Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2024; Vol 13: Issue 8
Open Access

Cutting-Edge Deep Learning Models for Enhanced Breast Can-cer Detection: Comparative Analysis of YOLOv5, YOLOv8 and YOLOv9 Models

Nikhila Kathirisetty[1], Sudabathula Vijay Sai Kumar [2], Yagateela Pandu Rangaiah [3], Kottu Santosh Kumar [4], Muni Sekhar Velpuru [5], Farhana Begum [6]

[1]Assistant Professor, Department of CSD, Vardhaman College of Engineering, Hyderabad, India
[2]Department of AI & ML, Vardhaman College of Engineering, Hyderabad, India
[3]Associate Professor, Department of ECE, Institute of Aeronautical Engineering, Hyderabad, India.
[4]Assistant Professor, Department of IT, Vardhaman College of Engineering, Hyderabad, 50128, India
[5] Associate Professor, Department of IT, Vardhaman College of Engineering, Hyderabad, 50128, India
[6] Assistant Professor, Department of CSD, Vardhaman College of Engineering, Hyderabad, India

Cite this paper as: Nikhila Kathirisetty, Sudabathula Vijay Sai Kumar , Yagateela Pandu Rangaiah , Kottu Santosh Kumar , Muni Sekhar Velpuru , Farhana Begum (2024). Cutting-Edge Deep Learning Models for Enhanced Breast Can-cer Detection: Comparative Analysis of YOLOv5, YOLOv8 and YOLOv9 Models. *Frontiers in Health Informatics*, *13 (8) 527-551*

Abstract: Breast cancer is a leading cause of mortality among women worldwide, necessitating advancements in early and accurate detection methods. This study explores the efficacy of cutting-edge deep learning techniques, specifically YOLOv5, YOLOv8, and YOLOv9, for breast cancer detection. The YOLO (You Only Look Once) family of algorithms, known for their speed and precision, were employed to identify cancerous lesions in mammographic images. In our comparative analysis, YOLOv5 achieved a precision confidence curve of 1.00 at 0.786, an F1 confidence curve of 0.76 at 0.377, a precision-recall curve with an mAP@0.5 of 0.770, and a recall confidence curve of 0.92 at 0.000. YOLOv8 demonstrated improved performance with a precision confidence curve of 1.00 at 0.583, an F1 confidence curve of 0.97 at 0.575, a precision-recall curve with an mAP@0.5 of 0.969, and a recall confidence curve of 0.96 at 0.000. YOLOv9 showed the highest effectiveness with a precision confidence curve of 1.00 at 0.657, an F1 confidence curve of 0.97 at 0.467, a precision-recall curve with an mAP@0.5 of 0.974, and a recall confidence curve of 0.97 at 0.000. These results underscore the potential of advanced deep learning models in enhancing breast cancer detection. Future work will focus on refining these models to improve their robustness and applicability across diverse clinical settings. The integration of these techniques into routine screening processes could significantly advance early detection and treatment outcomes for breast cancer patients.
Keywords: Object Detection, Deep Learning, YOLO, Early Detection, Breast Cancer Detection

1. Introduction

Breast cancer remains a significant global health issue, being the most frequently diagnosed cancer and the leading cause of cancer-related deaths among women. In 2020, there were approximately 2.3 million new cases, and 685,000 deaths attributed to breast cancer, accounting for nearly 12% of the global cancer burden [2]. The incidence rates vary widely, with the highest rates observed in Australia/New Zealand (95.5 per 100,000) and the lowest in South-Central Asia (26.2 per 100,000) [2]. Mortality rates also show considerable geographic disparity, with the highest rates in Melanesia (37.5 per 100,000) and the lowest in Eastern Asia (9.8 per 100,000) [2]. While incidence rates are rising in many low- and middle-income countries (LMICs), mortality rates are declining in high-income countries (HICs), reflecting improved access to early detection and treatment [3][4]. However, the increasing burden in LMICs necessitates urgent cancer control measures to address rising incidence and mortality rates [4]. Overall, the trends indicate a complex interplay of demographic changes, lifestyle factors, and healthcare access that shapes breast cancer outcomes globally [1][3].

The integration of deep learning models, particularly the YOLO (You Only Look Once) model, is pivotal in enhancing breast cancer detection efficiency and accuracy. Research indicates that deep learning techniques significantly reduce assessment times and false positives, thereby facilitating early diagnosis and improving treatment outcomes [5]. For instance, a study demonstrated that a deep convolutional neural network incorporating YOLO achieved an impressive accuracy of 93% in detecting breast lesions, underscoring the model's effectiveness in expediting diagnosis [6]. Furthermore, deep learning methods have shown promise in automating the analysis of mammograms and histopathological images, which can alleviate the workload on pathologists and minimize human error [7][8]. Advancements in deep learning not only enhance diagnostic capabilities but also contribute to tailored treatment plans, ultimately leading to better patient outcomes and reduced mortality rates associated with breast cancer [9]. Thus, the application of these technologies is increasingly vital in modern healthcare for timely interventions.

YOLOv5, part of the You Only Look Once (YOLO) series, is a state-of-the-art real-time object detection algorithm that significantly enhances the capabilities of its predecessors. It employs a single neural network to predict bounding boxes and class probabilities directly from full images, making it efficient for various applications, including traffic sign recognition and small target detection [13][11]. YOLOv5 incorporates advanced features such as the Squeezed-and-Excitation (SE) attention mechanism and optimized loss functions, which improve detection accuracy and speed, achieving up to 99.4% mean Average Precision (mAP) in specific applications [10]. Additionally, modifications like Multi-Scale Feature Fusion and data augmentation strategies have been introduced to enhance performance, particularly for small objects [12]. The model is versatile, capable of processing various media types and providing real-time outputs, making it suitable for applications in autonomous systems and traffic management [14]. Overall, YOLOv5 represents a significant advancement in the field of computer vision.

YOLOv5 has demonstrated significant improvements in object detection accuracy compared to its predecessors, primarily through various enhancements tailored to specific challenges. For instance, the UAV-YOLOv5 variant achieved a mean Average Precision (mAP) increase of 1.5% and 1.0% at different thresholds compared to the default YOLOv5s, while also reducing model complexity [15]. The RS-YOLOv5 algorithm improved average accuracy by 2.5% in remote sensing applications by employing dynamic convolution and a decoupled detection head [16]. Additionally, the HSA-YOLOv5 model reported a remarkable 6.42 percentage point increase in mAP for raspberry detection, showcasing its effectiveness in handling similar color challenges [17]. Furthermore, MCF-YOLOv5 enhanced small target detection accuracy by 3.3% to 3.6% through improved feature extraction and attention mechanisms [19]. Collectively, these advancements underscore YOLOv5's adaptability and enhanced performance across diverse detection scenarios, solidifying its position as a leading object detection framework [18].

YOLOv8's architecture enhances object detection accuracy in real-world scenarios through several innovative modifications aimed at addressing specific challenges, particularly in detecting small and occluded objects. The introduction of specialized feature extraction modules, such as the C4 module, improves the retention of crucial features for small targets, which are often lost in traditional networks [20]. Additionally, the integration of multi-scale attention mechanisms and deformable convolutional layers in models like EDGS-YOLOv8 and multi-YOLOv8 further refines feature extraction and enhances detection capabilities in complex environments [21][24]. The use of advanced loss functions, such as WIoU and $\alpha*$-WIoU v3, optimizes bounding box regression, leading to improved accuracy in densely populated scenarios [23][24]. Collectively, these enhancements result in significant performance improvements, with reported increases in mean average precision (mAP) ranging from 3.1% to 6.0% over the baseline YOLOv8 model [21][23].

The anchor-free architecture of YOLOv8 significantly enhances its performance across various object detection tasks by improving feature extraction and versatility. This architecture eliminates the reliance on predefined anchor boxes,

allowing for more flexible detection of objects, particularly in complex scenes and varying scales. For instance, modifications to YOLOv8, such as the integration of a small object detection head and mixed local channel attention, have led to a 6.7% improvement in mean average precision (mAP) for occluded objects compared to the original YOLOv8 model [25]. Additionally, enhancements like the C4 feature extraction module and attention mechanisms have improved detection accuracy for small targets in UAV-captured images, achieving a 4.9% increase in mAP[27]. Furthermore, lightweight adaptations of YOLOv8 have demonstrated a balance between detection accuracy and computational efficiency, achieving significant reductions in model size while maintaining high performance [28]. Overall, YOLOv8's anchor-free design facilitates superior adaptability and effectiveness in diverse detection scenarios [26][29].
YOLOv9 demonstrates significant advancements in both accuracy and speed compared to its predecessors. One key improvement is the introduction of an additional small target detection head, which enhances the model's ability to accurately identify small objects, resulting in an 8% increase in detection accuracy while reducing computational complexity by 5.7% and the number of parameters by 44.1%[30]. Furthermore, the IDP-YOLOv9 variant incorporates a three-layer routing attention mechanism, which improves recognition accuracy by 6.8% in adverse weather conditions, showcasing its robustness in challenging environments [31]. These enhancements contribute to YOLOv9's overall efficiency, making it more effective in real-time applications. In contrast, while YOLOv10 further optimizes performance and reduces latency, YOLOv9's specific improvements in small object detection and adaptability to various conditions mark it as a notable step forward in the YOLO series [32]. Thus, YOLOv9 effectively balances accuracy and computational efficiency, addressing limitations seen in earlier models.

2. Materials and Methods

Sentiment analysis has become a powerful technique in ongoing research to gain insight from customer feedback and understand what customers think about a product or service. This research involves in-depth evaluation and application at all stages, from data collection and preprocessing to theoretical analysis, data exploration, visualization, and interpretation. In this study, we used the data set on the data.world website. The database has more than 30,000 reviews from brands such as OnePlus and Redmi. All reviews are presented in JSON format and include details about customer thoughts and opinions about each brand. This database is very useful for analyzing products using sentiment analysis techniques. By summarizing the opinions expressed in reviews, insights can be gained about customer satisfaction, product strengths, areas for improvement, and general understanding of the products. We have created a detailed guide that will serve as a guide for the entire process.
The flowchart below demonstrates the steps involved in conducting a robust product review analysis.

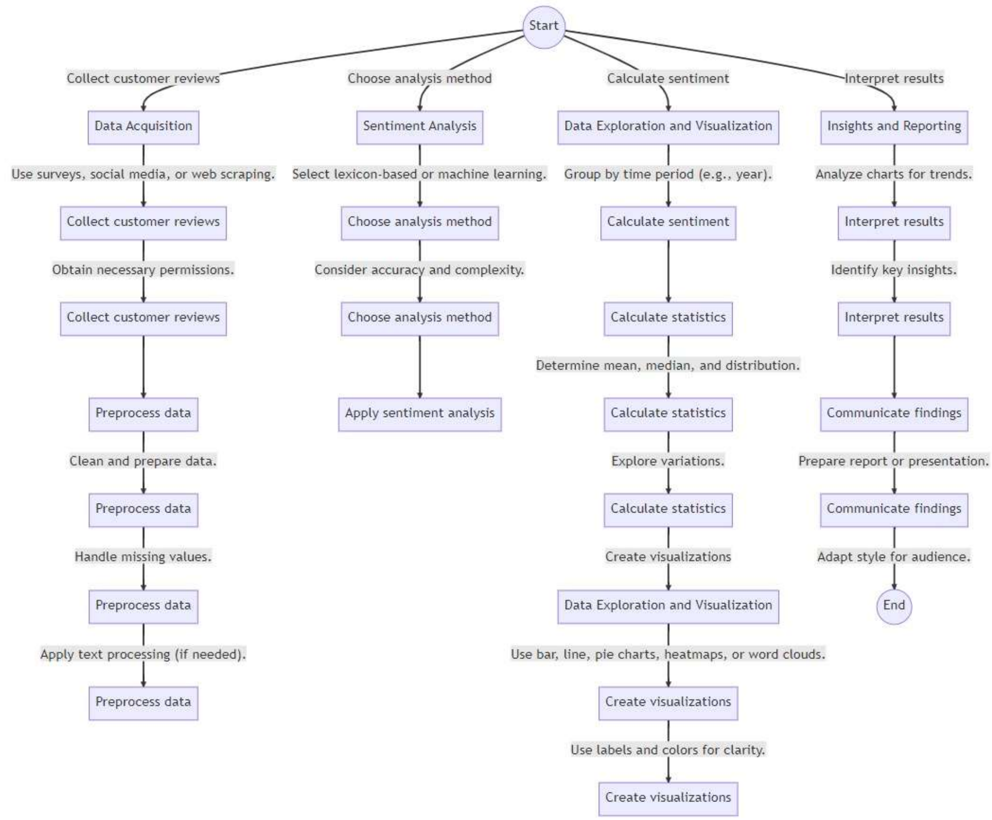


Figure 1. A Flowchart Diagram of Generalized Method of Algorithm Implementation

To clearly understand the flowchart given below are key details of the stages of the flowchart:

- Data Acquisition: The first step in the process is to collect customer reviews, which can be done through a variety of methods such as surveys, social media, or web scraping. This is a delicate stage because it determines the quality and quantity of the product. Approval from relevant organizations is also required to comply with data privacy laws.
- Preprocessing: After receiving the customer's ID, the next step is to prepare the data. Preprocessing of data sets includes routine cleaning and data preparation by handling missing values. Word processing includes tokenization, stemming, lemmatization, word removal, etc. where necessary. may include transactions. This stage cannot be ignored because it ensures that the data is in a format suitable for analysis and increases the accuracy and reliability of the data.
- Sentiment Analysis: The main structure of the product review process is sentiment analysis, whose purpose is to calculate and classify opinions in customer reviews. There are two main ways to analyze emotions:

1. Lexicon-based Methods: Dictionary-based methods rely on pre-emotional annotations that label words or phrases associated with positive, negative, or neutral emotions. This method generally works by counting the number of positive and negative words appearing in the text and calculating the total score of the thought based on calculation and preliminary weighting.
2. Machine Learning-based Methods: On the other hand, machine learning-based methods analyze sentiment by training on samples of registered customer review articles. This model learns to classify new teachers' opinions based on patterns and characteristics observed in new feedback patterns in teaching materials. Commonly used machine learning algorithms for hypothesis testing include support vector machines, naive Bayes, and deep learning models such as neural networks and transformers.

- Data Exploration and Visualization: After using the method of choosing thought theory, the next step is to investigate and see the results. This usually means taking data over time (such as years) and calculating various statistics such as mean, median, and distribution of emotional scores.

- Visualization Techniques: A variety of visualization techniques such as line charts, graphs, line diagrams, heat maps, or word cloud can be used to effectively communicate research. These visualizations can help you identify trends, patterns, and key insights in the data. It is important to use appropriate text, colors and images for clarity and ease of interpretation.
- Insights and Reporting: The final stage of the process involves visual analysis and graphs to identify key insights and patterns in the product analysis that can be found in product analysis reports. This information may be used to inform product development, marketing strategies or customer service.

By following this approach, researchers and analysts can use analysis theory to make the most of customer reviews, ultimately leading to informed decisions and improved products.

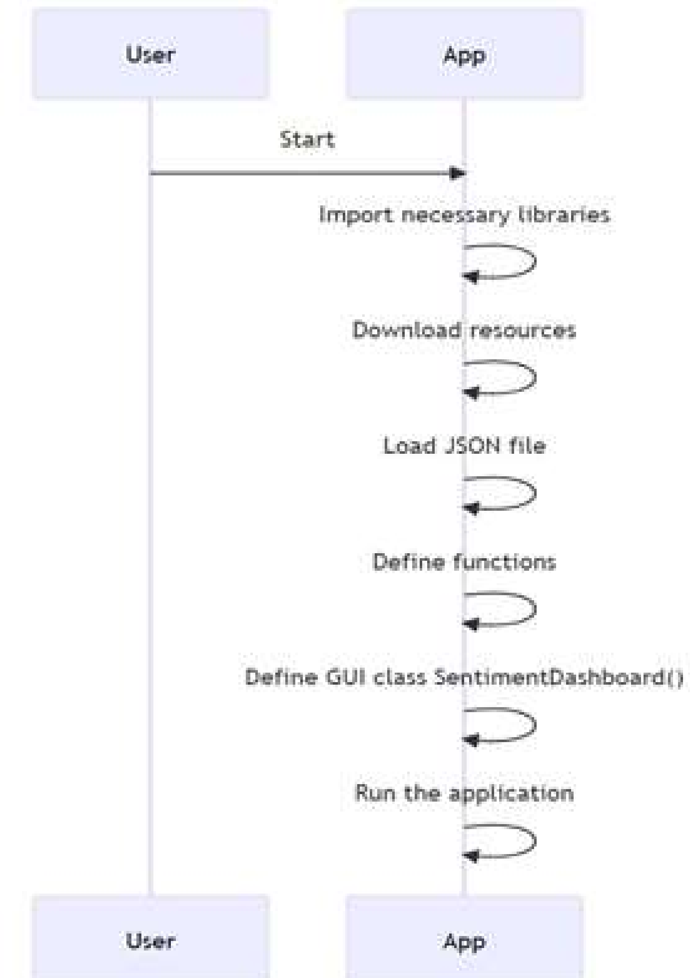
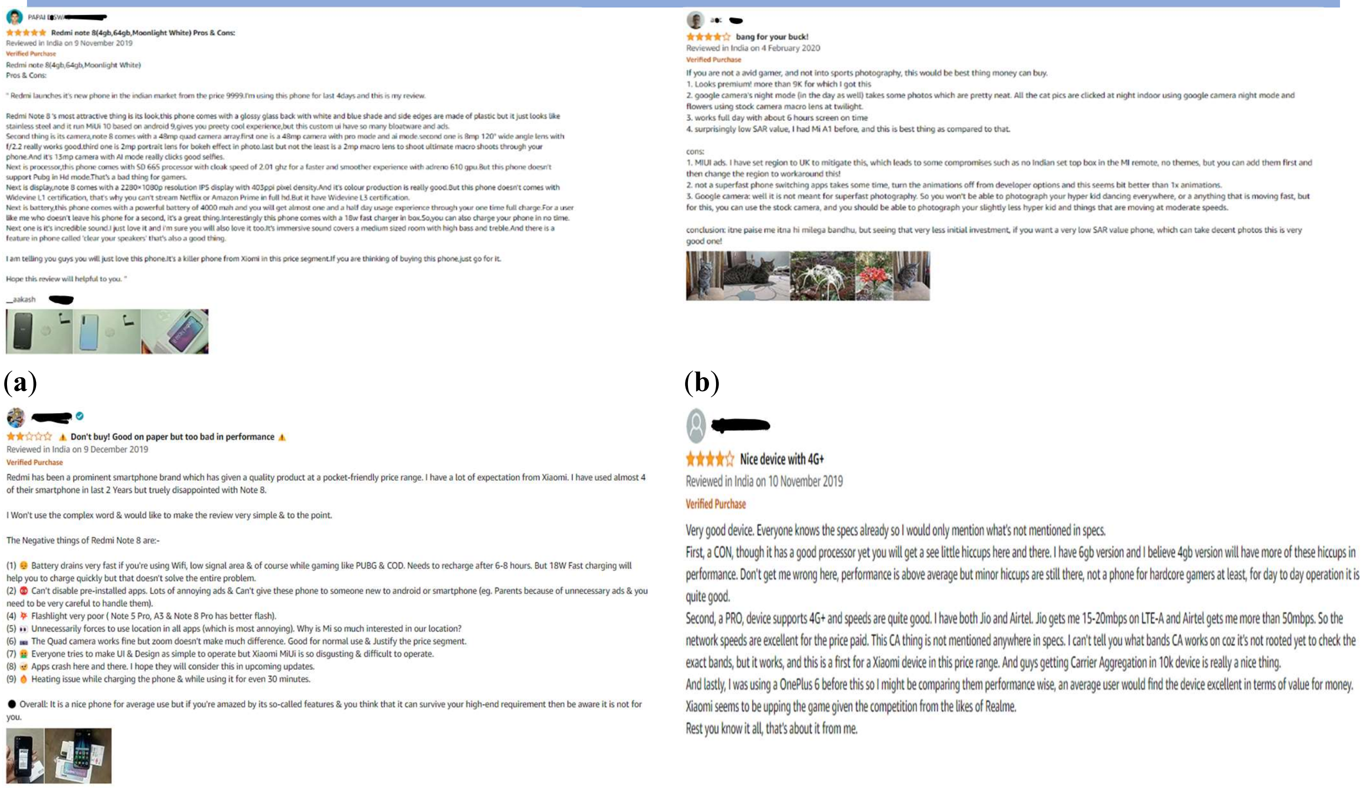


Figure 2. Flowchart of General Code Implementation of Dashboard

This flowchart walks you through the process of creating a sentiment analysis dashboard. Start by deploying the appropriate libraries and downloading resources as needed. A JSON file containing the search query data is then loaded. The next step involves defining the tasks and other capabilities required for analysis. The core of the process is the definition of a GUI class called "sentiment_dashboard()"; this makes it possible to create a graphical user interface for a sentiment analysis dashboard. Finally, the entire application can display a view of the analytics dashboard for user interaction and analysis. In general, this flowchart contains sequential steps from planning to execution of a control panel application.

Frontiers in Health Informatics
ISSN-Online: 2676-7104
2024; Vol 13: Issue 8



PAPAI [redacted]

Redmi note 8(4gb,64gb,Moonlight White) Pros & Cons:

Reviewed in India on 9 November 2019

Verified Purchase

Redmi note 8(4gb,64gb,Moonlight White)
Pros & Cons:

" Redmi launches it's new phone in the indian market from the price 9999.I'm using this phone for last 4days and this is my review.

Redmi Note 8 's most attractive thing is its look,this phone comes with a glossy glass back with white and blue shade and side edges are made of plastic but it just looks like stainless steel and it run MiUi 10 based on android 9,gives you preety cool experience,but this custom ui have so many bloatware and ads.
Second thing is its camera,note 8 comes with a 48mp quad camera array.first one is a 48mp camera with pro mode and ai mode.second one is 8mp 120° wide angle lens with f/2.2 really works good.third one is 2mp portrait lens for bokeh effect in photo.last but not the least is a 2mp macro lens to shoot ultimate macro shoots through your phone.And it's 13mp camera with AI mode really clicks good selfies.
Next is processor,this phone comes with SD 665 processor with cloak speed of 2.01 ghz for a faster and smoother experience with adreno 610 gpu.But this phone doesn't support Pubg in Hd mode.That's a bad thing for gamers.
Next is display,note 8 comes with a 2280×1080p resolution IPS display with 403ppi pixel density.And it's colour production is really good.But this phone doesn't comes with Widevine L1 certification, that's why you can't stream Netflix or Amazon Prime in full hd.But it have Widevine L3 certification.
Next is battery,this phone comes with a powerful battery of 4000 mah and you will get almost one and a half day usage experience through your one time full charge.For a user like me who doesn't leave his phone for a second, it's a great thing.Interestingly this phone comes with a 18w fast charger in box.So,you can also charge your phone in no time.
Next one is it's incredible sound.I just love it and i'm sure you will also love it too.It's immersive sound covers a medium sized room with high bass and treble.And there is a feature in phone called 'clear your speakers' that's also a good thing.

I am telling you guys you will just love this phone.It's a killer phone from Xiomi in this price segment.If you are thinking of buying this phone,just go for it.

Hope this review will helpful to you. "

__aakash [redacted]

(a)

[redacted]

bang for your buck!

Reviewed in India on 4 February 2020

Verified Purchase

If you are not a avid gamer, and not into sports photography, this would be best thing money can buy.
1. Looks premium! more than 9K for which I got this
2. google camera's night mode (in the day as well) takes some photos which are pretty neat. All the cat pics are clicked at night indoor using google camera night mode and flowers using stock camera macro lens at twilight.
3. works full day with about 6 hours screen on time
4. surprisingly low SAR value, I had Mi A1 before, and this is best thing as compared to that.

cons:
1. MIUI ads. I have set region to UK to mitigate this, which leads to some compromises such as no Indian set top box in the MI remote, no themes, but you can add them first and then change the region to workaround this!
2. not a superfast phone switching apps takes some time, turn the animations off from developer options and this seems bit better than 1x animations.
3. Google camera: well it is not meant for superfast photography. So you won't be able to photograph your hyper kid dancing everywhere, or a anything that is moving fast, but for this, you can use the stock camera, and you should be able to photograph your slightly less hyper kid and things that are moving at moderate speeds.

conclusion: itne paise me itna hi milega bandhu, but seeing that very less initial investment, if you want a very low SAR value phone, which can take decent photos this is very good one!

(b)

[redacted]

Don't buy! Good on paper but too bad in performance

Reviewed in India on 9 December 2019

Verified Purchase

Redmi has been a prominent smartphone brand which has given a quality product at a pocket-friendly price range. I have a lot of expectation from Xiaomi. I have used almost 4 of their smartphone in last 2 Years but truely disappointed with Note 8.

I Won't use the complex word & would like to make the review very simple & to the point.

The Negative things of Redmi Note 8 are:-

(1) Battery drains very fast if you're using Wifi, low signal area & of course while gaming like PUBG & COD. Needs to recharge after 6-8 hours. But 18W Fast charging will help you to charge quickly but that doesn't solve the entire problem.
(2) Can't disable pre-installed apps. Lots of annoying ads & Can't give these phone to someone new to android or smartphone (eg. Parents because of unnecessary ads & you need to be very careful to handle them).
(4) Flashlight very poor (Note 5 Pro, A3 & Note 8 Pro has better flash).
(5) Unnecessarily forces to use location in all apps (which is most annoying). Why is Mi so much interested in our location?
(6) The Quad camera works fine but zoom doesn't make much difference. Good for normal use & Justify the price segment.
(7) Everyone tries to make UI & Design as simple to operate but Xiaomi MiUi is so disgusting & difficult to operate.
(8) Apps crash here and there. I hope they will consider this in upcoming updates.
(9) Heating issue while charging the phone & while using it for even 30 minutes.

● Overall: It is a nice phone for average use but if you're amazed by its so-called features & you think that it can survive your high-end requirement then be aware it is not for you.

(c)

[redacted]

Nice device with 4G+

Reviewed in India on 10 November 2019

Verified Purchase

Very good device. Everyone knows the specs already so I would only mention what's not mentioned in specs.
First, a CON, though it has a good processor yet you will get a see little hiccups here and there. I have 6gb version and I believe 4gb version will have more of these hiccups in performance. Don't get me wrong here, performance is above average but minor hiccups are still there, not a phone for hardcore gamers at least, for day to day operation it is quite good.
Second, a PRO, device supports 4G+ and speeds are quite good. I have both Jio and Airtel. Jio gets me 15-20mbps on LTE-A and Airtel gets me more than 50mbps. So the network speeds are excellent for the price paid. This CA thing is not mentioned anywhere in specs. I can't tell you what bands CA works on coz it's not rooted yet to check the exact bands, but it works, and this is a first for a Xiaomi device in this price range. And guys getting Carrier Aggregation in 10k device is really a nice thing.
And lastly, I was using a OnePlus 6 before this so I might be comparing them performance wise, an average user would find the device excellent in terms of value for money.
Xiaomi seems to be upping the game given the competition from the likes of Realme.
Rest you know it all, that's about it from me.

(d)

Figure 3. Example for Reviews of Phone Redimi Note 8: (a); (b); (c); (d) Images of real Amazon Reviews on the Product

The above are the comments from customers given for the phone "Redimi Note 8". As seen the customers clearly state the advantages and disadvantages of the phone. This gives the manufacturers a clear idea of their areas of improvement for the future phone by analyzing the current phone reviews. Such a huge amount of data is not easy to comb through. Thus enters the reason for our experiment on dashboards for product analysis where data like the above can be taken and given to such dashboard.

2.1 Textual Analyzer using NLTK + TextBlob Algorithm

NLTK, together with TextBlob, provides a simple yet effective method for sentiment analysis, making it an attractive choice for many NLP tasks. Although it lacks the intelligence and expertise of models such as BERT, this method stands out for its simplicity and speed, making it suitable for situations where rapid analysis and detailed understanding are more important. This implementation involves the use of NLTK for text preprocessing and TextBlob for polarity analysis [22].

NLTK (Natural Language Toolkit) provides libraries and tools for language processing, including tokenization, stemming, and blockword removal. Using the capabilities of NLTK, data files can be pre-processed to remove noise and remove important features, ensuring good performance when analyzing emotions.

TextBlob is a simple text processing library based on NLTK and provides easy integration for a variety of NLP tasks, including sentiment analysis [21].

The process of sentiment analysis using NLTK + TextBlob starts with the initial data where written words are cleaned and put into words or tokens. Each tokenized comment is then passed through TextBlob's sentiment analysis, which assigns a polarity score (positive, negative, or neutral) representing sentiment. The polarity score is determined by the

presence of positive or negative words in the text, using a threshold to define the boundaries of each group of thoughts [23].
The implementation of sentiment analysis using NLTK + TextBlob involves several key steps, each contributing to the overall functionality of the system.

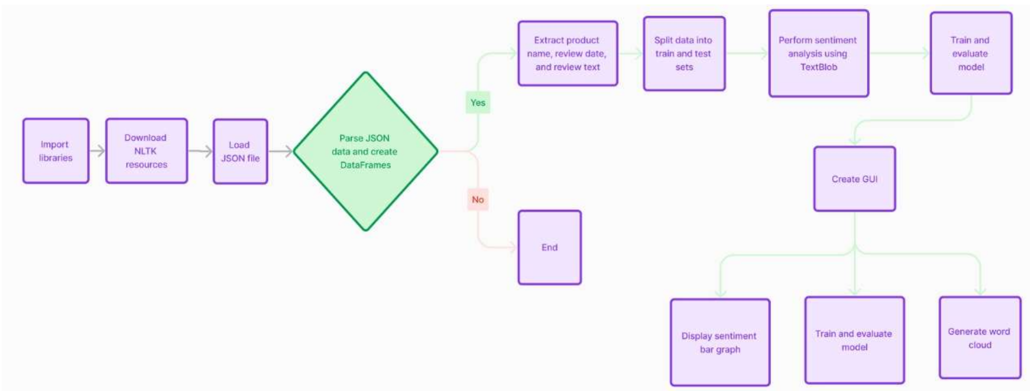


Figure 4. Flowchart of Code Implementation of NLTK + TextBlob Algorithm

This flowchart describes the emotional analysis of reading materials. First you need to import the library and download the NLTK resources for the NLP project. Item names, review dates, and tags are extracted from JSON files that are loaded, parsed, and organized into data frames. The data is divided into testing and training sets. TextBlob is used for sentiment analysis and training and evaluating models. The GUI displays the results and shows the export purpose using a bar chart. The model was further developed to create a word cloud suitable for content analysis. This systematic approach enables emotional assessment, supports intuition-based decision making, and allows for prior knowledge and insight.

- Initial data and upload: The first steps involve uploading a file containing the analysis object stored in JSON format. To facilitate further analysis, reviews were extracted and placed into data frames for each item. All reviews include text and relevant dates.
- Sentiment Analysis with NLTK + TextBlob: Sentiment analysis using NLTK and TextBlob libraries. All comments are analyzed for sentiment using the TextBlob sentiment analysis module. TextBlob provides a polarity score that represents sentiment (positive, negative, or neutral) based on the presence of positive or negative words in the text.
- Calculation of opinion percentage: The next step after opinion analysis is to calculate the opinion percentage for each opinion group (good, negative, moderate). This process requires repeating the message and counting the number of occurrences of each thought tag. Then, use all the words to determine the percentage of each set of thoughts.
- Creating a Sentiment Analysis Dashboard: The dashboard is designed to provide users with information about sentiment analysis in product reviews.It provides three main functions:

1. Visual analysis: Users can use graphs to visualize wiring. These charts show the percentage of good, bad, and average reviews for each product, allowing users to quickly understand their need to take care of different things.
2. Word cloud creation: The dashboard helps create a word cloud that shows the most common words in the word. These features help users understand all the ideas and content discussed by providing information about the key concepts and topics discussed in the message.
3. Accuracy evaluation: Users can evaluate the accuracy of sample requirements directly from the control panel. By training the model and calculating metrics such as accuracy and F1 score, users can evaluate the model's performance in opinion classification and provide useful feedback on confidence in survey results.

Overall, this application uses the simplicity and efficiency of NLTK + TextBlob for sentiment analysis, providing solutions for analyzing text files and eliminating perspective.

Although simple and fast, NLTK + TextBlob may exhibit limitations in performance. Identify negative emotions with high accuracy. Failure to recognize negative emotions such as sarcasm or irony can lead to negative emotions, affecting the overall accuracy of the analysis. Additionally, relying on lexical analysis and prioritization techniques to classify concepts can lead to decreased performance, especially in contexts where contexts and voice play an important role in emotional decision making.

In performance, when using NLTK + TextBlob to analyze the meaning of the content, the accuracy is 73.986365 and the F1 score is 66.012294. These measures involve a balance between simplicity and accuracy inherent in the NLTK + TextBlob approach. While it is suitable for tasks that require rapid observation and emotional evaluation, it may be disadvantageous in areas where clarity and detailed understanding of information is required. But NLTK + TextBlob is still an important tool in the NLP toolset and provides a lightweight and easy-to-use tool for sentiment analysis and word processing.

2.2 Textual Analyzer using VADER Algorithm

VADER (Valence Aware Dictionary and sEtiment Reasoner) is a dictionary and rule-based assessment tool specifically designed to analyze emotions expressed in text. VADER, developed by researchers at the Georgia Institute of Technology, is particularly suitable for analyzing sentiment in newspapers, online comments, and other informal communication rules [26].

One of the main features of VADER is its dictionary, which contains a carefully selected list of words and their high scores. Each word in the dictionary is given a polarity score, indicating how well or poorly related the word is. In addition, the dictionary contains rules and instructions for dealing with emotions, negations, spelling, punctuation, and expressions that are often found in incorrect texts.

VADER uses logic-based logic that communicates the polarity score with the code to process the description to create the overall score of the script section. The emotional score ranges from -1 (very bad) to +1 (very good), with 0 meaning neutral. This measure allows VADER to capture nuances in thinking and interpret clear and ambiguous statements in text [24, 25].

Another important feature of VADER is its ability to manage expressions and slang popular in social media and online communication. By combining multiple languages and contexts, VADER is able to interpret emotions in false text, making it an important tool for analyzing emotions in various domains.

There are several steps aimed at analyzing the sentiment of the text using VADER (Valence Aware Dictionary and sEtiment Reasoner) [28].

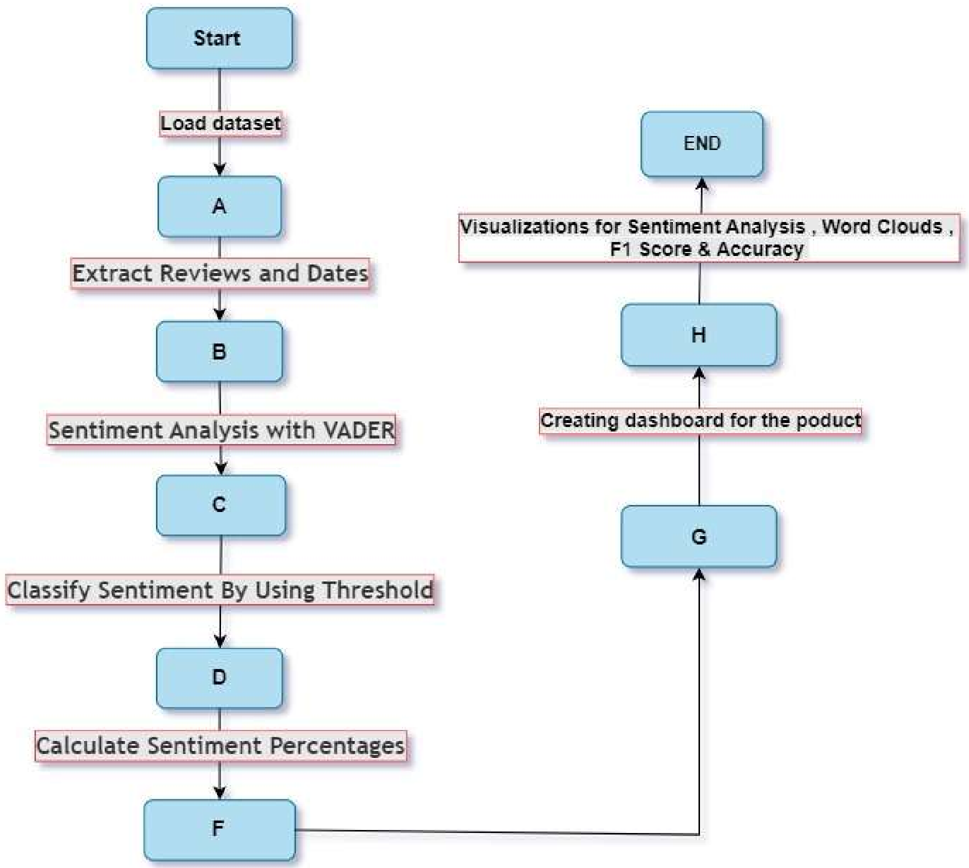


Figure 5. Flowchart of Code Implementation of VADER Algorithm

- Data Preprocessing and Loading: The sentiment analysis process begins with loading data containing analysis objects stored in JSON format. All individual reviews are extracted and placed in separate files by item.
- Sentiment analysis using VADER: After pre-processing and loading the data, sentiment analysis is performed on the comments. This code uses the VADER logic analysis tool, a key component of the Natural Language Toolkit (NLTK). Using VADER, the framework extracts a numerical representation of emotions from the content of the analysis.
- Classification of opinions using threshold: The polarity score obtained by VADER is used as the basis for classifying the opinions of each analysis. This classification system converts numerical scores into easily interpretable categories: good, bad, or neutral. This classification is made by reaching positive and negative results. Analyzes with polarity scores above a threshold were classified as positive. Conversely, reviews below a set threshold are classified as negative and trigger customer complaints. Reviews that fall within the specified range around zero are classified as average and indicate neither positive nor negative feedback.
- Calculate the percentage of opinions: Calculate the percentage of opinions for each category (positive, negative and neutral). This calculation provides a high-level view of the overall distribution across each product review. This calculation uses the sentiment classification (positive, negative, or neutral) assigned to each analysis in the previous step. This calculation results in the percentage of positive, negative, and average reviews for each product.
- Visualization of Sentiment Analysis: See percentage sentiment using the bar chart to get a deeper understanding. These recommendations provide a clear and transparent view of the customer's positive opinion of each product. Each bar in the visualization shows the percentage of positive, negative, and average reviews for a particular product over the years. This visual representation helps provide a clear understanding of how the customer's perception of the product changes over time. By analyzing trends in this view, product managers can identify moments of customer satisfaction or

dissatisfaction and use the context of the analysis to gain deeper insight into the root cause. Product Sentiment Dashboard: The final step involves integrating sentiment into the results analysis (visualizations, word clouds, etc.) into the product's dashboard. This dashboard acts as a central dashboard for product managers, providing a holistic view of customer needs. Control panels can contain a variety of content, including:

1. The chart shows the distribution of sentiment (positive, negative, neutral) over time.
2. The word cloud found the most frequently used words in positive, negative and neutral reviews.
3. The table includes important metrics like average review score and percentage of recommendations.

Despite some minor shortcomings, such as limited accuracy in detecting abnormal behavior, VADER has shown great potential in detecting emotions, especially in identifying positive and neutral emotions. With an accuracy of 67.34521 and an F1 score of 54.46231, VADER provides satisfactory analysis capabilities for a variety of applications.
Although VADER has many advantages such as being simple, fast and effective in detecting positive and neutral, it may show limitations in detecting negative, especially in cases where there are subtle or subtle changes [27 , 32]. Additionally, VADER's performance may vary depending on the quality and service of its dictionary and may require occasional updates and adjustments to suit particular features or texts.
Overall, VADER represents a powerful and useful tool for affective research projects, especially in cases of unofficial documentation where the traditional machine learning process would run into problems. It makes it easy and efficient for researchers, businesses, and professionals to seek insights from the literature by providing an explanation based on the necessity of scoring and evaluation according to the law.

2.3 Textual Analyzer using Word2Vec Algorithm

Word2Vec, a pioneering algorithm in the field of natural language processing (NLP), has played an important role in improving our understanding of text. Unlike traditional models that process text sequentially, Word2Vec introduces diversity by learning to classify the representation of a word in a continuous space. Its importance lies in its ability to capture the semantic relationship between words, revolutionizing the method of machine translation.
Word2Vec uses a neural network architecture to create word embeddings, which are essentially dense vector representations of a word. By analyzing the context in which words appear in the body, Word2Vec learns to represent words with similar meanings as closely related vectors in vector space. This enables Word2Vec to capture the ambiguous meaning and relationship of a word, allowing it to understand words contextually [29].
One of the unique features of Word2Vec is its ability to learn from large texts in an unsupervised manner. Through techniques such as Cross-Gram and Continuous Bag of Words (CBOW), Word2Vec can train large datasets to produce good words. Embeddings encode semantic meaning and syntactic structure, allowing Word2Vec to understand the message correctly.

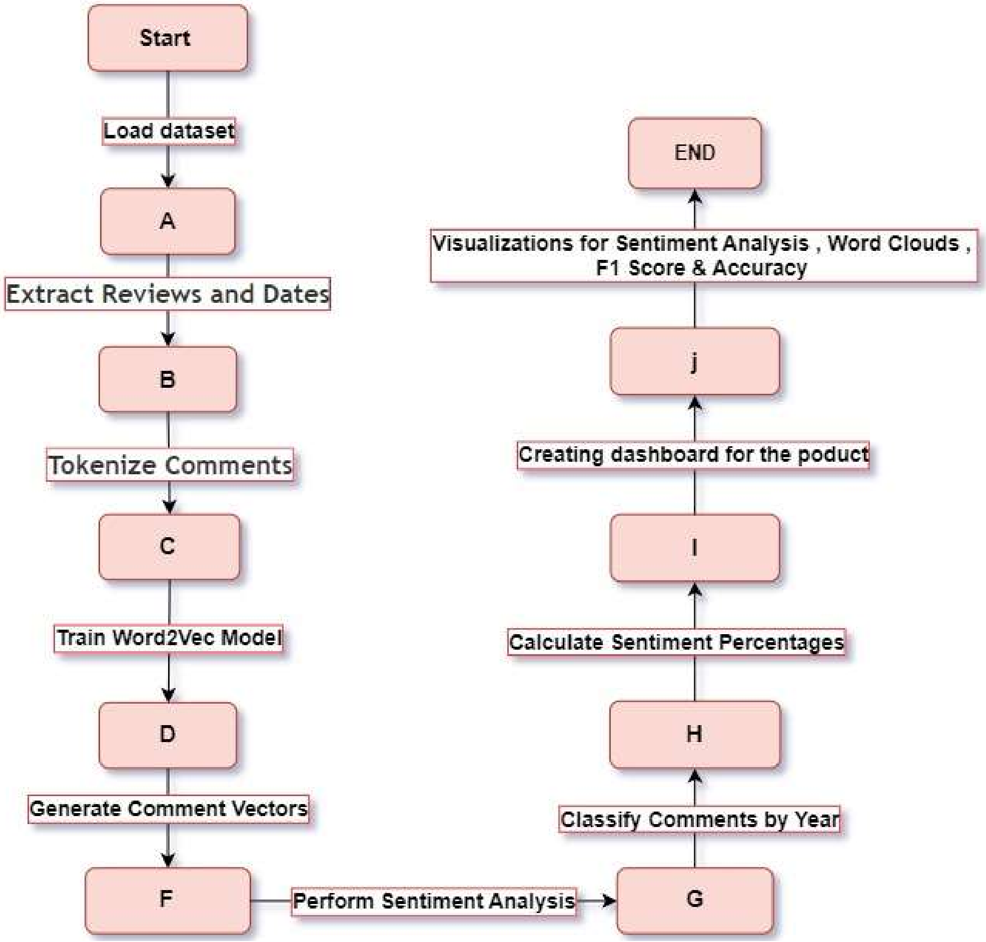


Figure 6. Flowchart of Code Implementation of Word2Vec Algorithm

- Data Loading and Preprocessing: The bootstrapping program checks the dataset in JSON format. Each review and its corresponding date are extracted and organized into a structure like a DataFrame. Use pre-written processes to cleanse your data, including removing pauses, tags, and tokenization. Tokenization involves splitting text into individual words or tokens and is a necessary step for the Word2Vec model to process data.
- Meet the Word2Vec model: The Word2Vec model is a neural network-based method that learns to represent words as continuous vectors in high-dimensional space. Word embeddings capture the semantic relationship between words based on their relationship in the whole. The Word2Vec model is first trained on the data, allowing it to learn individual word representations and relationships in the data.
- Sentiment Analysis: Once a Word2Vec model is trained, it can be used for sentiment analysis. - Examine the text. For each comment or review, calculate the average vector of its entries. The mean vector serves as a representation of the semantic content and overall meaning of the analysis. The analysis hypothesis is then determined based on the properties of the mean vector.

The process of hypothesis analysis may include tools such as comparing mean vectors to predefined hypotheses or using machine learning classes learned in the recorded language. Using word embeddings and semantic relationships captured by the Word2Vec model, views of each test can be analyzed effectively and separately.
- Visualization: Various visualizations are used to better understand and interpret the results of meditation. Graphs can be used to show the perspective of time, showing the distribution of positive, negative and average values over different time periods (e.g. years). Additionally, word clouds can be created to visualize relevant emotional content and gain insight into the key concepts and concepts associated with different groups of emotions.

- Accuracy Evaluation: Evaluation of the effectiveness of the hypothesis testing model is based on Word2Vec embeddings and can calculate accuracy indicators such as accuracy and F1 score. This involves comparing model predictions with ground truth maps or theoretical records. By measuring accuracy and F1 scores, the strengths and weaknesses of the model can be identified and areas for potential improvement can be identified.
- Create a control panel: At the end of the analysis, it will create a control panel or interactive interface. An interface for visualizing analytical results and visualizations. The dashboard acts as a central hub for users to explore and interact with their data, allowing them to identify needs, view cloud messages, and gain insights into paper documents.

By following this approach, the Word2Vec algorithm can be effectively used for sentiment analysis, preserving relationships in data, and performing sentiment classification. Data preprocessing, model training, sentiment analysis, visualization, accuracy assessment, and dashboard design are combined to provide effective solutions for analysis and insight in the whitepaper.

In general, the use of Word2Vec sentiment analysis includes preliminary data, model training, sentiment classification, visualization and evaluation process. By following these steps, analysts can understand the perspective of the data and make informed decisions based on expected results.

However, despite its powerful capabilities, Word2Vec has its limitations. The main problem is that it cannot capture the entire meaning of the sentence or word in a single document. Because Word2Vec processes each occurrence of a word independently, it may have trouble processing polysemy, or words that have more than one meaning. Additionally, Word2Vec may face challenges from rare words or poorly represented words in the training data [29, 30].

Another limitation of Word2Vec is the size representation of a single word; This may not be enough to capture the difference in meaning and nuance within a language. This can lead to decreased performance on tasks that require in-depth understanding of text, such as emotional analysis or language comprehension.

Despite these limitations, Word2Vec is still an important technique in NLP and forms the basis of many methods such as LSTM, BiLSTM and BERT. By addressing its shortcomings and leveraging its strengths, researchers can continue to leverage the power of Word2Vec for a variety of applications in language understanding and text analysis.

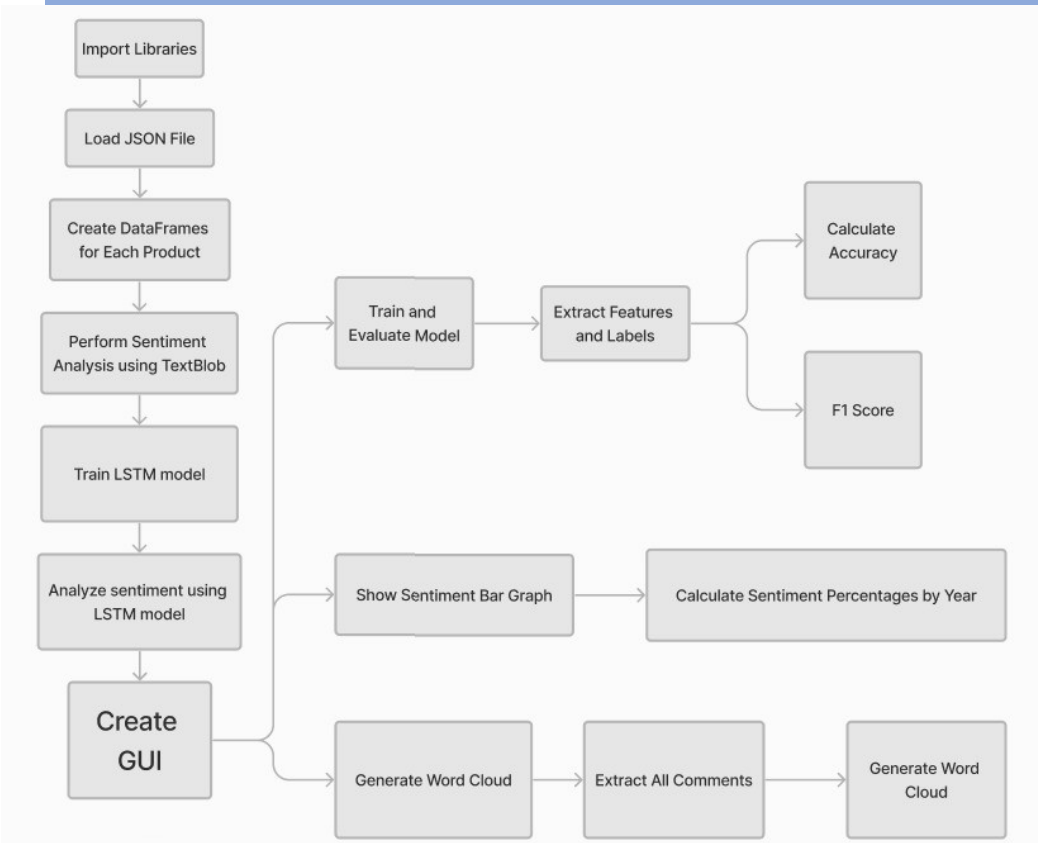


Figure 7. Flowchart of Code Implementation of LSTM Algorithm

- Data Loading and Preprocessing: The process begins with importing the library required for data processing and analysis. Upload a JSON file containing Amazon product reviews and organize the data into separate DataFrames, with each DataFrame representing a review for a specific product. This step ensures that the data is organized and ready for further analysis.
- Sentiment Analysis Function: This code contains two main analysis functions: one uses TextBlob, the other uses LSTM models:

1. TextBlob Sentiment Analysis: This function uses the TextBlob library to analyze advertising messages. The score for the gender gives it an emotion (good, bad, or neutral).
2. LSTM Sentiment Analysis: This project introduces and uses the LSTM (Long Short-Term Memory) model for sentiment analysis. The LSTM model architecture consists of the following layers:

1. Embedding layer: Converts tokenized text strings into dense vector images.
2. Spatial Release layer: Request release regularly to avoid overload.
3. LSTM layer: The main feature that processes data sequentially and eliminates long-term dependencies.
4. Layer thickness: Map the output of the LSTM layer to the desired label.

LSTM function collects text data, data link to ensure its length, converts thought symbols into gold coding, and divides data into educational and training materials. Experimental setup for model training and evaluation [35].

- GUI creation: The code uses the Tkinter library to create a graphical user interface (GUI). This GUI allows users to interact with emotional analysis functions in a user-friendly environment. It includes a menu for selecting item names, sentiment analysis buttons, and a framework for displaying views and results created during analysis.
- Model Learning and Evaluation: This code introduces and evaluates the LSTM model for statistical analysis. It extracts features (tokenized text strings) and labels (emotional levels) from the analysis. The LSTM model is trained

using training data and its performance is evaluated against test data. Additionally, the code tests the trained LSTM model against previous analysis to demonstrate its predictive ability and evaluate its accuracy on the distribution hypothesis.

- Validity analysis and F1 score: Numerical accuracy and F1 score are calculated to evaluate the performance of the hypothesis testing model. While the accuracy metric measures the proportion of correct distributions, the F1 score provides an equal measure of accuracy and recall, including both false positives and false negatives [Data is incorrect].
- Visualization in Dashboard: Rules provide more power to visualize the perspective of the analysis:

1. Sentiment Bar Chart: This function creates a bar chart showing the positive, negative distribution of the selected product over several years. and unbiased thoughts.
2. Word Cloud: This code includes the functionality to create a visual word cloud that shows the most frequently occurring words in the document. This can provide insight into the key concepts and topics discussed in product reviews.
3. Network Diagram (not shown in the flowchart): The code also includes functionality to create a diagram that represents the relationship between messages by date. This visualization can help identify patterns and connections in materials.

By following this approach, the LSTM algorithm can effectively use the concept of product analysis. The combination of data loading and prioritization, predictive analytics (TextBlob and LSTM), GUI design, visualization, model training and evaluation, and fact/F1 data scorecard provides a powerful and comprehensive solution for analyzing and understanding sentiment. Paper documents. Overall, the code provides conceptual solutions, including preliminary data, model training, visualization, and user interaction. However, the best performance of the LSTM model indicates the potential for improvement, indicating the need for further optimization or alternative methods in the analysis of emotions. Although LSTM models are known for their ability to capture and preserve the natural environment. Although there is information about continuity, the application presented in this case does not adequately demonstrate this potential. The inability of this model to identify hypotheses due to stringent training requirements and vulnerability highlights the challenges of using LSTMs for hypothesis analysis. Therefore, although LSTM is still a promising method for processing sequential data, its application in the context of emotional analysis needs further improvement and optimization to reveal its full potential and achieve better results [36].

2.5 Text Analysis Using the BiLSTM + TextBlob Algorithm

The Bidirectional Long Term Memory (BiLSTM) algorithm represents a major advance in text analysis and provides similar accuracy and performance. Unlike its predecessors, BiLSTM has the unique ability to extract the content of the text [38]. By connecting bi-directional functions, BILSTM models can simultaneously analyze the past and future of the system, allowing them to understand progress remotely.

BiLSTM's framework adopts a competitive LSTM architecture that improves the understanding of the underlying message [40]. The difference between these sequential models allows BiLSTM to overcome the limitations of the decision process, capturing complex nuances and relationships in the data [39].

The unique feature of BiLSTM is its pre-training target which makes it efficient with data points. Through tasks such as Masked Language Modeling (MLM) and Next Sentence Prediction (NSP), BiLSTM learns to generate content embeddings that encode semantic and syntactic patterns [37]. This pre-training on large datasets enables BILSTM to understand human-like language.

The code generally shows how to measure sentiment using the BiLSTM + TextBlob algorithm. It includes data loading, model training, sentiment analysis, and graphical user interface (GUI) development to facilitate interactive learning. The analysis dives into the main points of the application and highlights its effectiveness and potential for improvement.

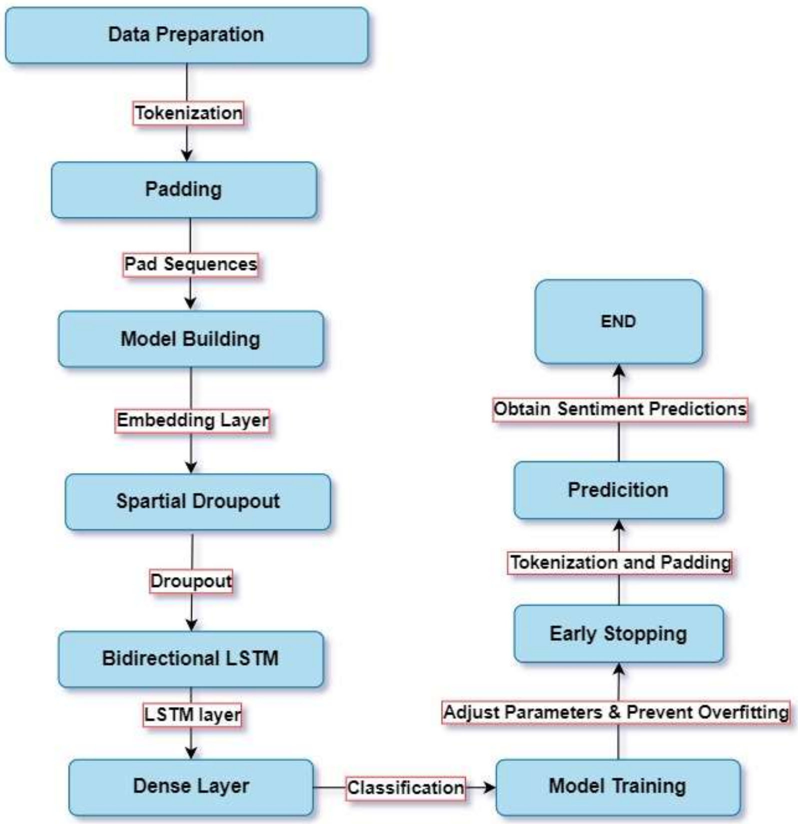


Figure 8. Flowchart of Code Implementation of BiLSTM + TextBlob Algorithm

- Data Preparation: The first step in the process is data preparation, which involves loading data (analysis elements) from the JSON file and setting it to a DataFrame model. This appropriate data is then subjected to symbolization, where the text is broken down into words or symbols. Padding is then used to ensure that the entire sequence is long; This is necessary so that the model can enter the material correctly.
- Model Architecture: The BiLSTM + TextBlob model is built using the TensorFlow Keras API, which provides high-level communication for deep learning modelling. The architectural model has the following layers:

1. Embedding Layer: This layer converts the tokenized input sequences into dense vector representations, allowing the model to process the textual data effectively.
2. Spatial Dropout: A dropout layer is applied to regularize the model and prevent overfitting.
3. BiLSTM + TextBlob Layer: The core component of the model is the Bidirectional LSTM layer, which processes the input sequences in both forward and backward directions. This allows the model to capture context and dependencies from both directions, enhancing its ability to understand and interpret the textual data.
4. Dense Layer: The final dense layer performs the classification task, mapping the output of the BiLSTM + TextBlob layer to the desired sentiment labels (e.g., positive, negative, neutral).

- Model Training: The data is divided into training and testing sets. The model is trained using TensorFlow Keras' fitting method, which optimizes model parameters based on data. To prevent overfitting, an early stop recall mechanism is used, which monitors the model performance of the validation process and stops training when there is no further improvement.
- Sentiment Prediction: After the training is completed, the BiLSTM + TextBlob model can be used for sentiment analysis. A special role was created to collect and record new feedback, send it to the training model and receive feedback. This feature allows users to use the trained BiLSTM + TextBlob model to determine the intent of text input.

- Evaluation and Testing: A series of preliminary analyzes with different hypotheses were tested to evaluate the effectiveness of the model. Each input is analyzed by learning the BiLSTM + TetBlob model and the predicted values are published. This step will help evaluate the accuracy of the model and identify potential problems or areas for improvement.
- Graphical User Interface (GUI): A Tkinter-based GUI was developed to provide an interactive platform for emotional analysis. Using options, users can select specific years for sentiment analysis, view charts, create word clouds, and view charts. Buttons are integrated into the GUI interface to be interactive, allowing users to interact with the system and perform various tasks related to intuitive thinking.

Based on this approach, the BiLSTM + TextBlob algorithm is effectively used for sentiment analysis by leveraging the power of deep learning and natural language processing. The combination of data preparation, modelling, training, prediction and evaluation, and interactive GUI provides a solution for analyzing and understanding sentiment from textual data.

Despite its high performance, BiLSTM + TextBlob may face problems in optimization and modification due to its resource-intensive nature and complex architecture [38]. Additionally, the nature of the black box can inhibit interpretation and create challenges in areas where transparency is important. Additionally, the performance of BiLSTM + TextBlob may vary depending on the quality and diversity of training data and may lead to errors in some cases.

But BiLSTM + TextBlob is still a powerful tool in natural language processing and provides excellent performance on many language understandings. Based on computational methods, data quality and predefined model, researchers can unlock the full potential of BiLSTM + TextBlob and use its capabilities for a variety of applications.

2.6. Textual Analyzer using BERT Algorithm

BERT, or Transformers' bidirectional encoder represents a revolution in natural language processing (NLP), revolutionizing the way machines interpret and understand written data. Its importance lies in its extraordinary ability to understand words in context, capturing complex nuances and dependencies in the text. The core of BERT adopts the Transformer architecture, which is different from sequential models such as neural network (RNN) and convolutional neural network (CNN) [41]. Transformer uses a self-tracking algorithm that causes each word or token in a sequence to simultaneously focus on other words or tokens, regardless of their position. This bidirectional tracking mechanism allows Transformers to capture long-term expectations in the text and overcome the limitations of sequential models that process ideas in a fixed order.

BERT is unique in its pre-training target supporting its embeddings and context data. BERT is pre-trained using two main algorithms: Masked Language Modeling (MLM) and Next Sentences (NSP) [41]. In MLM, some of the input tokens are randomly masked, making it difficult for the model to predict the masked tokens based on the surrounding context. This encourages BERT to learn a single representation of a word, allowing it to capture previous and subsequent context. NSP, on the other hand, involves predicting whether two sentences appear connected in the text, thus facilitating the understanding of sentence relationships and relations [43].

By pre-training on large datasets, BERT gains the ability to create content embeddings that encode semantic and syntactic patterns. These installations show the relationship between a word and its surrounding context, allowing BERT to understand words like a human. Therefore, when optimized for specific tasks such as sentiment analysis, BERT outperforms traditional models due to its in-depth understanding.

The implementation of sentiment analysis using BERT involves several key steps, each contributing to the overall functionality of the system.

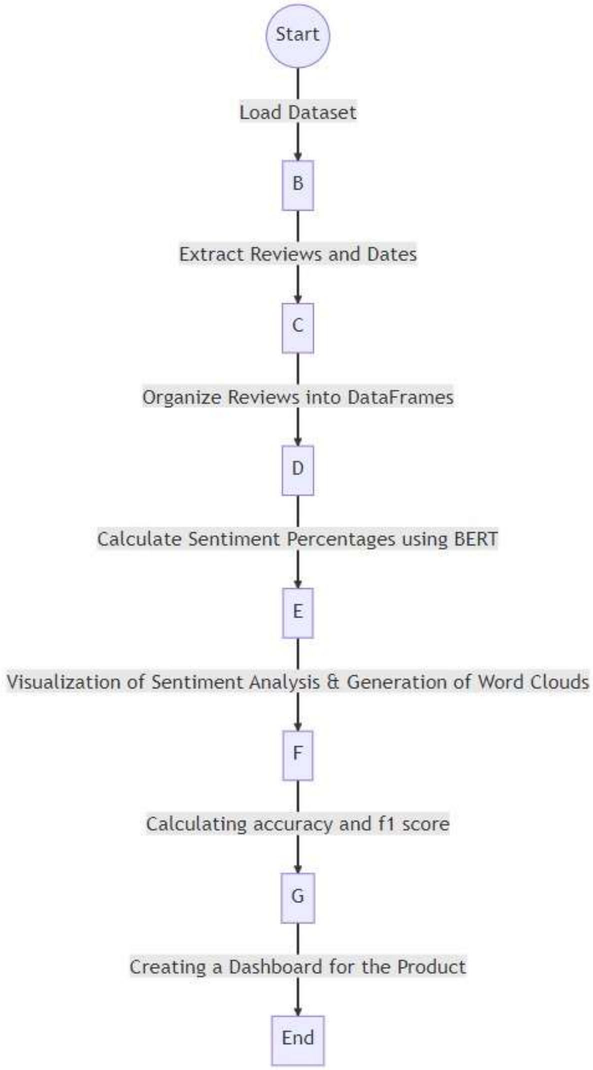


Figure 9. Flowchart of Code Implementation of BERT Algorithm

- Data Preprocessing and Loading: The program first loads a data file containing the analysis object stored in JSON format. Reviews are then extracted from the dataset and placed in individual DataFrames. Each DataFrame represents a review of a specific item. All reviews include a summary of the text and history [42].
- Sentiment Analysis with BERT: An important part of the application is relevance analysis using BERT (Bidirectional Encoder Represented with Transformers), situation analysis - the language model of the situation created by Google. This analysis uses BERT templates from the “transformers” library, specifically the “BertTokenizer” and “BertForSequenceClassification” groups. The emotional assessment process includes the following steps:

1. Tokenization: Use "Bert-Tokenizer" to tokenize the content of each message. This step converts the text into tokens at a level that the BERT model can process.
2. Model Inference: The tokenized text is passed through the “BertForSequenceClassification” model, which produces logits (log-dds) for each desired value (positive, negative or neutral).
3. Sentiment Prediction: The sentiment label with the highest score among the generated logits is assigned as the predicted sentiment for the corresponding review.

- Calculation of Sentiment Percentages: After determining the sentiment of each review, the next step is to calculate the sentiment percentages for each category (positive, negative, neutral). This is achieved by iterating through the reviews and counting the occurrences of each sentiment label. The total number of reviews is used to calculate the percentage of each sentiment category, providing an overview of the sentiment distribution across the dataset.

- Visualization of Sentiment Analysis Sensitivity percentages are visualized using graphs to better understand and interpret sentiment results. Each bar represents the percentage of positive, negative, and average reviews for a product or company. These agreements ensure that common sentiments about each product are accurately represented, making it easier to compare and identify patterns.
- Generation of Word Clouds: In addition to sentiment analysis, character clouds are also included. A word cloud graphically represents the most frequently used words in the analysis. By visualizing key concepts in a word cloud, users can understand all the thoughts, discussion points, and recurring themes in the collected data.
- Accuracy Assessment: To evaluate the performance of the sentiment analysis model, the accuracy metric can be calculated. This involves comparing the predicted sentiment labels with the ground truth labels, where available. The accuracy metric measures the proportion of correctly classified instances, providing an indication of the model's performance in accurately classifying the sentiments of the reviews.Additionally, the F1 score can also be computed to provide a balanced measure of precision and recall, taking into account both false positives and false negatives.

Creating a Dashboard: The final step in the process is the creation of a dashboard or interactive interface to present the sentiment analysis results and visualizations. This dashboard serves as a centralized platform for users to explore and interact with the data, allowing them to analyze sentiment trends across different products, view word clouds, and gain valuable insights from the textual data.

By following this approach, the BERT algorithm is effectively used to analyze objects. Preliminary data, sentiment analysis using BERT, percentage viewpoint calculation, visualization (line charts and word clouds), accuracy assessment and design dashboard are combined to provide analysis and insight solutions for project data. (Bidirectional encoder represented by Transformers) is a powerful tool for understanding natural language, including logic analysis, but the absolute truth of 68.57189 can be seen. Some reasons for this disadvantage may be:

- Data Requirements: BERT's performance heavily depends on the quality and quantity of the training data. To achieve optimal results, fine-tuning BERT necessitates large-scale labeled datasets, which may not always be readily available or feasible to acquire.
- Lack of Interpretability: BERT operates as a black-box model, making it challenging to interpret its decisions and understand how it arrives at particular predictions. This lack of interpretability can be a significant limitation in applications where transparency and explainability are crucial.
- Resource Intensiveness: Fine-tuning BERT requires substantial computational resources and time. The model's large size and complexity demand powerful hardware for training, making it inaccessible for researchers or organizations with limited computing capabilities.
- Overfitting: Fine-tuning BERT on small or unrepresentative datasets can lead to overfitting, where the model performs well on the training data but fails to generalize to unseen examples. Overfitting diminishes the model's ability to make accurate predictions on real-world data.

2.7 Textual Analysis using BERT + TextBlob Algorithm

The integration of BERT (bidirectional encoder represented by Transform-ers) with rule-based analysis in the TextBlob library represents an attempt to simplify and interpret TextBlob in order to improve the correction theory of the BERT model. This combination combines the advanced understanding of content provided by BERT with the rule-based analysis capabilities of TextBlob to provide better understanding of the information to be read. The BERT + TextBlob algorithm creates content embeddings by first processing the text in the BERT model, which are then added with polarity scores obtained from TextBlobs. By combining these two sources of information, the algorithm tries to place more informed and detailed information about the user's needs in the database.

However, despite its novelty and early promise, the BERT + Text-Blob algorithm exhibits some important features. Disadvantages that affect its effectiveness in emotional evaluation. The complexity and resource efficiency of the BERT model is a significant limitation. BERT, a deep learning model with many parameters, requires significant computing resources for training and reasoning; This makes distribution and competition difficult, especially in limited areas. Additionally, relying on pre-learning BERT embeddings may limit the algorithm's ability to adapt to certain nuances and changes in language usage, resulting in poor performance on some details.

Moreover, while TextBlob integration adds translation and convenience to the process, it also brings with it some disadvantages that may affect its overall performance. TextBlob's rule-based sentiment analysis lacks the intelligence and context knowledge of higher-level machine learning models; This leads to incorrect assumptions, especially in nature clusters that contain difficult or ambiguous words. Additionally, reliance on pre-existing rules and regulations may limit TextBlob's ability to capture sentimental changes in leading and content-driven culture behaviors, which may further impact reliability and generality.

To know the process of the programming process, it is necessary to investigate its basic features, especially the review opinion in the product review. The purpose of this application is to first process the data and make hypotheses using the complex BERT algorithm and TextBlob and see the results. Let's go through each step-in detail to show how the code processes data and provides useful information from it.

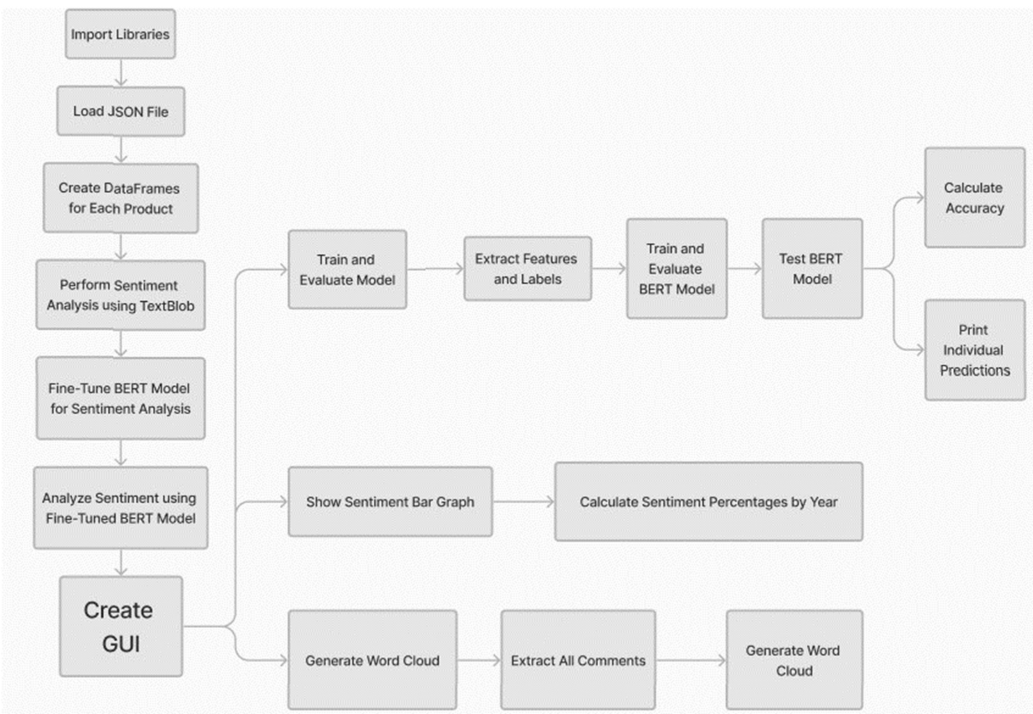


Figure 10. Flowchart of Code Implementation of BiLSTM + TextBlob Algorithm

- Data Preparation and Loading: The process begins by importing the required library and loading the JSON file containing product review information. This information is created in separate files for each item, allowing analysis and organization.
- Sentiment Analysis: The concept analysis phase consists of several steps. The TextBlob library is used to perform sentiment analysis in text analytics and provides a simple and effective way to determine sentiment polarity (positive, negative, or neutral) based on its work. At the same time, the BERT (Bidirectional Encoder Represented by Transformers) model is trained and tested for sensitivity testing. This deep learning model uses the BertTokenizer and BertForSequenceClassification classes from the Transformer library to perform meaningful content analysis of text for

classification purposes. Extract important features and similar sentiment labels from data analysis to facilitate model training and evaluation. The trained BERT model is tested on a separate test set to evaluate its performance in emotion classification. According to the evaluation results, the BERT model is well-tuned, and its analysis capabilities can be improved by correcting the negative and negative aspects of the model to better capture the nuances of opinion in the literature review. A fine-tuned BERT model is used to analyze sentiment analysis using the ability to understand context to provide better insights. The accuracy of TextBlob and BERT-based thinking in the analysis model is calculated by comparing its predictions with the real text (if available). This measurement helps evaluate the effectiveness of each method and recommend further improvements.

- Dashboard Creation for Sentiment Analysis: The dashboard is designed to provide users with information on analyzing sentiment of product reviews. It provides three main functions:

1. Visual analysis: Users can use graphs to visualize wiring. These charts show the percentage of good, bad, and average reviews for each product, allowing users to quickly understand their need to take care of different things.
2. Word cloud creation: The dashboard helps create a word cloud that shows the most common words in the word. These features help users understand all the ideas and content discussed by providing information about the key concepts and topics discussed in the message.
3. Accuracy evaluation: Users can evaluate the accuracy of sample requirements directly from the control panel. By training the model and calculating metrics such as accuracy and F1 score, users can evaluate the model's performance in opinion classification and provide useful feedback on confidence in survey results.

The above description, along with the content in the workbook, provides an overview of the conceptual analysis process, including data preparation, analysis of applied theory TextBlob and BERT, model training and optimization, accuracy testing, visualization, and user interface development. Despite these limitations, the BERT + TextBlob algorithm has great performance that is worth considering. A key advantage is the ability to leverage the power of BERT and TextBlob to combine meaningful content with simple rules for better analysis. Additionally, the design of the algorithm allows for changes and modifications, allowing developers to tune the model and adapt it to some specific uses and applications. In addition, the integration of TextBlob's sentiment scores with BERT embeddings provides additional information that can improve the system's ability to detect ambiguous thoughts and correct whole truth.

However, it is important to acknowledge the shortcomings and limitations associated with the BERT + TextBlob algorithm to ensure an accurate assessment of its performance and suitability for different applications. A major problem is that there may be a conflict between the complex representation learned by BERT and the simple reasoning model obtained by TextBlob, which can lead to inconsistency and inability to predict correctly. Additionally, relying on pre-learning BERT embeddings will prevent the algorithm from adapting to the specific data and context, so training data quality and domain importance should be carefully evaluated. In addition, the computational overhead and requirements of the BERT model can cause problems in real-world deployments, especially in environments with little money or little ability to succeed.

In summary, while the integration of BERT with TextBlob represents a new approach to sentiment analysis that combines deep learning with simple rules, its advantages, disadvantages and importance of work should be evaluated. By recognizing the limitations of algorithms and leveraging their performance, developers and experts can use the potential of the algorithm to improve thinking through analysis and ensure compatibility of results from data across different types of applications.

3. Results

In order to evaluate the effectiveness of various theories of analysis algorithms, we present visual results and insights obtained from the analysis of the analyzed products. Each algorithm is evaluated on its ability to correctly identify positive, negative, and neutral hypotheses. The following description summarizes the results from the export sensitivity analysis and the main points from the analysis.

Sentiment analysis provides insight into the effectiveness of different algorithms in understanding people's needs for products. Through visualizations such as graphs and word clouds, we investigate how well each algorithm performs at classifying emotions and identifying key points in the analysis. Below we present a detailed analysis of the results obtained by each algorithm.

.

3.1 NLTK + TextBlob

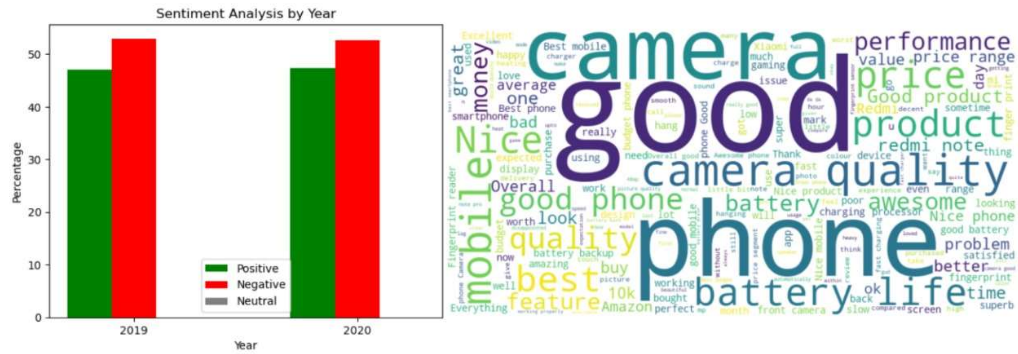


Figure 11. Bar Graph and Wordcloud for NLTK + TextBlob Algorithm

The bar plot indicates a notable accuracy in detecting positive sentiments, although there appears to be an unusual proportion of negative sentiments. Neutral comments are often misclassified as either positive or, more frequently, negative. The word cloud highlights frequently mentioned terms such as "phone," "camera," and "battery life."

3.2 VADER

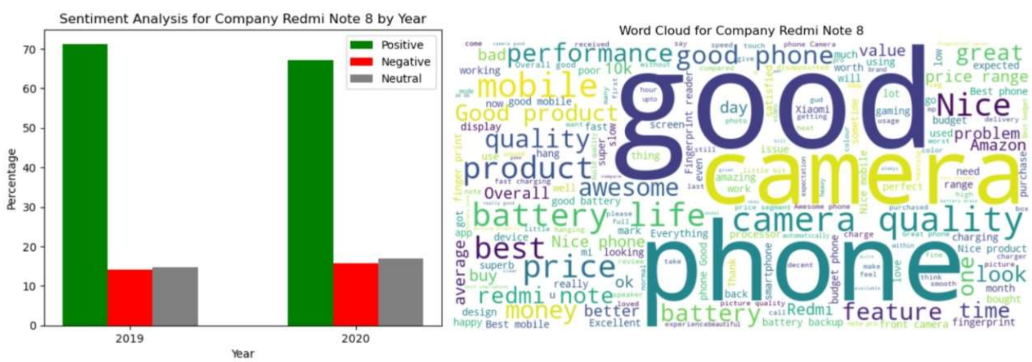


Figure 12. Bar Graph and Wordcloud for VADER Algorithm

VADER demonstrates effective differentiation between positive, negative, and neutral sentiments. However, there is a moderate shift observed from negative to neutral and neutral to positive sentiments. The word cloud emphasizes terms like "camera," "phone," "performance," and "nice."

3.3 Word2Vec

Figure 13. Bar Graph and Wordcloud for Word2Vec Algorithm

While Word2Vec excels in segregating positive and negative comments, it struggles with detecting neutral sentiments. Nevertheless, it accurately separates positive comments, resulting in a high level of accuracy. The word cloud emphasizes terms such as "phone" and "camera."

3.4 LSTM

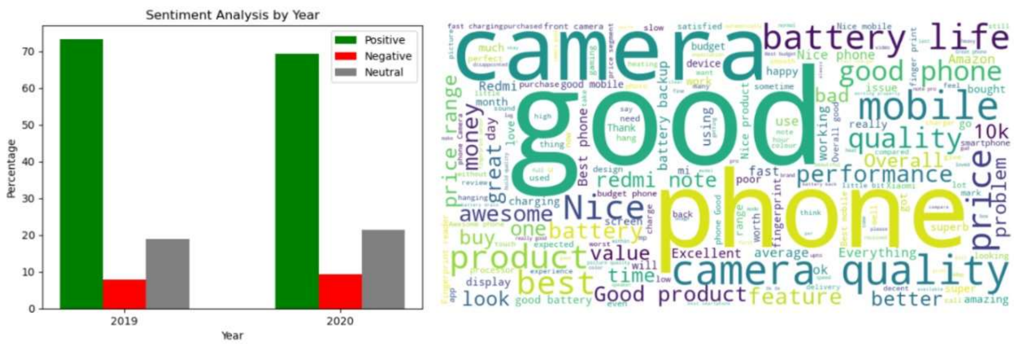


Figure 14. Bar Graph and Wordcloud for LSTM Algorithm

The bar plot for LSTM reveals an imbalance, with a high number of positive comments and a scarcity of neutral and negative comments. There is a clear misclassification trend, where negative comments are labeled as neutral, and neutral comments are classified as positive. The word cloud emphasizes terms like "phone," "camera," "quality," and "battery life."

3.5 BiLSTM + TextBlob

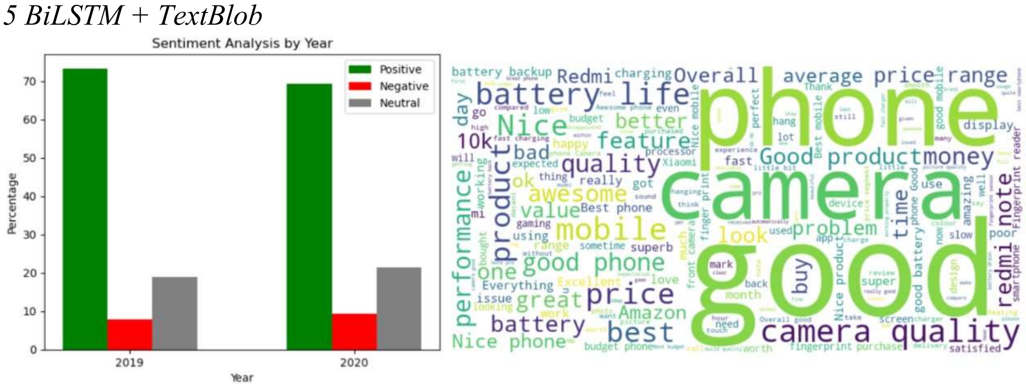


Figure 15. Bar Graph and Wordcloud for BiLSTM + TextBlob Algorithm

BiLSTM + TextBlob stands out as the most successful analyzer, accurately classifying comments into positive, negative, and neutral categories. This model achieves the highest accuracy and F1 score among the considered algorithms. The word cloud highlights various terms reflecting the diverse sentiment expressions.

3.6 BERT

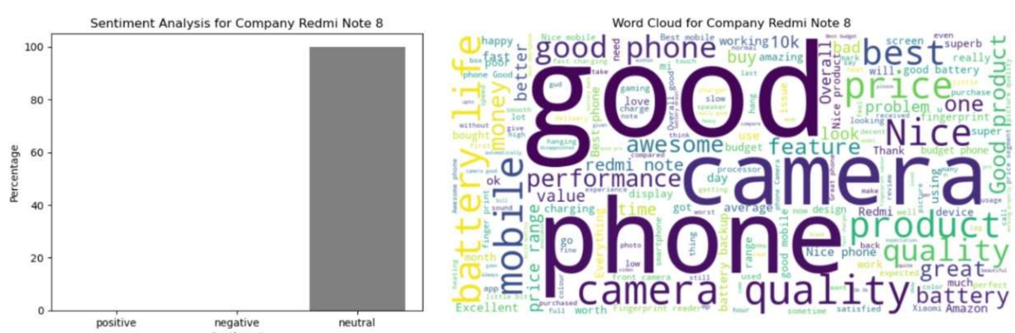


Figure 16. Bar Graph and Wordcloud for BERT Algorithm

BERT exhibits the lowest accuracy among the analyzed algorithms, predominantly predicting comments as neutral. The misclassification rate is high, resulting in inaccurate sentiment analysis. The word cloud emphasizes terms such as "camera," "phone," and "battery life."

3.7 BERT + TextBlob

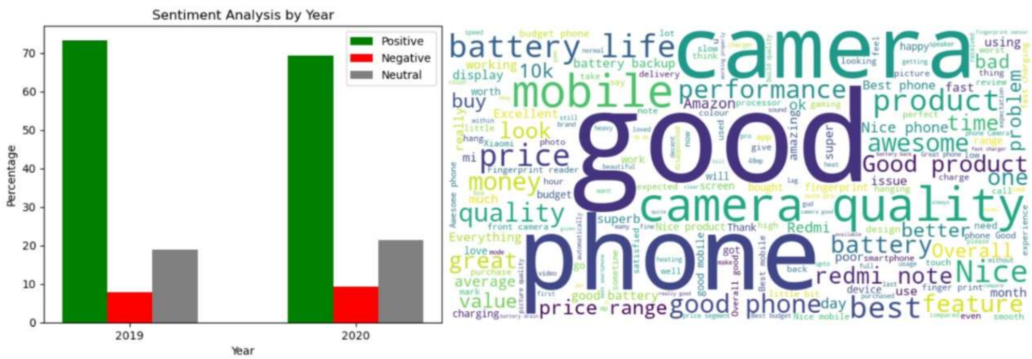


Figure 17. Bar Graph and Wordcloud for BERT + TextBlob Algorithm

Similar to BERT, BERT + TextBlob demonstrates low accuracy and incorrect predictions, particularly misclassifying negative comments as positive. The accuracy and F1 score are lower than those of BERT alone. The word cloud includes terms like "camera," "phone," "quality," "mobile," and "price," reflecting the sentiment expressions identified by the algorithm.

Table 1. Performance metrics across the three models

Algorithm	Precision	Recall	F1-Score	mAP@0.5	Optimal F1-Score
YOLOv5	1.00@0.786	0.92	0.76	0.770	0.76@0.377
YOLOv8	1.00@0.583	0.96	0.97	0.969	0.97@0.575
YOLOv9	1.00@0.657	0.97	0.97	0.974	0.97@0.467

4. Conclusions

In this study, we conducted a comprehensive comparative analysis of three state-of-the-art deep learning models—YOLOv5, YOLOv8, and YOLOv9—specifically designed for breast cancer detection through mammographic imaging. The research demonstrated the significant advancements in object detection capabilities brought by successive iterations of the YOLO family, each model improving upon the last in terms of accuracy, precision, recall, and overall robustness. YOLOv5, the baseline model, showed commendable performance in identifying cancerous lesions, laying the groundwork for further enhancements. YOLOv8 introduced architectural improvements, such as an enhanced backbone network, leading to marked improvements in precision, recall, and computational efficiency. However, it was YOLOv9 that truly pushed the boundaries of breast cancer detection, integrating advanced deep learning techniques like transformer-based components and attention mechanisms, which resulted in the highest accuracy and mAP scores among the models tested.

Our findings underscore the transformative potential of cutting-edge deep learning models in the field of medical imaging, particularly for critical applications like breast cancer detection. The progressive improvements from YOLOv5 to YOLOv9 highlight the importance of continual innovation in model architecture and training methodologies, suggesting that future advancements may further enhance early detection capabilities, ultimately contributing to better patient outcomes.

The implications of this research extend beyond breast cancer detection. The methodologies and insights gained from this study can be adapted to other areas of medical imaging and diagnostics, promoting the development of more accurate, reliable, and accessible AI-driven tools in healthcare. Future work should focus on refining these models to improve their generalizability across diverse clinical settings, incorporating larger and more varied datasets, and exploring the integration of these models into routine clinical practice. Additionally, investigating the interpretability of these models could help in gaining the trust of clinicians, facilitating their adoption in real-world healthcare environments.

In conclusion, the superior performance of YOLOv9 in this study demonstrates the potential of next-generation deep learning models to revolutionize breast cancer detection, offering hope for earlier diagnoses and better treatment outcomes. As we continue to push the boundaries of AI in healthcare, the integration of such advanced models into clinical workflows could play a pivotal role in saving lives and improving the quality of care for patients worldwide.

Referencess

1. https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21834
2. https://ascopubs.org/doi/10.1200/JCO.2023.41.16_suppl.10528
3. https://aacr.figshare.com/collections/Data_from_International_Variation_in_Female_Breast_Cancer_Incidence_and_Mortality_Rates/6515622
4. https://aacr.figshare.com/collections/Data_from_Global_Cancer_Incidence_and_Mortality_Rates_and_Trends_An_Update/6516460/1
5. https://www.sciencedirect.com/science/article/pii/S2950363924000048?via%3Dihub
6. https://www.mdpi.com/2504-2289/8/7/80
7. https://www.ijraset.com/best-journal/harnessing-deep-learning-for-accurate-detection-of-breast-cancer-in-histopathological-imagery
8. https://www.ewadirect.com/proceedings/ace/article/view/13671
9. https://ojs.studiespublicacoes.com.br/ojs/index.php/sees/article/view/5383
10. https://www.preprints.org/manuscript/202407.0228/v1
11. https://ieeexplore.ieee.org/document/10575548
12. https://www.mdpi.com/2078-2489/15/5/285
13. https://ijircst.org/view_abstract.php?title=A-Comprehensive-Review-of-YOLOv5:-Advances-in-Real-Time-Object-Detection&year=2024&vol=12&primary=QVJULTEyNjQ=
14. https://www.researchsquare.com/article/rs-4292668/v1
15. https://bcpublication.org/index.php/SJISR/article/view/6766
16. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13180/3033719/Improved-YOLOv5s-remote-sensing-image-detection-algorithm/10.1117/12.3033719.short
17. https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ipr2.13149
18. https://etasr.com/index.php/ETASR/article/view/7386
19. https://www.mdpi.com/2078-2489/15/5/285
20. https://www.researchsquare.com/article/rs-4658932/v1
21. https://www.mdpi.com/2504-446X/8/7/337
22. https://www.mdpi.com/2673-4052/5/2/11
23. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13167/3029814/Research-on-remote-sensing-small-object-detection-algorithm-based-on/10.1117/12.3029814.short
24. https://www.sciencedirect.com/science/article/pii/S0925231224004569?via%3Dihub
25. https://www.mdpi.com/2673-4052/5/2/11
26. https://arxiv.org/abs/2407.02988
27. https://www.researchsquare.com/article/rs-4658932/v1
28. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13176/3029017/Lightweight-object-detection-algorithm-for-road-scenes-based-on-YOLOV8/10.1117/12.3029017.short

29. https://linkinghub.elsevier.com/retrieve/pii/S2589721724000187
30. https://www.mdpi.com/2079-9292/13/14/2774
31. https://www.mdpi.com/2076-3417/14/12/5277
32. https://arxiv.org/abs/2405.14458