Your spouse needs professional help: Determining the Contextual Appropriateness of Messages through Modeling Social Relationships

David Jurgens*
jurgens@umich.edu

Agrima Seth*
agrima@umich.edu

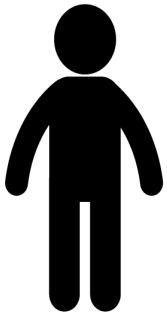
Jackson Sargent†
jacsarge@umich.edu

SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

BAB LAB LAB

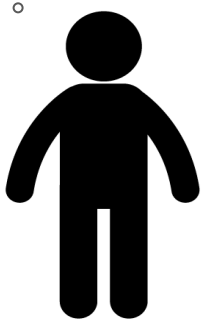
Athena Aghighi†
aaghighi@ucdavis.edu

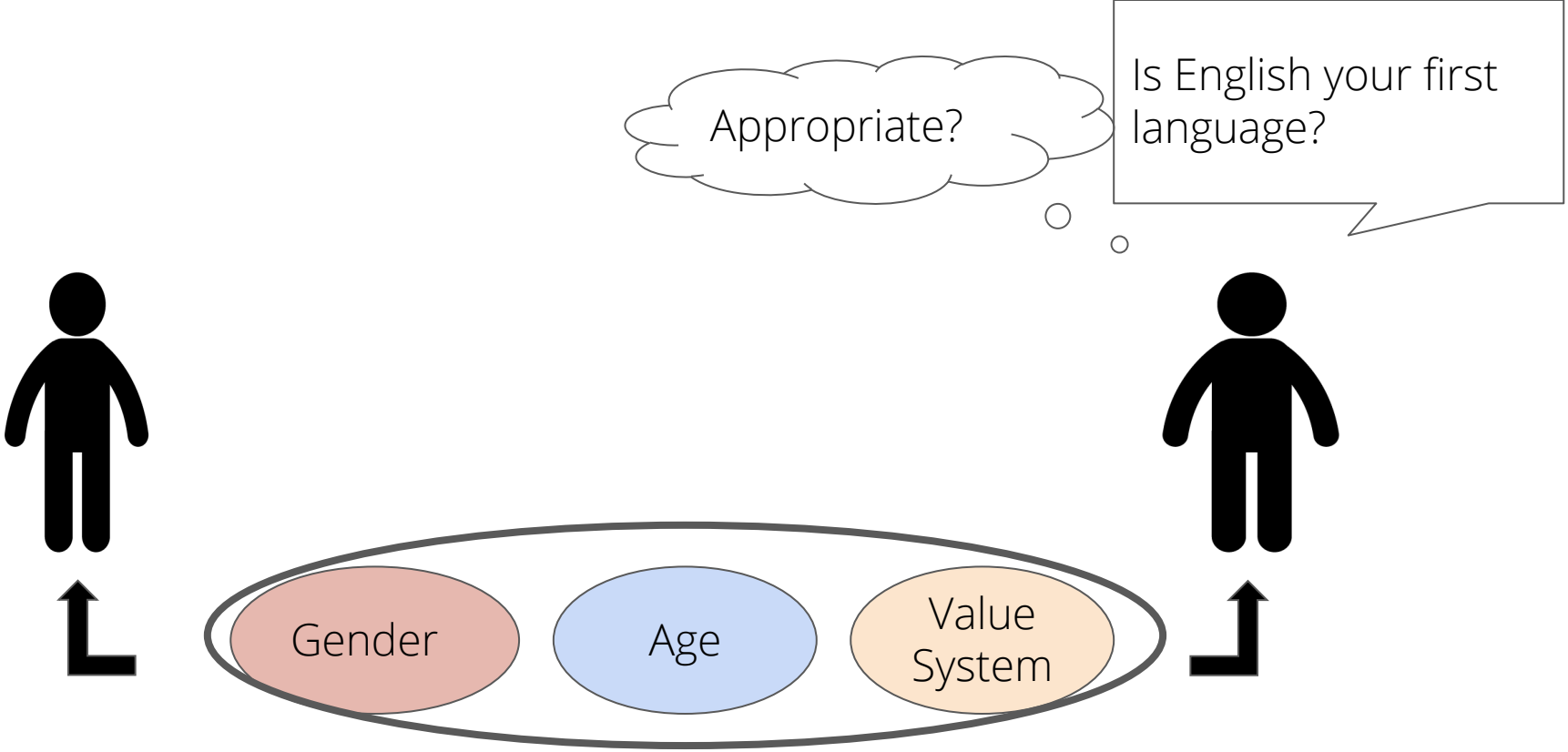
Michael Geraci†
megeraci@buffalo.edu



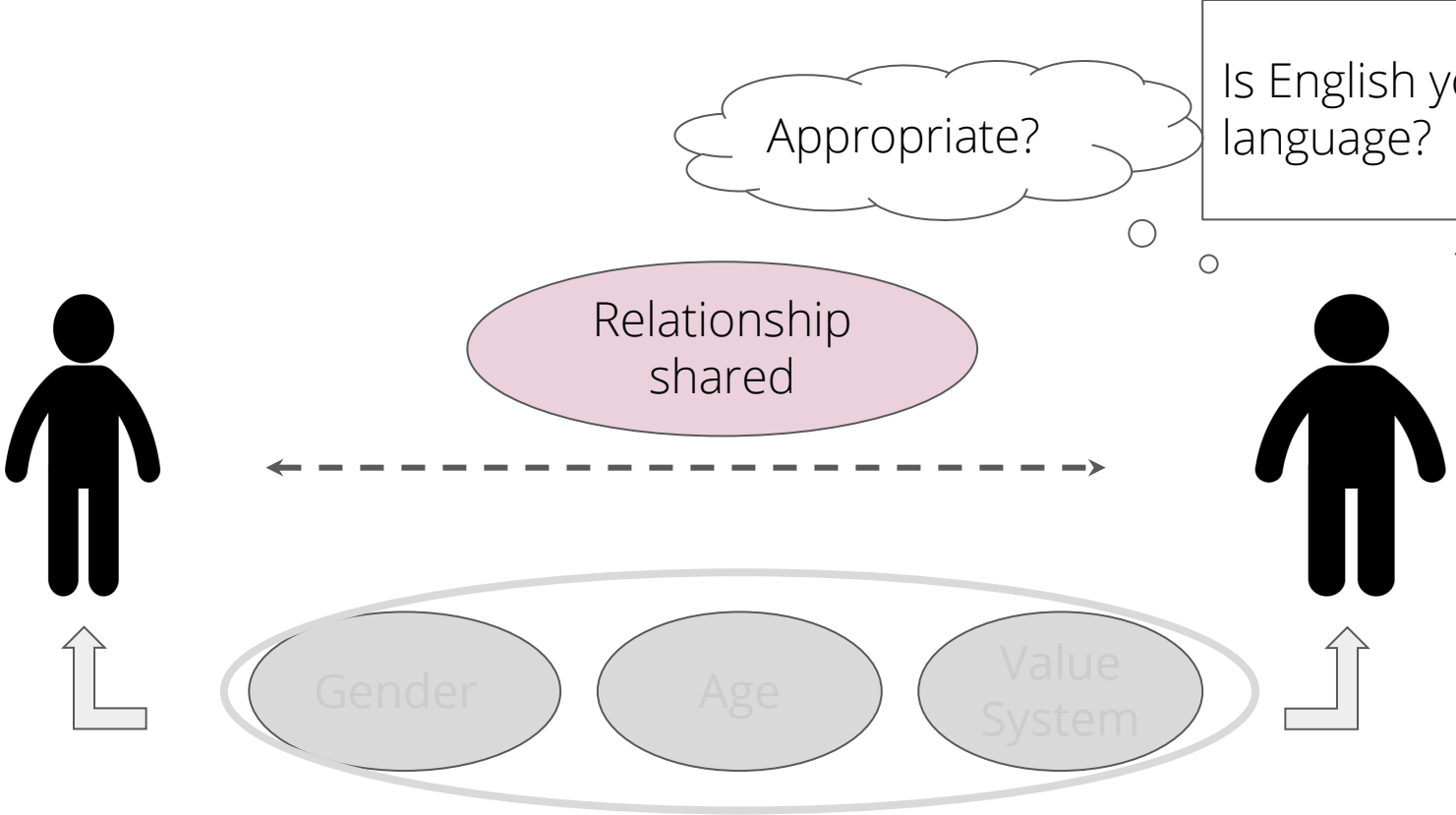
Appropriate?

Is English your first language?





Communication between individuals are a complex functions of multiple factors.



Communication between individuals are a complex functions of multiple factors.

Language has many ways to describe our relationship to others. Each relationship has its own norms!

Family

Social

Romance

Para-Social

Organizational

Peer-group

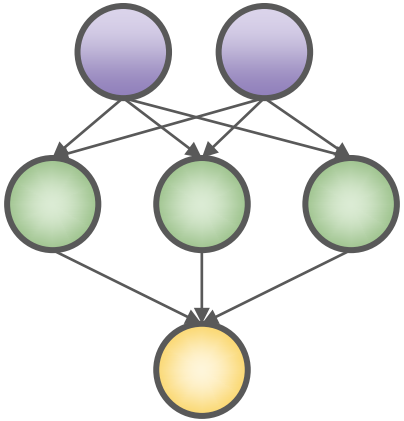
Role-based

Antagonist

Dataset Creation

Building a Model

Apply the model to understand



How much of the conversations are context-sensitive

Its usefulness to model other language behavior in conversations

Dataset Creation

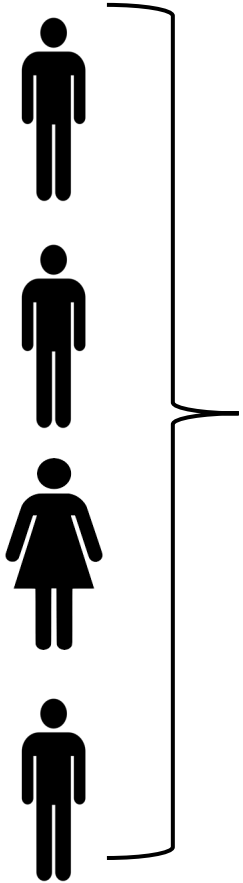
Building a Model

Apply the model to understand

How much of the conversations are context-sensitive

Its usefulness to model other language behavior in conversations

Dataset Creation (Phase 1)



Manually generated 401 sentences such that their appropriateness differs across relationships

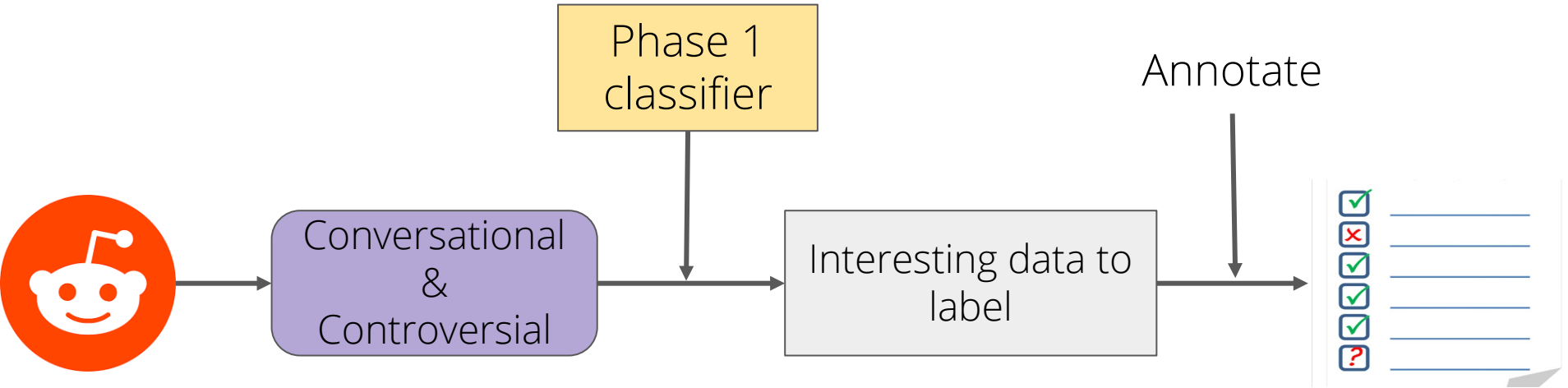
Phase 1 classifier

"You look super cute today"

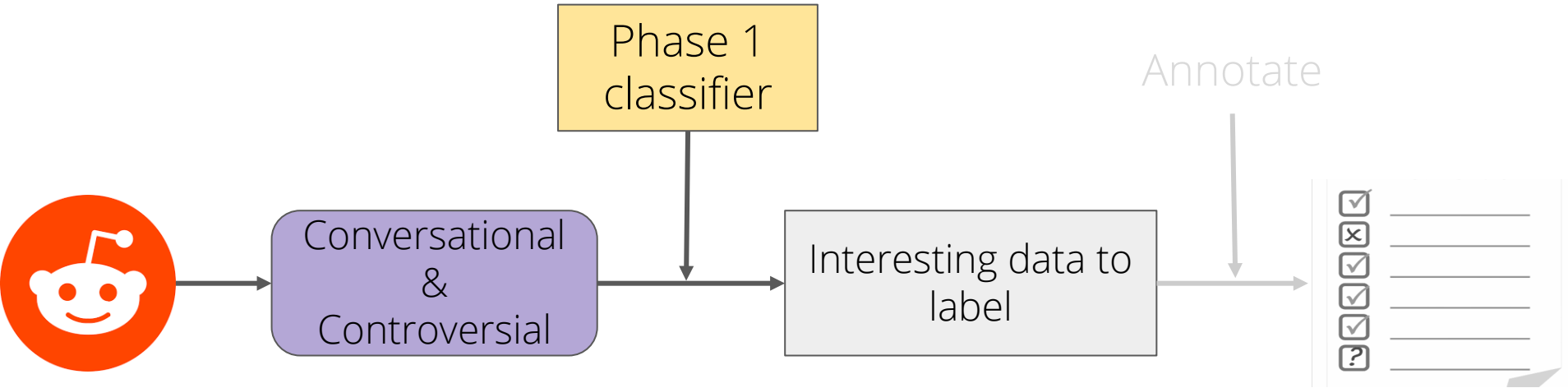
😊 Parent to their child

😬 Boss to their employee

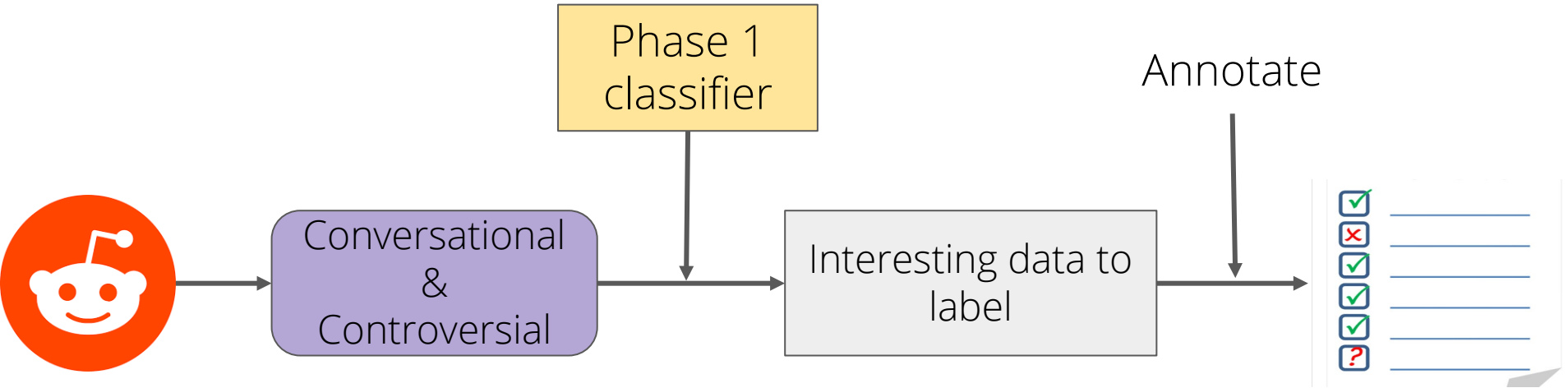
Dataset Creation (Phase 2)



Dataset Creation (Phase 2)



Dataset Creation (Phase 2)



Annotating appropriate is complex!

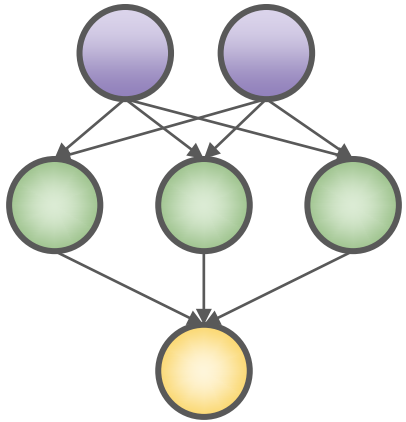
Measuring Annotation Quality

- Two trained annotators labeled 41 contexts, making 2,159 initial appropriateness judgments
- Initial agreement: Kripperdoff's α = 0.46 🤯
- Why did annotators disagree? (in order)
  - Disagreement on message meaning (most cases)
  - Mistakes
  - Disagreement on appropriateness
- Solution: Annotators adjudicate the meaning of each message and then re-annotated
  - Final agreement α = 0.92
- Caveat: One dataset—not necessarily expected to generalize

Dataset Creation

Building a Model

Apply the model to understand



How much of the conversations are context-sensitive

Its usefulness to model other language behavior in conversations

Can computational models identify contextual appropriateness?

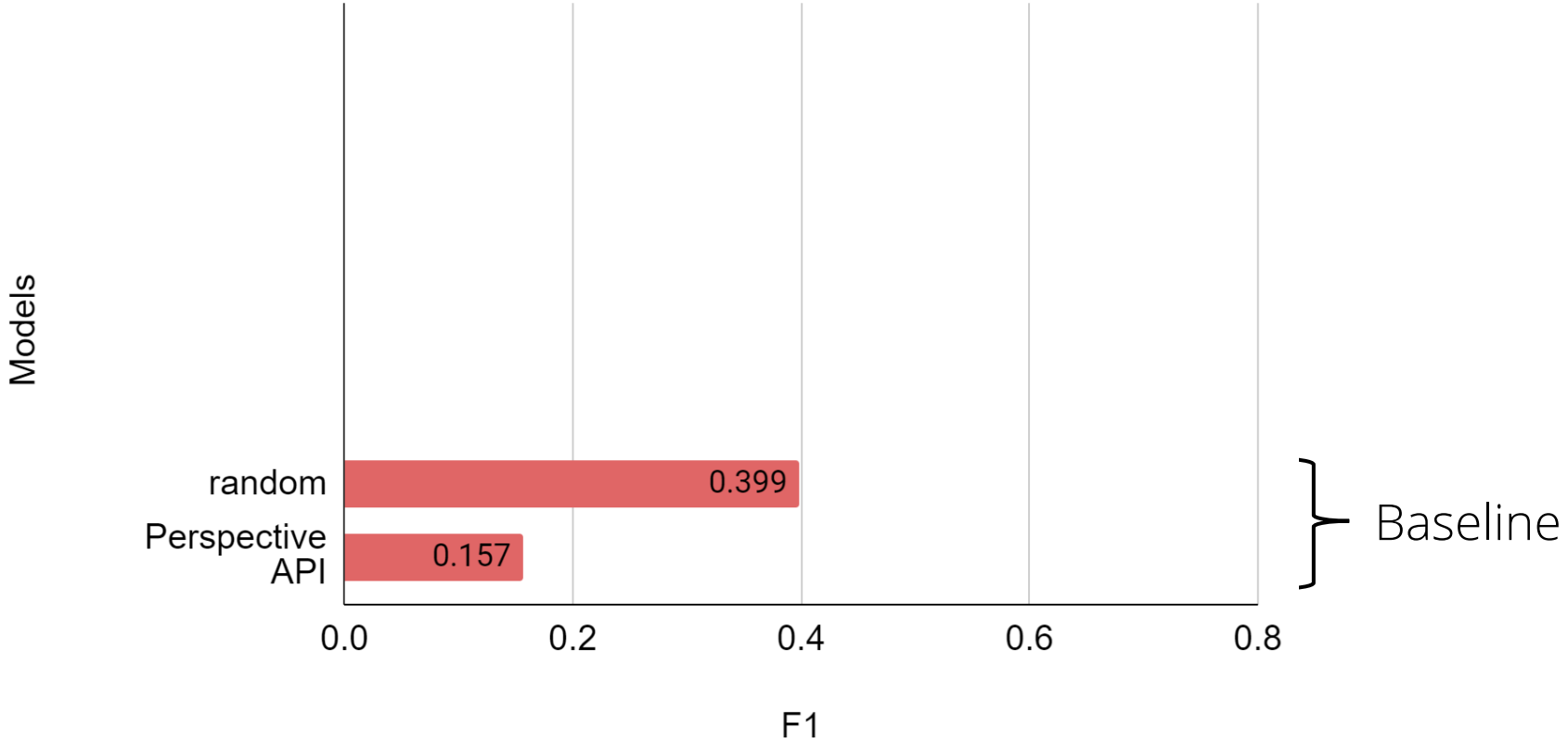
Task: Given the social relationship identify if it would be appropriate to say a particular message

Example: Would asking “Is English your first language?” be considered appropriate between work colleagues

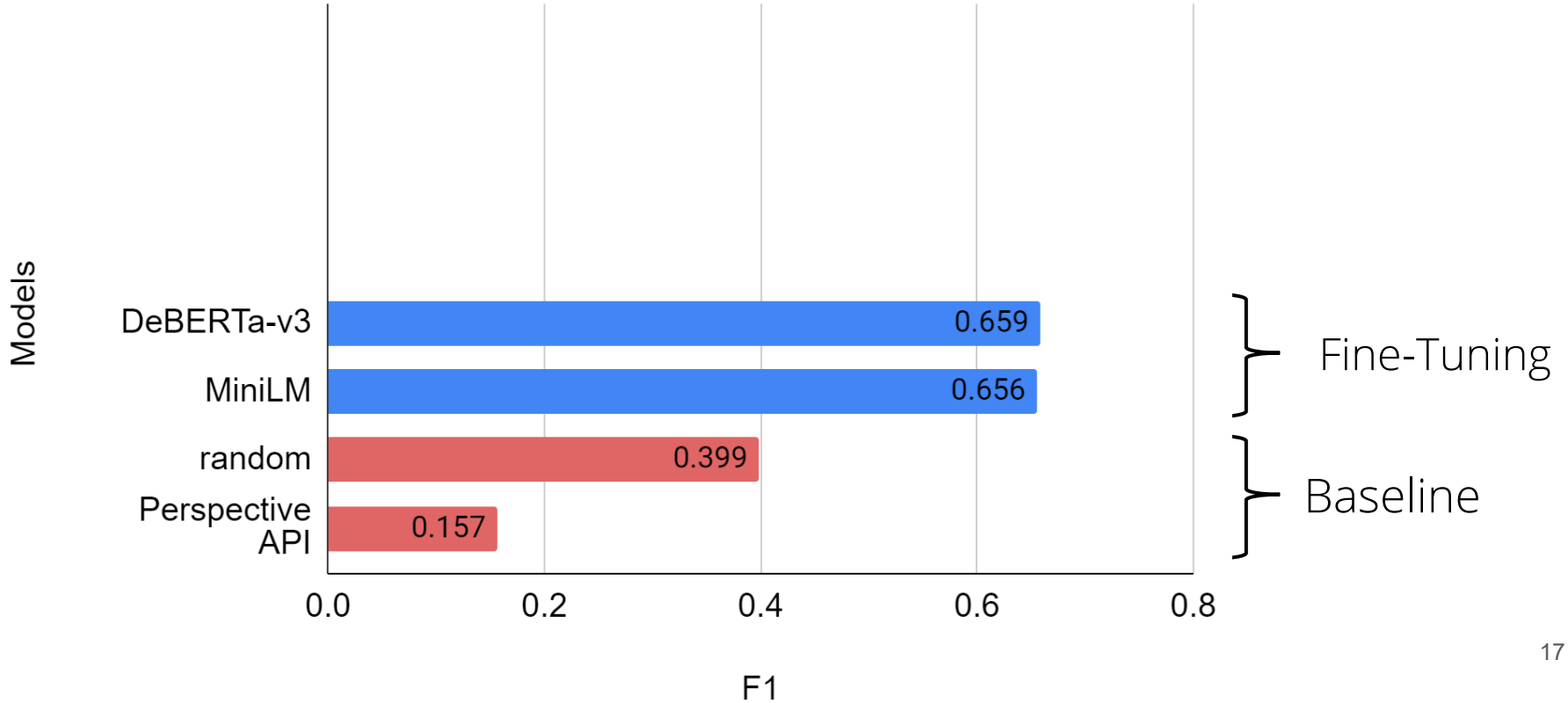
Prompt-based: ‘Is it appropriate for `person1` to say "`quote`" to `person2`, "yes" or "no"? [mask]’
Positive class: Inappropriate

Fine-tuning: Person1 saying “Quote" to person2 is inappropriate. *Label (1/0)*

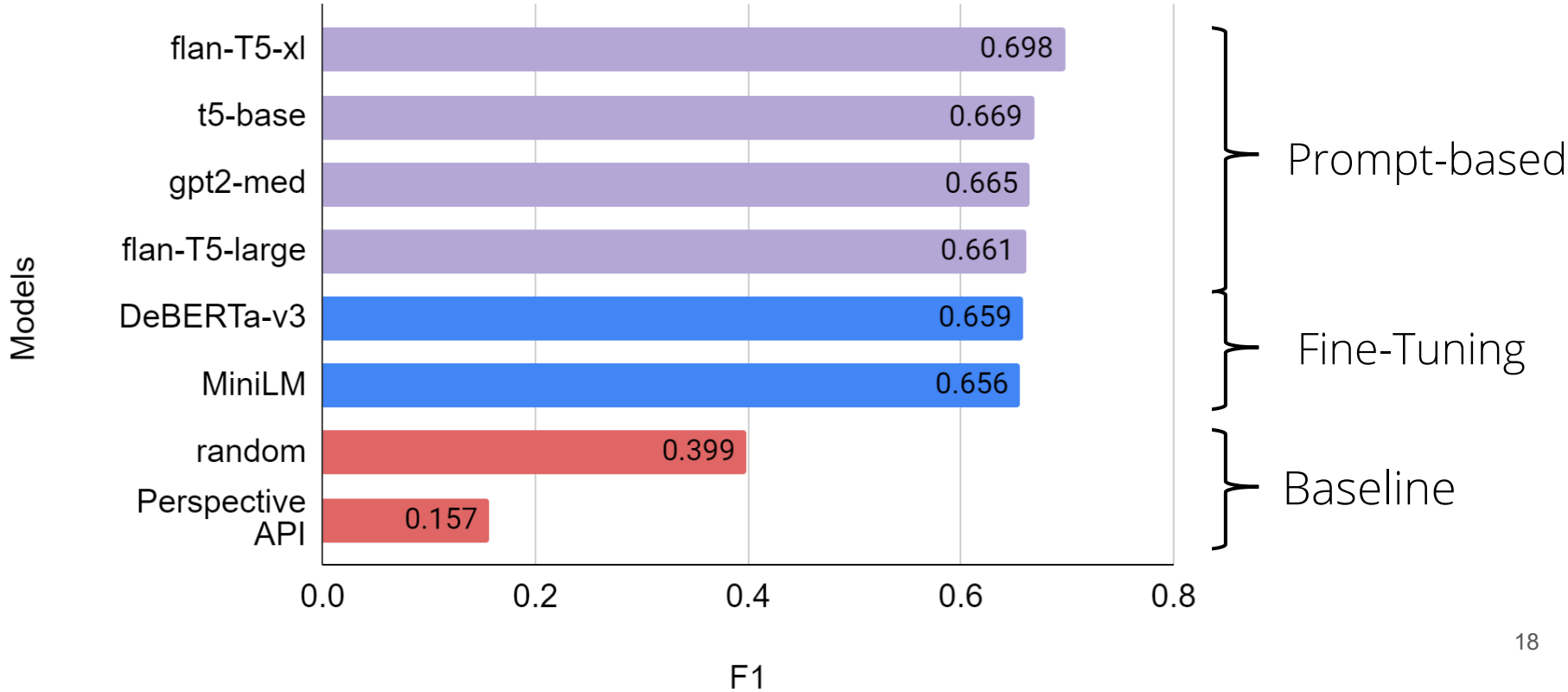
Can computational models identify contextual appropriateness?



Can computational models identify contextual appropriateness?



Can computational models identify contextual appropriateness?



Dataset Creation

Building a Model

Apply the model to understand

How much of the conversations are context-sensitive

Its usefulness to model other language behavior in conversations

How much of the conversations are sensitive to relationship-context?

"When you die do you want to be buried or cremated?"

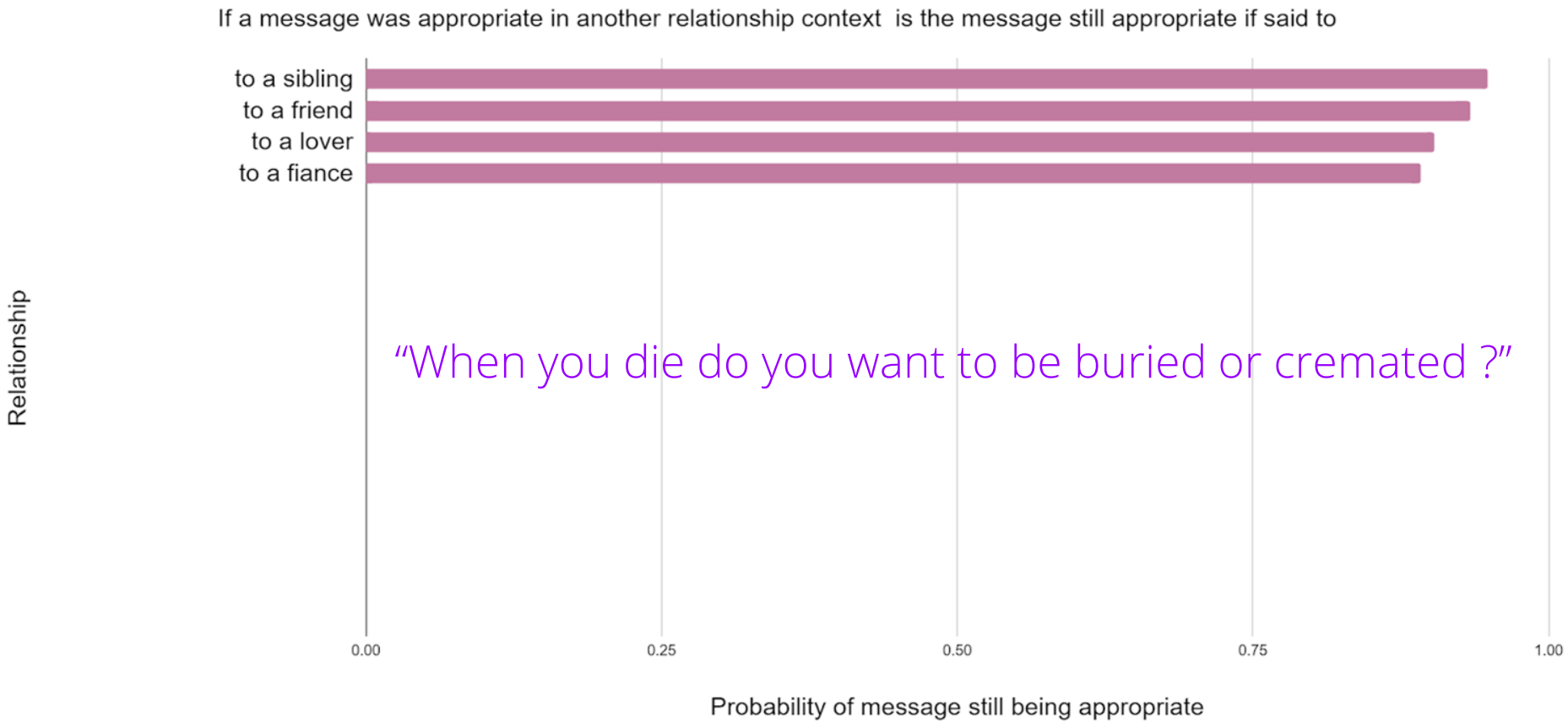
What if we switch context?

Still appropriate?

"When you die do you want to be buried or cremated?"

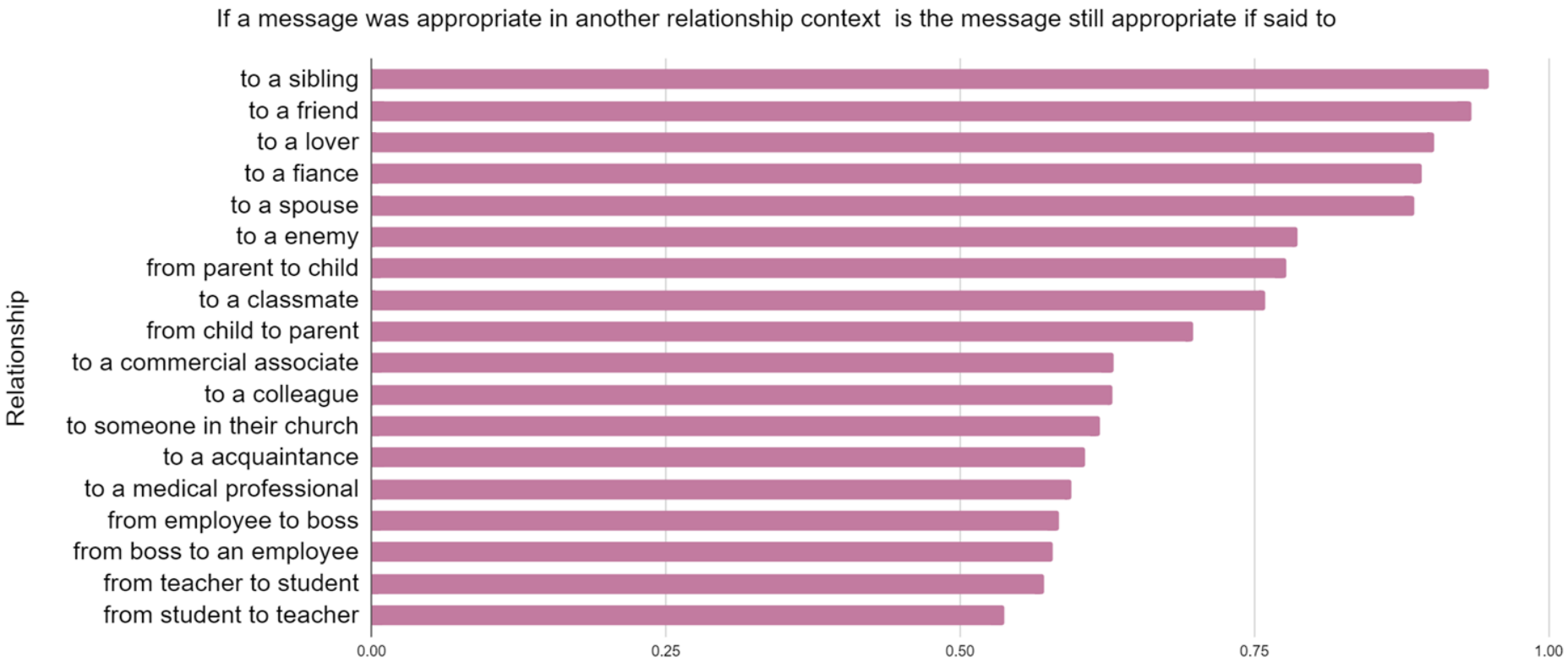
How much of the conversations are sensitive to relationship-context? (PRIDE; Tigunova et al 2021)

- Dialogue dataset with a known relationship between speaker and spoken to.
- Based on movie dialog but filtered down to generic messages
- We use a dataset of 47,801 messages across 18 relationships



"When you die do you want to be buried or cremated ?"

"When you die do you want to be buried or cremated ?"



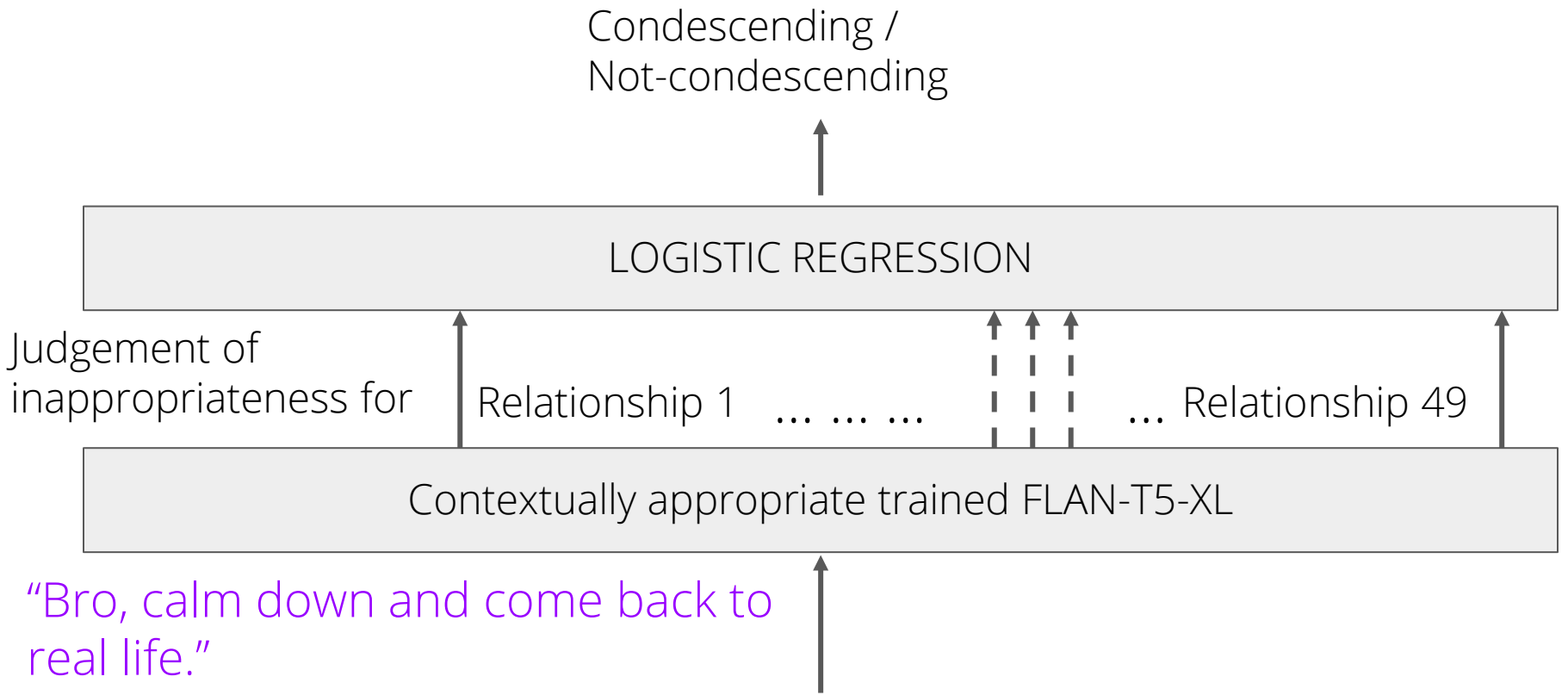
~1/5th of the conversations
change their appropriateness with context

Is appropriateness helpful for modeling other language behaviors? (Talkdown; Wang and Potts 2019)

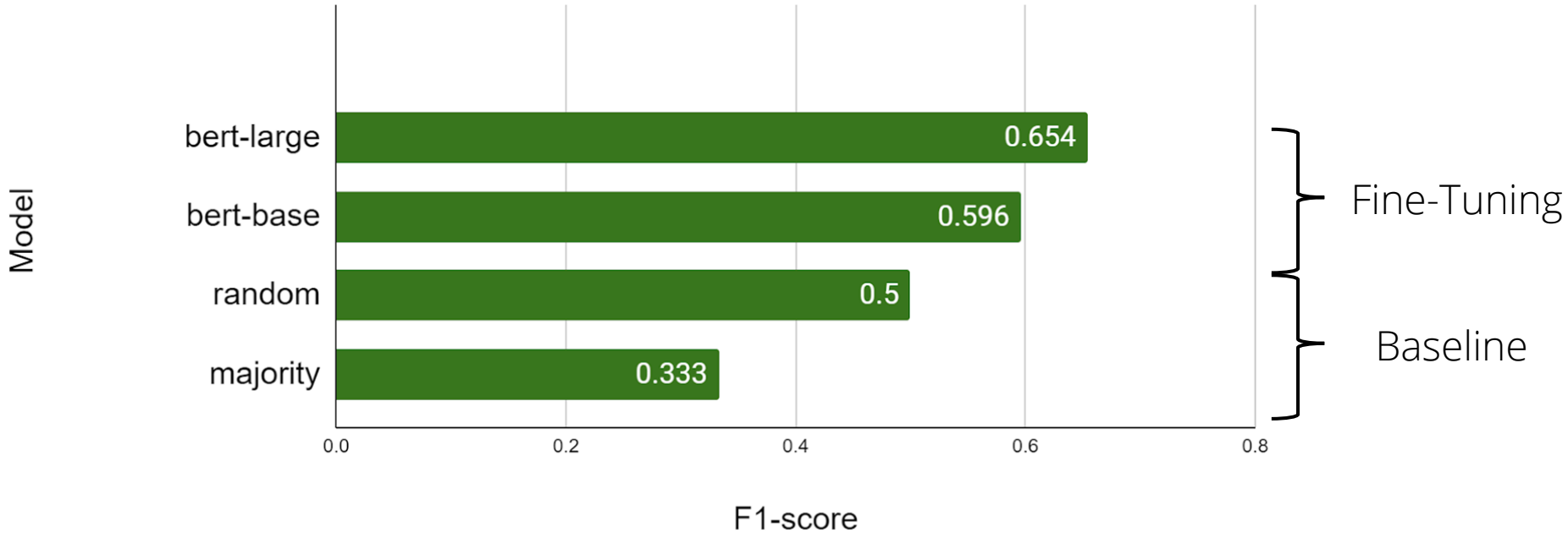
- Aim: Detect condescending utterances.
- Binary Classification task.
- Why:
  - Appropriateness and condescension are related
  - Condescension is inappropriate in hierarchical relationships.

    “Bro, calm down and come back to real life.”

Is appropriateness helpful for modeling other language behaviors?

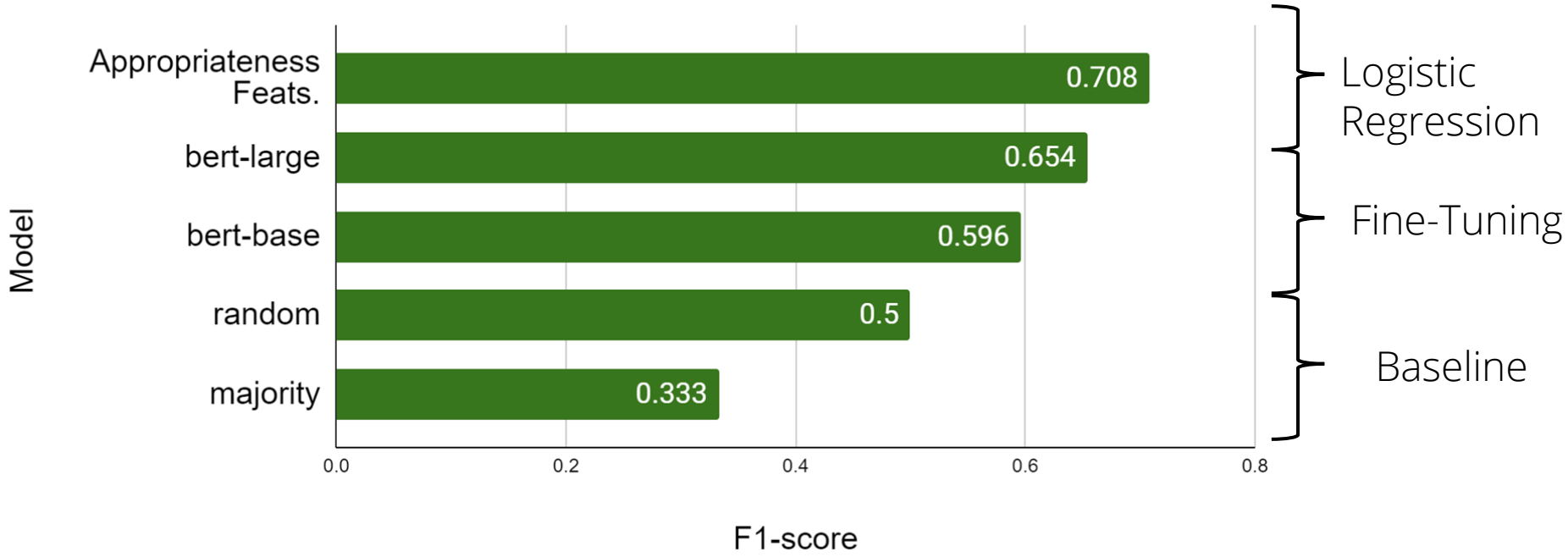


Is appropriateness helpful for modeling other language behaviors?



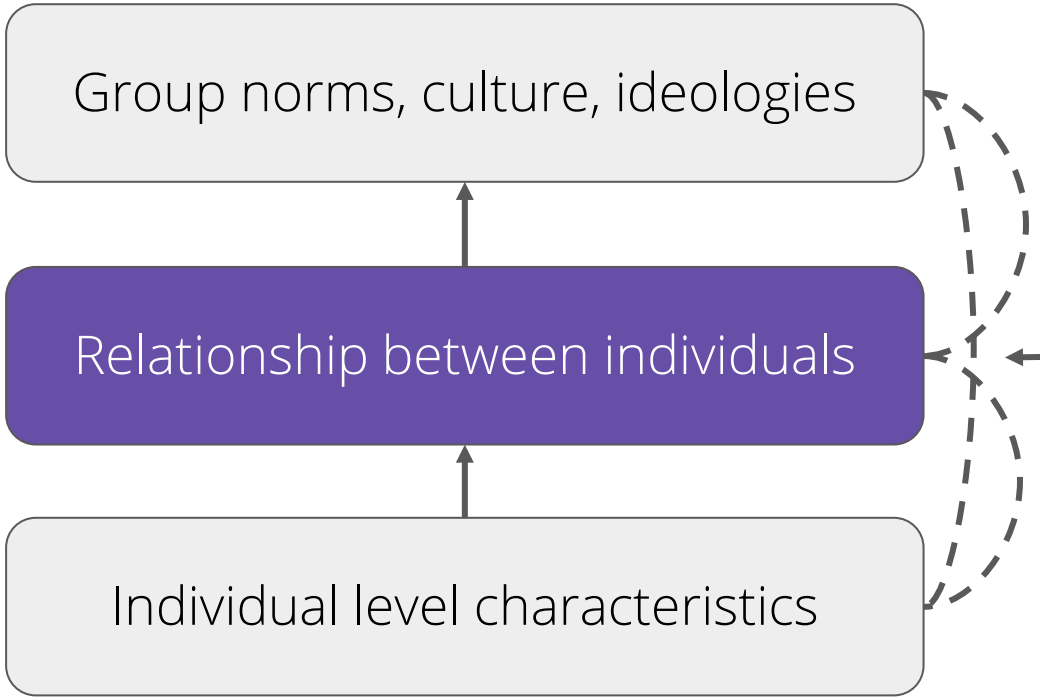
"Bro, calm down and come back to real life."

Is appropriateness helpful for modeling other language behaviors?



"Bro, calm down and come back to real life."

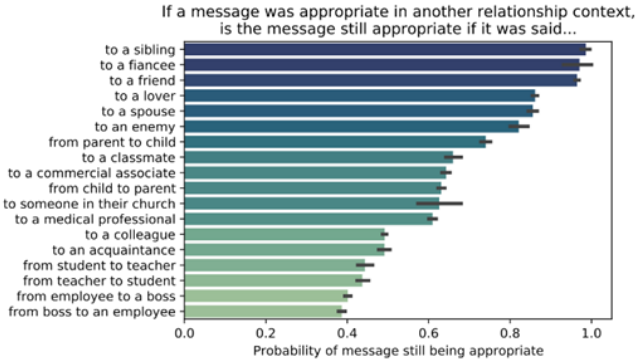
Limitations

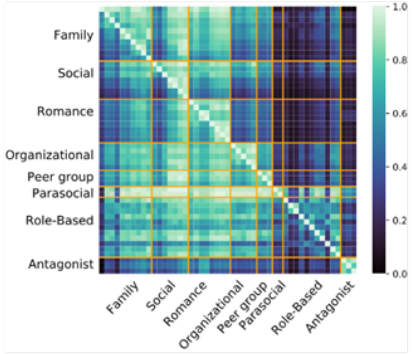


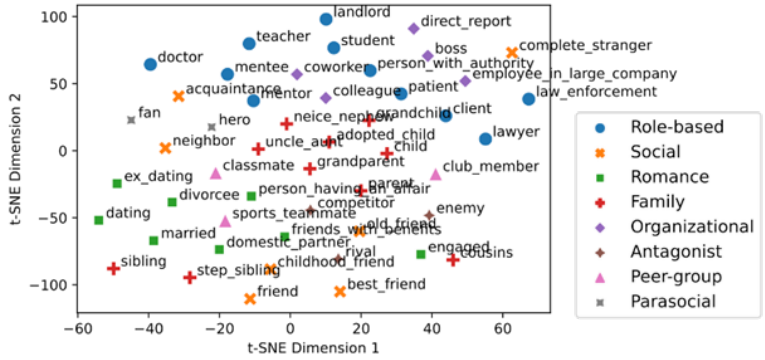
- More relationships to be analyzed; each with more nuance
- More #annotators; diverse viewpoints.

- This work focuses on this aspect

Thank you!!







- We introduce a new dataset of 12K appropriateness judgments across 49 relationships.
- An estimated 1 in 5 messages change their appropriateness in a new context
- Incorporating contextuality helps understand other antisocial behaviors

https://github.com/davidjurgens/contextual-appropriateness

SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

BAB LAB LAB

David Jurgens
jurgens
@umich.edu

Agrima Seth
agrima
@umich.edu

Jackson Sargent
jacsarge
@umich.edu

Athena Aghighi
aaghighi
@ucdavis.edu

Michael Geraci
megeraci
@buffalo.edu