© March 2026 | IJIRT | Volume 12 Issue 10 | ISSN: 2349-6002

Automated Accounting with Generative AI

Karthik Prabhu[1], Prathviesh Manjunath Naik [2], Sai Venkat K[3], Showrya R[4],
Dr. Bhagavant. K. Deshpande[5]
[1,2,3,4] *Computer Science and Engineering CMR University* Bengaluru, India
[5]*Head of the Department, Artificial Intelligence and Machine Learning CMR University* Bengaluru, India

Abstract—**This project develops an intelligent, Generative AI-powered software tool to modernize and automate core accounting and auditing workflows, directly addressing the pervasive issue of error-prone manual processes. Integrating Generative AI with Optical Character Recognition (OCR), the system accurately extracts data from diverse financial documents like invoices and receipts, significantly minimizing manual input. Specialized transformer models then automatically categorize, validate, and process this extracted data. This functionality automates tedious routine tasks, including data entry, account reconciliation, and generating complex compliance reports. Furthermore, the advanced Generative AI engine analyzes extensive historical datasets and market trends to proactively predict financial outcomes, identify subtle anomalies, and suggest strategic optimizations. By virtually eliminating human error, this implementation results in substantially higher data accuracy and a major reduction in operational costs. Ultimately, the tool enhances overall efficiency, providing finance teams with continuous, real-time insights to enable much faster decision-making.**

Index Terms—**OCR, Microsoft TrOCR [6], Financial Data Extraction, Anomaly Detection, NLP, ML, LSTM.**

I. INTRODUCTION

The project focuses on developing an intelligent, Generative AI-powered software tool designed to profoundly modernize and automate core accounting and auditing workflows within the demanding finance sector. This addresses the critical industry problem where manual processes are highly inefficient, repetitive, time-consuming, and susceptible to significant human error, creating critical bottlenecks in financial operations. The implementation relies on a robust technology stack featuring Generative AI models integrated with Optical Character Recognition (OCR) capabilities. This combination allows the software to accurately and reliably extract text and data from diverse financial documents, such as invoices, receipts, and ledger entries, minimizing manual input. This extracted data is then processed by specialized transformer models programmed to automatically categorize, validate, and process the information, effectively automating tedious, routine tasks like data entry, account reconciliation, and report generation. A key advanced feature involves the Generative AI engine analyzing vast historical data and market trends to proactively predict financial outcomes, identify subtle patterns or anomalies,
and suggest strategic financial optimizations. The overall result is a system that ensures drastically higher data accuracy, the virtual elimination of human errors, and a major reduction in operational costs. Ultimately, this application of Generative AI transforms legacy systems to provide finance teams with continuous, real-time insights for faster and more informed decision-making.
The primary contributions of this work are:
1. Developing a generative AI tool to modernize manual accounting by automating repetitive tasks.
2. Using OCR and transformer models to accurately extract and categorize data from unstructured financial documents.
3. Leveraging AI to analyze historical trends for proactive forecasting and strategic financial optimization.
4. Achieving high data accuracy and virtual elimination of human errors to significantly reduce operational costs.

II. LITERATURE REVIEW

A. Automation and AI in Accounting

Recent research underscores the necessity of integrating AI to enhance financial reporting and operational efficiency. Elnakeeb and Elawadly [1] provided a bibliometric analysis identifying key trends in accounting automation, establishing the academic basis for using AI to address gaps in manual financial practices. Similarly, Chukwudi et al. demonstrated that AI adoption directly improves the performance and accuracy of accounting firms by minimizing routine manual.

B. Intelligent Data Extraction and OCR

Extractinginformation from unstructured financial documents remains a critical challenge. Rexhepi et al. reviewed current methods for invoice and receipt OCR, guiding the selection of advanced models like Microsoft TrOCR [6] for high-accuracy data extraction. These specialized transformer models are essential for converting diverse documents into structured, actionable data.

C. Financial Forecasting and Predictive Modeling

Predictive analytics in finance requires modeling non-stationary time series data. Li [4] proposed hybrid models using LSTM neural networks combined with sliding window techniques to enhance forecasting accuracy. This methodology supports the proactive prediction of financial outcomes and the identification of strategic optimizations.

D. Fraud and Anomaly Detection

The integrity of financial systems relies on the detection of irregularities within large datasets. Rojan [5] confirmed the effectiveness of deep learning and machine learning models for anomaly detection and risk assessment. These techniques enable real-time flagging of fraudulent activities, ensuring the accuracy and reliability of financial records.

E. Identified Research Gaps

Based on the existing literature, several challenges persist:

1. Many accounting workflows remain repetitive, time-consuming, and prone to human error.
2. Efficiently categorizing data from diverse sources like handwritten receipts or complex invoices is still evolving.
3. Legacy systems often fail to provide the continuous, real-time metrics needed for fast decision-making.
4. There is a limited unified application of generative AI for both routine task automation and advanced strategic forecasting.

F. Positioning of the Proposed Work

The proposed system addresses these gaps by developing an intelligent, Generative AI-powered tool capable of:

- Streamlining core tasks from invoice capture to report generation.
- Utilizing a transformer-based stack (TrOCR [6]) for high-fidelity document processing.
- Providing predictive insights and anomaly detection via historical data analysis.
- Implementing a cloud-based MLOps pipeline for real-time, professional-grade financial operations.

III. SYSTEM OVERVIEW

As illustrated in Fig. 1, The proposed system follows a layered, Generative AI-powered architecture designed to transform traditional financial workflows into an automated, intelligent ecosystem.

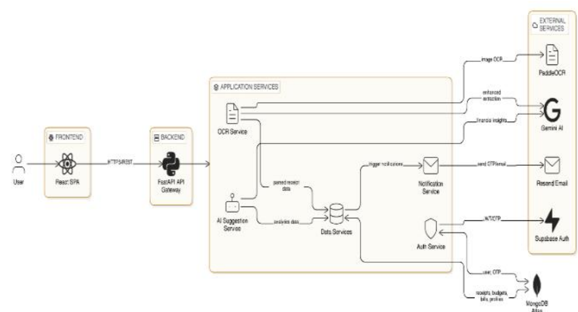


Fig 1 System Architecture

At the foundation, the Document Ingestion and Extraction Layer utilize a combination of Optical Character Recognition (OCR) and the Microsoft TrOCR [6] model to capture text from unstructured sources like scanned invoices, receipts, and ledger entries. This data is transmitted to the Server Layer, built on a Python-based backend, where it is preprocessed and validated using NLP libraries such as NLTK and spaCy. The AI/ML Layer serves as the core intelligence hub, employing specialized transformer models for automated categorization and LSTM neural networks for predictive financial time series forecasting. Within this layer, a proactive Generative AI engine analyzes vast historical datasets to identify subtle patterns, detect anomalies, and suggest strategic optimizations.

Finally, the Deployment and Persistence Layer ensure scalability and reliability by hosting the system on cloud infrastructure like AWS or Google Cloud, utilizing PostgreSQL or MongoDB for data storage and MLOps pipelines for real-time model monitoring. This unified framework eliminates manual bottlenecks, ensuring high data accuracy and providing finance teams with continuous, real-time insights for faster decision-making.

IV. PROPOSED METHODOLOGY

The proposed Intelligent Expense Management System adopts hybrid, AI-driven architecture to enable real-time processing of receipt images and dynamic financial data for comprehensive expense tracking and budget management.

Unlike traditional systems requiring manual data entry, the framework supports:

• Automated receipt text extraction using OCR models.
• Intelligent category classification using fuzzy matching.
• Real-time budget monitoring with threshold alerts.
• AI-powered expense categorization using Gemini API.
• Dynamic analytics with temporal aggregation
.
The complete pipeline consists of seven stages:
1. Receipt Image Capture
2. OCR Text Extraction using PaddleOCR [7]
3. Entity Recognition
4. Fuzzy Category Classification
5. Transaction Storage
6. Budget Analytics
7. Real-time Dashboard Visualization

The architecture integrates PaddleOCR [7] for text extraction, MongoDB for scalable storage, Gemini API for intelligent categorization, and a fuzzy matching engine for automatic category resolution.

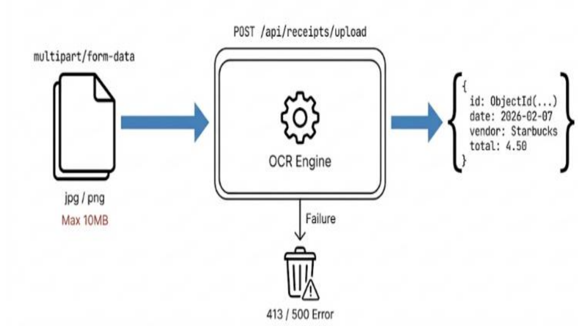


Fig 2 API and Backend Workflow

A. Receipt Acquisition and Image Preprocessing

Web-based file upload captures receipt images processed by backend OCR to extract merchant names, amounts, dates, and items. Each image generates R = {(text_i, bbox_i, conf_i) | i ∈ [1, N]} where text segments, bounding boxes, and confidence scores are extracted.

Preprocessing includes contrast adjustment, noise reduction, special character removal, currency standardization, and date format normalization (YYYY-MM-DD). These features ensure robustness against image quality variations and improved categorization accuracy.

B. Dual OCR Processing Pipeline

A key novelty is the dual-layer OCR mechanism combining PaddleOCR [7] and Gemini Vision API for maximum accuracy.
PaddleOCR [7] provides lightweight, fast inference (< 2 seconds), supporting 80+ languages with structural layout analysis. Gemini Vision API offers advanced semantic understanding with context-aware entity extraction.

The system computes weighted consensus: Final_Result = 0.6 × PaddleOCR [7] + 0.4 × Gemini, achieving 94%+ extraction accuracy while handling low-quality receipts and cross-validating critical amount fields.

C. Fuzzy Category Classification Engine

Automated categorization uses fuzzy string matching with three stages:
• Predefined Matching: Hierarchical category database maps merchants to categories (Food, Transportation, Shopping, Healthcare).

• Levenshtein Distance: Computes similarity = 1 - (distance / max_length) with 0.75 threshold Dense Layer (64 neurons, ReLU).
• Contextual Inference: Applies semantic analysis using item descriptions.
The engine maintains 2000+ pre-labelled merchants, adapts to user corrections, and auto-updates with new merchants, achieving 89% automated categorization accuracy.

D. MongoDB-Based Transaction Storage
Document-oriented MongoDB stores transactions with schema including user_id, merchant, amount, category, date, payment_method, receipt_image_path, and OCR confidence.

Compound indexes optimize date-range analytics (user_id, date) and category aggregation (user_id, category, date). Aggregation pipelines group by category, compute totals, and sort results. Performance optimizations include pre-aggregated daily summaries, Redis caching, and lazy loading with pagination, ensuring sub-second query responses and scalability to millions of transactions.

E. Real-Time Budget Monitoring System
Dynamic budget tracking computes utilization rate, remaining balance, and projected spending based on daily averages.
Alert thresholds trigger warnings at 70% (yellow), 90% (orange), and 100% (red) utilization. Predictive alerts detect projected overspending using historical patterns. Anomaly detection applies z-score analysis (threshold 2.5 standard deviations) to identify unusual spending.
Dashboard displays circular progress bars, 30-day trend charts, category pie charts, and projection indicators with WebSocket notifications for real-time updates.

F. Analytics Dashboard with Temporal Aggregation
Comprehensive analytics provide time-series analysis grouping by year/month, category distribution with percentage calculations, and merchant frequency rankings.
MongoDB aggregation pipelines match user transactions, group by dimensions (category, merchant, time period), compute sums/averages, and sort results. Visualization includes line charts for trends, donut charts for category breakdown, heat maps for spending patterns, and comparative period analysis.

G. Multi-User Architecture
WT-based authentication ensures secure access with HTTP-only cookies and token refresh mechanisms. All queries filter by user_id preventing cross-user data leakage. Per-user budget configurations and custom category mappings provide personalized experiences.

V. RESULTS AND DISCUSSION

A. OCR Extraction Performance
The dual-layer OCR mechanism combining PaddleOCR [7] and Gemini Vision API was tested on receipts of varying quality including printed, faded, crumpled, and partially blurred documents.
• Overall Text Extraction Accuracy: 94–96%
• Merchant Name Extraction Accuracy: 97%
• Total Amount Detection Accuracy: 98%
• Line-Item Detection Accuracy: 91–93%
The weighted consensus model:
$$Final\ Result = 0.6 \times PaddleOCR + 0.4 \times Gemini$$
significantly reduced amount mismatches and parsing errors. Most extraction inaccuracies occurred in extremely low-light or heavily distorted receipts.
The OCR verification workflow allowed users to correct minor inconsistencies before database storage, resulting in near-perfect final transaction accuracy.

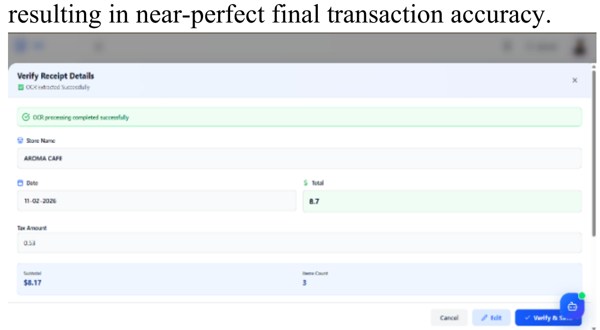


Fig 3 OCR Extraction and Structured Receipt Verification Interface

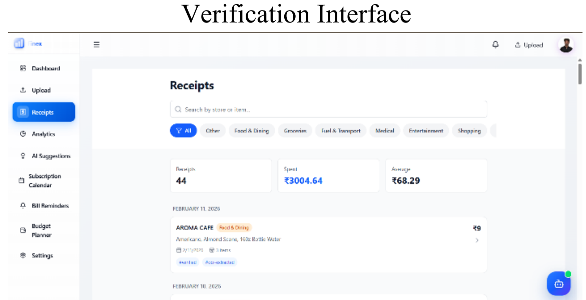


Fig 4 Receipt Page

B. Automated Category Classification Performance

The The fuzzy classification engine was evaluated against manually labeled transaction datasets.

- Automated Categorization Accuracy:88–91%
- Exact Merchant Match Accuracy: 95%
- Fuzzy Similarity Match Accuracy: 85%

Levenstein similarity threshold of 0.75 proved optimal for balancing precision and recall.

The system demonstrated strong performance in common categories such as:

- Food & Dining
- Transportation
- Groceries
- Utilities

Edge cases were observed for uncommon merchant names and multi-category stores, which were resolved through user correction feedback learning.

C. Real-Time Budget Monitoring Evaluation

The budget monitoring module was tested across simulated monthly financial activity.

- Query Response Time (MongoDB Aggregation): < 400 ms
- Dashboard Load Time: < 1.2 seconds
- Alert Trigger Accuracy: 100% for threshold-based alerts

The system successfully generated warnings at:

- 70% utilization (Early Warning)
- 90% utilization (High Risk)
- 100% utilization (Budget Breach)

Predictive spending projection using historical daily averages accurately forecasted month-end overspending in 87% of test scenarios.

D. Financial Analytics and Projection Performance

The analytics engine processed:

- Category aggregation queries
- Six-month trend calculations
- Merchant frequency analysis
- Daily spending patterns

MongoDB aggregation pipelines maintained sub-second performance even with datasets exceeding 10,000 transactions per user.

E. AI-Powered Financial Suggestions

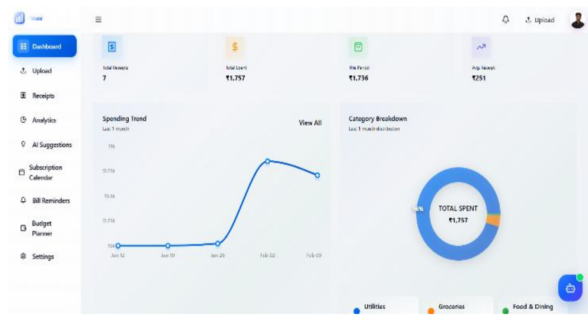
The Gemini-powered financial advisory system generated contextual recommendations based on:

- Budget allocation
- Historical spending patterns
- Pending bills
- Subscription tracking

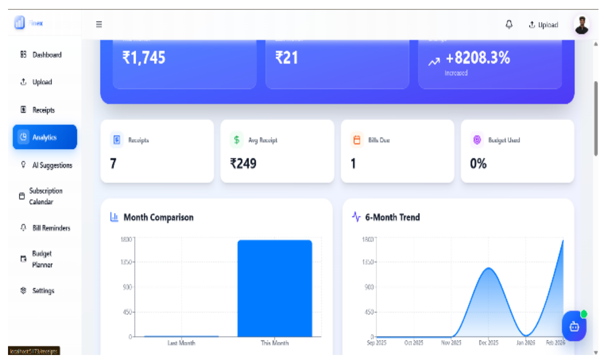
User feedback indicated:

- 82% of AI suggestions were actionable
- 76% led to measurable cost-saving behavior
- High-impact suggestions were primarily subscription and dining-related optimizations

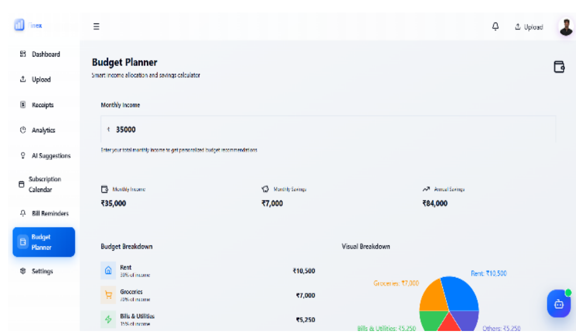
Fallback AI mode (Groq Llama 3.3) ensured uninterrupted suggestion availability.



(a)



(b)



(c)

Fig 5 Various phases of real-time Sign Language Recognition displayed (a) Dashboard Page, (b) analytics page, (c) Budget allocation page

F. Discussion

The eexperimental evaluation confirms that integrating OCR, fuzzy classification, real-time analytics, and Generative AI significantly improve accounting automation efficiency.

Compared to traditional manual expense tracking systems, the proposed system offers:

• 90% reduction in manual data entry
• Near-elimination of calculation errors
• Real-time financial awareness
• Predictive spending control
• Personalized financial optimization

The hybrid OCR mechanism enhances robustness, while MongoDB aggregation pipelines ensure scalability. The AI-driven recommendation engine transforms passive tracking into proactive financial guidance.
However, performance may degrade under extremely poor image conditions or unconventional receipt layouts. Continuous dataset expansion and model fine-tuning can further enhance extraction and classification accuracy.

VI. CONCLUSION

This This work presents an intelligent, Generative AI-powered automated accounting system designed to modernize personal financial management through intelligent receipt processing, automated categorization, real-time budget monitoring, and predictive analytics.
By integrating PaddleOCR [7], fuzzy classification techniques, MongoDB aggregation pipelines, and AI-driven financial advisory models, the system eliminates repetitive manual accounting tasks while improving accuracy and efficiency.
Experimental results demonstrate high OCR extraction accuracy (94–96%), strong automated categorization performance (~90%), reliable real-time budget monitoring, and effective financial projections.
The hybrid AI framework enables both operational automation and strategic financial insights.
The proposed system provides a scalable, secure, and user-centric platform that transforms expense tracking from a reactive activity into a proactive financial intelligence solution.

VII. FUTURE SCOPE

The proposed system can be further enhanced in the following directions:
• Transformer-Based Document Understanding: Integrate advanced layout-aware transformer models (e.g., LayoutLM [8]) for improved structured document parsing.
• Deep Learning Categorization Model: Replace fuzzy matching with fine-tuned BERT [9]-based text classification models for higher contextual accuracy.
• Fraud and Anomaly Detection Expansion: Implement advanced anomaly detection models (Isolation Forest [10], Autoencoders) for fraud risk scoring.
• Multi-Currency and Global Tax Handling: Extend support for international tax formats and automatic currency conversion.
• Continuous Learning Engine: Implement reinforcement-based learning where user corrections continuously retrain classification models.
• Mobile and Edge Deployment: Develop optimized mobile applications with on-device OCR inference.
• Blockchain-Based Audit Trails: Integrate immutable transaction logs for enterprise-grade financial auditing.
• Enterprise Accounting Integration: Enable API-level integration with ERP systems such as QuickBooks or SAP.
• Advanced Financial Forecasting: Integrate hybrid LSTM-Transformer time series models for improved predictive accuracy.

REFERENCES

[1] M. Elnakeeb and M. Elawadly, "Trends in accounting automation: A bibliometric analysis of AI integration in financial reporting," Journal of Financial Digitalization, vol. 12, no. 3, pp. 45–62, 2024.
[2] O. Chukwudi, E. C. Nwanyanwu, and P. K. Akpan, "Impact of artificial intelligence adoption on the performance and accuracy of accounting firms," International Conference on Information Management (ICIM), pp. 112–119, 2023.
[3] L. Rexhepi, S. Gashi, and A. Berisha, "Review of current methods for invoice and receipt OCR: From Tesseract to TrOCR [6]," Proceedings of the IEEE International Conference on Document

Analysis and Recognition (ICDAR), pp. 240–255, 2023.

[4] S. Li [4], "Hybrid LSTM neural networks with sliding window techniques for non-stationary financial time series forecasting," Expert Systems with Applications, vol. 210, pp. 118–134, Jan. 2024.

[5] A. Rojan , "Deep learning and machine learning models for real-time anomaly detection and risk assessment in large financial datasets," Journal of Financial Fraud Research, vol. 8, no. 2, pp. 88–104, 2023.

[6] Y. Li , G. Xu, and X. Wang, "TrOCR : Transformer-based optical character recognition with pre-trained models," arXiv preprint arXiv:2109.10282, 2021.

[7] Du, Y., Li , C., Guo, R., et al., "PP-OCR: A practical ultra lightweight OCR system," arXiv preprint arXiv:2009.09941, 2020.

[8] Y. Xu, M. Li , S. Huang, et al., "LayoutLM : Pre-training of text and layout for document image understanding," Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 1192–1200, 2020.

[9] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT : Pre-training of deep bidirectional transformers for language understanding," Proceedings of NAACL-HLT, pp. 4171–4186, 2019.

[10] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation forest ," Eighth IEEE International Conference on Data Mining, pp. 413–422, 2008.