Project Report On

SocialWave: Instagram Engagement Predictor

Submitted in partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology

in

Computer Science and Engineering

By

Ann Jacob (U2103037)

Aron Jude Maxwel (U2103049)

Bilna Bijoy (U2103064)

Cerin Saji (U2103065)

Under the guidance of

Ms. Anu Maria Joykutty

Dept. of Computer Science and Engineering
Rajagiri School of Engineering & Technology (Autonomous)
(Parent University: APJ Abdul Kalam Technological University)
Rajagiri Valley, Kakkanad, Kochi, 682039

April 2025

CERTIFICATE

This is to certify that the project report entitled ***"SocialWave: Instagram Engagement Predictor"*** *is a bonafide record of the work done by* ***Ann Jacob (U2103037), Aron Jude Maxwel (U2103049), Bilna Bijoy (U2103064), Cerin Saji (U2103065)****, submitted to the Rajagiri School of Engineering & Technology (RSET) (Autonomous) in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology (B. Tech.) in "Computer Science and Engineering"during the academic year 2024-2025.*

Project Guide
Ms. Anu Maria Joykutty
Asst. Professor
Dept. of CSE
RSET

Project Co-ordinator
Mr. Harikrishnan M.
Asst. Professor
Dept. of CSE
RSET

Head of the Department
Dr. Preetha K G
Professor
Dept. of CSE
RSET

ACKNOWLEDGMENT

We wish to express our sincere gratitude towards **Rev. Fr. Dr. Jaison Mulerickal CMI**, Principal of RSET, and **Dr. Preetha K G**, Head of the Department of Computer Science and Engineering, for providing us with the opportunity to undertake our project, ˝SocialWave: Instagram Engagement Predictor˝.

We are highly indebted to our project coordinators, **Mr. Harikrishnan M.**, Assistant Professor, Department of Computer Science and Engineering, **Ms. Meenu Mathew**, Assistant Professor, Department of Computer Science and Engineering, and **Dr. Varghese Paul**, Professor, Department of Computer Science and Engineering, for their valuable support.

It is indeed our pleasure and a moment of satisfaction for us to express our sincere gratitude to our project guide **Ms. Anu Maria Joykutty** for her patience and all the priceless advice and wisdom she has shared with us.

Last but not the least, We would like to express our sincere gratitude towards all other teachers and friends for their continuous support and constructive ideas.

Ann Jacob
Aron Jude Maxwel
Bilna Bijoy
Cerin Saji

Abstract

Social media has proved to be essential to spread content using platforms like Instagram, which in turn has become a crucial hub for businesses and individual creators to engage with their audiences. SocialWave focuses on the development of a tool designed to enhance content strategy for Instagram. The core functionality includes predicting the engagement rate of posts based on images and hashtags and generating optimized captions as well as recommending valid hashtags either through user prompts or automated image analysis. The project employs robust machine learning techniques, integrating algorithms for engagement prediction and image analysis. Caption generation and hashtag recommendation are facilitated by transformer models, enabling seamless text generation based on visual content and user input. The primary advantage is that it follows a holistic approach in empowering users to create more impactful posts and consequently increase their potential for virality. The combination of advanced machine learning algorithms and deep learning techniques ensures high accuracy and adaptability to evolving social media trends, distinguishing SocialWave from current market solutions.

Contents

Acknowledgment **i**

Abstract **ii**

List of Abbreviations **vi**

List of Figures **vii**

List of Tables **viii**

1 Introduction **1**

1.1 Background 1

1.2 Problem Definition 1

1.3 Scope and Motivation 2

1.4 Objectives 2

1.5 Challenges 3

1.6 Assumptions 3

1.7 Organization of the Report 3

2 Literature Survey **5**

2.1 [1] Predicting Instagram Post Popularity using Hashtags, Image Analysis, and User History 5

2.1.1 Introduction 5

2.1.2 Key Points 5

2.1.3 Conclusion 7

2.2 [2] Comparative analysis of deep learning image detection algorithms . . . 7

2.2.1 Overview 7

2.2.2 Algorithms Under Study 7

2.2.3 Performance Evaluation 8

2.2.4 Strengths and Limitations . . . 8
2.2.5 Conclusion and Future Directions . . . 8
2.3 [3] Prompt-Based Learning for Unpaired Image Captioning . . . 9
2.3.1 Introduction . . . 9
2.3.2 Methods . . . 9
2.3.3 Conclusion . . . 10
2.4 [4] Prompt-Based Learning for Unpaired Image Captioning . . . 11
2.4.1 Introduction . . . 11
2.4.2 Methodology . . . 11
2.4.3 Experiments and Results . . . 11
2.4.4 Conclusion and Future Work . . . 12
2.5 Summary and Gaps Identified . . . 12
2.5.1 Gaps Identified . . . 12
2.5.2 Summary . . . 12

3 System Design **14**
3.1 System Architecture . . . 14
3.2 Component Design . . . 16
3.3 Use Case Diagram . . . 19
3.4 Tools and Techniques . . . 19
3.4.1 Hardware Requirements . . . 19
3.4.2 Software Requirements . . . 20
3.5 Module Division and Work Breakdown . . . 20
3.6 Key Deliverables . . . 21
3.7 Project Timeline . . . 23

4 System Implementation **24**
4.1 Overview . . . 24
4.2 Technologies Used . . . 24
4.3 Implementation Details . . . 25
4.3.1 Data Collection & Web Scraping . . . 25
4.3.2 Engagement Rate Prediction & Classification . . . 25

4.3.3 Image-Based Analysis: Object Detection & Feature Extraction for Multi-modal Hashtag Recommendation 26
4.3.4 Caption Generation 26
4.3.5 Web Application Development 26

5 Results and Discussions 27
5.1 User-friendly Graphical Interface (GUI) 27
5.2 Caption Generation from Images and Prompts 28
5.3 Hashtag Recommendation 30
5.4 Engagement Rate Prediction 33

6 Future Scope and Conclusion 35
6.1 Future Scope 35
6.2 Conclusion 35

Appendix A: Presentation 39

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes 63

Appendix C: CO-PO-PSO Mapping 68

List of Abbreviations

- **CNN** - Convolutional Neural Network
- **COCO** - Common Objects in Context (dataset)
- **F1 Score** - Harmonic mean of precision and recall
- **LDA** - Latent Dirichlet Allocation
- **LSTM** - Long Short-Term Memory (neural network architecture)
- **LXMERT** - Learning Cross-Modality Encoder Representations from Transformers
- **PL-UIC** - Prompt-Based Learning for Unpaired Image Captioning
- **R-CNN** - Region-based Convolutional Neural Network
- **RPN** - Region Proposal Network
- **SSD** - Single Shot Detection
- **SVR** - Support Vector Regression
- **TF-IDF** - Term Frequency-Inverse Document Frequency
- **UIC** - Unpaired Image Captioning
- **ViT** - Vision Transformer
- **VL-PTM** - Vision-Language Pre-Trained Model
- **YOLO** - You Only Look Once

List of Figures

3.1 System Architecture diagram 14
3.2 Sequence Diagram 17
3.3 Use Case Diagram 19
3.4 Instagram scraping result for ”ieesb” 21
3.5 Hashtag generation based on image 22
3.6 Object detection module output 22
3.7 Gantt Chart 23

5.1 Landing Page 27
5.2 Authentication pages 27
5.3 Training loss reduction over epochs. 28
5.4 Caption and Hashtag Generation from Images 29
5.5 Caption and Hashtag Generation from Prompts 29
5.6 ViT Training Loss Reduction Over Epochs 30
5.7 Batch-wise Loss During Testing 31
5.8 LXMERT Model Evaluation Results 32
5.9 Engagement Rate Prediction Results 34

List of Tables

2.1 Summary of Literature Survey 13

Chapter 1

Introduction

1.1 Background

SocialWave is concerned with guiding strategic planning in creating content on Instagram, for a long time the major channel for sharing visual content. Instagram popularity has shot up with about 2 billion active users monthly in a way that makes it a vital tool for individuals, influencers, and brands in audiences and interaction. Nevertheless, optimizing posts to maximize reach and engagement remains a challenge due to the complex, unstable, and inspecial algorithms of the platform. Current options include social media analytics tools that mostly limit themselves to basic insights on follower growth, average post engagement, and demography. Although these metrics are valuable, they do not include any predictive model nor features for content optimization that pertain to specific posts. Such opens up a larger breach that creates an inherent need for stronger solutions that could integrate data-driven insights with enhancements to content strategy in order to facilitate maximizing audience reach by content creators. Informed by its innovation-the combination of engagement prediction and intelligent caption generation-SocialWave rides into the market as a universal tool combining post-analysis and content-generation aspects with machine learning and deep learning technologies.

1.2 Problem Definition

The SocialWave project's main purpose is to create an integrated tool that will examine the crucial issue of engagement optimization in relation to the post of a content creator on Instagram. The project attempts to fill the gap neighboring the present tools by predicting engagement rate features based on a combination of photos, hashtags and other features, allowing simple reproduction of captions based on photo content or user-entered prompts.

1.3 Scope and Motivation

The project revolves around the delivery of an all-inclusive web application meant for new Instagram content creators so as to optimize their posts for higher user engagement. The proposed tool features predictive analytics employing a regression model, for example, Random Forest, to forecast engagement levels based on image and hashtag input. Moreover, it integrates image detection algorithms that leverage deep learning methods, such as YOLO, in analyzing visual content in general. The project will also incorporate a Natural Language Process-oriented caption generation system that enables captions to be created directly from images or user prompts, making it an adaptable toolbox for content creation. The reason for Socialwave's birth has been necessitated by increasing appeals for advanced, data-informed tools, designed to go beyond the basic analytics of social media platforms to aid content creation. Most of the solutions available at the moment present one with post-performance statistics, with rather limited predictive capability and content-generation facilities. The dynamic nature of social media algorithms means that merely possessing historical metrics isn't sufficient- creators desire tools that help them visualize and conjure future engagement. This method of combining machine learning and deep learning gives SocialWave users insight into the metrics that drive user engagement so they can better tailor their posts' engagement potential. This project aims to fill existing gaps to make content creation easy for newly launching content creators targeting an audience with engaging creativity that works in resounding effect on their audience.

1.4 Objectives

1. Develop a machine learning model for predicting engagement rates in Instagram posts based on user engagement metrics and hashtags.

2. To implement deep learning algorithms for image detection, namely YOLO and Faster-RCNN, to analyze visual content in posts.

3. To create a robust caption generation system using a transformer-based model that can generate relevant captions from images or user-entered prompts.

4. To design and develop a user-friendly web application that integrates engagement prediction and caption generation features.
5. To assist content creators in optimizing their posts to increase reach and maximize the chances of virality through data-driven insights.

1.5 Challenges

Developing SocialWave involves overcoming several challenges, including the integration of multiple machine learning and deep learning models to ensure seamless and accurate predictions and analyses. Ensuring real-time performance for image detection and caption generation while maintaining the accuracy and scalability of the application adds further complexity. Additionally, scraping instagram data has proven to be complex due to various security measures implemented such as rate limiting and IP blocking.

1.6 Assumptions

1. Social media algorithms will not change drastically during the project's development and testing phases, allowing the predictive models to remain effective.
2. Users have access to a stable internet connection for seamless use of the web application.
3. The algorithms will function optimally with the hardware and software requirements outlined in the project's specifications.

1.7 Organization of the Report

The report begins with preliminary sections, including the Acknowledgment, Abstract, and lists of abbreviations, figures, and tables. Chapter 1 introduces the topic, covering the background, problem definition, objectives, and organization of the report. Chapter 2 reviews related research and identifies gaps. Chapters 3 and 4 detail the methodology and implementation, followed by Chapter 5, which presents results and discussions. Chapter 6 concludes the study with findings and future scope. The report concludes with References, a List of Publications, and appendices, including presentation, course outcomes, and a CO-PO-PSO mapping.

In summary, this chapter outlines the foundation of SocialWave, emphasizing its goal to help Instagram content creators optimize posts for engagement by integrating predictive analytics, image detection, and caption generation. It highlights the project's relevance, objectives, challenges, and assumptions, setting the stage for the detailed methodologies and technologies discussed in subsequent chapters.

Chapter 2

Literature Survey

2.1 [1] Predicting Instagram Post Popularity using Hashtags, Image Analysis, and User History

2.1.1 Introduction

The emergence of social media platforms, especially Instagram, has significantly changed the digital marketing and influencer landscape. As brands and individuals work to enhance their reach and effectiveness, it has become crucial to understand what drives post popularity. This literature review explores current research on Instagram post popularity, emphasizing the impact of hashtags, image content, user history, and machine learning methods in forecasting engagement rates. It also points out the shortcomings in existing studies, particularly concerning data variability and the incorporation of various predictive features.

2.1.2 Key Points

1. Popularity Prediction in Social Media

   Research has consistently highlighted the importance of post popularity on social media for effective marketing. Various metrics, including likes, comments, and shares, have been identified as key indicators of user engagement. These metrics not only show how well individual posts perform but also impact overall brand visibility and user perception. However, many studies have used localized datasets, which limits the applicability of their findings.

2. Influence of Hashtags

   Hashtags play a crucial role in increasing visibility and engagement on Instagram. Research has shown that using hashtags strategically can lead to higher interaction rates. While earlier studies often concentrated on the number of hashtags used, more recent research indicates the importance of considering factors like hashtag popularity and visibility. This paper builds on that concept by examining how various characteristics of hashtags influence engagement, filling a gap in the existing literature.

3. Image Content and Quality

   The visual aspect of Instagram plays a crucial role in the popularity of posts, as image content is key. Research indicates that visually appealing images are more likely to receive higher likes and comments. Various studies have employed both automated and manual assessments to gauge image quality. Nevertheless, the use of thorough image evaluation features in predictive models is still quite limited. This paper aims to address this gap by combining both manual and automated evaluations of image quality to improve prediction accuracy.

4. User History and Engagement Metrics

   The history of the user, such as parameters related to past engagement rate or follower dynamics, is extremely crucial for the prediction of future post performance. Studies have found that users are usually of high engagement with the same behavior. However, many existing works ignore that or use localized datasets that cannot capture global engagement trends. This study presents an opportunity to observe, through a global dataset, how user history applies to engagement variability.

5. Challenges with Data Variability

   The high data variability when using global datasets in studies poses a big challenge in the literature today. Many studies have focused on localized data, the reliability of which will show many regions that reflect diverse dynamics in global contexts. This research deals with this challenge by analyzing a global dataset rather than a localized one, and as such, it provides more relevant findings across different user populations and geographic areas.

6. Machine Learning Approaches

   The utilization of machine learning approaches, particularly regression models, is explored for enhancing the accuracy of popularity predictions. Support Vector Regression (SVR) has been shown to outperform conventional statistical methods when it comes to delivering better prediction accuracies. In this paper, we demonstrate how SVR is used to effectively predict engagement rates based on an extensive set of features comprising hashtags, image ratings, and the user's profile, thus adding knowledge to the rapidly advancing field of social media analytics.

2.1.3 Conclusion

Recent research has uncovered an intricate web of relationships among the different factors influencing user interaction through Instagram posts. Earlier studies have provided a foundation for understanding the nature of these dynamics, while this research has consolidated the existing findings with a broader array of features in a world context. Addressing problems on variance and multicollinearity in data is worthwhile for marketers and content creators wishing to optimize their brand ambassadorship on Instagram. Findings include the need for more robust predictive models integrating various engagement factors so that the dynamics of social media can be explained more finely in our contemporary digital world.

2.2 [2] Comparative analysis of deep learning image detection algorithms

2.2.1 Overview

Srivastava et al. (2021) authorized such a thorough comparative analysis of three object detection algorithms: Single Shot Detection, Faster Region-based Convolutional Neural Networks (Faster R-CNN), and You Only Look Once. This literature review outlines the important findings and contributions detailed in the paper.

2.2.2 Algorithms Under Study

1. Single Shot Detection (SSD): This model is known for its speed and efficiency, which uses a single deep neural network for the hint. The architecture is explained in the paper, where multiple feature maps are used to predict a bounding box and class score concurrently, making real-time processing possible.

2. Faster R-CNN: This algorithm builds on the original R-CNN method by introducing a Regions Proposal Network (RPN) that greatly reduces the computation in generating object proposals. The authors emphasize the accuracy and robustness of this method, especially in complex scenes.

3. You Only Look Once (YOLO): YOLO is distinguished by its unique way of treating object detection as a single regression problem, so it can process images in real time. The paper details YOLO's architecture, which predicts bounding boxes and class probabilities directly from full images, yielding remarkable speed and efficiency.

2.2.3 Performance Evaluation

Employing the Microsoft COCO dataset, the authors performed a detailed evaluation of the accuracy, precision, and F1 measures for SSD, Faster R-CNN, and YOLO. The results indicated that although speed is superior to the other two algorithms with YOLO-v3, the SSD was the best-balanced among speed and accuracy, thereby favoring applications that require real-time detection.

2.2.4 Strengths and Limitations

The paper provides detailed insights into the strengths and weaknesses of each algorithm:

- YOLO does very well in speed but may not work well in detecting small objects.
- Faster R-CNN provides very high accuracy at a cost-it is slow in doing this.
- SSD reports to be fast while providing moderate accuracy, making it suitable for a variety of applications. enditemize

2.2.5 Conclusion and Future Directions

The authors conclude that the choice of algorithm depends on the specific use case and requirements of the application. They suggest that future research could focus on enhancing the performance of these algorithms in real-time scenarios, particularly for small object detection and improving robustness against occlusions.

2.3 [3] Prompt-Based Learning for Unpaired Image Captioning

2.3.1 Introduction

Hashtag recommendation systems have gained immense popularity lately; they help viewers find content on social media platforms. With the abysmal increase in user-generated content on platforms like Twitter, Instagram, and Sina Weibo, effective hashtag prediction becomes important for optimizing user engagement and spreading information. Traditionally, however, a lot of research has focused on textual data, usually ignoring the potential of multimodal content encompassing text, images, and videos. The present literature review investigates the evolution of hashtag recommendation methodologies with an emphasis on the growing tendency toward enriching recommendation strategies with multimodal data and more sophisticated methods, especially transformer-based models.

2.3.2 Methods

1. Traditional Text-Based Approaches: Early models for recommending hashtags mainly relied on textual data. Techniques included:

   -IDF: This method used content similarity metrics to suggest hashtags based on existing user tweets (Zangerle et al.).

   - LDA: This approach employed topic modeling to pinpoint relevant hashtags based on the thematic content of microblogs.
   - Attention Mechanisms in LSTM: Models like TAB-LSTM incorporated attention mechanisms into LSTM architectures to enhance contextual awareness in hashtag predictions.

2. Image-Based Approaches: With the rise of visual content, researchers shifted their focus to image-driven models for hashtag recommendations:

   - Convolutional Neural Networks (CNNs): These were used to extract visual features pertinent to hashtags, as demonstrated in models like HARRISON.
   - Personalized Tag Recommendation: Such models took into account user preferences and visual context, thereby improving the accuracy of hashtag suggestions.

3. Multimodal Models: Acknowledging the significance of both text and images, several studies created multimodal approaches:

   - Co-Attention Mechanisms: Models like CoA and AMNN utilized co-attention networks to merge visual and textual features for hashtag recommendations.
   - Hybrid Neural Networks: These approaches sequentially extracted features from both images and texts before integrating them for better recommendation accuracy.

4. Transformer-Based Methods: The advent of transformer architectures represented a major leap forward in the field:

   - BERT and Its Variants: By leveraging self-attention mechanisms, transformer models such as LXMERT have excelled in capturing intricate interactions between text and images.
   - Cross-Modality Representation Learning: These models enable a deeper understanding of how visual and textual data relate to hashtag relevance, significantly enhancing prediction outcomes.

2.3.3 Conclusion

The literature on hashtag recommendation systems shows a clear trend towards integrating multimodal data, shifting from traditional text-focused models to more advanced transformer-based approaches. There is a growing acknowledgment of how visual content can improve recommendation accuracy, which has spurred the creation of hybrid models that leverage both text and visual features. The use of transformers like LXMERT marks a significant advancement in this field, providing strong frameworks for learning representations across different modalities. Future research will likely aim to refine these models further, investigate their applications on various social media platforms, and enhance their adaptability to different types of content. This progression highlights the increasing complexity and potential of hashtag recommendation systems in the digital landscape.

2.4 [4] Prompt-Based Learning for Unpaired Image Captioning

2.4.1 Introduction

"Prompt-Based Learning for Unpaired Image Captioning" introduces a new framework, PL-UIC, for generating image captions without using aligned image-caption pairs-Prompt-based learning for unpaired image captioning created by prompting the vision-language pre-trained models in unaligned image captioning processes. Traditional unpaired image captioning methods have difficulties capturing cross-domain correlations between visual and textual information. For this purpose, the authors propose two types of prompts:

2.4.2 Methodology

- **Semantic Prompt Generation:** Each image has been mapped to its semantic prompt via the CLIP model, followed by a layer of the feed-forward network. The output intends to increase the cosine similarity between the image features and text features. They contain special embeddings that mark the beginning and the end of the caption (like "SOS", "CLS", "EOS"). The Caption Generation Process uses the CNN- LSTM model, with the CNN modeling the image features and the LSTM translating these features to natural language captions. The semantic prompts are combined with image features, influencing the generation of this caption.

- **Metric Prompt Filtering:** The metric prompt serves as a criterion for filtering generated image-caption pairs based on their cosine similarity. Pairs that exceed a certain threshold are retained for further training, while others are discarded.

2.4.3 Experiments and Results

- Datasets: The model was evaluated using established datasets, COCO and Flickr30K, which serve as standard benchmarks for image captioning tasks

- Performance: The findings show that the proposed PL-UIC model significantly surpasses current UIC models, showcasing enhanced caption quality and relevance. By incorporating semantic and metric prompts, the model effectively utilizes a wealth of prior knowledge.

2.4.4 Conclusion and Future Work

The research concludes that prompt-based learning can substantially improve UIC, and the authors encourage further exploration of this approach in various applications requiring cross-domain knowledge extraction.

2.5 Summary and Gaps Identified

2.5.1 Gaps Identified

1. **Incomplete Feature Integration:**: Many current models do not fully integrate user history, hashtag trends, and image content, which limits their predictive capabilities.

2. **Issues with Small Object Detection:**: In image detection, models like YOLO still struggle with identifying small objects, which affects accuracy in intricate images.

3. **Narrow Dataset Usage:**: A lot of research depends on localized datasets that fail to reflect global engagement trends, making the findings less relevant across various user groups.

4. **Flexibility of Transformer Models:**: Although transformer models have shown potential, their use across different social media platforms and content types requires more investigation to improve their adaptability.

5. **Challenges in Cross-Modality Representation:** : Even with progress in multimodal and prompt-based learning, effectively capturing the interplay between text and visual content remains a challenging issue that calls for more sophisticated methods.

2.5.2 Summary

Title	Advantages	Disadvantages
[1] **Predicting Instagram Post Popularity using Hashtags, Image Analysis, and User History**	Integrates hashtags, image quality, and user history for prediction; uses global datasets for comprehensive analysis	Limited integration of user history features in previous studies; challenges with data variability in existing models
[2] **Comparative Analysis of Deep Learning Image Detection Algorithms**	Provides comparative insights on SSD, Faster R-CNN, and YOLO; highlights algorithm-specific strengths and weaknesses	YOLO struggles with small objects; Faster R-CNN is slower compared to other models
[3] **Hashtag Recommendation Systems**	Evolution from text-based to multimodal approaches; effective use of transformer-based models	Earlier models were text-centric and lacked multimodal integration; newer models need improved adaptability
[4] **Prompt-Based Learning for Unpaired Image Captioning**	Introduces PL-UIC framework leveraging VL-PTMs; significant performance improvement with semantic and metric prompts	Limited cross-domain correlation capture in traditional methods; prompt-based learning is still a developing area

Table 2.1: Summary of Literature Survey

Chapter 3

System Design

The System Design chapter describes the architecture and components of the SocialWave project. It explains how the different components interact and their roles in meeting the project's objectives. Visual aids such as Use Case Diagrams show how data flows through the system, offering a clear view of the overall design.

3.1 System Architecture

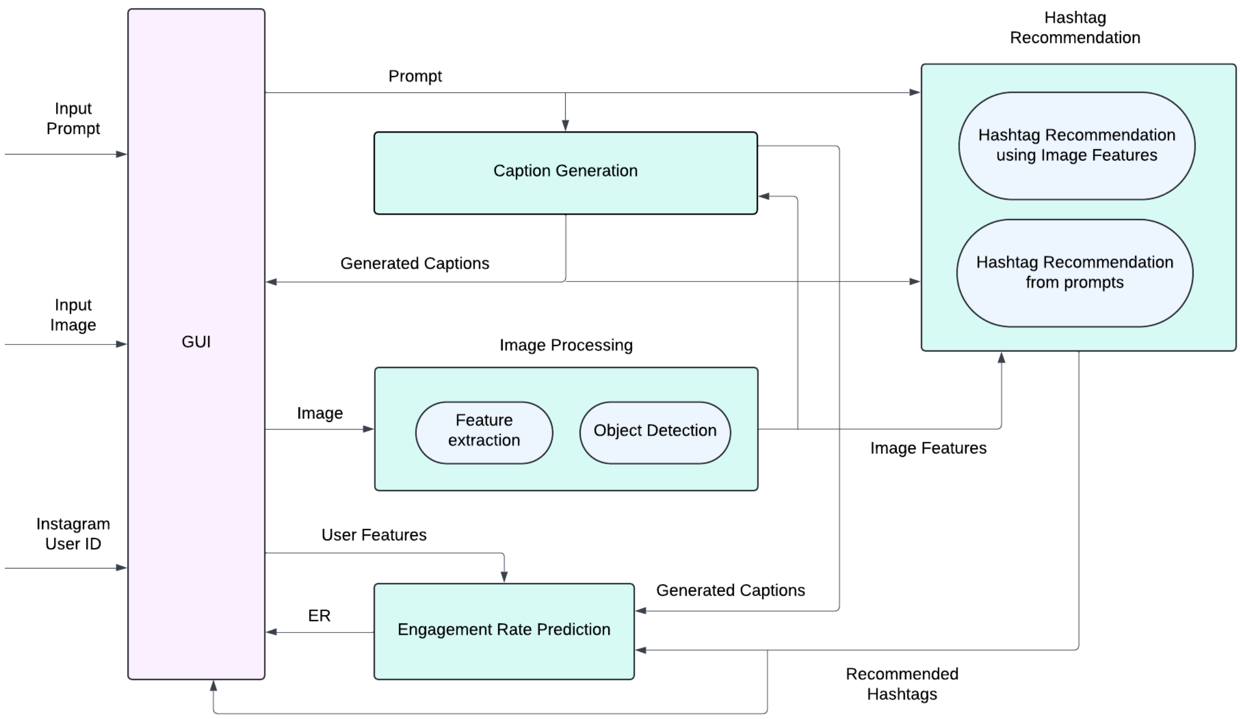


Figure 3.1: System Architecture diagram

The workflow can be summarized as follows:

- **Input Section:**

  - The system takes in two primary inputs: an image and a prompt (text provided by the user).
  - Additionally, the user's Instagram User ID is utilized for tailored engagement rate predictions.

- **GUI (Graphical User Interface):**
  - The GUI acts as the main interface for users, enabling them to submit inputs and view outputs (captions, hashtags, and engagement predictions).

- **Caption Generation:**
  - The system generates relevant captions for the uploaded image based on the input prompt.
  - These generated captions are then displayed through the GUI.

- **Image Processing:**
  - The input *image* undergoes:
    * *Feature Extraction:* This step identifies key characteristics from the image.
    * *Object Detection:* This process recognizes objects within the image to comprehend its content.

- **Hashtag Recommendation:**
  - Hashtags are suggested through two methods:
    * *Using Prompts:* Hashtags are recommended based on the prompt provided by the user.
    * *Using Image Features:* Hashtags are created based on the objects and features identified in the image.

- **Engagement Rate Prediction:**
  - The system leverages image features and user-specific data to forecast the Engagement Rate (ER) of the post.

- The predicted engagement rate and generated hashtags are sent back to the GUI for user evaluation.

- **Outputs:**
  - The system delivers the following outputs:
    * *Generated Captions:* Based on the provided prompt or generated from image features.
    * *Recommended Hashtags:* Suggested from both the prompt and image features.
    * *Predicted Engagement Rate:* Assists users in optimizing their Instagram posts for increased engagement.

3.2 Component Design

The sequence diagram illustrates the interaction flow among the different components of the SocialWave system. The steps are detailed as follows:

- **Components:**
  - **User:** The individual engaging with the system.
  - **User Interface (UI):** The front-end interface for user input and output display.
  - **Back-end API:** The intermediary that connects the UI with the system's processing modules.
  - **Image Analysis System:** Responsible for detecting objects and extracting features from images.
  - **Engagement Rate (ER) Prediction Model:** Predicts the engagement rate based on image and user data.
  - **Caption and Hashtag Generation Model:** Creates captions and hashtags based on user-provided prompts and image features.
- **Sequence of Events:**

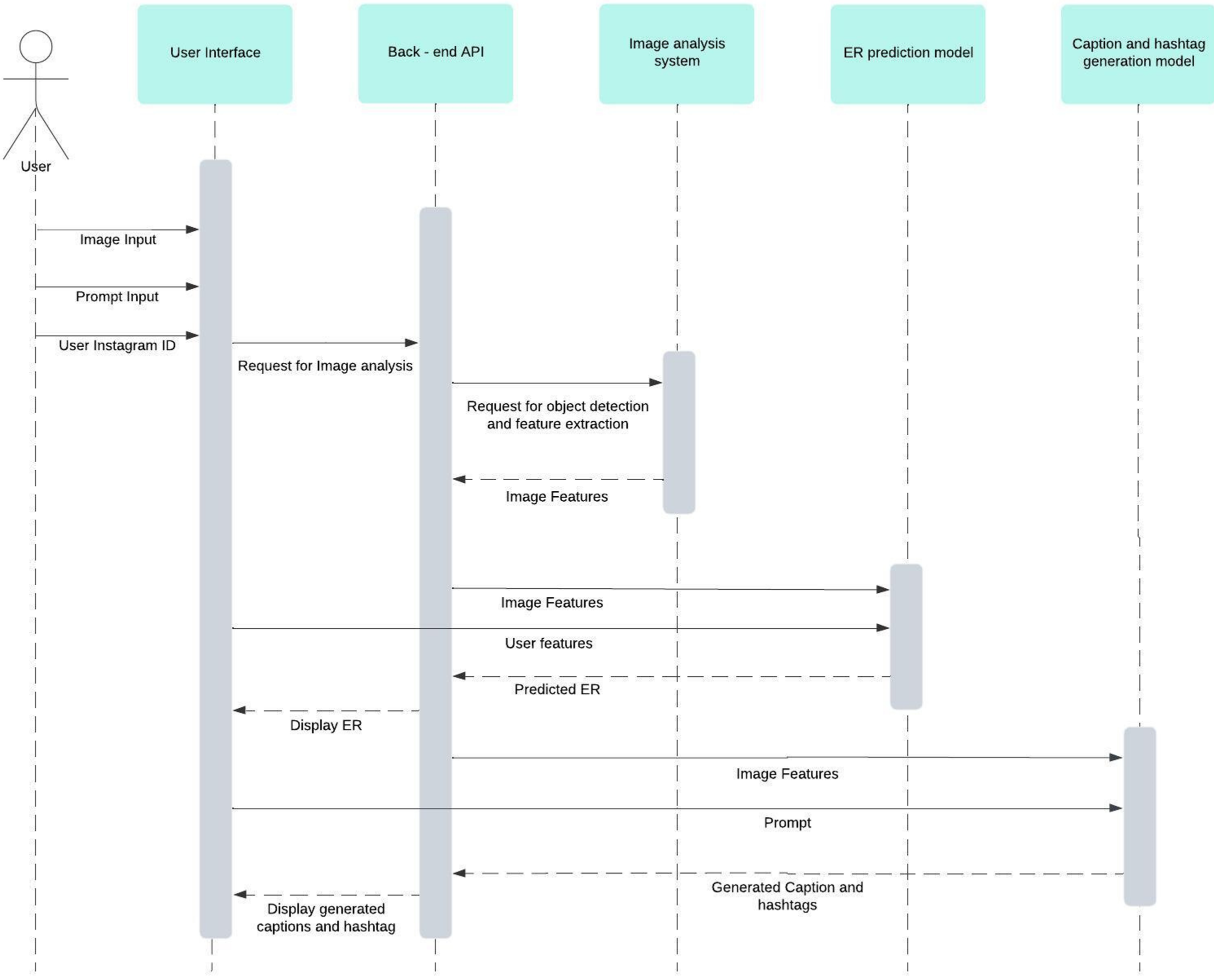


Figure 3.2: Sequence Diagram

1. **Input Gathering:**
   - The user submits inputs through the UI: an image, a prompt, and their Instagram User ID.
2. **Request for Image Analysis:**
   - The UI sends a request for image analysis to the Back-end API.
3. **Image Processing:**
   - The Back-end API forwards the request to the Image Analysis System, which performs:
     * *Object Detection:* Identifies objects within the image.
     * *Feature Extraction:* Extracts key characteristics from the image.
   - The extracted *Image Features* are returned to the Back-end API.
4. **Engagement Rate Prediction:**
   - The Back-end API sends the Image Features and User Features (derived from the Instagram User ID) to the ER Prediction Model.
   - The ER Prediction Model calculates the Predicted Engagement Rate and sends it back to the Back-end API.
   - The Back-end API then relays the predicted ER to the UI for display.
5. **Caption and Hashtag Generation:**
   - The Back-end API forwards the Image Features and Prompt to the Caption and Hashtag Generation Model.
   - The model generates captions and hashtags based on either the provided prompt or the extracted image features as per the user's requirements.
   - The generated captions and hashtags are sent back to the Back-end API and displayed on the UI.

- **Outputs:**
  - *Predicted Engagement Rate:* Displayed after analysis.
  - *Generated Captions and Hashtags:* Shown on the UI.

3.3 Use Case Diagram

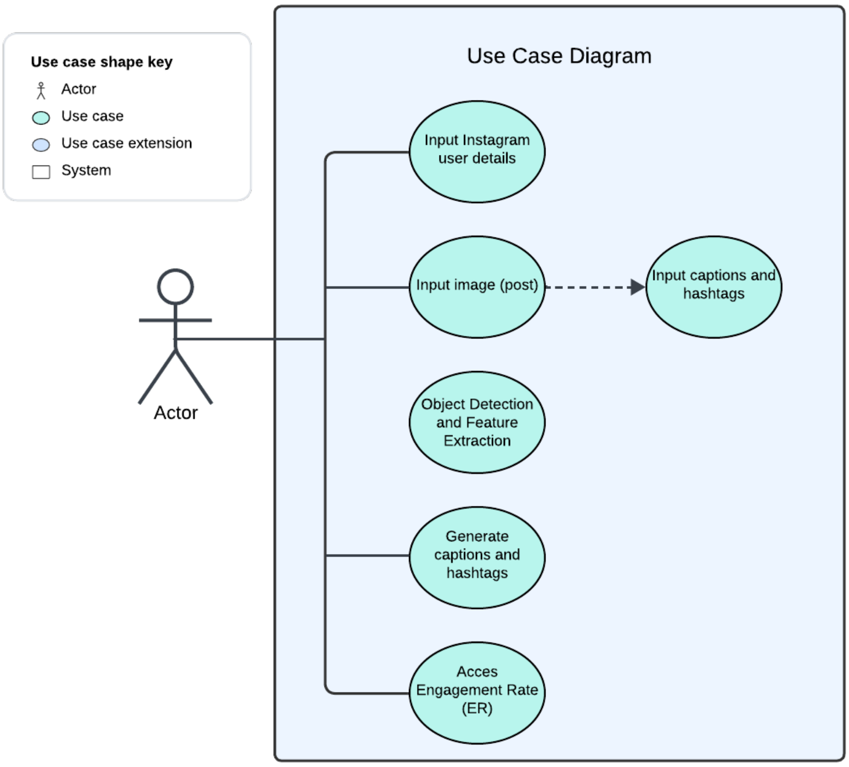


Figure 3.3: Use Case Diagram

3.4 Tools and Techniques

3.4.1 Hardware Requirements

- Processor: A modern multi-core processor (e.g., Intel Core i5) for efficient code compilation and development tasks.
- RAM: A minimum 8GB of RAM to handle multiple development tools and applications simultaneously without slowdowns. However, 16 would be recommended.
- Storage: A solid-state drive (SSD) with sufficient storage capacity (e.g., 256GB or higher) for storing project files, libraries, and development tools.

3.4.2 Software Requirements

- Python, Flask, HTML5, CSS3 and JavaScript and frameworks like React for front-end development.
- Libraries (TensorFlow and PyTorch) for training deep learning models.
- Pre-trained model (YOLOv5 and Faster-RCNN) for object detection and feature extraction tasks.

3.5 Module Division and Work Breakdown

- **Module 1: Caption Generation**
  - **Objective**: Generate meaningful and engaging captions for Instagram posts based on image content and contextual data.
  - **Work Allocation**: Ann Jacob
- **Module 2: Object Detection and ER Prediction**
  - **Objective**: Detect objects in images and predict the engagement level (likes, comments, shares) for an Instagram post based on image features and hashtags.
  - **Work Allocation**: Bilna Bijoy (ER Prediction using SVR, XGBoost)
- **Module 3: Hashtag Recommendation and UI Development**
  - **Objective**: Recommend relevant and high-performing hashtags to increase post visibility and reach using *ViT Classifier* and *LXMERT4Hashtag* on the *HARRISON* dataset.
  - **Additional Responsibilities**: Landing page UI, Front-end + Backend integration.
  - **Work Allocation**: Cerin Saji
- **Module 4: Web Scraping and ER Prediction**
  - **Objective**: Scrape images from Instagram to build the dataset and predict engagement levels using **Random Forest** and **Linear Regression**.

– **Additional Responsibilities**: UI Development (Login page, Signup page, Main page).

– **Work Allocation**: Aron Jude Maxwel

3.6 Key Deliverables

Expected Outputs : Results are obtained for the object detection module, Instagram scraping, caption generation, hashtag recommendation, and ER prediction module.

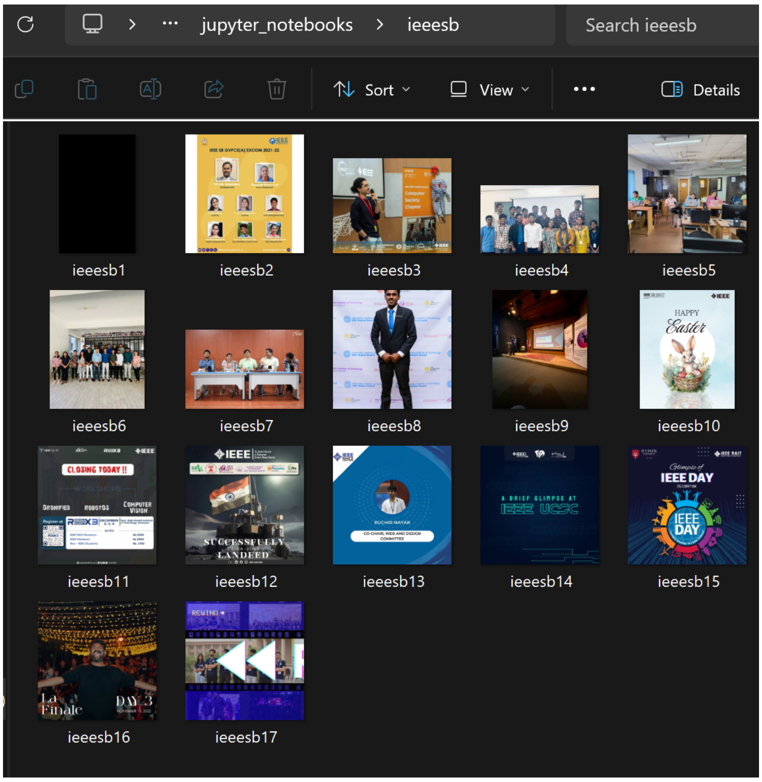


Figure 3.4: Instagram scraping result for ”ieesb”

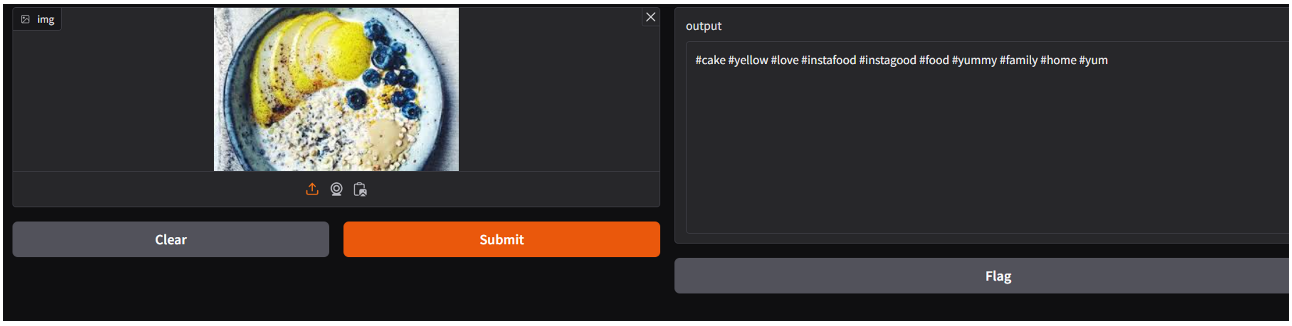


Figure 3.5: Hashtag generation based on image

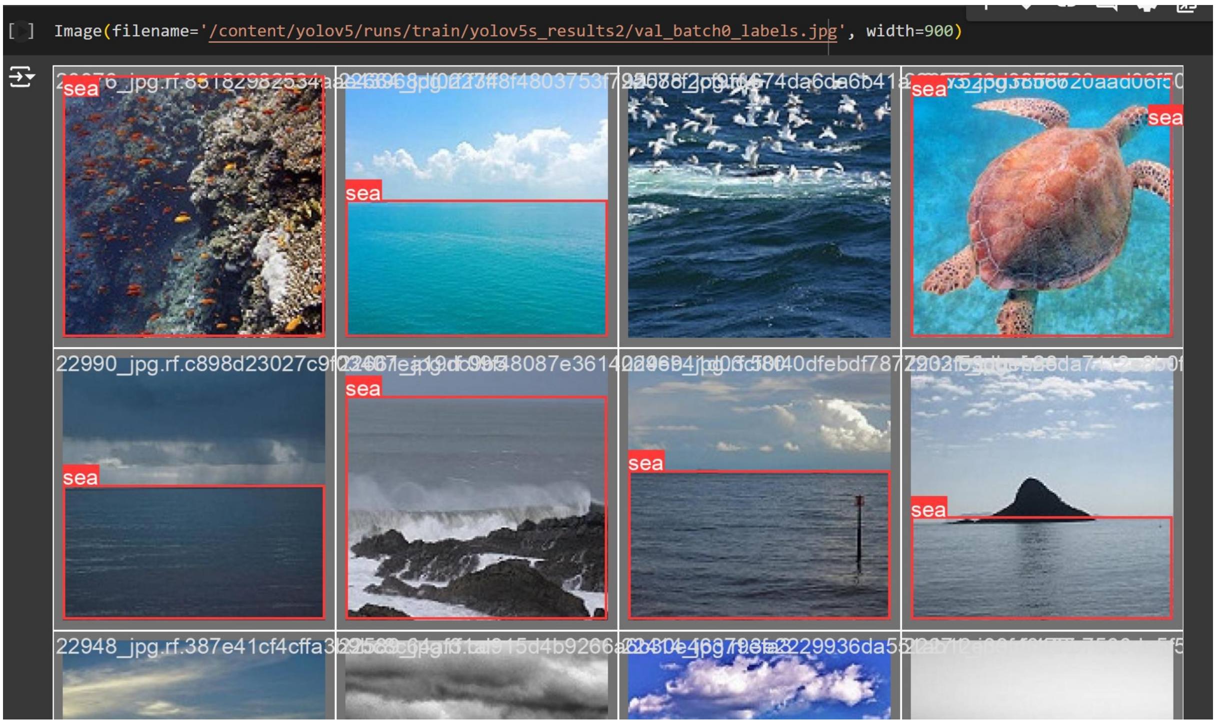


Figure 3.6: Object detection module output

3.7 Project Timeline

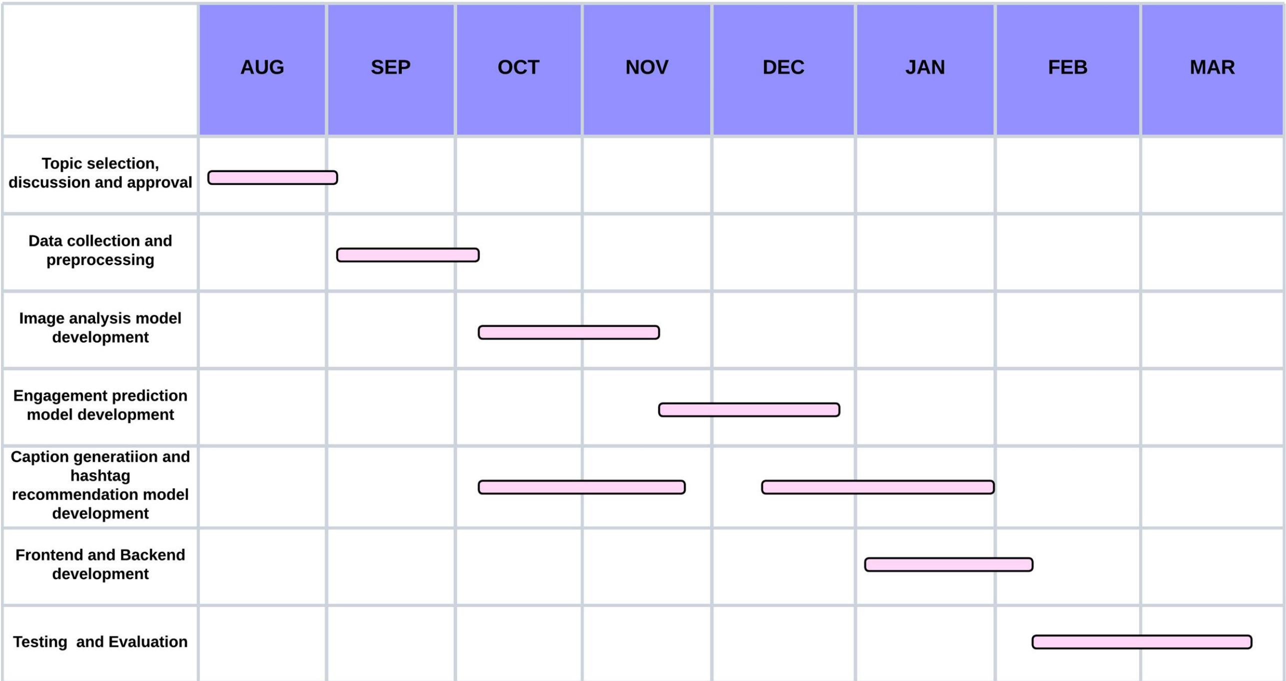


Figure 3.7: Gantt Chart

Chapter 4

System Implementation

4.1 Overview

The SocialWave project is a social media trend tracker that predicts engagement rates on Instagram posts and generates AI-powered captions. It is implemented as a web application using Flask as the backend, MongoDB for storing login credentials, and HTML, CSS, and JavaScript for the frontend. The system employs machine learning models for engagement rate prediction, deep learning models for caption and hashtag generation, and object detection models for analyzing image content.

4.2 Technologies Used

- **Frontend**: HTML, CSS, JavaScript
- **Backend**: Flask (Python-based web framework)
- **Database**: MongoDB (used exclusively for storing login credentials)
- **Web Scraping**: Selenium (Python) for extracting Instagram post data
- **Machine Learning Models**:
  - Random Forest Regressor (Best Performance)
  - Linear Regression
  - Support Vector Regression (SVR)
  - XGBoost
- **Deep Learning Models**:
  - Fine-tuned Microsoft Git Base Model for image-based caption generation

- LXMERT for hashtag generation (outperformed ViT Classifier)

- **Object Detection Models**:
  - YOLOv5 (initial model)
  - Faster R-CNN (final model, gave better results)

- **Prompt-Based Caption and Hashtag Generation**: Gemini API

- **Libraries & Frameworks**:
  - Data Processing: NumPy, Pandas
  - Machine Learning: Scikit-learn, XGBoost
  - Deep Learning: PyTorch, Transformers (Hugging Face)

4.3 Implementation Details

4.3.1 Data Collection & Web Scraping

The images were scraped from Instagram using Selenium. The datasets for the predicted engagement rate and hashtag quality calculation were obtained from Kaggle. The hashtag quality dataset contains popular hashtags and a popularity score.

4.3.2 Engagement Rate Prediction & Classification

The **initial engagement rate** (ER) is calculated as:

$$ER = \frac{\text{Likes} + \text{Comments}}{\text{Followers}} \times 100 \tag{4.1}$$

The **adjusted engagement rate** is obtained by incorporating **hashtag quality**. The **final engagement rate classification** is as follows, based on trial and error:

- **Low**: $ER < 5$
- **Moderate**: $5 \leq ER < 12$
- **High**: $12 \leq ER \leq 25$
- **Viral**: $ER > 25$

4.3.3 Image-Based Analysis: Object Detection & Feature Extraction for Multimodal Hashtag Recommendation

- **Object Detection and Feature Extraction**:
  - Initially implemented using YOLOv5 for detecting key elements in images.
  - Later replaced with Faster R-CNN, which provided better detection accuracy.
- **Hashtag Recommedation based on Image Features**:
  - Predicted using LXMERT, which outperformed the ViT Classifier by modeling image-text relationships more effectively.
  - Hashtag quality is integrated into the engagement rate model to improve post virality predictions.

4.3.4 Caption Generation

- **Image-Based Captions**: Generated using a fine-tuned Microsoft Git Base Model, trained on social media caption datasets.
- **Prompt-Based Captions**: Users can enter a text prompt to generate a caption using Gemini API, which provides high-quality, context-aware captions.

4.3.5 Web Application Development

- **Frontend**: Implemented using HTML, CSS, and JavaScript to provide a responsive UI.
- **Backend**: Implemented in Flask, exposing REST APIs for machine learning and deep learning model inference.
- **Database**: MongoDB securely stores user login credentials. It can be used to access previously generated captions, hashtags, and previous results.

Chapter 5

Results and Discussions

5.1 User-friendly Graphical Interface (GUI)

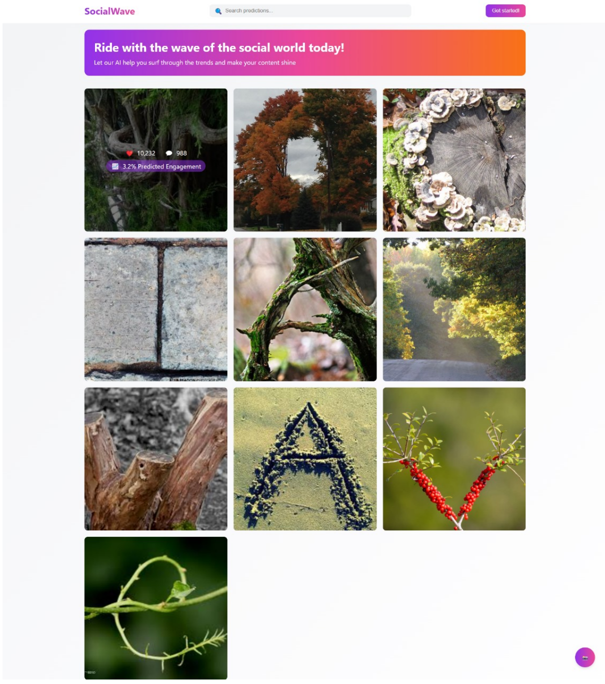


Figure 5.1: Landing Page

(a) Login Page

(b) Signup Page

Figure 5.2: Authentication pages

5.2 Caption Generation from Images and Prompts

Microsoft Git-base is a transformer-based language model designed for natural language generation (NLG) tasks. Originally developed for text completion and code generation, it leverages deep learning techniques to predict and generate coherent text sequences. By fine-tuning this model on Instagram captions, we enhanced its ability to understand social media language and produce engaging captions tailored to various contexts.

Data collection involved curating a dataset of Instagram posts, including images, captions, and engagement metrics. Additionally, popular hashtags and keywords were extracted to understand common linguistic patterns, ensuring the model was trained on relevant social media content.

The model fine-tuning process involved training the Microsoft Git-base model on a dataset of high-performing Instagram captions. Transfer learning was applied to enhance the model's ability to generate social media-friendly text. The training was conducted using PyTorch and the Hugging Face Transformers library, ensuring an optimized approach for generating high-quality captions.

```
Epoch 1/25 started...
Epoch 1/25 completed. Avg Loss: 0.0867 | Time elapsed: 48.36 sec

Epoch 2/25 started...
Epoch 2/25 completed. Avg Loss: 0.0640 | Time elapsed: 48.12 sec

Epoch 3/25 started...
Epoch 3/25 completed. Avg Loss: 0.0505 | Time elapsed: 48.12 sec

Epoch 4/25 started...
Epoch 4/25 completed. Avg Loss: 0.0437 | Time elapsed: 49.50 sec

Epoch 5/25 started...
Epoch 5/25 completed. Avg Loss: 0.0384 | Time elapsed: 49.56 sec

Epoch 6/25 started...
Epoch 6/25 completed. Avg Loss: 0.0349 | Time elapsed: 49.79 sec

Epoch 7/25 started...
Epoch 7/25 completed. Avg Loss: 0.0330 | Time elapsed: 48.75 sec

Epoch 8/25 started...
Epoch 8/25 completed. Avg Loss: 0.0318 | Time elapsed: 48.05 sec

Epoch 9/25 started...
...

Epoch 25/25 started...
Epoch 25/25 completed. Avg Loss: 0.0224 | Time elapsed: 48.11 sec

Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

Figure 5.3: Training loss reduction over epochs.

To further enhance the caption generation process, prompts were utilized with the Gemini API to generate creative and engaging captions. The Gemini API, leveraging

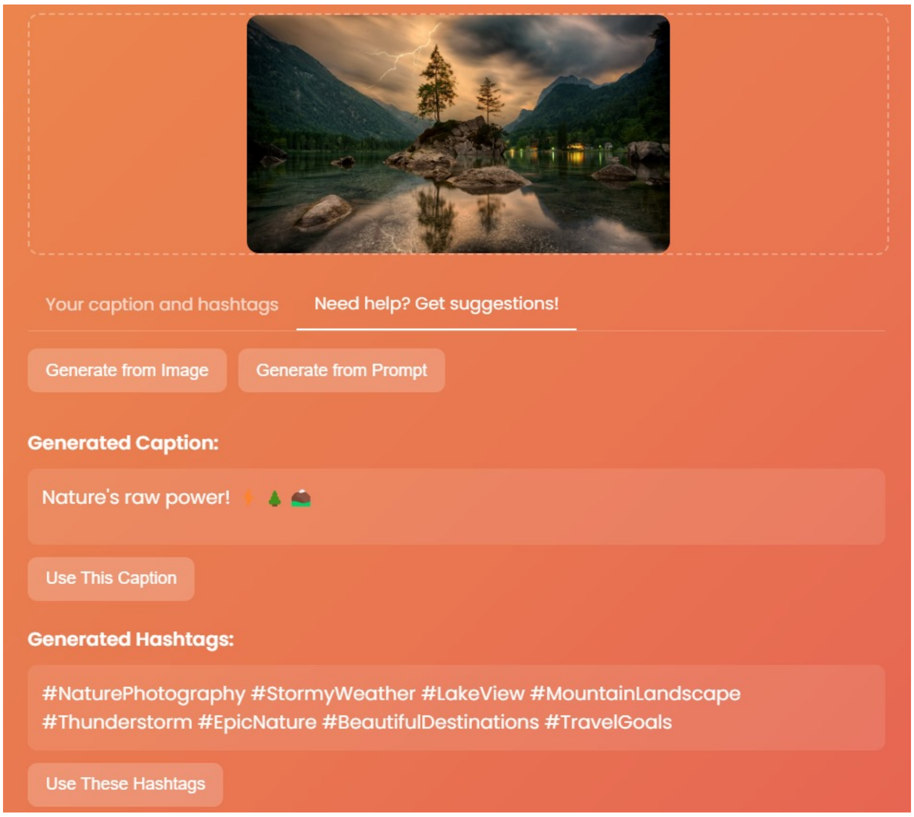


Figure 5.4: Caption and Hashtag Generation from Images

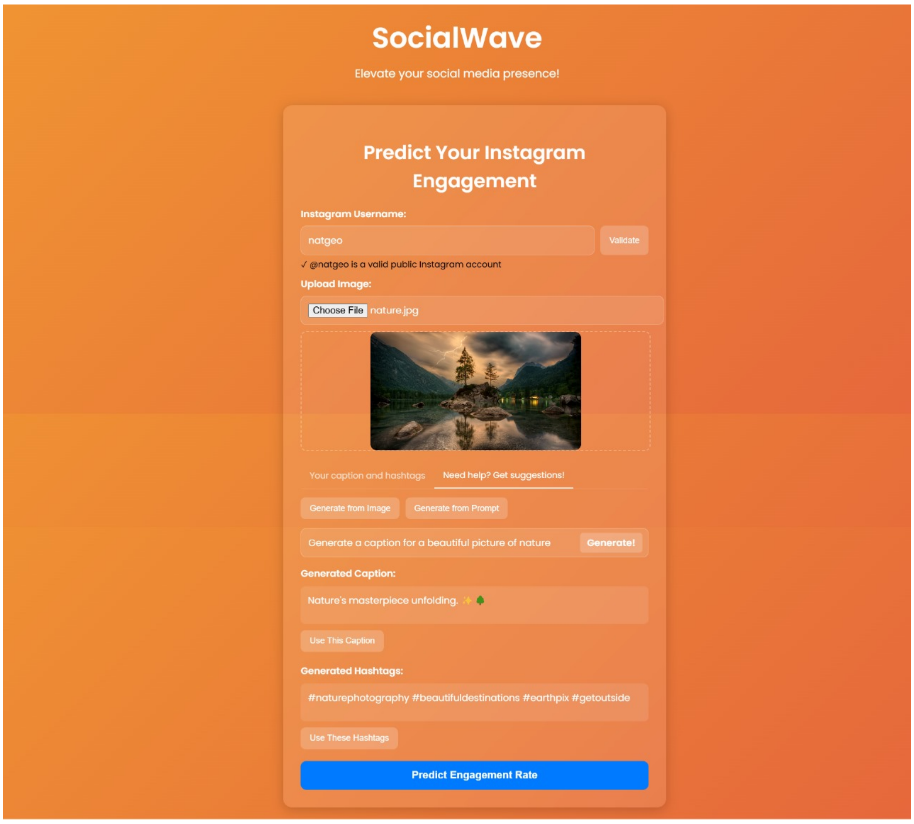


Figure 5.5: Caption and Hashtag Generation from Prompts

advanced natural language processing techniques, allowed for dynamic caption suggestions based on input prompts. This ensured that the generated captions were not only contextually relevant but also highly engaging for social media users.

5.3 Hashtag Recommendation

The models were trained on the HARRISON dataset, a novel benchmark for image hashtag recommendation. HARRISON consists of 57,383 Instagram photos, each annotated with an average of 4.5 hashtags (ranging from a minimum of 1 to a maximum of 10). The ground truth hashtags are selected from the 1,000 most frequently used hashtags, encoded based on their frequency ranking. The training process showed a progressive decrease in loss, indicating good convergence. One of the models, based on the Vision Transformer (ViT) architecture, achieved low loss values during training, with the final training loss decreasing to 0.0207 by the last epoch. However, despite good training metrics, its hashtag generation capabilities were limited. The model struggled with generating contextually relevant hashtags, leading to suboptimal results.

```
100%                                        7180/7180 [34:42<00:00,   6.02it/s]
<ipython-input-24-d93cbefc52e2>:8: FutureWarning: `torch.cuda.amp.GradScaler
  scaler = GradScaler()
/usr/local/lib/python3.10/dist-packages/torch/utils/data/dataloader.py:617:
  warnings.warn(
<ipython-input-24-d93cbefc52e2>:16: FutureWarning: `torch.cuda.amp.autocast(
  with autocast():  # Enable mixed precision
EPOCH 1/20, after current batch, Loss: 0.08918707072734833
EPOCH 2/20, after current batch, Loss: 0.04066097363829613
EPOCH 3/20, after current batch, Loss: 0.031434956938028336
EPOCH 4/20, after current batch, Loss: 0.02861868031322956
EPOCH 5/20, after current batch, Loss: 0.027552084997296333
EPOCH 6/20, after current batch, Loss: 0.027093520388007164
EPOCH 7/20, after current batch, Loss: 0.026879020035266876
EPOCH 8/20, after current batch, Loss: 0.02669500932097435
EPOCH 9/20, after current batch, Loss: 0.02641424722969532
EPOCH 10/20, after current batch, Loss: 0.02602866291999817
EPOCH 11/20, after current batch, Loss: 0.025623049587011337
EPOCH 12/20, after current batch, Loss: 0.02529928833246231
EPOCH 13/20, after current batch, Loss: 0.024686230346560478
EPOCH 14/20, after current batch, Loss: 0.0239713117480278
EPOCH 15/20, after current batch, Loss: 0.023834753781557083
EPOCH 16/20, after current batch, Loss: 0.02310897409915924
EPOCH 17/20, after current batch, Loss: 0.022464916110038757
EPOCH 18/20, after current batch, Loss: 0.02204717881977558
EPOCH 19/20, after current batch, Loss: 0.021241415292024612
EPOCH 20/20, after current batch, Loss: 0.020747294649481773
```

Figure 5.6: ViT Training Loss Reduction Over Epochs

Another model, [9] LXMERT4Hashtag, designed for multimodal learning, was also evaluated. The evaluation results demonstrated high precision and recall, making it a more effective choice for hashtag generation. The key evaluation metrics were a precision of 0.9987, recall of 0.9846, and an F1-score of 0.9916. The batch-wise evaluation during testing further confirmed the robustness of this approach, showing minimal loss values approaching zero by the final batch. The results indicate that the model successfully captures both visual and textual context to generate relevant hashtags.

```
Epoch 4/5, Batch 743/2870, Loss: 0.0038, Time: 0.10 seconds
Epoch 4/5, Batch 744/2870, Loss: 0.0064, Time: 0.10 seconds
Epoch 4/5, Batch 745/2870, Loss: 0.0031, Time: 0.10 seconds
Epoch 4/5, Batch 746/2870, Loss: 0.0024, Time: 0.10 seconds
Epoch 4/5, Batch 747/2870, Loss: 0.0044, Time: 0.10 seconds
Epoch 4/5, Batch 748/2870, Loss: 0.0036, Time: 0.11 seconds
Epoch 4/5, Batch 749/2870, Loss: 0.0075, Time: 0.11 seconds
Epoch 4/5, Batch 750/2870, Loss: 0.0071, Time: 0.10 seconds
Epoch 4/5, Batch 751/2870, Loss: 0.0036, Time: 0.10 seconds
Epoch 4/5, Batch 752/2870, Loss: 0.0015, Time: 0.11 seconds
Epoch 4/5, Batch 753/2870, Loss: 0.0012, Time: 0.11 seconds
Epoch 4/5, Batch 754/2870, Loss: 0.0043, Time: 0.11 seconds
Epoch 4/5, Batch 755/2870, Loss: 0.0025, Time: 0.11 seconds
Epoch 4/5, Batch 756/2870, Loss: 0.0028, Time: 0.11 seconds
Epoch 4/5, Batch 757/2870, Loss: 0.0044, Time: 0.11 seconds
Epoch 4/5, Batch 758/2870, Loss: 0.0025, Time: 0.11 seconds
Epoch 4/5, Batch 759/2870, Loss: 0.0031, Time: 0.10 seconds
Epoch 4/5, Batch 760/2870, Loss: 0.0030, Time: 0.10 seconds
Epoch 4/5, Batch 761/2870, Loss: 0.0021, Time: 0.10 seconds
Epoch 4/5, Batch 762/2870, Loss: 0.0049, Time: 0.11 seconds
Epoch 4/5, Batch 763/2870, Loss: 0.0044, Time: 0.10 seconds
Epoch 4/5, Batch 764/2870, Loss: 0.0022, Time: 0.11 seconds
Epoch 4/5, Batch 765/2870, Loss: 0.0030, Time: 0.10 seconds
Epoch 4/5, Batch 766/2870, Loss: 0.0034, Time: 0.11 seconds
...
Epoch 5/5, Batch 2868/2870, Loss: 0.0012, Time: 0.10 seconds
Epoch 5/5, Batch 2869/2870, Loss: 0.0009, Time: 0.10 seconds
Epoch 5/5, Batch 2870/2870, Loss: 0.0000, Time: 0.09 seconds
Epoch 5/5, Total Loss: 0.0015, Total Time: 298.80 seconds
```

Figure 5.7: Batch-wise Loss During Testing

While certain models demonstrated promising training results, their hashtag generation capabilities varied. The research indicates that single-modality models, like ViT, are not as well-suited for the multimodal nature of hashtag recommendation as they focus primarily on image classification without integrating textual context. On the other

```
Starting model evaluation...
Total batches: 2870
Processing batch 1/2870...
Processing batch 2/2870...
Processing batch 3/2870...
Processing batch 4/2870...
Processing batch 5/2870...
Processing batch 6/2870...
Processing batch 7/2870...
Processing batch 8/2870...
Processing batch 9/2870...
Processing batch 10/2870...
Processing batch 11/2870...
Processing batch 12/2870...
Processing batch 13/2870...
Processing batch 14/2870...
Processing batch 15/2870...
Processing batch 16/2870...
Processing batch 17/2870...
Processing batch 18/2870...
Processing batch 19/2870...
Processing batch 20/2870...
Processing batch 21/2870...
Processing batch 22/2870...
Processing batch 23/2870...
...
Precision: 0.9987
Recall: 0.9846
F1-score: 0.9916
Evaluation completed.
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

Figure 5.8: LXMERT Model Evaluation Results

hand, models like LXMERT, which combine vision and language processing, outperformed purely vision-based models significantly. The high precision, recall, and F1-score indicate that such approaches are highly effective in generating meaningful hashtags based on image content.

Based on the evaluation results, multimodal approaches like LXMERT4Hashtag are the preferred choice for SocialWave's hashtag generation module. By leveraging these advancements, SocialWave can provide high-quality hashtag recommendations, improving social media engagement and reach. (Refer figure 5.4)

Additionally, hashtag recommendations were generated using the Gemini API. By analyzing input prompts, the API suggested relevant and trending hashtags that could enhance post discoverability and engagement. This approach ensured that hashtag selection was data-driven and optimized for maximizing audience reach. (Refer figure 5.5)

5.4 Engagement Rate Prediction

Engagement rate is a key metric in SocialWave, representing user interaction with Instagram posts. It is calculated as:

$$\text{Engagement Rate} = \left(\frac{\text{Likes} + \text{Comments}}{\text{Followers}}\right) \times 100 \tag{5.1}$$

To enhance prediction accuracy, we introduced the **Adjusted Engagement Rate**, which incorporates hashtag quality, calculated using *TF-IDF* and cosine similarity for relevance and popularity scoring:

$$\text{AdjustedER} = \text{Engagement Rate} \times \left(1 + \frac{\text{Hashtag Quality Score}}{100}\right) \tag{5.2}$$

The adjustment factor accounts for the impact of hashtags on engagement by evaluating their popularity and contextual relevance to the post's caption. A higher hashtag quality score increases the predicted engagement rate, reflecting its role in content discoverability. Conversely, low-quality or irrelevant hashtags result in minimal adjustments.

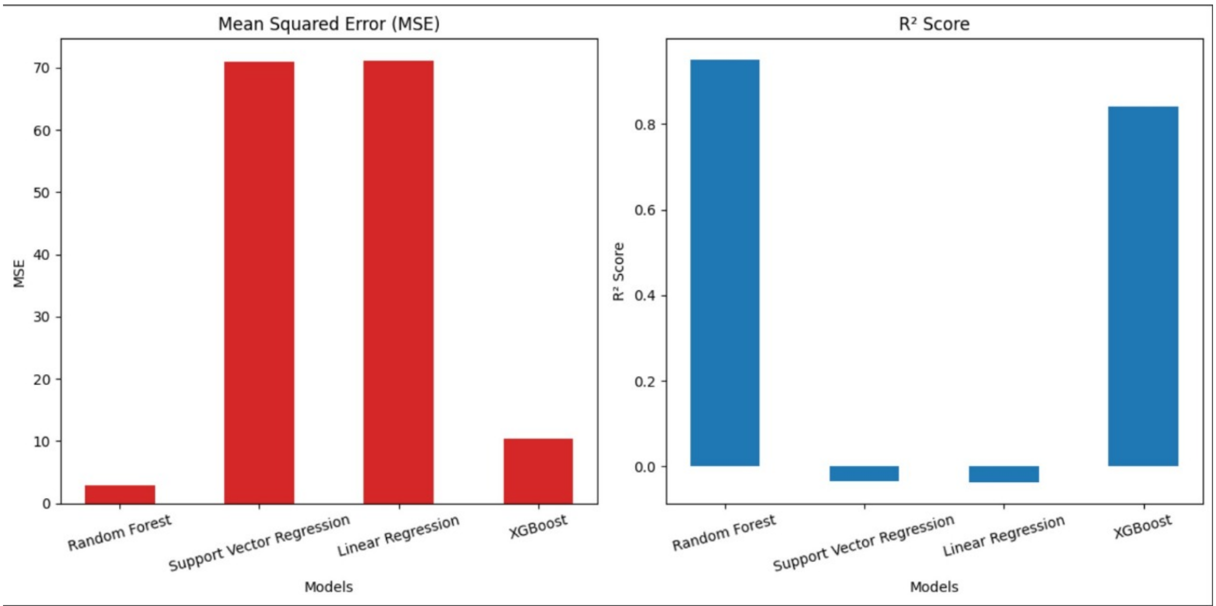
Experiments were conducted with **Random Forest, Support Vector Regression, Linear Regression, and XGBoost** for engagement rate prediction. Among these, **Random Forest** performed the best, achieving the lowest **Mean Squared Error (MSE)** and the highest **R^2 score**, indicating superior predictive accuracy. The decision-tree-based ensemble approach of Random Forest allows it to capture non-linear relationships between engagement factors, leading to better generalization.

By integrating both quantitative and qualitative factors, SocialWave ensures a more context-aware engagement prediction model. This enables content creators to optimize their posts by understanding how different factors influence engagement levels, ultimately improving their social media strategy.

```
sample_post = {
    'likes': 12000,
    'comments': 20,
    'followers': 10000,
    'caption': 'Loving the vibes of this beautiful sunset!',
    'hashtags': '#hiking #market #dogsofinstagram #motivationmonday #petstagram'
}

# Calculate the hashtag quality
hashtag_quality_score = hashtag_quality(sample_post['hashtags'], sample_post['caption'])
print(f"Hashtag Quality: {hashtag_quality_score}")

# Prepare input for prediction
sample_input_df = pd.DataFrame([sample_post], columns=['likes', 'comments', 'followers'])
predicted_engagement_rate = model.predict(sample_input_df)[0]


# Adjust engagement rate using hashtag quality
adjusted_engagement_rate = predicted_engagement_rate * (1 + (hashtag_quality_score / 100))
adjusted_engagement_rate = min(adjusted_engagement_rate, 100)

# Classify the engagement rate
engagement_category = classify_engagement_rate(adjusted_engagement_rate)

print(f"Predicted Engagement Rate (Adjusted): {adjusted_engagement_rate:.2f}%")
print(f"Engagement Category: {engagement_category}")
```

```
Mean Squared Error: 2.92552354916318
R^2 Score: 0.9573262315236633
Hashtag Quality: 58.8
Predicted Engagement Rate (Adjusted): 27.94%
Engagement Category: Viral
```

(a) ER Prediction Outputs



(b) Bar Plot of 4 ER prediction models

Figure 5.9: Engagement Rate Prediction Results

Chapter 6

Future Scope and Conclusion

6.1 Future Scope

SocialWave can extend its functionality beyond Instagram by integrating additional social media platforms such as Facebook, Twitter, LinkedIn, TikTok, and YouTube. This expansion will enable cross-platform content optimization, allowing users to maintain a consistent and effective social media strategy across multiple networks. Additionally, incorporating advanced generative AI models, such as transformer-based architectures, can enhance the quality and personalization of captions. Real-time social media trend analysis can further improve hashtag recommendations, ensuring that content remains relevant and engaging. Future iterations of SocialWave can also include real-time engagement tracking, providing users with instant feedback on post-performance. AI-powered predictive analytics can optimize content strategy by suggesting the best posting times based on audience activity and historical engagement patterns. Furthermore, the integration of improved deep learning models for object detection and scene analysis can extend SocialWave's capabilities to video content. Automated video captioning and thumbnail generation will enhance engagement and accessibility, making the platform more suitable for video-centric social media channels.

6.2 Conclusion

In conclusion, SocialWave is a powerful and innovative tool designed to help content creators maximize their social media engagement through AI-driven insights. By providing engagement predictions, caption and hashtag generation, and actionable recommendations, SocialWave enables users to refine their content strategies and enhance their digital presence. The platform's user-friendly interface and robust backend processing make it accessible and effective for a wide range of users, from influencers and marketers to businesses and casual social media users. As SocialWave continues to grow, expanding its functionality to include real-time trend tracking, cross-platform recommendations for Twitter and TikTok, and advanced competitor analysis will further strengthen its value. Additionally, the development of a mobile application will significantly improve accessibility and usability, ensuring that users can optimize their content strategy on the go. By integrating cutting-edge AI and social media analytics, SocialWave has the potential to become a leading tool in the social media optimization space, helping users stay ahead of trends, increase engagement, and build a stronger online presence.

References

[1] Purba, Kristo Radion, David Asirvatham, and Raja Kumar Murugesan. "Instagram post popularity trend analysis and prediction using hashtag, image assessment, and user history features." Int. Arab J. Inf. Technol. 18.1 (2021): 85-94.

[2] C. Qian, J. Tang, M. Penza, and C. Ferri, "Instagram popularity prediction via neural networks and regression analysis," in *Proceedings of the 2017 Conference*, pp. 2561–2570.

[3] A. Zohourian, H. Sajedi, and A. Yavary, "Popularity prediction of images and videos on Instagram," in *Proceedings of the 4th International Conference on Web Research*, 2018, pp. 111–117.

[4] M. Mazloom, R. Rietveld, S. Rudinac, M. Worring, and W. Dolen, "Multimodal popularity prediction of brand-related social media posts," in *Proceedings of the 24th ACM International Conference on Multimedia*, 2016, pp. 197–201.

[5] E. Zangerle, W. Gassler, and G. Specht, "Recommending #hashtags in Twitter," in *Proceedings of the Workshop on Semantic Adaptive Social Web*, 2011.

[6] Q. Zhang and Z. Huang, "Automatic hashtag recommendation for microblogs using topic-specific translation model," in *Proceedings of the ACM Conference on Recommender Systems*, 2012.

[7] Q. Zhang and Others, "Hashtag recommendation for multimodal microblog using co-attention network," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2017.

[8] D. Kao, K. Lai, and M. Chen, "An efficient and resource-aware hashtag recommendation using deep neural networks," in *Proceedings of the IEEE Conference on Multimedia and Expo*, 2019.

[9] H. Tan and M. Bansal, "LXMERT: Learning cross-modality encoder representations from transformers," in *Proceedings of the Empirical Methods in Natural Language Processing Conference (EMNLP)*, 2020.

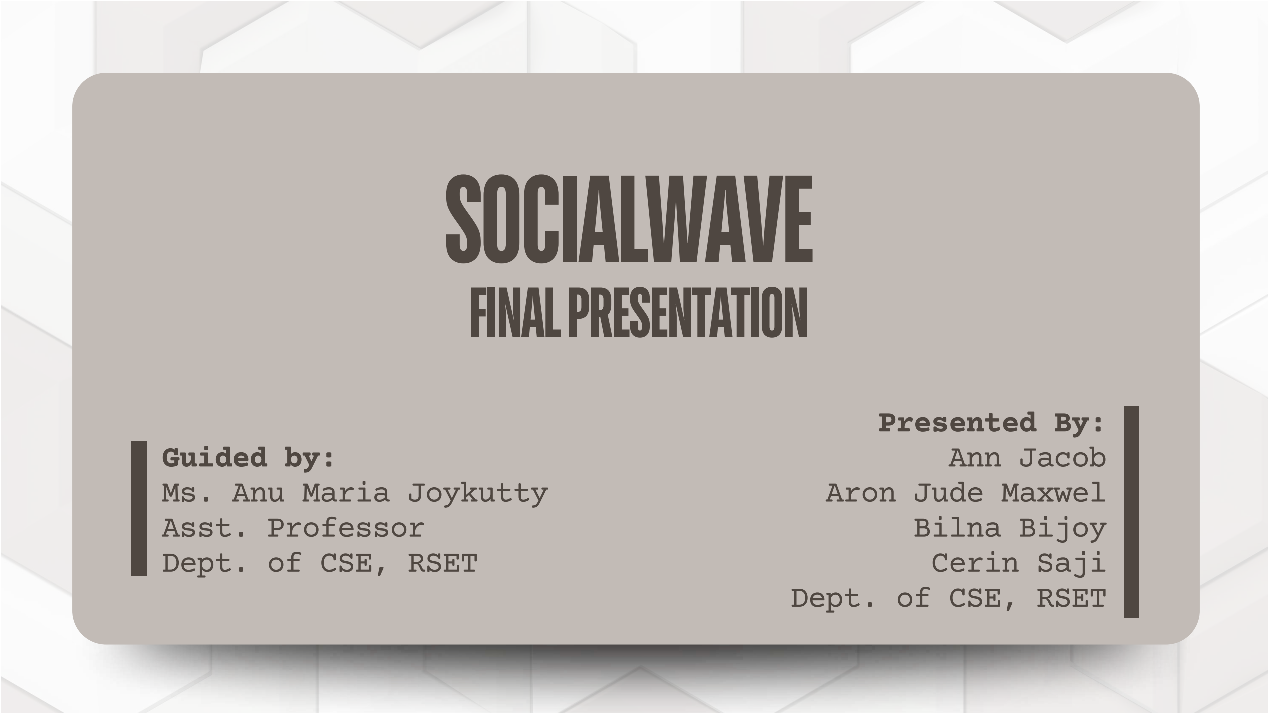
[10] M. M. Y. Khalil, Q. Wang, B. Chen, and W. Wang, "Cross-modality representation learning from transformer for hashtag prediction," *Journal of Big Data*, 2023.

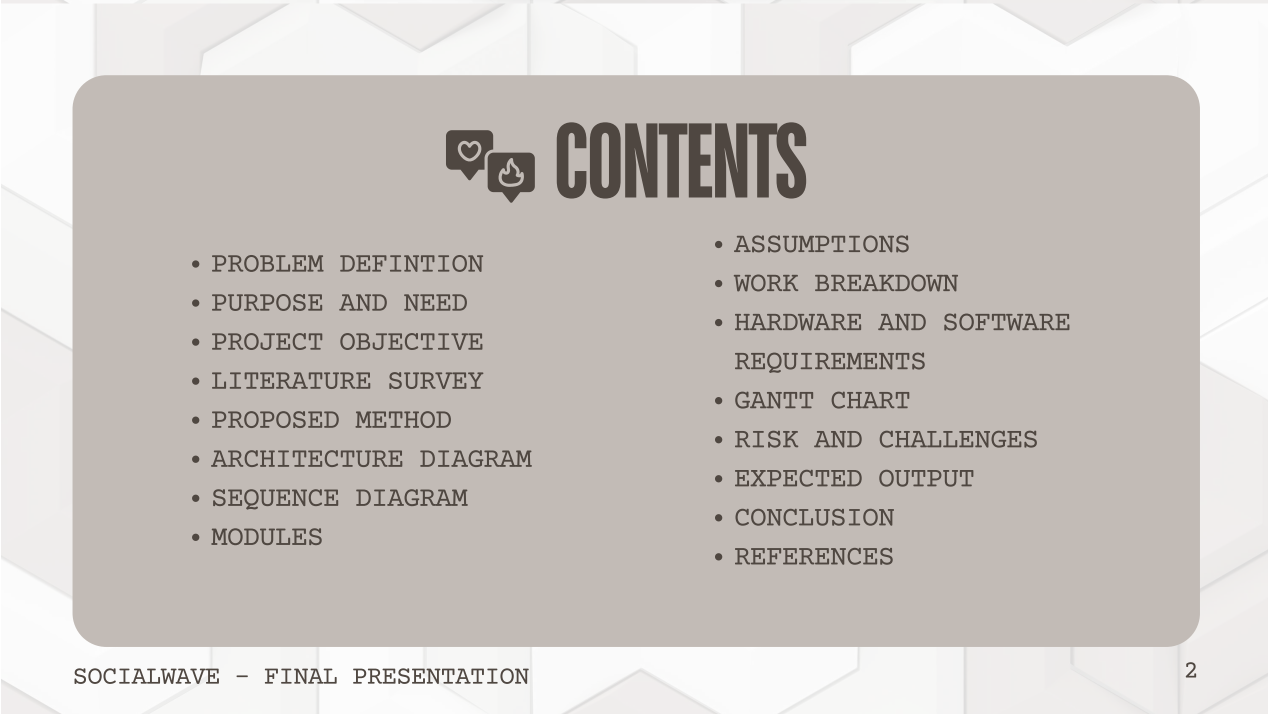
[11] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3156–3164.

[12] Y. Feng, L. Ma, W. Liu, and J. Luo, "Unsupervised image captioning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 4120–4129.

[13] I. Laina, C. Rupprecht, and N. Navab, "Towards unsupervised image captioning with shared multimodal embeddings," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2019.

Appendix A: Final Presentation


SOCIALWAVE
FINAL PRESENTATION
Guided by:
Ms. Anu Maria Joykutty
Asst. Professor
Dept. of CSE, RSET
Presented By:
Ann Jacob
Aron Jude Maxwel
Bilna Bijoy
Cerin Saji
Dept. of CSE, RSET


CONTENTS
• PROBLEM DEFINTION
• PURPOSE AND NEED
• PROJECT OBJECTIVE
• LITERATURE SURVEY
• PROPOSED METHOD
• ARCHITECTURE DIAGRAM
• SEQUENCE DIAGRAM
• MODULES
• ASSUMPTIONS
• WORK BREAKDOWN
• HARDWARE AND SOFTWARE REQUIREMENTS
• GANTT CHART
• RISK AND CHALLENGES
• EXPECTED OUTPUT
• CONCLUSION
• REFERENCES
SOCIALWAVE - FINAL PRESENTATION
2

PROBLEM DEFINTION

New and emerging content creators find it difficult to increase or predict the engagement rates of their posts as the creator space is highly saturated in the current age. They are also challenged when it comes to finding trending hashtags and captions that complement their posts well.

SOCIALWAVE - FINAL PRESENTATION 3

PURPOSE AND NEED

The proposed solution is a tool that enables Instagram content creators to optimize their posts for higher engagement by predicting engagement rates based on images and hashtags, and generating creative captions, helping them increase their reach and post virality.

SOCIALWAVE - FINAL PRESENTATION 4

PROJECT OBJECTIVE

1 **Develop an engagement prediction model:**
Use a machine learning technique (Random Forest) to predict the engagement rate of Instagram posts based on images and hashtags.

2 **Implement image analysis:**
Utilize an object detection algorithm (Faster-RCNN) to extract relevant image features.

3 **Build a caption generation and hashtag recommendation feature:**
Use NLP models to generate creative and contextually relevant captions and recommend hashtags based on uploaded images or user-provided prompts.

SOCIALWAVE - FINAL PRESENTATION 5

PROJECT OBJECTIVE

4 **Create a user-friendly web application:**
Provide an intuitive interface for content creators to upload images, enter hashtags, and receive engagement predictions and caption suggestions in real-time.

5 **Optimize posts for virality:**
Help content creators increase the likelihood of their posts going viral by providing insights into hashtag performance and content quality.

SOCIALWAVE - FINAL PRESENTATION 6

LITERATURE SURVEY

Paper	Description
IAJIT[1] Instagram Post Popularity Trend Analysis and Prediction using Hashtag, Image Assessment and User History Features (2021)	The paper analyzes Instagram post popularity trends using user metadata and image assessment, identifying key factors affecting engagement
Srivastava et al.[2] Comparative analysis of deep learning image detection algorithms (2021)	The paper analyzes YOLO's architecture and performance in object detection, comparing its efficiency and accuracy with other detection models.

SOCIALWAVE - FINAL PRESENTATION

7

LITERATURE SURVEY

Paper	Description
Khalil et al.[3] Cross-modality representation learning from transformer for hashtag prediction (2023)	This paper explores LXMERT for developing LXMERT4Hashtag, a cross-modality transformer model that predicts hashtags by learning from images and captions, improving post visibility and engagement.
Peipei Zhu.[4] Prompt-Based Learning for Unpaired Image Captioning (2024)	The paper presents a prompt-based learning method for unpaired image captioning using vison-language pre-trained model (VL-PTM).

SOCIALWAVE - FINAL PRESENTATION

8

PROPOSED METHOD

Random Forest for ER Prediction:

1. Collect a dataset of Instagram posts, including features such as generated hashtags, captions, and engagement metrics (likes, comments, followers).

2.Preprocess the data by cleaning, normalizing, and encoding categorical variables (e.g., hashtags, captions) into numerical formats suitable for analysis.

3.Split the dataset into training and testing sets to evaluate model performance.

4.Initialize the Random Forest algorithm with specified hyperparameters (number of trees, maximum depth, etc.).

SOCIALWAVE - FINAL PRESENTATION 9

PROPOSED METHOD

Random Forest for ER Prediction:

5.Fit the Random Forest model on the training dataset using features (hashtags, captions) to predict engagement rates.

6.Use the trained model to make predictions on the testing dataset and compare predicted engagement rates with actual values.

7.Evaluate model performance using metrics such as Mean Squared Error (MSE) and R-square values.

8.Perform feature importance analysis to identify which features (hashtags, captions) most influence engagement rates.

SOCIALWAVE - FINAL PRESENTATION 10

PROPOSED METHOD

Random Forest for ER Prediction:

9.Optionally tune hyperparameters and retrain the model to improve accuracy based on evaluation results.

10.Use the final model for predicting engagement rates of new Instagram posts based on their generated hashtags and captions.

SOCIALWAVE - FINAL PRESENTATION 11

PROPOSED METHOD

Generating Instagram Captions with Fine-Tuned Microsoft-git-base

- **Install Necessary Libraries:** Set up your environment by installing essential libraries such as Transformers, Accelerate, and Dataset
- **Load the Dataset:** Utilize a dataset containing Instagram post captions and corresponding images from the Hugging Face Hub.
- **Preprocess the Data:** Implement a custom dataset class to preprocess images and captions, including resizing images and tokenizing captions.
- **Create DataLoader:** Set up a DataLoader to facilitate batch processing during model training.

SOCIALWAVE - FINAL PRESENTATION 12

PROPOSED METHOD

- **Load Pre-trained GIT Model:** Initialize the pre-trained Microsoft GIT model for causal language modeling.
- **Fine-Tune the Model:** Train the model on your dataset, optimizing it to generate captions that align with Instagram's style and your specific content.
- **Generate Captions:** After fine-tuning, input images into the model to produce engaging and contextually relevant Instagram captions.

SOCIALWAVE - FINAL PRESENTATION 13

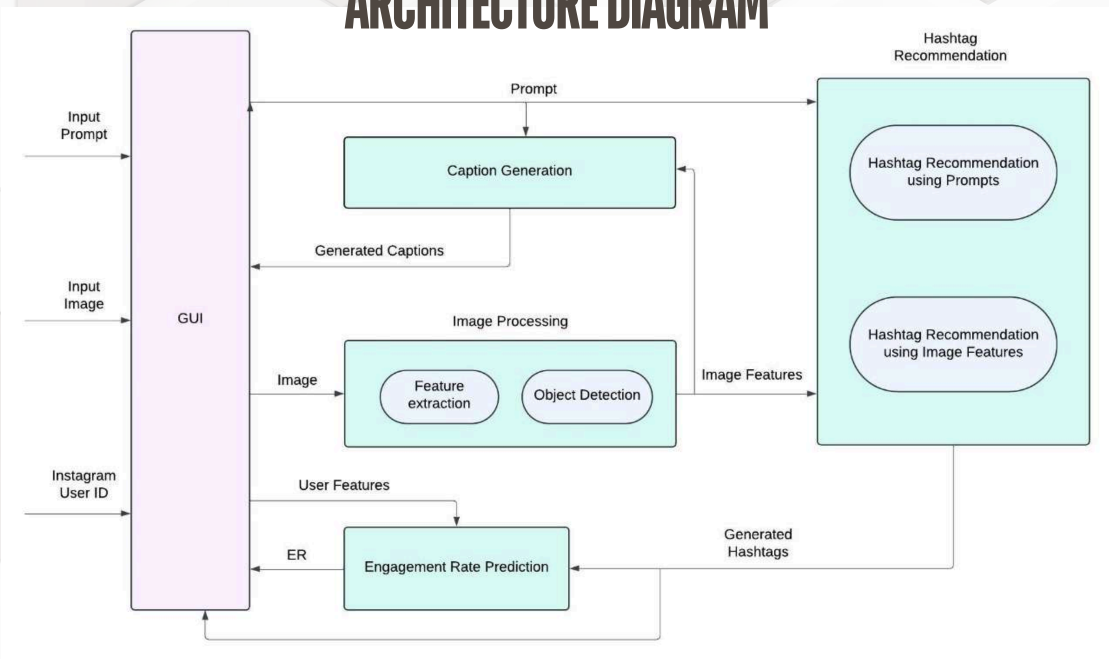
PROPOSED METHOD

LXMERT4Hashtag for Hashtag Recommendation

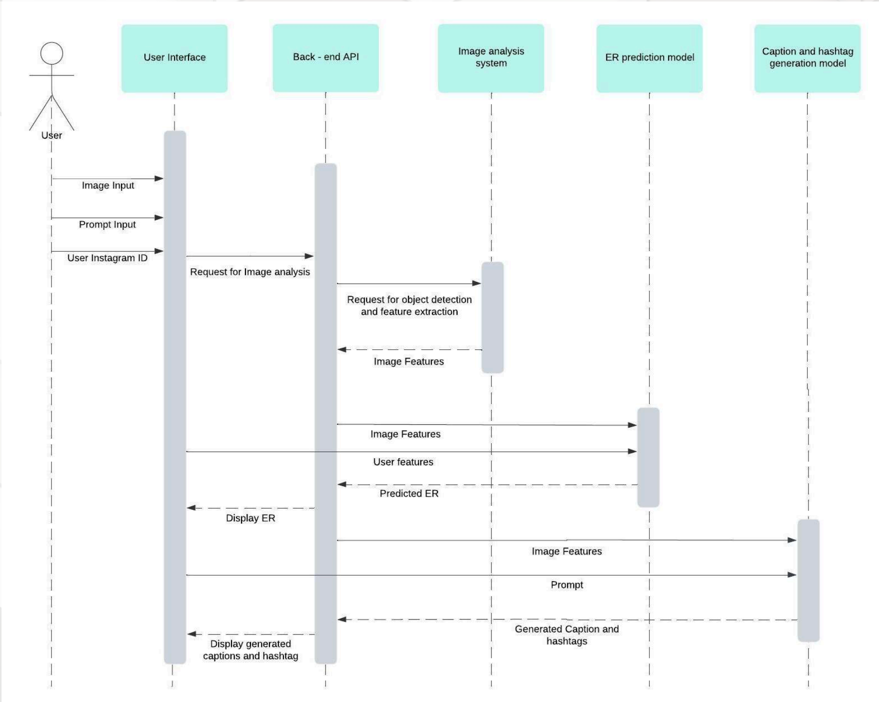
Algorithm 1 LXMERT for Cross-Modal Hashtag Prediction

```
procedure LXMERT4HASHTAG(captions, images, hashtags, max_sequence_length, batch_size, num_epochs)
    Initialize LXMERT model with a binary classification head
      • Load the pretrained LXMERT model, which includes the Text Encoder, Object-Relationship Encoder, and Cross-Modality Encoder.
      • Add a binary classification head.
    Tokenize captions and prepare images segments
    Split data into training and validation sets
    for epoch ← 1 to num_epochs do
        for batch ← 1 to (number_batches) do
            Perform Forward pass-through model to obtain logit
            Compute Binary Cross-Entropy Loss between logits and targets
            Backpropagate gradients and update model parameters
        end for
        Perform validation on the validation set and adjust hyperparameters.
    end for
    Initialize predicted_tags list
    for tweet in tweets do
        Tokenize caption and prepare image segments
        Perform Forward pass-through model to obtain predictions
        Threshold predictions to obtain predicted hashtags
        Append predicted hashtags to predicted_hashtags list
    end for
    return predicted_hashtags
end procedure
```

ARCHITECTURE DIAGRAM



SEQUENCE DIAGRAM



MODULES

1. IMAGE ANALYSIS MODULE

- Object Detection for Post Insights: The model detects and classifies objects in Instagram images, identifying key elements like products, people, or backgrounds to enhance content understanding.
- Content Relevance Evaluation: It extracts visual features from the post, aiding in analyzing how well the image aligns with the hashtags, captions, and audience engagement.

MODULES

2. HASHTAG GENERATION MODULE

- Hashtag Suggestion from Image Content: The module analyzes the visual content of images and automatically generates relevant hashtags that align with the image's subject, theme, and key features to boost visibility.
- Custom Hashtag Generation from Prompts: Users can input specific prompts, and the module generates tailored hashtags based on the given context or theme, enhancing engagement for targeted audiences.

MODULES

3. CAPTION GENERATION MODULE

- Automated Caption Creation from Images: The module analyzes visual elements within an image and generates relevant, engaging captions that reflect the content, context, and mood of the image.
- Personalized Caption Generation from Prompts: Based on user-provided prompts, the module crafts customized captions that align with the desired tone, theme, or message, allowing for more targeted and creative content.

SOCIALWAVE - FINAL PRESENTATION 19

MODULES

4. ENGAGEMENT RATE PREDICTION

- Hashtag and Caption Analysis: Predicts engagement using Random Forest by evaluating the effectiveness of hashtags and relevance of captions.
- User Feature Integration: Considers user-specific factors like follower count, posting frequency, and past engagement to enhance prediction accuracy.

SOCIALWAVE - FINAL PRESENTATION 20

ASSUMPTIONS

1	The Instagram accounts for which engagement rates are being predicted are public and allow access to necessary post and user data (e.g., post likes, comments, hashtags, and user followers).
2	Images provided to the system must be of good quality which includes a minimum resolution of 1080x1080 pixels for detailed object detection, free from blurring and excessive noise for clear identification.

HARDWARE REQUIREMENTS

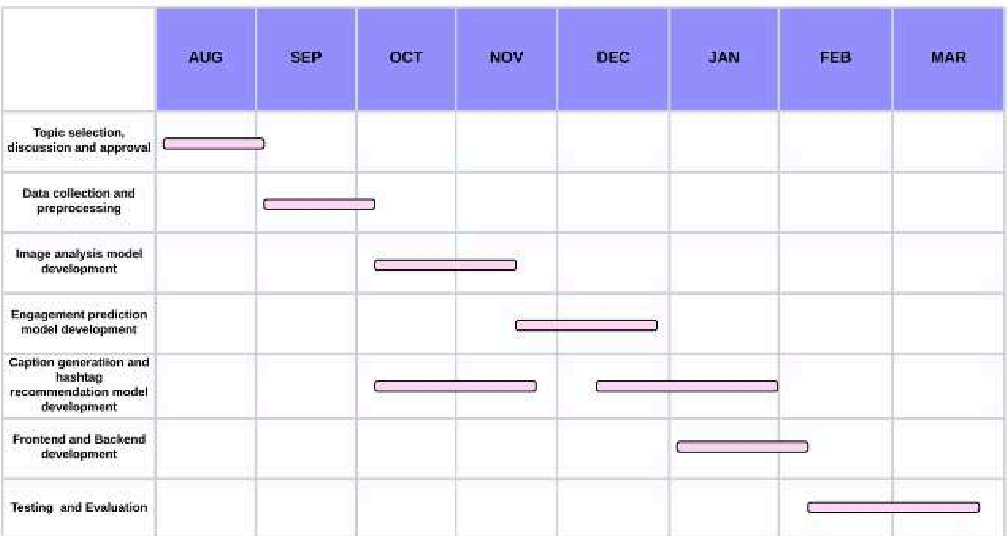
- Processor: A modern multi-core processor (e.g., Intel Core i5) for efficient code compilation and development tasks.
- RAM : A minimum 8GB of RAM to handle multiple development tools and applications simultaneously without slowdowns. However, 16 would be recommended.
- Storage: A solid-state drive (SSD) with sufficient storage capacity (e.g., 256GB or higher) for storing project files, libraries, and development tools.

SOFTWARE REQUIREMENTS

- Python, Flask, HTML5, CSS3 and JavaScript and frameworks like React for front-end development.
- Libraries (TensorFlow and PyTorch) for training deep learning models (LXMERT4Hashtag, Random Forest).
- Pre-trained model (Faster RCNN) for object detection and feature extraction tasks.

SOCIALWAVE - FINAL PRESENTATION 23

GANTT CHART



24

RISKS AND CHALLENGES

Handling and processing user-uploaded content (images, captions, hashtags) raises concerns about intellectual property rights and data privacy.	Virality can be influenced by many unpredictable factors, such as timing, external events, or platform algorithm changes, making it difficult to predict accurately.
Understanding hashtag meaning and context can be challenging, as hashtags may contain slang, abbreviations, or ambiguous meanings.	Users expect quick feedback when optimizing their social media posts, so delays could lead to frustration.

SOCIALWAVE - FINAL PRESENTATION 25

EXPECTED OUTPUT

1. Accurate Engagement Prediction

- Develop a machine learning model for predicting engagement rates of Instagram posts.
- Use likes,comments,images and hashtags as input to make predictions.
- Help content creators improve their posts to increase likes, comments, and shares.
- Provide actionable insights for optimizing post performance.

SOCIALWAVE - FINAL PRESENTATION 26

EXPECTED OUTPUT

2. Effective Image Caption Generation

- Implement an NLP-based caption generator for Instagram posts.
- Create contextually relevant, creative, and engaging captions.
- Generate captions based on:
  -Uploaded images.
  -User-provided prompts or text.
- Enhance content creation by providing high-quality, tailored captions.

EXPECTED OUTPUT

3. User-Friendly Web Application

- Build a fully functional, interactive web application.
- Key features:
  -Upload images and input hashtags.
  -Get real-time engagement rate predictions.
  -Generate captions instantly.
  -Ensure a seamless, intuitive, and easy-to-navigate UI.
- Enable smooth interaction with the machine learning models and caption generation system.

WORK BREAKDOWN

1 **ANN JACOB**
Caption Generation Module

2 **BILNA BIJOY**
Object Detection module, ER prediction module (SVR, XG Boost)

3 **CERIN SAJI**
Hashtag Recommendation using ViT Classifier and LXMERT4Hashtag on HARRISON, Landing page UI, Front-end + Backend integration

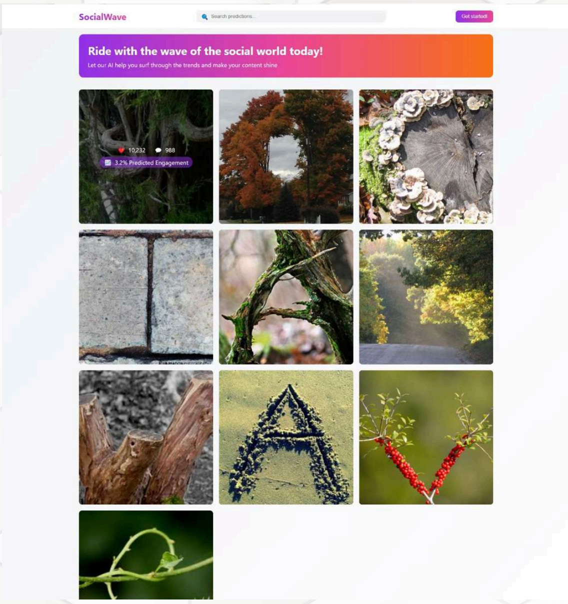
4 **ARON JUDE MAXWEL**
Web scraping images from Instagram for dataset, ER prediction module (Random forest, Linear regression), UI - Login page, signup page, main page

PAPER SUBMISSION

- Submitted paper to ASAR - International Conference on Advanced Research in Computer Science and Information Technology (ICARCSIT), New Delhi, India, April 15th, 2025
- 4th 2025 IEEE World Conference on Applied Intelligence and Computing: Submission (726) has been created.

OUTPUTS

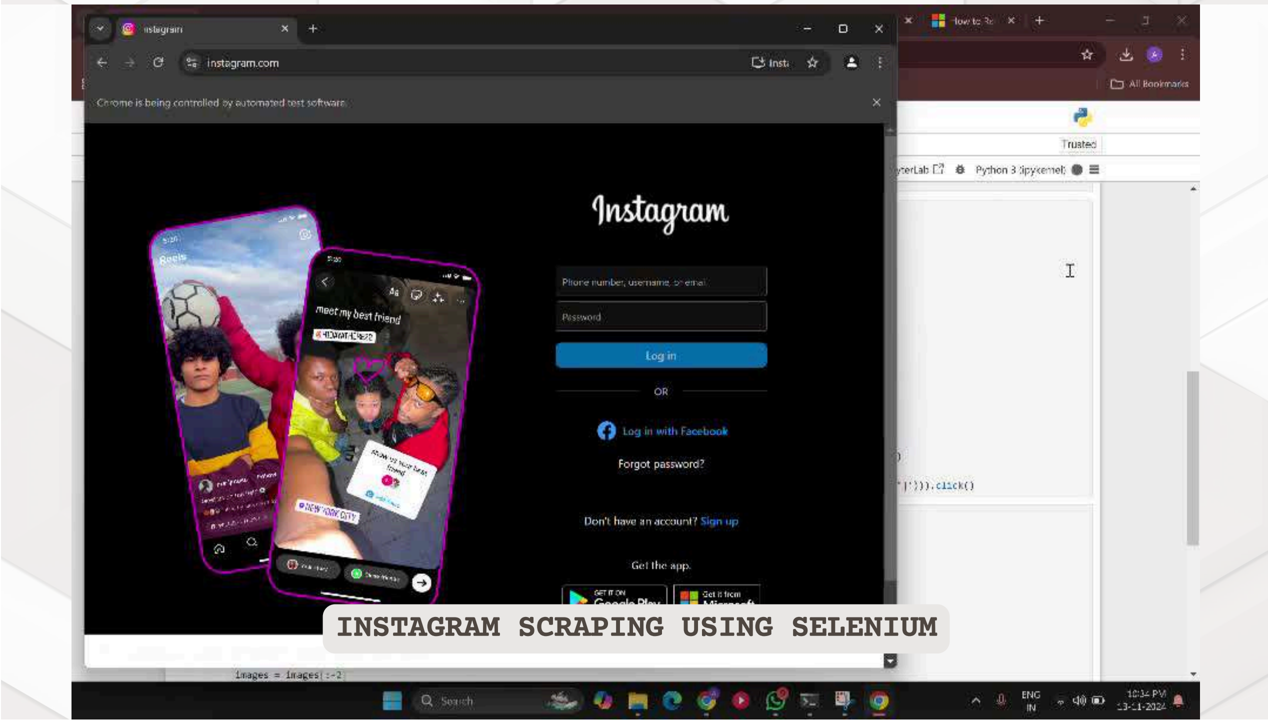
SOCIALWAVE - FINAL PRESENTATION 31

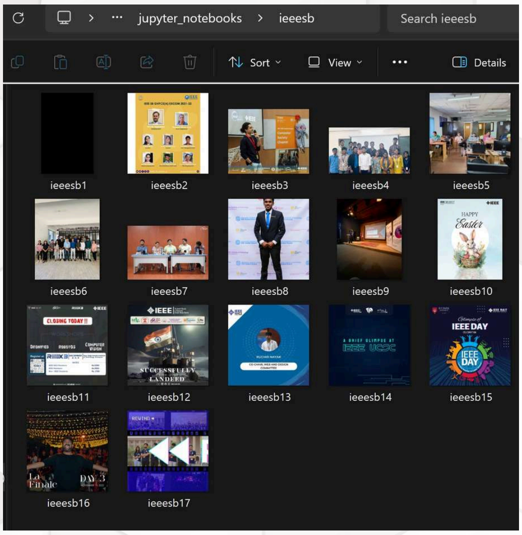


Landing page

SOCIALWAVE - FINAL PRESENTATION 32

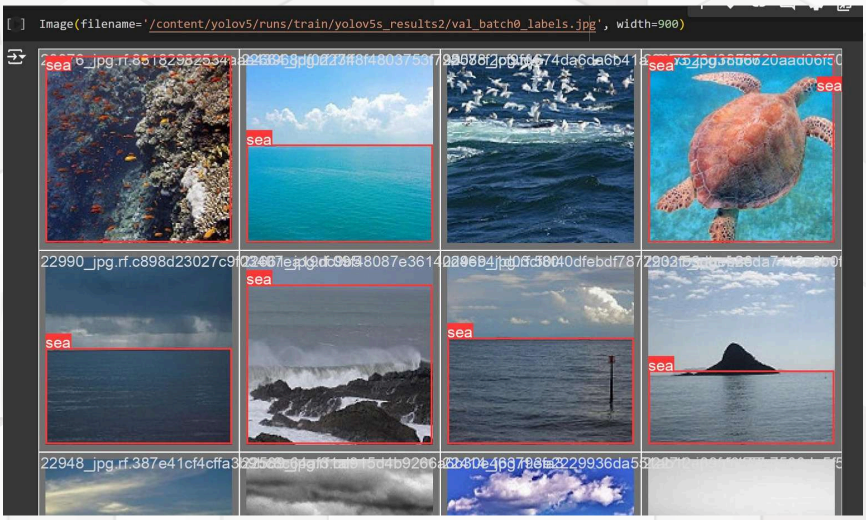

SocialWave
Login
Email
Enter your email
Password
Enter your password
Login
Don't have an account? Sign up
SocialWave
Sign Up
Full Name
Enter your full name
Email
Enter your email
Password
Enter your password
Confirm Password
Confirm your password
Sign Up
Already have an account? Login
Login and Signup Pages
SOCIALWAVE - FINAL PRESENTATION
33


instagram.com
Chrome is being controlled by automated test software.
Instagram
Phone number, username, or email
Password
Log in
OR
Log in with Facebook
Forgot password?
Don't have an account? Sign up
Get the app.
INSTAGRAM SCRAPING USING SELENIUM



Images obtained after scraping posts under the hashtag 'ieeesb'

Object detection using YOLOv5



Captions generation from Images

Caption generation from prompts

Hashtag Generation – Training using ViT Classifier

```
EPOCH 1/5, after current batch, Loss: 25.118423461914062
EPOCH 2/5, after current batch, Loss: 23.931997299194336
EPOCH 3/5, after current batch, Loss: 22.627573013305664
EPOCH 4/5, after current batch, Loss: 21.509244918823242
EPOCH 5/5, after current batch, Loss: 21.1131591796875
```

1st Test Training – Extremely High Loss

High loss obtained when trained considering Cross Entropy Loss

```
EPOCH 1/50, after current batch, Loss: 28.473480224609375
EPOCH 2/50, after current batch, Loss: 27.300024032592773
EPOCH 3/50, after current batch, Loss: 25.47911262512207
EPOCH 4/50, after current batch, Loss: 23.658483505249023
EPOCH 5/50, after current batch, Loss: 22.434818267822266
EPOCH 6/50, after current batch, Loss: 21.172727584838867
EPOCH 7/50, after current batch, Loss: 19.30763053894043
EPOCH 8/50, after current batch, Loss: 17.972003936767578
EPOCH 9/50, after current batch, Loss: 16.54920768737793
EPOCH 10/50, after current batch, Loss: 15.373448371887207
EPOCH 11/50, after current batch, Loss: 13.999801635742188
EPOCH 12/50, after current batch, Loss: 12.919532775878906
EPOCH 13/50, after current batch, Loss: 11.914688110351562
EPOCH 14/50, after current batch, Loss: 10.944092750549316
EPOCH 15/50, after current batch, Loss: 10.145915031433105
EPOCH 16/50, after current batch, Loss: 9.60278606414795
EPOCH 17/50, after current batch, Loss: 9.2628173828125
EPOCH 18/50, after current batch, Loss: 9.004796028137207
EPOCH 19/50, after current batch, Loss: 8.47618293762207
EPOCH 20/50, after current batch, Loss: 8.14021110534668
EPOCH 21/50, after current batch, Loss: 8.129434585571289
EPOCH 22/50, after current batch, Loss: 8.194571495056152
EPOCH 23/50, after current batch, Loss: 8.042525291442871
EPOCH 24/50, after current batch, Loss: 7.989798069000244
EPOCH 25/50, after current batch, Loss: 7.763183116912842
...
EPOCH 47/50, after current batch, Loss: 7.081780910491943
EPOCH 48/50, after current batch, Loss: 7.079284191131592
EPOCH 49/50, after current batch, Loss: 7.078474521636963
EPOCH 50/50, after current batch, Loss: 7.078038215637207
```

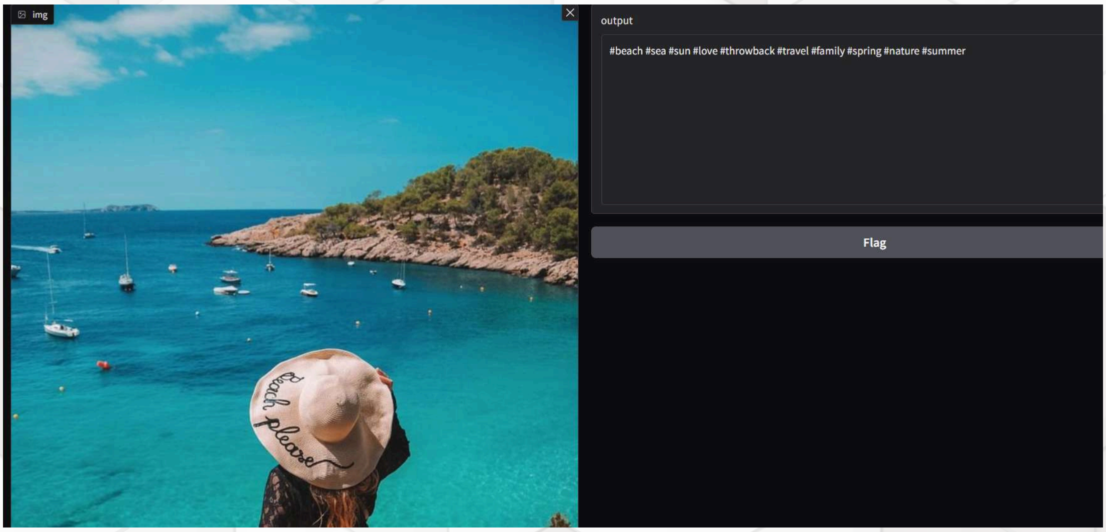
Hashtag Generation - Training using ViT Classifier

```
100% 7180/7180 [34:42<00:00,   6.02it/s]
<ipython-input-24-d93cbefc52e2>:8: FutureWarning: `torch.cuda.amp.GradScaler
  scaler = GradScaler()
/usr/local/lib/python3.10/dist-packages/torch/utils/data/dataloader.py:617:
  warnings.warn(
<ipython-input-24-d93cbefc52e2>:16: FutureWarning: `torch.cuda.amp.autocast(
  with autocast():  # Enable mixed precision
EPOCH 1/20, after current batch, Loss: 0.08918707072734833
EPOCH 2/20, after current batch, Loss: 0.04066097363829613
EPOCH 3/20, after current batch, Loss: 0.031434956938028336
EPOCH 4/20, after current batch, Loss: 0.02861868031322956
EPOCH 5/20, after current batch, Loss: 0.027552084997296333
EPOCH 6/20, after current batch, Loss: 0.027093520388007164
EPOCH 7/20, after current batch, Loss: 0.026879020035266876
EPOCH 8/20, after current batch, Loss: 0.02669500932097435
EPOCH 9/20, after current batch, Loss: 0.02641424722969532
EPOCH 10/20, after current batch, Loss: 0.02602866291999817
EPOCH 11/20, after current batch, Loss: 0.025623049587011337
EPOCH 12/20, after current batch, Loss: 0.02529928833246231
EPOCH 13/20, after current batch, Loss: 0.024686230346560478
EPOCH 14/20, after current batch, Loss: 0.0239713117480278
EPOCH 15/20, after current batch, Loss: 0.023834753781557083
EPOCH 16/20, after current batch, Loss: 0.02310897409915924
EPOCH 17/20, after current batch, Loss: 0.022464916110038757
EPOCH 18/20, after current batch, Loss: 0.02204717881977558
EPOCH 19/20, after current batch, Loss: 0.021241415292024612
EPOCH 20/20, after current batch, Loss: 0.020747294649481773
```

The loss significantly decreased when trained using BCE with Logits Loss since this is a multi-class classification problem

39

Hashtag Generation - Training using ViT Classifier



Hashtag Generation – fine-tuning LXMERT (LXMERT4Hashtag)

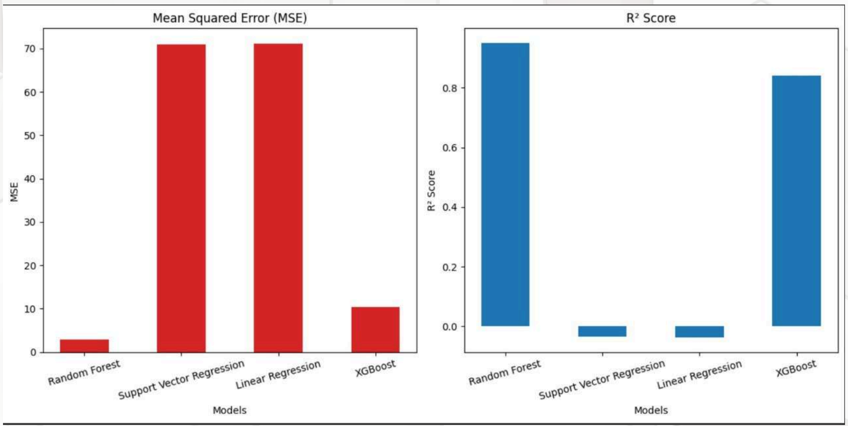
```
Epoch 4/5, Batch 743/2870, Loss: 0.0038, Time: 0.10 seconds
Epoch 4/5, Batch 744/2870, Loss: 0.0064, Time: 0.10 seconds
Epoch 4/5, Batch 745/2870, Loss: 0.0031, Time: 0.10 seconds
Epoch 4/5, Batch 746/2870, Loss: 0.0024, Time: 0.10 seconds
Epoch 4/5, Batch 747/2870, Loss: 0.0044, Time: 0.10 seconds
Epoch 4/5, Batch 748/2870, Loss: 0.0036, Time: 0.11 seconds
Epoch 4/5, Batch 749/2870, Loss: 0.0075, Time: 0.11 seconds
Epoch 4/5, Batch 750/2870, Loss: 0.0071, Time: 0.10 seconds
Epoch 4/5, Batch 751/2870, Loss: 0.0036, Time: 0.10 seconds
Epoch 4/5, Batch 752/2870, Loss: 0.0015, Time: 0.11 seconds
Epoch 4/5, Batch 753/2870, Loss: 0.0012, Time: 0.11 seconds
Epoch 4/5, Batch 754/2870, Loss: 0.0043, Time: 0.11 seconds
Epoch 4/5, Batch 755/2870, Loss: 0.0025, Time: 0.11 seconds
Epoch 4/5, Batch 756/2870, Loss: 0.0028, Time: 0.11 seconds
Epoch 4/5, Batch 757/2870, Loss: 0.0044, Time: 0.11 seconds
Epoch 4/5, Batch 758/2870, Loss: 0.0025, Time: 0.11 seconds
Epoch 4/5, Batch 759/2870, Loss: 0.0031, Time: 0.10 seconds
Epoch 4/5, Batch 760/2870, Loss: 0.0030, Time: 0.10 seconds
Epoch 4/5, Batch 761/2870, Loss: 0.0021, Time: 0.10 seconds
Epoch 4/5, Batch 762/2870, Loss: 0.0049, Time: 0.11 seconds
Epoch 4/5, Batch 763/2870, Loss: 0.0044, Time: 0.10 seconds
Epoch 4/5, Batch 764/2870, Loss: 0.0022, Time: 0.11 seconds
Epoch 4/5, Batch 765/2870, Loss: 0.0030, Time: 0.10 seconds
Epoch 4/5, Batch 766/2870, Loss: 0.0034, Time: 0.11 seconds
...
Epoch 5/5, Batch 2868/2870, Loss: 0.0012, Time: 0.10 seconds
Epoch 5/5, Batch 2869/2870, Loss: 0.0009, Time: 0.10 seconds
Epoch 5/5, Batch 2870/2870, Loss: 0.0000, Time: 0.09 seconds
Epoch 5/5, Total Loss: 0.0015, Total Time: 298.80 seconds
```

```
Starting model evaluation...
Total batches: 2870
Processing batch 1/2870...
Processing batch 2/2870...
Processing batch 3/2870...
Processing batch 4/2870...
Processing batch 5/2870...
Processing batch 6/2870...
Processing batch 7/2870...
Processing batch 8/2870...
Processing batch 9/2870...
Processing batch 10/2870...
Processing batch 11/2870...
Processing batch 12/2870...
Processing batch 13/2870...
Processing batch 14/2870...
Processing batch 15/2870...
Processing batch 16/2870...
Processing batch 17/2870...
Processing batch 18/2870...
Processing batch 19/2870...
Processing batch 20/2870...
Processing batch 21/2870...
Processing batch 22/2870...
Processing batch 23/2870...
...
Precision: 0.9987
Recall: 0.9846
F1-score: 0.9916
Evaluation completed.
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

ER prediction based on user-history features, hashtag quality, input / generated caption

```python
# Calculate the hashtag quality
hashtag_quality_score = hashtag_quality(sample_post['hashtags'], sample_post['caption'])
print(f"Hashtag Quality: {hashtag_quality_score}")

# Prepare input for prediction
sample_input_df = pd.DataFrame([sample_post], columns=['likes', 'comments', 'followers'])
predicted_engagement_rate = model.predict(sample_input_df)[0]


# Adjust engagement rate using hashtag quality
adjusted_engagement_rate = predicted_engagement_rate * (1 + (hashtag_quality_score / 100))
adjusted_engagement_rate = min(adjusted_engagement_rate, 100)

# Classify the engagement rate
engagement_category = classify_engagement_rate(adjusted_engagement_rate)

print(f"Predicted Engagement Rate (Adjusted): {adjusted_engagement_rate:.2f}%")
print(f"Engagement Category: {engagement_category}")
```

```
Mean Squared Error: 2.92552354916318
R^2 Score: 0.9573262315236633
Hashtag Quality: 58.8
Predicted Engagement Rate (Adjusted): 27.94%
```



Bar Plot of ER Models

CONCLUSION

SocialWave is on track to develop an innovative tool aimed at helping budding content optimize their posts for higher engagement and virality. Our initial progress in data collection, algorithm selection, and feature implementation sets a solid foundation for the next phases of development. As we move forward, we will focus on integrating these components into a user-friendly application and refining our models to ensure accurate and impactful results for content creators. We are confident that SocialWave will offer valuable insights and practical tools to enhance social media strategy and success.

REFERENCES

1. Purba, Kristo Radion, David Asirvatham, and Raja Kumar Murugesan. "Instagram post popularity trend analysis and prediction using hashtag, image assessment, and user history features." Int. Arab J. Inf. Technol. 18.1 (2021): 85-94.
2. Silva, Singhe, and Rasika Rajapaksha. "Predictive Analysis on Social Media Content to Become Viral." , International Conference on Advances in Technology and Computing (ICATC 2023)
3. Srivastava, S., Divekar, A.V., Anilkumar, C. et al. Comparative analysis of deep learning image detection algorithms. J Big Data 8, 66 (2021).
4. Carta, Salvatore, et al. “Popularity prediction of instagram posts." Information 11.9 (2020): 453.
5. Maulud, Dastan, and Adnan M. Abdulazeez. “A review on linear regression comprehensive in machine learning." Journal of Applied Science and Technology Trends 1.2 (2020): 140-147.

THANK YOU!

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes

Vision, Mission, Programme Outcomes and Course Outcomes

Institute Vision

To evolve into a premier technological institution, moulding eminent professionals with creative minds, innovative ideas and sound practical skill, and to shape a future where technology works for the enrichment of mankind.

Institute Mission

To impart state-of-the-art knowledge to individuals in various technological disciplines and to inculcate in them a high degree of social consciousness and human values, thereby enabling them to face the challenges of life with courage and conviction.

Department Vision

To become a centre of excellence in Computer Science and Engineering, moulding professionals catering to the research and professional needs of national and international organizations.

Department Mission

To inspire and nurture students, with up-to-date knowledge in Computer Science and Engineering, ethics, team spirit, leadership abilities, innovation and creativity to come out with solutions meeting societal needs.

Programme Outcomes (PO)

Engineering Graduates will be able to:

1. Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

2. Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

4. Conduct investigations of complex problems: Use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. Modern Tool Usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

6. The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. Individual and Team work: Function effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings.

10. Communication: Communicate effectively with the engineering community and with society at large. Be able to comprehend and write effective reports documentation. Make effective presentations, and give and receive clear instructions.

11. Project management and finance: Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team. Manage projects in multidisciplinary environments.

12. Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change.

Programme Specific Outcomes (PSO)

A graduate of the Computer Science and Engineering Program will demonstrate:

PSO1: Computer Science Specific Skills
The ability to identify, analyze and design solutions for complex engineering problems in multidisciplinary areas by understanding the core principles and concepts of computer science and thereby engage in national grand challenges.

PSO2: Programming and Software Development Skills
The ability to acquire programming efficiency by designing algorithms and applying standard practices in software project development to deliver quality software products meeting the demands of the industry.

PSO3: Professional Skills
The ability to apply the fundamentals of computer science in competitive research and to develop innovative products to meet the societal needs thereby evolving as an eminent researcher and entrepreneur.

Course Outcomes (CO)

Course Outcome 1: Model and solve real world problems by applying knowledge across domains (Cognitive knowledge level: Apply).

Course Outcome 2: Develop products, processes or technologies for sustainable and socially relevant applications (Cognitive knowledge level: Apply).

Course Outcome 3: Function effectively as an individual and as a leader in diverse teams and to comprehend and execute designated tasks (Cognitive knowledge level: Apply).

Course Outcome 4: Plan and execute tasks utilizing available resources within timelines, following ethical and professional norms (Cognitive knowledge level: Apply).

Course Outcome 5: Identify technology/research gaps and propose innovative/creative solutions (Cognitive knowledge level: Analyze).

Course Outcome 6: Organize and communicate technical and scientific findings effectively in written and oral forms (Cognitive knowledge level: Apply).

Appendix C: CO-PO-PSO Mapping

COURSE OUTCOMES:

After completion of the course the student will be able to

SL.NO	DESCRIPTION	Blooms' Taxonomy Level
CO1	Model and solve real world problems by applying knowledge across domains (Cognitive knowledge level:Apply).	Level 3: Apply
CO2	Develop products, processes or technologies for sustainable and socially relevant applications. (Cognitive knowledge level:Apply).	Level 3: Apply
CO3	Function effectively as an individual and as a leader in diverse teams and to comprehend and execute designated tasks. (Cognitive knowledge level:Apply).	Level 3: Apply
CO4	Plan and execute tasks utilizing available resources within timelines, following ethical and professional norms (Cognitive knowledge level: Apply).	Level 3: Apply
CO5	Identify technology/research gaps and propose innovative/creative solutions (Cognitive knowledge level:Analyze).	Level 4: Analyze
CO6	Organize and communicate technical and scientific findings effectively in written and oral forms (Cognitive knowledge level:Apply).	Level 3: Apply

CO-PO AND CO-PSO MAPPING

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO 1	PSO 2	PSO 3
CO1	2	2	2	1	2	2	2	1	1	1	1	2	3		
CO2	2	2	2		1	3	3	1	1		1	1		2	
CO3									3	2	2	1			3
CO4					2			3	2	2	3	2			3
CO5	2	3	3	1	2							1	3		
CO6					2			2	2	3	1	1			3

3/2/1: high/medium/low