City University of Hong Kong
Course Syllabus

offered by Department of Information Systems
with effect from Semester A 2021 / 2022

Part I   Course Overview

Course Title:	Big Data & Social Media Analytics
Course Code:	IS6941
Course Duration:	One Semester (13 weeks)
Credit Units:	3
Level:	P6
Medium of Instruction:	English
Medium of Assessment:	English
Prerequisites: *(Course Code and Title)*	Nil
Precursors: *(Course Code and Title)*	Nil
Equivalent Courses: *(Course Code and Title)*	Nil
Exclusive Courses: *(Course Code and Title)*	Nil

Part II   Course Details

1.   Abstract

This course aims to develop students' knowledge and skills to apply big data and social media analytics tools and techniques to enhance various business applications.

2.   Course Intended Learning Outcomes (CILOs)
(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs	Weighting (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	Describe the basic concepts and techniques for big data management and online social media analytics.	20%	✓	✓	
2.	Design and apply big data and social media analytics tools and techniques to enhance various business applications.	30%	✓	✓	✓
3.	Critically evaluate various big data and social media analytics technologies with respect to realistic business contexts.	30%		✓	✓
4.	Work productively as part of a team, and in particular, communicate and present big data and social media analytics solutions effectively in written and oral forms under a collaborative work environment.	20%	✓	✓	
		100%			

A1:   Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2:   Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to self-life problems.

A3:   Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Teaching and Learning Activities (TLAs)
(TLAs designed to facilitate students' achievement of the CILOs.)

TLA	Brief Description	CILO No.				Hours/week (if applicable)
		1	**2**	**3**	**4**	
TLA1: Lecture	Concepts and general knowledge of big data and social media analytics tools, techniques, and their business applications are illustrated. • *In-class discussion*: Students actively participate in discussions in lectures (e.g. face-to-face discussion, using mobile devices) and the lecturer provides feedback based on students' responses. • *Active Reflection*: At the beginning of every lecture, students are expected to actively reflect and summarize the main topics covered in the previous lecture and lecturer will provide feedback based on students' concerns and questions.	✓	✓	✓		
TLA2: Tutorial, Laboratory and Case studies	The tutorial and laboratory sessions cover the technical aspects of various big data and social media analytics tools and techniques, and their real-world applications. • *Tutorial exercises*: e.g. hands-on activities of utilizing big data management platforms such as Hadoop and Spark; hands-on activities of applying large-scale social network analysis software such as Pajek for social media analytics; hands-on activities of applying various Python-based APIs for online social media data management and analytics, etc. • *Case/Group project discussion*: Students will be given some case studies and a project relating to big data or social media analytics application under a realistic business context. Discussion on various aspects of a case or project for applying big data and social media analytics tools to improve the productivity and efficacy of a business application will be conducted.	✓	✓	✓	✓	
TLA3: On-Line Active Readings and Discussions	Students are encouraged to conduct self-reflection and sharing of concepts, techniques, and methods related to big data and social medial analytics methodologies, and their applications to real-world business contexts.	✓	✓		✓	

4. Assessment Tasks/Activities (ATs)
(ATs are designed to assess how well the students achieve the CILOs.)

Indicative of likely activities and tasks students will undertake to learn in this course. Final details will be provided to students in their first week of attendance in this course.

Assessment Tasks/Activities	CILO No.				Weighting	Remarks
	1	2	3	4		
Continuous Assessment: 60%						
AT1: Individual Tutorial Exercises Students need to complete and submit individual tutorial exercises that cover both conceptual exploration and hands-on skills of big data and social media analytics tools and techniques.	✓	✓			10%	
AT2: Team Project There will be a team project associated with the course. The assessment of the team project is based on the understanding of the concepts and techniques related to big data and social media analytics discussed in lectures, and the creativities of applying these tools and techniques to solve a real-world business problem. The project requires a project proposal, prototype system demonstration, and presentation, as well as a final report.	✓	✓	✓	✓	40%	
AT3: Individual Online Reflection and Discussion Exercises Students need to complete and submit individual self-reflection and collaborative discussion reports related to case studies of applying big data and social media analytics solutions to enhance various business applications.			✓	✓	10%	
Examination: 40% (duration: one 2-hour exam)						
AT4: Examination The final examination of 2 hours will be used to assess students' achievement on intended learning outcomes.	✓		✓		40%	
					100%	

Note: Students must pass BOTH coursework and examination in order to get an overall pass in this course.

5. Assessment Rubrics
(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
AT1: Individual Tutorial Exercises	Capability to describe the basic concepts and techniques for big data management and online social media analytics	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to design and apply big data and social media analytics tools and techniques to enhance various business applications.	High	Significant	Moderate	Basic	Not even reaching marginal levels

AT2: Team Project	Ability to describe the basic concepts and techniques for big data management and online social media analytics.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to design and apply big data and social media analytics tools and techniques to enhance various business applications.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to critically evaluate various big data and social media analytics technologies with respect to realistic business context.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to work productively as part of a team, and in particular, communicate and present big data and social media analytics solutions effectively in written and oral forms under a collaborative work environment.	High	Significant	Moderate	Basic	Not even reaching marginal levels
AT3: Individual Online Reflection and Discussion Exercises	Ability to critically evaluate various big data and social media analytics technologies with respect to realistic business context.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to work productively as part of a team, and in particular, communicate and present big data and social media analytics solutions effectively in written and oral forms under a collaborative work environment.	High	Significant	Moderate	Basic	Not even reaching marginal levels
AT4: Final Exam	Ability to describe the basic concepts and techniques for big data management and online social media analytics.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to design and apply big data and social media analytics tools and techniques to enhance various business applications.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to critically evaluate various big data and social media analytics technologies with respect to realistic business context.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	Ability to work productively as part of a team, and in particular, communicate and present big data and social media analytics solutions effectively in written and oral forms under a collaborative work	High	Significant	Moderate	Basic	Not even reaching marginal levels

	environment.					

Part III   Other Information (more details can be provided separately in the teaching plan)

1.   **Keyword Syllabus**
(An indication of the key topics of the course.)

- Online social media services (e.g. YouTube, Twitter, Facebook, GitHub, Pinterest, etc.) and their APIs.

- Social media analytics tools. Analytics development tools such as Python and social network analytics tools such as Pajek.

- Scraping and extracting conversational topics on Internet forums using Python.

- Analyzing social media contents (e.g. tweets) using sentiment analysis and entity recognition tools.

- Machine learning techniques for social media analytics such as Naïve Bayes classification for social link prediction, the K-Nearest Neighbour (KNN) method for social sentiment analysis, etc.

- Big data analytics platforms such as Apache Hadoop and Apache Spark.   The architecture of Hadoop versus that of Spark.   Hadoop Distributed File System (HDFS) versus Resilient Distributed Datasets (RDD).   The MapReduce model, constructing mapper and reducer using Scala.

- NoSQL and distributed SQL query engine.

- Parallel sentiment analysis with Hadoop and Spark.

2.   **Reading List**
2.1   **Compulsory Readings**
(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Chatterjee, Siddhartha and Krystyanczuk, Michal, Python Social Media Analytics: Analyze and visualize data from Twitter, YouTube, GitHub, and more (1st edition), Packt Publishing, 2017, ISBN: 978-1-78712-148-5.
2.	Nataraj Dasgupta. Practical Big Data Analytics: Hands-on techniques to implement enterprise analytics and machine learning using Hadoop, Spark, NoSQL and R (1st edition), Parkt Publishing, 2018, ISBN: 978-1783554393.

2.2   **Additional Readings**
(Additional references for students to learn to expand their knowledge about the subject.)

1.	Mohammed Guller. Big Data Analytics with Spark: A Practitioner's Guide to Using Spark for Large Scale Data Analysis (1st edition), Apress, 2015. ISBN-13: 978-1484209653.
2.	Russell Jurney. Agile Data Science 2.0: Building Full-Stack Data Analytics Applications with Spark (1st edition), O'Reilly, 2017. ISBN-13: 978-1491960110.
3.	Ganis, M. and Kohirkar, A., Social Media Analytics: Techniques and Insights for Extracting Business Value Out of Social Media, IBM Press, (1st edition, December, 2015), ISBN-13: 978-

	0133892567.
4.	Bali, R., Sarkar, D., Sharma, T., Learning Social Media Analytics with R: Transform data from social media platforms into actionable business insights (May 26, 2017), ISBN-13: 978-1787127524.
5.	Jeffrey Aven. Data analytics with spark using Python (1st edition), Addison-Wesley Professional, 2018, ISBN-13: 978-0134846019.

2.3 Online Resources

Course reading materials will be augmented by articles from journals and by whitepapers and other materials available on-line.   Below are some examples:

Lau, Raymond Y. K., Zhang, W. & Xu, W., Oct 2018, Parallel Aspect Oriented Sentiment Analysis for Sales Forecasting with Big Data, *Production and Operations Management* (27:10), pp. 1775-1794.

Lau, Raymond Y. K., Zhao, J. L., Chen, G. & Guo, X., Dec 2016, Big Data Commerce, *Information and Management* (53:8), pp. 929-933.

Lau, Raymond Y. K., Leon Zhao, J., Zhang, W., Cai, Y. & Ngai, E. W., Jun 2015, Learning context-sensitive domain ontologies from folksonomies: A cognitively motivated, method, *INFORMS Journal on Computing* (27:3), pp. 561-578.

Du, X., Ye, Y., Lau, R. Y. K. & Li, Y., Apr 2015, OpinionRings: Inferring and visualizing the opinion tendency of socially connected users, *Decision Support Systems* (75), pp. 11-34.

Fan, S., Lau, R. Y. K. & Zhao, J. L., Mar 2015, Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix, *Big Data Research* (2:1), pp. 28-32.

- Updated SYL template in July 2017.
- Revised course title and SYL content in January 2020.
- Changed course level from P5 to P6 and the course code in June 2021.