Workshop Report :

RAG with Gemini

Organized By:

IEEE Computer Society SCET

AND

IEEE WIE SCET

under IEEE SCET SB

Computer Engineering Department, SCET

➢ EVENT DETAILS:

Type: Workshop

Event Mode: Offline

Event Date: 22 January, 2025

Event Venue: Phase-I Lab, Computer Department

Event Accessibility: For IEEE members & non-IEEE members (students, faculties)

Event Speaker: Jay Thakkar , Google Developer Expert-TensorFlow

Number of Attendees: 40

Faculty Coordinator: Prof. (Dr.) Pariza Kamboj

Prof. Vandana Joshi

Student Coordinator: Fenil Patel

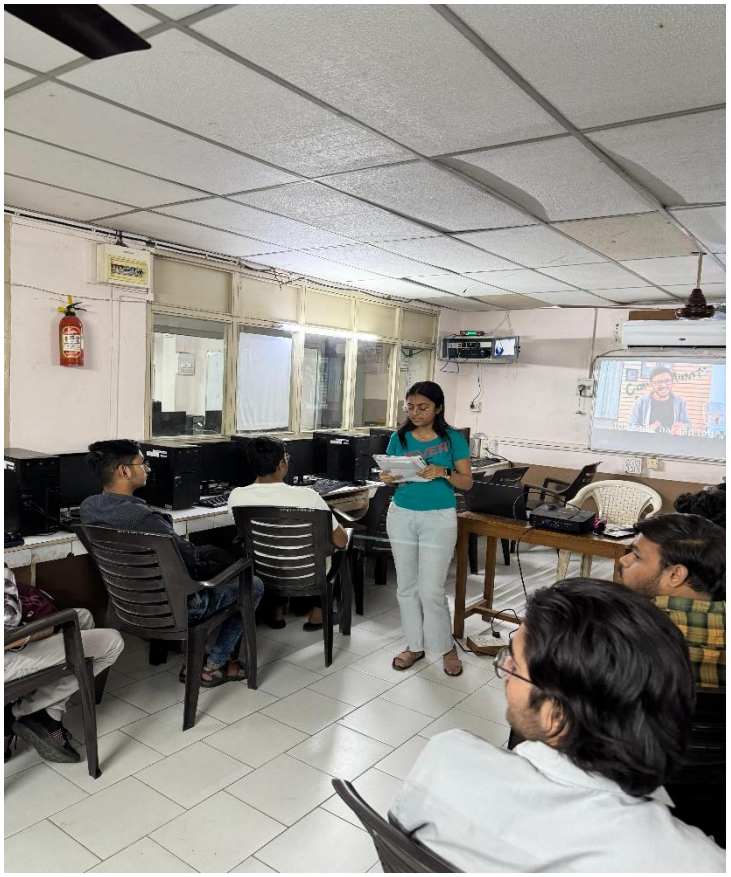
Devanshi Trivedi

EVENT POSTER

➢ INTRODUCTION:

The **IEEE Computer Society SCET SB** successfully organized an impactful workshop titled **"RAG-Retrieval Augmented Generation"** on 22nd January, 2025. The session was graced by Jay Thakkar who is The Google Developer Expert in TensorFlow along with that he is the TensorFlow user Group Surat's organizer and Master's Student at Symbiosis University , Pune

The event was attended by faculty members, IEEE student chapter representatives, and a large number of enthusiastic students.



- **WORKSHOP HIGHLIGHTS:**

The workshop began with a warm welcome of the speaker, Jay Thakkar by the organizing committee. The session then transitioned into an enlightening presentation that covered a wide range of topics related AI , Gen AI , LLM and RAG of the workshop:

1. **What is AI(Artificial Intelligence)**
   Speaker mentioned this as Artificial Intelligence (AI) is a branch of computer science that focuses on creating systems capable of performing tasks that typically require human intelligence.AI, which is designed for specific tasks like virtual assistants or recommendation systems, and general AI, which aims to perform a wide range of activities at a human level.

2. **Evaluation of AI :**
   The evaluation of Artificial Intelligence (AI) involves assessing its performance, accuracy, and impact in achieving specific tasks and solving real-world problems. This process includes testing AI models against predefined benchmarks to measure their efficiency, reliability, and adaptability, communication, and decision-making. Jay shared the way AI is becoming more and more advance by the data on which it's trained
3. **Introduction to RAG :**

The workshop begins with an introduction to **Retrieval-Augmented Generation (RAG) by Jay Thakkar** , an advanced AI framework that combines the power of information retrieval with text generation to create highly accurate and context-aware outputs. explore Gemini's role in improving RAG-based applications, and equip attendees with the knowledge and skills to implement these technologies effectively in real-world scenarios.

4. **Overview of Gemini :**
   The Expert demonstrate the **Gemini** is an advanced platform designed to elevate Retrieval-Augmented Generation (RAG) workflows by offering powerful capabilities and tools tailored for modern AI applications. Gemini streamlines the integration of retrieval and generation processes, enabling users to build intelligent, context-aware systems with greater efficiency..

5 . What are LLMs ?

Jay emphasized this in effective way **LLMs[Large Language Models]** play a pivotal role in both Retrieval-Augmented Generation (RAG) and Generative AI, serving as the core engines for understanding and generating human-like text. In RAG, LLMs are combined with retrieval mechanisms to ensure that generated outputs are not only coherent but also enriched with accurate and up-to-date information from external data sources.

1. **RAG** vs **NON-RAG**

Aspect	RAG	NON RAG
Methodology	Combines retrieval mechanisms with generative models to access external data.	Uses only pre-trained generative models without dynamic data retrieval.
Data Source	Integrates external, up-to-date information from databases or knowledge bases.	Relies solely on the information embedded during training, limited to static data.
Applications	Suitable for dynamic and knowledge-intensive tasks like question answering, chatbots, and document summarization.	Best for simpler, static tasks such as text generation, creative writing, or translation.

2. **Pricing Difference between RAG and NON RAG :**

- The Speaker Jay run GEN AI model with both method RAG and NON RAG for running this it at backend it count the price according to token and after running all task completely he calculate the price of RAG and NON RAG and the RAG is four times cheaper NON RAG , If you use more prompt than it's more cheaper


NEVER
GIVE
UP

- **Audience Engagement and Interaction**

The workshop was highly interactive, with Jay Thakkar encouraging participants to ask questions and share their experiences. The audience actively engaged in discussions, raising queries about the AI process for training, accuracy for hosting Gen AI model, and ways to maximize the productivity by Using AI efficiently. Jay's approachable demeanor and thoughtful responses ensured that all questions were addressed comprehensively. Several participants shared their aspirations and ideas for future Projects in AI, which added to the collaborative spirit of the session.

➢ CONCLUSION:

The workshop concluded with a vote of thanks delivered by the organizing team, expressing gratitude to Jay Thakkar for his insightful presentation and valuable guidance. The session left a lasting impression on the attendees, inspiring them to take full advantage of learning about Generative AI, AI, RAG [Retrieval Augmented Generation]

The event was a resounding success, serving as a stepping stone toward fostering a vibrant IEEE community within the institute. It not only educated participants about the understanding of Gen AI but also make them to create any kind of GEN-AI application and difference between RAG application and NON-RAG application

➢ Report Compiled by:

Fenil Patel,
Chairperson, IEEE Computer Society Chapter, Sarvajanik College of Engineering & Technology Student Branch