

USER MANUAL FOR COMPUTATIONAL APPROACH FOR DISEASE-GENE ASSOCIATIONS

STEPS TO RUN PROJECT:

To run the project, we require minimum of 16GB RAM system.

1. Install Anaconda Navigator.
2. Open Jupyter notebook from Anaconda Navigator.
3. Place the included notebook files and datasets with the names 'string_interactions.tsv' and 'Disease_Gene_Assocs.csv' in a folder.
4. Open the notebook file named 'Data Preprocessing.ipynb'.
5. Navigate to the 'Cell' menu, click on it then click on 'Run all Cells'.
6. On clicking 'Run all Cells ' two files named 'Dga_more.csv' and 'ppi_more.csv' will be generated.
7. Open the notebook file named 'GA_Disease_assoc.ipynb' and run all the cells as mentioned in step 5.
8. By default the parameters are set(population_size=965, mutation_percent=0.02, number_of_iterations=400, factor=3, k_value=10). If required adjust the parameters (population_size, num_parents_mating, mutation_percent, factor, k_value).
9. After all the cells are run successfully the Output will be generated at the last cell of the notebook file.