**HW4 – Epidemics**

**Submission Deadline:** 16.1.2021, 23:00

In this assignment, you will simulate and estimate the spread of an epidemic under different model assumptions and consider a number of vaccination strategies.

**Submission guidelines:**

You are required to submit your solution as a zip file:

* Python file with your functions implemented, the file’s name should be ID.py where ID is the student’s ID.
* A PDF file with answers to open questions (these are marked in a blue font).
* The zip file’s name should be ID.zip where ID is the student’s ID. For example, the student Moshe Moshe with an ID of 1234567, should submit a zip file “1234567.zip”, containing his implemented solution “1234567.py” and a pdf file.
  + Do not zip the directory where your solution is stored, only zip the required files!
* You are required to implement a function called ‘get\_name’ that returns your full name in English.
* You are required to implement a function called ‘get\_id’ that returns your ID number.
* You have to follow **the exact API described in the HW** (exact function names, parameters and returned types). Please avoid typos.
* Your code **should not contain any part of loading data**. You may include a ‘main part’ code block (using the if \_\_name\_\_ == "\_\_main\_\_": syntax). Within this code block you can load data and test your implementation.

**PART 1 - Epidemic Spread**

1. In this question you analyse the spread of an epidemic in a population. You are required to estimate the infection state by simulating the spread of a pathogen.

The population and the connections between individuals are represented by a *contact* network. Edge weights denote the number of interactions between the individuals in a time unit (e.g., a day).

Each individual (i.e., node) is assigned with two properties:

1. status - either ‘s’ (susceptible) or ‘i’ (infected).
2. mortality\_likelihood - a value in the [0,1] range. It reflects the probability of the individual to die from the disease in a time unit (while he is infected).

You should implement the function ‘epidmeic\_analysis’.

**The function should simulate an ‘epidemic evolution’ over time.** The function emulates the situation of the epidemic after a given number epochs (i.e., time intervals).

The function’s input parameters are:

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter name** | **Parameter type** | **Explanation** | **Default value** |
| ‘network’ | networkX object | The network to run the analysis over. The network contains edge information (infection probability) as well as node information (the status of the node - ‘S’, ‘I’ or ‘R’). |  |
| ‘model\_type’ | str | The epidemic model behavior. Can be either ‘SIS’ or ‘SIR’ | ‘SIS’ |
| ‘infection\_time’ | int | Number of epochs that an infectious person is infectious. | 2 |
| ‘p’ | float | Probability of an infectious node to pass the disease to another node upon a contact (in a single meeting) | 0.05 |
| ‘epochs’ | int | Number of time unit (e.g., days) to apply the simulation over | 20 |
| ‘seed’ | int | The seed to be used for simulation | Your ID |

The function should return a dictionary with three key-values:

* ‘infections\_total’ - int. Total number of infections that occur over all epochs.
* 'infectioius\_current’ - int. Number of infected individuals after ’epochs’.
* ‘mortality\_total’ - int. Total number of mortality cases in the population.
* ‘r\_0’ - estimated reproductive number after all ’epochs’.

Note that the ‘network’ you are given as input contains information about interactions (edges) between individuals (nodes). Each interaction (edge) is assigned with a number - that is the number of interaction between two users (during a time unit).

1. Use the function ‘epidmeic\_analysis’in order to analyse the spread of an epidemic over the following [two contact networks](https://moodle2.bgu.ac.il/moodle/mod/folder/view.php?id=1687284).

You should run the function per each network using 4 different setting as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Setting A** | **Setting B** | **Setting C** | **Setting D** |
| **‘model\_type’** | SIS | SIS | SIR | SIR |
| **‘infection\_time’** | 2 | 5 | 2 | 5 |
| **‘p’** | 0.05 | 0.1 | 0.05 | 0.1 |
| **‘epochs’** | 20 | 20 | 20 | 20 |

Report the results you get per each setting. Make sure you run the analysis and report the result via averaging **repeated simulations** (at least 10), in order to obtain meaningful statistics.

Can you conclude how the different parameters affect the population health in the networks provided?

**PART 2 - Vaccination Policy**

1. In this part you will compare three different vaccination policies:
2. Random - vaccinate a subset of the population, chosen at random.
3. Betweenness - vaccinate a subset of the population with the highest betweenness centrality measure.
4. Degree - vaccinate a subset of the population with the highest degree centrality measure.
5. Mortality - vaccinate a subset of the population with the highest mortality likelihood.

You need to re-run the simulation from Q1 but this time there is a finite number of vaccines available *prior* to to the beginning of the epidemic.

Write a function called ‘vaccination\_analysis’ with the following input parameters:

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter name** | **Parameter type** | **Explanation** | **Default value** |
| ‘network’ | networkX object | The network to run the analysis over. The network contains edge information (infection probability) as well as node information (the status of the node - ‘S’, ‘I’ or ‘R’). |  |
| ‘model\_type’ | str | The epidemic model behavior. Can be either ‘SIS’ or ‘SIR’ | ‘SIR’ |
| ‘infection\_time’ | int | Number of epochs that an infectious person “carries” the disease and risks his susceptible neighbors. | 2 |
| ‘p’ | float | Probability of an infectious node to pass the disease to another node he is in contact with (in a single meeting) | 0.05 |
| ‘epochs’ | int | Number of epochs (e.g., days) to apply the simulation over | 10 |
| ‘seed’ | int | The seed to be used for simulation | Your ID |
| ‘vaccines’ | int | Number of available vaccines | 1 |
| ‘policy’ | str | The required vaccination policy. Can be either ‘rand’, ‘betweenness’, ‘degree’ or ‘mortality’ | ‘rand’ |

The function should return the same dictionary as in part1.

1. Use the function ‘vaccination\_analysis’in order to analyse the spread of an epidemic over the following [two contact networks](https://moodle2.bgu.ac.il/moodle/mod/folder/view.php?id=1687284) when a limited number of vaccines are available.

You should run the function per each network using 4 different setting as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Setting A** | **Setting B** | **Setting C** | **Setting D** |
| **‘model\_type’** | SIS | SIS | SIS | SIS |
| **‘infection\_time’** | 3 | 3 | 3 | 3 |
| **‘p’** | 0.1 | 0.1 | 0.1 | 0.1 |
| **‘epochs’** | 30 | 30 | 30 | 30 |
| **‘vaccines’** | 20 | 20 | 20 | 20 |
| **‘policy’** | ‘rand’ | ‘betweenness’ | ‘degree’ | ‘mortality’ |

Report the results you get per each setting. Make sure you run the analysis and report the result via averaging **repeated simulations** (at least 10), in order to obtain meaningful statistics.

1. Which policy is the best one in terms of:
   1. Total number of infections
   2. The number of infected individuals after the specified epoch.
   3. The total number of mortality cases
2. What are the drawbacks of the vaccination policies suggested in this question? How can those be improved?

**Notes:**

For answering the questions in the best way, you are given [two contact networks](https://moodle2.bgu.ac.il/moodle/mod/folder/view.php?id=1687284). Obviously, you are encouraged to experiment with other networks and other settings.

Please make sure you install the required packages using the latest version.

If you use any package that is not mentioned in the list below, please include a “requirements.txt” file, stating which packages you used for your solution.

List of packages:

* Pandas
* NetworkX
* Numpy
* Scipy
* tqdm
* Pickle