UNIVERSITI TEKNOLOGI MARA

UCS551: INTRODUCTION TO DATA ANALYTICS AND APPLICATION

Course Name (English)	INTRODUCTION TO DATA ANALYTICS AND APPLICATION APPROVED
Course Code	UCS551
MQF Credit	3
Course Description	This course is an introductory course on data analytics and its application, which include basic data management, methods to explore and visualize data effectively; as well as classification and data clustering using selected machine learning techniques. During the course, the techniques will be illustrated using a range of data and case studies with respect to the study program.
Transferable Skills	Demonstrate ability to identify and articulate self skills, knowledge and understanding confidently and in a variety of contexts.
Teaching Methodologies	Lectures, Blended Learning, Lab Work, Discussion
CLO	CLO1 Demonstrate social skills in project related to data analytics CLO2 Demonstrate ethics and professionalism in task related to data analytics. CLO3 Analyze data with relevant techniques in data analysis and exploration. CLO4 Apply statistical methods and machine learning techniques on data related to substantive domain in Malaysia
Pre-Requisite Courses	No course recommendations

Topics
1. Introduction to data analytics 1.1) Definition of data analytics 1.2) The importance of data analytics 1.3) Framework and process of data analytics 1.4) Example of applications
2. Data Understanding 2.1) Data Type: structured and unstructured 2.2) Data Structure (vector, matrix, data frame, array, factor, list) 2.3) Level of Measurement 2.4) Univariate and Multivariate Data 2.5) Data representation with example
3. Data Management and Data Quality 3.1) Data Cleaning 3.2) Data Transformation 3.3) Data Sampling 3.4) Data Sub-setting and manipulating
4. Descriptive Analytics and Visualizations 4.1) Central tendency measures 4.2) Dispersion measure 4.3) Data visualization tools
5. Machine Learning Concept and Techniques 5.1) Concept of learning 5.2) Concept of Training and Testing 5.3) Cross Validation 5.4) Supervised and Unsupervised Learning

6. Machine Learning – Classification & Regression 6.1) Decision Tree 6.2) Random Forest 6.3) Logistic Regression 6.4) Naïve Bayes 6.5) Support Vector Machine 6.6) Neural Networks
7. Machine Learning – Clustering 7.1) K-Mean 7.2) K-Nearest Neighbor
8. Predictive Analytics for Unstructured Data 8.1) Handling unstructured data 8.2) Emerging technologies of data analytics

Assessment Breakdown	%
Continuous Assessment	100.00%

Details of Continuous Assessment

Assessment Type	Assessment Description	% of Total Mark	CLO
Case Study	n/a	20%	CLO4
Group Project	n/a	10%	CLO1
Group Project	project	20%	CLO4
Lab Exercise	n/a	10%	CLO2
Lab Exercise	lab assignment	20%	CLO3
Test	Test	20%	CLO3

Reading List

Recommended Text

- **Jared Dean 2014, *Big Data, Data Mining, and Machine Learning*, John Wiley & Sons [ISBN: 1118618041]**
- **Jonathan S. Walker 2017, *Data Analytics for Beginners*, Createspace Independent Publishing Platform [ISBN: 1973962861]**

Article/Paper List

This Course does not have any article/paper resources

Other References

- **Book EMC Education Services 2015, *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data,* , Wiley**
- **Book Stephanie D. H. Evergreen 2016, *Effective Data Visualization*, Sage Publications**
- **Book Hofmann, M. & Klinkenberg 2014, *RapidMiner: Data Mining Use Cases and Business Analytics Applications (Chapman & Hall/CRC Data Mining and Knowledge Discovery Series* http://ISBN: 9781482205497**