



Stony Brook  
University



# Author's Sentiment Prediction

---

Mohaddeseh Bastan, Mahnaz Koupaee, Youngseo Son, Richard Sicoli and Niranjan Balasubramanian

Stony Brook University

COLING2020

# Agenda

- ❖ Introduction
- ❖ PerSenT
  - Crowdsourcing
  - Quality Control
  - Dataset Characteristics
- ❖ Evaluation
- ❖ Conclusion

# Agenda

## ❖ **Introduction**

## ❖ PerSenT

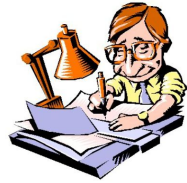
- Crowdsourcing
- Quality Control
- Dataset Characteristics

## ❖ Evaluation

## ❖ Conclusion

# Task definition

Input: a document in news domain



Output: Author's sentiment toward the main entity



# Input

## Target: Lisa Pratt

**Pratt** has a "leave no trace" ethic when it comes to the solar system and **she** isn't happy about the idea of red sports cars floating around aimlessly in space.

Despite his enthusiasm Davis didn't get the six figure job. Instead a more seasoned astrobiologist **Lisa Pratt** stepped into the Planetary Protection Officer role earlier this month.

**Pratt** says **she's** deeply concerned with how to "safely and ethically" look for life on Mars without accidentally killing any extraterrestrial life that might be lurking out there.

...

"I fully expect we will encounter life in our solar system " **she** said.

There might still be life on Mars and **Pratt** doesn't want us to kill it ...

**Pratt's** used to looking for life in really harsh places....

...

**She's** wandered into active gold mines in South Africa to discover what happens in super-hot waters flowing deep below the Earth's surface. **She's** also ventured to freezing cold mines ..

...

**She** says any bits of Earthly trash that get deposited could mess up our chances of a human future on Mars.

....

# Applications

Document Understanding

News analysis: Journalistic Bias

Sentiment Analysis



# Challenges: Irrelevant information

- Non-target information
- Non-author's sentiment

## Target: Lisa Pratt

Pratt has a "leave no trace" ethic when it comes to the solar system and **she isn't happy about the idea of red sports cars floating around aimlessly in space.**

**Despite his enthusiasm Davis didn't get the six figure job. ...**

...

# Challenges: Discourse Gap

- Long distance discourse connections

**Target: Lisa Pratt**

**Pratt** says **she's** deeply concerned with how to "safely and ethically" look for life on Mars **without** **accidentally killing any extraterrestrial life** that might be lurking out there.



# Paragraph Importance?

- Different levels of importance for different paragraphs

## Target: Lisa Pratt

..

Despite his enthusiasm Davis didn't get the six figure job. Instead a more seasoned astrobiologist **Lisa Pratt** stepped into the Planetary Protection Officer role earlier this month.

**Pratt** says **she's** deeply concerned with how to "safely and ethically" look for life on Mars without accidentally killing any extraterrestrial life that might be lurking out there.

...

"I fully expect we will encounter life in our solar system " **she** said.

...

**She's** wandered into active gold mines in South Africa to discover what happens in super-hot waters flowing deep below the Earth's surface. **She's** also ventured to freezing cold mines ..

...

# Agenda

❖ Introduction

❖ **PerSenT**

➤ Crowdsourcing

➤ Quality Control

➤ Dataset Characteristics

❖ Evaluation

❖ Conclusion

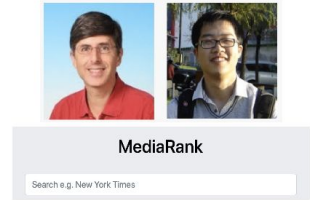
# PerSenT: Person SenTiment



News articles manually annotated for opinions, beliefs, emotions, sentiments, speculations on sentence level



TAC 2014 KBP English sentiment slot filling challenge dataset



Ranking news websites based on the news contents

# Agenda

- ❖ Introduction
- ❖ PerSenT
  - **Crowdsourcing**
  - Quality Control
  - Dataset Characteristics
- ❖ Evaluation
- ❖ Conclusion

# Crowdsourcing: Interface

- ❖ What is the paragraph-level sentiment?
- ❖ What is the document-level sentiment?
- ❖ Who's the main entity?

Sentiment Analysis Instructions (Click to expand)

Read the article below and determine each paragraph's view and the whole article's view towards **Edgar Martinez**. At the end, determine whether the article is about **Edgar Martinez** or not.

**Column: Baseball is all about numbers so is Hall of Fame**

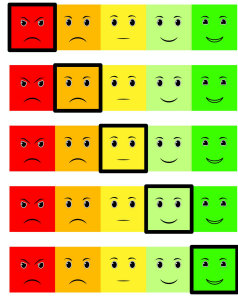
Consider the case of <b>Edgar Martinez</b> who baseball writers for years treated as a marginal candidate for induction. The fact he was a DH most of his career worked against him as did his relative lack of home runs and playing most of his career in relative obscurity in Seattle.	<input type="radio"/> Negative	<input type="radio"/> Slightly Negative	<input type="radio"/> Neutral	<input type="radio"/> Slightly Positive	<input type="radio"/> Positive
But <b>Martinez</b> surged in voting the last two years propelled by supporters that used WAR and OPS to make <b>his</b> case. <b>He</b> received 70.4 percent of the latest vote falling just 20 votes shy of joining Jones Thome Vladimir Guerrero and Trevor Hoffman in this year's class.	<input type="radio"/> Negative	<input type="radio"/> Slightly Negative	<input type="radio"/> Neutral	<input type="radio"/> Slightly Positive	<input type="radio"/> Positive
"We are trending up " <b>Martinez</b> tweeted. "Next year may be the year."	<input type="radio"/> Negative	<input type="radio"/> Slightly Negative	<input type="radio"/> Neutral	<input type="radio"/> Slightly Positive	<input type="radio"/> Positive
But is the fact that <b>Martinez</b> is tied for 32nd in history for OPS enough to get <b>him</b> in the Hall of Fame? Or is it more important to have the kind of impact <b>DiMaggio</b> had on a Yankees team that won a staggering nine World Series titles in his 13 years in pinstripes?	<input type="radio"/> Negative	<input type="radio"/> Slightly Negative	<input type="radio"/> Neutral	<input type="radio"/> Slightly Positive	<input type="radio"/> Positive
<b>Martinez</b> may indeed make it on <b>his</b> 10th and final year of eligibility. Next year's newcomers won't crowd ballots too much with Mariano Rivera looking to be the only first-ballot pick. There's a good chance both <b>Martinez</b> and Mussina pick up votes.	<input type="radio"/> Negative	<input type="radio"/> Slightly Negative	<input type="radio"/> Neutral	<input type="radio"/> Slightly Positive	<input type="radio"/> Positive
<b>The whole article's view towards Edgar Martinez is:</b>	<input type="radio"/> Negative	<input type="radio"/> Slightly Negative	<input type="radio"/> Neutral	<input type="radio"/> Slightly Positive	<input type="radio"/> Positive

Is this article about **Edgar Martinez** ?

Yes

No

# Crowdsourcing: Interface



paragraph-level sentiment

Taylor might be distancing herself from Kasich now that she's seeking endorsements from Ohio's conservatives for her gubernatorial campaign but that wasn't always the case.

Negative

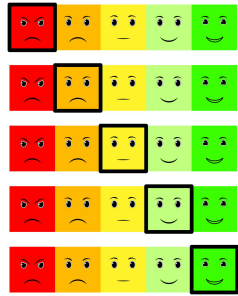
Slightly Negative

Neutral

Slightly Positive

Positive

# Crowdsourcing: Interface



document-level sentiment

The whole article's view  
towards **Mary Taylor** is:

Negative
Slightly Negative
Neutral
Slightly Positive
Positive

# Crowdsourcing: Interface

Target entity

Is this article about  
**Mary Taylor** ?

Yes

No



# Agenda

- ❖ Introduction
- ❖ PerSenT
  - Crowdsourcing
  - **Quality Control**
  - Dataset Characteristics
- ❖ Evaluation
- ❖ Conclusion

# Inter annotator Agreement?

Different quality control methods for annotation process is described in the paper

## **Weighted Fleiss's Kappa:**

Penalty for each dissimilar classification is based on the distance between two classes

### **Agreement:**

0.79 for paragraph-level annotations

0.81 for document-level annotations

Also in 90% of the instances, at least two out of three annotators agree with each other.

# Agenda

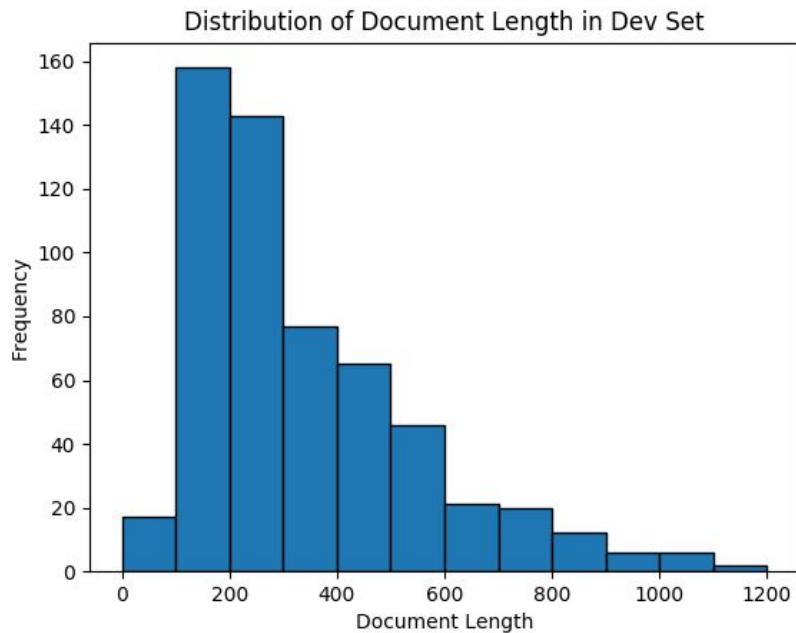
- ❖ Introduction
- ❖ PerSenT
  - Crowdsourcing
  - Quality Control
  - **Dataset Characteristics**
- ❖ Evaluation
- ❖ Conclusion

# Dataset Characteristics

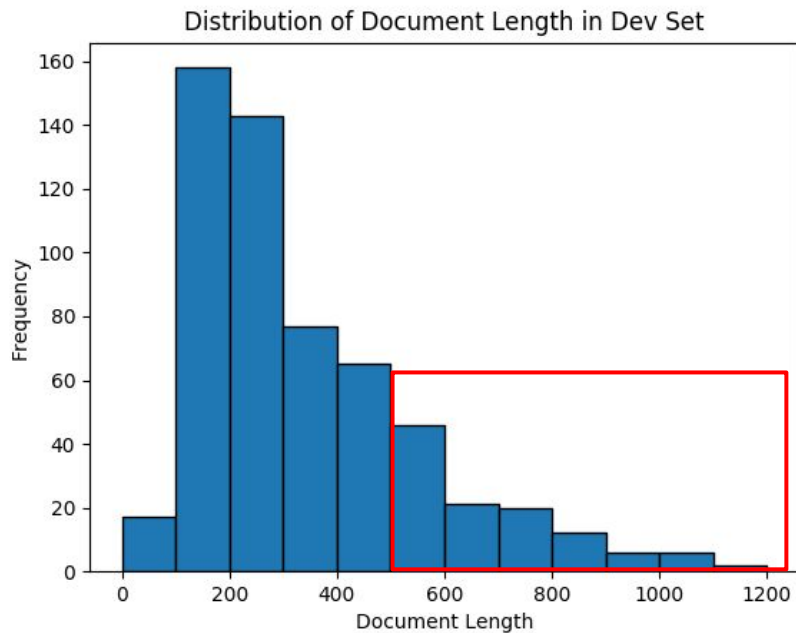
Few popular entities dominate news collection: 4 entities are main entity in nearly 800 articles

Dataset	Document	Paragraphs	Unique Entities
Train	3355	23953	2376
Dev	578	4169	417
Test Standard	579	4261	426
Test Frequent	827	5969	4
Total	5339	38352	3223

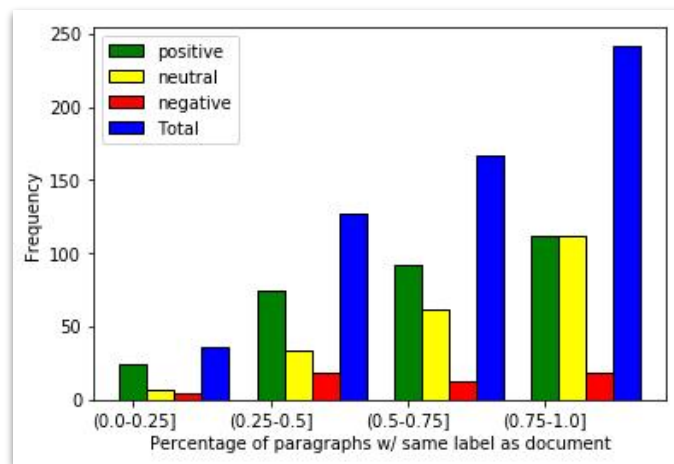
# Extremely long documents



# Extremely long documents



# Label agreement



# Agenda

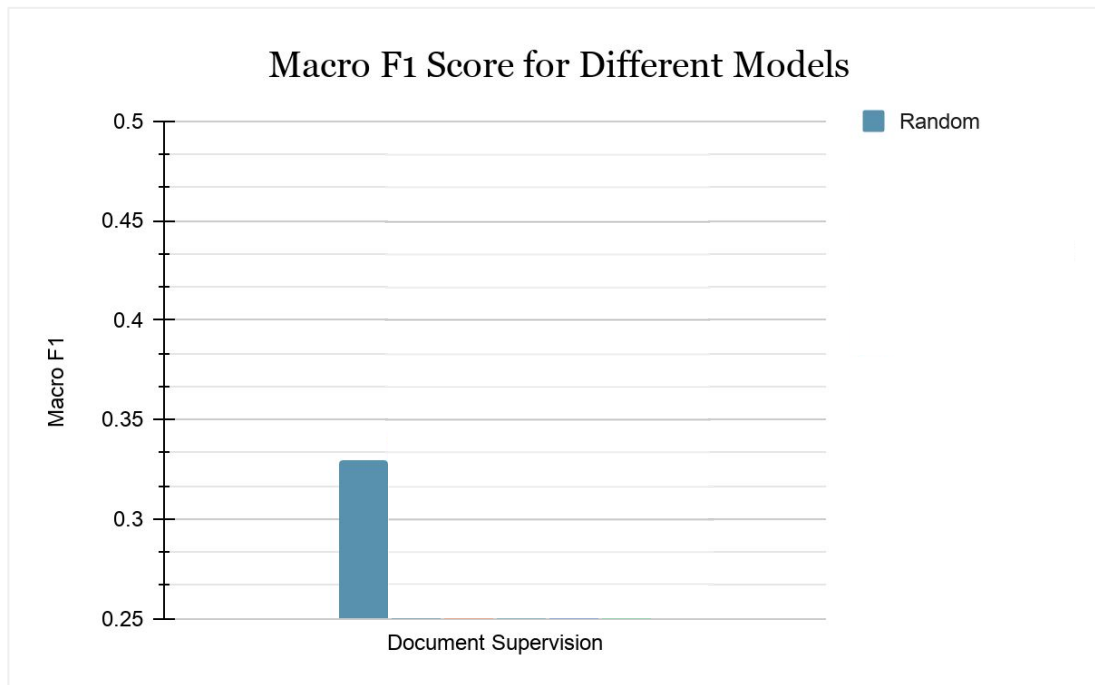
- ❖ Introduction
- ❖ PerSenT
  - Crowdsourcing
  - Quality Control
  - Dataset Characteristics
- ❖ **Evaluation**
- ❖ Conclusion



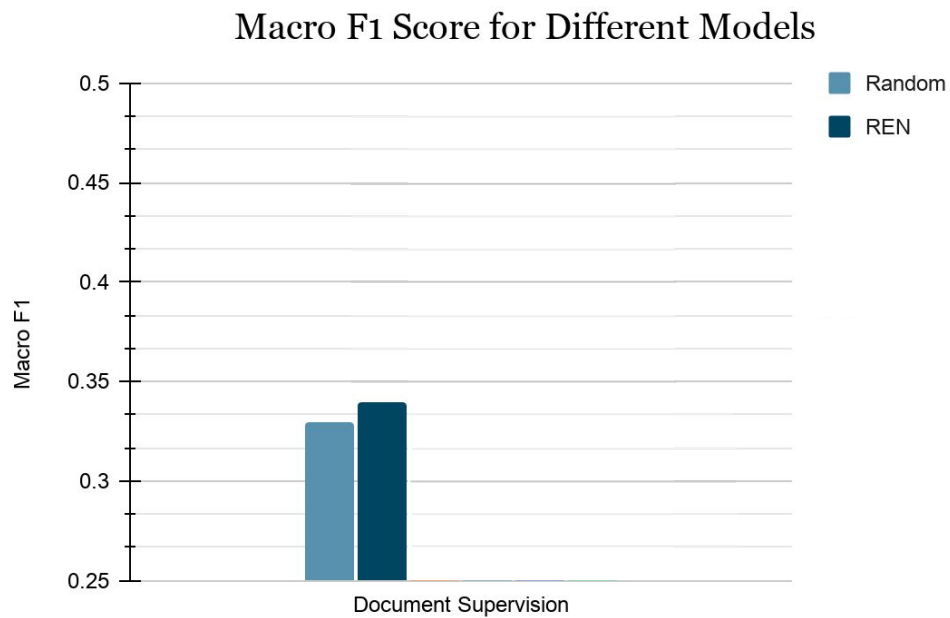
# Baselines

- ❖ Entity-based
  - Modified Recurrent Entity Network
  - Deep Averaging Network
- ❖ Discourse-based: aggregate information from different parts
  - Hierarchical Discourse-based LSTM
- ❖ Transformer model
  - BERT w/ top 3 layers fine-tuning.

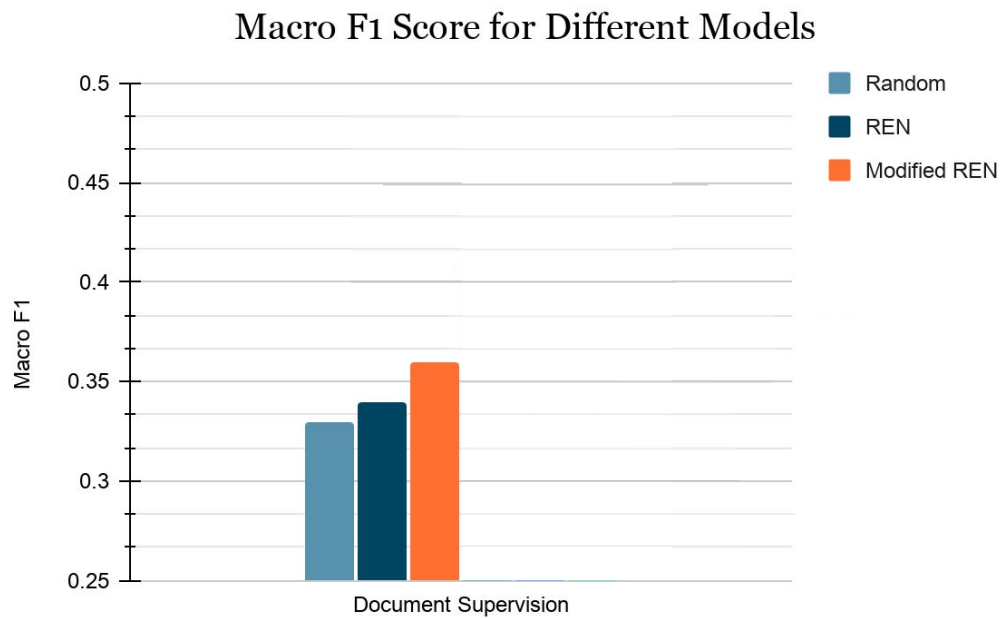
# Results



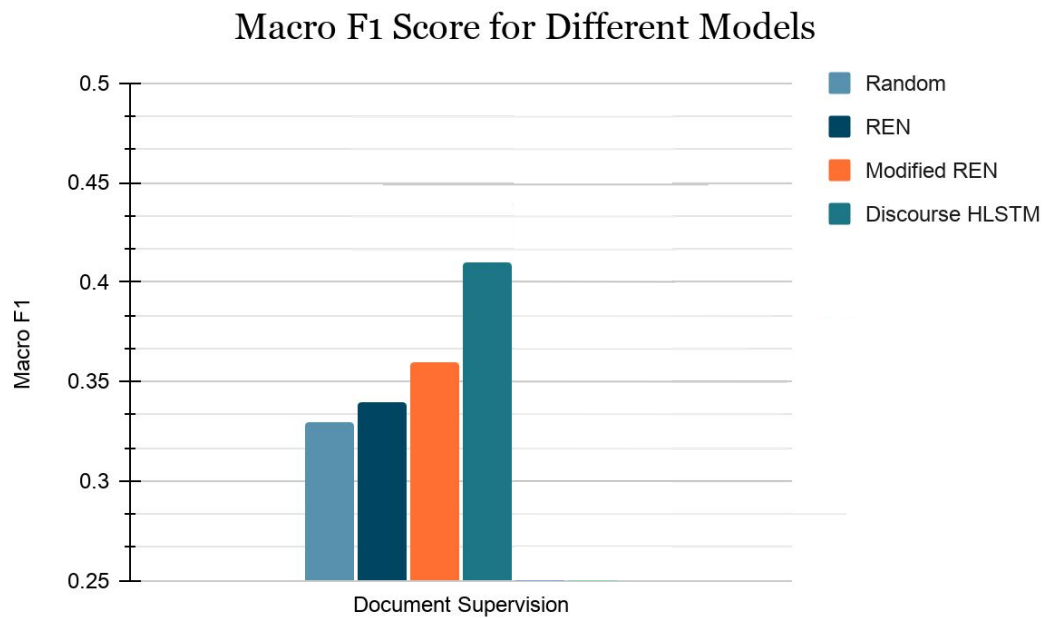
# Results



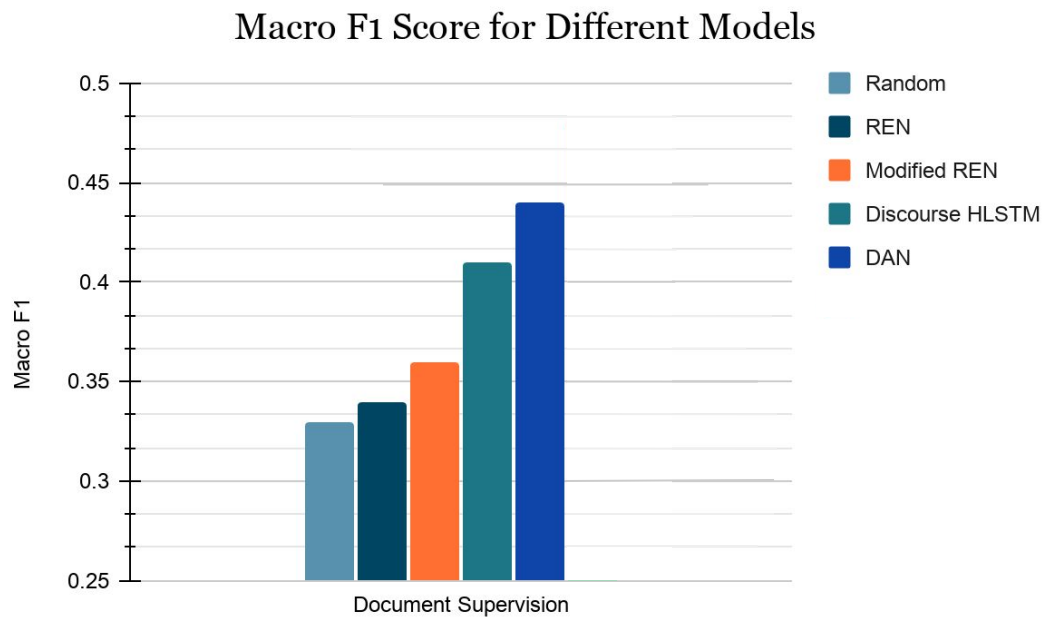
# Results



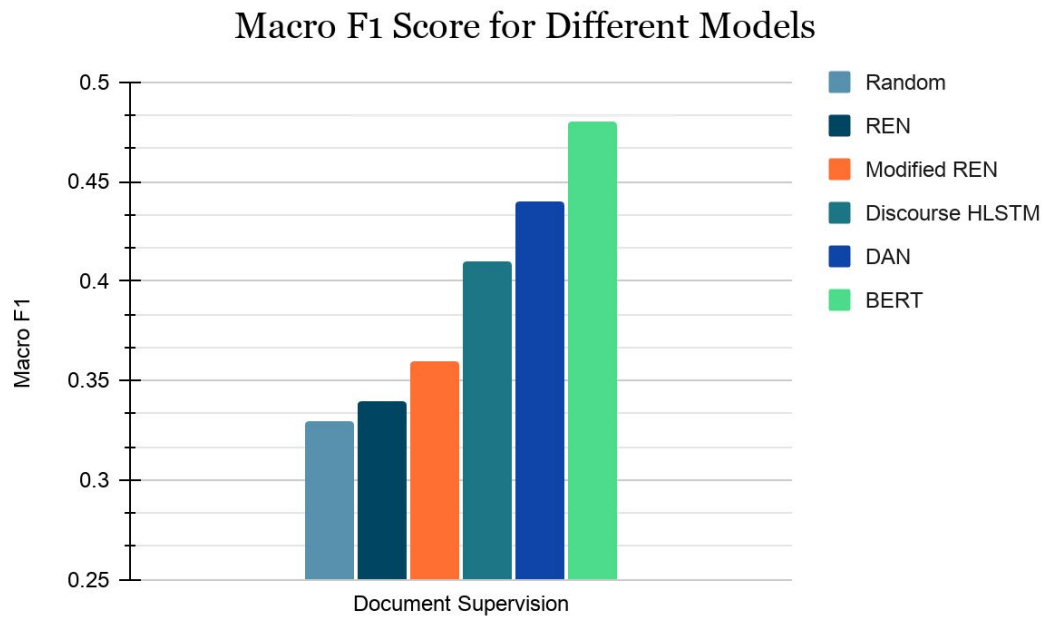
# Results



# Results



# Results



# Focused Representations

## Paragraph-level supervision

- Fine-tune BERT on paragraph-level task first.
- Fine-tune further on the document-level task.



# Focused Representations

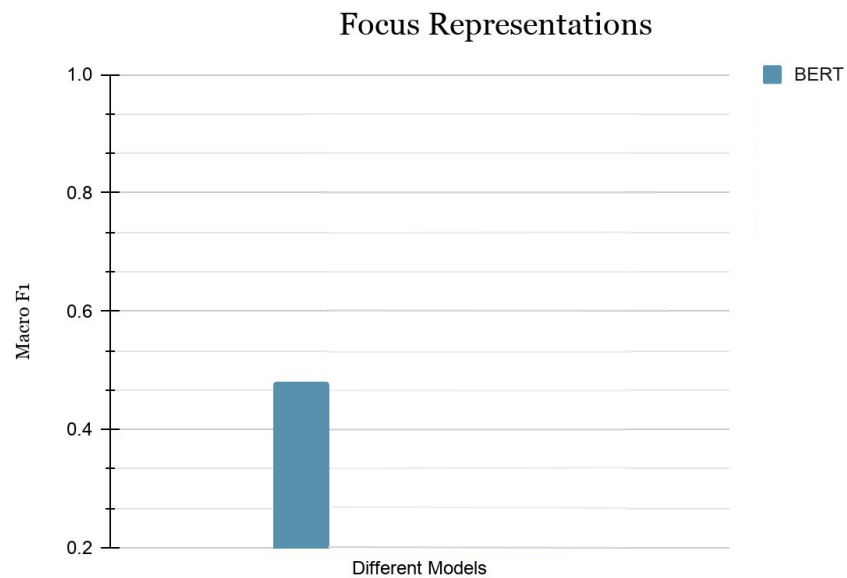
## Paragraph-level supervision

- Fine-tune BERT on paragraph-level task first.
- Fine-tune further on the document-level task.

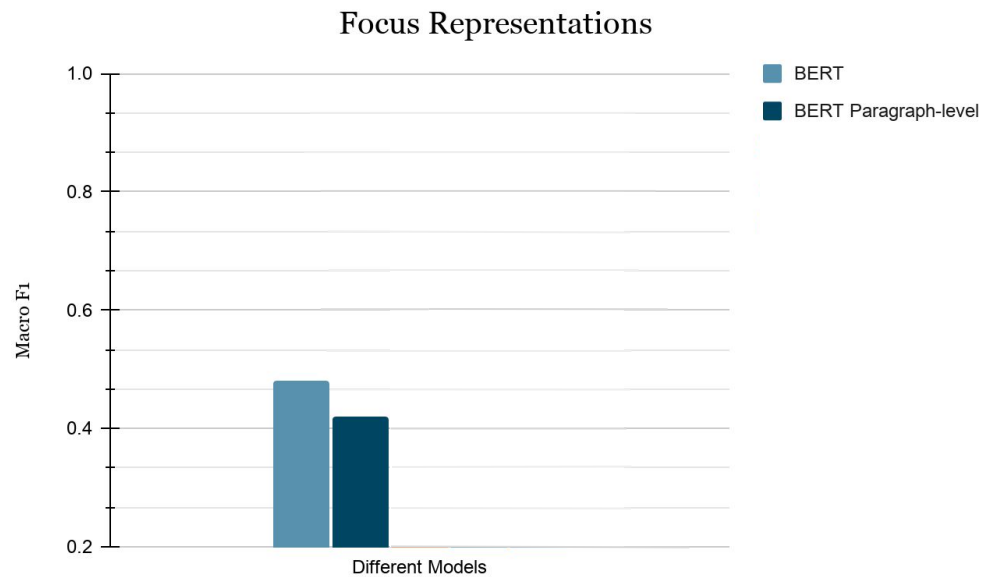
## Focusing towards the target entity

- Mask entities (target and non-target) and ask BERT to predict if MASK is target.
- Fine-tune pre-trained BERT on this ELM task.
- Sequential: Fine-tune further on the document-level task.
- Multi-task: Use ELM task while fine-tuning on document-level task.

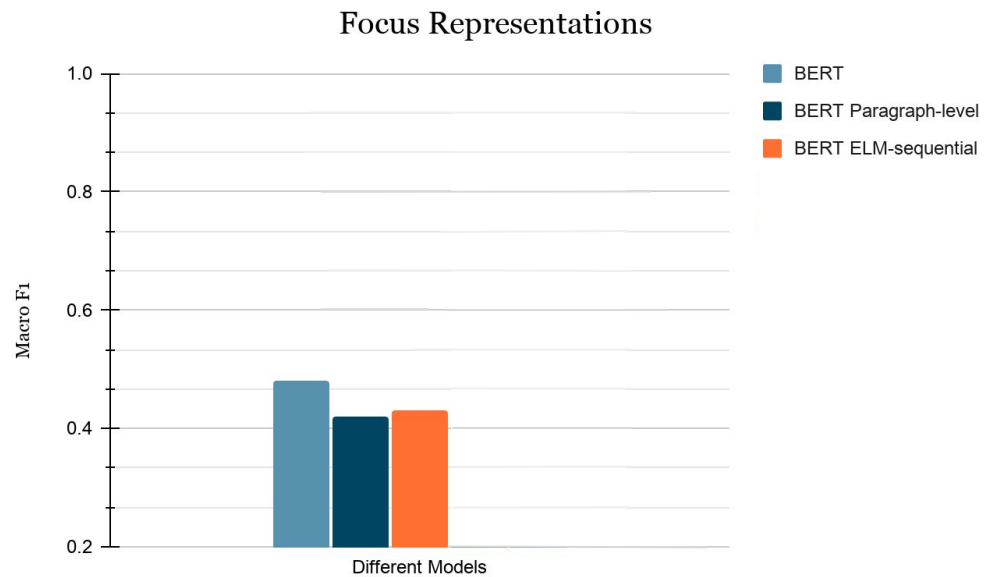
# Results



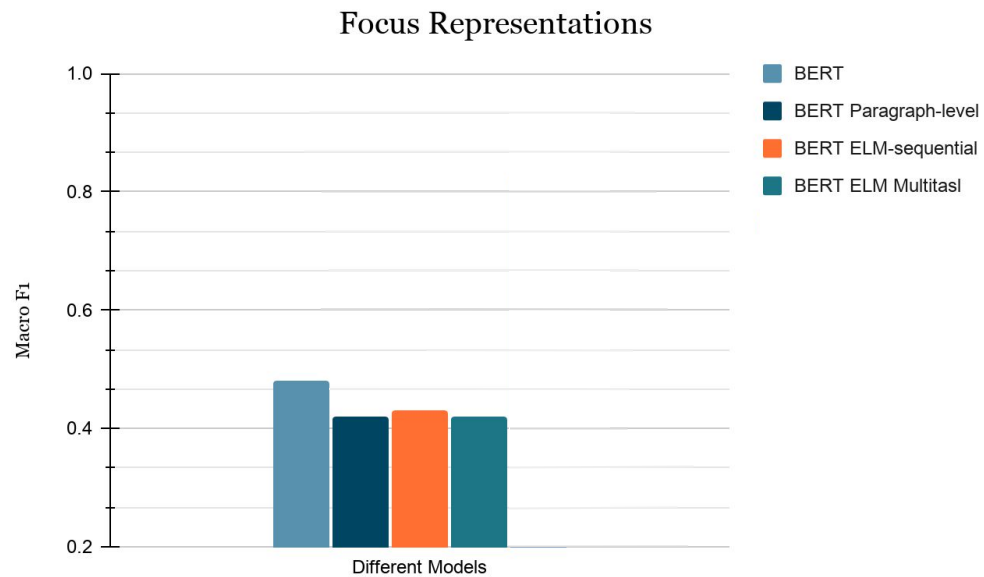
# Results



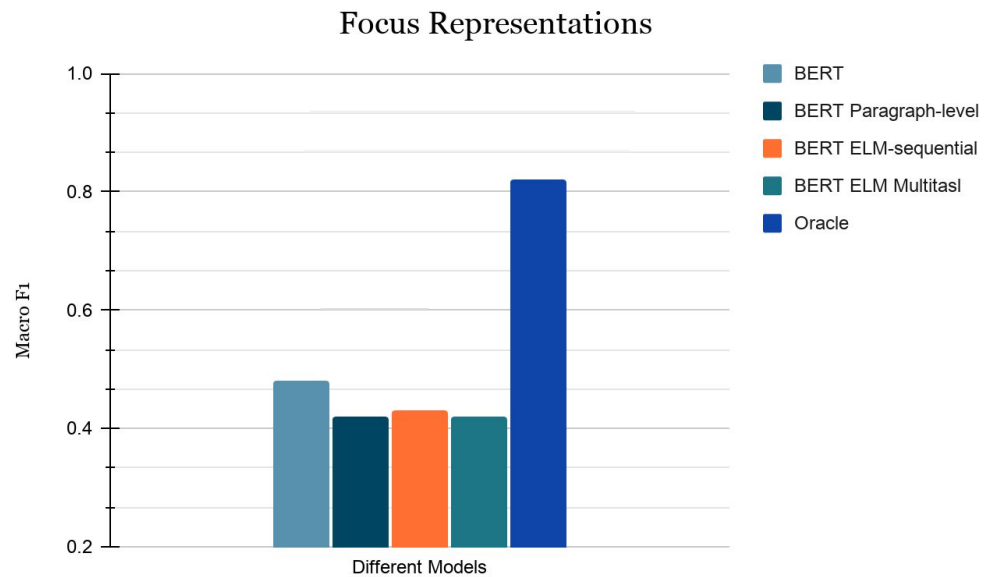
# Results



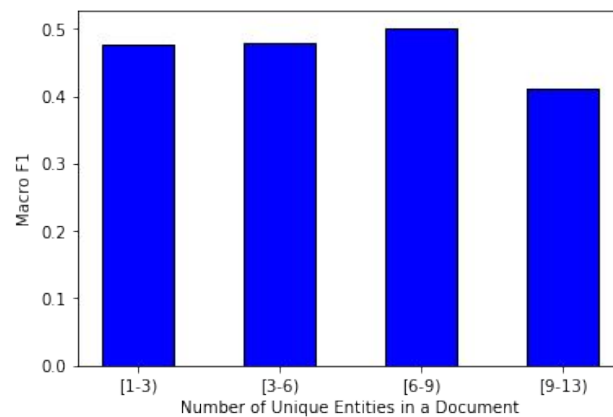
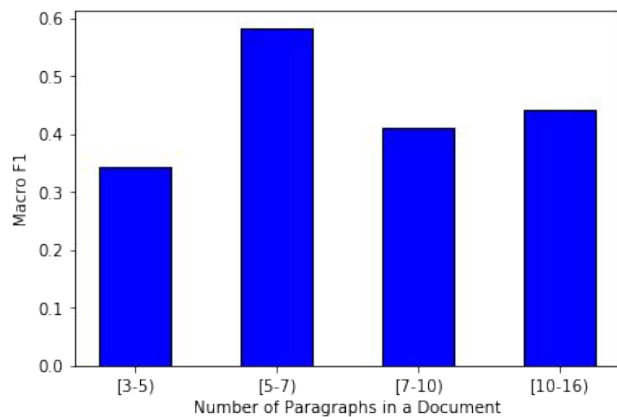
# Results



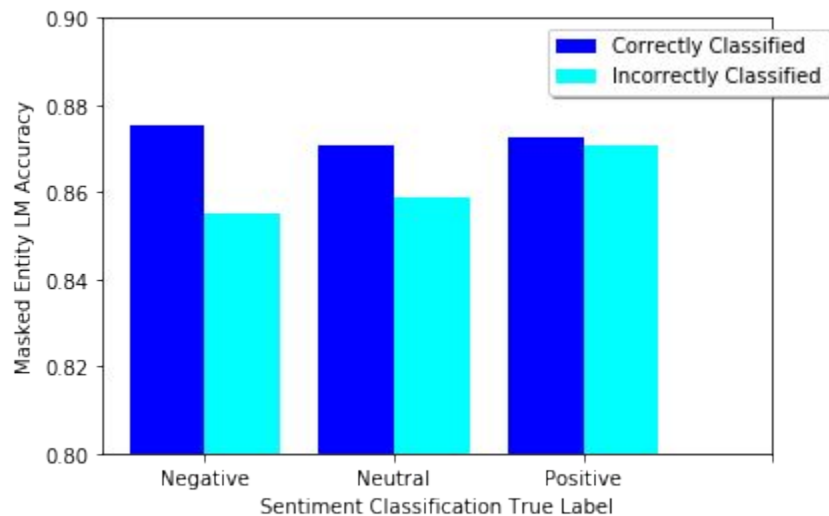
# Results



# Effect of # of paragraphs and unique entities

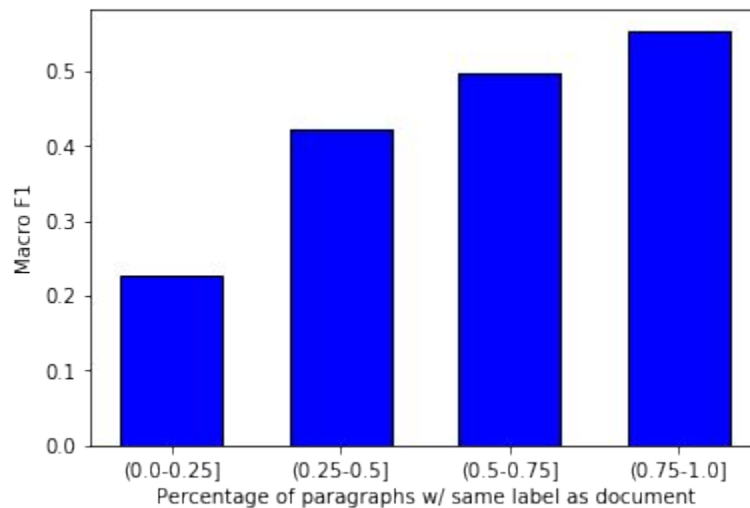


# Does Entity-LM performance correlate with task accuracy?

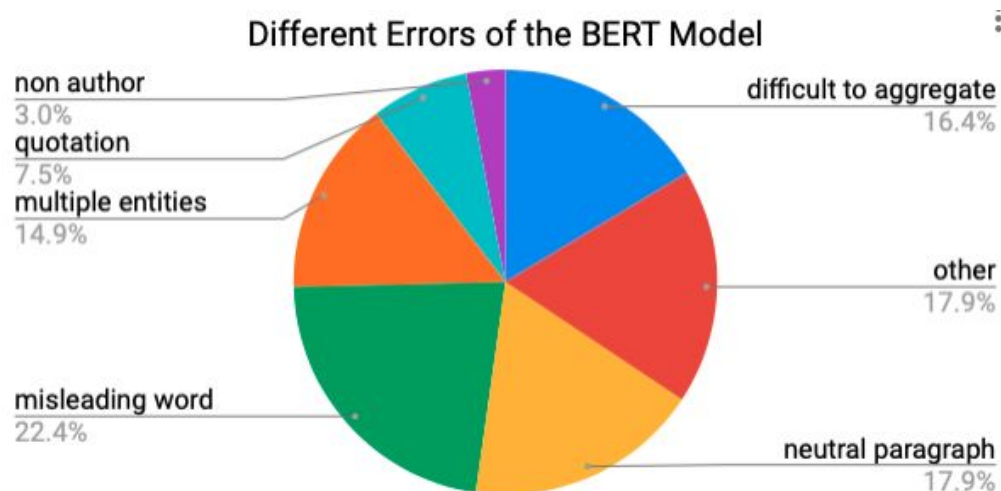




# How useful is paragraph-level agreement?



# Distribution of Errors in Different Categories



# Agenda

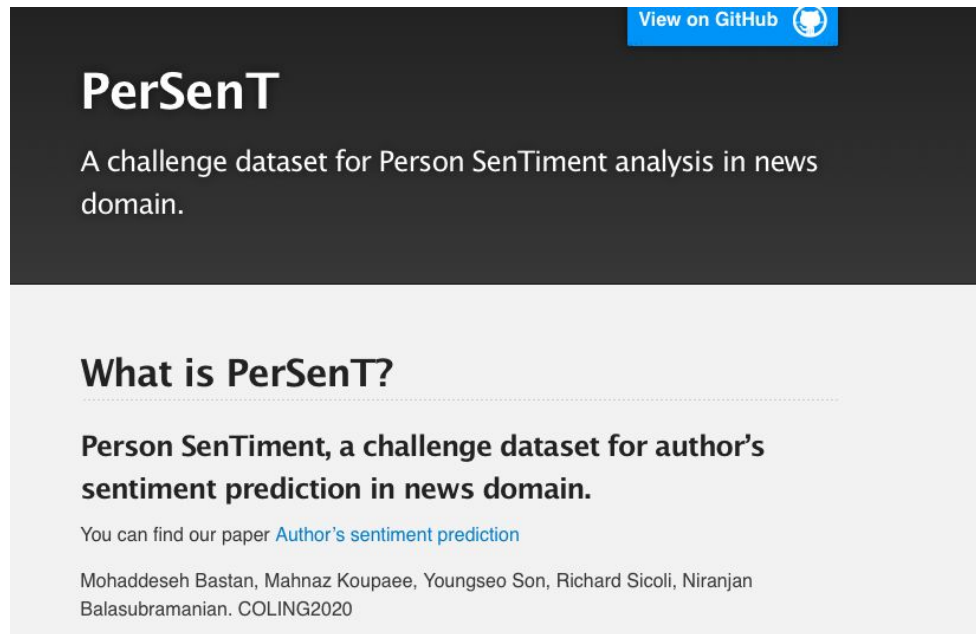
- ❖ Introduction
- ❖ PerSenT
  - Crowdsourcing
  - Quality Control
  - Dataset Characteristics
- ❖ Evaluation
- ❖ **Conclusion**

# Conclusion

- ❖ Introduced a new task: author sentiment inference in the news domain
- ❖ Released a challenge dataset
- ❖ Benchmarks suggest need for further research.
  - models that operate over global representation
  - models that aggregate local decisions

# Code and dataset is available in our website...

<https://stonybrooknlp.github.io/PerSenT/>



The image shows a screenshot of the GitHub repository page for PerSenT. At the top right, there is a blue button that says "View on GitHub" with the GitHub logo. Below this, the title "PerSenT" is displayed in a large, bold, white font. Underneath the title, a description reads: "A challenge dataset for Person SenTiment analysis in news domain." Below this, the section "What is PerSenT?" is followed by a dotted line. The text then states: "Person SenTiment, a challenge dataset for author's sentiment prediction in news domain." Below this, it says: "You can find our paper [Author's sentiment prediction](#)". At the bottom, the authors are listed: "Mohaddeseh Bastan, Mahnaz Koupaee, Youngseo Son, Richard Sicoli, Niranjan Balasubramanian. COLING2020".