



PDFSummarizer — Relatório do Projeto Final de Programação

Programa de Pós-Graduação em Informática — PUC-Rio

Aluno: Jorge Alfonso Rodriguez Mendez - **Matrícula** 2220963

Palavras-chave: Summarization, PDF, LangChain, OpenAI

Sumário

1. Descrição e Objetivos Gerais do Software	03
2. Tecnologias utilizadas	03
3. Especificação de Requisitos	03
4. Arquitetura.....	04
5. Modelo Funcional do Software.....	05
6. Manual de Utilização do Software	05
7. Cenários de Utilização	08
8. Referências	09

1. Descrição e Objetivos Gerais do Software

PDFSummarizer é um aplicativo de desktop que usa tecnologias como LangChain para resumir e consultar documentos em formato PDF. O seu principal objetivo é satisfazer as necessidades dos usuários que procuram obter de forma eficiente o conteúdo essencial de um documento, tirando partido de capacidades avançadas de síntese linguística. Além disso, o aplicativo foi desenvolvido para facilitar a formulação e execução de consultas, proporcionando uma interface intuitiva para solucionar dúvidas específicas sobre o conteúdo de arquivos PDF. Esta abordagem técnica avançada garante uma experiência ideal para usuários com requisitos específicos de análise de documentos.

Este projeto insere-se principalmente na área da Inteligência Artificial (IA), uma vez que a utilização de tecnologias como LangChain sugere a integração de capacidades avançadas de processamento de linguagem baseadas em algoritmos e modelos de aprendizagem automática.

A vida útil desta aplicação dependerá em grande parte das atualizações, da evolução tecnológica e das mudanças nas necessidades dos usuários. Para manter a competitividade e a utilidade ao longo do tempo, é crucial considerar o seguinte:

- **Atualizações e melhorias:** fornecer atualizações regulares ajuda a melhorar a funcionalidade, corrigir erros e se adaptar às tecnologias mais recentes.
- **Adaptabilidade:** para permitir fácil integração de novas tecnologias emergentes ou mudanças nas necessidades dos usuários.
- **Avaliação de Mercado:** para ajustar continuamente a aplicação de acordo com as novas demandas.

2. Tecnologias utilizadas

Este projeto foi implementado utilizando a linguagem de programação Python. Para o desenvolvimento da aplicação desktop foi utilizado o Tkinter, que é um bind da biblioteca gráfica Tcl/Tk projetada para a linguagem de programação Python. Tkinter é reconhecido como um padrão na criação de interfaces gráficas de usuário (GUIs) para aplicativos Python.

Na área de processamento de documentos PDF, foi integrado o LangChain, um framework projetado para o desenvolvimento de aplicações que dependem de modelos de linguagem. LangChain facilita a manipulação e compreensão avançadas de conteúdo textual, aproveitando modelos linguísticos especializados.

3. Especificação de Requisitos

Requisitos funcionais:

- O sistema deve ser capaz de processar arquivos PDF e gerar um resumo do conteúdo principal.
- Os usuários devem poder fazer consultas específicas sobre o conteúdo de documentos PDF.
- O aplicativo deve ser capaz de lidar com diferentes versões e estruturas de arquivos PDF.
- O aplicativo deve permitir salvar resumos e resultados de consultas para referência futura.

Requisitos não Funcionais:

- A aplicação deve ser eficiente e fornecer resultados resumidos e consultas rapidamente.
- A interface do usuário deve ser intuitiva e fácil de entender para usuários com diversos níveis de experiência.
- O aplicativo deve ser compatível com sistemas operacionais Windows.

Para rodar o projeto é imprescindível ter uma chave OpenAI e ter determinadas bibliotecas instaladas. Para simplificar o processo de instalação das bibliotecas, o projeto inclui um arquivo chamado “requirements.txt”, que lista as bibliotecas necessárias. O usuário pode instalá-los executando o seguinte comando no console:

```
pip install -r requirements.txt
```

Uma explicação detalhada deste procedimento é fornecida posteriormente na seção dedicada ao “Manual do Utilização do Software” para uma compreensão mais completa.

4. Arquitetura

O projeto adota uma estrutura de software simples composta por duas visualizações, que fazem uso de três classes principais: **Register_Key** (responsável por registrar a chave OpenAI e gerenciá-la para as vistas), **SummarizingPDF** (lida com operações de documentos, como carregamento, resumo e consultas) e **About** (representa a vista 'About'). A Figura 1 apresenta o diagrama de classes, detalhando as funcionalidades de cada uma.

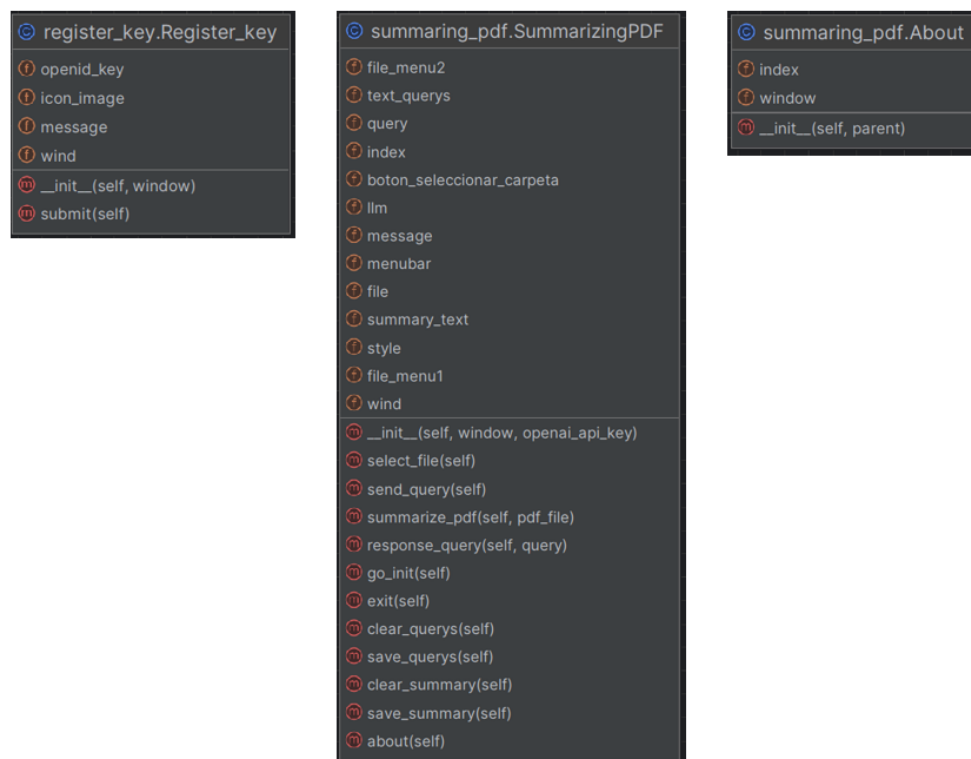


Fig.1 Diagrama de classes

5. Modelo Funcional do Software

Para discutir o modelo funcional, é mais eficaz explicá-lo por meio de um diagrama de sequência (Figura 2). Conforme observado, o primeiro passo do usuário é registrar a chave da OpenAI; em seguida, o sistema notifica se é válida ou não. Se for válida, o usuário passa para o carregamento de PDF, onde duas situações podem ocorrer: se houver um erro, o usuário é informado; caso contrário, um resumo do PDF carregado é gerado e exibido. Em seguida, o processo de consulta ao documento é iniciado, e, por último, há a opção de exportar o conteúdo. Após o registro da chave da OpenAI, o usuário sempre pode acessar informações "About".

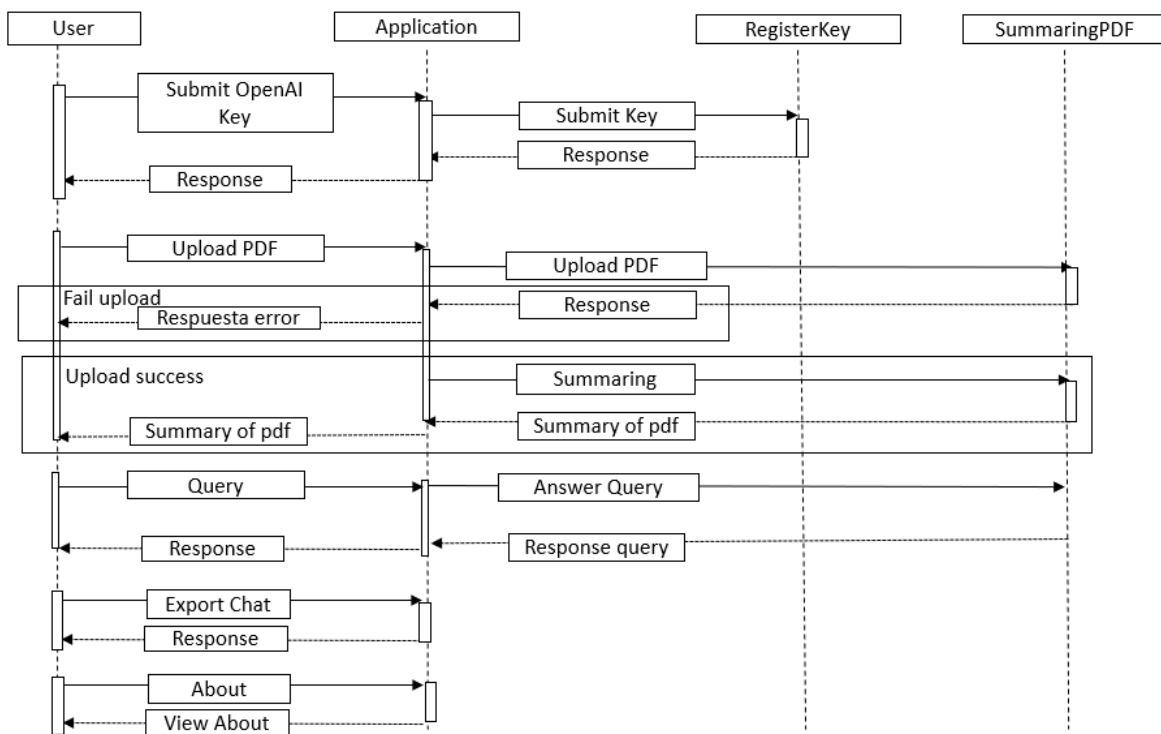


Fig.2 Diagrama de sequência

6. Manual de Utilização do Software

Para que um usuário possa utilizar este aplicativo de desktop, o primeiro passo seria baixar o projeto do GitHub a partir do seguinte endereço:

<https://github.com/Jarm1988/PDFSummarizing>

Após baixar o projeto, o próximo passo seria instalar as dependências necessárias. Para fazer isso, o usuário deve executar o seguinte comando a partir do console, garantindo que o console esteja na localização do projeto baixado:

```
pip install -r requirements.txt
```

Esse processo pode levar alguns minutos. Após a conclusão, o próximo passo seria executar o aplicativo. Para fazer isso, no mesmo console, o usuário deve inserir o seguinte comando:

```
python main.py
```

Isso iniciará o aplicativo e permitirá que o usuário comece a utilizá-lo.

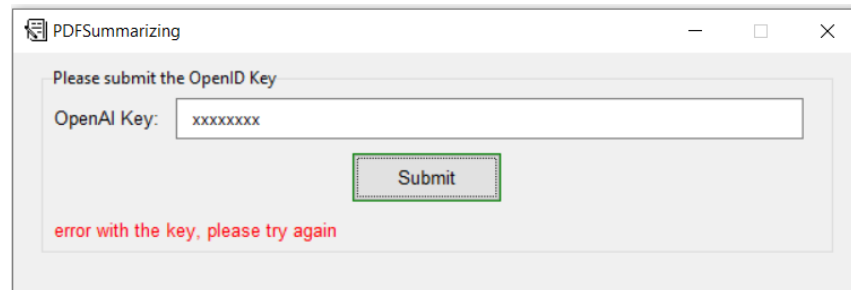


Fig.3 OpenAI Key invalida

Ao usuário é inicialmente apresentada uma tela para inserir a chave da OpenAI. Se essa chave não for válida, o sistema exibirá uma mensagem de erro (Fig. 3); no entanto, se a chave for válida, o aplicativo se abrirá mostrando uma nova interface dividida em três seções: a primeira se encarrega de carregar arquivos PDF, a segunda exibe um resumo se houver um PDF carregado, e a terceira permite realizar consultas (Fig. 4).

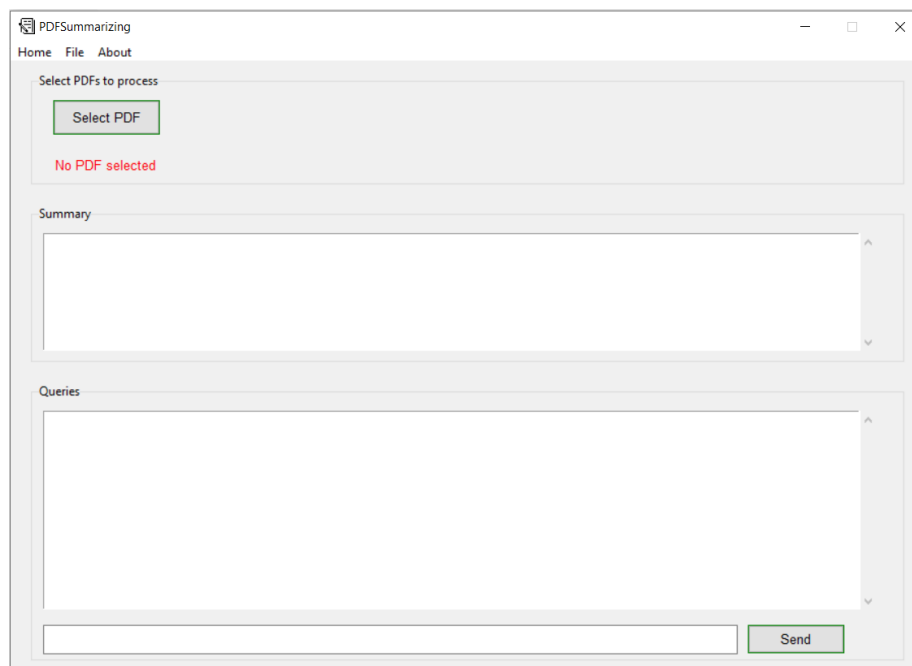


Fig.4 Janela de processamento de PDF

Nessa interface, existem três opções disponíveis na barra de menu: "Home", "File" e "About":

- **"Home"**: Essa opção permite que o usuário retorne à primeira página para inserir outra chave da OpenAI e fornece a opção de sair do aplicativo.

- **"File"**: Dentro dessa opção (Fig.5), existem funcionalidades como salvar resumos, salvar consultas com suas respostas e limpar o histórico de consultas.
- **"About"**: Essa seção fornece informações detalhadas sobre o projeto e dados de contato.

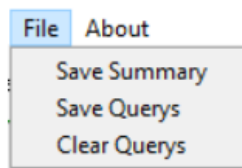


Fig.5 Menu "File"

Uma vez que o usuário está na janela de processamento de PDF, o primeiro passo é carregar um arquivo. Se a carga for bem-sucedida, automaticamente será gerado um resumo do mesmo (Fig.6), que será exibido na seção designada para esse fim. Em caso de falha na carga, será exibida uma mensagem de erro.

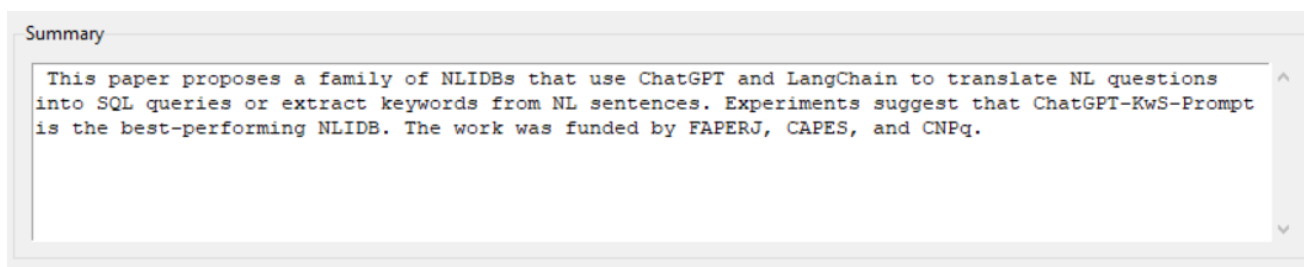


Fig.6 Resumo do documento

Após esta etapa, o usuário poderá fazer consultas à aplicação sobre o documento carregado (Fig.7). Uma vez que o PDF foi carregado e resumido, o usuário sempre terá acesso às funções anteriormente mencionadas de exportar e limpar.

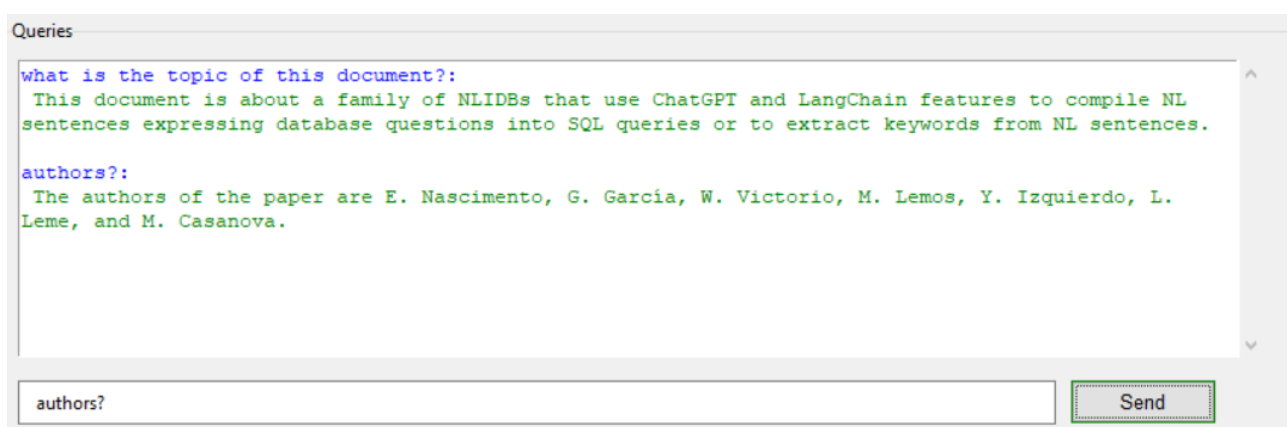


Fig.7 Consultas sobre o documento carregado

7. Cenários de Utilização

Cenários de uso de Sucesso

Cenário 1: Um pesquisador acadêmico com experiência em processamento de linguagem natural e conhecimentos avançados em Python encontra-se em um ambiente de pesquisa e precisa revisar vários documentos em PDF como parte de seu estudo. Deseja obter resumos automáticos e fazer perguntas específicas sobre o conteúdo para agilizar a análise. Consulta a documentação do software e instala a aplicação em sua máquina de trabalho. Abre a interface gráfica e fornece a chave da OpenAI. Carrega um PDF. Examina o resumo gerado automaticamente para obter uma visão geral do conteúdo. Utiliza a função de perguntas para fazer consultas específicas sobre detalhes do PDF. Avalia a precisão das respostas e utiliza os resultados para acelerar sua pesquisa.

Cenário 2: Um professor com conhecimentos básicos em programação e tecnologias educacionais trabalha em um ambiente educacional e deseja fornecer aos seus alunos uma ferramenta que permita resumir e compreender rapidamente documentos acadêmicos relevantes. Não possui experiência técnica avançada. Segue as instruções de instalação detalhadas na documentação do software. Abre a interface e insere a chave da OpenAI. Em sala de aula, mostra aos alunos como carregar um PDF e obter um resumo instantâneo. Os alunos utilizam a função de perguntas para obter informações específicas sobre o conteúdo do PDF.

Cenários com Problemas

Cenário 1: Estudante Inexperiente deseja usar o software para resumir um PDF como parte de um projeto de pesquisa. A estudante baixa o software e inicia a instalação sem consultar a documentação. Depois, a interface gráfica solicita a inserção da chave da OpenAI para ativar as funcionalidades do software. O estudante insere uma chave inválida e a interface informa imediatamente que a chave é inválida e impede a continuidade do processo. O estudante percebe que não pode prosseguir para a etapa de carregar o PDF até inserir corretamente a chave da OpenAI. A mensagem de erro na interface gráfica instrui o estudante a corrigir a chave da OpenAI para continuar. O estudante revisa as instruções, insere a chave correta e, ao fazer isso, a interface permite avançar para a etapa de carregar o PDF.

Cenário 2: Qualquer usuário, independente do conhecimento que tenha em tecnologia, pode estar trabalhando e pode aparecer o seguinte erro:

```
API error (Status Code 500): {
  "error": {
    "message": "Internal server error",
    "type": "auth_subrequest_error",
    "param": null,
    "code": "internal_error"
  }
}
```

Esse erro indica que ocorreu um erro interno no servidor OpenAI durante o processamento de sua solicitação. Esses tipos de erros geralmente são temporários e podem ocorrer por vários motivos, como problemas de infraestrutura do servidor, sobrecarga temporária ou problemas de processamento interno.

Para este tipo de erros só é recomendado aguardar uma recuperação do sistema.

8. Referências

- <https://medium.com/mlearning-ai/summarizing-and-querying-multiple-research-papers-with-langchain-fe0bf310926>
- <https://docs.python.org/es/3/library/tkinter.ttk.html>
- <https://tkdocs.com/tutorial/index.html>
- https://python.langchain.com/docs/get_started/introduction