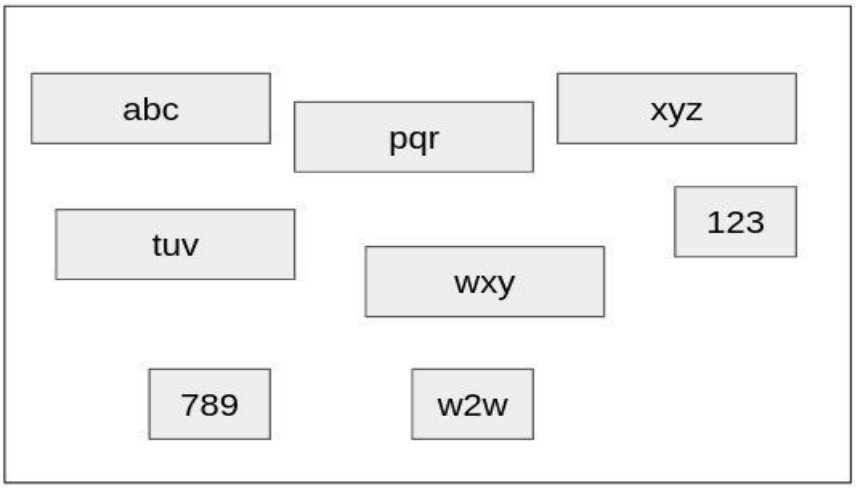
CV Data Engineer : Problem Statement

Problem Description:

Here is a CSV file with one column that contains the outputs from an optical character recognition (OCR) model. Each record contains a list of lists that describe the coordinates (of the bounding-box; 'bbox') of detected words and the words themselves. These words may appear in any random order.

The overall task is to generate a new csv from the given one, processing each record, where we need to rearrange the same information in a better order.

To elaborate, say we have an image with words detected by the OCR, like below:



The current list of lists contains entries like:
[[<bbox_coordinates>], ('tuv', <confidence_score>)],
[<bbox_coordinates>], ('pqr', <confidence_score>)],
[<bbox_coordinates>], ('w2w', <confidence_score>)],
[<bbox_coordinates>], ('123', <confidence_score>)], ... etc.]

You need to process these entries, such that the final output orders the words as if the location/position of the words were “flattened”; in that, the top left word appears first, then we take the word to its right, so on till we reach the top right word, and similarly continue with the second row, and then third, so that the last entry of the output is the bottom right word.

So, for the example above, we need to reorder the list of lists to give us the following order:
[[<bbox_coordinates>], (‘abc’, <confidence_score>)],
[<bbox_coordinates>], (‘pqr, <confidence_score>)],
[<bbox_coordinates>], (‘xyz’, <confidence_score>)],
[<bbox_coordinates>], (‘tuv’, <confidence_score>)],
[<bbox_coordinates>], (‘wxy’, <confidence_score>)],
[<bbox_coordinates>], (‘123’, <confidence_score>)],
[<bbox_coordinates>], (‘789’, <confidence_score>)],
[<bbox_coordinates>], (‘w2w’, <confidence_score>)]

The exact output format of the output CSV is for you to decide. Just that it should contain the same 1205 records for each output and conveniently highlight the sorted/flattened words and their bounding boxes. Make assumptions wherever needed and document them in your code.

You can use any tech stack/platform of your choice.
Please keep in mind that you should be able to explain your solution, in detail, during the followup codewalk round.

For technical questions please email mrinal@hammoq.com.
For other queries please contact shraddha@hammoq.com.