

Computational Gastronomy

Coding Assignment 1

You may use Python and Jupiter Notebook as an IDE for completing the assignments and documentation.

Notes: You are responsible for the backup of data as well as results, which will be used for evaluation.

Follow the rubric diligently while submitting. Name the files with the question numbers.

1. Identify an online recipe repository.
 - (a) Scrape (using libraries such as BeautifulSoup) any **10,000 recipes**. **Submit the raw data.** [5]
 - (b) Write a script for extracting information about the 'name of the ingredients' from the ingredients' section. **Store recipes in the form of a (Recipe ID)—(Ingredient Name) form.** [10]
2. Analyze the data obtained for the following.
 - (a) Find the number of unique ingredients. List them with their frequencies. **Submit the file.** [2]
 - (b) Plot the recipe size distribution for these recipes and the average size of the recipes (s). Properly label the axes. **Submit the file.** [3]
 - (c) Plot cumulative distribution of recipe size (label axes properly). **Submit the file.** [5]
3. For the data of recipes obtained in the above question (1):
 - (a) Plot the frequency-rank distribution. Scale and label axes properly. **Submit the file.** [5]
4. Implement the following copy-mutate algorithm for data obtained in question (1). [10]
 1. **Start** with Epoch = 0 and Number of Recipes Per Epochs = Round (Number of recipes/Number of Epochs)
 2. **Initiate the primordial cuisine** ($t = 0$), Nature Basket (NB) and the Kitchen Basket (KB):
 - Size of recipes = s (as obtained in 2(b), above)
 - Number of recipes at $t = 0$, $N_R^{t=0} = 100 - 300$ [adjust as needed to have unique recipes]
 - Size of Kitchen Basket = **[Choose so as to maintain the Kitchen-to-Nature Basket size ratio]**
 - $N_R^{t=0} \gg$ Size of KB
 3. Pick a random recipe from the primordial kitchen for the 'modification.'
 4. Pick an ingredient randomly from the chosen recipe & one from KB.
 5. Compare the ingredients.
 6. If the KB ingredient is not the same as recipe ingredient replace the latter with the former.
 7. **Repeat Steps 4—5** until the Recipe gets modified.
 8. Add the new 'modified/mutated' recipe to the Kitchen Basket if the same isn't there already.
 9. **Repeat 3—7** until the number of recipes reaches the desired number of the **next Epoch**.
 10. Compute and plot the recipe size as well as the frequency-rank statistics.
 11. Run at least 5 instances of cuisine evolution. **Compute and plot the recipe size as well as the frequency-rank statistics across the random instances. Submit the properly labelled file.**
5. Modify the above algorithm in question (4) to include the addition and deletion, over and above the modification, of ingredients. **Compute and plot the recipe size as well as the frequency-rank statistics** across the random instances. **Submit the file.** [10]
6. Improvise the above algorithm in question (5) to ascribe fitness values to each of the ingredients in the Nature Basket according to their frequency/popularity and picking a **fitter** ingredient at every instance of replacement. **Compute and plot the recipe size as well as the frequency-rank statistics** across the random instances. **Submit the file.** [10]