

Liverpool John Moores University

Title: APPLICATIONS IN BIG DATA & MACHINE LEARNING
Status: Definitive
Code: **7124COMP** (121347)
Version Start Date: 01-08-2021

Owning School/Faculty: Computer Science and Mathematics
Teaching School/Faculty: Computer Science and Mathematics

Team	Leader
Abir Hussain	Y

Academic Level: FHEQ7 **Credit Value:** 20 **Total Delivered Hours:** 55
Total Learning Hours: 200 **Private Study:** 145

Delivery Options

Course typically offered: Semester 2

Component	Contact Hours
Lecture	22
Practical	33

Grading Basis: 50 %

Assessment Details

Category	Short Description	Description	Weighting (%)	Exam Duration
Artefacts	AS1	Machine Learning Exercise 1	50	
Artefacts	AS2	Machine Learning Exercise 2	50	

Aims

The aim of this module is to complete the data science "process" by building data products that feedback and influence the real world using the results of that data science process. In doing so, students gain experience in bringing to fruition the data science process.

Learning Outcomes

After completing the module the student should be able to:

- 1 Design a data application using large scale data storage and machine learning tools
- 2 Construct a data application using large scale data storage and machine learning tools
- 3 Design a machine learning/analytics exercise for a given subject area
- 4 Develop a machine learning/analytics solution for a given subject area

Learning Outcomes of Assessments

The assessment item list is assessed via the learning outcomes listed:

Machine Learning Exercise 1	1	2
Machine Learning Exercise 2	3	4

Outline Syllabus

Review of large scale data storage platforms and machine learning tools e.g.Hadoop and Mahout

Recommendation Systems

Introduction to recommender systems

Representing recommender data

Production level recommendation systems

Distributing recommender computations

Clustering

Review of clustering

Representing data

Clustering algorithms reviewed

Evaluating and improving clustering quality

Clustering in production

Case studies in clustering

Classification

Introduction to classification

Training a classifier

Evaluating and tuning a classifier

Deploying a classifier

Case studies in classification

Emerging trends in applications of large scale data processing and machine learning

Learning Activities

Lectures will explore the practical application of the tools and techniques of data science to large scale data sets to produce tangible outcomes that feed back to the real world. Practical sessions will implement those ideas to provide experience in building useful applications using these ideas.

Notes

This module provides both theoretical and practical experience of large scale data storage considerations and the use of tools to support the processing of that data.