Project Report
on
Fake Review Detection

Submitted for the Partial Fulfillment of the Requirements for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

Nishchay Singh (UI21CS36)
Kunal Bhamoriya (UI21CS32)

Under the guidance of

Dr. Pradeep Kumar Roy

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY SURAT-394190

APRIL 2024

Indian Institute of Information Technology Surat
Computer Science and Engineering Department

CERTIFICATE

This is to certify that candidates **Nishchay Singh , Kunal Bhamoriya** bearing Roll No: **UI21CS36 , UI21CS32** of B.TECH. III, 6th Semester has successfully carried out the work on “Fake Review Detection” of ”Machine Learning” for the partial fulfillment of the degree of Bachelor of Technology (B.Tech.) in April, 2024.

Faculty Supervisor: Name: Dr Pradeep Kumar Roy
Faculty Supervisor: Name: Mr Vipul Kania

1. Member 1: Name: Sign: *Nishchay Singh*

2. Member 2: Name: Sign: *Kunal Bhamoriya*

DECLARATION

This is to certify that
(i) This report comprises our original work towards the degree of Bachelor of Technology in Computer Science and Engineering at Indian Institute of Information Technology (IIIT) Surat and has not been submitted elsewhere for a degree,
(ii) Due acknowledgement has been made in the text to all other material used.

Signature of Students:

Nishchay Singh

Kunal Bhamoriya

ACKNOWLEDGEMENTS

First of all, We would like to express our sincere and deep gratitude to our supervisors, **Dr. Pradeep Kumar Roy**, **Mr. Vipul Kania**, Faculty, Department of Computer Science and Engineering, for his kind and constant support for the completion of our Project. His valuable advice, critical criticism and active supervision encouraged me to sharpen our research methodology and was instrumental in shaping our professional outlook.

I would also like to thank the entire faculty for their constant encouragement and assistance they have provided us. Last, but definitely not the least, a very special note of thanks to our parents, our team-mates and friends who have helped us during this entire period directly or indirectly.

ABSTRACT

The emergence of products and services online has led to a reliance on review systems to aid consumers in making informed decisions. However, the rise of fake reviews, generated either by computers or humans, has introduced significant challenges to the authenticity of these systems. Fake reviews, whether positive or negative, can distort perceptions of products and services, impacting their popularity and relevance.

This study defines fake reviews as a form of opinion spamming and highlights their detrimental effects on online platforms. While existing research primarily focuses on supervised learning algorithms for fake review detection, this study proposes measures to efficiently detect fake reviews. Through the analysis of review samples, key distinctions between genuine and fake reviews are identified, informing the formulation of effective detection rules.

The study aims to contribute to the development of efficient detection models to combat the growing issue of fake reviews. By enabling consumers to make decisions based on legitimate reviews and empowering sellers and manufacturers to collect genuine feedback, the proposed measures seek to foster trust and integrity in online review systems.

Fake Review Detection

Introduction

With the prevalence of products and services being offered through the medium of the internet, the average consumer more often than not finds themselves confused with the products or services. This problem is usually negated with the help of a review system in which an existing user can leave a review of the product or service on the platform from which the product or service is available. These review systems have made lives easier for the general consumer as well as the sellers or manufacturers on the platforms as these reviews are useful for an average user to understand if the product is suitable for their needs and a manufacturer can understand the needs of the general public and make changes to their products and services in future accordingly.

In recent times it has been observed that many of these platforms have been flooded with reviews which are either generated by a computer or are fake. These fake reviews are sometimes posted by the sellers or the manufacturers to promote their products or services and spread false information. This process of posting a large quantity of fake reviews to sway opinions of general users is termed as opinion spamming. These fake opinions on platforms can be either positive or negative and these reviews or opinions can actually make a significant difference in the popularity or relevance of the product or the service. This problem has led to many platforms to integrate systems which can detect illegitimate reviews.

A **Fake Review**, as defined by Huy le and Ben kim, "is considered a type of opinion spamming"[1]. It is an illegal practice as sellers or manufacturers are hiring reviewers to post or generate untruthful reviews for their products or services.We can also define a fake review to be a untruthful opinion which would present a product or a service in a deceptive manner or try to portray untrustworthy information, opinions and experiences.

Existing work on this domain mainly focuses on supervised learning algorithms where the data is labeled beforehand. Our work aims to provide some measures to detect fake reviews with efficiency. We have analyzed a sample of reviews to understand the differences between an actual review and a fake review and on the basis of our findings we have formulated some rules that would be beneficial to review detection models.

We believe that our study will be helpful in designing and implementing models and measures to combat the problem of growing fake reviews. It will be beneficial to both the consumers as well as the seller or manufacturer as the consumer would be able to formulate their decisions based on legitimate reviews and the manufacturers and sellers would be able to collect true and actual data from the reviews.

Related Works

A lot of work on Fake Review Detection has been already done before this work as it has been a critical problem for digital sellers and manufacturers.some of the work that has been accomplished before this work has been listed below.

One study from ScienceDirect focuses on feature design for fake review detection, aiming to extract characteristics from review datasets that effectively distinguish between fake and genuine reviews [3]. Another study from the same source explores the creation and detection of fake reviews, experimenting with language models like ULMFiT and GPT-2 to generate fake reviews and assess methods to identify them [4].

A survey on Fake Reviews Detection from ResearchGate delves into various techniques, highlighting the importance of sentiment analysis and feature engineering in distinguishing between authentic and fake content.

SpringerOpen's research addresses the detection of fake reviewers within networks of buyers, exploring graph learning techniques to identify fake reviewers and contributing to more robust detection mechanisms [5].

Similarly, a systematic literature review from Springer provides insights into the chronological development of fake review detection models and research designs, outlining the challenges faced in this field and proposing avenues for further exploration.

Lastly, a study from IIETA focuses on the application of machine learning techniques for fake review detection, proposing methodologies for classifying and identifying fake reviews, thus contributing to the development of more effective detection algorithms.

In summary, these studies collectively contribute to the advancement of fake review detection methods by exploring various techniques, ranging from feature design and sentiment analysis to graph learning and machine learning algorithms. Their findings enhance our understanding of fake review detection and pave the way for more robust solutions to combat fraudulent activities online.

Proposed System

Importing Libraries is our first step as python libraries are collections of modules that contain useful codes and functions, eliminating the need to write them from scratch. There are tens of thousands of Python libraries that help machine learning developers, as well as professionals working in data science, data visualization, and more. Python is the preferred language for machine learning because its syntax and commands are closely related to English, making it efficient and easy to learn. Libraries used in this model for fake review detection are
Numpy : used for working with numerical data.
Matplotlib : used for understanding data graphically and statistically.
Nltk : It stands for Natural Language Toolkit, used for language processing, tokenizing, stemming and lemmatization.
Sci-Kit learn : It provides tools for classification and evaluation of models.

Collection of Data means pooling data by scraping, capturing, and loading it from multiple sources, including offline and online sources. In order to build intelligent applications, machine learning models need large amounts of structured training data. Gathering sufficient training data is the first step in solving any machine learning problem. Predictive models are only as good as the data from which they are built, so good data collection practices are crucial to developing high-performing models. The data needs to be error-free and contain relevant information for the task at hand. It is recommended to collect as much data as possible for good predictions. You can begin with small batches of data and see the result of the model. The most important thing to consider while data collection is diversity. Diverse data will help your model cover more scenarios. So, when focusing on how much data you need, you should cover all the scenarios in which the model will be used. There are various data collection techniques available. The oldest and most basic way is observation, but nowadays large companies have huge data collection which can be accessed by API's. For this model the dataset used is available on kaggle which is contributed by mexwell[2], the fake reviews dataset contains 20k fake reviews and 20k real product reviews. OR = Original reviews (presumably human created and authentic); CG = Computer-generated fake reviews.

Preprocessing and labeling is an integral step in Machine Learning as the

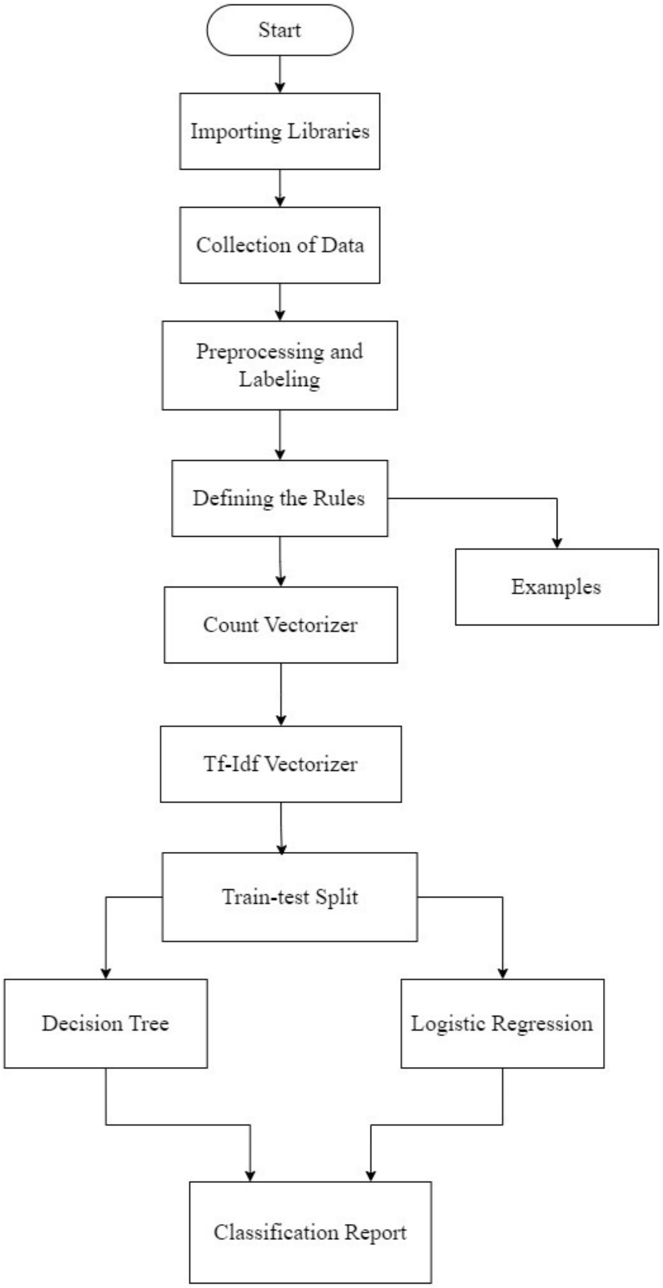


Figure 1: Flowchart depicting the working of Project.

quality of data and the useful information that can be derived from it directly affects the ability of our model to learn therefore, it is extremely important that we preprocess our data before feeding it into our model. Preprocessing includes handling null values, In any real-world dataset, there are always few null values. It doesn't really matter whether it is a regression, classification or any other kind of problem, no model can handle these NULL or NaN values on its own. Datasets also contain outliers, An outlier is a data point that is noticeably different from the rest. They represent errors in measurement, bad data collection, or simply show variables not considered when collecting the data. In machine learning, data labeling is the process of identifying raw data (images, text files, videos, etc.) and adding one or more meaningful and informative labels to provide context so that a machine learning model can learn from it. In this model labels 0 and 1 are added for Computer generated reviews and Original reviews.

CountVectorizer is a great tool provided by the scikit-learn library in Python. It is used to transform a given text into a vector on the basis of the frequency (count) of each word that occurs in the entire text. This is helpful when we have multiple such texts, and we wish to convert each word in each text into vectors. and each text sample from the document is a row in the matrix. The value of each cell is nothing but the count of the word in that particular text sample. We are using a bag of words i.e. Count Vectorizer a class in the sklearn library to extract the features from the sentences. It is used to count the frequency of the word that occurs in the sentence. It creates a vocabulary of all the unique words occurring in all the documents in the training set. Total vocabulary in the Dataset according to Bag of Words is 34450.

Tf-Idf Vectorizer stands for Term Frequency Inverse Document Frequency of records. It can be defined as the calculation of how relevant a word in a series or corpus is to a text. The meaning increases proportionally to the number of times in the text a word appears but is compensated by the word frequency in the corpus (data-set). Term Frequency: In document d, the frequency represents the number of instances of a given word t. Therefore, we can see that it becomes more relevant when a word appears in the text, which is rational. Since the ordering of terms is not significant, we can use a vector to describe the text in the bag of term models. For each specific term in the paper, there is an entry with the value being the term frequency.

$$\text{tf(t,d)} = \text{count of t in d,}$$
$$\text{df(t)} = \text{occurrence of t in documents}$$

Inverse Document Frequency mainly tests how relevant the word is. The key aim of the search is to locate the appropriate records that fit the demand. Since tf considers all terms equally significant, it is therefore not only possible to use the term frequencies to measure the weight of the term in the paper. First, find the document frequency of a term t by counting the number of documents containing the term:

$$\text{df(t)} = \text{Document frequency of a term t}$$
$$\text{N(t)} = \text{Number of documents containing the term t}$$
$$\text{idf(t)} = \text{log(N/ df(t))}$$

Train-Test Split is a process in which the data is split between training data and

testing data, The percentage of split depends upon the person who is training the model. One of the most important aspects of machine learning is training. The act of providing data to the algorithm training is vital in creating algorithms that perform tasks efficiently and effectively. The training can be time consuming and difficult to perform if using your own computer that's where the train test split comes in handy. A train test is the way of structuring your machine learning project so that you can test your hypothesis quickly and inexpensively. Basically it's a way to divide the training data so that you can try your algorithm to one half and evaluate the result on the other half. This split is effective only if the data is large, in case of small data training will provide less accurate results. In this model the dataset contains the 40k reviews and the train test split is 65-35%.
review_train, review_test, label_train, label_test = train_test_split(df['text_'], df['label'], test_size=0.35)

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome. In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches. the test are performed on the basis of features of the given dataset. It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions. It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure. The primary reason behind choosing the decision tree algorithm is that it mimics the human decision making ability. Working of decision trees is as follows. In a decision tree, for predicting the class of the given dataset, the algorithm starts from the root node of the tree. This algorithm compares the values of root attribute with the record (real dataset) attribute and, based on the comparison, follows the branch and jumps to the next node.For the next node, the algorithm again compares the attribute value with the other sub-nodes and moves further. It continues the process until it reaches the leaf node of the tree.

Logistic regression is a supervised machine learning algorithm that accomplishes binary classification tasks by predicting the probability of an outcome, event, or observation. The model delivers a binary or dichotomous outcome limited to two possible outcomes: yes/no, 0/1, or true/false. Logical regression analyzes the relationship between one or more independent variables and classifies data into discrete classes. It is extensively used in predictive modeling, where the model estimates the mathematical probability of whether an instance belongs to a specific category or not. For example, 0 – represents a negative class; 1 – represents a positive class.

Logistic regression is commonly used in binary classification problems where the outcome variable reveals either of the two categories (0 and 1). Logistic regression uses a logistic function called a sigmoid function to map predictions and their probabilities.

Classification Report visualizer displays the precision, recall, F1, and support scores for the model.
Precision: Accuracy of positive predictions.
Recall : Fraction of positives that were correctly identified.
F1 score : What percent of positive predictions were correct?
Support : It is the number of actual occurrences of the class in the specified dataset.

Result

The experimental results of our fake review detection study, including a detailed analysis of performance metrics, such as accuracy, precision, recall, and F1 score.In our case the task is a classification task so we have chosen (i) Decision Tree Algorithm and (ii) Logistic Regression.Additionally, we discuss the outcomes of the model by examining the confusion matrix and interpreting its implications.

For Decision Tree Algorithm ,
precision : 0.73 , recall : 0.72 , f1-score : 0.73 , accuracy : 73.91%.

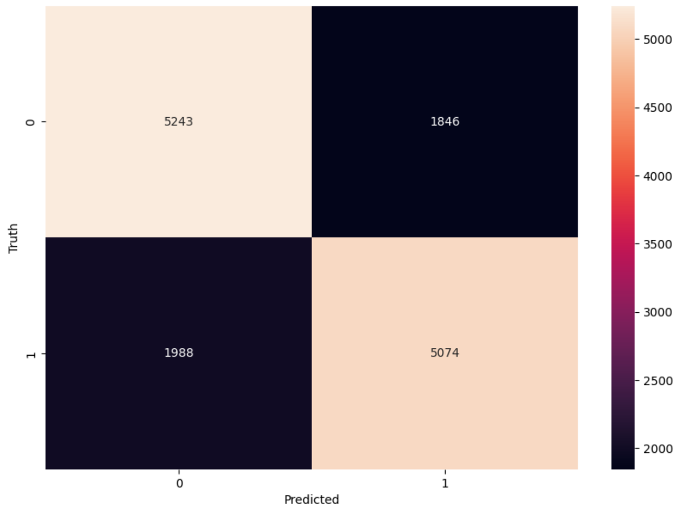


Figure 2: Confusion matrix of Decision Tree Algorithm.

For Logistic Regression Algorithm ,
precision : 0.87 , recall : 0.88 , f1-score : 0.86 , accuracy : 86.14%.

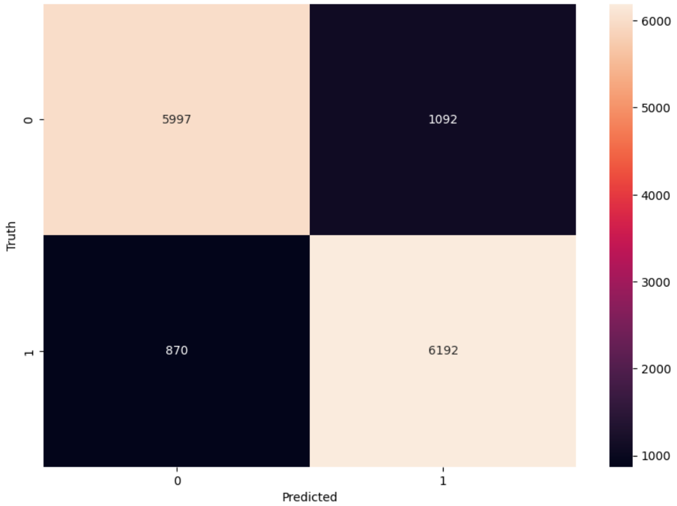


Figure 3: Confusion matrix of Logistic Regression Algorithm.

Resulting rules from sample data

The dataset that we have employed for our model is provided by mexwell[2] which is available on kaggle platform. But for analyzing reviews to better understand the various manners in which they are written and then to formulate some rules which could possibly detect fake reviews we have taken the liberty to sample reviews from various sources. After our analysis some rules that we have formulated are:

Typos and Excellent Language may show a fake review , for a normal user it is highly likely that they would commit a mistake while writing a review. But a computer Generated review is most of the time perfect.
Example: Love it, a great upgrade from the original. I've had mine for a couple of years.
The above example exhibits excellent language and there are no mistakes.

Irrelevant information in a review is also an indicator of an illegitimate review. A legitimate review of a product always contains relevant information while it is possible that a fake review may contain bogus information and details.
Example: I WANTED DIFFERENT FLAVORS BUT THEY ARE NOT.
The above example review is for a product from the home and decor section but focuses on the flavors which are not there in furniture.

A long and grammatically perfect review which also contains details and information which aren't much useful might also be a fake review.
Example: OMG, First reason why I chose this style is because it is a nice fit for a person with a narrow foot. The front of the shoe is slightly wider than the back, so

that it is a little smaller in the heel. The front is very soft and gives good support. The straps are very comfortable and the material is very soft and soft. The color is very nice and the material is very comfortable. The only reason I didn't buy this style was because it is very short for a woman, but I am glad I did. I am usually a size 6, but ordered a medium. The length is perfect. The fabric is soft, but the color is very bright and vivid. It is a cute color to wear for a costume. I wish it had the support that I need. I would recommend this style for anyone.I really like these shoes. I have had a few pairs of these in the past. They are very comfortable and are very comfortable. I usually wear a 9.5 and these fit perfectly. They are very comfortable and the cushioning is very nice. I have never had a pair of these in a store and this one was on sale. I would recommend these shoes.I ordered these in a size small and they fit perfectly. I bought them for my mom, who wears a medium. She loves them. They are great for the weather, and she likes them so much that she has been wearing them every day. They are very comfy and fit great. She is so happy with them.
The above example is very long for a normal user to spend time writing and also the grammar is perfect.

Rating and review mismatch, in a case where a review has a certain rating but the sentiment that the review itself conveys is not corresponding to the rating the review can be classified as a fake review.
Example: Not what I am accustomed to. The only reason I gave it 4 stars is because I just can
The above example has been marked for 2 star rating while in the review itself it is marked for 4 stars.

Use of Incorrect Emoji can also serve as an indicator of a fake review. In a case where an emoji does not match with the sentiment of the review itself ,the review is most likely a fake review.

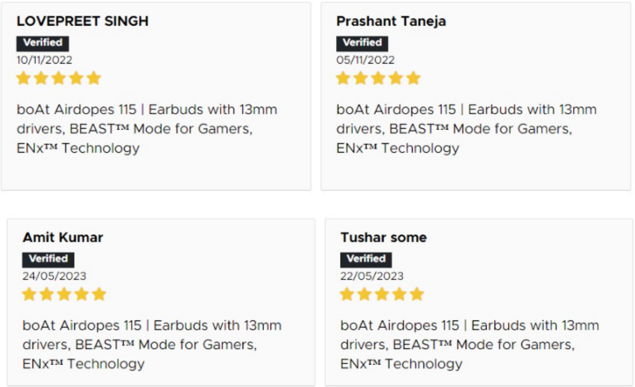
Incorrect product or service name in the review text is also a giveaway of a fake review. An actual reviewer would not make a mistake of writing a review with the wrong product name.
Example: A little heavy for my liking. The only problem is that it's kind of hard to put on. If you want to get a solid feel, the wallet is great.
The above example does not make sense as it talks about fitting the user while the product described is infact a wallet.

Typos and improper Language, while no Typos and perfect Language depicts a computer generated review in the same manner a review with too many Typos and poor Language may have been a review written by a person who is being paid to manually flood reviews section of products[6].
Example: Absolutely adorable! And excellent price. We have had the wooden ones

for a few months now and they
The above example does not make sense when we look at its last part.

Extremely positive or negative reviews are also found to be a way to opinion spam a product review page. As in case of most products both positive and negative aspects are present.
Example: I love pyrex quality and this set is the best. I will be purchasing more in the future.
The above example review is overly positive about the product and tries to depict it the best in the segment.

Some Examples of fake review on websites are:



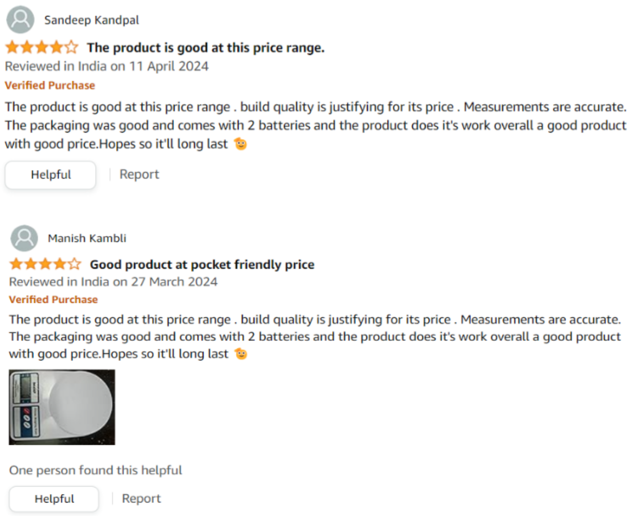
LOVEPREET SINGH
Verified
10/11/2022
★★★★★
boAt Airdopes 115 | Earbuds with 13mm drivers, BEAST™ Mode for Gamers, ENx™ Technology

Prashant Taneja
Verified
05/11/2022
★★★★★
boAt Airdopes 115 | Earbuds with 13mm drivers, BEAST™ Mode for Gamers, ENx™ Technology

Amit Kumar
Verified
24/05/2023
★★★★★
boAt Airdopes 115 | Earbuds with 13mm drivers, BEAST™ Mode for Gamers, ENx™ Technology

Tushar some
Verified
22/05/2023
★★★★★
boAt Airdopes 115 | Earbuds with 13mm drivers, BEAST™ Mode for Gamers, ENx™ Technology

Figure 4: All of the above reviews are from different users but have the same content.



Sandeep Kandpal

★★★★☆ **The product is good at this price range.**

Reviewed in India on 11 April 2024

Verified Purchase

The product is good at this price range . build quality is justifying for its price . Measurements are accurate. The packaging was good and comes with 2 batteries and the product does it's work overall a good product with good price.Hopes so it'll long last 🙂

Manish Kambli

★★★★☆ **Good product at pocket friendly price**

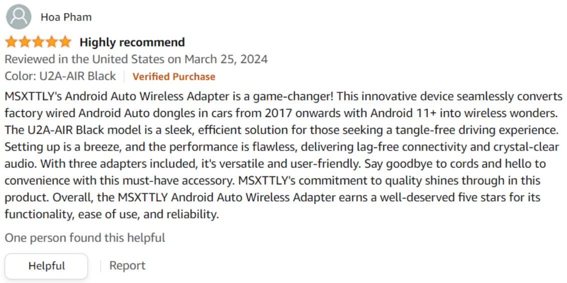
Reviewed in India on 27 March 2024

Verified Purchase

The product is good at this price range . build quality is justifying for its price . Measurements are accurate. The packaging was good and comes with 2 batteries and the product does it's work overall a good product with good price.Hopes so it'll long last 🙂

One person found this helpful

Figure 5: Above two reviews are from different users but both of them contain exactly the same content.



Hoa Pham

★★★★★ **Highly recommend**

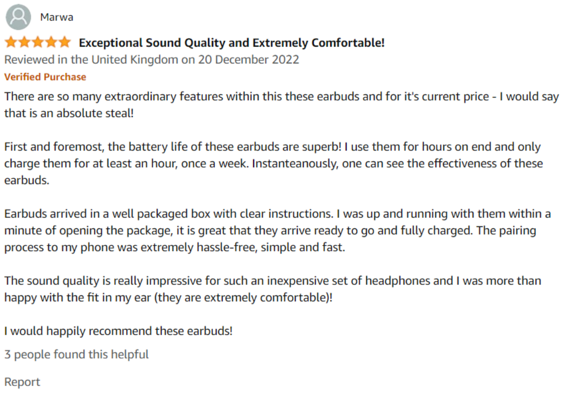
Reviewed in the United States on March 25, 2024

Color: U2A-AIR Black | Verified Purchase

MSXTTLY's Android Auto Wireless Adapter is a game-changer! This innovative device seamlessly converts factory wired Android Auto dongles in cars from 2017 onwards with Android 11+ into wireless wonders. The U2A-AIR Black model is a sleek, efficient solution for those seeking a tangle-free driving experience. Setting up is a breeze, and the performance is flawless, delivering lag-free connectivity and crystal-clear audio. With three adapters included, it's versatile and user-friendly. Say goodbye to cords and hello to convenience with this must-have accessory. MSXTTLY's commitment to quality shines through in this product. Overall, the MSXTTLY Android Auto Wireless Adapter earns a well-deserved five stars for its functionality, ease of use, and reliability.

One person found this helpful

Figure 6: The review is too positive and looks like a sales pitch.



Marwa

★★★★★ **Exceptional Sound Quality and Extremely Comfortable!**

Reviewed in the United Kingdom on 20 December 2022

Verified Purchase

There are so many extraordinary features within this these earbuds and for it's current price - I would say that is an absolute steal!

First and foremost, the battery life of these earbuds are superb! I use them for hours on end and only charge them for at least an hour, once a week. Instanteanously, one can see the effectiveness of these earbuds.

Earbuds arrived in a well packaged box with clear instructions. I was up and running with them within a minute of opening the package, it is great that they arrive ready to go and fully charged. The pairing process to my phone was extremely hassle-free, simple and fast.

The sound quality is really impressive for such an inexpensive set of headphones and I was more than happy with the fit in my ear (they are extremely comfortable)!

I would happily recommend these earbuds!

3 people found this helpful

Figure 7: The review is extremely positive and is possibly not legitimate.

Conclusion

The prevalence of fake reviews poses a significant challenge for both consumers and businesses alike. While review systems have greatly simplified decision-making for consumers and provided valuable feedback for sellers, the influx of fake reviews has undermined the trustworthiness of these platforms.

To address this issue, our study focused on developing efficient measures for detecting fake reviews. By analyzing a sample of reviews, we identified key differences between genuine and fake reviews, which enabled us to formulate rules beneficial for review detection models.

Our research contributes to the ongoing efforts to combat fake reviews by providing insights into the characteristics of deceptive reviews and proposing practical solutions. We believe that our findings will aid in the design and implementation of more effective detection algorithms, benefiting both consumers and businesses.

Moving forward, it is crucial to continue advancing fake review detection methods to maintain the integrity of online review systems. Collaboration between researchers, platform providers, and regulatory bodies will be essential in developing robust solutions to mitigate the impact of fake reviews on online commerce.

By enhancing our understanding of fake review detection and implementing measures to address this issue, we can create a more transparent and trustworthy online environment for consumers and businesses alike. Together, we can work towards ensuring that online reviews accurately reflect the quality of products and services, fostering confidence and informed decision-making among consumers.

References

[1] Huy Le , Ben Kim .(2020). ‘Detection of fake reviews on social media using Machine learning Algorithms’. https://web.archive.org/web/20201107223232/https://iacis.org/iis/2020/1_iis_2020_185-194.pdf

[2] Mexwell . ‘Fake Reviews Dataset’.https://www.kaggle.com/datasets/mexwell/fake-reviews-dataset

[3] Research on false review detection Methods: A state-of-the-art review https://www.sciencedirect.com/science/article/pii/S1319157821001993

[4] Creating and detecting fake reviews of online products https://www.sciencedirect.com/science/article/pii/S0969698921003374

[5] Machine Learning Approaches for Fake Reviews Detection: A Systematic Literature Review https://journals.riverpublishers.com/index.php/JWE/article/view/18579/19105

[6] Patrick Collinson .(July 2023). ‘Beat the fakes: how to find online reviews you can trust’ https://www.theguardian.com/money/2023/jul/15/fakes-online-reviews-trust