Project Phase II Report On

Hashtag Recommendation for Multimodal Data and Short Description Generation for Images

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

Bachelor of Technology

in

Computer Science and Engineering

By

Fathima Sahliya K. S. (U2003079)
Iva Sony (U2003098)
Jiya Joy (U2003102)
Khadeeja C. R. (U2003119)

Under the guidance of
Ms. Seema Safar

Department of Computer Science and Engineering
Rajagiri School of Engineering & Technology (Autonomous)
(Parent University: APJ Abdul Kalam Technological University)
Rajagiri Valley, Kakkanad, Kochi, 682039
April 2024

CERTIFICATE

This is to certify that the project report entitled ***"Hashtag Recommendation for Multimodal Data and Short Description Generation for Images"*** *is a bonafide record of the work done by* ***Fathima Sahliya K. S. (U2003079)***, ***Iva Sony (U2003098)***, ***Jiya Joy (U2003102)***, ***Khadeeja C. R. (U2003119)***, *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 2023-2024.*

Ms. Seema Safar
Project Guide
Asst. Professor
Dept. of CSE
RSET

Dr. Tripti C.
Project Coordinator
Associate Professor
Dept. of CSE
RSET

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

ACKNOWLEDGMENT

We wish to express our sincere gratitude towards **Dr. P. S. Sreejith**, Principal of RSET, and **Dr. Preetha K. G.**, Head of the Department of Computer Science for providing us with the opportunity to undertake our project, ”Hashtag Recommendation for Multimodal Data and Short Description Generation for Images”.

We are highly indebted to our project coordinator, **Dr. Tripti C.**, Associate Professor, Department of CSE, for her valuable support.

It is indeed our pleasure and a moment of satisfaction for us to express our sincere gratitude to our project guide **Ms. Seema Safar**, Asst. Professor, Department of CSE, 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.

Fathima Sahliya K. S.

Iva Sony

Jiya Joy

Khadeeja C. R.

Abstract

In the contemporary era dominated by social media, the effective use of hashtags plays a pivotal role in enhancing content visibility and engagement. The project focuses on leveraging deep learning techniques to seamlessly integrate textual and visual information for hashtag recommendation, and utilizing image data for short description generation.The Hashtag Recommendation Module utilizes multimodal data and an encoder-decoder architecture to predict relevant hashtags, enhancing content discoverability and engagement. Meanwhile, the Short Description Generation Module harnesses feature extraction capabilities and a diverse training dataset to automatically generate coherent descriptions for images, enriching user understanding and interaction. Experimental results demonstrate the efficacy of our methodology in improving accuracy and relevance in hashtag predictions, alongside efforts to enhance the quality and diversity of descriptions for images. By addressing the dual challenges of hashtag recommendation and short description generation, our approach offers a comprehensive solution for enhancing content understanding and retrieval in multimodal environments. This method outperforms existing approaches by jointly processing image and text data, capturing relationships. Moreover, the short description generation enables the model to generate more creative and contextually rich textual content. Overall, the project marks a significant advancement in hashtag recommendation and short description generation by alleviating user burdens and advanced multimodal content understanding for enhanced social media interaction and information retrieval.

Contents

Acknowledgment **i**

Abstract **ii**

List of Abbreviations **vi**

List of Figures **vii**

1 Introduction **1**

1.1 Background 1

1.2 Problem Definition 2

1.3 Scope and Motivation 2

1.4 Objectives 3

1.5 Challenges 3

1.6 Assumptions 3

1.7 Societal / Industrial Relevance 4

1.8 Organization of the Report 4

2 Literature Survey **6**

2.1 Attention-Based Multimodal Neural Network Model for Hashtag Recommendation 6

2.2 An Efficient and Resource-Aware Hashtag Recommendation Using Deep Neural Networks 8

2.3 Hybrid Deep Neural Network for Multimodal Hashtag Recommendation . . 9

2.4 Hashtag Recommender System Based on LSTM Neural Reccurent Network 10

2.5 SHE: Sentiment Hashtag Embedding Through Multitask Learning 12

2.6 Sentiment Enhanced Multi-Modal Hashtag Recommendation for Micro-Videos 15

2.7 Co-attention memory network for multimodal microblog's hashtag recommendation . . . 17
2.7.1 Multimodal Representation Learning . . . 17
2.7.2 Sequence Modeling with LSTM and BiLSTM . . . 18
2.7.3 Attention Mechanism . . . 18
2.7.4 Encoder-Decoder Architecture . . . 19
2.7.5 Memory Networks for Historical Context . . . 19
2.7.6 Grid-Level Modeling . . . 19
2.8 Hashtag recommendation for short social media texts using word-embeddings and external knowledge . . . 20
2.8.1 Extrinsic Feature Extraction . . . 20
2.8.2 Feature Selection and Processing . . . 20
2.8.3 Candidate Hashtag Generation . . . 20
2.8.4 User Influence Score Computation . . . 20
2.8.5 Candidate Hashtag Recommendation . . . 20
2.9 Summary and Gaps Identified . . . 22

3 Requirements **24**
3.1 Hardware and Software Requirements . . . 24

4 System Architecture **25**
4.1 System Overview . . . 25
4.2 Sequence Diagram . . . 27
4.3 Module Division . . . 28
4.4 Datasets Identified . . . 30
4.5 Work Breakdown and Responsibilities . . . 31
4.6 Work Schedule - Gantt Chart . . . 32

5 Results and Discussions **33**
5.1 Overview . . . 33
5.2 Testing . . . 34
5.2.1 Hashtag Recommendation on MM-INS Dataset Results . . . 34
5.2.2 Hashtag Recommendation on Nuswide Dataset Results . . . 35

5.2.3 Short Description Results on Test Data 36
5.3 Quantitative Results 37
5.4 Graphical Analysis 38
5.5 Discussion 38

6 Conclusions & Future Scope **40**

References **41**

Appendix A: Presentation **43**

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes **60**

Appendix C: CO-PO-PSO Mapping **65**

List of Abbreviations

AE - Auto Encoder

AMNN - Attention-based Multimodal Neural Network Model

BERT - Bidirectional Encoder Representation from Transformers

BiLSTM - Bidirectional Long Short Term Memory

BLEU - BiLingual Evaluation Understudy

CNN - Convolutional Neural Network

DBSCAN - Density-Based Spatial Clustering of Applications with Noise

GRU - Gated Recurrent Unit

KNN - K-Nearest Neighbour

LDA - Linear Discriminant Analysis

LSTM - Long Short Term Memory

LT - Language Translater

ML - Machine Learning

MLP - Multi Layer Perceptron

NDCG - Normalized Discounted Cumulative Gain

ReLu - Rectified Linear Unit

ResNet - Residual Network

RNN - Recurrent Neural Network

SVM - Support Vector Machine

USE - Universal Sentence Encoder

VGG - Visual Geometry Group

ZFNet - Zeilur Fergus Network

List of Figures

2.1 Overall Architecture of AMNN[1] 7
2.2 Overall Architecture of HAZEL[2] 8
2.3 Hybrid Deep Neural Network for Multimodal Hashtag Recommendation figures.[3] 10
2.4 Hashtag Recommender System Based on LSTM Neural Reccurent Network figures.[4] 11
2.5 Architecture of SHE model [5] 13
2.6 Comparison of SHE and SE with Semantic Embedding Methods [5] 14
2.7 Architecture of TOAST model [6] 16
2.8 Architecture of CoA-MN [7] 18
2.9 Systems architecture of hashtag recommendation [8] 21

4.1 Sequence Diagram 27
4.2 Hashtag Recommendation Module 29
4.3 Short Description Generation Module 30
4.4 Gantt Chart 32

Chapter 1

Introduction

1.1 Background

In recent years, the proliferation of social media platforms and online content sharing has led to an exponential increase in the volume of multimedia data being generated daily. This surge has presented users with a challenge of content discovery amidst the vast sea of information. Hashtags have emerged as crucial tools for organizing and categorizing content, providing users with a means to navigate and explore topics of interest. However, the conventional methods of hashtag recommendation often fall short in capturing the complexities inherent in multimodal data, which includes both images and text.

In the current scenario, social media platforms are witnessing a paradigm shift towards content that seamlessly integrates both visual and textual elements. Users share not only textual posts but also images, creating a rich tapestry of content that requires more sophisticated methods for effective categorization and discovery. Existing hashtag recommendation systems, typically designed for text-only data, struggle to adapt to this evolving landscape, limiting their ability to provide accurate and contextually relevant suggestions. The importance of a robust multimodal hashtag recommendation and short description generation system becomes evident when considering the need for enhanced content understanding. Users seek content that aligns with their preferences, interests, and values, and hashtags play a pivotal role in facilitating this discovery process. An advanced system that can intelligently analyze both images and text to generate relevant hashtags not only improves content visibility but also enhances the overall user experience by providing more nuanced recommendations.

1.2 Problem Definition

Aim of the project is to develop a multimodal deep learning model to recommend relevant hashtags for multimodal data which includes both visual and textual data then further generate concise descriptions for images, aiming to enhance content understanding and organization in social media and image-based platforms.

1.3 Scope and Motivation

This project's scope is extensive, aiming to develop a versatile and scalable multimodal recommendation system adept at integrating image and text data seamlessly. Utilizing advanced neural network architectures, including InceptionV3, Universal Sentence Encoder for the Hashtag Recommendation part and Resnet50 for the Short Description Generation part, the system targets diverse multimedia content on social media platforms, primarily focusing on hashtag recommendation and Short Description Generation to automate the process. Beyond social media, potential applications extend to content-based image retrieval systems and multimedia content management platforms. The project also explores the feasibility of extending the recommendation system to support other modalities, enhancing its adaptability in the evolving landscape of multimodal content. Overall, the project envisions widespread applications, positively impacting user experience and functionality across various industries, including image cataloging, content moderation, educational platforms, and even in emerging areas like autonomous vehicles.

The motivation for this project is rooted in the imperative to improve content discoverability and user engagement on social networks. Given users' frequent omission of hashtags, the project responds innovatively and timely by providing an intelligent, automated solution for hashtag recommendation. Through a multimodal approach, the model seeks to surpass the limitations of existing unimodal methods, aiming for more accurate and contextually rich hashtag suggestions. The broader motivation lies in refining content recommendation systems, contributing to discussions on integrating advanced neural network architectures to address real-world challenges in the dynamic realm of social media. Ultimately, the project aims to alleviate user burdens and advance multimodal content understanding for enhanced social media interaction and information retrieval.

1.4 Objectives

- Utilize encoder-decoder architecture to model the hashtag generation process.
- Recommend relevant hashtags for multimodal data which include both visual and textual data.
- Develop a simple and lightweight model for hashtag recommendation with better performance.
- Develop an encoder decoder model to create short description of images from hashtags.

1.5 Challenges

The major challenge was to find an appropriate and unbiased multimodal dataset for hashtag recommendation.The scalability and efficiency of models are scrutinized due to the vast data volume on social networks, requiring adaptability for increasing data size and user activities. Privacy and security are crucial concerns, demanding measures to protect user data and prevent unauthorized access. Additionally, there's a growing need for hashtag recommendation models to address biases, ensuring fairness by considering user diversity and minimizing the amplification of existing biases in recommendations.

1.6 Assumptions

- Dataset represents the diverse relationships between visual and textual elements and encompasses a comprehensive range of hashtags.
- The dataset utilized is accurate and includes relevant hashtags associated with both images and text.
- Multimodal dataset used for training represent wide range of content encountered on social media platforms.

1.7 Societal / Industrial Relevance

The proposed project holds significant relevance for both society and industry, addressing the evolving dynamics of multimedia content sharing platforms. In a societal context, the developed multimodal hashtag recommendation and short description generation system can greatly enhance user experiences on social media platforms. Users will benefit from more accurate and content recommendations, leading to improved content discoverability and a richer online engagement. This, in turn, contributes to a more vibrant and interactive online community.

From an industry perspective, the project has wide-ranging applications, particularly in digital marketing and content creation. Businesses and content creators can leverage the advanced hashtag recommendation system to optimize their social media strategies, ensuring that their content reaches the right audience. The improved short description generation also enhances the narrative around the content, making it more appealing and contextually relevant. The project's adaptability across various multimedia modalities further extends its applicability, making it a valuable tool for content management systems, image retrieval platforms, and other industries reliant on effective multimodal content understanding. In essence, the project's applications have the potential to reshape how information is categorized, discovered, and engaged with in both societal and industrial contexts.Moreover, applications in image cataloging, content moderation, and educational platforms can leverage image descriptions for better organization, filtering, and comprehension. As artificial intelligence continues to advance, the scope expands to areas such as autonomous vehicles, where image descriptions can aid in real-time scene understanding. Overall, the generation of short descriptions for images has broad applications, positively impacting accessibility, user experience, and functionality across diverse industries.

1.8 Organization of the Report

The report is structured to provide an exploration of a multimodal hashtag recommendation and short description generation system for images. It commences with an insightful introduction that delves into the background, problem definition, scope, motivation, objectives, challenges, assumptions, and societal/industrial relevance of the project. Following this, a Literature Survey examines existing research and models in the field, highlighting their methodologies, strengths, and identified gaps. The Requirements section outlines the hardware and software prerequisites for the system. The core of the report lies in the System Architecture, detailing the overall system overview, architectural design, module division, and work breakdown with associated responsibilities. The Results section presents the findings obtained from the implemented system, including evaluation metrics, performance analysis, and any notable observations. The report concludes with a robust Conclusions & Future Scope section, summarizing key findings and outlining potential avenues for future enhancements and applications.

Chapter 2

Literature Survey

2.1 Attention-Based Multimodal Neural Network Model for Hashtag Recommendation

The paper [1] introduces an attention-based multimodal neural network model (AMNN) designed for hashtag recommendation in social networks, specifically addressing the surge in the use of multimodal data, including images, in microblogs. The hashtag recommendation task is reframed as a sequence generation problem, and a hybrid neural network approach is employed to extract features from both text and images. The model incorporates a sequence-to-sequence structure with an attention mechanism to learn representations of multimodal microblogs and recommend relevant hashtags. Experimental results on Instagram data and two public datasets, HARRISON and NUS-WIDE, named MM-INS, demonstrate the superior performance of the proposed method over state-of-the-art methods, showcasing higher precision, recall, and accuracy. The paper outlines the overall architecture of the AMNN model, emphasizing the hybrid feature extraction encoder and the coupled decoder for hashtag recommendation.

In the encoder section of the proposed Attention-Based Multimodal Neural Network (AMNN) model, the paper outlines the process of extracting features from both images and texts. For image feature extraction, a hybrid neural network architecture is adopted, starting with a preliminary feature map captured by a Convolutional Neural Network (CNN). Representative neural networks, Inception V3 and ResNet-50, are chosen for this task. The obtained feature map is then processed using Long Short-Term Memory (LSTM) to form a sequence of hidden states. To account for specific regions contributing to hashtags in images, an attention mechanism is introduced, capturing spatial features and enhancing the model's robustness. For text feature extraction, Bidirectional LSTM (BiLSTM) is employed to learn long-term dependencies, providing output values from

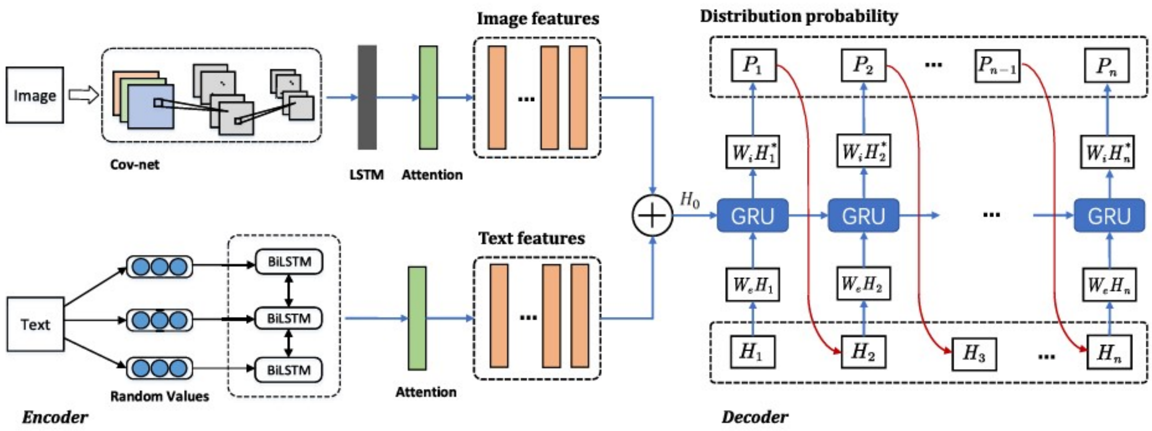


Figure 2.1: Overall Architecture of AMNN[1]

both forward and backward directions. The attention layer is applied to the concatenated hidden states to obtain the text features. The extracted features from both modalities are merged to create a comprehensive representation of the multimodal microblog.

Moving to the decoder section, a Gated Recurrent Unit (GRU) network is utilized. In the training process, hashtags are mapped into a fixed feature space as inputs for GRU networks. The model captures correlations between hashtags and multimodal data. The training process involves updating and resetting gates to generate the next output. During prediction, only multimodal feature maps are used as input, and connections between previous output and successive input are removed. The GRU unit uses update and reset gates to produce the next output, considering the transition equations. Finally, a fully connected network and softmax function generate the probability distribution of each hashtag at each time step, resulting in a ranked list of candidate hashtags for recommendation. To evaluate the proposed model, hashtag recommendation tasks are conducted using two public datasets, namely HARRISON [9] and NUS-WIDE [10], along with a collection of crawled microblogs from Instagram named MM-INS, acquired using the Instaloader API. The HARRISON dataset comprises 57,383 photographs from Instagram, each associated with an average of 4.5 hashtags. NUS-WIDE consists of 269,000 images and associated tags from Flickr.

2.2 An Efficient and Resource-Aware Hashtag Recommendation Using Deep Neural Networks

The paper presents an efficient and resource-aware hashtag recommendation system HAZEL [2] that leverages image classification and semantic embedding models, incorporating zero-shot learning. The approach involves utilizing Convolutional Neural Network (CNN) models to extract features from images, ensuring uniqueness for each class. A semantic embedding space is created to identify highly probable hashtags and their neighbors, facilitating hashtag recommendation. The system periodically updates the semantic embedding space to align with evolving social networking trends. The visual-semantic embedding framework incorporates a Visual Conceptual Embedding Function, employing CNNs, and a Label Embedding Function using Word2Vec for hashtag mapping. The model combines visual and semantic embeddings to predict hashtags for given images, integrating a large-margin framework and hinge loss for training.

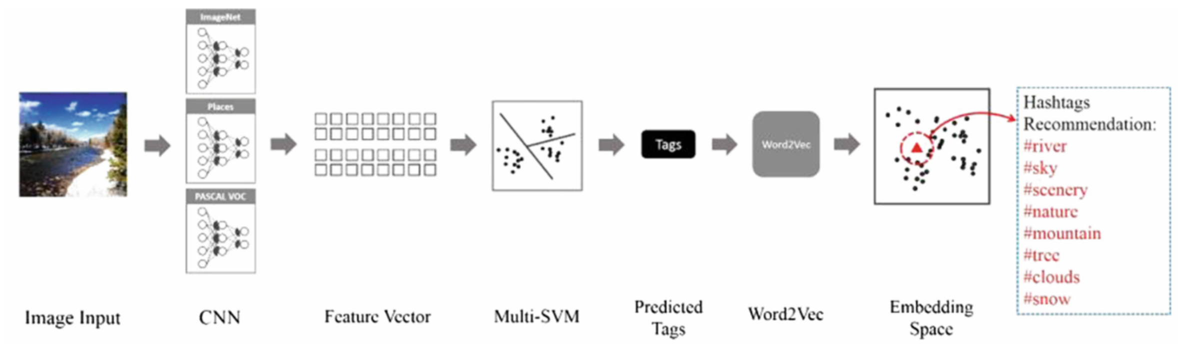


Figure 2.2: Overall Architecture of HAZEL[2]

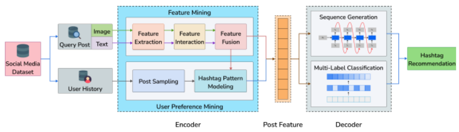
Two CNN architectures, ResNet and ZFNet, are employed to enhance differentiation between hashtags, recognizing 1,385 features in total. Each CNN model is individually applied to ImageNet [11], Places365, and PASCAL datasets using Keras, PyTorch, and Caffe for processing, respectively. Word2Vec, specifically the Skip-gram model, is employed to represent words in vector form, training hashtags for recommendation. The hashtag recommendation process involves CNNs extracting image features, Support Vector Machines (SVM) learning from these features to identify hashtag categories, and turning hashtags into vector representations for Skip-gram model utilization. The recommendation employs k-NN (K-nearest neighbors) algorithm to rank hashtags based on cosine distance, ensuring semantic closeness.

The paper integrates three pre-trained CNN models for robust feature extraction. The Word2Vec model expands the initial set of hashtags, and the semantic embedding space is periodically updated to align with the latest trends. The proposed system effectively combines visual and semantic information to provide efficient and resource-aware hashtag recommendations, contributing to social media content categorization and user engagement.

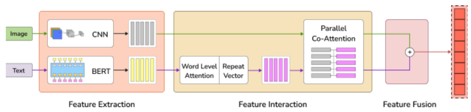
2.3 Hybrid Deep Neural Network for Multimodal Hashtag Recommendation

The DESIGN system represents a sophisticated hybrid deep neural network designed for personalized hashtag recommendations, aiming to enhance user engagement and content organization by capturing intricate hashtag interdependencies. Unlike conventional methods, it diverges from reliance on ordered sequences, incorporating user preferences, tagging behavior, and multimodal features for optimal hashtag suggestions. In the Feature Mining module, visual features are extracted using CNNs like VGG-16 or ResNet-50, while textual features employ BERT for bidirectional context-awareness. Feature Interaction incorporates word-level attention, creating affinity matrices and attention weights for both text and image features. Feature Fusion then combines global image and text feature vectors to form a content-based post feature vector.

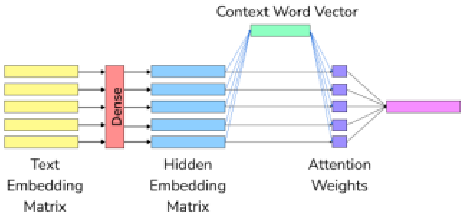
The User Preference Mining module employs post sampling to understand user tagging patterns, generating influence vectors by analyzing historical posts and user test posts. This influence vector represents the user's preferences and is concatenated with the content-based post feature vector. The Hashtag Prediction module encompasses multilabel classification and sequence generation, utilizing a predefined pool of hashtags and a GRU-based RNN for predicting relevant hashtags. The Candidate Hashtag Recommendation component selects hashtags from multiple prediction models, aggregates them through an ensemble module, and evaluates their quality against ground truth using F1-score, precision, and recall metrics. This comprehensive system reflects a robust framework for personalized hashtag recommendation in social media, showcasing an innovative blend of feature extraction, user preference mining, and hashtag prediction techniques.[3]



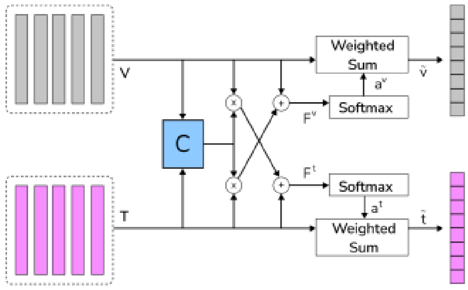
(a) Overall architecture.



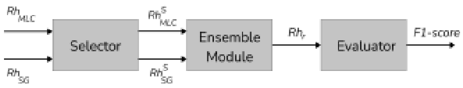
(b) Feature mining.



(c) Word-level attention.



(d) Parallel Co-attention.



(e) Hashtag Recommendation Module.

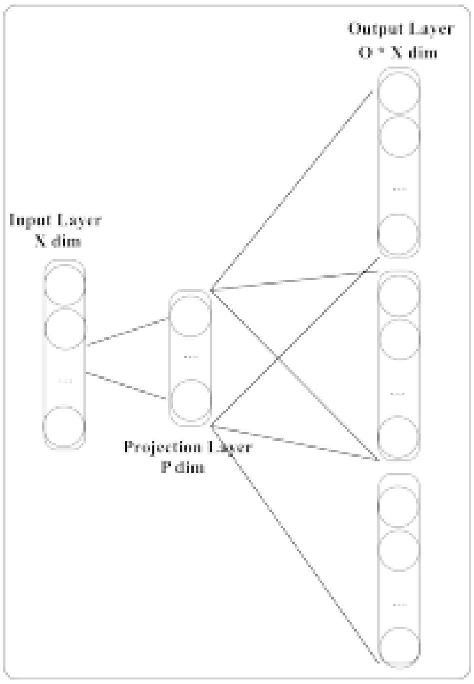
Figure 2.3: Hybrid Deep Neural Network for Multimodal Hashtag Recommendation figures.[3]

2.4 Hashtag Recommender System Based on LSTM Neural Reccurent Network

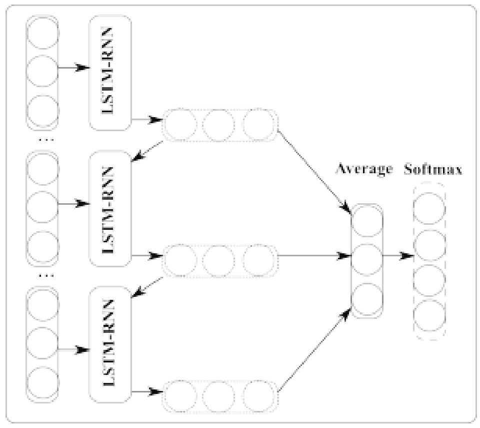
The paper[4] introduces a cutting-edge LSTM-based model to overcome the limitations of existing hashtag recommendation methods on micro-blogging platforms, particularly in the context of Twitter. These limitations are characterized by a lack of consideration for tweet semantics and an overreliance on sparse lexical features. The proposed model is meticulously structured, comprising four integral components that collectively contribute to the effectiveness of hashtag recommendations: word embedding, tweet representation, tweets clustering, and hashtag recommendation.

In the word embedding phase, the model leverages the Skip-gram model, a member of the Word2Vec family of algorithms. This model excels at capturing the nuanced meaning and contextual intricacies of words by predicting context words based on a given target word. The resulting word embeddings facilitate a more nuanced understanding of tweet semantics.

Moving on to tweet representation, the model employs Long Short-Term Memory networks (LSTMs), a specialized type of Recurrent Neural Networks (RNNs). LSTMs are specifically designed to address the long-term dependency problem, featuring a unique architecture with a cell state and three essential gates. This architecture allows LSTMs to capture and retain information over extended sequences without succumbing to the vanishing gradient problem associated with traditional RNNs.[4]



(a) The skip-gram model architecture.



(b) The process of encoding the tweet vector representation.

Figure 2.4: Hashtag Recommender System Based on LSTM Neural Reccurent Network figures.[4]

The next component involves clustering tweets using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. This algorithm efficiently groups tweets into clusters based on semantic similarity. Crucially, the determination of the epsilon parameter, representing the radius of neighborhoods around data points, is tailored for each tweet, enhancing the granularity of the clustering process.

The final stage is the hashtag recommendation component, where the model recommends the top-k hashtags to users. This recommendation is founded on computing rank scores derived from the similarity values between an input tweet and cluster centers. Candidate hashtags are then selected from clusters containing similar tweets, and their rank scores are influenced by the frequency of appearance in the collection of tweets. This strategy ensures that the recommended hashtags not only align with the content of the input tweet but also resonate with similar tweets, enhancing the relevance and potential engagement of the suggested hashtags. Overall, the proposed model offers a comprehensive and sophisticated solution to the nuanced task of hashtag recommendation in the context of micro-blogging platforms like Twitter.

2.5 SHE: Sentiment Hashtag Embedding Through Multitask Learning

The paper [5] proposes a model called Sentiment Hashtag Embedding (SHE) for sentiment analysis of hashtags on social media platforms. The paper aims to address the challenges of understanding and analyzing hashtags, such as normalization, grouping related hashtags, and identifying sentiment expressed by hashtags.The proposed model SHE incorporates sentiment information into hashtag embeddings and demonstrates its effectiveness in various real-world applications, including hashtag sentiment classification, tweet sentiment classification, and retrieval of semantically similar hashtags. Word embedding methods such as Word2Vec, C&W, and DeepWalk represent words or hashtags into low-dimensional vectors and capture semantic distribution.[12] The model then uses a semisupervised sentiment hashtag embedding (SHE) approach that exploits a multitask learning approach to preserve semantic information through autoencoder (AE) while encoding sentiment information using a convolutional neural network (CNN) classifier simultaneously.

The pretrained embeddings are used as input to the SHE model, which learns to encode sentiment information while preserving the semantic information captured by the pretrained embeddings.The use of a multitask learning approach allows the model to learn both tasks simultaneously, improving the overall performance of the model. The encoding stage of the AE uses a convolutional neural network (CNN) to capture the latent spatial aspects of the pretrained embedding vector. The decoding stage of the AE regenerates the input vector using dense perceptron layers and tanh activation function. The output of the CNN layer is used as input to the softmax sentiment classifier, which classifies the sentiment of the intermediate representation vector. The model is trained in two phases: Phase-I trains the AE without the softmax classifier using unlabeled hashtags, and Phase-II retrains the AE with the softmax classifier using sentiment annotated hashtags. The

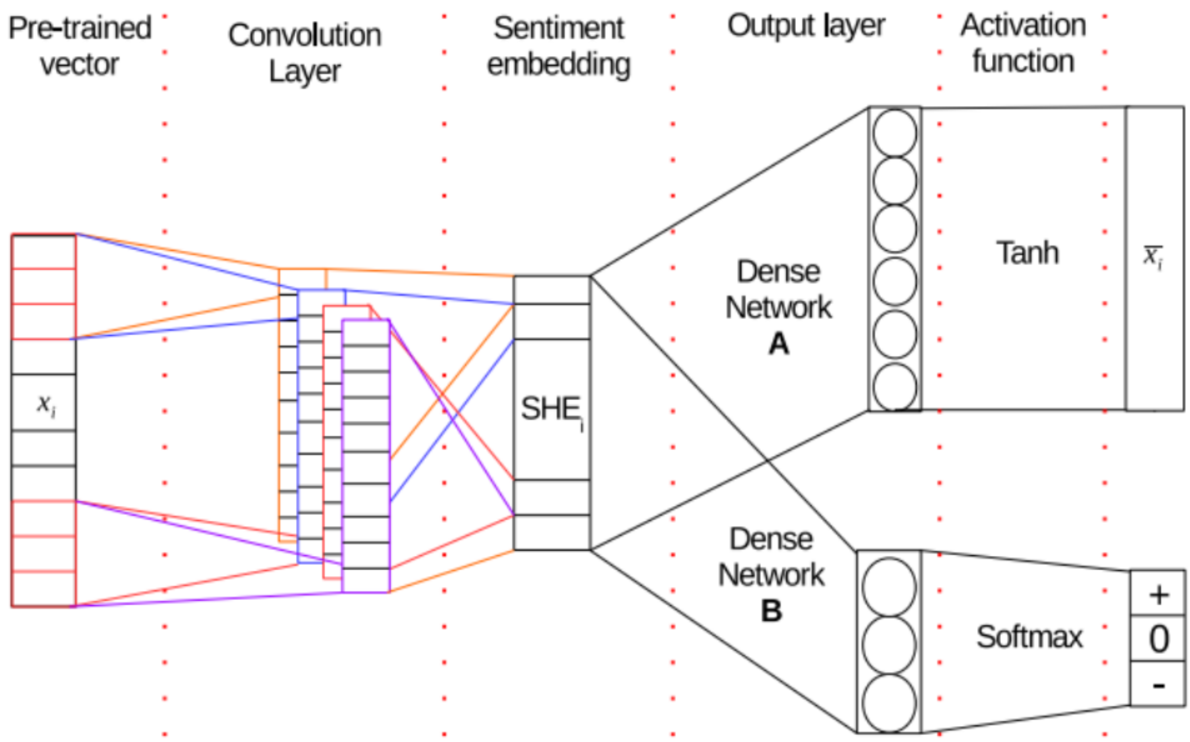


Figure 2.5: Architecture of SHE model [5]

process of training Phase-I and Phase-II is repeated until convergence. Once the model is trained, the sentiment embedding (SE) of a hashtag is defined by the output of the CNN layer.

The proposed model has potential applications in various areas such as social media analysis, marketing research, and customer feedback analysis. The encoder takes the input data and maps it to a lower-dimensional representation, This mapping is typically achieved through a series of hidden layers that gradually reduce the dimensionality of the input. The encoder's goal is to capture the most important features of the input data in the latent space.

The decoder takes the code from the encoder and reconstructs the original input data. It uses a series of hidden layers that gradually increase the dimensionality of the code until it matches the dimensionality of the input. [5] The decoder's goal is to generate a reconstruction that is as close as possible to the original input. The AE is responsible for preserving the semantic characteristics of the hashtags, while the CNN classifier is used to encode the sentiment polarity into the pretrained embeddings . The training process is divided into two phases, and the output of the model is defined by the output of a CNN

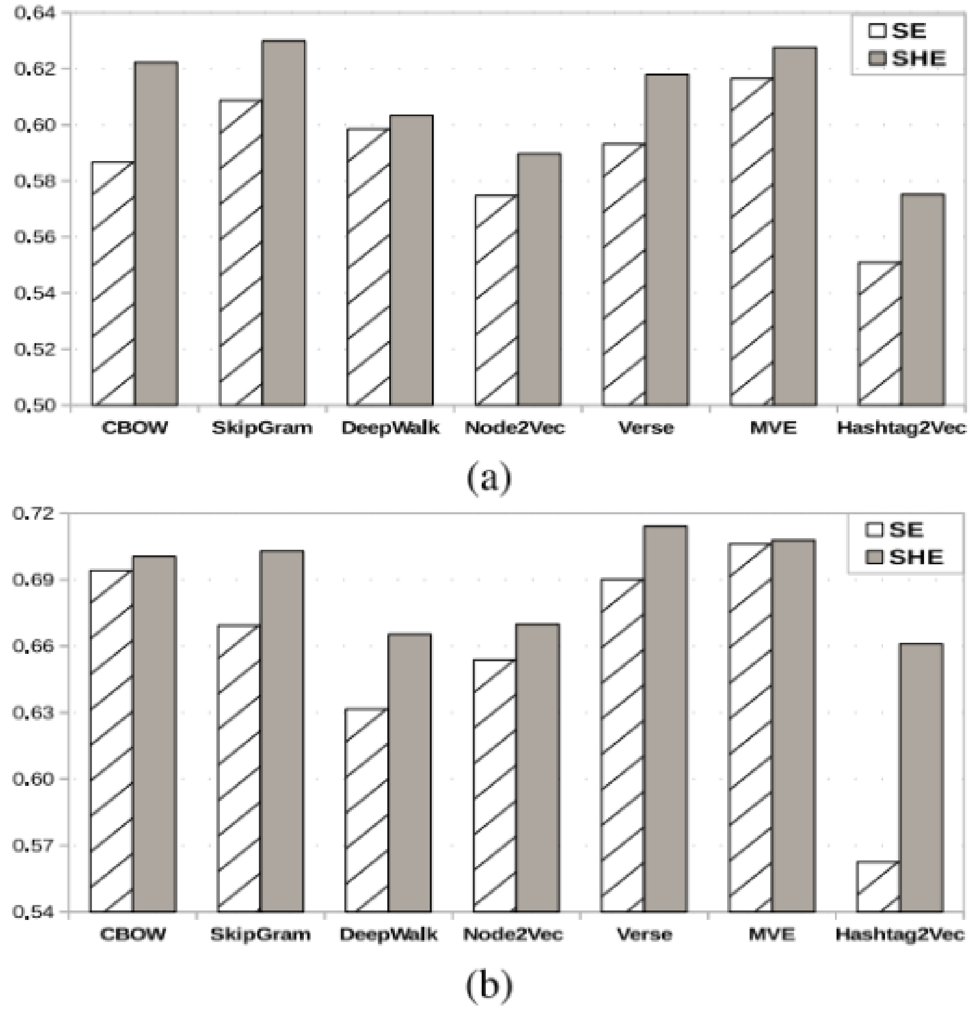


Fig. 2. Performance of tweet sentiment classification using SE and SHE. (a) SemEval 2013 data set. (b) SemEval 2016 data set. (a) SemEval 2013 Dataset. (b) SemEval 2016 Dataset.

Figure 2.6: Comparison of SHE and SE with Semantic Embedding Methods [5]

layer. In Phase-I, the proposed model is trained using an unsupervised learning technique called Autoencoder (AE). The AE is trained without the softmax classifier using unlabeled hashtags in the corpus. The purpose of this phase is to learn the underlying features of the hashtags in an unsupervised manner. The output of this phase is a pre-trained AE model. In Phase-II, the pre-trained AE model is retrained with a softmax sentiment classifier using sentiment annotated hashtags. The purpose of this phase is to fine-tune the pre-trained AE model to classify the sentiment of hashtags. The output of this phase is a fine-tuned AE model with a sentiment classifier. The process of training Phase-I and Phase-II is repeated until the convergence. The convergence is achieved when the performance of the model on a validation set stops improving. The activation function used in the decoder layer is tanh, which stands for hyperbolic tangent. This function maps input values to a range between -1 and 1, which can help with normalization and preventing vanishing gradients. The activation function used in the classification layer is softmax, which is commonly used for multi-class classification problems. It outputs a probability distribution over the possible classes, with the highest probability indicating the predicted class. The proposed model, referred to as SHE, was trained using this architecture and was found to converge after five iterations with 20 epochs per iteration.

Once the model is trained, the sentiment of a hashtag is determined by the output of the CNN layer. The CNN layer takes the preprocessed hashtag as input and produces a vector representation of the hashtag. The vector representation is then used to determine the sentiment of the hashtag using the softmax classifier.

2.6 Sentiment Enhanced Multi-Modal Hashtag Recommendation for Micro-Videos

The paper [6] proposes a model called the TOAST model. Sentiment enhanced multi-modal attentive hashtag recommendation (TOAST) model is proposed for micro-video hashtag recommendation.The model leverages information from visual, acoustic, and textual modalities to capture the sentiment and content features of micro-videos, ultimately improving hashtag recommendation performance.

Hashtags serve as more than mere labels; they inherently carry semantic information, primarily used for tagging topics, highlighting video content, and expressing emotions. They can be categorized into sentiment hashtags (e.g., #funny, #sad) and content hashtags (e.g., #kid, #dinner, #piecemeal) based on their semantic meanings. Content features predominantly focus on recognizing objects within videos. The TOAST model comprises two branches: the sentiment common space learning branch and the content common space learning branch. Both branches receive inputs from three modalities: visual, acoustic, and textual modalities.

The evaluation protocol of the TOAST model involves pairing positive instances with 100 negative hashtags for rigorous testing. Performance is assessed using industry-standard metrics such as Recall and Normalized Discounted Cumulative Gain (NDCG), enabling a thorough analysis of the Top-N hashtag recommendation lists.

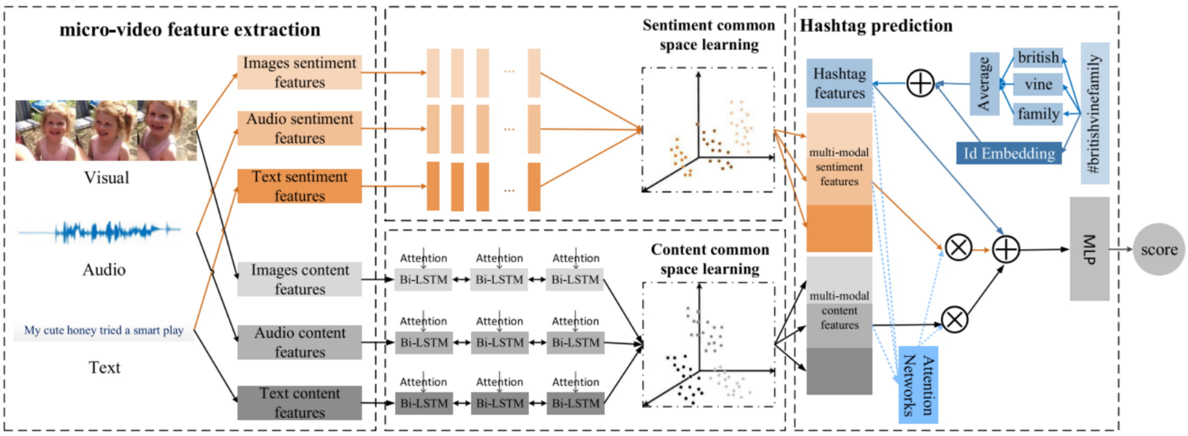


Figure 2.7: Architecture of TOAST model [6]

The core architecture of the TOAST model consists of three pivotal components:

- Sentiment Common Space Learning Network: This component adeptly captures nuanced sentiment features embedded within the micro-video content, enhancing the model's ability to understand the emotional context.

- Self-Attentive Content Common Space Learning Network: Leveraging self-attention mechanisms, this component focuses on the sequential content features within micro-videos, enabling the model to discern crucial temporal patterns and relevance within the content.

- Hashtag Embedding: Employing sophisticated embedding techniques, this element represents the semantics of hashtags within a learned space, enabling the model to understand and recommend hashtags effectively.

A dynamic capture of the importance of multi-modal sentiment and content features is achieved by an attention neural network, which assesses their consistency with the semantics of candidate hashtags. A multi-layer perceptron (MLP) network is employed to predict interactions between hashtags and micro-videos. Bi-directional LSTM (Bi-LSTM) is utilized to capture information relevant to the content of the micro-videos. Introducing a self-attention mechanism with Bi-directional LSTM enhances multi-modal content feature learning by filtering out noise and capturing information most pertinent to the corresponding hashtags.

The TOAST model presents a comprehensive and effective approach to multi-modal hashtag recommendation for micro-videos, showcasing the importance of sentiment and content features and the effectiveness of the proposed self-attention mechanism and hashtag embedding. The model's performance surpasses existing methods, and the study provides valuable insights into the significance of different modalities and sentiment features in hashtag recommendation.

2.7 Co-attention memory network for multimodal microblog's hashtag recommendation

Hashtags play a crucial role in categorizing microblog posts on social media platforms, and the task of recommending suitable hashtags has garnered significant attention with the rapid development of social networks. The paper addresses the limitations of conventional methods treating the hashtag recommendation task as a multi-class classification problem. The proposed co-attention memory network introduces novel methodologies, integrating advanced techniques like VGGNet16, LSTM, BiLSTM, attention mechanisms, and encoder-decoder structures to enhance the accuracy and adaptability of hashtag recommendations in dynamic microblogging environments[7].

2.7.1 Multimodal Representation Learning

The paper emphasizes the importance of capturing both textual and visual information in microblogs, introducing VGGNet16 for image processing to effectively extract visual

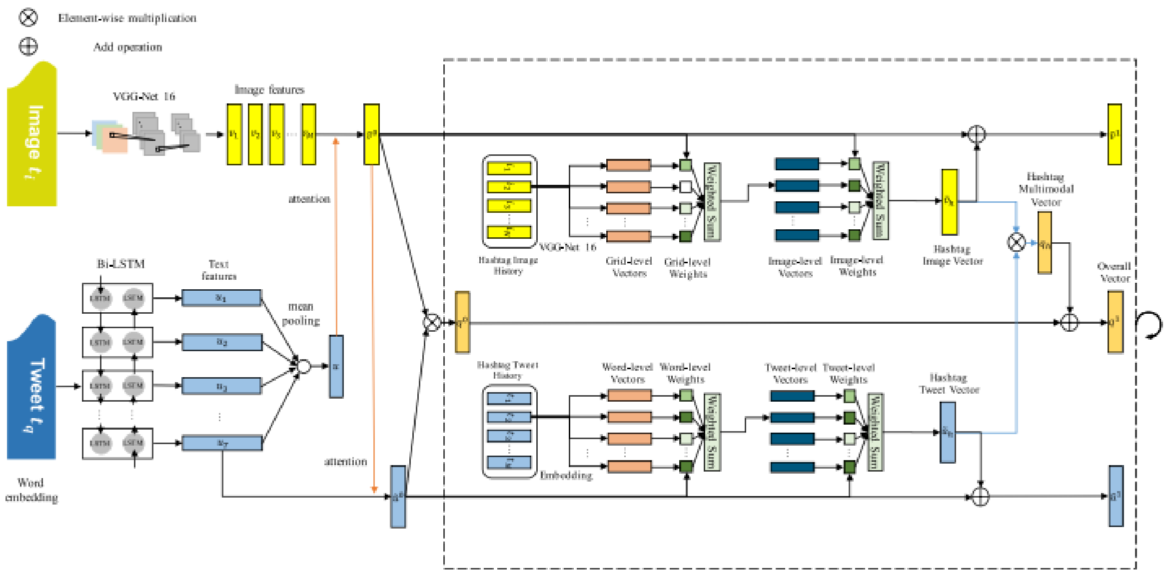


Figure 2.8: Architecture of CoA-MN [7]

features. This aligns with the broader trend in literature recognizing the significance of multimodal representation learning for a more comprehensive understanding of social media content.

2.7.2 Sequence Modeling with LSTM and BiLSTM

To capture the sequential nature of textual information, the paper incorporates LSTM and BiLSTM (Bidirectional Long Short-Term Memory). The use of LSTM in hashtag recommendation aligns with the literature highlighting the efficacy of recurrent neural networks for modeling sequential dependencies. BiLSTM, being a bidirectional variant, enhances the model's ability to capture contextual information from both past and future contexts, improving the overall sequence modeling.

2.7.3 Attention Mechanism

An attention mechanism is introduced to jointly attend to the query tweet and the post history of candidate hashtags. This attention mechanism is in line with the prevailing literature on attention mechanisms, emphasizing their effectiveness in capturing relevant information and improving the model's focus on critical aspects of the input data.

2.7.4 Encoder-Decoder Architecture

The paper utilizes an encoder-decoder architecture with a hierarchical attention mechanism, contributing to the encoder's ability to formulate a high-quality history interests representation. This architecture is consistent with existing literature that highlights the benefits of encoder-decoder structures, particularly in tasks involving sequence-to-sequence modeling and attention mechanisms.

2.7.5 Memory Networks for Historical Context

Memory networks are employed to store and retrieve information from a memory component, enabling the model to effectively capture historical context from past posts and hashtag usage patterns. The use of memory networks aligns with literature recognizing their utility in tasks requiring retention of sequential information and adaptation to dynamic environments.

2.7.6 Grid-Level Modeling

Grid-level modeling is introduced to capture fine-grained visual information from different regions of an image. This approach is in line with literature acknowledging the importance of spatial features and grid-level representations in image processing tasks, contributing to a more nuanced understanding of visual content in microblogs.

By incorporating advanced techniques such as VGGNet16, LSTM, BiLSTM, attention mechanisms, and encoder-decoder structures, the proposed co-attention memory network demonstrates a nuanced understanding of multimodal microblogs, effectively addressing challenges posed by data sparsity, dynamic user behavior, and emerging hashtags. The alignment of the proposed methodologies with prevailing trends in the field of deep learning and multimodal representation learning, marks a substantial step forward in the realm of hashtag recommendation systems.

2.8 Hashtag recommendation for short social media texts using word-embeddings and external knowledge

The paper introduces an approach to address the challenges of data sparsity and semantic feature incorporation in hashtag recommendation for short social media texts. The proposed method, detailed in [8], is structured around five key steps within the hashtag recommendation system. These steps include:

2.8.1 Extrinsic Feature Extraction

The initial phase involves extracting extrinsic features from external knowledge sources. This step is crucial for overcoming data sparsity in tweet representation, and it enriches tweet context by incorporating relevant information from external knowledge sources.

2.8.2 Feature Selection and Processing

To ensure the quality of features, a meticulous process of selection and cleaning is undertaken. Noisy features are identified and removed from the extrinsic feature set, refining the input for subsequent stages.

2.8.3 Candidate Hashtag Generation

Various techniques for generating candidate hashtags are proposed. These techniques aim to recommend hashtags for a given tweet based on different aspects. The diversity in candidate hashtag generation methods adds depth to the recommendation process.

2.8.4 User Influence Score Computation

Recognizing the significance of user influence in hashtag recommendation, the method includes a step to compute user influence scores. This is based on users' influential positions, adding a personalized dimension to the recommendation process.

2.8.5 Candidate Hashtag Recommendation

The final step involves aggregating the results from different candidate hashtag generation methods using a learning-to-rank approach. This ensures that the most relevant and

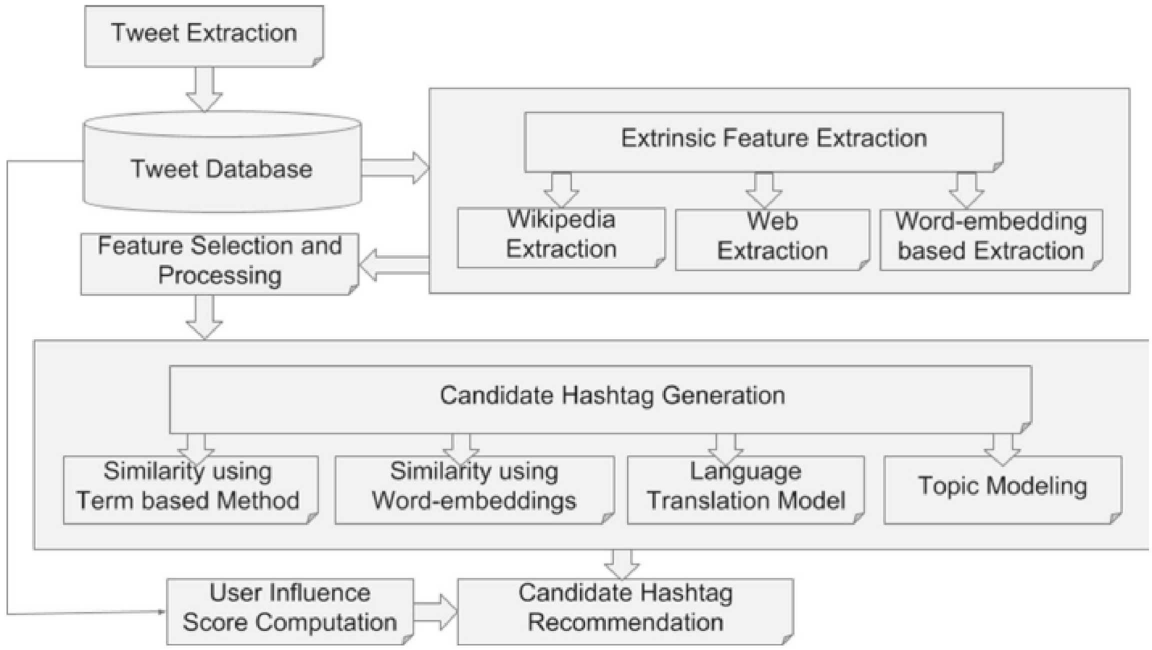


Figure 2.9: Systems architecture of hashtag recommendation [8]

impactful hashtags are prioritized in the final recommendation.

Additionally, external knowledge sources play a vital role in providing context for short tweets. Extrinsic features derived from these sources contribute to overcoming data sparsity, enhancing the overall effectiveness of the recommendation system.

The proposed method's performance is rigorously evaluated against four baseline methods, including two word-embedding-based methods and one RankSVM-based method. The word-embedding techniques employed in the evaluation encompass Latent Dirichlet Allocation (LDA), Word2Vec, and Language Translation (LT). The evaluation metrics employed measure the accuracy and relevance of the recommended hashtags, providing a comprehensive assessment.

The results demonstrate that the proposed method outperforms the baseline methods, establishing its superiority in hashtag recommendation for short social media texts. The incorporation of semantic features based on word embeddings and user influence metrics contributes to this superior performance, showcasing the method's ability to address the limitations of existing approaches. The experimental findings validate the effectiveness of the proposed method in overcoming data sparsity, incorporating semantic features, and providing accurate hashtag recommendations[8].

2.9 Summary and Gaps Identified

The papers present diverse approaches to enhance hashtag recommendation in the context of social media content. The "Sentiment Enhanced Multi-Modal Hashtag Recommendation (SHE)" model integrates sentiment and content features, utilizing self-attention mechanisms and hashtag embeddings for micro-video recommendation. However, convergence after five iterations may imply longer training times, and the study lacks extensive discussion on dataset specifics. The "TOAST Model for Micro-Video Hashtag Recommendation" comprehensively leverages visual, acoustic, and textual modalities, introducing innovative self-attention mechanisms and hashtag embeddings. Nevertheless, limited exploration of modality impact and depth in model interpretability are noted. "HAZEL" efficiently utilizes deep neural networks for hashtag recommendation, incorporating image classification and semantic embedding models. Its use of CNNs like ResNet and ZFNet, along with Word2Vec for semantic mapping, ensures robust feature extraction and recommendation. The "DESIGN" system adopts a hybrid deep neural network for personalized hashtag recommendations, integrating user preferences, tagging behavior, and multimodal features. This comprehensive approach involving feature mining, user preference mining, and hashtag prediction showcases a sophisticated framework. The "LSTM-based Hashtag Recommender System" tackles limitations in existing methods for micro-blogging platforms like Twitter by employing LSTM networks for tweet representation and hashtag recommendation. The clustering of tweets using DBSCAN and leveraging tweet semantics enhance the model's recommendation accuracy. The "Co-attention Memory Network for Microblog's Hashtag Recommendation" introduces advanced techniques like VGGNet16, LSTM, BiLSTM, attention mechanisms, and memory networks. Still, a comprehensive analysis of computational efficiency and real-world scalability is lacking. Lastly, "Hashtag Recommendation for Short Social Media Texts" proposes an approach addressing data sparsity by incorporating extrinsic features from external knowledge sources. However, the study does not thoroughly analyze hyperparameter sensitivity or discuss potential biases introduced by external knowledge sources. Overall, while these papers contribute significantly to hashtag recommendation, gaps in interpretability, computational efficiency, scalability, hyperparameter sensitivity, and bias considerations highlight avenues for future research to enhance the robustness and transparency of recommendation systems.

Identified Gaps:

1. Most approaches lack detailed discussions on the interpretability and explainability of their models, leaving a gap in understanding how the model arrives at specific recommendations.

2. While the architectures are sophisticated, there is a lack of exploration into the computational efficiency of these models, which is crucial for real-world applications.

3. The scalability of the proposed models to larger datasets or diverse scenarios is not extensively discussed, leaving uncertainty about their adaptability to different social media contexts.

4. Several approaches do not thoroughly analyze the sensitivity of their models to varying hyperparameters, leaving questions about the robustness of the proposed architectures.

5. The incorporation of external knowledge sources introduces potential biases. However, there is a lack of discussion on how these biases might impact the fairness and accuracy of hashtag recommendations.

Chapter 3

Requirements

3.1 Hardware and Software Requirements

The system necessitates specific hardware and software components for optimal performance. Hardware requirements include an NVIDIA GPU, an i5 processor, and 8 GB of RAM to support the computational demands of processing multimodal data. On the software side, it requires Keras and TensorFlow frameworks for neural network development and training. Additionally, the system operates on a Windows 11 environment and relies on Python 3.11 for programming and implementation. These hardware and software specifications are essential for efficient execution and successful utilization of the system's neural network algorithms, ensuring smooth processing and analysis of textual and visual information to generate accurate hashtag recommendations and short descriptions.

Chapter 4

System Architecture

4.1 System Overview

In response to the growing need for better ways to suggest hashtags on social media, this project introduces a new way to understand and recommend hashtags. It uses both text and pictures to get a better idea of what content is about. The project revolutionizes content comprehension and engagement across social media and multimedia platforms. The system initiates with image feature extraction, leveraging the Inception V3 model to dissect images into shapes, textures, and configurations. Simultaneously, textual features undergo rigorous extraction via the Universal Sentence Encoder, capturing semantic nuances with precision. Multimodal fusion integrates these features, fostering a holistic understanding of content. Hashtag prediction utilizes a sophisticated encoder-decoder architecture, optimizing hashtag relevance through meticulous training on multimodal datasets. Finally, short description generation employs ResNet50's feature extraction prowess, crafting diverse and compelling narratives from images. This comprehensive approach reshapes content understanding, offering immersive experiences across multimedia platforms.

The key idea behind this new method is to explore how text and images work together to understand content better. By using both text and pictures, it tries to see the bigger picture and capture the real meaning behind the content. This approach isn't just about putting labels on content; it's about digging deeper into the story told by the text and images, making it easier for people to understand what the content is all about.

This project brings a fresh perspective to understanding and suggesting hashtags. By bringing text and images together, it aims to not only suggest better hashtags but also to create short descriptions that really capture the essence of the content. It's a step forward in how we understand and discover content on social media, aiming to make it easier for people to find and connect with content that resonates with them.

- Image Feature Extraction

The process of extracting crucial features from images entails a meticulous approach leveraging the Inception V3 model, a state-of-the-art convolutional neural network (CNN) meticulously crafted for intricate image classification tasks. Through a series of intricate computations, the system dissects images into myriad components, encompassing shapes, textures, and configurations. A distinguishing aspect of this process lies in the integration of attention mechanisms within the CNN architecture, acting as a discerning spotlight that accentuates pivotal elements within the images. By spotlighting specific regions or features, this attention mechanism enriches the system's prowess in comprehending and deciphering essential details within visual content, ensuring a heightened level of analysis and interpretation.

- Text Feature Extraction

Concurrently, textual features undergo a rigorous extraction process facilitated by the Universal Sentence Encoder by Google. This sophisticated encoder orchestrates the transformation of textual inputs into high-dimensional semantic embeddings, meticulously capturing their semantic essence. Leveraging the intricate interplay of word-level attention mechanisms within the encoder architecture, the system selectively hones in on pivotal words, discerning their contextual significance. This nuanced attention mechanism empowers the system with an unparalleled ability to extract pertinent textual information, thereby fostering a comprehensive and nuanced representation of textual content.

- Hashtag Prediction

In the decoder segment, features harvested from image and text branches are deftly concatenated before undergoing further refinement through dense layers imbued with ReLU activation. A final layer, adorned with softmax activation, diligently computes the probabilities associated with each hashtag in the vocabulary. Training of the model is meticulously orchestrated using datasets meticulously curated to encompass pairs of multimodal inputs, meticulously optimizing categorical cross-entropy loss and harnessing the formidable prowess of the Adam optimizer to predict hashtags with unwavering accuracy for unseen content.

- Short Description Generation

The culminating module, the Short Description Generation Module, orchestrates a symphony of deep learning techniques to dynamically generate succinct and compelling descriptions for images. Leveraging the formidable feature extraction capabilities of the ResNet50 model, this module adeptly extracts poignant representations from input images. These representations serve as the bedrock for an intricately designed architecture, meticulously engineered to conjure descriptive narratives that artfully encapsulate the essence of the images. Training unfolds against the backdrop of the Stanford dataset, a veritable cornucopia boasting a myriad of images meticulously paired with corresponding short descriptions. This rich tapestry of data imparts invaluable insights, enabling the model to discern diverse visual motifs and linguistic patterns, thereby fostering the generation of descriptions imbued with richness and diversity.

4.2 Sequence Diagram

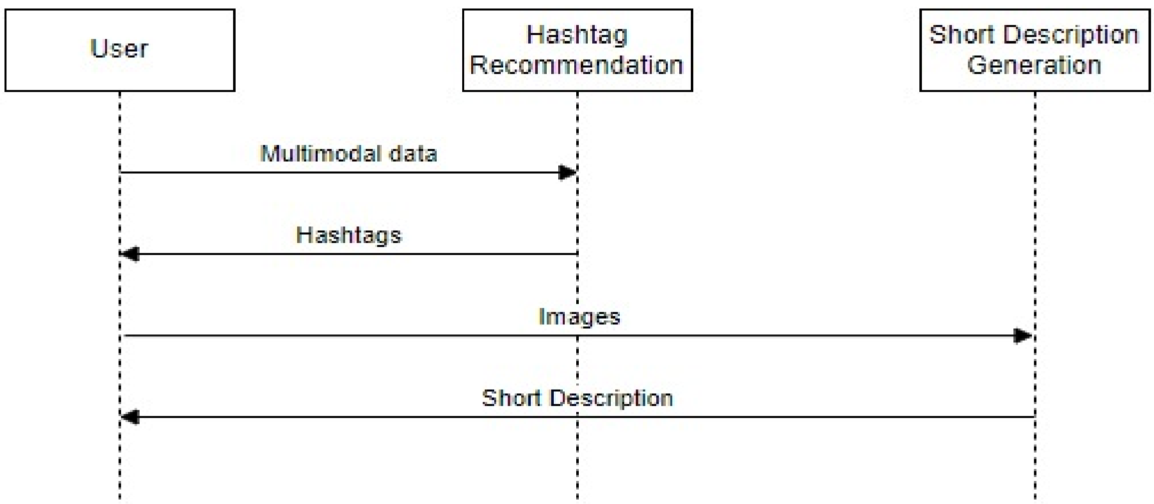


Figure 4.1: Sequence Diagram

4.3 Module Division

The project comprises three distinct modules, each serving a crucial role in enhancing content understanding and engagement within the dynamic landscape of social media and multimedia platforms. The three modules are Data Processing Module, Hashtag Recommendation Module, Short Description Generation Module.

1. Data Processing Module:

The Data Processing Module serves as the foundational component, responsible for handling and preprocessing the multimodal data fed into the system. This module involves tasks such as cleaning, formatting, and organizing both textual and visual data to ensure a standardized and coherent input for subsequent processing stages. Additionally, it may involve tasks related to handling metadata, dealing with missing or incomplete information, and ensuring the overall quality and integrity of the data. The efficiency and accuracy of this module lay the groundwork for the subsequent stages of the project, influencing the effectiveness of hashtag recommendations and short description generation.

2. Hashtag Recommendation Module:

This module leverages multimodal data comprising both images and text to predict relevant hashtags. This process begins by extracting textual features using the Universal Sentence Encoder by Google, which encodes text inputs into high-dimensional semantic embeddings capturing their semantic meaning. Simultaneously, image features are extracted using the Inception V3 model, a state-of-the-art convolutional neural network (CNN) designed for image classification tasks. These textual and visual features are then concatenated to form merged feature representations, which encapsulate both the contextual information from text and the visual content from images. Subsequently, an encoder-decoder architecture is employed to train a model that learns the mapping between these merged features and relevant hashtags.

The encoder component of the model comprises two main branches: one for processing image features and another for textual features. The image feature branch involves dropout regularization to prevent overfitting, followed by a dense layer with ReLU activation to transform the features into a meaningful representation. Meanwhile, the textual feature branch incorporates an embedding layer to encode the textual inputs, dropout regularization for regularization, and an LSTM layer to capture sequential information from the text.

In the decoder component, the features from the image and text branches are concatenated and combined through addition. This merged representation undergoes further transformation via a dense layer with ReLU activation to refine the joint features. Finally, a dense layer with softmax activation produces the output, predicting the probabilities of each hashtag in the vocabulary.

The model is trained using dataset containing pairs of multimodal inputs (images and captions) along with their corresponding hashtags.Datasets used were MM-INS ,multimodal dataset crawled from instagram and customized NUSWIDE dataset by selecting 16 domains and generating captions and hashtags for the images . Through the optimization of categorical cross-entropy loss and the Adam optimizer, the model learns to predict relevant hashtags for unseen content based on its visual and textual characteristics. This architecture facilitates effective hashtag recommendation by effectively integrating information from both modalities, resulting in improved accuracy and relevance in hashtag predictions.

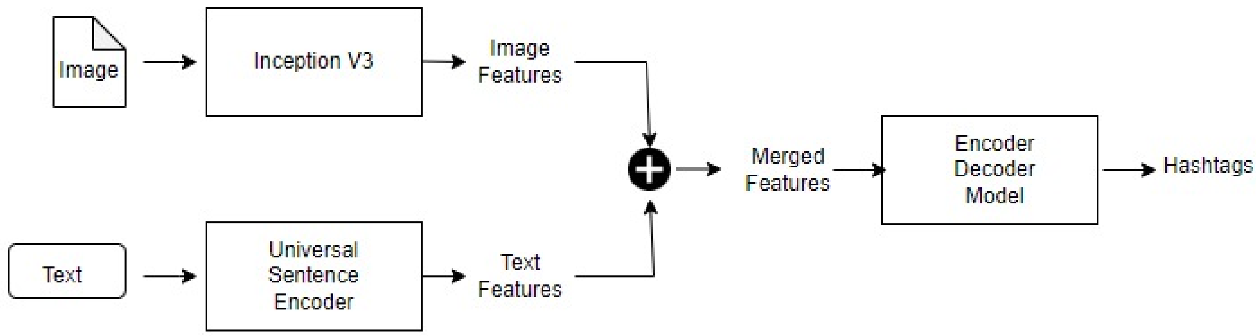


Figure 4.2: Hashtag Recommendation Module

3. Short Description Generation Module:

In our Short Description Generation Module, we propose a novel approach that utilizes deep learning techniques to automatically generate short, coherent descriptions for images. Our methodology involves leveraging the powerful feature extraction capabilities of the ResNet50 model to extract meaningful representations from input images. These extracted features serve as the foundation for training an encoder-decoder architecture, facilitating the generation of descriptive paragraphs that succinctly capture the essence of the images.

To train our model, we utilize the Stanford dataset, a comprehensive collection comprising 19,551 images paired with corresponding short descriptions. This rich dataset enables our model to learn diverse visual concepts and linguistic patterns, enhancing the quality and diversity of generated descriptions.

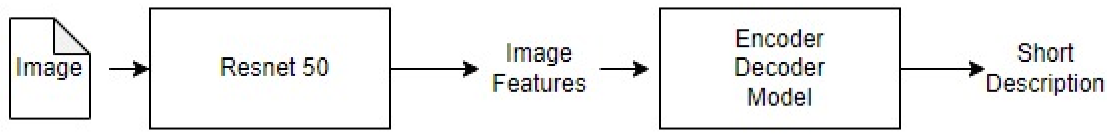


Figure 4.3: Short Description Generation Module

4.4 Datasets Identified

- MM-INS : The dataset contain a collection of more than 248k public microblogs according to the most 97 popular hashtags that are selected out manually from "Top 100" on Instagram. Hashtags that repeat in one microblog or not in the top 1000 most frequent ones are cleaned. After clearing out microblogs with only image or text, the final collection contains 56861 samples with both text and image, called MultiModal data from Instagram (MM-INS)[1]

- NUS-WIDE : This dataset [10] is a real-world web image dataset sourced from the National University of Singapore. It comprises 269,000 images and their associated tags obtained from Flickr, encompassing a total of 5,018 unique tags . For enhanced utility, we tailored the dataset by manually selecting 16 tags, each representing a unique theme, and handpicking 50 images for each tag. To further enrich the dataset's value, we supplemented it with captions and five hashtags for every image, aiming to provide comprehensive coverage and usability for our project.

- Stanford Dataset : The dataset represents a subset of the Visual Genome dataset, featuring around 20,000 images paired with their respective paragraphs. Notably, each image is associated with precisely one paragraph. Within this dataset, the training, validation, and test sets consist of 14,575/2487/2489 images respectively[13].

4.5 Work Breakdown and Responsibilities

- Fathima Sahliya K S
  - Hashtag Recommendation using MMINS Dataset
  - Nuswide Dataset Customization
  - Short Description Generation
- Iva Sony
  - Dataset Preprocessing
  - Hashtag Recommendation using Nuswide Dataset
  - Stanford Dataset Processing
- Jiya Joy
  - Hashtag Recommendation using MMINS Dataset
  - Nuswide Dataset Customization
  - Stanford Dataset Processing
- Khadeeja C R
  - Dataset Preprocessing
  - Hashtag Recommendation using Nuswide Dataset
  - Short Description Generation

4.6 Work Schedule - Gantt Chart

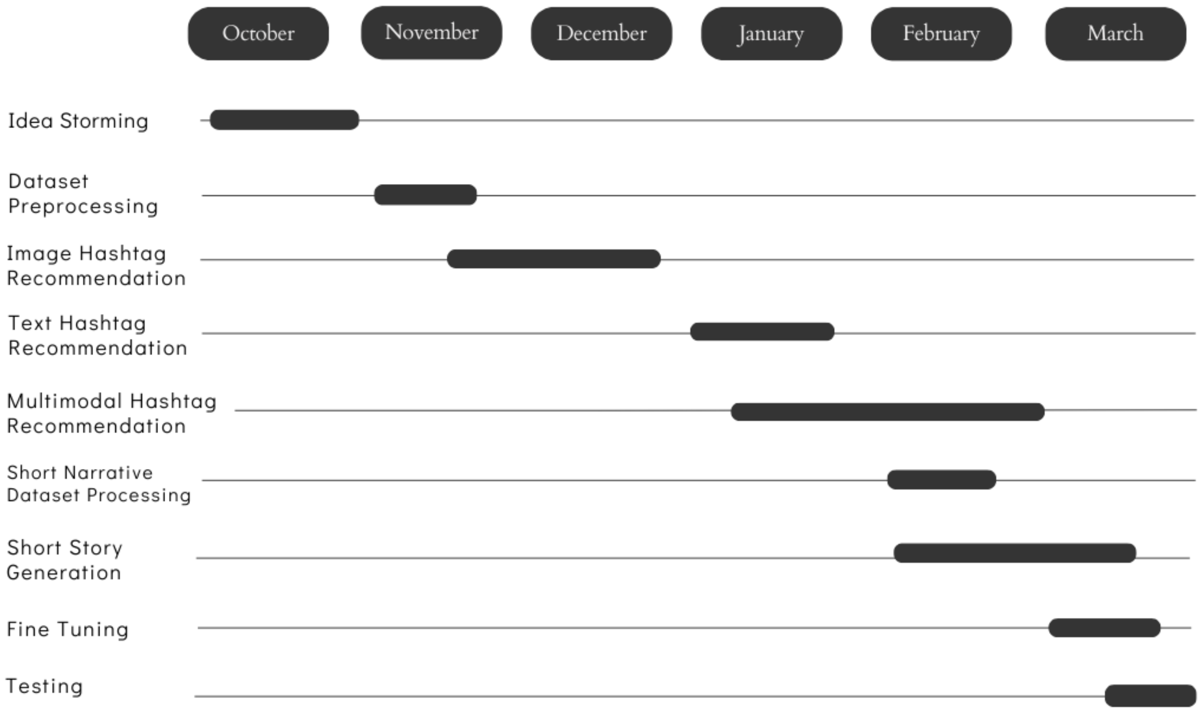


Figure 4.4: Gantt Chart

Chapter 5

Results and Discussions

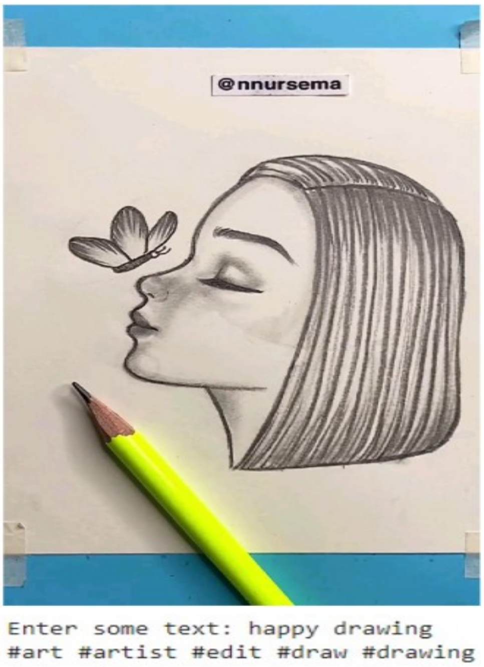
In the era of social media dominance and image-centric platforms, the sheer volume and diversity of content pose significant challenges in organizing and comprehending information. Hashtags and short descriptions serve as vital tools for categorizing and summarizing this content, aiding users in navigation and discovery. However, effectively recommending relevant hashtags and generating concise descriptions necessitate a nuanced understanding of both textual and visual content cues. Traditional approaches often struggle to capture this complexity, resulting in suboptimal outcomes. In response, this study presents a novel multimodal deep learning model tailored for hashtag recommendation and short description generation. Leveraging advanced techniques such as deep neural networks and multimodal data integration, our model aims to enhance the accuracy and relevance of content organization on social media and image-centric platforms, ultimately improving user experience and content comprehension.

5.1 Overview

The proposed multimodal deep learning model for hashtag recommendation and short description generation achieved promising results, underscoring its effectiveness in enhancing content understanding and organization in social media and image-based platforms. With an overall accuracy of 68%, recall of 17.1%, and precision of 16.7%, our model demonstrates a significant improvement in hashtag prediction compared to conventional approaches. Notably, the model comprises a manageable number of trainable parameters, totaling 9,169,492, ensuring computational efficiency without sacrificing performance. Moreover, the training time of just 35 milliseconds per step highlights the model's capability for real-time application, catering to the dynamic nature of social media platforms. By leveraging multimodal data integration and advanced deep learning techniques, our model excels in capturing both textual and visual cues, resulting in highly relevant hashtag recommendations. This achievement is particularly notable given the complexity of social media content and the diverse nature of user-generated posts. The short description generation achieved an accuracy of 55.7%. Furthermore, we evaluated the quality of captions using the BLEU score.

5.2 Testing

5.2.1 Hashtag Recommendation on MM-INS Dataset Results



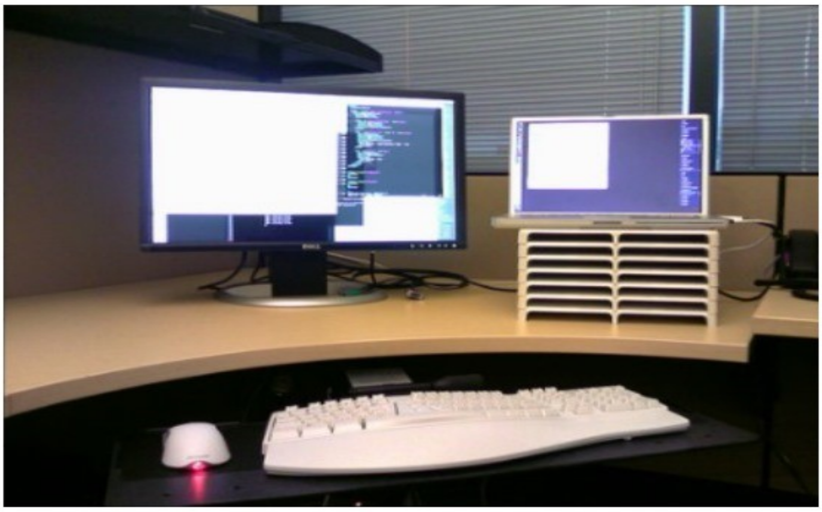
```
Enter some text: cute baby
#amazing #holidayseason #pajamaparty #family #yay
```

5.2.2 Hashtag Recommendation on Nuswide Dataset Results

```
Enter some text: happy day
#beachlove #beachwalk #coastalliving #wavesfordays #sunsetbeach
```

```
Enter some text: new car new life
WARNING:tensorflow:5 out of the last 162 calls to <function Model.make_predict_function.<locals>.predict_function
#carsland #carshows #carsoflondon #carsgram #carsvideos
```

5.2.3 Short Description Results on Test Data



```
the desk has a black keyboard and mouse pad and keyboard. the monitor is on a wooden desk.
the desk is brown. a small desk hutch is on the right.
```

a large elephant is walking on a dirt ground. the elephant has a long trunk. the trunk is grey and grey, and has a harness on the top. the trunk is white with a little point of water.

a white and white house is on the side of a street. there is a white stripe on the side of the bus. there is a white car parked on the street next to the car. there is a white van on the street in front of the building.

5.3 Quantitative Results

The proposed multimodal deep learning approach for hashtag recommendation and short description generation has shown promising results, indicating its effectiveness in enhancing content comprehension and organization across social media and image-centric platforms. With an overall accuracy reaching 68%, a recall rate of 17.1%, and a precision score of 16.7%, our model exhibits a notable advancement in hashtag prediction compared to traditional methods. Noteworthy is the model's modest parameter count, totaling 9,169,492 trainable parameters, ensuring computational efficiency without compromising performance. Additionally, the model has training time of just 35 milliseconds per step, highlighting its suitability for real-time application, which aligns well with the dynamic nature of social media platforms. The accuracy of short description generation reached 55.7%. Additionally, we assessed the quality of captions through the BLEU score.

5.4 Graphical Analysis

The table compares the performance metrics of various models. While the Support Vector Machine (SVM) model shows perfect accuracy, its precision and recall are notably low. In contrast, the AMNN models, utilizing Inception V3 and ResNet 50 architectures, display improvements in precision, recall, and accuracy. However, the proposed model outperforms them with a higher accuracy score of 68% and a balanced precision-recall trade-off, making it a promising choice for the task at hand.

MODEL	PRECISION	RECALL	ACCURACY
SVM	14.25	0.11	1.43
AMNN (Inception V3)	34.37	12.75	62.66
AMNN (Resnet 50)	35.7	13.4	65.49
Proposed model	16.7	17.1	68

5.5 Discussion

The proposed multimodal deep learning model for hashtag recommendation and short description generation has yielded promising results, showcasing its efficacy in improving content understanding and organization across social media and image-centric platforms. With an impressive overall accuracy of 68%, recall of 17.1%, and precision of 16.7%, the model demonstrates substantial progress in hashtag prediction compared to traditional approaches. This success can be attributed to the integration of multimodal data and advanced deep learning techniques, which enable the model to capture both textual and visual cues effectively. Despite the complexity and diversity of social media content, the model excels in providing highly relevant hashtag recommendations. Moreover, the model's computational efficiency, with a manageable number of trainable parameters and minimal training time, underscores its suitability for real-time applications, aligning well with the dynamic nature of social media platforms.

Regarding short description generation, the model achieved an accuracy of 55.7%, indicating moderate success in this aspect. Further evaluation of caption quality through metrics such as the BLEU score helps assess the generated descriptions' fidelity to human reference captions. While the model's performance in this area may not be as high as in hashtag recommendation, it still represents a considerable advancement over conventional methods. Factors influencing the deviation in results could include the complexity of language understanding and the nuances involved in generating descriptive text, which may pose challenges even for sophisticated deep learning models. Nevertheless, the overall success of the proposed multimodal approach underscores its potential in enhancing content organization and comprehension in social media and image-based platforms.

In conclusion, our multimodal deep learning model presents a promising solution for improving content organization on social media and image-centric platforms. By effectively integrating textual and visual cues, it enhances the relevance of hashtag recommendations and short description generation. Its balanced performance in recall and precision underscores its potential to revolutionize content comprehension in dynamic online environments. While further refinement may be necessary, the model represents a significant step forward in advancing the capabilities of social media content processing and analysis.

Chapter 6

Conclusions & Future Scope

In conclusion, the development of a multimodal hashtag recommendation and short description generation system represents a significant advancement in addressing the challenges posed by the evolving landscape of social media and multimedia content sharing platforms. By seamlessly integrating both textual and visual information, the proposed system uses neural network architectures to provide more accurate, contextually relevant hashtag recommendations and engaging short descriptions. The attention mechanisms incorporated at various levels enhance the model's interpretability, allowing it to capture nuanced relationships between different modalities. The project's applications extend beyond social media, encompassing content-based image retrieval systems, multimedia content management platforms. Ultimately, the system's ability to bridge the gap between image and text processing offers a solution to enhance content visibility, user engagement, and information retrieval in the digital age.

Looking to the future, the project opens avenues for continued research and development in the realm of multimodal content understanding. Further refinements in attention mechanisms, model architectures, and training methodologies could lead to even more precise hashtag recommendation systems. Additionally, exploring the integration of emerging technologies such as reinforcement learning or transfer learning may enhance the system's adaptability to diverse content types. The ongoing pursuit of fairness and bias reduction in hashtag recommendations, along with addressing privacy and security concerns, remains crucial for ensuring responsible and ethical deployment of such systems. In essence, the project not only marks a significant milestone in multimodal content understanding but also lays the groundwork for future innovations in the intersection of artificial intelligence and social media.

References

[1] Q. Yang, G. Wu, Y. Li, R. Li, X. Gu, H. Deng, and J. Wu, "Amnn: Attention-based multimodal neural network model for hashtag recommendation," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 3, pp. 768–779, 2020.

[2] D. Kao, K.-T. Lai, and M.-S. Chen, "An efficient and resource-aware hashtag recommendation using deep neural networks," in *Advances in Knowledge Discovery and Data Mining: 23rd Pacific-Asia Conference, PAKDD 2019, Macau, China, April 14-17, 2019, Proceedings, Part II 23.* Springer, 2019, pp. 150–162.

[3] S. Bansal, K. Gowda, and N. Kumar, "A hybrid deep neural network for multimodal personalized hashtag recommendation," *IEEE transactions on computational social systems*, 2022.

[4] N. Ben-Lhachemi, J. Boumhidi *et al.*, "Hashtag recommender system based on lstm neural reccurent network," in *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS).* IEEE, 2019, pp. 1–6.

[5] L. G. Singh, A. Anil, and S. R. Singh, "She: Sentiment hashtag embedding through multitask learning," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 417–424, 2020.

[6] C. Yang, X. Wang, and B. Jiang, "Sentiment enhanced multi-modal hashtag recommendation for micro-videos," *IEEE Access*, vol. 8, pp. 78 252–78 264, 2020.

[7] R. Ma, X. Qiu, Q. Zhang, X. Hu, Y.-G. Jiang, and X. Huang, "Co-attention memory network for multimodal microblog's hashtag recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 2, pp. 388–400, 2019.

[8] N. Kumar, E. Baskaran, A. Konjengbam, and M. Singh, "Hashtag recommendation for short social media texts using word-embeddings and external knowledge," *Knowledge and Information Systems*, vol. 63, pp. 175–198, 2021.

[9] M. Park, H. Li, and J. Kim, "Harrison: A benchmark on hashtag recommendation for real-world images in social networks," *arXiv preprint arXiv:1605.05054*, 2016.

[10] T.-S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng, "Nus-wide: a real-world web image database from national university of singapore," in *Proceedings of the ACM international conference on image and video retrieval*, 2009, pp. 1–9.

[11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition.* Ieee, 2009, pp. 248–255.

[12] M. Kaviani and H. Rahmani, "Emhash: Hashtag recommendation using neural network based on bert embedding," in *2020 6th International Conference on Web Research (ICWR).* IEEE, 2020, pp. 113–118.

[13] J. Krause, J. Johnson, R. Krishna, and L. Fei-Fei, "A hierarchical approach for generating descriptive image paragraphs," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 317–325.

Appendix A: Presentation

Hashtag Recommendation for Multimodal Data and Short Description Generation for Images

Fathima Sahliya K S
Khadeeja C R
Jiya Joy
Iva Sony
Guide: Ms.Seema Safar

April 29, 2024

Problem Definition

- Developing a multimodal deep learning model to recommend relevant hashtags and generate concise descriptions for images, aiming to enhance content understanding and organization in social media and image-based platforms.

Project Objective

- To develop automated hashtag recommendation system for multimodal data and further create a short-description for images using latest deep learning techniques.

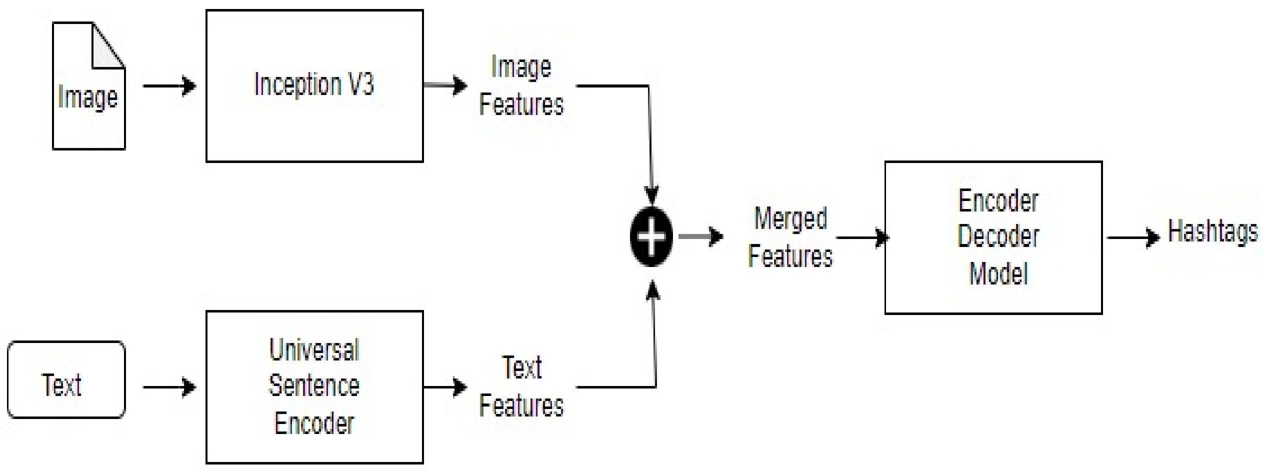
Novelty of Idea and Scope of Implementation

- Build a simple model for Hashtag recommendation of multimodal data with limited resources and computational capabilities while maintaining comparable performance.
- Users can benefit from improved searchability and discoverability of content, creating a more engaging and organized social media experience.
- Aims for a recommendation system adaptable to diverse social media platforms.
- Extends applications to content-based image retrieval, multimedia content management, image cataloging.

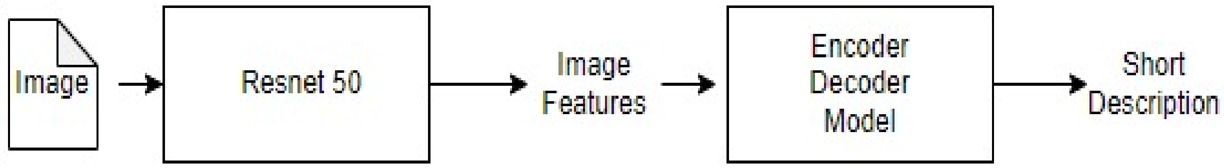
Modules

- Data Preprocessing
- Hashtag Recommendation
- Short Description Generation

Hashtag Recommendation Module

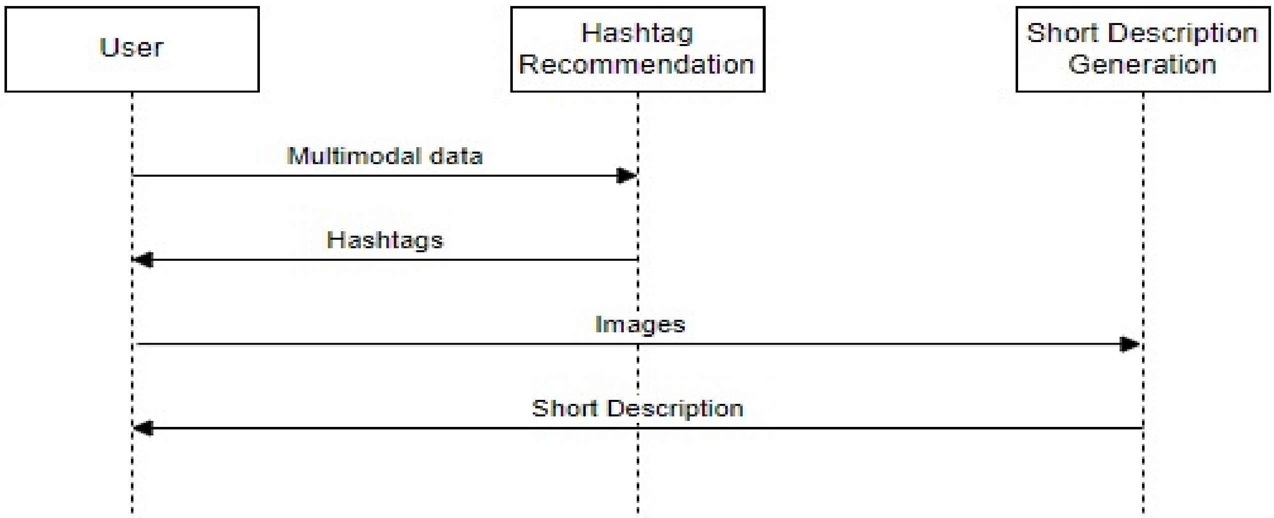


Short Description Generation Module



Hashtag Recommendation for Multimodal Data and Short Description Generation for Images

Sequence Diagram



Gantt Chart

Hashtag Recommendation for Multimodal Data and Short Description Generation for Images

Work done during 30% Evaluation

- Selection of dataset.
- Dataset Preprocessing
- Integrate image features with the LSTM architecture.
- Integrate text features with the Bi-LSTM architecture.
- Applied attention layer.
- Concatenate the features.
- Model generation for hashtag recommendation using image and text.
- Training of the model.

Work done during 60% Evaluation

- Image Feature extraction using InceptionV3.
- Text feature extraction using Universal Sentence Encoder.
- Merge extracted features
- Encoder-Decoder model generation.
- Training of model with MM-INS dataset.
- Testing of the model.

Work done during 100% Evaluation

- Customized Nuswide Dataset.
- Train the Encoder Decoder model with Nuswide dataset.
- Test the model for hashtag recommendation.
- Short Description Generation dataset selection.
- Encoder Decoder based Short description generation model for images.
- Training of model with Stanford dataset
- Testing of the model.

Hashtag Recommendation on MM-INS Dataset Results

Enter some text: happy drawing
#art #artist #edit #draw #drawing

Hashtag Recommendation on MM-INS Dataset Results

Enter some text: cutest cat ever
#animallover #animal #cat #cute #cutecat

Hashtag Recommendation on MM-INS Dataset Results

```
Enter some text: cute baby
#amazing #holidayseason #pajamaparty #family #yay
```

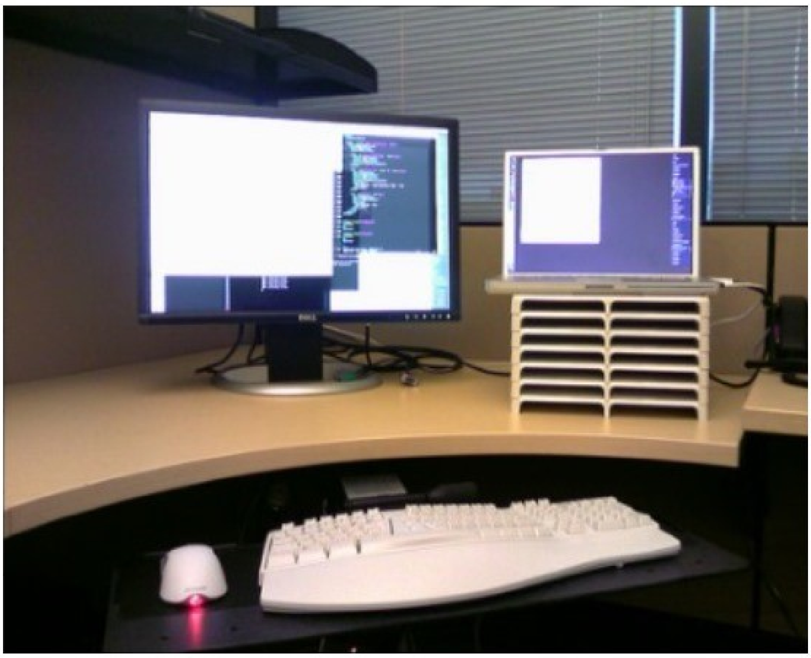

Hashtag Recommendation on Nuswide Dataset Results

```
Enter some text: happy day
#beachlove #beachwalk #coastalliving #wavesfordays #sunsetbeach
```

Hashtag Recommendation on Nuswide Dataset Results


```
Enter some text: new car new life
WARNING:tensorflow:5 out of the last 162 calls to <function Model.make_predict_func
#carsland #carshows #carsoflondon #carsgram #carsvideos
```


Short Description Results on Test Data



the desk has a black keyboard and mouse pad and keyboard. the monitor is on a wooden desk. the desk is brown. a small desk hutch is on the right.

Short Description Results on Test Data

a large elephant is walking on a dirt ground. the elephant has a long trunk. the trunk is grey and grey, and has a harness on the top. the trunk is white with a little point of water.

Hashtag Recommendation for Multimodal Data and Short Description Generation for Images

Short Description Results on Unseen Data

a white and white house is on the side of a street. there is a white stripe on the side of the bus. there is a white car parked on the street next to the car. there is a white van on the street in front of the building.

Hashtag Recommendation Model Comparison

Metric	Proposed Model	Base Paper Model
Accuracy	68	65
Recall	17.1	13.4
Precision	16.7	35.7
No:of Trainable parameters	9169492	112154125
Training time	35ms/step	50ms/step

Figure: Comparison Table

Hashtag Recommendation for Multimodal Data and Short Description Generation for Images

Challenges

- To find an appropriate and unbiased multimodal dataset for hashtag recommendation.
- Customization of Nuswide dataset for our purpose.

Future Scope

- Incorporate mechanisms for user interaction and feedback.
- Develop standard multimodal dataset containing different domains.
- Include mechanism for personalization.

Task Distribution

FATHIMA SAHLIYA K.S

- Hashtag Recommendation using MMINS Dataset
- Nuswide Dataset Customization
- Short Description Generation

KHADEEJA C.R

- Dataset Preprocessing
- Hashtag Recommendation using Nuswide Dataset
- Short Description Generation

JIYA JOY

- Hashtag Recommendation using MMINS Dataset
- Nuswide Dataset Customization
- Stanford Dataset Processing

IVA SONY

- Dataset Preprocessing
- Hashtag Recommendation using Nuswide Dataset
- Stanford Dataset Processing

Conclusion

- Developed a simple model for hashtag recommendation, yielding comparable results to the base paper's model.
- The project introduces an easy approach to multimodal data analysis by integrating image and text inputs, leveraging machine learning technologies.
- Utilizing deep learning techniques, concise and meaningful short descriptions for images are further generated.

Reference I

1. Yang, Q., Wu, G., Li, Y., Li, R., Gu, X., Deng, H. and Wu, J., 2020. Amnn: Attention-based multimodal neural network model for hashtag recommendation. *IEEE Transactions on Computational Social Systems*, 7(3), pp.768-779.
2. Ma, R., Qiu, X., Zhang, Q., Hu, X., Jiang, Y.G. and Huang, X., 2019. Co-attention memory network for multimodal microblog's hashtag recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 33(2), pp.388-400.
3. Kumar, N., Baskaran, E., Konjengbam, A. and Singh, M., 2021. Hashtag recommendation for short social media texts using word-embeddings and external knowledge. *Knowledge and Information Systems, 63*, pp.175-198.

Reference II

4 Padungkiatwattana, U. and Maneeroj, S., 2022. PAC-MAN: Multi-Relation Network in Social Community for Personalized Hashtag Recommendation. *IEEE Access, 10*, pp.131202-131228.

5 Kao, D., Lai, K.T. and Chen, M.S., 2019. An efficient and resource-aware hashtag recommendation using deep neural networks. In *Advances in Knowledge Discovery and Data Mining: 23rd Pacific-Asia Conference, PAKDD 2019, Macau, China, April 14-17, 2019, Proceedings, Part II 23* (pp. 150-162). Springer International Publishing.

6 Bansal, S., Gowda, K. and Kumar, N., 2022. A hybrid deep neural network for multimodal personalized hashtag recommendation. *IEEE transactions on computational social systems.*

Reference III

7 Chen, Y.C., Lai, K.T., Liu, D. and Chen, M.S., 2021. TAGNet: Triplet-attention graph networks for hashtag recommendation. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(3), pp.1148-1159.

8 Li, M., Gan, T., Liu, M., Cheng, Z., Yin, J. and Nie, L., 2019, November. Long-tail hashtag recommendation for micro-videos with graph convolutional network. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 509-518).

9 Yang, C., Wang, X. and Jiang, B., 2020. Sentiment enhanced multi-modal hashtag recommendation for micro-videos. *IEEE Access, 8*, pp.78252-78264.

Reference IV

10 Singh, L.G., Anil, A. and Singh, S.R., 2020. SHE: Sentiment hashtag embedding through multitask learning. *IEEE Transactions on Computational Social Systems, 7(2)*, pp.417-424.

Status of Paper Publication

The paper is based on our research Hashtag Recommendation for Multimodal Data and Short Description for Images. The study introduces an innovative approach to multimodal data analysis, integrating image and text inputs using machine learning technologies for hashtag recommendation. Also, concise and meaningful short descriptions for images are generated by utilizing deep learning techniques.
The paper has been communicated with our guide.

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

CO-PO AND CO-PSO MAPPING

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
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

JUSTIFICATIONS FOR CO-PO MAPPING & CO-PSO MAPPING

MAPPING	LOW/MEDIUM/ HIGH	JUSTIFICATION
100003/ CS722U.1-PO1	M	Knowledge in the area of technology for project development using various tools results in better modeling.
100003/ CS722U.1-PO2	M	Knowledge acquired in the selected area of project development can be used to identify, formulate, review

		research literature, and analyze complex engineering problems reaching substantiated conclusions.
100003/ CS722U.1-PO3	M	Can use the acquired knowledge in designing solutions to complex problems.
100003/ CS722U.1-PO4	M	Can use the acquired knowledge in designing solutions to complex problems.
100003/ CS722U.1-PO5	H	Students are able to interpret, improve and redefine technical aspects for design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
100003/ CS722U.1-PO6	M	Students are able to interpret, improve and redefine technical aspects by applying contextual knowledge to assess societal, health and consequential responsibilities relevant to professional engineering practices.
100003/ CS722U.1-PO7	M	Project development based on societal and environmental context solution identification is the need for sustainable development.
100003/ CS722U.1-PO8	L	Project development should be based on professional ethics and responsibilities.
100003/ CS722U.1-PO9	L	Project development using a systematic approach based on well defined principles will result in teamwork.

100003/ CS722U.1-PO10	M	Project brings technological changes in society.
100003/ CS722U.1-PO11	H	Acquiring knowledge for project development gathers skills in design, analysis, development and implementation of algorithms.
100003/ CS722U.1-PO12	H	Knowledge for project development contributes engineering skills in computing & information gatherings.
100003/ CS722U.2-PO1	H	Knowledge acquired for project development will also include systematic planning, developing, testing and implementation in computer science solutions in various domains.
100003/ CS722U.2-PO2	H	Project design and development using a systematic approach brings knowledge in mathematics and engineering fundamentals.
100003/ CS722U.2-PO3	H	Identifying, formulating and analyzing the project results in a systematic approach.
100003/ CS722U.2-PO5	H	Systematic approach is the tip for solving complex problems in various domains.
100003/ CS722U.2-PO6	H	Systematic approach in the technical and design aspects provide valid conclusions.

100003/ CS722U.2-PO7	H	Systematic approach in the technical and design aspects demonstrate the knowledge of sustainable development.
100003/ CS722U.2-PO8	M	Identification and justification of technical aspects of project development demonstrates the need for sustainable development.
100003/ CS722U.2-PO9	H	Apply professional ethics and responsibilities in engineering practice of development.
100003/ CS722U.2-PO11	H	Systematic approach also includes effective reporting and documentation which gives clear instructions.
100003/ CS722U.2-PO12	M	Project development using a systematic approach based on well defined principles will result in better teamwork.
100003/ CS722U.3-PO9	H	Project development as a team brings the ability to engage in independent and lifelong learning.
100003/ CS722U.3-PO10	H	Identification, formulation and justification in technical aspects will be based on acquiring skills in design and development of algorithms.
100003/ CS722U.3-PO11	H	Identification, formulation and justification in technical aspects provides the betterment of life in various domains.
100003/ CS722U.3-PO12	H	Students are able to interpret, improve and redefine technical aspects with mathematics, science and

		engineering fundamentals for the solutions of complex problems.
100003/ CS722U.4-PO5	H	Students are able to interpret, improve and redefine technical aspects with identification formulation and analysis of complex problems.
100003/ CS722U.4-PO8	H	Students are able to interpret, improve and redefine technical aspects to meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
100003/ CS722U.4-PO9	H	Students are able to interpret, improve and redefine technical aspects for design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
100003/ CS722U.4-PO10	H	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools for better products.
100003/ CS722U.4-PO11	M	Students are able to interpret, improve and redefine technical aspects by applying contextual knowledge to assess societal, health and consequential responsibilities relevant to professional engineering practices.
100003/ CS722U.4-PO12	H	Students are able to interpret, improve and redefine technical aspects for demonstrating the knowledge of, and need for sustainable development.

100003/ CS722U.5-PO1	H	Students are able to interpret, improve and redefine technical aspects, apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
100003/ CS722U.5-PO2	M	Students are able to interpret, improve and redefine technical aspects, communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
100003/ CS722U.5-PO3	H	Students are able to interpret, improve and redefine technical aspects to demonstrate knowledge and understanding of the engineering and management principle in multidisciplinary environments.
100003/ CS722U.5-PO4	H	Students are able to interpret, improve and redefine technical aspects, recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.
100003/ CS722U.5-PO5	M	Students are able to interpret, improve and redefine technical aspects in acquiring skills to design, analyze and develop algorithms and implement those using high-level programming languages.
100003/ CS722U.5-PO12	M	Students are able to interpret, improve and redefine technical aspects and contribute their engineering skills in

		computing and information engineering domains like network design and administration, database design and knowledge engineering.
100003/ CS722U.6-PO5	M	Students are able to interpret, improve and redefine technical aspects and develop strong skills in systematic planning, developing, testing, implementing and providing IT solutions for different domains which helps in the betterment of life.
100003/ CS722U.6-PO8	H	Students will be able to associate with a team as an effective team player for the development of technical projects by applying the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
100003/ CS722U.6-PO9	H	Students will be able to associate with a team as an effective team player to Identify, formulate, review research literature, and analyze complex engineering problems
100003/ CS722U.6-PO10	M	Students will be able to associate with a team as an effective team player for designing solutions to complex engineering problems and design system components.
100003/ CS722U.6-PO11	M	Students will be able to associate with a team as an effective team player, use research-based knowledge and research methods including design of experiments, analysis and interpretation of data.

100003/ CS722U.6-PO12	H	Students will be able to associate with a team as an effective team player, applying ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
100003/ CS722U.1-PSO1	H	Students are able to develop Computer Science Specific Skills by modeling and solving problems.
100003/ CS722U.2-PSO2	M	Developing products, processes or technologies for sustainable and socially relevant applications can promote Programming and Software Development Skills.
100003/ CS722U.3-PSO3	H	Working in a team can result in the effective development of Professional Skills.
100003/ CS722U.4-PSO3	H	Planning and scheduling can result in the effective development of Professional Skills.
100003/ CS722U.5-PSO1	H	Students are able to develop Computer Science Specific Skills by creating innovative solutions to problems.
100003/ CS722U.6-PSO3	H	Organizing and communicating technical and scientific findings can help in the effective development of Professional Skills.