

Rob & Rik: Overview of NLP/CL conferences

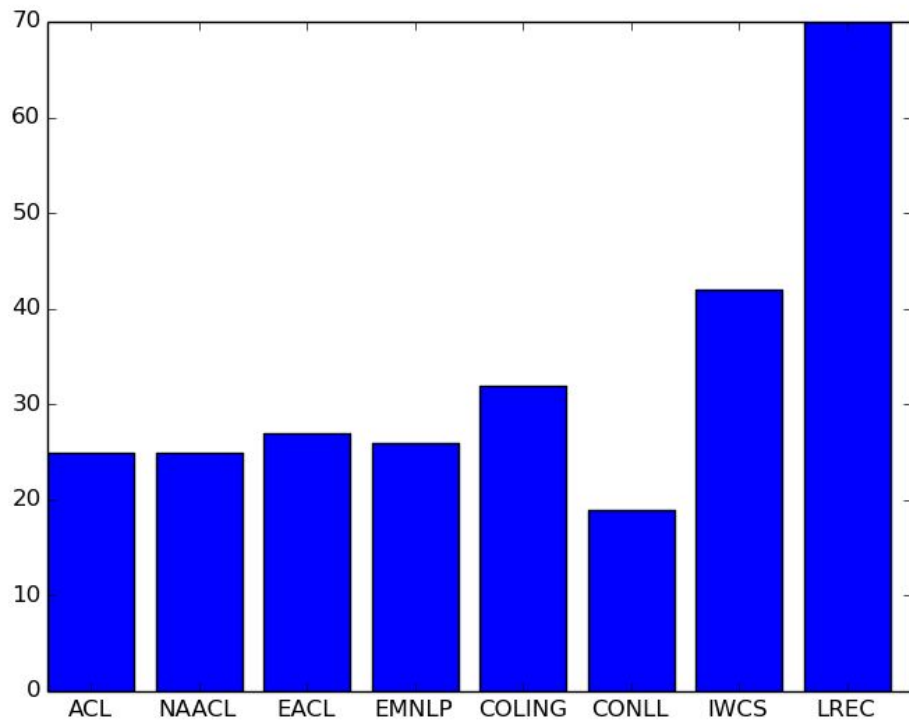


Experiences ACL & EMNLP

Main conferences

- ACL
- NAACL
- EACL
- EMNLP
- COLING
- CONLL
- IWCS
- LREC

Acceptance rates



Local

Tabu Dag (general linguistics): June?

CLIN: Friday 26 January 2018 (Deadline submission: November 3rd, 2017)

Deadlines

LREC: 25 Sept (monday) <https://twitter.com/lrec2018>

CLIN: 3 Nov

NAACL: 15 Dec & 10 Jan

COLING: 16 March

Shared tasks

- SemEval
 - 4 RUG teams competing (Shared Task course)
 - 2 Task 1: Affect in Tweets
 - 2 Task 2: Multilingual Emoji Prediction
- CLEF
- CONLL
- Workshops
- CLIN!

Other useful information

Anthology

<http://www.aclweb.org/anthology/>

Corpora list

<https://mailman.uib.no/listinfo/corpora>

EMNLP 2017



Copenhagen

WNUT workshop

- New (even noisier) domains
- Bias
- Smaller languages/dialects

Interesting Keynotes:

- Miles Osborne
- Dirk Hovy

Shared Task WNUT

- Emerging and Rare entity recognition
- Difficult!



Saturday

- Keynote on robotics
- Syntax sessions
- Embeddings session
- Posters parallel with oral presentations!
- Tobias Horsmann

Sunday

Program seemed less interesting

Sunday



Monday

Keynote Dan Jurafsky



Does This Vehicle Belong to You?

Processing the Language of Policing for Improving Police-Community Relations

Best papers

Bringing Structure into Summaries: Crowdsourcing a Benchmark Corpus of Concept Maps

Natural Language Does Not Emerge ‘Naturally’ in Multi-Agent Dialog

Depression and Self-Harm Risk Assessment in Online Forums

Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints

In general

- Bias (gender, race, etc.)
- Asian languages
- Neural networks, embeddings and neural networks
- Data selection/semi-supervised/auxiliary tasks/multilingual

ACL 2017 Vancouver



July 31st - August 4th

Disclaimer: it's been almost 2 months

Presidential Address Joakim Nivre

- Three topics
 - Diversity
 - arXiv pre-publishing
 - Good methodology in science

- Not much of his own opinion, mainly survey results
- Kind of boring
- Gertjan's picture occurred on the slides twice

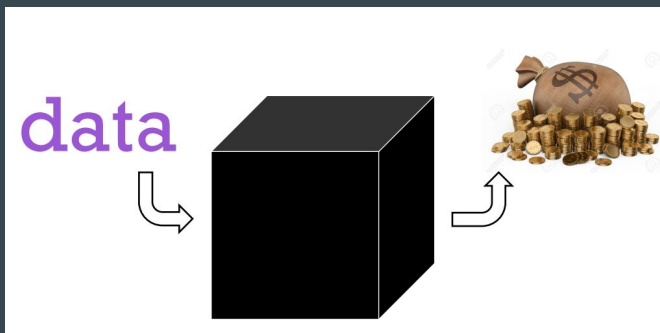


Talk Noah Smith

“Squashing computational linguistics”



- New ways of using representation learning
- Seems to be a fan of deep learning models to “squash” everything
- biLSTM, RNN, etc



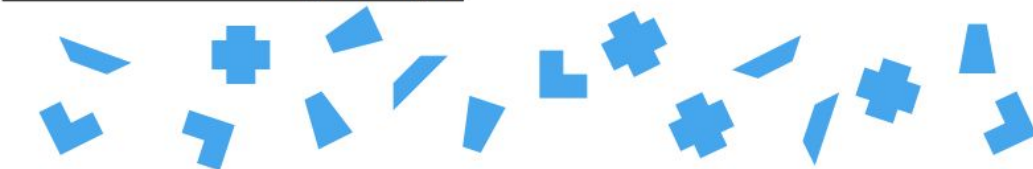
Talk Noah Smith

Linguistic Structure Prediction



training objective

output (structure)



part representations



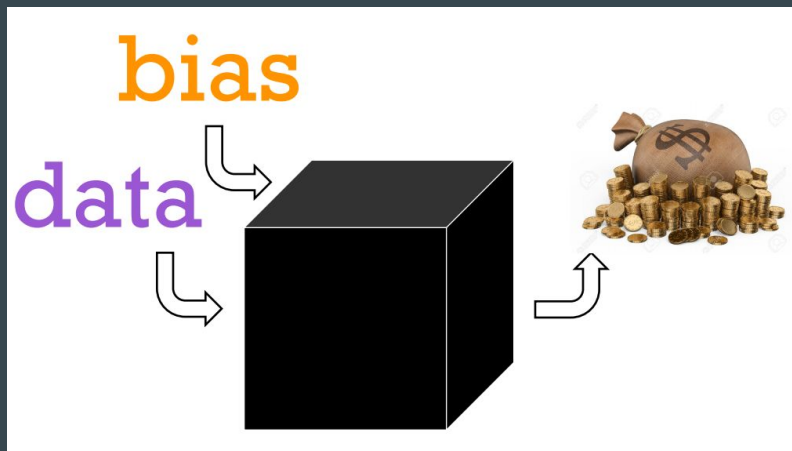
input representation



input (text)

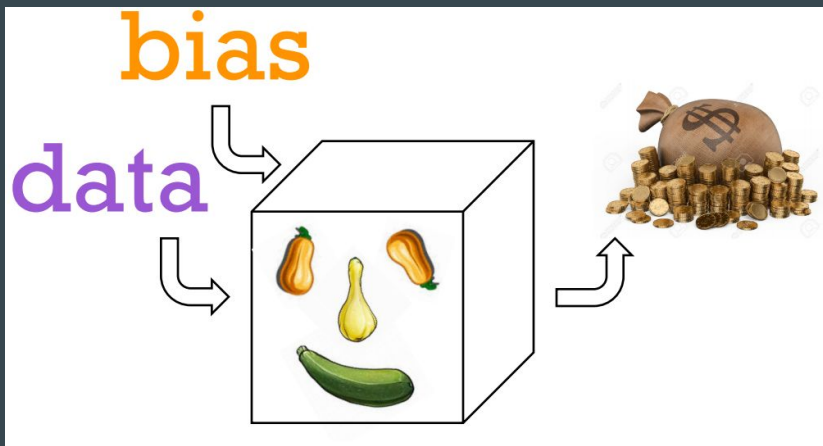
Talk Noah Smith

- Inductive bias - what does your model assume?
- Lists results on three semantic structure prediction tasks
- They share input representations (parts) -> multitask learning



Talk Noah Smith

- Inductive bias - what does your model assume?
- Lists results on three semantic structure prediction tasks
- They share input representations (parts) -> multitask learning



Talk Mirella Lapata

“Translating from Multiple Modalities to Text and Back”

- Her opinion about deep learning:



Talk Mirella Lapata

“Translating from Multiple Modalities to Text and Back”

- Her opinion about deep learning:



Talk Mirella Lapata

“Translating from Multiple Modalities to Text and Back”

- Her opinion about deep learning:



Talk Mirella Lapata

“Translating from Multiple Modalities to Text and Back”

- Her opinion about deep learning:



Talk Mirella Lapata

- Structure is coming back
- Neural nets can learn structure even though they don't know what it is
- Neural nets are here to stay, but they can still benefit from structure
- Especially multi-modal tasks
 - Language to code, movie summarization



General observations

- Everyone is using sequence-to-sequence models
 - Lots of talks were very similar, despite being on separate topics
 - Posters were even worse (except Rob, of course)
- People are getting tired of only seeing sequence-to-sequence models
- ACL is dominated by Asian males between 25-40
- Lots of complaints about quality of reviews and accepted papers
 - Reviewers play safe too often (e.g. scores of 3 on 1-5 scale)
 - There are just too many papers...

arXiv prepublishing

- Does it have a positive effect on the field?
 - It speeds up research
 - arXiv is here to stay
- Are authors supposed to be aware of them?
 - Reviewers need guidelines
 - They can mention arXiv papers, but it can not influence their review
- Should we cite them?
 - Yes, but at your own risk
- Are reviewers biased if they already saw the paper on arXiv?
 - Possibly, yes
 - Best solution: lobby arXiv to allow anonymous publishing

The logo for arXiv.org, consisting of the text "arXiv.org" in white, sans-serif font, centered within a solid red rectangular background.

arXiv.org

Opinion

- For ACL 2018: focus on papers offering linguistic insight, not too heavy on sequence-to-sequence or other deep learning methods
 - Visualization of *how* the models learn
- Reviewers can be more strict on comparison of previous work
 - Often unclear what data sets and baselines are used
 - Deliberately choosing weakish baseline to improve on (MT)
- Surprisingly few people put all their code online
 - Some papers include empty links