

Monitoring Leveraged Loan Covenants – A GenAI Approach

DISSERTATION

Submitted in partial fulfillment of the requirements of the

Degree : MTech in Artificial Intelligence and Machine Learning

By

Gokul K

2022AC05398

Under the supervision of

Anand Narayanan, Vice President at Wells Fargo

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE

Pilani (Rajasthan) INDIA

March, 2025

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
WORK-INTEGRATED LEARNING PROGRAMS DIVISION
FIRST SEMESTER 2024-25

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my mentor Mr. Anand Narayanan for his invaluable guidance, insightful feedback, and continuous support throughout the development of this dissertation. His expertise in investment Banking and Machine Learning has been instrumental in shaping the direction of my research.

I extend my sincere appreciation to Prof. Vandana Jha my evaluator, for her valuable time, constructive feedback, and encouragement, which helped refine my work and enhance the quality of dissertation.

A heartfelt thanks to my friends and family for their unwavering support, patience, and encouragement throughout the journey. Their belief in me kept me motivated during the challenging times. Finally, I acknowledge all the researchers and scholars whose work inspired and guided my research on Large Language Models for the Loan Covenants.

GOKUL K
ID: 2022AC05398

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

I SEMESTER 24-25

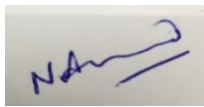
AIMLCZG628T DISSERTATION

CERTIFICATE

This is to certify that the Dissertation entitled Monitoring Leveraged Loan
Covenants – A GenAI Approach

and submitted by Mr. Gokul K ID No. 2022AC05398

in partial fulfillment of the requirements of AIMLCZG628T Dissertation, embodies
the work done by him under my supervision.



Signature of the Supervisor

Anand Narayanan
Vice President, Wells Fargo

Place: Bangalore

Date: 01-03-2025

Name
Designation

ABSTRACT

In the world of finance, leveraged loan covenants play a vital role in ensuring that borrowers stick to agreed terms, protecting the interests of lenders. However, traditional methods of monitoring these covenants are manual, time-consuming, and prone to errors, leaving room for missed deadlines and compliance risks. This project aims to bridge that gap by harnessing the power of Generative AI to streamline and enhance the covenant-monitoring process.

The project begins with retrieving publicly available covenant agreements from the SEC Edgar site. Using advanced AI models, these documents are analysed to extract key terms, milestones, and deadlines - data that is then securely stored in a centralized database. To make this information actionable, an intuitive React-based dashboard is developed, enabling real-time tracking of upcoming milestones and deadlines.

By combining AI's data processing capabilities with a user-friendly interface, this solution not only saves time but also empowers financial institutions to make informed decisions and reduce operational risks. Ultimately, this project highlights how technology can revolutionize financial risk management and make complex processes simpler and more efficient.

Key Words: Leveraged Loans, Covenant Monitoring, Generative AI, SEC Edgar, Financial Risk Management, Dashboard Development

LIST OF SYMBOLS & ABBREVIATIONS

| | |
|--------------|------------------------------------|
| AI | Artificial Intelligence |
| SEC | Securities and Exchange Commission |
| GenAI | Generative Artificial Intelligence |
| NLP | Natural Language Processing |
| ML | Machine Learning |
| RAG | Retrieval Augmented Generation |
| LLM | Large Language Model |
| UI | User Interface |
| PDF | Portable Document Format |

LIST OF FIGURES

| | | |
|-----|------------------------------------|---|
| 3.1 | Overall Architecture Diagram | 5 |
| 4.1 | User Interface Screen | 8 |

TABLE OF CONTENTS

| CHAPTER NO. | TITLE | PAGE NO. |
|----------------|---------------------------------------|-------------|
| | COVER PAGE | i |
| | ACKNOWLEDGEMENT | ii |
| | CERTIFICATE FROM SUPERVISOR | iii |
| | ABSTRACT | iv |
| | LIST OF SYMBOLS & ABBREVIATIONS | v |
| | LIST OF FIGURES | vi |
| | TABLE OF CONTENTS | vii |
| 1 | INTRODUCTION | 1 |
| | 1.1 OVERVIEW | 1 |
| | 1.2 PROBLEM STATEMENT | 3 |
| | 1.3 OBJECTIVES | 3 |
| 2 | LITERATURE SURVEY | 4 |
| 3 | SYSTEM ARCHITECTURE | 7 |
| 4 | IMPLEMENTATION AND RESULTS | 9 |
| 5 | CONCLUSION | 13 |
| | DIRECTIONS FOR FUTURE WORK | 14 |
| | BIBLIOGRAPHY / REFERENCES | 15 |
| | APPENDICES | 17 |

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Loans are a major driver of economic growth, helping businesses expand, create jobs, and fund innovations. Leveraged loans, which are often given to companies already carrying significant debt, are a high-risk, high-reward tool for growth. To ensure these loans are handled responsibly, they include covenants - rules that borrowers must follow to keep their finances in check. These covenants might require a company to avoid taking on more debt or to maintain specific financial ratios. By doing so, they protect lenders from losing money and help borrowers stay financially stable. If these rules aren't followed or monitored carefully, both sides can face major financial trouble, which could also affect the broader economy [1][2]

Traditionally, monitoring covenants is a tedious and time-consuming job. Analysts manually read through long, complex documents to find the important terms and deadlines. This method is slow and prone to human error, especially when handling a large number of agreements. Missing key details can lead to costly mistakes [1]. Some banks have adopted rule-based systems to automate parts of this process. While helpful, these systems are rigid and can't adapt to complex or unusual agreements. They follow fixed rules, making them ineffective when loan documents are written in different styles or contain ambiguous language. This limits their usefulness in today's dynamic financial world [2][3]

Currently, most banks use a mix of manual work and basic tools to manage covenants. They often store loan documents in software systems and use simple alerts to track deadlines. These systems are better than manual-only processes but don't provide a complete view of all the rules and risks involved. Because the tools aren't fully connected, banks often miss how different parts of a covenant might

affect one another. This can lead to errors, penalties, and even damaged relationships with borrowers. As agreements become more complicated, these traditional tools struggle to keep up, making it clear that smarter, more integrated solutions are needed [3][4]

Generative AI (GenAI) offers a powerful solution to these problems. Unlike traditional systems, GenAI can process large volumes of complex documents quickly and accurately. Using natural language processing (NLP), it can read and understand loan agreements, extracting critical terms, deadlines, and milestones [4]. This technology isn't just faster; it's also more flexible. It can handle different document formats and even complex legal language. When combined with tools like dashboards, GenAI allows banks to see all important covenant details in one place, enabling better decision-making. Research shows that using AI reduces errors and improves efficiency, making it ideal for managing today's financial complexities [5]

This project introduces a GenAI-powered system to automate covenant monitoring. The process starts by collecting loan agreements from sources like the SEC Edgar database. Advanced AI models analyze these agreements, extracting key details like rules, deadlines, and milestones. This information is stored in a central database, which is connected to a React-based dashboard. The dashboard gives banks a clear, real-time view of all covenant details, helping them manage risks more effectively. Unlike manual methods, this system is fast, accurate, and scalable, making it a smarter way to monitor loan covenants in today's complex financial environment [1][5]

1.2 PROBLEM STATEMENT

Monitoring loan covenants, particularly in leveraged loans, is a critical task for financial institutions to ensure borrower compliance and manage risks effectively. However, current methods rely heavily on manual processes and rule-based systems, which are time-consuming, error-prone, and unable to scale with the increasing complexity and volume of financial agreements. Manual reviews often result in missed compliance deadlines or overlooked terms, while rule-based systems fail to handle the nuances of unstructured legal documents and evolving financial conditions. These inefficiencies lead to increased operational costs, compliance risks, and reputational damage for lenders. Therefore, there is a pressing need for an automated, intelligent solution that can enhance the speed, accuracy, and scalability of covenant monitoring. GenAI offers a promising approach to address these challenges. By leveraging advanced NLP and ML, GenAI can process complex legal agreements, extract critical information, and provide actionable insights. The integration of GenAI with user-friendly tools like dashboards can empower financial institutions to proactively manage risks, reduce operational inefficiencies, and improve compliance in real-time.

1.3 OBJECTIVES

- To retrieve covenant agreements from publicly available sources like SEC Edgar
- To utilize Generative AI models to extract key terms, milestones, and deadlines from the retrieved documents
- To design and develop a centralized database for storing extracted covenant information
- To create a React-based dashboard UI for real-time tracking and monitoring of covenant milestones

CHAPTER 2

LITERATURE SURVEY

Valeri V. Nikolaev et.al [6] discuss that private information covenants play a vital role in improving transparency between borrowers and lenders. By requiring borrowers to disclose key financial details, such as projections and interim reports, these covenants allow lenders to monitor financial health more effectively and reduce uncertainty. While the benefits of such transparency are evident, the associated costs—like compliance efforts and potential unintended consequences—remain less understood. Exploring these tradeoffs sheds light on the delicate balance between ensuring accountability and managing the practical challenges of financial oversight, opening new opportunities for refining monitoring practices

Richard Carrizosa et.al [7] identifies some loan agreements include covenants that require borrowers to share private financial information, such as future projections and monthly historical reports, with lenders. These covenants are often included in situations where they enhance monitoring and improve lenders' oversight. Research suggests that such covenants are linked to frequent contract amendments, showing their active role in adapting agreements as conditions change. Additionally, lenders may use this private information in secondary loan markets, leveraging insights to make informed decisions. Interestingly, the two types of covenants—projected and historical—serve different purposes, with distinct factors influencing their use and varying effects on financial agreements. This highlights how tailored financial disclosures can strengthen lender-borrower relationships and improve loan management strategies

Ervin L. Black et.al [8] proposed debt covenants and federal monitoring both serve to limit banks' decision-making freedom, but they may interact in interesting

ways. During a period of increased federal monitoring (1979–1984), research suggests that banks reduced their reliance on debt covenants in loan agreements. This shift indicates that bank shareholders may have leveraged the overlap between regulatory oversight and debt covenants to minimize agency costs. Notably, the study found a decline in both the number of debt issues with such covenants and the overall debt subject to them during this time. In contrast, no similar decline was observed for non-banking firms or in subsequent periods for banks, highlighting how external monitoring can influence covenant usage in unique ways for financial institutions. This sheds light on the dynamic relationship between regulatory scrutiny and contractual provisions in shaping financial practices

Retrieval-Augmented Generation (RAG) is a natural language processing (NLP) approach that enhances text generation by retrieving external knowledge before generating responses. Introduced by Lewis et al. [9], RAG integrates retrieval-based and generative models, improving factual accuracy and reducing hallucinations. Unlike traditional generative models that rely only on pre-trained parameters, RAG dynamically retrieves documents from a knowledge base, enriching response generation [10]. Studies indicate its effectiveness in domains such as question answering, summarization, and conversational AI, where retrieved context significantly improves coherence and factual correctness [11]. The integration of dense retrieval mechanisms like DPR [12] enhances retrieval efficiency, making RAG a critical development in knowledge-intensive AI applications. Its adaptability in leveraging large-scale knowledge bases ensures it remains a vital tool for reducing misinformation and improving AI-driven text generation

The effectiveness of RAG is dependent on its retrieval mechanism, which influences the relevance and coherence of generated responses. Sparse retrieval techniques like TF-IDF and BM25, compared to dense retrieval models like FAISS

and DPR, impact knowledge augmentation and retrieval precision [12]. Khattab & Zaharia [13] introduced ColBERT, a late-interaction retrieval method that enhances document selection accuracy, improving RAG's overall performance. Challenges such as incorrect retrieval, leading to misinformation, have been addressed through reinforcement learning with human feedback and prompt engineering techniques. Hybrid retrieval methods integrating structured (knowledge graphs) and unstructured data sources have been explored for improving domain-specific applications, particularly in medical and legal AI systems [14]. Advancements in retrieval mechanisms and filtering techniques continue to shape RAG's ability to generate reliable and context-aware responses across diverse fields

RAG has been widely applied in domains such as finance, healthcare, and legal industries, where accurate and contextual information retrieval is crucial. Financial NLP applications leverage RAG for analyzing economic policies and summarizing regulatory documents [11]. In healthcare, RAG supports medical question answering and clinical decision-making, ensuring alignment with validated medical literature [14]. Additionally, RAG enhances explainable AI, improving governance and compliance monitoring in corporate environments. Future research focuses on refining RAG's integration with multi-modal data, enabling cross-lingual and multi-source retrieval for improved contextual understanding. Advancements in indexing techniques and continual learning models aim to enhance retrieval latency and output accuracy, solidifying RAG's role as a cornerstone in AI-driven knowledge augmentation. As AI evolves, optimizing RAG for domain-specific applications remains a key research priority

CHAPTER 3

SYSTEM ARCHITECTURE

This chapter deals with the overall system architecture from downloading the publically available covenant document from SEC Edgar site and load the covenant agreement into the project UI. The Color patterns are followed in the architecture diagram to depict the novelty of the project which is shown in the figure 3.1.

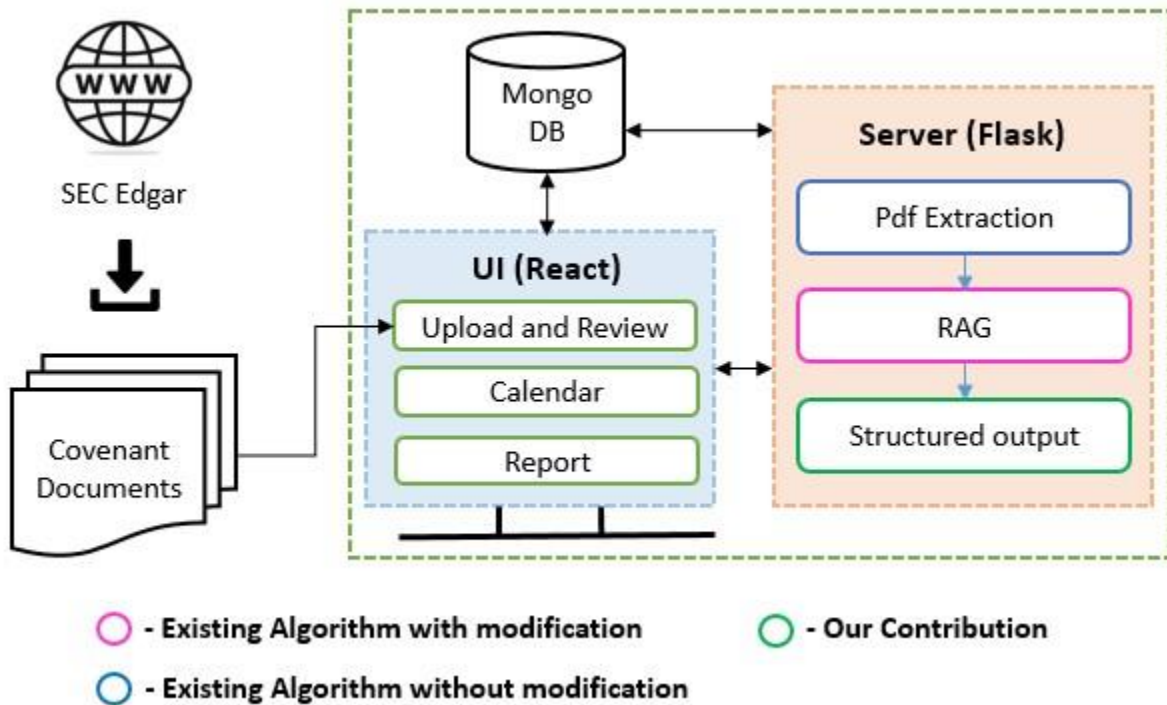


Figure 3.1 Overall Architecture Diagram

The system is designed to help users manage and track loan covenants by using an easy-to-navigate interface and robust AI tools. Users can upload documents from various sources, including SEC EDGAR or local files, through a React-based interface built with TypeScript. The interface provides three main features: document upload and review, a calendar to track important covenant deadlines, and reports to display extracted information in a structured format. The uploaded data is stored in a MongoDB database, which acts as the central storage system for all the

documents and extracted details. MongoDB ensures efficient data retrieval and supports scalability as the system grows. The system is designed to be simple, with an intuitive layout, allowing users to easily upload files, track deadlines, and access reports. Overall, the system focuses on providing a user-friendly experience while ensuring accurate and efficient tracking of loan covenants.

The backend of the system, built with Flask, handles the heavy lifting of processing the uploaded covenant documents. After a user uploads a document, the system employs Retrieval-Augmented Generation (RAG) to extract relevant information from the PDFs. RAG retrieves key content from the document and uses generative AI to ensure the extracted data is both accurate and contextually relevant. This method allows the system to efficiently capture details like financial ratios, compliance requirements, and other covenant terms. The extracted data is then structured and stored in the MongoDB database, allowing easy retrieval and display in the user interface. This integration of AI helps automate the extraction process and enhances accuracy, reducing manual errors. The structured information is made available to users in the form of reports, allowing them to track covenant compliance effectively. The backend ensures the system runs smoothly and efficiently handles large volumes of documents.

CHAPTER 4

IMPLEMENTATION AND RESULTS

4.1 IMPLEMENTATION

The implementation of the Monitoring Leveraged Loan Covenants – A GenAI Approach involves several key stages that ensure efficient retrieval, processing, and presentation of covenant agreements. The system begins with the retrieval of covenant agreements from publicly available sources such as the SEC Edgar database. A web scraping mechanism extracts structured data, and the extracted documents are stored in MongoDB for efficient retrieval. The pre-processing stage involves cleaning the extracted text using regular expressions (re) and structuring financial terms using Natural Language Processing (NLP) with spaCy. The text is further tokenized and chunked into manageable segments (`CHUNK_SIZE = 450`) for effective retrieval and processing. This structured approach ensures that the system accurately captures financial terms and prepares them for further analysis.

The system integrates Retrieval-Augmented Generation (RAG), which combines retrieval-based and generative models to enhance covenant extraction. The retrieval module, powered by SentenceTransformers, ensures that the most relevant clauses are fetched with high semantic similarity. The generative module, leveraging OpenAI's LLMs and integrated via the ollama API, reformulates extracted clauses into structured summaries. This ensures that the extracted covenant details are both contextually relevant and easily interpretable. Additionally, PyPDF2 is used to process loan agreements in PDF format, enabling seamless text extraction from complex financial documents. By combining retrieval and generative models, the system effectively identifies and structures covenant clauses, significantly improving financial risk analysis.

The backend, built using Flask, handles document processing, retrieval, and real-time notifications. API endpoints facilitate document uploads, retrieval, and database interactions. MongoDB serves as the storage solution, allowing efficient management of extracted covenant data. The system employs PyTorch-based embeddings for optimized query performance, ensuring low latency and high responsiveness. The React-based frontend with Type Script provides a user-friendly interface where users can upload documents, review extracted clauses, and track key milestones. The calendar view displays covenant deadlines, enabling users to monitor compliance in real-time. Additionally, the report generation feature provides downloadable insights, ensuring that financial analysts can make data-driven decisions efficiently. This comprehensive approach enhances the usability and effectiveness of covenant monitoring.

4.2 RESULTS:

The proposed system has effectively streamlined the monitoring of leveraged loan covenants by automating document analysis and compliance tracking. Traditional manual covenant monitoring is labour intensive and requires financial analysts to sift through extensive legal documents, a process that is prone to errors and inefficiencies. By implementing Generative AI techniques, this system has significantly reduced the time required for document review, transforming a multi-hour manual process into an automated workflow that completes within seconds. The use of Retrieval-Augmented Generation (RAG) has enhanced the accuracy and contextual relevance of extracted covenant clauses, making compliance assessment more reliable and structured.

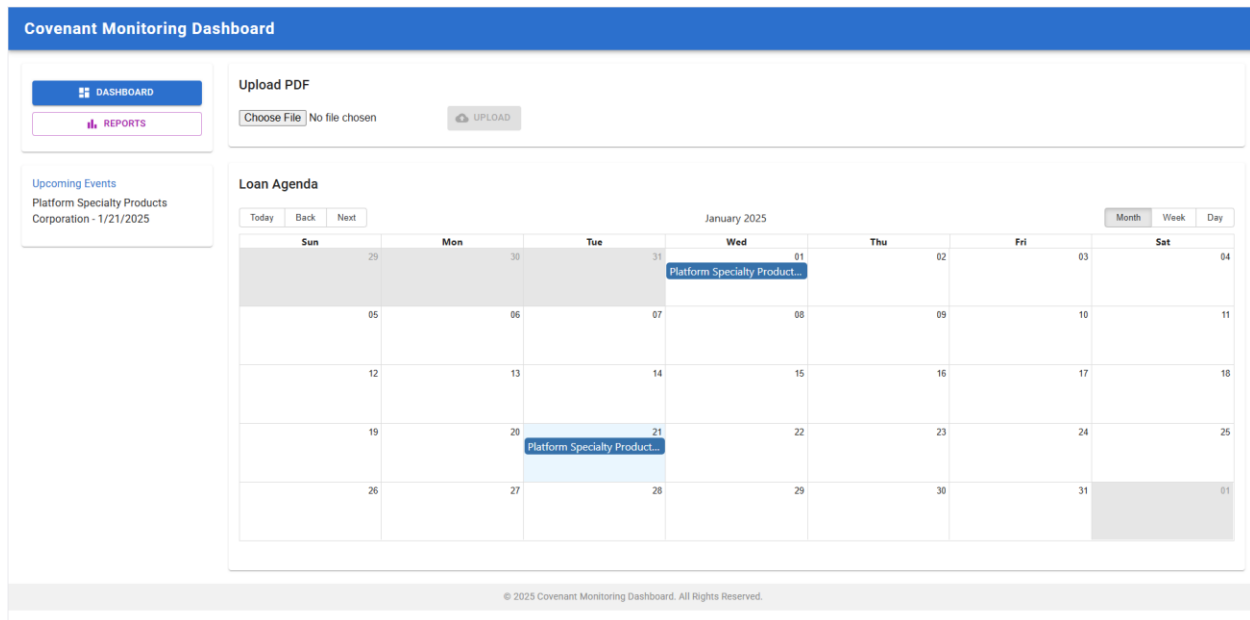


Figure 4.1 User Interface Screen

The screen on the fig 4.1 is the User Interface developed for the active tracking of the extracted covenants. The monitor provides upload functionality to upload the new covenant agreement into the system and then server process the pdf document and return the extracted details for the user review. Then the user can review the extracted details and edit or confirm the details to store it in the database. Then the covenant event will be appearing in the calendar. The report option allows the user to download the upcoming events. To enhance usability, the system allows users to review the parsed data and confirm its accuracy. This step ensures that only refined and relevant information is saved in the centralized MongoDB database. The database serves as a secure and organized repository, storing extracted covenant information for easy retrieval and analysis in the future. Another key feature implemented so far is the calendar functionality integrated into the UI. Once the user confirms the data, the system updates the calendar with critical covenant deadlines and milestones. This interactive feature provides users with a clear and accessible way to track important dates, improving the overall monitoring experience.

In terms of operational efficiency, the system provides real-time tracking of covenant milestones through an intuitive dashboard. The integration of a calendar-based monitoring feature allows financial professionals to proactively manage obligations and mitigate compliance risks. The system's ability to extract and structure financial clauses enables organizations to minimize errors and improve decision-making. Compared to traditional manual efforts, the reduction in time complexity and the elimination of repetitive tasks have led to increased productivity and a more scalable approach to financial monitoring. By ensuring consistent accuracy and structured reporting, the system demonstrates the potential of AI-driven automation in modernizing financial risk management and regulatory compliance

CONCLUSION

The implementation of an AI-driven system for monitoring leveraged loan covenants presents a transformative approach to automating financial compliance. Traditional covenant monitoring has long been a labour intensive and error-prone process, requiring financial professionals to manually review complex legal agreements to ensure regulatory adherence. This research project introduces a Generative AI-powered solution that significantly improves efficiency by automating key aspects of covenant extraction, tracking, and analysis. By leveraging Retrieval-Augmented Generation (RAG), the system accurately identifies critical clauses and presents structured insights, thereby reducing reliance on manual labour and minimizing the risks associated with human error. The seamless integration of natural language processing (NLP) techniques ensures that financial institutions can process large volumes of loan agreements in a fraction of the time required for conventional review methods. Additionally, the system provides real-time tracking capabilities, enabling organizations to proactively manage compliance obligations and reduce operational bottlenecks. The introduction of an intuitive, interactive dashboard further enhances usability, allowing analysts to review, edit, and confirm extracted information efficiently. This shift from manual document analysis to AI-powered automation demonstrates the potential of machine learning and NLP in streamlining regulatory workflows, ensuring greater accuracy, and enhancing financial risk assessment. While the system has proven effective in reducing time complexity and improving scalability, continuous refinements in retrieval accuracy, interpretability, and multi-document processing will further strengthen its role in financial governance. This research underscores the growing importance of AI in regulatory compliance and sets the stage for broader adoption of intelligent automation in financial risk management.

DIRECTIONS FOR FUTURE WORK

While the current system provides a robust solution for monitoring leveraged loan covenants, future enhancements will focus on expanding its capabilities. One major area for improvement is multi-document processing, allowing the system to analyse multiple loan agreements simultaneously for cross-referencing covenant terms. Additionally, real-time risk analysis can be integrated to provide alerts on potential covenant breaches based on market conditions and borrower financials. The incorporation of explainable AI (XAI) methods will enhance transparency, ensuring that AI-generated insights are interpretable and auditable by financial analysts. Lastly, extending support for multiple languages and jurisdictions will enable global financial institutions to leverage the system for broader compliance monitoring. These improvements will further refine the system's ability to handle complex financial documents, making it an indispensable tool for modern financial risk management.

BIBLIOGRAPHY / REFERENCES

1. R. Y. Maragheh et al., "LLM-TAKE: Theme-Aware Keyword Extraction Using Large Language Models," in 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 2023, pp. 4318-4324, doi: 10.1109/BigData59044.2023.10386476
2. D. Baviskar, S. Ahirrao, V. Potdar, and K. Kotecha, "Efficient Automated Processing of the Unstructured Documents Using Artificial Intelligence: A Systematic Literature Review and Future Directions," IEEE Access, vol. 9, pp. 72894-72936, 2021, doi: 10.1109/ACCESS.2021.3072900
3. Acuity Knowledge Partners, "Machine Learning and Covenant Monitoring," Acuity Insights, 2021. Available: www.acuitykp.com.
4. Virtusa Insights, "Mitigating Risks with AI-Led Covenant Management," 2022. Available: www.virtusa.com
5. D. Baviskar et al., "AI and Financial Risk Management: Enhancing Monitoring Capabilities," IEEE Transactions on Systems, Man, and Cybernetics, vol. 51, no. 5, pp. 3425-3441, 2023, doi: 10.1109/TSMC.2023.3054267
6. Valeri V. Nikolaev, "Discussion of Borrower private information covenants and loan contract monitoring", Journal of Accounting and Economics, Volume 64, Issues 2–3, 2017, Pages 340-345, ISSN 0165-4101, <https://doi.org/10.1016/j.jacceco.2017.08.002>
7. Richard Carrizosa, Stephen G. Ryan, "Borrower private information covenants and loan contract monitoring, Journal of Accounting and Economics", Volume 64, Issues 2–3, 2017, Pages 313-339, ISSN 0165-4101, <https://doi.org/10.1016/j.jacceco.2017.05.004>
8. Ervin L. Black, Thomas A. Carnes, Michael Mosebach, Susan E. Moyer, "Regulatory monitoring as a substitute for debt covenants", Journal of Accounting and Economics, Volume 37, Issue 3, 2004, Pages 367-391, ISSN 0165-4101, doi.org/10.1016/j.jacceco.2004.01.001
9. P. Lewis, E. Perez, A. Piktus, Fabio Petroni, et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks", Advances in Neural Information Processing Systems, vol. 33, pp. 9459-9474, 2020
10. G. Izacard and E. Grave, "Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering", Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, pp. 874-880, 2021

11. K. Guu, K. Lee, Z. Tung, P. Pasupat, and M. Chang, "REALM: Retrieval-Augmented Language Model Pre-Training", Proceedings of the 37th International Conference on Machine Learning, vol. 119, pp. 3929-3938, 2020
12. V. Karpukhin, B. Oguz, S. Min, et al., "Dense Passage Retrieval for Open-Domain Question Answering", Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 6769-6781, 2020
13. O. Khattab and M. Zaharia, "ColBERT: Efficient and Effective Passage Search via Contextualized Late Interaction over BERT", Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 39-48, 2020
14. F. Petroni, T. Rocktäschel, P. Lewis, et al., "Language Models as Knowledge Bases?", Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 2463-2473, 2019

APPENDICES

A.1 System Requirements

The project was developed and tested using the following system configurations:

- Hardware: 16GB RAM, Intel Core i7 Processor, 512GB SSD, 4GB Graphics Card
- Operating System: Windows 11 / Ubuntu 20.04
- Software Dependencies: Python 3.8+, Flask, React.js, MongoDB, OpenAI API
- Development Tools: Jupyter Notebook, VS Code, Postman

A.2 Data Sources

The system utilizes publicly available financial data for covenant analysis, primarily from:

- SEC Edgar Database – Official financial filings and loan agreements
- Financial Institution Reports – Sample documents from banks and regulatory bodies

A.3 AI Models and Techniques Used

- Natural Language Processing (NLP) for text parsing and extraction
- Retrieval-Augmented Generation (RAG) for contextual analysis
- Sentence Transformers for document similarity assessment
- PyPDF2 for PDF text extraction

A.4 Limitations

- The system currently supports English-language documents only.
- Financial jargon and unconventional legal terminologies may require further model fine-tuning.
- Real-time updates on financial markets are not integrated in the current version

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
WORK-INTEGRATED LEARNING PROGRAMS DIVISION
FIRST SEMESTER 2024-25

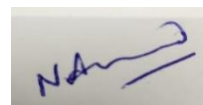
REPORT CHECK LIST

Check list of items for the Final report

- | | | |
|------|--|---|
| a) | Is the Cover page in proper format? | Y |
| b) | Is the Title page in proper format? | Y |
| c) | Is the Certificate from the Supervisor in proper format? | Y |
| d) | Is Abstract included in the Report? Is it properly written? | Y |
| e) | Does the Table of Contents page include chapter page numbers? | Y |
| f) | Does the Report contain a summary of the literature survey? | Y |
| i. | Are the Pages numbered properly? | Y |
| ii. | Are the Figures numbered properly? | Y |
| iii. | Are the Tables numbered properly? | Y |
| iv. | Are the Captions for the Figures and Tables proper? | Y |
| v. | Are the Appendices numbered? | Y |
| g) | Does the Report have Conclusion / Recommendations of the work? | Y |
| h) | Are References/Bibliography given in the Report? | Y |
| i) | Have the References been cited in the Report? | Y |
| j) | Is the citation of References / Bibliography in proper format? | Y |



GOKUL K
Student ID: 2022AC05398
01-03-2025



Anand Narayanan
Supervisor
01-03-2025