*Load a .txt or .pdf file using LangChain's TextLoader and split it using RecursiveCharacterTextSplitter. Print the total number of document chunks created. *

```
pip install langchain langchain-community pypdf
```

```
Collecting dataclasses-json<0.7,>=0.5.7 (from langchain-community)
   Downloading dataclasses_json-0.6.7-py3-none-any.whl.metadata (25 kB)
Collecting pydantic-settings<3.0.0,>=2.4.0 (from langchain-community)
   Downloading pydantic_settings-2.10.1-py3-none-any.whl.metadata (3.4 kB)
Collecting httpx-sse<1.0.0,>=0.4.0 (from langchain-community)
   Downloading httpx_sse-0.4.1-py3-none-any.whl.metadata (9.4 kB)
Requirement already satisfied: numpy>=1.26.2 in /usr/local/lib/python3.11/dist-packages (from langchain-community) (2.0.2)
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langcha
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-comm
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-communi
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-com
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-c
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-comm
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain-com
Collecting marshmallow<4.0.0,>=3.18.0 (from dataclasses-json<0.7,>=0.5.7->langchain-community)
   Downloading marshmallow-3.26.1-py3-none-any.whl.metadata (7.3 kB)
Collecting typing-inspect<1,>=0.4.0 (from dataclasses-json<0.7,>=0.5.7->langchain-community)
   Downloading typing_inspect-0.9.0-py3-none-any.whl.metadata (1.5 kB)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in /usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.72->la
Requirement already satisfied: typing-extensions>=4.7 in /usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.72->
Requirement already satisfied: packaging>=23.2 in /usr/local/lib/python3.11/dist-packages (from langchain-core<1.0.0,>=0.3.72->langcha
Requirement already satisfied: httpx<1,>=0.23.0 in /usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.17->langchain) (0.28.1
Requirement already satisfied: orjson>=3.9.14 in /usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.17->langchain) (3.11.1)
Requirement already satisfied: requests-toolbelt>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.17->langchain)
Requirement already satisfied: zstandard>=0.23.0 in /usr/local/lib/python3.11/dist-packages (from langsmith>=0.1.17->langchain) (0.23.
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langcha
Requirement already satisfied: pydantic-core==2.33.2 in /usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langchai
Requirement already satisfied: typing-inspection>=0.4.0 in /usr/local/lib/python3.11/dist-packages (from pydantic<3.0.0,>=2.7.4->langc
Collecting python-dotenv>=0.21.0 (from pydantic-settings<3.0.0,>=2.4.0->langchain-community)
   Downloading python_dotenv-1.1.1-py3-none-any.whl.metadata (24 kB)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (3
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2->langchain) (2025.8.
Requirement already satisfied: greenlet>=1 in /usr/local/lib/python3.11/dist-packages (from SQLAlchemy<3,>=1.4->langchain) (3.2.3)
Requirement already satisfied: anyio in /usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsmith>=0.1.17->langchain)
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.11/dist-packages (from httpx<1,>=0.23.0->langsmith>=0.1.17->lan
Requirement already satisfied: h11>=0.16 in /usr/local/lib/python3.11/dist-packages (from httpcore==1.*->httpx<1,>=0.23.0->langsmith>=
Requirement already satisfied: jsonpointer>=1.9 in /usr/local/lib/python3.11/dist-packages (from jsonpatch<2.0,>=1.33->langchain-core<
Collecting mypy-extensions>=0.3.0 (from typing-inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain-community)
   Downloading mypy_extensions-1.1.0-py3-none-any.whl.metadata (1.1 kB)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-packages (from anyio->httpx<1,>=0.23.0->langsmith>=0.1.1
Downloading langchain_community-0.3.27-py3-none-any.whl (2.5 MB)
                                    ━━━━━━━━━━━━━━━━ 2.5/2.5 MB 43.3 MB/s eta 0:00:00
Downloading pypdf-5.9.0-py3-none-any.whl (313 kB)
                                    ━━━━━━━━━━━━━━━━ 313.2/313.2 kB 27.9 MB/s eta 0:00:00
Downloading dataclasses_json-0.6.7-py3-none-any.whl (28 kB)
Downloading httpx_sse-0.4.1-py3-none-any.whl (8.1 kB)
Downloading pydantic_settings-2.10.1-py3-none-any.whl (45 kB)
                                    ━━━━━━━━━━━━━━━━ 45.2/45.2 kB 4.5 MB/s eta 0:00:00
Downloading marshmallow-3.26.1-py3-none-any.whl (50 kB)
                                    ━━━━━━━━━━━━━━━━ 50.9/50.9 kB 4.8 MB/s eta 0:00:00
Downloading python_dotenv-1.1.1-py3-none-any.whl (20 kB)
Downloading typing_inspect-0.9.0-py3-none-any.whl (8.8 kB)
Downloading mypy_extensions-1.1.0-py3-none-any.whl (5.0 kB)
Installing collected packages: python-dotenv, pypdf, mypy-extensions, marshmallow, httpx-sse, typing-inspect, pydantic-settings, datac
Successfully installed dataclasses-json-0.6.7 httpx-sse-0.4.1 langchain-community-0.3.27 marshmallow-3.26.1 mypy-extensions-1.1.0 pyda
```

```python
from langchain_community.document_loaders import TextLoader, PyPDFLoader
from langchain.text_splitter import RecursiveCharacterTextSplitter

# ---- Step 1: File path ----
file_path = "/content/ijepav1n3_1.pdf"  # or "example.pdf"

# ---- Step 2: Load document ----
if file_path.endswith(".txt"):
    loader = TextLoader(file_path, encoding="utf-8")
elif file_path.endswith(".pdf"):
    loader = PyPDFLoader(file_path)
else:
    raise ValueError("Only .txt or .pdf supported")

documents = loader.load()
```

```
documents = loader.load()

# ---- Step 3: Split into chunks ----
text_splitter = RecursiveCharacterTextSplitter(
    chunk_size=500,
    chunk_overlap=50
)
chunks = text_splitter.split_documents(documents)

# ---- Step 4: Print number of chunks ----
print(f"Total number of document chunks: {len(chunks)}")
```

⇥ Total number of document chunks: 25