

Class 8: Homework

Big Data Application Development
Spring 2018



Homework

Class 8

A. Analytics Application

1. Research your project. **Each team member should upload a list of 5 papers relevant to your project and a short summary for one paper which includes your thoughts on how the paper is connected to your project.**

It is useful to understand the state of the art before you begin a project. To do this, we read recently published papers. IEEE and ACM conferences and journals from 2015 to present are a good source.

Please do not choose papers on a Hadoop technology or other tool. Instead, choose papers related to your project thesis - the paper does not have to be in the same domain as your project. For example, do not choose a paper about Spark MLlib, do choose a paper about using big data in healthcare to solve some problem.

Where do I find a paper?

- Try using GoogleScholar to find papers.
- Try googling 'IEEE Big Data Analytics' for example - this brings up a bunch of conferences to pick from. Let me know if you have trouble finding a paper.
- Ask the professor if you get stuck.

Each team member should find at least 5 related papers. List all five papers - title, authors, link where you found it - at the top of your document. Below this list, add a summary (a few paragraphs) about the paper. This summary should include your thoughts on how the paper relates to your project. Upload to NYU Classes. (In a future homework, you will add the team summaries to the 'Related Work' section of your project paper.)

Coordinate with your teammates to ensure each member reads a different paper. Share what you learn with your teammates.

Note: The MapReduce, HDFS, and other papers already assigned cannot be used for this assignment.

Some places to look for papers:

ACM KDD Conference: <http://www.kdd.org/kdd2016/>

Google Scholar: <https://scholar.google.com/>

ACM DL: <http://dl.acm.org/>

Please provide the following:

- 5 papers - title, authors, link to paper
- For the summarized paper, provide:
 - Paper title
 - Paper authors
 - Link to the paper
 - Paper abstract
 - Your summary

Homework

Class 8

A. Analytics Application (continued)

2. Draw *initial* diagrams using PowerPoint, Visio, etc. to describe your project. Include the software architecture (Big Data tools you'll use), the data flows, and anything else you think is important to show. This is a first draft, you will refine it in the coming weeks. **All team members should upload the diagrams.**
3. Create a list of tasks for your analytics project - you must use the TaskList.xlsx/numbers template in Resources. Assign team members to tasks, and assign a due date to each task. This can be just a simple table showing tasks, team member name, and target completion date. Try to identify milestones – that will help you know if you are on or off track. The next page has some suggested tasks, feel free to add to it. **All team members should upload the schedule.**
5. Create a rendering of how you picture your final visualization or UI looking. **All team members should upload the rendering.**