IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14234

Analyzing Customer Feedback using NLP

Mrs.N.Malathi[1] , S.Rajesh[2], T.Pavan Datta[3], M.Naga Sai Krishna[4], M.Trinadh[5]

Associate Professor of CSE-Data Science, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India.[1]

B. Tech CSE-Data Science, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India.[2-5]

Abstract: Customer feedback analysis plays a crucial role in helping organizations improve their products and services. Using Natural Language Processing (NLP) techniques, this project converts unstructured customer feedback into meaningful insights. The analysis utilizes tools such as NLTK, along with models like Bag of Words (BoW) and advanced deep learning frameworks such as Transformers. The process starts with data preprocessing steps like tokenization, removal of stop words, and lemmatization, efficiently handled by the NLTK library. The Bag of Words model transforms text into numerical data for sentiment classification and topic identification, though it lacks the ability to grasp context. To overcome this, Transformers are employed, offering contextual understanding and accurate sentiment detection. By combining traditional methods like BoW with the sophisticated capabilities of Transformers, this project ensures precise and scalable analysis of customer feedback. This integration enables companies to address user concerns promptly and enhance customer satisfaction.

Keywords: Automated Review System, Semantic Analysis, Natural Language Processing (NLP), NLTK (Natural Language Toolkit), Bag of Words (BoW), Transformers, Tokenization

I. INTRODUCTION

In the digital era, customer reviews significantly influence consumer decisions and business strategies. Manual analysis of large volumes of reviews is inefficient, making automated sentiment analysis essential. This project aims to develop an automated sentiment analysis system to classify customer reviews as positive, negative, or neutral, integrating rule-based and deep learning- based Natural Language Processing (NLP) techniques for enhanced accuracy and contextual understanding.

The system processes user-input text through various NLP techniques, including tokenization, stop-word removal, stemming/lemmatization, and part-of-speech tagging. Sentiment analysis is performed using VADER (Valence Aware Dictionary and Sentiment Reasoner) and RoBERTa (Robustly Optimized BERT Pretraining Approach). VADER, a lexicon-based model, is efficient for short, informal text, while RoBERTa, a transformer-based model, excels in capturing nuanced sentiment in complex textual data. Data visualization tools like Seaborn and Matplotlib illustrate sentiment trends. The comparative analysis highlights that while VADER offers quick, interpretable results, RoBERTa provides superior contextual sentiment detection. This system benefits various domains such as e-commerce, customer feedback monitoring, and brand reputation management by providing actionable insights and improving decision-making.

II. METHODOLOGY

This project adopts a structured research methodology integrating traditional and advanced NLP techniques. Amazon reviews serve as the primary dataset, undergoing preprocessing steps such as tokenization, stop- word removal, stemming/lemmatization, and POS tagging for effective analysis. Sentiment classification employs rule-based (VADER) and deep learning-based (RoBERTa) approaches.

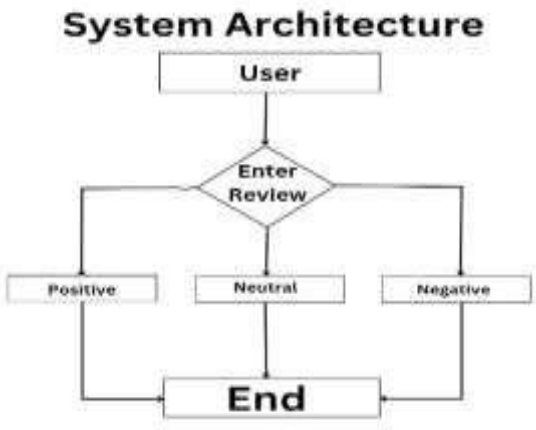
VADER assigns polarity scores using a predefined lexicon, while RoBERTa, trained on large datasets, offers nuanced sentiment classification.User input is processed through both models, with sentiment scores compared for effectiveness. Data visualization presents sentiment trends, and the evaluation compares predicted sentiment with actual review ratings using accuracy, precision, recall, and F1-score metrics. The study reveals that VADER is efficient for short texts but struggles with context-heavy content, while RoBERTa excels in complex text analysis, ensuring a comprehensive sentiment analysis framework for various applications.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14234



III. MODELING AND ANALYSIS

The accuracy of VADER and RoBERTa is evaluated by comparing predicted sentiment scores with actual review ratings. VADER is effective for short texts but less accurate for context-heavy content, while RoBERTa provides nuanced sentiment analysis in longer reviews. Accuracy metrics such as precision, recall, and F1-score are used for evaluation, highlighting RoBERTa's superior performance in complex sentiment classification.

IV. RESULTS AND DISCUSSION

The analysis will tells us that while VADER is effective for straightforward reviews, it struggles with context, whereas RoBERTa excels in sentiment detection. Sentiment trends align with review ratings, with RoBERTa providing balanced sentiment distribution. Combining both models enhances accuracy, making the system suitable for real-world applications like customer feedback monitoring and business intelligence.

The accuracy of VADER and RoBERTa is evaluated by comparing predicted sentiment scores with actual review ratings. VADER is effective for short texts but less accurate for context-heavy content, while RoBERTa provides nuanced sentiment analysis in longer reviews. Accuracy metrics such as precision, recall, and F1-score are used for evaluation

V. RESULTS

Our model receives the example review and uses Robertta and Vader sentiment scoring to determine whether our product is good, awful, or neutral.

Website Interface

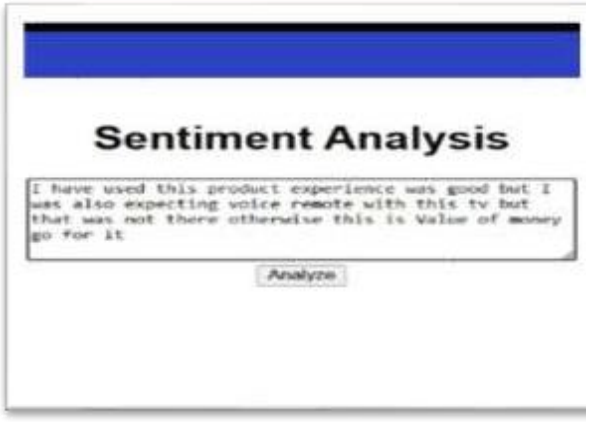
- Our model takes the user's sample review from any e-commerce website as input .

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14234



- The model will gives you all the positive,neutral and negative scores

VI. CONCLUSION

Sentiment analysis classifies text as positive, negative, or neutral by analyzing linguistic patterns and contextual meanings. VADER uses a predefined lexicon for quick sentiment detection, while RoBERTa leverages deep learning for nuanced classification. Integrating both models balances speed and accuracy, providing comprehensive sentiment analysis. Future work may include fine-tuning RoBERTa with domain-specific data and developing hybrid models for dynamic sentiment analysis.

REFERENCES

[1]. International Journal on "Wielding Neural Networks to Interpret Facial Emotions in Photographs with Fragmentary Occlusion", on American Scientific Publishing Group (ASPG) Fusion: Practice and Applications(FPA) ,Vol. 17, No. 01, August, 2024, pp. 146-158.

[2]. International Journal on "Prediction of novel malware using hybrid convolution neural network and long short-term memory approach", on International Journal of Electrical and Computer Engineering (IJECE),Vol. 14, No. 04, August, 2024, pp. 4508-4517.

[3]. International Journal on "Cross-Platform Malware Classification: Fusion of CNN and GRU Models",on International Journal of Safety and Security Engineering (IIETA),Vol. 14, No. 02, April, 2024, pp. 477-486

[4]. International Journal on "Enhanced Malware Family Classification via Image-Based Analysis Utilizing a Balance-Augmented VGG16 Model,onInternational information and Engineering Technology Association (IIETA),Vol. 40, No. 5, October, 2023, pp. 2169-2178

[5]. International Journal on "Android Malware Classification Using LSTM Model, International information and Engineering Technology Association (IIETA) Vol. 36, No. 5, (October, 2022), pp. 761 – 767. Android Malware Classification Using LSTM Model | IIETA.

[6]. International Journal on "Classification of Image spam Using Convolution Neural Network", Traitement du Signal, Vol. 39, No. 1, (February 2022), pp. 363-369 .

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14234

[7]. International Journal on "Medical Image Classification Using Deep Learning Based Hybrid Model with CNN and Encoder", International information and Engineering Technology Association (IIETA), Revue d'Intelligence ArtificiellVol. 34, No. 5, (October, 2020), pp. 645 – 652.

[8]. International Journal on "Prediction of Hospital Re-admission Using Firefly Based Multi-layer Perception, International information and Engineering Technology Association (IIETA) Vol. 24, No. 4, (sept, 2020), pp. 527 – 533.

[9]. International Journal on "Energy efficient intrusion detection using deep reinforcement learning approach",Journal of Green Engineering (JGE),Volume-11, Issue-1,January 2021.625-641.

[10]. International Journal on "Classification of High Dimensional Class Imbalance Data Streams Using Improved Genetic Algorithm Sampling", International Journal of Advanced Science and Technology, Vol. 29, No. 5, (2020), pp. 5717 – 5726.

[11]. Dr. M. Ayyappa Chakravarthi etal. published Springer paper "Machine Learning-Enhanced Self-Management for Energy-Effective and Secure Statistics Assortment in Unattended WSNs" in Springer Nature (Q1), Vol 6, Feb 4th 2025

[12]. Dr. M. Ayyappa Chakravarthi etal. published Springer paper "GeoAgriGuard AI-Driven Pest and Disease Management with Remote Sensing for Global Food Security" in Springer Nature (Q1), Jan 20th 2025.

[13]. Dr. M. Ayyappa Chakravarthi etal. presented and published IEEE paper "Machine Learning Algorithms for Automated Synthesis of Biocompatible Nanomaterials" , ISBN 979-8- 3315-3995-5, Jan 2025.

[14]. Dr. M. Ayyappa Chakravarthi etal. presented and published IEEE paper "Evolutionary Algorithms for Deep Learning in Secure Network Environments" ISBN:979-8-3315- 3995-5, Jan 2025.

[15]. Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Time Patient Monitoring Through Edge Computing: An IoT-Based Healthcare Architecture" in Frontiers in Health Informatics (FHI), Volume 13, Issue 3, ISSN- Online 2676-7104, 29th Nov 2024.

[16]. Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Amalgamate Approaches Can Aid in the Early Detection of Coronary heart Disease" in Journal of Theoretical and Applied Information Technology (JATIT) , Volume 102, Issue 19, ISSN 1992-8645, 2nd Oct 2024.

[17]. Dr. Ayyappa Chakravarthi M, etal, published Scopus paper "The BioShield Algorithm: Pioneering Real-Time Adaptive Security in IoT Networks through Nature-Inspired Machine Learning" in SSRG (Seventh Sense Research Group)-International Journal of Electrical and Electronics Engineering (IJEEE), Volume 11, Issue 9, ISSN 2348-8379, 28th Sept 2024.

[18]. Ayyappa Chakravarthi M, Dr M. Thillaikarasi, Dr Bhanu Prakash Battula, published SCI paper "Classification of Image Spam Using Convolution Neural Network" in International Information and Engineering Technology Association (IIETA)- "Traitement du Signal" Volume 39, No. 1

[19]. Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published Scopus paper "Classification of Social Media Text Spam Using VAE-CNN and LSTM Model" in International Information and Engineering Technology Association (IIETA)-Ingénierie des Systèmes d'Information (Free Scopus) Volume 25, No. 6.

[20]. Ayyappa Chakravarthi M, Dr. M.Thillaikarasi, Dr. Bhanu Praksh Battula,published Scopus paper "Social Media Text Data Classification using Enhanced TF_IDF based Feature Classification using Naïve Bayesian Classifier" in International Journal of Advanced Science and Technology (IJAST) 2020

[21]. Ayyappa Chakravarthi M. presented and published IEEE paper on "The Etymology of Bigdata on Government Processes" with DOI 10.1109/ICICES.2017.8070712 and is Scopus Indexed online in IEEE digital Xplore with Electronic ISBN: 978-1-5090-6135-8, Print on Demand(PoD) ISBN:978-1-5090-6136-5, Feb'2017.

[22]. Subba Reddy Thumu & Geethanjali Nellore, Optimized Ensemble Support Vector Regression Models for Predicting Stock Prices with Multiple Kernels. Acta Informatica Pragensia, 13(1), x–x. 2024.

[23]. Subba Reddy Thumu, Prof. N. Geethanjali. (2024). "Improving Cryptocurrency Price Prediction Accuracy with Multi-Kernel Support Vector Regression Approach". International Research Journal of Multidisciplinary Technovation 6 (4):20-31.

[24]. Dr syamsundararaothalakola et.al.published Scopus paper "An Innovative Secure and Privacy-Preserving Federated Learning Based Hybrid Deep Learning Model for Intrusion Detection in Internet-Enabled Wireless Sensor Networks " in IEEETransactions on Consumer Electronics 2024.

[25]. Dr syamsundararaothalakola et.al.published Scopus paper "Securing Digital Records: A Synerigistic Approach with IoT and Blockchain for Enhanced Trust andTransparency " in International Journal of Intelligent Systems and Applications in Engineering 2024.

[26]. Dr syamsundararaothalakola et.al.published Scopus paper "A Model for Safety Risk Evaluation of Connected Car Network " inReview of Computer Engineering Research2022.

[27]. Dr syamsundararaothalakola et.al.published Scopus paper "An Efficient Signal Processing Algorithm for Detecting Abnormalities in EEG Signal Using CNN" in Contrast Media and Molecular Imaging 2022.