硕士学位论文答辩

基于情感的虚假新闻检测方法研究

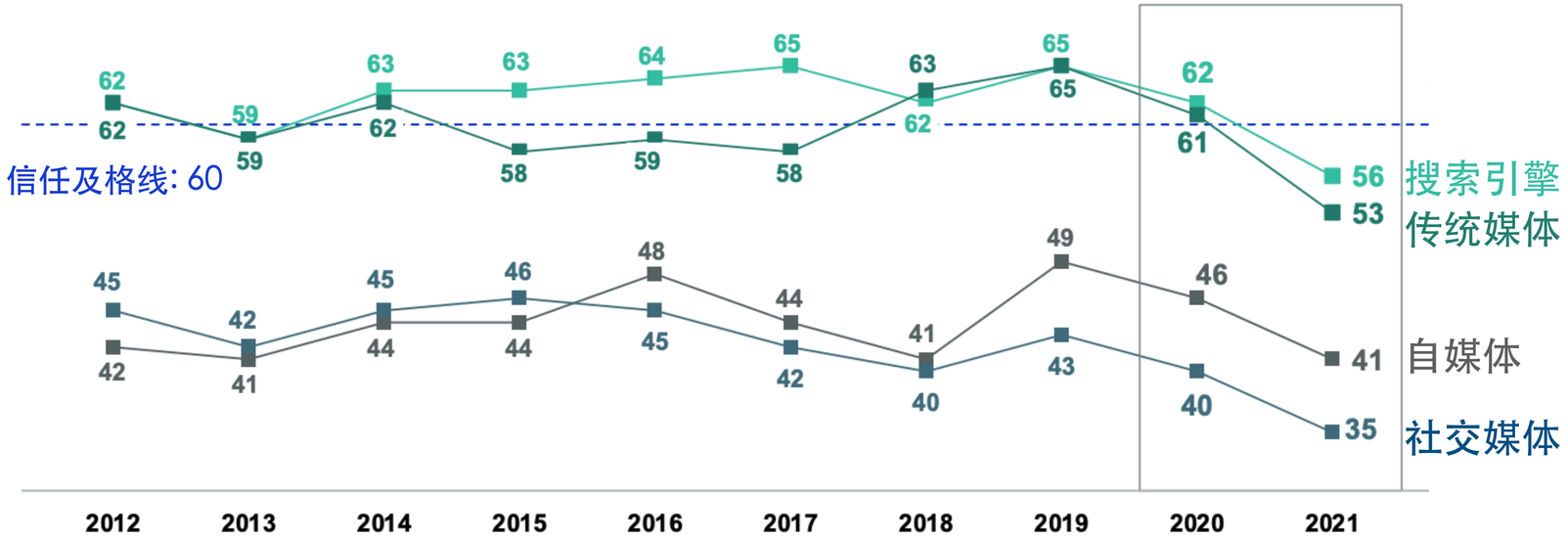
学生姓名：张雪遥
指导教师：曹娟 研究员
培养类别：2019级硕士生
所在部门：数字内容合成与伪造检测实验室

2022年5月25日

目录

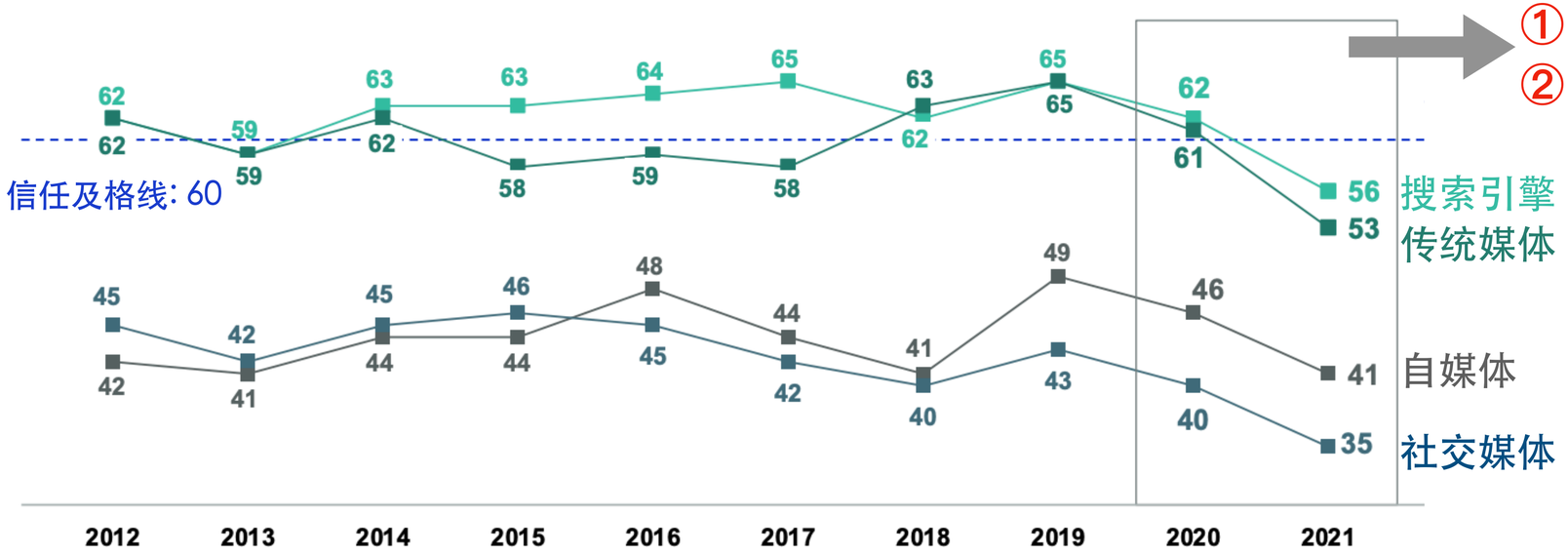
目录

当今的社交媒体上存在巨大的信任危机！



爱德曼公司《2021年全球信任度调查报告》

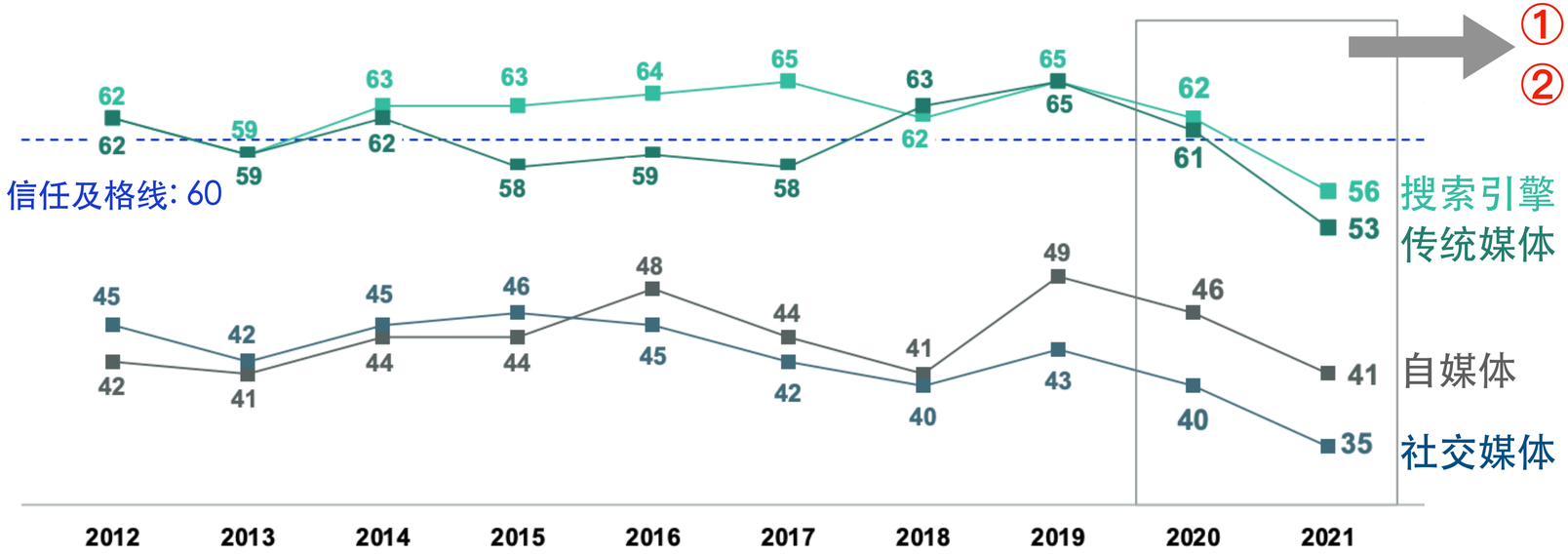
当今的社交媒体上存在巨大的信任危机！



爱德曼公司《2021年全球信任度调查报告》

① 人们对于多种信息媒介的信任度都在持续下降
② 社交媒体始终是人们信任度最低的信息来源

当今的社交媒体上存在巨大的信任危机！



① 人们对于多种信息媒介的信任度都在持续下降
② 社交媒体始终是人们信任度最低的信息来源

爱德曼公司《2021年全球信任度调查报告》

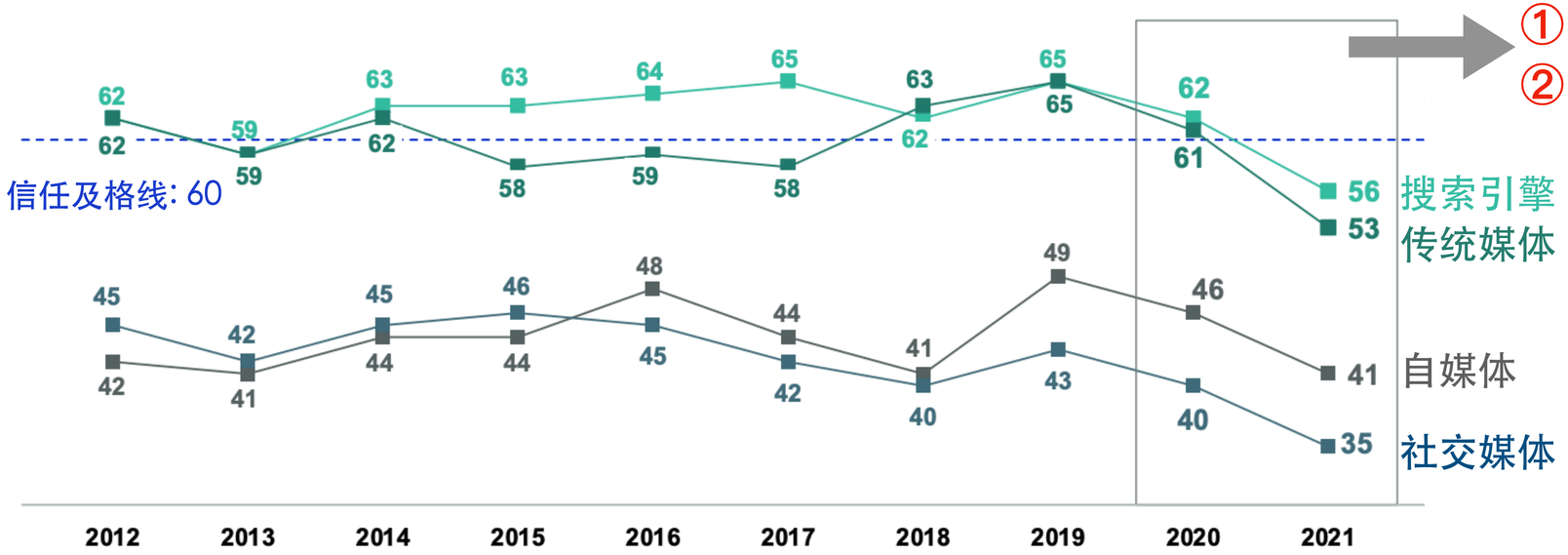
'Fake news' is the number one worry for internet users worldwide

Seventy-one percent of people who use the internet recognised at least one of the three main internet-related risks — the biggest concern was fake news, ahead of fraud and cyberbullying.

虚假新闻已成为全球互联网用户的头号担忧

2020年 - 劳氏基金会《世界风险调查》

当今的社交媒体上存在巨大的信任危机！



爱德曼公司《2021年全球信任度调查报告》

① 人们对于多种信息媒介的信任度都在持续下降
② 社交媒体始终是人们信任度最低的信息来源

SOCIAL SCIENCE

The spread of true and false news online

Soroush Vosoughi,[1] Deb Roy,[1] Sinan Aral[2]*

We investigated the differential diffusion of all of the verified true and false news stories distributed on Twitter from 2006 to 2017. The data comprise ~126,000 stories tweeted by ~3 million people more than 4.5 million times. We classified news as true or false using information from six independent fact-checking organizations that exhibited 95 to 98% agreement on the classifications. Falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information, and the effects were more pronounced for false political news than for false news about terrorism, natural disasters, science, urban legends, or financial information. We found that false news was more novel than true news, which suggests that people were more likely to share novel information. Whereas false stories inspired fear, disgust, and surprise in replies, true stories inspired anticipation, sadness, joy, and trust. Contrary to conventional wisdom, robots accelerated the spread of true and false news at the same rate, implying that false news spreads more than the truth because humans, not robots, are more likely to spread it.

2018年《科学》期刊

虚假新闻比客观真相的传播速度更快、影响力更大、涉足领域更广

'Fake news' is the number one worry for internet users worldwide

Seventy-one percent of people who use the internet recognised at least one of the three main internet-related risks — the biggest concern was fake news, ahead of fraud and cyberbullying.

虚假新闻已成为全球互联网用户的头号担忧

2020年 - 劳氏基金会《世界风险调查》

虚假新闻严重危害了人类社会的各个领域！

新华访谈：反对历史虚无主义，解读涉党史类谣言

联合国教科文组织：新冠疫情导致了信息疫情

Promoting Reliable Information and Reducing Harmful Misinformation About COVID-19

COVID-19 is still a major public health issue, and we are committed to helping people get authoritative information, including vaccine information. We continue to remove harmful COVID-19 misinformation and prohibit ads that try to exploit the pandemic for financial gain. Since the start of the pandemic through June:

- We removed more than 20 million pieces of content from Facebook and Instagram globally for violating our policies on COVID-19-related misinformation.
- We have removed over 3,000 accounts, pages, and groups for repeatedly violating our rules against spreading COVID-19 and vaccine misinformation.
- We displayed warnings on more than 190 million pieces of COVID-related content on Facebook that our third-party fact-checking partners rated as false, partly false, altered or missing context, collaborating with 80 fact-checking organizations in more than 60 languages around the world. When they rate a piece of content with one of these ratings, we add a prominent label warning people before they share it and show it lower in people's feed.

Facebook：有超过2000万个和疫情有关的虚假新闻帖子被删除，有超过1.9亿个存在错误信息的帖子被标注了警示信息(2020.6-2021.8)

美国国家经济研究局：“财经类虚假新闻极大影响了美国证券交易市场”

Social Media and Financial News Manipulation

74 Pages • Posted: 31 Aug 2018 • Last revised: 15 Sep 2021

Shimon Kogan
IDC Herzliya - Arison School of Business; University of Pennsylvania - The Wharton School
Tobias J. Moskowitz
Yale University, Yale SOM; AQR Capital; National Bureau of Economic Research (NBER)
Marina Niessner
University of Pennsylvania - The Wharton School

Date Written: September 15, 2021

Abstract

We dissect an undercover SEC investigation into the manipulation of financial news on social media to study the indirect effects of market manipulation. While fraudulent news had a direct impact on retail trading and prices, revelation of the fraud caused market participants to discount all news, including legitimate news, from these platforms. The results highlight the indirect consequences of fraud and its spillover effects that reduce the social network's impact on information dissemination, especially for small firms. The effect appears to dissipate over time, becoming insignificant a year later. The results highlight the importance of social capital for financial activity.

危害国家安全　危害社会生活　危害经济市场

虚假新闻严重危害了人类社会的各个领域！

新华访谈：反对历史虚无主义，解读涉党史类谣言

联合国教科文组织：新冠疫情导致了信息疫情

Promoting Reliable Information and Reducing Harmful Misinformation About COVID-19

COVID-19 is still a major public health issue, and we are committed to helping people get authoritative information, including vaccine information. We continue to remove harmful COVID-19 misinformation and prohibit ads that try to exploit the pandemic for financial gain. Since the start of the pandemic through June:

- We removed more than 20 million pieces of content from Facebook and Instagram globally for violating our policies on COVID-19-related misinformation.
- We have removed over 3,000 accounts, pages, and groups for repeatedly violating our rules against spreading COVID-19 and vaccine misinformation.
- We displayed warnings on more than 190 million pieces of COVID-related content on Facebook that our third-party fact-checking partners rated as false, partly false, altered or missing context, collaborating with 80 fact-checking organizations in more than 60 languages around the world. When they rate a piece of content with one of these ratings, we add a prominent label warning people before they share it and show it lower in people's feed.

Facebook：有超过2000万个和疫情有关的虚假新闻帖子被删除，有超过1.9亿个存在错误信息的帖子被标注了警示信息(2020.6-2021.8)

美国国家经济研究局：“财经类虚假新闻极大影响了美国证券交易市场”

Social Media and Financial News Manipulation

74 Pages • Posted: 31 Aug 2018 • Last revised: 15 Sep 2021

Shimon Kogan
IDC Herzliya - Arison School of Business; University of Pennsylvania - The Wharton School
Tobias J. Moskowitz
Yale University, Yale SOM; AQR Capital; National Bureau of Economic Research (NBER)
Marina Niessner
University of Pennsylvania - The Wharton School

Date Written: September 15, 2021

Abstract

We dissect an undercover SEC investigation into the manipulation of financial news on social media to study the indirect effects of market manipulation. While fraudulent news had a direct impact on retail trading and prices, revelation of the fraud caused market participants to discount all news, including legitimate news, from these platforms. The results highlight the indirect consequences of fraud and its spillover effects that reduce the social network's impact on information dissemination, especially for small firms. The effect appears to dissipate over time, becoming insignificant a year later. The results highlight the importance of social capital for financial activity.

危害国家安全　危害社会生活　危害经济市场

虚假新闻严重危害了人类社会的各个领域！

新华访谈：反对历史虚无主义，解读涉党史类谣言

联合国教科文组织：新冠疫情导致了信息疫情

研制出能自动化检测虚假新闻的技术迫在眉睫

Facebook：有超过2000万个和疫情有关的虚假新闻帖子被删除，有超过1.9亿个存在错误信息的帖子被标注了警示信息(2020.6-2021.8)

美国国家经济研究局：“财经类虚假新闻极大影响了美国证券交易市场”

危害国家安全　危害社会生活　危害经济市场

目录

相关术语定义

虚假新闻

- 本研究中沿用的定义为：**虚假新闻是指故意捏造并可被证实为假的消息** [1-2]，且我们关注的是在社交媒体 (如微博、推特）上发布的**互联网在线新闻**。

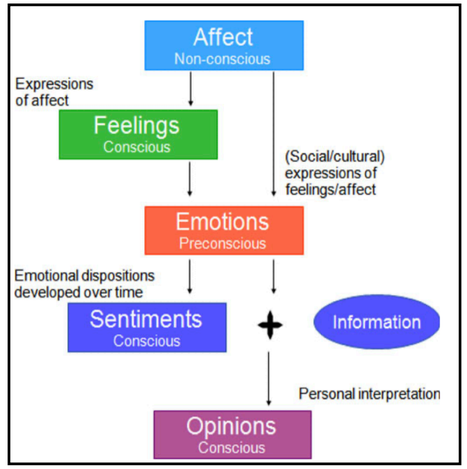
	Authenticity	Intention	News?
Fake news	False	Bad	Yes
False news	False	Unknown	Yes
Satire news	Unknown	Not bad	Yes
Disinformation	False	Bad	Unknown
Misinformation	False	Unknown	Unknown
Rumor	Unknown	Unknown	Unknown

本图源自[1]

“虚假新闻”相关术语之间的联系

情感

- 相关术语：情感 (affection)，情感态度 (sentiment)，情绪 (emotion)，感受 (feeling)等 [3].
- 本研究使用“情感”来泛指**各种粒度、各个层面的综合情感**。例如：在研究点一中，我们建模的“情感特征”中既包含了细粒度的情绪特征 (emotion)，也包含了正负极性的情感特征 (sentiment).



本图源自[3]

“情感”相关术语之间的联系

[1] Xinyi Zhou, et al. A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Comput. Surv., 2020.
[2] Kai Shu, et al. Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 2017.
[3] Myriam Munezero, et al. Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. IEEE Transactions on Affective Computing, 2014.

社会心理学的相关研究：情感与虚假新闻存在内在关联

1947年：*The Psychology of Rumor* [1]

MOTIVES IN RUMOR MONGERING 37

It is important to note here the complex purpose that rumor serves. By permitting one to slap at the thing one hates it *relieves* a primary emotional urge. But at the same time—in the same breath—it serves to *justify* one in feeling as he does about the situation, and to *explain* to himself and to others why he feels that way. Thus rumor rationalizes while it relieves. "Why shouldn't I dislike Russia? It came to our aid only at the cost of an enormous bribe. . . ." "Why shouldn't I feel panicky? Our fleet was wiped out at Pearl Harbor. . . ." "Why shouldn't I distrust the Jews? They are so clannish. . . ." "Why shouldn't I feel superior to my neighbor? I don't indulge in his irregularities of living. . . ."

① 假新闻能够释放人们心中的情感冲动
② 假新闻能够让人产生对当前处境的共鸣
③ 传播假新闻其实是一种向他人表达情感与倾诉自我的方式

1991年：*Inside Rumor: A Personal Journey* [2]

Personal Anxiety

The third variable that I noted, personal anxiety, is without a direct counterpart in the basic law of rumor. G. W. Allport and Postman (1947) did raise the possibility that ambiguity might be "induced . . . by some emotional tensions that make the individual unable or unwilling to accept the facts set forth in the news" (p. 33). By *personal anxiety,* as I shall use the term, I mean an affective state—acute or chronic—that is produced by, or associated with, apprehension about an impending, potentially disappointing outcome (Rosnow, 1980). Here, the hypothesis is that rumors persist not only because they play on cognitive unclarity, but also because they give vent or expression to emotional tensions attributable to the nature of anticipated outcomes (cf. G. W. Allport & Postman, 1947; Ambrosini, 1983; Festinger, 1957; Hart, 1916; Jung, 1910, 1959; Loewenberg, 1943).

对[1]中提出的"谣言公式"，应融入"个人焦虑" (Personal Anxiety) 作为因子

谣言 ~ 重要性 × 模糊度 × 个人焦虑

[1] Allport G W, Postman L. The psychology of rumor. Henry Holt, 1947.
[2] Ralph L Rosnow. Inside rumor: A personal journey. American psychologist, 1991.

计算机科学的现有研究：基于情感信息的虚假新闻检测

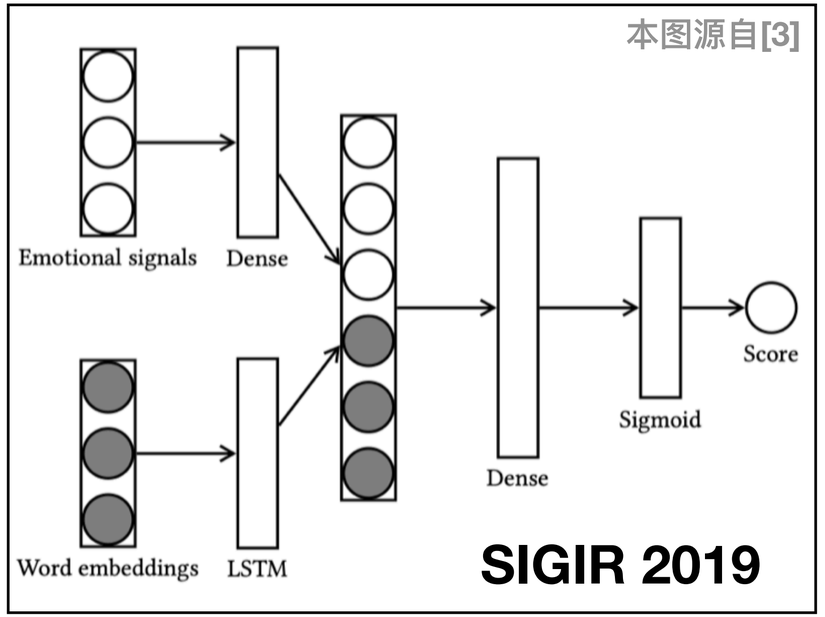
- **最早的工作（WWW 2021）**：2011年，Castillo等人在实验中发现了**与情感相关的特征**(如正负向情感词、感叹号等）对**新闻可信度评估**很重要 [1].
- 后来，陆续有学者在**新闻文本**中挖掘不同的**情感特征**，来辅助**虚假新闻的检测** [2-3].

模型

J48 决策树

实验发现

与情感相关的特征更靠近决策树模型的根结点



模型

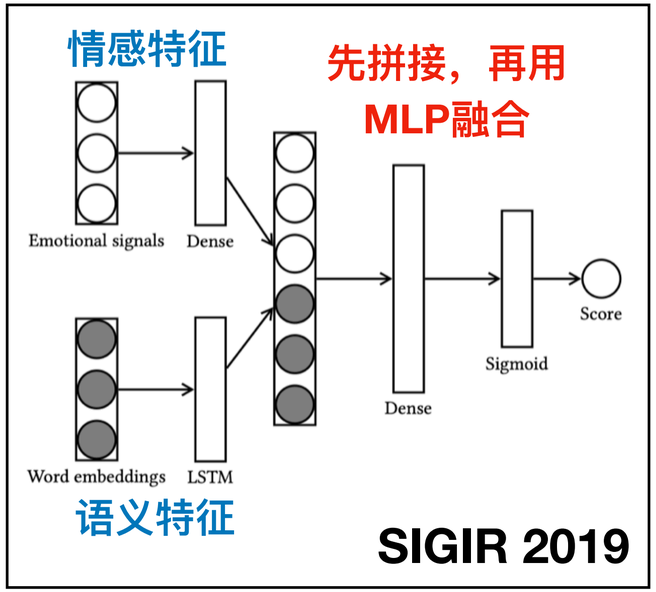
LSTM

实验发现

额外融入情感信号，能增强LSTM的检测效果

[1] Carlos Castillo, et al. Information credibility on twitter. WWW 2011.
[2] Oluwaseun Ajao, et al. Sentiment Aware Fake News Detection on Online Social Networks. IEEE ICASSP 2019.
[3] Anastasia Giachanou, et al. Leveraging Emotional Signals for Credibility Detection. ACM SIGIR 2019.

总结：现有工作利用情感信息的范式



1. 抽取新闻发布者情感

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

抽取的情感特征
(Emotion Features)

虚假新闻检测器
(Fake News Detector)

抽取的检测特征
(Detection Features)

特征融合
(Fusion)

可信度预测
(Prediction)

2. 情感与其他特征的融合

现有研究的局限性（本文的研究思路）

现有研究的局限性（本文的研究思路）

现有研究

从新闻原文中抽取情感特征

只建模新闻发布者的情感，
忽略了假新闻对读者的情绪煽动

现有研究的局限性（本文的研究思路）

现有研究

从新闻原文中抽取情感特征

只建模新闻发布者的情感，
忽略了假新闻对读者的情绪煽动

研究点一

基于双重情感的虚假新闻检测

建模社区群体的情感，
并挖掘双重情感之间的联系

现有研究的局限性（本文的研究思路）

现有研究

从新闻原文中抽取情感特征

只建模新闻发布者的情感，
忽略了假新闻对读者的情绪煽动

研究点一

基于双重情感的虚假新闻检测

建模社区群体的情感，
并挖掘双重情感之间的联系

只把情感作为辅助特征，
忽略了可以增强模型自身对情感的表征

现有研究的局限性（本文的研究思路）

现有研究

从新闻原文中抽取情感特征

只建模新闻发布者的情感，
忽略了假新闻对读者的情绪煽动

研究点一

基于双重情感的虚假新闻检测

建模社区群体的情感，
并挖掘双重情感之间的联系

只把情感作为辅助特征，
忽略了可以增强模型自身对情感的表征

研究点二

情感偏好增强的虚假新闻即时检测

对模型的学习过程加以引导，
增强其对情感的偏好

目录

研究点一：基于双重情感的虚假新闻检测

新闻原文 惊异、愤怒

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

社区评论 愤怒、厌恶

- 又是强拆！
- ...
- 杀人者不可饶恕！
- ...
- 该死的政府...

假新闻示例1

新闻原文 无明显情感

被扣押的孟晚舟女士在加拿大出庭时，脚上正穿着一双浅蓝色的鸿星尔克运动鞋

社区评论 赞许、质疑

- 不得不佩服这些企业家！
- ...
- 你看清了？是鸿星尔克？
- ...
- 新时代绝对英雄👍

假新闻示例2

研究动机：假新闻的发布者往往会煽动起群众的激烈情绪

研究点一：基于双重情感的虚假新闻检测

新闻原文 惊异、愤怒

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

社区评论 愤怒、厌恶

- 又是强拆！
- ...
- 杀人者不可饶恕！
- ...
- 该死的政府...

假新闻示例1

新闻原文 无明显情感

被扣押的孟晚舟女士在加拿大出庭时，脚上正穿着一双浅蓝色的鸿星尔克运动鞋

社区评论 赞许、质疑

- 不得不佩服这些企业家！
- ...
- 你看清了？是鸿星尔克？
- ...
- 新时代绝对英雄👍

假新闻示例2

现有工作

新闻发布者情感

双重情感

社区群体情感

本研究

研究动机：假新闻的发布者往往会煽动起群众的激烈情绪

数据分析：假新闻中的双重情感究竟有何特别?

数据分析：假新闻中的双重情感究竟有何特别?

- **新闻发布者情感**
  - 真、假新闻的情感种类，情感强度，以及情感词表达上有显著区别 [1]

[1] Chuan Guo, Juan Cao, **Xueyao Zhang**, et al. Exploiting Emotions for Fake News Detection on Social Media. arXiv:1903.01728, 2019.
[2] Soroush Vosoughi, et al. The Spread of True and False News Online. Science, 2018.

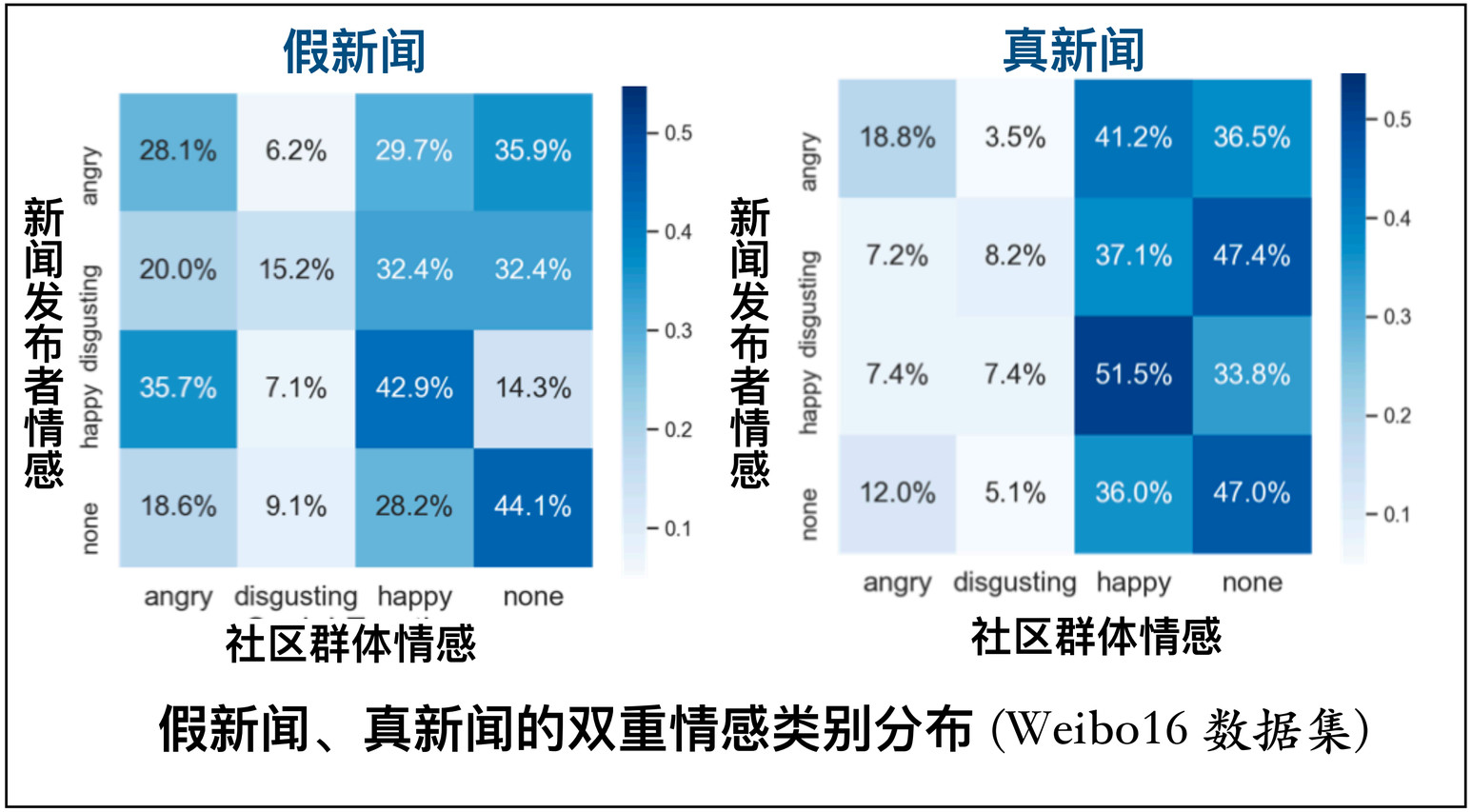
数据分析：假新闻中的双重情感究竟有何特别?

- **新闻发布者情感**
  - 真、假新闻的情感种类，情感强度，以及情感词表达上有显著区别 [1]
- **社区群体情感**
  - 真、假新闻会激起读者不同种类的情感 [2]

[1] Chuan Guo, Juan Cao, **Xueyao Zhang**, et al. Exploiting Emotions for Fake News Detection on Social Media. arXiv:1903.01728, 2019.
[2] Soroush Vosoughi, et al. The Spread of True and False News Online. Science, 2018.

数据分析：假新闻中的双重情感究竟有何特别?

- **新闻发布者情感**
  - 真、假新闻的情感种类，情感强度，以及情感词表达上有显著区别 [1]
- **社区群体情感**
  - 真、假新闻会激起读者不同种类的情感 [2]
- **双重情感之间的关系**



假新闻、真新闻的双重情感类别分布 (Weibo16 数据集)

卡方检验结果：双重情感类别与新闻的可信度显著相关 (p<0.01)

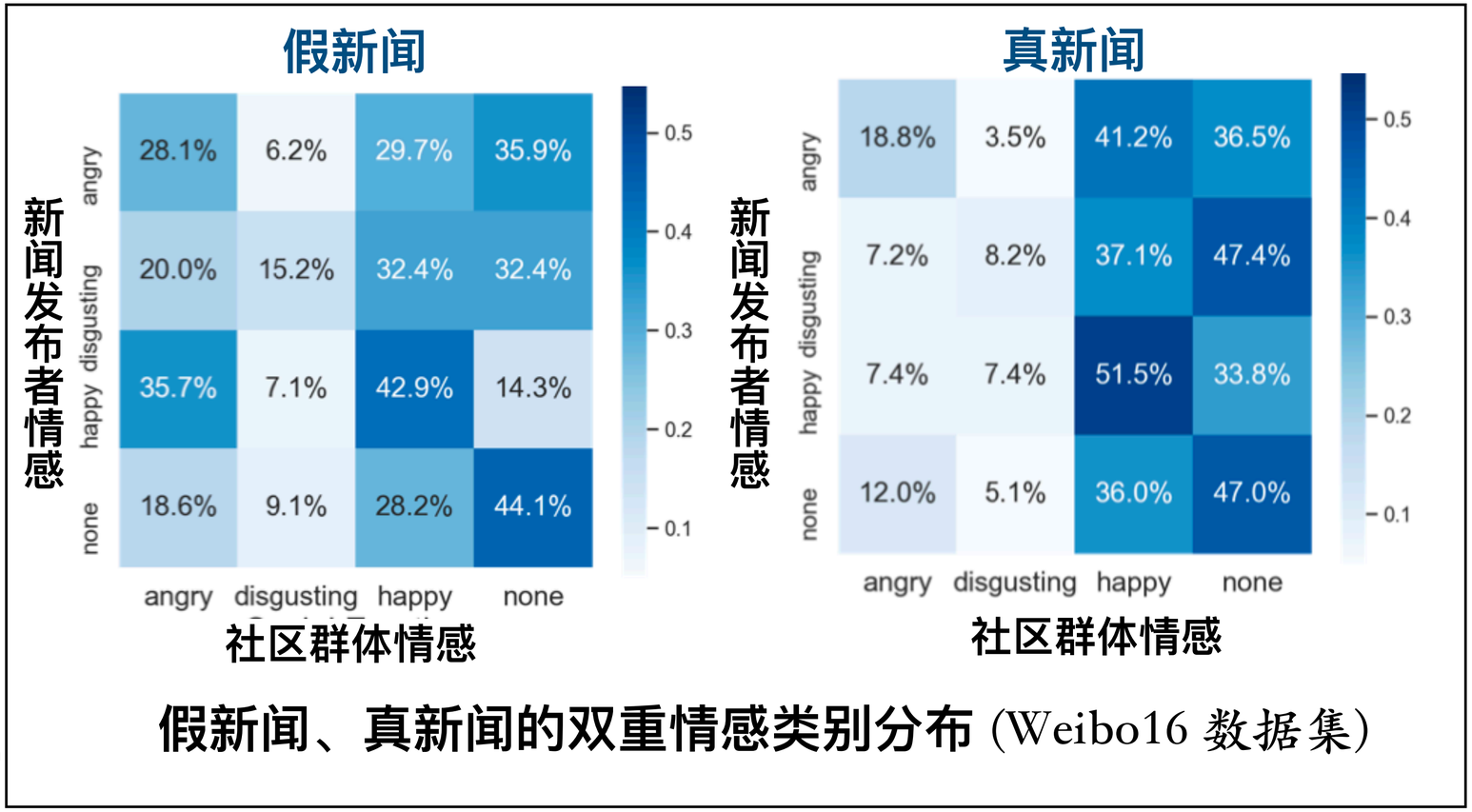
[1] Chuan Guo, Juan Cao, **Xueyao Zhang**, et al. Exploiting Emotions for Fake News Detection on Social Media. arXiv:1903.01728, 2019.
[2] Soroush Vosoughi, et al. The Spread of True and False News Online. Science, 2018.

数据分析：假新闻中的双重情感究竟有何特别?

- **新闻发布者情感**
  - 真、假新闻的情感种类，情感强度，以及情感词表达上有显著区别 [1]
- **社区群体情感**
  - 真、假新闻会激起读者不同种类的情感 [2]
- **双重情感之间的关系**

结论

双重情感间的共鸣与分歧在真、假新闻之间显著不同



假新闻、真新闻的双重情感类别分布 (Weibo16 数据集)

卡方检验结果：双重情感类别与新闻的可信度显著相关 ($p<0.01$)

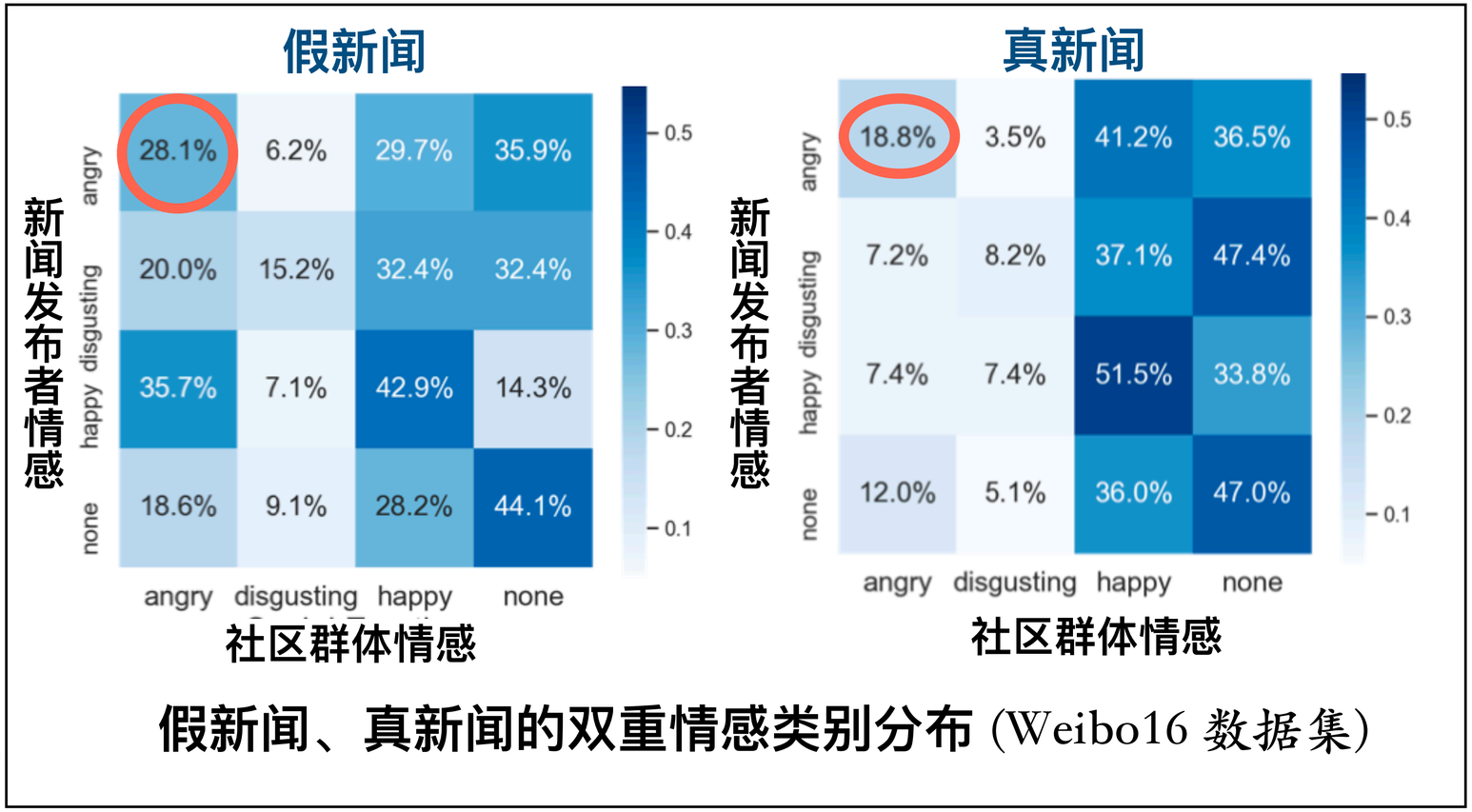
[1] Chuan Guo, Juan Cao, **Xueyao Zhang**, et al. Exploiting Emotions for Fake News Detection on Social Media. arXiv:1903.01728, 2019.
[2] Soroush Vosoughi, et al. The Spread of True and False News Online. Science, 2018.

数据分析：假新闻中的双重情感究竟有何特别?

- **新闻发布者情感**
  - 真、假新闻的情感种类，情感强度，以及情感词表达上有显著区别 [1]
- **社区群体情感**
  - 真、假新闻会激起读者不同种类的情感 [2]
- **双重情感之间的关系**

结论

双重情感间的共鸣与分歧在真、假新闻之间显著不同



假新闻、真新闻的双重情感类别分布 (Weibo16 数据集)

卡方检验结果：双重情感类别与新闻的可信度显著相关 ($p<0.01$)

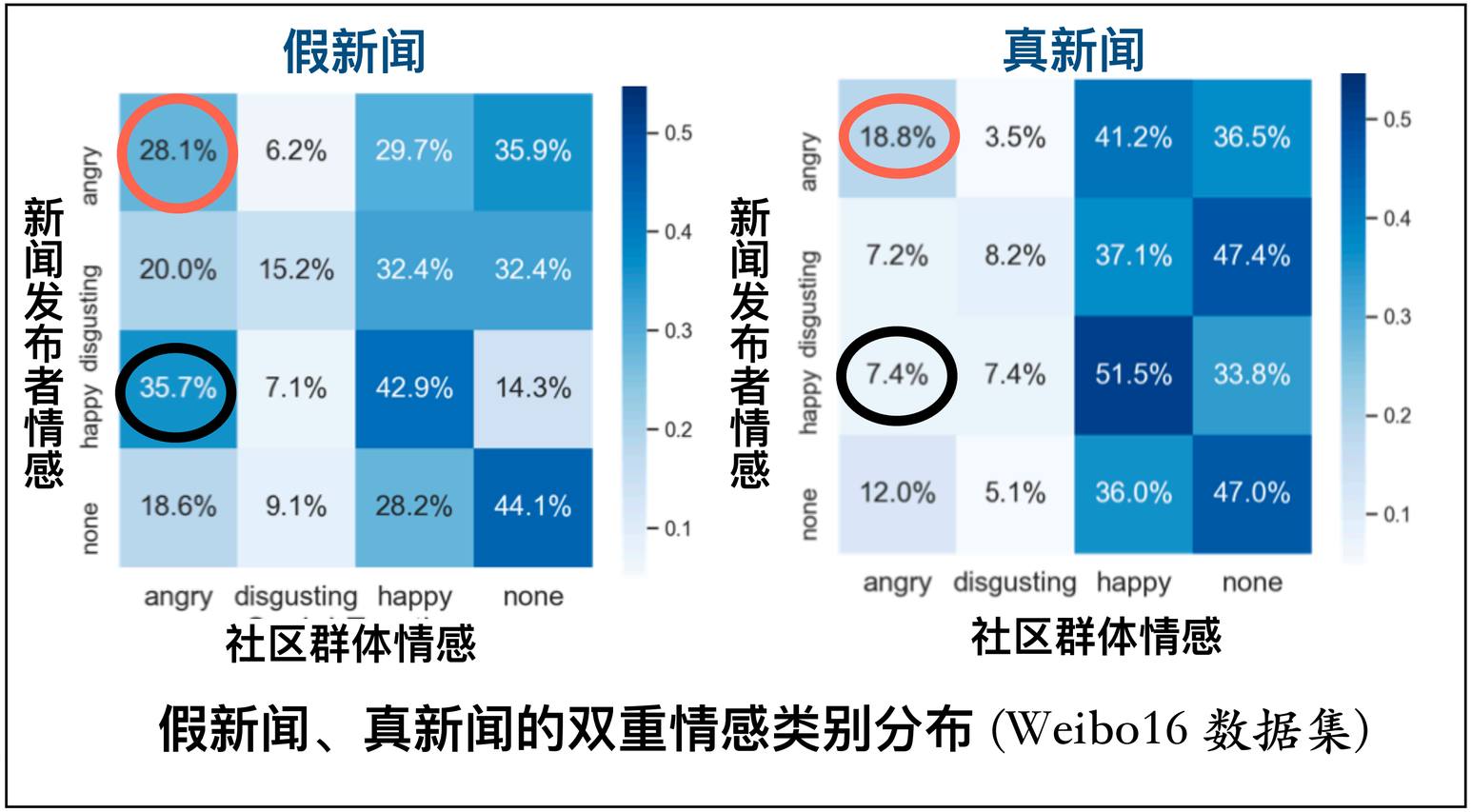
[1] Chuan Guo, Juan Cao, **Xueyao Zhang**, et al. Exploiting Emotions for Fake News Detection on Social Media. arXiv:1903.01728, 2019.
[2] Soroush Vosoughi, et al. The Spread of True and False News Online. Science, 2018.

数据分析：假新闻中的双重情感究竟有何特别?

- **新闻发布者情感**
  - 真、假新闻的情感种类，情感强度，以及情感词表达上有显著区别 [1]
- **社区群体情感**
  - 真、假新闻会激起读者不同种类的情感 [2]
- **双重情感之间的关系**

结论

双重情感间的共鸣与分歧在真、假新闻之间显著不同

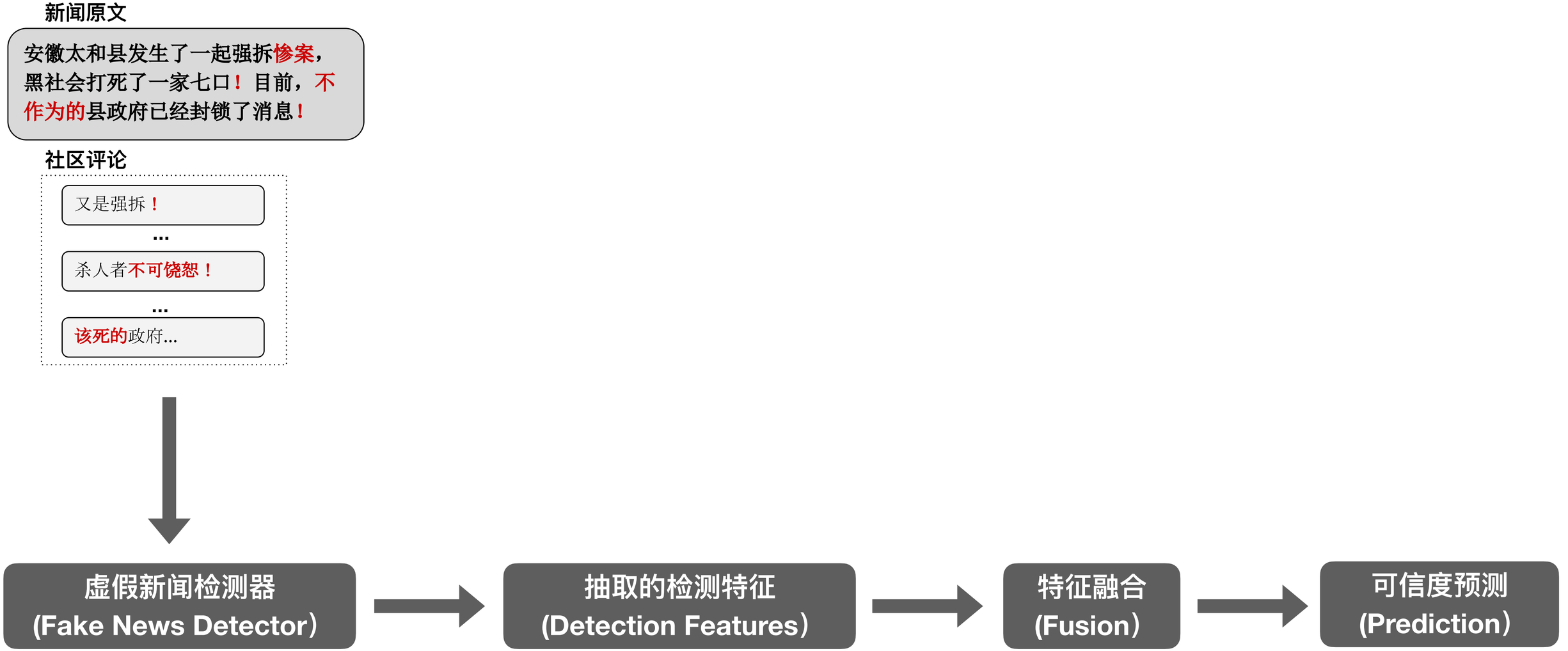


假新闻、真新闻的双重情感类别分布 (Weibo16 数据集)

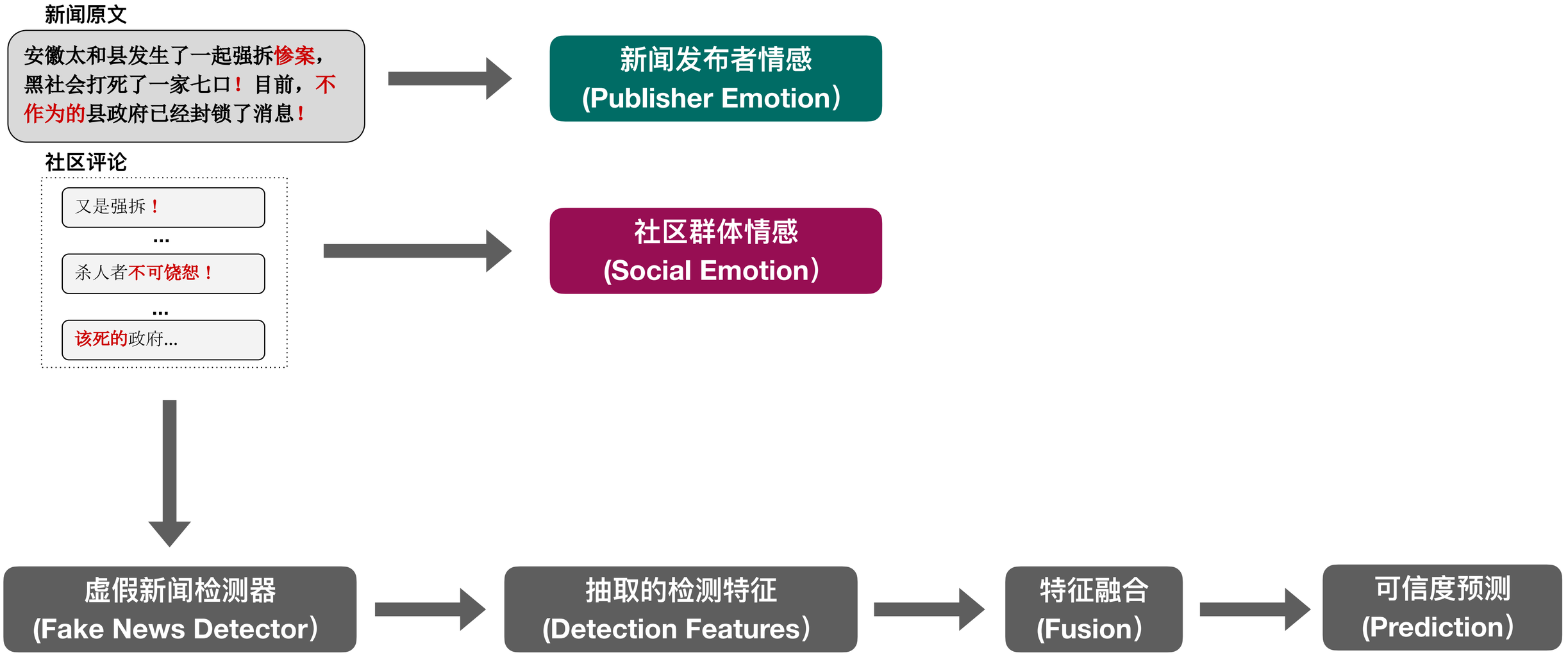
卡方检验结果：双重情感类别与新闻的可信度显著相关 ($p<0.01$)

[1] Chuan Guo, Juan Cao, **Xueyao Zhang**, et al. Exploiting Emotions for Fake News Detection on Social Media. arXiv:1903.01728, 2019.
[2] Soroush Vosoughi, et al. The Spread of True and False News Online. Science, 2018.

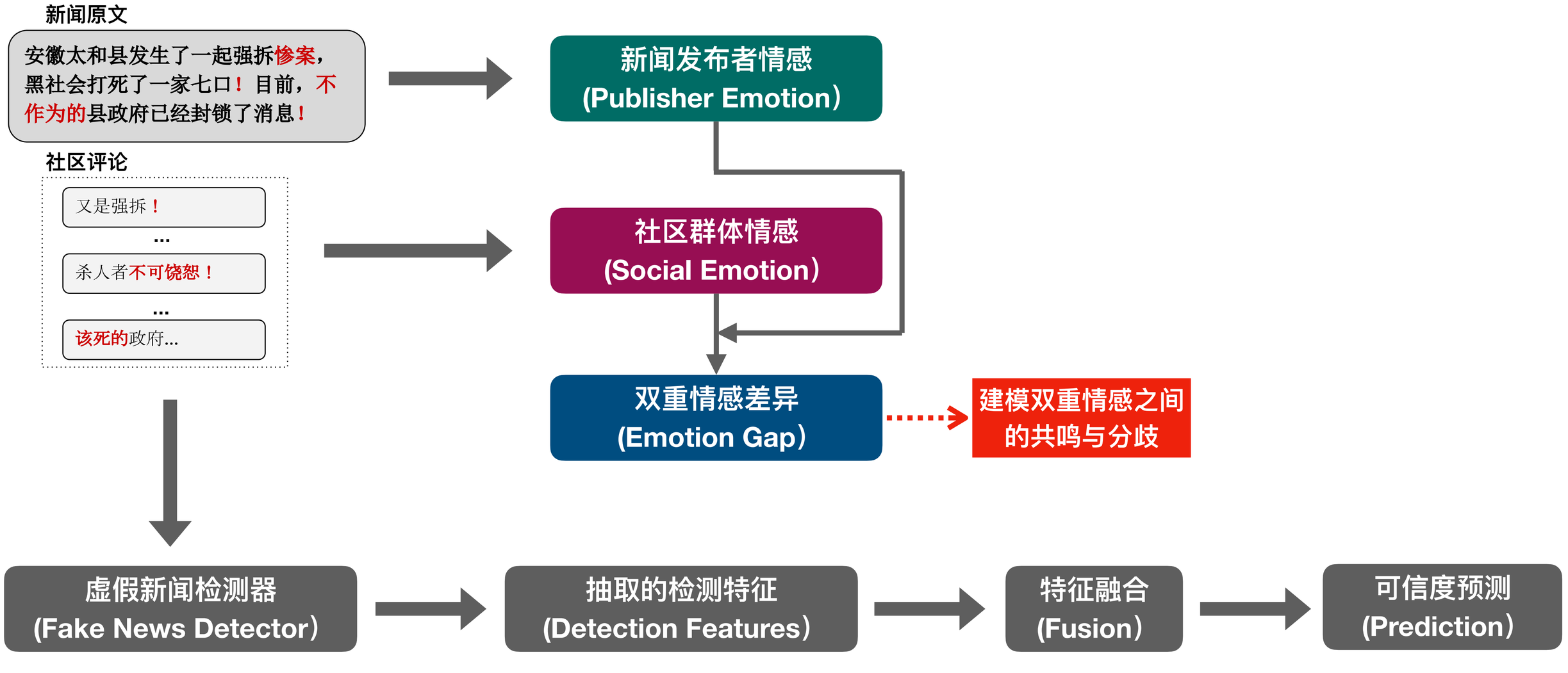
方法设计：基于双重情感特征的虚假新闻检测框架



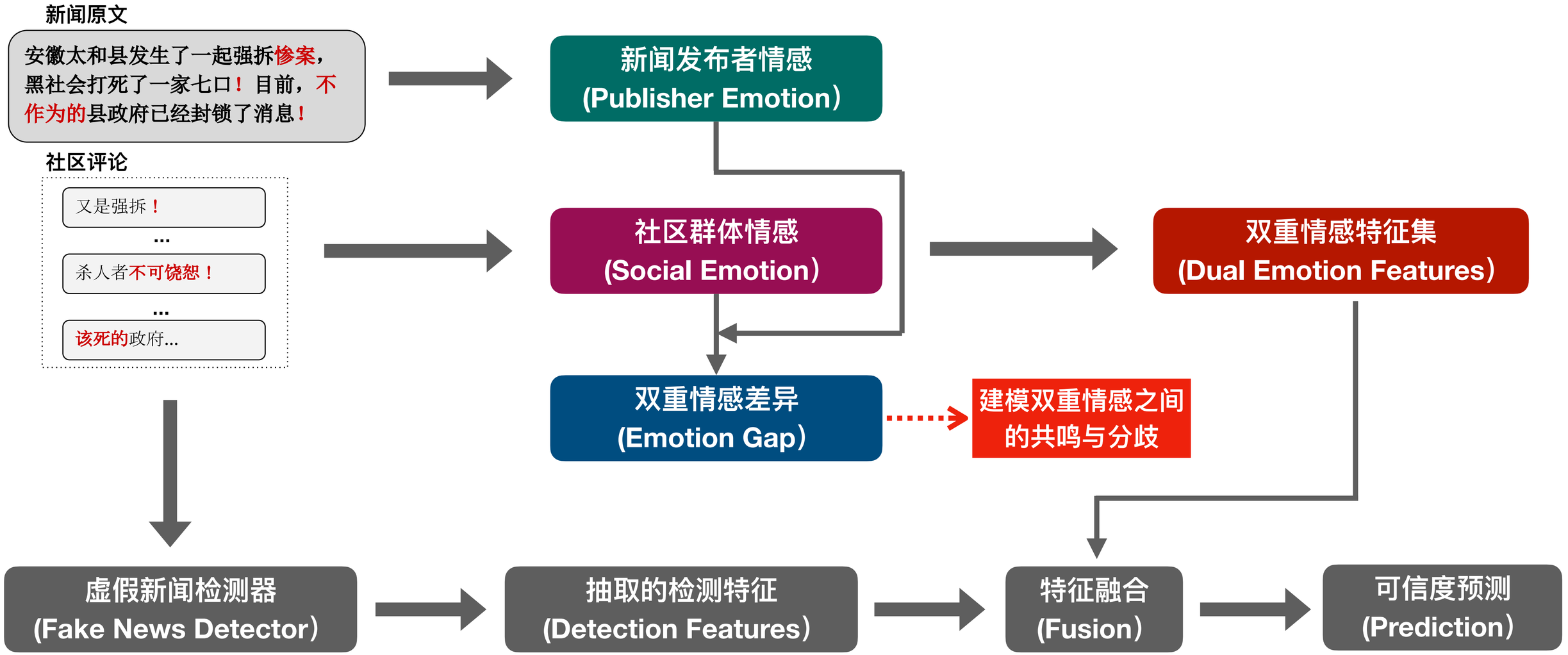
方法设计：基于双重情感特征的虚假新闻检测框架



方法设计：基于双重情感特征的虚假新闻检测框架



方法设计：基于双重情感特征的虚假新闻检测框架



难点一： 如何抽取新闻发布者情感?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

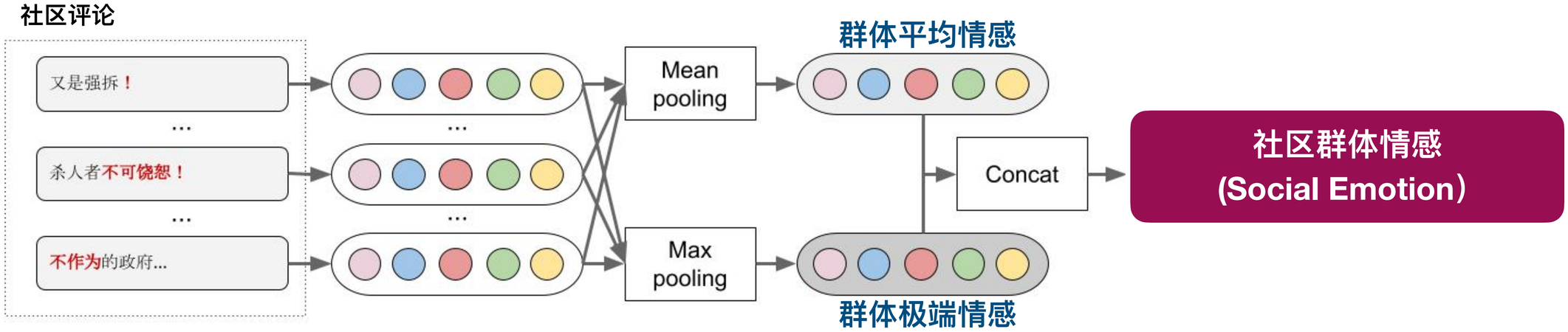
→ 新闻发布者情感 (Publisher Emotion)

解决方案：利用情感词典、预训练情感模型抽取新闻发布者的情感特征

特征类型	利用资源	情感信号的粒度	特征维度（中/英语料）
情感类别	预训练的情感模型 (Baidu AI)	句子级别	8 / 16
情感词	专家情感词典（大连理工情感词典、HowNet情感词典）	词级别	21 / 8
情感强度		词级别	21 / 8
情感极性		句子级别	3 / 6
辅助情感特征集 (包括表情符号、标点符号、人称代词等)	专家情感词典 (HowNet情感词典）、维基百科等	词级别、符号级别	13 / 14

总计：66 / 52

难点二：如何抽取社区群体情感?



解决方案 分别获取评论中的群体平均情感与群体极端情感

难点三：如何捕捉双重情感间的差异?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

社区评论

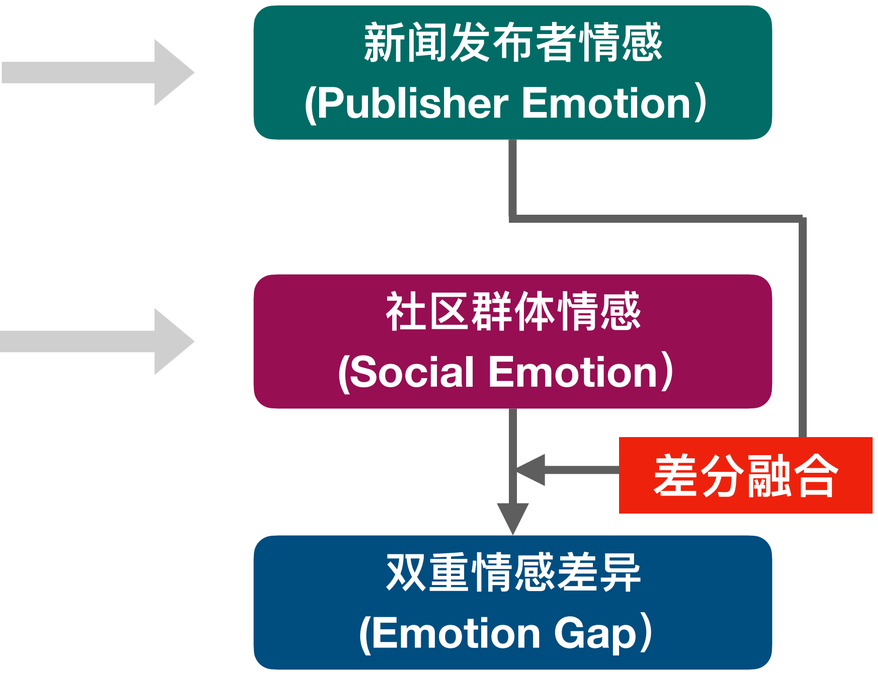
又是强拆！

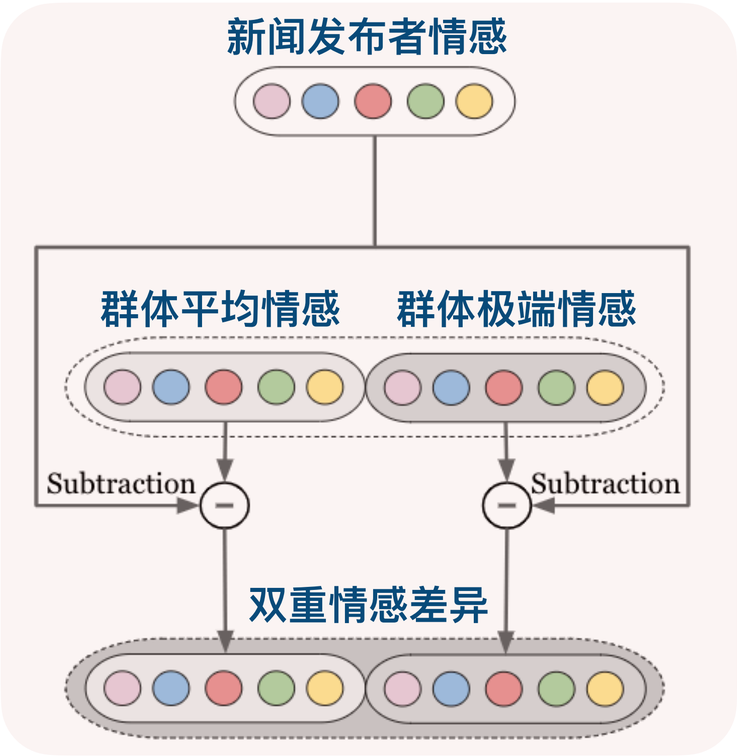
...

杀人者不可饶恕！

...

该死的政府...





解决方案 利用差分的方式捕捉双重情感之间的差异（共鸣/分歧）

实验评估设计

- **评估一：双重情感特征集自身的有效性**
  - 其是否优于现有的情感特征集?
  - 其构建过程中所使用的五个类型的情感特征是否均为有效?
- **评估二：基于双重情感特征的虚假新闻检测框架的有效性**
  - 双重情感特征集能否提升现有检测器的检测效果?
- **评估三：双重情感特征集中各部件的有效性**
  - 新闻发布者情感、社区群体情感与双重情感差异，是否均可以提升虚假新闻检测器的效果?
  - 哪个部件对于检测器的增强效果更好?

实验数据集

- **中文数据集**
  - Weibo-16 [1]
  - Weibo-20 **(本研究首次提出)**
    - 基于Weibo-16，增加了2014年4月-2018年11月的新闻数据
    - 假新闻来源：微博社区管理中心
    - 真新闻来源：AI识谣系统
- **英文数据集**
  - RumourEval-19 [2]

	新闻可信度	Weibo-16		Weibo-20		RumourEval-19	
		新闻数量	评论数量	新闻数量	评论数量	新闻数量	评论数量
训练集	假	1,386	789,841	1,896	749,141	79	1,135
	真	1,410	482,226	1,920	516,795	144	1,905
	待查证	-	-	-	-	104	1,838
	总计	2,796	1,272,067	3,816	1,265,936	327	4,878
验证集	假	463	255,833	632	137,941	19	824
	真	470	146,948	640	185,087	10	404
	待查证	-	-	-	-	9	212
	总计	933	402,781	1,272	323,028	38	1,440
测试集	假	463	224,795	633	245,216	40	689
	真	471	179,942	641	149,260	31	805
	待查证	-	-	-	-	10	181
	总计	934	404,737	1,274	394,476	81	1,675
总计	假	2,312	1,270,469	3,161	1,132,298	138	2,648
	真	2,351	809,116	3,201	851,142	185	3,114
	待查证	-	-	-	-	123	2,231
	总计	4,663	2,079,585	6,362	1,983,440	446	7,993

三个数据集的统计情况

[1] Jing Ma, et al. Detecting rumors from microblogs with recurrent neural networks. IJCAI 2016.
[2] Genevieve Gorrell, et al. SemEval-2019 Task 7: RumourEval, Determining Rumour Veracity and Support for Rumours. SemEval@NAACL-HLT 2019.

评估一：双重情感特征集自身的有效性

(评估指标为Macro F1检测值)

特征来源	情感特征	Weibo-16	Weibo-20	RumourEval-19
新闻原文	Emoratio (Ajao 等, 2019)	0.553	0.532	0.185
	EmoCred (Giachanou 等, 2019)	0.564	0.548	0.253
	新闻发布者情感	0.571	0.569	0.290
用户评论	**社区群体情感**	0.692	0.603	0.296
新闻原文、用户评论	**双重情感差异**	0.716	0.617	0.332
	双重情感特征集	**0.728**	**0.648**	**0.337**

在五层MLP模型上，仅使用情感特征作为输入

移除的情感信号	Weibo-16	Weibo-20	RumourEval-19
-	**0.728**	**0.648**	**0.337**
情感种类	0.679	0.576	0.193
情感词	0.715	0.635	0.239
情感强度	0.725	0.640	0.216
情感极性	0.723	0.633	0.245
辅助情感特征集	0.653	0.612	0.307

移除双重情感特征集中某个特定类型的情感信号

1. **建模情感信号方式的有效性**：(1) 新闻发布者情感优于 Emoratio [1] 与 EmoCred [2]; (2) 无论移除哪一种类型的情感信号，双重情感特征集的Macro F1值均会有所下降
2. **建模社区群体情感、双重情感的重要性**：双重情感差异 > 社区群体情感 > 新闻发布者情感
3. **双重情感特征集的有效性**：联合使用新闻发布者情感、社区群体情感以及双重情感差异后的检测性能最好

[1] Oluwaseun Ajao, et al. 2019. Sentiment Aware Fake News Detection on Online Social Networks. IEEE ICASSP 2019.
[2] Anastasia Giachanou, et al. 2019. Leveraging Emotional Signals for Credibility Detection. ACM SIGIR 2019.

评估一：双重情感特征集自身的有效性

(评估指标为Macro F1检测值)

特征来源	情感特征	Weibo-16	Weibo-20	RumourEval-19
新闻原文	Emoratio (Ajao 等, 2019)	0.553	0.532	0.185
	EmoCred (Giachanou 等, 2019)	0.564	0.548	0.253
	新闻发布者情感	0.571	0.569	0.290
用户评论	**社区群体情感**	0.692	0.603	0.296
新闻原文、用户评论	**双重情感差异**	0.716	0.617	0.332
	双重情感特征集	**0.728**	**0.648**	**0.337**

在五层MLP模型上，仅使用情感特征作为输入

移除的情感信号	Weibo-16	Weibo-20	RumourEval-19
-	**0.728**	**0.648**	**0.337**
情感种类	0.679	0.576	0.193
情感词	0.715	0.635	0.239
情感强度	0.725	0.640	0.216
情感极性	0.723	0.633	0.245
辅助情感特征集	0.653	0.612	0.307

移除双重情感特征集中某个特定类型的情感信号

1. **建模情感信号方式的有效性**：(1) 新闻发布者情感优于 Emoratio [1] 与 EmoCred [2]; (2) 无论移除哪一种类型的情感信号，双重情感特征集的Macro F1值均会有所下降
2. **建模社区群体情感、双重情感的重要性**：双重情感差异 > 社区群体情感 > 新闻发布者情感
3. **双重情感特征集的有效性**：联合使用新闻发布者情感、社区群体情感以及双重情感差异后的检测性能最好

[1] Oluwaseun Ajao, et al. 2019. Sentiment Aware Fake News Detection on Online Social Networks. IEEE ICASSP 2019.
[2] Anastasia Giachanou, et al. 2019. Leveraging Emotional Signals for Credibility Detection. ACM SIGIR 2019.

评估一：双重情感特征集自身的有效性

(评估指标为Macro F1检测值)

特征来源	情感特征	Weibo-16	Weibo-20	RumourEval-19
新闻原文	Emoratio (Ajao 等, 2019)	0.553	0.532	0.185
	EmoCred (Giachanou 等, 2019)	0.564	0.548	0.253
	新闻发布者情感	0.571	0.569	0.290
用户评论	**社区群体情感**	0.692	0.603	0.296
新闻原文、用户评论	**双重情感差异**	0.716	0.617	0.332
	双重情感特征集	**0.728**	**0.648**	**0.337**

在五层MLP模型上，仅使用情感特征作为输入

移除的情感信号	Weibo-16	Weibo-20	RumourEval-19
-	**0.728**	**0.648**	**0.337**
情感种类	0.679	0.576	0.193
情感词	0.715	0.635	0.239
情感强度	0.725	0.640	0.216
情感极性	0.723	0.633	0.245
辅助情感特征集	0.653	0.612	0.307

移除双重情感特征集中某个特定类型的情感信号

1. **建模情感信号方式的有效性**：(1) 新闻发布者情感优于 Emoratio [1] 与 EmoCred [2]; (2) 无论移除哪一种类型的情感信号，双重情感特征集的Macro F1值均会有所下降
2. **建模社区群体情感、双重情感的重要性**：双重情感差异 > 社区群体情感 > 新闻发布者情感
3. **双重情感特征集的有效性**：联合使用新闻发布者情感、社区群体情感以及双重情感差异后的检测性能最好

[1] Oluwaseun Ajao, et al. 2019. Sentiment Aware Fake News Detection on Online Social Networks. IEEE ICASSP 2019.
[2] Anastasia Giachanou, et al. 2019. Leveraging Emotional Signals for Credibility Detection. ACM SIGIR 2019.

评估一：双重情感特征集自身的有效性

(评估指标为Macro F1检测值)

特征来源	情感特征	Weibo-16	Weibo-20	RumourEval-19
新闻原文	Emoratio (Ajao 等, 2019)	0.553	0.532	0.185
	EmoCred (Giachanou 等, 2019)	0.564	0.548	0.253
	新闻发布者情感	0.571	0.569	0.290
用户评论	**社区群体情感**	0.692	0.603	0.296
新闻原文、用户评论	**双重情感差异**	0.716	0.617	0.332
	双重情感特征集	**0.728**	**0.648**	**0.337**

在五层MLP模型上，仅使用情感特征作为输入

移除的情感信号	Weibo-16	Weibo-20	RumourEval-19
-	**0.728**	**0.648**	**0.337**
情感种类	0.679	0.576	0.193
情感词	0.715	0.635	0.239
情感强度	0.725	0.640	0.216
情感极性	0.723	0.633	0.245
辅助情感特征集	0.653	0.612	0.307

移除双重情感特征集中某个特定类型的情感信号

1. **建模情感信号方式的有效性**：(1) 新闻发布者情感优于 Emoratio [1] 与 EmoCred [2]; (2) 无论移除哪一种类型的情感信号，双重情感特征集的Macro F1值均会有所下降
2. **建模社区群体情感、双重情感的重要性**：双重情感差异 > 社区群体情感 > 新闻发布者情感
3. **双重情感特征集的有效性**：联合使用新闻发布者情感、社区群体情感以及双重情感差异后的检测性能最好

[1] Oluwaseun Ajao, et al. 2019. Sentiment Aware Fake News Detection on Online Social Networks. IEEE ICASSP 2019.
[2] Anastasia Giachanou, et al. 2019. Leveraging Emotional Signals for Credibility Detection. ACM SIGIR 2019.

评估二：基于双重情感特征的虚假新闻检测框架的有效性

模型	Weibo-16				Weibo-20			
	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$
BiLSTM (Graves 等, 2005)	0.822	0.807	0.754	0.860	0.667	0.660	0.710	0.610
+ Emoratio (Ajao 等, 2019)	0.810	0.794	0.738	0.851	0.632	0.628	0.665	0.592
+ EmoCred (Giachanou 等, 2019)	0.778	0.766	0.711	0.820	0.666	0.659	0.709	0.609
+ 双重情感特征集	**0.838**	**0.826**	**0.781**	**0.871**	**0.702**	**0.701**	**0.714**	**0.689**
BERT (Devlin 等, 2019)	0.845	0.824	0.762	0.886	0.712	0.708	0.743	0.672
+ Emoratio (Ajao 等, 2019)	0.837	0.857	0.780	0.894	0.724	0.719	0.757	0.681
+ EmoCred (Giachanou 等, 2019)	0.867	0.849	0.797	**0.901**	0.728	0.725	0.752	**0.699**
+ 双重情感特征集	**0.873**	**0.867**	**0.837**	0.896	**0.734**	**0.734**	**0.773**	0.692
HSA-BLSTM (Guo 等, 2018)	0.855	0.849	0.819	0.879	0.778	0.776	0.796	0.686
+ Emoratio (Ajao 等, 2019)	0.872	0.863	0.829	0.898	0.774	0.771	0.796	0.663
+ EmoCred (Giachanou 等, 2019)	0.861	0.854	0.822	0.886	0.781	0.777	0.806	0.646
+ 双重情感特征集	**0.913**	**0.908**	**0.885**	**0.930**	**0.808**	**0.805**	**0.827**	**0.694**

模型	Macro F1 值	RMSE(↓)	$F1_{fake}$	$F1_{real}$	$F1_{unv}$
BiLSTM (Graves 等, 2005)	0.269	0.804	0.500	0.222	0.083
+ Emoratio (Ajao 等, 2019)	0.275	0.823	0.463	0.160	**0.200**
+ EmoCred (Giachanou 等, 2019)	0.311	0.797	0.456	0.295	0.182
+ 双重情感特征集	**0.340**	**0.752**	**0.580**	**0.337**	0.104
BERT (Devlin 等, 2019)	0.272	0.808	0.533	0.105	0.176
+ Emoratio (Ajao 等, 2019)	0.271	0.857	0.406	0.240	0.167
+ EmoCred (Giachanou 等, 2019)	0.308	0.833	0.367	**0.367**	0.189
+ 双重情感特征集	**0.346**	**0.778**	**0.557**	0.244	**0.238**
NileTMRG (Enayet 等, 2017)	0.309	0.770	0.557	0.245	0.125
+ Emoratio (Ajao 等, 2019)	0.331	**0.754**	**0.571**	0.280	**0.143**
+ EmoCred (Giachanou 等, 2019)	0.307	0.786	0.296	0.500	0.125
+ 双重情感特征集	**0.342**	**0.754**	0.565	**0.565**	0.100

中文数据集
Weibo-16、Weibo-20

英文数据集
RumourEval-19

1. **检测框架的有效性：**在三个实验数据集上，双重情感特征集的引入均能够显著提升各类模型的检测指标
2. **双重情感特征集的兼容性与鲁棒性：**引入Emoratio或EmoCred时更容易使模型过拟合，而双重情感特征集则拥有更好的泛化性，这证明了不能只关注新闻原文中的情感，更要关注用户评论的情感以及这二者之间的关系

评估二：基于双重情感特征的虚假新闻检测框架的有效性

模型	Weibo-16				Weibo-20			
	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$
BiLSTM (Graves 等, 2005)	0.822	0.807	0.754	0.860	0.667	0.660	0.710	0.610
+ Emoratio (Ajao 等, 2019)	0.810	0.794	0.738	0.851	0.632	0.628	0.665	0.592
+ EmoCred (Giachanou 等, 2019)	0.778	0.766	0.711	0.820	0.666	0.659	0.709	0.609
+ 双重情感特征集	**0.838**	**0.826**	**0.781**	**0.871**	**0.702**	**0.701**	**0.714**	**0.689**
BERT (Devlin 等, 2019)	0.845	0.824	0.762	0.886	0.712	0.708	0.743	0.672
+ Emoratio (Ajao 等, 2019)	0.837	0.857	0.780	0.894	0.724	0.719	0.757	0.681
+ EmoCred (Giachanou 等, 2019)	0.867	0.849	0.797	**0.901**	0.728	0.725	0.752	**0.699**
+ 双重情感特征集	**0.873**	**0.867**	**0.837**	0.896	**0.734**	**0.734**	**0.773**	0.692
HSA-BLSTM (Guo 等, 2018)	0.855	0.849	0.819	0.879	0.778	0.776	0.796	0.686
+ Emoratio (Ajao 等, 2019)	0.872	0.863	0.829	0.898	0.774	0.771	0.796	0.663
+ EmoCred (Giachanou 等, 2019)	0.861	0.854	0.822	0.886	0.781	0.777	0.806	0.646
+ 双重情感特征集	**0.913**	**0.908**	**0.885**	**0.930**	**0.808**	**0.805**	**0.827**	**0.694**

中文数据集
Weibo-16、Weibo-20

模型	Macro F1 值	RMSE(↓)	$F1_{fake}$	$F1_{real}$	$F1_{unv}$
BiLSTM (Graves 等, 2005)	0.269	0.804	0.500	0.222	0.083
+ Emoratio (Ajao 等, 2019)	0.275	0.823	0.463	0.160	**0.200**
+ EmoCred (Giachanou 等, 2019)	0.311	0.797	0.456	0.295	0.182
+ 双重情感特征集	**0.340**	**0.752**	**0.580**	**0.337**	0.104
BERT (Devlin 等, 2019)	0.272	0.808	0.533	0.105	0.176
+ Emoratio (Ajao 等, 2019)	0.271	0.857	0.406	0.240	0.167
+ EmoCred (Giachanou 等, 2019)	0.308	0.833	0.367	**0.367**	0.189
+ 双重情感特征集	**0.346**	**0.778**	**0.557**	0.244	**0.238**
NileTMRG (Enayet 等, 2017)	0.309	0.770	0.557	0.245	0.125
+ Emoratio (Ajao 等, 2019)	0.331	**0.754**	**0.571**	0.280	**0.143**
+ EmoCred (Giachanou 等, 2019)	0.307	0.786	0.296	0.500	0.125
+ 双重情感特征集	**0.342**	**0.754**	0.565	**0.565**	0.100

英文数据集
RumourEval-19

1. **检测框架的有效性：**在三个实验数据集上，双重情感特征集的引入均能够显著提升各类模型的检测指标
2. **双重情感特征集的兼容性与鲁棒性：**引入Emoratio或EmoCred时更容易使模型过拟合，而双重情感特征集则拥有更好的泛化性，这证明了不能只关注新闻原文中的情感，更要关注用户评论的情感以及这二者之间的关系

评估三： 双重情感特征集中各部件的有效性

模型	Weibo-16	Weibo-20	RumourEval-19
BiLSTM (Graves 等, 2005)	0.807	0.660	0.269
+ 新闻发布者情感	0.809	0.681	0.310
+ 社区群体情感	0.818	0.693	0.322
+ 双重情感差异	0.811	0.693	0.336
+ 新闻发布者情感 + 社区群体情感	0.820	0.697	0.325
+ 双重情感特征集	**0.826**	**0.701**	**0.340**
BERT (Devlin 等, 2019)	0.824	0.708	0.272
+ 新闻发布者情感	0.850	0.722	0.312
+ 社区群体情感	0.856	0.730	0.339
+ 双重情感差异	0.858	0.731	0.338
+ 新闻发布者情感 + 社区群体情感	0.859	0.733	0.341
+ 双重情感特征集	**0.867**	**0.734**	**0.346**
HSA-BLSTM (Guo 等, 2018)	0.849	0.776	-
+ 新闻发布者情感	0.876	0.779	-
+ 社区群体情感	0.892	0.792	-
+ 双重情感差异	0.901	0.800	-
+ 新闻发布者情感 + 社区群体情感	0.896	0.799	-
+ 双重情感特征集	**0.908**	**0.805**	-
NileTMRG (Enayet 等, 2017)	-	-	0.309
+ 新闻发布者情感	-	-	0.311
+ 社区群体情感	-	-	0.325
+ 双重情感差异	-	-	0.337
+ 新闻发布者情感 + 社区群体情感	-	-	0.330
+ 双重情感特征集	-	-	**0.342**

评估指标为Macro F1值

1. **各部件的有效性**：（1）任意一个部件均有效；（2）比起利用MLP隐式地学习“新闻发布者情感”与“社区群体情感”之间的联系，显式地引入“双重情感情感差异”更为有效
2. **建模社区群体情感、双重情感的重要性**：“双重情感差异”与“社区群体情感”的检测效果更好
3. **双重情感特征集的必要性**：同时使用三个部件对现有模型的提升最大

评估三： 双重情感特征集中各部件的有效性

模型	Weibo-16	Weibo-20	RumourEval-19
BiLSTM (Graves 等, 2005)	0.807	0.660	0.269
+ 新闻发布者情感	0.809	0.681	0.310
+ 社区群体情感	0.818	0.693	0.322
+ 双重情感差异	0.811	0.693	0.336
+ 新闻发布者情感 + 社区群体情感	0.820	0.697	0.325
+ 双重情感特征集	**0.826**	**0.701**	**0.340**
BERT (Devlin 等, 2019)	0.824	0.708	0.272
+ 新闻发布者情感	0.850	0.722	0.312
+ 社区群体情感	0.856	0.730	0.339
+ 双重情感差异	0.858	0.731	0.338
+ 新闻发布者情感 + 社区群体情感	0.859	0.733	0.341
+ 双重情感特征集	**0.867**	**0.734**	**0.346**
HSA-BLSTM (Guo 等, 2018)	0.849	0.776	-
+ 新闻发布者情感	0.876	0.779	-
+ 社区群体情感	0.892	0.792	-
+ 双重情感差异	0.901	0.800	-
+ 新闻发布者情感 + 社区群体情感	0.896	0.799	-
+ 双重情感特征集	**0.908**	**0.805**	-
NileTMRG (Enayet 等, 2017)	-	-	0.309
+ 新闻发布者情感	-	-	0.311
+ 社区群体情感	-	-	0.325
+ 双重情感差异	-	-	0.337
+ 新闻发布者情感 + 社区群体情感	-	-	0.330
+ 双重情感特征集	-	-	**0.342**

评估指标为Macro F1值

1. **各部件的有效性**：（1）任意一个部件均有效；（2）比起利用MLP隐式地学习“新闻发布者情感”与“社区群体情感”之间的联系，显式地引入“双重情感情感差异”更为有效
2. **建模社区群体情感、双重情感的重要性**：“双重情感差异”与“社区群体情感”的检测效果更好
3. **双重情感特征集的必要性**：同时使用三个部件对现有模型的提升最大

案例分析：双重情感特征集对现有模型错误的纠正

新闻原文 愤怒、质疑

这是著名的云南怒江傈僳族女孩飞索渡江求学！修一座桥40万，当地县政府却说没钱！！记者高天俊质问：为何书记的奥迪车70多万？

社区评论 愤怒、厌恶

一边是庄严，一边是无耻。

...

钱都用书记身上了嘛😒😒

...

快把你的奥迪卖掉吧！

新闻原文 惊讶、开心

印度一妇女一次产下11胞胎，创世界记录，一支足球队就此诞生！我和我的小伙伴们都惊呆了，这才是天使😄 妈妈真的好伟大👍

社区评论 惊讶

😱小伙伴们这次彻底惊呆了

...

妈呀，一群天使降临呐！

...

我的天！伟大的妈

新闻原文 无明显情感

【顶花黄瓜或有猫腻，生长剂含避孕药】据青岛新闻网，顶着鲜艳黄花的黄瓜是用植物生长调节剂，甚至是避孕药“扮嫩”的。

社区评论 愤怒、厌恶

还让不让人活了！！？？😡

...

满身的吐槽点

...

...别让农民太有文化

各模型判定结果

模型	假	真
BiLSTM	0.39	**0.61**
+ Emoratio	0.35	**0.65**
+ EmoCred	0.43	**0.57**
+ 双重情感特征集	**0.77**	0.23

模型	假	真
BiLSTM	0.36	**0.64**
+ Emoratio	0.40	**0.60**
+ EmoCred	0.37	**0.63**
+ 双重情感特征集	**0.72**	0.28

模型	假	真
BiLSTM	0.48	**0.52**
+ Emoratio	0.41	**0.59**
+ EmoCred	0.35	**0.65**
+ 双重情感特征集	**0.73**	0.27

极度不确定

启示：双重情感为何有效?

1. **对情感信号的建模：专家先验知识**（情感词典） + **大规模预训练模型中蕴含的知识**
2. **双重情感的差分融合：**极大贴合了**虚假新闻中双重情感的呈现模式**（拥有独特的共鸣与分歧）

新闻原文

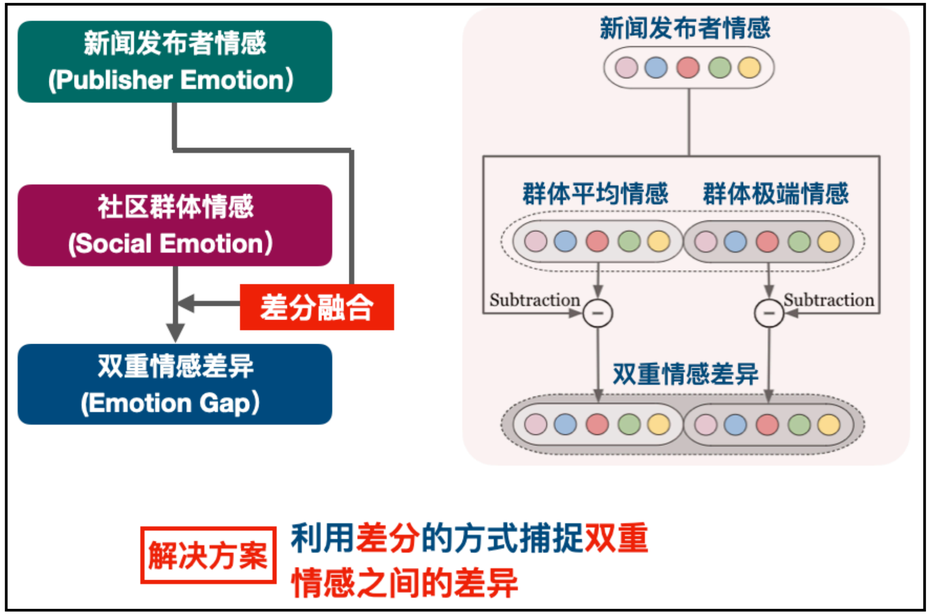
安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

→ 新闻发布者情感 (Publisher Emotion)

解决方案 利用情感词典、预训练情感模型抽取新闻发布者的情感特征

特征类型	利用资源	情感信号的粒度	特征维度（中/英语料）
情感类别	预训练的情感模型 (Baidu AI)	句子级别	8 / 16
情感词	专家情感词典（大连理工情感词典、HowNet情感词典）	词级别	21 / 8
情感强度		词级别	21 / 8
情感极性		句子级别	3 / 6
其他辅助特征 (包括表情符号、标点符号、人称代词等)	专家情感词典 (HowNet情感词典）、维基百科等	词级别、符号级别	13 / 14

总计：66 / 52



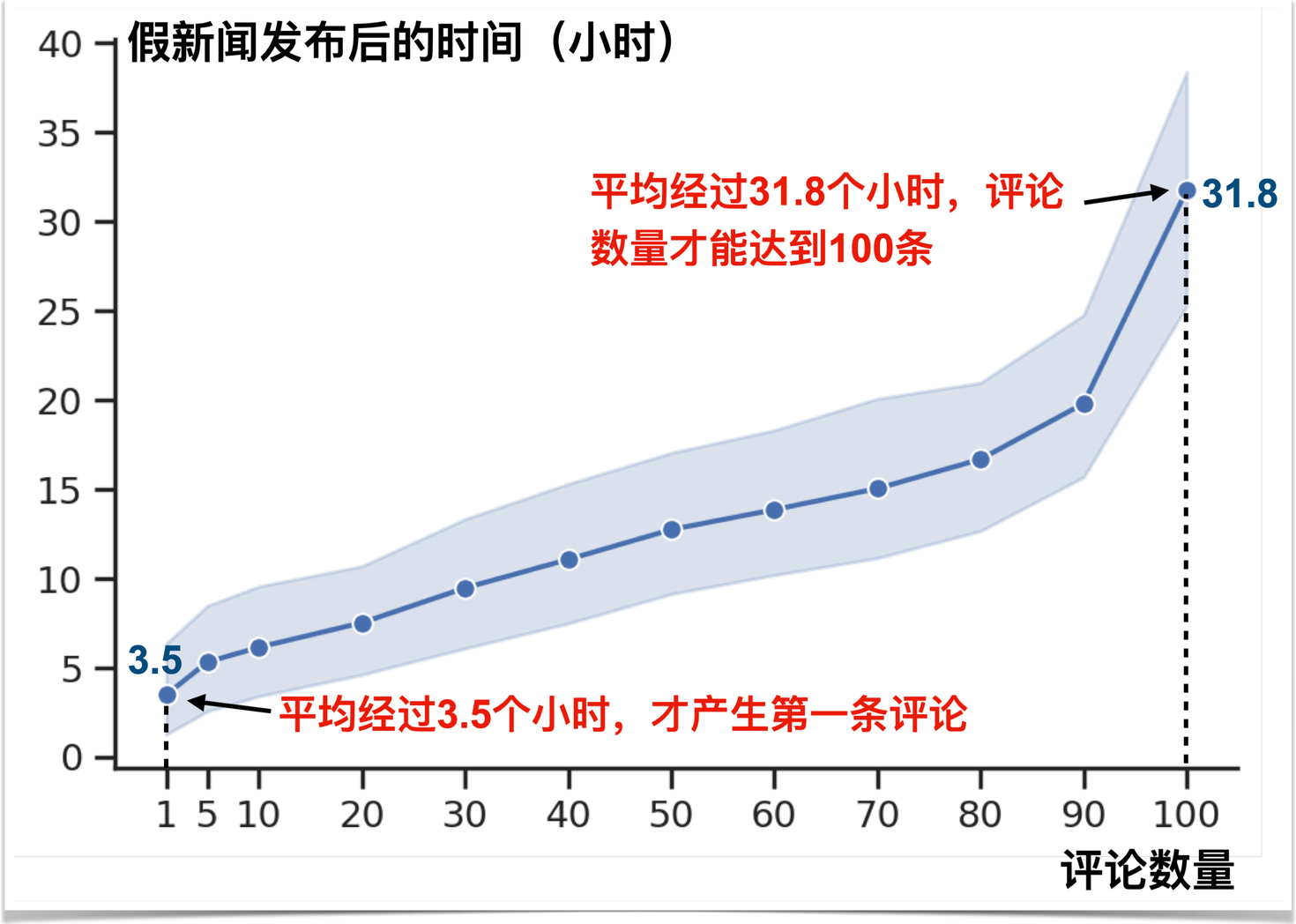
解决方案 利用差分的方式捕捉双重情感之间的差异

小结：基于双重情感的虚假新闻检测

- **贡献总结**
  - **问题贡献：**提出并确认了**双重情感信号在真、假新闻之间具有显著区别**
  - **方法贡献：**双重情感特征集的**性能增益显著、泛化性好、兼容性强**
  - **数据集贡献：****新的中文数据集** (Weibo-20).
- **研究成果**
  - **论文：****Xueyao Zhang**, et al. Mining Dual Emotion for Fake News Detection. WWW 2021.
  - **专利：**《一种基于双重情感的舆情检测方法及系统》. 曹娟；**张雪遥**；盛强；谢添；李锦涛.

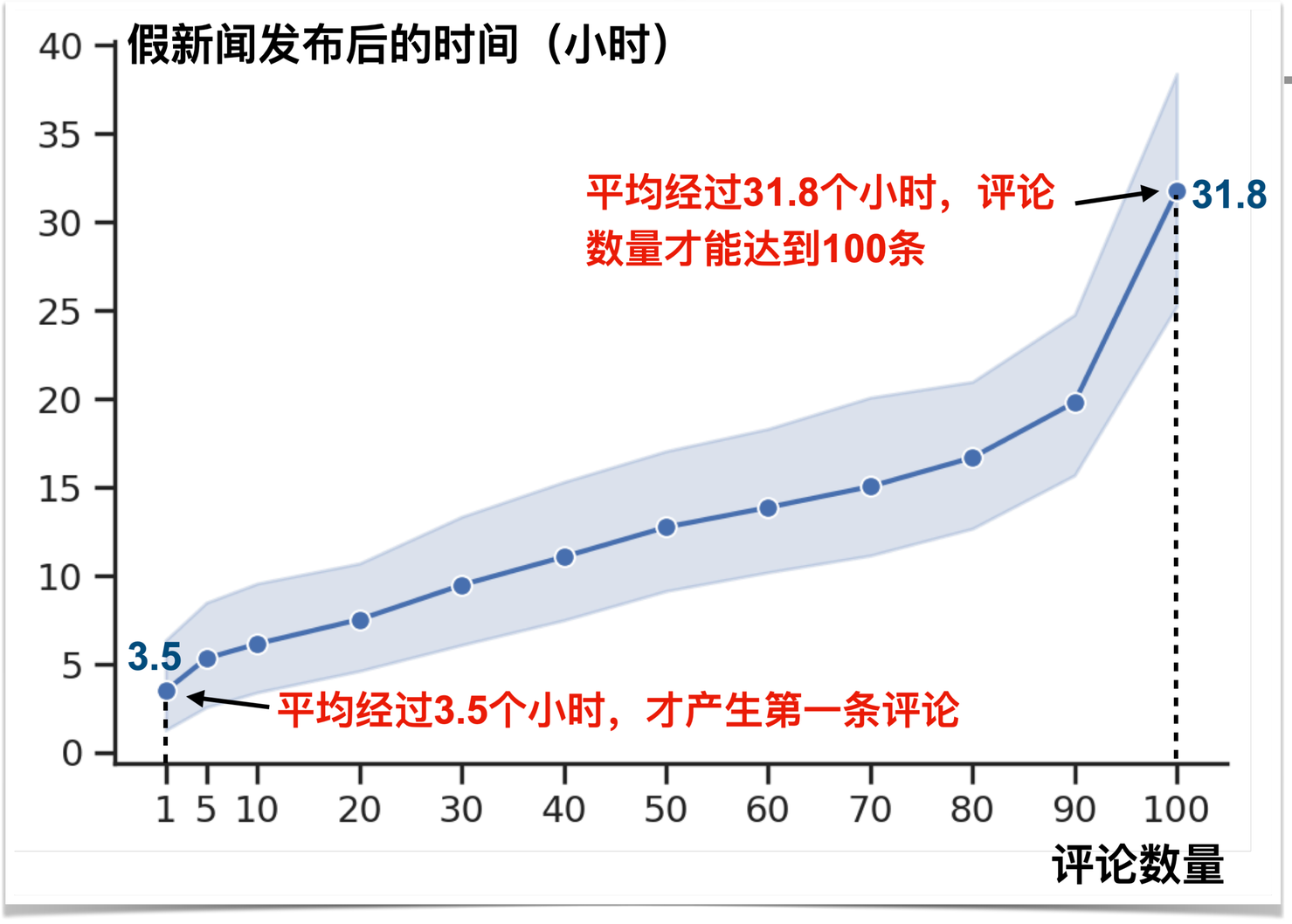
目录

研究问题：从延时检测到即时检测



假新闻自发布后的评论数量变化趋势（Weibo-16 数据集）

研究问题：从延时检测到即时检测

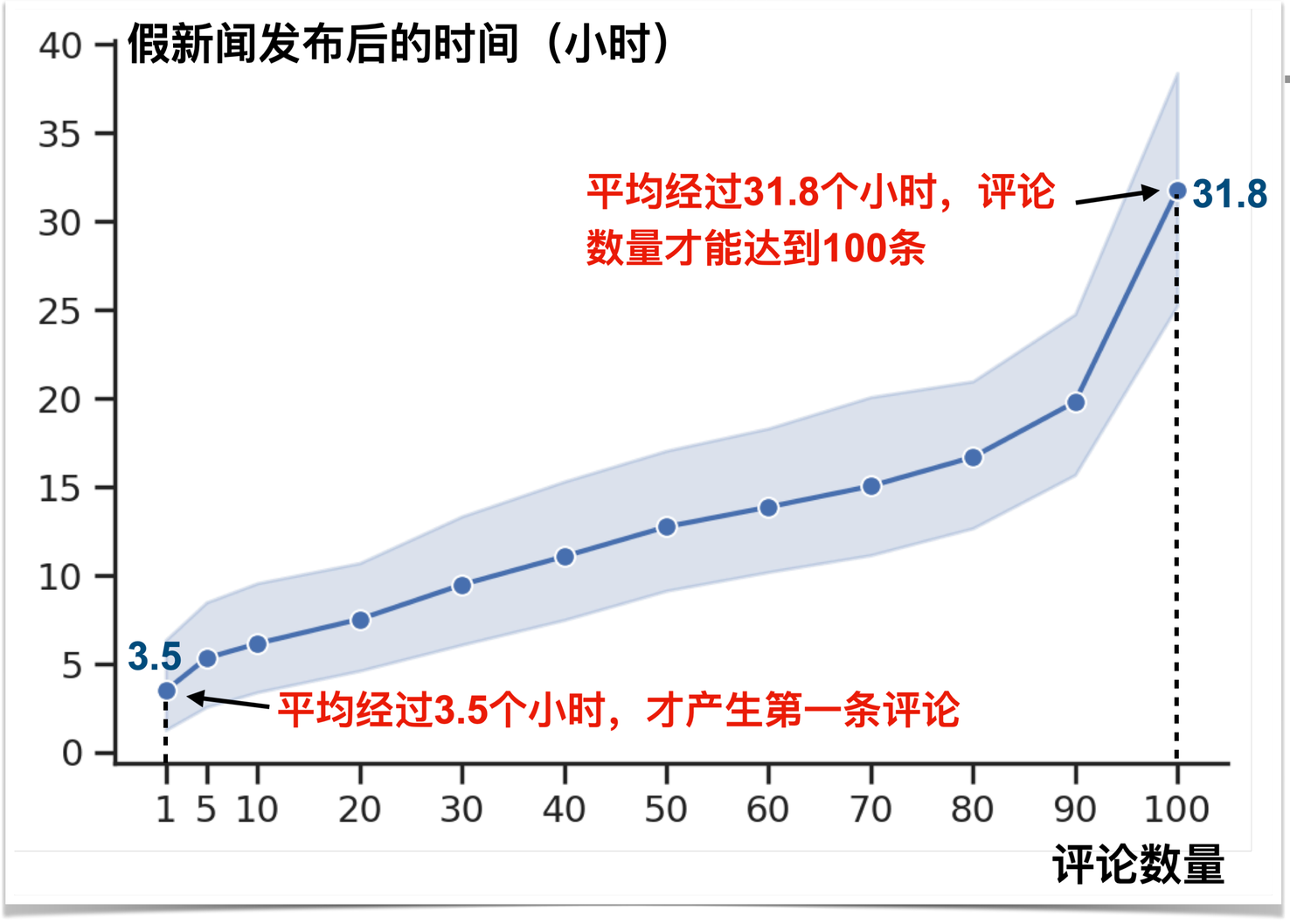


假新闻自发布后的评论数量变化趋势（Weibo-16 数据集）

结论

依靠新闻评论的**延时**检测，无法满足虚假新闻防治的**即时性**需求

研究问题：从延时检测到即时检测



假新闻自发布后的评论数量变化趋势（Weibo-16 数据集）

结论

依靠新闻评论的**延时**检测，无法满足虚假新闻防治的**即时性**需求

即时检测

定义 新闻**一经发布**就进行的检测

难点
1. 如何更好地表征**新闻内容** [1]
2. 如何融入**外部知识** [2]

[1] Xinyi Zhou, et al. Fake news early detection: A theory-driven model. Digital Threats, 2020.
[2] Qiang Sheng, et al. Zoom out and observe: News environment perception for fake news detection. ACL, 2022.

如何在即时检测的场景中利用情感信息?

新闻原文

惊异、愤怒

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

如何在即时检测的场景中利用情感信息?

新闻原文

惊异、愤怒

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

已知方案 **抽取新闻发布者情感，作为辅助特征**

模型	Weibo-16
BiLSTM	0.807
+ 新闻发布者情感	0.809
BERT	0.824
+ 新闻发布者情感	0.850

Weibo-16数据集上的Macro F1值

如何在即时检测的场景中利用情感信息?

新闻原文

惊异、愤怒

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

为何情感信号为不同模型带来的增益差别如此之大?

已知方案 抽取新闻发布者情感，作为辅助特征

模型	Weibo-16	
BiLSTM	0.807	0.2% ↑
+ 新闻发布者情感	0.809	
BERT	0.824	2.6% ↑
+ 新闻发布者情感	0.850	

Weibo-16数据集上的Macro F1值

如何在即时检测的场景中利用情感信息?

新闻原文

惊异、愤怒

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

为何情感信号为不同模型带来的增益差别如此之大?

模型自身已经拥有一定的情感表征能力

不同模型对于情感的表征能力不同

已知方案 抽取新闻发布者情感，作为辅助特征

模型	Weibo-16	
BiLSTM	0.807	0.2% ↑
+ 新闻发布者情感	0.809	
BERT	0.824	2.6% ↑
+ 新闻发布者情感	0.850	

Weibo-16数据集上的Macro F1值

如何在即时检测的场景中利用情感信息?

新闻原文

惊异、愤怒

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

已知方案 抽取新闻发布者情感，作为辅助特征

模型	Weibo-16	
BiLSTM	0.807	0.2% ↑
+ 新闻发布者情感	0.809	
BERT	0.824	2.6% ↑
+ 新闻发布者情感	0.850	

Weibo-16数据集上的Macro F1值

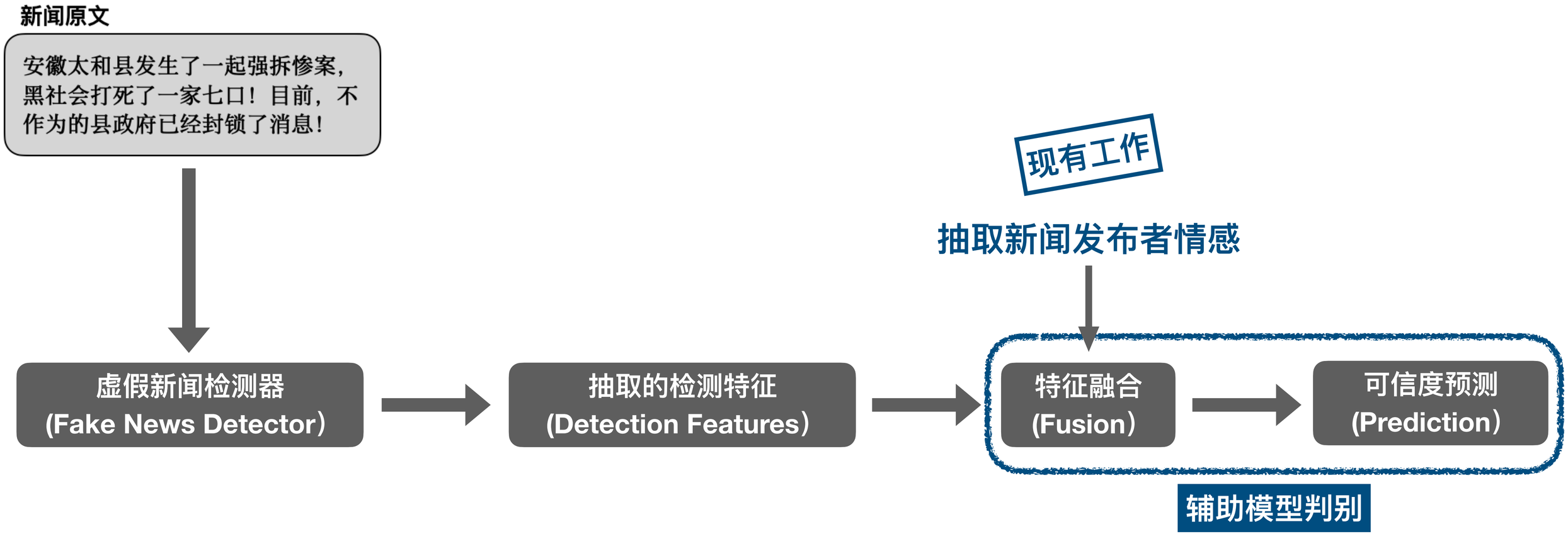
为何情感信号为不同模型带来的增益差别如此之大?

模型自身已经拥有一定的情感表征能力

不同模型对于情感的表征能力不同

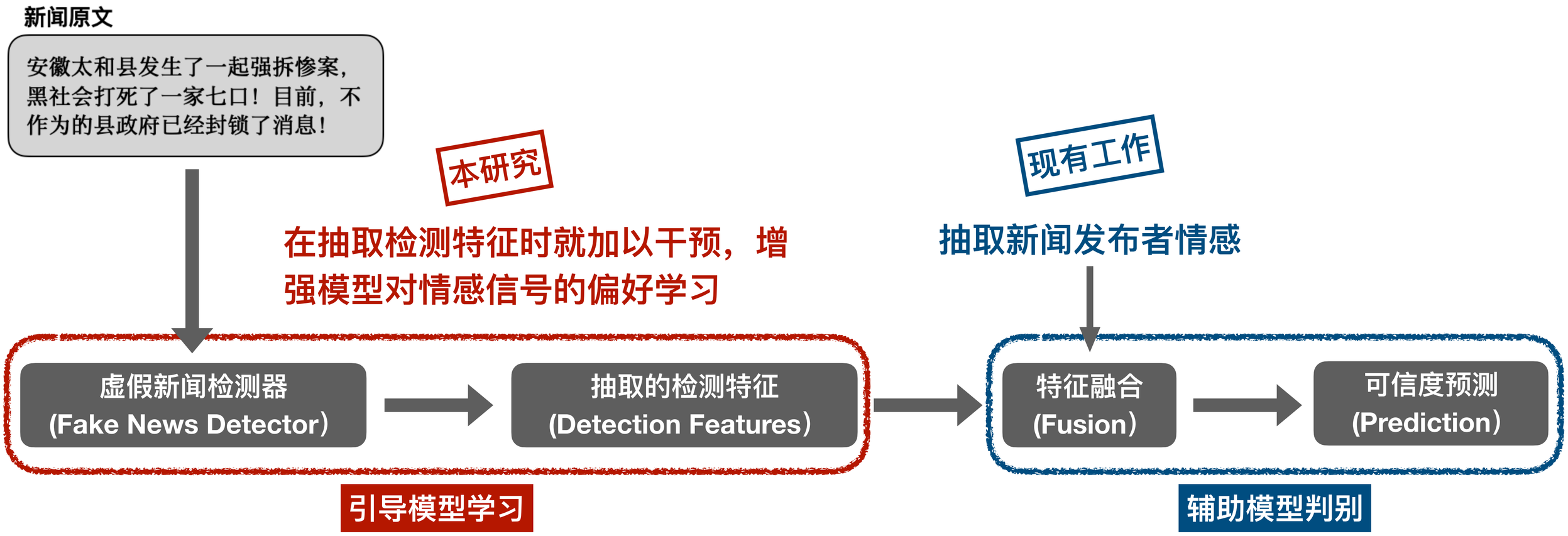
模型自身建模情感的潜能仍可挖掘

研究点二：情感偏好增强的虚假新闻即时检测



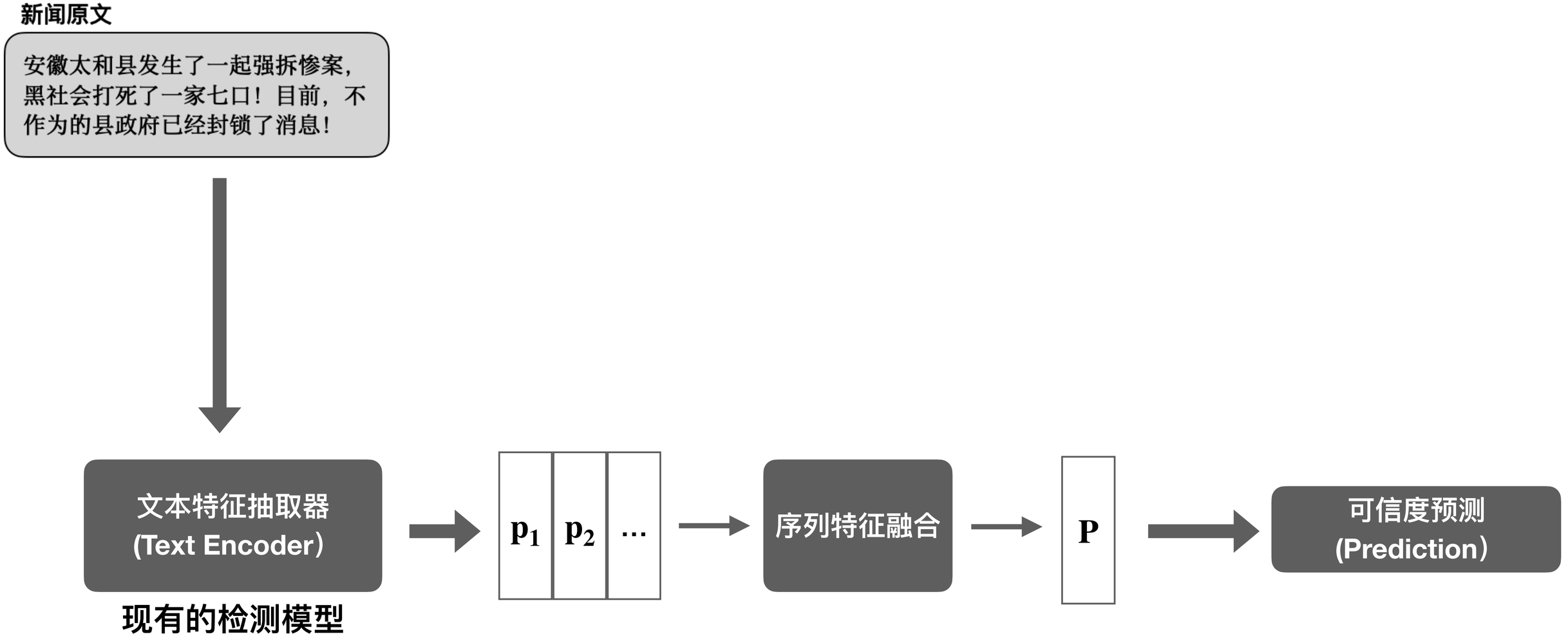
研究动机：增强模型对情感信号的建模，将会提升其检测假新闻的能力

研究点二：情感偏好增强的虚假新闻即时检测

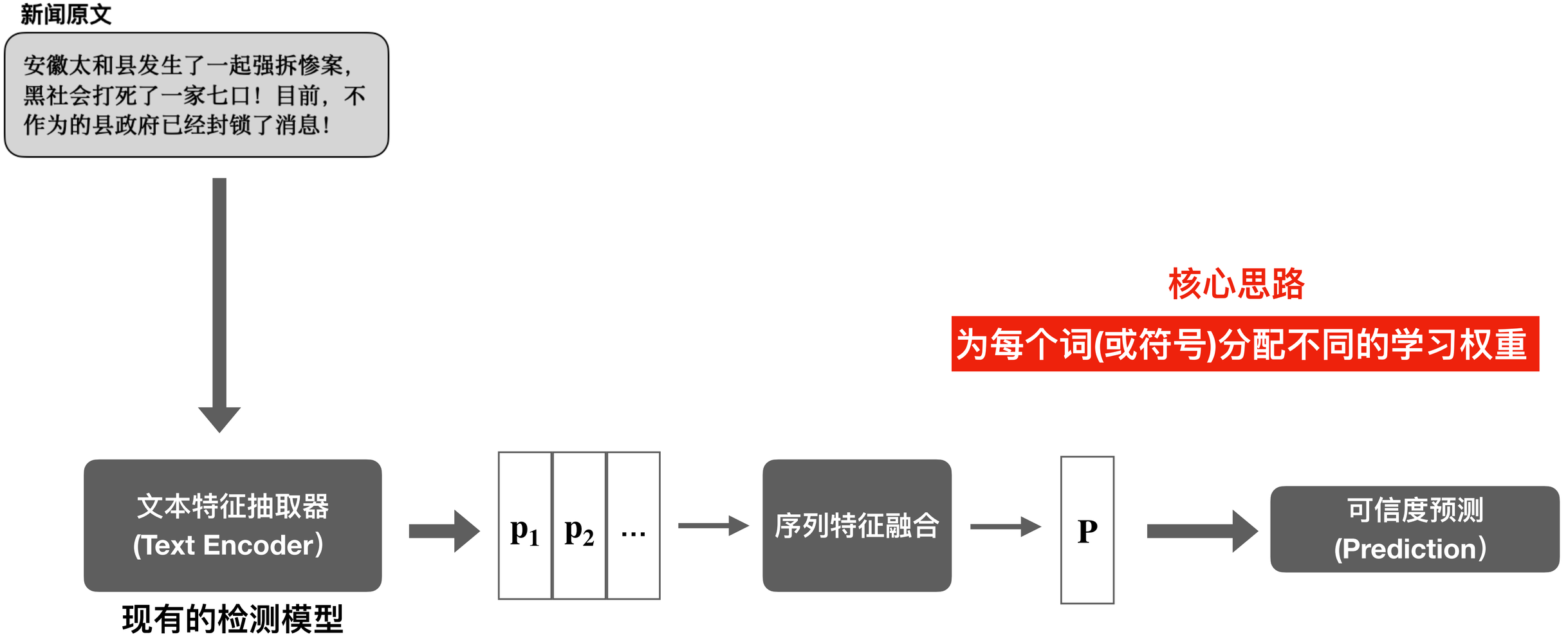


研究动机：增强模型对情感信号的建模，将会提升其检测假新闻的能力

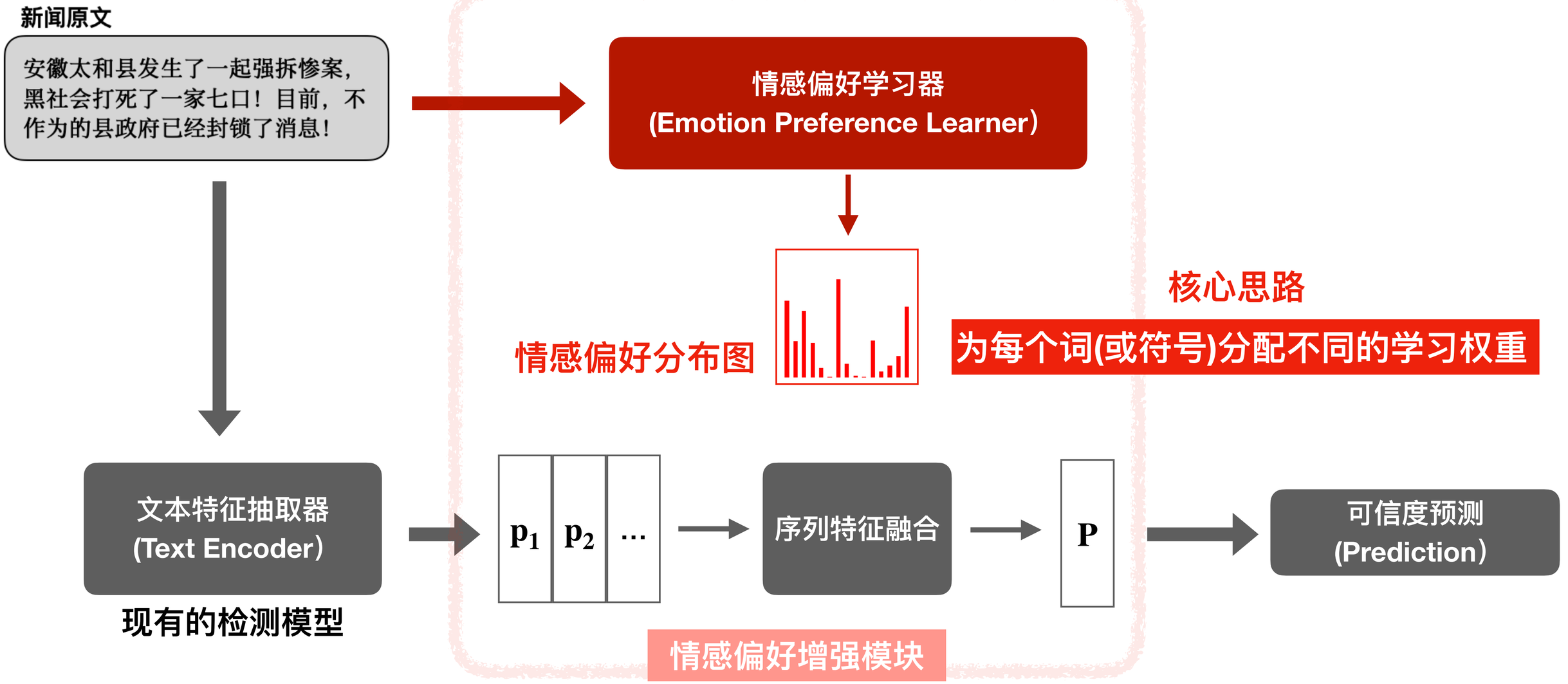
方法设计：情感偏好增强的虚假新闻检测框架 (EmoPref)



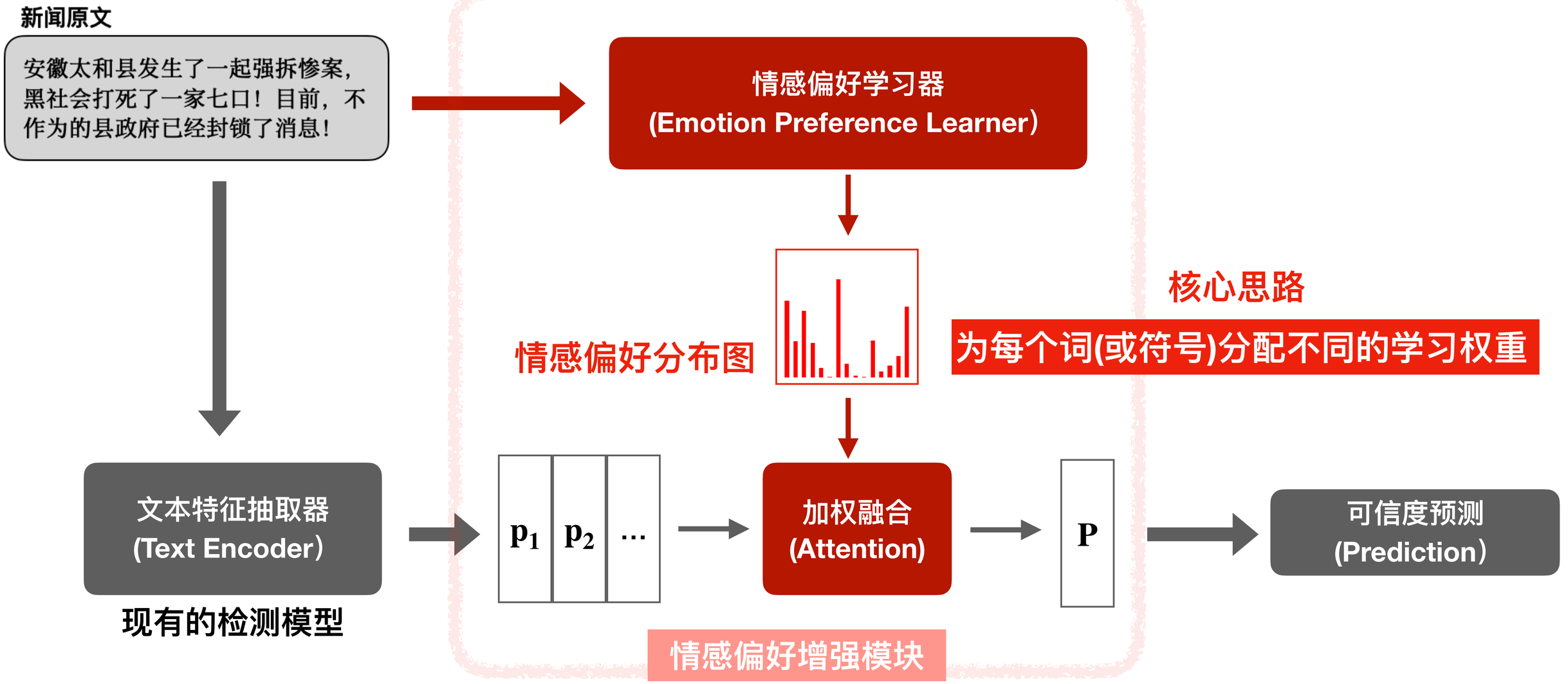
方法设计：情感偏好增强的虚假新闻检测框架 (EmoPref)



方法设计：情感偏好增强的虚假新闻检测框架 (EmoPref)



方法设计：情感偏好增强的虚假新闻检测框架 (EmoPref)

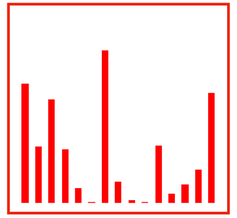


核心问题：如何设计情感偏好学习器?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

情感偏好学习器



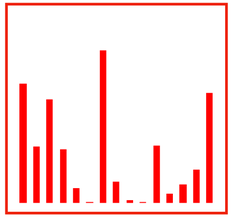
情感偏好分布图

核心问题：如何设计情感偏好学习器?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

情感偏好学习器



情感偏好分布图

偏好！

情感词

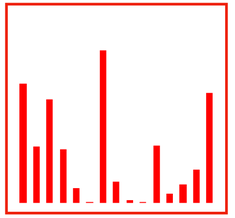
惨案 ！ 不作为的

核心问题：如何设计情感偏好学习器?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

情感偏好学习器



情感偏好分布图

偏好！

远离！

情感词

惨案 ！ 不作为的

实体词

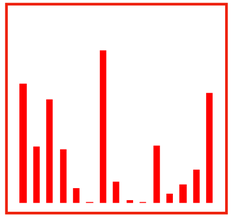
安徽太和县 县政府 黑社会

核心问题：如何设计情感偏好学习器?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

情感偏好学习器



情感偏好分布图

偏好！

远离！

如何利用?

情感词

惨案 ！ 不作为的

实体词

安徽太和县 县政府 黑社会

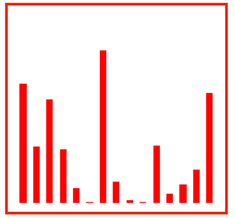
发生 一起 强拆 … 封锁 消息

核心问题：如何设计情感偏好学习器?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

情感偏好学习器



情感偏好分布图

偏好！

远离！

如何利用?

情感词

惨案 ！ 不作为的

实体词

安徽太和县 县政府 黑社会

发生 一起 强拆 … 封锁 消息

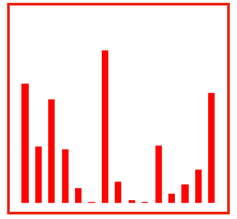
每个词都可能蕴含着情感信号

核心问题：如何设计情感偏好学习器?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

情感偏好学习器



情感偏好分布图

偏好！

如何利用?

远离！

情感词

惨案 ！ 不作为的

实体词

安徽太和县 县政府 黑社会

发生 一起 强拆 … 封锁 消息

每个词都可能蕴含着情感信号

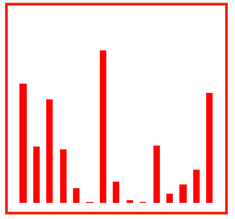
需要建模词与词之间的关联性

核心问题：如何设计情感偏好学习器?

新闻原文

安徽太和县发生了一起强拆惨案，黑社会打死了一家七口！目前，不作为的县政府已经封锁了消息！

情感偏好学习器



情感偏好分布图

偏好！

如何利用?

远离！

情感词

惨案 ！ 不作为的

发生 一起 强拆 … 封锁 消息

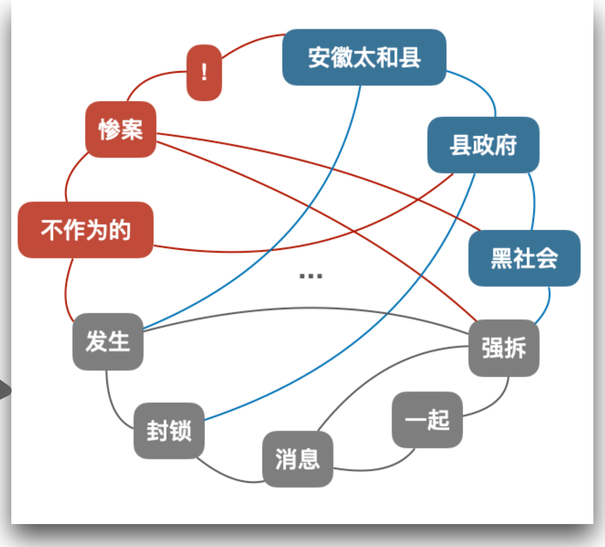
实体词

安徽太和县 县政府 黑社会

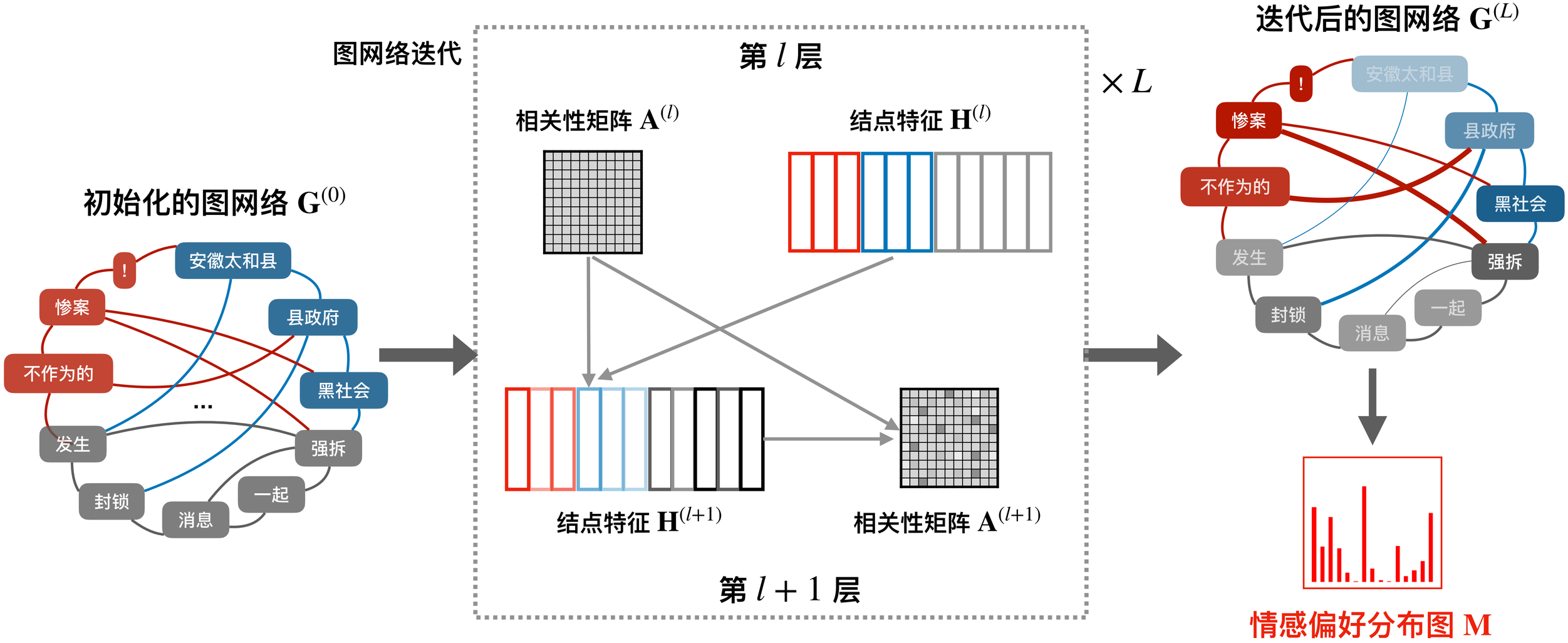
每个词都可能蕴含着情感信号

需要建模词与词之间的关联性

异构图网络



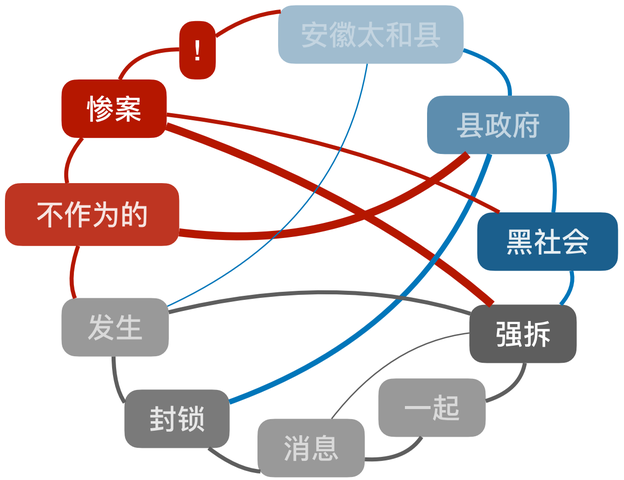
方法设计：情感偏好分布图的生成流程

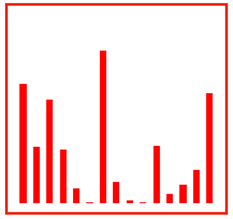


图网络构建 | 异构图卷积操作 | 相关性矩阵迭代 | 情感偏好分布图的读出

难点：情感偏好分布图的读出

迭代后的图网络 $\mathbf{G}^{(L)}$





情感偏好分布图 M

情感偏好分布图的读出

- **主要思想：**

在分配每个词符的情感偏好权重时：

(1) 将其与所有**情感词符**之间的相关性累积起来，作为需要**增强**的一项；

(2) 而将其与所有**实体词符**之间的相关性累积起来，作为需要**减弱**的一项。

情感词符的数量　　实体词符的数量

$$m_i = \sum_{k=1}^{n_S} \mathbf{A}_S^{(L)}(i,k) - \sum_{k=1}^{n_E} \mathbf{A}_E^{(L)}(i,k)$$

词符m_i的情感偏好权重

所有结点与情感（实体）词符结点之间构成的相关性矩阵$\mathbf{A}_S^{(L)}$（$\mathbf{A}_E^{(L)}$）

实验评估设计

- **评估一：EmoPref框架的有效性**
  - 使模型在学习过程中偏好情感信号是否能够有利于其检测假新闻？
- **评估二：EmoPref框架中情感偏好学习器设计的有效性**
  - 使学习器感知到情感词、实体词的先验知识是否必要？
  - 增强模型对情感词的偏好是否必要？抑制模型对实体词的偏好是否必要？
- **评估三：对EmoPref框架中生成的情感偏好分布图的探讨分析**
  - 在情感偏好增强的虚假新闻检测中，模型会更关注哪些词符？
  - 虚假新闻样本的情感偏好分布图会反映出哪些信息？

实验数据集

- **中文数据集（和研究点一相同）**
  - Weibo-16 [1]
  - Weibo-20
- **英文数据集**
  - Twitter [2]：是虚假新闻即时检测任务上，规模最大的英文数据集之一

	训练集	验证集	测试集	总计
假	3,419	1,140	1,140	5,699
真	5,406	1,802	1,802	9,010
总计	8,825	2,942	2,942	14,709

Twitter数据集的统计情况

[1] Jing Ma, et al. Detecting rumors from microblogs with recurrent neural networks. IJCAI 2016.
[2] Qiang Sheng*, **Xueyao Zhang***, et al. Integrating pattern- and fact-based fake news detection via model preference learning. CIKM 2021.

评估一：EmoPref框架的有效性

模型	Weibo-16				Weibo-20				Twitter			
	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$
BiLSTM (Graves 等, 2005)	0.822	0.807	0.754	0.860	0.667	0.660	0.710	0.610	0.767	0.732	0.829	0.635
***w/* EmoPref**	**0.849**	**0.850**	**0.802**	**0.898**	**0.709**	**0.709**	**0.715**	**0.702**	**0.793**	**0.788**	**0.822**	**0.754**
BERT (Devlin 等, 2019)	0.845	0.824	0.762	0.886	0.712	0.708	0.743	0.672	0.782	0.746	0.842	0.650
***w/* EmoPref**	**0.883**	**0.869**	**0.825**	**0.913**	**0.740**	**0.739**	**0.761**	**0.717**	**0.803**	**0.774**	**0.865**	**0.682**
EANN-Text (Wang 等, 2018)	0.836	0.819	0.760	0.878	0.692	0.690	0.717	0.663	0.770	0.725	**0.837**	0.614
***w/* EmoPref**	**0.877**	**0.867**	**0.821**	**0.915**	**0.740**	**0.740**	**0.727**	**0.752**	**0.798**	**0.788**	0.834	**0.741**
BERT-pEmo (Zhang 等, 2021)	0.869	0.850	0.794	0.906	0.728	0.722	0.762	0.682	0.794	0.762	0.850	0.675
***w/* EmoPref**	**0.909**	**0.896**	**0.847**	**0.945**	**0.746**	**0.744**	**0.768**	**0.720**	**0.804**	**0.776**	**0.855**	**0.697**

研究点一中的“BERT+新闻发布者情感”模型

1. **EmoPref的有效性**：无论把哪一种基线模型融入到EmoPref框架中均能得到提升
2. **情感偏好增强学习的有效性**：EmoPref 能够为BERT-pEmo带来提升，说明仅把情感信号作为一种辅助特征仍是不够的，通过情感偏好增强的学习，能够进一步挖掘情感的作用

评估一：EmoPref框架的有效性

模型	Weibo-16				Weibo-20				Twitter			
	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$	准确率	Macro F1 值	$F1_{fake}$	$F1_{real}$
BiLSTM (Graves 等, 2005)	0.822	0.807	0.754	0.860	0.667	0.660	0.710	0.610	0.767	0.732	0.829	0.635
***w/* EmoPref**	**0.849**	**0.850**	**0.802**	**0.898**	**0.709**	**0.709**	**0.715**	**0.702**	**0.793**	**0.788**	**0.822**	**0.754**
BERT (Devlin 等, 2019)	0.845	0.824	0.762	0.886	0.712	0.708	0.743	0.672	0.782	0.746	0.842	0.650
***w/* EmoPref**	**0.883**	**0.869**	**0.825**	**0.913**	**0.740**	**0.739**	**0.761**	**0.717**	**0.803**	**0.774**	**0.865**	**0.682**
EANN-Text (Wang 等, 2018)	0.836	0.819	0.760	0.878	0.692	0.690	0.717	0.663	0.770	0.725	**0.837**	0.614
***w/* EmoPref**	**0.877**	**0.867**	**0.821**	**0.915**	**0.740**	**0.740**	**0.727**	**0.752**	**0.798**	**0.788**	0.834	**0.741**
BERT-pEmo (Zhang 等, 2021)	0.869	0.850	0.794	0.906	0.728	0.722	0.762	0.682	0.794	0.762	0.850	0.675
***w/* EmoPref**	**0.909**	**0.896**	**0.847**	**0.945**	**0.746**	**0.744**	**0.768**	**0.720**	**0.804**	**0.776**	**0.855**	**0.697**

研究点一中的“BERT+新闻发布者情感”模型

1. **EmoPref的有效性**：无论把哪一种基线模型融入到EmoPref框架中均能得到提升
2. **情感偏好增强学习的有效性**：EmoPref 能够为BERT-pEmo带来提升，说明仅把情感信号作为一种辅助特征仍是不够的，通过情感偏好增强的学习，能够进一步挖掘情感的作用

评估二：情感偏好学习器设计的有效性

模型	Weibo-16		Weibo-20		Twitter	
	准确率	Macro F1 值	准确率	Macro F1 值	准确率	Macro F1 值
BiLSTM (Graves 等, 2005)	0.822	0.807	0.667	0.660	0.767	0.732
w/ EmoPref$_{GCN}$	0.819	0.806	0.665	0.663	0.768	0.737
w/ EmoPref$_{emo\uparrow}$	0.834	0.827	0.690	0.687	0.785	0.772
w/ EmoPref$_{entity\downarrow}$	0.831	0.822	0.694	0.692	0.779	0.758
***w/* EmoPref**	**0.849**	**0.850**	**0.709**	**0.709**	**0.793**	**0.788**
BERT (Devlin 等, 2019)	0.845	0.824	0.712	0.708	0.782	0.746
w/ EmoPref$_{GCN}$	0.850	0.837	0.701	0.703	0.785	0.734
w/ EmoPref$_{emo\uparrow}$	0.872	0.855	0.732	0.729	0.798	0.768
w/ EmoPref$_{entity\downarrow}$	0.865	0.846	0.738	0.730	0.790	0.757
***w/* EmoPref**	**0.883**	**0.869**	**0.740**	**0.739**	**0.803**	**0.774**
EANN-Text (Wang 等, 2018)	0.836	0.819	0.692	0.690	0.770	0.725
w/ EmoPref$_{GCN}$	0.839	0.830	0.689	0.688	0.772	0.747
w/ EmoPref$_{emo\uparrow}$	0.874	0.865	0.728	0.726	0.783	0.755
w/ EmoPref$_{entity\downarrow}$	0.854	0.840	0.708	0.700	0.778	0.742
***w/* EmoPref**	**0.877**	**0.867**	**0.740**	**0.740**	**0.798**	**0.788**
BERT-pEmo (Zhang 等, 2021)	0.869	0.850	0.728	0.722	0.794	0.762
w/ EmoPref$_{GCN}$	0.854	0.836	0.740	0.728	0.802	0.761
w/ EmoPref$_{emo\uparrow}$	0.888	0.872	0.736	0.732	0.796	0.765
w/ EmoPref$_{entity\downarrow}$	0.902	0.889	0.734	0.730	0.800	0.771
***w/* EmoPref**	**0.909**	**0.896**	**0.746**	**0.744**	**0.804**	**0.776**

三种消融实验：

- EmoPref$_{GCN}$：完全不依赖情感词、实体词的先验知识，并使用同构的GCN实现EmoPref
- EmoPref$_{emo\uparrow}$：只增强对情感词符的利用
- EmoPref$_{entity\downarrow}$：只减弱对实体词符的利用

1. **先验知识的重要性**：**EmoPref$_{GCN}$ 并未带来性能提升（因为此时模型依然主要依靠新闻可信度标签的监督信息，来学习对不同词符的利用权重）**
2. **“增强情感”与“削弱实体”都很重要**：**(1) 无论是EmoPref$_{emo\uparrow}$还是EmoPref$_{entity\downarrow}$ 都会带来性能提升；(2) 这二者之间并不存在明显的优劣关系**
3. **“增强情感”与“削弱实体”拥有互补性**：**完整的EmoPref框架效果最好**

评估三：对情感偏好分布图的探讨分析

研究点一中辅助情感特征集建模的一些情感信号

	类型	词符
偏好词集	标点符号	， 。 ！ ： ？ 【 】 “ ” （ ） … @ < # ~ > ； /
	否定词相关	不是、没有、不行、不够
	程度词相关	可能、一定、非常、有点、一些
	人称代词	我们、他们、你们
	其他	发现、这样、觉得、公开、作为、小心、发展、注意、爱心、不过
非偏好词集	事实证据相关	称、视频、链接、网页、全文、表示、图、调查、据、爆料
	实体相关	中国、北京、警察、地、车、上海、官员、央视、美国、政府
	人称代词	他、它、你
	其他	的、被、就、也、和、已、会、等、去、做、女、太、想、什么、某、死、过、死亡、时、里、秒、更、把、遭、晚、社会

发布者自我表达的词语

与特定新闻要素相关的词语

BiGRU模型通过EmoPref框架学到的（非）偏好词集

评估三：对情感偏好分布图的探讨分析

研究点一中辅助情感特征集建模的一些情感信号

	类型	词符
偏好词集	标点符号	， 。 ！ ： ？ 【 】 “ ” （ ） … @ < # ~ > ； /
	否定词相关	不是、没有、不行、不够
	程度词相关	可能、一定、非常、
	人称代词	我们、他们、你们
	其他	发现、这样、觉得、公开、作为、小心、发展、注意、爱心、不过
非偏好词集	事实证据相关	称、视频、链接、网页、全文、表示、图、调查、据、爆料
	实体相关	中国、北京、
	人称代词	他、它、你
	其他	的、被、就、 某、死、
		过、死亡、时、里、秒、更、把、遭、晚、社会

复数形式的人称代词——在拥有明显假新闻模式的帖子中，经常包含对某些群体的讨论，或者旨在怂恿读者行动起来

发布者自我表达的词语

单数形式的人称代词——常出现在具体事件描述的帖子中，这些帖子通常会与特定的人或事物有关，因此模型便很难从中学到一些模式共性

与特定新闻要素相关的词语

BiGRU模型通过EmoPref框架学到的（非）偏好词集

#	新闻原文
案例 1	山东 泗水 一群城管追着老大爷 撞 ， 直 致 老大爷所有 鸡蛋 都打碎在地上 ， 老大爷 无奈 的 坐 在那里 ， 城管 撞 了人 就 跑了 ， 白发 苍苍 应该八十 左右 了 ， 卖点 鸡蛋 也 挣 不了 几个钱 ， 又 何必 咄咄逼人 ？ 城管 就 没有 孤立无援 的 时候 。 如果城管只 会 欺压 百姓 ， 那 还要 城管有什么用 ？ 你们 这样 欺压 弱势群体 ， 迟早 会 遭 报应 的 。 **新闻可信度**：假 **模型判定结果**：BiLSTM（假），BiLSTM w/ EmoPref（假）
案例 2	【 浙大 学生 溺亡 ， 为拍 疯狂 毕业 照跳 西湖 】 6 月 29 日 ， 浙大 男生小辛 （ 化名 ） ， 和 同学 到 杭州 西湖 北里湖 孤山 “ 空谷传音 ” 景点 附近 水域 ， 请 同学 拍摄 自己 下水 游泳 的 照片 ， 他 从 北山路西泠桥 旁 纵身 跳入 西湖 ， 游 往 对岸 的 孤山公园 荷花池 ， 游 到 湖 中心 时 溺水 ， 此前 ， 他 已经拿到了一所美国大学 的 博士全额 奖学金 **新闻可信度**：假 **模型判定结果**：BiLSTM（真），BiLSTM w/ EmoPref（假）
案例 3	上海 谁 有兴趣 养狗 ？ 免费 。 金毛 ， 泰迪 ， 萨摩 ， 各种各样 。 有 杀狗场 被 捣毁 ， 无人领养 就要 安乐死 让这些 可爱 的 小生 命 陪伴 吧 ， 实在 没 条件 养 的 转发一下 ， 善莫大焉 。 13918487772 ， 周聊 。 **新闻可信度**：假 **模型判定结果**：BiLSTM（真），BiLSTM w/ EmoPref（假）

三则Weibo-20数据集上的假新闻

红色 代表增强模型偏好的词符

蓝色 代表削弱模型偏好的词符

颜色越深，代表增强（削弱）模型偏好的程度越大

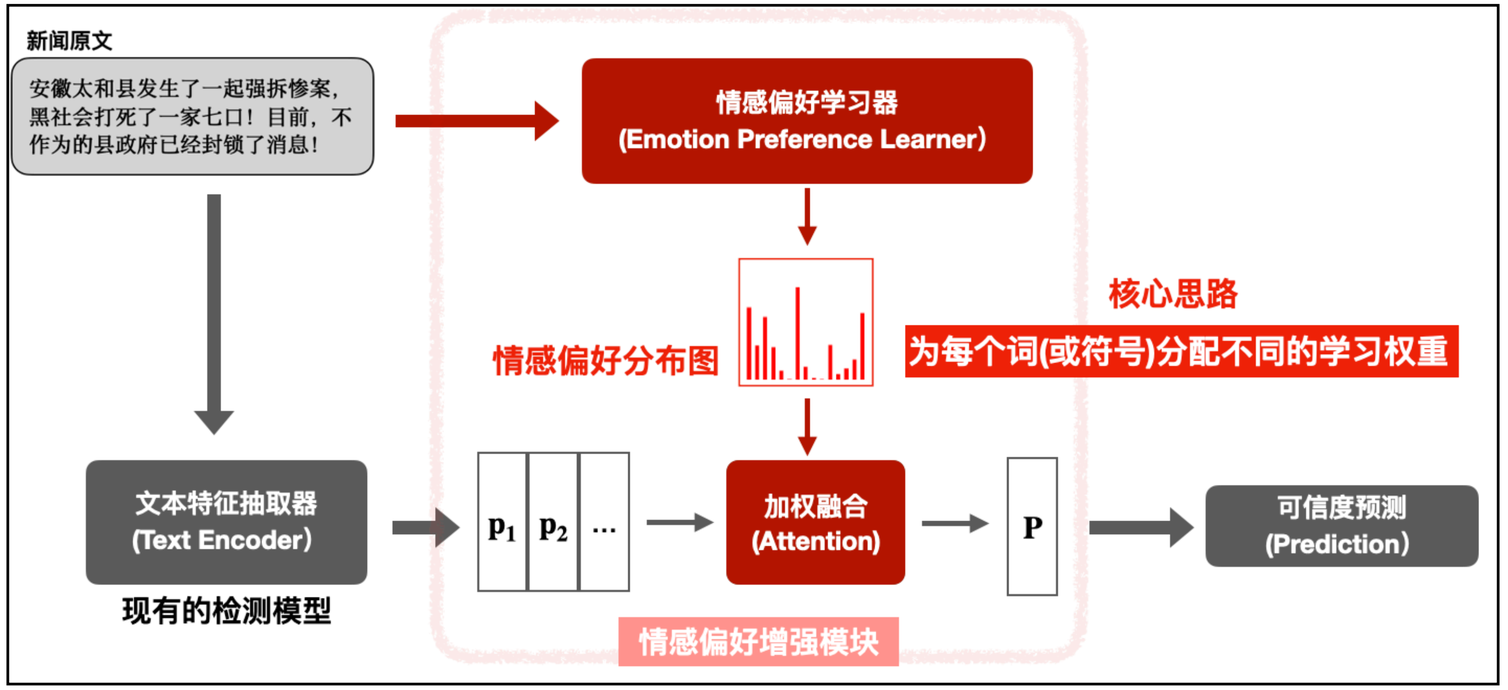
案例1 (情感信号非常强烈)
BiLSTM 已能够判对

案例2 (情感信号弱，实体信号强)
在EmoPref的引导下才能判对

案例3 (情感信号强，实体信号强)
在EmoPref的引导下才能判对

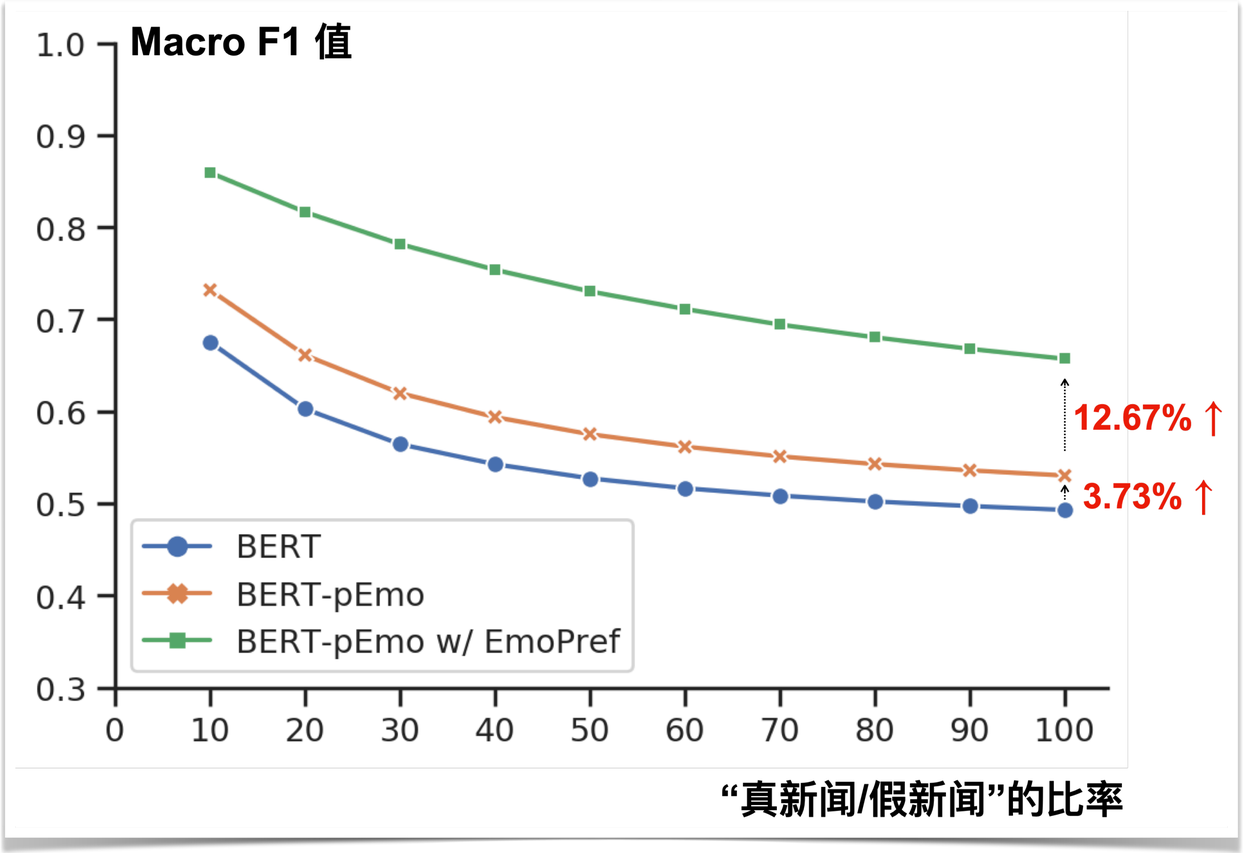
启示：EmoPref为何有效?

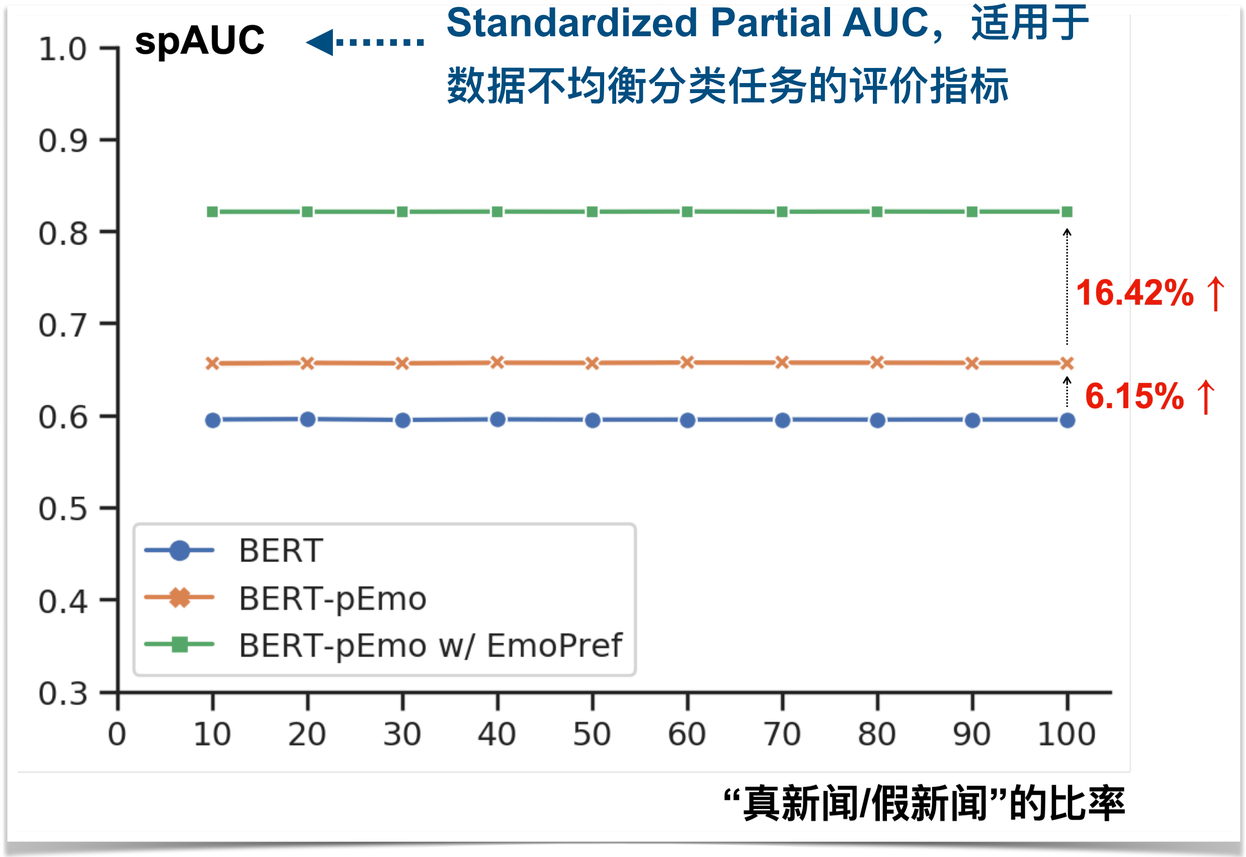
1. **对先验知识的利用：**情感词、实体词的**先验知识**能够被**有效建模与传播**。
2. **对数据的“有偏式”学习：**“为每个词符分配不同的学习权重”的设计思想启示我们，需要引导模型**有偏好地学习**一部分代表性数据，引导模型**有选择地关注**一些具有特殊含义的信号



目录

离线数据测试（一）：在不均衡数据下的性能表现





模型 BERT-pEmo w/ EmoPref (训练集为Weibo-20)

数据 由AI识谣平台导出，真:假的倾斜比从10:1到100:1

1. 数据极度不均衡，将会创造异常困难的业务场景
2. BERT-pEmo w/ EmoPref 模型具有优异的泛化性

线上系统评估（二）：与线上模型的性能比较

- 模型
  - BERT-pEmo w/ EmoPref
- 训练数据（与线上模型相同）
  - AI识谣平台中的95277条新闻数据（47415条假新闻，45162条真新闻）
- **评估结果**
  - **假新闻的F1值提高了6.5%**
  - **真新闻的F1值提高了1.4%**
  - **Macro F1值提高了3.9%**

BERT-pEmo w/ EmoPref 模型与平台上的多个方法共同决策出新闻的可信度

目录

本文工作总结

已有研究

从新闻原文中抽取情感特征

只建模新闻发布者的情感，
忽略了假新闻对读者的情绪煽动

研究点一

基于双重情感的虚假新闻检测

建模社区群体的情感，
并挖掘双重情感之间的联系

只把情感作为辅助特征，
忽略了可以增强模型自身对情感的表征

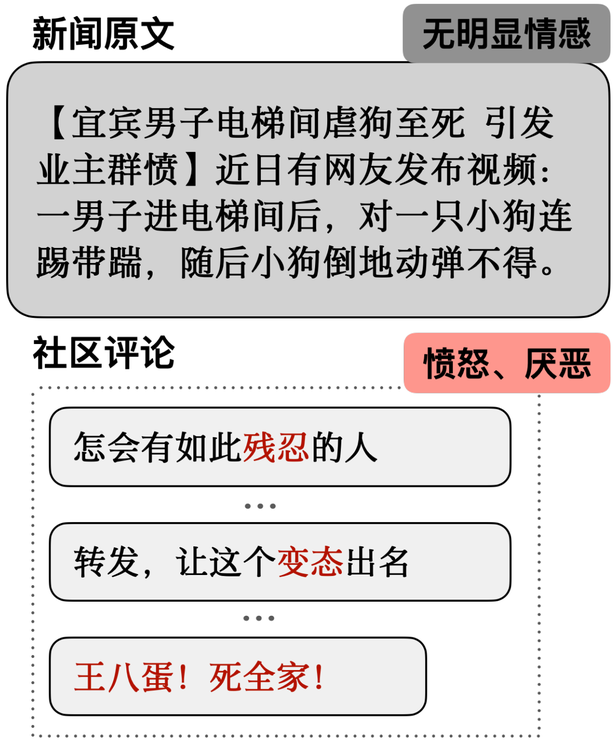
研究点二

情感偏好增强的虚假新闻即时检测

对模型的学习过程加以引导，
增强其对情感的偏好

本文的局限性与未来展望

- **情感信号与外部证据的联合检测**
  - 既要利用基于情感检测方法的及时性、便捷性，也要兼顾基于外部证据检测的可靠性、可解释性
- **多模态虚假新闻检测任务中的情感利用**
  - 新闻的图片、视频、音乐等媒介中都存在丰富的情感信号
  - 多个模态间的情感信号是否存在关联？如何融合多个模态的情感信号？
- **从延时检测到即时检测的迁移学习**
  - 模型的训练数据字段完整，但测试数据字段不完整



一则 Weibo-20 数据集上的真新闻

基于情感的检测方法，理论上很难判断该新闻的真实性

个人研究成果

学术论文:

- **[CCF-A；代表作]** **Xueyao Zhang**, et al. Mining Dual Emotion for Fake News Detection. **WWW 2021 (长文；Oral；一作).**
- **[CCF-B；代表作]** Qiang Sheng*, **Xueyao Zhang***, et al. Integrating Pattern- and Fact-based Fake News Detection via Model Preference Learning. **CIKM 2021 (长文；Oral；共同一作).**
- **[CCF-A]** Qiang Sheng, Juan Cao, **Xueyao Zhang**, et al. Zoom Out and Observe: News Environment Perception for Fake News Detection. **ACL 2022 (长文).**
- **[CCF-A]** Qiang Sheng, Juan Cao, **Xueyao Zhang**, et al. Article Reranking by Memory-enhanced Key Sentence Matching for Detecting Previously Fact-checked Claims. **ACL 2021 (长文).**

专利:

- 曹娟；**张雪遥**；盛强；谢添；李锦涛.《一种基于双重情感的舆情检测方法及系统》.
- 曹娟；盛强；**张雪遥**；钟雷；谢添.《引述句和辟谣模式句引导的“谣言-辟谣文章”匹配方法及系统》.
- 曹娟；盛强；**张雪遥**；钟雷；谢添.《基于模式信息和事实信息的联合虚假新闻检测方法》.
- 曹娟；盛强；**张雪遥**.《基于新闻环境信息建模的虚假新闻检测方法》.

研究项目:

- 网络谣言检测与舆论引导算法研究（国家自然科学基金-新疆联合基金 U1703261）
- 面向网络空间的事件全生命周期监测（国家重点研发计划 2017YFC0820604）

硕士学位论文答辩

敬请各位老师批评指正!

论文评阅意见与修改情况

- **评阅结果（4位评阅人）**
  - 总体评价：4票优秀
  - 答辩建议：3 票“同意答辩”，1 票“修改后答辩（论文需通过小的修改后答辩）”

论文评阅意见与修改情况

问题1：研究点一中，是否有必要额外地建模双重情感差分？建议作者在消融实验中补充“新闻发布者情感+社区群体情感”的消融实验。

修改情况：

- 在论文的2.3.4.3 节中，增补了“【检测模型】+【新闻发布者情感+社区群体情感】”的消融实验。
- 由表 2.8 中的实验结果可知：
  1. 同时引入“新闻发布者情感”和“社区群体情感”，能够比单独使用这二者具有更好的检测 Macro F1 值；
  2. 但在此基础上引入“双重情感差异”（即： 使用完整的“双重情感特征集”）能够进一步提升检测性能。由此可知：**额外地显式建模“新闻发布者”与“社区群体情感”的差分，能够比单独依赖 MLP 学习带来更好的检测结果**。

论文评阅意见与修改情况

问题2：研究点一中，RumourEval-19的实验结果为何远低于两个中文数据集?

修改情况：

- 论文重新修改了对于RumourEval-19检测指标很低的解释：
  1. 一方面，相关文献[1]的分析指出，该数据集的训练集、验证集、测试集的**类别分布差距很大**，例如：训练集中的待查证新闻的比例，远多于验证集和测试集中的比例（表 2.2），这也造成了模型对于待查证新闻的检测效果较差（如表2.7 所示，待查证新闻的F1值远低于真、假新闻）；
  2. 另一方面，由于 RumourEval-19的**数据规模非常小**，这也造成了模型学习和检测指标结果的极大波动性。
- 对于两个中文数据集，由于其训练集、验证集、测试集的类别分布都较为均衡（都约为1:1），且数据规模都相对较大，因此更有利于机器学习模型展现出良好的检测能力。

[1] Li Q, et al. Eventai at semeval-2019 task 7: Rumor detection on social media by exploiting content, user credibility and propagation information. SemEval, 2019.

论文评阅意见与修改情况

问题3：文中的即时检测和实时检测有什么区别？有无客观的标准（如时间节点等），来界定即时检测和早期检测、延时检测？

修改情况：

- 已在1.2.2.2节中作出补充说明。具体地：
  1. “实时检测（Real-time Detection）”一般代表的是软、硬件系统具有极快的响应时间，更多针对的是**对系统或算法的运行性能的约束**；
  2. 本文中的“即时检测（Instant Detection）”代表的是“新闻一经发布就进行的检测”，它**约束的是事件的发展阶段**。
- 在领域内的已有研究中，暂无客观标准来明确给出“即时检测–早期检测–延时检测”之间的界限。本文是从“**对数据、信息或知识的利用程度**”的来区别这三者，是一种启发式的定性划分。

论文评阅意见与修改情况

其余的论文写作、格式规范问题：

包括：（1）国内外研究现状的总结章节；（2）系统章节对本文工作的介绍；（3）公式3.2的文字解释；（4）建议将“本研究”改为“本文”；（5）个别文字可再凝练。比如p52, "在本章中，本研究针对的是...问题" 可考虑改为 “本章针对问题”；（6）论文中的脚注信息大部分实际上是引用，应该采用引用的格式；（7）论文中很多地方用楷体来进行专用名词的描述，但是哪些词用楷体，哪些词不用楷体不是很明确，建议讨论是否有必要性用字体来表示这些概念；（8）建议文中将相关的"证明"改成"验证"或者"表明"等。

修改情况：

- **均已完成修改。**