Instructions for the Intern

We need to automate the extraction of data from patient assessment forms and integrate it into a SQL database. This task will involve using Optical Character Recognition (OCR) to parse image files (JPEG or PDF), extract structured and unstructured data, and store it in a database in JSON format.

Intern Assignment Overview

1. **Develop a Python-based OCR Script**
    a. Extract text from patient assessment forms (handwritten and printed).
    b. Identify and extract key data points.
    c. Convert the extracted data into a structured JSON format.
2. **Store Extracted Data in a SQL Database**
    a. Design a database schema.
    b. Insert JSON data into relevant tables.
3. **Provide a Public GitHub Repository**
    a. Include all code, documentation, and setup instructions.

Step-by-Step Instructions

1. Data Extraction Using OCR

- Use Tesseract OCR or EasyOCR for text recognition.
- Ensure handwritten and printed text is accurately captured.
- Preprocess the image (grayscale, thresholding) to improve OCR accuracy.

2. Identify and Extract Key Data Points

- Extract the following fields:
    - **Patient Details**: Name, DOB
    - **Treatment Details**: Date, Injection (Yes/No), Exercise Therapy (Yes/No)
    - **Difficulty Ratings (0-5)**: Bending, Putting on Shoes, Sleeping, etc.
    - **Patient Changes**:
        - Since last treatment
        - Since the start of treatment
        - Last 3 days (Good/Bad)
    - **Pain Symptoms (0-10)**:
        - Pain, Numbness, Tingling, Burning, Tightness
    - **Medical Assistant (MA) Inputs**:

- Blood Pressure, HR, Weight, Height, SpO2, Temperature, Blood Glucose, Respirations

3. Convert Extracted Data to JSON

Format Example:

```
{
 "patient_name": "John Doe",
 "dob": "01/05/1988",
 "date": "02/06/2025",
 "injection": "Yes",
 "exercise_therapy": "No",
 "difficulty_ratings": {
  "bending": 3,
  "putting_on_shoes": 1,
  "sleeping": 2
 },
 "patient_changes": {
  "since_last_treatment": "Not Good",
  "since_start_of_treatment": "Worse",
  "last_3_days": "Bad"
 },
 "pain_symptoms": {
  "pain": 2,
  "numbness": 5,
  "tingling": 6,
  "burning": 7,
  "tightness": 5
 },
 "medical_assistant_data": {
  "blood_pressure": "120/80",
  "hr": 80,
  "weight": 67,
  "height": "5'7",
  "spo2": 98,
  "temperature": "98.6",
  "blood_glucose": 115,
  "respirations": 16
 }
}
```

4. Store JSON Data in SQL Database

- Design a patients table with structured fields.
- Store JSON in a forms_data table for flexibility.

Sample SQL Schema:

```
CREATE TABLE patients (
  id SERIAL PRIMARY KEY,
  name VARCHAR(255),
  dob DATE
);

CREATE TABLE forms_data (
  id SERIAL PRIMARY KEY,
  patient_id INT REFERENCES patients(id),
  form_json JSONB,
  created_at TIMESTAMP DEFAULT NOW()
);
```

- Insert JSON data into forms_data.

5. Submit via GitHub

- The intern must push the complete project (code + documentation) to a **public GitHub repository**.
- Provide:
  - README.md with setup instructions.
  - OCR implementation script.
  - SQL schema and database scripts.
  - Sample JSON output.

Submission Requirements

1. **Working OCR script** (Python-based)
2. **Structured JSON output**
3. **Database schema and data storage implementation**
4. **GitHub public repository** with documentation
5. **Submit the repository link before the interview**

This assignment is the first step in the hiring process. It will help evaluate problem-solving skills, coding ability, and database integration knowledge.

One sample copy and blank copy attached in this doc bellow.

Date : _____ **MA Initials :**

INJECTION : YES NO **Exercise Therapy : YES NO**

Functional Assessment Questionnaire

Patient Name : Wesley Ingram **DOB :** 21/05/89

Rate on a scale from 0-5 (5 being the highest) how difficult it is to do the following tasks:

Bending or Stooping: 0 1 2 3 4 5

Putting on shoes: 0 1 2 3 4 5

Sleeping: 0 1 2 3 4 5

Standing for an hour: 0 1 2 3 4 5

Going up or down a flight of stairs: 0 1 2 3 4 5

Walking through a store: 0 1 2 3 4 5

Driving for an hour: 0 1 2 3 4 5

Preparing a meal: 0 1 2 3 4 5

Yard work: 0 1 2 3 4 5

Picking up items off the floor: 0 1 2 3 4 5

Patient Changes since last treatment:

Not Good

Patient changes since the start of treatment:

Not Good

Describe any functional changes within the last three days (good or bad):

Bad

Rate pain symptoms on a scale of 0-10 (10 being the highest):

Pain: 2 **Numbness:** 5 **Tingling:** 6 **Burning:** 7 **Tightness:** 6

**To Be Completed by MA:

Blood Pressure: 113 ***HR:*** 12 ***Weight:*** 60 ***Height:*** 7'

Date : _____ **MA Initials :**

INJECTION : YES NO **Exercise Therapy : YES NO**

Functional Assessment Questionnaire

Patient Name : _____ DOB : _____

Rate on a scale from 0-5 (5 being the highest) how difficult it is to do the following tasks:

Bending or Stooping: 0 1 2 3 4 5

Putting on shoes: 0 1 2 3 4 5

Sleeping: 0 1 2 3 4 5

Standing for an hour: 0 1 2 3 4 5

Going up or down a flight of stairs: 0 1 2 3 4 5

Walking through a store: 0 1 2 3 4 5

Driving for an hour: 0 1 2 3 4 5

Preparing a meal: 0 1 2 3 4 5

Yard work: 0 1 2 3 4 5

Picking up items off the floor: 0 1 2 3 4 5

Patient Changes since last treatment:

Patient changes since the start of treatment:

Describe any functional changes within the last three days (good or bad):

Rate pain symptoms on a scale of 0-10 (10 being the highest):

Pain: _____ Numbness: _____ Tingling: _____ Burning: _____ Tightness: ____

**To Be Completed by MA:

Blood Pressure: _____ HR: _____ Weight: _____ Height: _____

Program Number: _____ Treatment Number: _____ Placement: _____

SpO2: _____ Temperature: _____ Blood Glucose: _____ Respirations: _____