IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

E-Shopping Sentiment Analysis System Using Machine Learning and Web-Based Visualization

Aishwarya Jain *, Santosh Gaikwad**

* Department of Computer Science and Application, JSPM University, Pune, India

**Associate Professor, Faculty of Science and Technology, JSPM University, Pune, India

I. Abstract

E-commerce has expanded at an extraordinary pace over the course of the last decade; that should come as no surprise, but rather a testament to how important customer reviews are in influencing consumer purchase choices. Most frequently when buying a smartphone through Amazon or Flipkart, consumers will go directly to the "Reviews" tab before scrolling to check to see if they found any value in an additional review prior to selecting a model of smartphone from Amazon.

The objective of this paper is to describe a sentiment analysis system that was specifically designed for analyzing Amazon mobile phone reviews. The objective of the stakeholder analysis system is very simple: to accept raw customer feedback, analyse it through a fully automated process, and provide results to allow a customer to make an informed decision as to whether the devices are worthy purchases.

The technical aspects of the system rely on the following technologies: Python/Flask is used for the backend; TextBlob is used for sentiment analysis and processing; and Chart.JS is used to render charts/graphs that users will actually see on their screens. The review data used in the sentiment analysis system is taken from CSV files and is representative of many different brands of smartphones (e.g. Samsung, Apple, OnePlus, Vivo, Oppo, Realme, Motorola, Nokia). Every review is sorted into one of 3 categories: Positive, negative and neutral. Each of these categories is then broken into each unique individual feature — battery, camera, display, processor and software — all through various ways including pie charts, histograms and comparative charts.

The end result of this process creates a single source of good and bad for two very distinct sets of people. For someone looking to purchase a new mobile device, they are able to see what other people believe about their choices without having to sift through hundreds of separate opinions to find out. For a manufacturer, it will show them specific portions of the products that people are unhappy with, and this feedback can be very difficult to obtain from a star rating alone.

Keywords — Sentiment Analysis, Natural Language Processing, Machine Learning, Flask, Amazon Reviews, Mobile Review Analysis, Data Visualization.

I. INTRODUCTION

As consumers increasingly use the internet for product research and read reviews of products prior to making a purchase decision as compared to relying solely on the opinion of sales associates; shopping has changed significantly. Online shopping (e-commerce) is based on the reviews of past customers' experience with specific products. The number of reviews associated with a specific item gives you some level of comfort about making the purchase, especially when you are comparing two items; for example, the number of reviews will tell you that most people have had a positive experience with that product.

That's kind of a problem, though—there is way too many reviews to go through. For example, A popular model of a smartphone can gather more than ten thousand opinions on various platforms. No one has time for that and truth be told, even if they did, they'd come away with more questions than when they started. One person loves the battery. Another says it dies by noon. They say the camera, and then a guy says that the camera is not good. With everyone saying different things, it gets really tough to have your head around what you are supposed to think about the camera. The camera gets a lot of reviews and that can be confusing.

This is the issue that this project was designed to address. The project includes a system that automatically analysis consumer reviews, rather than just presenting them all to a user and hoping for the best. When a consumer requests to locate a specific brand or model of product, all the user has to do is enter that brand or model of product into the database, and the consumer will receive consumer sentiment/rating information (i.e., ratings/reviews can be sorted out by product feature, etc.) on that product. The below will provide a high-level overview of the processes occurring in the background when the consumer accesses this application.

1. The user enters a phone brand and model name
2. The system pulls all relevant reviews from the dataset
3. Each review is analyzed for sentiment
4. Ratings for each feature such as camera, battery, and screen will be evaluated separately.
5. The information will be provided via a straightforward web-based interface using graphs and tables.

It's not about putting a thumbs-up or thumbs-down rating on a review. The objective is to provide actionable information that people can use to determine the viability of purchasing the mobile device.

II. RELATED WORK

For many years, there has been extensive research on developing approaches using Natural Language Processing and machine learning tools for automating the sentiment analysis process of customer reviews on online marketplaces like Amazon and Flipkart, where there is plenty of data available and applications are obvious.

Pang and Lee's work on sentiment classification highlighted the reliability of algorithms such as Naive Bayes and Support Vector Machines in sorting customer reviews into positive, negative, and neutral categories. Hence, they established that the automation of classification was not only a theoretical possibility but also had practical applications.

Following that groundwork, subsequent researchers extended their focus towards developing a 'features analysis' approach; analysing not just whether a review was overall positive or negative, but also examining the specific features consumers mentioned in their reviews. In the case of smartphones for example, researchers dissected all feedback into independent components (or parameters) including, battery life, camera quality, display quality and CPU speed. This represents a significant improvement because a smartphone may possess an excellent camera and have very poor battery life, and thus any measure reflecting a single overall sentiment would not provide this kind of detail to users.

Although there has been much work done on the various systems to date, the common problem across them all is that the results are hard for a non-specialist to understand. The output of these systems consists mainly of text with a conventional technical style of presentation, with few (if any) graphical representations to help with interpretation. As such, a non-specialist who has been sent a report which contains output generated from any of these systems will struggle to understand its contents without having had previous experience in statistical analysis (data analysis) of any kind.

This project has been specifically designed to bridge this gap in classification and provides an online dashboard with features such as a comparison table, pie chart and histogram to present the results in a way that everyone can access, not just researchers.

The proposed system improves these limitations by integrating:

- Web-based dashboard
- Feature comparison table
- Pie chart visualization
- Histogram analysis
- Real-time mobile search system

Table 1: Comparative Analysis of Existing and Proposed Systems

System	Sentiment Analysis	Visualization	Feature Comparison
Traditional Systems	Yes	No	No
ML-Based Systems	Yes	Limited	Limited
Proposed System	Yes	Yes	Yes

III. Literature Survey

Sentiment analysis has been extensively researched to the extent that there is a wealth of existing work available to consult and use. The foundational work of Pang and Lee represents one of the first significant contributions to this area: their research into opinion mining helped to demonstrate the ability to apply various classifiers — including Naive Bayes and Support Vector Machines — on text data to produce valid predictions about the sentiment expressed in that text. This was a pivotal moment in the history of sentiment analysis, as it made able the transition of sentiment analysis from an area of theoretical curiosity to an area where researchers could actually begin developing applications, thus leading to a surge in research activity — i.e., product reviews and political comments posted on social media.

On the e-commerce side, academic studies demonstrated what most "normal" internet shoppers have known from their own experience — reviews are a powerful influencer in shaping purchasing decisions. Customers read the reviews prior to purchasing a product; they trust reviews more than they trust branding messages; and customers frequently change their mind about purchasing a product based on something someone else (i.e., another customer) wrote about it in an online review. In light of the above, it is surprising that although the amount of influence generated by reviews seems quite large, the tools and applications developed for reviews are mostly limited. Most review-related applications provide nothing more than an average rating (e.g., a star rating) for products and fail to differentiate between the reason behind negative reviews (e.g., non-delivery of goods or problems with the actual product) or provide a means for differentiating between long-time users' and first-time users' opinions of a product. There is no distinction between feedback from long-term users and opinions from someone who used the product once. That level of detail simply does not exist in most consumer-facing tools.

Research rooted in Twitter data brought its own lessons to the field. Working with short, informal, real-time text forced researchers to get creative with preprocessing and feature extraction — problems that carry over directly into product review analysis, where language is similarly unstructured and inconsistent.

Looking across all of this work, the same limitation keeps coming up. The results reported from high volume sentiment analysis systems almost never show evidence that their abilities to portray results of non-technical users are being considered. Very little of the output is visual in nature, and at a feature level there would be very few of these types of breakdowns for specific products/brands. This project is aimed directly at filling this gap.

IV. PROBLEM STATEMENT

Finding smartphone reviews is not the problem. There are millions of them, spread across dozens of platforms, covering virtually every model on the market. The problem is that none of the existing tools help users make sense of all that information in any meaningful depth.

Examining all reviews on one's own is simply not feasible. In the same manner, taking just a few samples arbitrarily and making conclusions about the entire product based on them is absolutely unprofessional because even the most prejudiced sample cannot provide a full understanding of the product when customers find it completely satisfactory. What one should do is look for some means of automatically analysing the great volume of data available and presenting the results in the most convenient way.

This can be considered the main challenge that should be addressed and which is reflected in the following statement: what methods could be applied for turning a number of opinions concerning some particular model of a smartphone into meaningful data about its qualities? This is what the author proposes to do.

V. OBJECTIVES

1. This project was developed using a well-defined vision behind all of the decisions made:
2. Sourcing reviews for smartphones from CSV datasets including several different brands
3. We can assess the sentiment towards an object by analysing large volumes of text using NLP techniques
4. Evaluate every single review and classify it as positive, negative or neutral
5. Do not stop at analysing general sentiment but provide ratings for individual features
6. Create a simple and understandable user interface for web application using Flask
7. Represent results graphically using pie chart and histogram instead of just numbers

VI. TECHNOLOGIES USED:

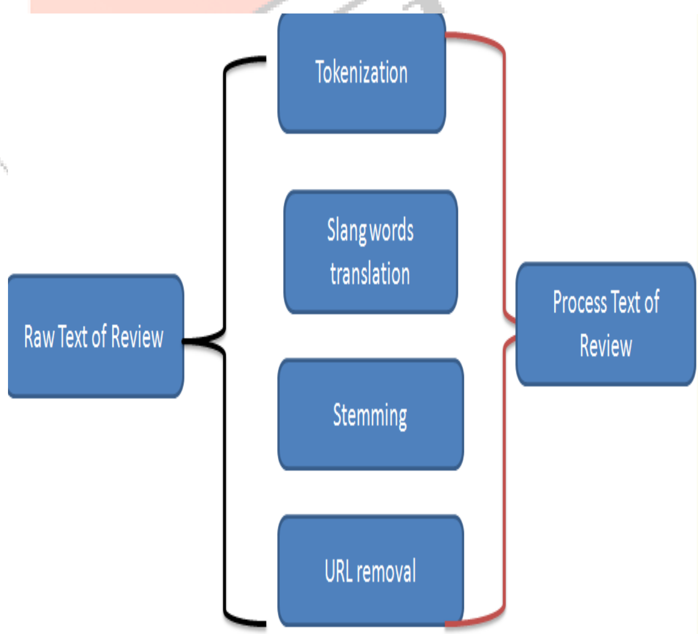
Technology	Purpose
Python	Backend Programming
Flask	Web Framework
Pandas	Data Processing
TextBlob	Sentiment Analysis
HTML	Frontend Structure
CSS	UI Styling
Chart.js	Graph Rendering
CSV Dataset	Review Storage

Table 1 Technology used

VII. METHODOLOGY

A. Data Preprocessing

The data collected from reviews usually isn't always ready to use. Prior to the actual analysis process, some basic data cleaning will occur to get everything prepared for analysis. In instances where there are any missing rows, these will be removed since missing data might lead to misleading conclusions. A case transformation is performed on all review text in lower case form so that "Battery" and "battery" will be considered as one single word and not two different words. All unimportant data, including punctuation and symbols, are removed because they serve no purpose other than adding to the noise.



B. Sentiment Analysis

Now, once the data set is cleaned, the TextBlob algorithm kicks into action. This algorithm will read the review and output a polarity value between -1 and +1. A value above 0 implies the review is a positive review; while a value below 0 is a negative review. A value of zero indicates a neutral sentiment. Although that may appear to be too simplistic, when applied to the reviews available in this data set — which comprise of "good camera," "quick draining battery," and "crisp display," amongst others — it does work fine.

C. Feature-Based Analysis

However, overall sentiment alone provides only partial insights. For more refined analysis, the system analyses each review based on keywords that relate to certain characteristics of the phones:

Feature	Keywords
Camera	camera, selfie, photo
Battery	battery, charging
Display	display, screen
Processor	gaming, processor

Table 2 Feature-Based Analysis

VIII. ADVANTAGES

1. Many reviews are looked at and explained in a time just a few minutes so you can save a lot of time that you would have spent reading reviews for hours.
2. The charts and graphs are easy to understand you do not need to be good with technology to see the results and know what they mean.
3. Our site looks at each feature of a product one by one this gives you useful information than just giving a product one rating. For example, if a phone has a camera but the battery does not last long this tells you more, about the phone than just a single rating.
4. The site is easy to use anyone can use it without knowing about data or how the site works.
5. Our site works with different kinds of smartphones you do not need to change anything you can just switch phones and keep looking.
6. Although much data is analysed within the time frame during which the website is working efficiently and fast, you will not need to spend extra time loading information.

IX. LIMITATIONS

1. The reliability of the system depends on the reliability of the input data – if there is low-quality data in the dataset, it will be reflected in the output as well.
2. Fake and incentivized reviews present a real problem. If manufactured feedback enters the dataset, it will silently pull the results in the wrong direction.
3. TextBlob is a basic sentiment tool. It handles straightforward opinions well but struggles with anything more subtle or layered.
4. Sarcasm and irony go completely undetected. A review that sounds positive on the surface but means the opposite will be misclassified every time.
5. There is no live data connection. The system depends on static CSV files, which means the analysis reflects past reviews rather than current ones.

X. RESULT AND DISCUSSION

A. Feature Comparison

The total rating of the iPhone 6 was 4.1 out of 5 from 13,256 unique review sources. Two of the four main areas that reviewers evaluated received good marks; however, the highest score went to "performances" which was rated 4.9 (out of 5) and was 0 negative votes from all the reviewers, and others like the operating system, and/or the camera performance were rated fairly good.

The battery and display were rated the lowest; the battery received a score of 3.6 out of 5, but scored even lower respectively at 3.4 out of 5. The number of negative reviews regarding battery was very large at 250 – making it the most complaining feature among all groups of reviewers.

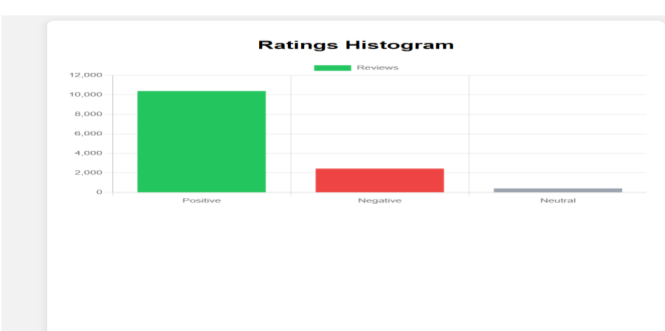
Parameter	Positive	Negative	Neutral	Rating	Sentiment
Battery	297	146	14	3.6	Positive
Camera	308	76	16	4.1	Positive
Display	412	250	23	3.4	Positive
Processor	58	0	1	4.9	Positive
Software	475	91	7	4.2	Positive

Table 3 Feature Comparison

Overall, the iPhone 6 has an overall good reception due to its reviews; but if the quality of the battery and/or display are important to your purchase, consider that when deciding on your purchase.

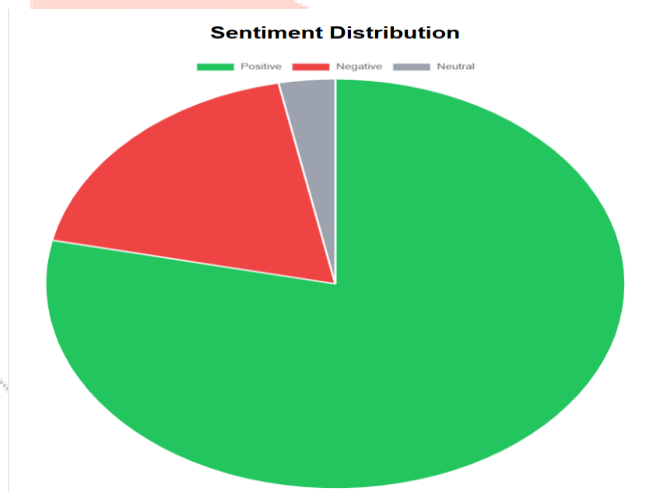
B. Sentiment Distribution

From the histogram, it's apparent that the iPhone 6 has been extremely popular among buyers as a whole. You can see how much higher the green, positive side of the chart is than the back, or negative, side. Approximately 10,000 of the reviews were positive compared to approximately 2,000 reviews were negative. The neutral reviews were almost not even noticeable. There are certainly some customers who have complaints about their purchases; however, this is to be expected on any product sold in such high quantities. All in all, most of the people who purchased the iPhone 6 have expressed happiness with their purchase based on all the data that is available; therefore, it is difficult to disagree with this statement.



C. Rating Histogram

The pie chart depicts the positive disproportion of sentiment toward iPhone 6 — represented in green and predominantly across the area of the chart — making it easy to tell that there is a positive experience for a majority of those reviewed; negative sentiment (in red) makes up no more than a quarter of the total aggregate sentiment; and neutral sentiment (in grey) appears so much smaller than either positive or negative that it hardly qualifies as an expression of either — it seems as if this small segment has been inadvertently placed at the end of the pie chart in order to highlight and give prominence to the other segments. Taken together, this evidence clearly reveals that among individuals that submitted reviews of the iPhone 6, the vast majority provided very positive reviews, the smaller but still significant number of individuals provided reviews of less than positive experience and nearly no individuals were expressing a neutral opinion on their review of the iPhone 6.



XI. FUTURE SCOPE

Although the system works adequately for what it currently does, there are certainly a number of significant updates that will enable it to do even more than what it is currently doing. The most important of these updates would be replacing the TextBlob algorithm with an algorithm based on Bert LSTM, since the ML models that use this type of net (BERT & Double LSTM) would be able to provide much better results when it comes to being able to interpret context, identify sarcasm, and processing natural, "messy" language that people actually use when they write real reviews.

When it comes to replacing static CSV files with the Amazon Product API, this would also have a significant effect on the way the system works today, since it would allow the system to retrieve up-to-date reviews and instead, be working with older review information. With regard to the multilingual support and the fact that many consumers worldwide buy a mobile device and write reviews in a language other than English would also be a significant area to revisit, as there are a lot of people who have reviewed products using their native language, and when those reviews are not included in the analysis, it does not provide the complete view of the marketplace.

Additionally, adding features such as built-in voice search, built-in recommendation engine, automatic updates to the dashboard would also change the overall function of the tool from something that is considered static to something that can be perceived as being more alive, and thus, provide a very different user experience. These enhancements to the tool will make it much more valuable for consumers that are looking for private and/or unbiased reviews from other people to help them make the best possible decision when purchasing a product, as well as, to brands who want to consistently be aware of what their customers are thinking and feeling over a longer time.

XII. CONCLUSION

Smartphone purchasing should be a straightforward experience; however, if you have ever spent hours sifting through customer rave and grumble reviews on Amazon and still find yourself clueless about your new smartphone purchase, you understand how incredibly frustrating the whole process can be; that frustration actually drove the creation of this product.

The tool being created allows a user to enter in their brand and model of smartphone and then provides a complete breakdown of how customers have rated the battery, camera, display, processor, and software within seconds! The visual presentation for all of these items/create a clear description for the user, and no technical expertise are needed to read/understand/use the data provided.

After putting this tool through its paces with the iPhone 6, testing approximately 13,000 reviews, the tool worked as designed—the processor and software received top marks while the battery and display provided the lowest ratings, and that type of detailed information would have taken a user hours to research manually.

On the sentiment analysis side, TextBlob has been reliable in extracting the sentiment from that large of a data set; on occasion and depending upon the author's use of language, TextBlob sometimes misses sarcasm or has difficulty extracting sentiment from emotions more nuanced in their presentation. That presents the biggest potential limitation of the tool, and it is the most logical area for future improvement.

Overall, the project accomplished its mission of taking a dizzying mountain of customer reviews and presenting them as a truly useful tool for consumers when they are making smarter informed purchasing decisions for smartphones.

REFERENCES

[1] B. Pang and L. Lee wrote a paper called "Opinion Mining and Sentiment Analysis" in the journal Foundations and Trends in Information Retrieval. This paper was published in 2008. It has 135 pages.

[2] S. Bird, E. Klein and E. Loper wrote a book called Natural Language Processing with Python. This book was published by O'Reilly Media in 2009.

[3] J. Leskovec, A. Rajaraman and J. Ullman wrote a book called Mining of Massive Datasets. This book was published by Cambridge University Press in 2014.

[4] A. Go, R. Bhayani and L. Huang wrote a report called "Twitter Sentiment Classification using Distant Supervision". This report was published by Stanford University in 2009.

[5] B. Liu wrote a paper called "Sentiment Analysis and Opinion Mining". This paper was published by Morgan and Claypool Publishers in 2012.

[6] A. Agarwal, B. Xie, I. Vovsha, O. Rambow and R. Passonneau wrote a paper called "Sentiment Analysis of Twitter Data". This paper was presented at the Workshop on Language in Social Media in 2011.

[7] M. Hu and B. Liu wrote a paper called "Mining and Summarizing Customer Reviews". This paper was presented at the ACM SIGKDD International

Conference on Knowledge Discovery and Data Mining in 2004.

[8] R. Feldman wrote a paper called "Techniques and Applications for Sentiment Analysis". This paper was published in the journal Communications of the ACM in 2013.

[9] V. Vapnik wrote a book called The Nature of Statistical Learning Theory. This book was published by Springer in 1995.

[10] T. Mikolov, K. Chen G. Corrado and J. Dean wrote a paper called "Efficient Estimation of Word Representations in Vector Space". This paper was presented at the International Conference on Learning Representations in 2013.

[11] J. Devlin, M. Chang, K. Lee and K. Toutanova wrote a paper called "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". This paper was presented at the NAACL-HLT conference in 2019.

[12] F. Pedregosa and others wrote a paper called "Scikit-learn: Machine Learning in Python". This paper was published in the Journal of Machine Learning Research in 2011.

[13] W. McKinney wrote a paper called "Data Structures for Statistical Computing in Python". This paper was presented at the Python in Science Conference in 2010.

[14] The Python Software Foundation has a website with information about the Python Language Reference and Documentation. You can find it at https://www.python.org.

[15] The Pallets Projects has a website with information about the Flask Web Framework Documentation. You can find it at https://flask.palletsprojects.com.

[16] The TextBlob Development Team has a website with information about TextBlob: Simplified Text Processing Documentation. You can find it at https://textblob.readthedocs.io.

[17] The Chart.js Development Team has a website with information about Chart.js Open Source HTML5 Charts Documentation. You can find it at https://www.chartjs.org.

[18] The Amazon Customer Reviews Dataset is available on the Kaggle Open Dataset Repository. You can find it at https://www.kaggle.com.

[19] D. Maynard and M. Greenwood wrote a paper called "Who Cares about Sarcastic Tweets? Investigating the Impact of Sarcasm on Sentiment Analysis". This paper was presented at the Language Resources and Evaluation Conference in 2014.

[20] S. Poria, E. Cambria, R. Bajpai and A. Hussain wrote a paper called "A Review of Affective Computing: From Unimodal Analysis, to Multimodal Fusion". This paper was published in the journal Information Fusion in 2017.

.