

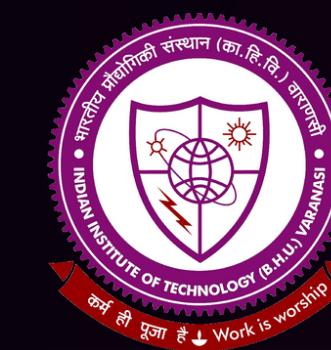
# Intelligence Gtral Shift

**COPSWEEK2025**

Team: “CodeWeavers”

*from*

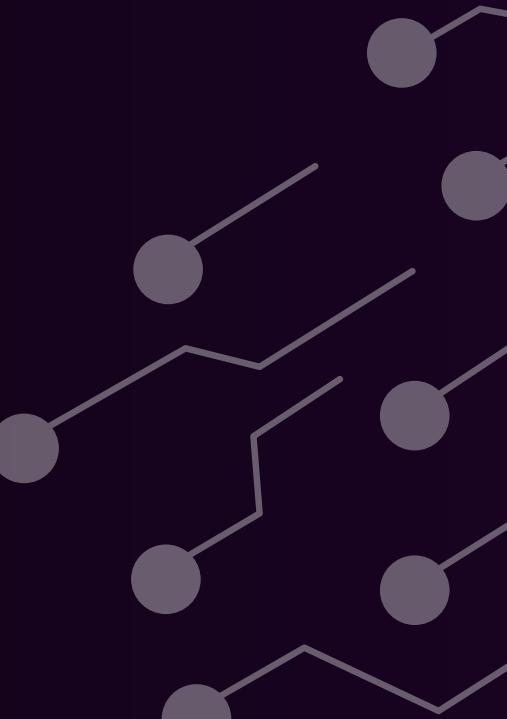
IIT BHU





## TEAM

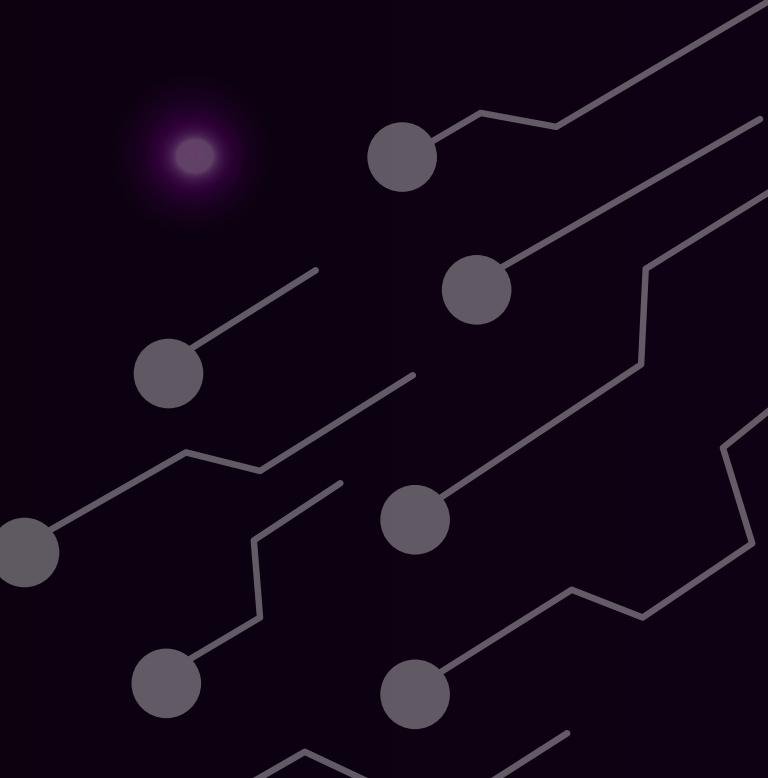
**ADITYA SAMANT  
VANSH SHRIVAS  
ABHINANDAN ROSHAN  
ADARSH PANDEY**



# CONTENTS

1. INTRODUCTION
2. TECH STACK
3. SYSTEM ARCH.
4. FEATURES
5. WORKFLOW

6. CODE WALKTHROUGH
7. FUTURE SCOPES
8. DEMO & CONCLUSIONS



# INTRODUCTION

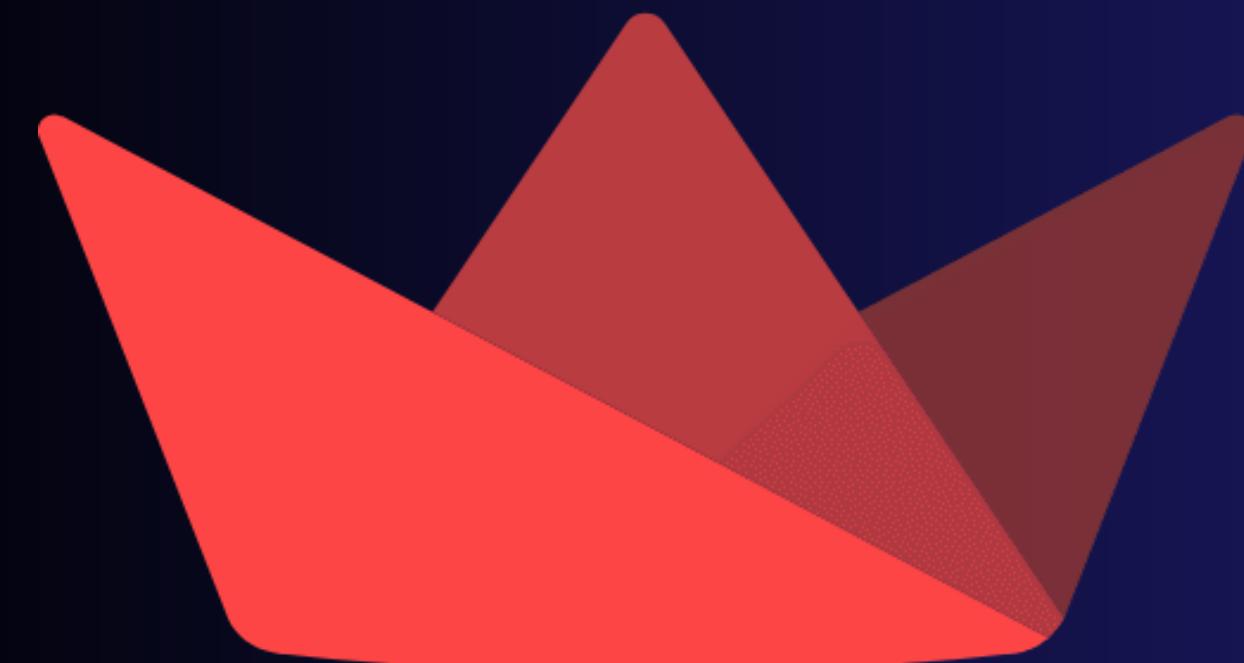
In today's fast-paced world, extracting key insights from lengthy documents is time-consuming and inefficient. Professionals, students, and researchers often struggle to quickly grasp essential information from dense legal texts, medical reports, or general documents.

To solve this, we built an AI-powered document summarizer that provides concise, category-specific summaries for general PDFs, legal documents, and medical reports.

Using Python, LangChain, and Streamlit, our web app intelligently processes uploaded documents, understands their context, and delivers structured, relevant summaries. With dedicated interfaces for each document type, it ensures precise and meaningful extraction of critical information, enhancing productivity and decision-making.

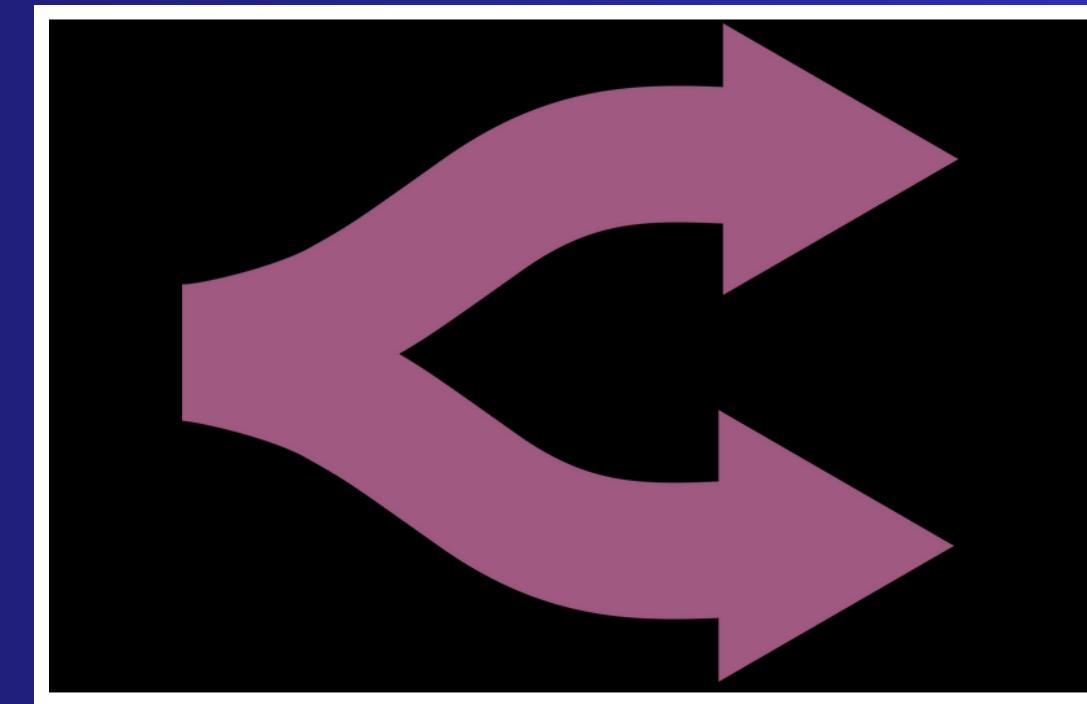
# SummarAize

# TECHSTACK

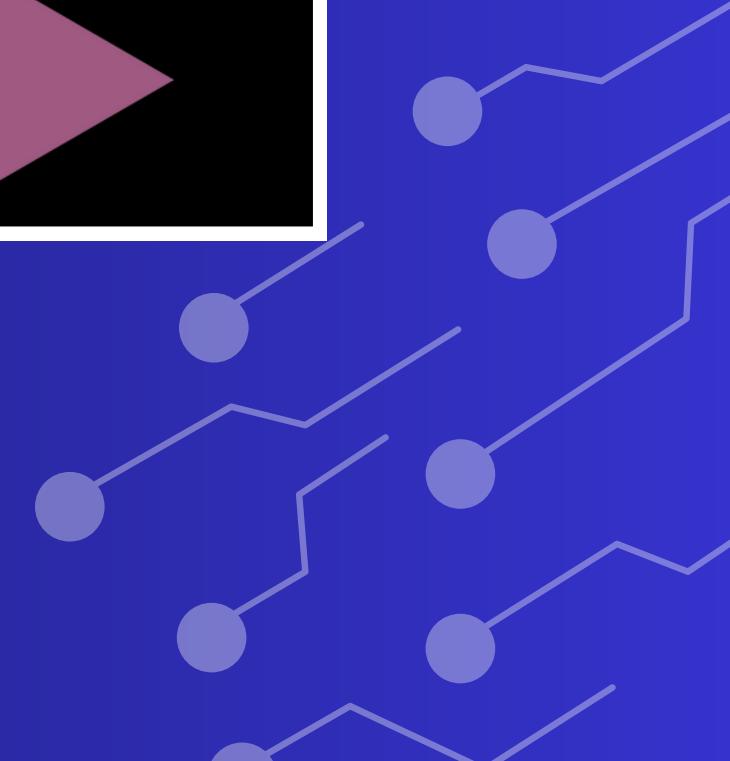


STREAMLIT

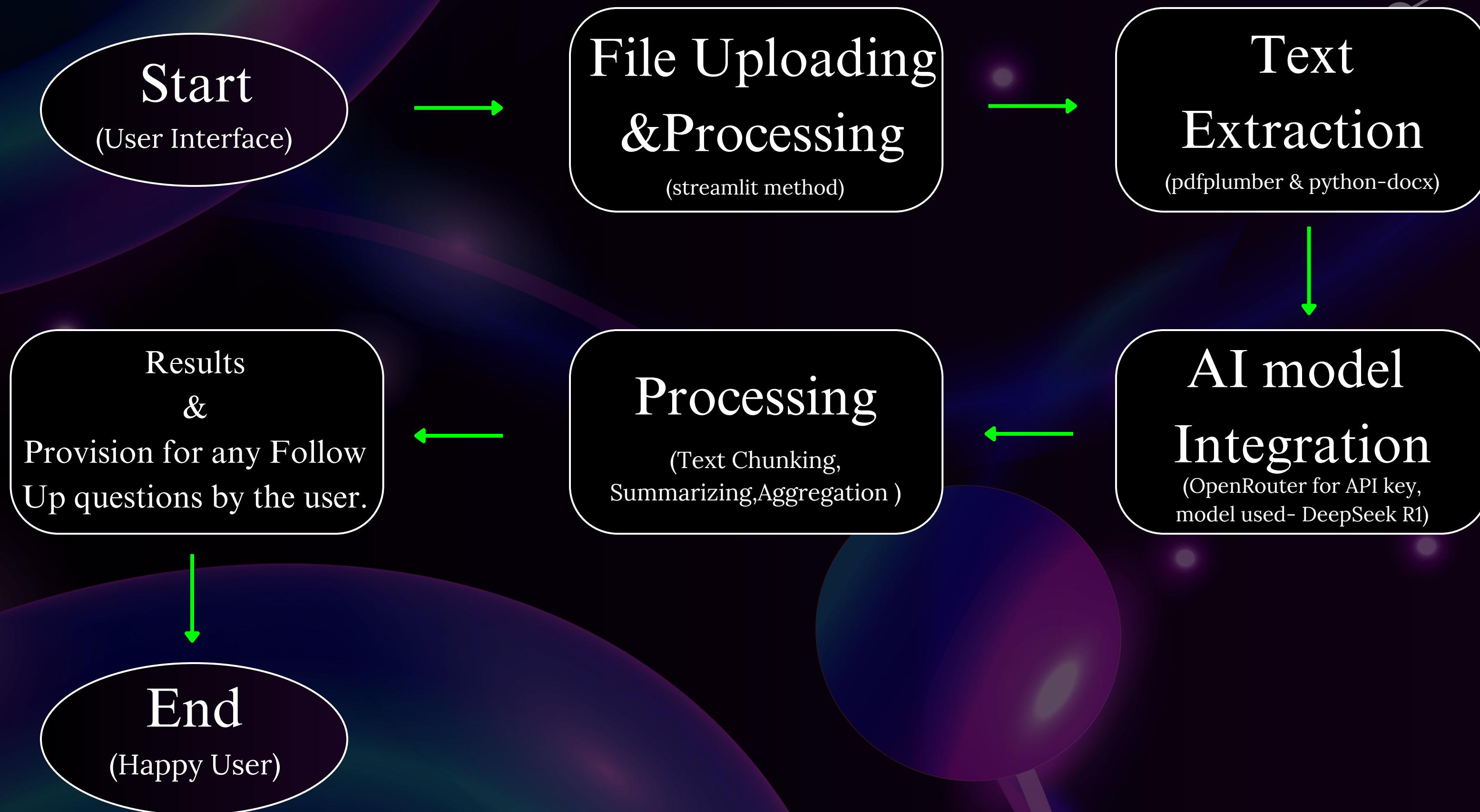
LANGCHAIN



OPENROUTER



# SYSTEM ARCHITECTURE



# FEATURES

- ✓ **AI-Powered Summarization** - Uses advanced NLP models to generate concise and meaningful summaries.
- ✓ **Category-Specific Processing** - Three dedicated interfaces for general PDFs, legal documents, and medical reports, ensuring accurate summarization based on context.
- ✓ **User-Friendly Interface** - Built with Streamlit, offering a clean and interactive experience for document upload and summary retrieval.
- ✓ **Efficient Handling of Medium-Sized Files** - Optimized processing for moderately large documents without compromising accuracy.
- ✓ **Multi-Format Support** - Accepts PDFs and DOC files, making it versatile for different types of users.
- ✓ **Fast & Automated** - Eliminates manual effort by instantly extracting key insights from lengthy documents.

# SummarAIze

Transform your documents into clear, actionable insights with AI-powered summaries

## General

### General Documents

- Quick PDF summaries
- Key insights extraction
- Comprehensive analysis

## Legal

### Legal Documents

- Legal point extraction
- Section highlights
- Terms interpretation

## Medical

### Medical Reports

- Report interpretation
- Key findings summary
- Treatment overview

# Homepage

# WORKFLOW

## SummarAize

Transform your documents into clear, actionable insights with AI-powered summaries

General

Legal

Medical

General Documents

- Quick PDF summaries
- Key insights extraction
- Comprehensive analysis

Legal Documents

- Legal point extraction
- Section highlights
- Terms interpretation

Medical Reports

- Report interpretation
- Key findings summary
- Treatment overview



## Medical Report Summarization

Upload a medical report to get a summary and ask questions about it.

Upload any PDF/DOCX



Drag and drop file here

Limit 200MB per file • PDF, DOCX

Browse files

# Drag & Drop Files

## Follow-up Questions

Ask a follow-up question based on the summary:

What is a Mental State examination to be conducted?

## Answer

Okay, let me try to figure out how to answer the question about what a Mental State Examination is, based on the provided summary. The user wants a detailed and simple explanation without medical jargon.

From the summary, the Mental State Examination was mentioned under Test Results with the date 20 June 2015. It showed disorientation, memory loss, and inability to retain information. In the medical context, a Mental State Examination (MSE) is a standard way to assess a patient's current mental abilities. The MSE typically evaluates things like appearance, behavior, mood, thoughts, perceptions, cognitive functions, and insight.

But the summary specifically mentions disorientation (time, place, date, age), memory deficits, and inability to perform basic tasks. So the answer should focus on the parts of the MSE that check cognitive functions like memory, orientation, attention, and problem-solving. The example from the summary includes checking if the patient knows current date, time, place, their age, and simple arithmetic. Also,

# Ask Follow-Ups

## Summary/Explanation



Drag and drop file here  
Limit 200MB per file • PDF, DOCX

Browse files

Medical\_Sample.pdf 301.9KB X

### Medical Summary/Explanation:

Okay, I need to summarize this medical report based on the user's instructions. Let's start by reading through the document carefully to extract the required information: diagnosis, symptoms, test results, and treatment plans.

First, the diagnosis section clearly states Dementia and Stroke. The patient has a history of hypertension, hyperlipidemia, strokes in 2005, heart issues like cardiomyopathy, cardiac failure, chronic renal disease, and a confirmed stroke in 2010 via CT and MRI scans.

Symptoms mentioned include cognitive decline, disorientation (can't tell time, place, date), memory issues (wrong age, address, Prime Minister), inability to perform simple arithmetic, recognize money, manage personal care (incontinent, needs help bathing/toileting), and lack of awareness about his medical conditions and medications.

Test results are the CT and MRI brain scans confirming the 2010 stroke. They also did mental state exams showing disorientation and cognitive impairments.

Treatment plans aren't explicitly mentioned, but the report says there's no treatment to reverse dementia, implying management of symptoms. Further tests or medications aren't detailed here.

The prognosis states that the patient is unlikely to regain capacity, with expected deterioration. So the summary needs to include all these points concisely, without extra explanations, just facts. Diagnosis:

- Dementia
- Stroke



# CODE WALKTHROUGH (I)

[GitHub Repository](#)

## HomePage Settings

```
1 import streamlit as st
2
3 # API_KEY = st.secrets["API_KEY"]
4
5 st.set_page_config(page_title="SummarAize",
6                     layout="wide")
7
8 # Redirect to Home page
9 st.markdown("""
10     <meta http-equiv="refresh" content="0;
11     url=/Home">
12     <script>window.location.href = "/Home"</
13     script>
14     """, unsafe_allow_html=True)
```

## Text Extraction

```
✓ def extract_text(uploaded_file):
    filetype=uploaded_file.name.split(".")[-1]
    if filetype=="pdf":
        | return extract_text_from_pdf(uploaded_file)
    elif filetype=="docx":
        | return extract_text_from_docx(uploaded_file)
    else:
        | return print("!!Incompatible filetype!!")

✓ def extract_text_from_pdf(uploaded_file):
    with pdfplumber.open(io.BytesIO(uploaded_file.read())) as pdf:
        return "\n" .join([page.extract_text() for page in pdf.
                           pages if page.extract_text()])
    '''the above code line runs a loop to go through every pdf page
    and the \n extracts text and joins them up'''

✓ def extract_text_from_docx(uploaded_file):
    doc=docx.Document(uploaded_file)
    return "\n" .join([para.text for para in doc.paragraphs])
```

\*\*Only selected code snippets have been presented in the slides....

# CODE WALKTHROUGH ( II )

## Prompt and Response

```
✓ def summarize_text(text, category):
    prompts = {
        "general": f"Summarize and explain this document in simple words and do
not show any internal reasoning, chain-of-thought, or explanation of
your process:\n\n{text}",
        "legal": f"Summarize this legal document (this may be a legal notice
too), highlighting key laws, sections involved, and judgements (and
further actions to be taken by the receiver in case this is a legal
notice.; if this is not a legal notice do not involve anything related
to it) and do not show any internal reasoning, chain-of-thought, or
explanation of your process:\n\n{text} and try to keep it simple and
easy to understand. Try to identify the laws and sections involved and
what the receiver should do based on the judgements and what the
judgements mean.",
        "medical": f"Summarize this medical document/report by focusing on
diagnoses, symptoms, tests results, and further treatment plans (if
mentioned) and do not show any internal reasoning, chain-of-thought, or
explanation of your process:\n\n{text} and try to keep it simple and
easy to understand. Try to identify the disease and the treatment plan
and what the patient should do based on the test results and what the
test results mean. If there are any symptoms, mention them too. If
there are any medications, mention them too. If there are any further
tests, mention them too. If there are any surgeries, mention them too.",
    }
    response = client.chat.completions.create(
        model="deepseek/deepseek-r1:free",
        messages=[...],
        temperature=1.0,
        max_completion_tokens=128000,
    )
    output = response.choices[0].message.content
    return output
```

## Code for Follow-UPS

```
# ask follow up question
def ask_followup(summary, question):
    prompt = f"Based on the following summary:\n\n{summary}\n\n .Please answer
the following Question: {question} and please do not show any internal
reasoning, chain-of-thought, or explanation of your process. Answer it in
detailed and simple words without too much legal jargons."
    response = client.chat.completions.create(
        model="deepseek/deepseek-r1:free",
        messages=[
            {"role": "system", "content": "You are an AI assistant that
provides thoughtful answers based on the given summary."},
            {"role": "user", "content": prompt}
        ],
        temperature=1.0,
        max_completion_tokens=128000,
    )
    output = response.choices[0].message.content
    return output
```

\*\*Only selected code snippets have been presented in the slides....

# FUTURE SCOPE & IMPROVEMENTS

- 🚀 **Upgrade to Faster AI Models** – Optimize summarization speed by integrating more efficient AI models beyond DeepSeek AI, ensuring quicker response times without compromising accuracy.
- 📁 **Support for More Document Types** – Extend compatibility to include **scanned images (OCR-based processing)**, **spreadsheets**, and **PowerPoint presentations** for broader usability.
- ⚡ **Performance Optimization** – Enhance processing capabilities for **handling larger and more complex documents** while maintaining efficiency.
- 🌐 **Improved UI/UX** – Upgrade the **Streamlit interface** for a more polished, user-friendly experience with better navigation and design.
- 📊 **Interactive Insights & Visualization** – Implement **keyword extraction**, **sentiment analysis**, and **graphical summaries** for deeper document understanding.
- ☁️ **Cloud Integration & API Deployment** – Enable **Google Drive**, **OneDrive**, and **Dropbox integration**, along with an **API** for seamless integration into other platforms.
- 🔒 **Enhanced Security & Privacy** – Strengthen **data encryption** and **access control** to ensure confidential document processing.

# **DEMO**

A demo video has been submitted  
alongwith the ppt, do check it out.

**Thank**  
**you**

“Code**Weavers**”