Aug 10th, 12:00 AM

Use of Emotions in Fake Review Detection

Mehmet Akgul
McMaster University, akgulm@mcmaster.ca

Use of Emotions in Fake Review Detection

Emergent Research Forum (ERF)

Mehmet Akgul
McMaster University
akgulm@mcmaster.ca

Abstract

With the availability of information technologies, the number of online reviews is increasing day by day. As consumers utilize online reviews in their purchasing decisions, they need to know the genuineness of the reviews as non-genuine, i.e. fake, reviews result in both monetary and time losses. Furthermore, businesses also suffer financial loss due to fake reviews and face challenges in retaining consumers' trust. Recent studies show that almost one-third of online reviews are fake, and the consumer spending due to fake online reviews is $152 billion. As a result of its huge impact, it is vital for organizations, especially online review platforms, to mitigate fake reviews. In this study, we concentrated on the review content, and by utilizing text analytics, we proposed utilizing the emotional content of online reviews in fake online review detection. We believe such utilization will enable organizations to increase the efficiency of fake review detection systems.

Keywords

Online Reviews, Fake Reviews, Emotions

Introduction

The volume of e-commerce is increasing each day, and the total amount of e-commerce transactions reached $4.2 trillion in 2021 (Verdon, 2021). While making purchases, consumers also listen to the other consumers' experiences, as 93% of the consumers stated that online reviews affect their purchasing decision (Ismagilova, 2020). With the availability of information technologies, more individuals post online reviews on various platforms: for example, by the end of 2021, the number of reviews on TripAdvisor reached almost 1 billion for close to 8 million businesses (Tripadvisor, 2021).

However, the ease of posting and consuming online reviews doesn't come without a cost, and both businesses and consumers are facing challenges due to fake online reviews. According to Canada Competition Bureau (2017), almost one-third of online reviews are fake, and the consumer spending due to fake online reviews is $152 billion (CHEQ, 2021). Fake online reviews not only result in financial losses but also erode the trust between the consumers and businesses (Vignolo, 2019; Rohrlich, 2020).

As a result of the high cost of fake reviews, it is high importance for organizations and online review platforms to take preventive and corrective active actions to mitigate fake online reviews. Within the extant literature, many data mining methods with various review and review text related features have been proposed to detect fake reviews (Mohawesh et al., 2021). However, the emotional content of reviews is generally ignored in those methods to the best of the authors' knowledge. In this study, we propose to utilize the emotional content of reviews in fake online review detection in the hope of increasing the efficiency of such systems.

Literature Review

Online Reviews

Online reviews are "a type of product information created by individuals based on their personal usage experiences" (Mangold and Smith, 2012). They are one of the resources consumers use to make purchase decisions (Weathers et al., 2015), and they are considered to be more effective than traditional advertising (Z. Zhang et al., 2014).

As a result of its effectiveness, some organizations manipulate online reviews to affect the consumers' decisions (Dellarocas, 2006). Moreover, some organizations provide incentives in exchange for reviews (i.e. incentivized reviews) and some of those reviews are found to have higher star ratings compared to normal reviews (Petrescu et al., 2018). Therefore, due to its impact on consumers and businesses, it is important for organizations, especially online review platforms, to determine the genuineness of online reviews.

Fake Reviews

Fake online reviews are defined as "any positive, neutral, or negative review that is not an actual consumer's honest and impartial opinion and does not reflect a genuine experience of a product, service, or business." (CHEQ, 2021). As suggested, fake reviews are not only positive, but they can also be negative, and businesses are posting fake reviews either for themselves or their competitors (Luca and Zervas, 2016).

Businesses can benefit from the fake reviews as it is found that a one-star increase can result in up to a 9% percent increase in revenue for restaurants (Luca, 2016), and a $250,000 spending on fake reviews may generate more than $5 million in sales (Marciano, 2021). As a result of these monetary benefits, some businesses tend to purchase fake reviews on different markets, as buying 1000 amazon reviews can cost as low as $11000 (Dean, 2021).

While some organizations can benefit from fake reviews, it has side effects on businesses and consumers. Fake reviews not only cost monetary losses but also result in distrust between consumers and businesses, which is hard to measure: 85% of the consumers are suspicious of the truthfulness of online reviews (Vignolo, 2019). As the trust in online reviews erodes, consumers may tend to ignore online reviews resulting in a loss of their effectiveness in the consumer decision-making process. Businesses also suffer from fake reviews: if the consumers feel that posted reviews are suspicious, 28% of the consumers tend not to trust the brand (Rohrlich, 2020), and 40% of consumers tend not to use the products/services provided by those businesses (Vignolo, 2019).

As it has detrimental effects on businesses and consumers, it is important to understand the nature of fake reviews and take preventive and corrective actions. Considering the huge amount of reviews, organizations can utilize some technological tools, such as data mining, as decision support to identify fake reviews and then take corrective actions such as flagging the review as fake to alert the consumers.

Fake review detection

From the data mining perspective, fake review detection is a classification task where the objective is to predict whether a given review is fake or not. In their survey, Mohawesh et al. (2021) identified that researchers used various data mining methods such as decision trees, support vector machines, and neural networks to determine if a review is fake by using various properties of the review and review text such as frequency of words, sentiment, type of word (e.g., noun, adverb) or other textual features obtained using dictionaries such as LIWC. However, to the best of our knowledge, previous studies haven't explored the use of specific emotions in fake review detection.

Emotions in online reviews

Emotion "refers to the psychological processes and cognitive patterns that an individual derives from events or thoughts" (Xu,2020). Plutchik (1980) categorized the emotions into eight primary emotions: anger, fear, joy, sadness, disgust, anticipation, surprise, and trust. He asserted that any emotion could be represented by a combination of these primary emotions.

To vent negative feelings or express positive feelings, consumers write online reviews (Yoo and Gretzel, 2008): they tend to write more about the advantages of products/services when they feel positive, and they tend to complain more when they have negative emotions (Xu, 2020). Previous research identified that consumers could identify emotional words faster compared to non-emotional ones (Guo et al., 2020), and emotions have been used for online review analyses, such as product judgements and review helpfulness. For example, Hu et al. (2014) found that review sentiment significantly affects sales. Other researchers investigated the specific emotions and found that anger, fear and sadness in reviews are associated with review helpfulness (Ren and Hong, 2019). We infer that review helpfulness can also be used as accepting the review and hence believing the review content. Hence we believe that emotions can also play a role in fake review identification. To that end, in this study, we conducted an exploratory analysis to identify if there are differences in the emotions between fake and truthful reviews.

Service types

Depending on information availability and the assessment content, services are categorized into search, experience and credence services (Hsieh et al., 2005). For the search services, consumers are able to obtain information about the service and evaluate them before purchasing (Hsieh et al., 2005). On the other hand, for experience services, consumers can evaluate them only after consumption (Hsieh et al., 2005). Some examples of experience services are hotels and restaurants (H. Zhang et al., 2014). As consumers are making decisions while reading online reviews, they are taking some risks based on their assessments, and it has been identified that the risk associated with experience services is higher than the search services (Zhang et al., 2021). Hence for this study, we selected experience services as our context, and we selected hotel reviews since the data is publicly available.

Methodology

We use the IBM Watson Natural Language Understanding (NLU) online service for emotion detection in online reviews. The NLU service is available through API, and we used the Python environment to connect to the NLU services. NLU is able to detect five emotions from a given text: sadness, joy, fear, disgust and anger, and for each emotion, the service provides a score between [0,1] with the constraint that the sum of the scores of all emotions are 1. For a given text, the NLU service can provide emotions at two levels: the document level and the entity level. While the document-level emotions provide the overall emotion of a given text, entity-level emotions provide emotions for the specific entities detected in the given text. Since we are interested in the overall emotions in fake review detection, we opted to use the document-level emotion calculation for the online reviews.

Preliminary Analyses

The dataset for the Hotel reviews is publicly available and can be downloaded from https://myleott.com/op-spam.html. The dataset consists of 1600 fake and truthful reviews for the 20 most popular Hotels in Chicago. The reviews are divided into two groups depending on review polarity: positive reviews (400 fake and 400 truthful reviews) and negative reviews (400 fake and 400 truthful reviews) (Ott et al., 2011; Ott et al., 2013). For each review, we computed the emotions using the IBM NLU service as stated above. Table 1 presents the mean values for the sadness, joy, fear, disgust and anger emotion scores for all the reviews in the dataset.

	Count	Sadness	Joy	Fear	Disgust	Anger
Fake	800	0.179	0.490	0.068	0.088	0.079
Truthful	800	0.192	0.455	0.075	0.072	0.081

Table 1. Mean values for reviews' emotions for the whole dataset

To compare the differences in the means, we applied a two-sample t-test. We found that except for the anger, there is a significant difference in the means of emotions in fake and truthful reviews at the 5% level and that fake reviews are more positive (higher joy, less sadness and fear value) compared to truthful reviews.

It may be the case that the emotional content of fake and truthful reviews may differ depending on the valence (e.g., positive or negative review) of the review. Table 2 presents the emotion scores for the fake and truthful reviews grouped by the polarity of the reviews. Similar to the first analysis, we applied a two-sample t-test to the reviews. For the negative reviews, we found that only disgust is significantly different between fake and truthful reviews at the 5% level and that fake reviews are more extreme (higher disgust value) compared to truthful reviews. For the positive reviews, except for the disgust, there is a significant difference in the means of all emotions in fake and truthful reviews at the 5% level. Similar to the negative reviews, we observe that fake positive reviews are more extreme (higher joy value) compared to truthful reviews. This behavior (posting more disgust (joy) for negative (positive) reviews) may be explained by the objective of the fake reviews, which is to affect consumers' decisions. Hence to be more effective, fake review writers may post more extreme reviews to emphasize the negative or positive experiences related to the service.

	Count	Sadness	Joy	Fear	Disgust	Anger
Negative - Fake	400	0.231	0.294	0.086	0.121	0.111
Negative - Truthful	400	0.232	0.307	0.091	0.092	0.107
Positive - Fake	400	0.125	0.686	0.049	0.056	0.048
Positive - Truthful	400	0.151	0.603	0.059	0.053	0.056

Table 2. Mean values for reviews' emotions grouped by review valance

We also checked the predictive power of emotions in detecting whether a review is fake or not. By using Rapid Miner software, we applied various data mining algorithms, and we present the accuracy and F-measure results in Table 3. For the analysis, we take the review truth status (e.g., fake or truthful) as the dependent variable and the review emotions (sadness, joy, fear, disgust and anger) as the predictor variables. Our analysis reveals that Gradient Boosted Trees can be a good candidate for fake review prediction as it has the highest accuracy and F measure, i.e. 65%.

Model	Accuracy	F Measure	Model	Accuracy	F Measure
Naive Bayes	0.57	0.60	Decision Tree	0.50	
Generalized Linear Model	0.57	0.51	Random Forest	0.53	0.18
Logistic Regression	0.57	0.51	**Gradient Boosted Trees**	**0.65**	**0.65**
Fast Large Margin	0.57	0.46	Support Vector Machine	0.60	0.56

Table 3. Accuracy and F measures scores of the DM algorithms for fake review detection

Discussion and Conclusion

With the vast amount of information on the internet, on top of the information load, fake reviews bring in additional challenges to consumers. Consumers not only face with financial losses but also time losses; they also impact the trust between the consumers and products/services. In this study, we explored if the emotional context of reviews can be used in the identification of fake reviews. Our preliminary analysis shows that fake reviews tend to be more extreme compared to truthful reviews: for positive reviews, they are more joyful, and for negative reviews, they contain more disgust compared to truthful ones. Our initial analysis also suggests that some data mining algorithms can be used in fake review detection. This paper doesn't claim that emotions are the best candidates for fake review detection; rather, we contend that they can be included in fake review detection systems along with other variables. We believe this study is one of the few studies exploring the use of emotions in fake review identification, and we hope that the use of emotions will enable organizations to increase the efficiency of fake review detection systems.

REFERENCES

Canada Competition Bureau. 2017. *Fake Online Reviews: Too Good to Be True! Canada.Ca*. (https://www.canada.ca/en/competition-bureau/news/2017/03/fake_online_reviewstoogoodtobetrue.html).

C.H.E.Q.2021.*Fake Online Reviews*. (https://f.hubspotusercontent00.net/hubfs/5228455/Research/Fake%20Online%20Reviews%202021.pdf).

Dean, G. 2021. "Websites Are Selling Fake Reviews 'in Bulk' to Amazon Merchants, a Report Found. One Site Offered 1,000 Reviews for $11,000," *Business Insider*. (https://www.businessinsider.com/fake-amazon-reviews-for-sale-buy-merchants-amazons-choice-2021-2).

Dellarocas, C. 2006. "Strategic manipulation of internet opinion forums: Implications for consumers and firms," *Management science* (52:10), pp. 1577–1593.

Guo, J., Wang, X., and Wu, Y. 2020. "Positive Emotion Bias: Role of Emotional Content from Online Customer Reviews in Purchase Decisions," *Journal of Retailing and Consumer Services* (52), p. 101891.

Hsieh, Y. C., Chiu, H. C., and Chiang, M. Y. 2005. "Maintaining a Committed Online Customer: A Study across Search-Experience-Credence Products," *Journal of Retailing* (81:1), pp. 75–82.

Hu, N., Koh, N. S., and Reddy, S. K. 2014. "Ratings Lead You to the Product, Reviews Help You Clinch It? The Mediating Role of Online Review Sentiments on Product Sales," *Decision Support Systems* (57), pp. 42–53.

Ismagilova, E., Slade, E. L., Rana, N. P., and Dwivedi, Y. K. 2020. "The Effect of Electronic Word of Mouth Communications on Intention to Buy: A Meta-Analysis," *Information Systems Frontiers* (22:5), pp. 1203–1226.

Luca, M. 2016. "Reviews, Reputation, and Revenue: The Case of Yelp Com," Harvard Business School NOM Unit Working Paper, pp. 12–016.

Luca, M., and Zervas, G. 2016. "Fake it till you make it: Reputation, competition, and Yelp review fraud," *Management Science* (62:12), pp. 3412–3427.

Mangold, W. G., and Smith, K. T. 2012. "Selling to Millennials with Online Reviews," *Business Horizons* (55:2), pp. 141–153.

Marciano, J. 2021. "Fake Online Reviews Cost $152 Billion a Year. Here's How e-Commerce Sites Can Stop Them," *World Economic Forum*. (https://www.weforum.org/agenda/2021/08/fake-online-reviews-are-a-152-billion-problem-heres-how-to-silence-them/).

Mohawesh, R., Xu, S., Tran, S. N., Ollington, R., Springer, M., Jararweh, Y., and Maqsood, S. 2021. "Fake Reviews Detection: A Survey," *IEEE Access* (9), pp. 65771–65802.

Ott, M., Cardie, C., and Hancock, J. T. 2013. "Negative Deceptive Opinion Spam," in *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 497–501.

Ott, M., Choi, Y., Cardie, C., and Hancock, J. T. 2011. *Finding Deceptive Opinion Spam by Any Stretch of the Imagination*. (https://doi.org/10.48550/ARXIV.1107.4557).

Petrescu, M., O'Leary, K., Goldring, D., and Mrad, S. B. 2018. "Incentivized Reviews: Promising the Moon for a Few Stars," *Journal of Retailing and Consumer Services* (41), pp. 288–295.

Plutchik, R. 1980. *Emotion, a psychoevolutionary synthesis*, New York: Harper & Row.

Ren, G., and Hong, T. 2019. "Examining the Relationship between Specific Negative Emotions and the Perceived Helpfulness of Online Reviews," *Information Processing & Management* (56:4), pp. 1425–1438. (https://doi.org/10.1016/j.ipm.2018.04.003).

Rohrlich, J. 2020. *What Fake Reviews Can Mean for Your Business*. Bazaarvoice. (https://www.bazaarvoice.com/blog/what-fake-reviews-can-mean-for-your-business/).

Tripadvisor. 2021. *Tripadvisor's Year in Review: All the Good in 2021*. (https://ir.tripadvisor.com/news-releases/news-release-details/tripadvisors-year-review-all-good-2021).

Verdon, J. 2021. *Global E-Commerce Sales to Hit $4.2 Trillion as Online Surge Continues, Adobe Reports,* Forbes. (https://www.forbes.com/sites/joanverdon/2021/04/27/global-ecommerce-sales-to-hit-42-trillion-as-online-surge-continues-adobe-reports/?sh=47c636d550fd).

Vignolo, J. 2019. *Fake Reviews Are a Fast-Growing Fraud Concern That Leads to Brand Abandonment*, Sift Blog. (https://blog.sift.com/2019/fake-reviews-are-a-fast-growing-fraud-concern/).

Weathers, D., Swain, S. D., and Grover, V. 2015. "Can Online Product Reviews Be More Helpful? Examining Characteristics of Information Content by Product Type," *Decision Support Systems* (79), pp. 12–23.

Xu, X. 2020. "Examining the Role of Emotion in Online Consumer Reviews of Various Attributes in the Surprise Box Shopping Model," *Decision Support Systems* (136), p. 113344.

Yoo, K. H., and Gretzel, U. 2008. "What Motivates Consumers to Write Online Travel Reviews?," *Information Technology &Tourism* (10:4), pp. 283–296.

Zhang, H., Sun, J., Liu, F., and Knight, J. G. 2014. "Be Rational or Be Emotional: Advertising Appeals, Service Types and Consumer Responses," *European Journal of Marketing*.

Zhang, Z., Wang, X., and Wu, R. 2021. "Is the Devil in the Details? Construal-Level Effects on Perceived Usefulness of Online Reviews for Experience Services," *Electronic Commerce Research and Applications* (46), p. 101033.

Zhang, Z., Zhang, Z., and Law, R. 2014. "Positive and Negative Word of Mouth about Restaurants: Exploring the Asymmetric Impact of the Performance of Attributes," *Asia Pacific Journal of Tourism Research* (19:2), pp. 162–180.