

Subject Code	COMP6434
Subject Title	Big Data Analytics and Artificial Intelligence
Credit Value	3
Level	6
Normal Duration	1 semester
Pre-requisite/ Co-requisite/ Exclusion	Nil
Role and Purposes	<p>The objectives of this subject are for students to:</p> <ol style="list-style-type: none"> 1. learn about the concepts and principles of mobile computing; 2. explore both theoretical and practical issues of mobile computing; 3. introduce students the concept and challenge of big data (3 V's) 4. apply skills and tools to manage and analyze big data in the fintech context. 5. understand the fundamental concepts of artificial intelligence; 6. understand the contemporary techniques in machine learning; 7. appreciate the effectiveness of hybridization of different artificial intelligence techniques. <p>This subject will contribute to the achievement of the DFintech program outcomes by</p> <ul style="list-style-type: none"> • allowing students to acquire the ability to conduct original applied research in tech-related business areas. (Outcome 3)
Subject Learning Outcomes	<p>Upon completion of the subject, students will be able to:</p> <ol style="list-style-type: none"> a. grasp the concepts and features of mobile computing technologies and applications; b. have a good understanding of how the underlying wireless and mobile communication networks work, their technical features, and what kinds of applications they can support; c. understand the concept and challenge of big data; d. collect, manage, store, query, and analyze various form of big data; e. understand the impact of big data for business decisions and strategy. f. understand the history, development and various applications of artificial intelligence; g. appreciate how uncertainty is being tackled in the knowledge representation and reasoning process; h. have some basic understanding on the skills and techniques in machine learning, such as decision tree induction, artificial neural networks, and genetic algorithm.

<p>Subject Synopsis/ Indicative Syllabus</p>	<p>Topic 1. Overview of Mobile Computing</p> <ul style="list-style-type: none"> • Motivations, concepts, challenges, and applications of mobile computing; • Relationship with distributed computing, Internet computing, ubiquitous/pervasive computing; • Mobile computing models and architectures. <p>Topic 2. Introduction to and collection of Big Data</p> <ul style="list-style-type: none"> • The 3 V's, their challenges and application domains. • Eventual Consistency and NoSQL systems MongoDB, Google BigTable • Data Visualization: Data types and dimensions; Visual encoding and perception <p>Topic 3. Large-Scale Data Analytics Systems</p> <ul style="list-style-type: none"> • Auto-Parallel Data Programming; MapReduce, Hive, and Parallel Databases • Sentiment Analysis <p>Topic 4. Machine Learning Systems for Big Data</p> <p>Topic 5. Artificial Intelligence (AI), its roots and scope</p> <ul style="list-style-type: none"> • Overview of AI application areas; • Expert systems, natural language understanding and semantics, planning and robotics, and machine learning. <p>Topic 6. Artificial Intelligence as representation and search</p> <ul style="list-style-type: none"> • The Propositional Calculus and Predicate Calculus; using inference rules to produce predicate calculus expressions; • Strategies and structures for state space search; heuristic search; recursion-based search; admissibility, monotonicity and informed-ness of search algorithms. <p>Topic 7. Knowledge representation and reasoning</p> <ul style="list-style-type: none"> • Rule-based production systems; case-based reasoning systems and model based reasoning systems; • Reasoning under uncertain situations: stochastic methods and fuzzy expert systems. <p>Topic 8. Machine Learning</p> <ul style="list-style-type: none"> • Decision tree induction algorithms; • Artificial neural networks; • Genetic algorithms.
<p>Teaching/Learning Methodology</p>	<p>The course will be offered in a mode that combines seminars, case study, team presentations, and group discussions.</p>

Assessment Methods in Alignment with Intended Learning Outcomes <i>(Note 4)</i>	Specific assessment methods/tasks	% weighting	Intended subject learning outcomes to be assessed (Please tick as appropriate)							
			a.	b.	c.	d.	e.	f.	g.	h.
	Continuous Assessment*	100%								
	1. Class participation	20 %	√	√	√	√	√	√	√	√
	2. Group project & presentation	20 %	√	√	√	√	√	√	√	√
	3. Lab exercise/Assignment	20%	√	√	√	√	√	√	√	√
	4. Individual Assessment (e.g. Test)	40%	√	√	√	√	√	√	√	√
	Total	100 %								
<p><i>*Weighting of assessment methods/tasks in continuous assessment may be different, subject to each subject lecturer.</i></p> <p>To pass this subject, students are required to obtain Grade D or above in the Continuous Assessment components.</p> <p>Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes: the various methods are designed to ensure that all students taking this subject –</p> <ol style="list-style-type: none"> 1. Class participation aims to stimulate students to be exposed to various new potential applications of the smart city and urban informatics technologies in different business areas. 2. Lab exercise is designed to encourage students to understand the relevant knowledge and to practice, in order to enrich their hands-on experience with various software tools. 3. The group project is designed to enhance students’ ability to acquire the understanding and using different knowledge, principles, techniques, tools to solve a real problem through team. 4. Individual assessment is used to assess individual students’ ability to have an overall understanding of the inter-relationship of the various technologies and their individual characteristics. 										
Student Study Effort Expected	Class contact:									
	▪ Lectures		25 Hrs.							
	▪ Laboratories		5 Hrs.							

	Other student study effort:	
	<ul style="list-style-type: none"> ▪ Preparation for the class 	30 Hrs.
	<ul style="list-style-type: none"> ▪ Preparation for Projects/Assignments 	50 Hrs.
	<ul style="list-style-type: none"> ▪ Preparation for Individual assessment 	10 Hrs.
	Total student study effort	120 Hrs.
Reading List and References	<p><u>Mobile Computing</u></p> <ol style="list-style-type: none"> 1. Martyn Mallick, “Mobile and Wireless Design Essentials”, Wiley Publishing, 2003. 2. Yu-Kwong Ricky Kwok, Vincent K.N. Lau, “Wireless Internet and Mobile Computing: Interoperability and Performance”, Wiley-IEEE Press, 2007. (Google Book) 3. D.P. Agrawal and Q.-A. Zeng, “Introduction to Wireless and Mobile Systems”, 2nd edition, Thomson Learning, 2006. 4. Reza B’Far, “Mobile Computing Principles: Designing and Developing Mobile Applications with UML and XML”, Cambridge University Press, 2005. 5. J. Schiller, “Mobile Communications”, 2nd edition, Pearson Education, 2003. <p><u>Big Data Analytics</u></p> <ol style="list-style-type: none"> 1. How Vertica Was the Star of the Obama Campaign, and Other Revelations 2. Cohen et al. “MAD Skills: New Analysis Practices for Big Data”, 2009 3. Ullman, Rajaraman, Mining of Massive Datasets, Chapter 2 4. Stonebraker et al., “MapReduce and Parallel DBMS’s: Friends or Foes?”, Communications of the ACM, January 2010. 5. Dean and Ghemawat, “MapReduce: A Flexible Data Processing Tool”, Communications of the ACM, January 2010. 6. Rick Cattell, “Scalable SQL and NoSQL Data Stores”, SIGMOD Record, December 2010 (39:4) 7. Elmagarmid, et. al. "Duplicate Record Detection: A Survey" 8. Koudas, et. al. "Record Linkage: Similarity Measures and Algorithms" 9. Chapter 3 of A Handbook of Statistical Analyses Using R 10. Gregory Park on overfitting to the leaderboard in a Kaggle Competition 11. Xindong Wu et al., Top 10 Algorithms in Data Mining, Knowledge and Information Systems, 14(2008), 1: 1-37. (read C4.5) 12. Ullman, Rajaraman, Mining of Massive Datasets , Chapter 1 13. Pedro Domingos, A Few Useful Things to Know about Machine Learning, CACM 55(10), 2012 14. Pat Hanaran, Tools for Data Enthusiasts 15. Jeffrey Heer, Michael Bostock, Vadim Ogievetsky, A Tour through the Visualization Zoo, Communications of the ACM, Volume 53 Issue 6, June 2010 16. Howard Wen, "Big Ethics for Big Data", O'Reilly Media <p><u>Artificial Intelligence</u></p> <ol style="list-style-type: none"> 1. George F. Luger, Artificial Intelligence: Structures and Strategies for Complex Problem Solving, 6th Edition, Addison Wesley, 2009. 	

Reference Books:

1. Sankar K. Pal and Simon C. K. Shiu, Foundations of Soft Case-Based Reasoning, John Wiley, 2004.
2. Michael Negnevitsky, Artificial Intelligence: A Guide to Intelligent Systems. 2nd edition, Addison Wesley, 2005.