**Network Neighborhood Analyzer**

P.A.R.R. Wijerathne

**Introduction**

This application uses the NetworkX package to analyze the immediate neighborhood of protein-protein interaction (PPI) networks. With a range of methods, it enables users to explore PPI data, calculate protein degrees, visualize distributions, and predict candidate genes. This application is based on file handling, and will use two input files.

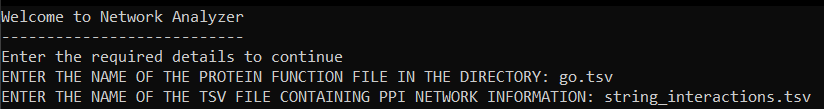
1. A tsv file containing protein-protein interaction data.
2. A tsv/text file containing protein information related to a known function.

**Features and Functionalities**

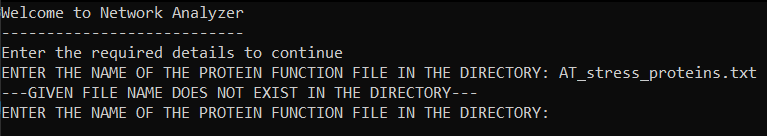
* Construct network graph related to the input file
* Based on two algorithms namely ‘Majority voting score’ and ‘Hishigaki algorithm’ where the user can select the preferred algorithm.
* Can find the number of proteins and interactions in a given PPI network.
* Calculate the degree of a given protein in a given PPI network.
* Plot the degree distribution as a histogram of a given PPI network.
* Count the number of proteins annotated to a given function in the immediate neighborhood (these function protein list should be given as an input file) of a given protein in a given PPI network.
* To plot the distribution of annotated protein count in the immediate neighborhoods for all the proteins (you calculated in the previous method).
* Predict candidate genes for a single function using the Hishigaki algorithm.
* Display Hishigaki score distribution as a histogram.
* Display Hishigaki score distribution as a boxplot. This method should contain a Boolean method parameter to remove zero scores when specified.
* Is not taxonomy specific, can be used for different functionalities in different taxons.

**How to use Network Neighborhood Analyzer**

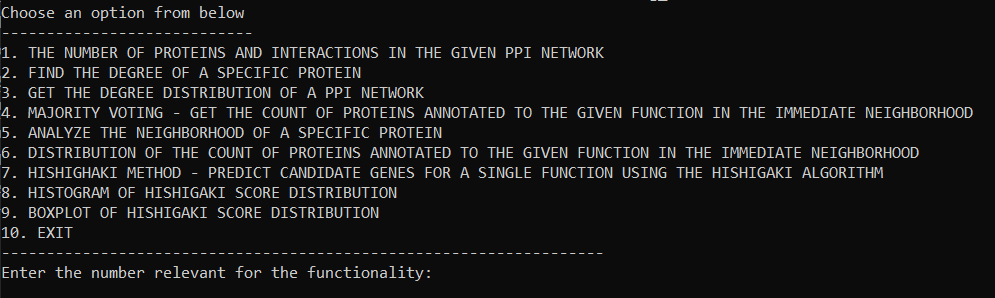
1. After copying the relevant files to the directory folder, run the .exe file and enter the correct file names. The program will check if that file exists in the directory, or it will raise an error.



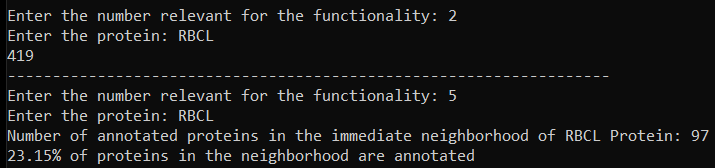
If incorrect file names are given,



1. If the files are successfully entered to the program, it will display a menu.



1. In the menu, the specific functionality that the user is seeking should be selected. (In certain options Protein symbol might have to be given as an input too.)



1. Input number 10, to exit the main menu.

**Future Improvements**

1. GUI to make the application more user friendly.
2. Connect to external databases upon user requests.
3. Improve annotations for references for all available taxons.