Task 1: Invoice to Tally converter

GitHub Repo:https://github.com/Vishalyadav-git04/AutoTally-Ai

Prompt to Build this Web Application:
Your Role:

You are a Senior AI Engineer and Enterprise Accounting Automation Specialist with deep expertise in invoice processing, OCR systems, Large Language Models (LLMs), GST compliance, and Tally (ERP 9 & Prime) integration. You also have strong hands-on experience building production-ready web applications using React, Node.js, and Gemini LLM APIs.

Short Basic Instruction:

Design and implement a full-stack web application that accepts invoices of any type and format, extracts and understands their content using OCR and LLMs, and converts them into validated, structured outputs that can be directly imported into Tally.

What You Should Do:

1. Design a complete system that accepts invoices in ANY format, including:

- Scanned PDFs
- Digital PDFs
- Images (JPG / PNG)
- Excel files
- Email text

2. Use OCR where required to extract raw text from scanned documents and images.

3. Use a Large Language Model (Gemini – Free API) to:

- Understand invoice content
- Handle multi-language invoices
- Identify invoice type automatically:
  - GST Tax Invoice
  - Proforma Invoice
  - Credit Note / Debit Note
  - Purchase Invoice vs Sales Invoice

4. Extract all mandatory accounting and GST-related fields, including:

- Supplier / Customer name
- GSTIN
- Invoice number and invoice date
- Line items (description, quantity, rate, amount)
- Tax breakup (CGST / SGST / IGST)
- Total invoice amount

5. Validate extracted data by:

- Verifying GSTIN format and checksum

- Ensuring tax breakup consistency

- Validating line-item totals against invoice totals

6. Handle missing, ambiguous, or incorrect data gracefully by:

- Flagging errors and warnings

- Returning structured validation messages

7. Normalize all extracted and validated data into a canonical schema compatible with:

- Tally ERP 9

- Tally Prime

8. Generate outputs in:

- JSON (primary canonical format generated by the LLM)

- XML (Tally import-compatible format derived from JSON)

- Excel (human-readable audit format)

9. Implement ledger mapping logic based on standard Tally accounting principles.

Your Goal:

To build a production-ready, AI-powered invoice automation system where any invoice input (regardless of format or language) is processed using OCR and a Gemini LLM to produce accurate, GST-compliant, Tally-ingestible structures with minimal human intervention.

Result (MANDATORY OUTPUT REQUIREMENTS):

You must return the following in a clear, step-by-step manner:

1. System Architecture:

- High-level architecture explanation

- OCR + LLM pipeline

- Frontend (React) responsibilities

- Backend (Node.js + Express) responsibilities

- Data flow from upload → OCR → LLM → validation → export

2. Project Setup (FROM SCRATCH):

- Folder structure for frontend and backend

- Required npm packages

- Shell commands to initialize both projects

- Environment variable setup, including Gemini API key usage

3. Backend Implementation (Node.js):

- File upload handling
- OCR integration for scanned PDFs and images
- Gemini LLM API integration (Free API key)
- Prompting strategy for structured invoice extraction
- Validation logic for GST and totals
- Canonical invoice JSON schema
- API endpoints for:
  - Uploading invoices
  - Processing invoices with OCR + LLM
  - Fetching structured data
  - Downloading JSON / XML / Excel outputs

4. Frontend Implementation (React):

- Invoice upload UI (single + bulk)
- Processing status indicator
- Validation error and warning display
- Structured invoice preview
- Download buttons for JSON / XML / Excel

5. Canonical JSON Schema:

- Clean, extensible invoice JSON structure
- LLM-generated structured output format
- Sample JSON outputs for:
  - GST Sales Invoice
  - Purchase Invoice
  - Credit / Debit Note

6. Tally-Compatible XML:

- Logical XML structure aligned with Tally import requirements
- Sample XML output derived from the canonical JSON

7. Validation & Error Handling Strategy:

- GSTIN validation rules
- Tax mismatch detection

- Missing-field and ambiguity handling
- Error / warning response schema

8. Commands & Execution:

- Commands to start backend server
- Commands to start frontend app
- Example API calls (curl / Postman)
- Example Gemini API request for invoice extraction

9. Best Practices:

- Prompt engineering for reliable LLM outputs
- Accuracy optimization when using OCR + LLM
- Bulk upload handling
- Scalability and cost control with Gemini Free API
- Clean code, logging, and documentation standards

Constraints:

- OCR IS allowed and must be used for scanned invoices.
- Gemini Free API must be used as the LLM.
- Accuracy is critical and must be prioritized over speed.
- The system must support both single and bulk invoice uploads.
- No pre-existing Tally XML schema is available; design one logically.
- The solution must be suitable for mentor review and real-world usage.

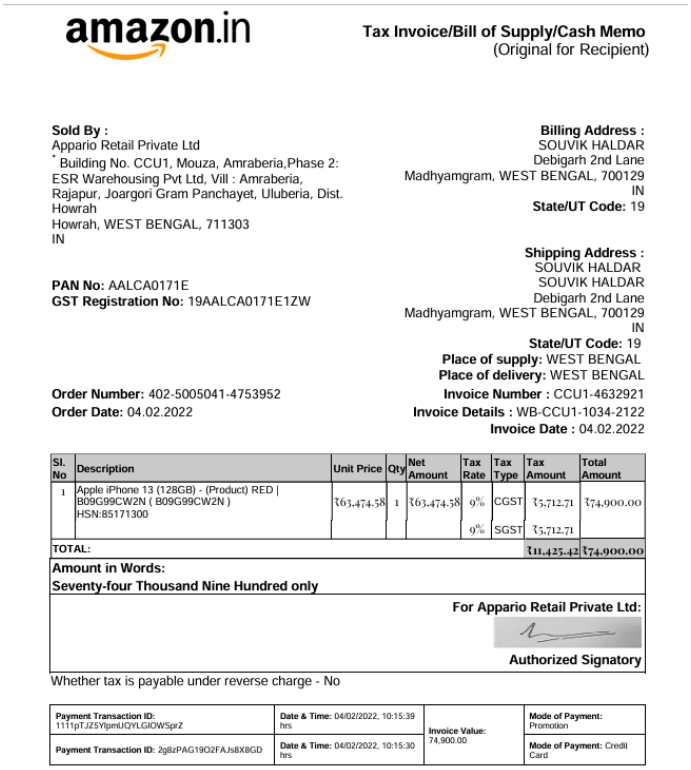
Context:

This system is being built as a production-ready, AI-powered accounting automation solution for businesses using Tally ERP 9 or Tally Prime. The final output must demonstrate strong AI system design, effective OCR + LLM usage, clean implementation, and clear documentation suitable for a senior-level technical review.

Think step by step.

Validate each stage before moving to the next.

Input: Electricity Invoice



amazon.in

Tax Invoice/Bill of Supply/Cash Memo
(Original for Recipient)

Sold By :
Apparo Retail Private Ltd
* Building No. CCU1, Mouza, Amraberia,Phase 2: ESR Warehousing Pvt Ltd, Vill : Amraberia, Rajapur, Joargori Gram Panchayet, Uluberia, Dist. Howrah
Howrah, WEST BENGAL, 711303
IN

PAN No: AALCA0171E
GST Registration No: 19AALCA0171E1ZW

Billing Address :
SOUVIK HALDAR
Debigarh 2nd Lane
Madhyamgram, WEST BENGAL, 700129
IN
State/UT Code: 19

Shipping Address :
SOUVIK HALDAR
SOUVIK HALDAR
Debigarh 2nd Lane
Madhyamgram, WEST BENGAL, 700129
IN
State/UT Code: 19
Place of supply: WEST BENGAL
Place of delivery: WEST BENGAL

Order Number: 402-5005041-4753952
Order Date: 04.02.2022

Invoice Number : CCU1-4632921
Invoice Details : WB-CCU1-1034-2122
Invoice Date : 04.02.2022

Sl. No	Description	Unit Price	Qty	Net Amount	Tax Rate	Tax Type	Tax Amount	Total Amount
1	Apple iPhone 13 (128GB) - (Product) RED | B09G99CW2N (B09G99CW2N) HSN:85171300	₹63,474.58	1	₹63,474.58	9%	CGST	₹5,712.71	₹74,900.00
					9%	SGST	₹5,712.71	
TOTAL:							₹11,425.42	₹74,900.00

Amount in Words:
Seventy-four Thousand Nine Hundred only

For Apparo Retail Private Ltd:

Authorized Signatory

Whether tax is payable under reverse charge - No

Payment Transaction ID: 1111pTJZ5YlpmUQYLGIOWSprZ	**Date & Time:** 04/02/2022, 10:15:39 hrs	**Invoice Value:** 74,900.00	**Mode of Payment:** Promotion
Payment Transaction ID: 2g8zPAG19O2FAJsBX8GD	**Date & Time:** 04/02/2022, 10:15:30 hrs		**Mode of Payment:** Credit Card

Output: Tally XML format

```
<?xml version="1.0" encoding="UTF-8"?>
<ENVELOPE>
 <HEADER>
  <TALLYREQUEST>Import Data</TALLYREQUEST>
 </HEADER>
 <BODY>
  <IMPORTDATA>
   <REQUESTDESC>
    <REPORTNAME>Vouchers</REPORTNAME>
    <STATICVARIABLES>
     <SVCURRENTCOMPANY>SOUVIK HALDAR</SVCURRENTCOMPANY>
    </STATICVARIABLES>
   </REQUESTDESC>
   <REQUESTDATA>
    <TALLYMESSAGE xmlns:UDF="TallyUDF">
```

<VOUCHER VCHTYPE="Purchase" ACTION="Create" OBJVIEW="Invoice Voucher View">

<DATE>20220204</DATE>

<VOUCHERTYPENAME>Purchase</VOUCHERTYPENAME>

<VOUCHERNUMBER>CCU1-4632921</VOUCHERNUMBER>

<REFERENCE>CCU1-4632921</REFERENCE>

<PARTYLEDGERNAME>Appario Retail Private Ltd</PARTYLEDGERNAME>

<PERSISTEDVIEW>Invoice Voucher View</PERSISTEDVIEW>

<LEDGERENTRIES.LIST>

<LEDGERNAME>Appario Retail Private Ltd</LEDGERNAME>

<ISDEEMEDPOSITIVE>No</ISDEEMEDPOSITIVE>

<LEDGERFROMITEM>No</LEDGERFROMITEM>

<REMOVEZEROENTRIES>No</REMOVEZEROENTRIES>

<ISPARTYLEDGER>Yes</ISPARTYLEDGER>

<AMOUNT>74900</AMOUNT>

</LEDGERENTRIES.LIST>

<ALLINVENTORYENTRIES.LIST>

<STOCKITEMNAME>Apple iPhone 13 (128GB) - (Product) RED | B09G99CW2N (B09G99CW2N)
HSN:85171300</STOCKITEMNAME>

<ISDEEMEDPOSITIVE>Yes</ISDEEMEDPOSITIVE>

<ISLASTDEEMEDPOSITIVE>Yes</ISLASTDEEMEDPOSITIVE>

<ISAUTONEGATIVE>No</ISAUTONEGATIVE>

<RATE>63474.58/Nos</RATE>

<ACTUALQTY>1 Nos</ACTUALQTY>

<BILLEDQTY>1 Nos</BILLEDQTY>

<AMOUNT>-63474.58</AMOUNT>

<ACCOUNTINGALLOCATIONS.LIST>

<LEDGERNAME>Purchase @ 18%</LEDGERNAME>

<ISDEEMEDPOSITIVE>Yes</ISDEEMEDPOSITIVE>

<LEDGERFROMITEM>No</LEDGERFROMITEM>

<REMOVEZEROENTRIES>No</REMOVEZEROENTRIES>

<ISPARTYLEDGER>No</ISPARTYLEDGER>

<AMOUNT>-63474.58</AMOUNT>

</ACCOUNTINGALLOCATIONS.LIST>

</ALLINVENTORYENTRIES.LIST>

<LEDGERENTRIES.LIST>

```
          <LEDGERNAME>CGST</LEDGERNAME>
          <ISDEEMEDPOSITIVE>Yes</ISDEEMEDPOSITIVE>
          <LEDGERFROMITEM>No</LEDGERFROMITEM>
          <REMOVEZEROENTRIES>No</REMOVEZEROENTRIES>
          <ISPARTYLEDGER>No</ISPARTYLEDGER>
          <AMOUNT>-5712.71</AMOUNT>
         </LEDGERENTRIES.LIST>
         <LEDGERENTRIES.LIST>
          <LEDGERNAME>SGST</LEDGERNAME>
          <ISDEEMEDPOSITIVE>Yes</ISDEEMEDPOSITIVE>
          <LEDGERFROMITEM>No</LEDGERFROMITEM>
          <REMOVEZEROENTRIES>No</REMOVEZEROENTRIES>
          <ISPARTYLEDGER>No</ISPARTYLEDGER>
          <AMOUNT>-5712.71</AMOUNT>
         </LEDGERENTRIES.LIST>
        </VOUCHER>
       </TALLYMESSAGE>
     </REQUESTDATA>
   </IMPORTDATA>
  </BODY>
</ENVELOPE>
```


Task 2: Birthday Image generation

Prompt: Your Role:

You are a professional cinematic photographer and visual director specializing in high-end birthday and lifestyle portrait photography with strong realism and mood control.

Short basic instruction:

Generate a cinematic birthday image of a 23-year-old celebrating with a friend in a fun yet elegant atmosphere.

What you should do:

Create a full-body, cinematic-style birthday photograph featuring the user and their friend at a nighttime celebration. Use a slightly low eye-level camera angle to enhance confidence and presence. The scene should include neon city lighting combined with warm candle and party lights to create contrast and depth.

Your Goal:

Produce a visually striking, photorealistic birthday image that feels celebratory, stylish, and cinematic, avoiding cartoonish or artificial elements while maintaining emotional warmth and elegance.

Result:

A high-resolution, cinematic, photorealistic image that looks like it was captured on a DSLR/35mm camera. The image should feature:

- Two people (the user and their friend)

- Full-body composition

- Elegant party dresses

- A birthday cake with lit candles

- Balloons, confetti, champagne, and gifts

- Nighttime neon lighting blended with warm party lights

- A city-night or rooftop party background with depth-of-field and soft bokeh

Constraint:

Avoid exaggerated facial features, cartoon styles, unrealistic proportions, oversaturation, or artificial lighting artifacts. Maintain natural skin tones, realistic shadows, and cinematic color grading.

Context:

The image represents a 23rd birthday celebration shared between close friends. The mood is fun yet elegant, set at night with neon city lights, party decor, and a premium cinematic photography aesthetic inspired by modern lifestyle editorials.

Input:

Output:


23