

Problem Statement

PS 17 : Business Contract Validation- To classify content within the Contract Clauses & to determine deviations from Template & highlight them.

Unique Idea Brief(Solution)

1. **Automated Named Entity Recognition (NER):** - The tool employs state-of-the-art NER models from the Huggingface Transformers library to identify and classify key entities within contracts. Entities such as dates, parties, legal terms, and obligations are automatically highlighted, making it easier to review the document.
2. **Deviation Detection:** - One of the unique features of this tool is its ability to detect deviations from standard contract templates. By comparing the uploaded contract against predefined templates, the tool highlights discrepancies and potential issues, ensuring compliance and consistency.
3. **Dual Format Support:** - The application supports both PDF and DOCX formats, allowing users to upload contracts in their preferred format. This flexibility ensures broader usability across different user preferences and organizational requirements.
4. **Efficiency and Accuracy:** - By automating the extraction and highlighting processes, the tool significantly reduces the time required for contract review. This not only enhances efficiency but also minimizes the risk of human error, leading to more accurate contract analysis.

Features Offered

1. Automated Clause Classification:

- The tool automatically classifies different sections of the contract into predefined categories such as terms, conditions, obligations, and rights. This helps users quickly understand the structure and content of the contract.

2. Named Entity Recognition (NER):

- Leveraging advanced NLP models from the Huggingface Transformers library, the tool identifies and highlights key entities within the contract. This includes dates, parties involved, monetary values, legal terms, and more.

3. Interactive Highlighting:

- The tool provides an interactive interface where users can upload contracts and view highlighted content directly in their browser. Key entities and clauses are highlighted in different colors for easy identification.

4. Multi-format Support:

- Users can upload contracts in both PDF and DOCX formats. The tool handles text extraction and processing for both formats, providing flexibility and convenience.

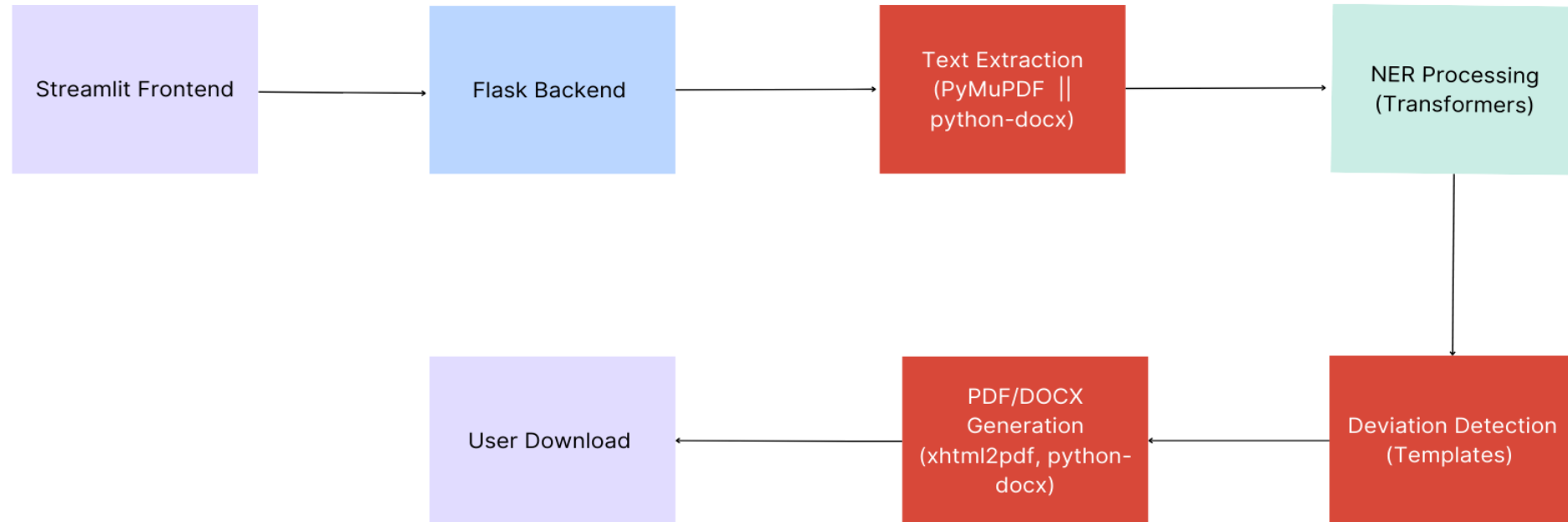
5. Downloadable Results:

- After processing the contract, users can download the highlighted contract as a PDF or DOCX file. This allows for easy sharing and record-keeping.

Processflow

1. Upload Contract:
2. Extract Text:
3. Named Entity Recognition (NER):
4. Highlight Entities and Clauses:
5. Deviation Detection:
6. Generate Summarized View:
7. Interactive Review:
8. Download Highlighted Contract:

Architecture Diagram



Technologies used

- Streamlit: For creating the interactive user interface.
- Flask: Serving as the backend framework for handling HTTP requests and processing data.
- PyMuPDF (fitz): For extracting text from PDF files.
- python-docx: For extracting text from DOCX files and generating highlighted DOCX documents.
- Transformers (Huggingface): For Named Entity Recognition (NER) to identify and classify entities within the contract.
- xhtml2pdf: For generating highlighted PDF documents from HTML content.
- Jinja2: For rendering HTML templates in the Flask backend.
- pandas: For data manipulation and processing tasks within the application

Team members and contribution

Name : Harsh Raj

Individual Project

Mentor : Prakash Kalingrao Aithal

Conclusion

The Business Contract Validation Tool provides a powerful and efficient solution for automating the review and analysis of business contracts. By leveraging advanced Natural Language Processing (NLP) techniques, such as Named Entity Recognition (NER), the tool significantly reduces the time and effort required for contract review, enhances accuracy, and ensures compliance with predefined standards.

Key Benefits:

Efficiency: Automates the tedious and time-consuming process of contract review, allowing users to focus on more strategic tasks.

Accuracy: Minimizes human errors by accurately identifying and classifying key entities and clauses within contracts.

Compliance: Ensures contracts adhere to organizational standards and legal requirements by detecting deviations from standard templates.

User-Friendly: Offers an intuitive interface that makes it accessible for users with varying levels of technical expertise.