**Astra Zeneca LLRR Documentation**

-Shravwan, Madhaesh, Naveen, Jithin

Overview

The provided code is a simple Python application that uses the Tesseract OCR (Optical Character Recognition) engine to extract text from an image. The extracted text is then processed and inserted into a MySQL database. The application is built using the Tkinter library for the graphical user interface.

Components

**1. ImageReader Class**

The ImageReader class is responsible for configuring the Tesseract OCR engine based on the operating system (OS) and extracting text from an image.

**1.1. OS Enum**

Enumerates the supported operating systems (only Windows is currently supported).

**1.2. Language Enum**

Enumerates the supported languages for text extraction.

**1.3. ImageReader Class**

Methods:

\_init\_(self, os: OS): Initializes the ImageReader object and sets the Tesseract OCR command path for Windows.

extract\_text(self, image: str, lang: str) -> str: Takes an image file path and a language, then returns the extracted text using Tesseract OCR.

**2. User Interface (Tkinter)**

The Tkinter GUI allows users to input the file path of an image, extract text, and view the extracted text.

**2.1. Entry Field and Button**

Users can input the file path of the image in a text entry field.

Clicking the "Extract Text" button triggers the text extraction process.

**2.2. Result Display**

The extracted text is displayed in a label widget.

**3. Database Interaction**

The extracted text is processed and specific data points are inserted into a MySQL database named "astra." The database table is created if it doesn't exist.

**3.1. insert\_into\_database Function**

Processes the extracted text to obtain relevant data points.

Inserts the data into the "PATIENTDATA" table in the MySQL database.

Dependencies

Pillow (PIL): Python Imaging Library for image processing.

Pytesseract: Python wrapper for Tesseract OCR.

Tkinter: GUI library for creating the user interface.

MySQL Connector: Python driver for MySQL database.

Notes

The code assumes that the image contains specific medical data formatted in a certain way.

It creates a database table "PATIENTDATA" with specific columns for storing the extracted data.

Important: This documentation assumes the correct installation and configuration of Tesseract OCR, MySQL Connector, and other dependencies.