G.H. RAISONI COLLEGE OF ENGINEERING AND MANAGEMENT, WAGHOLI, PUNE.

Department of AI & AIML

AI content detector

Guide:

Prof. Rachna Sable

Name of Projectees

Devashri Bhosale A19

Devang Pagare A18

Pratik Jade A70

Eshan Kasliwal A22

Engineering   Management   Law   Schools   Other Courses

NAGPUR   PUNE   JALGAON   AMRAVATI   AHMEDNAGAR   RAIPUR   CHHINDWARA
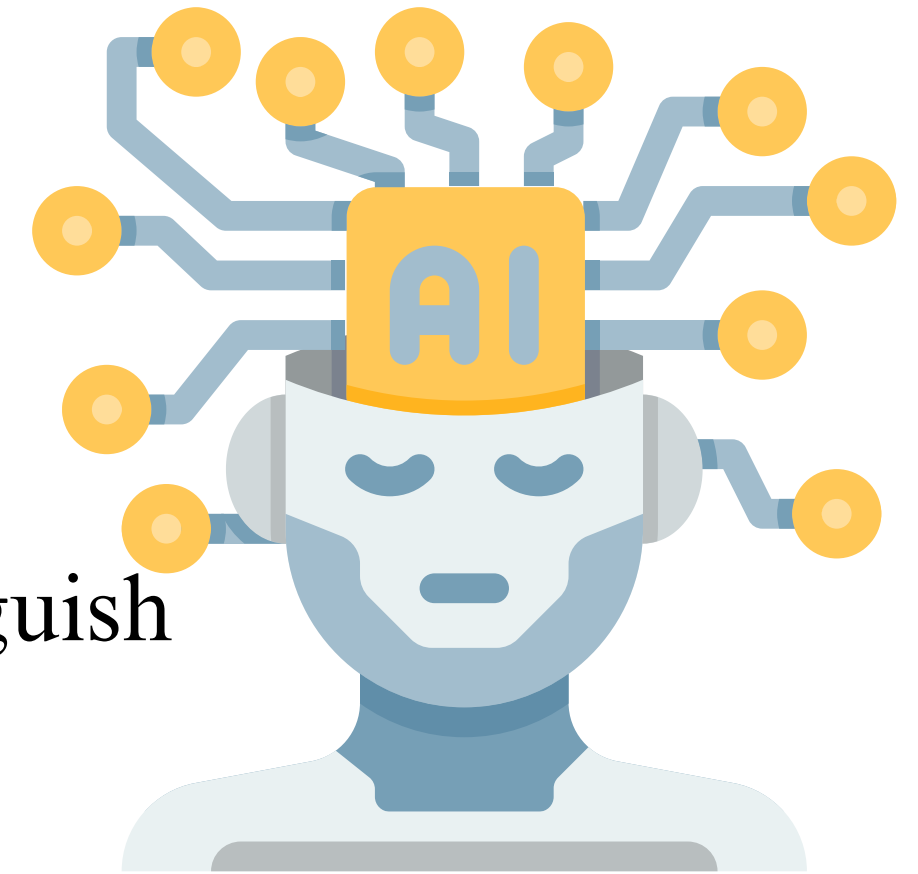
RAISONI GROUP OF INSTITUTIONS

Contents

- Introduction
- Justifications for Selecting the Title
- Problem Statement
- Literature Survey
- Block Diagram
- Expected Result
- Work plan
- References

Introduction

- Over the decades, Natural Language Processing (NLP) has evolved into a significant field of research aimed at enhancing computer ability to generate and comprehend human language.
- While ChatGPT brings revolutionary social value, it also raises concerns about the potential misuse of AI-generated content.
- Consequently, an important question arises: how can one distinguish between content generated by ChatGPT and that created by humans?

Problem Statement

The project aims to leverage AI and ML technologies to distinguish between human and AI-generated content.

Justification For Selecting The Title

- The title "AI Content Detector" is apt because it accurately reflects the purpose of the tool.
- As AI language rapidly growing, it is becoming more difficult to distinguish between content generated by AI and humans. This raises concerns about the potential for AI-generated content to be misused, such as for plagiarism or fake news.
- The title "AI Content Detector" highlights the urgency of developing methods to accurately identify AI-generated content. This is important for ensuring transparency and trust in an increasingly AI-driven world.
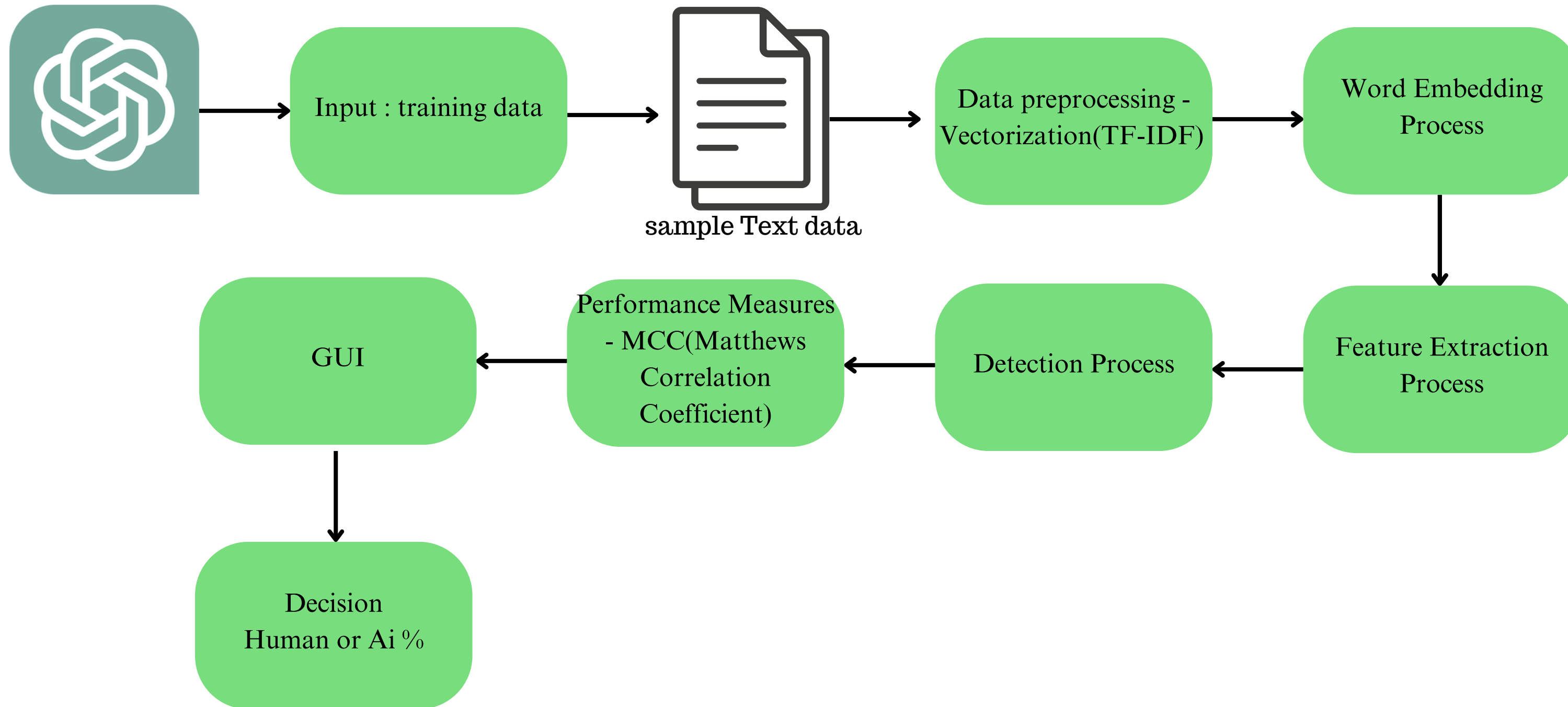
Literature Survey

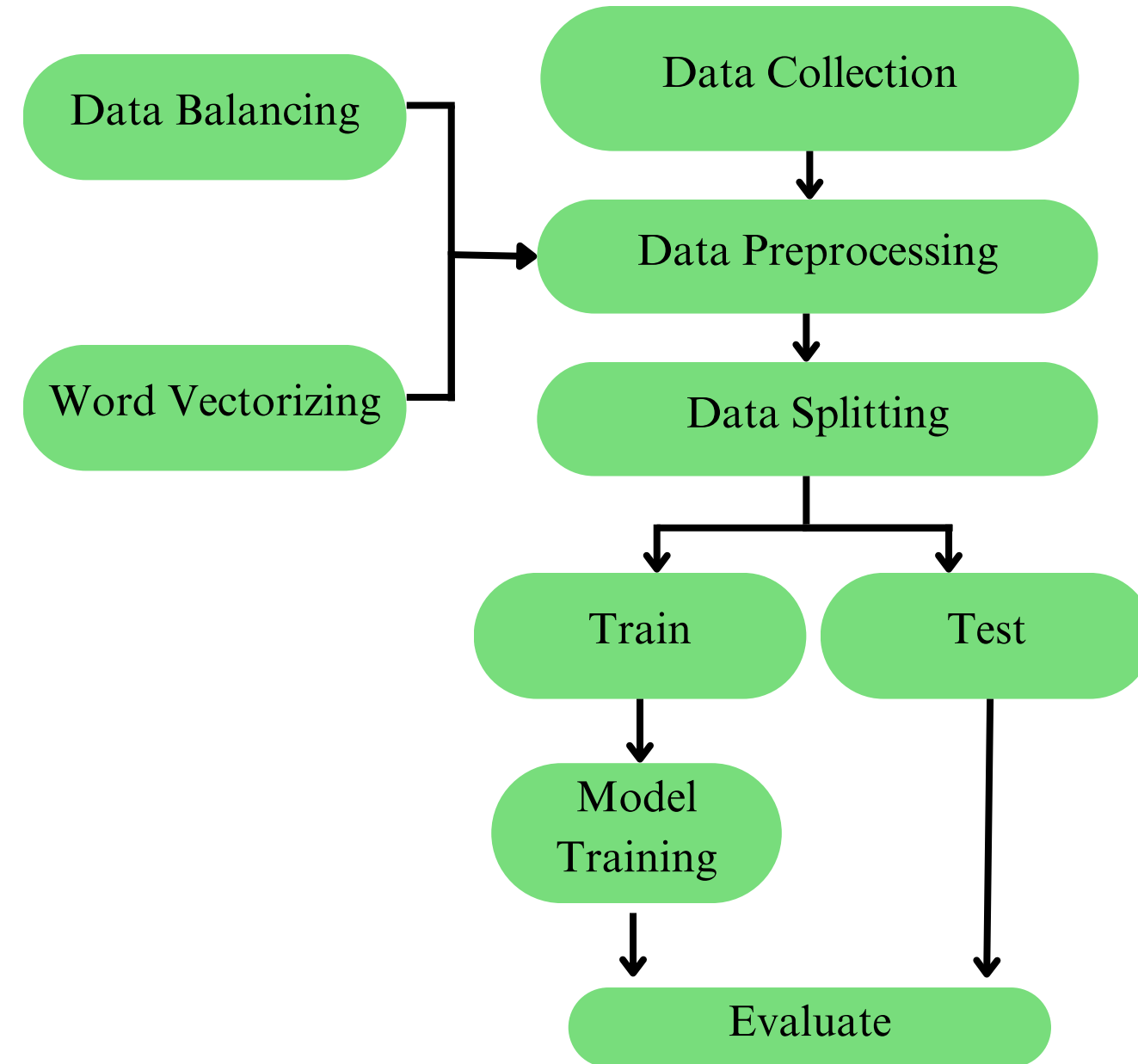| Sr. No.	Paper Title and its Author	Details of Publication	Date of Publication	Findings	Models used	Future Scope
1	CHATGPT OR HUMAN? DETECT AND EXPLAIN. Sandra Mitrovic´ 1 , Davide Andreoletti2 , and Omran Ayoub2	Cornell University	30 Jan, 2023	https://arxiv.org/abs/2301.13852	Transformer-based ML model SHAP for Explaining Model's Decisions	Some regularization techniques may improve the performance
2	Distinguishing Human-Written and ChatGPT-Generated Text Using Machine Learning Hosam Alamleh; Ali Abdullah S. AlQahtani; AbdElRahman ElSaid	IEEE	31 May, 2023	https://ieeexplore.ieee.org/document/10137767	LR, DT,SVM,NN & RF	To improve accuracy and performance
4	Distinguishing Human Generated Text From ChatGPT Generated Text Using Machine Learning	Cornell University	26 May, 2023	https://arxiv.org/abs/2306.01761	TF-IDF vectorizer & tree classifier, LSTM for detecting ChatGPT text	More sophisticated models for transperancy and accountability
5	DetectGPT: Zero-Shot Machine-Generated Text Detection using Probability Curvature	Cornell University	26 Jan, 2023	https://arxiv.org/abs/2301.11305	Probability Curvature	Identification of audio and video files through future models

Existing research has several shortcomings:

- Experiments are carried out for AI-generated text is on a **limited set of data.**
- Previously used to train language models and/or provide **no guarantee that they were actually written by humans.**
- The methodological aspects of the research are **not always described in detail** and are thus not available for replication.
- Testing whether the AI-generated and further translated text can influence the **accuracy of the detection tools is not discussed at all.**
- A limited number of measurable metrics is used to **evaluate the performance.**
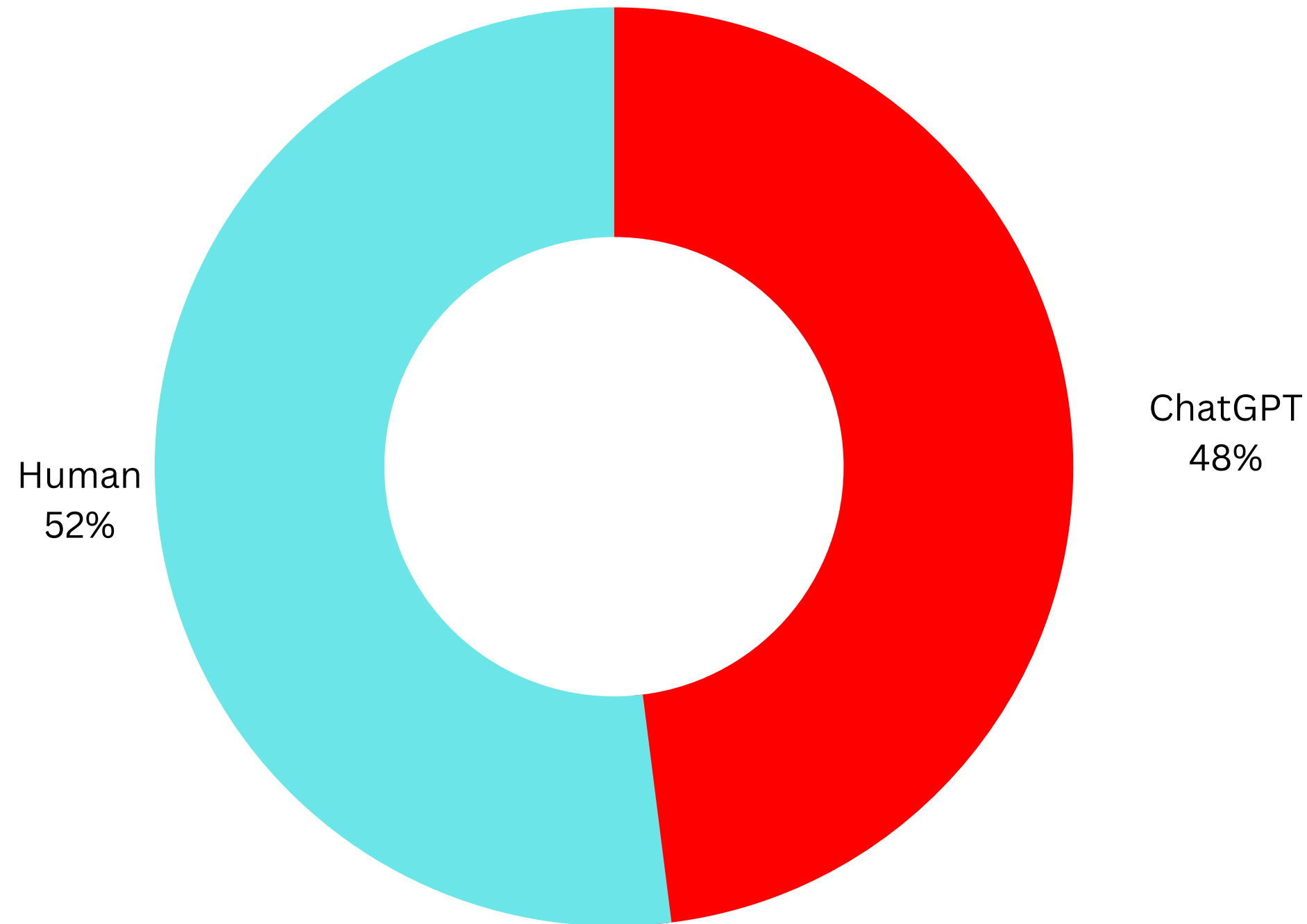
Block Diagram

Input : training data

sample Text data

Data preprocessing - Vectorization(TF-IDF)

Word Embedding Process

GUI

Performance Measures - MCC(Matthews Correlation Coefficient)

Detection Process

Feature Extraction Process

Decision
Human or Ai %

Block Diagram

overview of the process

Block Diagram

Dataset distribution

Human
52%

ChatGPT
48%

Expected Result

- Real-time Content Detection
- Content Categorization
- Accuracy and Precision
- User-Friendly Interface
- Performance Metrics
- Error Handling and Robustness
- Scalability

AI Detector

Enter Text

Clear | Detect Text

Your Text is AI/GPT Generated

31% AI Generated

Engineering | **Management** | **Law** | **Schools** | **Other Courses**
NAGPUR · PUNE · JALGAON · AMRAVATI · AHMEDNAGAR · RAIPUR · CHHINDWARA

RAISONI
GROUP OF INSTITUTIONS

System Requirements

- Windows 10,11
- Internet
- Python
- Jupiter Notebook
- CPU - Ryzen 5 5600G
- Ram - 16GB
- SKlearn, Tensorflow, Numpy, Pandas
- Integrated GPU

Work – Plan

MonthsActivities	June'23	July'23	August'23	Sept'23	Oct'23	Nov'23
Literature Reviews	✔	✔				
Requirement Analysis		✔				
Designing		✔	✔			
Experimental Analysis			✔	✔	✔	
Module wise Implementation				✔	✔	
Testing and Debugging					✔	✔
Preparation of Project Report						✔

References

https://ieeexplore.ieee.org/document/10137767

https://ieeexplore.ieee.org/document/9983964

https://arxiv.org/pdf/2301.13852.pdf

https://arxiv.org/pdf/2306.01761.pdf

https://www.zerogpt.com/

https://detectgpt.ericmitchell.ai/

https://github.com/jhroy/auditGPTZero

https://huggingface.co/datasets/artem9k/ai-text-detection-pile

Thank you!