Document Uploading Task

(9-10-2024)

Downloading different types of files for document uploading

1. .PDF Extension files (2 Files)

2. .txt Extension files (2 Files)

3. .docx Extension Files (2 Files)

4. .csv Extension Files (2 Files)

5. .xlsx Extension Files (2 Files)

6. .PPT Extension Files (2 Files)

Code Explored:

> Install Necessary Libraries

```
!pip install PyPDF2
!pip install python-docx
!pip install openpyxl
!pip install python-pptx
```

> Import neccesary Libraries

```
import os
import pandas as pd
import PyPDF2
import docx
import csv
import pptx
import openpyxl
from google.colab import files
from PIL import Image
from openpyxl import load_workbook
from pptx import Presentation
```

```python
# Dictionary to store lists of uploaded files
uploaded_files = {
    'csv': [],
    'pdf': [],
    'docx': [],
    'txt': [],
    'xlsx': [],
    'pptx': [],
    'images': []
}


# Supported file extensions
supported_extensions = {
    'csv': '.csv',
    'pdf': '.pdf',
    'docx': '.docx',
    'txt': '.txt',
    'xlsx': '.xlsx',
    'pptx': '.pptx',
    'images': ['.jpg', '.jpeg', '.png']
}


# Function to upload CSV files
def upload_csv():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.csv'):
            uploaded_files['csv'].append(filename)
```

```python
# Function to upload PDF files
def upload_pdf():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.pdf'):
            uploaded_files['pdf'].append(filename)


# Function to upload DOCX files
def upload_docx():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.docx'):
            uploaded_files['docx'].append(filename)


# Function to upload TXT files
def upload_txt():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.txt'):
            uploaded_files['txt'].append(filename)


# Function to upload XLSX files
def upload_xlsx():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.xlsx'):
            uploaded_files['xlsx'].append(filename)
```

```python
# Function to upload PPT files
def upload_ppt():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.pptx'):
            uploaded_files['pptx'].append(filename)


#Function to upload image files
def upload_image():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith(('.jpg', '.jpeg', '.png')):
            uploaded_files['images'].append(filename)


# Call the functions to upload files
upload_csv()
upload_pdf()
upload_docx()
upload_txt()
upload_xlsx()
upload_ppt()
upload_image()


#You can access the lists of uploaded files as follows:
print(uploaded_files)
```

Updated Code for uploading files and checking all the file extensions working correctly or not.

Code:

```python
# Dictionary to store lists of uploaded files
uploaded_files = {
    'csv': [],
    'pdf': [],
    'docx': [],
    'txt': [],
    'xlsx': [],
    'pptx': [],
    'images': []
}
# Supported file extensions
supported_extensions = {
    'csv': '.csv',
    'pdf': '.pdf',
    'docx': '.docx',
    'txt': '.txt',
    'xlsx': '.xlsx',
    'pptx': '.pptx',
    'images': ['.jpg', '.jpeg', '.png']
}
```

```python
# Function to check if a file is supported
def is_supported(filename):
    for file_type, extensions in supported_extensions.items():
        # Single extension case
        if isinstance(extensions, str) and filename.endswith(extensions):
            return True
        # List of extensions case (for images)
        elif isinstance(extensions, list) and any(filename.endswith(ext) for ext in extensions):
            return True
    return False


# Function to upload CSV files
def upload_csv():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.csv'):
            uploaded_files['csv']. append(filename)
        elif not is_supported(filename):
            print (f"File type not supported: {filename}")


# Function to upload PDF files
def upload_pdf():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.pdf'):
            uploaded_files['pdf']. append(filename)
        elif not is_supported(filename):
            print (f"File type not supported: {filename}")
```

```python
# Function to upload DOCX files
def upload_docx():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.docx'):
            uploaded_files['docx']. append(filename)
        elif not is_supported(filename):
            print (f"File type not supported: {filename}")


# Function to upload TXT files
def upload_txt():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.txt'):
            uploaded_files['txt'].append(filename)
        elif not is_supported(filename):
            print(f"File type not supported: {filename}")


# Function to upload XLSX files
def upload_xlsx():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.xlsx'):
            uploaded_files['xlsx'].append(filename)
        elif not is_supported(filename):
            print(f"File type not supported: {filename}")
```

```python
# Function to upload PPT files
def upload_ppt():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith('.pptx'):
            uploaded_files['pptx'].append(filename)
        elif not is_supported(filename):
            print(f"File type not supported: {filename}")


# Function to upload image files
def upload_image():
    uploaded = files.upload()
    for filename in uploaded.keys():
        if filename.endswith(('.jpg', '.jpeg', '.png')):
            uploaded_files['images'].append(filename)
        elif not is_supported(filename):
            print(f"File type not supported: {filename}")


# Call the functions to upload files
upload_csv()
upload_pdf()
upload_docx()
upload_txt()
upload_xlsx()
upload_ppt()
upload_image()
```

You can access the lists of uploaded files as follows:

print(uploaded_files)

Outputs:

☐ **save_water__poster.jpg**(image/jpeg) - 146985 bytes, last modified: 10/9/2024 - 100% done

☐ **artificial_intelligence.pptx**(application/vnd.openxmlformats officedocument.presentationml.presentation) - 1189683 bytes, last modified: 10/9/2024 - 100% done

Saving save_water__poster.jpg to save_water__poster.jpg

Saving artificial_intelligence.pptx to artificial_intelligence.pptx

☐ **Chat_with_MultiplePDFs_Mistral_7B_Instruct1.ipynb**(n/a) - 279731 bytes, last modified: 8/19/2024 - 100% done

Saving Chat_with_MultiplePDFs_Mistral_7B_Instruct1.ipynb to Chat_with_MultiplePDFs_Mistral_7B_Instruct1.ipynb

File type not supported: Chat_with_MultiplePDFs_Mistral_7B_Instruct1.ipynb

☐ **file_example_XLSX_50.xlsx**(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 7360 bytes, last modified: 10/9/2024 - 100% done

☐ **file_example_XLSX_100.xlsx**(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 9299 bytes, last modified: 10/9/2024 - 100% done

Saving file_example_XLSX_50.xlsx to file_example_XLSX_50.xlsx

Saving file_example_XLSX_100.xlsx to file_example_XLSX_100.xlsx

☐ **business-operations-survey-2022-business-finance.csv**(text/csv) - 680782 bytes, last modified: 10/1/2024 - 100% done

☐ **annual-enterprise-survey-2023-financial-year-provisional.csv**(text/csv) - 8065547 bytes, last modified: 10/1/2024 - 100% done

Saving business-operations-survey-2022-business-finance.csv to business-operations-survey-2022-business-finance.csv

Saving annual-enterprise-survey-2023-financial-year-provisional.csv to annual-enterprise-survey-2023-financial-year-provisional.csv

☐ **artificial_intelligence.txt**(text/plain) - 92488 bytes, last modified: 10/9/2024 - 100% done

☐ **data analytics.txt**(text/plain) - 129447 bytes, last modified: 10/9/2024 - 100% done

Saving artificial_intelligence.txt to artificial_intelligence.txt

Saving data analytics.txt to data analytics.txt

☐ **Four-Steps-to-Forgiveness-William-Fergus-Martin.docx**(application/vnd.openxmlformats-officedocument.wordprocessingml.document) - 277019 bytes, last modified: 10/9/2024 - 100% done

Saving Four-Steps-to-Forgiveness-William-Fergus-Martin.docx to Four-Steps-to-Forgiveness-William-Fergus-Martin.docx