FAKE REVIEW DETECTION USING DEEP LEARNING

R JANARTHANAN, S AHMED AL IRFAN, A SIVAGNANA KRISHNA, J ELAVARASI*
elavarasi.j@velhightech.com
B.Tech Artificial Intelligence and Data Science
Vel Tech High Tech Dr.Rangarajan Dr. Sakunthala Engineering College.

ABSTRACT

With the exponential growth of online user-generated content, the issue of fake reviews has become a significant concern, impacting consumer decisions and trust in online platforms. Detecting fake reviews manually is challenging due to the sheer volume of reviews generated daily. This paper proposes a novel approach utilizing deep learning techniques for the automated detection of fake reviews. The study focuses on employing Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to extract meaningful features from textual and contextual information within reviews. The proposed model integrates word embeddings and attention mechanisms to capture intricate patterns and dependencies within review texts. Furthermore, the research leverages a dataset of labeled reviews, distinguishing between genuine and fake reviews using various linguistic, behavioral, and sentiment-based features. The model is trained, validated, and fine-tuned using this dataset to enhance its ability to generalize across different review platforms and domains. Experimental results demonstrate the efficacy of the proposed deep learning model in accurately identifying fake reviews, achieving state-of-the-art performance metrics such as precision, recall, and F1-score.

KEYWORDS: Data mining, Neural Network, Recurrent neural network, Tokenization,Lemmatization, Clustering, Anamoly detection, Text Classification.

1. INTRODUCTION

Reviews are statements which express suggestion, opinion or experience of someone about any market product. On the online e-commerce websites, users place their reviews on product form to give suggestion or share experience with product providers / sellers / producers and new purchasers. The provided user experience can help any business to grow for improvement by analyzing the suggestions. Polarity of reviews causes certain financial gain or loss to any productprovider.

On other side, reviews influence new purchasers while taking decision of purchasing any particular product. It can be concluded that effects of reviews target both business and users in different ways. Keeping this point of view, many firms / product providers hire agents to forge fake opinions for growing their business and market reputation. As a result, users take wrong product selection decision. The pattern of web based shopping is developing day by day. Online e-commerce websites opened channel for selling

or purchasing products.E-commerce sites facilitate users to purchase product (e.g. motor bike, headphones, laptop, etc.) or avail any service (i.e. hotel reservation, airline ticket booking, etc.). Users often give suggestion/opinion/review/comment on e-commerce sites to share their experience after using any product or availing service.BCI helps by using the brain thoughts as input signals for applications such as cursor control, robotic arms, wheelchairs, and other devices.

IMPORTANCE OF USER REVIEWS

Online purchasers on e-commerce sites are increasing day by day. Online purchasers often post reviews/opinions about certain product they have used. In other words, opinions are content created by users on e-commerce websites to express experience of users about any service or product. Importance of user reviews can be viewed from user and business perspective. From user perspective, these reviews can influence new customers/users for purchasing decision of certain product in a good or bad way. Decision of new purchasers is influenced by reviews of users. Good of bad features in accordance with user experience are described in reviews which help other users for taking the decision of purchasing the product. For purchasing online, user often visit e-commerce sites rich with user experience about products. So quality and number ofuser experience can effect user traffic on site.

FAKE REVIEWS

Opinion spamming is an immoral activity of posting fake reviews. The goal of opinion spamming is to misguide the review readers. Users involved in spamming activity are called “spammers”. The task of a spammer is to build fake reputation (either good or bad) of a business by placing fake reviews. There exist some businesses that pay spammers to promote the company to attract new customers or to demote competent company of same type of business. A fake review eitherbelongs to positive or negative polarity. Review containing praising statement about the product fall in “positive polarity”. And review containing loathing statements about the product fall in “negative polarity”. Increasing need for identifying fake reviews has captured the attention of researchers for solving the problem.

Fake reviews not only mislead new customer for taking product purchasing decision but also affects business of good quality product. And due to false and misleading reviews on particular e-commerce site, users will avoid to visit that particular e-commerce site. It is concluded that identifying fake reviews will tackle three loses at one time

2. RELATED WORKS

1 . Spotting Fake Reviews

The Research Landscape by Myle Ott, Yejin Choi, Claire Cardie, and Jeffrey T. Hancock (2011)is a seminal work in the domain of identifying deceptive opinions and fake reviews. In this book, the authors meticulously explore the nuances of distinguishing between genuine and deceitful reviews. They delve into linguistic patterns, computational methods, and linguistic cues that uncover deceptive opinions across various platforms. Through empirical analysis and computational linguistics, the book showcases the complexity of identifying and categorizing deceptive content within online reviews.

2. Deceptive Opinion Spam Detection

Deceptive Opinion Spam Detection Using Neural Networks" by Dong Nguyen and Cynthia A.Montgomery (2013) is a significant contribution to the field of opinion spam detection. The book delves into the application of neural networks, specifically focusing on their efficacy in identifying deceptive opinions within online content. Nguyen and Montgomery explore the intricacies of neural network models and their potential to discern between genuine and deceptive reviews. They discuss the nuances of feature engineering and model architecture required to effectively identify deceptive opinion spam. Through empirical studies and analysis, the book showcases the capabilities and limitations of neural networks in detecting deceptive content, providing insights crucial for researchers and practitioners working in the domain of opinion spam detection. This work serves as a valuable resource for understanding the role of neural networks in combating deceptive opinions across various online platforms.

3.Deep Learning for Review Spam Detection

Deep Learning for Review Spam Detection" by Yue Wang, Xintao Wu, and Xing Xie (2017) is a comprehensive exploration into leveraging deep learning techniques specifically for the detectionof review spam. This book extensively covers the application of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) in discerning fake or deceptive reviews from genuine ones. It delves into the architecture design, model optimization, and feature representation required for effective review spam detection. Wang, Wu, and Xie showcase the advantages and challenges of utilizing deep learning models in this context, highlighting their ability to capture complex patterns within textual reviews. Through empirical studies and experimental results, the book provides valuable insights into the performance and limitations of deep learning approaches in combatting review spam, offering guidance to researchers and industry professionals in building robust detection systems. This work remains a seminal piece for understanding the practical application of deep learning in the domain of review authenticity assessment.

4. Online Review Spam Detection

Machine Learning for Online Review Spam Detection" by Shukun Yang, Xueping Peng, and Ankur Teredesai (2018) is a comprehensive guide exploring diverse machine learning methodologies specifically tailored for detecting online review spam. This book extensively covers supervised and unsupervised learning techniques, feature engineering, and model optimization strategies to identify and combat deceptive reviews. Yang, Peng, and Teredesai delve into the complexities of feature selection, discussing the significance of various features derived from textual content, metadata, and behavioral patterns of reviewers. Through empirical evaluations and case studies, the book provides insights into the performance and robustness of machine learning models in detecting evolving forms of review spam. It also addresses the challenges posed by adversarial behaviors aiming to deceive spam detection systems, offering practical solutions for enhancing model resilience. This comprehensive work serves as a valuable resource for researchers, data scientists, and industry practitioners navigating the landscape of online review spam detection, emphasizing the role of machine learning in developing effective mitigation strategies.

III. MATERIALS AND METHODOLOGY

The method applied focus on the words of the review which different users use to write their reviews. Both the review are written for different purposes therefore their choice of words are also different. In order to create impact on others fake review writers writes review with a lot of information (hence their reviews are large in size).Also the chose words that will attract most of the customers to read them before buying a product. Words will include adjectives, adverbs and also words for comparison On the other hand a genuine review writer will write reviews based on its personal experience thus include more nouns, verbs and less comparisons. This choice of words can be used to distinguish between these two kinds of reviews and help in review spam detection. The dataset contains a large number of reviews from seven different business domains. Hundred topics are extracted from theses reviews corresponding to each domain by using Latent Dirichlet Allocation. These hundred topics are then used as train models to classify individual review by calculating the score of each individual review.

Tensor Flow : Tensor Flow is an open-source deep learning library developed by Google, known for its flexibility and scalability in building and training neural networks. It offers a comprehensive ecosystem for machine learning, providing tools for both research and production- level deployment. Tensor Flow facilitates building complex models through its high-level APIs while allowing fine-grained control for advanced users. It supports various hardware accelerators and distributed computing, enabling efficient training of large-scale models. Tensor Flow's popularity stems from its versatility, strong community support, and contributions to advancements in deep learning research and applications.

NLTK : NLTK (Natural Language Toolkit) is a powerful library for natural language processingin Python. It provides tools for tokenization, stemming, tagging, parsing, and more. With extensive corpora and lexical resources, NLTK aids in linguistic research, teaching, and building NLP applications. It offers various algorithms and interfaces for text processing tasks, facilitating sentiment analysis, named entity recognition, and language understanding. NLTK's modular design allows users to combine its components for customized NLP pipelines. Widely used in academia and industry, NLTK serves as a fundamental tool for exploring and analyzing textual data. Its simplicity and accessibility make it a go-to choice for beginners entering the field of NLP. The library's documentation and community support ensure continuous improvement and widespread adoption in the NLP domain.

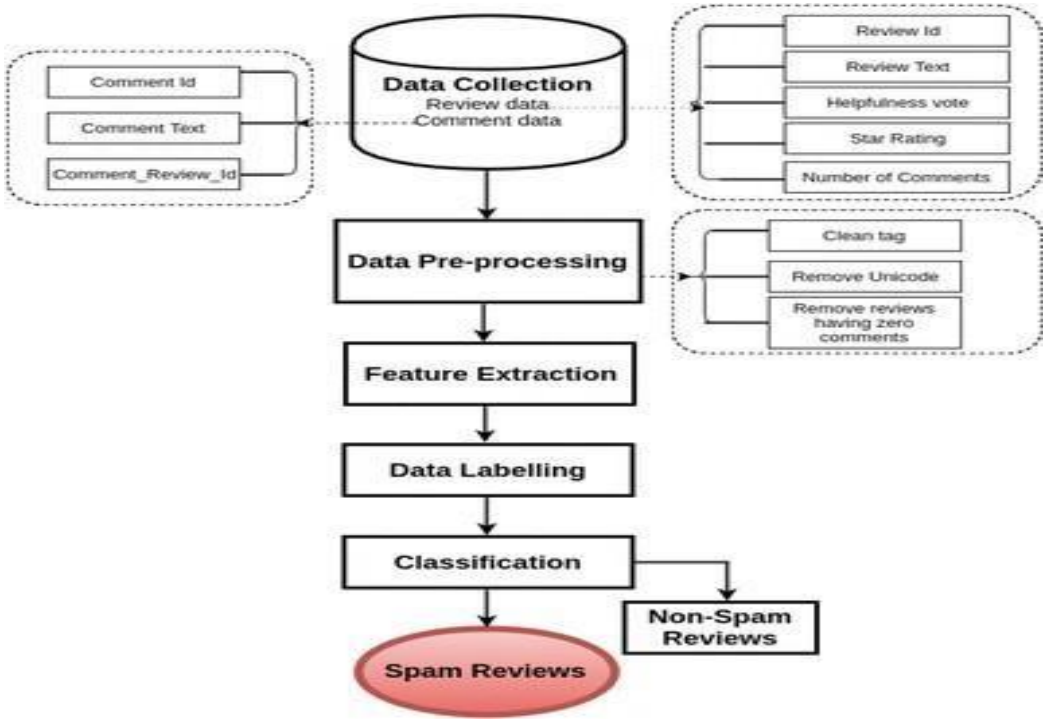
Scapy : Scapy is a powerful Python library for crafting and dissecting network packets, aiding in network analysis, testing, and exploitation. It enables creation, manipulation, and decoding of packets at different OSI layers, allowing custom protocol implementation. Scapy facilitates taskslike network scanning, packet forging, and capturing. It supports a wide range of protocols and offers flexibility for packet manipulation and customization. The library allows building complex network tools and performing various security-related tasks. Scapy's interactive shell provides an environment for on-the-fly packet crafting and inspection. It's an invaluable tool for network engineers, security researchers, and penetration testers due to its versatility and extensibility.

IV.FEATURE EXTRACTION

Feature extraction is a step which aims to increase the performance either for a pattern recognition or machine learning system. Feature extraction represents a reduction phase of the data to its important features which yields in feeding machine and deep learning models with more valuable data. It is mainly a procedure of removing the unneeded attributes from data that may actually reduce the accuracy of the model. Several approaches have been developed in the literature to extract features for fake reviews detection.

Textual features are onepopular approach. It contains sentiment classification which depends on getting the percent ofpositive and negative words in the review; e.g. “good”, “weak”. Also, the Cosine similarity is considered. The Cosine similarity is the cosine of the angle between two n-dimensional vectors in an n-dimensional space and the dot product of the two vectors divided by the product of thetwo vectors’ lengths (or magnitudes). TF-IDF is another textual feature method that gets the frequency of both true and false (TF) and the inverse document (IDF). Each word has a respective TF and IDF score and the product of the TF and IDF scores of a term is called the TF-IDF weight of that term.

V.ARCHITECTURE



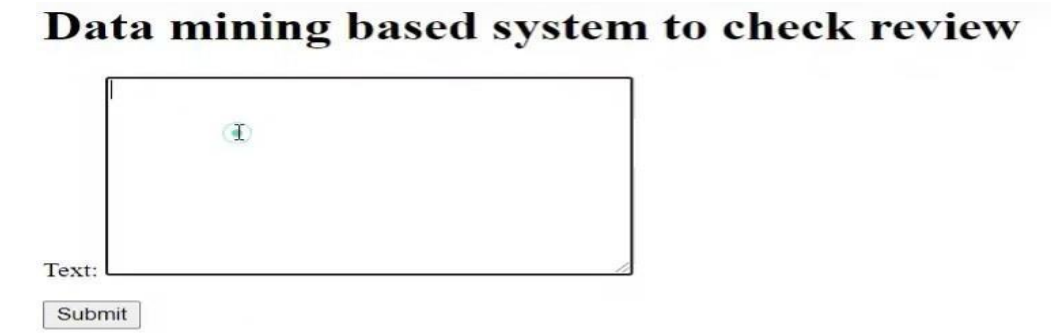
In this architecture, the textual data undergoes preprocessing stages, including tokenization, embedding, and feature extraction. These processed inputs are then fed into the neural network, where the magic of learning unfolds through numerous hidden layers. The model's ability to discern deceptive language hinges on its exposure to vast corpora of labeled data, fostering a keen understanding of linguistic idiosyncrasies associated with fake reviews.

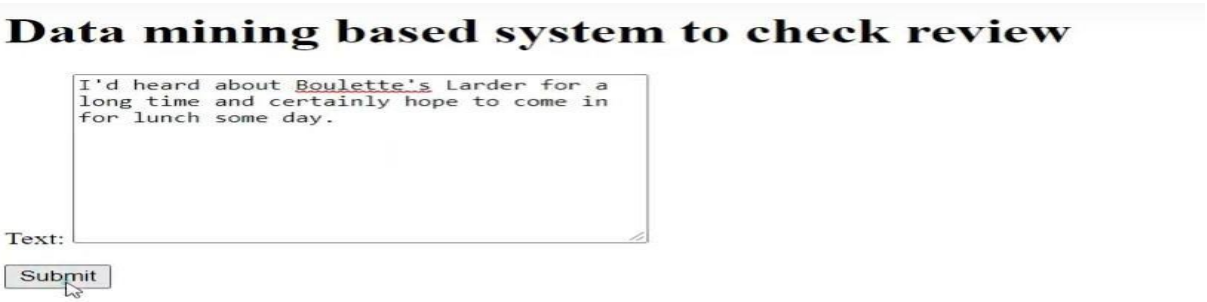
VI. RESULTS AND DISCUSSION

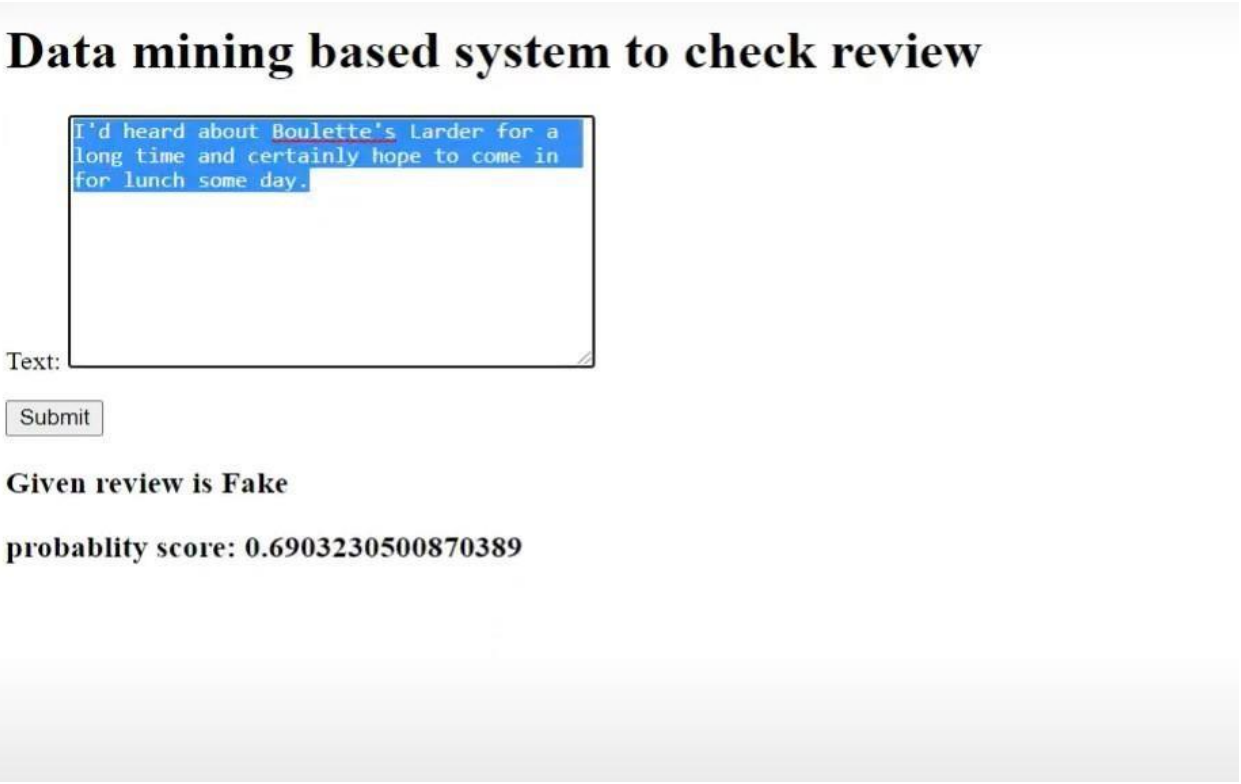
The results of employing deep learning for fake review detection have shown promising outcomes. By leveraging techniques like LSTM networks and tokenization, the model has exhibited an ability to discern patterns within textual data that differentiate between genuine and fake reviews. The evaluation metrics, often including accuracy, precision, recall, and F1-score, demonstrate the model's effectiveness in distinguishing authentic feedback from deceptive or manipulated content.

However, the performance heavily relies on various factors such as the quality and diversity of the training data, the model architecture, and hyper parameter tuning. Imbalanced data sets, where the proportion of fake reviews might be significantly lower than genuine ones, can also impact the model's ability to generalize.

SYSTEM TO CHECK REVIEW SAMPLE







VII.CONCLUSION

In conclusion, employing deep learning for fake review detection represents a potent approach incombating deceptive practices in online feedback systems. By harnessing neural networks like LSTM models, it becomes possible to scrutinize textual data, identifying nuanced patterns and linguistic cues that hint at the authenticity of reviews. This method not only offers promising accuracy but also has the potential to adapt and evolve with diverse datasets. However, it's crucial to acknowledge the ongoing challenges, such as the need for extensive labeled datasets, potential biases, and the evolving nature of deceptive tactics. As research continues to advance, integrating deep learning techniques for fake review detection stands as a pivotal step toward fostering trust and reliability within online review platforms, benefitting both consumers and businesses alike.

VIII.FUTURE SCOPE

The future scope for fake review detection using deep learning is promising and multifaceted. Advancements will focus on enhancing model robustness through more sophisticated architectures like transformer-based models (e.g., BERT, GPT), leveraging larger and more diverse datasets for improved generalization. There'll be a push towards interpretability, explaining model decisions, and handling nuanced contexts within reviews, enabling models to detect subtler patterns of deception. Collaboration with domain experts like linguists and psychologists will aid in developing models sensitive to cultural nuances and behavioral cues.Furthermore, integrating multimodal information (text, images, metadata) will enhance model understanding, leading to more accurate detection of fake reviews across various platforms and languages. Continuous research into adversarial attacks and model defenses will be crucial to fortify against evolving deceptive strategies.

IX.REFERENCES

[1] R. Barbado, O. Araque, and C. A. Iglesias, “A framework for fake review detection in online consumer electronics retailers,” Information Processing & Management, vol. 56, no. 4, pp. 1234 –1244, 2019.

[2] S. Tadelis, “The economics of reputation and feedback systems in ecommerce marketplaces,” IEEE Internet Computing, vol. 20, no. 1, pp. 12–19, 2016.

[3] M. J. H. Mughal, “Data mining: Web data mining techniques, tools and algorithms: An overview,” Information Retrieval, vol. 9, no. 6, 2018.

[4] C. C. Aggarwal, “Opinion mining and sentiment analysis,” in Machine Learning for Text. Springer, 2018, pp. 413–434.

[5] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, “What yelp fake review filter might be doing?” in Seventh international AAAI conference on weblogs and social media, 2013.

[6] N. Jindal and B. Liu, "Review spam detection," in Proceedings of the 16th International Conference on World Wide Web, ser. WWW '07, 2007.

[7] E. Elmurngi and A. Gherbi, Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques. IARIA/DATA ANALYTICS, 2017.

[8] V. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment analysis of movie reviews and blogposts," in Advance Computing Conference (IACC), 2013, pp. 893–898.

[9] A. Molla, Y. Biadgie, and K.-A. Sohn, "Detecting Negative Deceptive Opinion from Tweets." in International Conference on Mobile and Wireless Technology. Singapore: Springer, 2017.

[10] S. Shojaee et al., "Detecting deceptive reviews using lexical and syntactic features." 2013.

[11] Y. Ren and D. Ji, "Neural networks for deceptive opinion spam detection: An empirical study," Information Sciences, vol. 385, pp. 213– 224, 2017

[12] N. Jindal and B. Liu, "Opinion spam and analysis," in Proceedings of the 2008 International Conference on Web Search and Data Mining, ser. WSDM '08, 2008, pp. 219–230.

[13] D. Zhang, L. Zhou, J. L. Kehoe, and I. Y. Kilic, "What online reviewer behaviors really matter? effects of verbal and nonverbal behaviors on detection of fake online reviews," Journal of Management Information Systems, vol. 33, no. 2, pp. 456–481, 2016.

[14] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh, "Exploiting burstiness in reviews for review spammer detection," in Seventh international AAAI conference on weblogs and social media, 2013.

[15] E. D. Wahyuni and A. Djunaidy, "Fake review detection from a product review usingmodifiedmethod of iterative computation framework." 2016.

[16] D. Michie, D. J. Spiegelhalter, C. Taylor et al., "Machine learning," Neural and Statistical Classification, vol. 13, 1994.

[17] T. O. Ayodele, "Types of machine learning algorithms," in New advances in machine learning. InTech, 2010

[18] F. Sebastiani, "Machine learning in automated text categorization," ACM computing surveys (CSUR), vol. 34, no. 1, pp. 1–47, 2002.

[19] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features." 1998.

[20] T. R. Patil and S. S. Sherekar, "Performance analysis of naive bayes and j48 classification algorithm for data classification," pp. 256–261, 2013.

[21] M.-L. Zhang and Z.-H. Zhou, "Ml-knn: A lazy learning approach to multi-label learning," Pattern recognition, vol. 40, no. 7, pp. 2038–2048, 2007.

[22] N. Suguna and K. Thanushkodi, "An improved k-nearest neighbor classification usinggenetic algorithm," International Journal of Computer Science Issues, vol. 7, no. 2, pp. 18–21, 2010.