



# Computer Engineering Department

**Government College of Engineering & Research Avasari Khurd Savitribai Phule Pune University**

Final Year of Computer Engineering 2024-25 Semester I

Project Synopsis

* Project Group ID : 6
* Title of the Project : Product Review and Visual Categorization using Deep Learning
* Team Members : 4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr.** | **Roll No** | **Name** | **Email** | **Mobile** | **Sign** |
| 1 | 21121036 | Aditya Pawar | aditya.pawar122020@gco eara.ac.in | 9146601460 |  |
| 2 | 21121007 | Atmaram Kambli | atmaram.kambli122020@ gcoeara.ac.in | 7768075680 |  |
| 3 | 21121004 | Swastik Gupta | swastik.gupta122020@gc oeara.ac.in | 9028830937 |  |
| 4 | 22221078 | Atharva Pawar | pawaratharv22221078@g coeara.ac.in | 8888482905 |  |

Date: Name & Signature Name & Signature Guide Project Coordinator

# Abstract

This research aims to develop a robust deep learning model capable of accurately determining the sentiment expressed in product reviews while simultaneously categorizing accompanying images into predefined visual categories. By leveraging the power of deep neural networks, we propose to extract meaningful features from both textual and visual data to enhance the understanding of consumer opinions and preferences. The model will be trained on a large dataset of product reviews and corresponding images, enabling it to learn complex patterns and correlations between text and image content. The expected outcome is a system that can provide valuable insights for businesses by identifying product strengths, weaknesses, and customer sentiment trends, as well as facilitating image-based product search and recommendation systems.

* + **Keywords :** Product review sentiment, visual categorization, deep learning, natural language processing, computer vision, sentiment analysis, image classification.

## Background and Motivation



Online platforms have transformed the way consumers make purchasing decisions, with product reviews and accompanying images becoming crucial sources of information. Extracting valuable insights from these vast amounts of data presents a significant challenge. Traditional methods for sentiment analysis and image categorization have limitations in handling complex linguistic patterns and visual features.

The increasing availability of large-scale datasets and advancements in deep learning have opened new avenues for addressing these challenges. By combining natural language processing and computer vision techniques, we can develop sophisticated models capable of understanding the nuances of human language and interpreting visual content.

This research is motivated by the need for a comprehensive approach to analyze product reviews and images. By accurately determining sentiment and categorizing visual content, businesses can gain a deeper understanding of customer preferences, identify product strengths and weaknesses, and make data-driven decisions to improve product offerings and marketing strategies. Additionally, this research contributes to the development of advanced recommendation systems that can leverage both textual and visual information to provide personalized product suggestions.

Ultimately, the goal is to create a powerful tool that empowers businesses to harness the wealth of information available in online product reviews and images, leading to enhanced customer satisfaction and business growth.

## Problem Definition and Objectives



**Problem Definition**

Extracting meaningful insights from the vast volume of product reviews and accompanying images available online is a complex challenge. Existing methods for sentiment analysis and image categorization often fall short in accurately capturing the nuances of human language and visual content. Furthermore, there is a lack of effective approaches for jointly analyzing textual and visual information to understand the relationship between product reviews and corresponding images

## Objectives

* 1. Develop a deep learning model capable of accurately classifying the sentiment expressed in product reviews.
  2. Create a robust image categorization system that can effectively classify product images into predefined visual categories.
  3. Design a multi-modal framework that can jointly analyze product reviews and images to enhance sentiment analysis and image categorization performance.
  4. Evaluate the proposed models on a large-scale dataset to demonstrate their effectiveness and generalizability.
  5. Explore potential applications of the developed models in various domains, such as product recommendation, market research, and customer feedback.

## Literature Survey



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Title | Source and Author | Years | Techniques | Merits | Demerits |
| Current State of Text Sentiment Analysis from Opinion to Emotion Mining | ACM Computing Surveys; Yadollahi, A. et al. | 2017 | Survey of text-based sentiment analysis techniques | Comprehensive overview from opinion to emotion mining | Focuses only on text, not multimodal |
| A survey of recent architectures of deep convolutional neural networks. | Artificial Intelligence Review; Khan, A. et al. | 2019 | Survey of deep CNN architectures for image categorization | Detailed review of CNN developments for image analysis | Limited to image modality, doesn't cover text |
| Deep Residual Learning for Image Recognition. | IEEE CVPR; He, K. et al. | 2015 | ResNet architecture for image categorization | Significant improvement in deep network training for images | Focused solely on image recognition |
| GLoRIA: A Multimodal Framework | AAAI; Yu Wang, Q.Xing et al. | 2015 | Progressive training and domain image. | Robust approach for image sentiment | Does not incorporate textual data |

## Methodology



Our project employs a dual-stream approach, incorporating both text-based sentiment analysis and image-based product categorization. The system architecture consists of two main pipelines that converge into a unified interface for user interaction.

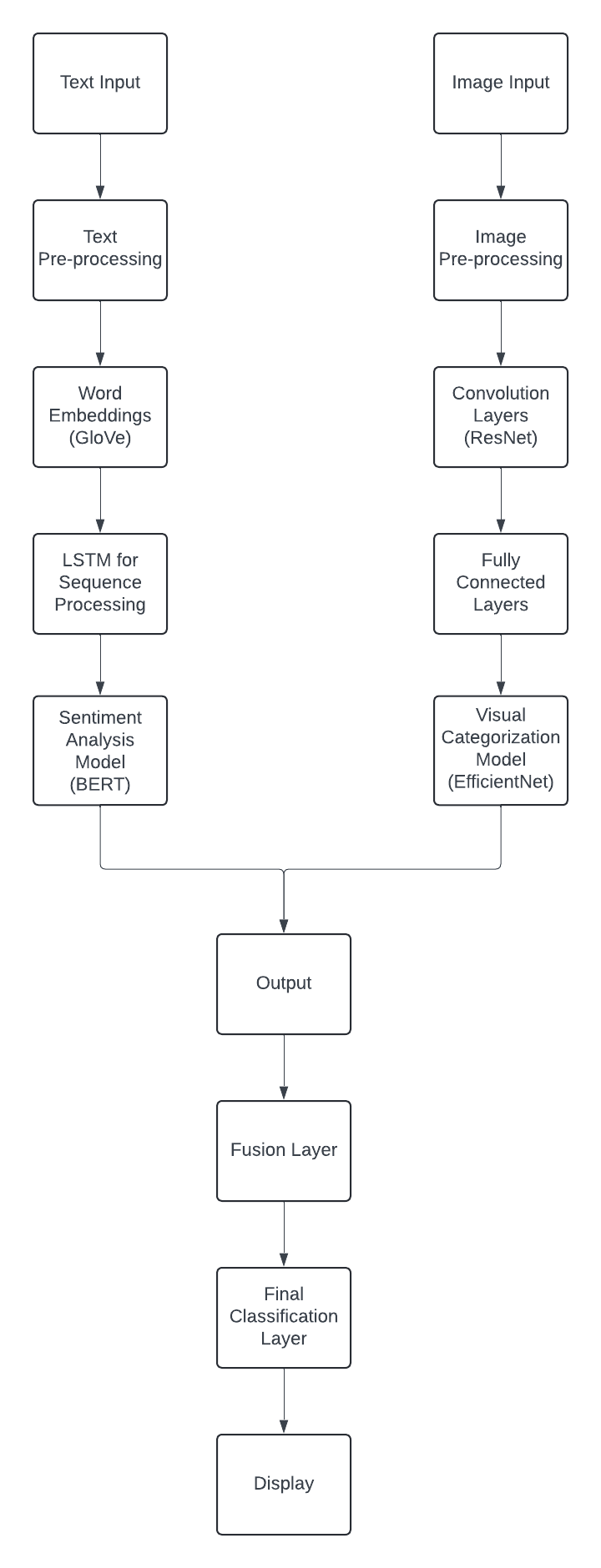


Fig 4.1 System Architecture Diagram



1. Data Ingestion Layer
   * Text input module for product reviews: Handles the input of textual data, allowing users to submit product reviews either through a user interface or via API calls. This module may include features like batch uploading of reviews or real-time input.
   * Image upload module for product images: Manages the upload of product images. It supports various image formats and can handle both single image uploads and bulk uploads.

2 . Preprocessing Layer

* Text cleaning and tokenization for reviews: Processes the raw text data by removing noise (e.g., special characters, HTML tags), correcting spelling, and tokenizing the text into words or subwords. It may also include steps like lowercasing and removing stop words.
* Image resizing and normalization: Standardizes all input images to a consistent size and format required by the visual categorization model. This includes resizing images, adjusting color channels, and normalizing pixel values.

1. Deep Learning Models Layer
   * Sentiment Analysis model (BERT-based): A fine-tuned BERT model that analyzes the preprocessed text to determine the sentiment of the product review. It classifies the sentiment as positive, negative, or neutral.
   * Visual Categorization model (CNN-based): A convolutional neural network, likely based on a proven architecture like ResNet, that processes the normalized images to categorize products into predefined classes.
2. Output Processing Layer



* + Sentiment classification post-processing: Processes the output from the sentiment analysis model, potentially applying threshold adjustments, confidence scoring, or ensemble methods to improve accuracy.
  + Category prediction post-processing: Handles the output from the visual categorization model, possibly implementing techniques like top-k predictions or confidence thresholding to provide more useful results.

1. User Interface Layer
   * Results display and visualization: Presents the processed results in a user- friendly format. This could include sentiment scores, category predictions, and confidence levels, possibly with visualizations like charts or highlighted text.
   * User interaction handling: Manages user inputs, such as submitting new reviews or images, adjusting settings, or requesting additional analyses.

## Software and Hardware Requirements



**Hardware Requirements**

* + **Processor:** Multi-core processor (Intel Core i5 or equivalent) for efficient computation.
  + **Memory (RAM):** Minimum 8GB RAM, preferably 16GB or more for handling large datasets and complex models.
  + **Storage:** Sufficient storage capacity (SSD recommended) to accommodate datasets, models, and intermediate results.
  + **GPU:** A dedicated GPU (NVIDIA GeForce GTX or RTX series) is highly recommended for accelerating deep learning computations.

## Software Requirements

* + **Operating System:** Windows 10, macOS, or Linux (Ubuntu preferred).
  + **Deep Learning Framework:** TensorFlow, PyTorch, or Keras for model development and training.
  + **Programming Language:** Python as the primary programming language for data preprocessing, model development, and evaluation.
  + **Development Environment:** Integrated Development Environment (IDE) like PyCharm, Jupyter Notebook, or Visual Studio Code.

## Libraries:

* + - NumPy, Pandas, Matplotlib for data manipulation and visualization.
    - OpenCV for image processing.
    - NLTK for natural language processing.
    - Scikit-learn for machine learning algorithms.
    - TensorBoard for model visualization and debugging.
  + **Database:** For storing processed data and model artifacts (optional).

## References



1. Lu, Yaojie., Liu, Qing., Dai, Dai., Xiao, Xinyan., Lin, Hongyu., Han, Xianpei., Sun, Le., & Wu, Hua. (2022). Unified Structure Generation for Universal Information Extraction. , 5755-5772 . <http://doi.org/10.48550/arXiv.2203.12277>
2. Manek, Asha S.., Shenoy, P.., Mohan, M.., & Venugopal, K. R.. (2016). Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. World Wide Web , 20 , 135 - 154

. <http://doi.org/10.1007/s11280-015-0381-x>

1. Khan, Asifullah., Sohail, A.., Zahoora, Umme., & Qureshi, Aqsa Saeed. (2019). A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review , 53 , 5455 - 5516

. <http://doi.org/10.1007/s10462-020-09825-6>

1. Wankhade, M.., Rao, A. C.., & Kulkarni, Chaitanya. (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review , 55 , 5731 - 5780 . <http://doi.org/10.1007/s10462-022-10144-1>
2. Yadollahi, A.., Shahraki, Ameneh Gholipour., & Zaiane, Osmar R. (2017). Current State of Text Sentiment Analysis from Opinion to Emotion Mining. ACM Computing Surveys (CSUR) , 50 , 1 - 33 . <http://doi.org/10.1145/3057270>
3. He, Kaiming., Zhang, X.., Ren, Shaoqing., & Sun, Jian. (2015). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) , 770-778 . <http://doi.org/10.1109/cvpr.2016.90>
4. Huang, Shih-Cheng., Shen, Liyue., Lungren, M.., & Yeung, Serena. (2021). GLoRIA: A Multimodal Global-Local Representation Learning Framework for Label-efficient Medical Image Recognition. 2021 IEEE/CVF International

Conference on Computer Vision (ICCV) , 3922-3931



. <http://doi.org/10.1109/ICCV48922.2021.00391>

1. Quanzeng., Luo, Jiebo., Jin, Hailin., & Yang, Jianchao. (2015). Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks. , 381-388 . <http://doi.org/10.1609/aaai.v29i1.9179>