University of Pretoria Yearbook 2020

MSc eScience (Coursework) (02250193)

Minimum duration of study	2 years
Total credits	180
NQF level	09

Programme information

The curriculum for the MSc (eScience) coursework degree programme comprises 180 credits of coursework and a research component. One of the key features of the curriculum is a capstone project that runs parallel with coursework modules in the first year of study. During the capstone project, students will go through the entire cycles of solving a real-world data science problem, collecting and processing real-world data, designing methods to solve the problem, and implementing a solution. The capstone project and coursework prepare the student for the mini-dissertation problem supervised by an expert.

Admission requirements

- An honours degree in either statistics, mathematics, computer science, physics, or related fields; and demonstrable knowledge of basic principles of probability and statistics, computing, calculus and linear algebra;

  or

- Passing an entrance evaluation designed by the academic advisory committee of the programme within the consortium. An average of 65% or more at honours level is the minimum for consideration. Admission will be competitive and an honours average of at least 70% is highly recommended.
- Student numbers are limited to a maximum of 30 over the consortium (6 at the University of Pretoria).
- Admission is additionally dependent on availability of supervisor/s and/or projects within the participating departments.
- Historical performance during prior studies will also be considered in selecting students. Specific attention will be given to modules repeated and duration of study.
- The research proposal of applicants should be in line with the research focus of the participating departments.
- Any further additional entrance requirements as specified by the head of department in consultation with the departmental postgraduate selection committee and the academic advisory committee of the programme.
- Applications must include full academic transcripts from undergraduate to current level. SAQA evaluation documentation is required for international applications and on request for students with qualifications from South African institutions.

Other programme-specific information

Candidates are required to familiarise themselves with the General Regulations regarding the maximum period of

registration and other requirements for master's degrees.

Promotion to next study year

The progress of all master's candidates is monitored biannually by the supervisor and the postgraduate coordinator. A candidate's study may be terminated if the progress is unsatisfactory or if the candidate is unable to finish his/her studies during the prescribed period.

Subject to exceptions approved by the Dean, on recommendation of the head of department, and where applicable, a student may not enter for the master's examination in the same module more than twice.

Pass with distinction

The MSc degree is conferred with distinction to candidates who obtain a final average mark of at least 75%, with a minimum of 65% in each module, and a mark of at least 75% for the mini-dissertation from each of the members of the examination panel. Where a member of the examination panel awards a mark of less than 75% for the mini-dissertation, that member of the examination panel must offer, in writing, support for his/her decision, or indicate in writing that he/she supports the examination committee's decision to confer the degree with distinction.

Curriculum: Year 1

Minimum credits: 90

Choose 4 modules to the value of 60 credits from the list of electives.

Core modules

Research methods and capstone project in data science 801 (NEP 801)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

Scientific writing styles; layouts for assignments, projects, theses or publications; research methodologies; scientific assignments; integration of all the aforementioned content items for a capstone project in data science.

Data privacy and ethics 802 (NEP 802)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

Technical processes of data collection, storage, exchange and access; Ethical aspects of data management; Legal and regulatory frameworks in South Africa and in relevant jurisdictions; Data policies; Data privacy; Data ownership; Legal liabilities of analytical decisions and discrimination; and the Technical and algorithmic approaches to enhance data privacy, and relevant case studies.

Elective modules

Adaptive computation and machine learning 803 (NEP 803)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content
Introduction: Basic concepts. Supervised learning setup: Least means squares, logistic regression, perceptron, exponential family, generative learning algorithms, Gaussian discriminant analysis, naïve Bayes, support vector machines, model selection and feature selection. Learning theory: bias/variance tradeoff, union and Chernoff/Hoeffding bounds, VC dimension, worst case (online) learning. Unsupervised learning: clustering, k-means, expectation maximisation, mixture of Gaussians, factor analysis, principal components analysis, independent components analysis. Reinforcement learning and control: Markov decision processes, Bellman equations, value iteration and policy iteration, Q-learning, value function approximation, policy search, reinforce, partially observable Markov decision problems.

Data visualisation and exploration 804 (NEP 804)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

Data and image models; visualisation attributes (colour) and design (layout); exploratory data analysis; interactive data visualisation; multidimensional data; graphical perception; visualisation software (Python & R); and types of visualisation (animation, networks and text).

Large-scale computing systems and scientific computing 805 (NEP 805)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

Introduction to scientific computing architectures in Python, introduction to distributed systems, introduction to distributed databases, introduction to parallelism, large-data computation and storage models, introduction to well-known distributed systems architectures, and programming large-data applications on open-source infrastructures for data processing and storage systems.

Mathematical foundations of data science 806 (NEP 806)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

High-dimensional space, best-fit subspaces and singular value decomposition, random walks and Markov chains, statistical machine learning, clustering, random graphs, topic models, matrix factorisation, hidden Markov models, graphical models, wavelets, and sparse representations.

Special topics in data science 807 (NEP 807)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

Specialised and applied concepts and trends in data science.

Statistical foundations of data science 808 (NEP 808)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

An understanding of multivariate statistics, hypothesis testing and confidence intervals. The ability to model data using well-known statistical distributions as well as the ability to handle data that is both continuous and categorical. The ability to perform statistical modelling including multivariate linear regression and adjust for multiple hypotheses. Forecasting, extrapolation, prediction and modelling using statistical methods. Bayesian statistics, an understanding of bootstrapping and Monte Carlo simulation.

Large-scale optimisation for data science 809 (NEP 809)

Module credits	15.00
Prerequisites	No prerequisites.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Semester 1 or Semester 2

Module content

Introduction to convex optimisation, subgradient methods, decomposition and distributed optimisation, proximal and operator splitting methods, conjugate gradients, and nonconvex problems.

Curriculum: Final year

Minimum credits: 90

Fundamental modules

Mini-dissertation: eScience 800 (NEP 800)

Module credits	90.00
Prerequisites	Completion of the coursework programme.
Language of tuition	Module is presented in English
Department	Statistics
Period of presentation	Year

Module content

This is the research component of the MSc (eScience) degree and comprises a mini-dissertation which develops the research skills and bridges the gap between theory and practice.

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.