Master of Science in Data Science (MSc DS)

PostGraduate Program

FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY

Prepared By: Graduate Committee
1440-1441 2019-2020

Introduction

The Faculty of Computers and Information Technology at the University of Tabuk proposes a new Master of Science program in Data Science. Bearing in mind the rapid growth of this filed, and the lack of such programs in the region, the program is one of the first programs in in this field in Saudi Arabia. It is designed to provide students with a comprehensive knowledge in Data Science and to develop their thought of the key and innovative technologies in data science related fields, e.g., machine learning, data mining, statistical modelling and text analytics. Hence the program prepares students for leadership positions in data science in the region and make them capable to compete for jobs in high-demand in the field. The program adopts a project-based method that helps the students to apply the concepts learned in the core courses into a project during their last two semesters.

To meet the declared objectives and the growing demands of highly qualified individuals in data analytics and information management, the students are going to be involved into building deep knowledge of the field through a set and well described courses in the field. In addition, they are going to apply recent techniques for analysing big data across different disciplines of data science through programming techniques and data tools through course work. An AI lab plus our new Centre of Industrial Innovative and Robotics, with all its valuable equipment, will be available for our Master program in Data Science.

The program curriculum will cover a variety of topics that provide students with a solid understanding to state-of-the-art in field. Examples include, but are not limited to, Data Science and their applications, Machine Learning, Deep Learning, Text Analytics, Data Visualization, etc. The curriculum also contains a project that enables students to apply core course materials to a practical project in an actual environment.

Vision of the Faulty of Computing and Information Technology

A distinguished faculty and pioneer in education and research contributing to the service of the community in the field of computers and information technology locally and regionally

Mission of the Department of Information Systems

To prepare a qualified scientific cadres in the various disciplines of information systems through innovative education, which builds an analytical abilities and skills that can serve their society.
.

Master of Science in Data Science program

Degree Title

Master of Science in Data Science.

Mission

Given the growing needs in industrial, the mission of the Master of Science in Data Science program is to graduate students with a solid knowledge in data science and preparing them to compete successfully for jobs in high-demand A.I. industry.

Educational Objectives

The Objectives of the Master of Science in Data Science are:

- o demonstrate knowledge of statistical data analysis techniques utilized in decision making.
- o apply principles of Data Science to the analysis of business problems.
- o prepare students to understand and apply data science techniques and algorithms in achieving organizations missions
- o to be able to solve problems involving large, diverse data sets from different application domains
- o develop programming skills in data analytics, visualization, learning and mining.
- o to predestine substantially trained specialists in data science for the Saudi industry.
- o Prepare students for career advancement in all areas of information science and technology

Learning Outcomes

At the end of the program, graduates should be able to:

1. develop an effective Big Data solution in a real environment
2. apply computing theory, and algorithms, as well as mathematical and statistical models, to appropriately formulate and use data analyses.
3. carry out a cost-benefit analysis
4. use appropriate models of analysis, from results, and investigate potential issues
5. obtain a clean data process and transformation data

Program Design

The Master program of Science in Data Science at the University of Tabuk is a course-based program along with a project to be submitted at the end of the program during two semesters.
The students are required to complete a minimum of 47 credit hours by both courses and a project.

Detailed Program Design

With the approval of a supervising professor, qualified students may be admitted to the program. The students of the Master of Science in Data Science must complete 12 courses, 2 research projects (Research Project 1 and Research Project 2) and an internship in one of well-known organizations
In particular, students must complete minimum 47 credit hours, including:
1. At least 39 course credits that include:
    o 36 course credits of core courses.
    o 3 elective course credit.
2. 8 Project credits: Research Project 1 (4 credits) and Research Project 2 (4 credits).

During the first and second semester of the program time the students will focus on courses. From the third semester students will focus on their research projects.

Admission Requirements

Applicants must have an earned bachelor's degree in the field of information technology, engineering, computer science or engineering fields (i.e. Electrical or Electronics Engineering), mathematics, statistics and information systems from a regional and/or national college (or its equivalent from a recognized institution) prior to enrolment at the program. In addition, the following should be satisfied for a complete application

- Minimum very good academic standing.

- Successful completion of at least one course in advanced programming (e.g., Java/C/C++/C#) Object Oriented, Databases, and mathematics (e.g., linear algebra, calculus, etc)

Additional requirements:

- o A portfolio indicating relevant work experience and skills in Computer Science and information technology.

- o Work experience in IT-related jobs is preferred.

Program Prerequisites

Applicants of the program are required to have background knowledge in areas such as programming, mathematics, and computer systems. The Graduate Studies Committee will determine which, if any, of the following prerequisite courses will need to be taken.

Applicants, who fulfill some of below listed prerequisites while lacking others, may undertake them at the undergraduate level, but prior to program begin. Applicants are required in this case to complete these prerequisites with a grade of "Good" or better (from the first time of enrollment) and a GPA not less than "Good" and the final decision about admission to program under these terms are at the sole discretion of the Graduate Studies committee and the department council. Prerequisite courses do not count toward the required credits for the Master's degree.

Accordingly, all Applicants must satisfy the following **"breadth"** requirements demonstrating proficiency at the undergraduate level in the "breadth" requirements which include:

- Programming Languages.

- Theoretical Foundations of Computer Science and Information Technology.

- Mathematical Foundations

Each breadth contains the followings:

1. Programming Languages

    - Programming Languages: e.g. Java, C++, C# and advanced topics in programming e.g. object oriented.

2. Theoretical Foundations of Computer Science and Information Technology.

    - Data Structures and Databases.

3. Mathematical Foundations

    - Calculus
    - Linear Algebra
    - Statistics and Probability Theory

An applicant for admission should minimally have a bachelor's or equivalent degree in computer science, information technology, computer information systems, Computer engineering, information systems,

mathematics, statistics and engineering from a recognized college or university, must have the necessary academic preparation to complete graduate level courses, and must be able to score competitively on the on the TOEFL / IELTS test. Condition admission is not offered and a transitional or pre-master's program is not available.

Besides, the department of Computer Science may grant provisional approval to candidates with a B.Sc. in a different area with the requirement that they succeed in B.Sc. level courses of the required prerequisites with a minimum grade of "Good" from the first time.

In addition to the conditions enumerated in the 13ththrough 18th articles(in particular) and all articles (in general) of *the unified law organizing the graduate studies in Saudi universities*, and those also enumerated in *the Regulations and operating procedures for the Postgraduate Study booklet (University of Tabuk)*, applicant has to fit and/or submit the following criteria:

1. **Required documents to be submitted**:

   - Official transcript and graduation statement from each institution attended. Official transcripts are defined as original documents issued by the institution that bear the actual signature of the Registrar.
   - Two (2) letters of recommendation. The letters should be from academic staff members in recognized universities and from those who have taught the student in his/her junior and senior years.
   - General Aptitude Test (GAT) for Bachelor's degree holders (KSA) with a minimum score of 55.
   - A recent resume or vitae.
   - Personal Statement (Application Letter).
   - A formal approval from applicant's employer (for those who are employed).
   - Any additional departmental or university admission requirements

2. **Grade Point Requirement:**

   The university requires at least a scholastic average of "Very Good" in the bachelor program.

3. **English Proficiency:**

   Proficiency in English is an admission requirement for the program. Proficiency tests must be taken within the last two (2) years of the date of the application submission.

   Accepted tests:
   - Test of English as a Foreign Language (TOEFL) with a minimum score of 450 (PBT), 160 (CBT), 55 (IBT).
   - International English Language Testing System (IELTS) with a minimum score of 4.5.
   - Standardized Test for English Proficiency (STEP) with a minimum score of 75%.
   - English Proficiency Test (Level Four). The test is provided by the Deanship of the Community

Service and Continuing Education at the University of Tabuk).

4. **Admission Exam:**

All applicants must pass a written admission exam.

Degree Requirements

Courses

Students must satisfy the requirements listed in *Program Design* Section.

Each student must submit an approved project, based on appropriate research experience, in accordance with the regulations of the both the university and the department. The project must be produced under the direction and with the approval of the student's research supervisor, who must be a member of the faculty staff. The handbook(s) of the Deanship of Postgraduate Studies should be consulted for rules, procedures, and deadlines for preparation and submission of the final approved copy of the project.
Each candidate must also pass a project examination to evaluate the student knowledge and contribution of research in terms of a viva examination.

Academic Standards

Students must maintain accumulative Grade Point Average (GPA) not less than "Good" on the graduate degree courses.

Time Limit
The entire work for the Master's degree can be completed within a period of 2 years. But, all articles of the unified law organizing the graduate studies in Saudi universities, and those also enumerated in the Regulations and operating procedures for the Postgraduate Study booklet (University of Tabuk), control time limit for acquiring Master's degree.

Program Duration

2 years (4 Semesters): Full-time
Fees

18,000 SAR per semester.
72,000 SAR per program.

Master of Science in Data Science: Study Plan

MSc in DS Full Time –2 Years Plan

Semester One

المقابل اللغوى		Prerequisite	Credits		Total Credits	المقـــــــرر
Course Title	Code/No		Pr.	Th.		
Fundamentals of Data Science	CSC640	-	-	٣	٣	أساسيات علم البيانات
Computational Mathematics	CSC641	-	-	٣	٣	الرياضيات الحاسوبية
Programming for Data Science	CSC650	-	-	٣	٣	البرمجة لعلم البيانات
Advanced Topics in Databases	CIS680	-	-	٣	٣	مواضيع متقدمة في قواعد البيانات
			-	١٢	١٢	المجموع

Semester Two

المقابل اللغوى		Prerequisite	Credits		Total Credits	المقـــــــرر
Course Title	Code/No		Pr.	Th.		
Machine Learning	CSC603	CSC641	-	٣	٣	التعلم الآلي
Research Methods	CIS608	-	-	٣	٣	طرق بحث
Data Mining	CSC617	CSC641	-	٣	٣	تنقيب البيانات
Data Visualization	CIS651	CSC650	–	٣	٣	عرض البيانات
				١٢	١٢	المجموع

Semester Three

المقابل اللغوى		Prerequisite	Credits		Total Credits	المقــــــرر
Course Title	Code/No		Pr.	Th.		
Information Retrieval	CSC613	CSC641	-	٣	٣	أنظمة استرجاع المعلومات
Big Data Analytics	CIS641	CSC640	-	٣	٣	تحليل البيانات الضخمة
Seminar in Data Science	CIS670	Completion of 18 credit hours		3	3	سمنار في علم البيانات
Research Project 1	CIS624	Completion of 21 credit hours	–	٤	٤	مشروع بحثي ١
				13	١٣	المجموع

Semester Four

المقابل اللغوى		Prerequisite	Credits		Total Credits	المقــــــرر
Course Title	Code/No		Pr.	Th.		
Data Science Professional and Ethical issues	CIS690	-	–	٣	٣	قضايا علم البيانات المهنية والأخلاقية
Elective Course	-	Completion of 30 credit hours	-	٣	٣	مقرر اختياري
Research Project 2	CIS625	CIS624	–	٤	٤	مشروع بحثي ٢
				10	10	المجموع

Elective Courses

المقابل اللغوى		Prerequisite	Credits		Total Credits	المقــــــرر
Course Title	Code/No		Pr.	Th.		

Text Analytics	CIS684	Completion of 30 credit hours				
Neural Networks and Deep Learning	CSC604		–	٣	٣	الشبكات العصابية والتعلم العميق
Bioinformatics	CSC615		–	٣	٣	البيانات الحيوية
Parallel Computing	CSC618		–	٣	٣	الحوسبة المتوازية
Decision Support Systems	CIS682		–	٣	٣	أنظمة دعم القرار
Selected topics in Data Science	CIS622		-	٣	٣	مواضيع مختارة في علم البيانات

Appendix A: Statistics about the Faculty of Computers and Information Technology

Academic Year: 2019/2020   -   1440/1441 H

Faculty Members Statistics

PhD Holders (Professors – Associate Professors – Assistant Professors) Lecturers – Teaching Assistants	
Professors	1
Associate Professors	12
Assistant Professors	32
Master Degree Holders (Lecturers)	6
BSc Degree Holders (Teaching Assistants)	15 (The majority of them are Master Degree Holders)

Current Laboratories

Lab Type	Number of Labs	Capacity / Lab
General Computer Lab	13	30
Logic Design Lab	2	30
Network Lab	1	30

Core Courses:

Code	Course Title	Credits	Prerequisite
CSC640	Fundamental of Data Science	3	None

Description

The Data Science course concentrates on techniques and methods needed to in the Data Science project life-cycle, which includes data collection, data management and data preprocessing, analysis, presentation, as well as operationalization. This class aims at giving students an introduction to all phases of the process of data process and using modern tools and real data; they will gain hands-on experience of the process. The course includes topics such as data formats, cleaning, and loading; data analysis, data governance, data storage in no-relational and relational store; topping up using cluster computing; and data visualization. They will also store and access various data through using suitable data management tools, database, control accessibility of data that is sensitive, and implement conversions of data in different formats. Finally, students will be capable of presenting the results of data science project using reports and visualizations to be used as a foundation of operationalization.

Code	Course Title	Credits	Prerequisite
CIS 680	Advanced Topics in Databases	٣	None

Description

This course emphases on advanced topics in database systems. The course begins with an overview of the associated database analysis, design, and methodology issues. Then, it highlights transaction management, query processing, distributed DBMSs and replication, and object DBMSs. In addition, the course emphases on web technology and DBMSs, semi-structured data and XML. In Particular, the related business intelligence topics such data warehousing concepts, OLAP, and data mining also covered and emphasized.

Code	Course Title	Credits	Prerequisite
CSC 641	Computational Mathematics	3	None

Description

This course provides students with theories, fundamental concepts that are needed to study data science. The course will include statistics, probability and probability distributions, calculus, integrals and their applications. The course also will cover topics of linear algebra that are required in data science like matrices, vector spaces, subspaces, bases and dimension, Eigenvectors Eigen spaces, and linear transformations. The course also will include some of the major concepts in and multi-variate calculus like Lagrange Multipliers and Constrained Optimization problems. Through the course students will acquire cognitive skills through thinking and problem solving with special focus on applications of the different techniques of data science related topic.

Code	Course Title	Credits	Prerequisite
CSC650	Programming for Data Science	3	None

Description

This course introduces students to the dominant programming languages for AI/ML and deep learning. It dives deeply into programming tools, libraries, and frameworks for building research projects. This includes reading and loading datasets, preprocessing data, understanding structure using statistical summaries and data visualization, data manipulation, and cleaning techniques. This course also guides students through learning and implementing popular ML and deep learning frameworks such as TensorFlow, Keras, and PyTorch for symbolic math, used to perform differential programming and linear algebra.

Code	Course Title	Credits	Prerequisite
CSC603	Machine Learning	3	CSC641

Description

This course introduces the fundamental concepts and functioning of machine learning. It covers theoretical foundations as well as essential algorithms for supervised and unsupervised learning. The practical part will focus on the application of machine learning to a range of real-world problems. The topics include: linear and logistic regression, naïve Bayes classifier, k-NN, decision tree, random forest, support vector machine, clustering, dimensionality reduction, and perceptron.

Code	Course Title	Credits	Prerequisite
CSC617	Data Mining	3	CSC641

Description

In this course we explore data mining interdisciplinary field brings together techniques from databases, statistics, machine learning, and information retrieval. We will discuss the main data mining methods currently used, including data warehousing and data cleaning, clustering, classification, association rules mining, query flocks, text indexing and searching algorithms, how search engines rank pages, and recent techniques for web mining. Designing algorithms for these tasks is difficult because the input data sets are very large, and the tasks may be very complex. One of the main focuses in the field is the integration of these algorithms with relational databases and the mining of information from *semi-structured* data.

Code	Course Title	Credits	Prerequisite
CIS608	Research Methods	3	None

Description

The course aims to familiarize students with the fundamental concepts of research and the importance of research and its methodologies, including theory of science and qualitative and quantitative methods. Also, the course aims to give students skills for understanding the structure of a research paper, critical reading of research paper, developing a research proposal for a master's project and writing a research manuscript.

Students will use these theoretical concepts of research to begin to critically review literature relevant to the field of artificial intelligent and its applications.

Code	Course Title	Credits	Prerequisite
CIS651	Data Visualization	3	CSC650

Description

The course discusses the use of computer-supported, interactive and visual representations of data in order to amplify cognition, help people reason effectively about information, find patterns and meaning in the data, and easily explore the datasets from different perspectives in particular in data-intensive environment. The course covers the principles of data visualization, techniques and methods needed to provide clear illustrations of data. Specific techniques to display certain types of data such as text or time series data are also covered. Students get practical experience on how to employ and evaluate data visualization software tools and programming libraries, and learn the skills needed to convert raw datasets into meaningful, interactive, dynamic, and insightful graphical dashboards

Code	Course Title	Credits	Prerequisite
CIS641	Big Data Analytics	3	CSC640

Description

This course gives an overview of the Big Data phenomenon, focusing then on extracting value from the Big Data using predictive analytics techniques. It also focuses and includes fundamental of big data storage and management related issues, the main Big Data tools and technologies such as e.g. Hadoop & Spark, the potential use in a corporate environment, and the use of predictive analytics on big data.

Code	Course Title	Credits	Prerequisite
CSC613	Information Retrieval	3	CSC641

Description

This course studies the basic principles and practical algorithms used for information retrieval and text mining. It covers the tasks of indexing, searching, and recalling data, particularly text or other unstructured forms. The contents also includes: statistical characteristics of text, several important retrieval models, text categorization, recommendation system, clustering, information extraction, etc. The course emphasizes both the above applications and solid modeling techniques (e.g., probabilistic modeling) that can be extended for other applications.

Code	Course Title	Credits	Prerequisite
CIS670	Seminar in Data Science	3	completion of 18 credit hours

Description

The Seminar in Data Science course is intended to provide students with the latest research and practical experience in the field of data science and how its applied in the real environment. It also aims to give students the nature of daily work of data scientist. Student will attend actual or virtual talks by renowned data scientists and write a review report on such a talk, speakers are from both private sector and academia will be invited along the course. Every student will be given a real case study and work on it during the course to apply the knowledge they gain during the course and present the finding.

Code	Course Title	Credits	Prerequisite
CIS690	Data Science Professional and Ethical issues	3	-

Description

The objective of this module is to highlights the professional and ethical issues in Data Science field. Thus, this course provides a framework to analyze such issues and examine the ethical and privacy implications of collecting and managing big data. Explore the broader impact of the data science field on modern society and the principles of fairness, accountability and transparency

Elective Courses:

Code	Course Title	Credits	Prerequisite
CIS684	Text Analytics	3	completion of 30 credit hours

Description

Text Analytics are the methods and techniques used to extract useful knowledge from text to support decision making. This field includes a collection of research from the natural language processing, databases, data mining, and machine learning communities. The aim of this course is to be a primer for text analytics theory and practice. After taking this course, students will have an understanding of how to independently obtain, parse, and analyze textual information for organizations that want to extract valuable insights. Topics discussed in the course include obtaining data sets, understanding data formats, duplicate detection, cleaning data sets, tagging, indexing and search, evaluating algorithms, classification, clustering, topic modeling and entity resolution. Time permitting we may discuss advanced topic such as relation extraction, slot filling, knowledge graphs, knowledge base construction, the semantic web, question answering or other cutting edge topics..

Code	Course Title	Credits	Prerequisite
CSC615	Bioinformatics	3	completion of 30 credit hours

Description

This course provides an introduction to key concepts and methods in bioinformatics. Emphasis will be put on efficient algorithms and techniques used in common applications for the analysis of genetic sequences. Topics covered: comparison and alignment of two or more sequences, indexing and searching of sequence databases, motif discovery, searching with sequence patterns, gene prediction as well as mapping and assembly of data from genome sequencing. Necessary basic knowledge of molecular biology will be communicated throughout

Code	Course Title	Credits	Prerequisite
CSC616	Self-driving Vehicles	3	completion of 30 credit hours

Description

Self-driving cars, have rapidly become one of the most transformative technologies to emerge. They depend on Deep Learning algorithms and they create new opportunities in the mobility sector. This course is an introduction to the practice of deep learning through the applied theme of building a self-driving car. It is open to beginners and is designed for those who are new to machine learning, but it can also benefit advanced researchers in the field looking for a practical overview of deep learning methods and their application.

Code	Course Title	Credits	Prerequisite
CSC618	Parallel Computing	3	completion of 30 credit hours

Description

This course provides an introduction to the field of computational linguistics, also called natural language processing (NLP) - the creation of computer programs that can understand and generate natural languages (such as English). We will use natural language understanding as a vehicle to introduce the three major subfields of NLP: syntax (which concerns itself with determining the structure of an utterance), semantics (which concerns itself with determining the explicit truth-functional meaning of a single utterance), and pragmatics (which concerns itself with deriving the context-dependent meaning of an utterance when it is used in a specific discourse context). The course will introduce both linguistic (knowledge-based) and statistical approaches to NLP, illustrate the use of NLP techniques and tools in a variety of application areas, and provide insight into many open research problems.

Code	Course Title	Credits	Prerequisite
CIS682	Decision Support Systems	3	completion of 30 credit hours

Description

The objective of this course to provide learns with fundamental knowledge on decision support systems for managers and IS developers. This course explores topics in computer-based Decision Support Systems with a practical focus on the application of information technology to the solution of management problems. Topics include Management Support Systems, decision making systems, Data .mining for business and intelligent system.

Code	Course Title	Credits	Prerequisite
CSC 604	Neural Networks and Deep Learning	3	completion of 30 credit hours

Description

This course offers a broad introduction to neural networks and deep learning. It also explores the applications and theories relevant to problem solving using deep learning. The course applies deep learning algorithms to real-life problems in diverse domains such as computer vision, natural language processing, sequence modelling, and more. It covers neural networks, backpropagation, optimization (SGD, RMSprop, and Adam), autoencoders, convolutional neural networks, inception, residual networks, RNNs, LSTM, dropout, batch normalization, Xavier initialization, transfer learning, generative adversarial networks, and deep reinforcement learning.

Code	Course Title	Credits	Prerequisite
CIS622	Selected topics in Data Science	3	completion of 30 credit hours

Description

The objective of this module is to highlights the up to date issues/topics in Data Science field. Thus, the main purpose of this course is to highlight and investigate "selected/special topics" in Data Science that are not covered in the other offered courses. Such topics might be interrelated to one or more Data Science disciplines. However, the module should cover/give profound understanding for "the selected topics" including the associated perceptions, techniques, models, and tools.

Research Projects:

Code	Course Title	Credits	Prerequisite
CIS624	Research Project 1	4	completion of 21 credit hours

Description

This course provides students with an opportunity to gather the knowledge and skills learned from the program coursework and conduct a research project with industrial applications. Students are expected to conduct a review of research literature and develop a set of hypotheses for a research project in data science. A research project explaining the hypotheses and alternative remedies to the problem must be submitted to the faculty supervisor at the end of the semester. Students are evaluated based on their research project and oral presentation.

Code	Course Title	Credits	Prerequisite
CIS625	Research Project 2	4	CIS624

Description

The research outlined in the CIS624 proposal must be completed during this course. The final report of the research findings and recommendations should be submitted to the advisor and the results presented. The results should have direct practical applications and / or be available for publication in refereed publications. Students are evaluated based on the submitted research and oral presentation.