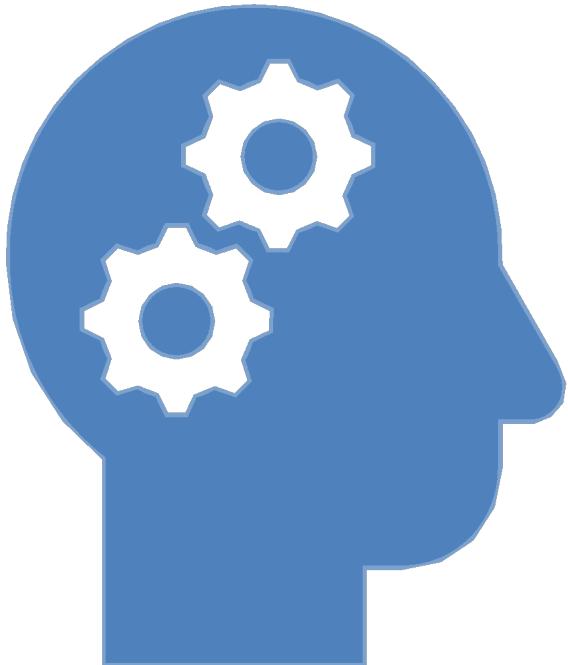


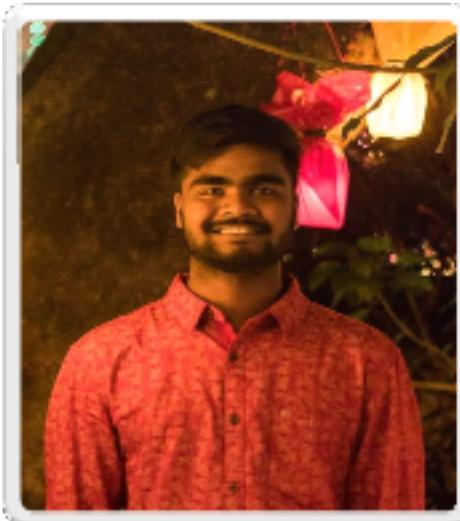
AI-Powered Career Guidance Chatbot using Retrieval- Augmented Generation (RAG)



TEAM-02



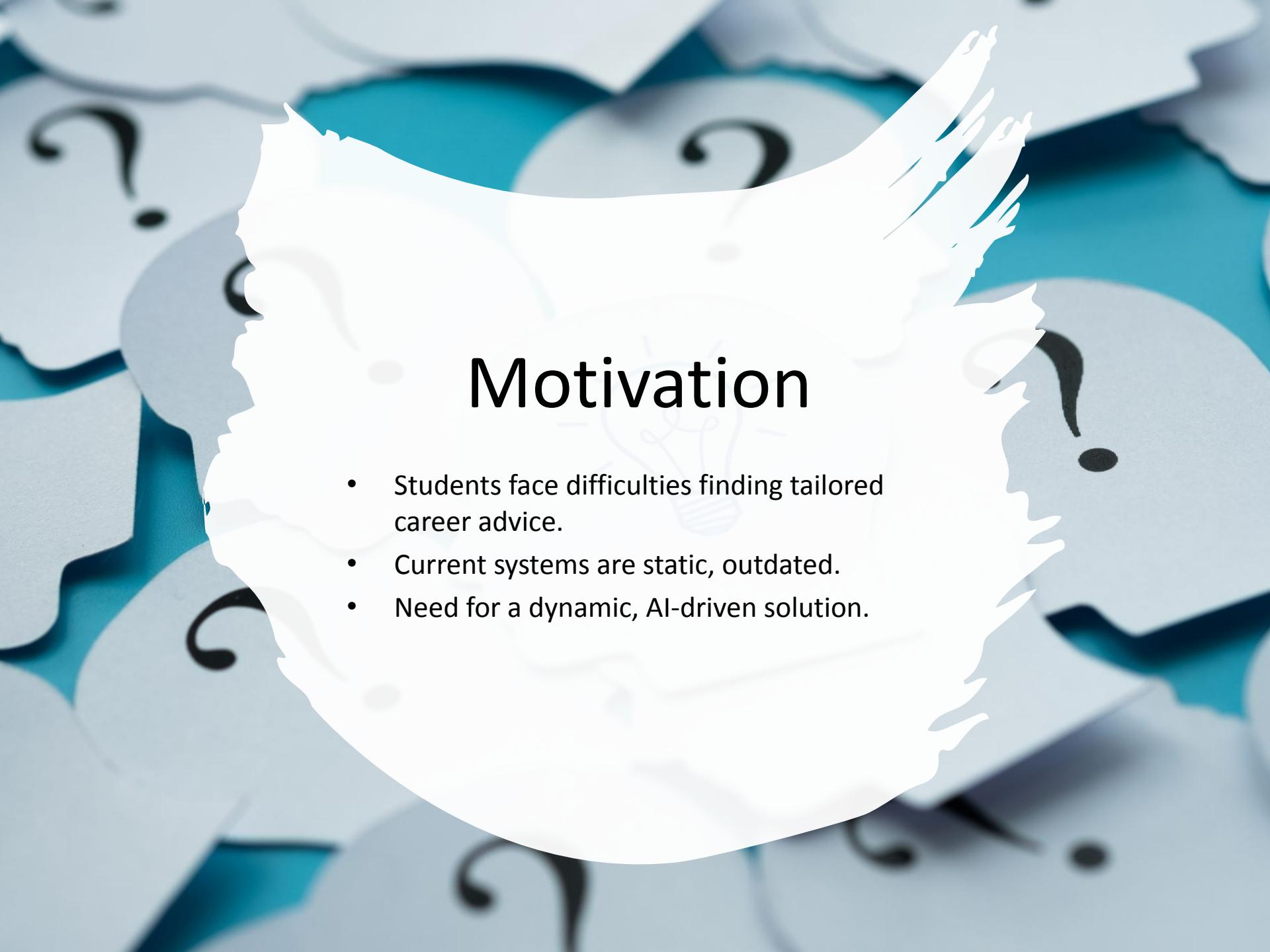
Simran Sattar



**Sandeep Raj
Katipagala**



Nertila Cahani



Motivation

- Students face difficulties finding tailored career advice.
- Current systems are static, outdated.
- Need for a dynamic, AI-driven solution.

Objective

- Build a RAG-based system for personalized career advice.
- Retrieve relevant articles + generate human-like responses.
- Minimize hallucinations, ensure factual consistency.

Dataset Overview



Dataset: career_articles.json



50+ curated, cleaned career articles.



Covers: Data Science, Cybersecurity, Marketing, Finance, Engineering.

Knowledge Base Creation

Used all-MiniLM-L6-v2 to embed articles.

FAISS indexing for scalable, semantic search.

Retrieval (R in RAG)



User query embedded and matched with top-k articles.



Keyword boosting enhances domain relevance.

Generation (AG in RAG)



Retrieved articles form Chain-of-Thought prompts.



Phi-2, Mistral-7B, Falcon-7B generate responses.



Strategy improves reasoning and reduces hallucination.

Complete RAG Flow

User Query

Retrieval Module
(FAISS Index + Optional Keyword Filtering)

Chain-of-Thought Prompt

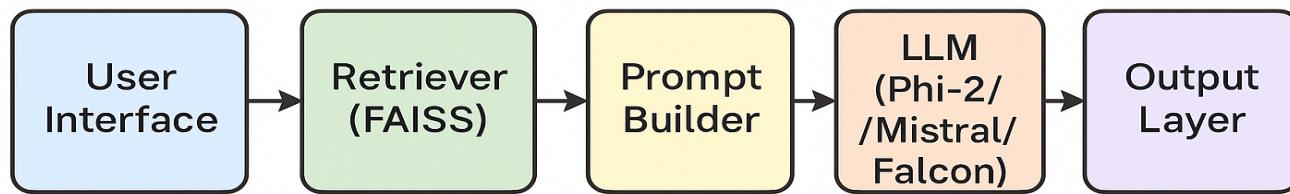
- Build Context
- Format the Prompt

LLM Generation

(Huggingface Phi-2 / Mistral / etc.)

Output Response
(Career Advice)

System Architecture Diagram



Models Used



Phi-2: Lightweight,
efficient.



Mistral-7B: Fluent,
descriptive text.



Falcon-7B: Balanced
factual grounding +
creativity.

Evaluation Metrics



BLEU Score: Measures
n-gram overlap.



ROUGE Score: Recall
and phrase overlap.



Manual Hallucination
Checks.

Key Results



BLEU > 0.6 for all
models.



Phi-2: Least
hallucination.



Mistral-7B: Best
fluency.

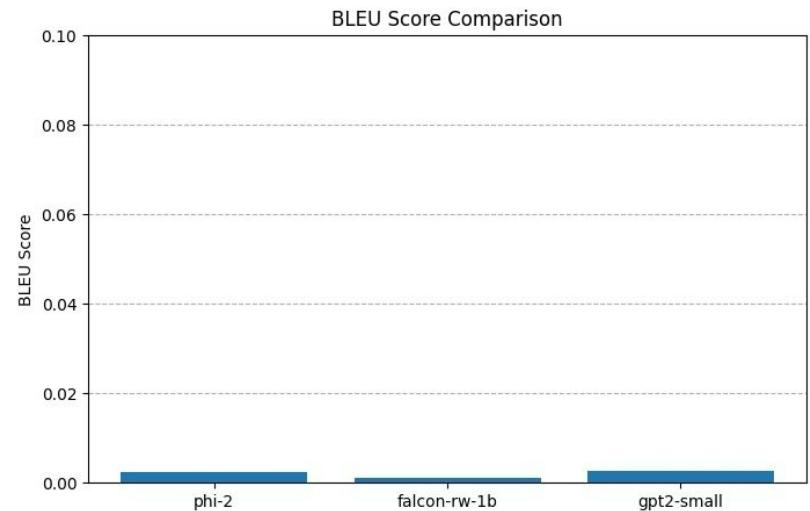
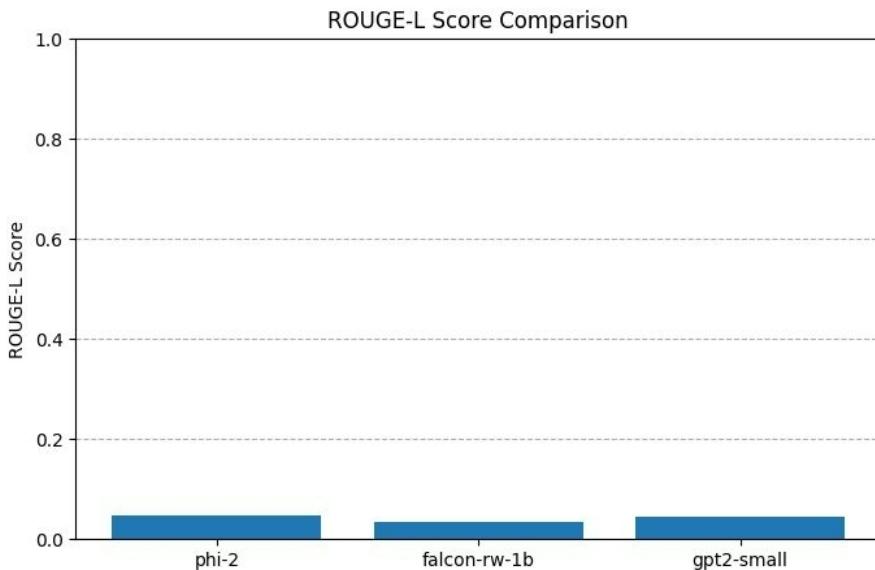


Retrieval improved
factuality.

Metrics

<input type="checkbox"/> Model	BLEU Score	ROUGE-1	ROUGE-L
Phi-2	0.002358	0.060516	0.048098
Falcon-RW-1B	0.001065	0.037290	0.035097
GPT2-Small	0.002612	0.054304	0.044022

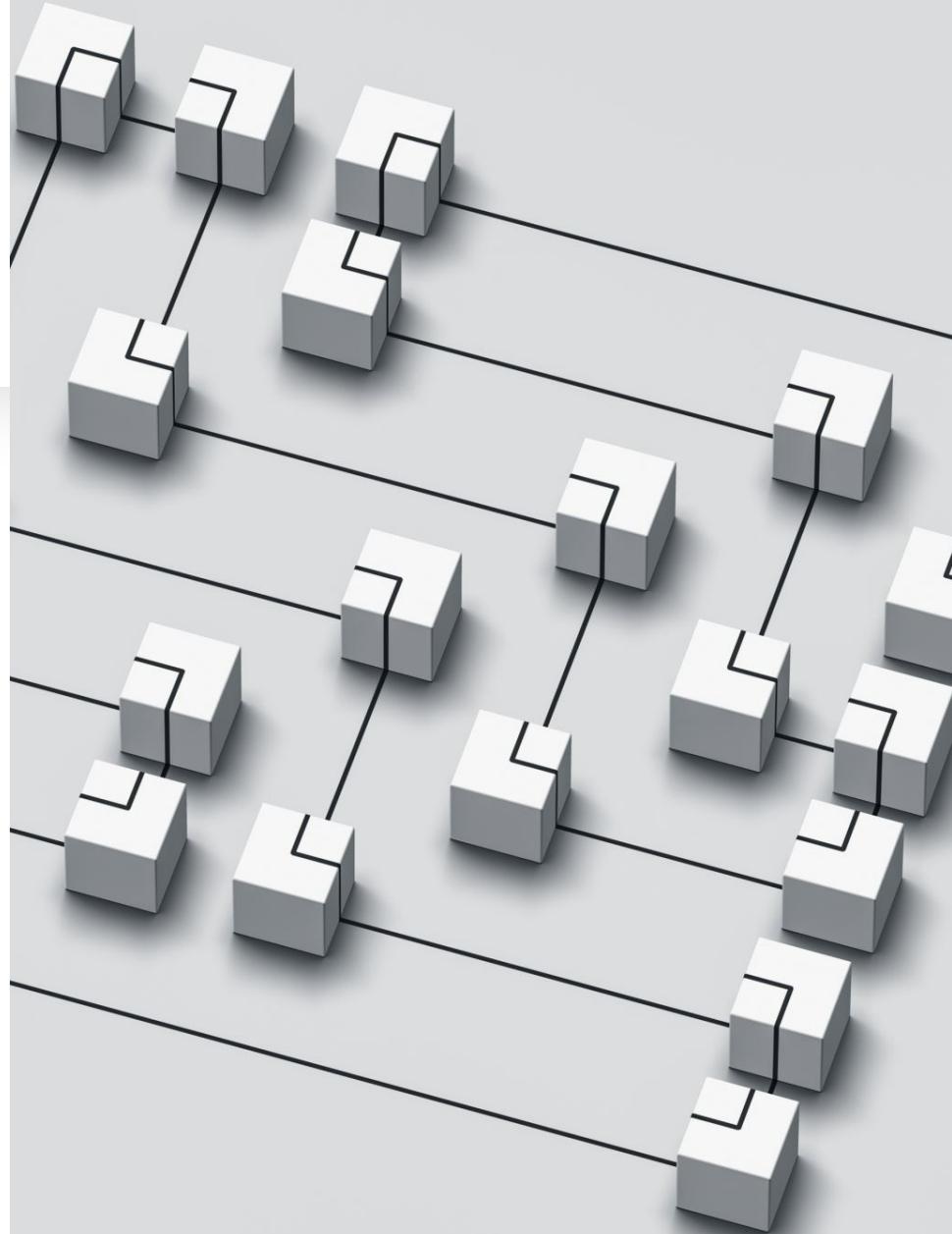
Score Comparison



"Phi-2 achieved the best ROUGE scores, while GPT2-small slightly outperformed in BLEU score."

Conclusion and Future Work

- Built a modular, scalable RAG chatbot.
- Future Plans:
 - Multi-turn conversation support.
 - Expand dataset to 500+ articles.
 - Implement Dense Passage Retrieval (DPR).



Thank You

- Thank you for your attention!

