Memory-Guided Multi-View Multi-Domain Fake News Detection

Yongchun Zhu, Qiang Sheng, Juan Cao, Qiong Nan, Kai Shu, Minghui Wu, Jindong Wang, and Fuzhen Zhuang

TKDE2022

code: https://github.com/ICTMCG/M3FEND

Reported by Xiaoke Li

Introduction

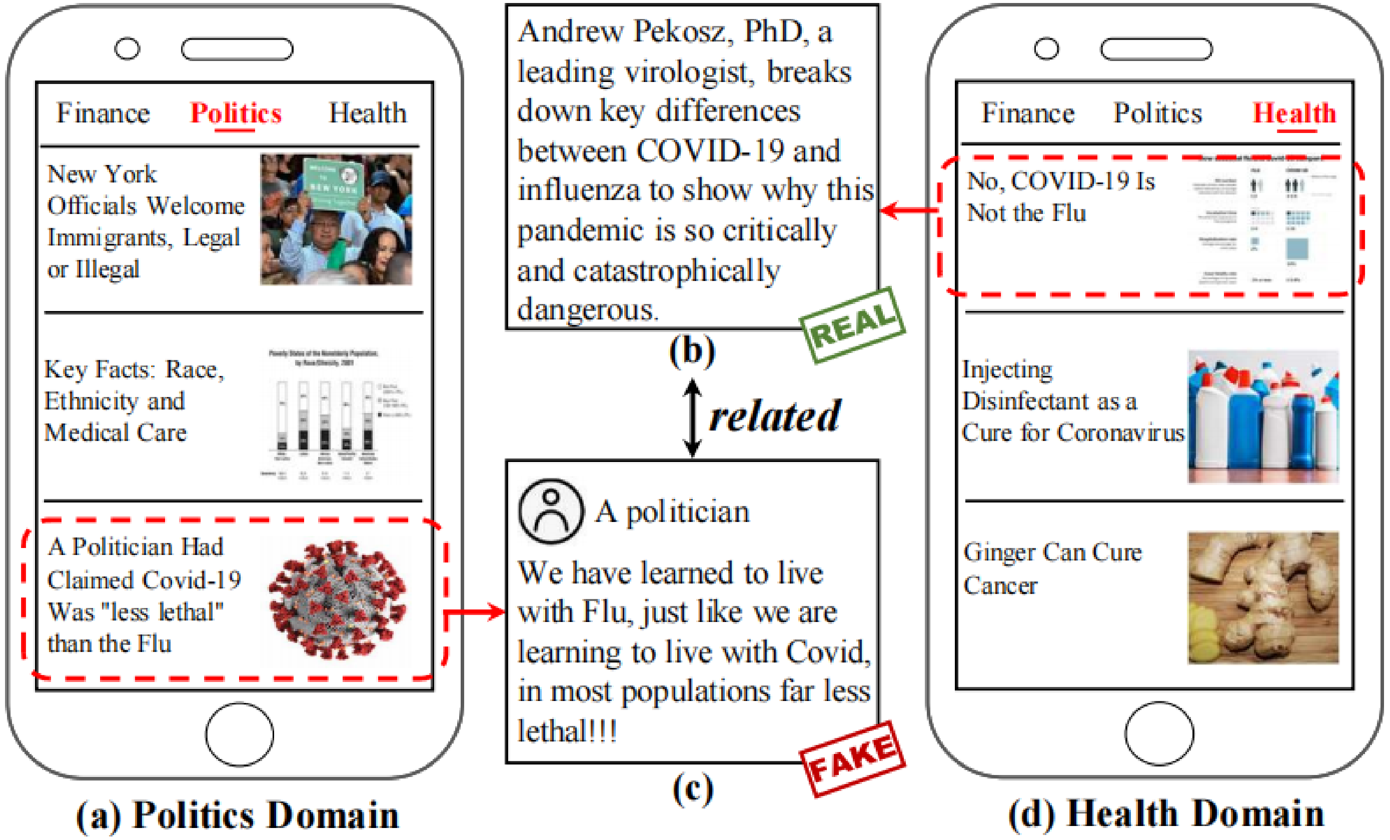
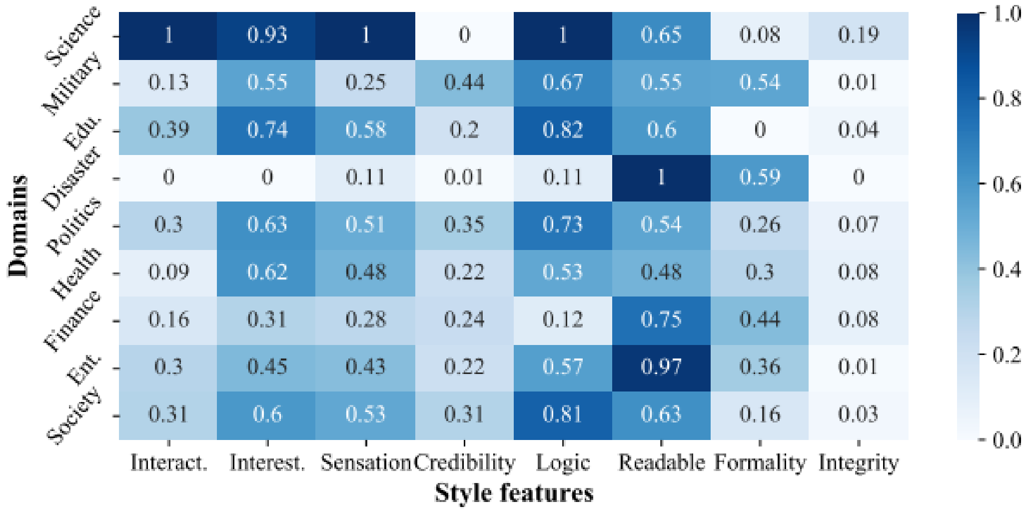


Fig. 1. An example of a real-world news platform with ***multiple news domains***. The news distributions vary from domain to domain, leading to the challenge of ***domain shift***. However, a news piece is a mixture of diverse elements which makes it relate to multiple news domains, e.g., the political news (c) is also related to the health news (b), leading to the challenge of ***domain labeling incompleteness***.

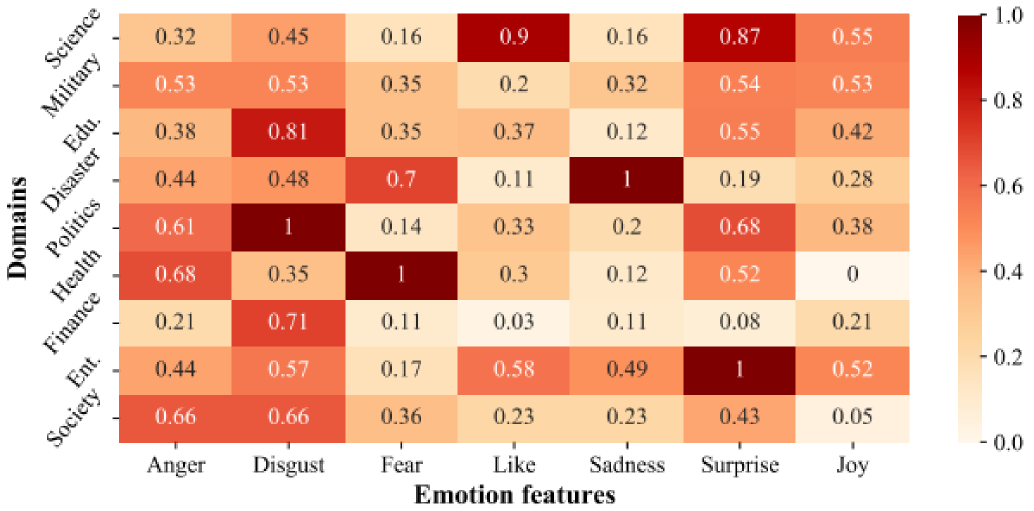
Method

(a) Top 20 words in the nine domains

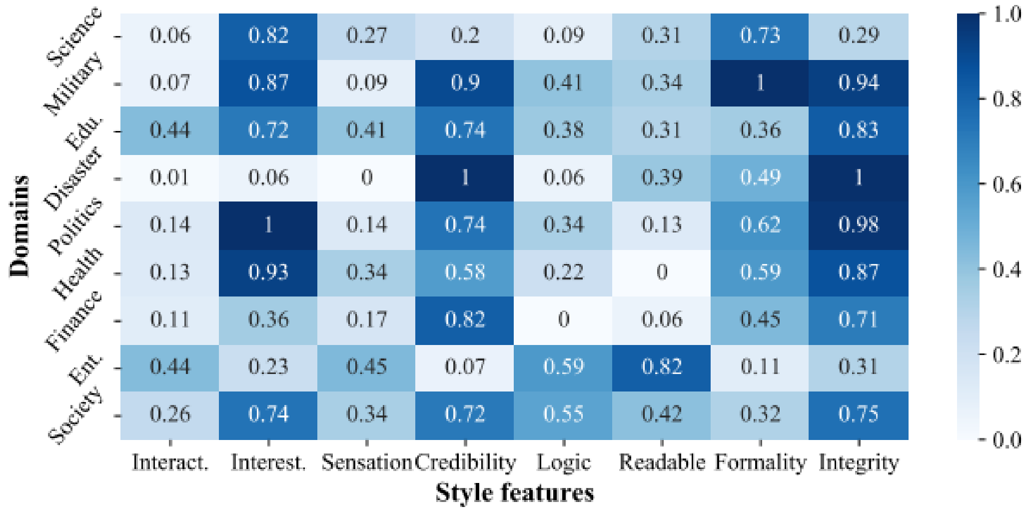
Method



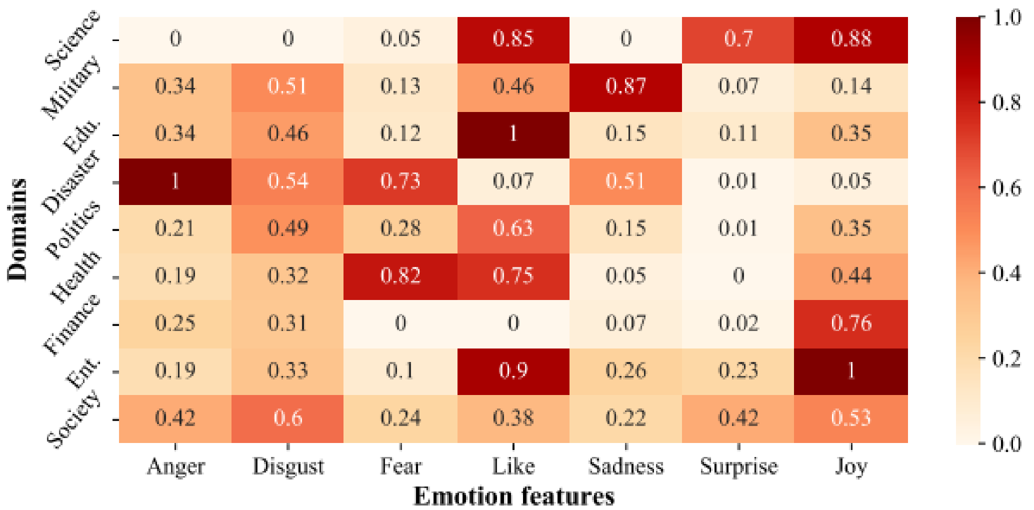
(b) Style features of fake news



(c) Emotion features of fake news

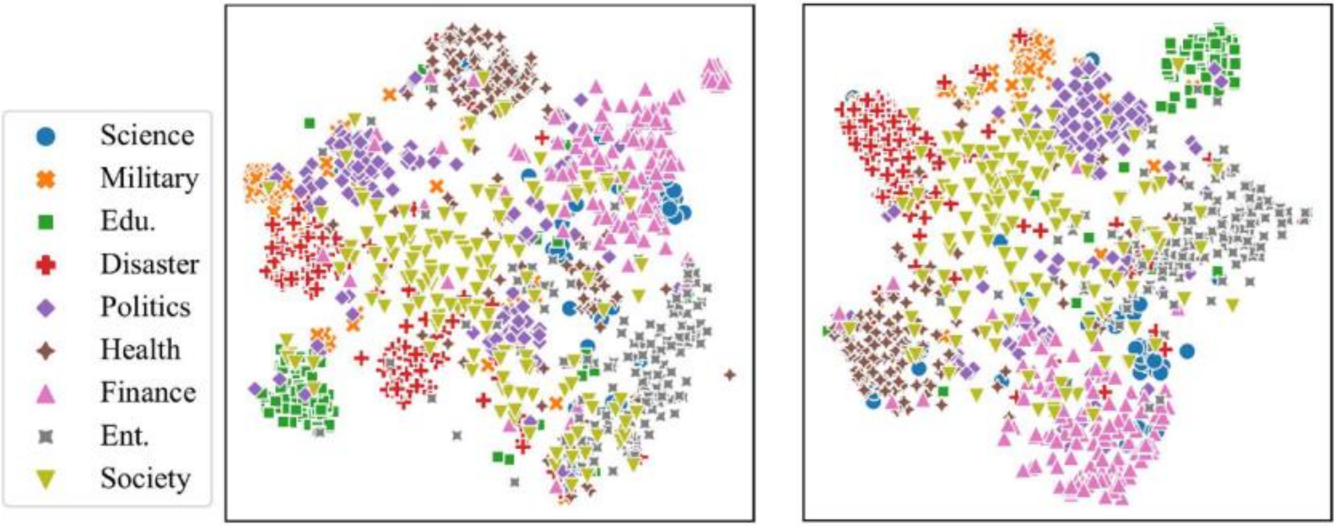


(d) Style features of real news



(e) Emotion features of real news

Method



(a) TextCNN (b) BiGRU

Fig. 3. Visualizations of unclear domain boundaries with the domain classification task using t-SNE on the Ch-9.

TABLE 1
Accuracy on Ch-9 for domain classification.

	Science	Military	Edu.	Disaster	Politics	Health	Finance	Ent.	Society
TextCNN	0.42	0.57	0.75	0.74	0.77	0.82	0.83	0.84	0.78
BiGRU	0.56	0.62	0.87	0.69	0.77	0.78	0.84	0.77	0.79

Method

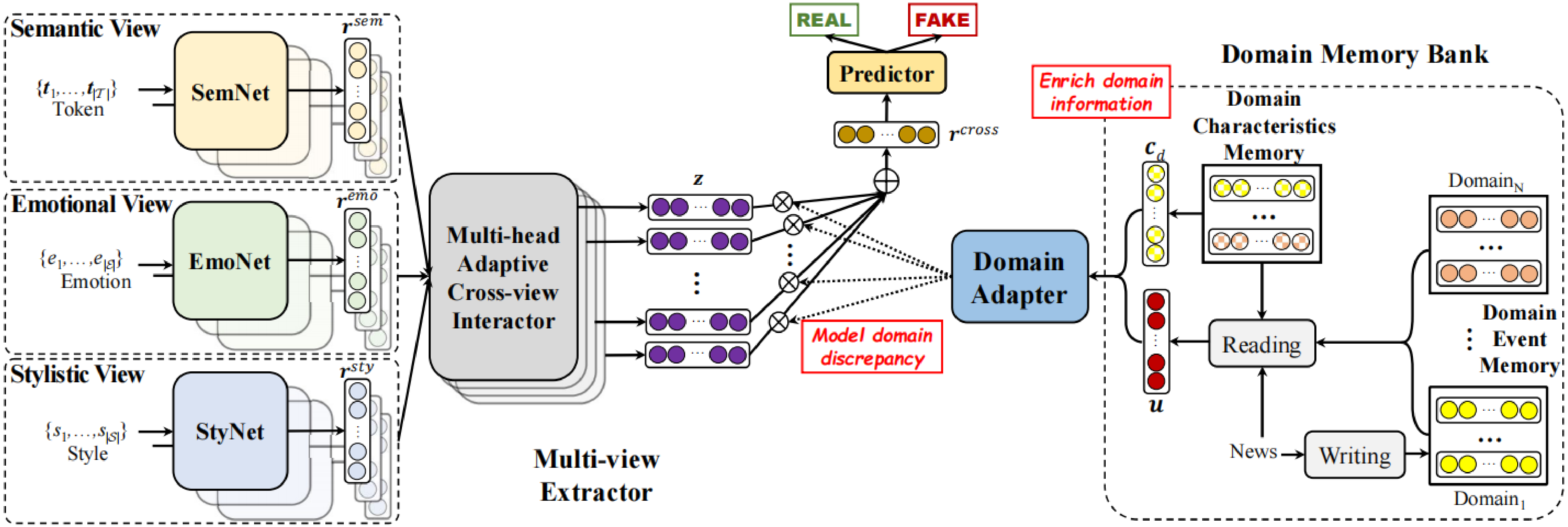


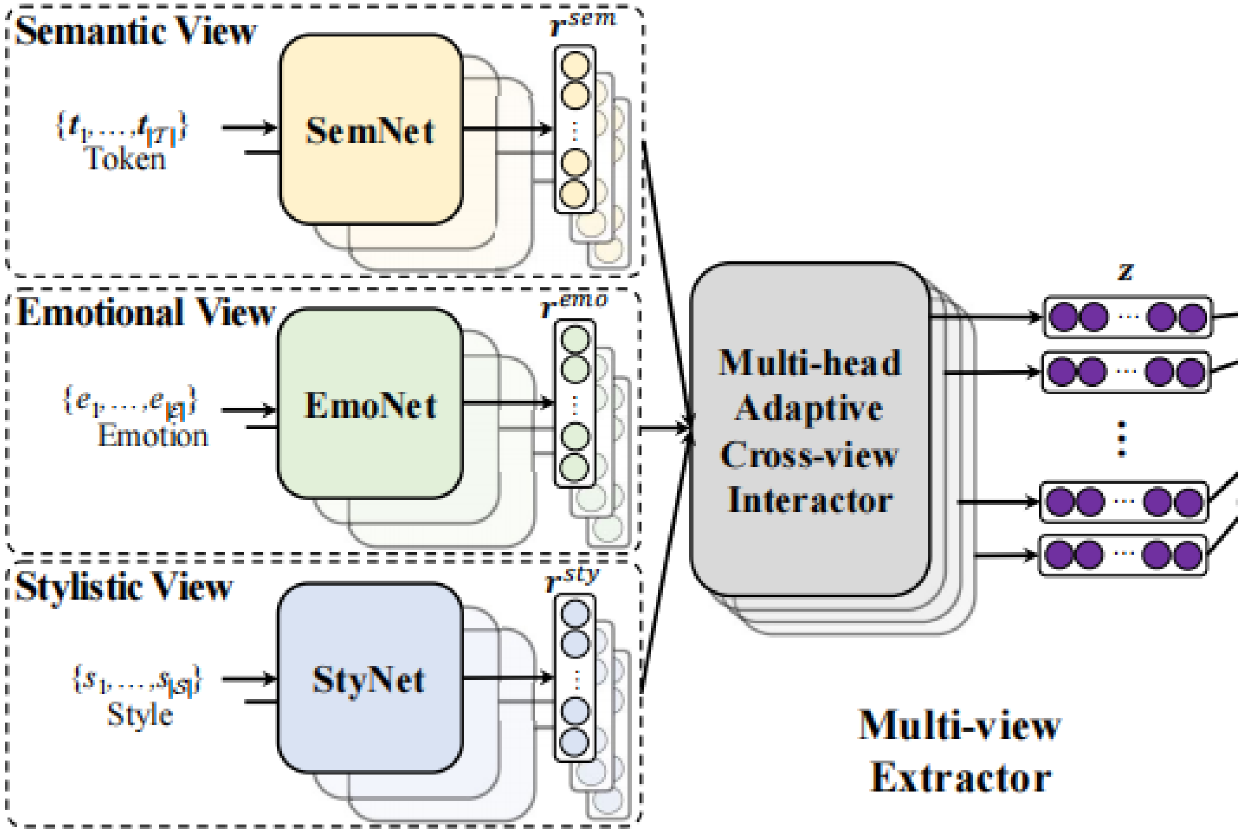
Fig. 4. Overall architecture of the Memory-guided Multi-view Multi-domain Fake News Detection Framework (M^3FEND). The model consists of a Multi-view Extractor, a Domain Memory Bank, a Domain Adapter, and a Predictor. The Multi-view Extractor aims to extract multi-view representations and model cross-view interactions. The Domain Memory Bank stores and provides enriched domain information. The Domain Adapter aggregates discriminative cross-view representations for news in different domains. The Predictor uses the aggregated representations for the final prediction.

Method

$$\mathcal{T} = \{t_1, \ldots, t_{|\mathcal{T}|}\}$$

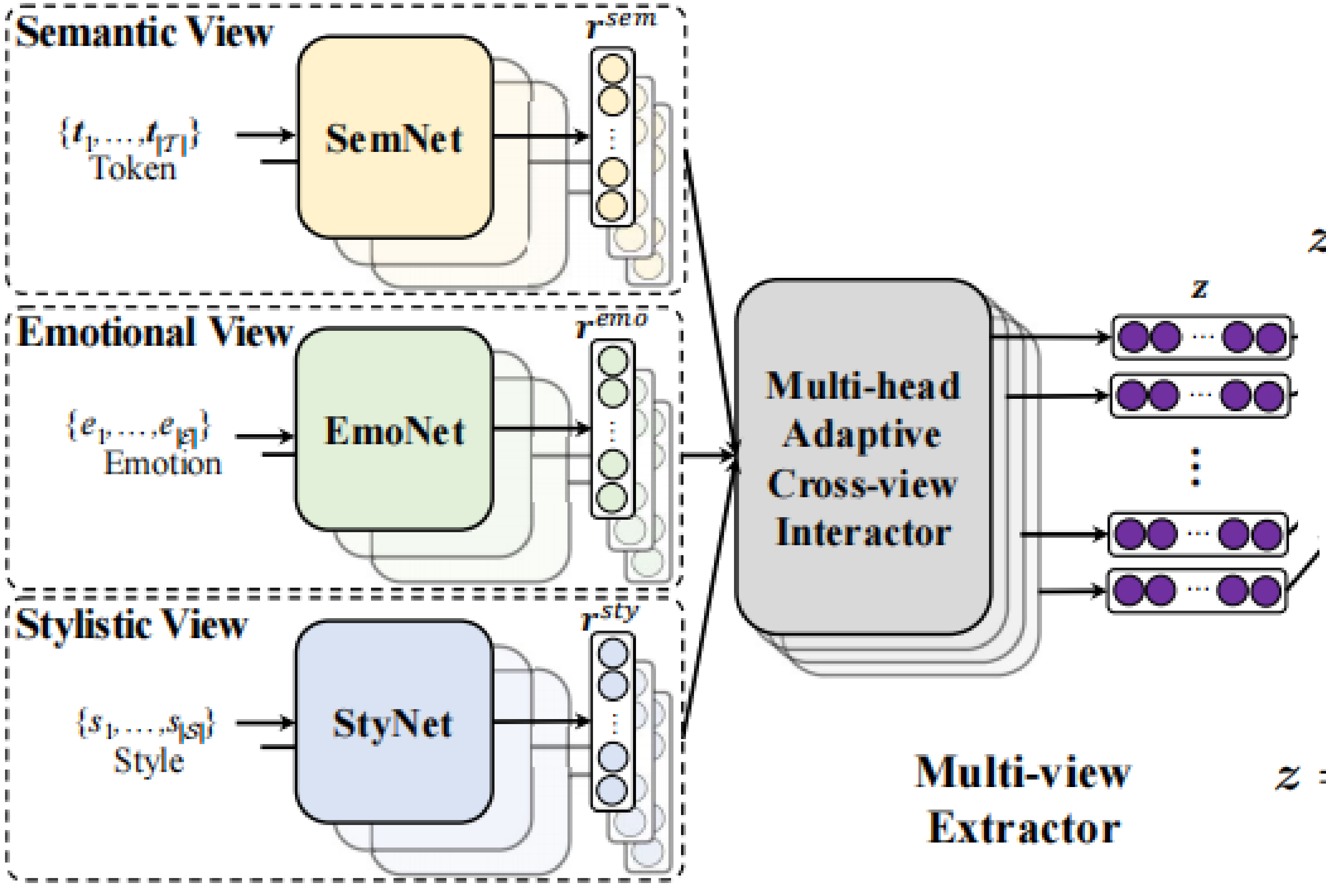
$$\mathcal{E} = \{e_1, \ldots, e_{|\mathcal{E}|}\}$$

$$\mathcal{S} = \{s_1, \ldots, s_{|\mathcal{S}|}\}$$



$$\boldsymbol{r}^{sem} = \mathrm{SemNet}(\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{|\mathcal{T}|}\}). \quad (1) \qquad \boldsymbol{r}^{emo} = \mathrm{EmoNet}(\{e_1, \ldots, e_{|\mathcal{E}|}\}). \quad (2) \qquad \boldsymbol{r}^{sty} = \mathrm{StyNet}(\{s_1, \ldots, s_{|\mathcal{S}|}\}). \quad (3)$$

Method



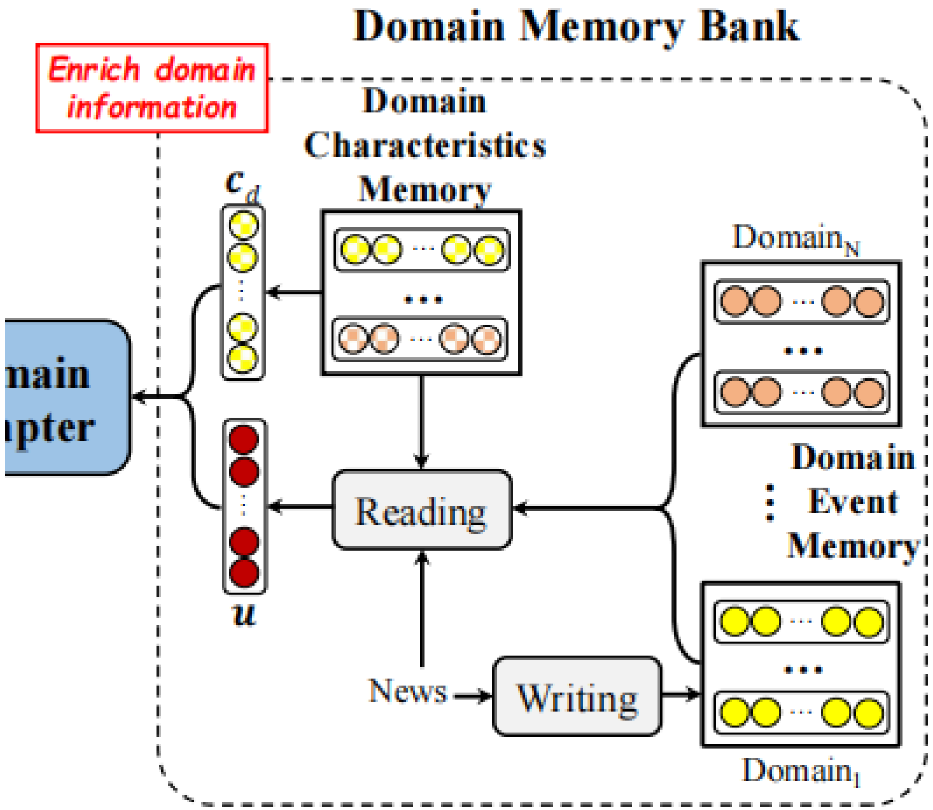
$$\{r_i^{sem}\}_{i=1}^{k_{sem}} \quad \{r_i^{emo}\}_{i=1}^{k_{emo}} \quad \{r_i^{sty}\}_{i=1}^{k_{sty}}$$

$$\boldsymbol{z} = \exp\left[\sum_{i=1}^{k_{sem}} a_i^{sem} \ln \boldsymbol{r}_i^{sem} + \sum_{j=1}^{k_{emo}} a_j^{emo} \ln \boldsymbol{r}_j^{emo} + \sum_{q=1}^{k_{sty}} a_q^{sty} \ln \boldsymbol{r}_q^{sty}\right], \tag{4}$$

$$\boldsymbol{z} = \prod_{i=1}^{k_{sem}} (\boldsymbol{r}_i^{sem})^{a_i^{sem}} \odot \prod_{j=1}^{k_{emo}} (\boldsymbol{r}_j^{emo})^{a_j^{emo}} \odot \prod_{q=1}^{k_{sty}} (\boldsymbol{r}_q^{sty})^{a_q^{sty}}, \tag{5}$$

$$\{z_i\}_{i=1}^{H}$$

Method



$$u = \sum_{i=1}^{O} v_i c_i,$$

$$\mathcal{C} = \{c_i\}_{i=1}^{N} \qquad \mathcal{M}j = \{m_i\}_{i=1}^{Q}$$

$$n = [\mathcal{G}(\{t_1, \ldots, t_{|\mathcal{T}|}\}); \{e_1, \ldots, e_{|\mathcal{E}|}\}; \{s_1, \ldots, s_{|\mathcal{S}|}\}] \in \mathbb{R}^I.$$

$$o_j = \mathrm{softmax}(nWg(\mathcal{M}_j)/\tau)\mathcal{M}_j, \qquad (6)$$

$$\mathcal{D} = [o_1, \ldots, o_N] \in \mathbb{R}^{N \times I},$$

$$v = \mathrm{softmax}(nVg(\mathcal{D})), \qquad (7)$$

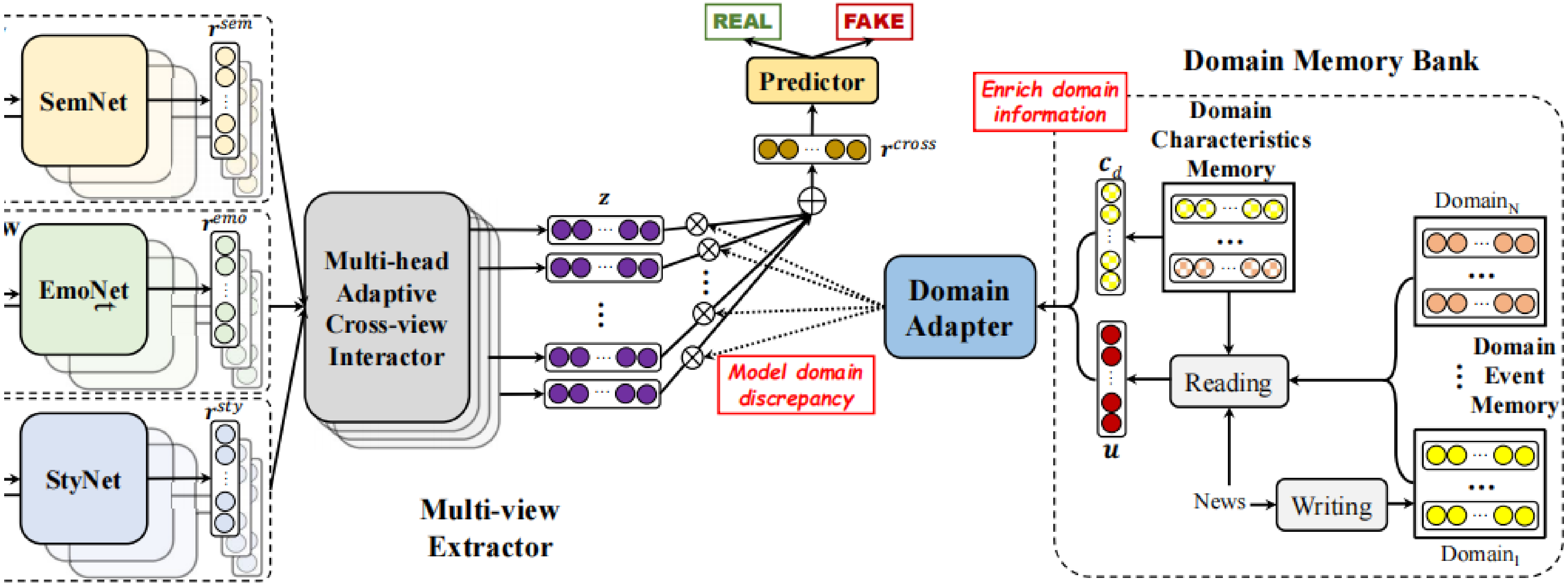
$$m_i = m_i - \beta erase_i + \beta add_i,$$

$$add_i = sim_i \cdot n,$$

$$erase_i = sim_i \cdot m_i, \qquad (8)$$

$$sim = \mathrm{softmax}(nWg(\mathcal{M}_d)/\tau).$$

Method



$$r = \sum_{i=1}^{H} w_i z_i, \quad w = \mathrm{softmax}(f([c_d, u])), \tag{9}$$

$$\mathcal{L} = -y \log \hat{p} - (1-y)\log(1-\hat{p}). \tag{11}$$

$$\hat{p} = \mathrm{Sigmoid}(\mathrm{MLP}(r)). \tag{10}$$

Experiments

TABLE 2
Data Statistics of Ch-9

Domain	Science	Military	Edu.	Disasters	Politics
#Real	143	121	243	185	306
#Fake	93	222	248	591	546
Total	236	343	491	776	852
Domain	**Health**	**Finance**	**Ent.**	**Society**	**All**
#Real	485	959	1,000	1,198	4,640
#Fake	515	362	440	1,471	4,488
Total	1,000	1,321	1,440	2,669	9,128

TABLE 3
Data Statistics of En-3

Domain	GossipCop	PolitiFact	COVID	All
#Real	16,804	447	4,750	22,001
#Fake	5,067	379	1,317	6,763
Total	21,871	826	6,067	28,764

TABLE 4
Emotion Features.

Feature	Description
Emotional Category	The probabilities that the given text contains certain emotions obtained from publicly available emotion classifiers.
Emotional Lexicon	The overall emotion score that is aggregated from scores of each word and the whole text across all the emotions.
Emotional Intensity	The overall intensity scores which is extracted from the existing emotion dictionaries annotated with similar process as Emotional Lexicon.
Sentiment Score	The degree of the positive or negative polarity of the whole text calculated by using sentiment dictionaries or public toolkits.
Auxiliary Features	The frequency of emoticons, punctuations, sentimental words, personal pronoun, and uppercase letters.

Experiments

TABLE 5
Results on the En-3 dataset. * ($p \leq 0.05$) and ** ($p \leq 0.005$) indicate paired t-test of M^{3}FEND vs. the best baseline.

	Method	Gossip.	Polit.	COVID	overall F1	overall Acc	overall AUC
single	BiGRU	0.7666	0.7722	0.8885	0.7958	0.8668	0.8840
single	TextCNN	0.7786	0.8011	0.9040	0.8079	0.8692	0.9023
single	RoBERTa	0.7810	0.8583	0.9288	0.8184	0.8802	0.9108
mixed	BiGRU	0.7479	0.7339	0.7448	0.7501	0.8321	0.8504
mixed	TextCNN	0.7519	0.7040	0.8322	0.7679	0.8362	0.8674
mixed	RoBERTa	0.7823	0.7967	0.9014	0.8101	0.8744	0.9058
mixed	StyleLSTM	0.8007	0.7937	0.9252	0.8285	0.8826	0.9250
mixed	DualEmo	0.8056	0.7868	0.9019	0.8270	0.8818	0.9251
multi	EANN	0.7937	0.7558	0.8836	0.8123	0.8743	0.9053
multi	MMoE	0.8022	0.8477	0.9379	0.8361	0.8920	0.9265
multi	MoSE	0.7981	**0.8576**	0.9326	0.8318	0.8885	0.9252
multi	EDDFN	0.8067	0.8505	0.9306	0.8378	0.8912	0.9263
multi	MDFEND	0.8080	0.8473	0.9331	0.8390	0.8936	0.9237
	M^{3}FEND	**0.8237****	0.8478	**0.9392**	**0.8517****	**0.8977***	**0.9342***

TABLE 6
Results on the Ch-3 dataset. * ($p \leq 0.05$) and ** ($p \leq 0.005$) indicate paired t-test of M^{3}FEND vs. the best baseline.

	Method	Politics	Health	Ent.	overall F1	overall Acc	overall AUC
single	BiGRU	0.8469	0.8335	0.7913	0.8402	0.8411	0.9213
single	TextCNN	0.8514	0.9041	0.8423	0.8846	0.8850	0.9521
single	RoBERTa	0.8137	0.8924	0.8434	0.8691	0.8697	0.9420
mixed	BiGRU	0.8384	0.8577	0.8687	0.8733	0.8741	0.9402
mixed	TextCNN	0.8579	0.8716	0.8683	0.8833	0.8838	0.9493
mixed	RoBERTa	0.8300	0.8955	0.8862	0.8911	0.8915	0.9566
mixed	StyleLSTM	0.8298	0.8924	0.8896	0.8912	0.8917	0.9564
mixed	DualEmo	0.8362	0.8968	0.9020	0.8977	0.8980	0.9605
multi	EANN	0.8405	0.9189	0.8974	0.9038	0.9042	0.9644
multi	MMoE	**0.8779**	0.9215	0.8800	0.9048	0.9052	0.9629
multi	MoSE	0.8564	0.9023	0.8872	0.8978	0.8985	0.9572
multi	EDDFN	0.8440	0.9235	0.8748	0.8965	0.8970	0.9614
multi	MDFEND	0.8555	0.9419	0.9103	0.9205	0.9208	0.9750
	M^{3}FEND	0.8618	**0.9479***	**0.9304****	**0.9308****	**0.9311****	**0.9759**

Experiments

TABLE 7

Results on the Ch-6 dataset. * ($p \leq 0.05$) and ** ($p \leq 0.005$) indicate paired t-test of M^{3}FEND vs. the best baseline.

	Method	Edu.	Disaster	Health	Finance	Ent.	Society	overall F1	overall Acc	overall AUC
single	BiGRU	0.7697	0.7191	0.8451	0.8247	0.8026	0.8015	0.8266	0.8270	0.8979
	TextCNN	0.7805	0.4388	0.9012	0.7671	0.7930	0.8654	0.8494	0.8499	0.9195
	RoBERTa	0.8175	0.7584	0.8909	0.8498	0.8549	0.8304	0.8576	0.8580	0.9288
mixed	BiGRU	0.8253	0.7938	0.8626	0.8254	0.8604	0.8206	0.8491	0.8501	0.9249
	TextCNN	0.8593	0.8240	0.8832	0.8646	0.8659	0.8641	0.8776	0.8783	0.9483
	RoBERTa	0.8664	0.8515	0.9100	0.8700	0.8872	0.8634	0.8872	0.8877	0.9494
	StyleLSTM	0.8565	0.8374	0.9080	0.8766	0.8957	0.8546	0.8844	0.8851	0.9489
	DualEmo	0.8472	0.8352	0.9055	0.8951	0.9043	0.8642	0.8904	0.8909	0.9579
multi	EANN	0.8613	0.8657	0.9150	0.8621	0.8871	0.8791	0.8919	0.8925	0.9605
	MMoE	0.8625	0.8777	0.9260	0.8546	0.8882	0.8655	0.8894	0.8900	0.9563
	MoSE	0.8569	0.8588	0.9118	0.8639	0.8904	0.8757	0.8913	0.8918	0.9533
	EDDFN	0.8780	0.8734	0.9280	0.8456	0.8819	0.8716	0.8917	0.8921	0.9544
	MDFEND	0.8826	0.8781	0.9430	0.8749	0.9095	0.8940	0.9093	0.9097	0.9694
	M^{3}FEND	**0.8836**	**0.8824**	**0.9515***	**0.8997***	**0.9296****	**0.9043****	**0.9208****	**0.9211****	**0.9762***

Experiments

TABLE 8
Results on the Ch-9 dataset. * ($p \leq 0.05$) and ** ($p \leq 0.005$) indicate paired t-test of M^{3}FEND vs. the best baseline.

	Method	Science	Military	Edu.	Disaster	Politics	Health	Finance	Ent.	Society	overall F1	overall Acc	overall AUC
single	BiGRU	0.5175	0.3365	0.7416	0.7293	0.8588	0.8373	0.8137	0.7992	0.7918	0.8103	0.8103	0.8902
single	TextCNN	0.4074	0.3365	0.8059	0.4388	0.8482	0.8819	0.8215	0.7973	0.8615	0.8369	0.8370	0.9094
single	RoBERTa	0.7463	0.7369	0.8146	0.7547	0.8044	0.8873	0.8361	0.8513	0.8300	0.8477	0.8477	0.9226
mixed	BiGRU	0.7269	0.8724	0.8138	0.7935	0.8356	0.8868	0.8291	0.8629	0.8485	0.8595	0.8598	0.9309
mixed	TextCNN	0.7254	0.8839	0.8362	0.8222	0.8561	0.8768	0.8638	0.8456	0.8540	0.8686	0.8687	0.9381
mixed	RoBERTa	0.7777	0.9072	0.8331	0.8512	0.8366	0.9090	0.8735	0.8769	0.8577	0.8795	0.8797	0.9451
mixed	StyleLSTM	0.7729	0.9187	0.8341	0.8532	0.8487	0.9084	0.8802	0.8846	0.8552	0.8820	0.8821	0.9471
mixed	DualEmo	0.8323	0.9026	0.8362	0.8396	0.8455	0.8905	**0.9053**	0.8944	0.8569	0.8846	0.8846	0.9541
multi	EANN	0.8225	0.9274	0.8624	0.8666	0.8705	0.9150	0.8710	0.8957	0.8877	0.8975	0.8977	0.9610
multi	MMoE	**0.8755**	0.9112	0.8706	0.8770	0.8620	0.9364	0.8567	0.8886	0.8750	0.8947	0.8948	0.9547
multi	MoSE	0.8502	0.8858	0.8815	0.8672	0.8808	0.9179	0.8672	0.8913	0.8729	0.8939	0.8940	0.9543
multi	EDDFN	0.8186	0.9137	0.8676	0.8786	0.8478	0.9379	0.8636	0.8832	0.8689	0.8919	0.8919	0.9528
multi	MDFEND	0.8301	0.9389	0.8917	**0.9003**	**0.8865**	0.9400	0.8951	0.9066	0.8980	0.9137	0.9138	0.9708
	M^{3}FEND	0.8292	**0.9506****	**0.8998**	0.8896	0.8825	**0.9460**	0.9009	**0.9315****	**0.9089****	**0.9216****	**0.9216****	**0.9750***

Experiments

TABLE 9
Relative improvement over the online baseline.

Improvement on	SPAUC	AUC	F1
EANN	2.12%	0.67%	0.33%
EDDFN	-0.37%	-2.02%	-3.34%
MDFEND	2.82%	0.74%	1.85%
M^3FEND	**5.50%**	**2.89%**	**4.49%**

TABLE 10
Results of ablation study.

	Ch-3	Ch-6	Ch-9	En-3
M^3FEND	**0.9308**	**0.9208**	**0.9216**	**0.8517**
w/o SemView	0.8202	0.8161	0.8249	0.6573
w/o EmoView	0.9195	0.9136	0.9147	0.8403
w/o StyView	0.9255	0.9178	0.9177	0.8472
w/o Interactor	0.9217	0.9169	0.9173	0.8398
w/o Memory	0.9237	0.9182	0.9176	0.8501
w/o Adapter	0.9172	0.9169	0.9157	0.8367

Experiments

TABLE 11
A case of the distribution of predicted domain label.

Target News		Trump nearly fainted during his speech and cancelled his subsequent trip. A symptom of COVID-19?
Domain	**Similarity v**	**Representative Example**
Science	0.02	NASA used the Nuclear Spectroscopy Telescope to photo the spiral galaxy 1068 in the Cetus.
Military	0.04	U.S. sends 35 medical ships.
Edu.	0.01	A student admitted to Harvard University.
Disaster	0.02	The US "World Journal" reported a five-level fire in a restaurant.
Politics	0.33	US deaths from COVID-19 exceed 100k.
Health	0.21	The animal experiment of Oxford's COVID-19 vaccine failed.
Finance	0.12	Pfizer's stocking price rose 15%, boosted by the company's COVID-19 vaccine news.
Ent.	0.09	10 more people tested positive for COVID-19 in Italian Serie A.
Society	0.16	A COVID-19 carrier refused security check at the airport.

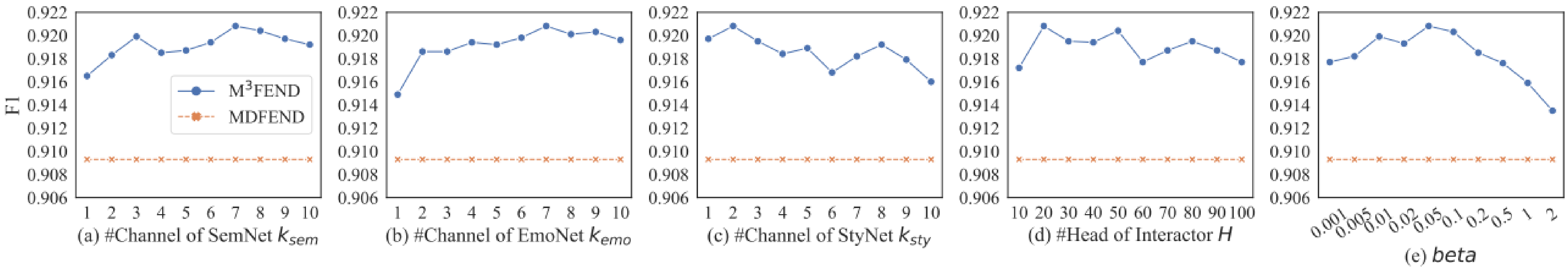


Fig. 7. Performance (F1) of M^{3}FEND with various hyperparameters.

Experiments

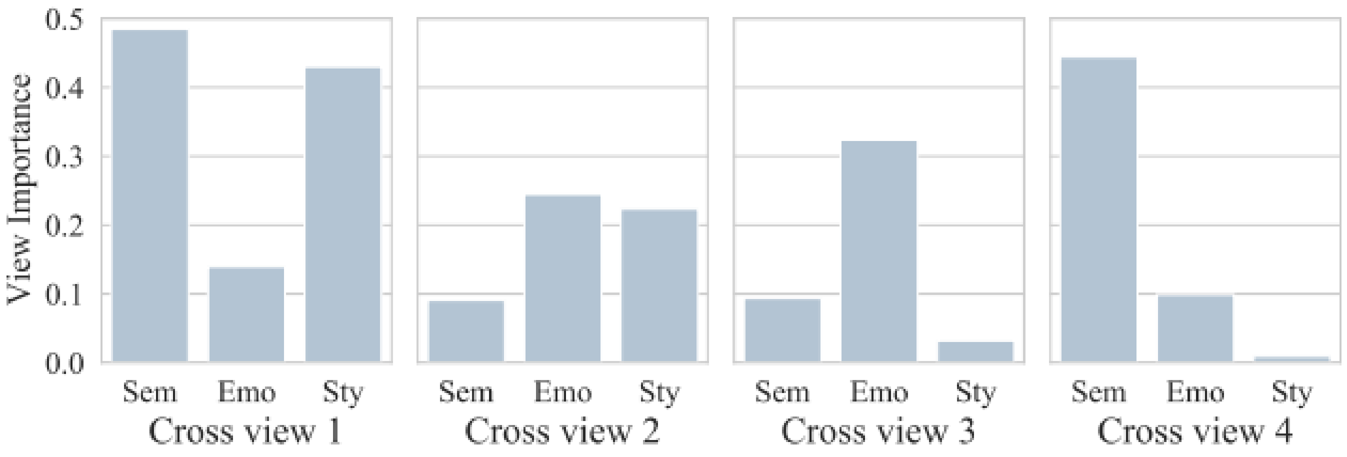


Fig. 5. Each figure indicates importances of different views in a cross-view interaction.

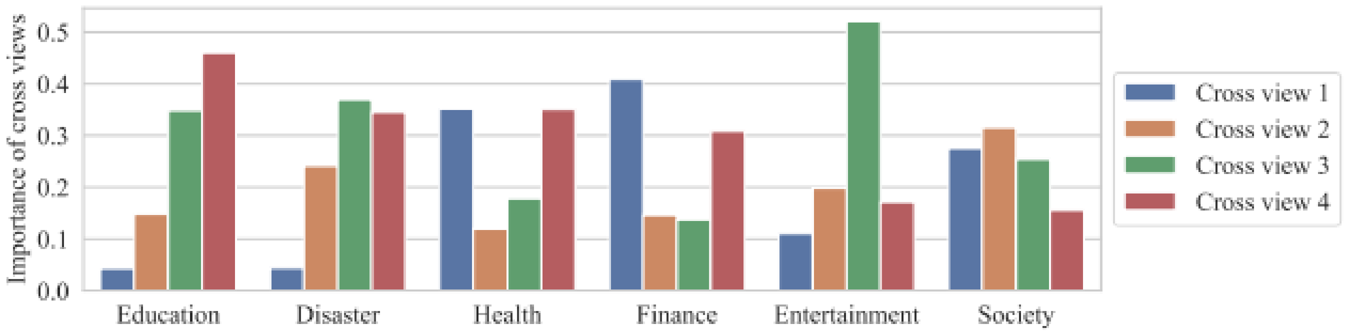


Fig. 6. Various importances of four cross-view interactions for different domains.

Thanks