Title (Units): **COMP7930 Big Data Analytics (3,2,1)**

Course Aims: To introduce the basic knowledge of big data analytics; to learn the techniques and tools for big data analytics; to conduct application case studies to show the usage of big data analytics.

Prerequisite: Basic knowledge in linear algebra, probability and statistics, and basic database concepts.

Course Intended Learning Outcomes (CILOs):
Upon successful completion of this course, students should be able to:

No.	Course Intended Learning Outcomes (CILOs)
	Knowledge
1	Identify and distinguish big data analytics applications
2	Describe big data analytics tools
3	Explain big data analytics techniques
4	Present cases involving big data analytics in solving practical problems
	Professional Skill
5	Conduct big data analytics using system tools
6	Suggest appropriate solutions to big data analytics problems

Calendar Description: This course aims to introduce the basic knowledge of big data analytics as well as the common data analytics techniques and tools. Furthermore, their potential applications to a variety of domains such as business, finance/banking industry, and health care are shown via case studies.

Teaching and Learning Activities (TLAs):

CILOs	Type of TLA
1-3	Students will learn the concepts and techniques via lectures, in-class discussions, quizzes, and assignments.
4-6	Students will learn the skills via guided laboratories and mini-projects.

Assessment:

No.	Assessment Methods	Weighting	CILOs to be addressed	Description of Assessment Tasks
1	Continuous assessment	40%	1-6	Continuous assessments are designed to measure how well the students have learned the concepts and techniques in big data analytics as well as the skills for solving real-world big-data problems.
2	Examination	60%	1-4, 6	Final examination questions are designed to see how far students have achieved their intended learning outcomes. Analysis based questions will be used to assess the understanding of big data analytics concepts and techniques. Problem solving questions will be used to assess the students' ability in tackling applications in big data analytics.

Assessment Rubrics:

Excellent (A)

- Achieve all the six CILOs, demonstrating a good mastery of both the theoretical and practical aspects of big data analytics
- Have a thorough understanding of big data analytics concepts and techniques, and be able to explain and highlight the key points of these concepts
- Able to conduct big data analytics according to application requirements, using system tools such as MapReduce/Hadoop

	• Able to draw on a variety of relevant knowledge and concepts and apply them to solve big data analytics problems
Good (B)	• Achieve all the six CILOs, demonstrating a good understanding of both the theoretical and practical aspects of big data analytics • Have a good understanding of big data analytics concepts and techniques • Able to conduct big data analytics according to most of the application requirements, using system tools such as MapReduce/Hadoop • Able to make use of relevant knowledge and concepts and apply them to solve big data analytics problems
Satisfactory (C)	• Achieve most of the six CILOs, demonstrating a basic level of understanding of the theoretical and practical aspects of big data analytics • Have a basic understanding of big data analytics concepts and techniques • Able to conduct big data analytics according to some of the application requirements, using system tools such as MapReduce/Hadoop • Able to make use of some relevant knowledge and concepts and apply them to solve big data analytics problems
Fail (F)	• Achieve less than three of the CILOs, and have little understanding of the theoretical and practical aspects of big data analytics • Unable to provide solutions to simple problems which require basic understanding of big data analytics concepts • Unable to conduct big data analytics using system tools • Unable to apply big data analytics concepts and techniques to solve practical problems

Course Content and CILOs Mapping:

Content		**CILO No.**
I	Background mathematics	1,3
II	Overview of big data analytics	1,2,3,4
III	Technologies and tools for big data analytics	1,2,3,4,5
IV	Theory and methods for big data analytics	3,4,6
V	Big data analytics applications	1,2,4,5,6
VI	Case studies: big data in banking	2,4,5,6

References:

- Jure Leskovec, Anand Rajaraman and Jeffrey David Ullman, Mining of Massive Datasets, 3rd Edition, Cambridge University Press, 2020.
- Vladimir Shikhman and David MÜller, Mathematical Foundations of Big Data Analytics, Springer Gabler, 1st Edition, 2021.
- Ron Bekkerman, Scaling up Machine Learning: Parallel and Distributed Approaches, Cambridge University Press, 2012.
- Tom White, Hadoop: The Definitive Guide, O′ Reilly Media, 4th Edition, 2015.
- Bill Franks, Taming The Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics, Wiley, 2012.
- Michael Minelli, Michele Chambers, and Ambiga Dhiraj, Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, Wiley, 2013.
- Frank J. Ohlhorst, Big Data Analytics: Turning Big Data into Big Money, Wiley an d SAS Business Series, 2013.
- Arvind Sathi, Big Data Analytics: Disruptive Technologies for Changing the Game, MC Press, 2013.

Course Content:

Topic

I. Background mathematics

II. Overview of big data analytics

III. Technologies and tools for big data analytics
   A. Introduction to MapReduce/Hadoop
   B. Data analytics using MapReduce/Hadoop
   C. Data visualization techniques

IV. Theory and methods for big data analytics
   A. Selected machine learning and data mining methods for big data (such as dimensionality reduction and locality sensitive hashing)
   B. Statistical analysis techniques (such as conjoint analysis and correlation analysis)
   C. Search strategies in AI

V. Big data analytics applications
   A. PageRank and spam analysis
   B. Social network analysis
   C. Recommendation systems

VI. Case studies: big data in banking
   A. Fraud analytics
   B. Customer segmentation