Case Study: Multilingual Invoice Decoder with Anomaly Detection and AI Insights

Problem Faced by Real World:

In the fast-paced business world, managing and processing invoices can be a tedious task, especially when dealing with multilingual invoices or invoices in non-standard formats. Businesses often face challenges in:

- Detecting discrepancies or anomalies in invoices, which can lead to financial losses or errors.
- Extracting key financial insights from complex invoice data, especially when presented in different languages or image formats.
- Automating the analysis to reduce human intervention and improve accuracy in invoice management processes.

What This Project Achieves:

This project provides an advanced solution for businesses to automatically analyze and detect anomalies in invoices, while also extracting insights through AI-powered financial analysis. The key outcomes include:

- **Multilingual Support**: The ability to handle invoices in different languages.
- **Anomaly Detection**: Automated detection of any unusual financial entries in the invoices, helping businesses identify potential fraud, data entry mistakes, or irregular transactions.
- **AI-driven Insights**: The ability to generate summaries and financial insights from the invoice data through the Gemini AI model.
- **Speech Integration**: A speech recognition interface to allow users to provide input via voice, further enhancing accessibility and usability.

Approach Used:

1. **Multilingual Invoice Processing**: The project uses a **Google Gemini AI model** for natural language processing and understanding of invoice data. The AI helps generate relevant insights from text or image-based invoices.
2. **Anomaly Detection**: **Isolation Forest** from **scikit-learn** is used for anomaly detection. The model identifies outliers in numerical data by evaluating patterns and deviations in the financial entries, flagging them as anomalies.
3. **Speech Recognition Integration**: The **SpeechRecognition** library is used to allow users to interact with the system via voice input, making the system more accessible for different user preferences.
4. **Data Preprocessing and Feature Engineering**: Using **Pandas** for data manipulation and cleaning, the system processes invoice data (whether in CSV format or as an image) and prepares it for analysis.

Tech Stack:

- **Streamlit**: The framework used for building the interactive web application, allowing users to upload invoices, visualize data, and interact with the system.

- **Google Gemini AI**: Used for natural language processing and generating insights from text-based and image-based invoices.
- **scikit-learn**: For machine learning, particularly the **Isolation Forest** model used for anomaly detection in the financial data.
- **Pandas & NumPy**: For data manipulation and numerical analysis.
- **Pillow (PIL)**: Used for handling image uploads and displaying invoices in image format.
- **SpeechRecognition**: For converting speech input into text, making the system more accessible.

Key Features and Functionalities:

1. **Invoice Upload**: Users can upload invoices either as images (JPG, PNG) or CSV files. The system processes and extracts data from both formats.
2. **Anomaly Detection**: The system automatically checks for anomalies in numerical data (e.g., unexpected large sums or outliers), alerting users to potential issues.
3. **AI Insights**: Upon uploading an invoice, the Gemini AI model generates a financial analysis based on the data, offering insights into spending patterns, trends, and overall financial health.
4. **Speech Recognition**: Users can provide instructions or prompts through voice commands, making the interface more intuitive.
5. **Data Visualization**: Invoices, once uploaded, are displayed in the app for further analysis, with financial trends shown in a tabular format.

Output:

Invoice Analysis Options

Input Prompt:

what is the address of the company

Use Speech Input

Upload an Invoice (Image or CSV)

Drag and drop file here

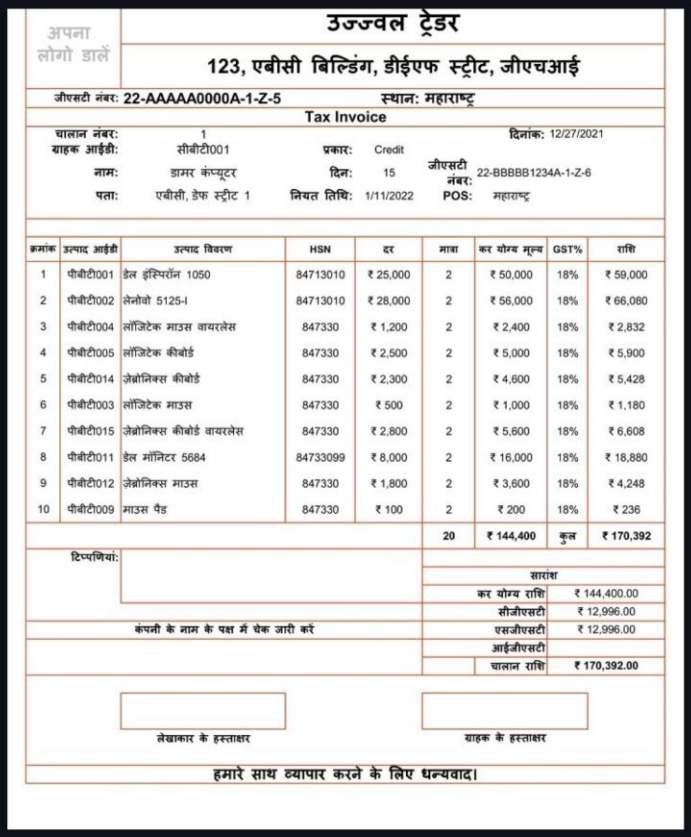
Limit 200MB per file • JPG, JPEG, PNG, CSV

Browse files

GST-Invoice-Template-In-Hi… ×

127.7KB

Analyze Invoice



अपना लोगो डालें

उज्ज्वल ट्रेडर

123, एबीसी बिल्डिंग, डीईएफ स्ट्रीट, जीएचआई

जीएसटी नंबर: 22-AAAAA0000A-1-Z-5 स्थान: महाराष्ट्र

Tax Invoice

चालान नंबर: 1 दिनांक: 12/27/2021

ग्राहक आईडी: सीबीटी001 प्रकार: Credit

नाम: डामर कंप्यूटर दिन: 15 जीएसटी नंबर: 22-BBBBB1234A-1-Z-6

पता: एबीसी, डेफ स्ट्रीट 1 नियत तिथि: 1/11/2022 POS: महाराष्ट्र

क्रमांक	उत्पाद आईडी	उत्पाद विवरण	HSN	दर	मात्रा	कर योग्य मूल्य	GST%	राशि
1	पीबीटी001	डेल इंस्पिरॉन 1050	84713010	₹ 25,000	2	₹ 50,000	18%	₹ 59,000
2	पीबीटी002	लेनोवो 5125-I	84713010	₹ 28,000	2	₹ 56,000	18%	₹ 66,080
3	पीबीटी004	लॉजिटेक माउस वायरलेस	847330	₹ 1,200	2	₹ 2,400	18%	₹ 2,832
4	पीबीटी005	लॉजिटेक कीबोर्ड	847330	₹ 2,500	2	₹ 5,000	18%	₹ 5,900
5	पीबीटी014	ज़ेब्रोनिक्स कीबोर्ड	847330	₹ 2,300	2	₹ 4,600	18%	₹ 5,428
6	पीबीटी003	लॉजिटेक माउस	847330	₹ 500	2	₹ 1,000	18%	₹ 1,180
7	पीबीटी015	ज़ेब्रोनिक्स कीबोर्ड वायरलेस	847330	₹ 2,800	2	₹ 5,600	18%	₹ 6,608
8	पीबीटी011	डेल मॉनिटर 5684	84733099	₹ 8,000	2	₹ 16,000	18%	₹ 18,880
9	पीबीटी012	ज़ेब्रोनिक्स माउस	847330	₹ 1,800	2	₹ 3,600	18%	₹ 4,248
10	पीबीटी009	माउस पैड	847330	₹ 100	2	₹ 200	18%	₹ 236
					20	₹ 144,400	कुल	₹ 170,392

टिप्पणियां:

सारांश

कर योग्य राशि	₹ 144,400.00
सीजीएसटी	₹ 12,996.00
एसजीएसटी	₹ 12,996.00
आईजीएसटी	
चालान राशि	₹ 170,392.00

कंपनी के नाम के पक्ष में चेक जारी करें

लेखाकार के हस्ताक्षर

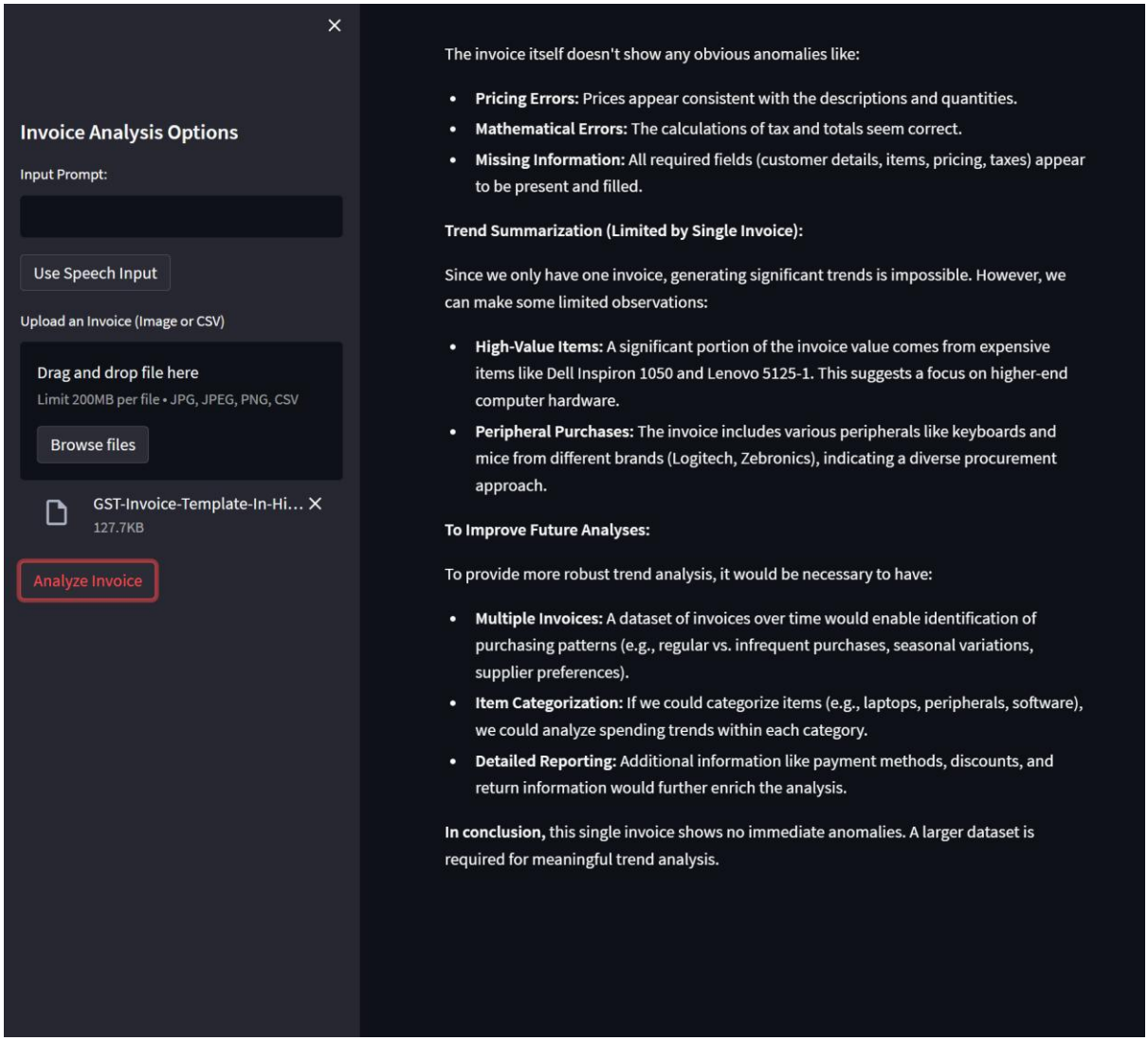
ग्राहक के हस्ताक्षर

हमारे साथ व्यापार करने के लिए धन्यवाद।

Uploaded Invoice

Response from AI

The company's address is 123, एबीसी बिल्डिंग, डीईएफ स्ट्रीट, जीएचआई, महाराष्ट्र. (This translates approximately to 123, ABC Building, DEF Street, GHI, Maharashtra).



The invoice itself doesn't show any obvious anomalies like:

- **Pricing Errors:** Prices appear consistent with the descriptions and quantities.
- **Mathematical Errors:** The calculations of tax and totals seem correct.
- **Missing Information:** All required fields (customer details, items, pricing, taxes) appear to be present and filled.

Trend Summarization (Limited by Single Invoice):

Since we only have one invoice, generating significant trends is impossible. However, we can make some limited observations:

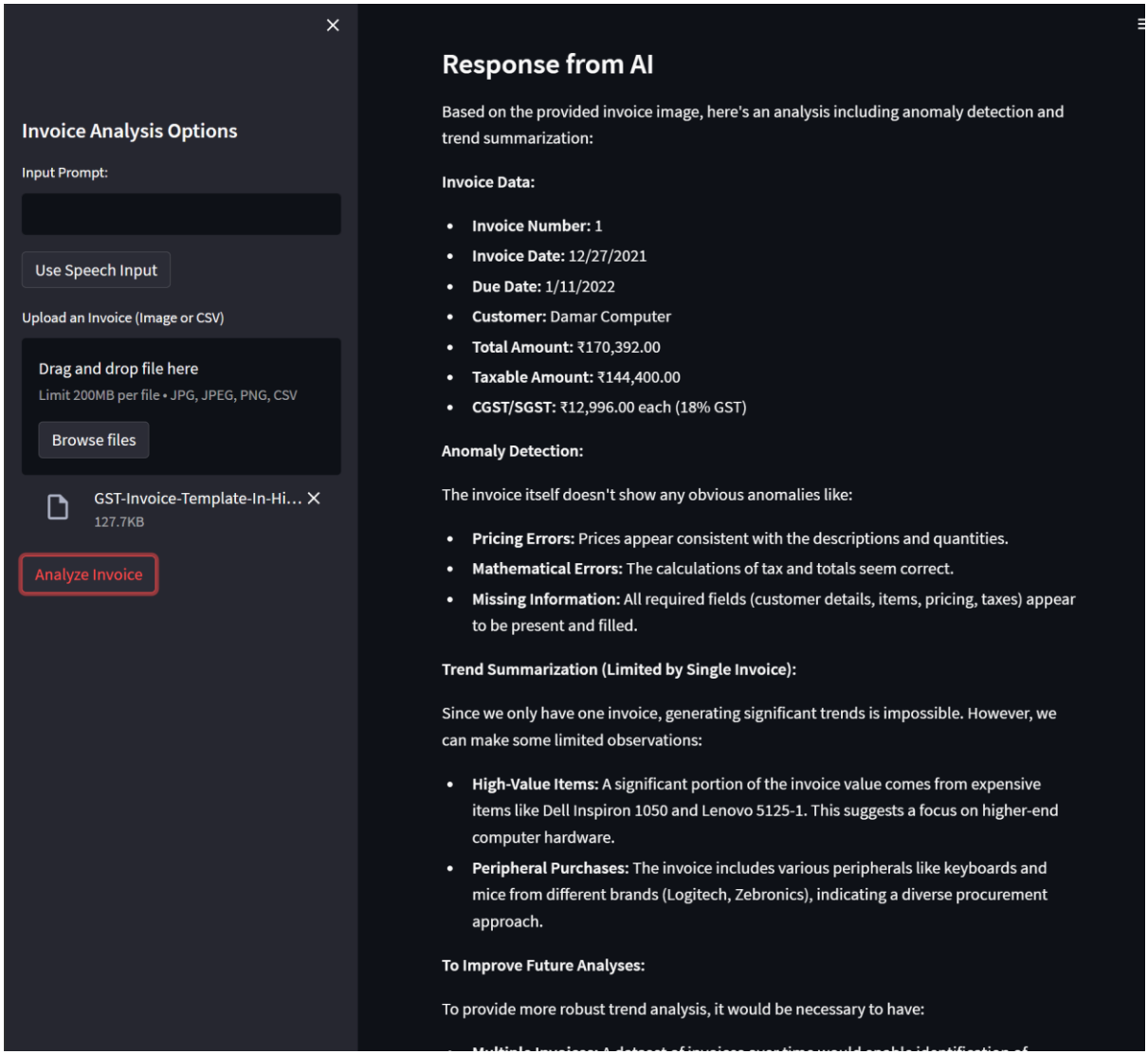
- **High-Value Items:** A significant portion of the invoice value comes from expensive items like Dell Inspiron 1050 and Lenovo 5125-1. This suggests a focus on higher-end computer hardware.
- **Peripheral Purchases:** The invoice includes various peripherals like keyboards and mice from different brands (Logitech, Zebronics), indicating a diverse procurement approach.

To Improve Future Analyses:

To provide more robust trend analysis, it would be necessary to have:

- **Multiple Invoices:** A dataset of invoices over time would enable identification of purchasing patterns (e.g., regular vs. infrequent purchases, seasonal variations, supplier preferences).
- **Item Categorization:** If we could categorize items (e.g., laptops, peripherals, software), we could analyze spending trends within each category.
- **Detailed Reporting:** Additional information like payment methods, discounts, and return information would further enrich the analysis.

In conclusion, this single invoice shows no immediate anomalies. A larger dataset is required for meaningful trend analysis.



Conclusion:

This project solves a significant problem in the real world by automating the invoice analysis process, detecting anomalies in financial data, and providing actionable insights using AI. Businesses can save valuable time and reduce errors in invoice management. By leveraging advanced technologies like AI, machine learning, and speech recognition, this tool offers a modern solution to a common challenge in financial operations.