Mining Dual Emotion for Fake News Detection

Xueyao Zhang†
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences
zhangxueyao19s@ict.ac.cn

Juan Cao* †
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences
caojuan@ict.ac.cn

Xirong Li*
Key Lab of Data Engineering and Knowledge Engineering, Renmin University of China
Beijing, China
xirong@ruc.edu.cn

Qiang Sheng†
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences
shengqiang18z@ict.ac.cn

Lei Zhong†
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences
zhonglei18s@ict.ac.cn

Kai Shu
Illinois Institute of Technology
Chicago, Illinois, USA
kshu@iit.edu

WWW'21

210903 Chia-Chun Ho

Outline

Introduction

Related Work

Methodology

Experiments

Conclusion and Future Work

Comments

Introduction

Sentimental or Emotional signals in existing works

- Some works find a relationship between news veracity and the sentiments of the posted text, and append a sentimental feature (ratio of number of negative and positive words) to help text-only detectors.
- Some works extract richer emotional features from the contents based on emotional lexicons for detection.
- Most existing works leverage the emotional signals of the fake news conveyed be the publishers but rarely focus on the emotions of news comments aroused in the crowds.

Introduction

Sentimental or Emotional signals in existing works

- For spreading in the crowd virally, fake news often evokes high-arousal or activating emotions of the crowds.
- Therefore, in additional to emotions of news contents, it's necessary to explore whether
  - emotions of news comments and
  - the relationship between the two emotions
- are helpful for fake news detection.

Introduction

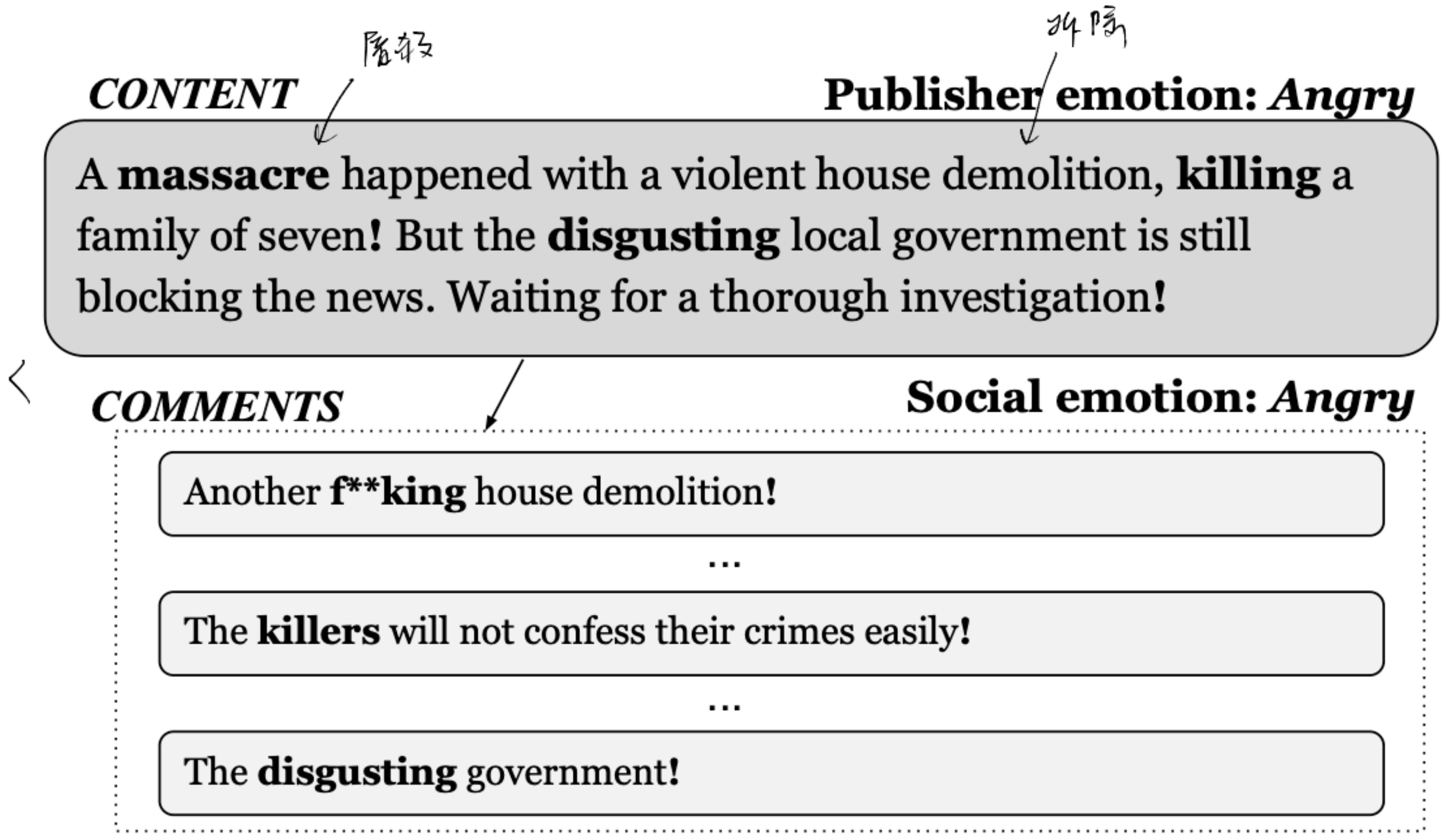
Define the two emotions

- Publisher emotion
  - The emotions conveyed by publishers of the news pieces.
- Social emotion
  - The emotions aroused in the crowd facing to the news pieces.
- And adopt dual emotion as general term of these two emotions.

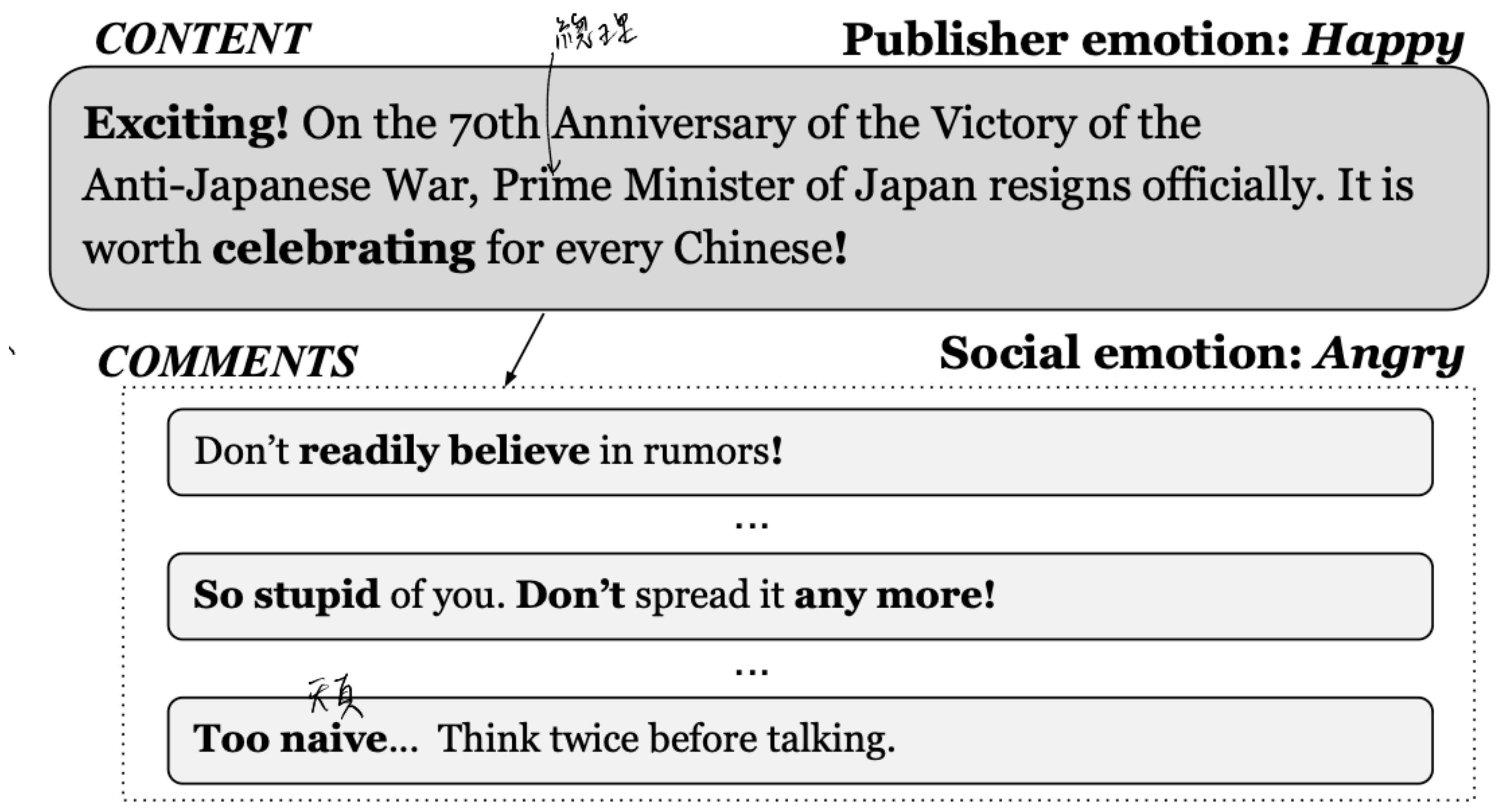
Introduction

Dual emotion appearances

- Emotion resonances
  - Publisher emotion is same or similar to social emotion.
- Emotion dissonances
  - Publisher emotion is different from the social emotion.
- The data observation statistical findings highlight that the relationship in dual emotion can be indicative of the news veracity and should be considered when modeling.



(a) Emotion resonance in a fake news piece: the *publisher emotion* and *social emotion* are both *angry*.



(b) Emotion dissonance in a fake news piece: the *publisher emotion* is *happy* while the *social emotion* is *angry*.

Introduction

Dual Emotion Features

- To model the dual emotion and emotion resonances and dissonances for fake news detection.
- Proposed *Dual Emotion Features* to represent publisher emotions, social emotion and the similarity and difference of the dual emotion jointly.
- Besides, it's convenient to implement and plug the features into existing fake news detectors as an enhancement.

Introduction

Contributions of Dual Emotion Features

- Propose and verify that the dual emotion (publisher / social emotion) signal is distinctive between fake and real news.

- First proposed the feature set, *Dual Emotion Features*, to comprehensively represent dual emotion and the relationship between the two kinds of emotions, and exhibit how to plug it into the existing fake news detectors as a complement and enhancement.

*2020 CIKM - Truth be Told: Fake News Detection Using User Reactions on Reddit

Related Work
of fake news detection

- In recent years, researchers begin to utilize deep learning models.
  - GRU-based ('16), CNN-based ('17) models.
- Beyond news content, social contexts such as following are emphasized as well.
  - Texts of comments and reposts ('16-'18)
  - Viewpoints ('16) and stances of the crowd ('18)
  - User credibility ('19)

Related Work

Discovering the distinctive emotional signals

- The approaches as mentioned beginning, some works only leverage the emotion signals of fake news contents but ignore the emotions of fake news comments and the relationship between the two emotions.

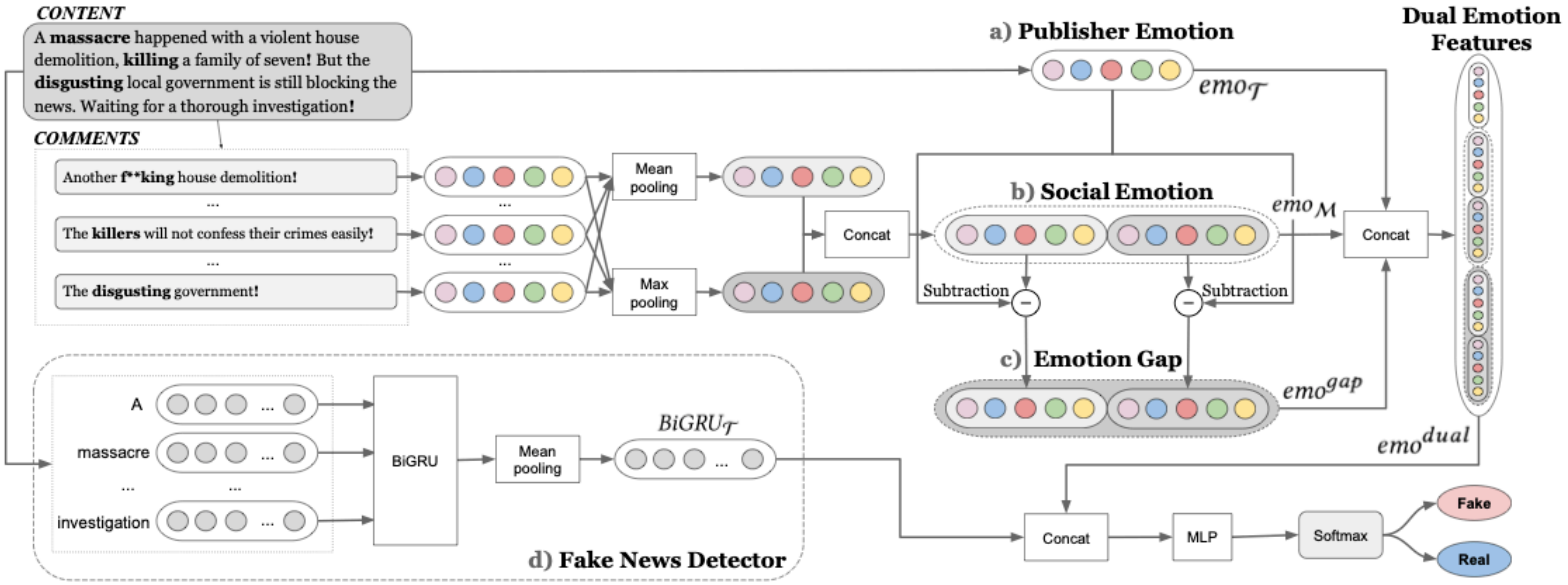
- Recently, has propose an adaptive fusion network for fake news detection, modeling emotion embeddings from the contents and the comments.

  - This work focuses on adaptively fusing various features by advanced deep learning models.

  - Do not explore the specific distinction of dual emotion signals between fake and real news.

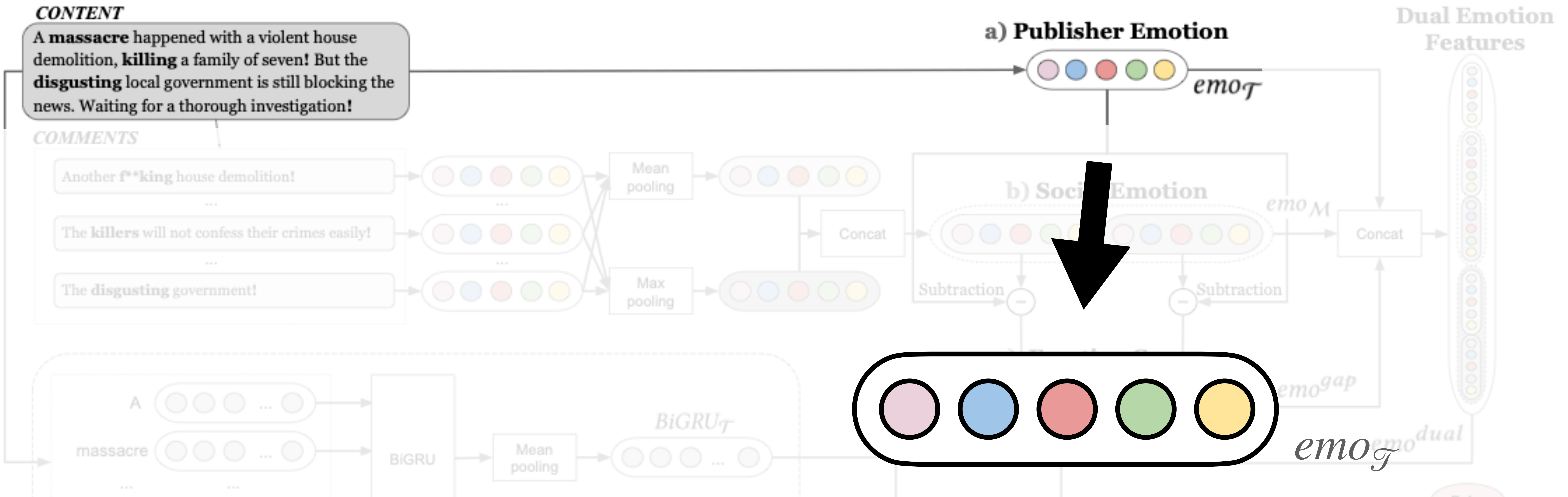
Methodology

Overall framework of using Dual Emotion Features for fake news detection



Methodology

Publisher Emotion



$$emo_{\mathcal{T}} = emo_{\mathcal{T}}^{cate} \oplus emo_{\mathcal{T}}^{lex} \oplus emo_{\mathcal{T}}^{int} \oplus emo_{\mathcal{T}}^{senti} \oplus emo_{\mathcal{T}}^{aux}$$

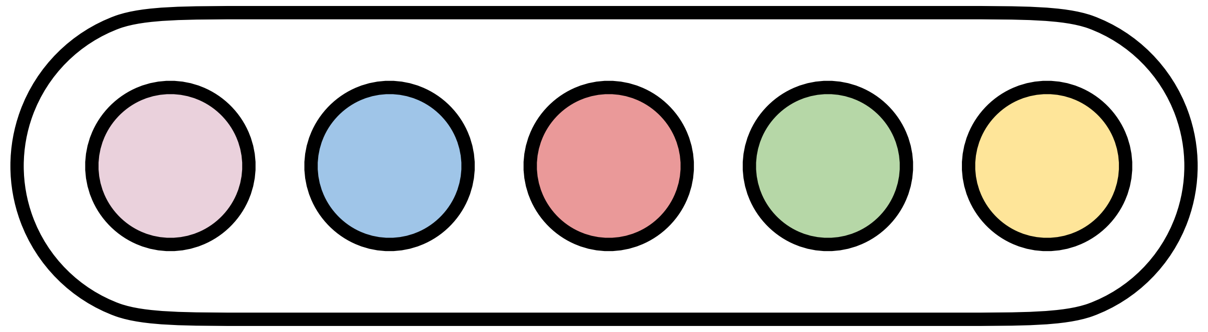
Methodology

Publisher Emotion

- Given the input sequence of the textual content with length L, $\mathcal{T} = [t_1, t_2, \ldots, t_L]$, where t_i is the i^{th} word in the text, the goal is to extract emotion features $emo_{\mathcal{T}}$ from $\mathcal{T}$.

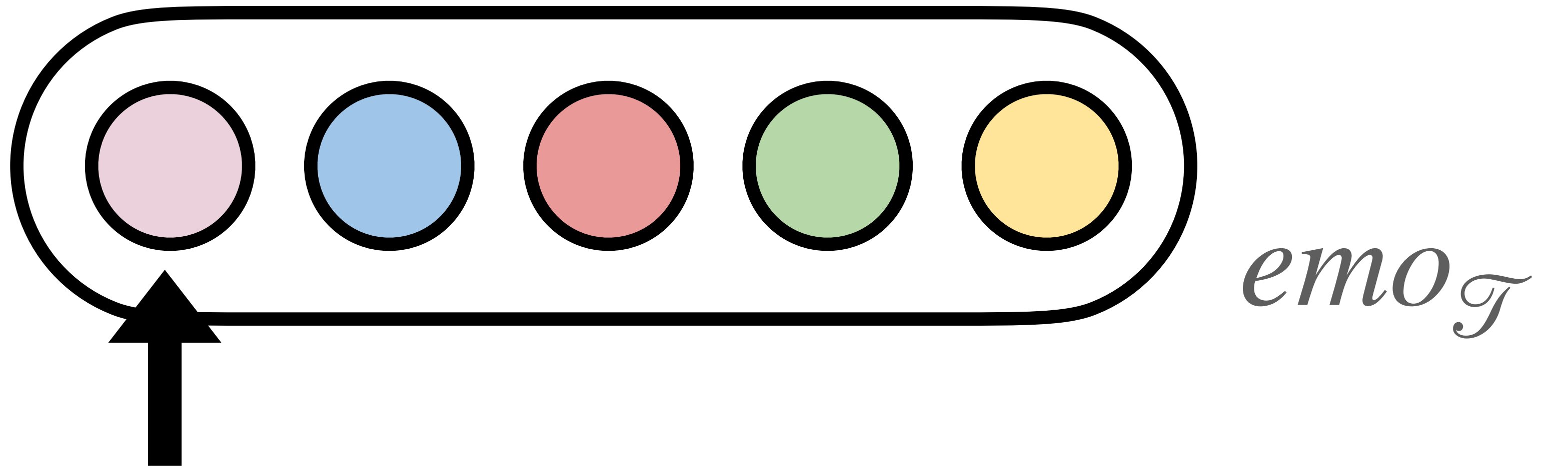
- To comprehensively represent the Publisher Emotion, use variety of features extracted from news contents.

  - Emotion category
  - Emotion lexicon
  - Emotion intensity
  - Sentiment score
  - Other auxiliary features



$emo_{\mathcal{T}}$

$$emo_{\mathcal{T}} = emo_{\mathcal{T}}^{cate} \oplus emo_{\mathcal{T}}^{lex} \oplus emo_{\mathcal{T}}^{int} \oplus emo_{\mathcal{T}}^{senti} \oplus emo_{\mathcal{T}}^{aux}$$



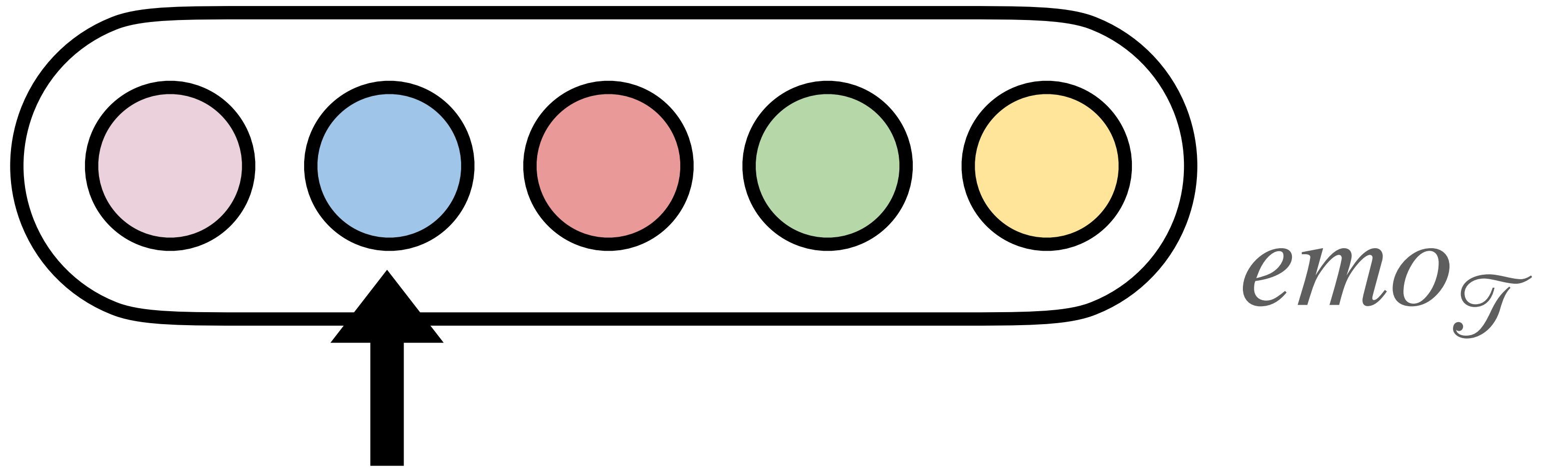
Methodology

Emotion Category

- Use public emotion classifiers (will intro later) to get emotion category features.
- Usually, the output of an emotion classifier is the probabilities that the given text contains certain emotions.
- So obtain the emotion category features $emo_{\mathcal{T}}^{cate} = f(\mathcal{T})$.

Methodology

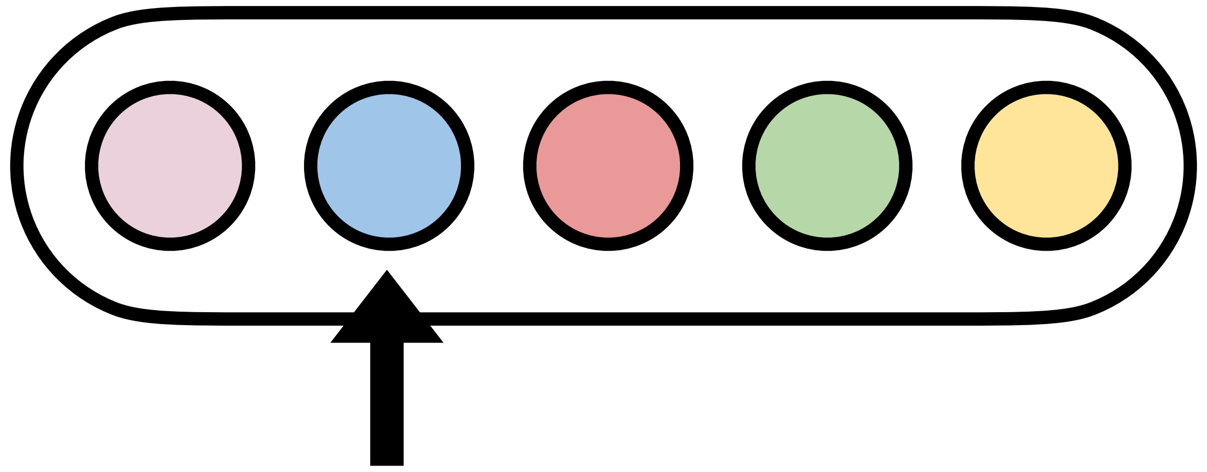
Emotion Lexicon



- Usually, a piece of text conveys specific emotions by using several specific words (which are generally included in emotional lexicons).
- The approach is dependent on the existing emotion dictionaries annotated by experts.
- Assume that there're d_e kinds of emotions: $E = \{e_1, e_2, \ldots, e_{d_e}\}$
- The dictionary provides a list of emotional words $\mathscr{E}_e = \{w_{e,1}, w_{e,2}, \ldots, w_{e,L_e}\}$

Methodology

Emotion Lexicon



$emo_{\mathcal{T}}$

- Given $\mathcal{T}$, gradually aggregate the scores of each word and the whole text across all the emotions for rich representation.

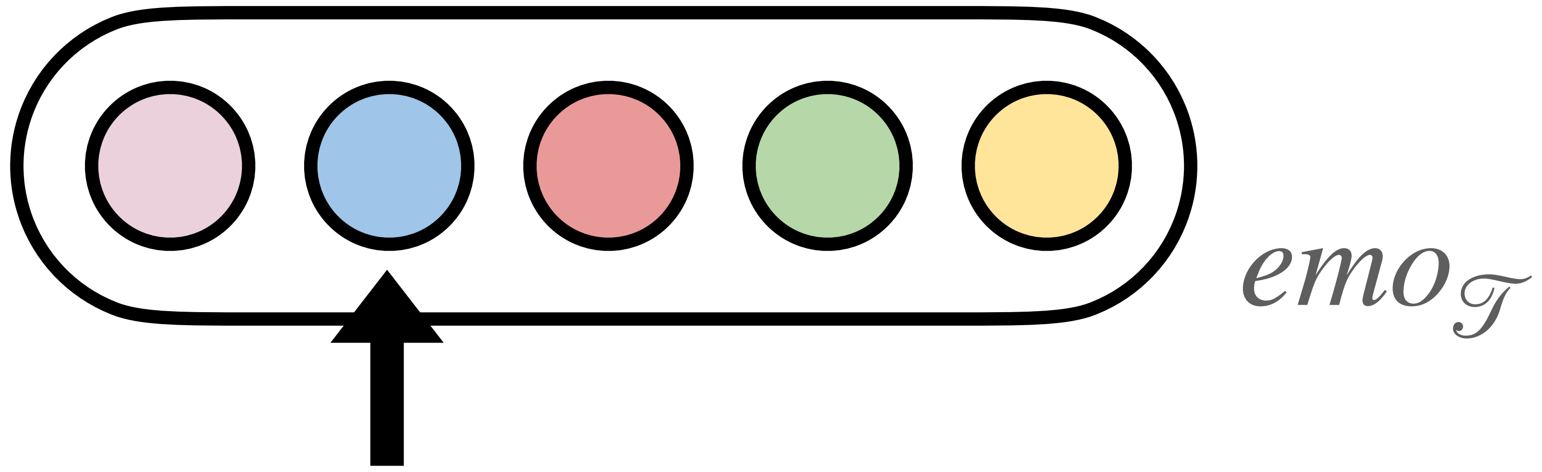
- For one of the emotion e, firstly calculate the word-level scores $s(t_i, e)$, if t_i is in $\mathcal{E}_e$, consider not only its occurrence frequency but also its contextual words.

- $$s\left(t_i, e\right) = \frac{1_{\mathcal{E}_e}\left(t_i\right) \times \mathrm{neg}\left(t_i, w\right) \times \mathrm{deg}\left(t_i, w\right)}{L}$$

- Example: "I am <u>not very joyful</u> today." $s(\text{joyful}, e_{happy}) = -1 \times 2 \times (1/6) = -1/3$

Methodology

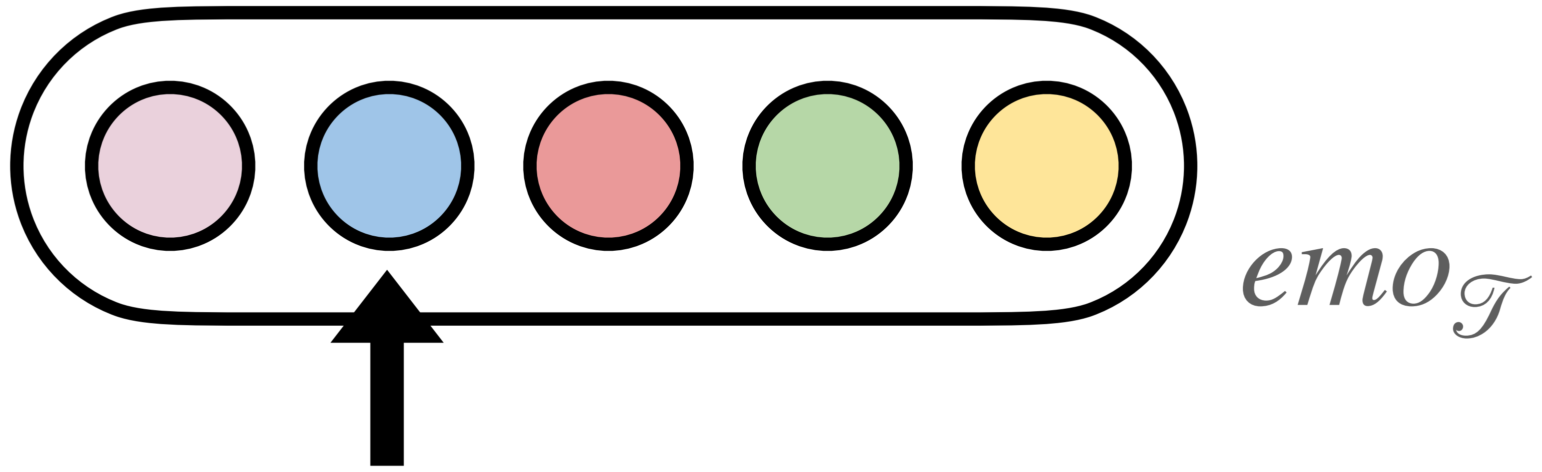
Emotion Lexicon



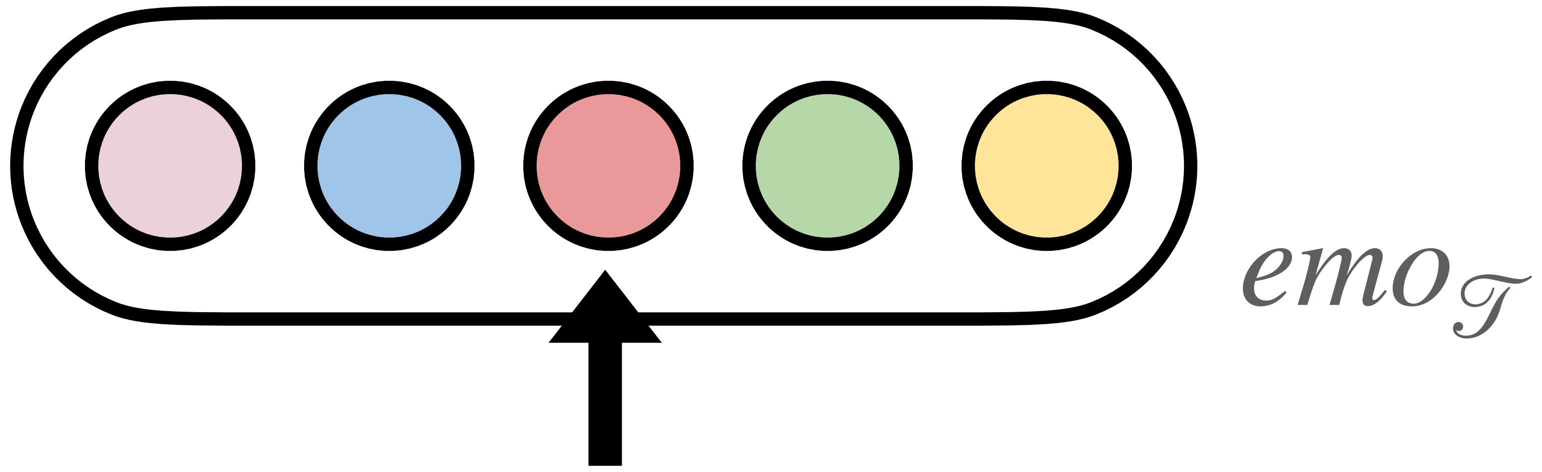
- $s\left(t_i, e\right) = \dfrac{1_{\mathscr{E}_e}\left(t_i\right) \times \text{neg}\left(t_i, w\right) \times \text{deg}\left(t_i, w\right)}{L}$
- $\text{neg}\left(t_i, w\right) = \prod_{j=i-w}^{i-1} \text{neg}\left(t_j\right)$, $\text{deg}\left(t_i, w\right) = \prod_{j=i-w}^{i-1} \text{deg}\left(t_j\right)$
- Example: "I am <u>not very joyful</u> today." $\text{neg}(\text{not}) = -1$, $\text{deg}(\text{very}) = 2$
- $s(\text{joyful}, e_{happy}) = -1 \times 2 \times (1/6) = -1/3$

Methodology

Emotion Lexicon



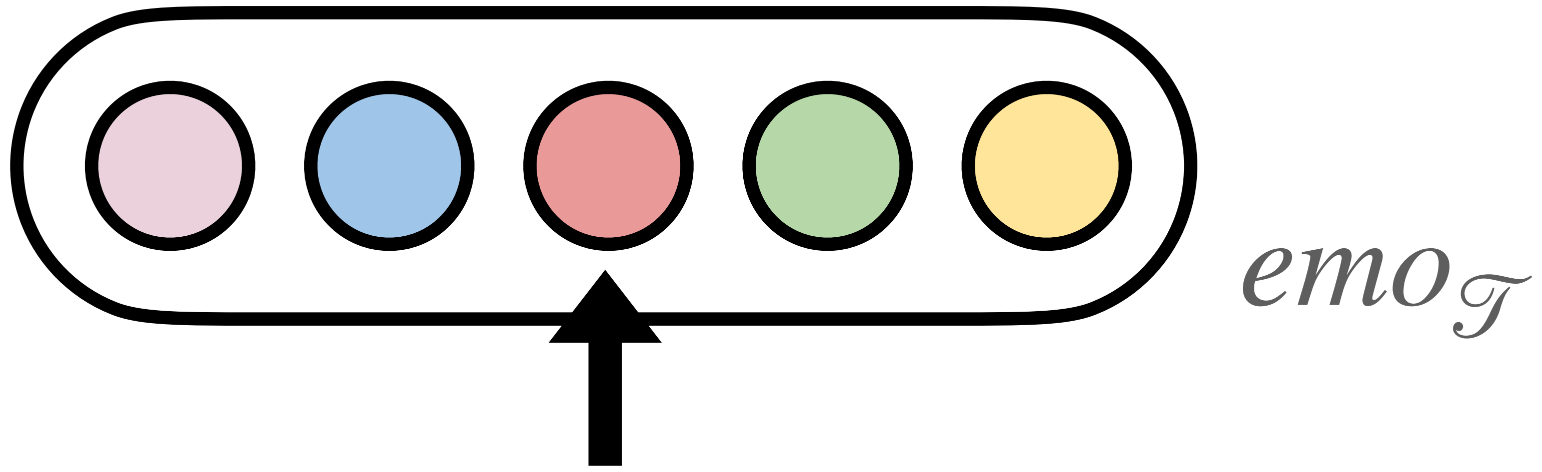
- Then summing the scores of each word in the text, as follow:
- $$s(\mathscr{T}, e) = \sum_{i=1}^{L} s\left(t_i, e\right), \quad \forall e \in E$$
- Finally, the emotional lexicon features $emo_{\mathscr{T}}^{lex}$ are obtained by concatenating all scores of the d_e emotions.
- $$emo_{\mathscr{T}}^{lex} = s\left(\mathscr{T}, e_1\right) \oplus s\left(\mathscr{T}, e_2\right) \oplus \cdots \oplus s\left(\mathscr{T}, e_{d_e}\right)$$



Methodology

Emotion Intensity

- As for emotional lexicons, also consider the emotional intensity of the lexicons.
- For example, when expressing the emotion happy, the word “ecstatic” owns higher intensity than “joyful”.
- The extracting process in similar to that of the emotional lexicon features, except for that include the intensity score.
- $$s'(\mathscr{T}, e) = \sum_{i=1}^{L} s'\left(t_i, e\right) = \sum_{i=1}^{L} \mathrm{int}\left(t_i\right) \times s\left(t_i, e\right), \quad \forall e \in E$$



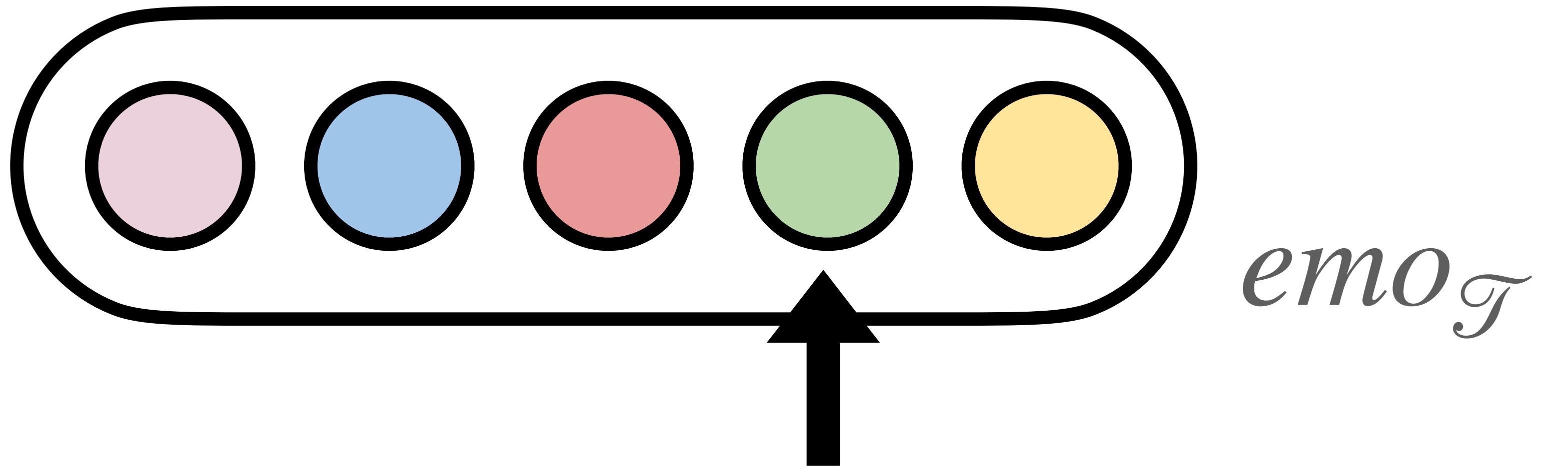
Methodology

Emotion Intensity

- $$s'(\mathscr{T}, e) = \sum_{i=1}^{L} s'\left(t_i, e\right) = \sum_{i=1}^{L} \mathrm{int}\left(t_i\right) \times s\left(t_i, e\right), \quad \forall e \in E$$
- $\mathrm{int}(t_i)$ denotes the intensity score of the word t_i.
- If t_i is in the dictionary, $\mathrm{int}(t_i)$ can be calculated according to emotion dictionary, otherwise $\mathrm{int}(t_i) = 0$
- $$emo_{\mathscr{T}}^{int} = s'\left(\mathscr{T}, e_1\right) \oplus s'\left(\mathscr{T}, e_2\right) \oplus \cdots \oplus s'\left(\mathscr{T}, e_{d_e}\right)$$

Methodology

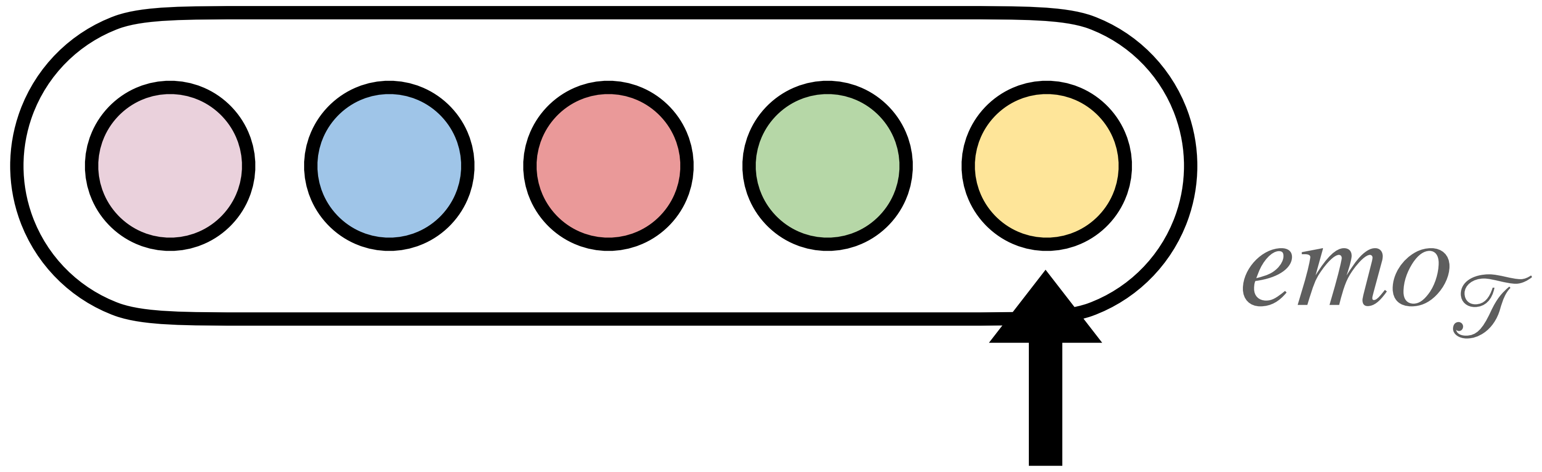
Sentiment Score



- Also consider the coarse-grained sentiment score of the text.
- Usually, the sentiment score is a positive or negative value, which represents the degree of the positive or negative polarity of the whole text.
- It can be calculated by using sentiment dictionaries or public toolkits.
- Then can get sentiment score feature $emo_{\mathcal{T}}^{senti}$.

Methodology

Other Auxiliary Features



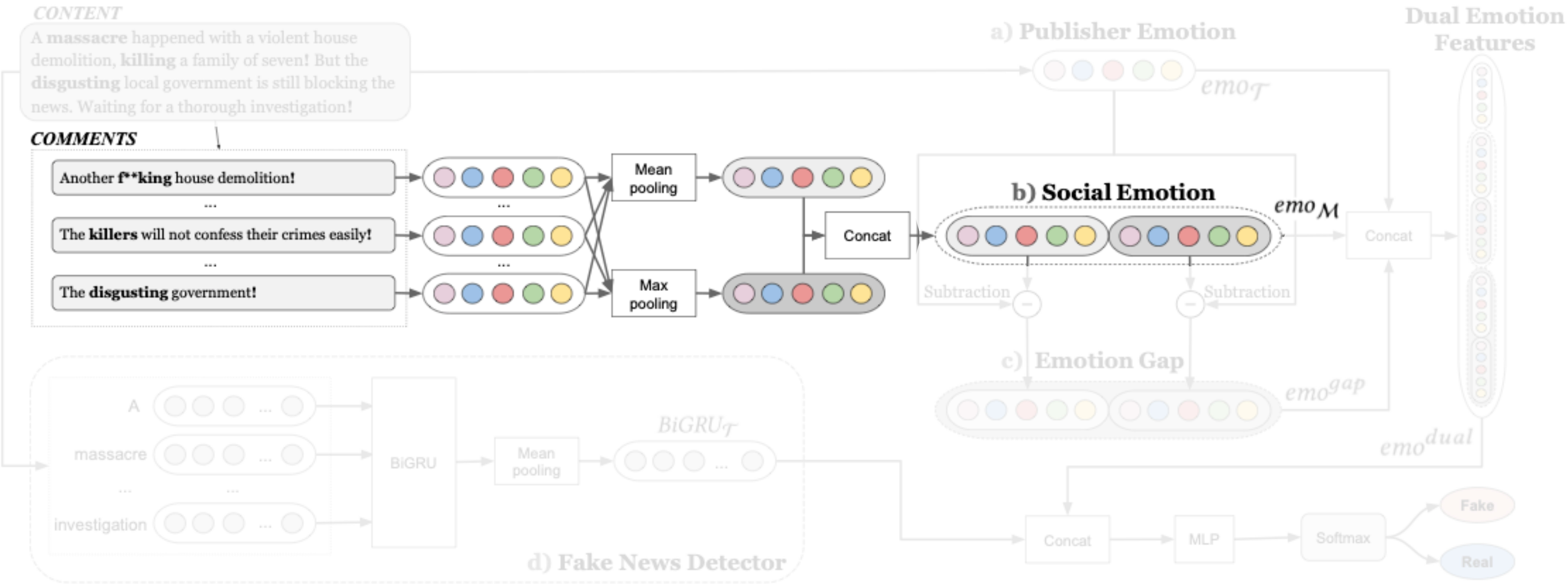
- Considering that the above features don't explicitly exploit the information beyond emotion dictionaries.
- Design a set of auxiliary features to capture the emotional signals behind the non-word elements.
- Then can extract the other auxiliary features $emo_{\mathcal{T}}^{aux}$.

Type	Features
Emoticons	The frequency of happy emoticons
	The frequency of angry emoticons
	The frequency of surprised emoticons
	The frequency of sad emoticons
	The frequency of neutral emoticons
Punctuations	The frequency of exclamation mark
	The frequency of question mark
	The frequency of ellipsis mark
Sentimental Words	The frequency of positive sentimental words
	The frequency of negative sentimental words
	The frequency of degree words
	The frequency of negation words
Personal Pronoun	The frequency of pronoun first
	The frequency of pronoun second
	The frequency of pronoun third
Others (For English corpus)	The frequency of uppercase letters

Table 1: Auxiliary Feature List

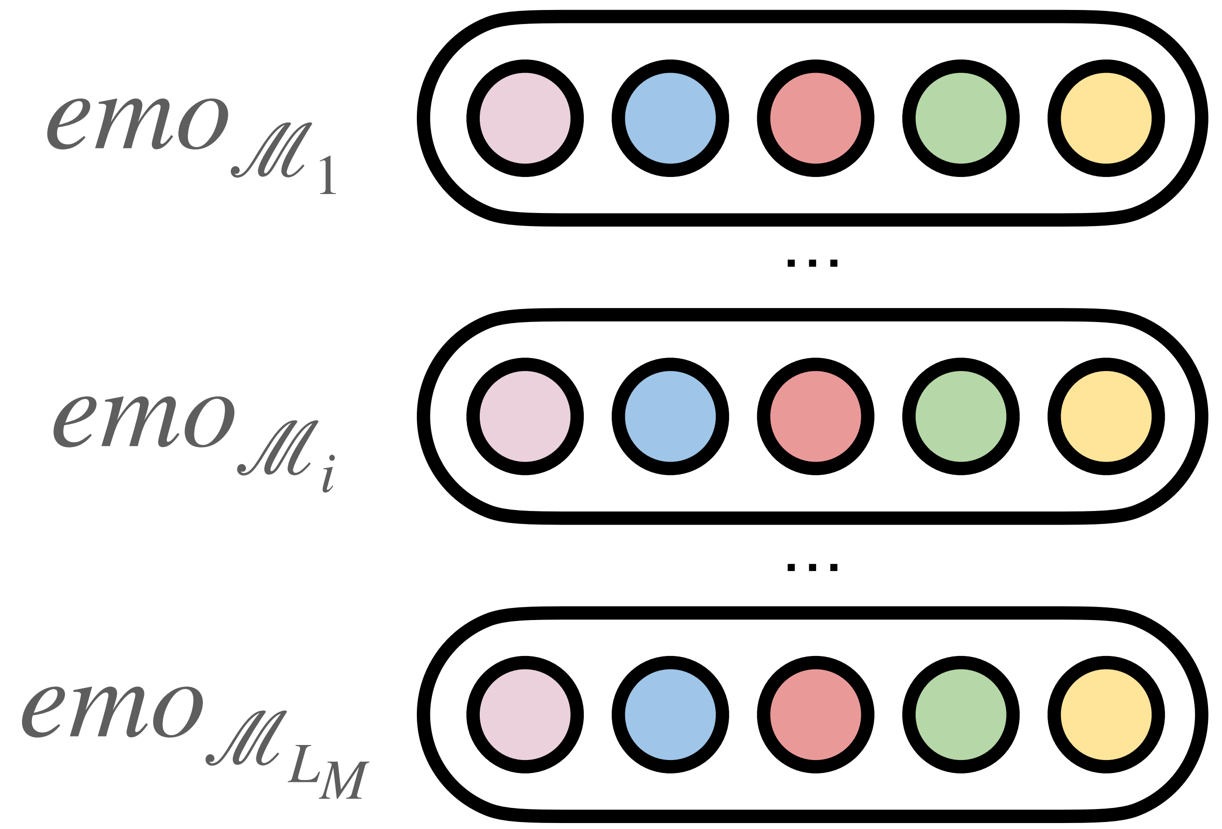
Methodology

Social Emotion



Methodology

Social Emotion



- Extract Social Emotion from the comments of a news piece and then aggregate them as the whole representation.

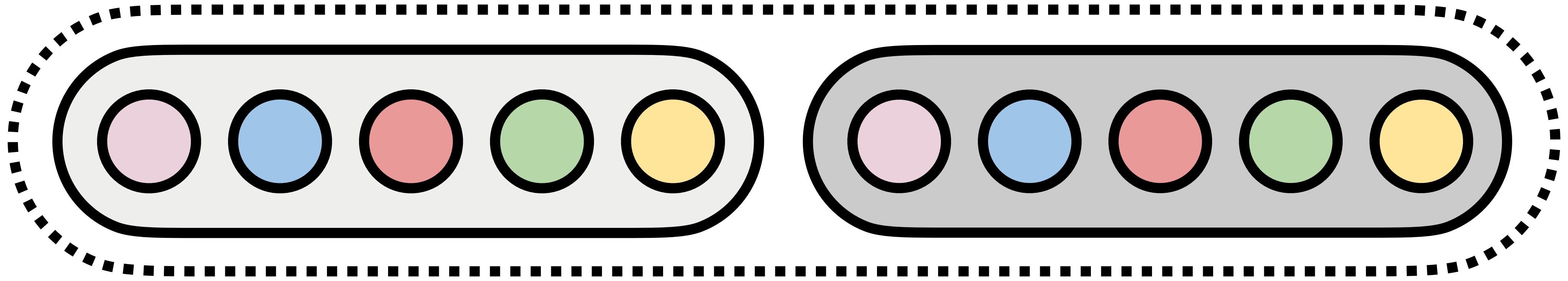
- Given a set of comments of a news piece: $\mathscr{M} = [\mathscr{M}_1, \mathscr{M}_2, \ldots, \mathscr{M}_{L_{\mathscr{M}}}]$

- As for $\mathscr{M}_i$, can get $emo_{\mathscr{M}_i}$, then stack the transposed emotion vector of every comment to obtain the whole emotion vector of comments $\widehat{emo_{\mathscr{M}}}$.

- $\widehat{emo_{\mathscr{M}}} = emo_{\mathscr{M}_1}^{\top} \oplus emo_{\mathscr{M}_2}^{\top} \oplus \cdots \oplus emo_{\mathscr{M}_{L_{\mathscr{M}}}}^{\top}$

Methodology

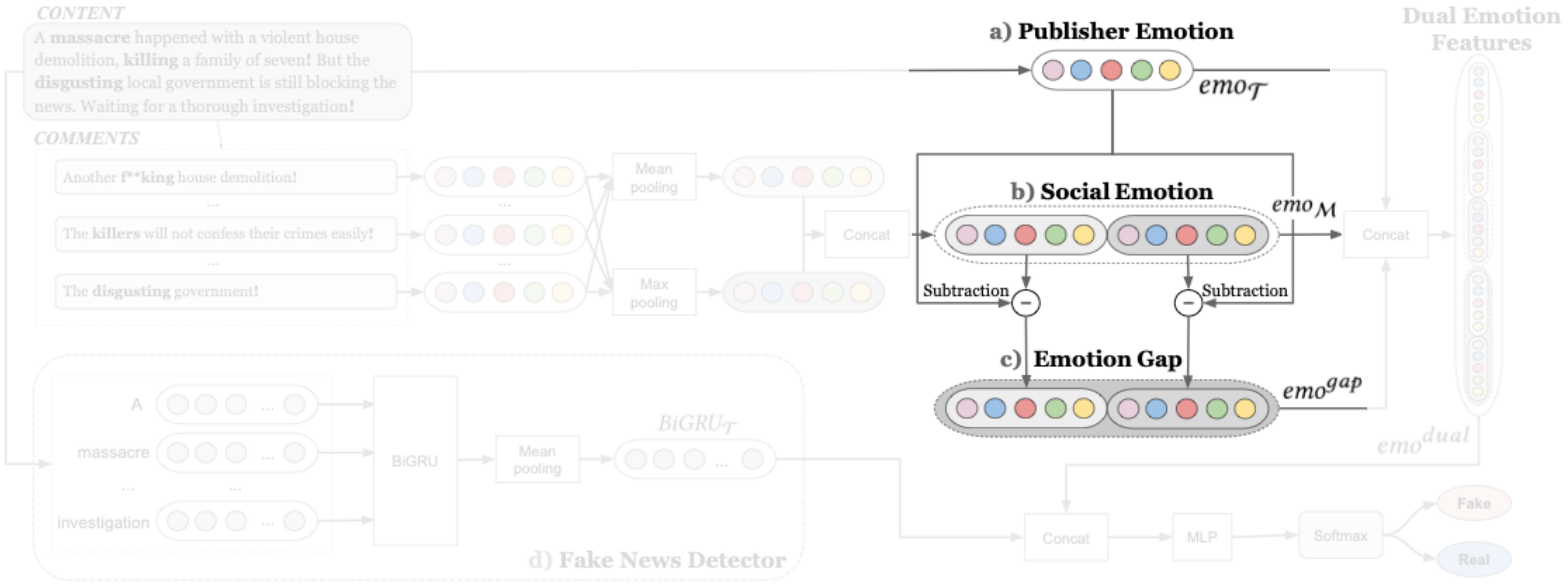
Social Emotion



$$emo_{\mathcal{M}} = emo_{\mathcal{M}}^{mean} \oplus emo_{\mathcal{M}}^{max}$$

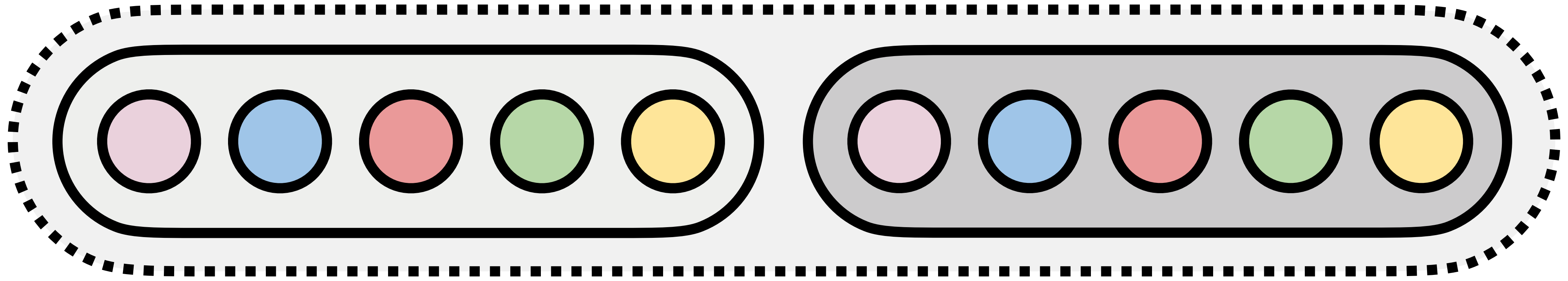
- After getting $\widehat{emo_{\mathcal{M}}}$, consider two aggregators to generate the social emotion of the whole comment list:
  - Mean pooling for representing the average emotional signals
    - $emo_{\mathcal{M}}^{mean} = mean(\widehat{emo_{\mathcal{M}}})$
  - Max pooling for capturing the extreme emotional signals
    - $emo_{\mathcal{M}}^{max} = max(\widehat{emo_{\mathcal{M}}})$
- Then concatenate them as Social Emotion: $emo_{\mathcal{M}} = emo_{\mathcal{M}}^{mean} \oplus emo_{\mathcal{M}}^{max}$

Methodology
Emotion Gap



Methodology

Emotion Gap



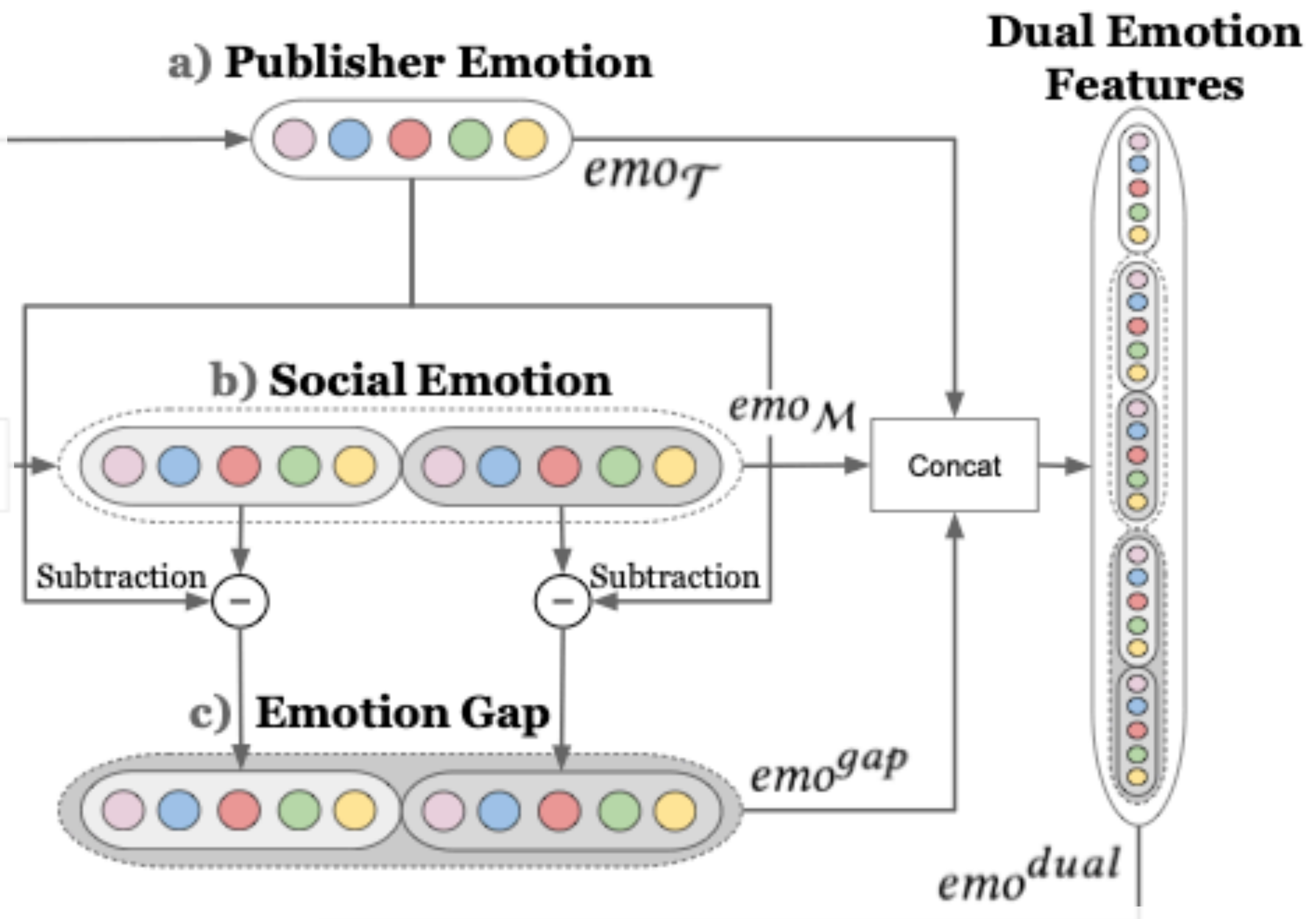
$$emo^{gap} = (emo_{\mathscr{T}} - emo_{\mathscr{M}}^{mean}) \oplus (emo_{\mathscr{T}} - emo_{\mathscr{M}}^{max})$$

- To model the resonances and dissonance of dual emotion, propose Emotion Gap emo^{gap}.
- It's designed as the subtraction between $emo_{\mathscr{T}}$ & $emo_{\mathscr{M}}$, emo^{gap} is concatenated by difference of $emo_{\mathscr{T}}$ & $emo_{\mathscr{M}}^{mean}$ and difference $emo_{\mathscr{T}}$ & $emo_{\mathscr{M}}^{max}$.
  - $emo^{gap} = (emo_{\mathscr{T}} - emo_{\mathscr{M}}^{mean}) \oplus (emo_{\mathscr{T}} - emo_{\mathscr{M}}^{max})$
- By this means, it can measure the differences (i.e., dissonances) between the dual emotion.
- For emotions resonances, the values in the Emotion Gap vector are tiny (nearly zero).

Methodology

Dual Emotion Features

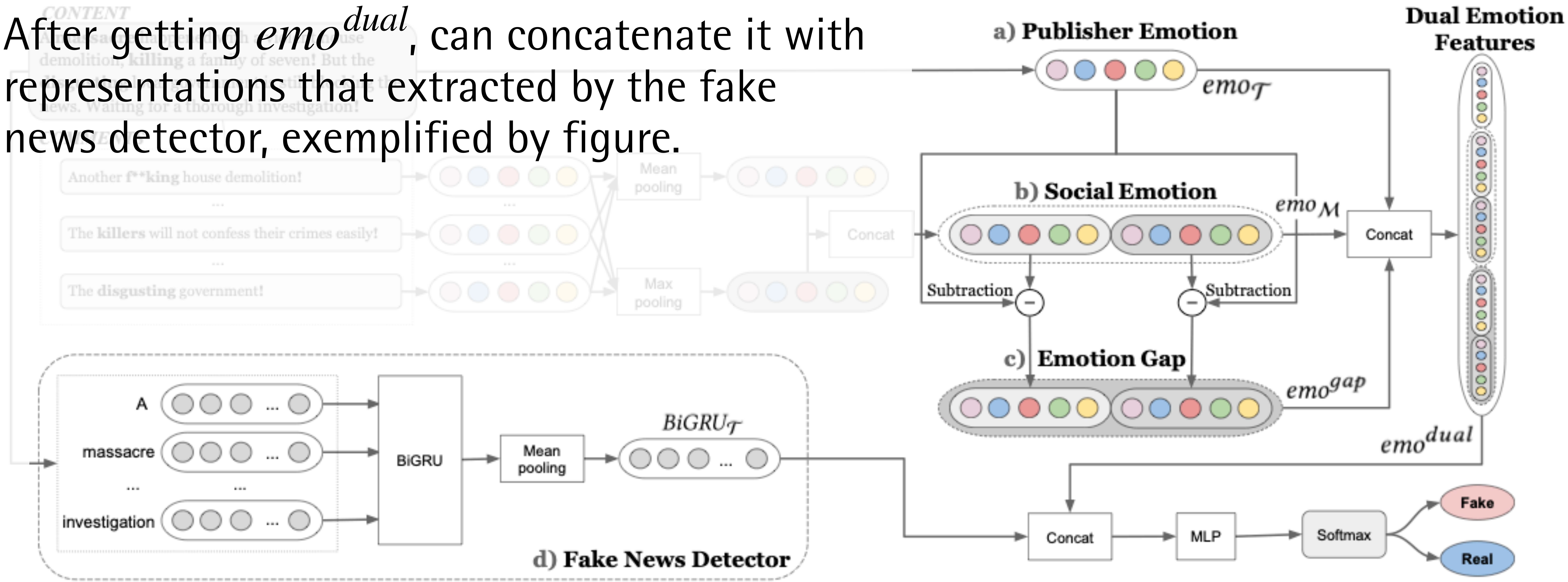
- Finally, Dual Emotion Features are concatenated by the Publisher Emotion, the Social Emotion and the Emotion Gap.
  - $emo^{dual} = emo_{\mathcal{T}} \oplus emo_{\mathcal{M}} \oplus emo^{gap}$



Methodology

Dual Emotion Features

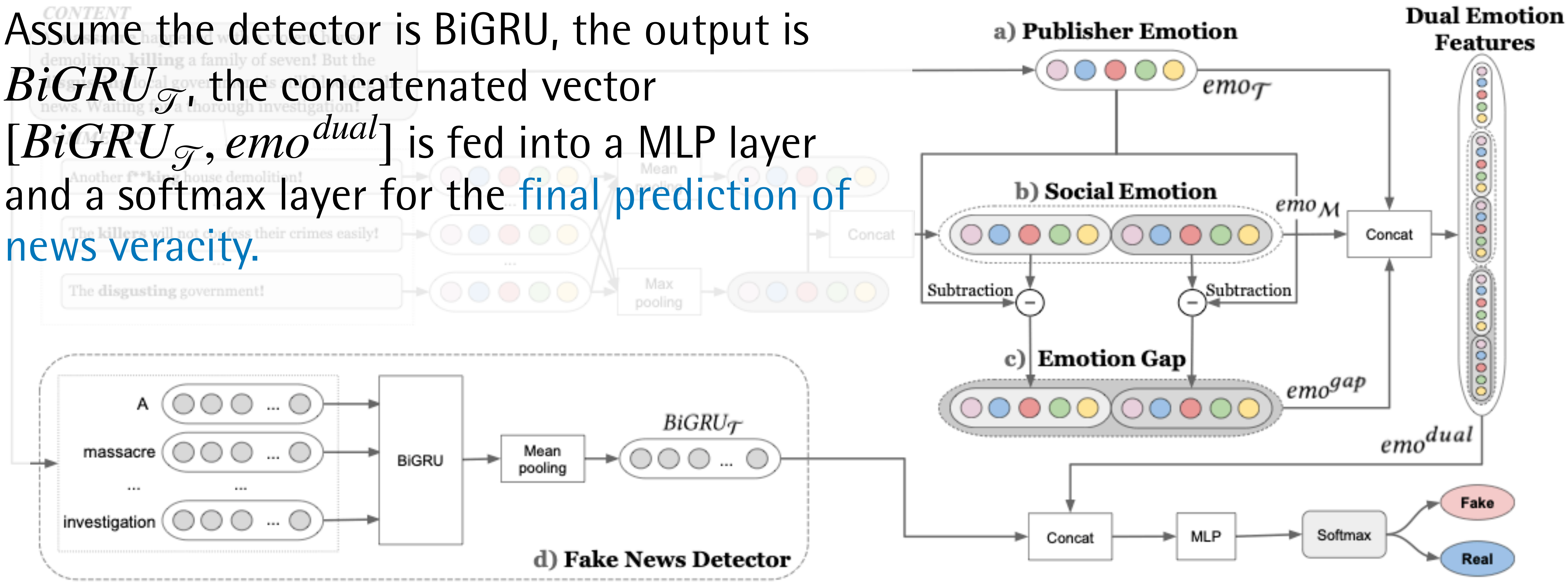
- After getting emo^{dual}, can concatenate it with representations that extracted by the fake news detector, exemplified by figure.



Methodology

Dual Emotion Features

- Assume the detector is BiGRU, the output is $BiGRU_{\mathcal{T}}$, the concatenated vector $[BiGRU_{\mathcal{T}}, emo^{dual}]$ is fed into a MLP layer and a softmax layer for the final prediction of news veracity.



$$\hat{y} = \mathrm{Softmax}(\mathrm{MLP}([BiGRU_{\mathcal{T}}, emo^{dual}]))$$

Experiments

Datasets

- RumourEval-19 (Twitter, Reddit)
- Weibo-16 (2010.12-2014.04)
    - Deduplication on the subset of fake news based on a clustering algorithm based on text similarity.
- Weibo-20 (2010.12-2018.11)
    - Further collect news from the same source with Weibo-20.

	Veracity	RumourEval-19		Weibo-16		Weibo-20	
		#pcs	#com	#pcs	# com	#pcs	#com
Training	Fake	79	1,135	801	649,673	1,896	749,141
	Real	144	1,905	1,410	482,226	1,920	516,795
	Unverified	104	1,838	-	-	-	-
	Total	327	4,878	2,211	1,131,899	3,816	1,265,936
Validating	Fake	19	824	268	222,149	632	137,941
	Real	10	404	470	146,948	640	185,087
	Unverified	9	212	-	-	-	-
	Total	38	1,440	738	369,097	1,272	323,028
Testing	Fake	40	689	286	193,740	633	245,216
	Real	31	805	471	179,942	641	149,260
	Unverified	10	181	-	-	-	-
	Total	81	1,675	757	373,682	1,274	394,476
Total	Fake	138	2,648	1,355	1,065,562	3,161	1,132,298
	Real	185	3,114	2,351	809,116	3,201	851,142
	Unverified	123	2,231	-	-	-	-
	Total	446	7,993	3,706	1,874,678	6,362	1,983,440

Experiments

Experiment Setup

- Emotion classifiers
  - Adopt pretrained models of NVIDIA (en) & Baidu AI (ch).
  - Randomly sampled 100 instance to ensure the robustness (NVIDIA / Baidu Acc: 87 / 83%).
- Emotion lexicon and intensity features
  - NRC Emotion lexicon (en) & Affective Lexicon Ontology (ch)
- Calculate sentiment scores by NLTK (en) & HowNet (ch).
- Auxiliary features from Wikipedia, HowNet, Oxford / Cambridge Dictionary.

Experiments

Baselines

- Select 2 baseline emotion features to evaluate the effectiveness of Dual Emotion Features.
    - **Emoratio** (ICASSP'19): It's calculated by the ratio of count of negative emotional words and count of positive emotional words.
    - **EmoCred** (SIGIR'19): Utilize emotional lexicon and intensity features of content texts, calculated based on the lexicons' occurrence frequency.

Experiments

Fake News Detectors

- For testing the ability of the emotional features to help the text-based fake news detectors (especially those do not explicitly model the emotional signals).
    - **BiGRU**: As word embedding using GloVe (en) Chinese Word Vectors (ch).
    - **BERT** (ECAI'20): Adopted to represent semantic signals when detecting fake news. Fine-tune the pretrained models for task.
    - **NileTMRG** (SemEval@ACL'17): For RumourEval-19, linear SVM and uses text features, social features, and use comments stance features.
    - **HSA-BLSTM** (CIKM'18): For Weibo datasets, proposed a hierarchical attention neural network and utilize not only the contents of news pieces but also comments.

Experiments

Evaluation Questions

- EQ1: Are Dual Emotion Features more effective than baseline features when used alone for fake news detection? How effective are the different types of features in Dual Emotion Features?

- EQ2: Can Dual Emotion Features help improve the performance of text-based fake news detectors?

- EQ3: How robust do the fake news detection models with Dual Emotion Features in real-world scenarios?

- EQ4: How effective are the components of Dual Emotion Features, including the publisher emotion, social emotion, and emotion gap?

Experiments

Evaluation Questions

- EQ1: Are Dual Emotion Features more effective than baseline features when used alone for fake news detection? How effective are the different types of features in Dual Emotion Features?
- EQ2: Can Dual Emotion Features help improve the performance of text-based fake news detectors?
- EQ3: How robust do the fake news detection models with Dual Emotion Features in real-world scenarios?
- EQ4: How effective are the components of Dual Emotion Features, including the publisher emotion, social emotion, and emotion gap?

Experiments

Effectiveness of Dual Emotion Features

Source	Emotion Features	R-19	W-16	W-20
Content	Emoratio	0.185	0.553	0.524
	EmoCred	0.253	0.564	0.542
	Publisher Emotion	0.290	0.571	0.573
Comments	**Social Emotion**	0.296	0.692	0.754
Content, Comments	**Emotion Gap**	0.332	0.716	0.746
	Dual Emotion Features	**0.337**	**0.728**	**0.759**

- Publisher Emotion is more effective than EmoCred and Emoratio, especially on the RumorEval.
  - It's reveals the effectiveness of Dual Emotion Features in modeling emotional signals.
- Moreover, observe that more improvements of Social Emotion and Emotion Gap,
- Using Dual Emotion Features can further obtain enhancements on the three datasets.
- *Clarified that RumourEval-19 are rather worse than the two Chinese datasets.
  - Because of the amount of news pieces is small and there is a relatively low inter-annotator agreement for the dataset.

Experiments

Effectiveness of different types features

Removed type	R-19	W-16	W-20
Emotion Category	0.193	0.679	0.686
Emotion Lexicon	0.239	0.715	0.745
Emotional Intensity	0.216	0.725	0.750
Sentiment Score	0.245	0.723	0.743
Other Auxiliary Features	0.307	0.653	0.722
Dual Emotion Features	**0.337**	**0.728**	**0.759**

- The macro F1 scores of Dual Emotion Features all decrease regardless of the removed type of emotion features.
- Thus, it reveals the necessity of using five types of emotion features jointly.

Experiments
Evaluation Questions

- EQ1: Are Dual Emotion Features more effective than baseline features when used alone for fake news detection? How effective are the different types of features in Dual Emotion Features?
- EQ2: Can Dual Emotion Features help improve the performance of text-based fake news detectors?
- EQ3: How robust do the fake news detection models with Dual Emotion Features in real-world scenarios?
- EQ4: How effective are the components of Dual Emotion Features, including the publisher emotion, social emotion, and emotion gap?

Experiments

Performance Comparison: RumourEval-19

Models	Macro F1 score	RMSE	F1 score		
			Fake News	Real News	Unverified News
BiGRU	0.269	0.804	0.500	0.222	0.083
+ Emoratio	0.275	0.823	0.463	0.160	**0.200**
+ EmoCred	0.311	0.797	0.456	0.295	0.182
+ **Dual Emotion Features**	**0.340**	**0.752**	**0.580**	**0.337**	0.104
BERT	0.272	0.808	0.533	0.105	0.176
+ Emoratio	0.271	0.857	0.406	0.240	0.167
+ EmoCred	0.308	0.833	0.367	**0.367**	0.189
+ **Dual Emotion Features**	**0.346**	**0.778**	**0.557**	0.244	**0.238**
NileTMRG	0.309	0.770	0.557	0.245	0.125
+ Emoratio	0.331	**0.754**	**0.571**	0.280	**0.143**
+ EmoCred	0.307	0.786	0.296	0.500	0.125
+ **Dual Emotion Features**	**0.342**	**0.754**	0.565	**0.565**	0.100

- After using Dual Emotion Features, the three fake news detectors are improve a lot.

Experiments

Performance Comparison: Weibo

Models	Weibo-16				Weibo-20			
	Macro F1 score	Accuracy	F1 score		Macro F1 score	Accuracy	F1 score	
			Fake	Real			Fake	Real
BiGRU	0.807	0.822	0.754	0.860	0.839	0.839	0.839	0.839
+ Emoratio	0.794	0.810	0.738	0.851	0.850	0.850	0.854	0.846
+ EmoCred	0.766	0.778	0.711	0.820	0.829	0.829	0.836	0.821
+ **Dual Emotion Features**	**0.826**	**0.838**	**0.781**	**0.871**	**0.855**	**0.855**	**0.857**	**0.852**
BERT	0.824	0.845	0.762	0.886	0.900	0.900	0.900	0.900
+ Emoratio	0.837	0.857	0.780	0.894	0.901	0.901	0.900	0.902
+ EmoCred	0.849	0.867	0.797	**0.901**	0.902	0.902	0.901	0.903
+ **Dual Emotion Features**	**0.867**	**0.873**	**0.837**	0.896	**0.915**	**0.915**	**0.913**	**0.918**
HSA-BLSTM	0.849	0.855	0.819	0.879	0.913	0.913	0.912	0.914
+ Emoratio	0.863	0.872	0.829	0.898	0.920	0.920	0.920	0.920
+ EmoCred	0.854	0.861	0.822	0.886	0.903	0.903	0.902	0.905
+ **Dual Emotion Features**	**0.908**	**0.913**	**0.885**	**0.930**	**0.932**	**0.932**	**0.932**	**0.933**

- Observed that Dual Emotion Features outperforms Emoratio & EmoCred on any models.
- However, when using Emoratio & EmoCred on BiGRU, sometimes the metrics even decrease, it reveals that Emoratio & EmoCred are more likely to be overfitted.

Experiments

Evaluation Questions

- EQ1: Are Dual Emotion Features more effective than baseline features when used alone for fake news detection? How effective are the different types of features in Dual Emotion Features?
- EQ2: Can Dual Emotion Features help improve the performance of text-based fake news detectors?
- EQ3: How robust do the fake news detection models with Dual Emotion Features in real-world scenarios?
- EQ4: How effective are the components of Dual Emotion Features, including the publisher emotion, social emotion, and emotion gap?

Experiments

Real-world scenarios

- Split dataset temporally.
- Such a scenario can somehow expose the drawback of existing techniques and it requires a model of higher generalizability to cope with novel instances.
- Under this hard setting, Dual Emotion Features still outperform others, reveals the effectiveness and generalization ability to some extent.

Models	Macro F1	Acc.	F1 score	
			Fake	Real
BiGRU	0.839	0.839	0.839	0.839
+ Emoratio	0.850	0.850	0.854	0.846
+ EmoCred	0.829	0.829	0.836	0.821
+ Dual Emotion Features	**0.855**	**0.855**	**0.857**	**0.852**
BERT	0.900	0.900	0.900	0.900
+ Emoratio	0.901	0.901	0.900	0.902
+ EmoCred	0.902	0.902	0.901	0.903
+ Dual Emotion Features	**0.915**	**0.915**	**0.913**	**0.918**
HSA-BLSTM	0.913	0.913	0.912	0.914
+ Emoratio	0.920	0.920	0.920	0.920
+ EmoCred	0.903	0.903	0.902	0.905
+ Dual Emotion Features	**0.932**	**0.932**	**0.932**	**0.933**

Models	Macro F1	Acc.	F1 score	
			Fake	Real
BiGRU	0.680	0.681	0.694	0.666
+ Emoratio	0.628	0.632	0.665	0.592
+ EmoCred	0.659	0.666	0.709	0.609
+ Dual Emotion Features	**0.701**	**0.702**	**0.714**	**0.689**
BERT	0.722	0.728	0.762	0.682
+ Emoratio	0.719	0.724	0.757	0.681
+ EmoCred	0.725	0.728	0.752	**0.699**
+ Dual Emotion Features	**0.734**	**0.734**	**0.773**	0.692
HSA-BLSTM	0.776	0.778	0.796	0.686
+ Emoratio	0.771	0.774	0.796	0.663
+ EmoCred	0.777	0.781	0.806	0.646
+ Dual Emotion Features	**0.805**	**0.808**	**0.827**	**0.694**

Table 7: Results on *Weibo-20* (temporal data split). Acc. is short for Accuracy.

Experiments

Evaluation Questions

- EQ1: Are Dual Emotion Features more effective than baseline features when used alone for fake news detection? How effective are the different types of features in Dual Emotion Features?
- EQ2: Can Dual Emotion Features help improve the performance of text-based fake news detectors?
- EQ3: How robust do the fake news detection models with Dual Emotion Features in real-world scenarios?
- EQ4: How effective are the components of Dual Emotion Features, including the publisher emotion, social emotion, and emotion gap?

Experiments

Ablation Study

- Observed that adding Dual Emotion Features into detectors all obtain the highest macro F1 scores.
- Exhibits that adopting Social Emotion or Emotion Gap improves the macro F1 scores more than Publisher Emotion on any models.
- Concludes that Social Emotion and Emotion Gap matter more when detecting fake news.

Models		R-19	W-16	W-20	W-20(t)
BiGRU+	Publisher Emotion	0.310	0.809	0.842	0.681
	Social Emotion	0.322	0.818	0.847	0.693
	Emotion Gap	0.336	0.811	0.849	0.693
	Dual Emotion Features	**0.340**	**0.826**	**0.855**	**0.701**
BERT+	Publisher Emotion	0.312	0.850	0.889	0.705
	Social Emotion	0.339	0.856	0.911	0.730
	Emotion Gap	0.338	0.858	0.906	0.731
	Dual Emotion Features	**0.346**	**0.867**	**0.915**	**0.734**
Nile TMRG+	Publisher Emotion	0.311	-	-	-
	Social Emotion	0.325	-	-	-
	Emotion Gap	0.337	-	-	-
	Dual Emotion Features	**0.342**	-	-	-
HSA-BLSTM+	Publisher Emotion	-	0.876	0.915	0.779
	Social Emotion	-	0.892	0.922	0.792
	Emotion Gap	-	0.901	0.926	0.800
	Dual Emotion Features	-	**0.908**	**0.932**	**0.805**

Table 8: Ablation study of the three components of *Dual Emotion Features*. The evaluation metric is macro F1 scores. R-19: RumourEval-19, W-16: Weibo-16, W-20: Weibo-20, and W-20(t): temporally split Weibo-20.

Experiments

Case Study

Content **Publisher emotion: *Angry***

Black Lives Matter **THUGS** Blocking Emergency Crews From Reaching Hurricane Victims.

Comments **Social emotion: *Angry***

Run over their **asses**.

...

When thugs block emergency vehicles, vehicles should **run over** the **thugs**

...

And why don't we have military support there to enforce the law?

	Fake	Real	Unverified
BiGRU	0.33	**0.61**	0.06
BiGRU + Emoratio	0.35	**0.57**	0.08
BiGRU + EmoCred	0.27	**0.64**	0.09
BiGRU + **Dual Emotion Features**	**0.65**	0.21	0.14

Content **Publisher emotion: *Joyful***

Believe it or not, this is a shark on the freeway in #Houston, #Texas #HurricaneHarvy...

Comments **Social emotion: *Joyful***

Fresh water? **Nice** try.

...

Oh my!

...

It's trying to ask you for directions...must be a female shark!! 😆

	Fake	Real	Unverified
BiGRU	0.31	**0.50**	0.19
BiGRU + Emoratio	0.36	**0.56**	0.08
BiGRU + EmoCred	0.40	**0.54**	0.06
BiGRU + **Dual Emotion Features**	**0.65**	0.22	0.13

Content **Publisher emotion: *None***

Category 6? If Hurricane Irma Becomes The Strongest Hurricane In History, It Could Wipe Entire Cities Off The Map.

Comments **Social emotion: *Sad***

Watch what the power of **prayer** does

...

Get out of there people. Florida, South Carolina.

...

Pray to god the only chance you have...

	Fake	Real	Unverified
BiGRU	0.47	**0.52**	0.01
BiGRU + Emoratio	0.40	**0.59**	0.01
BiGRU + EmoCred	0.28	**0.58**	0.14
BiGRU + **Dual Emotion Features**	**0.63**	0.17	0.20

- It exhibits using Emoratio & EmoCred do not help BiGRU detect rightly for the three cases.
- It reveals that mining dual emotion additionally sometimes is a remedy for incompetence of only using semantics for detecting fake news.

Conclusion and Future Work

- Bring a new concept of dual emotion (i.e., publisher emotion, social emotion).
- Uncover the relationship between dual emotion signals (emotion gap).
- Proposed Dual Emotion Features expose the distinctive emotional signals for detecting fake news.
- The proposed features can be easily plugged into existing fake news detectors as enhancement.
- In future work, authors plan to leverage the multi-modal information to capture the emotions more precisely for dual emotion representation.

Comments

of Dual Emotion Features

- Focus on emotion signal to deal with fake news detection.
- Redesign the emotion feature by 5 kinds of statistical feature.
  - Most features are from existing methods or toolkits.
- Discover the relationship between posts & comments.
- Easy to plugged into existing fake news detector as enhancement.