



TEQSA ID PRV:14323
CRICOS Provider Code: 03866C

MIT633 BIG DATA

SYDNEY INSTITUTE OF HIGHER EDUCATION > PROGRAMS > MIT633 BIG DATA

Unit Outline

| | |
|---------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Important Update: | Our aim is to provide you with an optimal learning experience, regardless of how this unit is delivered. Teaching will be delivered in line with the most current COVID Safe health guidelines. This may include a mix of online and face-to-face. Please check the learning management system for announcements and updates. Thank you for your flexibility and commitment to studying with Sydney Institute of Higher Education. |
| Enrolment Modes: | Year 2, Semester 2. |
| Credit Point(s): | 12.5 |
| EFTSL Value: | 0.125 |
| Prerequisites: | MIT631 Data Analytics |
| Typical study commitment: | Students will on average spend 10 hours per week over the teaching period undertaking the teaching, learning and assessment activities for this unit. |
| Scheduled learning activities: | 4 timetabled hours per week, 6 personal study hours per week. |
| Other resource requirements: | Students will need access to lab computers or will need their own laptops in order to carry out lab exercises and assignments. Students will need to use Apache Hadoop and Apache Hbase. |

Unit description

This unit covers the key concepts, applications, architectures, and processes that are widely used in big data applications to collect, integrate, analyse and present data, often in different formats and from various sources. It also covers technologies that are commonly used in industry such as NoSQL, Map-Reduce and Hadoop. Students will gain understanding of the challenges faced by organisations for managing large volumes of data.

Unit learning outcomes (ULO)

On the successful completion of this unit student will be able to:

| | |
|------|----------------------------------------------------------------------------|
| ULO1 | Evaluate different big data concepts, tools, techniques, and applications. |
| ULO2 | Analyse and visualise data using available big data tools. |
| ULO3 | Design appropriate repository structure for storing big data. |
| ULO4 | Design big data solutions using Map-reduce techniques. |
| ULO5 | Create solutions for data storage and manipulation. |

Topics to be included in the unit

| | |
|-----|----------------------------------------------|
| 1. | Introduction to big data |
| 2. | Non-structured data |
| 3. | Organizing, storing, and processing big data |
| 4. | Finding similar items |
| 5. | Link Analysis |
| 6. | Mining Social Network Graphs |
| 7. | Recommendation Systems |
| 8. | Data analytics in big data |
| 9. | Data visualization |
| 10. | Map-reduce framework |
| 11. | Hadoop |
| 12. | Big data applications & Revision |

Assessment

| Assessment Description | Grading and weighting (% total mark for unit) | Due date |
|------------------------------------------------------|--------------------------------------------------|-----------------|
| Assessment 1: Class Participation | 10% | Weeks 1-12 |
| Assessment 2: Design Report and Presentation (Group) | 30% | Week 8 |
| Assessment 3: Practical Project | 30% | Week 12 |
| Assessment 4: Final Exam | 30% | Final exam week |