

AI

RESEARCH PAPER SUMMARIZER

PDF Text Extraction and Summarization

INTRODUCTION

A web-based application built using Django and Python.

01

UPLOAD A PDF

Users can easily upload any research paper in PDF format through a user-friendly interface.

02

EXTRACT ITS CONTENT

The system reads and processes the PDF, extracting key sections like the abstract, introduction, and conclusion.

03

CONCISE SUMMARY

An AI-powered summarization algorithm generates a clear, concise, and well-structured summary of the paper's core findings.

WHY IS IT USEFUL?

- Saves time by summarizing long research papers into 150-200 words.
- Highlights important points, making it easier for researchers and students.
- **Powered by** Huggingface's T5-Small transformer model for summarization.

PRIMARY GOALS:

- Automate the summarization of research papers.
- Ensure summaries maintain context and key insights.
- Provide an easy-to-use web interface for users.

SECONDARY GOALS:

- Ensure fast processing and high accuracy.
- Optimize chunking for long documents.
- Implement a smooth user experience with a loading animation.

TECHNOLOGY STACK

BACKEND:

- Django (Python framework for web development)
- Transformers (Huggingface library for T5-Small model)

FRONTEND:

- Django (Python framework for web development)
- Transformers (Huggingface library for T5-Small model)

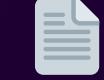
OTHER TOOLS:

- PyMuPDF (for PDF text extraction)
- Google Colab (for model testing and fine-tuning)
- Postman (for API testing)

WORKFLOW OVERVIEW

- User Upload: User uploads a PDF and adds a title.
- PDF Extraction: PyMuPDF extracts text from the uploaded file.
- Text Chunking: The text is divided into smaller chunks for summarization.
- Summarization: T5-Small generates a concise summary for each chunk.
- Output Display: The final summary is displayed on the web interface.

KEY FEATURES

-  PDF Upload: Upload any research paper in PDF format.
-  Accurate Summarization: Uses AI to summarize while preserving meaning.
-  Efficient Processing: Processes large documents in seconds.
-  User-Friendly UI: Simple web form with smooth navigation.
-  Page Loader Animation: Displays "Reading, Analyzing, Summarizing" while processing.

SCREENSHOTS

Home Page



MAIN PAGE

A screenshot of a web browser window displaying a form titled "AI Research Paper Summarizer". The URL in the address bar is "127.0.0.1:8000/form/". The form has a dark grey header with the title and a light blue gradient background. It contains two input fields: one for "Research Paper Title" with a placeholder "Enter paper title" and another for "Upload PDF File" with a "Choose File" button and a message "No file chosen". Below these is a green button with a star icon and the text "Summarize Paper". The browser interface includes standard navigation buttons at the top right.

① 127.0.0.1:8000/form/

AI Research Paper Summarizer

Research Paper Title

Enter paper title

Upload PDF File

Choose File No file chosen

⭐ Summarize Paper

RESULT PAGE

The screenshot shows a web browser window with the URL `127.0.0.1:8000/form/` in the address bar. The main content is a form titled "AI Research Paper Summarizer". The form has two input fields: "Research Paper Title" with a placeholder "Enter paper title" and "Upload PDF File" with a "Choose File" button and a message "No file chosen". Below the form is a green button labeled "Summarize Paper" with a star icon. At the bottom, there is a summary card with the title "Summary for: use of mobile devices" and a detailed description of the research findings.

127.0.0.1:8000/form/

AI Research Paper Summarizer

Research Paper Title

Enter paper title

Upload PDF File

Choose File No file chosen

Summarize Paper

Summary for: use of mobile devices

the data were collected from 20 students by means of a semi-structured interview . some subjects manifested heightened awareness relating to the advantageous role of mobile devices in their learning endeavors, their ability to reach for suitable tools and retrieve necessary information .

MODEL AND PERFORMANCE

MODEL

Huggingface T5-Small (lightweight, fast, and efficient)

PERFORMANCE

- Input Chunk Size: 3000 characters
- Output: 150 words per chunk
- Processing Time: 10-15 seconds per document

CHALLENGES AND SOLUTIONS

CHALLENGES

- Inconsistent Summaries
- Slow Processing in Django
- Memory Management for PDFs
- UI Delay during processing

SOLUTION

- Standardized chunk size and consistent model usage
- Optimized model pipeline and environment setup
- Temporary storage and cleanup after processing
- Added smooth page loader for better UX

ACHIEVEMENTS AND BENEFITS

-  Simplified research paper reading.
-  Accurate and concise summaries.
-  Improved user experience with smooth loading and clean UI.
-  Consistent results across Django and Colab environments

CONCLUSION

- Efficient, AI-driven research paper summarization.
- Enhanced user experience with fast, accurate summaries.
- Scope for further improvement and scalability.

 THANK YOU! 

This presentation was proudly presented by:

SYED USMAN ALI (GL)

SWEETI

MUHAMMAD SAAD

IQRA ARSHAD

TAHREEM RIAZ

MUSKAN