Interpretable Natural Language Processing

Anton Kolonin

akolonin@aigents.com

Facebook: akolonin

Telegram: akolonin







autonio.foundation



SingularityNET

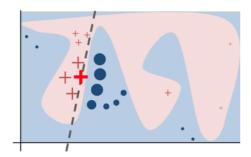
Why do we need Interpretable NLP (AI)?

Certain kinds of confusing classification errors due to errors in models may be better avoided in advance rather than found excuses later.

Explaining sentiment analysis model by LIME¹ method

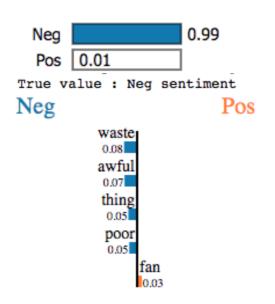
LIME (Local interpretable model agnostic explanations)

- Model agnostic
- Learns local behaviour of black box model around an instance for a candidate linear interpretable model



1 Toy example showing LIME learning linear behaviour around class boundary

Prediction Probabilities



➤ **LIME** technique implemented on black box sentiment analysis classifier trained on IMDB movie dataset². In the image, Lime estimates the model's decision making criteria for prediction on a test instance.

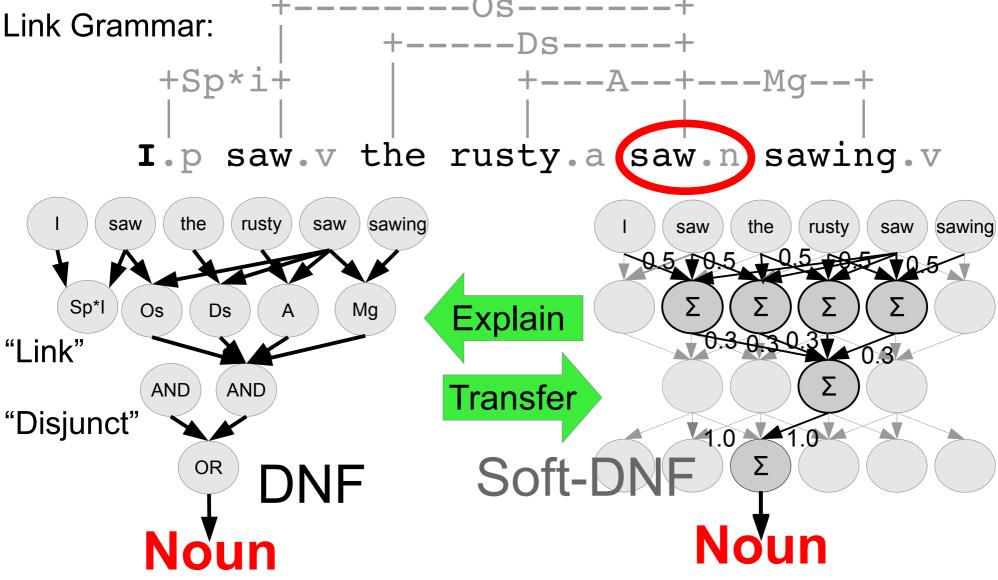
Text with highlighted words

Being a long-time fan of Japanese film, I expected more than this. I can't really be bothered to write to much, as this movie is just so poor. The story might be the cutest romantic little something ever, pity I couldn't stand the awful acting, the mess they called pacing, and the standard "quirky" Japanese story. If you've noticed how many Japanese movies use characters, plots and twists that seem too "different", forcedly so, then steer clear of this movie. Seriously, a 12-year old could have told you how this movie was going to move along, and that's not a good thing in my book.lbr /llbr /lFans of "Beat" Takeshi: his part in this movie is not really more than a cameo, and unless you're a rabid fan, you don't need to suffer through this waste of film.lbr /llbr /l2/10

- 1 Source: https://arxiv.org/abs/1602.04938
- 2 Source: https://ai.stanford.edu/~amaas/data/sentiment/

Bridging the Symbolic-Subsymbolic gap

in NLP between distributed representations and formal grammars with ontologies



https://arxiv.org/abs/1401.3372

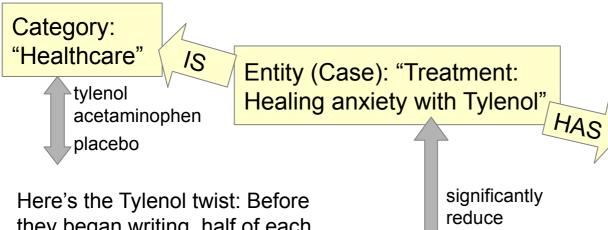
https://arxiv.org/pdf/2006.00093.pdf

Aigents® "Deep Patterns" - Text Mining

Classification

Case/Relationship Extraction

Property Attribution Entity Extraction



Here's the Tylenol twist: Before they began writing, half of each group received acetaminophen while the other half swallowed a placebo. Even among those people who wrote about death, the Tylenol takers set bail at roughly \$300—a sign that acetaminophen may significantly reduce feelings of existential anxiety, explains study lead author Daniel Randles, a PhD candidate in UBC's department of... psychology.

significantly reduce feelings study

"acetaminophen may significantly reduce feelings of existential anxiety, explains study lead author Daniel Randles" Brand: Tylenol

Substance:

acetaminophen

Reliability: medium

Effect: positive

Diagnosis: Anxiety

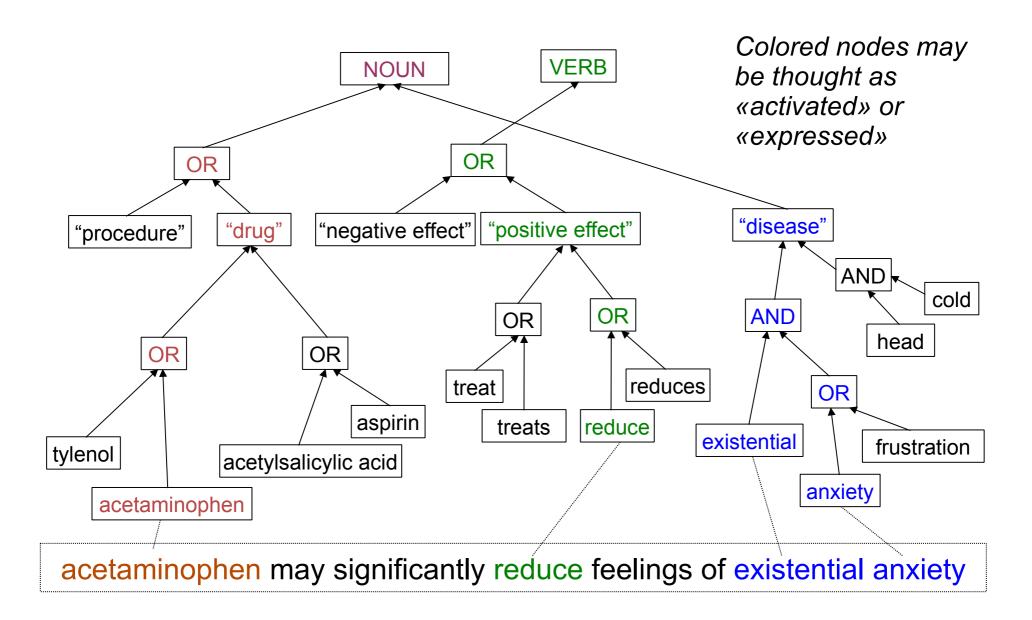
Reporter: Daniel Randles

acetaminophen may reduce anxiety explains

acetaminophen may significantly reduce feelings of existential anxiety, explains study lead author Daniel Randles.

https://ieeexplore.ieee.org/document/7361868?arnumber=7361868 https://github.com/aigents/aigents-java

Aigents® "Deep Patterns" - Language Model



https://ieeexplore.ieee.org/document/7361868?arnumber=7361868 https://github.com/aigents/aigents-java

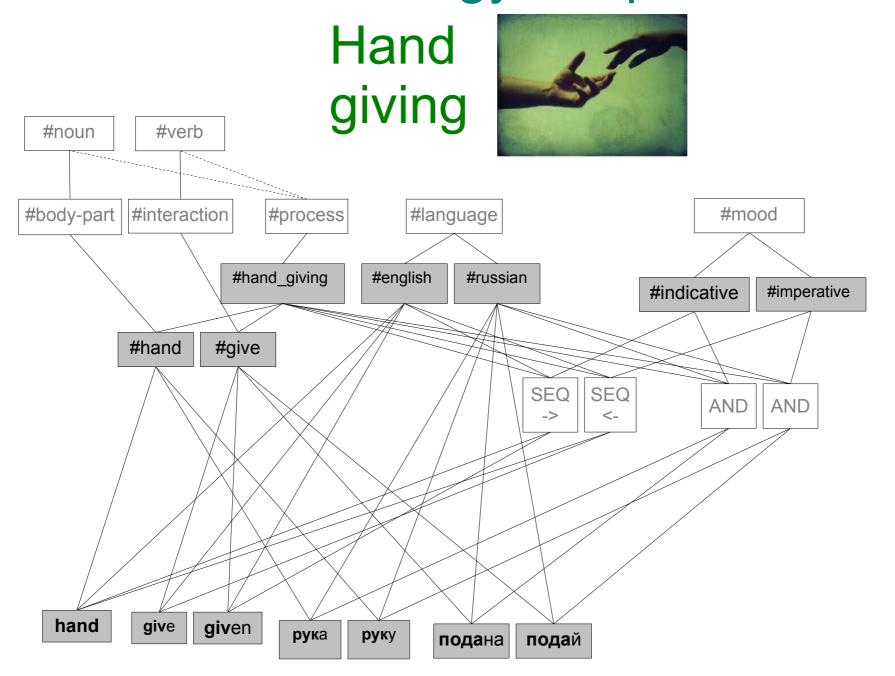
Aigents® "Deep Patterns" - Text Mining

```
<set> := <disjunctive-set> | <conjunctive-set> | <M-skip-N-gram>
<disjunctive-set> := { <pattern> * }
                                        Variables may have domain restrictions
<conjunctive-set> := ( <pattern> * )
                                          in ontology and/or refer to other
                                              patterns as subgraphs
<N-gram> := [ <pattern> * ]
<pattern> := <token> | <regexp> | <variable> | <set>
                 Example:
{[$description catheter] [$coating coating] [$inner-diameter
   {diameter inner-diameter}] [$tip tip] [$pattern pattern]}
"Convey Guiding Catheter. Unique hydrophilic coating.
   Small atraumatic soft tip. Ultra-thin 1 × 2 flat wire braid pattern"
{ coating : "hydrophillic", description : "convey guiding",
```

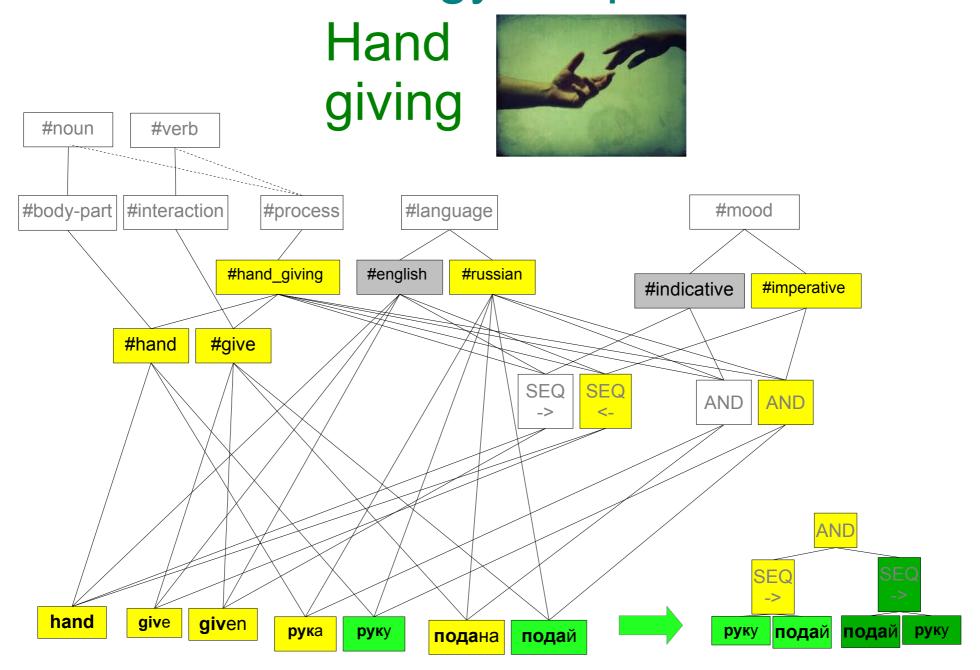
https://ieeexplore.ieee.org/document/7361868?arnumber=7361868 https://github.com/aigents/aigents-java

pattern: "ultra-thin 1 × 2 flat wire braid", tip: "soft" }

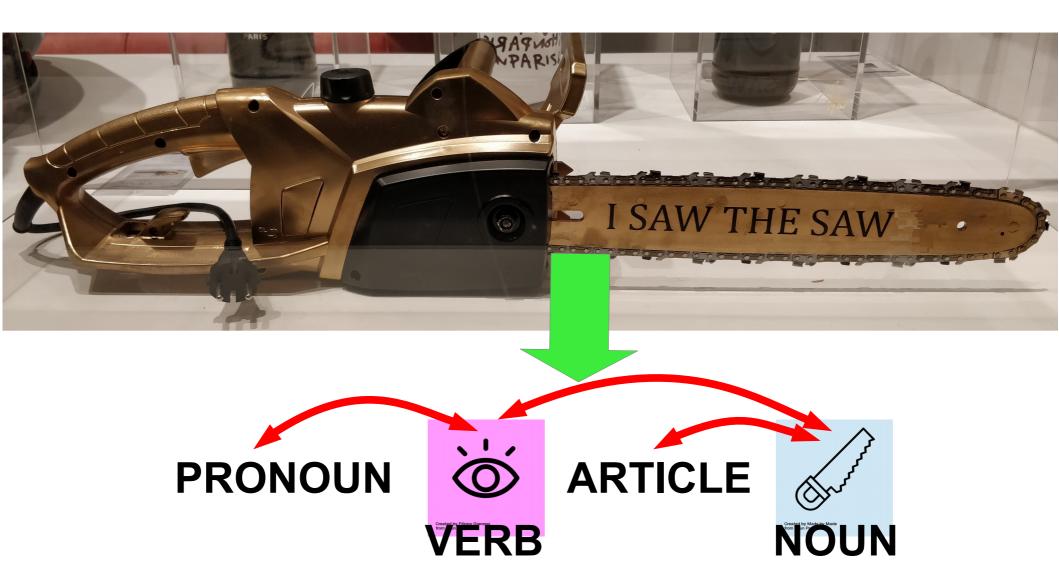
Grammar & Ontology Graph - Structure



Grammar & Ontology Graph - Production



Unsupervised Link Grammar Learning

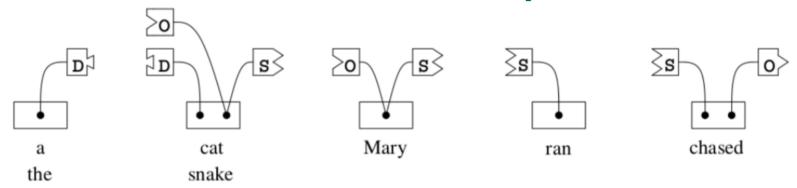


https://www.springerprofessional.de/unsupervised-language-learning-in-opencog/15995030 https://www.springerprofessional.de/en/programmatic-link-grammar-induction-for-unsupervised-language-le/17020348

Project goal and applications

- Grammar learning from scratch programmatically
- Grammar extension/customization for specific domains
- Building dictionaries and patterns for NLP applications
- Parsing texts for NLP applications
- Grammar checking (more than spell checking)

Link Grammar Connectors & Disjuncts as a fundamental particles

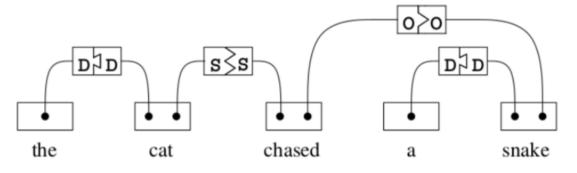


An illustration of Link Grammar connectors and disjuncts. The connectors are the jigsaw-puzzle-shaped pieces; connectors are allowed to connect only when the tabs fit together. A disjunct is the entire (ordered) set of connectors for a word. As lexical entries appearing in a dictionary, the above would be written as

```
a the: D+;
cat snake: D- & (S+ or O-);
Mary: O- or S+;
ran: S-;
chased S- & O+;
```

Note that although the symbols ''&'' and ''or'' are used to write down disjuncts, these are **not** Boolean operators, and do **not** form a Boolean algebra. They do form a non-symmetric compact closed monoidal algebra. The diagram below illustrates puzzle pieces, assembled to form a parse:

B. Goertzel, L. Vepstas, 2014

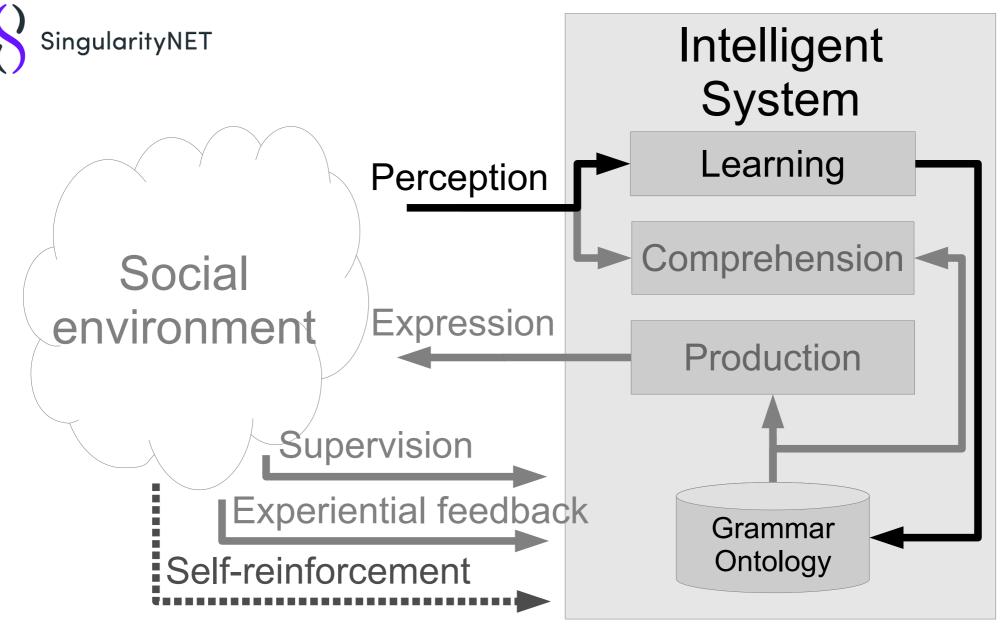




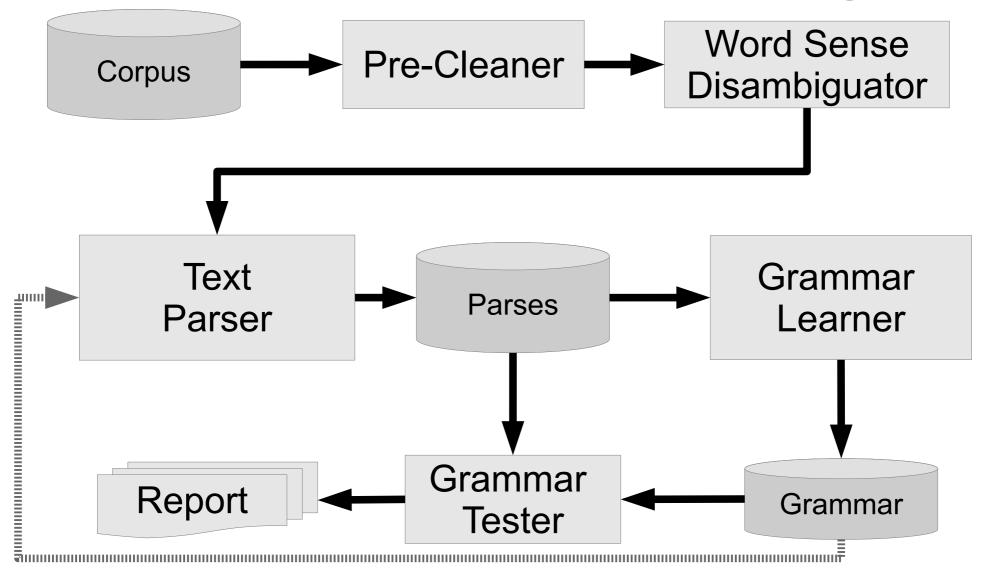
Constraints of the currently explored approach

- Controlled corpora
- Using Link Grammar formalism
- Relying on MST parses
- No account for morphology
- Self-reinforcement with F1 on parses
- Test against training data

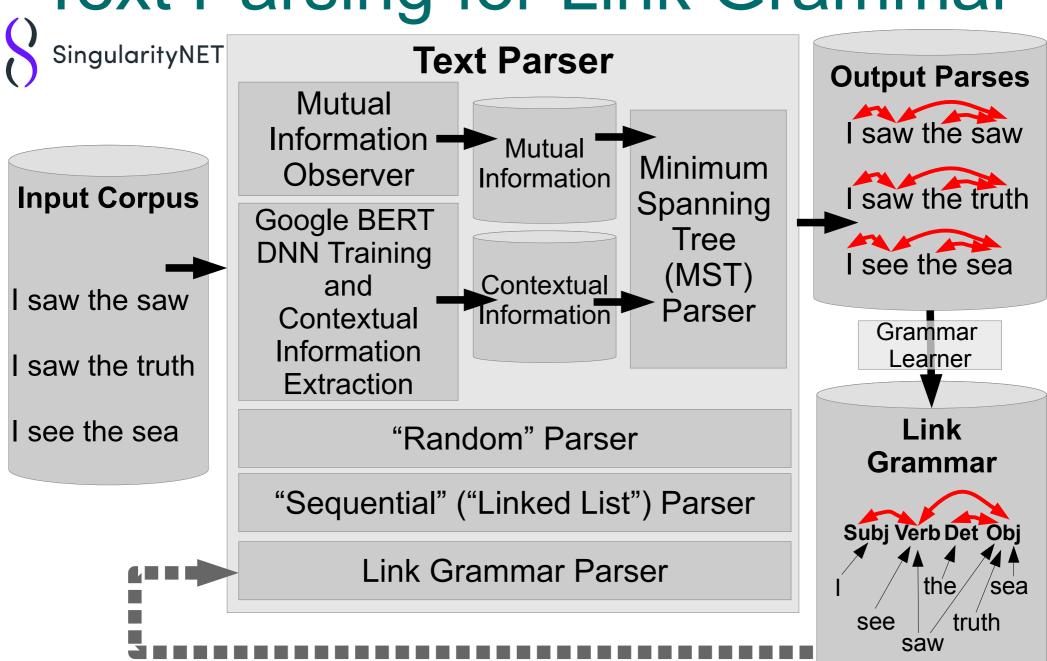
Language Learning Environment



Unsupervised language learning pipeline with OpenCog

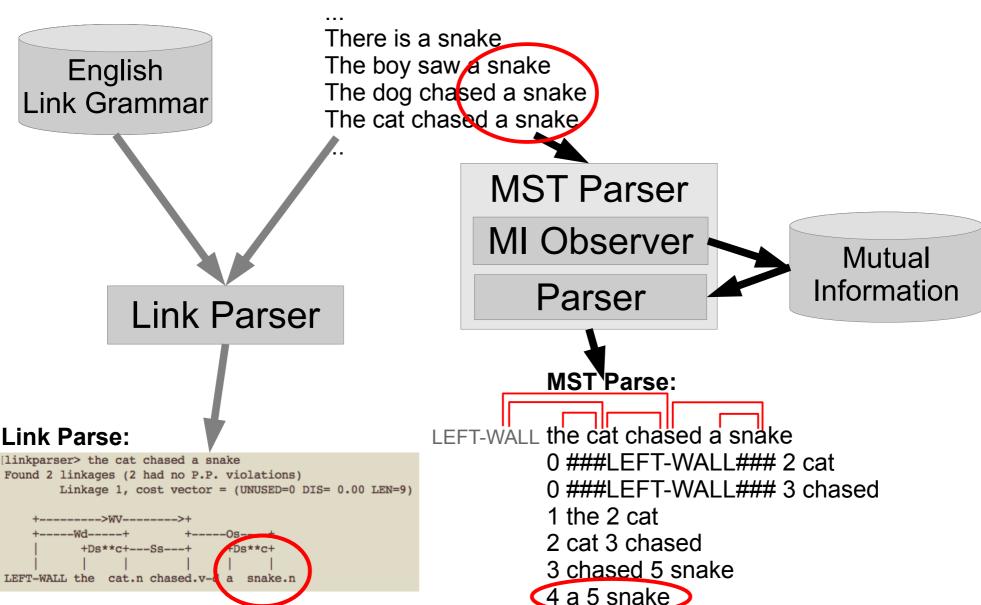


Text Parsing for Link Grammar

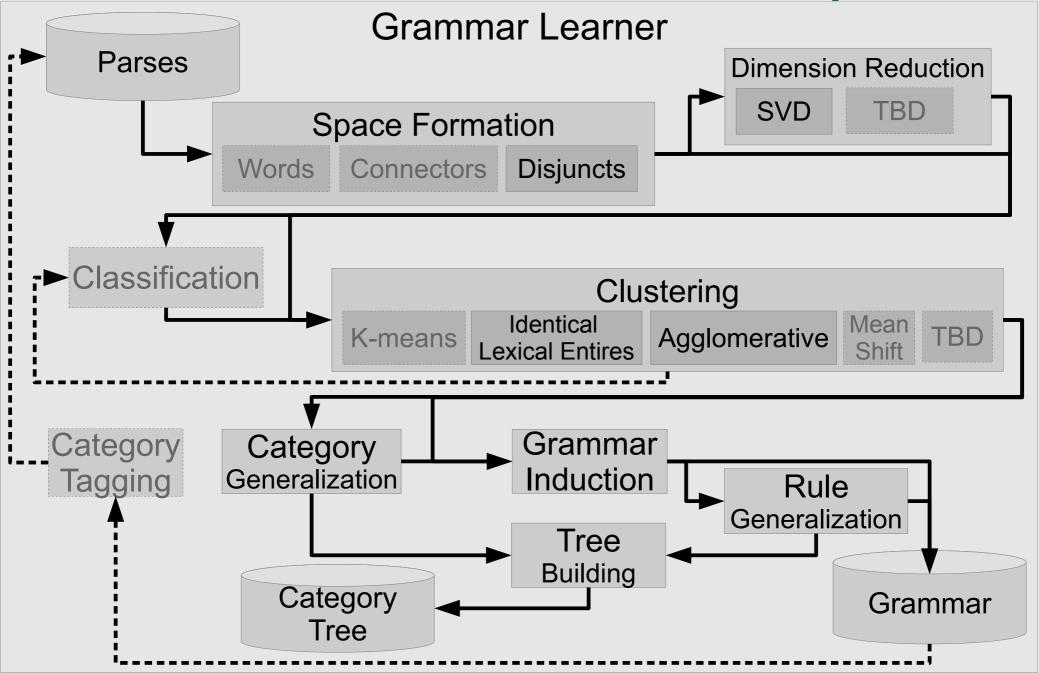


MST Parses vs. Link Parses





Link Grammar Learner Pipeline





Corpora in Use

Corpus	Total words	Unique words	Occurrences per word	Total sentences	Average sentence length
POC-English	388	55	7	88	4
Child-Directed Speech	124185	3399	37	38181	4
Gutenberg Children	2695151	54054	50	207130	13

- POC-English Proof-of-Concept corpus made of artificially selected sentences on limited number of topics ("small world").
- Child Directed Speech (CDS) corpus obtained from subsets of the CHILDES corpus – a collection of English communications directed to children with limited lexicon and grammar complexity https://childes.talkbank.org/derived/
- Gutenberg Children (GC) compendium of books for children contained within Project Gutenberg (https://www.gutenberg.org), following the selection used for the Children's Book Test of the Babi CBT corpus https://research.fb.com/down-loads/babi/

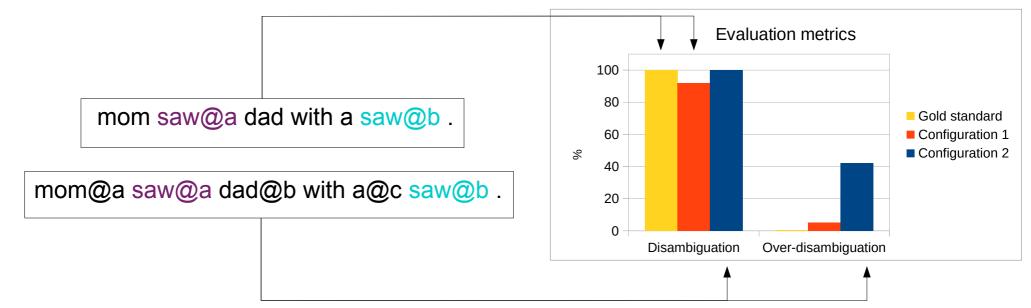
Word-Sense Disambiguation

Using AdaGram¹ we disambiguate our POC-English corpus without supervision.

Two ambiguous words in corpus, with only two senses each:



After parameter tuning, we found two promising results:

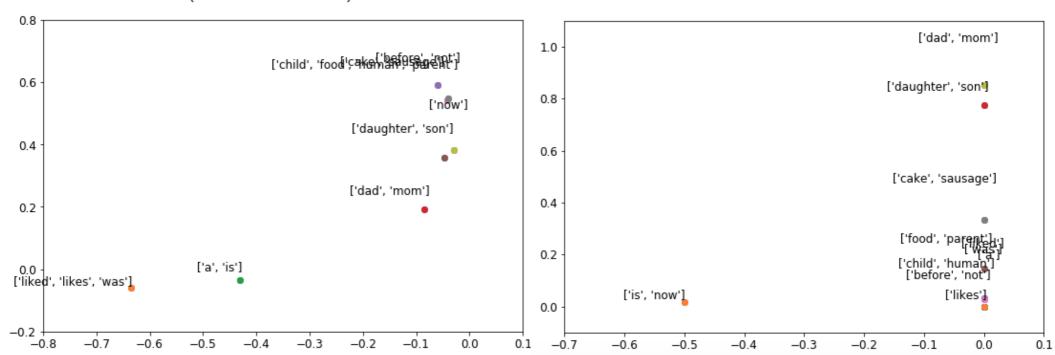


¹ https://github.com/glicerico/AdaGram/tree/take_sentences

Learning Grammatical and Semantic Categories in Vector Spaces of Connectors and Disjuncts

POC-English (Connectors)

POC-English (Disjuncts)

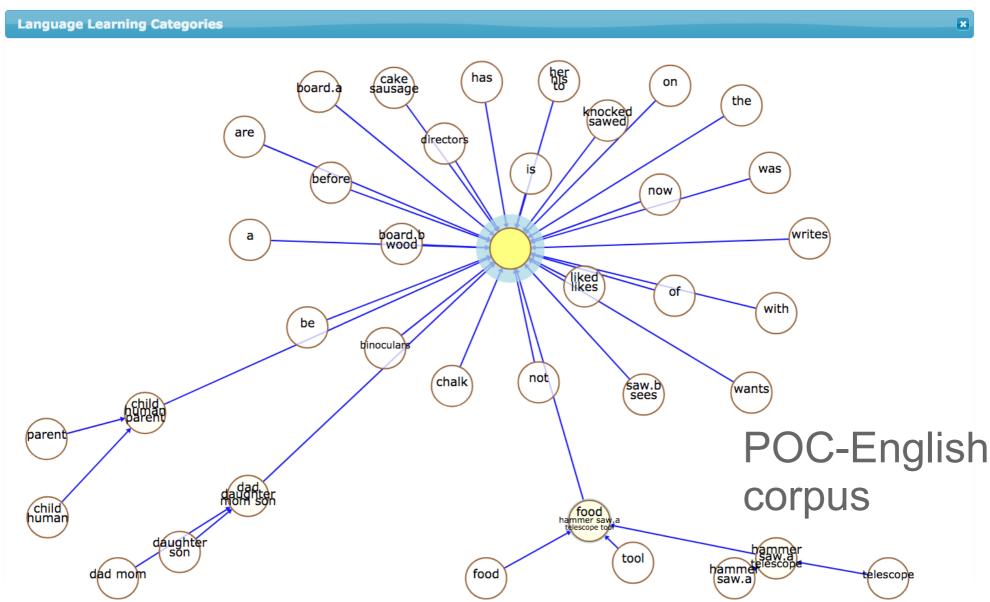


https://github.com/singnet/language-learning

https://www.youtube.com/watch?v=cwgtcOfA3KI

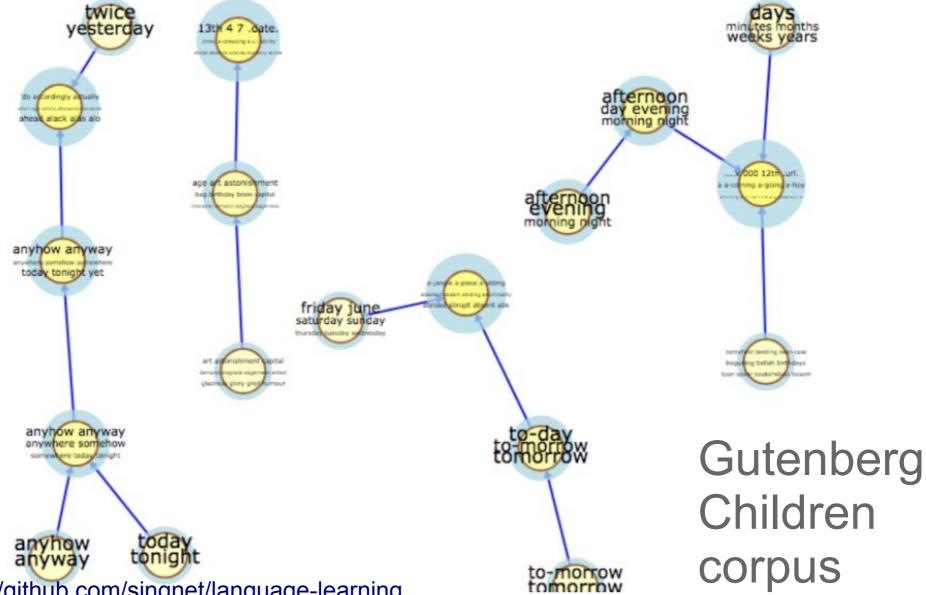
https://www.springerprofessional.de/en/programmatic-link-grammar-induction-for-unsupervised-language-le/17020348

Grammar Ontology from Parses





Learning Ontology from Parses

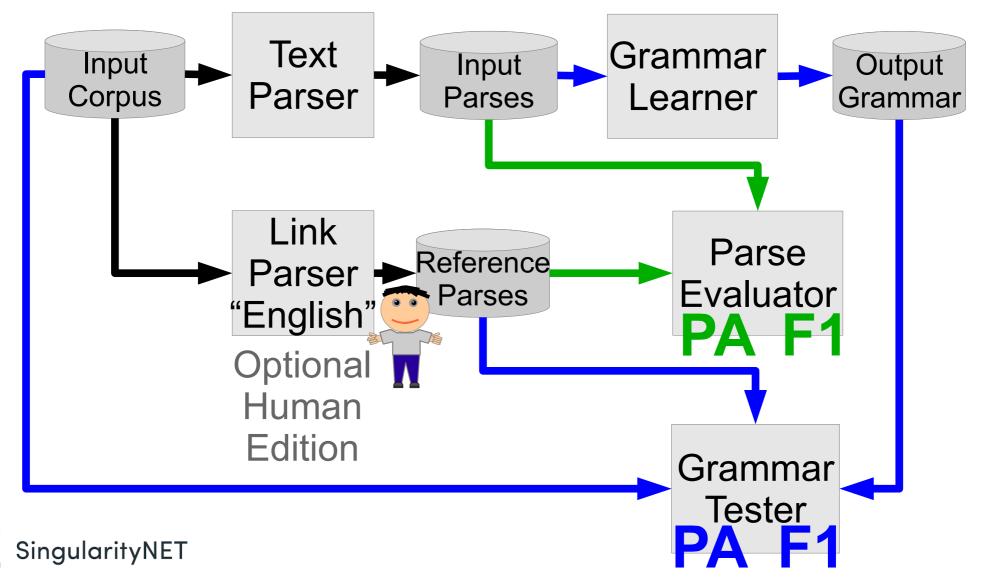


https://github.com/singnet/language-learning https://www.youtube.com/watch?v=cwgtcOfA3KI

https://www.springerprofessional.de/en/programmatic-link-grammar-induction-for-unsupervised-language-le/17020348

Copyright © 2022 Anton Kolonin, Aigents®, SingularityNET Foundation, Auitonio Foundation

Quality-Assessment with on Parses and Grammar



F1 Results Across the Corpora

Corpus	Parses	Parses F1	Clustering	Parse- Ability	Grammar F1
POC-English	Manual	1.00	ILE	100%	1.00
POC-English	Manual	1.00	ALE-400	100%	1.00
POC-English	MST	0.71	ILE	100%	0.72
POC-English	MST	0.71	ALE-400	100%	0.73
Child-Directed Speech	LG-English	1.00	ILE	99%	0.98
Child-Directed Speech	LG-English	1.00	ALE-400	99%	0.97
Child-Directed Speech	MST	0.68	ILE	71%	0.45
Child-Directed Speech	MST	0.68	ALE-400	82%	0.50
Gutenberg Children	LG-English	1.00	ILE	63%	0.65
Gutenberg Children	LG-English	1.00	ALE-500	69%	0.66
Gutenberg Children	MST	0.52	ILE	93%	0.50
Gutenberg Children	MST	0.52	ALE-500	99%	0.53

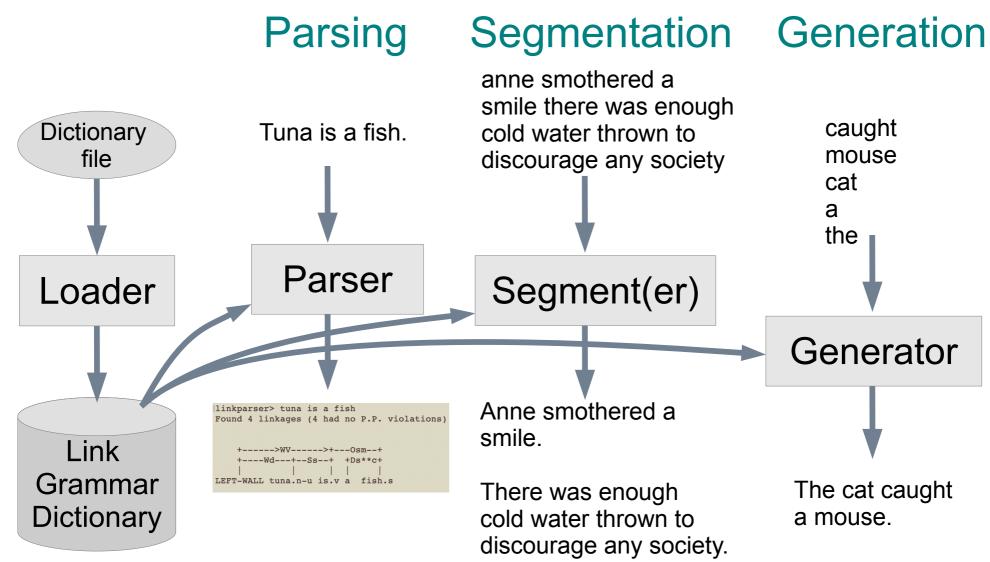
https://www.springerprofessional.de/unsupervised-language-learning-in-opencog/15995030 https://www.springerprofessional.de/en/programmatic-link-grammar-induction-for-unsupervised-language-le/17020348 Copyright © 2022 Anton Kolonin, Aigents®, SingularityNET Foundation, Auitonio Foundation

Conclusions and Next Steps

- Grammars can be induced from parses
 - Better parses => better grammars (Pearson between F1 on parses and F1 on grammar ≥ 0.9)
- MST-Parsing and BERT-Milking can't get parses better than "sequential" ("linked list")
- "Curriculum learning" is a next try for:
 - Parses better than "sequential"
 - Better grammars for larger corpora
 - Incremental Grammar Learning



Link Grammar – beyond parsing



Code: https://github.com/aigents/aigents-java-nlp

Segmentation: https://ieeexplore.ieee.org/document/9303220

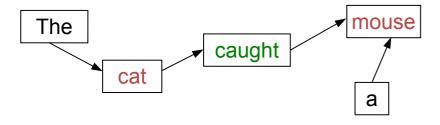
Generation: https://arxiv.org/abs/2105.00830

Grammatical Generation - Methodology

Generator determines what sentences can be formed from a given list of words via valid Link Grammar rules:

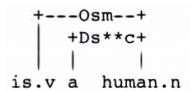
- 1) Given a list of words, the Generator determines a subset of all orderings of those words that satisfies initial checks of the planarity and connectivity metarules.
- 2) For each ordering in the subset, the Generator determines if that ordering is valid; specifically, it ensures that every pair of consecutive words can be connected via links part of the Dictionary objects. To do so, the Generator uses the connects() function, which returns a boolean value indicating whether its two parameters left and right can be linked together.

Planarity metarule: links do not cross Connectivity metarule: links and words of a sentence must form a connected graph that can be completely traversed via one path



```
Algorithm 2: CONNECTS
 Input: A pair of strings left and right, representing the two words to potentially be
 Output: An boolean value indicating whether left and right can be connected via valid
          Link Grammar rules
 Obtain leftList, the list of rules corresponding with left (i.e. the rule when left is a verb,
  the rule when left is a gerund, etc.), from the global Dictionary variables dict and
  hyphenated
 Obtain rightList in a similar manner
 for leftRule in leftList do
     for rightRule in rightList do
         Split leftRule and rightRule into lists of Disjuncts ld and rd
         for l in ld do
            for r in rd do
                Replace all instances of '-' in l with '+' and vice versa
                if l = r then
                 return true
                else
                 continue
                end
            end
        end
     end
 end
 return false
```

connects () is not always applicable. For instance, when the determiner "a" is present in the phrase "is a human," the links are not "is" \rightarrow "a" and "a" \rightarrow "human" but rather "is" \rightarrow "human" and "a" \rightarrow "human" as shown in this Link Grammar parse:



https://arxiv.org/abs/2105.00830

Grammatical Generation - Results

Our algorithm was primarily tested on 92 sentences with words all part of SingularityNET's "small world" POC-English corpus. For this purpose, we have used the Link Grammar dictionary (automatically inferred from high quality Link Grammar parses created by SingularityNET's ULL pipeline) containing 42 total words and 5 total word clusters.

When tested on the same 92 sentences while using the complete Link Grammar dictionary for English, the algorithm achieved the following results. The decrease in "Single correct generated sentence" and increase in "Multiple sentences with one correct" is a direct result of the increased grammatical and semantic ambiguity from using Link Grammar instead of "small world" grammar. Since the "small world" grammar was created from the "small world" corpus itself, each of the words in the corpus contains only a subset of the grammatical or semantic contexts that Link Grammar does.

Our NLG architecture was also tested on 54 sentences part of Charles Keller's production of Lucy Maud Montgomery's "Anne's House of Dreams" as found in the Gutenberg Children corpus and performed as follows.

Metric	Result
Single correct generated sentence	62/92
Multiple sentences with one correct	30/92
Multiple sentences with none correct	0/92
No generated sentences	0/92
Too many results*	0/92
Accuracy	1.000
Total runtime	18 min, 46 sec
Average runtime per sentence	0 min, 12 sec

Metric	Result
Single correct generated sentence	8/92
Multiple sentences with one correct	57/92
Multiple sentences with none correct	0/92
No generated sentences	0/92
Too many results*	27/92
Accuracy	0.707
Total runtime	115 min, 6 sec
Average runtime per sentence	1 min, 15 sec

^{* &}quot;Too many results" is defined as over 25 generated sentences.

Metric	Result
Single correct generated sentence	1/54
Multiple sentences with one correct	53/54
Multiple sentences with none correct	0/54
No generated sentences	0/54
Accuracy	1.000
Total runtime	141 min, 51 sec
Average runtime per sentence	2 min, 37 sec

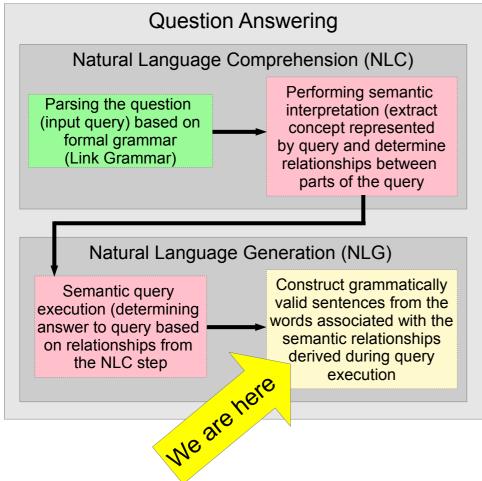
- •POC-English Proof-of-Concept corpus made of artificially selected sentences on limited number of topics ("small world").
- •Gutenberg Children (GC) compendium of books for children contained within Project Gutenberg (https://www.gutenberg.org), following the selection used for the Children's Book Test of the Babi CBT corpus https://research.fb.com/down-loads/babi/

https://github.com/aigents/aigents-java-nlp

https://arxiv.org/abs/2105.00830

Grammatical Generation Applications and Future Work

Applications



https://github.com/aigents/aigents-java-nlp https://arxiv.org/abs/2105.00830

Limitations

Grammatical ambiguity: same word may have different roles in a sentence

```
"I saw the saw."
```

First "saw" – verb, second "saw" – noun (different sets of grammar rules for each instance of "saw" – semantic/word sense disambiguation)

Subject-object ambiguity: a specific case of grammatical ambiguity which refers to the potential interchangeability of the subject and object in a sentence

```
["mouse", "a", "the", "caught", "cat"]
```

Result 1: "The cat caught a mouse."
Result 2: "The mouse caught a cat."

Both results are grammatically valid, but "The cat caught a mouse" is more contextually valid. Implementing grammatical and semantic disambiguation to solve these issues will be a product of our future work, along with extending the algorithm's generation capabilities to languages other than English (including those that require heavy morphology usage, such as Russian).

Link Grammar – Question Answering

Context: Identity and Relationships

A mom is a human. A dad is a human. A mom is a parent. A dad is a parent. A son is a child. A daughter is a child. A son is a human. A daughter is a human. Mom is a human now. Dad is a human now. Mom is a parent now. Dad is a parent now. Son is a child now. Daughter is a child now. Son is a human now. Daughter is a human now. Mom was a daughter before. Dad was a son before. Mom was not a parent before. Dad was not a parent before.

Question: mom daughter

Answer: [Mom was a daughter before., Daughter was a mom before.]

Question: dad human

Answer: [Dad is a human.]

Table 1. Results when tested on 60 queries from SingularityNET's POC-English corpus.

Metric			Result	ts	
	Ours	BERT	ELECTRA	DistilBERT	RoBERTa
BLEU	0.878	0.639	0.712	0.604	0.767
WVCS	0.944	0.606	0.741	0.595	0.799
WER	0.645	0.924	0.550	1.095	0.150
TER	0.166	0.381	0.342	0.457	0.245

https://www.youtube.com/watch?v=MKIOqO9FRq0

https://link.springer.com/chapter/10.1007/978-3-030-93758-4_22

Interpretable Sentiment Mining and Topic Matching in Aigents®



N-grams for Sentiment



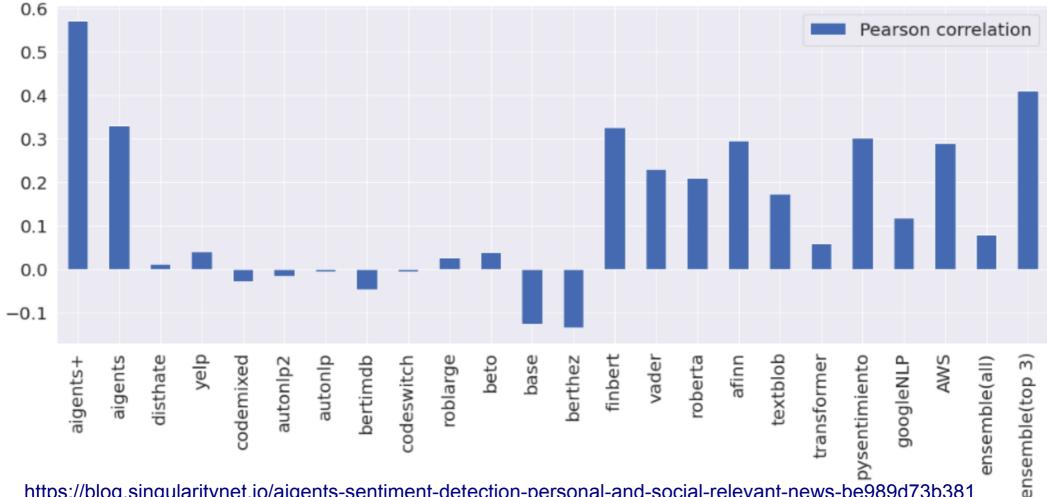


https://blog.singularitynet.io/aigents-sentiment-detection-personal-and-social-relevant-news-be989d73b381

Sentiment Analysis – Models' Fight

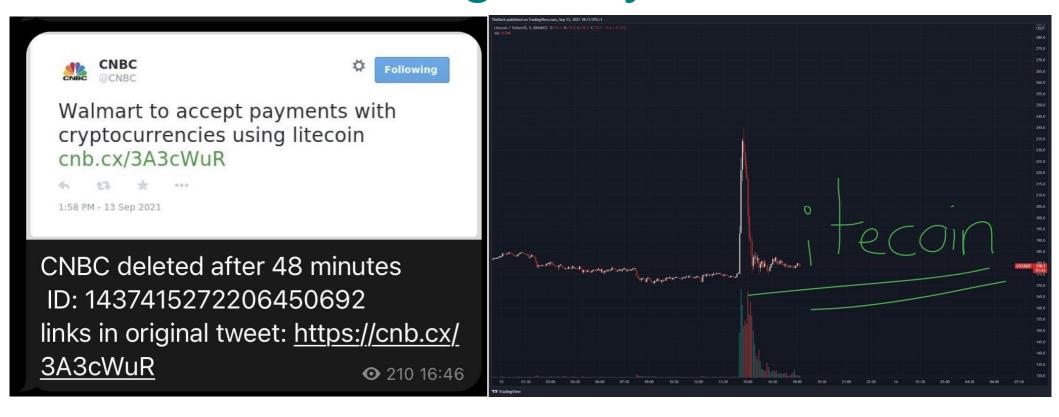
Out-of-the-box Aigents® "interpretable" model competes with Bert model fine-tuned on financial data

Average correlation across all models



https://blog.singularitynet.io/aigents-sentiment-detection-personal-and-social-relevant-news-be989d73b381

Sentiment and Behavioral Patterns Mining in Autonio Foundation and SingularityDAO

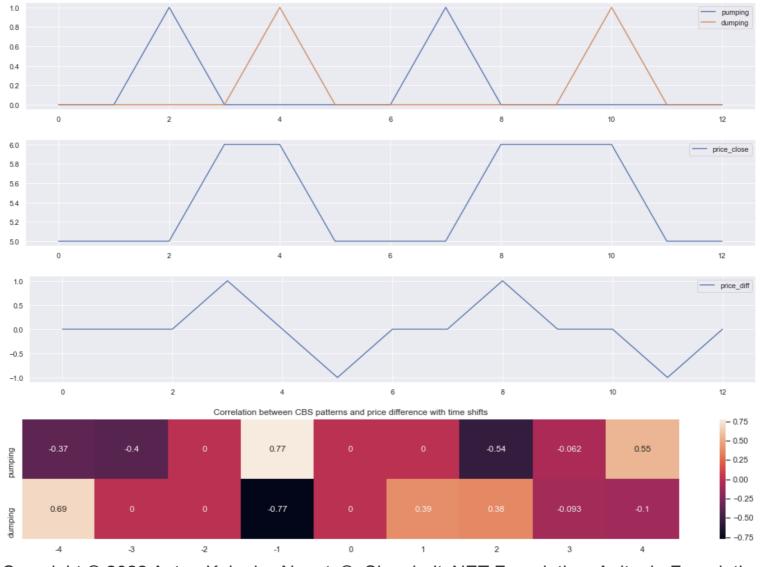




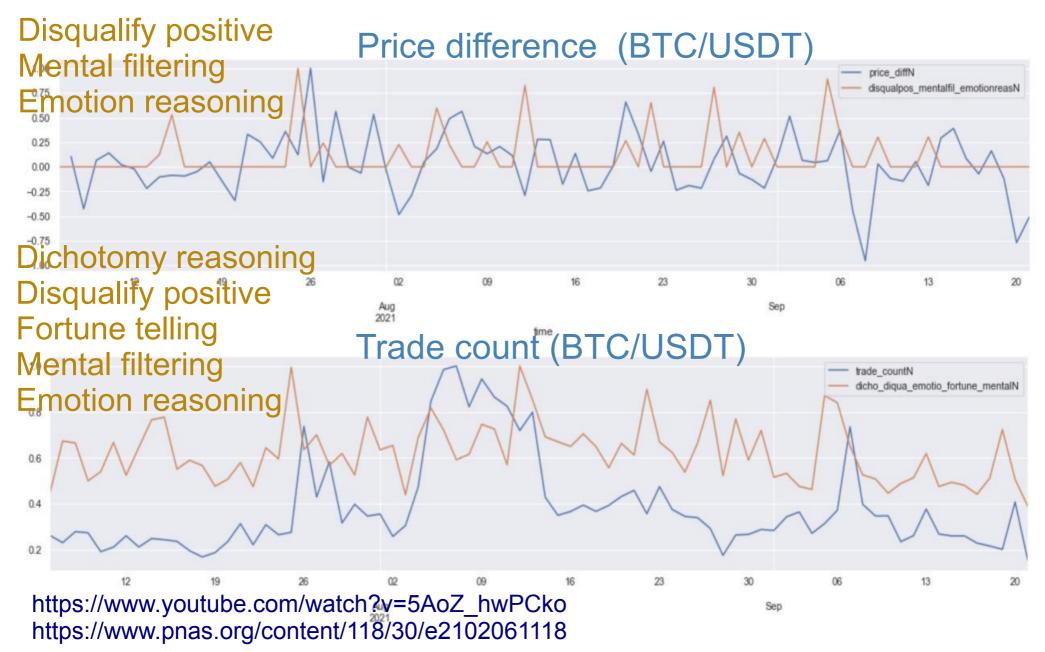


Sentiment for Market Price

Reverse temporal correlation as causal connection between sentiment patterns and price change



Cognitive Distortions for FinTech



Sentiment for Market Price

Tracking Twitter and Reddit sentiment for BTC price

medifficion, pos 0.065 0.19 0.12 0.045 0.13 0.075 0.13 0.095 0.102 0.011 0.006 0.011 0.005 0.024 0.14 0.005 0.024 0.15 0.015 0.13 0.16 0.022 0.016 0.31 0.17 0.0029 0.081 meddiff-Crypto/Deformery, pos 40.03 0.082 0.28 0.02 0.0045 0.28 0.02 0.0045 0.22 0.098 0.024 0.012 0.14 0.005 0.088 0.018 0.13 0.12 0.049 0.0045 0.1 0.074 0.08 0.24 0.012 0.14 0.007 0.007 0.007 0.007 0.007 0.007 0.007 0.007 0.007 0.007 0.007 0.007 0.008 0.022 0.007 0.007 0.007 0.007 0.008 0.027 0.013 0.007 0.007 0.008 0.027 0.013 0.007 0.007 0.008 0.027 0.007 0.008					(Correlation	between ov	erall sentir	ment (pos) a	and price dif	ference wit	th time shift	S			
meddit/CryptoMoonShots_pos	reddit/Bitcoin_pos	-0.065	0.19	-0.12	-0.0045	-0.13	-0.079	-0.13	-0.052	0.029	-0.017	0.14	-0.07	0.076	-0.011	-0.064
meddit/CryptoMoonShols, pos	reddit/btc_pos	0.14	-0.055	0.24	0.14	0.038	0.15	0.13	0.16	0.026	0.18	0.16	0.31	0.17	-0.0029	0.081
reddiff/degecoin_pos	reddit/CryptoCurrency_pos	-0.093	-0.082	-0.26	0.15	-0.054	-0.00062	-0.18	0.095	0.22	-0.096	0.28	-0.02	-0.0045	-0.26	-0.2
reddff/ethereum_pos 0.18 0.22 0.26 0.036 0.13 0.055 0.12 0.14 0.0066 0.39 0.32 0.38 0.27 0.28 0.22 0.25 0.26 0.047 0.16 0.28 0.29 0.2 0.17 0.088 0.2 0.054 0.42 0.53 0.099 0.17 0.17 0.019 0.14 0.016 0.097 0.13 0.013 0.16 0.12 0.2 0.35 0.26 0.26 0.000 0.000 0.17 0.000 0.0	reddit/CryptoMoonShots_pos	-0.24	0.055	0.068	-0.18	0.13	0.12	-0.049	0.0045	0.1	0.074	-0.08	-0.24	-0.012	-0.14	-0.017
Beddil/Ethtrader_pos 0.66	reddit/dogecoin_pos	0.57	-0.42	-0.13	-0.15	0.75	-0.38	-0.52	0.52	-0.18	-0.13	-0.028	0.24	0.42	-0.35	-0.13
meddil/SatoshiStreetBets	reddit/ethereum_pos	0.18	0.22	0.26	-0.036	-0.13	-0.055	0.12	0.14	0.0066	-0.39	-0.32	-0.38	-0.27	-0.28	-0.22
witter/alantonop_os 0.052 0.077 0.17 0.019 0.14 -0.16 -0.07 0.055 -0.091 0.16 0.011 -0.13 -0.061 -0.26 0.038 witter/adam3us_pos 0.019 0.047 0.0015 0.063 -0.1 -0.4 0.13 0.077 -0.069 -0.046 0.16 0.17 0.14 0.17 0.15 witter/AndreCroripTech_pos -0.22 0.953 -0.014 0.1 0.23 0.13 0.056 -0.089 0.065 0.01 -0.16 0.064 0.025 -0.055 0.08 witter/Britocin_pos 0.08 -0.28 0.29 -0.36 0.076 -0.091 0.17 -0.01 -0.03 0.14 0.014 -0.02 -0.085 witter/Bitocin_pos 0.076 -0.033 -0.32 -0.21 -0.034 0.1 -0.22 -0.11 -0.08 -0.05 -0.14 0.01 -0.02 -0.03 0.01 -0.03 -0.01 -0.03 -0.04 <td< td=""><td>reddit/ethtrader_pos</td><td>-0.66</td><td>-0.26</td><td>-0.047</td><td>-0.16</td><td>0.28</td><td>0.29</td><td>-0.2</td><td>-0.17</td><td>0.088</td><td>0.2</td><td>-0.54</td><td>0.42</td><td>-0.53</td><td>0.099</td><td>-0.17</td></td<>	reddit/ethtrader_pos	-0.66	-0.26	-0.047	-0.16	0.28	0.29	-0.2	-0.17	0.088	0.2	-0.54	0.42	-0.53	0.099	-0.17
witter/adam3us_pos 0.019 -0.047 -0.015 0.063 -0.1 -0.4 0.13 -0.077 -0.069 -0.046 0.16 0.17 -0.14 0.17 0.15 witter/AndreCronjeTech_pos -0.22 0.95 0.34 -0.059 0.45 0.59 0.24 -0.99 -0.86 0.29 0.85 0.22 -0.85 0.18 0.12 witter/Apmplianc_pos 0.022 -0.033 0.14 0.1 0.23 0.13 0.056 -0.089 0.065 -0.01 -0.16 0.064 0.025 -0.055 witter/Bitcoin_pos 0.05 0.058 -0.22 -0.043 0.1 -0.24 -0.24 -0.44 -0.044 -0.017 -0.044 -0.017 -0.044 -0.007 -0.044 -0.053 -0.053 -0.022 -0.036 -0.044 -0.051 -0.044 -0.053 -0.044 -0.053 -0.044 -0.053 -0.024 -0.024 -0.03 -0.014 -0.076 -0.03 -0.024 -0.035 <td>reddit/SatoshiStreetBets_pos</td> <td>-0.12</td> <td>0.04</td> <td>-0.32</td> <td>-0.067</td> <td>-0.1</td> <td>0.095</td> <td>0.097</td> <td>-0.13</td> <td>0.013</td> <td>0.16</td> <td>0.12</td> <td>0.2</td> <td>0.35</td> <td>0.26</td> <td>0.2</td>	reddit/SatoshiStreetBets_pos	-0.12	0.04	-0.32	-0.067	-0.1	0.095	0.097	-0.13	0.013	0.16	0.12	0.2	0.35	0.26	0.2
twitter/Andre Cronje Tech_Dos -0.22 0.95 0.94 -0.059 0.45 0.59 0.24 -0.99 -0.86 0.29 0.85 0.22 -0.85 0.18 0.12 twitter/Aprompliano_pos -0.22 -0.053 -0.14 0.1 0.23 0.13 0.056 -0.089 0.065 -0.01 0.16 0.064 0.025 -0.05 -0.085 twitter/Birance_pos 0.059 -0.058 -0.2 -0.043 0.1 0.22 -0.17 -0.11 -0.085 -0.05 -0.14 0.014 -0.023 -0.13 0.14 0.014 -0.023 -0.16 0.004 -0.05 -0.04 0.014 -0.023 -0.16 0.022 -0.014 -0.014 -0.07 0.033 -0.22 -0.035 0.048 0.15 0.11 -0.038 -0.12 -0.033 -0.027 -0.11 -0.06 0.16 0.15 -0.11 -0.08 0.076 0.082 -0.87 -0.91 -0.6 0.15 -0.1	twitter/aantonop_pos	-0.052	0.077	0.17	0.019	0.14	-0.16	-0.07	0.055	-0.091	0.16	0.011	-0.13	-0.061	-0.26	0.038
twitter/APompiano_pos 0.22 0.053 0.14 0.1 0.23 0.13 0.056 0.089 0.065 -0.01 -0.16 0.064 0.025 -0.05 -0.085 witter/Darrysilbert_pos 0.08 -0.28 0.29 -0.36 0.076 -0.091 0.17 0.24 0.45 0.44 0.17 0.0087 0.26 0.0023 0.13 witter/Bircoin_pos 0.059 -0.058 -0.2 -0.043 0.1 0.22 -0.17 -0.11 -0.084 0.055 -0.05 -0.14 0.014 -0.023 -0.16 witter/Bircoin_pos 0.016 -0.28 -0.033 -0.12 -0.034 -0.19 -0.03 -0.14 -0.038 -0.12 -0.14 0.07 -0.053 witter/Bircoin_pos 0.76 0.82 -0.87 -0.91 -0.6 0.97 0.64 -0.15 -0.41 0.58 0.95 0.61 0.99 -0.03 witter/Coin_pos 0.14 0.12 0.31 0.	twitter/adam3us_pos	0.019	-0.047	-0.0015	0.063	-0.1	-0.4	0.13	-0.077	-0.069	-0.046	0.16	0.17	-0.14	0.17	0.15
bwitter/barrysilbert_pos 0.08 0.28 0.29 4.036 0.076 -0.091 0.17 0.24 0.45 0.44 0.17 0.0087 0.26 0.0023 0.13 Witter/Biroince_pos 0.059 -0.058 -0.2 -0.043 0.1 0.22 -0.17 -0.11 -0.084 0.055 -0.05 -0.14 0.014 -0.023 -0.16 Witter/Bitcoin pos 0.016 -0.26 -0.033 -0.32 -0.21 -0.24 -0.27 -0.19 -0.03 -0.14 0.038 -0.12 0.14 0.07 0.053 Witter/Bitcoin pos -0.075 0.22 -0.038 0.12 -0.055 0.048 0.15 0.11 0.36 0.26 0.16 0.28 -0.076 0.03 -0.027 0.011 0.06 0.06 0.06 0.097 0.64 -0.15 -0.41 0.58 0.95 0.61 0.99 0.06 0.074 Witter/Coinbase_pos 0.15 0.13	twitter/AndreCronjeTech_pos	-0.22	0.95	0.94	-0.059	0.45	0.59	0.24	-0.99	-0.86	0.29	0.85	0.22	-0.85	0.18	0.12
twitter/binance_pos 0.059 -0.058 -0.2 -0.043 0.1 0.22 -0.17 -0.11 -0.084 0.055 -0.05 -0.14 0.014 -0.023 -0.16 twitter/Bitcoin_pos 0.016 -0.26 -0.033 -0.32 -0.21 -0.24 -0.27 -0.19 -0.03 -0.14 0.007 0.053 twitter/BitcoinMagazine_pos 0.075 0.22 -0.038 0.12 -0.035 0.048 0.15 0.11 0.36 0.26 0.16 0.28 -0.076 0.03 -0.027 twitter/CoinMagazine_pos 0.76 0.82 -0.87 -0.91 -0.6 0.15 0.11 0.36 0.26 0.16 0.02 -0.011 0.01 0.01 0.01 0.01 0.01 0.02 0.01 0.02 0.01 0.01 0.01 0.01 0.01 0.02 0.01 0.01 0.01 0.00 0.01 0.01 0.00 0.02 0.01 0.01 0.01 <td>twitter/APompliano_pos</td> <td>-0.22</td> <td>-0.053</td> <td>-0.14</td> <td>0.1</td> <td>0.23</td> <td>0.13</td> <td>0.056</td> <td>-0.089</td> <td>0.065</td> <td>-0.01</td> <td>-0.16</td> <td>0.064</td> <td>0.025</td> <td>-0.05</td> <td>-0.085</td>	twitter/APompliano_pos	-0.22	-0.053	-0.14	0.1	0.23	0.13	0.056	-0.089	0.065	-0.01	-0.16	0.064	0.025	-0.05	-0.085
britter/Bitcoin_pos 0.016 -0.26 -0.033 -0.32 -0.21 -0.24 -0.27 -0.19 -0.03 -0.14 0.038 -0.12 0.14 0.07 0.053 witter/Bitcoin/Magazine_pos -0.075 0.22 -0.038 0.12 -0.035 0.048 0.15 0.11 0.36 0.26 0.16 0.28 -0.076 0.03 -0.027 twitter/Brian_armstrong_pos 0.76 0.82 -0.87 -0.91 -0.6 0.97 0.64 -0.15 -0.41 0.58 0.95 0.61 0.9 0.96 0.074 witter/BTCTN_pos 0.14 0.12 0.31 0.18 -0.06 0.16 0.15 -0.2 -0.011 -0.011 0.005 0.027 -0.011 -0.13 witter/Dinbase_pos 0.15 0.13 0.29 0.4 -0.35 0.22 -0.65 -0.78 -0.27 0.37 0.57 0.31 0.03 0.02 -0.15 witter/Coinpase_pos 0.22 -	twitter/barrysilbert_pos	0.08	-0.28	0.29	-0.36	0.076	-0.091	0.17	0.24	0.45	0.44	0.17	0.0087	0.26	0.0023	0.13
witter/BitcoinMagazine pos -0.075 0.22 -0.038 0.12 -0.035 0.048 0.15 0.11 0.36 0.26 0.16 0.28 -0.076 0.03 -0.027 witter/Brian_armstrong_pos 0.76 0.82 -0.87 -0.91 -0.6 0.97 0.64 -0.15 -0.41 0.58 0.95 0.61 0.9 0.96 0.074 witter/Coinbase_pos -0.15 0.13 0.29 0.4 -0.35 0.22 -0.65 -0.78 -0.27 0.37 0.57 0.31 0.37 -0.89 -0.15 witter/CoinDesk_pos -0.22 -0.2 -0.088 -0.23 -0.23 -0.13 0.048 -0.075 0.083 0.09 -0.31 -0.06 -0.044 -0.03 0.12 0.24 -0.71 0.14 -0.38 -0.11 -0.45 0.086 0.29 -0.11 0.071 0.12 witter/CoinDesk_pos -0.21 -0.098 -0.17 -0.058 -0.067 -0.097 -	twitter/binance_pos	0.059	-0.058	-0.2	-0.043	0.1	0.22	-0.17	-0.11	-0.084	0.055	-0.05	-0.14	0.014	-0.023	-0.16
witter/brian_armstrong_pos 0.76 0.82 0.87 0.91 0.6 0.97 0.64 -0.15 -0.41 0.58 0.95 0.61 0.9 0.96 0.074 witter/EDTCTN_pos 0.14 0.12 0.31 0.18 -0.06 0.16 0.15 -0.2 -0.011 -0.11 0.0058 -0.16 0.027 -0.011 -0.13 witter/coinbase_pos -0.15 0.13 0.29 0.4 -0.35 0.22 -0.65 -0.78 -0.27 0.37 0.57 0.31 0.37 -0.89 -0.15 witter/coinpecko_pos -0.22 -0.2 -0.088 -0.23 -0.23 -0.13 0.048 -0.075 0.083 0.09 -0.31 -0.06 -0.024 -0.27 0.13 0.046 -0.024 -0.32 -0.1 -0.14 -0.088 -0.11 -0.041 -0.048 -0.11 -0.04 -0.024 -0.22 -0.1 0.11 -0.024 -0.024 -0.22 -0.11 -0.02	twitter/Bitcoin_pos	0.016	-0.26	-0.033	-0.32	-0.21	-0.24	-0.27	-0.19	-0.03	-0.14	0.038	-0.12	0.14	0.07	0.053
hwitter/BTCTN_pos 0.14 0.12 0.31 0.18 -0.06 0.16 0.15 -0.2 -0.011 -0.11 0.0058 -0.16 0.027 -0.011 -0.13 hwitter/coinbase_pos -0.15 0.13 0.29 0.4 -0.35 0.22 -0.65 -0.78 -0.27 0.37 0.57 0.31 0.37 -0.89 -0.15 hwitter/coinbase_pos -0.22 -0.2 -0.088 -0.23 -0.23 -0.13 0.048 -0.075 0.083 0.09 -0.31 -0.06 -0.024 -0.32 -0.1 hwitter/coindwarketCap_pos -0.31 0.24 -0.063 0.12 0.24 0.71 0.14 -0.38 -0.11 -0.45 0.086 0.29 -0.11 0.071 0.12 hwitter/CoinMarketCap_pos -0.21 -0.098 -0.17 -0.058 -0.067 -0.0097 -0.13 -0.14 -0.056 0.048 -0.12 -0.2 0.11 0.01 -0.008 -0.11 -0.03	twitter/BitcoinMagazine_pos	-0.075	0.22	-0.038	0.12	-0.035	0.048	0.15	0.11	0.36	0.26	0.16	0.28	-0.076	0.03	-0.027
twitter/coinbase_pos -0.15 0.13 0.29 0.4 -0.35 0.22 -0.65 -0.78 -0.27 0.37 0.57 0.31 0.37 -0.89 -0.15 twitter/CoinDesk_pos -0.22 -0.2 -0.088 -0.23 -0.23 -0.13 0.048 -0.075 0.083 0.09 -0.31 -0.06 -0.024 -0.1 twitter/CoinDesk_pos -0.31 0.24 -0.063 0.12 0.24 0.71 0.14 -0.38 -0.11 -0.45 0.086 0.29 -0.11 0.071 0.12 twitter/CoinMarketCap_pos -0.21 -0.098 -0.17 -0.058 -0.067 -0.0097 -0.13 -0.14 -0.056 0.048 -0.12 -0.12 0.11 -0.34 twitter/CoinMarketCap_pos -0.21 -0.098 -0.11 -0.13 -0.15 -0.083 -0.15 -0.088 0.2 -0.13 -0.11 -0.12 0.11 -0.14 -0.02 -0.083 -0	twitter/brian_armstrong_pos	0.76	0.82	-0.87	-0.91	-0.6	0.97	0.64	-0.15	-0.41	0.58	0.95	0.61	0.9	0.96	0.074
twitter/CoinDesk_pos -0.22 -0.2 -0.088 -0.23 -0.13 0.048 -0.075 0.083 0.09 -0.31 -0.06 -0.024 -0.32 -0.1 twitter/CoinDesk_pos -0.31 0.24 -0.063 0.12 0.24 0.71 0.14 -0.38 -0.11 -0.45 0.086 0.29 -0.11 0.071 0.12 twitter/CoinMarketCap_pos -0.21 -0.098 -0.17 -0.058 -0.067 -0.0097 -0.13 -0.14 -0.056 0.048 -0.12 -0.2 -0.12 0.11 -0.34 twitter/Cointelegraph_pos 0.073 -0.098 -0.1 0.13 -0.15 -0.083 -0.15 0.038 0.2 -0.13 -0.11 -0.25 0.17 0.19 -0.063 twitter/cz_binance_pos -0.2 -0.28 0.26 -0.32 0.018 0.14 0.17 -0.16 -0.0067 -0.019 0.061 -0.26 -0.46 0.021 0.037	twitter/BTCTN_pos	0.14	0.12	0.31	0.18	-0.06	0.16	0.15	-0.2	-0.011	-0.11	0.0058	-0.16	0.027	-0.011	-0.13
twitter/coingecko_pos -0.31 0.24 -0.063 0.12 0.24 0.71 0.14 -0.38 -0.11 -0.45 0.086 0.29 -0.11 0.071 0.12 twitter/CoinMarketCap_pos -0.21 -0.098 -0.17 -0.058 -0.067 -0.0097 -0.13 -0.14 -0.056 0.048 -0.12 -0.2 -0.12 0.11 -0.34 twitter/Cointelegraph_pos 0.073 -0.098 -0.1 0.13 -0.15 -0.083 -0.15 0.038 0.2 -0.13 -0.11 -0.25 0.17 0.19 -0.036 twitter/capinance_pos -0.2 -0.28 0.26 -0.32 0.018 0.14 0.17 -0.16 -0.0067 -0.019 0.061 -0.26 -0.46 0.021 0.037 twitter/capinance_pos -0.23 -0.76 0.18 0.93 -0.51 0.091 0.88 0.25 0.96 0.73 0.98 -0.061 0.31 1	twitter/coinbase_pos	-0.15	0.13	0.29	0.4	-0.35	0.22	-0.65	-0.78	-0.27	0.37	0.57	0.31	0.37	-0.89	-0.15
twitter/CoinMarketCap_pos	twitter/CoinDesk_pos	-0.22	-0.2	-0.088	-0.23	-0.23	-0.13	0.048	-0.075	0.083	0.09	-0.31	-0.06	-0.024	-0.32	-0.1
twitter/Cointelegraph_pos 0.073 0.0098 0.1 0.13 0.15 0.083 0.15 0.038 0.2 0.13 0.11 0.25 0.17 0.19 0.036 twitter/cz_binance_pos 0.2 0.28 0.26 -0.32 0.018 0.14 0.17 -0.16 -0.0067 -0.019 0.061 -0.26 -0.46 0.021 0.037 twitter/leanmusk_pos 0.23 0.76 -0.75 0.18 0.93 -0.51 0.091 0.88 0.25 0.96 0.73 0.98 -0.061 0.31 1 twitter/ErikVoorhees_pos 0.15 0.085 -0.17 -0.14 0.4 0.0099 -0.2 0.15 -0.15 -0.14 -0.23 0.025 -0.37 0.06 -0.013 twitter/maxkeiser_pos -0.062 0.16 0.033 0.094 0.041 -0.091 -0.02 -0.23 0.15 0.22 -0.019 0.27 0.17 -0.00036 -0.17 twitter/naval_pos	twitter/coingecko_pos	-0.31	0.24	-0.063	0.12	0.24	0.71	0.14	-0.38	-0.11	-0.45	0.086	0.29	-0.11	0.071	0.12
twitter/cz_binance_pos	twitter/CoinMarketCap_pos	-0.21	-0.098	-0.17	-0.058	-0.067	-0.0097	-0.13	-0.14	-0.056	0.048	-0.12	-0.2	-0.12	0.11	-0.34
twitter/elonmusk_pos	twitter/Cointelegraph_pos	0.073	-0.0098	-0.1	0.13	-0.15	-0.083	-0.15	0.038	0.2	-0.13	-0.11	-0.25	0.17	0.19	-0.036
twitter/ErikVoorhees_pos 0.15 0.085 -0.17 -0.14 0.4 0.0099 -0.2 0.15 -0.15 -0.14 -0.23 0.025 -0.37 0.06 -0.013 twitter/maxkeiser_pos -0.062 0.16 0.033 0.094 0.041 -0.091 -0.02 -0.23 0.15 0.22 -0.019 0.27 0.17 -0.00036 -0.17 twitter/naval_pos -0.036 0.46 -0.12 0.35 0.77 0.23 0.67 0.46 -0.082 0.13 -0.41 -0.62 -0.66 0.65 0.15 twitter/RaoulGMI_pos -0.25 -0.088 0.041 -0.39 -0.031 -0.14 0.24 0.32 -0.048 -0.11 -0.0011 -0.0067 0.16 0.045 -0.24 twitter/SatoshiLite_pos -0.4 -0.21 -0.52 -0.22 0.091 0.25 0.078 -0.19 -0.76 0.66 0.27 0.3 0.56 0.62 0.027 twitter/VitalikButerin_pos -1 1 1 1 -1 -1 1 1 1 1 1 1 1 1 1 1 1 1	twitter/cz_binance_pos	-0.2	-0.28	0.26	-0.32	0.018	0.14	0.17	-0.16	-0.0067	-0.019	0.061	-0.26	-0.46	0.021	0.037
twitter/maxkeiser_pos	twitter/elonmusk_pos	-0.23	-0.76	-0.75	0.18	0.93	-0.51	0.091	0.88	0.25	0.96	0.73	0.98	-0.061	0.31	1
twitter/naval_pos	twitter/ErikVoorhees_pos	0.15	0.085	-0.17	-0.14	0.4	0.0099	-0.2	0.15	-0.15	-0.14	-0.23	0.025	-0.37	0.06	-0.013
twitter/RaoulGMI_pos	twitter/maxkeiser_pos	-0.062	0.16	0.033	0.094	0.041	-0.091	-0.02	-0.23	0.15	0.22	-0.019	0.27	0.17	-0.00036	-0.17
twitter/SatoshiLite_pos -0.4 -0.21 -0.52 -0.22 0.091 0.25 0.078 -0.19 -0.76 0.66 0.27 0.3 0.56 0.62 0.027 twitter/VitalikButerin_pos -1 1 1 -1 1 1 -1 1	twitter/naval_pos	-0.036	0.46	-0.12	0.35	0.77	0.23	0.67	0.46	-0.082	0.13	-0.41	-0.62	-0.66	0.65	0.15
twitter/VitalikButerin_pos -1 1 1 -1 -1 1 1 -1 -1 1 1 1 1 -1 -1 1	twitter/RaoulGMI_pos	-0.25	-0.088	0.041	-0.39	-0.031	-0.14	0.24	0.32	-0.048	-0.11	-0.0011	-0.0067	0.16	0.045	-0.24
	twitter/SatoshiLite_pos	-0.4	-0.21	-0.52	-0.22	0.091	0.25	0.078	-0.19	-0.76	0.66	0.27	0.3	0.56	0.62	0.027
	twitter/VitalikButerin_pos	-1	1	1	-1	-1	1	1	-1	-1	1	1	1	-1	-1	1
-/ -0 -0 -4 -3 -2 -1 U 1 2 3 4 5 6 /		-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7

CBS patterns for Market Price

Tracking Twitter and Reddit Cognitive Behavioral Schemata patterns for BTC buy/sell volume

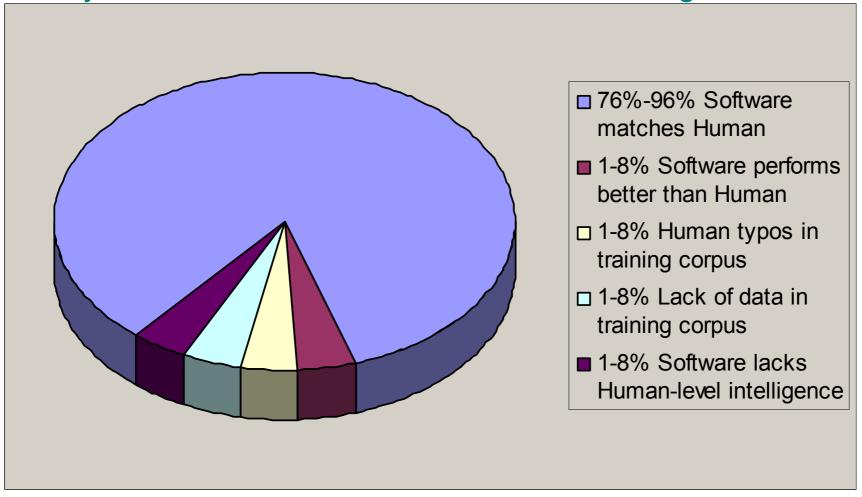
					Correlati	on betweer	n CBS patte	rns and se	ll volume wi	th time shift	ts (days)				
catastrophizing	-0.048	-0.00051	0.062	0.031	0.098	0.29	0.26	0.046	-0.048	0.0064	0.02	0.079	0.015	-0.18	-0.15
dichotoreasoning	-0.35	-0.33	-0.089	-0.015	-0.058	-0.039	0.071	0.11	0.22	0.22	0.34	0.41	0.53	0.32	0.24
disqualpositive	-0.00061	-0.036	-0.027	-0.013	0.12	-0.046	0.21	0.18	0.12	-0.094	-0.031	0.086	-0.051	-0.054	0.1
emotionreasoning	-0.074	-0.094	-0.068	-0.047	-0.068	-0.04	-0.013	0.098	0.22	0.22	0.19	0.26	0.29	0.3	0.31
fortunetelling	-0.04	-0.21	-0.003	0.099	0.097	-0.13	-0.096	-0.064	0.018	-0.062	-0.09	-0.12	0.0043	0.038	0.1
labeling	-0.31	-0.27	-0.13	0.048	-0.081	0.016	0.12	0.13	0.13	0.15	-0.032	0.12	0.083	0.084	-0.13
magnification	-0.072	-0.16	-0.066	-0.016	0.016	0.015	-0.031	0.0075	0.1	0.095	0.046	0.064	0.11	-0.16	-0.22
mentalfiltering	0.032	0.025	0.082	0.016	0.092	0.42	0.31	0.022	-0.055	0.06	0.02	0.041	0.03	0.11	0.072
mindreading	-0.19	-0.25	-0.0032	0.088	-0.018	-0.15	-0.11	-0.17	0.052	-0.054	-0.084	-0.014	0.16	0.083	0.11
overgeneralizing	-0.12	-0.12	-0.03	0.0073	-0.079	-0.1	0.025	0.054	0.088	0.025	0.17	0.11	0.17	0.049	0.081
personalizing	-0.16	-0.086	0.098	0.051	-0.03	0.054	0.023	-0.024	-0.0038	0.094	0.14	0.33	0.38	0.26	0.3
shouldment	-0.21	-0.24	-0.091	0.085	-0.0073	-0.01	-0.0058	-0.016	-0.012	0.13	0.21	0.26	0.3	0.22	0.24
cds	-0.31	-0.33	-0.053	0.058	-0.0038	0.052	0.12	0.062	0.15	0.18	0.2	0.35	0.44	0.25	0.24
	-7	-6	-5	4	-3	-2	-1	0	1	2	3	4	5	6	7

#Catastrophizing: Exaggerating the importance of negative events
distortions['catastrophizing'] = "will fail, will go wrong, will end, will be impossible, will not happen...
#Mental Filtering: Paying too much attention to negative details instead of the whole picture
distortions['mentalfiltering'] = "I see only, all I see, all I can see, can only think, nothing good...

Why do we need Interpretable NLP (AI)?

Errors and blind spots in training data cause incorrect models. Explainable AI enables to locate errors in the models.

Interpretable AI makes it possible to fix the models with no overtraining. Ideally, we would like it to trace it back to training data and fix that...



http://webstructor.net/papers/Kolonin-HP-ACA-IC-text.pdf

Welcome to the Interpretable Natural Language Processing Community and Series of Workshops

Telegram English https://t.me/internlp

Telegram Russian https://t.me/agibots

Workshop 2022 (Seattle, WA, USA & Online) https://aigents.github.io/inlp/

Workshop 2021 https://aigents.github.io/inlp/2021/

Thank You and Welcome!

Anton Kolonin

akolonin@aigents.com

Facebook: akolonin

Telegram: akolonin









autonio.foundation

