Programming assignment

Your task is to create an algorithm for identifying the country an address belongs to.

For example, when giving the algorithm the address string “Nieuwendammerkade 26A-5, Amsterdam, Netherlands” as input, the algorithm should be able to return a variable containing “Netherlands” as the country.

Your method should be robust to variations, as not all addresses are written in the same way, for example if the address omits the country as in “Nieuwendammerkade 26A-5, Amsterdam” we would still like to resolve it to “Netherlands”.

You can assume that the address will always contain a city.

A simple approach for this problem would be to index each city to its corresponding country. Then given an address string, split it into tokens and search for each token in the index, when a match is found return the corresponding country. This is a simple method, with many shortcomings, can you identify them? Can you improve it?

You’ll be provided with datasets of addresses as well as cities with corresponding countries, to build and validate your algorithm. You can find the datasets here: https://drive.google.com/file/d/1W5zbLpKfu-3l394TDioRKIhPfxy1wAgk/view?usp=sharing

You may use any publicly available data and open source tools and libraries. It should take approximately 2-3 hours to complete the assignment.

Requirements for the algorithm

- Fast computation time, potentially scalable and extendable to cover the entire world
- Self-contained: no use of external resources (e.g. external API endpoints) during computation
- Accepts as input an address, and returns a hash as output.
- Structure the code with a separation of business logic and data logic.

Result of the assignment

- Source code of your algorithm
- Any required data and description of how to update the data to support more countries
- Short description of the approach and limitations of the implementation, and possible room for improvement in the future.

- Instructions on how to run and test your algorithm on our side

Evaluation criteria

- Ability to transform functional requirements into a working system
- Code quality: correctness, readability, documentation
- Code structure: flexibility to encompass new feature requests and separation of concerns
- Testing of the solution

Good luck!