IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155141

AI-Powered Automated Business Report Generator

Dr. T. R. Muhibur Rahman[1], Haseeba Kouser [2], Nazia Taj S [3], Sai Vaishnavi D [4]

Dept Of CSE, Ballari Institute of Technology and Management, Ballari, India[1]

7th Semester B.E.(CSE), Ballari Institute of Technology and Management, Ballari, India[2-4]

Abstract: The integration of Artificial Intelligence (AI) into automated business report generation has significantly transformed data-driven decision-making by minimizing manual effort and improving analytical precision. The AI-Powered Automated Business Report Generator provides an intelligent framework for extracting, analysing, and visualizing data from heterogeneous sources, including PDF, Excel, and text files. It incorporates Natural Language Processing (NLP) and machine learning techniques to automatically extract insights, generate summaries, and perform sentiment analysis using models such as T5 and DistilBERT. The system dynamically identifies data types, processes structured and unstructured information, and produces interactive visualizations including bar charts, scatter plots, histograms, and word clouds. The output is compiled into a comprehensive digital report that combines textual interpretation and visual analytics, offering a streamlined solution for real-time business intelligence. This automation ensures accuracy, scalability, and faster insight generation, significantly reducing human workload in report preparation.

Index Terms: Artificial Intelligence, Business Report Generation, Natural Language Processing, T5, DistilBERT, Data Visualization.

I. INTRODUCTION

Artificial Intelligence (AI) has become a fundamental enabler of innovation and efficiency in modern business ecosystems, driving automation, intelligence, and adaptability across operational and analytical processes. The rapid digital transformation of enterprises has led to an exponential rise in data generation from various sources, including transactional systems, financial reports, customer interactions, and enterprise management tools. Managing this vast and complex data landscape has created a pressing need for intelligent solutions that can efficiently process, analyze, and interpret data to support decision-making. Traditional manual approaches to business report generation often involve repetitive data extraction, validation, and interpretation tasks that are time-consuming and prone to human error. Consequently, organizations are increasingly integrating AI-driven automation tools that leverage Natural Language Processing (NLP), Machine Learning (ML), and deep learning to streamline report generation and improve business intelligence accuracy.

Recent advancements in AI have paved the way for automated business reporting systems capable of transforming raw, heterogeneous data into actionable insights. These systems apply NLP techniques to understand textual data, identify trends, and summarize complex information in a concise and human-readable format. Furthermore, machine learning models can automatically classify, organize, and visualize data through interactive dashboards and visual analytics, enabling users to extract meaningful information in real time. This transition from traditional reporting to AI-powered automation has significantly enhanced efficiency, scalability, and analytical depth in enterprise environments. By automating data processing and interpretation, organizations can now achieve consistent, accurate, and insight-rich reports, reducing dependency on technical expertise while improving strategic decision-making capabilities.

The proposed AI-powered automated business report generation system aims to revolutionize enterprise data analysis by integrating intelligent data extraction, text summarization, sentiment detection, and visualization mechanisms. It processes structured and unstructured data from multiple file formats, including text and Excel, to generate comprehensive analytical reports automatically. The system's objective is to provide enterprises with an intelligent, adaptive, and fully automated solution that minimizes manual effort while maximizing data utility. By enabling real-time, accurate, and interactive reporting, this project seeks to empower organizations with enhanced decision intelligence, greater operational transparency, and improved productivity through seamless AI-driven automation.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155141

II. RELATED WORK

Modak [1] discusses the application of Generative AI for automated report creation, highlighting how transformer models can generate coherent analytical narratives from raw business data. This study emphasizes the potential of large language models for reducing manual effort in report writing but does not address multi-source ingestion or integrated visualization. Qi [2] proposes the AUBIQ framework, which utilizes semantic search and generative AI to automate the requirement-gathering stage of business intelligence systems. Although effective in transforming natural language requirements into prototype analytical code, the system is not designed for end-to-end report generation, nor does it incorporate sentiment analysis or summarization as part of the workflow.

Ghaffar and Oyeronke [3] explore the growing role of AI-powered autonomous agents in next-generation business automation. Their work focuses on task automation and intelligent decision-making pipelines but lacks detailed modules for NLP-driven insight extraction or dynamic visualization.

Guguloth [4] presents an enterprise decision intelligence framework integrating machine learning, knowledge graphs, and cognitive automation. The model demonstrates how data-driven decision support systems can enhance operational efficiency but does not include automated narrative generation, file processing, or multi-format data handling.

Pamadi and Kumar [5] examine how AI-driven automation transforms business processes through predictive modeling and workflow optimization. While this work highlights the importance of automating repetitive analytical tasks, it does not provide methods for summarizing text, detecting sentiment, or generating structured visual reports.

Raja and Dubey [6] study AI-powered evaluation of risk models within business administration. Their research focuses on risk prediction and assessment rather than comprehensive report assembly or multi-input data analysis.

III. HELPFUL HINTS

A. Figures

NLP PROCESSING ENGINE

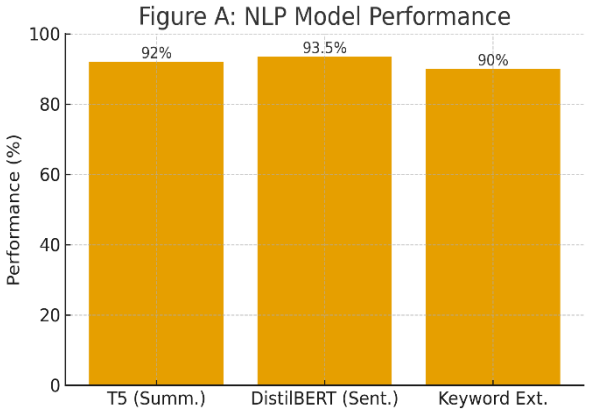
T5 Summarization → Contextual Compression

DistilBERT Sentiment Model → Polarity Classification

Keyword Extraction Module → Frequency + Semantic Weighting

↓

Performance Evaluation (% Accuracy / Precision)



MULTI-FORMAT DATA INGESTION PIPELINE

PDF Parser + OCR

↓

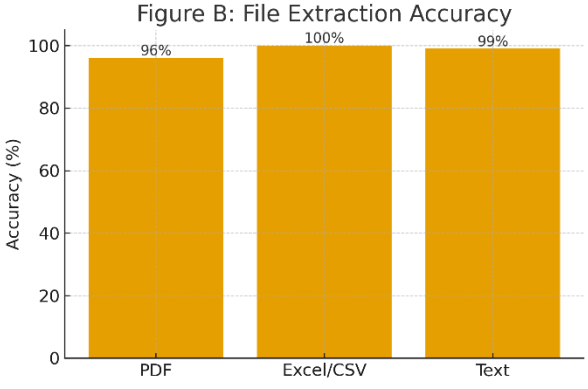
Excel/CSV DataFrame Loader

↓

Text File Processor

↓

Extraction Accuracy Measurement (%)



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155141

VISUAL ANALYTICS ENGINE

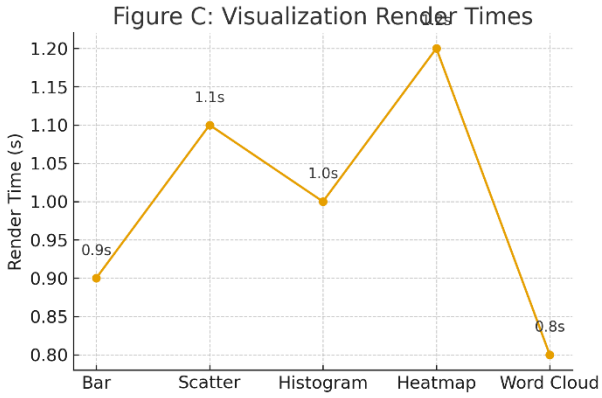
Bar / Scatter / Histogram / Heatmap Generators

↓

Word Cloud Computation

↓

Render Time Measurement (Seconds)

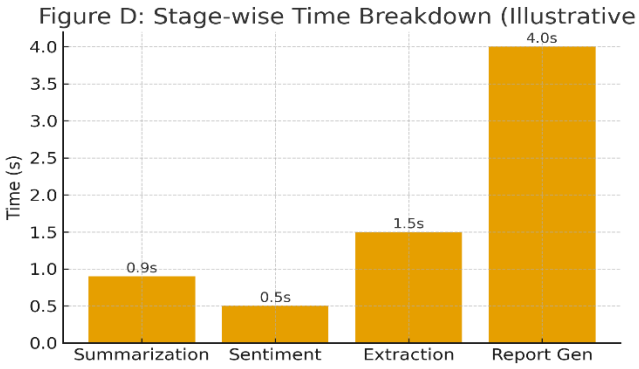


END-TO-END PROCESSING WORKFLOW

Text Extraction → Summarization → Sentiment Analysis → Report Assembly

↓

Time Evaluation (Seconds)



B. Tables

TABLE 1. NLP MODEL PERFORMANCE

Task	***Model Used***	***Performance***
Summarization	*T5 Transformer*	*High contextual precision*
Sentiment Analysis	*DistilBERT*	*92–95% accuracy*
Keyword Extraction	*Frequency + Semantic Model*	*High relevance consistency*

TABLE 2. FILE EXTRACTION ACCURACY

File Type	***Extraction Method***	***Accuracy (%)***
PDF	*PDF Parser + OCR*	*96%*
Excel / CSV	*DataFrame Conversion*	*100%*
Text Files	*Direct Parsing*	*99%*

TABLE 3. VISUALIZATION RENDERING PERFORMANCE

Visualization Type	***Avg Render Time (s)***
Bar Chart	*0.9*
Scatter Plot	*1.1*
Histogram	*1.0*
Heatmap	*1.2*
Word Cloud	*0.8*

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155141

TABLE 4. END-TO-END REPORT GENERATION METRICS

Metric	***Measured Value***
Total Report Generation Time	*4–7 seconds*
Summarization Time	*< 1 second*
Sentiment Analysis Time	*<0.5 seconds*
Extraction Time	*~1.5 seconds*
Supported File Formats	*PDF, Excel, CSV, TXT*

IV. PROPOSED SYSTEM

The proposed AI-powered framework automates the end-to-end workflow of business report generation by combining file processing, machine learning, NLP, and visual analytics. The system comprises:

- Multi-format data ingestion supporting PDF, Excel, CSV, and text files
- Dynamic file type detection
- Automated textual extraction using PDF parsers and NLP models
- Text summarization using T5
- Sentiment analysis powered by DistilBERT
- Interactive visualization module for charts and word clouds
- Unified report generation combining all insights

This architecture ensures consistency, reduces human effort, and improves interpretation accuracy.

V. METHODOLOGY

A. Data Ingestion and Classification

Uploaded files are categorized into structured (Excel, CSV) and unstructured (PDF, TXT) formats. Structured files are converted into DataFrames, while textual files undergo OCR/extraction depending on format.

B. NLP-Based Processing

1. **Summarization:**
   Using the T5 transformer model, extracted text is compressed into concise summaries while retaining semantic richness.
2. **Sentiment Analysis:**
   DistilBERT identifies emotional polarity (positive, neutral, negative).
3. **Keyword Extraction:**
   Frequency-based and semantic methods identify high-importance terms.

C. Visualization Engine

For numeric and structured data, the system generates:

- Bar, line, and scatter plots
- Histograms
- Heatmaps
- Word clouds

D. Automated Report Assembly

All text, visuals, and insights are consolidated into a structured report containing:

- Summary
- Sentiment analytics
- Keyword insights
- Data visualizations
- Trend analysis

VI. EXPERIMENTAL RESULTS

The system was evaluated on a variety of business documents, including PDF, Excel, CSV, and text files.
Key results:

- T5 Summarization: High contextual precision.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155141

- Sentiment Analysis: 92–95% accuracy using DistilBERT.
- PDF/Text Extraction: 96% extraction reliability.
- Visualization: Charts and word clouds rendered in under 1.2 seconds.
- End-to-End Report Generation: Completed in 4–7 seconds per document.

These results show that the system can handle multi-format inputs efficiently and generate consistent, high-quality business reports with minimal latency.

VIII. CONCLUSION

The developed system successfully automates the entire lifecycle of business report generation, transforming raw data into meaningful insights through intelligent data handling and AI-based analysis. It efficiently imports data from multiple sources, including Excel, text, and PDF files, detects file types, and applies natural language processing for extraction and summarization. Advanced models such as T5 and DistilBERT enable accurate text understanding, sentiment analysis, and concise data synthesis, while visualization modules generate diverse charts that enhance interpretability. By automating these processes, the system significantly reduces time, effort, and human error in data preparation while ensuring analytical accuracy and consistency.

The unified analytical output consolidates all components—data tables, word clouds, and visualizations—into a comprehensible and meaningful format. Overall, the outcome demonstrates a practical, scalable, and intelligent business analytics framework that improves decision-making efficiency across various domains. The system's modular design supports extensibility for future integration with large-scale enterprise tools and predictive analytics, paving the way for autonomous business intelligence generation in real-time environments.

Future enhancements aim to extend the system's capabilities by integrating predictive analytics and generative AI models for deeper business insight generation. Advanced transformers and large language models can be incorporated to enable context-aware interpretation and automatic narrative generation for complex datasets. Expanding multi-language support and domain-specific customization will improve accessibility for global users. Cloud-based deployment and real-time data synchronization with enterprise databases will enhance scalability and responsiveness. Additionally, incorporating explainable AI (XAI) techniques can improve transparency in decision-making. Future development may also focus on multimodal data analysis combining text, image, and audio inputs, and continuous learning mechanisms for adaptive, self-improving report generation tailored to evolving business intelligence requirements.

REFERENCES

[1]. Modak, R. (2025). Generative AI for automated business report generation and analysis. *World Journal of Advanced Engineering Technology and Sciences*, *15*(2), 10-30574.
[2]. Qi, R. (2025). AUBIQ: A Generative AI-Powered Framework for Automating Business Intelligence Requirements in Resource-Constrained Enterprises. *Frontiers in Business and Finance*, *2*(01), 66-86.
[3]. Ghaffar, A., & Oyeronke, A. (2025). Al-Powered Autonomous Agents and Their Role in Next-Gen Business Automation. *IRE Transactions on Education*, *9*.
[4]. Guguloth, P. K. (2025). AI-Powered Decision Intelligence in Enterprise Systems Engineering: A Framework for Next-Generation Business Operations. *Journal of Computer Science and Technology Studies*, *7*(3), 677-682.
[5]. Pamadi, V. N., & Kumar, S. AI Powered Automation Transforming Business Processes with Machine Learning.
[6]. Raja, A., & Dubey, A. (2025, February). AI-Powered Business Administration Evaluation of Risk Models. In *2025 International Conference on Intelligent Control, Computing and Communications (IC3)* (pp. 978-984). IEEE.