Republic of the Philippines
BATANGAS STATE UNIVERSITY
Pablo Borbon Main II, Batangas City, Philippines 4200
COLLEGE OF INFORMATICS AND COMPUTING SCIENCES
Tel. No. 425-0139 local 2223 Email: batstateu.cics@gmail.com

CURRICULUM
Master of Science in Data Science (MSDS)

Academic Year 2019-2020
Reference CMOs: CMO No. 07, s. 2010

Curriculum Description

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms, and systems to extract value from data. Data scientists combine a range of skills—including statistics, computer science, and business knowledge—to analyze data collected from the web, smartphones, customers, sensors, and other sources. In which, this program focuses on creating or innovating new solutions to real-world challenges through research. The course follows the traditional face-to-face, with an added blended learning method through online interactions.

Program Objectives

The alumni of Master of Science in Data Science, about three to five years after graduation shall:

1. Practiced as high-level specialist in solving complex data science problems leading to improvements and innovations, while taking into consideration the environmental, social, and economical requirements.
2. Assumed high level leadership position in industry, academe, government, or private sector with consideration to social and ethical responsibility.
3. Engaged in lifelong learning through further studies, research, certifications, promotions, and other personal and professional development activities.

Program Outcomes

1. Demonstrate a comprehensive and broad understanding of data science principles and apply advanced knowledge in the specific discipline;
2. Analyze, synthesize, create and evaluate the challenges in data science practice;
3. Design components, devices, and systems to meet specified engineering needs under real–world constraints;
4. Communicate effectively the technical knowledge, both orally and in writing, on complex data science activities;
5. Function effectively as an individual, a team member, or as a leader in diverse work environments;
6. Contribute to the generation, dissemination and preservation of knowledge, methodologies, techniques, and processes;
7. Engage in continuous professional development and lifelong learning endeavors;
8. Conduct oneself within professional and ethical standards; and
9. Perform independent scientific research that results in innovation with application.

Curriculum Components

	Courses	Units	Total
	A. Core Courses		**9 units**
	Fundamentals of Data Science	3	
	Big Data and Cloud Computing	3	
	Data Visualization and Storytelling	3	
	B. Specialization Courses		**9 units**
	Machine Learning and Neural Networks	3	
	Mathematical and Computational Theories for Data Science	3	
	Seminars in Data Science	3	
	C. Elective Courses		**6 units**
	Industry Internship	3	
	Data Privacy and Ethical Practice for Data Science	3	
	Biomedical Modeling for Health Data	3	
	Health Informatics	3	
	Business Economics and Strategies	3	
	Time-Series Data Analysis and Forescasting for Business and Finace	3	
	Computation and Complexity	3	
	High Performance Computing	3	
	Advanced Statistical Models and Methods for Data Science	3	
	Cross-Platform Applications Design and Development	3	
	Special Topics in Data Science 1	3	
	Special Topics in Data Science 2	3	
	D. Thesis Courses		**6 units**
	Thesis 1	3	
	Thesis 2	3	

SUMMARY	
Courses	**Number of Units**
Core Courses	9
Specialization Courses	9
Elective Courses	6
Capstone Project	6
TOTAL	30

PROGRAM OF STUDY

FIRST YEAR					
FIRST SEMESTER					
Code	**Course Title**	**Units**	**Lec**	**Lab**	**Prerequisite**
MSDS 500	Fundamentals of Data Science	3	3	-	-
MSDS 501	Big Data and Cloud Computing	3	3	-	-
MSDS 502	Data Visualization and Storytelling	3	3	-	-
	TOTAL	**9**	**9**	**-**	
SECOND SEMESTER					
Code	**Course Title**	**Units**	**Lec**	**Lab**	**Prerequisite**
MSDS 503	Machine Learning and Neural Networks	3	3	-	-
MSDS 504	Mathematical and Computational Theories for Data Science	3	3	-	-
	Elective Course 1	3	3	-	-
	TOTAL	**9**	**9**	**-**	
MIDTERM					
MSDS 505	Seminars in Data Science	3	3	-	-
	Elective Course 2	3	3	-	-
	Total	**6**	**6**		
SECOND YEAR					
FIRST SEMESTER					
Code	**Course Title**	**Units**	**Lec**	**Lab**	**Prerequisite**
MSDS 520	Thesis 1	3	3	-	-
	***Comprehensive Examination*				
	TOTAL	**6**	**6**	**-**	
SECOND SEMESTER					
Code	**Course Title**	**Units**	**Lec**	**Lab**	**Prerequisite**
MSDS 521	Thesis 2	3	3	-	-
	TOTAL	**3**	**3**	**-**	

COURSE DESCRIPTION

Core Courses	
MSDS 500	**Fundamentals of Data Science**
	This course provides the initial knowledge to develop a strong foundation in Data Science for both professional and research works. The course covers overall program offered, the history, tools, methods, and current trends and technologies across diverse fields of disciplines aligned for the course. The introductory course would also extend its contents to inspire the future graduates to pursue greater heights in their careers after graduation.
MSDS 501	**Big Data and Cloud Computing**
	The course enables the student to systematically collect, process, and analyze large and complex real-world structured and unstructured data. The course also includes data mining, warehousing, and business intelligence.
MSDS 502	**Data Visualization and Storytelling**
	Using state-of-the-art technology, the students can learn to visualize large-scale data and initiate appropriate analysis to generate new knowledge. The course helps the student to develop easy and reliable visualization of large-scale data in a much simpler format for a wide variety of audience. The visualized data is then analyzed using proper methods. Also, students are encouraged to formulate new methods through guided research. Furthermore, the program does not limit students to just analyze or apply the usual data science methods. This course also provides both theoretical and empirical approach to extend their capabilities to deliver and discuss new knowledge to a variety of large audiences to yield greater outcomes from pre-existing discoveries.
Specialization Courses	
MSDS 503	Machine Learning and Neural Networks
	This course covers the principles and importance of machine learning and neural networks in solving real-world problems. Students will also learn to build and train their own effective neural network models, predict future outcomes from data patterns, and even improve pre-existing algorithms in the field of machine learning and neural networks. However, this course is not limited to the given description and may be extended to other forms of practice.
MSDS 504	Mathematical and Computational Theories for Data Science
	The course tackles the essential mathematical and computational theories of data science. Using advanced tools and technology, the course also aims to demystify the field of mathematics in real-life applications of data science. Furthermore, the instructor shall guide the student to formulate a new set of theories to improve the overall body of knowledge in the field of Data Science.
MSDS 505	Seminars in Data Science
	Seminars are small, discussion-based activities that feature intense study of one topic. Seminars help students reach beyond the basics of a topic and get at the deeper implications and relationships involved in the area of study. Seminars allow students to learn from one another and expose them to new ideas and to the thoughts of people with different backgrounds. To some extent in a seminar course, students do assigned reading and then, under the facilitator guidance and direction, grapple aloud with the ideas they've read. They learn to form arguments and support them with facts; they learn to communicate coherently and courteously with those who disagree. This course helps to ensure that students acquire the skills needed to become successful professionals in their fields.
Thesis	
MSDS 520	Thesis 1
	Students are required to conduct research on topics approved by the Academic Panel for the Master of Science in Data Science Program. In the research proposal stage, students are required to identify important issues that involves Data Science. Based on the identified issues, students will propose the problems to be researched for the MS study. They are required to explain how their research findings will give impacts and reform the aspects of Data Science and program curriculum either formally or informally.
MSDS 521	Thesis 2
	Students are tasked to present their research output based on the topics approved by the Academic Panel for the Master of Science in Data Science. Also, the accomplished research is required to be presented in a peer-reviewed conference or published on local or international journals.

Elective Courses	
MSDS 506	**Industry Internship**
	This course prepares the student to further elevate their acquired knowledge and apply it in a realistic working environment. The students are given the chance to choose from a wide-variety of existing and future industry partners both local and international that can nourish them with the essential skills and knowledge through rigorous practice. The immersion program guarantees maximum satisfaction for students to experience actual practice rather than just the typical classroom setup.
MSDS 507	**Data Privacy and Ethical Practice for Data Science**
	Together with the advancing technology and demand for data, several questions arise. This course aims to answer and mitigate future problems related to privacy and improper handling of data. Also, the course aims to generate new knowledge through research to improve the ethical adaptability and acceptance of emerging technologies for modern societies through data science.
MSDS 508	**Biomedical Modeling for Health Data**
	This course aims to equip students with fundamental concepts and skills relevant to biomedical data science, with an emphasis on bioinformatics, a sub-discipline of this broader field, through examples of mining and modeling of genomic and proteomic data. More specifically, bioinformatics encompasses the analysis of gene sequences, macromolecular structures, and functional genomics data on a large scale. It represents a major practical application for modern techniques in data mining and simulation. Specific topics to be covered include sequence alignment, large-scale processing, next-generation sequencing data, comparative genomics, phylogenetics, biological database design, geometric analysis of protein structure, molecular-dynamics simulation, biological networks, mining of functional genomics data sets, and machine learning approaches for data integration.
MSDS 509	**Health Informatics**
	This course provides the essential tools and concepts to analyze and interpret data related to health. The course also pursues the application of data science to the fields of genomics, microbiology, biotechnology, and biochemistry, including software and research methodologies.
MSDS 510	**Business Economics and Strategies**
	Provides real-life strategies for improving the business sector through the use of economics and IT. This is a research based course provides the students to come up with their potential solutions and test it in an actual scenario.
MSDS 511	**Time-Series Data Analysis and Forescasting for Business and Finace**
	The course introduces the fundamental to advance concepts to further analyze time-series data in the time domain and to provide the students with experience in analysing cases with relation to business and finance.
MSDS 512	**Computation and Complexity**
	This course is an introduction to the theory of computational complexity and standard complexity classes. One of the most important insights to have emerged from Theoretical Computer Science is that computational problems can be classified according to how difficult they are to solve. This classification has shown that many computational problems are impossible to solve, and many more are impractical to solve in a reasonable amount of time. To classify problems in this way, one needs a rigorous model of computation, and a means of comparing problems of different kinds. This course introduces these ideas, and shows how they can be used.
MSDS 513	**High Performance Computing**
	HPC is the use of powerful processors, networks and parallel supercomputers to tackle problems that are very computationally or data-intensive. The goal of this course is to provide students the solid foundations for developing, analyzing, and implementing parallel and locality-efficient algorithms and other theoretical underpinnings. To give a practical feeling for how algorithms map to and behave on real systems, we will supplement algorithmic theory with hands-on

	exercises on modern HPC systems. Data science involves the manipulation, processing and analysis of data to extract knowledge, and HPC provides the power that underpins it.
MSDS 514	**Advanced Statistical Models and Methods for Data Science**
	This course further examines statistics and data analysis beyond an introductory course. Topics include t-tools and permutation-based alternatives including bootstrapping, multiple-group comparisons, analysis of variance, linear regression, model checking, and refinement. Statistical computing and simulation-based emphasis is covered as well as basic programming in the R statistical package. Thinking statistically, evaluating assumptions, and developing tools for real-life applications are emphasized. This course is heavily focused more on statical modeling and other related methods.
MSDS 515	**Cross-Platform Applications Design and Development**
	A professional course that delivers the design and development of industry standard applications using a wide variety of technology stacks. Using software engineering methodologies and technology awareness, students can build and develop their own infrastructure design together with a working application that can solve IT-based solutions in an industrial setting.