

1. PDF File Ingestion

Task: Create an API to upload PDF files and validate them.

Explanation:

PDF ingestion involves uploading a PDF document into a system for further processing.

- **Steps:**
 1. Build a REST API (e.g., using Flask or FastAPI) to allow file uploads.
 2. Validate the uploaded file (e.g., check file type, size, and content).
 3. Handle errors such as unsupported formats or corrupted files.
- **Tools:** Use Python libraries like Flask, FastAPI, or Django.
- **Sample Code Snippet:**

python

Copy code

```
from fastapi import FastAPI, File, UploadFile
```

```
app = FastAPI()
```

```
@app.post("/upload/")
```

```
async def upload_pdf(file: UploadFile = File(...)):
```

```
    if file.content_type != "application/pdf":
```

```
        content = await file.read()
```

```
        # Save or process file
```

```
        return {"filename": file.filename, "status": "Uploaded successfully"}
```

```
    return {"error": "Invalid file format"}
```