Course Outline

1. Document Information

Degree Program	Computer Science
Course Number	CS 534
Course Title	Big Data Management and Analytics
Semester Hours	3
Course Coordinator	Dunren Che
Revision Term	Fall 2020
Latest Revision	Spring 2021

2. Catalog Description

This course provides comprehensive and in-depth discussions of big data management and analytics. Main subjects include computation and programming models, management and analytics algorithms, and platforms/frameworks especially designed for big data. The objective of this course is to equip students with the ability to understand, use, and build big data management and analytics systems or tools.

3. Textbooks

- Lin, J. & Dyer, C. (2010). Data-Intensive Text Processing with MapReduce. Morgan & Claypool.

4. References

- Anand, X. & Ullman, J. (2010). Mining of Massive Datasets, Rajaaraman. Cambridge.
- Lam, C. (2011). Hadoop in Action. Manning.

5. Course Learning Outcomes

- Understanding the key features and issues of Big Data.
- To learn the important approaches to Big Data management.
- To learn the computation models and frameworks of Big Data.
- To study the data mining methods designed/customized for Big Data.

6. Assessment of the Contribution to Student Outcomes

Outcome	1	2	3	4	5	6	7
Assessed	X	X			X		

7. Prerequisites by Topic

CS 430 with a grade of C or better or graduate standing.

8. Major Topics Covered in the Course

1. Course Introduction {1 lecture}
2. Introduction to Big Data {2 lectures}
3. Big Data Collection {2 lectures}
4. Big Data Storage Systems {5 lectures}
5. Big Data Computation Models {5 lectures}
6. Big Data Management {5 lectures}
7. Big Data Mining {5 lectures}
8. Learning Insights from Big Data {5 lectures}
9. Big Data Visualization {5 lectures}
10. Crowdsourcing {5 lectures}

Total hours: 40 lecture hours plus extra seminar hours