**Virus-Human PPIs Centrality Analysis**

This repository contains Python code for analyzing virus-interacting human proteins in the context of the human protein-protein interaction (PPI) network. The workflow calculates network centrality measures for virus-interacting proteins and performs statistical comparisons with non-interacting proteins.

## Features

* Constructs a human PPI network from user-provided interactome data.
* Labels proteins as virus-interacting or non-interacting.
* Calculates centrality measures: Degree, Closeness, Betweenness.
* Applies log2 transformation to degree, closeness centralities and betweenness.
* Performs Wilcoxon rank-sum tests to compare virus-interacting vs non-interacting proteins.
* Saves results in Excel files for downstream analysis.

## Usage

1. **Prepare your data files in a folder**:

* Interactome\_data.xlsx – Human PPI network
* VIRUS.xlsx – Virus-interacting human proteins (replace VIRUS with the actual virus name)

1. **Update the file paths in centrality\_analysis.py**:

* data\_folder = "path\_to\_your\_data\_folder"
* virus\_file = os.path.join(data\_folder, "VIRUS.xlsx")

1. **Run the script**:

* python centrality\_analysis.py

1. **Outputs**:

* centrality\_results\_<VIRUS>.xlsx – Centrality measures for all proteins
* wilcoxon\_test\_results\_<VIRUS>.xlsx – Statistical comparison results

## Requirements

* Python 3.8+
* pandas
* numpy
* networkx
* scipy
* openpyxl (for Excel I/O)

## Notes

* Replace VIRUS.xlsx with the specific virus file you want to analyze.
* The code is generalized and works for any virus if the corresponding Excel file is provided.
* This workflow can be extended to analyze multiple viruses by providing multiple virus files in the same folder.