

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**Course Code :23CS0303**

**Project Title : AI – DRIVEN RESEARCH PAPER SIMPLIFIER**

**Semester & Year : V & III**

# Student’s Name & Id : SANJEEVNI DHIR & 11023210006

**:** **NIKUNJ NANDA & 11023210036**

**: DEV KUMAR & 11023210045**

# 

**Name of Supervisor :**

**Supervisor Signature & Date:** **Project Committee Project Coordinator**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | SRM University, Sonepat, Haryana.  3rd Year Project  **Computer Science Department** |
|  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Program:** | B.Tech -CSE |  |  |
| **Department :** | Computer Science & Engineering | **Session:** | **2023-25** |

|  |  |
| --- | --- |
| **Project Title:** | **AI – DRIVEN RESEARCH PAPER SIMPLIFIER** |
| **Project Abstract :**    **AI – DRIVEN RESEARCH PAPER SIMPLIFIER**  The exponential growth of scientific publications across diverse disciplines has made it increasingly difficult for researchers, students, and professionals to efficiently extract relevant knowledge. Traditional research papers often contain dense technical jargon, lengthy explanations, and complex structures that hinder accessibility and comprehension, especially for non-experts. To address this challenge, we propose an **AI–Driven Research Paper Simplifier**, an intelligent system that leverages Natural Language Processing (NLP) and Machine Learning techniques to automatically summarize and simplify academic articles. The system is designed to analyze research papers, extract key sections such as objectives, methodology, results, and conclusions, and present them in a concise, easy-to-understand format without compromising the scientific integrity of the content. Furthermore, the model incorporates domain-specific knowledge to preserve technical accuracy while reducing linguistic complexity. The proposed solution aims to assist students, early-career researchers, policymakers, and interdisciplinary professionals in quickly grasping essential insights from research literature. By improving accessibility, comprehension, and knowledge dissemination, the AI–Driven Research Paper Simplifier has the potential to accelerate learning, support informed decision-making, and foster cross-disciplinary collaboration. | |

|  |
| --- |
| **Goals & Objectives:**  **Goal:** To develop an AI-powered system that automatically simplifies, summarizes, and makes research papers more accessible to a wider audience while maintaining technical accuracy.  **Objectives:**   1. **Automated Summarization** – Extract the most important points from research papers (problem statement, methodology, results, and conclusions). 2. **Text Simplification** – Reduce technical jargon and linguistic complexity to make content understandable for non-experts without losing core meaning. 3. **Domain Awareness** – Integrate domain-specific NLP models to ensure scientific accuracy is preserved during simplification. 4. **User-Friendly Presentation** – Provide concise summaries in different formats (short abstracts, key points, visual highlights, etc.) to suit diverse user needs. 5. **Cross-Disciplinary Support** – Enable learners, researchers, and policymakers from various backgrounds to quickly grasp insights from research papers. 6. **Scalability & Efficiency** – Design the system to handle a large volume of academic papers across multiple fields. 7. **Enhanced Accessibility** – Bridge the gap between experts and learners by democratizing access to complex scientific knowledge. |
| **Project Scope:**  The **AI–Driven Research Paper Simplifier** aims to build an intelligent system that can automatically analyze, summarize, and simplify academic research papers across various domains. The project will employ **Natural Language Processing (NLP), Machine Learning (ML), and Deep Learning techniques** to extract key insights and present them in a concise, easy-to-understand format.  **In Scope:**   * Collecting and preprocessing research papers (PDF, text, or online repositories). * Automatic identification of important sections (abstract, methodology, results, conclusion). * Generation of concise summaries tailored to different user levels (beginner, intermediate, expert). * Simplification of technical jargon into layman-friendly language. * User interface (web/app) for uploading papers and retrieving simplified summaries. * Multi-domain support (science, engineering, medicine, social sciences, etc.). * Integration of visualization (e.g., keyword clouds, bullet-point highlights, or graphs).   **Out of Scope (initial phase):**   * Translation into multiple languages (may be included in future versions). * Voice-based interaction (text-to-speech or voice assistants). * In-depth plagiarism or citation analysis.   **Expected Outcomes:**   * A functional prototype that can read and simplify research papers with high accuracy. * Increased accessibility of academic content for students, early researchers, and policymakers. * A scalable system that can expand to multilingual and cross-disciplinary applications. |

|  |
| --- |
| **Tools / Technologies to be used in Project:**   1. **Programming Languages**    * **Python** – Core language for AI, NLP, and backend development.    * **JavaScript (React.js / Next.js)** – For frontend web application. 2. **Frameworks & Libraries**    * **NLP & AI:** o **NLTK, SpaCy** – Text preprocessing and linguistic analysis.      + **Transformers (Hugging Face)** – Pre-trained models like BERT, SciBERT, GPT, or T5 for summarization and simplification.      + **OpenAI / LLaMA models** – For advanced text understanding and summarization.    * **Machine Learning:** o **Scikit-learn** – Classical ML tasks.      + **PyTorch** – Deep learning model training and fine-tuning. 3. **Data Handling & Storage**    * **Pandas, NumPy** – Data preprocessing and manipulation.    * **MongoDB** – For storing processed papers and summaries. 4. **Backend Technologies**    * **FastAPI** – To build APIs for model inference. 5. **Frontend Technologies**    * **React.js / Next.js** – For interactive, user-friendly interface.    * **TailwindCSS** – For styling and responsive UI. 6. **Deployment & Cloud Services**    * **Docker** – Containerization of models and services.    * **IBM Cloud** – For hosting models, storage, and scalability.    * **Streamlit** – For quick prototyping and demo interface. 7. **Chatbot Integration**    * **IBM Watson Assistant** – To provide a chatbot for user interaction and answering research-related queries. **8. Other Tools**    * **Git & GitHub** – Version control and collaboration.    * **Jupyter Notebook / Google Colab** – For model experimentation and testing. • **VS Code** – Primary development environment |
|  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Supervisor Information** |  | |  | | | |
| Name: |  | |  | | | |
| Address: |  | |  | | | |
| Designation: |  | |  |  | | |
| Ph. (Res.) |  | Office: |  | | Mobile: |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Group Students:** | |  |  |  |  |
| No. | Uni. ID | Name | Email | Ph (Mob.) | Signature |
| 1 | 11023210006 | **SANJEEVNI DHIR** | 11023210006@Stu.srmuniversity.ac.in | 9310169588 | Sanjeevni Dhir |
| 2 | 11023210006 | **NIKUNJ**  **NANDA** | 11023210036@stu.srmuniversity.ac.in | 9899601598 | Nikunj Nanda |
| 3 | 11023210045 | **DEV**  **KUMAR** | 11023210045@stu.srmuniversity.ac.in | 9821947903 | Dev Kumar |