



BFSI- OCR of Bank Statements

Presented by Jasneet Arora



Table of Contents

- Problem Statement
- Introduction
- Technology Stack
- Architecture
- Pipeline
- Outcomes & Deliverables
- Dashboard Snippets
- Challenges & Solutions
- Future Scope





Problem Statement



Manual Document Processing:

Time-consuming, error-prone, and inefficient workflows.



Data Extraction Challenges:

Difficulty extracting structured data from diverse formats, especially from scanned documents or images.



Limited Insights:

Traditional methods lack interactive tools for in-depth analysis and querying of financial data.



Inefficient Data Interpretation:

Slow and cumbersome analysis hinders timely decision-making.



Need for Automation:

A need for an automated system to streamline data extraction and analysis from bank statements



Introduction



Project Overview: Develops a robust OCR system for automated extraction and analysis of bank statement data in various formats.



Goal: Streamline financial data interpretation, providing valuable insights to financial institutions and individuals.



Modular Development: Key modules ensure seamless Cloudinary API integration, accurate OCR, comprehensive analysis, and system integration.

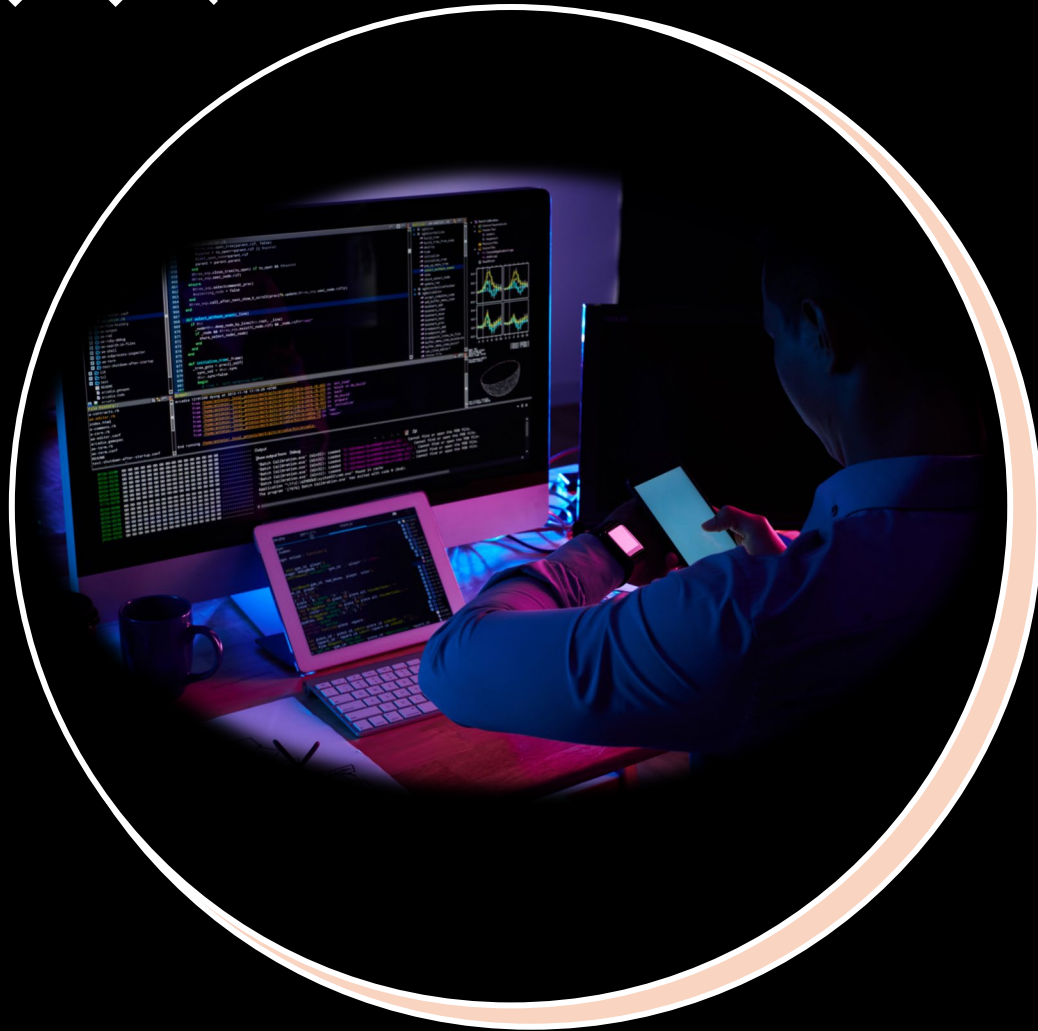


Target Users: Banks, financial institutions, and individuals seeking efficient bank statement analysis.





Technology Stack

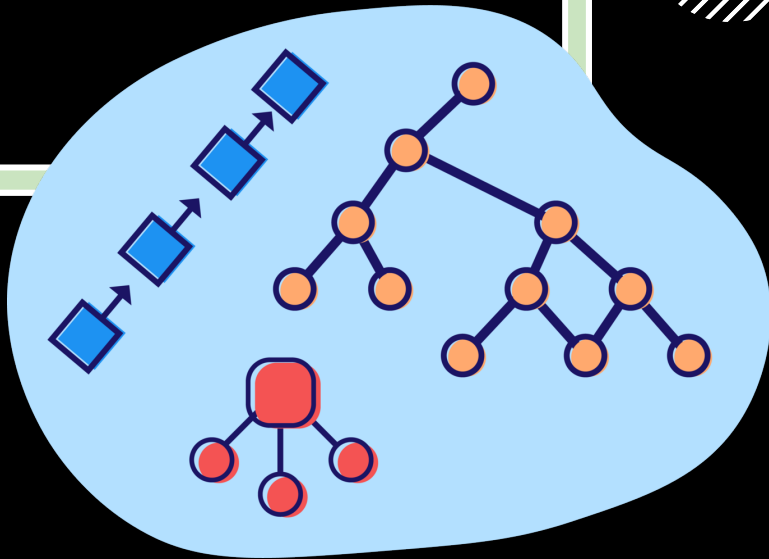


- **Primary Language:** Python for core development and integrations.
- **OCR Engine:** Meta-Llama/Llama-3.2-11B-Vision-Instruct-Turbo for accurate text extraction.
- **Web Framework:** Streamlit for creating an interactive and user-friendly interface.
- **Data Processing:** Pandas for data manipulation, cleaning, and organization.
- **Visualization:** Plotly for generating dynamic and interactive charts and graphs.
- **Cloud Storage:** Cloudinary for secure and scalable image management and retrieval.
- **Other Libraries:** PyMuPDF for PDF processing, Pillow for image handling.





Architecture



Data Input: Users can manually upload images or fetch them from Cloudinary.

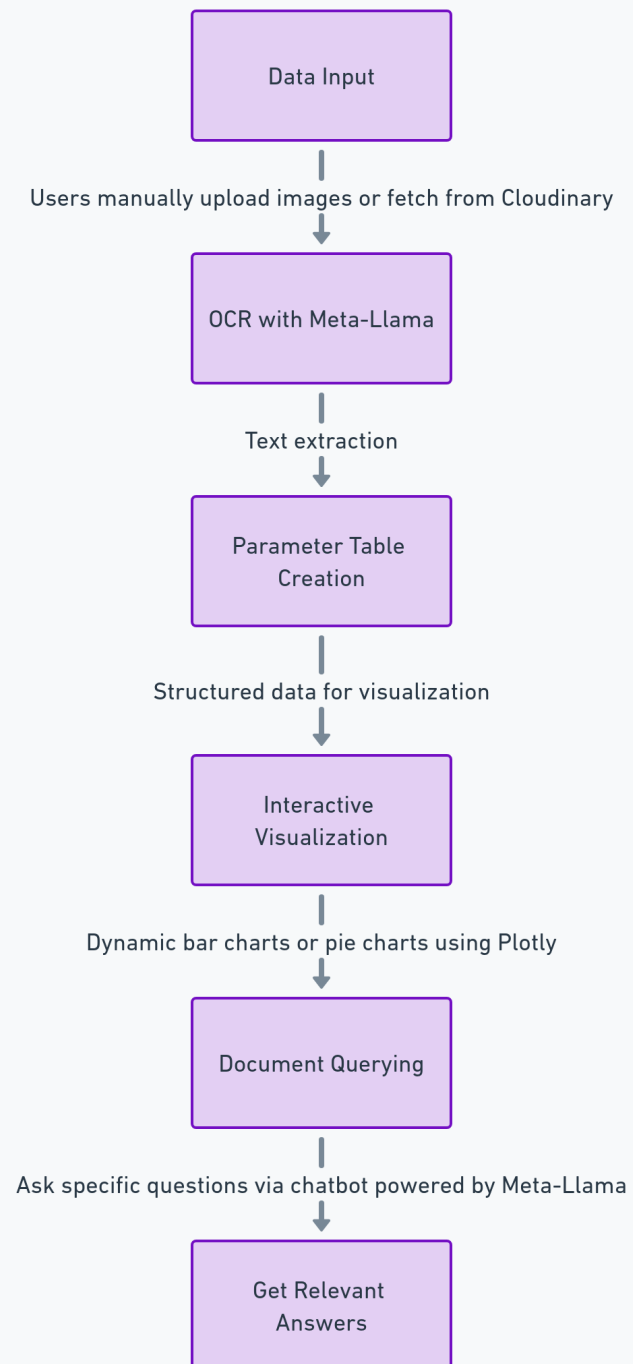
OCR with Meta-Llama: The uploaded/fetched images are passed to the meta-llama/Llama-3.2-11B-Vision-Instruct-Turbo model for text extraction.

Parameter Table Creation: The extracted text data is then structured into a parameter table for visualization and analysis.

Interactive Visualization: Plotly uses the parameter table to create dynamic bar charts or pie charts, as selected by the user.

Document Querying: A chatbot interface, also powered by the Meta-Llama model, allows users to ask specific questions about the document content and get relevant answers.

Pipeline



Outcomes & Deliverables



Unified Interface: A single, user-friendly Streamlit interface for uploading, visualization, and querying.



OCR and Data Extraction: The system extracts text and provides structured, organized data for analysis.

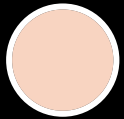


Dynamic Visualization: Interactive Plotly charts (bar/pie) help users explore and understand financial data.

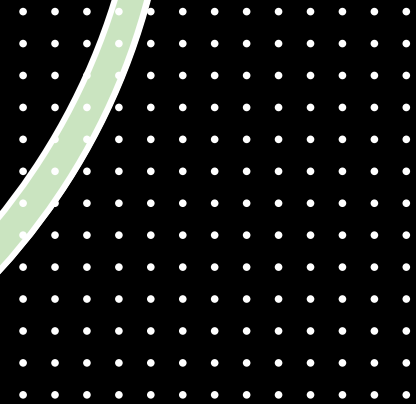


Querying Feature: The AI-powered chatbot allows users to gain deeper insights into document content.





Dashboard Snippets



Example: Bank Statements



Financial Document Analyzer

Select Document Type

Bank Statement

Select Data Source

☒ Fetch from Cloudinary

☐ Upload Files

Select Graph Type

Bar Chart


Clear All Data


Bank Statement Analysis


Number of Images to fetch


7


Fetch Images

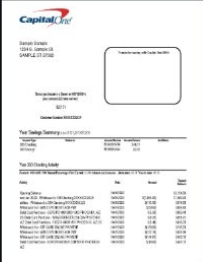

image65.jpg



image7.jpg


image146.jpg


image25.jpg


image89.jpg


image32.jpg


image33.jpg

Updated Images Using Cloudinary

Example: Bank Statements

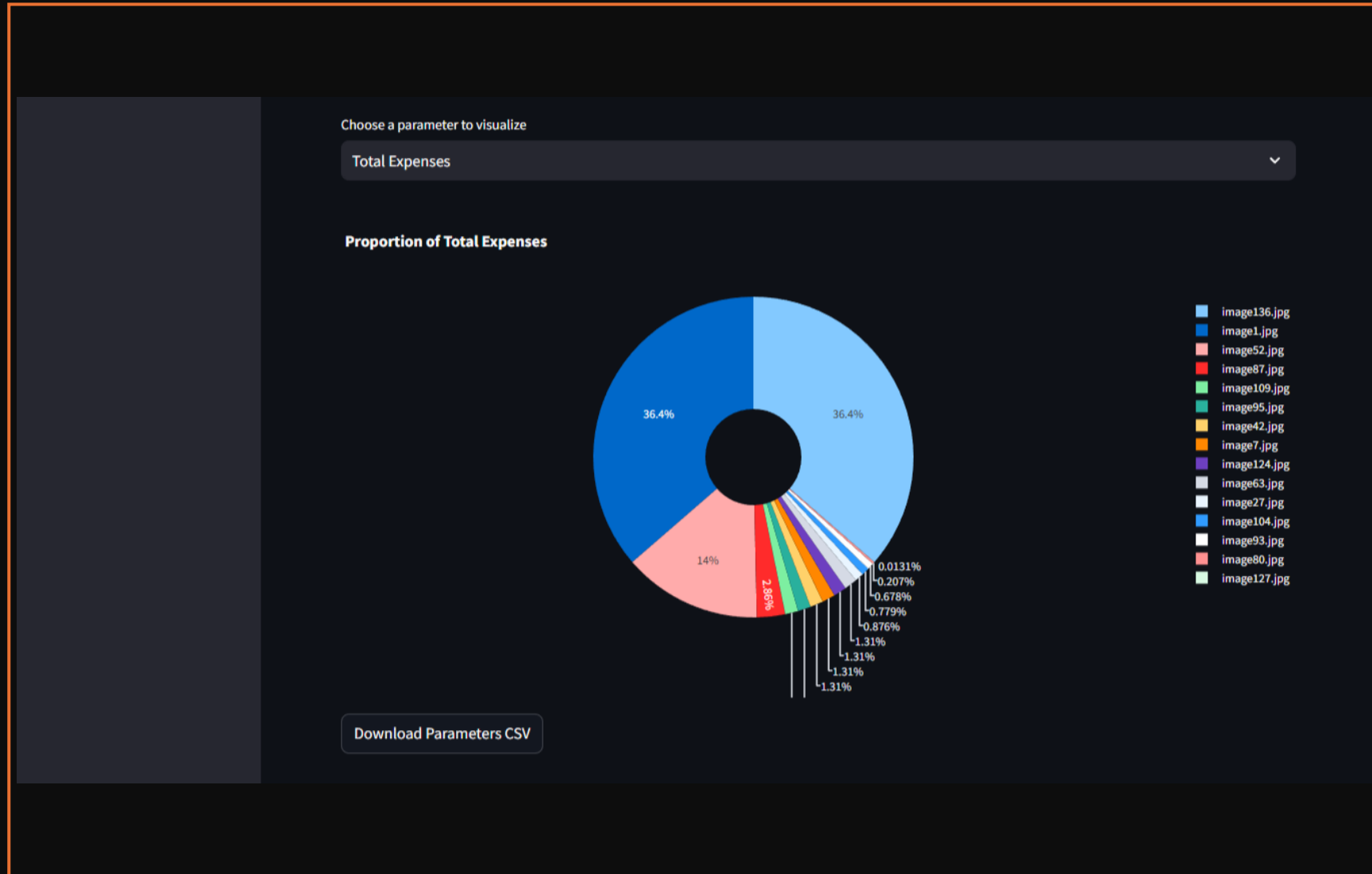


Extracted Parameters

	Parameter	Value	Document
0	Total Revenue	100,000	image109.jpg
1	Total Expenses	75,000	image109.jpg
2	Gross Profit	25,000	image109.jpg
3	Net Profit	20,000	image109.jpg
4	Operating Expenses	5,000	image109.jpg
5	Total Revenue	100,000	image95.jpg
6	Total Expenses	75,000	image95.jpg
7	Gross Profit	25,000	image95.jpg
8	Net Profit	20,000	image95.jpg
9	Operating Expenses	5,000	image95.jpg

Parameter Extraction

Example: Bank Statements



Visualization: Pie Chart

Example: Bank Statements



Ask a Question About the Document

Select Image to Query

Document 5

Enter your question about the document:

How do I increase the gross profit?

To increase the gross profit, you can consider the following strategies:

1. **Reduce COGS (Cost of Goods Sold):** Lowering the cost of producing or purchasing your products can directly contribute to an increase in gross profit. This could involve negotiating better prices with suppliers, streamlining production processes, or implementing cost-saving measures.
2. **Increase Selling Price:** If possible, consider raising the selling price of your products without compromising their quality or value. This will directly increase revenue and, if COGS remains constant, lead to a higher gross profit margin.
3. **Improve Efficiency:** Enhancing operational efficiency can help reduce waste, minimize errors, and optimize resource utilization. This can lead to a decrease in COGS and an increase in gross profit.
4. **Invest in Product Development:** Developing new products or improving existing ones can increase their value and potentially command higher prices. This can contribute to an increase in gross profit if the cost of development is managed effectively.
5. **Optimize Inventory Management:** Managing inventory levels effectively can help reduce holding costs and minimize waste. This can contribute to a decrease in COGS and an increase in gross profit.
6. **Implement Lean Manufacturing:** Implementing lean manufacturing principles can help eliminate waste and improve efficiency in production. This can lead to a decrease in COGS and an increase in gross profit.
7. **Reduce Waste:** Identifying and eliminating unnecessary waste in production can help reduce COGS and increase gross profit.
8. **Invest in Technology:** Investing in technology can help streamline processes, reduce errors, and improve efficiency. This can lead to a decrease in COGS and an increase in gross profit.
9. **Develop Strategic Partnerships:** Forming strategic partnerships with suppliers or other businesses can help reduce costs and improve efficiency. This can contribute to a decrease in COGS and an increase in gross profit.
10. **Focus on Quality Control:** Ensuring high-quality products can help command higher prices and reduce returns or defects. This can contribute to an increase in gross profit if the cost of quality control is managed effectively.

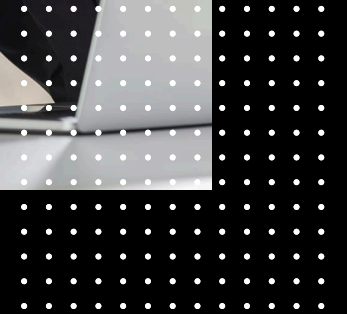
By implementing these strategies, you can potentially increase your gross profit and improve your overall financial performance.

Ask a Question About the Document

Challenges & Solutions

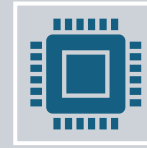


- **Document Variability:** Handling diverse document formats and layouts.
 - **Solution:** Custom prompts provided to the Meta-Llama model for consistent data extraction.
- **Accuracy in Extraction:** Ensuring consistency and accuracy of extracted data.
 - **Solution:** Data organized into a parameter table, using a specific format for accuracy and consistency.
- **Integration Complexity:** Coordinating multiple technologies (Cloudinary, OCR, visualization) was challenging.
 - **Solution:** Used a modular architecture for seamless integration and extensive testing for reliability.





Future Scope



Real-time Deployment: Deploy the system in a production environment for real-time financial document analysis.



Performance Optimization: Optimize the system to handle larger datasets and various document formats more efficiently.



Enhanced Features: Add features like multi-language support and automated report generation for broader applicability.



