**INST414 – Medium Post Grading Rubric, Assignment 3**

Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Topic\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |
| --- | --- | --- | --- |
| **Item** | **Item Name** | **Max** | **Earned** |
| 1 | Identify a non-obvious insight you want to extract from your data, describing what decision(s) this insight might inform. | 10 |  |
| 2 | Describe the source of your network data. Where did it come from, and how did you collect it? | 10 |  |
| 3 | Define the features you're using when measuring similarity, and which similarity metric you will use. | 10 |  |
| 4 | For each of your three (3) chosen query items, identify the top 10 most similar items in your dataset and list them. | 15 |  |
| 5 | Describe any software you used to facilitate your analysis (e.g., SKLearn, NetworkX, or similar). | 10 |  |
| 6 | Describe how you cleaned up this data, common bugs you think others might encounter, and how you fixed them, etc. | 5 |  |
| 7 | Include figures or tables summarizing your findings. | 5 |  |
| 8 | Conclude with a discussion of the limitations of your analysis. What’s missing? How might it be biased? | 10 |  |
| 9 | Include a link to one of your GitHub repositories that contains the code you have developed for this assignment. |  |  |
| 10 | Excellence | 10 |  |
| 11 | Length | 5 |  |
|  |  |  |  |
|  | TOTAL | 100 |  |

Other Comments: