

## Instructions for Project Report 2 for MGS 655

The purpose of the report is to provide a comprehensive summary of everything that you did in your second project for MGS 655.

Your writeup should be at most 2 page(s) with normal fonts (10 - 12 font in Arial, Helvetica, Times, etc) and margins (one inch on all sides). You can use single or double column format depending on what is relevant for the project.

The organization should be roughly as follows:

1. **Team Composition and Contributions [Not more than 2 sentences]:** Who is doing the project? If participants are from different batches, please indicate so here. How was the work split up among the partners if relevant.
2. **Connection to CCR Cluster:** Describe (in not more than 2 sentences) how you set up the environment for your project.
3. **Data sets:** Explain how you replicated the data - manually, wrote scripts? This section should clearly indicate whether you merge all the data together, how many articles you are working with and why you chose this number.
4. **Design Issues:** Please provide descriptions of the design of your Mapper, Reducer and Main program (if any).
5. **Empirical Results and Discussions:** Present a brief summary of your empirical results of your algorithm. Graphs should be legible with X and Y- axis clearly marked. Present ONLY the main result(s) in the 2 page document. If you have more graphs/plots you would like to share, please use an Appendix [Appendix: Should not be more than one page.]. Discuss your results and observations. PLEASE DO NOT PASTE GRAPHS AND CHARTS AND EXPECT ME TO UNDERSTAND THEM. Please provide answers to the following questions:
  - How many ?unique? trigrams were identified by your program?
  - Does a change in the number of mappers and reducers affect the run time of your program? [Please refer to the problem statement to study two specific cases]
  - Did you have substantially long wait times before your job was processed?
6. **Practical Experiences** What were the challenges encountered during the project? Were you able to solve all of them? If not, what was the work around that you adopted? Is that fair? What would have liked to do if you had infinite resources available?
7. **Appendix:** You are allowed one page of the appendix. Use it only if you need it. No extra credits for the appendix. You can provide charts, tables, and anything else that does not fit into the main project report.
8. **Submission Instructions:** The project report, and code (if relevant) should be submitted. There will be separate links for submission of each of the following items. Please upload accordingly.