

# University of Pretoria Yearbook 2020

## MIT Big Data Science (Coursework) (12254017)

**Minimum duration of study** 2 years

**Total credits** 180

**NQF level** 09

### Programme information

This degree programme is presented in English only.

Also consult G Regulations G.30 to G.54

The curriculum is determined in consultation with the programme organiser.

A student will have to apply to the Dean of the Faculty of Engineering, Built Environment and Information Technology if he/she requires more than three years to complete the degree.

### Admission requirements

- A BIS Hons degree specialising in any of the following: Library Science/Information Science/Multimedia/Publishing; or MA (Development Communication) degree; or any applicable honours degree, and an average of at least 65%.
- The departmental Postgraduate Committee reserves the right to make a thorough assessment of the applicant's academic transcript and to decide if the applicant is suitable for postgraduate studies.
- The applicant must determine the research group/focus area and a potential supervisor and confirm the potential supervisor's availability (see departmental website for research groups and heads of research groups), and submit an abstract on the proposed research idea of approximately 800 to 1200 words.

### Other programme-specific information

#### Discontinuation of studies

The Dean may, on the recommendation of the admissions committee, cancel the studies of a student who fails more than one module. A module may only be repeated once.

#### Deregistration of modules

Deregistration of modules is only allowed before the early deadline.

### Examinations and pass requirements

A minimum semester mark of 40% is required in order to be admitted to the final examinations in all the prescribed modules of the degree. A final mark of 50% is required to pass all coursework modules and the mini-dissertation.

## Pass with distinction

The degree is conferred with distinction on students who have obtained at least 75% for the mini-dissertation and a minimum of 75% weighted average final mark for the coursework modules.

# Curriculum: Year 1

## Core modules

### Introduction to big data science 800 (MIT 800)

<b>Module credits</b>	5.00
<b>Contact time</b>	2 day workshop
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Quarter 1

#### Module content

This is the first and introductory module for the MIT degree in Big Data Science. Big Data and Data Science will be defined and students will be exposed to different application domains within the participating academic departments in the MIT degree. These departments include: Computer Science, Electrical, Electronic and Computer Engineering (EECE), Informatics, Information Science, Mathematics and Applied Mathematics, Statistics, and Health Science departments. The presentation of this module will be in the format of a two-day workshop.

### Introduction to machine and statistical learning 801 (MIT 801)

<b>Module credits</b>	15.00
<b>Prerequisites</b>	First year level higher education modules in Computer Science, Mathematics and Statistics.
<b>Contact time</b>	16 contact hours per semester
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Semester 1

#### Module content

In this module students will be exposed to different categories of machine and statistical learning algorithms that can be used to manipulate big data, identify trends from the data, modelling trends for prediction purposes as well as modelling for the detection of hidden knowledge. Students will be exposed to various machine and statistical learning algorithms/methods and they will learn how to make the right choice with regard to these. Learning, in a supervised and unsupervised mode will be covered. Furthermore students will develop a practical understanding of methods that can aid the learning process, such as, new developments in regression and classification, probabilistic graphical models, numerical Bayesian and Monte Carlo methods, neural networks, decision trees, deep learning and other computational methods. This module also includes a visualisation component focusing on the encoding of information, such as patterns, into visual objects.

### Introduction to data platforms and sources 802 (MIT 802)

<b>Module credits</b>	5.00
-----------------------	------

<b>Prerequisites</b>	First year level higher education modules in Computer Science and Statistics.
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Quarter 2

#### Module content

Students will obtain hands-on experience on the following technologies such as: Python, Spark, Hadoop, R and SAS, Streaming, Data fusion, Distributed file systems; and Data sources such as social media and sensor data.

### Introduction to Information Ethics for Big Data Science 803 (MIT 803)

<b>Module credits</b>	5.00
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Quarter 1

#### Module content

The focus in this module is on Information Ethics and its place within the disciplines of Ethics and Philosophy. The following topics will be covered: Information Ethics and PAPAS (privacy, accuracy, property, access, security); Information ethics and the life cycle of big data; Information ethical dilemmas within big data in different disciplines, e.g. science, technology, engineering and mathematics (STEM), health sciences, economics and management sciences, social sciences and the humanities; and Case studies.

### Introduction to mathematical optimization for big data science 804 (MIT 804)

<b>Module credits</b>	5.00
<b>Prerequisites</b>	First year level higher education modules in Computer Science, Mathematics and Statistics.
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Quarter 2

#### Module content

In this module students will be introduced to Mathematical Optimization through gaining knowledge about the theory and algorithms to solve optimisation problems. Topics will include: Linear programming, unconstrained optimization, equality constrained optimization, general linearly and nonlinearly constrained optimization, quadratic programming, global optimization, Theory and algorithms to solve these problems.

### Big data 805 (MIT 805)

<b>Module credits</b>	10.00
-----------------------	-------

<b>Prerequisites</b>	First year level higher education modules in Computer Science.
<b>Contact time</b>	10 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Semester 2

#### Module content

This module focuses on tools for Big Data processing. The focus is on the 3 V- characteristics of Big Data namely volume, velocity and variety. Students will learn about the different architectures available for Big Data processing. The map-reduce algorithm will be studied in detail as well as graphical models for Big Data. The module will include a significant component of practical work (hands-on) where students will be exposed to real use cases that are or can be implemented on Big Data platforms.

### Big data management 806 (MIT 806)

<b>Module credits</b>	10.00
<b>Prerequisites</b>	First year level higher education modules in Computer Science.
<b>Contact time</b>	6 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Quarter 4

#### Module content

Big data management is the governance, administration and organization of large volumes of both structured and unstructured data. Aspects included in big data management are: big data as organizational asset, harnessing big data as disruptive technology for competitive advantage, big data quality and accessibility; management strategies for large and fast-growing internal and external data, big data infrastructure and platform management, and big data policy, strategy and compliance.

### Research methods for big data science 809 (MIT 809)

<b>Module credits</b>	5.00
<b>Contact time</b>	6 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Semester 2

#### Module content

Similar to MIT 862; which has the following description: Research methodologies applicable to the IT field as preparation for the mini-dissertation for the A Stream students.

### Elective modules

### Big data science elective 801 (COS 801)

<b>Module credits</b>	5.00
<b>Prerequisites</b>	No prerequisites.
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Computer Science
<b>Period of presentation</b>	Semester 2

#### Module content

Example courses, amongst others, may include: Cyber-security, Digital Forensics, Deep Machine Learning, Image and sound analysis, Feature extraction, and Graph Modelling. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

### Big data science elective 802 (COS 802)

<b>Module credits</b>	5.00
<b>Prerequisites</b>	No prerequisites.
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Computer Science
<b>Period of presentation</b>	Semester 2

#### Module content

Example courses, amongst others, may include: Cyber-security, Digital Forensics, Deep Machine Learning, Image and sound analysis, Feature extraction, and Graph Modelling. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

### Big data science elective 801 (ERZ 801)

<b>Module credits</b>	5.00
<b>Prerequisites</b>	No prerequisites.
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Electrical, Electronic and Computer Engineering
<b>Period of presentation</b>	Semester 2

#### Module content

Example courses may include: Intelligent systems and Internet of Things. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

### Big data science elective 802 (ERZ 802)

<b>Module credits</b>	5.00
<b>Prerequisites</b>	No prerequisites.
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Electrical, Electronic and Computer Engineering
<b>Period of presentation</b>	Semester 2

#### Module content

Example courses may include: Intelligent systems and Internet of Things. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

### Big data science elective 801 (INF 801)

<b>Module credits</b>	5.00
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Informatics
<b>Period of presentation</b>	Semester 2

#### Module content

See existing electives from MIT modules in Stream A and B. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

### Big data science elective 802 (INF 802)

<b>Module credits</b>	5.00
<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Informatics
<b>Period of presentation</b>	Semester 2

#### Module content

See existing electives from MIT modules in Stream A and B. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

### Big data science elective 820 (INL 820)

<b>Module credits</b>	5.00
<b>Language of tuition</b>	Module is presented in English

**Department** Information Science

**Period of presentation** Semester 2

### Module content

Five credits of an elective module can be drawn from Information Science. A module in Research Data Management (RDM) is available as an elective. The following topics would typically be covered: Open Science and the dependency on open (big) data, The research process and the life cycle of big data (data management plans to publishing derivative data sets, licensing and legal implications); managing (curating) big vs long tail data; solving problems with research data vs the business value of big data (data-intensive decisionmaking); managing data as an asset (also data citation); issues and challenges involved in the management of big data (principles and best practices for effective big data governance); trusted data repositories; data stewardship frameworks for big data; and the data steward's toolbox.

## Statistics elective 801 (STK 801)

**Module credits** 5.00

**Prerequisites** As determined by the Department of Statistics.

**Contact time** 5 contact hours

**Language of tuition** Module is presented in English

**Department** Statistics

**Period of presentation** Semester 2

### Module content

Five 5 credits of an elective course can be drawn from the Department of Statistics. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

## Statistics elective 802 (STK 802)

**Module credits** 5.00

**Prerequisites** As determined by the Department of Statistics.

**Contact time** 5 contact hours

**Language of tuition** Module is presented in English

**Department** Statistics

**Period of presentation** Semester 2

### Module content

Five 5 credits of an elective course can be drawn from the Department of Statistics. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

## Big data science elective 801 (WTW 801)

**Module credits** 5.00



<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Mathematics and Applied Mathematics
<b>Period of presentation</b>	Semester 2

#### **Module content**

Five 5 credits of an elective course can be drawn from Mathematics and Applied Mathematics. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

### **Big data science elective 802 (WTW 802)**

**Module credits** 5.00

<b>Contact time</b>	5 contact hours
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	Mathematics and Applied Mathematics
<b>Period of presentation</b>	Semester 2

#### **Module content**

Five 5 credits of an elective course can be drawn from Mathematics and Applied Mathematics. In addition to study-leader approval, elective course selection may be subject to course pre-requisites, course availability, and internal departmental regulations as decided by the Head of the Department.

## Curriculum: Final year

### Core modules

#### Mini dissertation in big data science 807 (MIT 807)

<b>Module credits</b>	90.00
<b>Prerequisites</b>	All the core modules must be passed
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Year

#### Module content

Students may choose a supervisor/co-supervisor from any of the participating departments, which includes, but are not limited to: Electrical, Electronic & Computer Engineering (EECE), Computer Science, Informatics, Information Science, Mathematics and Applied Mathematics, Statistics and Faculty of Health Science departments (Computational biology, Family Medicine, Radiology).

#### Big data science project 808 (MIT 808)

<b>Module credits</b>	20.00
<b>Prerequisites</b>	All the core modules must be passed
<b>Contact time</b>	8 contact hours per semester
<b>Language of tuition</b>	Module is presented in English
<b>Department</b>	School of Information Technology
<b>Period of presentation</b>	Semester 1

#### Module content

This module provides the opportunity to students for demonstrating the application of the theoretical Big Data Science knowledge gained in the core part of this degree. Students are expected to identify and work with a collaborator who is taking ownership for the project. This collaborator can either be an industry partner or a researcher within one of the participating departments. Projects will be based on the entire big data lifecycle as discussed in this degree programme. This includes the gathering of data of a significant size as well as a final technical report describing the process followed and the deliverables. Depending on the complexity of the project, students can apply to work in groups with a maximum of two members. The proposed project will be subject to approval by the Department Computer Science.

The information published here is subject to change and may be amended after the publication of this information. The [General Regulations \(G Regulations\)](#) apply to all faculties of the University of Pretoria. It is expected of students to familiarise themselves well with these regulations as well as with the information contained in the [General Rules](#) section. Ignorance concerning these regulations and rules will not be accepted as an excuse for any transgression.