

22AIE201 FUNDAMENTALS OF AI

Project Report on

ClarifAI - A Smart Research Paper Summarizer

Submitted by

Team No.: 12

Team Members:

Baranidharan S	–	CB.SC.U4AIE23015
Hygrevan P U	–	CB.SC.U4AIE23050
Pranav R	–	CB.SC.U4AIE23054
Sai Nivedh V	–	CB.SC.U4AIE23062

in partial fulfillment for the award of the degree of

**BACHELOR OF TECHNOLOGY
IN
CSE(AI)**



Amrita School of Artificial Intelligence
AMRITA VISHWA VIDYAPEETHAM
COIMBATORE - 641 112 (INDIA)
November 2024



AMRITA SCHOOL OF ARTIFICIAL INTELLIGENCE
AMRITA VISHWA VIDYAPEETHAM COIMBATORE – 641 112 (INDIA)

BONAFIDE CERTIFICATE

This is to certify that the report of the project entitled “**ClarifAI – A Smart Research Paper Summarizer**” submitted by **Team – 12**, for the award of the Degree of Bachelor of Technology in the “CSE(AI)” is a bonafide record of the work carried out by us, under our guidance and supervision at Amrita School of Artificial Intelligence, Coimbatore.

Date of Submission: 11/11/2024

Dr. Abhishek S
Project Supervisor

Dr. K.P.Soman
Professor and Head CEN



AMRITA SCHOOL OF ARTIFICIAL INTELLIGENCE
AMRITA VISHWA VIDYAPEETHAM COIMBATORE – 641 112 (INDIA)

DECLARATION

We hereby declare that this project entitled “**ClarifAI – A Smart Research Paper Summarizer**” is the record of the original work done by us under the guidance of Dr. Abhishek S, Centre for Computational Engineering and Networking, Amrita School of Artificial Intelligence, Coimbatore. To the best of our knowledge this work has not formed the basis for the award of any degree/diploma/ associate ship/fellowship/or a similar award to any candidate in any University.

Place: Coimbatore

Date: 11.11.2024

Signature of Faculty

Names

Roll Numbers

Baranidharan S

CB.SC.U4AIE23015

Hygrevan P U

CB.SC.U4AIE23050

Pranav Rajeshkumar

CB.SC.U4AIE23054

Sai Nivedh S

CB.SC.U4AIE23062

TABLE OF CONTENTS

ABSTRACT.....	5
INTRODUCTION.....	6
PROPOSED APPROACH.....	8
EXPERIMENTS & RESULTS.....	11
DISCUSSION.....	14
CONCLUSION.....	15
REFERENCES.....	16

LIST OF FIGURES

Fig 2.1. Overall Pipeline

Fig 3.1. Index page of ClarifAI

Fig 3.2. Completion of File upload

Fig 3.3. Output – Summarized content of Input Research Paper

Fig 3.4. Query request to Chatbot

ABSTRACT

ClarifAI - A Research Paper Summarizer offers an innovative approach to making complex academic research accessible and engaging through AI-powered summarization and visualization. As the volume of research publications grows, professionals and students face challenges in digesting intricate information quickly. ClarifAI addresses this by transforming research papers into concise, visually enriched summaries via an interactive dashboard, employing Retrieval-Augmented Generation (RAG), Natural Language Processing (NLP), and Large Language Models (LLMs).

RAG enhances content by integrating additional context from external sources, while NLP identifies core concepts and extracts essential insights. LLMs generate thematic visuals and informative graphs, providing a unique layer of engagement that deepens understanding. Unlike traditional summarizers, ClarifAI's interface enables users to explore summaries interactively, drilling down into detailed visual analyses that support cognitive retention.

Implemented with a modular architecture and evaluated on readability, engagement, and expert accuracy metrics, ClarifAI demonstrates improved user comprehension of research content. Future improvements will enhance visual relevance and expand to multi-modal research types. Ultimately, ClarifAI exemplifies the power of AI to democratize access to knowledge, making complex research more approachable and insightful.

1. INTRODUCTION

In today's fast-paced academic and professional landscape, the sheer volume of research available can be both overwhelming and challenging to navigate. Researchers, students, and professionals often spend significant time and effort sifting through complex content to uncover core insights, trends, and relevant information. Traditional summarization tools fall short in this regard, as they generally offer only basic, text-based summaries that lack interactivity, intellectual depth, and visual elements that aid comprehension.

This project aims to address these limitations by developing an AI-Based Research Summarizer Tool that transforms how users consume and engage with research material. This tool leverages advanced natural language processing (NLP) techniques and retrieval-augmented generation (RAG) methods to create coherent, accessible, and visually enriched summaries of complex research documents. Users can upload entire documents or portions of text, and the tool will produce a detailed summary, highlighting fundamental concepts, key findings, and relevant contextual information.

1.1. Key features of the tool

Customizable Summarization: Allows users to adjust the summary length and focus on specific areas of interest, whether it be methodologies, conclusions, or data interpretations.

Enhanced Summaries: Integrates extended facts, simplified language, and keyword extraction to deepen understanding and make the content more approachable.

Visual Enhancements: Adds relevant images, diagrams, and graphical data representations generated by AI to support and reinforce the summarized content.

Interactive Chatbot: Utilizes a RAG-powered chatbot to facilitate Q&A sessions, context-aware explanations, and keyword searches, providing users with a conversational interface to explore the content further.

To achieve these functionalities, the tool combines multiple AI components and frameworks. Upon document upload, it uses a large language model (LLM) to parse the document and establish a graph-based knowledge structure using Neo4j, which supports retrieval and interaction. Groq is then used for summarizing the parsed content, while image generation models like Stable Diffusion produce visuals aligned with key points in the summary.

This project delivers an innovative solution that does not just summarize research but also enhances user understanding through a combination of text, visuals, and interactivity. By doing so, it empowers users to navigate academic and professional research more effectively, enabling them to save time and focus on analysis and application.

1.2. Related Works

Neo4j Documentation:

Neo4j is a graph database management system that provides powerful tools for structuring and querying complex data relationships. The Neo4j documentation covers the use of Cypher, a specialized query language that allows for efficient retrieval and organization of connected data points, making it instrumental in ClarifAI's capability to provide contextually relevant information. [1]

GROQ Documentation:

GROQ (Graph-Relational Object Queries) offers a flexible, intuitive querying syntax, particularly useful for structured data retrieval. The GROQ documentation was essential in setting up precise query mechanisms within ClarifAI, allowing for efficient and adaptable extraction of structured information tailored to user inquiries. [2]

Large Language Models (LLM) Documentation:

Large Language Models (LLMs), including models available through Hugging Face, underpin the generative functionalities within ClarifAI, enabling summarization, image generation, and data visualization. Documentation for these models provided essential

guidance on fine-tuning, tokenization, and prompt engineering, enabling ClarifAI to deliver accurate, engaging, and contextually aligned outputs.[3]

2. PROPOSED APPROACH

The AI-Based Research Summarizer Tool employs a robust and systematic methodology, combining advanced natural language processing, visual generation, and interactive querying to deliver comprehensive, accessible summaries of academic research content. This methodology comprises several sequential stages, ensuring that the content is not only summarized but also enriched with visual aids and interactive features for a more engaging user experience.

2.1. Document Upload and Parsing

Objective: To efficiently process and parse complex academic documents, enabling downstream tasks such as summarization and RAG modeling.

Process: Users upload documents in PDF or DOCX formats. The uploaded document is parsed using a large language model (LLM) optimized for document parsing and understanding. This LLM extracts key sections, concepts, and relationships within the text, which become the basis for generating the knowledge structure and summaries.

2.2. Graph-Based Knowledge Modeling with Neo4j for RAG

Objective: To represent parsed content in a graph format that enables efficient retrieval and contextual understanding.

Process: The parsed content is transformed into a graph structure using Neo4j, a graph database designed for complex data relationships. This graph-based representation aids in creating a Retrieval-Augmented Generation (RAG) model, where nodes represent key concepts and facts, and edges represent relationships between them. This RAG model allows for contextual querying and supports knowledge-driven responses in the Q&A chatbot.

2.3. Summarization via Groq

Objective: To generate a user-customizable summary that distills the main ideas while preserving essential details and extended facts.

Process: Once the parsed content is structured, it is sent to Groq, an NLP engine that generates high-quality, concise summaries. Groq allows users to set summary length and focus areas, such as specific sections or concepts within the research. It also extracts keywords and includes extended facts where relevant, offering users a deep yet manageable overview of the document's core themes.

2.4. Visual Content Generation with Stable Diffusion

Objective: To enhance the summarized text with relevant visuals, supporting a more intuitive understanding of the material.

Process: Key points and concepts identified during the summarization phase are used as prompts for Stable Diffusion, a generative model for creating custom images. This model generates relevant illustrations, diagrams, or other visuals that align with the summarized content. These visuals are then integrated into the final summary, providing users with visual anchors that help clarify complex ideas.

2.5. Interactive Dashboard with Graphical Visualization

Objective: To provide a user-friendly interface for engaging with the summarized content, visuals, and interactive graph.

Process: A backend server consolidates the text summary, visual content, and knowledge graph structure, displaying them on an interactive dashboard. This dashboard presents a unified view of the research content, including a graphical visualization that illustrates the relationships between key concepts. The interactive graph enables users to explore specific themes, helping them understand the document's structure and main ideas visually.

2.6. RAG-Powered Chatbot for Q&A and Contextual Assistance

Objective: To enable users to engage in conversational Q&A related to the research content, providing context-aware answers to their inquiries.

Process: The tool integrates a chatbot powered by the Neo4j-based RAG model, specifically designed to answer questions related to the research document. When users query the content, the chatbot retrieves relevant information from the graph, offering responses that are contextually accurate and directly tied to the document. This chatbot facilitates in-depth understanding, allowing users to clarify doubts or gain insights into specific sections of the research.

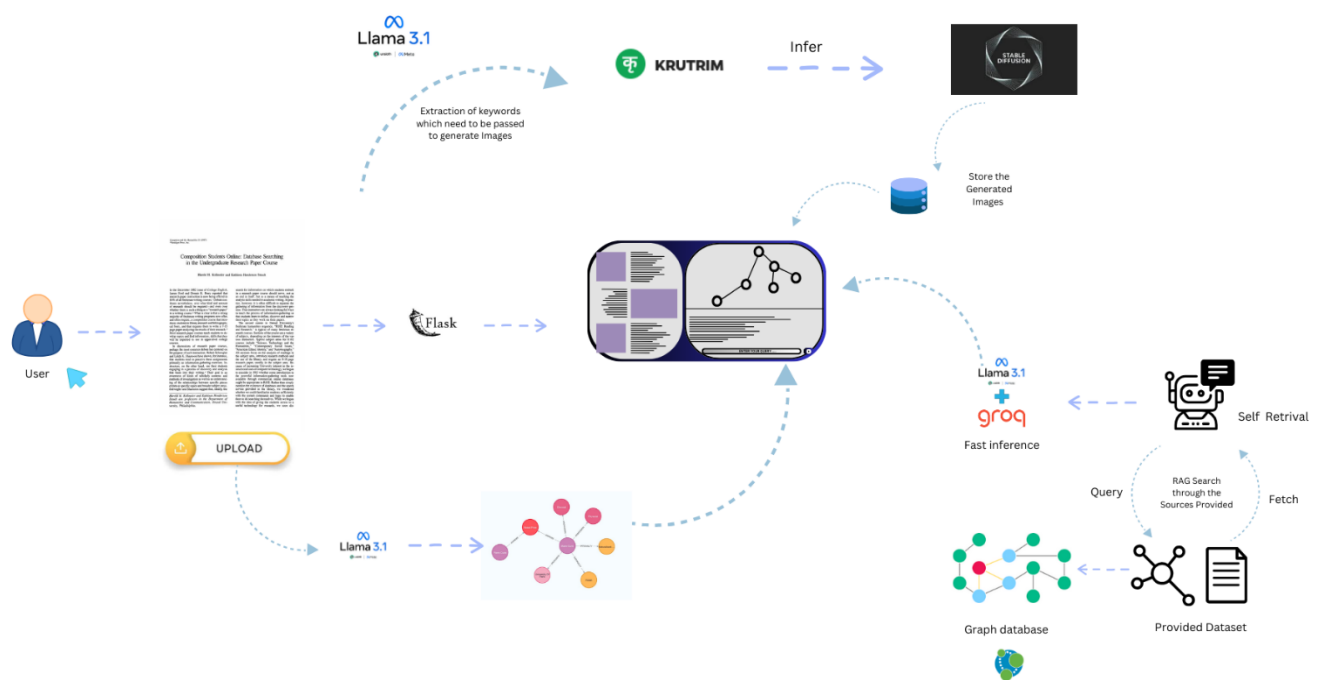


Fig 2.1. Overall Pipeline

3. EXPERIMENTS & RESULTS



Fig 3.1. Index page of ClarifAI

A clean and simple file uploader interface for ClarifAI, featuring a drag-and-drop area with browsing options. The design has a soft gradient background, giving it a modern, professional look.

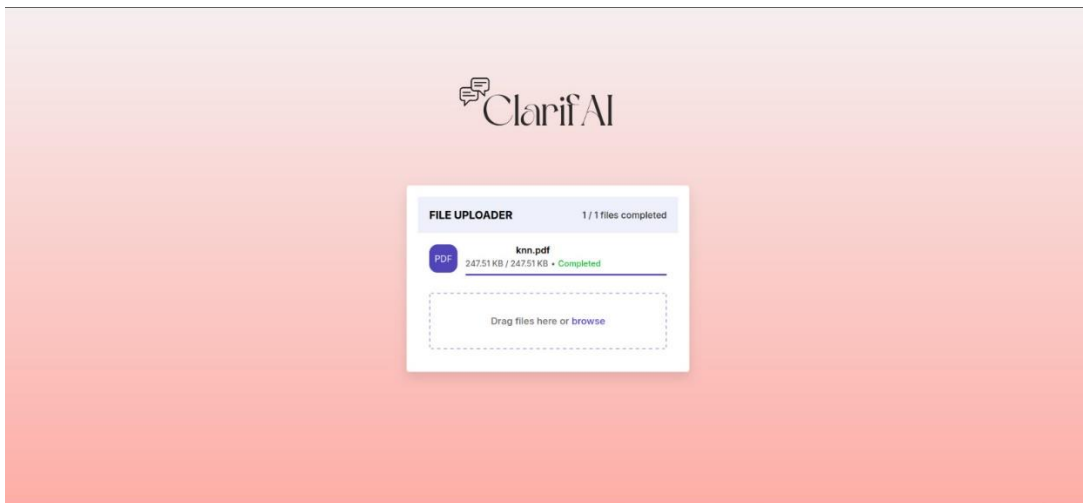


Fig 3.2. Completion of File upload

The ClarifAI file uploader shows a completed upload of "knn.pdf," with file size and progress status. A simple and intuitive interface with a progress bar and drag-and-drop functionality.

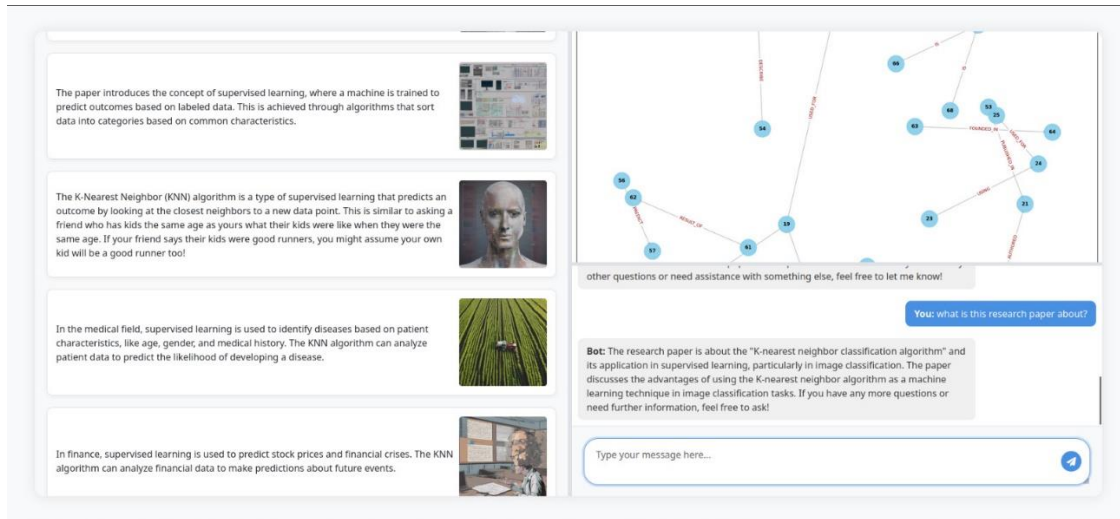


Fig 3.3. Output – Summarized content of Input Research Paper

This is a chat interface for a research summarization platform. It provides summaries of the "K-Nearest Neighbor (KNN)" algorithm applied in various fields, with text blocks and relevant images on the left. On the right, a graphical network visualizes concepts, and a chat area allows users to ask questions about the research.

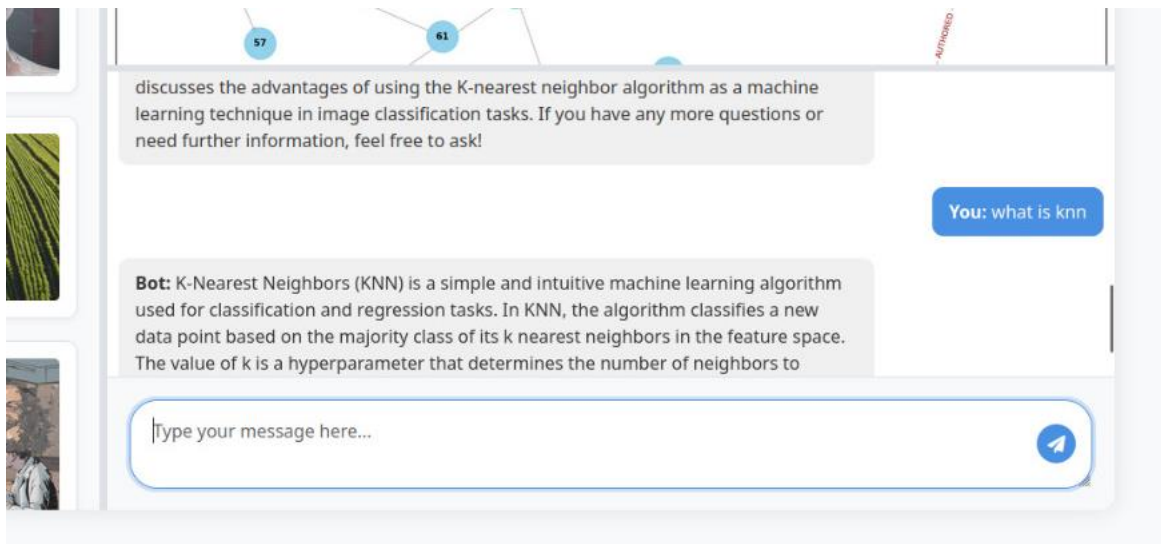


Fig 3.4. Query request to Chatbot

In this image of a chat interface, a user asks "what is knn" and receives a response explaining that K-Nearest Neighbors (KNN) is a machine learning algorithm used for classification and regression tasks, where classification is based on the majority class of k nearest neighbors in the feature space.

4. DISCUSSION

The AI-Based Research Summarizer Tool has the potential to transform the way academic content is processed, summarized, and engaged with by leveraging advanced NLP, knowledge graphs, and generative visual content. This tool provides an intuitive and comprehensive interface for navigating complex research documents. While it has shown promising results, several areas remain for improvement and enhancement. Currently, the tool faces limitations in processing time, particularly with large documents, which could be mitigated by implementing parallel processing and research-optimized NLP models. The visual content generation, though effective, is computationally intensive, and improvements such as reusable visuals or simplified image models could enhance efficiency. Scalability also poses a challenge; moving to a cloud-based infrastructure would better accommodate growing user demands, supporting larger datasets and multiple concurrent users. Customizing summaries based on user preferences could add significant value, allowing users to prioritize specific sections of documents. Real-time processing would further enrich the user experience by offering immediate insights and partial summaries as documents are processed, while supporting additional formats like LaTeX would broaden accessibility across various academic disciplines.

Looking ahead, expanding the tool's capabilities to support multiple languages could make it globally accessible, catering to a more diverse user base. The integration of AI-driven citation and reference management would streamline the research process by automating citation organization. Additionally, linking with research databases like Google Scholar, PubMed, and arXiv would facilitate seamless document import, optimizing the workflow for researchers. Adaptive summarization, influenced by user feedback, would allow the tool to provide increasingly personalized summaries over time. The potential for real-time collaboration would make it possible for multiple users to analyze a document simultaneously, promoting academic teamwork and interactive discussions. Lastly, integration with virtual research assistants could further support researchers in navigating academic content and suggesting relevant literature, reducing search time and fostering academic collaboration.

5. CONCLUSION

The AI-Based Research Summarizer Tool has demonstrated significant promise in transforming the way academic content is processed, understood, and interacted with. By integrating cutting-edge technologies like natural language processing, knowledge graph modeling, and generative visual content, the tool offers users an efficient and user-friendly way to digest complex research documents. Through its ability to generate concise summaries, provide visual aids, and enable interactive exploration, it has successfully addressed key challenges in academic research, including the time-consuming task of reading and understanding large volumes of literature.

While the tool has proven effective in its current form, several opportunities for improvement and future development have been identified. Optimizing processing time, enhancing visual content generation, and expanding the system's scalability are key areas that could significantly improve user experience and system performance. Additionally, further enhancements, such as multilingual support, adaptive summarization, and integration with research databases, could greatly expand the tool's reach and usability across global academic communities.

The tool's future implications, such as incorporating real-time collaboration features, AI-driven citation management, and integration with virtual research assistants, point toward a more dynamic, interactive, and intelligent academic research environment. These innovations will not only streamline research workflows but also foster collaboration, making the research process more efficient, effective, and accessible.

6. REFERENCES

1. Neo4j Inc. (n.d.). Neo4j Documentation. Retrieved from <https://neo4j.com/docs/>
2. Sanity.io (n.d.). GROQ Documentation. Retrieved from <https://www.sanity.io/docs/groq>
3. Hugging Face (n.d.). Transformers Documentation. Retrieved from <https://huggingface.co/docs/transformers>