

Simone Papandrea

✉ papandrea.simone@gmail.com

Alessandro Raganato

✉ raganato@di.uniroma1.it

Claudio Delli Bovi

✉ dellibovi@di.uniroma1.it

Online Demo: <https://goo.gl/yXq84e>
Also available as
Chrome Extension!

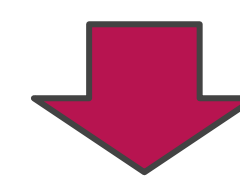
- **Right-click** on text and select:
 Disambiguate
- When the disambiguation is ready, the text will be **highlighted**
- **Mouse over** it to see the sense tags

The human brain is quite proficient at word-sense disambiguation. The fact that natural language is formed in a way that requires so much of it is a reflection of that neurologic reality. In computer science and the information technology that it enables, it has been a long-term challenge to develop the ability in computers to do natural language processing and machine learning.

Disambiguate

Clear

training data: SemCor+OMSTI with Hybrid preprocessor. Models include surrounding words, POS tags of surroundings words, and local collocations as features. All models use WordNet 3.0 as sense inventory.



Results

The human **brain** is quite **proficient** at word-sense disambiguation . The fact that natural language is formed **in a way** that **requires** so much of it is a reflection of that neurologic reality . In computer science and the information technology that it enables , it has been a long-term **challenge** to **develop** the ability in computers to do natural language processing and machine learning .

(ADJECTIVE) proficient%5:00:00:skilled:00
having or showing knowledge and skill and aptitude; "adept in handicrafts"; "an adept juggler"; "an expert job"; "a good mechanic"; "a practiced marksman"; "a proficient engineer"; "a lesser-known but no less skillful composer"; "the effect was achieved by skillful retouching"

(NOUN) brain%1:08:00::
that part of the central nervous system that includes all the higher nervous centers; enclosed within the skull; continuous with the spinal cord

(VERB) develop%2:36:01::
make something new, such as a product or a mental or artistic creation; "Her company developed a new kind of building material that withstands all kinds of weather"; "They developed a new technique"

(VERB) require%2:42:00::
require as useful, just, or proper; "It takes nerve to do what she did"; "success usually requires hard work"; "This job asks a lot of patience and skill"; "This position demands a lot of personal sacrifice"; "This dinner calls for a spectacular dessert"; "This intervention does not postulate a patient"

(ADVERB) in_a_way%4:02:00::
from some points of view; "She was right in a way"

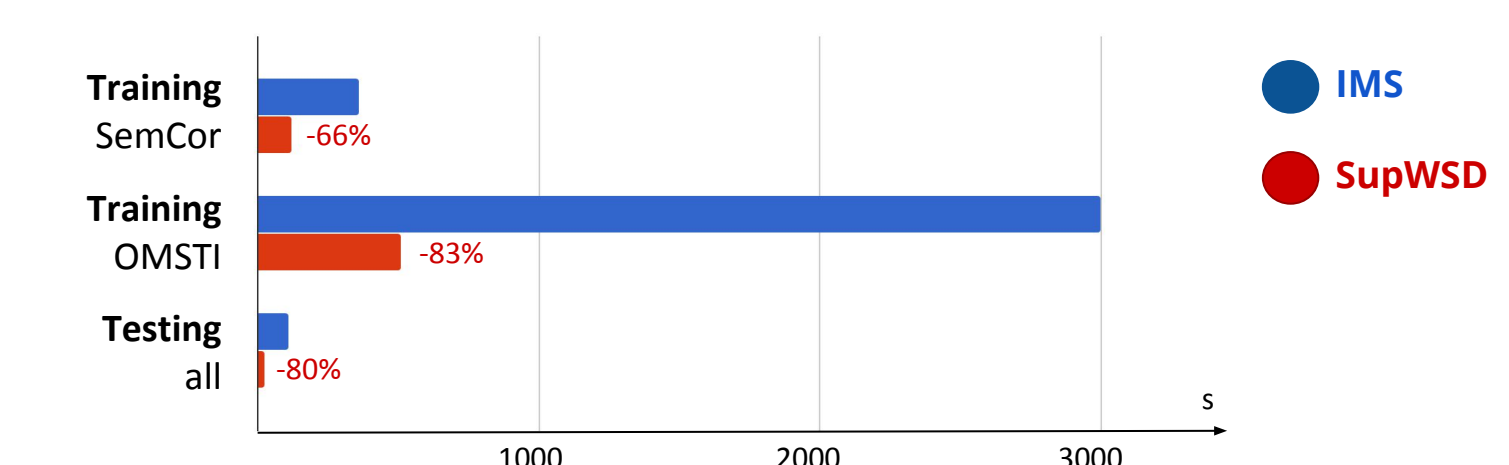
(NOUN) challenge%1:26:00::
a demanding or stimulating situation; "they reacted irrationally to the challenge of Russian power"

Experiments

- F-scores (%) of different models in five all-words WSD datasets.

Tr. Corpus	System	Senseval-2	Senseval-3	SemEval-07	SemEval-13	SemEval-15
SemCor	IMS	70.9	69.3	61.3	65.3	69.5
	SUPWSD	71.3	68.8	60.2	65.8	70.0
	IMS+emb	71.0	69.3	60.9	67.3	71.3
	SUPWSD+emb	72.7	70.6	63.1	66.8	71.8
	IMS _s +emb	72.2	70.4	62.6	65.9	71.5
	SUPWSD _s +emb	72.2	70.3	63.3	66.1	71.6
	Context2Vec	71.8	69.1	61.3	65.6	71.9
SemCor + OMSTI	MFS	65.6	66.0	54.5	63.8	67.1
	IMS	72.8	69.2	60.0	65.0	69.3
	SUPWSD	72.6	68.9	59.6	64.9	69.5
	IMS+emb	70.8	68.9	58.5	66.3	69.7
	SUPWSD+emb	73.8	70.8	64.2	67.2	71.5
	IMS _s +emb	73.3	69.6	61.1	66.7	70.4
	SUPWSD _s +emb	73.1	70.5	62.2	66.4	70.9
	Context2Vec	72.3	68.2	61.5	67.2	71.7
	MFS	66.5	60.4	52.3	62.6	64.2

- Speed comparison for both training and testing:


Toolkit: <https://goo.gl/ACYijt>

- The **SupWSD toolkit** (version 1.0.0, June 2017) is an open-source Java toolkit for supervised WSD (also including pre-trained English models on SemCor and OMSTI) which enables you to:

- add a **customized parser** for the input (XML, free text, etc.)
- add a **customized preprocessing module** to the pipeline (e.g. NER)
- add a **new feature** to the classifier (with default value, cut-off value, extractor)
- add a **new classifier** to the pipeline

Java API: <https://goo.gl/pU6HN5>

- The **SupWSD Java API** (version 1.1.0, June 2017) is a Java binding to an HTTP RESTful service that gives you programmatic access to SupWSD:

```
import it.si3p.supwd.SupWSD;
import it.si3p.supwd.SupSense;
import it.si3p.supwd.SupResult;
import it.si3p.supwd.SupWSDException;

public class Main {
    public static void main(String[] args) throws SupWSDException {
        String text = "The human brain is quite proficient at word-sense disambiguation.";
        SupWSD swd = SupWSD.getInstance();
        for (SupSense sense : swd.getSenses(text, true)) {
            System.out.println("Word: " + sense.getWord() + "\tLemma: " + sense.getLemma() + "\tPOS: "
                + sense.getPOS());
            for (SupResult result : sense.getResults()) {
                System.out.println("SenseKey " + result.getKey() + "\tProbability: " + result.getProbability());
            }
        }
    }
}
```

Sense keys
(WordNet)Confidence
(probability)