https://github.com/Sukumaru/TIMORIS/blob/main/design\_documents/SRS.md

Homework 1 – Project design and data selection

DUE: Wednesday, Feb 12, 2025, at 11:59pm

TOTAL 30 points

PROBLEMS: 3

##############################################################################

PROBLEM 1: (10 points) Select and validate data & choose first feature to implement

1. (5 points) Identify data to use for the project and identify or compute relevant statistics on that data to validate that the data meets the project requirements for input and describe the data format and requirements in the Software Requirements Specification (SRS) document. You should also add the link to the data with a brief description and the validation stats in the datasets.md document.

In this part of the assignment, you describe the input data that you will use to answer the biological question(s) you have in mind – meaning all data you want to process. Keep in mind that as you build the tool is not efficient to use large datasets so you will use a small (representative) subset of the data or simulated data for the tool development and only at the end you will use the full data to validate that the tool can be used to answer the question(s) it was built to answer on full scale datasets.

* + To account for that, post url for the SRS and datasets.md documents, the SRS can be changed to a .md format
  + SRS DOCUMENT URL: <https://github.com/Sukumaru/TIMORIS/blob/main/design_documents/SRS.md>
  + Datasets.md DOCUMENT URL: <https://github.com/Sukumaru/TIMORIS/blob/main/design_documents/datasets.md>

1. (5 points) Now that you identified the data you need, the next step is to identify the first basic output that you would like to compute and the steps to compute it – some of the steps may involve data download or data simulation and/or pre-processing so that the data is in the format needed to compute the output

Add that as an activity in WBS document and break it into tasks as you describe the steps of the computation.

* + To account for that, post the url to the WBS document
  + WBS DOCUMENT URL: <https://github.com/Sukumaru/TIMORIS/blob/main/design_documents/WBS.md>

An example for the Metabolic pathway analysis project would be to compute a score for one pathway metabolite – so I would have to figure out the following:

* what is the structure I use for the pathway data?
  + list of reaction objects with two lists of metabolites left and right
* how do I identify a metabolite from a pathway?
  + I would go to the first rection’s left side and take the first element (metabolite ID) in that list
* what is the structure I use for the metabolite data?
  + matrix with metabolites as rows (row labels are metabolite IDs) and condition and control columns and average metabolite levels for the two groups as the values in the matrix
* how do I find a metabolite in the metabolite data?
  + I would subset the data matrix using the metabolite ID
* what is the score I want to compute?
  + Log 2 ratio between condition and control average metabolite level
* What would be the algorithm to compute the score?
  + Implement the math formula for Log 2 ratio and use the matrix values retrieved by subsetting the metabolite data matrix

##############################################################################

PROBLEM 2: (10 points) Break project into modules and tasks

1. (5 points) Break project into modules and describe each module scope and content, meaning what will contain and what connects the content to make it a module

Add that to the DDS document below the goals and milestones you can add a Modules and their dependencies diagram – you can change the format of the DDS document if you want to make it .docx or .md or another format.

If you keep it a draw.io document, please make sure to also save the DDS document into an image format (e.g. tool\_DDS\_wmodules.png in addition to having the tool\_DDS\_wmodules.drawio) so that when the link to the document is accessed the browser can display the image.

* + To account for that, post the Design Document Specification url
  + DDS URL:
  + <https://github.com/Sukumaru/TIMORIS/blob/main/design_documents/DDS.md>

1. (5 points) Looking forward: Add another least 3 activities to the Work breakdown structure document and break them into actions and sub-actions and tasks and describe each task – including the task deliverable and completion criteria
   * To account for that, post the url to the WBS document
   * WBS DOCUMENT URL: <https://github.com/Sukumaru/TIMORIS/blob/main/design_documents/WBS.md>
   * – I should see in the GitHub repository history the difference to the WBS document entries for Problem 1B, or mark them with a comment.

##############################################################################

PROBLEM 3: (10 points) GitHub repository audit

1. (5 points) Add modules to your codebase with basic documentation (comments) and define functions for your main functionality and implement at least one function
   * REPO URL: <https://github.com/Sukumaru/TIMORIS>
   * URL for the code file with the implemented function: <https://github.com/Sukumaru/TIMORIS/blob/main/src/detection/spatial_analysis.py>
2. (5 points) You should have a project that builds in your GitHub repository
   * To account for that, post a screenshot of the build result
   * BUILD SCREENSHOT:

A screenshot of a computer

Description automatically generated

##############################################################################