" EXTRACTING SIGNIFICANT INFORMATION USING LARGE LANGUAGE MODELS(LLM) "

A SEMINAR REPORT SUBMITTED TO

THE NATIONAL INSTITUTE OF ENGINEERING

(An Autonomous Institute under VTU, Belagavi)

In partial fulfillment of the requirements for Major Project Presentation seventh semester

Bachelor of Engineering In Computer Science & Engineering

Submitted By

RAKSHITH ARYA (4NI20CS075)
SHASHANK SHANDILYA (4NI20CS096)
SHASHANK K R (4NI20CS095)
VISHAL M V (4NI20CS124)

Under The Guidance Of

Dr. Shabana Sultana
Professor
Department of CS & E

THE NATIONAL INSTITUTE OF ENGINEERING
Mysore-570 008
2023-24

THE NATIONAL INSTITUTE OF ENGINEERING

(An Autonomous institution, affiliated to VTU)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
Mysore -570008

CERTIFICATE

Certifies that the seminar work entitled ***"EXTRACTING SIGNIFICANT INFORMATION USING LARGE LANGUAGE MODELS(LLM)"*** is a work carried out by Rakshith Arya, ***Shashank Shandilya, Shashank K R, Vishal M V*** *bearing* ***4NI120CS075, 4NI120CS096, 4NI120CS094, 4NI120CS124*** *respectively* in partial fulfillment for the seminar prescribed by National Institute of Engineering, Autonomous Institution under Vishvesvaraya Technological University, Belgaum for the academic year 2023-2024, Computer Science & Engineering. It is certified that all correction/suggestions indicated for Internal Assessment have been incorporated. The Seminar report has been approved as it satisfies the academic requirements in respect of the seminar work prescribed for the Eight Semester.

Signature of the Guide **Signature of the H.O.D**

(Dr. Shabana Sultana) **(Dr. Anitha R)**

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who helped me in completing the seminar successfully.

We express our profound thanks to **Dr. Rohini Nagapadma, Principal, NIE, Mysore** for all the support and encouragement.

We are grateful to **Dr**. **Anitha R**, Prof. and Head of the Department of Computer Science and Engineering, NIE, Mysore for her support and encouragement facilitating the progress of this work

We sincerely extend my thanks to Project guide **Dr. Shababna Sultana**, Professor Department of Computer Science and Engineering, NIE, Mysore for their valuable guidance and support for this seminar.

Finally we thank our family and friends for being a constant source of inspiration and advice.

ABSTRACT

This abstract explores the application of machine learning techniques for extracting significant information from complex datasets. In an era marked by an overwhelming influx of data, the need to identify and prioritize relevant information has become imperative. This study delves into the utilization of advanced machine learning algorithms to sift through vast datasets and discern patterns, trends, and critical insights. By employing techniques such as natural language processing, clustering, and classification, the research aims to develop a robust framework for automated information extraction.

The focus is on enhancing the efficiency and accuracy of information retrieval, enabling swift decision-making processes in various domains, including business, healthcare, and research. The study evaluates the performance of diverse machine learning models in discerning meaningful information from noisy or unstructured data, contributing to the evolution of intelligent systems capable of distilling key insights. The findings from this research promise to advance the frontier of information extraction methodologies, facilitating a more streamlined and effective utilization of data in the ever-expanding landscape of information-driven decision-making.

TABLE OF CONTENTS

Serial No	Topic	Page No
1	Introduction	1
1.1	what is information extraction?	1
1.2	what are large language models (LLM)	2
2	Literature Survey	4
3	Existing and Proposed system	5
3.1	Existing System	5
3.2	Proposed System	6
4	System Architecture	7
4.1	Edge Computing System Architecture	7
4.1.1	Overview	7
4.1.2	Components of the System	8
5	Conclusion and Future Work	11
6	Referrences	12

LIST OF FIGURES

Serial No	Topic	Page No
1	Proposed System Architecture	7