Problem Statement

“Business Contract Validation- To classify content within the Contract Clauses & to determine deviations from Template & highlight them.”

Unique Idea Brief (Solution)

Objective: Automatically review contracts to classify clauses and identify deviations from a standard template.

Key Components

* Clause Classification: Use AI to categorize contract sections (e.g., payment terms, confidentiality).
* Deviation Detection: Compare each clause with a standard template to find differences.
* Highlighting and Reporting: Highlight deviations in the contract and generate a summary report.

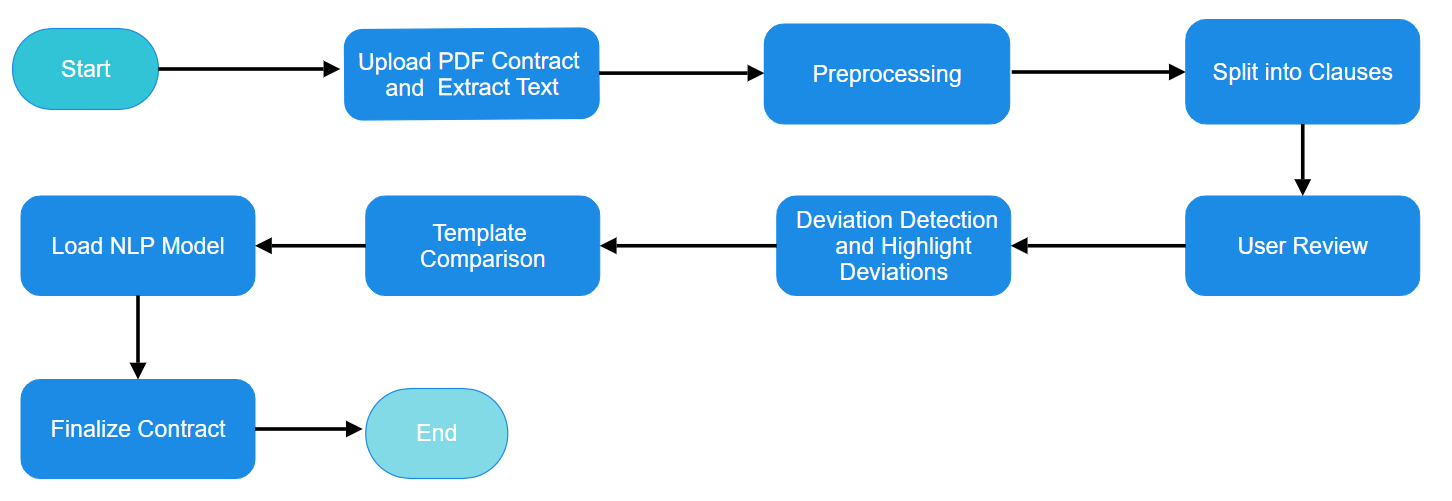
Implementation

1. Collect Data: Gather standard templates and sample contracts.
2. Train Model: Teach AI to recognize and categorize contract clauses.
3. Develop System: Build tools to process, compare, and highlight contracts.
4. Test and Validate: Ensure the system works with various contracts.
5. Deploy: Make the tool available as a web application or software.
6. This tool will speed up contract reviews, ensure template compliance, and reduce errors.

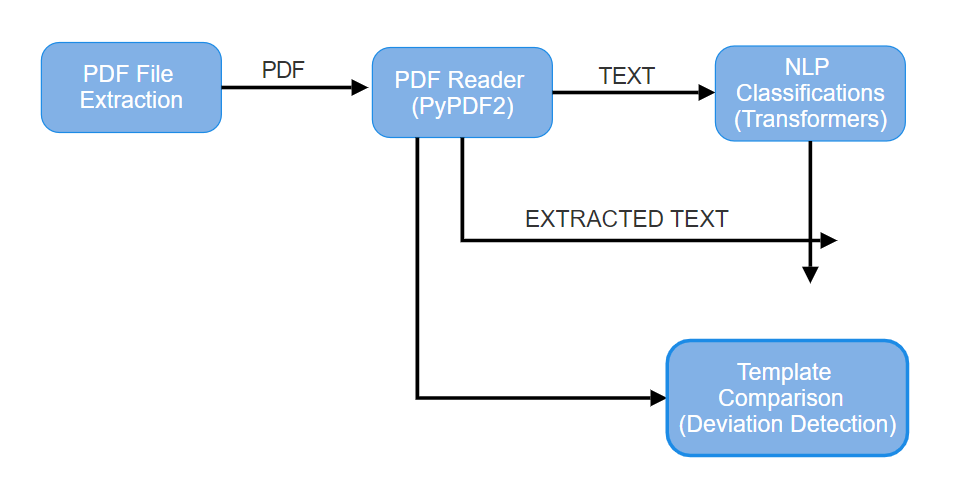
Features Offered

* Clause Classification: Automatically categorize different sections of contracts.
* Deviation Detection: Identify and highlight differences from standard templates.
* Highlighting: Visually mark deviations within the contract document.
* Reporting: Generate summary reports of all deviations and recommended corrections.
* Template Management: Store and manage standard contract templates for comparisons.

Process flow



Architecture Diagram





Technologies used

* Programming Languages: Python
* PDF Processing: PyPDF2, PyMuPDF
* Natural Language Processing (NLP): Transformers (Hugging Face)
* Machine Learning Framework: PyTorch (or optionally TensorFlow)

Team members and contribution:

1. Vedant Dange: Works on integrating the entire system, ensuring seamless interaction between the PDF processing pipeline, NLP models
2. Srushti Yawale: Tasked with implementing and fine-tuning the NLP models and Machine learning, Documentation.
3. Sohel Khan: Focuses on developing and optimizing the PDF processing pipeline,
4. Anshul Suryawanshi: Integrates the NLP models into PyTorch or TensorFlow frameworks.

Conclusion

The project aimed to automate and improve the process of validating business contracts using machine learning and PDF processing technologies. By extracting text from PDF contracts, the system uses advanced natural language processing (NLP) models like BERT to classify contract content and clauses accurately. This automation reduces manual effort, speeds up contract review times, and enhances accuracy in identifying deviations from a predefined template. The project showcases the potential of AI to handle large volumes of contracts efficiently, ensuring they align closely with established standards. Future enhancements could focus on refining the NLP models for even better accuracy, integrating with more robust PDF handling tools, and ensuring compliance with legal requirements to further enhance contract management practices.