

# Sarcasm Detection

Aman Sinha

Café TAL  
May 4, 2021

# Understanding language

- Natural language is complex, surface semantic meaning and intended meaning often differ.
- There are a lot sources text data, such as social platform, OTTs, commercial platforms, etc.
- Some forms in which these occur includes plain text, combined form (text + images)
- There are many applications related to sentiment analysis such as - recommendation, AI assistants, and so on.
- Sarcasm/satire can often cause problem for these systems.

# Sentiment Analysis

R:   **Awesome** !!! Everybody does it. Thats the  
**greatest** reason to do something.

# Sarcasm Detection

R:   Awesome !!! Everybody does it. Thats the  
greatest reason to do something.

# Sarcasm Detection

C2 Hes gonna have a **short** career if he keeps **smoking** . **Not** **good** for your health

R: **Awesome** !!! Everybody does it. **Thats** the **greatest** reason to do something.

# Sarcasm Detection

C1 This is the **greatest** video in the history of college football.

C2 Hes gonna have a **short** career if he keeps **smoking** . **Not** **good** for your health

R: **Awesome** !!! Everybody does it. **Thats** the **greatest** reason to do something.

# C-Net: Contextual Network for Sarcasm Detection

**Amit Kumar Jena, Aman Sinha\* and Rohit Agarwal\***

Indian Institute of Technology (Indian School of Mines) Dhanbad, India  
{amitjena40, amansinha091, agarwal.102497}@gmail.com

## **Abstract**

Automatic Sarcasm Detection in conversations is a difficult and tricky task. Classifying an utterance as sarcastic or not in isolation can be futile since most of the time the sarcastic nature of a sentence heavily relies on its context. This paper presents our proposed model, C-Net, which takes contextual information of a sentence in a sequential manner to classify it as sarcastic or non-sarcastic. Our model showcases competitive performance in the Sarcasm Detection shared task organised on CodaLab and achieved 75.0% F1-score on the Twitter dataset and 66.3% F1-score on Reddit dataset.

utterance is sarcastic or not. For this, the baseline models were first created using traditional machine learning algorithms like logistic regression, SVM etc. which were trained to classify utterances without considering their contextual information. Sequence models like vanilla RNN and LSTM were trained similarly. Then different types of word embeddings (ELMo and Glove) and sentence embedding (DeepMoji) to capture emotional states in the sentences were also experimented to detect incongruities within the text. The latest state-of-the-art transformer based models like BERT, XLNet and RoBERTa were also used for classifying sentences in isolation. Our investigations for creating systems

# Dataset

- Reddit, Twitter

Dataset	Train	Test
Reddit	4400	1800
Twitter	5000	1800

Dataset\* Statistics [2]

Turns	Message
<i>Context<sub>1</sub></i>	This is the greatest video in the history of college football.
<i>Context<sub>2</sub></i>	Hes gonna have a short career if he keeps smoking . Not good for your health
<i>Response</i>	Awesome !!! Everybody does it. Thats the greatest reason to do something.

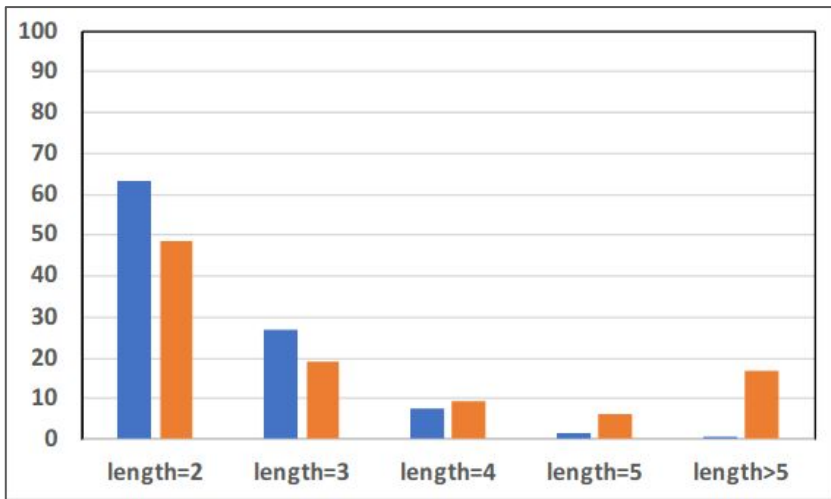


# Dataset

- Reddit, Twitter

Dataset	Train	Test
Reddit	4400	1800
Twitter	5000	1800

Dataset\* Statistics [2]



Length refers to size of context set; blue denotes *Reddit* and orange denotes *Twitter*

# Investigation

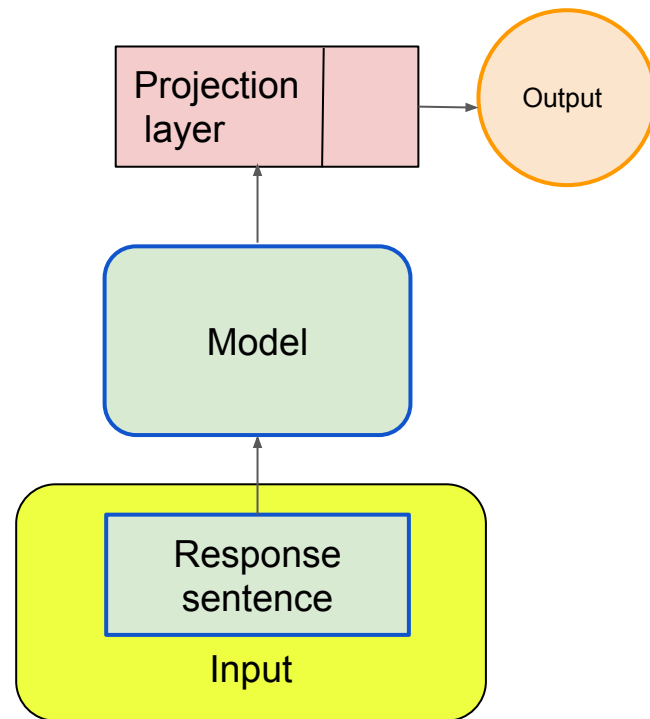
1. Is response only effective?
2. How does context help?
  - i. Partial context (fixed size context set)
  - ii. Complete context set

# Investigation:1

1. Is response only effective?

# Investigation:1

<i>Response Only Set</i>	Twitter	Reddit
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>



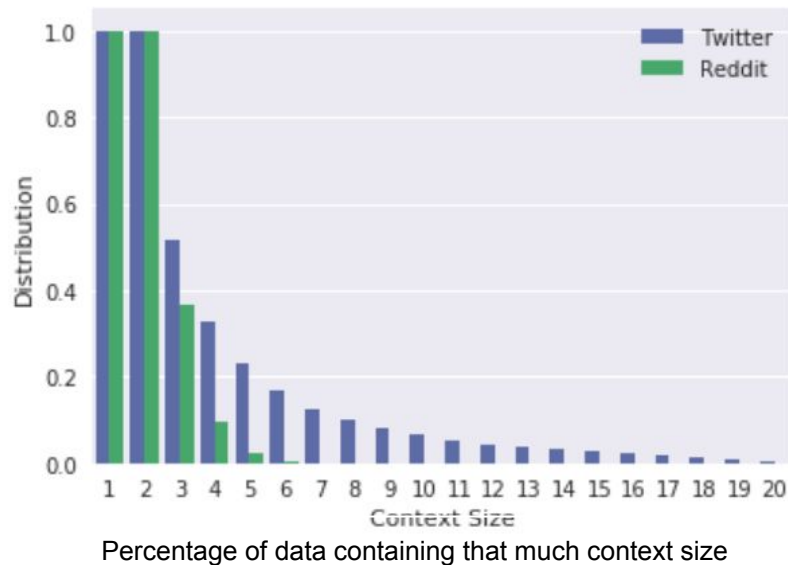
❖ Is response only effective?

# Investigation:2.i

- ❖ How does context help?
  - Partial context (fixed size context set)

Contribution: C-Net

- Training using Pseudo labeling
- Simple Exponential Smoothing



- ❖ Partial context (fixed size context set)

# Data Instance

**C1:** Talk shit, get hit.

**C2:** I like watching people swing and miss.

**C3:** He didn't miss, the big guy dodged it.

**C4:** Soooo he missed...Because the other guy dodged it

**R:** No, no... he didn't miss, it's just that his face wasn't where his fist ended up because the face moved.

SARCASTIC!!

❖ Partial context (fixed size context set)

# Pseudo Labelling

**C1:**

Talk shit, get hit.

SARCASTIC!!

**C2:**

I like watching people swing and miss.

SARCASTIC!!

**C3:**

He didn't miss, the big guy dodged it.

SARCASTIC!!

**C4:**

Soooo he missed...Because the other guy dodged it

SARCASTIC!!

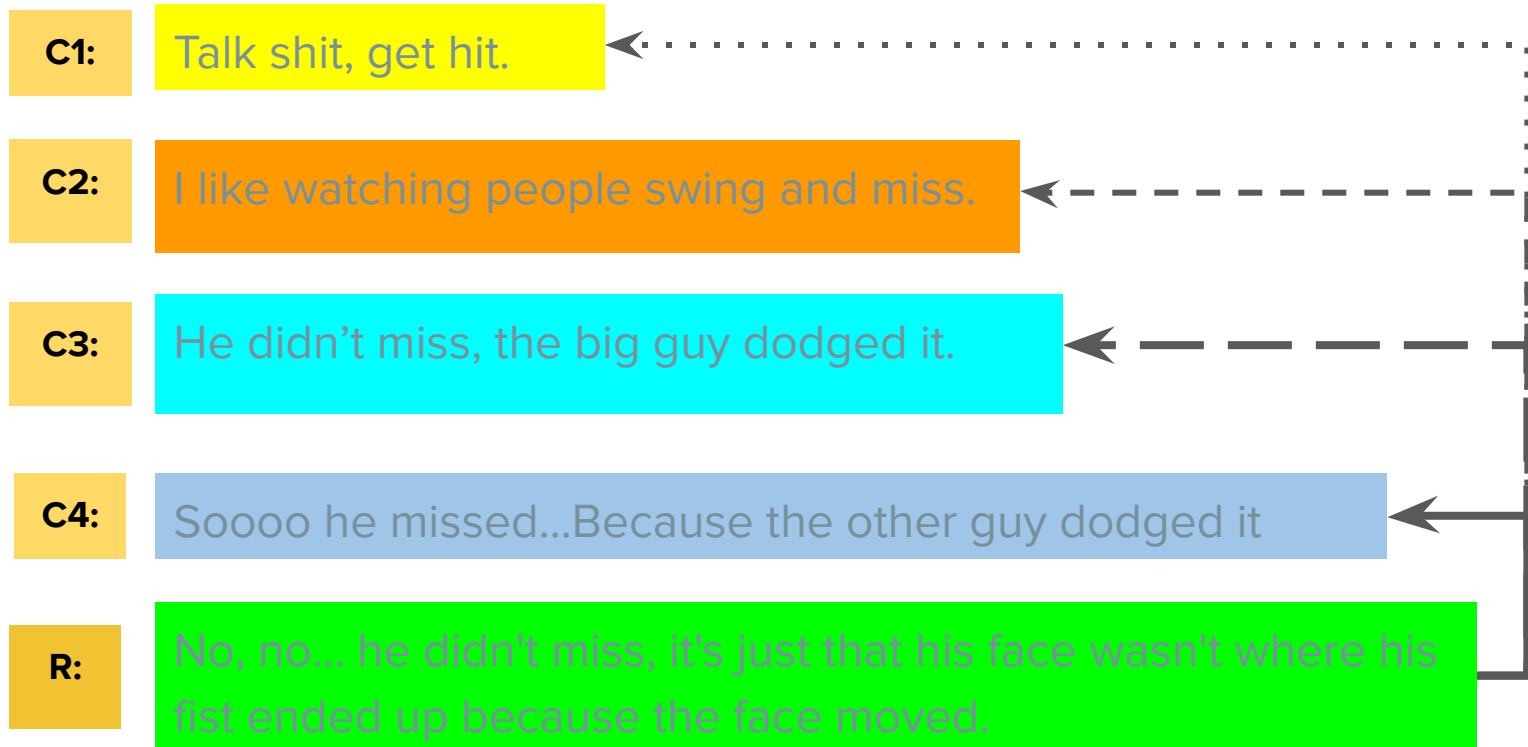
**R:**

No, no... he didn't miss, it's just that his face wasn't where his fist ended up because the face moved.

SARCASTIC!!

❖ Partial context (fixed size context set)

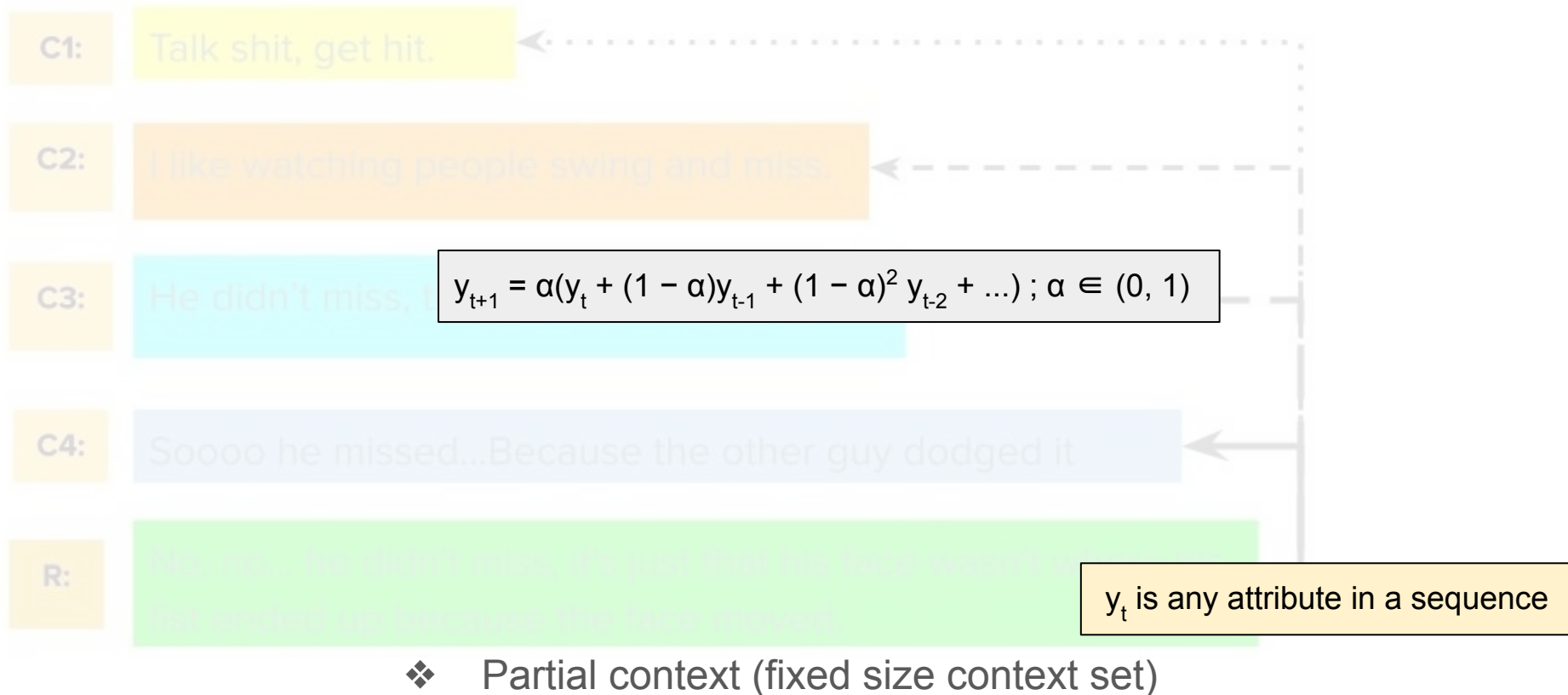
# Simple Exponential Smoothing?



❖ Partial context (fixed size context set)



# Simple Exponential Smoothing?



# Simple Exponential Smoothing?

C1: Talk shit, get hit.

C2: I like watching people swing and miss.

C3: He didn't miss, it

$$y_{t+1} = \alpha(y_t + (1 - \alpha)y_{t-1} + (1 - \alpha)^2 y_{t-2} + \dots) ; \alpha \in (0, 1)$$

$$P_{\text{sarcastic}} = \alpha(p_R + (1 - \alpha)p_{C4} + (1 - \alpha)^2 p_{C3} + (1 - \alpha)^3 p_{C2} + (1 - \alpha)^4 p_{C1})$$

C4: Soooo he missed...Because the other guy dodged it

R:

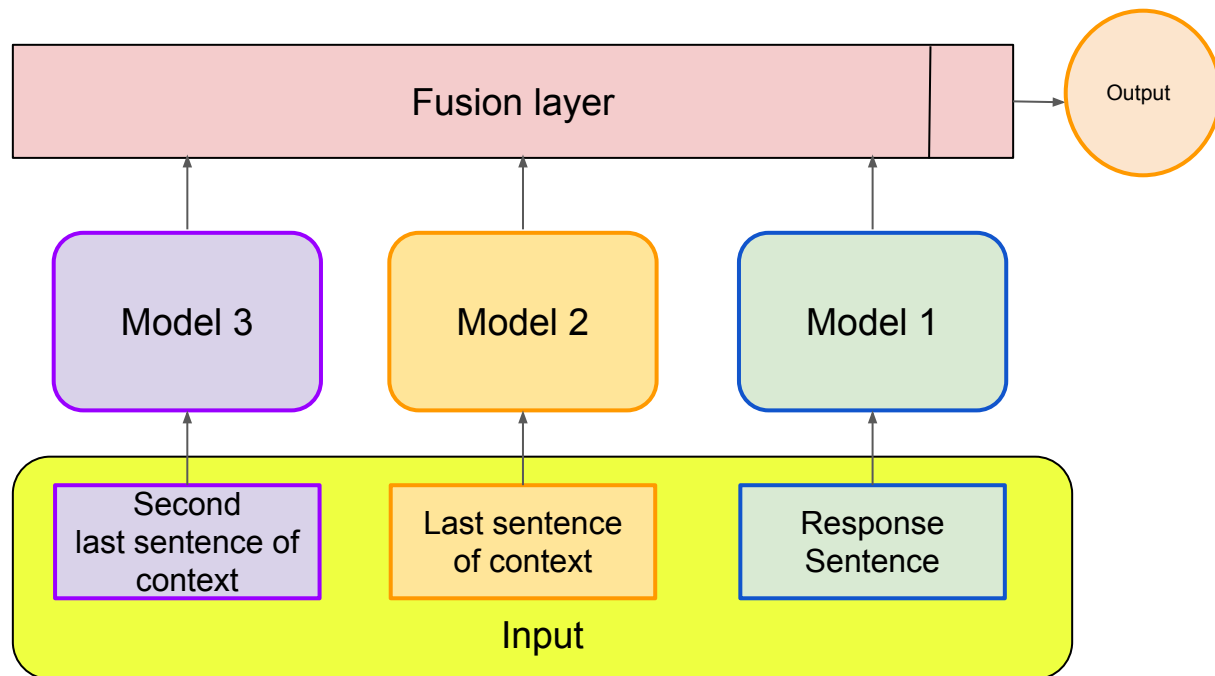
No, no... he didn't miss, it's just that his face wasn't a  
bit wrinkled up because the face moved.

$p_i$  is individual model output

❖ Partial context (fixed size context set)

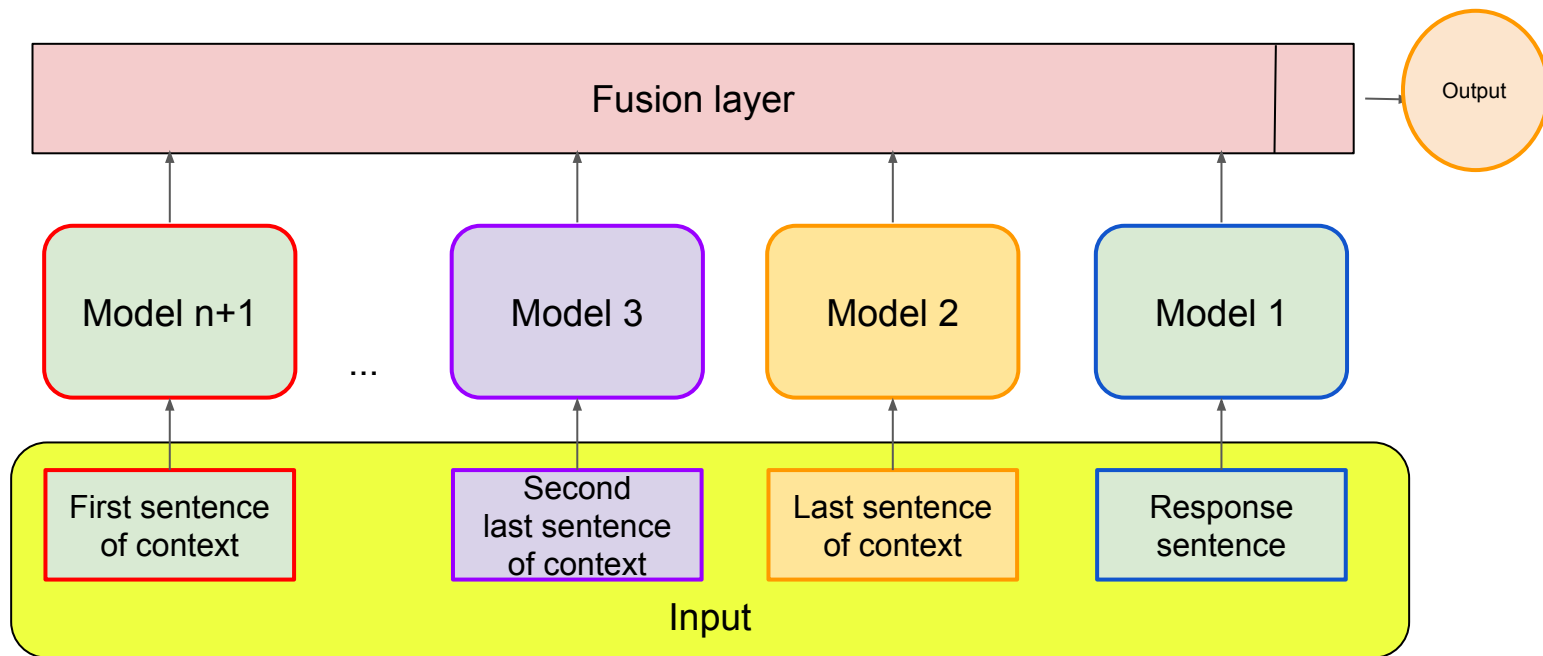
## Investigation:2.i

<i>Fixed Context Set</i>	Twitter	Reddit
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>



❖ Partial context (fixed size context set)

## Investigation:2.i (extended)



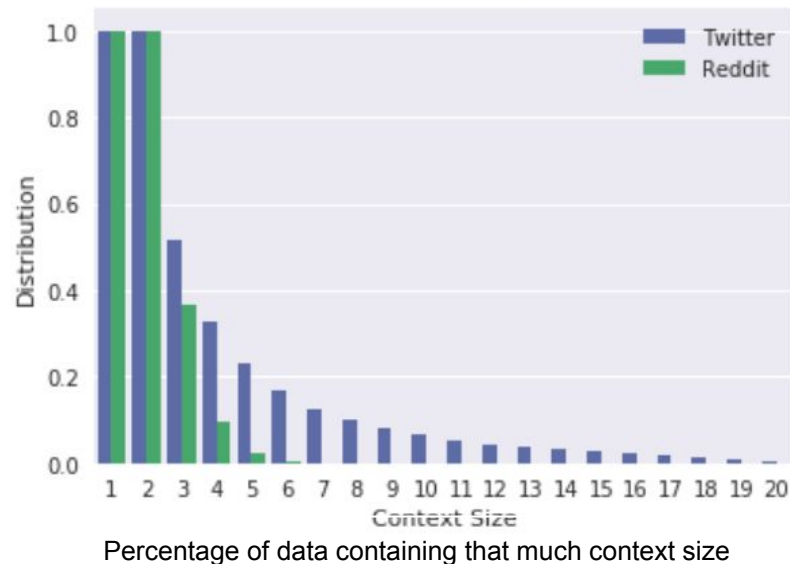
❖ Partial context set

# Investigation:2.ii

## 2. How does context help?

### ii. Complete context set

- Limitation of C-Net - variable context set length



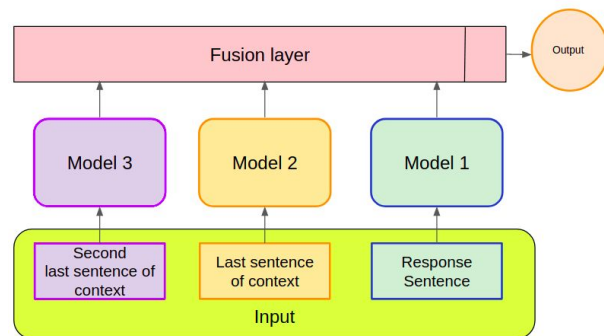
❖ Complete context set

# Investigation:2.ii

## 2. How does context help?

### ii. Complete context set

- Limitation of C-Net - variable context set length
- Trained with Pseudo Labelling
- Separation marker - fixed C-Net setup



<i>Complete Context Set</i>	Twitter	Reddit
Time-stamping	0.710	0.500

❖ Complete context set

# Analysis

Method	Twitter	Reddit
<i>Response Only Set</i>		
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
{ Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>
<i>Fixed Context Set</i>		
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>
<i>Complete Context Set</i>		
Time-stamping	0.710	0.500

- RNN, Bi-LSTM are ineffective to capture from sarcasm just with the response.

# Analysis

Method	Twitter	Reddit
<i>Response Only Set</i>		
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>
<i>Fixed Context Set</i>		
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>
<i>Complete Context Set</i>		
Time-stamping	0.710	0.500

- RNN, Bi-LSTM are ineffective to capture from sarcasm just with the response.
- Contextual /pretrained embedding such as Elmo, Deepmoji are effective.



# Analysis

Method	Twitter	Reddit
<i>Response Only Set</i>		
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>
<i>Fixed Context Set</i>		
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>
<i>Complete Context Set</i>		
Time-stamping	0.710	0.500

- RNN, Bi-LSTM are ineffective to capture from sarcasm just with the response.
- Contextual /pretrained embedding such as Elmo, Deepmoji are effective.
- Bert based models are more effective for sarcasm detection (Response-only set).

# Analysis

Method	Twitter	Reddit
<i>Response Only Set</i>		
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>
<i>Fixed Context Set</i>		
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>
<i>Complete Context Set</i>		
Time-stamping	0.710	0.500

- RNN, Bi-LSTM are ineffective to capture from sarcasm just with the response.
- Contextual /pretrained embedding such as Elmo, Deepmoji are effective.
- Bert based models are more effective for sarcasm detection (Response-only set).
- Ro/BERT are more effective compared to XLNet can be attributed to the fact that XLNet uses permutation LM while training.

# Analysis

Method	Twitter	Reddit
<i>Response Only Set</i>		
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>
<i>Fixed Context Set</i>		
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>
<i>Complete Context Set</i>		
Time-stamping	0.710	0.500

- SES is effective, compared to LR, with C-Net.

# Analysis

Method	Twitter	Reddit
<i>Response Only Set</i>		
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>
<i>Fixed Context Set</i>		
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>
<i>Complete Context Set</i>		
Time-stamping	0.710	0.500

- SES is effective, compared to LR, with C-Net.
- Fixed context is better than complete context.

# Analysis

Method	Twitter	Reddit
<i>Response Only Set</i>		
Logistic Regression	0.685	0.622
Naive Bayes	0.673	0.626
SGD Classifier	0.668	0.626
XGBoost	0.672	0.617
SVM	0.632	0.334
Vanilla RNN	0.478	0.463
Bi-LSTM	0.497	0.481
DeepMoji	0.679	0.633
ELMo	0.684	0.544
ELMo+DeepMoji	0.681	0.518
XLNet (base-cased)	0.712	0.598
BERT (base-uncased)	<b>0.733</b>	0.671
RoBERTa (base)	0.680	<b>0.678</b>
<i>Fixed Context Set</i>		
C-Net+LR	0.747	0.650
C-Net+SES	<b>0.750</b>	<b>0.663</b>
<i>Complete Context Set</i>		
Time-stamping	0.710	0.500

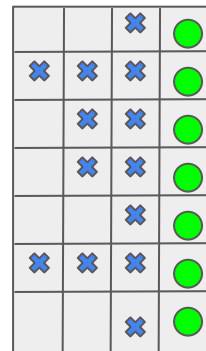
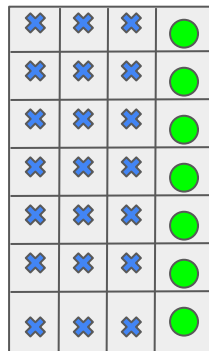
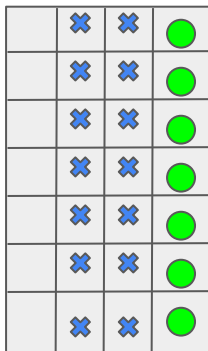
- For Twitter, model F1 range: (0.478, 0.750)
- For Reddit, model F1 range: (0.334, 0.678)
- Adding context helps in case of Twitter compared to Reddit (C-Net are BERT based models).

# Takeaways

- Context helps in modeling sarcasm in better way.
- Partial context is better then Complete context.
- Data characteristics effects the modelling of sarcasm,
  - Twitter has a character limit, vs Reddit no limit
  - On twitter more markers(hashtags), emojis, URLs, gifs vs on Reddit more text.

# Prospects

- How much context is necessary - fixed vs all vs variable?



● : Response

× : Context

# Prospects

- What is the better way to get the include context - SES vs learnable?

$$P_{\text{sarcastic}} = \alpha(p_R + (1 - \alpha)p_{C4} + (1 - \alpha)^2 p_{C3} + (1 - \alpha)^3 p_{C2} + (1 - \alpha)^4 p_{C1})$$

$$P_{\text{sarcastic}} = w_0 p_R + w_1 p_{C4} + w_2 p_{C3} + w_3 p_{C2} + w_4 p_{C1}$$

$p_i$  is individual C-Net component output;  $\alpha \in (0, 1)$ ;  $w_i$  are learnable weights



# Reference

- [1] Amit Kumar Jena, Aman Sinha, and Rohit Agarwal. 2020. C-net: Contextual network for sarcasm detection. In Proceedings of the Second Workshop on Figurative Language Processing, Seattle, WA, USA.
- [2] Debanjan Ghosh, Avijit Vajpayee, and Smaranda Muresan. 2020. A report on the 2020 sarcasm detection shared task. In Proceedings of the Second Workshop on Figurative Language Processing, pages 1 – 11, Online. Association for Computational Linguistics.

Thank you for your attention!