://github.com/allenai/semEval-2019-task-10/blob/master/docs/gettingStarted.md>

TASK 11 NORMALIZATION OF MEDICAL CONCEPTS IN CLINICAL NARRATIVE

- <https://competitions.codalab.org/competitions/19350>
- This task focus on Named Entity Normalization (NEN). NEN involves linking named entity to standardized concept. For instance, heart attack, MI, and Myocardial infarction refer to the same general concept.
- Participation requests will be approved upon issues with i2b2 being solved.

TASK 12 TOPONYM RESOLUTION IN SCIENTIFIC PAPERS

- <https://competitions.codalab.org/competitions/19948>
- Toponym resolutions aims to assign location names mentioned in texts with geographic coordinates.
- Task 1 involves toponym detection.
- Task 2 emphasizes on the disambiguation of the toponyms. Participants need to build a resolver that choose an GeoNames (a database for geospatial location) ID corresponding to the expected place.
- Task 3 evaluates the resolver in task 2.
- Join the google group on the webpage above for details.

- Thank you!