AI-BASED IDENTIFICATION AND CORRECTION OF INAPPROPRIATE LANGUAGE

A Project Report (Phase I)-ECE18R498

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

of

Bachelor of Technology

in

ELECTRONICS & COMMUNICATION ENGINEERING

by

ALLAKA HARI SANKAR	**(9920005003)**
KODE VIVEK SRINIVAS	**(9920005038)**
SUNDU LEELA KRISHNA	**(9920005058)**
MD. GHOUSEE ALI SIDDIQ	**(9920005182)**

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

SCHOOL OF ELECTRONICS, ELECTRICAL AND BIOMEDICAL TECHNOLOGY

KALASALINGAM ACADEMY OF RESEARCH AND EDUCATION

(Deemed to be University)

KRISHNANKOIL-626126, TN (INDIA)

April - 2024

KALASALINGAM ACADEMY OF RESEARCH AND EDUCATION
(Deemed to be University)

SCHOOL OF ELECTRONICS, ELECTRICAL AND BIOMEDICAL TECHNOLOGY
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BONAFIDE CERTIFICATE

Certified that this project report on **"AI BASED IDENTIFICATION AND CORRECTION OF INAPPROPRIATE LANGUAGE"** is the work of **"A. HARI SANKAR (9920005003), K. VIVEK SRINIVAS (9920005038), S. LEELA KRISHNA (9920005058), MD. GHOUSEE ALI SIDDIQ (9920005182)"**

who carried out the project work under my supervision.

SIGNATURE
Dr. V. Muneeswaran, M.E., Ph.D.,
Supervisor
Associate Professor
Department of ECE
Kalasalingam Academy of
Research and Education
Krishnankoil-626126.

SIGNATURE
Dr.A. Muthukumar, M.E., Ph.D.,
Head of the Department
Department of ECE
Kalasalingam Academy of
Research and Education
Krishnankoil-626126

Project Final Review Viva-voce held on ……

Internal Examiner

External Examiner

MAJOR DESIGN EXPERIENCE INFORMATION

Student Details	:	ALLAKA HARI SANKAR KODE VIVEK SRINIVAS SUNDU LEELA KRISHNA MOHAMMAD GHOUSEE ALI SIDDIQ
Project Supervisor	:	DR.V. MUNEESWARAN
Project Title	:	AI-BASED IDENTIFICATION AND CORRECTION OF INAPPORIATIVE LANGUAGE
Program Concentration Area	:	Artificial Intelligence and Machine learning
Subject(s) as Pre-requisite	:	ECE18R379- Deep Learning and Implementation with Keras, ECE18R375- Statistical Inference and Machine Learning
Constraints	:	Data Quality and Diversity Algorithmic Challenges Computational Resources
Project Related to	:	Social Impact and Cultural Sensitivity
Standard(s) used in this project	:	IEEE- P3168 Standard for Robustness Evaluation Test Methods for a Natural Language Processing Service that uses Machine Learning
Sustainable Development Goal indicators	:	SDG 4.7

DECLARATION BY THE STUDENTS

We certify that,

a. the work "AI BASED IDENTIFICATION AND CORRECTION OF INAPPROPRIATE LANGUAGE" contained in this B.Tech. project report is original and authentic and has been done by us under the guidance of my supervisor Dr. V. Muneeswaran, Associate Professor, **ECE Department** during a period from November 2023 to April 2024.
b. the work has not been submitted to any other Institute for any degree or diploma.
c. We have followed the guidelines provided by the department in preparing the report.
d. We have conformed to the norms and guidelines given in the Ethical Code of Conduct of the University.
e. Whenever we have used materials (data, theoretical analysis, figures, and text) from other sources, we have given due credit to them by citing them in the text of the report and giving their details in the references. Further, we have taken permission from the copyright owners of the sources, whenever necessary.

The matter presented in this project report has not been submitted by us for the award of any other degree elsewhere. It is submitted by us in partial fulfillment of the requirements for the award of the **Bachelor of Technology in Electronics & Communication Engineering** to the Department of Electronics & Communication Engineering, Kalasalingam Academy of Research and Education (Deemed to be University) Tamilnadu.

Project Final Review Viva-voce held on

Signature of Candidate(s)

A.HARI SANKAR (9920005003)

K.VIVEK SRINIVAS (9920005038)

S.LEELA KRISHNA (9920005058)

MD.G.ALI SIDDIQ (99200050182)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date:

Signature of Supervisor(S)
DR.V.MUNEESWARAN
Associate Professor-ECE

Head of the Department
Department of Electronics & Communication Engineering,
Kalasalingam Academy of Research and Education (Deemed to be University) TN

ACKNOWLEDGEMENT

We are indebted to Kalasalingam University Founder and Chairman "Kalvivallal" **Thiru T. Kalasalingam**, "Illayavallal" **Dr. K. Sridharan**, Chancellor, **Dr. Shasi Anand**, Vice President, **Dr. S. Narayanan**, Vice-Chancellor and **Dr. V. Vasudevan**, Registrar, Kalasalingam Academy of Research and Education for funding my project work.

In preparing this report, we were in contact with many people, researchers, academicians, and practitioners. They have contributed to our understanding and thoughts. We wish to express our sincere appreciation to our supervisor, **Dr. V. MUNEESWARAN**, Associate Professor-ECE for encouragement, guidance, criticism, and friendship. We are also very thankful to **Dr. A. MUTHUKUMAR**, HoD- ECE, and Project Coordinator **Dr. RAJAKRISHNAMOORTHI**, Professor-ECE for his guidance, advice, and motivation. Without his continued support and interest, this project report would not have been the same as presented here. We wish to express our sincere thanks also to Nanochip Solutions, Bangalore for providing the problem statement and to the industrial experts **Mr. MANJUNATH**, **Mr. MANORANJAN,** and **Ms. VIJAYASHREE** for providing technical inputs for the implementation of the project. Librarians at Kalasalingam Academy of Research and Education also deserve special thanks for their assistance in supplying the relevant literature.

We thank all the teaching and non-teaching faculty of the ECE department for their help in completing this project work. We are also grateful to all my family members, friends, and others who have aided us on various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space.

ABSTRACT

In the moment's digital age, the proliferation of online communication has led to an increased need for automated systems to identify and combat unhappy language and detest speech, and offensive content. This study explores the effectiveness of four distinct machine learning algorithms – Decision Trees, Random Forest, Long Short-Term Memory(LSTM), and BERT – in predicting whether a given text contains offensive language. Decision Trees and Random Forest are traditional machine knowledge ways that calculate on structured point engineering, while LSTM and BERT represent slice-edge deep knowledge styles suitable for handling unstructured text data. By comparing these algorithms, this disquisition aims to give perceptivity into their separate strengths and sins in addressing the evolving challenges of online content temperance. Our findings reveal the delicacy, perfection, recall, and F1-score of each model, slipping light on their capability to discern unhappy language. ultimately, this study contributes to the development of robust AI-d- predicated content filtering systems, abetting platforms, and social networks in maintaining safe and dutiful online surroundings by automatically relating and flagging offensive text. analogous systems have the eventuality to foster further inclusive and responsible digital communities, icing a healthier online discourse for all stoners

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

LSTM	-	Long short-term memory
DT	-	Decision tree
RF	-	Random forest

CHAPTER 1

INTRODUCTION

1.1 Problem Background

In a period defined by digital communication, the rise of online content poses a growing challenge. This study investigates the performance of four different machine learning algorithms, slipping light on their capability to describe offensive language.
By comparing traditional styles like Decision Trees and Random Forests with slice-edge approaches like LSTM and BERT, we aim to advance the development of robust AI-d- predicated content filtering systems. These systems have the power to foster safer and more inclusive digital communities, making online discourse healthier for everyone.

1.2 Problem Statement

In today's digital generation, there are lots of people who knowingly or unknown they are using inappropriate language, hate speech, and offensive content. This leads to a challenge for all online platforms to create a safe and respectful online environment. This issue evaluates the effectiveness of machine learning algorithms in identifying offensive language, providing valuable insights for content moderation

1.3 Project Goals and Objectives

This is to estimate and compare the effectiveness of four distinct machine learning algorithms, Decision Trees, Random Forest, Long Short-Term Memory (LSTM), and BERT, in predicting the presence of offensive language in online text. Through the analysis of delicacy, perfection, recall, and F1-score criteria, the study aims to give perceptivity into the strengths and sins of these algorithms for content temperance purposes. ultimately, the design seeks to contribute to the development of AI- -predicated content filtering systems that can help online platforms create safer and more dutiful digital surroundings by automatically detecting and flagging offensive text.

CHAPTER 2

LITERATURE REVIEW

2.1 Related Work

(1)"Vaticination of Twitter Communication omission" Author Alisa Gazizullina; Manuel Mazzara Published Year 2019

Social media platforms like Twitter have become important tools for individuals to establish their character and promote ideas on a global scale. Twitter, distinct from other social media sources, is famed for its capacity to circulate real-time textual information. It serves as a vital platform for participating opinions, ideas, and breaking news through brief and compelling statements that reach millions worldwide. still, the immense reach of Twitter also poses implicit pitfalls, as posting unhappy content can negatively affect the public image and sequestration of individualities, including celebrities, politicians, and ordinary druggies. This paper aims to address the need for visionary measures to cover druggies' individualities and reports on Twitter by fastening the automatic identification of potentially vulnerable dispatches. Our primary idea is to develop a classifier able to prognosticate whether a stoner is likely to cancel a particular post in the future. To achieve this, we employ intermittent Neural Networks (RNNs), which influence environment-grounded information within tweets for the effective bracket. also, our exploration contributes to the field by constructing an expansive set of features, including Twitter metadata, stoner information, and chitter textbook, to train classical machine learning algorithms on Twitter data. This work trials to enhance our understanding of post omission geste on Twitter and offers precious perceptivity into securing druggies' online individualities. 2)" Detest Speech & Offensive Language Detection Using ML & NLP" Author V V S Sasank; Dory Ratna Harshitha Adidela Published Year 2022. To restore peace and harmony in this cross-cultural Internet period, it's of utmost significance for every citizen to bear and spread brotherhood. Under the given circumstances of 5G elaboration citizens have taken their part in the internet veritably seriously thereby utmost of the

(2)" Detest Speech & Offensive Language Detection Using ML & NLP" Author V V S Sasank; Dory Ratna Harshitha Adidela Published Year 2022.

To restore peace and harmony in this cross-cultural Internet period, it's of utmost significance for every citizen to bear and spread brotherhood. Under the given circumstances of 5G elaboration citizens have taken their part on the internet veritably seriously thereby utmost of the netizens spend their time condemning, judging, and trolling other netizens, public numbers for that matter. Because of the consequences in an unbiased society involving race, gender, or religion, the challenge of automatically detecting hate speech and reprehensible language in social media material is critical. still, exploration in this field is substantially concentrated on several languages, which limits its applicability to certain groups. The use of harsh language on social media platforms, as well as the consequences that this has, has become a serious problem in ultramodern culture. Automatic ways to fete and deal with this kind of content are necessary due to the large volume of content produced every day. Machine Learning & Natural Language processing have slice-edge algorithms and classifiers that have advantaged humanity in insolvable ways. Hence, our trouble in this design is to make use of this impeccable technology to produce an effective system that automatically detects hate speech and obnoxious language from the Twitter dataset.

3) Rodrigues,L.F., Naldi,M.C., & Mari,J.F.(2020). Comparing convolutional neural networks and preprocessing ways for HEp- 2 cell bracket in immunofluorescence images. Computers in biology and drug, 116, 103542.

surroundings and conditions. Tracking and understanding changes in modules and connections in a software design is difficult, but indeed more so when the software goes through several types of changes. The typical complexity and size of software also make it harder to grasp software elaboration patterns. In this paper, we present an interactive matrix-ground visualization fashion that, combined with vitality, depicts how software designs evolve. For illustration, it shows which new modules and couplings are added and removed over time. Our general visualization supports dynamic and weighted digraphs and is applied in the environment of software elaboration. assaying source law changes is important to determine the software's structural association and identify quality issues over time. To demonstrate our approach, we explore open-source depositories and bandy

some of our findings regarding these evolving software designs Automated image brackets have become a critical element in colorful medical and natural operations, enabling the effective analysis of complex visual data. One similar area of significance is the bracket of HEp- 2 (mortal Epithelial type 2) cell patterns in immunofluorescence images, which plays a pivotal part in diagnosing autoimmune conditions. The study presented in" Comparing convolutional neural networks and preprocessing ways for HEp-2 cell bracket in immunofluorescence images" by Rodrigues, Naldi, and Mari (2020) addresses this important sphere. Immunofluorescence imaging offers precious perceptivity into cellular structures and can help in the identification of distinct patterns associated with autoimmune conditions. The paper explores the operation of Convolutional Neural Networks (CNNs), an important deep literacy fashion, and preprocessing ways to enhance the delicacy of the HEp- 2 cell pattern bracket. This exploration aims to contribute to the advancement of automated individual tools, which can potentially revise the effectiveness and delicacy of autoimmune complaint discovery. In this prolusion, we give an overview of the significance of the HEp- 2 cell pattern type, illuminate the connection of CNNs, and emphasize the study's ideal of comparing these neural networks with preprocessing ways. The paper's findings hold the pledge of perfecting medical opinion and treatment through the integration of slice-edge technology into the field of immunofluorescence image analysis.

4) **Wang, Y., & Shan, S (2021). Accurate complaint discovery quantification of iris-predicated retinal images using arbitrary imputation image classifier fashion. Microprocessors and Microsystems, 80, 103350.**

The paper named" Accurate Disease Detection Quantification of Iris- predicated Retinal Images Using Random Recrimination Image Classifier fashion" by Wang and Shan, published in Microprocessors and Microsystems in 2021, addresses a critical aspect of medical image analysis. The study focuses on enhancing complaint discovery and quantification in retinal images through an innovative approach known as the Random Recrimination Image Classifier fashion. Retinal conditions pose significant risks to vision and overall optic health, making early and precise discovery vital for effective treatment and operation. Traditional styles constantly encounter challenges in achieving high delicacy and responsibility in complaint discovery from retinal images. To address this

issue, the authors introduce a new methodology predicated on the Random Recrimination Image Classifier fashion. This paper not only presents an advanced fashion but also provides perceptivity into the implicit operations and advancements in the field of medical image analysis. The disquisition aims to contribute to the development of more accurate and effective tools for the early opinion and quantification of retinal conditions.

5) Mitra, A (2020). Sentiment Analysis Using Machine Learning Approaches (dictionary predicated on movie review dataset). Journal of Ubiquitous Computing and Communication Technologies (UCCT), 2 (03), 145- 152.

In the ever-evolving terrain of natural language processing and sentiment analysis, Mitra's 2020 paper named" Sentiment Analysis Using Machine Learning Approaches (dictionary predicated on movie review dataset)" published in the Journal of Ubiquitous Computing and Communication Technologies (UCCT) delves into the operation of machine knowledge ways for sentiment analysis. Sentiment analysis, the process of sapient and quantifying emotional tone within text data, has gained elevation due to its wide-ranging operations, from request disquisition to social media monitoring. Mitra's disquisition focuses specifically on sentiment analysis using a dictionary-predicated approach, using a dataset comprised of movie reviews. The choice of the dataset is significant as movie reviews constantly contain rich and nuanced expressions of sentiment, making them a precious resource for studying sentiment analysis methodologies. The paper is set against the background of the adding connection of sentiment analysis in the digital age, where understanding public opinion and sentiment on various motifs is vital. By employing machine knowledge ways and a specialized dictionary, the study seeks to contribute to the advancement of sentiment analysis methodologies, slipping light on the practical operations of analogous ways in the realm of movie reviews and potentially extending their avail to broader disciplines of text analysis. This paper's findings hold a connection for researchers, practitioners, and suckers interested in the ever-expanding field of sentiment

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Existing system

System for online content temperance generally relies on traditional machine literacy ways similar to Decision Trees and Random Forest to identify and combat unhappy language and descent content. still, while these styles have shown some effectiveness, they may struggle with the nuanced and environment-dependent nature of obnoxious language. As a result, there's a growing need to incorporate advanced deep literacy styles like Long Short- Term Memory (LSTM) to handle unshaped textbook data and ameliorate the delicacy of content filtering systems

3.2 Disadvantages:

1. Limited environment Understanding: Struggles to grasp nuanced and environment-dependent obnoxious language, leading to false cons or negatives.

2. Homemade point Engineering: Requires labor- ferocious point engineering, making it less adaptable to evolving online language trends.

3. Scalability Issues: May not gauge effectively to handle the vast quantum of stoner-generated content in real time.

4. Difficulty with Multimodal Data: shy for content that combines textbook with images, audio, or videotape.

5. Model Interpretability: Lack of translucency in model decision- timber can hamper explaining and perfecting temperance issues.

3.3 Proposed System

The system leverages advanced machine learning algorithms, including Decision Trees, Random Forest, and BERT, to effectively identify and combat obnoxious language in online content. Through rigorous evaluation, this exploration aims to assess the strengths

and sins of these algorithms in addressing the challenges of content temperance. By furnishing perceptivity into their delicacy, perfection, recall, and F1- score, this study contributes to the development of robust AI-- AI-grounded content filtering systems, fostering safer and further regardful digital communities.

3.4 Advantages

1 **Decision Trees:** Transparent, interpretable, and stoner-friendly for understanding and conforming content temperance criteria.

2 **Random Forest**: Robust ensemble system that reduces overfitting and improves delicacy in relating obnoxious content.

3 **BERT:** Contextual understanding enables nuanced discovery, reducing false cons and perfecting content filtering delicacy.

4 **Comprehensive perceptivity**: Evaluation criteria give a holistic view of algorithm performance, abetting nonstop refinement.

5 **Safer Digital Communities**: These algorithms inclusively enhance content temperance, contributing to regardful and inclusive online spaces.

3.5 work Flow of Proposed system

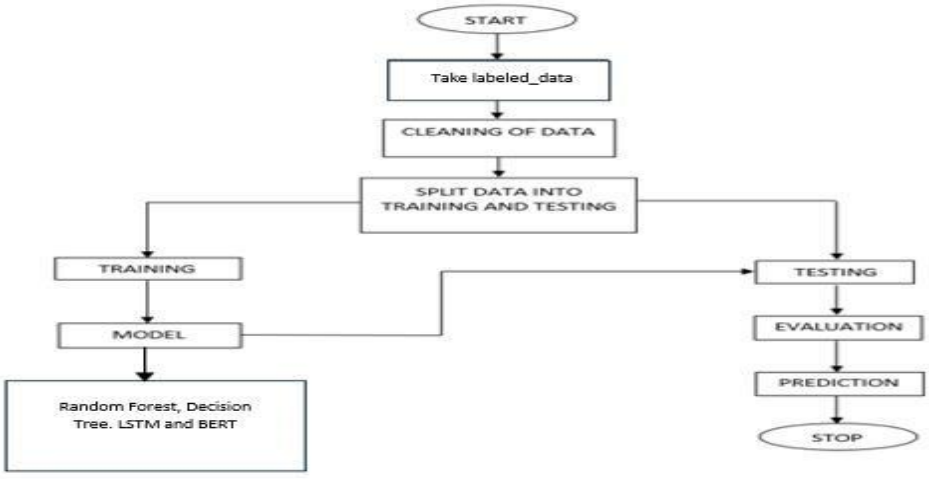


Figure 3.1 Workflow of Proposed System

CHAPTER 4

STANDARDS AND PROTOCOLS INVOLVED IN THE PROPOSED WORK

4.1 Functional and non-functional conditions

Requirement analysis is a very critical process that enables the success of a system or software design to be assessed. Conditions are generally resolved into two types Functional and non-functional conditions. All these functionalities need to be inescapably incorporated into the system as a part of the contract. These are represented or stated in the form of input to be given to the system, the operation performed and the affair anticipated. They're principally the conditions stated by the stoner which one can see directly in the final product, unlike the non-functional conditions.

Examples of functional conditions:

1) Authentication of stoner whenever he/she logs into the system

2) System arrestment in case of a cyber-attack

3) A verification dispatch is transferred to the stoner whenever he/ she registers for the first time on some software system.

Non-functional conditions: These are principally the quality constraints that the system must satisfy according to the design contract. The precedence or extent to which these factors are enforced varies from one design to another. They're also called nonbehavioral requirements.

They principally deal with issues like:

- Portability
- Security
- Maintainability
- trustability
- Scalability
- Performance
- Reusability
- flexibility

Examples of non-functional conditions:

1) Emails should be transferred with a quiescence of no less than 12 hours from such an exertion.

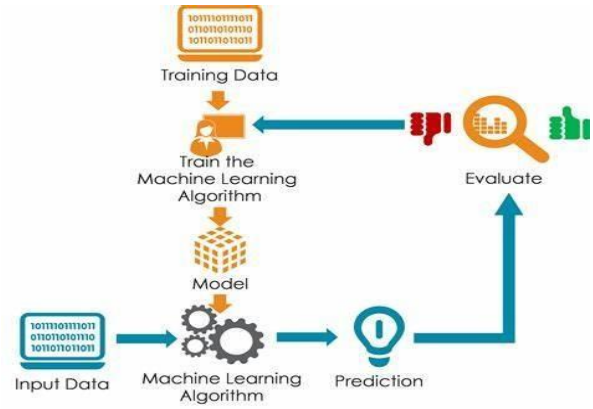
2) The processing of each request should be done within 10 seconds

3) The point should load in 3 seconds whenever contemporaneous druggies are> 10000

4.2 Hardware Requirements

Processor	**- I7/ Intel Processor**
Hard Disk	- 160 GB
Key Board	- Standard Windows Keyboard
Mouse	- Two or Three-Button Mouse
Monitor	- SVGA
RAM	- 8 GB

4.3 Architecture



4.1 Architecture

4.4 Software Requirements

Operating System	-Windows 11
Garçon side Script	-HTML, CSS & JS
Programming Language	-Python
Libraries	-Django, Pandas, Numpy
IDE/ Workbench	-PyCharm
Technology	-Python3.6

CHAPTER 5

SYSTEM DESIGN

5.1 Introduction of Input Design:

In design, the formulators must consider the input bias analogous to PC, MICR, OMR, etc. Well-designed input forms and defenses have the following parcels

• It should serve a specific purpose effectively analogous to storing, recording, and repossessing the information.

• It ensures proper completion with delicacy.

• It should concentrate on the user's attention, consistency, and simplicity.

• All these objects are attained using the knowledge of introductory design principles regarding –

o What are the inputs demanded for the system?

o How end stoners respond to different rudiments of forms and defenses.

objectives for Input Design:

The objects of input design are –

• To design data entry and input procedures

• To reduce input volume

• To design source documents for data internee or contrive other data internee styles

• To design input data records, data entry defenses, user interface defenses, etc.

• To use evidence checks and develop effective input controls.

Output Design:

The design of affairs is the most important task of any system. During affair design, formulators identify the type of labor demanded and consider the necessary affair controls and prototype report layouts.

Objectives of Output Design:

The objects of input design are

• To develop affair design that serves the conscious purpose and eliminates the product of unwanted affairs.

• To develop the affair design that meets the end user's conditions.

• To deliver the applicable volume of affairs.

• To form the affair in an applicable format and direct it to the right person.

• To make the affair available on time for making good opinions.

5.2 UML Diagrams

5.2.1 Use Case Diagram:

A use-case illustration in the Unified Modeling Language(UML) is a type of behavioral illustration defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality handed by a system in terms of actors, their pretensions(represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. The roles of the actors in the system can be depicted

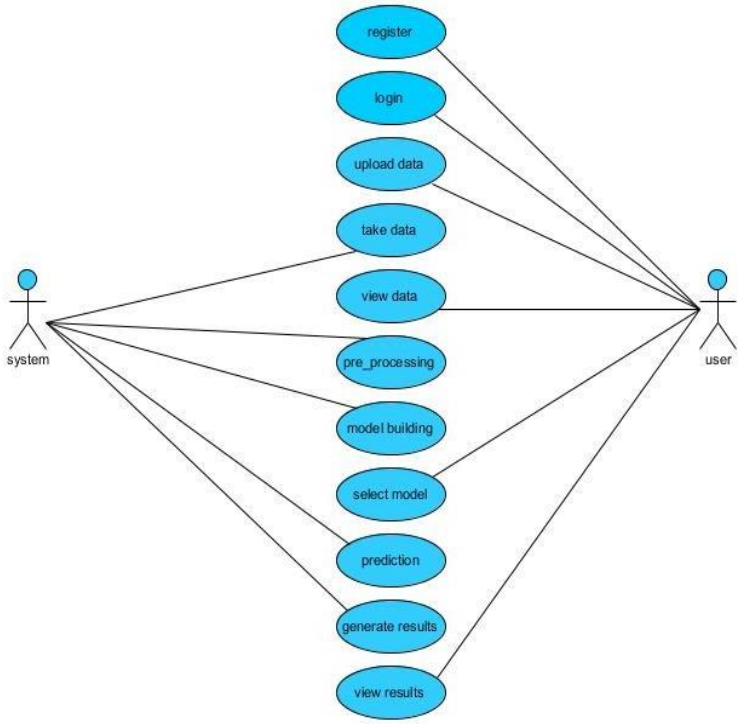


Figure5.1 Use Case Diagram

5.2.2 Class Diagram:

In software engineering, a class illustration in the Unified Modelling Language(UML) is a type of stationary structure illustration that describes the structure of a system by showing the system's classes, attributes, operations(or styles), and the connections among the classes. It explains which class contains information.

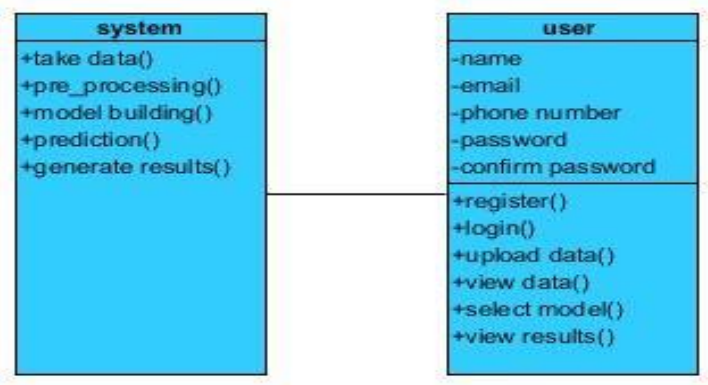


Figure5.2 Class Diagram

5.2.3 Sequence Diagram:

A sequence illustration in Unified Modelling Language(UML) is a kind of commerce illustration that shows how processes operate with one another and in what order. Sequence plates are occasionally called event plates, event scripts, and timing plates.

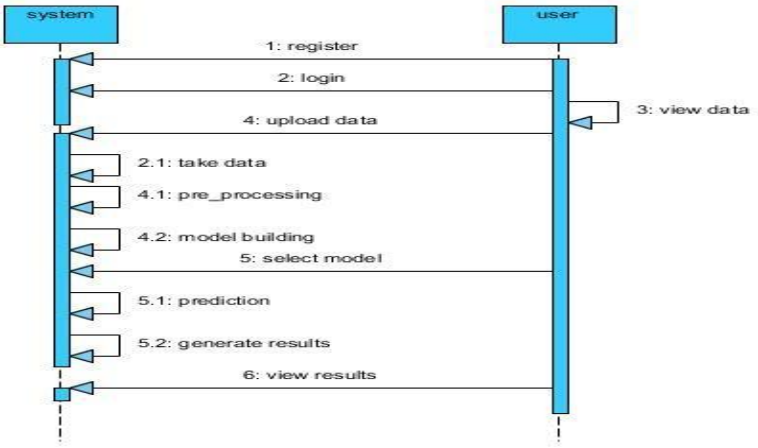


Figure5.3 Sequence Diagram

5.2.4 Collaboration Diagram:

In collaboration illustration, the system call sequence is indicated by some numbering fashion as shown below. We've taken the same order operation system to describe the collaboration illustration. we have taken the same order management system to describe the collaboration diagram

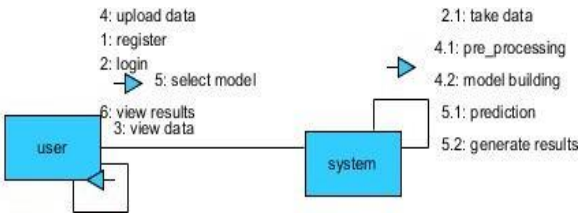


Figure5.4 Collaboration Diagram

5.2.5 Deployment Diagram:

Deployment illustration represents the deployment view of a system. It's related to the element illustration. Because the factors are stationed using the deployment plates. A deployment illustration consists of bumps. Bumps are nothing but physical tackles used to emplace the operation.

Figure5.5 Deployment Diagram

5.2.6 Activity Diagram:

Activity plates are graphical representations of workflows of accretive conditioning and conduct with support for choice, replication, and concurrency. In the Unified Modelling Language, exertion plates can be used to describe the business and functional step-by-step workflows of factors in a system.

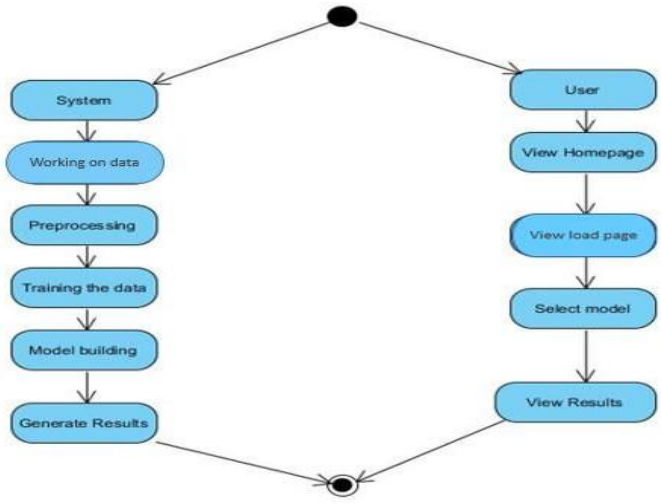


Figure5.6 exertion Diagram

5.2.7 Element Diagram:

An element illustration, also known as a UML element illustration, describes the association and wiring of the physical factors in a system. element plates are frequently drawn to help model perpetration details and double-check that every aspect of the system's needed functions is covered by planned development.

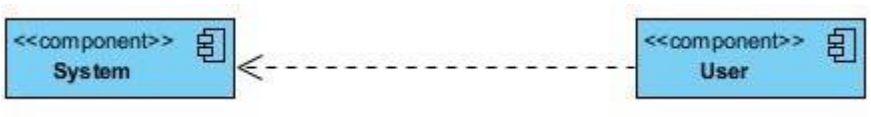


Figure5.7 Element Diagram

5.2.8 ER Diagram:

A reality–relationship model(ER model) describes the structure of a database with the help of an illustration, which is known as Entity Relationship Diagram(ER Diagram). An ER model is a design or design of a database that can latterly be enforced as a database. The main factors of the E-R model are the reality set and the relationship set. A reality set is a group of analogous realities and these realities can have attributes. In terms of DBMS, reality is a table or trait of a table in the database, so by showing the relationship among tables and their attributes, ER illustration shows the complete logical structure of a database.

Let’s have a look at a simple ER illustration to understand this conception.

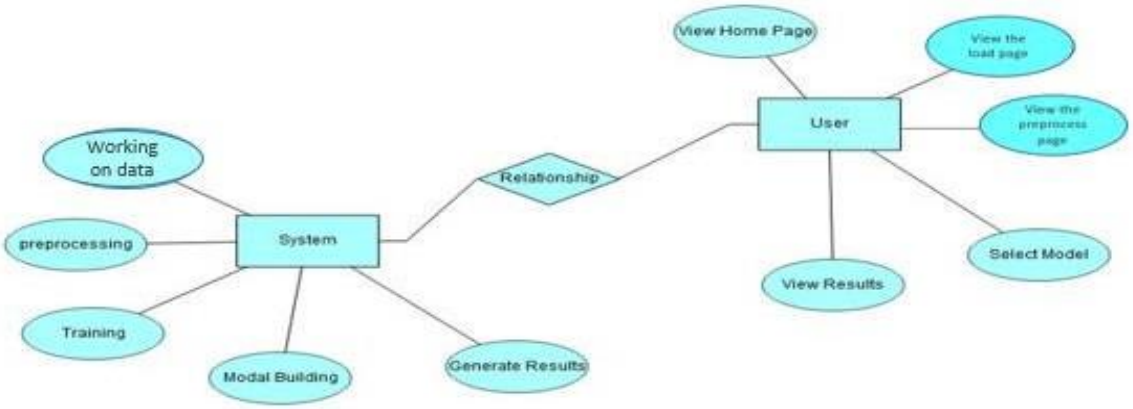


Figure5.8 ER Diagram

5.3 DFD Diagram:

A Data Flow Diagram(DFD) is a traditional way to fantasize the information flows within a system. The purpose of a DFD is to show the compass and boundaries of a system as a whole. It may be used as a dispatch tool between a systems critic and any person who plays a part in the system and acts as the starting point for redesigning a system.

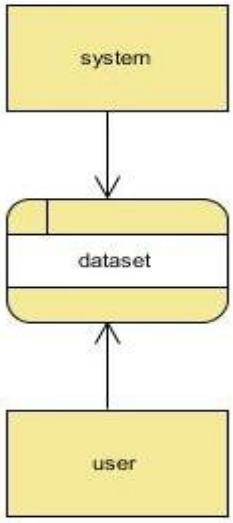


Figure5.9 DFD Diagram

CHAPTER 6

RESULTS AND OUTCOMES

6.1 RESULTS AND OUTCOMES

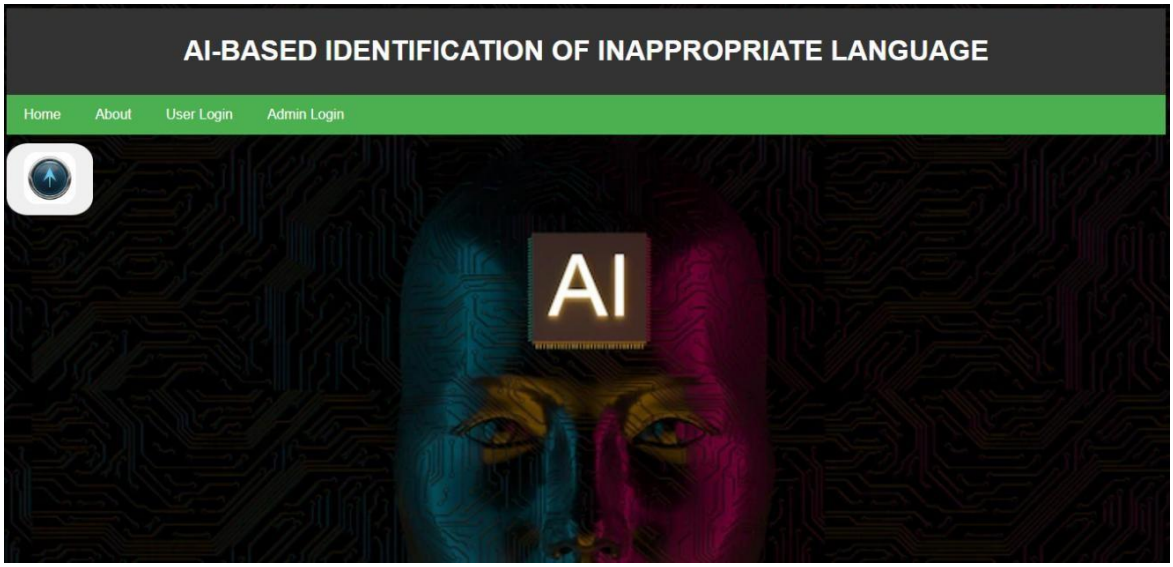


Figure 6.1 Sample screenshot of the AI tool

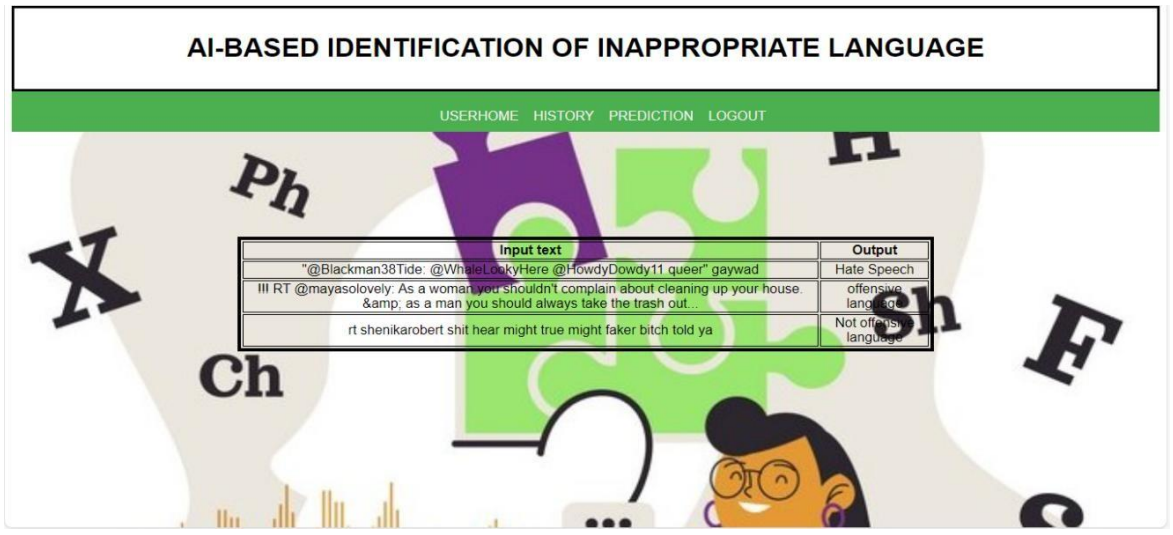


Figure 6.2 Sample screenshot of the output

CHAPTER 7

CONCLUSION AND FUTURE ENHANCEMENT

7.1 CONCLUSION

This study has handed precious perceptivity into the effectiveness of four distinct machine learning algorithms – Decision Trees, Random Forest, LSTM, and BERT – in detecting obnoxious language within online content. Each algorithm displayed its unique strengths and sins, with Decision Trees and Random Forest showcasing the significance of structured point engineering, while LSTM and BERT demonstrated the power of deep literacy ways in handling unshaped textbook data. By assessing delicacy, perfection, recall, and F1- score, we've gained a comprehensive understanding of their prophetic capabilities. The findings of this exploration hold a significant pledge for the development of robust AI- grounded content temperance systems, offering the eventuality to produce further inclusive and regardful online spaces, thereby contributing to a healthier and further responsible digital converse for all druggies.

7.2 FUTURE ENHANCEMENT:

This study could involve the disquisition of mongrel models that combine the strengths of both traditional machine learning algorithms like Decision Trees and Random Forest, with deep literacy ways similar to LSTM and BERT. also, the exploration could claw into multilingual and cross-cultural aspects of obnoxious language discovery to make these systems more inclusive and adaptable to different online communities. nonstop monitoring and updating of the models to keep up with evolving language trends and arising forms of online importunity would also be pivotal. likewise, enriching the interpretability of deep literacy models like BERT to understand their decision- making processes can enhance trust and translucency in content temperance systems. Incipiently, collaboration with social media platforms and content providers for real- world perpetration and evaluation would be a precious avenue for unborn exploration, icing the practical connection and effectiveness of these AI- driven results in maintaining regardful online converse.

CHAPTER 8

REFERENCES

REFERENCES

- Gaydhani, A., Doma, V., Kendre, S., & Bhagwat, L. (2018). Detecting hate speech and offensive language on twitter using machine learning: An n-gram and tfidf based approach. arXiv preprint arXiv:1809.08651.

- Akhter, M. P., Jiangbin, Z., Naqvi, I. R., Abdelmajeed, M., Mehmood, A., & Sadiq, M. T. (2020). Document-level text classification using single-layer multisize filters convolutional neural network. IEEE Access, 8, 42689-42707.

- Chillapalli, J., Gite, S., Saini, B., Kotecha, K., & Alfarhood, S. (2023). A Review of Diagnostic Strategies for Pulmonary Embolism Prediction in Computed Tomography Pulmonary Angiograms. IEEE Access.

- Leong, Y. S. (2022). Micro calcification detection in mammography for early breast cancer diagnosis using deep learning technique/Leong Yew Sum (Doctoral dissertation, Universiti Malaya).

- Sum, L. Y. (2022). Microcalcification Detection in Mammography for Early Breast Cancer Diagnosis Using Deep Learning Technique (Doctoral dissertation, University of Malaya (Malaysia)).

- Heidari, M., Mirniaharikandehei, S., Khuzani, A. Z., Danala, G., Qiu, Y., & Zheng, B. (2020). Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. International journal of medical informatics, 144, 104284.

- Poloni, K. M., de Oliveira, I. A. D., Tam, R., Ferrari, R. J., & Alzheimer's Disease Neuroimaging Initiative. (2021). Brain MR image classification for Alzheimer's disease diagnosis using structural hippocampal asymmetrical attributes from directional 3-D log-Gabor filter responses. Neurocomputing, 419, 126-135.

- Rodrigues, L. F., Naldi, M. C., & Mari, J. F. (2020). Comparing convolutional neural networks and preprocessing techniques for HEp-2 cell classification in immunofluorescence images. Computers in biology and medicine, 116, 103542.

DrillBit

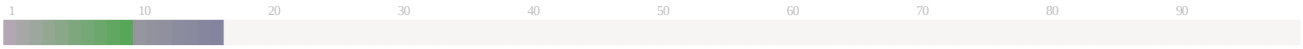
The Report is Generated by DrillBit Plagiarism Detection Software

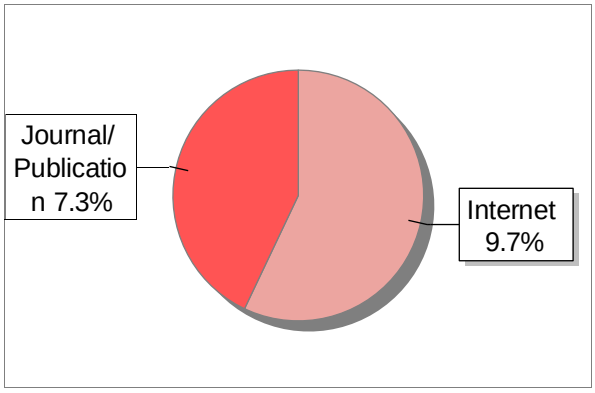
Submission Information

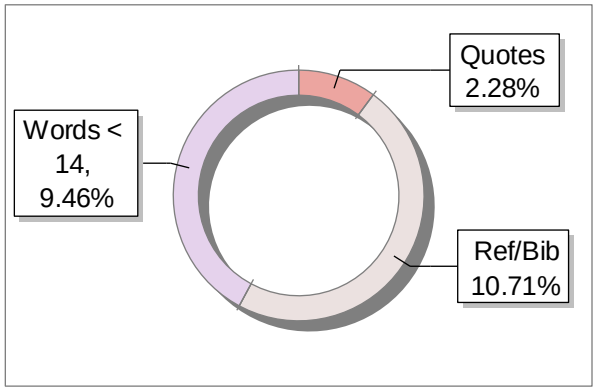
Author Name	Siddiq
Title	Revised version
Paper/Submission ID	1089359
Submitted by	v.muneeswaran@klu.ac.in
Submission Date	2023-11-08 10:50:37
Total Pages	20
Document type	Project Work

Result Information

Similarity **17 %**







Exclude Information

Quotes	Not Excluded
References/Bibliography	Not Excluded
Sources: Less than 14 Words Similarity	Not Excluded
Excluded Source	**0 %**
Excluded Phrases	Not Excluded

A Unique QR Code use to View/Download/Share Pdf File

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)
(Established Under Section 3 of UGC Act,1956)
Accredited with A+ Grade by NAAC in the Second Cycle
Eachanari, Coimbatore - 641021, TamilNadu.

E-NEXA

ELECTTRICCO

Faculty of Engineering
Department of Electrical and Electronics Engineering

TEPL
TALLY ACADEMY

ELECTTRICCO 2K24

CERTIFICATE OF MERIT

This is to certify that Dr./Mr./Ms. MOHAMMAD GHOUSEE ALI SIDDIQ of KALASALINGAM ACADAMY OF R&E has won II PRIZE - PROJECT PRESENTATION. AI BASED IDENTIFICATION & CORRECTION OF INAPPROPRIATE LANG in the National Level Technical fest Conducted by the Department of Electrical and Electronics Engineering, Faculty of Engineering, Karpagam Academy of Higher Education,Coimbatore on 08.02.2024

T.Selva ganapahy
EVENT CO-ORDINATOR

Dr. M. Vaigundamoorthi
HOD,DEPT OF EEE

Dr. A.Amudha
Dean , FOE

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)
(Established Under Section 3 of UGC Act,1956)
Accredited with A+ Grade by NAAC in the Second Cycle
Eachanari, Coimbatore - 641021, TamilNadu.

E-NEXA

ELECTTRICCO

Faculty of Engineering
Department of Electrical and Electronics Engineering

TEPL
TALLY ACADEMY

ELECTTRICCO 2K24

CERTIFICATE OF MERIT

This is to certify that Dr./Mr./Ms. SUNDU LEELA KRISHNA of KALASALINGAM ACADEMY OF R&E has won II - PRIZE - PROJECT PRESENTATION (LANG.) - AI BASED IDENTIFICATION & CORRECTION OF INAPPROPRIATE in the National Level Technical fest Conducted by the Department of Electrical and Electronics Engineering, Faculty of Engineering, Karpagam Academy of Higher Education,Coimbatore on 08.02.2024

T.Selva ganapahy	Dr. M. Vaigundamoorthi	Dr. A.Amudha
EVENT CO-ORDINATOR	HOD,DEPT OF EEE	Dean , FOE

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)
(Established Under Section 3 of UGC Act,1956)
Accredited with A+ Grade by NAAC in the Second Cycle
Eachanari, Coimbatore - 641021, TamilNadu.

E-NEXA

ELECTTRICCO

Faculty of Engineering
Department of Electrical and Electronics Engineering

TEPL TALLY ACADEMY

ELECTTRICCO 2K24

CERTIFICATE OF MERIT

This is to certify that Dr./Mr./Ms. KODE VIVEK SRINIVAS of KALASALINGAM ACADEMY OF R&E has won II PRIZE - PROJECT PRESENTATION. AI BASED IDENTIFICATION & CORRECTION OF INAPPROPRIATE LANG, in the National Level Technical fest Conducted by the Department of Electrical and Electronics Engineering, Faculty of Engineering, Karpagam Academy of Higher Education,Coimbatore on 08.02.2024

T.Selva ganapahy	Dr. M. Vaigundamoorthi	Dr. A.Amudha
EVENT CO-ORDINATOR	HOD,DEPT OF EEE	Dean , FOE

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)
(Established Under Section 3 of UGC Act,1956)
Accredited with A+ Grade by NAAC in the Second Cycle
Eachanari, Coimbatore - 641021, TamilNadu.

E-NEXA

ELECTTRICCO

TEPL
TALLY ACADEMY

Faculty of Engineering
Department of Electrical and Electronics Engineering

ELECTTRICCO 2K24

CERTIFICATE OF MERIT

This is to certify that Dr./Mr./Ms. A. HARISHANKAR. of KALASALINGAM ACADEMY OF R&E has won II - PRIZE - PROJECT PRESENTATION. A.I BASED IDENTIFICATION & CORRECTION OF INAPPROPRIATE LAND. in the National Level Technical fest Conducted by the Department of Electrical and Electronics Engineering, Faculty of Engineering, Karpagam Academy of Higher Education,Coimbatore on 08.02.2024

T.Selva ganapahy	Dr. M. Vaigundamoorthi	Dr. A.Amudha
EVENT CO-ORDINATOR	HOD,DEPT OF EEE	Dean , FOE