Language in Social Context
A Closer Look at Roles, Persuasion and Bias

Diyi Yang
School of Interactive Computing
Georgia Tech

Language Interaction Grows Exponentially

- between human and human
- between human and agents/machines

Language includes both

content and social information

Language in Social Context

Ideational
Semantics

Interpersonal
Semantics

“Systemic Functional Linguistics”:
Relations between language and its functions in social settings

by Michael Halliday

Language in Social Context

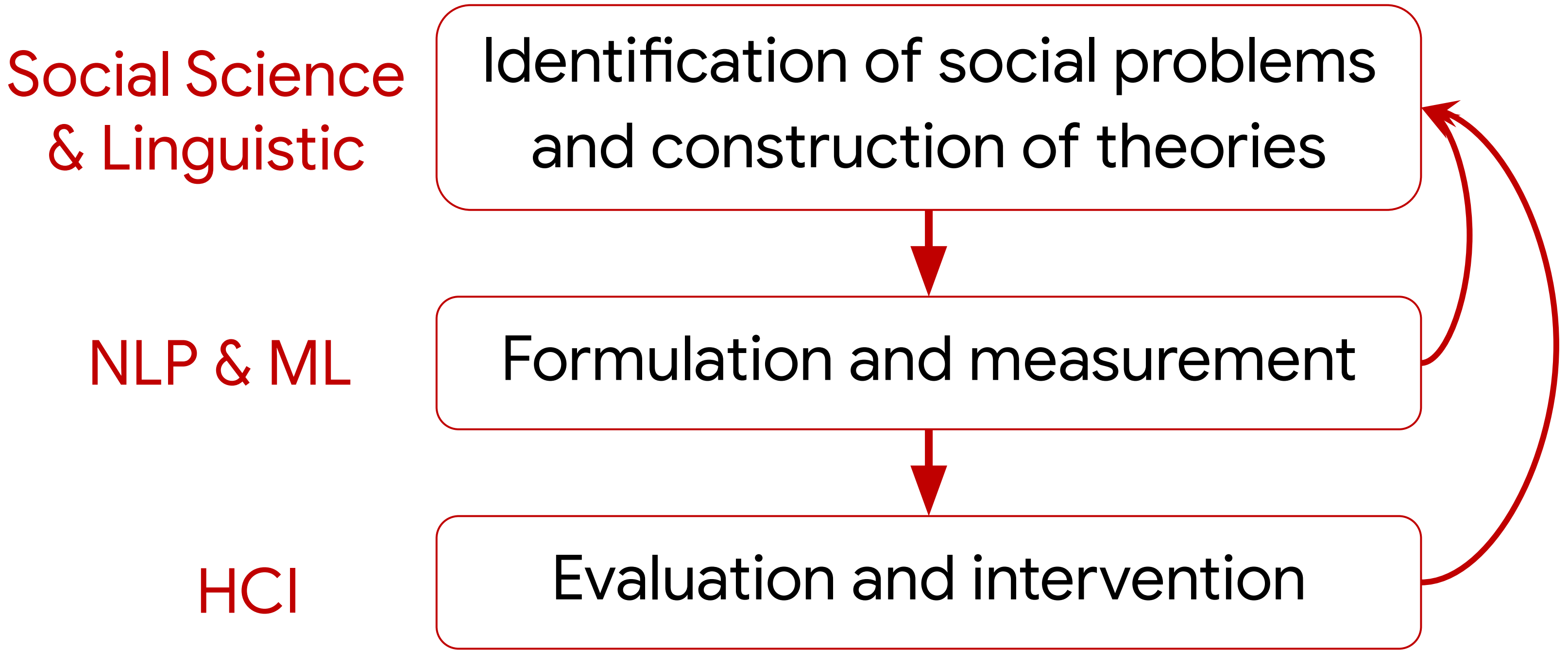
Ideational Semantics

Social roles, influence, identity, structures, culture, ideology, bias

Interpersonal Semantics

Emotion, opinion, argument, persuasion, politeness

“Systemic Functional Linguistics”, by Michael Halliday


Social Science
& Linguistic
Identification of social problems
and construction of theories
NLP & ML
Formulation and measurement
HCI
Evaluation and intervention

Language in Social Context

Ideational Semantics

Social roles, influence, identity, structures, culture, ideology, **bias**

Interpersonal Semantics

Emotion, opinion, argument, **persuasion**, politeness

“Systemic Functional Linguistics”, by Michael Halliday

This Talk

1. Computational Social Roles
2. Model Persuasion in Language
3. Neutralize Subjectively Biased Text

Computational Social Roles

Diyi Yang, Robert Kraut, Tenbroeck Smith, Elijah Mayfield, and Dan Jurafsky. *Seekers, Providers, Welcomers, and Storytellers*: Modeling Social Roles in Online Health Communities. CHI, 2019. Best paper honorable mention

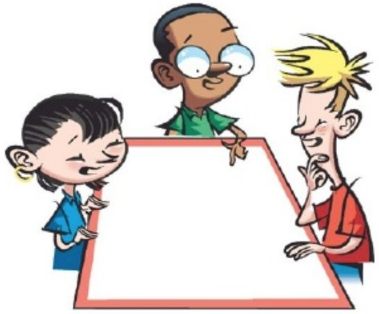
Diyi Yang, Zheng Yao, Joseph Seering, and Robert Kraut. The Channel Matters: Self-disclosure, Reciprocity and Social Support in Online Cancer Support Groups. CHI, 2019. Best paper honorable mention

Roles that People Play in Everyday Life

Self

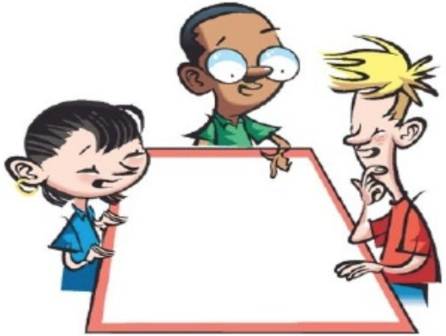
Family

Friend



Teammate

Social Roles Occur during Mass Collaboration



Group Project

Company Teams

Volunteer Editors Work on Wikipedia

Practical Use of Social Roles

- Team formation & success
- Expert-to-task allocation
- Quality control
- Conflict management
- Avoid social loafing

Scientific Understanding of Social Roles

- Name division of labour and facilitate social cooperation (Kozlowski and Klein, 2000; Kittur and Kraut, 2010)
- Understand how individuals contribute to teams (Belbin, 1981)
- Linguistics of role language
- Computer science on automated role recognition

Prior Work: Theoretical Modeling of Roles

- Basic roles, structural roles, functioning group roles, value roles (Biddle, 1979)
- Roles in teams (Belbin, 1993; Parker, 1990; Mumford et al., 2006)

Rich taxonomies of prescriptive roles (the norm of *ought*), but sometimes vague for practical use

Prior Work: Computational Modeling of Roles

- Structural roles look at network structure (Zhao et al., 2013; Henderson et al., 2012)
- Behavioral roles examine content of interaction (Bamman et al., 2013, 2014; Maki et al., 2017)

Descriptive roles (the norm of *is*) for specific contexts, lacking generalizability and theory consideration

Five Facets Social Role Framework

A cluster of interaction patterns regulated by expectations adopted by people in a context to achieve specific goals

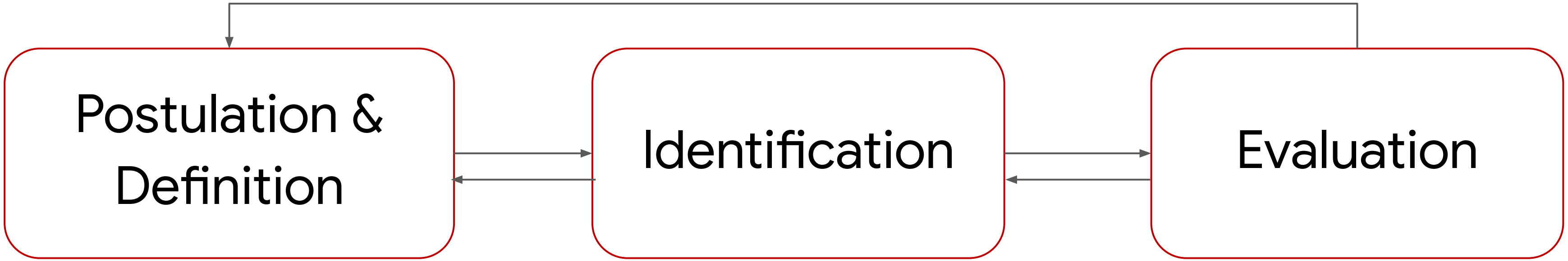
"The role concept centers upon behaviors that are characteristic of persons in a context."
— Bruce J Biddle 1979

"A social role is a comprehensive pattern of behavior and attitudes, constituting a strategy for coping with a recurrent set of situations" — Ralph H Turner 1990

Generic Methodology for Role Identification

A repeated cycle of role postulation, definition, identification and evaluation.

— *A version of the Scientific Method*



Case Studies in Different Socially Important Contexts

Wikipedia
Yang et al., 2016

Cancer Survivor Network
Yang et al., 2019

Lending Platform
Sun, Kraut and Yang, 2019

KIVA
Loans that change lives

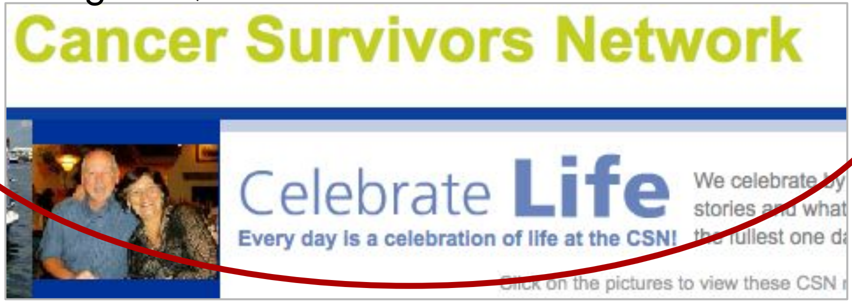
Massive Open Online Courses
Yang et al., 2015

NovoED
Learn. Collaborate. Innovate.

Case Studies in Different Socially Important Contexts

Wikipedia
Yang et al., 2016

Cancer Survivor Network
Yang et al., 2019



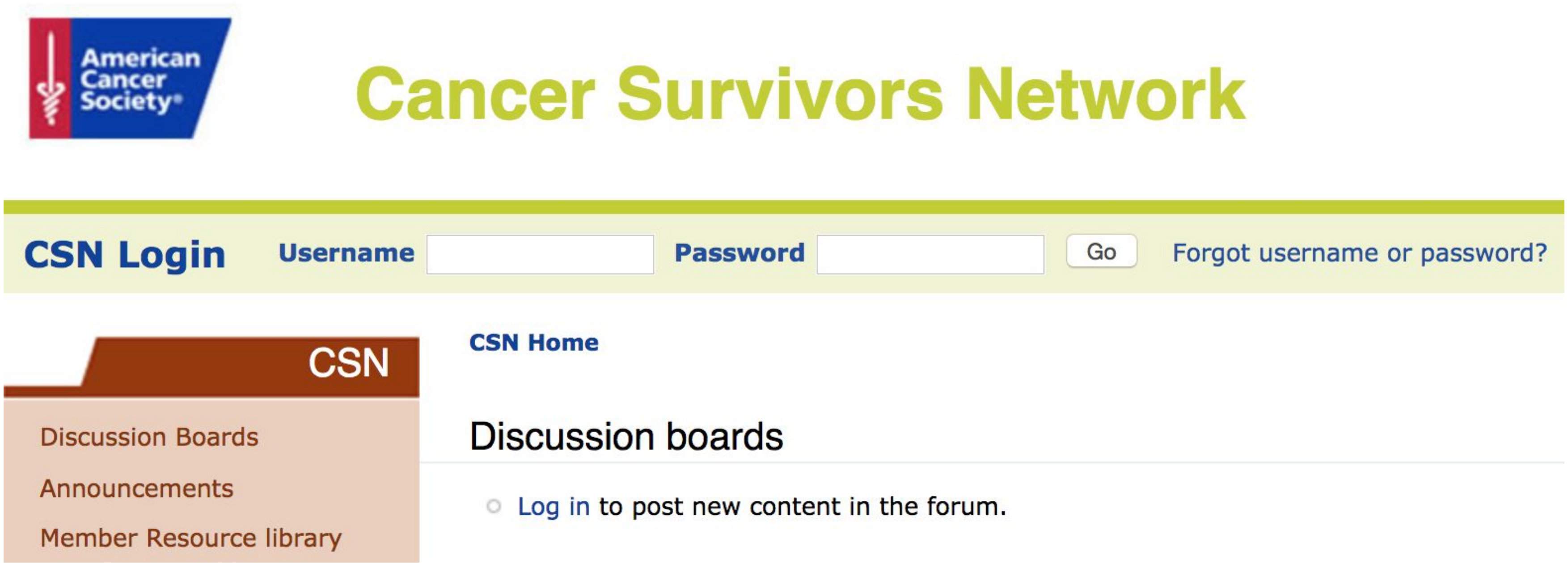
Lending Platform
Sun, Kraut and Yang, 2019

KIVA
Loans that change lives

Massive Open Online Courses
Yang et al., 2015

NovoED
Learn. Collaborate. Innovate.

Cancer Survivor Network - An Online Cancer Support Group



Cancer Survivor Network - An Online Cancer Support Group

Cancer Survivors Network

13-year data since 2005

66K users

140K threads and 1.3M replies

I was diagnosed with Invasive Ductal Carcinoma grade 2. I'm told I will need chemo. I don't understand. Any words of that will help me wrap my head around this nightmare?

Since you are a triple positive they can put you on hormones and the chance of recurrence is low. Listen to your chemo nurse ...

It gives me faith that you can have cancer and live a full life. Sorry to hear. God bless you. Stay strong please!

This conversation has been paraphrased.

28%

of Internet users have used online support group for medical information (Fox 2009)

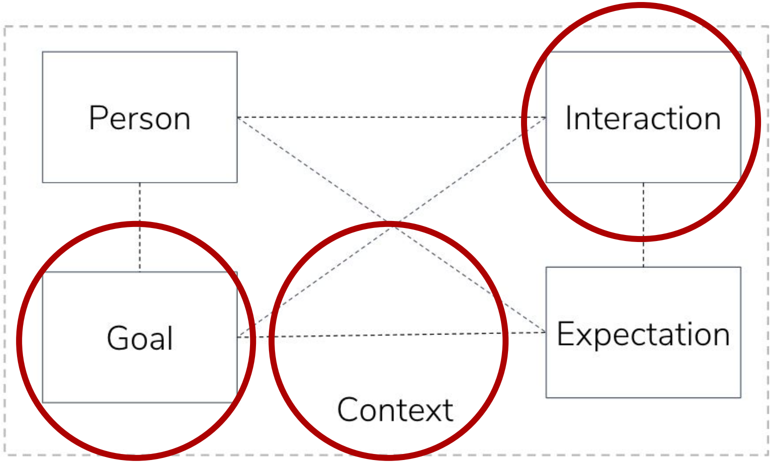
Modeling Social Roles to Better Support Patient

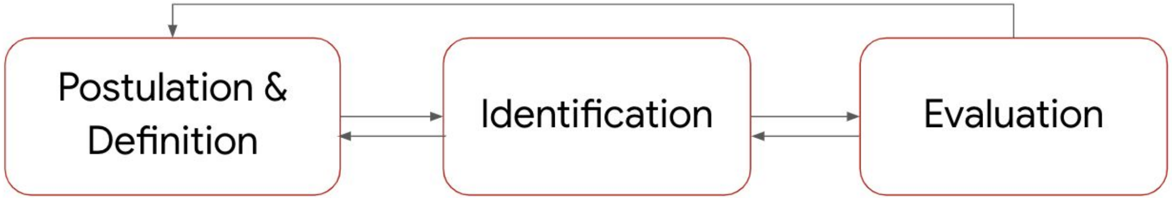
Receive Timely Help

Match with Support Providers

Connect with Similar Peers

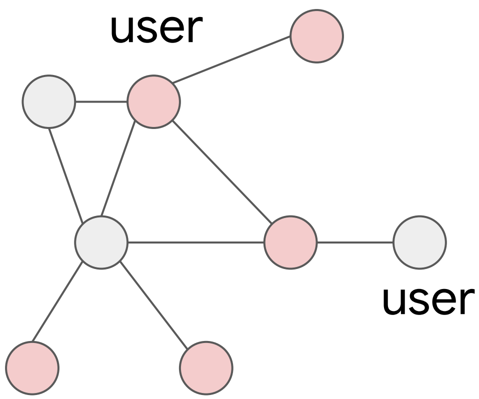
Modeling Social Roles on CSN





Postulation & Definition: The Facet of Interaction

Network-based Measures



user-user reply network

Content-based Measures

- ✓ # seek request, # comments
- ✓ Emotional aspects: *anger, sadness*
- ✓ Social concerns: *friend, family, social*
- ✓ Religious orientation: *religious, death*
- ✓ Self-focus: *I, you, he/she*
- ✓ Topics modeling

Postulation & Definition: The Facet of Goal

Since you are a triple positive they can put you on hormones and the chance of recurrence is low. Listen to your chemo nurse …

Informational Support

It gives me faith that you can have cancer and live a full life. Sorry to hear that. God bless you . Please stay strong!

Emotional Support

Automatic Measurement of the Facet of Goal

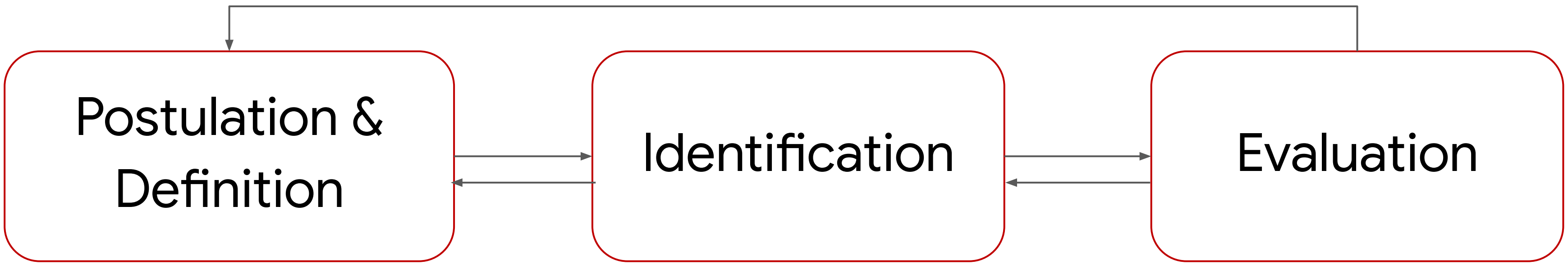
1. Seek emo support (r=0.64)
2. Provide emo support (r=0.75)
3. Provide empathy (r=0.72)
4. Provide appreciation (r=0.67)
5. Provide encouragement (r=0.64)

Reasonable correlations between model predictions and human judgements

6. Seek info support (r=0.73)
7. Provide info support (r=0.79)

Regression models trained on human-annotated data (Yang et al., 2017), with features from LIWC and word embeddings

Role Identification Methodology



The facets of interaction, goal, context

Gaussian mixture models

Modeling Social Roles via Mixture Model

Intuition: a user is a mixture of different social roles

$$p(x) = \sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(x|\mu_k, \Sigma_k), \; \sum_k \pi_k = 1$$

- ✓ User X is represented by the aforementioned features
- ✓ Select the number of roles K quantitatively and qualitatively

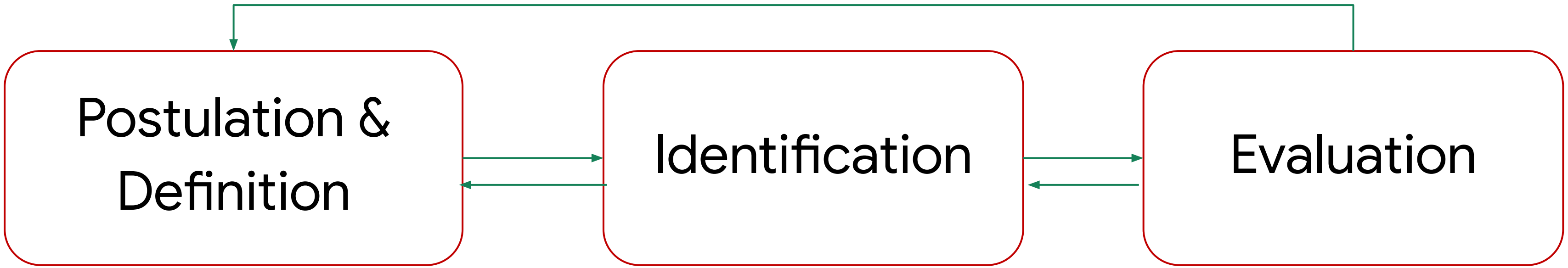
Roles Identified by Our Model (trained on 66K users)

Emotional Support Provider	Private Support Provider
Newcomer Welcomer	All-round Expert
Informational Support Provider	Newcomer Member
Story Sharer	Knowledge Promoter
Informational Support Seeker	Private Networker
Private Communicator	

Example Roles Identified by Our Model

Role Name	%	Typical Behaviors Listed in Importance
Emotional Support Provider	33.3	Provide emo support, appreciation, empathy, encouragement & empathy, info support, # subforums a user participated
Welcomer	15.9	Out-degree in forum, # replies, % of talk to newcomers, provide encouragement & empathy
Informational Support Provider	13.3	Provide info support, empathy in the forum, use words related to symptoms, anxiety, and drugs related words
Story Sharer	10.2	Initialize threads, positive and negative self-disclose, seek emo support, negative self-disclosure, seek info support
Informational Support Seeker	8.9	Initialize threads, seek info support, negative self-disclosure, seek emo support, use words related to disease and symptoms

Role Identification Methodology



The facets of interaction, goal, context

Gaussian mixture models

1. **Quantitative measures**
2. **Qualitative measures**
3. **Survey role holders**
4. **Downstream applications**

Role Evaluation: Qualitative Measures

Work with 6 moderators on CSN to assess the derived roles

*" It seems very **comprehensive** and there are so many different examples, so I feel like it is **covered very well** with your different roles and labels. "*

The identified roles were comprehensive

Role Evaluation: Qualitative Measures

Work with 6 moderators on CSN to assess the derived roles

*“The one that I think did not emerge is the **policeman,** these people complain to moderators when some people are doing things wrong.”*

Model failed to capture the *“defenders”*

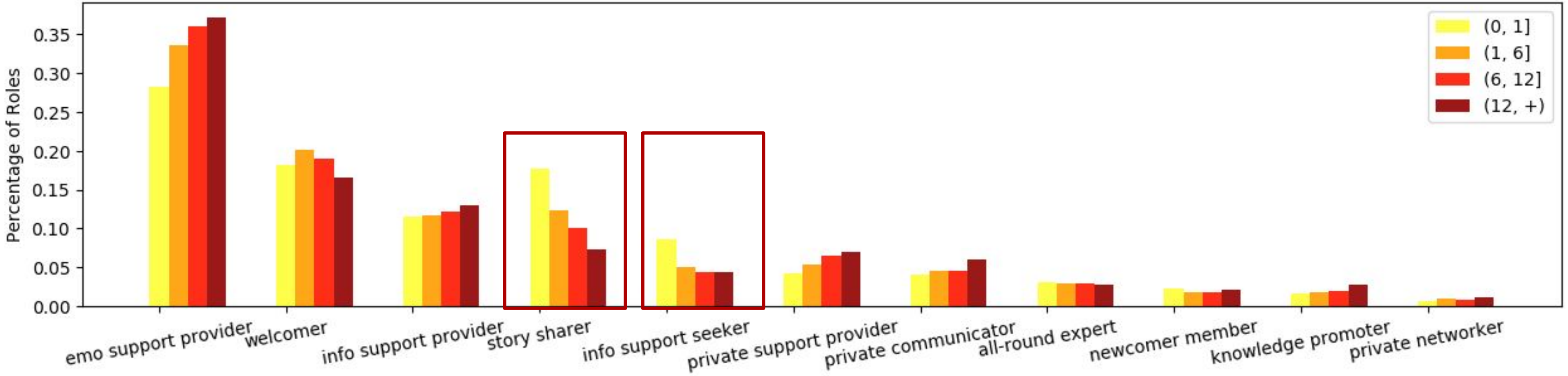
Role Evaluation: Qualitative Measures

Work with 6 moderators on CSN to assess the derived roles

*“there are **not a lot of them**, but they kind of stick in your memories since they are telling others what to do.”*

Model failed to capture the *“defenders”*

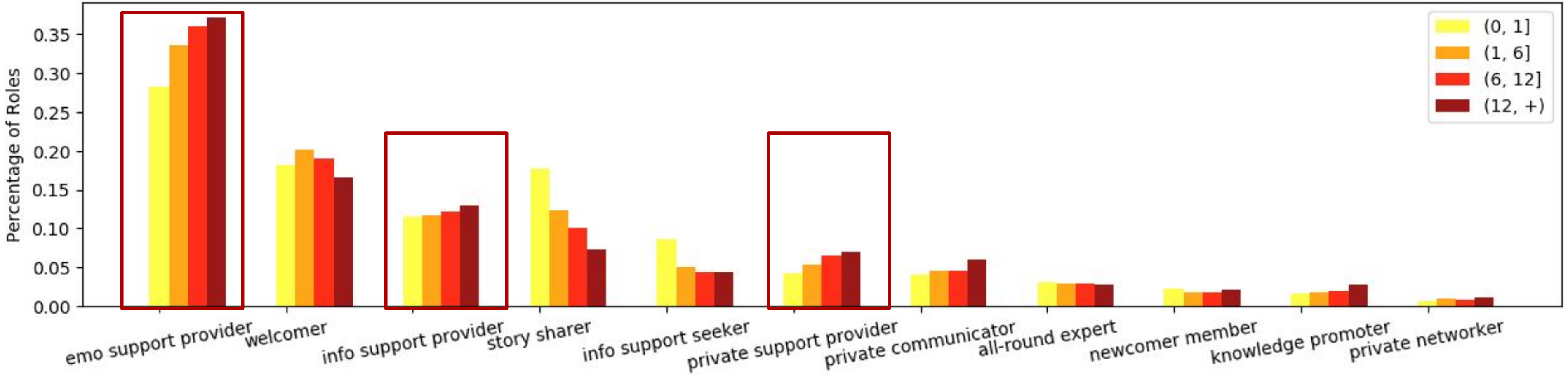
Roles People Enact Change with Time on CSN



Among those who stay on CSN for at least 12 months:

✓ Support seekers and story sharer decline with tenure

Roles People Enact Change with Time on CSN



Among those who stay on CSN for at least 12 months:

- ✓ Support seekers and story sharer decline with tenure
- ✓ Support providers increase with tenure

From Roles Seeking Sources to Ones Offering Help

12 interviews of users on Cancer Survivor Network

*I **initially stayed** because information was important, but **over time,** I found talking with people who had similar experiences **is more helpful***

From Roles Seeking Sources to Ones Offering Help

12 interviews of users on Cancer Survivor Network

*I'm **now** looking for **helping** people **who are seeking for advice.***

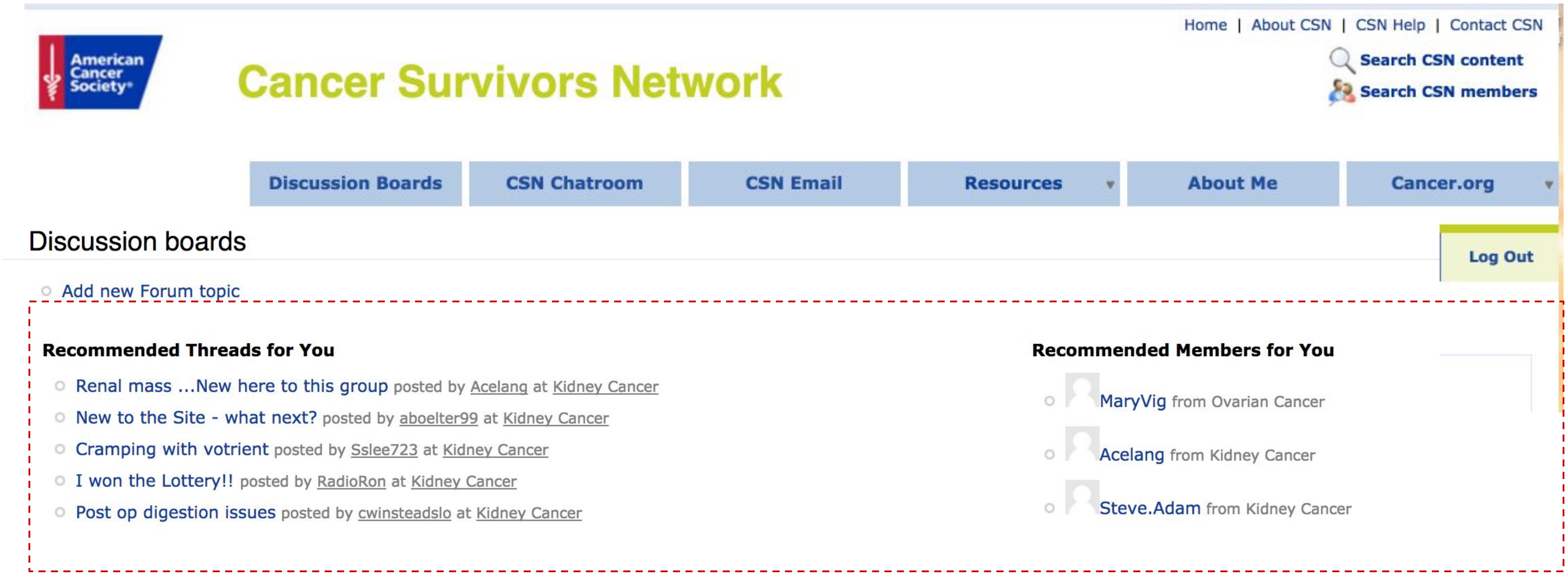
Automated Intervention to Improve Interaction

Provide Timely Help

Match with Support Providers

Connect with Similar Peers

Our Deployed Recommender System* on CSN



* Feature-based matrix factorization (Yang et al., 2014a; Yang et al., 2014b; Yang et al., 2014c)

Our Deployed Recommender System* on CSN

Add to favorites | Manage your favorites

Renal mass ...New here to this group

Acelang

Posts: 19
Joined: Apr 2017

Apr 18, 2017 - 8:46 pm

Hi I'm new here but very no nervous about findings on a CT scan I had done in the er .. a 5 cm renal mass was found along with enlarged lymph nodes my concern is 6 month prior this tumor was 2.6cm I don't see a doctor until next week but I'm worried because I have a lump in my side and back ...Have night sweats fatigue and back and stabbing pains in my stomach where the lump is

How relevant is this thread to you?:

Your rating: None

Add new comment 1083 reads Report as Inappropriate

icemantoo

Thats a lot of growth

Apr 18, 2017 - 9:14 pm

Ace,

Do what you have to do to get this taken care of asap. It is either very aggressive or was measured 2 different ways like Ultrasound and CT. It shoild have been addressed at 2.6 cm as that is too large for watch and wait. At 5 cm a surgery only result is still very possable and if so you would need no further tratments other than scans.

You might be interested in...

right kidney mass???
atos posted at Kidney Cancer

Renal Cell Cancer
gerard posted at Kidney Cancer

Help eliminate cancer
drrhorho posted at Kidney Cancer

20 Year Old with Kidney Cancer
veg4you posted at Kidney Cancer

HELP! JUST DIAGNOSED WITH RENAL CANCER! NEED A FRIEND TO TALK TOO!
TOM659 posted at Kidney Cancer

* Feature-based matrix factorization (Yang et al., 2014a; Yang et al., 2014b; Yang et al., 2014c)

People (newcomers & oldtimers) On CSN Use Our Intervention

Over 11,000 people have signed up and are using our intervention on Cancer Survivor Network since 2018

https://csn.cancer.org/forum

Recommendation Increase Reading outside Favorite Forum

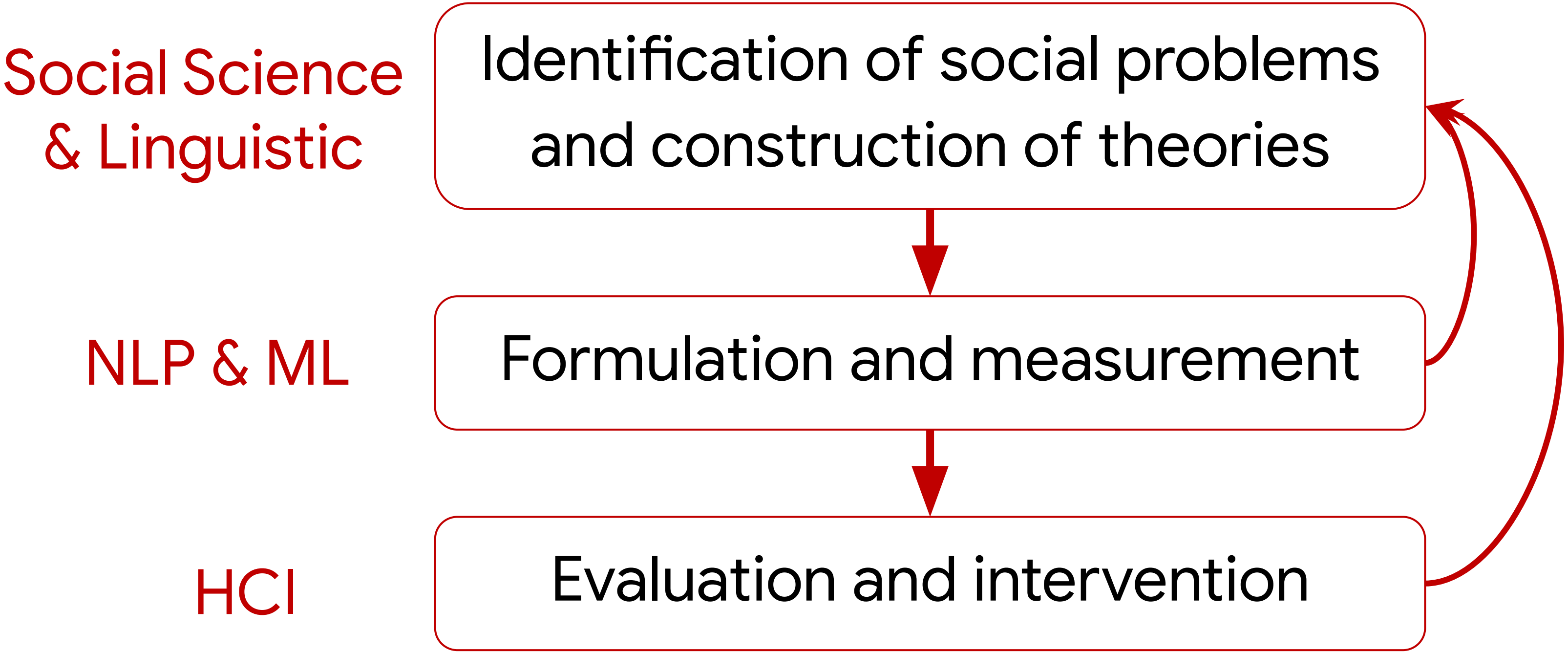
Recommender Setting (Result till May, 2018)	Displayed	Hits	Hit Ratio (%)	Improvement over control
Recommend recent threads in favorite forum	6718	150	2.23	
Recommend based on history & restricted to favorite forum	7002	138	1.97 (ns)	-3%
Recommend based on history	6615	238	3.60 ***	+61%

Recommendations based on history and user roles are ongoing

Explain Recommendation via Social Roles

“ Here are some newcomers you might want to say hi ”

“ Here are some information experts you could reach out to”

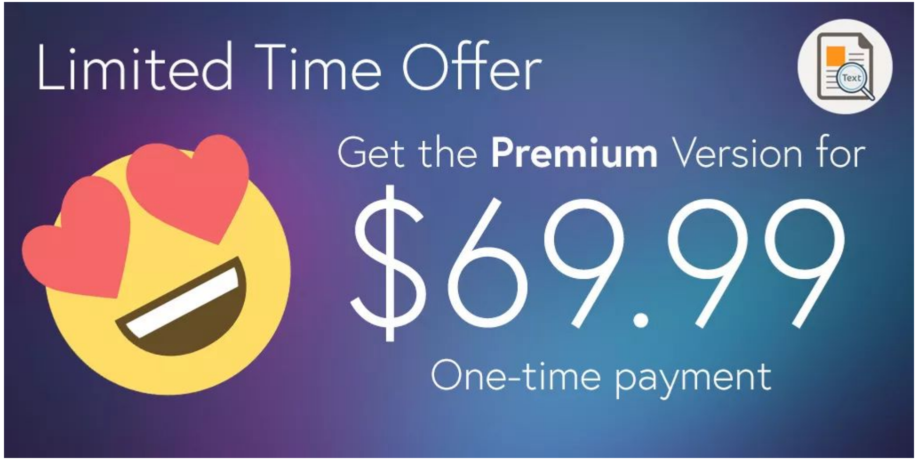

Social Science
& Linguistic
Identification of social problems
and construction of theories
NLP & ML
Formulation and measurement
HCI
Evaluation and intervention

This Talk

1. Identify Social Roles in Online Health Communities
2. **Model Persuasion in Language**
3. Neutralize Subjectively Biased Text

2 Modeling Persuasion in Language on Crowdfunding Platforms

Diyi Yang, Jiaao Chen, Zichao Yang, Dan Jurafsky, and Eduard Hovy. “Modeling Persuasive Strategies via Semi-supervised Neural Nets on Crowdfunding Platforms”. NAACL 2019


Limited Time Offer
Get the **Premium** Version for
$69.99
One-time payment


Tasty Flake Breakfast Cereal
Everyone is enjoying it!


Coca-Cola
open a Coke,
open happiness


iPhone 6
Bend it like Beckham

Loan Advocacy Requests on Crowdfunding Site

93% funded

Only 34 hours left! $50 to go

Total loan: $775

Powered by 27 lenders

Lillyana Maria

Medellín, Colombia / Sewing

$25 ▾

KIVA

loans that change lives

What Makes Language Persuasive?

“I am the first lender on this woman-lead group loan in Burma. This loan will be utilized to repair her old duck farm and enable her to purchase nutritious duck feed that can help boost duck egg production. This way, she will be able to support her children’s education in the future.”

This request persuaded 3 out of 50 people to lend

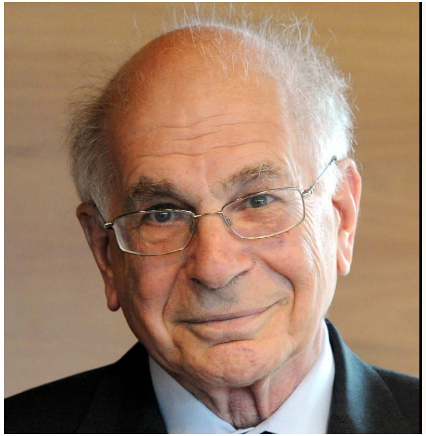
Kahneman: Thinking, Fast and Slow

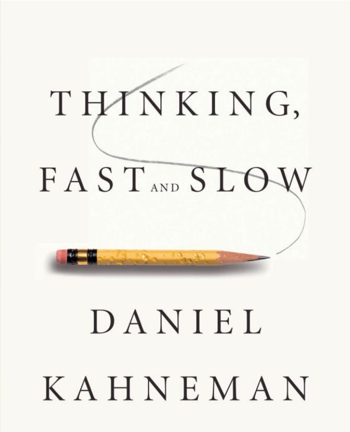
System 1: Think Fast

“Only a few left?”

“Experts recommended?”

“People who I like are using it?”





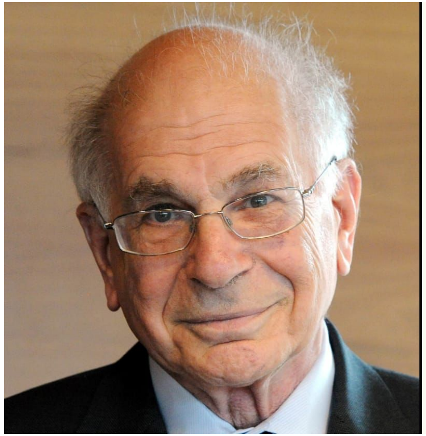
Kahneman: Thinking, Fast and Slow

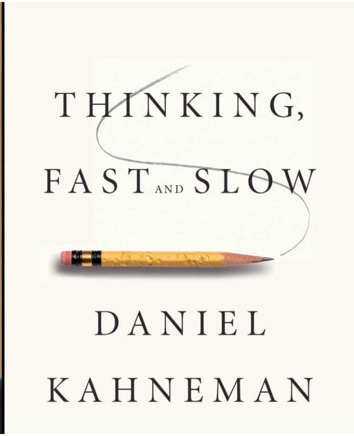
System 1: Think Fast

System 2: Think Slow

“Are the facts correct?”

“Are the conclusions warranted?”





Translating into Measurable Language Cues

Scarcity - people value an item more when it becomes rare or urgent

"This loan is going to expire in 35 minutes, please help!"

(Cialdini, 1987; Chiaken, 1980)

Translating into Measurable Language Cues

Scarcity

Emotion - messages full of emotional valence make people care

“The picture of widow Bunisia holding her baby in front of her meager home brings tears to my eyes.”

(Cialdini, 1987; Chiaken, 1980)

Translating into Measurable Language Cues

Scarcity - people value an item more when it becomes rare or urgent

Emotion - messages full of emotional valence make people care

Identity - people like their group/identity more over others

Commitment - we like to convince others we made the correct choice

Concreteness - providing concrete facts or evidence

Impact - emphasizing the importance or bigger impact

(Cialdini, 1987; Chiaken, 1980)

Translating into Measurable Language Cues

Scarcity
Emotion
Identity
Commitment
Concreteness
Impact

System 1: Heuristic Processing (think fast)

System 2: Systematic Processing (think slow)

(Cialdini, 1987; Chiaken, 1980)

To Predict Persuasive Strategies w/ Limited Data

Sheila ✉ Mar 5, 2017 - 6:21 pm PST

Send a Kiva Card

Joined Aug 23, 2013

I am the first lender on this woman-lead group loan in Burma. This loan will be utilized to repair her old duck farm and enable here to purchase nutritious duck feed that can help boost duck egg production. This way, she will be able to support her children's education in the future.

commitment · concreteness · impact

Classical Semi-supervised Setting

Limited Labeled Sentences

Commitment	I am the first lender on this woman-lead group loan in Burma.
Concreteness	This loan will be utilized to repair her old duck farm and enable here to purchase nutritious duck feed that can help ...
Impact	This way, she will be able to support her children's education.
...	...

Extra Unlabeled Sentence

?	Who's the cutest 82-year-old you've ever seen who needs funds in 6 days?
?	She's still actively working, and needs funds to place her orders..
?	Look at that smile, adorable!
...	...

(Zhu, Lafferty, and Rosenfeld, 2003; Kingma et al., 2014; Chapelle, Scholkopf, Zien, 2006)

Semi-supervised Setting w/ Document Supervision

Labeled Docs
+
Limited Labeled Sentences

Doc-label: 9/50

Commitment	I am the first lender on this woman-lead group loan in Burma.
Concreteness	This loan will be utilized to repair her old duck farm and enable here to purchase nutritious duck feed that can help ...
Impact	This way, she will be able to support her children's education.

Doc-label: 3/50

...	...

Labeled Docs
+
Extra Unlabeled Sentences

Doc-label: 14/100

?	Who's the cutest 82-year-old you've ever seen who needs funds in 6 days?
?	She's still actively working, and needs funds to place her orders.
?	Look at that smile, adorable!

Doc-label: 27/100

...	...

(Obquab et al, 2015; Pinheiro and Collobert, 2015)

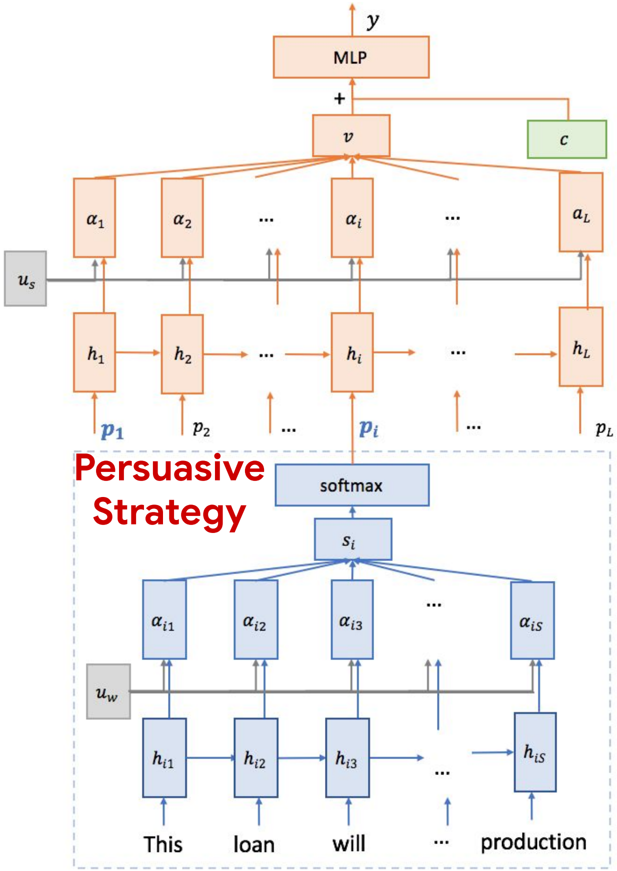
Semi-supervised Net

Sent/Doc encoder

Semi-supervised objective

$$l = \gamma \sum_{d} (y_d - \bar{y}_d)^2 - \beta \sum (- g_i \log p_i)$$

(Bahdanau et al., 2014; Yang et al., 2016)



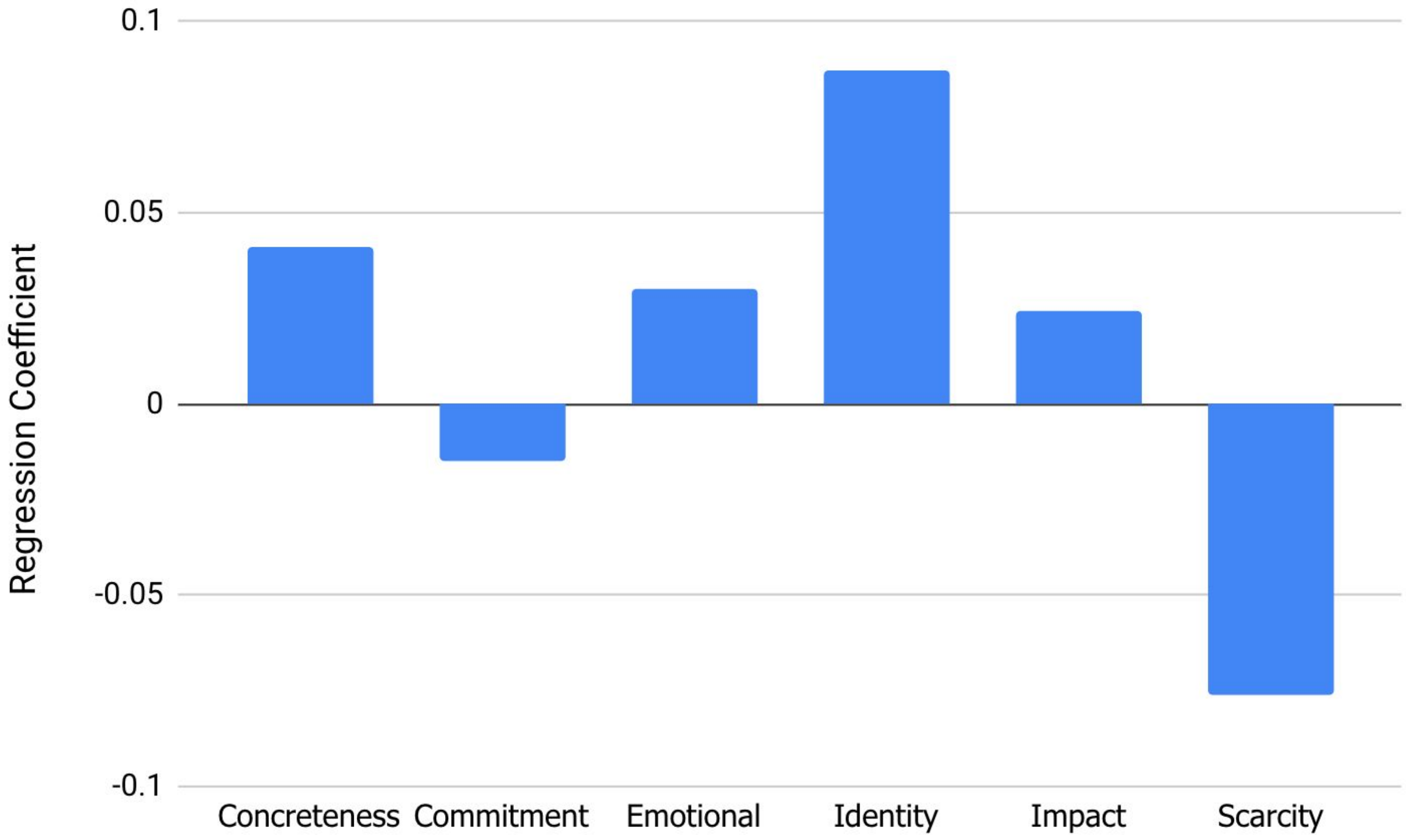
Document Encoder

Sentence Encoder

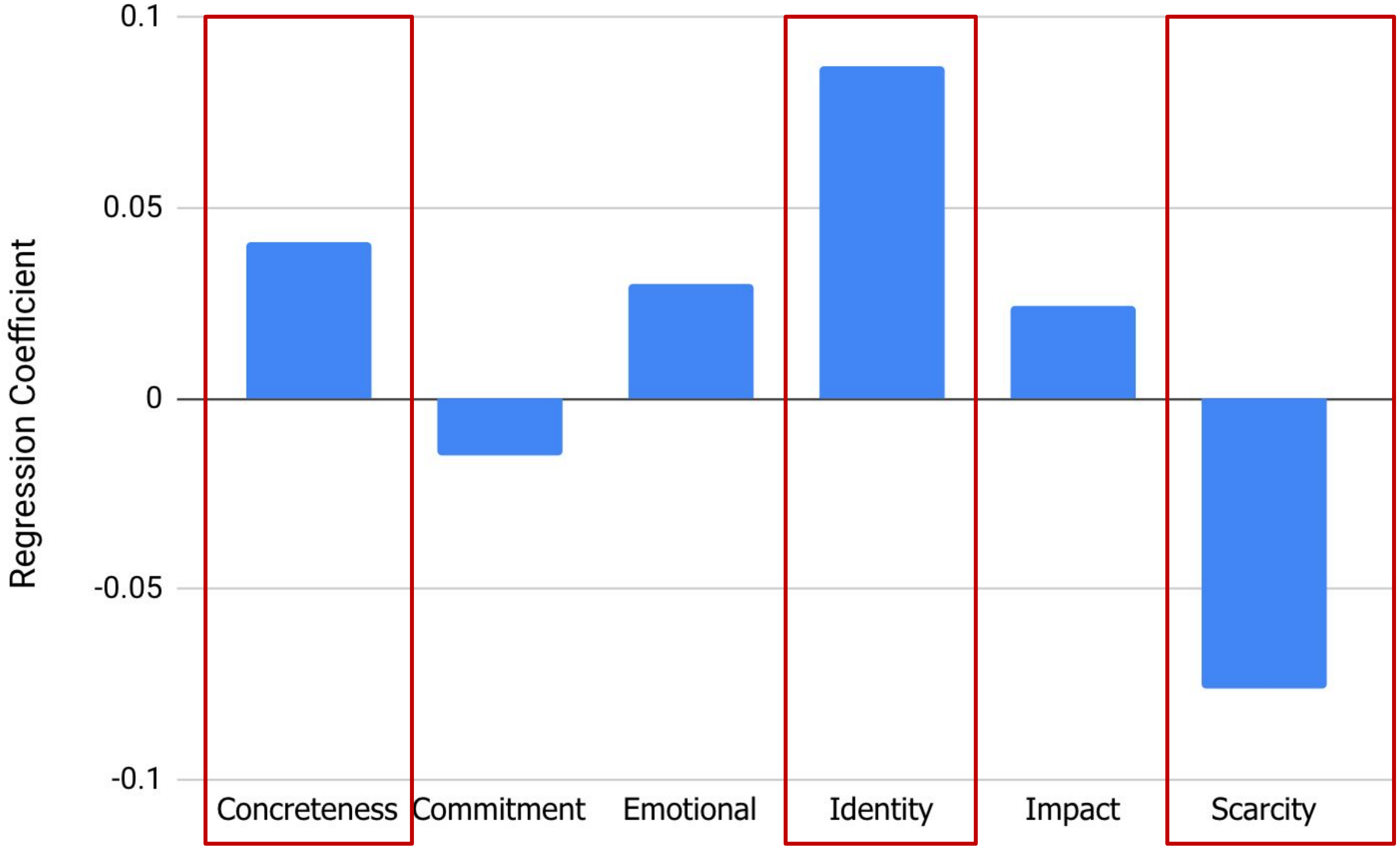
Results on Predicting Persuasion Strategies

Model	Sentence-Level		Doc-Level
	Accuracy	F1	RMSE
Sentence Only (SVM)	0.34	0.17	
Sentence Only (GRU)	0.51	0.47	
Supervised Sentence + Doc (Hierarchical Net)	0.48	0.43	1.15
Semi-supervised Sentence + Doc (Semi + Hierarchical Attention Net)	**0.57**	**0.52**	1.04

To Explain the Influence of Persuasion Strategies

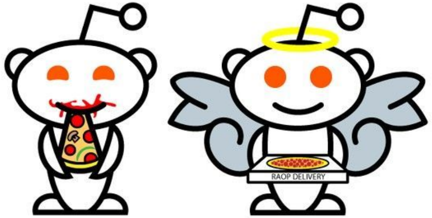


Identity and Concreteness Matter, Not Scarcity



Generalizing Model to Random Acts of Pizza

Random Acts of Pizza on Reddit (r/RAOP)



50

[REQUEST] I battled through a cold, took 4 exams and gave 2 presentations all in 1 week. I just found out I have all A's because of my hard work!

submitted 2 days ago by OriginalWF **Small Fish (15)**

My wife and I have been sick for the past week while working and taking care of our daughter. I just found out that I passed all my exams and I have straight A's in all my classes, which is a first for my entire college career. I want to celebrate but we just payed rent, which means we don't have the ability to go out.

If anyone could help with a pizza, I'd appreciate it a bunch!

2 comments share save hide report

Top Effective Strategies on /r/RAOP:
Emotion, Scarcity, Concreteness

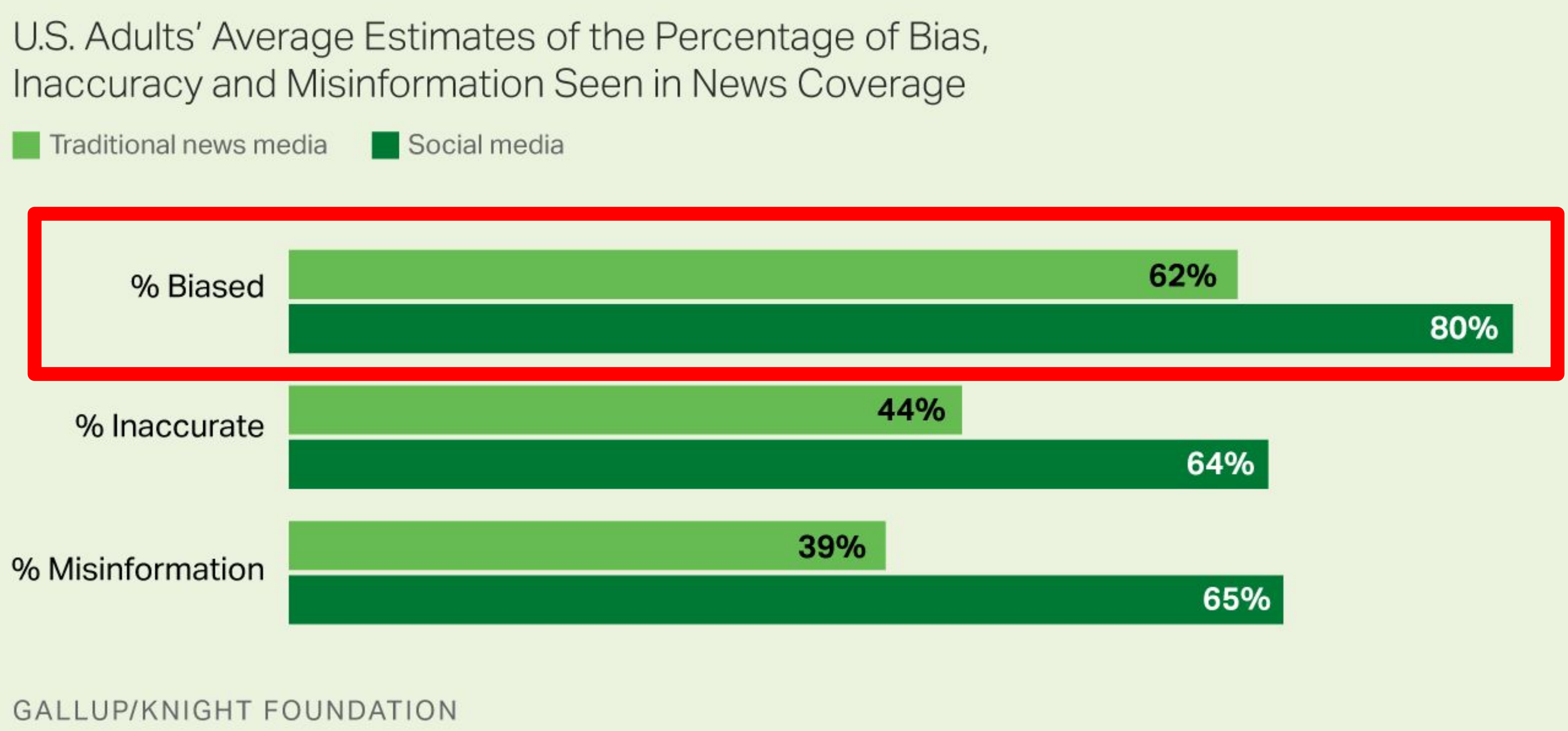
This Talk

1. Identify Social Roles in Online Health Communities
2. Model Persuasion in Language
3. **Neutralize Subjectively Biased Text**

3 Neutralizing Subjectively Biased Text

Reid Pryzant, Richard Diehl Martinez, Nathan Dass, Sadao Kurohashi, Dan Jurafsky, Diyi Yang. Automatically Neutralizing Subjective Bias in Text. AAAI 2020.

Slides credit to Reid Pryzant



American Views: Trust, Media and Democracy

What Makes One Headline Biased and Another Neutral?

Can Suggest Less Biased Alternatives to Text?

?

Three Types of Subjectivity Bias

Framing bias

✓ *Most of the gameplay is **pilfered from** ddr*

Epistemological bias

✓ *Developing a new downtown **will bring back** our arts.*

Demographic bias

✓ *A lead programmer usually spends **his career**...*

Neutralizing Subjectivity Bias

based on

Framing bias

✓ *Most of the gameplay is **pilfered from** ddr*

which its promoters hope

Epistemological bias

✓ *Developing a new downtown **will bring back** our arts.*

their careers

Demographic bias

✓ *A lead programmer usually spends **his career**...*

A Large Scale Wikipedia Neutrality Corpus

The first parallel corpus of biased language, with 180,000 sentence pairs

Text with subjective bias	Corresponding neutral point of view
Kathy Kirby, 1960's **blonde singing legend**	Kathy Kirby, 1960's **singer**
Go is **the deepest** game in the world	Go is **one of the deepest game** in the world
The authors' **expose** on nutrition studies	The authors' **statements** on nutrition studies
Marriage is a **holy union** of individuals	Marriage is a **personal union** of individuals

Method 1: Concurrent

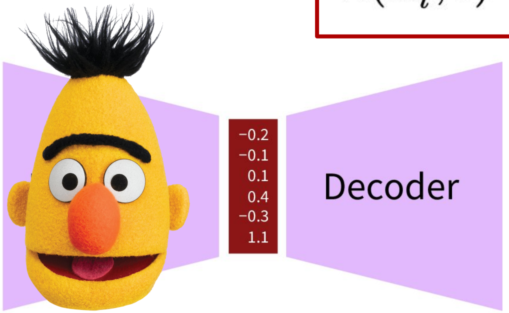
Seq2seq (with copy mechanism (See et al. 2017))
+ Bert
+ Pretraining on the unbiased corpus
+ Token-weighted loss function

$$\mathcal{L}(s,t) = -\sum_{i=1}^{m} \boxed{\lambda(w_i^t, s)} \log p(w_i^t | s, w_{<i}^t) + c$$

$$\lambda(w_i^t, s) = \begin{cases} \alpha & : \; w_i^t \notin s \\ 1 & : \; \text{otherwise} \end{cases}$$

The first **man** who saw the need for a clear definition of probability was Laplace.

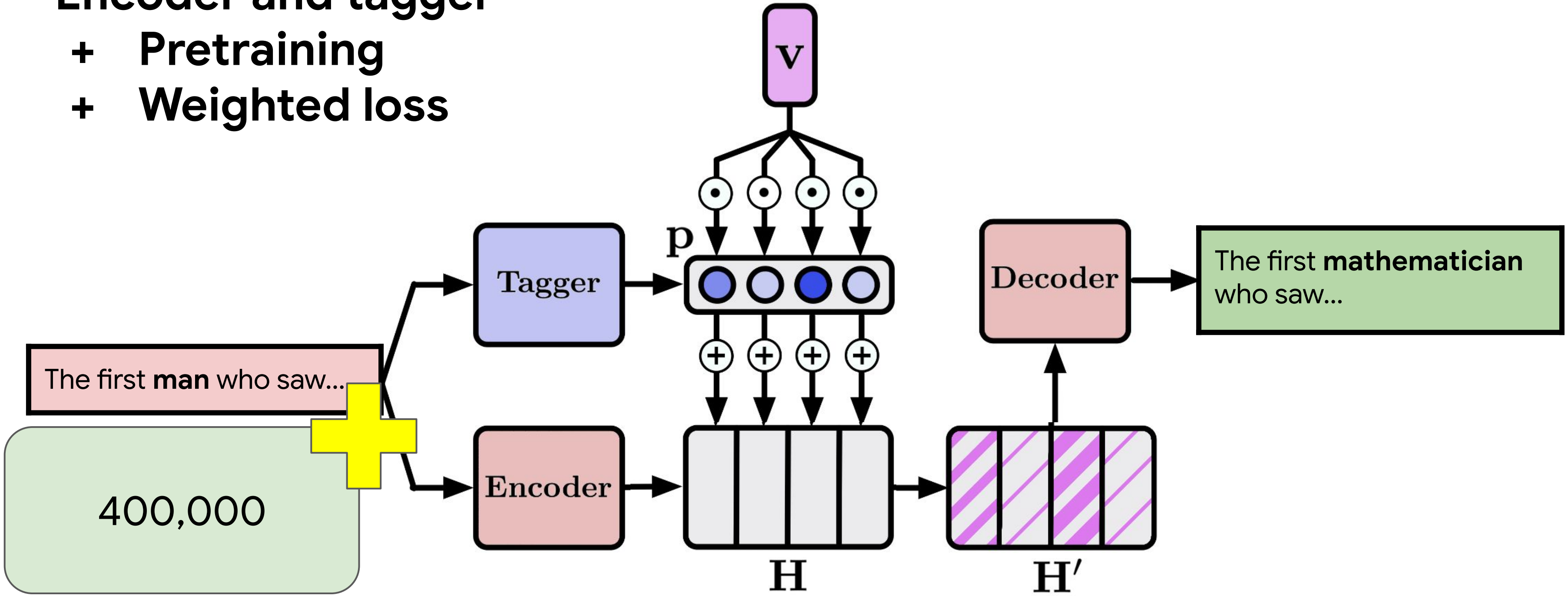
Neutral Text



Loss weight (lambda):	1.0	1.0	1.5	1.0
Token	The	first	**mathematician**	who

Method 2: Modular

Encoder and tagger
- **+ Pretraining**
- **+ Weighted loss**



Method 1: Concurrent

PROS

- Straightforward
- Easy to use

CONS

- Opaque
- Uncontrollable

Method 2: Modular

PROS

- Controllable
- Interpretable

CONS

- Complicated
- Harder to train

Neutralizing Biased Text Results

Method	BLEU	Accuracy	Fluency	Meaning	Bias
Source Copy	91.33	0.00	-	-	-
Detector (always delete biased word)	92.43*	38.19*	-0.253	1.108	-0.324
Detector (predict substitution from biased word)	92.51	36.57*	-0.233	1.139	-0.327
Delete Retrieve (ST) (Li et al. 2018)	88.46*	14.50*	-0.209	1.294	-0.456
Back Translation (ST) (Prabhumoye et al. 2018)	84.95*	9.92*	-0.359	1.126	-0.390
Transformer (MT) (Vaswani et al. 2017)	86.40*	24.34*	-0.259	0.905	-0.458
Seq2Seq (MT) (Luong, Pham, and Manning 2015)	89.03*	23.93	-0.423	1.294	-0.436

Fluency: how much more fluent is the output compared to the source?

⇒ **higher is better**

Meaning: how well does the output preserve the meaning of the source?

⇒ **higher is better**

Bias: how much more biased is the output compared to the source?

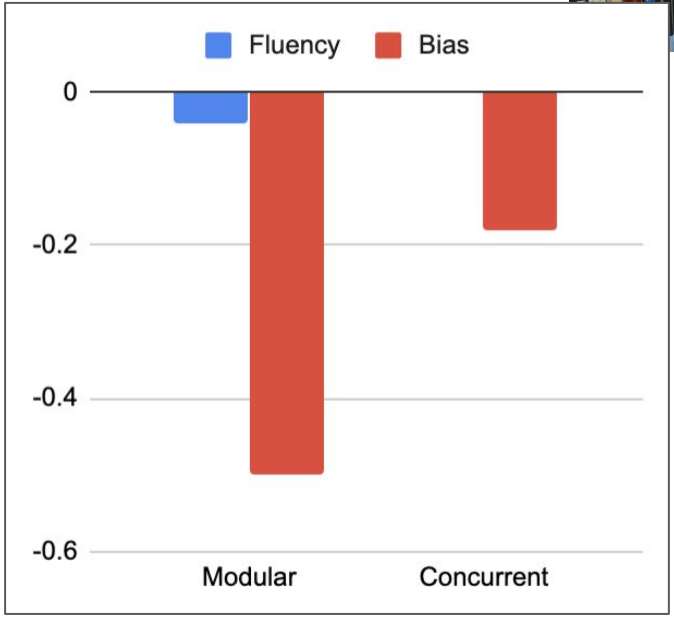
⇒ **Lower is better**

Neutralizing Biased Text Results

Method	BLEU	Accuracy	Fluency	Meaning	Bias
Source Copy	91.33	0.00	-	-	-
Detector (always delete biased word)	92.43*	38.19*	-0.253	1.108	-0.324
Detector (predict substitution from biased word)	92.51	36.57*	-0.233	1.139	-0.327
Delete Retrieve (ST) (Li et al. 2018)	88.46*	14.50*	-0.209	1.294	-0.456
Back Translation (ST) (Prabhumoye et al. 2018)	84.95*	9.92*	-0.359	1.126	-0.390
Transformer (MT) (Vaswani et al. 2017)	86.40*	24.34*	-0.259	0.905	-0.458
Seq2Seq (MT) (Luong, Pham, and Manning 2015)	89.03*	23.93	-0.423	1.294	-0.436
Base	89.13	24.01	-	-	-
+ *loss*	90.32*	24.10	-	-	-
+ *loss* + *pretrain*	92.89*	34.76*	-	-	-
+ *loss* + *pretrain* + *detector* (MODULAR)	93.52*	**46.80***	-0.078	0.996	**-0.467**
+ *loss* + *pretrain* + *BERT* (CONCURRENT)	**93.94**	45.87	**0.132**	**0.758**	-0.423
Target copy	100.0	100.0	-0.077	1.128	-0.551

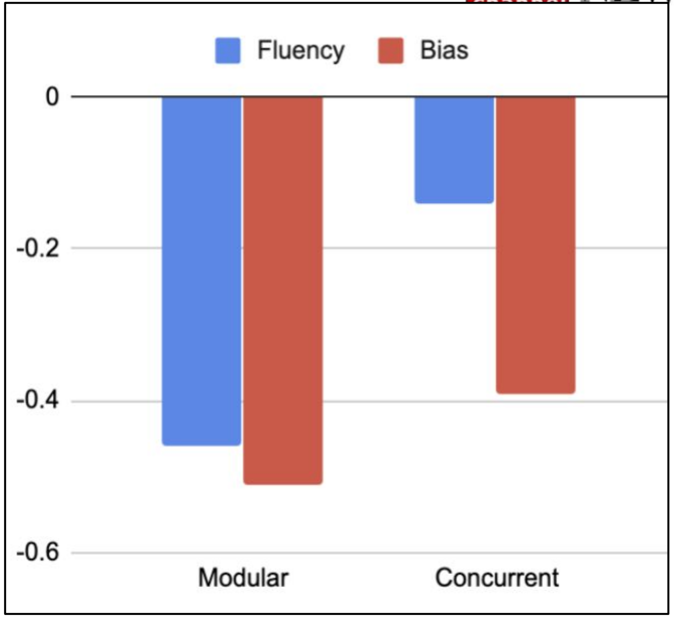
Neutralizing **Political Books**

	The Ideological Books Corpus
Original	Activists have filed a lawsuit...
Modular	Critics of it have filed a lawsuit ...
Concurrent	Critics have filed a lawsuit ...



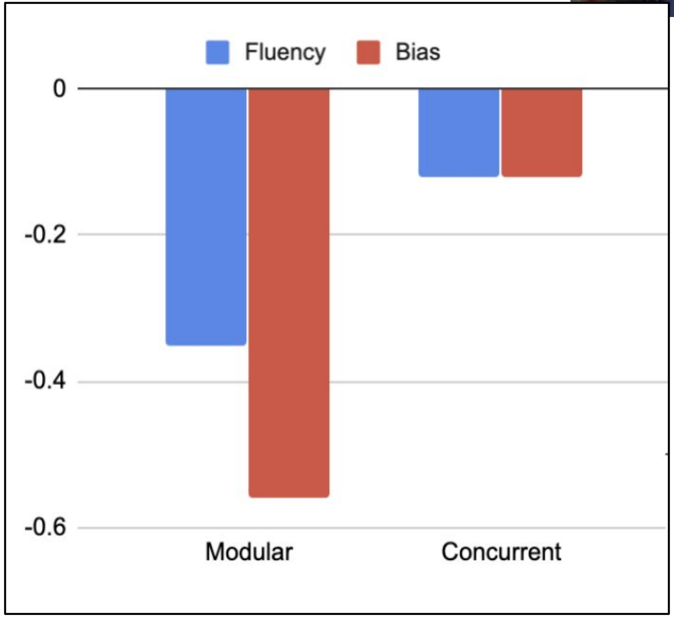
Neutralizing **News Headlines**

	News Headlines
Original	Zuckerberg claims Facebook can...
Modular	Zuckerberg stated Facebook can...
Concurrent	Zuckerberg says Facebook can...



Neutralizing **Campaign Speeches**

	Campaign speeches
Original	This includes amazing Americans like...
Modular	This includes Americans like...
Concurrent	This includes some Americans like...



Summary of This Talk

1. Computational Social Roles

2. Model Persuasion in Language

3. Neutralize Subjectively Biased Text

Language in Social Context

A Closer Look at Roles, Persuasion and Bias

Diyi Yang

Comments? diyi.yang@cc.gatech.edu