A Hybrid Roberta-Large and Vader-Based Deep Learning Framework for Fake Reviews Detection in Amazon Dataset

Zaid Z. Abdulshaheed and Ahmed J. Obaid
Faculty of Computer Science and Mathematics, University of Kufa, 54001 Najaf, Iraq
zaidz.alrwaziq@student.uokufa.edu.iq, ahmedj.aljanaby@uokufa.edu.iq

Keywords: Fake Review Detection, Amazon Reviews, RoBERTa-Large, VADER, BiGRU, Deep Learning, Sentiment Analysis.

Abstract: Because online shopping has grown so quickly, what people say about products in online reviews now has a huge impact on what people buy. Unfortunately, fake reviews are a big issue, as they can trick shoppers, harm a company's good name, and make the market unfairly skewed. This research presents a way to find these fake reviews in Amazon reviews specifically, and it uses both the words in the review plus how positive or negative the review feels. The system works by merging RoBERTa-Large (which understands the context of words) with VADER (which measures emotional tone), and then using a Transformer encoder, a BiGRU layer, a way of focusing on important information (an attention mechanism), and finally, layers of fully connected classifiers. Amazon reviews were used for this because of their sheer number, the fact they come from the actual world, and how often they're used by others studying fake review spotting. Testing showed the RoBERTa-Large, VADER and BiGRU combination did very well; The proposed model achieved 95.45% accuracy and 0.9898 ROC-AUC across roughly 4 million Amazon reviews. What this shows is that combining a deep understanding of what words mean in their context with the positivity or negativity expressed in the review improves finding fake reviews, and gives a reliable method for looking at lots of reviews on big online shops.

1 INTRODUCTION

E-commerce has grown extremely quickly lately, and as a result, how businesses do things has changed a lot from the old ways to being done on the internet. Online shopping sites let people look at items, read what other buyers think, and then decide what to buy with a better understanding of things. What people say in their reviews of products are hugely important for what customers do; they give you an idea of how good the product is, what it's like to use, and have a big effect on whether or not someone buys it [1], [2] However, even though reviews are so important, they aren't always what they seem. Because people are depending on customers to create the content, lots of fake reviews are appearing, designed to trick you. These can be complimentary to push certain items, or bad to harm companies that are competing. This kind of dishonesty lowers how much people trust online sites and has a bad effect on both the people buying and the businesses [3]-[5]. Finding fake reviews is now a central difficulty for those in data mining, natural language processing (NLP), and machine learning. Older ways of finding them often don't work when people sending out the fake ones get clever, blending in fake with real opinions or acting like normal customers. It's still difficult and complex to locate fake reviews within massive amounts of information [6]-[8]. The set of Amazon reviews is often used to test techniques for identifying them as it's a massive, actual-world data source. It contains millions of opinions on a huge range of products and a good deal of thorough description, details of use and user experience. Because of this, it's a good way to assess how well 'systems' for detection perform in normal circumstances. This project tries to deal with the problem of fake reviews by using a blend of looking at the meaning of the text and what emotions are in it [5], [9]-[11]. This project tries to deal with the problem of fake reviews by using a blend of looking at the meaning of the text and what emotions are in it. The research employs complicated deep learning, including aspects which assess positivity or negativity, to hopefully make identifying misleading reviews both more accurate and consistent. The performance of this new system will be examined using Amazon review data to show how it operates in a realistic situation.

2 RELATED WORK

Loads of people are now focused on spotting fake reviews, and for good reason: they really affect how we shop online and whether we trust what we read. Lots of different ways to do this have been suggested in the past, from older style machine learning, to deep learning, to methods using transformers, looking at the feeling in the writing (sentiment analysis), and combinations of these. This part of the work looks over those previous studies to find out what the main methods are, what each does well, where they fall short, and how the work I'm presenting in this thesis will make detecting fake reviews better. Sikarwar and Tiwari found in 2020 that a Linear SVC model correctly identified whether Amazon reviews were positive or negative 91% of the time. Nguyen in 2023 got around 87.9% accuracy with TF-IDF and SVC together. What Nguyen and Gupta and Sharma (2023) discovered, [12], [13] is that Linear SVC is a good way to do sentiment analysis on Amazon reviews. After that, researchers started to use more complex "deep learning" methods, in order to get at, and pinpoint, finer details within the review text. Hajek et al. (2020) got 89.8% accuracy and an AUC of 96.5 by mixing n-gram characteristics, Word2Vec embeddings, and information about emotional tone with CNNs and DFFNNs. Also, Li et al. (2023) found that LSTMs did a better job than older SVM methods, hitting roughly 91% accuracy [14], [15] Newer research has looked at techniques built around transformers, or which blend transformer approaches with other methods. For instance, Salminen et al. (2022) introduced fake RoBERTa and got 96.6% of questions right (as in, were 96.6% accurate) [16]. He et al. (2024) used characteristics of the network itself to get 86% accuracy and a 93% Area Under the Curve (AUC), [17]. Hashmi and Yayilgan (2024) merged Fast Text with a group of multiple models (an ensemble), and that got 93 to 94% accuracy [18]. Finally, Geetha et al. (2025) improved on a DeBERTa model and managed to achieve up to 98% accuracy, along with a 97% AUC [19]. The following Table 1 summarizes the research works completed on the Amazon Dataset and their final results.

3 RESEACH GAP

Past research has shown some good progress in finding fake reviews, but there are still problems. Older machine learning methods, for example Support Vector Machines with TF-IDF or document-term matrices, rely heavily on textual features people have specifically designed. These don't really grasp the more subtle meanings within the review's wording [12], [13], [20]. Deep learning techniques like CNNs, DFFNNs and LSTMs have gotten better at identifying fakes by picking up on complicated patterns in the text. Yet a lot of these systems still concentrate on the text itself, and don't use a review's positive or negative feeling, or the emotional coloring, as extra ways to spot a dishonest review [14], [15]. RoBERTa and DeBERTa, both built on the transformer idea, are very good at spotting fake reviews.

Table 1: Summary of related research works based on amazon dataset.

T	Ref. No	Year of Study	Model / Method Name	Dataset Version	No. of Samples	Accuracy Result	
						ACC	AUC
1	[12]	2020	Linear SVC	Amazon Reviews	~	91%	91%
2	[15]	2020	CNN/DFFNN	Amazon + TripAdvisor	~	89.8%	96.5%
3	[16]	2022	fakeRoBERTa	Amazon	(40K)	96.6%	96.6%
4	[13]	2023	SVC+TF-IDF	Amazon + TripAdvisor	~	87.9%	~
5	[20]	2023	Linear SVC	Amazon Reviews	~	91%	90%
6	[14]	2023	LSTM	Amazon	(1M)	91%	~
7	[17]	2024	Random Forest	Amazon Product	~	86%	93%
8	[18]	2024	FastXCatStack	Amazon	(150K)	94%	~
9	[21]	2024	SVM+TF-IDF	Amazon	(40K)	93%	~
10	[19]	2025	MBO-DeBERTa	Amazon	(21K)	98%	97%
11	~	2026	RoBERTa-Large +VADER +BiGRU	Amazon	4M	0.9545	0.9898

However, finding them is still hard, as fakes are commonly written much like actual, honest opinions. To cover their tracks, those posting fake reviews might sprinkle them among legitimate ones, or even act like typical customers, and this makes it harder to tell what's what. Plus, detecting fake reviews in the mess of information you get from the real world is complicated by things like inaccurate tags on reviews, having far more real than fake ones (or vice versa), spammers constantly changing their techniques, and genuine and deceptive reviews
using a lot of the same kinds of wording and characteristics [22]-[24].

We still need a really strong way to mix understanding what words mean in their situation with figuring out how people feel about something. This project fills that need by suggesting a combined deep learning system. This method looks at what the review actually says and also whether it's positive, negative, or neutral. When we put together the meaning of the words with the reviewer's opinion, the system is much better at finding fake Amazon reviews; it's both more you can count on and does the job well.

4 METHODOLOGY

To find fake reviews, this system does everything in order. First it gathers and gets the review data ready, then it pulls out important characteristics (features), builds a system to work with those characteristics (a model), and finally sees how well the system does. Essentially, it's designed to spot deceptive Amazon reviews by looking at what the review says and also figuring out if it's saying something good, bad, or somewhere in between. The process has these parts: we get the data ready, clean up the text, do a VADER analysis to understand the sentiment, get RoBERTa-Large embeddings (a way of representing the text's meaning), combine these pieces of information, use a combined deep learning method for categorizing, and ultimately evaluate how good the model is.

4.1 Deep learning

is now one of the best ways to work with natural language, as it'tches up on meaningful patterns from lots of text by itself. Older, standard machine learning needed people to create the important qualities (like TF-IDF or Bag-of-Words) but deep learning models are able to get at the more complicated meanings and how things are situated within the text. This is really useful for finding fake reviews because fake and real ones can be almost the same in the way they are written and organized[25].

4.1.1 A Bidirectional Gated Recurrent Unit (BiGRU)

Is a deep learning model that repeats as it processes things in order, like text. A standard GRU is designed to understand how words relate to each other, but is less to compute than LSTM. BiGRU takes this further by reading the sequence of words in both directions, from beginning to end and end to beginning. This lets the model understand what came before and what comes after a word in a sentence. That's helpful for finding hidden things in the text of a review. and can be used for identifying hidden patterns in review text[26].

4.2 Sentiment Analysis

Is another key method for this work, as fake reviews will often have over the top feelings expressed or strange patterns in whether they're positive or negative? VADER (Valence Aware Dictionary and sEntiment Reasoner) is a sentiment analysis tool using word lists and rules, and it's used a lot with text that people have written. It gives scores showing if something is positive, negative, neutral or balanced emotionally. We use VADER's combined score in this research because it gives a single number for the overall positive or negative feeling of a review.

4.2.1 RoBERTa-Large Embedding

RoBERTa-Large creates representations of words that are tied to how they're used in a sentence. Essentially, a word's meaning is understood from the words around it. This is crucial for categorizing text, as a single word can mean different things in different situations. RoBERTa-Large, unlike older ways of representing words with a single meaning, can find much more complex connections between meanings and create a more detailed picture of the text. Because of this, it works well for looking at Amazon reviews and getting lots of specific detail about the context [27]. We chose each of the techniques for a different type of information it adds to our model. Deep learning is the overall method for the system to learn, BiGRU understands the order of words in a review, VADER determines if the review is positive or negative, and RoBERTa-Large gets at the context and meaning of words. Using these parts together gives

the system a much better basis to accurately identify fake reviews.

4.3 Dataset

Many researchers in sentiment analysis and spotting fake reviews use the Amazon 2013 review dataset, and this study is no exception. It's made up of reviews customers left for items on Amazon [28], [29]. Each review has three parts: a Label (what type of review it is), a Title (a quick summary of what the customer thought), and the Text (the entire review itself). With 4 million reviews in total, it's quite large and therefore a good size for building and testing machine learning and deep learning programs. On average, reviews are around 74.17 words and 405.14 characters long, meaning there's a good amount of writing in each one to actually analyze. The following Figure 1 shows the Word Cloud representation in amazon dataset.

Figure 1: Word cloud representation of amazon dataset.

5 PROPOSED WORKFLOW

Here's how our new system for spotting fake reviews works. It's intended to get better at correctly identifying real or fake Amazon reviews and does this by combining the understanding of words in context from RoBERTa-Large with the emotional tone measurements from VADER. The system we're suggesting has several key steps. First it gets the review data ready, then pulls out the important characteristics (features) of those reviews, combines those characteristics, uses them to teach a 'learning' program, and finally checks how well that program does. Figure 2 shows how the whole system is put together and how all these steps link up to identify fake reviews.

This system is meant to find fake reviews on Amazon by using a combined deep learning method. Its core concept is to bring together two kinds of information: what the review says in its context and the emotional leaning of the review. Understanding the meaning of the review text is helped by contextual This system is meant to find fake reviews on Amazon by using a combined deep learning method. Its core concept is to bring together two kinds of information: what the review says in its context and the emotional leaning of the review. Understanding the meaning of the review text is helped by contextual features, and sentiment features show the emotional tone. When both of these are used together, each review gets a much better description and fake reviews are more easily identified. First, the system gets the review text ready. It puts the review's Title and Text together into one section, called full_text. This is significant as the title is a quick take on what the user thinks, and the text has all the details. Then this combined review text has its issues cleared up: it's cleaned, missing parts are handled, and unnecessary symbols are Figure 3.

Following this preparatio n, the system gets two main types of characteristics. FiIt uses RoBE RTa-Large to create contextual embeddings from the review text - RoBERTa-Large understands the meaning of words as they'and creates a 1024-number 'embedding vector' for each review Figure 3 shows RoBERTa embedding extraction.

It also uses the VADER sentiment analyzer to find a single, combined sentiment score for each review, which represents its overall positive or negative feeling Figure 4 shows VADER sentiment score.

These RoBERTa embeddings and VADER scores are then joined to make a combined, 'hybrid' way of representing each review. This combination is fed to the system's deep learning classifier. This classifier is made of a Transformer encoder layer, a Bidirectional GRU layer, and something called an attention mechanism. The Transformer encoder finds how the different characteristics relate to each other, the BiGRU learns patterns in the sequence of words going both forwards and backwards (Table 2).

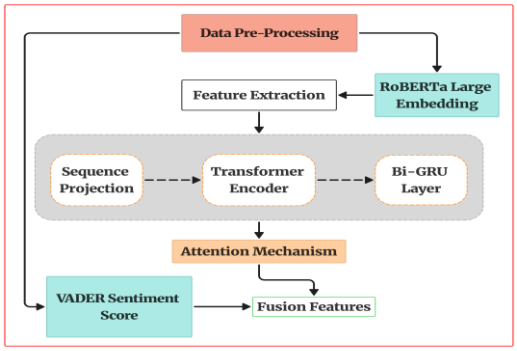

Figure 2: General proposed system.

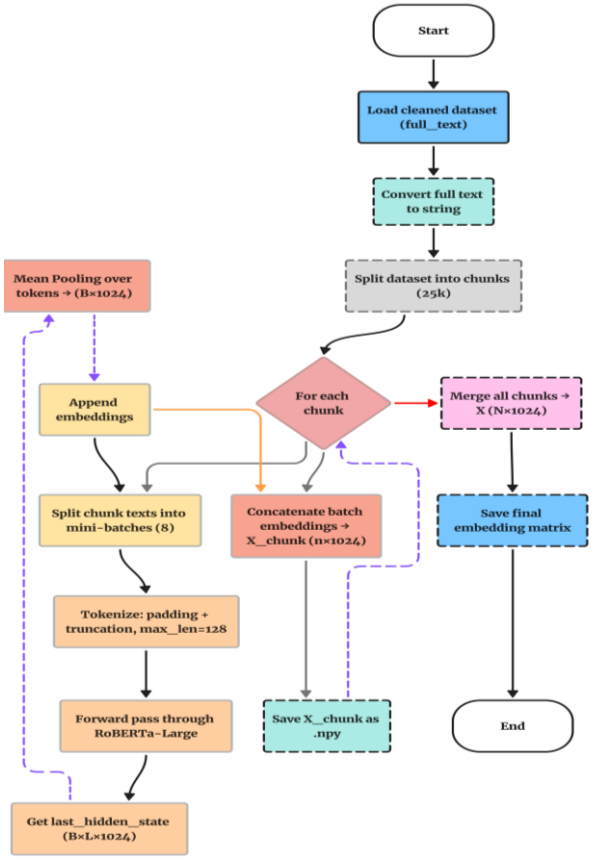

Figure 3: RoBERTA embedding extraction.

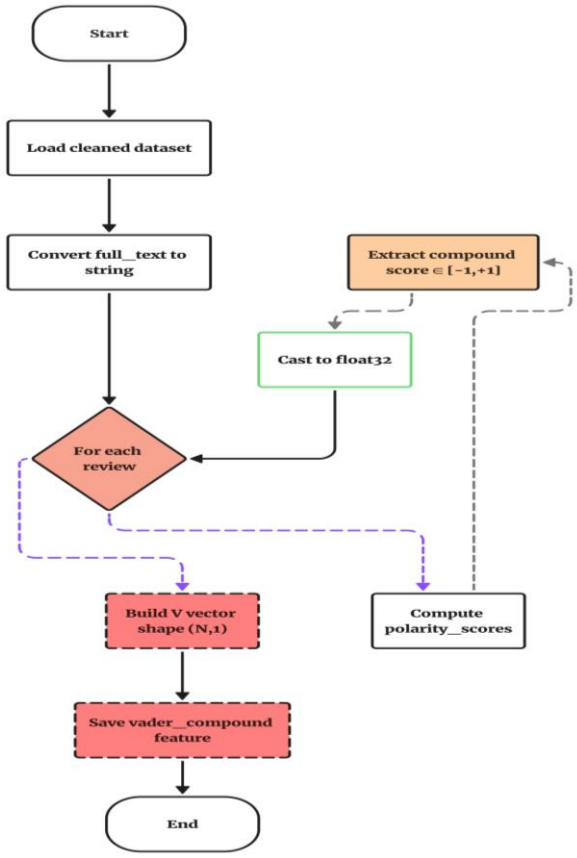

Figure 4: VADER sentiment score.

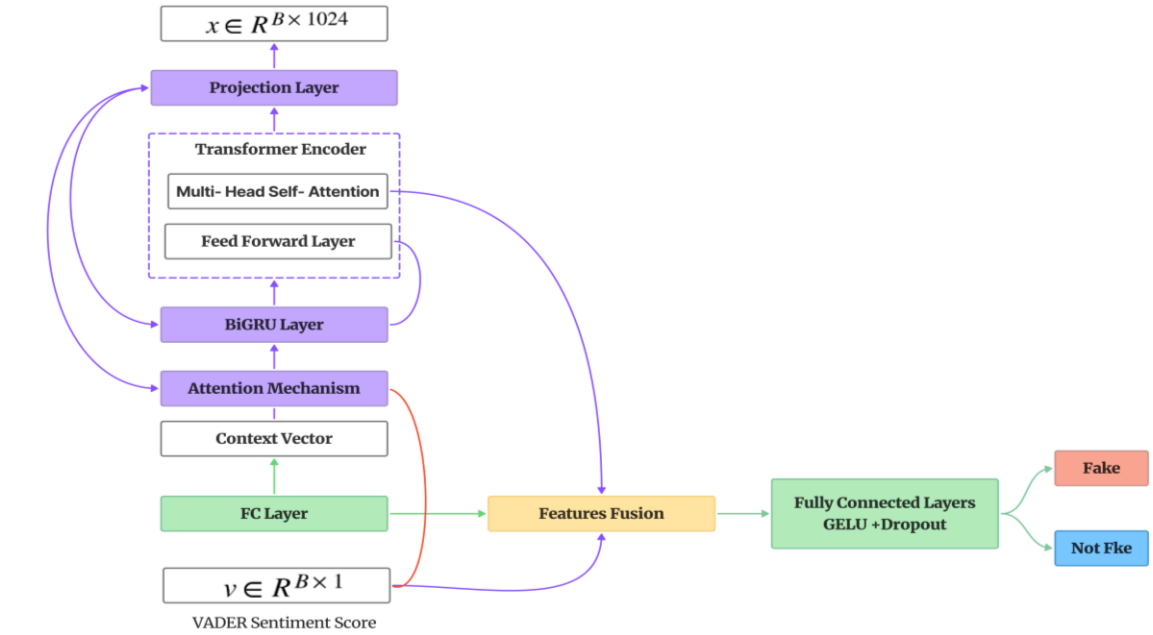

Figure 5: Hybrid transformer–BiGRU–attention classification with VADER fusion.

Table 2: Performance results of the proposed model.

Test Performance				
Metric	Value			
Test Accuracy	0.9545			
Test ROC-AUC	0.9898			
Test Loss	0.1250			
Total Samples	400,000			
Positive Samples	200,000			
Negative Samples	200,000			
Classification Result				
Class	Precision	Recall	F1-Score	Support
Negative (0)	0.9555	0.9533	0.9544	200,000
Positive (1)	0.9534	0.9556	0.9545	200,000
Macro Avg	0.9545	0.9545	0.9545	400,000
Weighted Avg	0.9545	0.9545	0.9545	400,000

The attention mechanism zeroes in on the most important parts of what the system has learned. if each review what the system has learned. Then, final layers of fully connected calculations are used to say if each review is fake or real Figure 5 shows the architecture of the proposed hybrid deep learning model. In essence, the system aims to be better at detecting fake reviews by combining a deep understanding of what's being said with information about the emotions expressed. This blended approach means the model gains from both semantic (meaning) and emotional characteristics, and is more reliable when looking at lots of Amazon reviews.

6 RESULTS AND DISCUSSION

We'll explain here what we did in the experiments and how we decided whether our system was successful at spotting fake reviews. We ran our model on a bunch of Amazon reviews and compared its performance to the results of more standard machine learning and deep learning methods. The evaluation's main goal was to understand how effectively our combined model could distinguish between fake and genuine reviews. Accuracy and AUC, both typical methods for assessing a classification's correctness, are what we used to do this. how the model performed. Accuracy simply tells you what percentage of all the reviews were labelled correctly. AUC on the other hand, assesses how well the model can tell the difference between fake and real reviews, when you change how sure it has to be before calling a review one or the other. Both of these measurements are frequently used for spotting fake reviews and sorting text in general. The following Figure 6 showing Confusion Matrix of the proposed model.

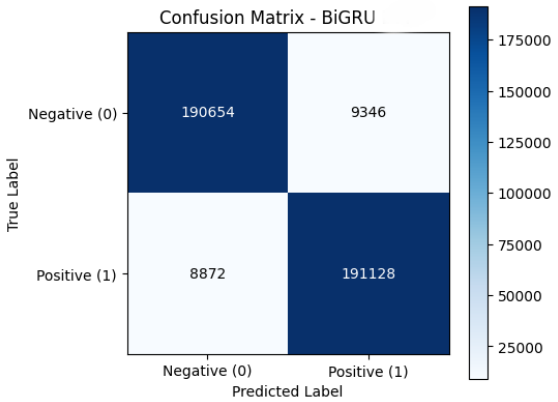

Figure 6: Confusion matrix of the proposed system.

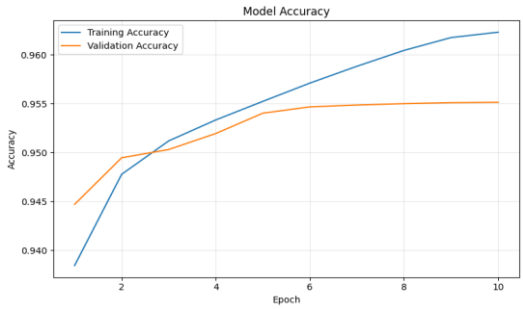

Figure 7: Training ACC valus.

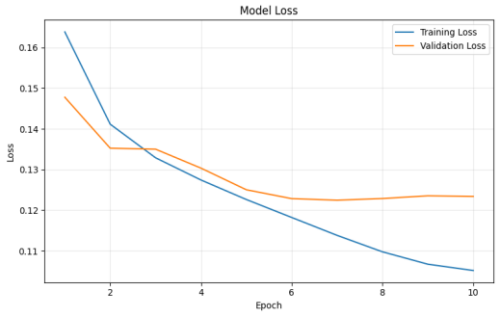

Figure 8: Traning loss function values.

Looking at Figures 7 and 8, it's obvious the model was successfully taught. As training went on (through each 'epoch'), the 'loss' got consistently smaller, and the accuracy steadily rose, meaning it got better at categorizing things. Plus, training and validation outcomes weren't far apart, which is good - it means the model will likely perform well on new data and hasn't memorized the training data. Essentially, the model learned in a reliable way and did a very good job of finding fake reviews.

Table: 3 Comparison with Baseline Models.

N0.	Model / Method Name	Dataset Version	No. of Samples	ACC	AUC
1	TF-IDF +SVM	Amazon	(4M)	0.91	0.96
2	Word2Vec+ LightGBM	Amazon	(4M)	0.87	0.94
3	FastText +XGBoost	Amazon	(4M)	0.88	0.94
4	GloVe +LightGBM	Amazon	(4M)	0.79	0.87
5	RNN+ Word2Vec	Amazon	(4M)	0.92	0.92
6	RCNN+VADER	Amazon	(4M)	0.93	0.98
7	BiLSTM+GLoVe+ VADER	Amazon	(4M)	0.92	0.97
8	RoBERTa-Large +VADER +BiGRU	Amazon	(4M)	0.9545	0.9898

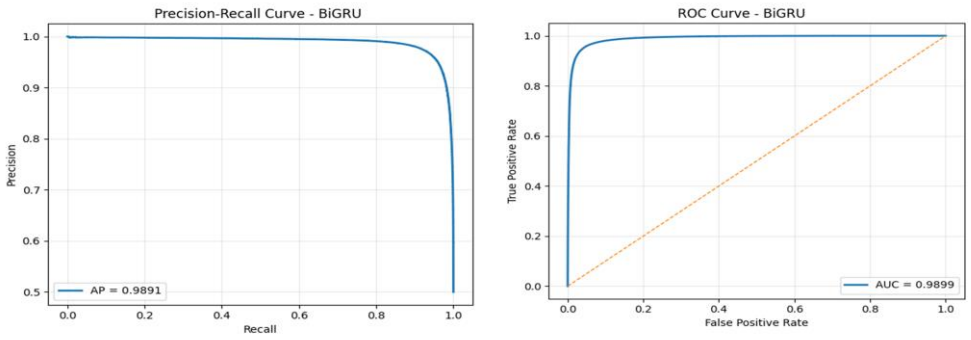

Figure: 9 AUC curve.

The AUC curve illustrates how well our model tells the difference between false and real reviews (Table 3). The better the model is at classifying, regardless of where you set the dividing line between "fake" and "real", the closer the AUC value will be to one (or 100%). Our model got a really good AUC score, proving it's good at clearly showing which reviews are which; it does a good job of keeping fakes and not fake reviews in separate piles. The following Figure 9 showed the AUC curve.

When we tested it, our combined (or 'hybrid') model did a better job than lots of the usual machine learning and deep learning approaches. It's better because it merges information about the situation around words, which RoBERTa-Large gets, with the feelings expressed in the text, from VADER. Then, BiGRU and attention mechanisms within the system help it to highlight the important bits of what people write in their reviews. So, all in all, our method is more successful at spotting fake reviews than if you just use one piece of information or one of the standard models on its own.

7 CONCLUSIONS

The paper details a deep learning system to spot phony Amazon reviews. It does this by putting together the way the review text is used (from RoBERTa-Large) with the emotional feelings expressed in it (using VADER). Then, Transformer learning, BiGRU and attention mechanisms are applied to these combined characteristics to make the system even better at deciding if a review is fake. Essentially, by understanding what is being said and how someone feels about it, this system can identify fake reviews much better, and in doing so, make people feel more secure when they're shopping online.

8 FUTURE WORK

There are quite a few ways we could improve this system for spotting fake reviews in the future. For a start, we could add more details about how reviewers act: how active they are, how often they review, what ratings they tend to give, and when they actually post their reviews. These kinds of details might show something is off, even when the review itself doesn't give it away. Also, we could get a fuller picture by using details about the products and the reviewers themselves, alongside what's in the review and how positive or negative it is. That is to say, by looking at the review from all angles, the system should get better at separating honest opinions from fakes.

REFERENCES

[1] M. Bramer, Principles of Data Mining, 1st ed. London, U.K.: Springer, 2007, p. 344, [Online]. Available: https://doi.org/10.1007/978-1-84628-766-4.

[2] D. Cao, X. He, L. Nie, X. Wei, X. Hu, S. Wu, and T.-S. Chua, "Cross-platform app recommendation by jointly modeling ratings and texts," ACM Transactions on Information Systems (TOIS), vol. 35, no. 4, pp. 1-27, Jul. 2017, [Online]. Available: https://doi.org/10.1145/3017429.

[3] M. A. Sharaf, M. A. Cheema, and J. Qi, "Preface," Springer Verlag, 2015, [Online]. Available: https://doi.org/10.1007/978-3-319-19548-3.

[4] 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, Apr. 2016, [Online]. Available: https://doi.org/10.1080/07421222.2016.1205907.

[5] N. Jindal and B. Liu, "Opinion Spam and Analysis," in Proceedings of the 2008 International Conference on Web Search and Data Mining (WSDM 2008), New York, NY, USA: ACM, 2008, pp. 219-230, [Online]. Available: https://doi.org/10.1145/1341531.1341560.

[6] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "What Yelp Fake Review Filter Might Be Doing?," in Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media (ICWSM 2013), 2013, pp. 409-418, [Online]. Available: https://doi.org/10.1609/icwsm.v7i1.14389.

[7] L. Akoglu, R. Chandy, and C. Faloutsos, "Opinion Fraud Detection in Online Reviews by Network Effects," in Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media (ICWSM 2013), 2013, pp. 2-11, [Online]. Available: https://doi.org/10.1609/icwsm.v7i1.14380.

[8] B. Hooi, N. Shah, A. Beutel, S. Günnemann, L. Akoglu, M. Kumar, and C. Faloutsos, "Birdnest: Bayesian inference for ratings-fraud detection," in Proceedings of the 2016 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, Jun. 2016, pp. 495-503, [Online]. Available: https://doi.org/10.1137/1.9781611974348.56.

[9] J. J. McAuley and J. Leskovec, "From amateurs to connoisseurs," in Proceedings of the 22nd International Conference on World Wide Web, New York, NY, USA: ACM, May 2013, pp. 897-908, [Online]. Available: https://doi.org/10.1145/2488388.2488466.

[10] J. McAuley, C. Targett, Q. Shi, and A. van den Hengel, "Image-Based Recommendations on Styles and Substitutes," in Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, New York, NY, USA: ACM, Aug. 2015, pp. 43-52, [Online]. Available: https://doi.org/10.1145/2766462.2767755.

[11] L. He, X. Wang, H. Chen, and G. Xu, "Online Spam Review Detection: A Survey of Literature," Human-Centric Intelligent Systems, vol. 2, no. 1-2, pp. 14-30, Jun. 2022, [Online]. Available: https://doi.org/10.1007/s44230-022-00001-3.

[12] S. Singh Sikarwar and N. Tiwari, "Analysis the Sentiments of Amazon Reviews Dataset by Using Linear SVC and Voting Classifier," International Journal of Scientific & Technology Research, vol. 9, no. 3, pp. 6288-6291, 2020.

[13] L. N. Nguyen, "An Empirical Study on Fake Review Detection," TRA VINH University Journal of Science, Jul. 2023, [Online]. Available: https://doi.org/10.35382/tvujs.13.6.2023.2107.

[14] J. Li, Q. Pan, and Y. Wang, "Sentiment analysis applied on Amazon reviews," Applied and Computational Engineering, vol. 44, no. 1, pp. 26-32, Mar. 2024, [Online]. Available: https://doi.org/10.54254/2755-2721/44/20230079.

[15] P. Hajek, A. Barushka, and M. Munk, "Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining," Neural Computing and Applications, vol. 32, no. 23, pp. 17259-17274, Dec. 2020, [Online]. Available: https://doi.org/10.1007/s00521-020-04757-2.

[16] J. Salminen, C. Kandpal, A. M. Kamel, S. Gyo Jung, and B. J. Jansen, "Creating and detecting fake reviews of online products," Journal of Retailing and Consumer Services, vol. 64, Jan. 2022, [Online]. Available: https://doi.org/10.1016/j.jretconser.2021.102771.

[17] S. He, B. Hollenbeck, G. Overgoor, D. Proserpio, and A. Tosyali, "Detecting fake-review buyers using network structure: Direct evidence from Amazon," Proceedings of the National Academy of Sciences of the United States of America, vol. 119, no. 47, Art. no. e2211932119, Nov. 2022, [Online]. Available: https://doi.org/10.1073/pnas.2211932119.

[18] E. Hashmi and S. Y. Yayilgan, "A robust hybrid approach with product context-aware learning and explainable AI for sentiment analysis in Amazon user reviews," Electronic Commerce Research, 2024, [Online]. Available: https://doi.org/10.1007/s10660-024-09896-5.

[19] S. Geetha, E. Elakiya, R. S. Kanmani, and M. K. Das, "High performance fake review detection using pretrained DeBERTa optimized with Monarch Butterfly paradigm," Scientific Reports, vol. 15, no. 1, Dec. 2025, [Online]. Available: https://doi.org/10.1038/s41598-025-89453-8.

[20] V. Gupta, "Sentiments Analysis of Amazon Reviews Dataset By using Machine Learning," International Journal for Research in Applied Science and Engineering Technology, vol. 11, no. 1, pp. 951-957, Jan. 2023, [Online]. Available: https://doi.org/10.22214/ijraset.2023.48678.

[21] M. Tabany and M. Gueffal, "Sentiment Analysis and Fake Amazon Reviews Classification Using SVM Supervised Machine Learning Model," Journal of Advances in Information Technology, vol. 15, no. 1, pp. 49-58, 2024, [Online]. Available: https://doi.org/10.12720/jait.15.1.49-58.

[22] Y. Pan and L. Xu, "Detecting Fake Online Reviews: An Unsupervised Detection Method With a Novel Performance Evaluation," International Journal of Electronic Commerce, vol. 28, no. 1, pp. 84-107, Jan. 2024, [Online]. Available: https://doi.org/10.1080/10864415.2023.2295067.

[23] R. Gupta, V. Jindal, and I. Kashyap, "Recent state-of-the-art of fake review detection: a comprehensive review," The Knowledge Engineering Review, vol. 39, p. e8, Nov. 2024, [Online]. Available: https://doi.org/10.1017/S0269888924000067.

[24] Y. Liu, X. Ao, Z. Qin, J. Chi, J. Feng, H. Yang, and Q. He, "Pick and choose: A GNN-based imbalanced learning approach for fraud detection," in Proceedings of the Web Conference 2021, Apr. 2021, pp. 3168-3177, [Online]. Available: https://doi.org/10.1145/3442381.3449989.

[25] A. Khamparia and K. M. Singh, "A systematic review on deep learning architectures and applications," Expert Systems, vol. 36, no. 3, Jun. 2019, [Online]. Available: https://doi.org/10.1111/exsy.12400.

[26] Y. Sun, J. Zhang, Z. Yu, Y. Zhang, and Z. Liu, "The Bidirectional Gated Recurrent Unit Network Based on the Inception Module (Inception-BiGRU) Predicts the Missing Data by Well Logging Data," ACS Omega, vol. 8, no. 30, pp. 27710-27724, Aug. 2023, [Online]. Available: https://doi.org/10.1021/acsomega.3c03677.

[27] C. Özkurt, "Comparative Analysis of State-of-the-Art QA Models: BERT, RoBERTa, DistilBERT, and ALBERT on SQuAD v2 Dataset," Chaos and Fractals, Jul. 2024, [Online]. Available: https://doi.org/10.69882/adba.chf.2024073.

[28] J. McAuley and J. Leskovec, "Hidden factors and hidden topics: Understanding rating dimensions with review text," in RecSys 2013 - Proceedings of the 7th ACM Conference on Recommender Systems, 2013, pp. 165-172, [Online]. Available: https://doi.org/10.1145/2507157.2507163.

[29] N. Shrestha and F. Nasoz, "Deep Learning Sentiment Analysis of Amazon.Com Reviews and Ratings," International Journal on Soft Computing, Artificial Intelligence and Applications, vol. 8, no. 1, pp. 01-15, Feb. 2019, [Online]. Available: https://doi.org/10.5121/ijscai.2019.8101.