KING'S OWN INSTITUTE*

Success in Higher Education

ICT731 DATA MINING T326 BRIEF

All information in the Subject Outline is correct at the time of approval. KOI reserves the right to make changes to the Subject Outline if they become necessary. Any changes require the approval of the KOI Academic Board and will be formally advised to those students who may be affected by email and via Moodle.
Information contained within this Subject Outline applies to students enrolled in the trimester as indicated

1. General Information

1.1 Administrative Details

Associated HE Award(s)	Duration	Level	Subject Coordinator
Master of Information Technology (MIT)	1 trimester	Postgraduate	Dr Rachid Hamadi. rachid.hamadi@koi.edu.au P: +61 (2) 9283 3583 L: 7-11, 11 York Street. Consultation: via Moodle or by appointment.

1.2 Core/Elective

This subject is
- a core subject for the Master of Information Technology (MIT) Data Analytics
- an elective subject for the Master of Information Technology (MIT) General

1.3 Subject Weighting

Indicated below is the weighting of this subject and the total course points

Subject Credit Points	Total Course Credit Points
4	Master of Information Technology (64 Credit Points)

1.4 Student Workload

Indicated below is the expected student workload per week for this subject

No. Timetabled Hours/Week*	No. Personal Study Hours/Week**	Total Workload Hours/Week***
3 hours/week plus supplementary online material	7 hours/week	10 hours/week

* Total time spent per week at lectures and tutorials
** Total time students are expected to spend per week in studying, completing assignments, etc.
*** Combination of timetable hours and personal study

1.5 Mode of Delivery Classes will be face-to-face or hybrid. Certain classes will be online (e.g., special arrangements).

1.6 Pre-requisites ICT713 Advanced Database Design and Development
ICT724 Intelligent Systems

1.7 General Study and Resource Requirements

*AUSTRALIAN INSTITUTE OF BUSINESS AND MANAGEMENT PTY LTD © ABN: 72 132 629 979 CRICOS 03171A

- Students are expected to attend classes with the weekly worksheets and subject support material provided in Moodle. Students should read this material before coming to class to improve their ability to participate in the weekly activities.
- Students will require access to the internet and their KOI email and should have basic skills in word processing software such as MS Word, spreadsheet software such as MS Excel and visual presentation software such as MS PowerPoint.
- Computers and WIFI facilities are extensively available for student use throughout KOI. Students are encouraged to make use of the campus Library for reference materials.

Software resource requirements specific to this subject: Office 365, MS Imagine, MS Excel, Python, Spyder and Jupyter Notebook, RapidMiner.

1.8 Academic Advising

Academic advising is available to students throughout teaching periods including the exam weeks. As well as requesting help during scheduled class times, students have the following options:

- Consultation times: A list of consultation hours is provided on the homepage of Moodle where appointments can be booked.
- Subject coordinator: Subject coordinators are available for contact via email. The email address of the subject coordinator is provided at the top of this subject outline.
- Academic staff: Lecturers and Tutors provide their contact details in Moodle for the specific subject. In most cases, this will be via email. Some subjects may also provide a discussion forum where questions can be raised.
- Head of Program: The Head of Program is available to all students in the program if they need advice about their studies and KOI procedures.
- Vice President (Academic): The Vice President (Academic) will assist students to resolve complex issues (but may refer students to the relevant lecturers for detailed academic advice).

2. Academic Details

2.1 Overview of the Subject

This subject provides a comprehensive introduction to data mining techniques, equipping students with practical skills to extract insights from data using Python and other tools like Postgres SQL and MS Excel. Topics include data preparation, warehousing, visualization, and advanced machine learning algorithms such as clustering, regression, classification, and text mining, with a focus on their application to real-world problems. Ethical and privacy challenges in data mining are also addressed, ensuring responsible use of these techniques. Students will gain hands-on experience through practical assessments aligned with industry standards, preparing them for data-driven decision-making roles in business and technology.

2.2 Graduate Attributes for Postgraduate Courses

Graduates of postgraduate courses from King's Own Institute will gain the graduate attributes expected from successful completion of a postgraduate degree under the Australian Qualifications Framework (2nd edition, January 2013). Graduates at this level will be able to apply advanced body of knowledge from their major area of study in a range of contexts for professional practice or scholarship and as a pathway for further learning.

King's Own Institute's generic graduate attributes for a master's level degree are summarised below:

	KOI Postgraduate Degree Graduate Attributes	**Detailed Description**

*AUSTRALIAN INSTITUTE OF BUSINESS AND MANAGEMENT PTY LTD © ABN: 72 132 629 979 CRICOS 03171A

	Knowledge	Current, comprehensive and coherent knowledge, including recent developments and applied research methods
	Critical Thinking	Critical thinking skills to identify and analyse current theories and developments and emerging trends in professional practice
	Communication	Communication and technical skills to analyse and theorise, contribute to professional practice or scholarship, and present ideas to a variety of audiences
	Research and Information Literacy	Cognitive and technical skills to access and evaluate information resources, justify research approaches and interpret theoretical propositions
	Creative Problem Solving Skills	Cognitive, technical and creative skills to investigate, analyse and synthesise complex information, concepts and theories, solve complex problems and apply established theories to situations in professional practice
	Ethical and Cultural Sensitivity	Appreciation and accountability for ethical principles, cultural sensitivity and social responsibility, both personally and professionally
	Leadership and Strategy	Initiative, leadership skills and ability to work professionally and collaboratively to achieve team objectives across a range of team roles Expertise in strategic thinking, developing and implementing business plans and decision making under uncertainty
	Professional Skills	High level personal autonomy, judgement, decision-making and accountability required to begin professional practice

Across the courses, these skills are developed progressively at three levels:

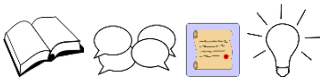
- **Level 1 Foundation –** Students learn the skills, theories and techniques of the subject and apply them in stand-alone contexts
- **Level 2 Intermediate –** Students further develop skills, theories and techniques of the subject and apply them in more complex contexts, beginning to integrate the application with other subjects
- **Level 3 Advanced** – Students have a demonstrated ability to plan, research and apply the skills, theories and techniques of the subject in complex situations, integrating the subject content with a range of other subject disciplines within the context of the course

Generally, skills gained from subjects in the Graduate Certificate and Graduate Diploma are at levels 1 and 2 while other subjects in the Master's degree are at level 3.

2.3 Subject Learning Outcomes

Listed below, are key knowledge and skills students are expected to attain by successfully completing this subject:

Subject Learning Outcomes	**Contribution to Course Graduate Attributes**

a) Analyse the value, rationale and applications of data mining for organisations	
b) Critically evaluate and recommend different data preparation methods and strategies	
c) Apply various data mining methods and models to provide results to enhance business decision making	
d) Design a predictive model using data, text, and web mining techniques	
e) Research and evaluate ethical issues related to data mining	

2.4 Subject Content and Structure

Below are details of the subject content and how it is structured, including specific topics covered in lectures and tutorials. Reading refers to the text unless otherwise indicated.

Weekly Planner:

Week (beginning)	Topic covered in each week's lecture	Reading(s)	Expected work as listed in Moodle
Week 1 26 October	Fundamentals of data mining	Chs. 1, 2 (Jamsa) Ch. 1 (Han)	Tutorial on Python language. Formative not graded
Week 2 02 November	Data Warehousing	Ch. 3 (Jamsa) Ch. 3 (Han)	Tutorial on data warehousing. Discussion questions and case study using Postgres SQL and Python Summative graded
Week 3 09 November	Data Pre-processing and Cleansing	Ch. 9 (Jamsa) Ch. 2 (Han)	Tutorial on data preparation. Discussion questions and case study plus lab exercise on data preparation. Summative graded
Week 4 16 November	Data Visualisation and Dimensionality Reduction	Chs. 2, 4 (Jamsa)	Lab exercise on data visualisation using MS Excel and Python. Exercises on dimensionality reduction using PCA. Summative graded
Week 5 23 November	Association rules	Ch. 13 (Jamsa)	Tutorial on association rules. Discussion questions and case study plus lab exercise on Apriori and FP-growth.

Week (beginning)	Topic covered in each week's lecture	Reading(s)	Expected work as listed in Moodle
			Summative graded
Week 6 30 November	Clustering	Ch.10 (Jamsa) Chs. 8, 9 (Han)	Tutorial on clustering methods. Discussion questions and case study plus lab exercise on k-means and Hierarchical clustering. Summative graded **Assessment 2 due: Report**
Week 7 07 December	Regression – linear regression	Ch. 12 (Jamsa)	Tutorial on regression models. Discussion questions on linear regression. Summative graded
Week 8 14 December	Classification – logistic regression	Ch. 11 (Jamsa) Ch. 7 (Han)	Tutorial on classification models. Discussion questions and case study plus lab exercise on logistic regression. Summative graded
Week 9 04 January	Classification – decision tree	Ch. 11 (Jamsa) Ch. 6 (Han)	Tutorial on classification models. Discussion questions and case study plus lab exercise on decision trees. Summative graded **Assessment 3 due: Quiz**
Week 10 11 January	Text Classification – Naïve Bayes	Chs 11, 14 (Jamsa)	Tutorial on text classification models. Discussion questions and case study plus lab exercise on Bayes classifier. Summative graded
Week 11 18 January	Data mining ethics	Ch. 16 (Jamsa)	Tutorial on Data mining ethics Summative graded **Assessment 4 Report due**
Week 12 25 January	Revision + Viva Presentations	All chapters	Revision **Assessment 4 viva**
Week 13 01 February	Study review week and Final Exam Week		

Week (beginning)	Topic covered in each week's lecture	Reading(s)	Expected work as listed in Moodle
Week 14 08 February	- Final Exam Week - Continuing students - enrolments for T127 open		Please see the exam timetable for the exam date, time and location
Week 15 15 February	- Student Vacation begins - New students - enrolments for T127 open		
Week 16 22 February	f) Results Released g) Review of Grade Day for T326 – see Sections 2.6 and 3.2 below for relevant information. h) Certification of Grades NOTE: More information about the dates will be provided at a later date through Moodle/KOI email.		
T127 01 March 2027			
Week 1 01 March	Week 1 of classes for T127		

2.5 Teaching Methods/Strategies

Briefly described below are the teaching methods/strategies used in this subject:

- *Lectures* (1 hour/week) are conducted in seminar style and address the subject content, provide motivation and context and draw on the students' experience and preparatory reading.
- *Tutorials* (2 hours/week) include class discussion of case studies and research papers, practice sets and problem-solving and syndicate work on group projects. Tutorials often include group exercises and so contribute to the development of teamwork skills and cultural understanding. Tutorial participation is an essential component of the subject and contributes to the development of many of the graduate attributes (see section 2.2 above). Tutorial participation contributes towards the assessment in many subjects (see details in Section 3.1 for this subject). Supplementary tutorial material such as case studies, recommended readings, review questions etc. will be made available each week in Moodle.
- *Online* teaching resources include class materials, readings, model answers to assignments and exercises and discussion boards. All online materials for this subject as provided by KOI will be found in the Moodle page for this subject. Students should access Moodle regularly as material may be updated at any time during the trimester
- *Other contact* - academic staff may also contact students either via Moodle messaging, or via email to the email address provided to KOI on enrolment.

2.6 Student Assessment

Assessment is designed to encourage effective student learning and enable students to develop and demonstrate the skills and knowledge identified in the subject learning outcomes. Assessment tasks during the first half of the study period are usually intended to maximise the developmental function of assessment (formative assessment). These assessment tasks include weekly tutorial exercises (as indicated in the weekly planner) and low stakes graded assessments (as shown in the graded assessment table). The major assessment tasks where students demonstrate their knowledge and skills (summative assessment) generally occur later in the study period. These are the major graded assessment items shown in the graded assessment table.

Final grades are awarded by the Board of Examiners in accordance with KOI's Assessment and Assessment

*AUSTRALIAN INSTITUTE OF BUSINESS AND MANAGEMENT PTY LTD © ABN: 72 132 629 979 CRICOS 03171A

Appeals Policy. The definitions and guidelines for the awarding of final grades are:

- *HD High distinction* (85-100%): an outstanding level of achievement in relation to the assessment process.
- *D Distinction* (75-84%): a high level of achievement in relation to the assessment process.
- *C Credit* (65-74%): a better than satisfactory level of achievement in relation to the assessment process.
- *P Pass* (50-64%): a satisfactory level of achievement in relation to the assessment process.
- *F Fail* (0-49%): an unsatisfactory level of achievement in relation to the assessment process.
- *FW:* This grade will be assigned when a student did not submit any of the compulsory assessment items.

Provided below is a schedule of formal assessment tasks and major examinations for the subject.

Assessment Type	**When assessed**	**Weighting**	**Learning Outcomes Assessed**
Assessment 1: Weekly tutorials	Weeks (2 – 11)	20% (2% per week)	a, b, c, d, e
Assessment 2: Individual report and video presentation on data pre-processing and preliminary analysis (2000–2500 words)	Week 6	30%	a, b, e
Assessment 3: Quiz (In-class)	Week 9	15%	a, b
Assessment 4: Report + Viva Examination (in-class)	Report: Week 11; Viva: Week 12	Report: 15% Viva Examination: 20% Total: 35%	a, b, c, d, e

Requirements to Pass the Subject:

To gain a pass or better in this subject, students must gain a *minimum of 50%* of the total available subject marks.

2.7 Prescribed and Recommended Readings

Provided below, in formal reference format, is a list of the prescribed and recommended readings.

Prescribed Text:

Han, J., Pei, J. and Tong, H. (2022) *Data mining : concepts and techniques. 4th edn.* Morgan Kaufmann. ISBN: 978-0128117613.

Jamsa, Kris., 2020. *Introduction to Data Mining and Analytics, Jones & Bartlett Learning*, LLC. ISBN: 9781284180909.

Recommended Readings:

Eldén, L., 2019. *Matrix methods in data mining and pattern recognition (Second edition.),* Society for Industrial and Applied Mathematics. ISBN: 9781611975864.

Gori, M., Betti, A., & Melacci, S. (2023). Machine Learning: A constraint-based approach. Elsevier. ISBN: 9780323898591.

Grus, J., 2019. *Data science from scratch: first principles with Python (Second edition.),* O'Reilly Media. ISBN:9781492041085.

Olson, David L., 2018. *Data Mining Models, 2nd edition, Business Expert Press.* ISBN: 9781948580496.

Rao, G. S., & Muneeswari, G. (2024). A Review: Machine Learning and Data Mining Approaches for Cardiovascular Disease Diagnosis and Prediction. EAI Endorsed Transactions on Pervasive Health and Technology, 10.DOI: https://doi.org/10.4108/eetpht.10.5411.

Tsui, KL., Chen, V., Jiang, W., Yang, F., Kan, C. (2023). Data Mining Methods and Applications. In: Pham, H. (eds) Springer Handbook of Engineering Statistics. Springer Handbooks. Springer, London. DOI: https://doi.org/10.1007/978-1-4471-7503-2_38.

Suggested Periodicals:

- Journal of Management Analytics: https://www.tandfonline.com/loi/tjma20
- International Journal of Big Data Intelligence: https://www.inderscience.com/jhome.php?jcode=ijbdi
- MIS Quarterly: https://www.misq.org/
- International Journal of Information Management: https://www.journals.elsevier.com/international-journal-of-information-management

Useful Websites:

The following industry websites are useful introductory sources covering a range of information useful for this subject.

- Kanade, V. (2022). Data Mining: Definition, techniques, and Tools, Spiceworks. Available at: https://www.spiceworks.com/tech/big-data/articles/what-is-data-mining/
- What is data mining? IBM. Available at: https://www.ibm.com/topics/data-mining
- W3Schools (2024). Data Science Tutorial. https://www.w3schools.com/datascience/default.asp.
- W3Schools (2024). Machine Learning Tutorial. https://www.w3schools.com/python/python_ml_getting_started.asp.
- W3Schools (2024). Pandas Tutorial. https://www.w3schools.com/python/pandas/default.asp.

Conference/ Journal Articles:

Students are encouraged to read peer reviewed journal articles and conference papers. Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: articles, theses, books, abstracts and court opinions, from academic publishers, professional societies, online repositories, universities and other web sites.

*AUSTRALIAN INSTITUTE OF BUSINESS AND MANAGEMENT PTY LTD © ABN: 72 132 629 979 CRICOS 03171A