

Q1.

b. Approach:

- Create a database of ingredients and classify them on the following basis:
  1. Basic (spices, oil, vinegar, etc.)/Non-basic
  2. Vegan/Vegetarian/Non-vegetarian
  3. Local/Non-local
  4. Expensive/Cheap
  5. Cuisine: Continental/Mediterranean/Oriental, etc.
  6. Anything more that is relevant
- The classification could be done manually or better deploy an ML algorithm to predict these in case many new ingredients are added.
- Do a search of the article and create a term document matrix with TF – IDF weighting (or just normal count weighting since this is just a single article we are using). Keywords related to the above categories such as ‘vegan’ (signifying Vegan), ‘boutique’ (signifying Expensive), etc. would be included in this.
- Create an algorithm that sorts the input ingredients based on these priorities (and in the following order):
  1. Vegan/Vegetarian/Non-vegetarian depending on the occurrence of keywords suggesting these
  2. The type of cuisine depending on the occurrence of keywords suggesting these
  3. Occurrence of an ingredient in the article with the highest repeated at the top
  4. Expensive/Cheap depending on the occurrence of keywords suggesting these
  5. Local/Non-local
  6. Non-basics first and basics at last
- Return the created ranked list of ingredients (could add descriptions for each from the web)
- My code is not complete owing to the lack of time and my inexperience.

Q2. Code is attached in the Github project.

Q3.

- a. I want to gain professional experience in the field of data science, explore new areas and ideas in data science, and also learn and develop the advanced skills necessary in this field. I think Spoonshot is an interesting and unique startup utilizing data science in a novel field, and I am very interested in working for and inputting my analytical skills in this company.
- b. Data science is one of the most sought-after fields at present, with many applications and domains, and is the most interesting field to me. I enjoy coding, research and analysis (in the same order) which are among the core aspects of this field, and I want to add to this field in some positive way.

- c. I don't have such a problem in my mind right now since I'm just a sophomore student and a novice who is just properly beginning to explore the field of data science.