Detection of False Online Advertisements with DCNN

Hen-Hsen Huang, Yu-Wei Wen and Hsin-Hsi Chen
Department of Computer Science and Information Engineering
National Taiwan University, Taipei, Taiwan
{hhhuang, ywwen}@nlg.csie.ntu.edu.tw; hhchen@ntu.edu.tw

ABSTRACT
In addition to opinion spam, the overstated or unproven information in false advertisements could also mislead customers while making purchasing decisions. A false-advertisement judgement system aims at recognizing and explaining the illegal false advertisements. In this paper, we incorporate the convolutional neural network (CNN) with word embeddings and syntactic features in the system. The recognition experiments show that Dependency-based CNN (DCNN) achieves F-scores of 86.77%, 93.18%, and 87.46% in the cosmetics, food, and drug datasets, respectively. Moreover, the explanation of illegality experiments shows the F-scores of 56.19%, 50.36%, and 62.06% in the three datasets. Our judgement system can contribute to different roles in the online advertising.

Categories and Subject Descriptors
•**Computing methodologies** → Natural language processing
•**Applied computing** → Law, social and behavioral sciences

Keywords
Convolutional neural network; opinion spam detection; overstated advertisement identification

1. INTRODUCTION
As online advertising becomes a major channel for marketing, malicious advertising such as email spam and fake reviews are unavoidably widespread on the Internet [1-3, 6]. In false advertising, a type of malicious advertising, overstated or unproven information is embedded in order to promote the products.

To prevent the customers from being misled by false advertisements, the authorities are involved in advertisement examination. Tang et al. [7] and Tang and Chen [8] addressed the topic of false advertisement detection. Their models recognize the illegal advertisements in the cosmetics and food domains. In their preliminary studies, the Naive Bayes and SVM classifiers are trained with only word features (e.g., unigrams).

In this paper, we propose a robust false-advertisement judgement system that not only detects illegal advertisements, but also explains the reasons of illegality. Convolutional neural network (CNN) classifiers with linguistic features are developed on the cosmetics, food, and drug datasets. Our judgement system can contribute to different roles in the online advertising. Firstly, it can be deployed to aid the authorities screening advertisements widespread on the Internet. Secondly, the electronic commerce can employ it to regulate the product information provided by manufacturers. Thirdly, it can serve as an advertisement checker that highlights the problematic expressions in an advertisement draft.

WWW 2017 Companion, April 3-7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3054233

2. DEVELOPMENT OF DATASETS
The limitation of training data is one of barriers to false advertisement detection. It is impractical to annotate a dataset by crowdsourcing because the determination of false advertisements requires professionals with legal training. Fortunately, the Department of Health, Taipei City Government monthly publishes the cases of illegal advertisements in the government's open data collection [1]. Each case of illegal advertisements consists of a number of fields, including the date of judgement, the product name, the manufacturer, the source of the advertisement, the exact statements expressing false information, and the corresponding act this case violates. We collect all the cases in the period from July 2009 to June 2016, and extract the overstatements and the violated act of each case. The overstatements are specifically annotated at the sentence level. Accordingly, we split the overstatements into sentences and regard them as the "positive" instances for training and testing the false advertisement detector.

The "negative" instances for training our detector are the sentences in legal advertisements from the cosmetics, food, and drug domains. Most existing advertisements are legal ones, so the cosmetics, food, and drug advertisements on the Internet are the materials of our "negative" instances. We crawl the advertisements from various sources including a cosmetic news website, UrCose [2], and electronic commerce websites, UDN Shopping[3] and Yahoo Shopping[4]. We split the advertisements into sentences and regard them as negative (legal) instances. Some sentences appear in both positive set and negative set. We remove them from the negative set since they have been published as false advertisements by the government. The numbers of the false and the legal instances from each domain are shown in Table 1.

Table 1. Statistics of the datasets

Domain	False Ads	Legal Ads	Total
Cosmetics	24,470	115,847	140,317
Food	15,726	11,644	27,370
Drug	2,090	2,273	4,363

3. FALSE ADVERTISEMENT DETECTION
False advertisement detection is a task of two-way sentence classification. We explore the effectiveness of the convolutional neural network (CNN) in this task. Sentence classifier based on CNN achieves improved performances in some NLP tasks [4]. Based on Kim's model, the Dependency-based CNN (DCNN) proposed by Ma et al. [5] further takes the syntactic features such as ancestor and sibling paths into account. We also implement the SVM-based false advertisement detector proposed by Tang and

[1] http://health.gov.taipei/Default.aspx?tabid=295
[2] https://www.urcosme.com
[3] https://shopping.udn.com/mall/Cc1a00.do
[4] https://tw.buy.yahoo.com

Chen [7] as the baseline. The Chinese word segmenter and the parser of Stanford CoreNLP[5] are performed for feature extraction. Five-fold cross-validation is used.

Table 2 summarizes experimental results. For each model in each domain, the performance is measured in accuracy (A), precision (P), recall (R), and F-score (F). Overall, DCNN achieves the highest accuracies and F-scores in all the three domains, and the CNN performs the second. In terms of precision, the baseline model, SVM, is slightly better than CNN and DCNN in all domains. On the other hand, the recalls of SVM are greatly inferior to those of the other two models. In our practical scenario, for example, the detector automatically screens online advertisements, recognizes the suspicious ones, and suggests them to relevant officials for determination. From this viewpoint, the recall is more important.

Table 2. Performances of false advertisement detection

Domain	Model	A (%)	P (%)	R (%)	F (%)
Cosmetics	SVM	93.88	**97.57**	66.58	79.15
	CNN	95.40	89.34	83.59	86.37
	DCNN	**95.55**	90.12	**83.66**	**86.77**
Food	SVM	83.78	**94.27**	76.43	84.42
	CNN	91.92	93.53	92.32	92.92
	DCNN	**92.24**	93.99	**92.39**	**93.18**
Drug	SVM	80.90	**90.99**	66.74	77.00
	CNN	87.92	87.77	**86.89**	87.33
	DCNN	**88.10**	88.37	86.56	**87.46**

4. EXPLANATION OF ILLEGALITY

Our judgement system not only detects illegal advertisements, but also tries to explain the reasons of illegality. The explanation of illegality is more useful for the officials to impose a penalty or for the advertiser to revise the advertisement. As described in Section 2, the act an illegal advertisement article violates is available in the government's open data. Table 3 shows the frequently violated acts in the three domains.

Given a false advertisement, our judgement system suggests the best elaborative act. We regard this task as a sentence classification task so that the CNN and the DCNN classifiers described in Section 3 are trained to predict the most probable act for each false advertisement. The SVM classifier is the baseline. Five-fold cross-validation is performed.

Table 4 shows the performances of explanation of illegality. Macro averaged F-scores are reported. For 2-way classification in cosmetics domain, both CNN and DCNN achieve F-scores around 56%. For 3-way classification in the other two domains, both CNN and DCNN achieve F-scores much better than the baseline model. Compared to false advertisement detection (Section 3), explanation of illegality is a more challenging task. Error analysis reveals some confusion between acts. For instance, both Article 28 of Act Governing Food Sanitation and Article 14 of Health Food Control Act forbid the overstated or unproven information in advertisements, and the difference is that the former applies to general food while the latter applies to health food, a special genre of food products. It is hard to distinguish these two acts without the knowledge of the food products.

Table 3. Frequently violated acts in the datasets

Domain	Violated Act	#
Cosmetics	Law for the Control of Cosmetic Hygiene, Art. 24-1	13,339
	Law for the Control of Cosmetic Hygiene, Art. 24-2	10,069
Food	Act Governing Food Sanitation, Art. 19	9,252
	Act Governing Food Sanitation, Art. 28	3,892
	Health Food Control Act, Art. 14	476
Drug	Pharmaceutical Affairs Law, Art. 65	163
	Pharmaceutical Affairs Law, Art. 66	1,562
	Pharmaceutical Affairs Law, Art. 69	84

[5] http://stanfordnlp.github.io/CoreNLP/

Table 4. Macro F-scores of explanation of illegality

Domain	SVM	CNN	DCNN
Cosmetics	36.30%	56.75%	56.19%
Food	37.72%	49.76%	50.36%
Drug	38.44%	60.32%	62.06%

5. CONCLUSION

In this paper, we propose a false online advertisement judgement system by using the convolutional neural network with syntactic features. Experimental results show that the DCNN models achieve the best F-scores in both detection and explanation tasks in almost all these three domains. This system fits a wide range of applications that are helpful to customers, government departments, e-commerce managers, manufacturers, and advertisers.

6. ACKNOWLEDGMENTS

This research was partially supported by Ministry of Science and Technology, Taiwan, under grants MOST-104-2221-E-002-061-MY3 and MOST-105-2221-E-002-154-MY3.

7. REFERENCES

[1] Yu-Ren Chen and Hsin-Hsi Chen. 2015. Opinion Spammer Detection in Web Forum. In *Proceedings of the 38th Annual ACM SIGIR Conference*. ACM, 759-762.

[2] Yu-Ren Chen and Hsin-Hsi Chen. 2015. Opinion Spam Detection in Web Forum: A Real Case Study. In *Proceedings of 24th International World Wide Web Conference*. ACM, 173-183.

[3] Nitin Jindal and Bing Liu. 2008. Opinion Spam and Analysis. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*. ACM, 219-230.

[4] Yoon Kim. 2014. Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. ACL, 1746–1751.

[5] Mingbo Ma, Liang Huang, Bing Xiang, and Bowen Zhou. 2015. Dependency-based Convolutional Neural Networks for Sentence Embedding. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics*. ACL, 174–179.

[6] Yafeng Ren, Donghong Ji, and Hongbin Zhang. 2014. Positive Unlabeled Learning for Deceptive Reviews Detection. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. ACL, 488–498.

[7] Yi-jie Tang and Hsin-Hsi Chen. 2014. FAdR: A System for Recognizing False Online Advertisements. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. ACL, 103–108.

[8] Yi-jie Tang, Cong-kai Lin, and Hsin-Hsi Chen. 2012. Advertising Legality Recognition. In *Proceedings of the 24th International Conference on Computational Linguistic*. ICCL, 1219–1228.