CMP461 Big Data

Course Information

Dr. Lydia Wahid

Instructor Info

- *Course Instructor:* Dr. Lydia Wahid
- *Email:* LydiaWahid@Outlook.com
- *Office hours:* TBA

Objectives of the Course

- Learn the concepts of Big Data and Big Data Analytics.
- Learn Big Data and Big Data Analytics techniques and how to solve problems whose solutions are enabled by technology that can support the analysis of Big Data datasets.
- Acquire the necessary knowledge to develop working code for real-world Big Data applications.

Learning Outcomes (LOs)

1. Identify and Describe the fundamentals of Big Data and Big Data Analytics. (NARS mapping: A.1)
2. Apply MapReduce processing technique and Design its different components that suits the Big Data problem at hand. (NARS mapping: C.1)
3. Acquire the necessary knowledge of the different Big Data processing frameworks and know how to use them to build a Big Data application. (NARS mapping: B.3)
4. Understand Big Data Predictive Analytics and know how to apply them on Big Data problems. (NARS mapping: C.2)
5. Understand Big Data Descriptive Analytics and know how to apply them on Big Data problems. (NARS mapping: C.2)

Learning Outcomes (LOs)

6. Understand Big Data Text Analytics and know how to apply them on Big Data problems. (NARS mapping: C.2)
7. Understand the concepts and strategies created to achieve cost-effective and highly scalable storage solutions for Big Data. (NARS mapping: A.3)
8. Acquire the necessary knowledge of the Big Data storage technologies. (NARS mapping: B.2)
9. Get acquainted with real-world Big Data applications and know how to apply the learnt techniques to design those applications. (NARS mapping: C.2)
11. Develop a hands-on experience by developing a real-world Big Data application through the term project. (NARS mapping: C.1)

Grades Distribution

- Two-Semester System:
  - Final Exam → 70 marks
  - Midterm Exam → 10 marks
  - Labs → 4 marks
  - Lecture Quizzes → 4 marks
  - Project → 12 marks

  - Total → 100 marks

Lecture Attendance

- For on campus weeks, the attendance will be recorded normally by calling students names during Lecture time.

- For online weeks, there will be a quiz after each online Lecture to ensure that students have watched the online material, from such quiz the attendance will be recorded. (These quizzes have a total of 4 marks as indicated in the grades distribution)

Exams Policy

- Midterm and Final exams will be held on campus.

- Midterm and Final exams will be **Open-Material**:
  - You can get only hardcopy material.
  - **No** electronic devices are allowed.
  - **No** borrowing of any material from your colleagues during exam time is allowed.

References

- Thomas Erl, Wajid Khattak, and Paul Buhler. *Big Data Fundamentals: Concepts, Drivers & Techniques (1st. ed.)*. Prentice Hall Press, USA, 2016

- C.S.R., P., et al. *Big Data Analytics: Systems, Algorithms, Applications*. Springer Nature, 2019.

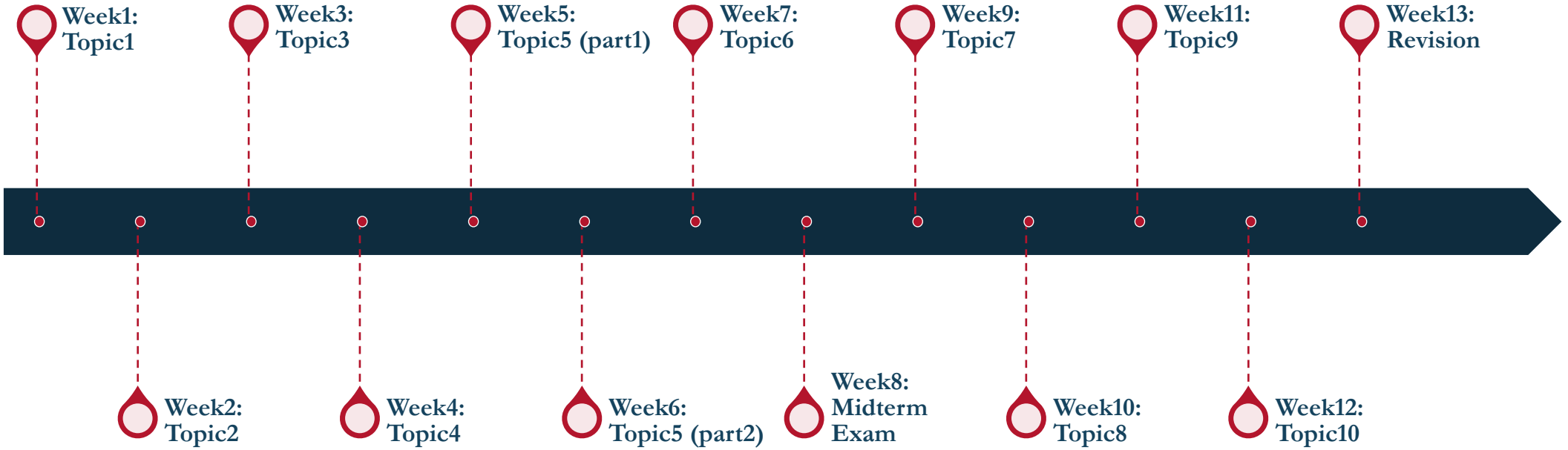
- Schintler, L. and McNeely, C. *Encyclopedia of Big Data*. Springer International Publishing, 2020.

Topics

1. Fundamentals of Big Data
2. Fundamentals of Big Data Analytics
3. Big Data Processing Frameworks
4. Big Data Processing Techniques
5. Big Data Analytics – Predictive Analytics
6. Big Data Analytics – Descriptive Analytics
7. Big Data Analytics – Text Analytics
8. Big Data storage concepts and strategies
9. Big Data storage technologies
10. Big Data real-world applications

Lectures Schedule

Odd weeks Online



Even weeks on campus

Sections Schedule

Week2:	Lab1 – Introduction to R
Week4:	Lab2 – Integrating Hadoop and R + MapReduce
Week6:	Lab3 – Predictive Big Data Analytics
Week7:	Sheet1
Week8:	Midterm Exam
Week10:	Lab4 – Descriptive Big Data Analytics
Week11:	Sheet2
Week12:	Final Project Delivery

Project

Week6:	**Project Proposal Delivery**
Week7:	**Project Proposal Feedback from TA**
Week12:	**Final Project Delivery**

Note:

- Project Proposal :
  - All teams will receive feedback from the TA within a maximum of 1 week from the delivery date.

Thank You