Book Search Database

**Idea:**

I wanted to design a system that can take an image of a bookshelf with many different titles and authors, then organize them based on highest rating and returns a list to the user. The idea for this project comes from my habit of buying an excessive amount of books and not knowing which to pick up next form my shelf. In the future, I would love to adapt this project to be able to sort books based on recent popularity or which ones are most similar to ones you have rated well in the past.

**Progress:**

I have incorporated the pytesseract, cv2, beautifulsoup and other libraries to allow me to use functions like taking text from an image and navigating files on my desktop. Temporally I used print statements to output one title as a test to ensure this part of the code was function properly before moving on. My initial test for this stage involved typing “The Hunger Games” into a word document, then screenshotting it and saving it as a jpg to be processed by my program.

A white background with blue text

Description automatically generated

The program passed the test, and I moved on to testing with a single book cover. Unfortunately, because books often have elaborate fonts, the program only read certain parts of the book title. My solution for this was to grayscale the picture before processing using the cv2 library.

A screen shot of a computer code

Description automatically generated

Original Image: Grayscaled Image:

A hand holding a book

Description automatically generated A book cover with text on it

Description automatically generated

Making this adjustment helped my program’s ability to pick out the words from the image substantially.

**Current Challenges:**

My plan for compare each book rating was to use Goodreads’ API and integrate the information directly into my program, however unfortunately Goodreads made their API private, so I had to consider other methods. I attempted integrating HTML scraping; however, it was unsuccessful in finding the ratings of multiple different book titles.

A computer screen with text on it

Description automatically generated

Currently I am researching other ways to get the same information, possibly from a different website