THE CYPRUS INTERNATIONAL INSTITUTE OF MANAGEMENT

COURSE UNIT DESCRIPTION

Course Unit Title	**Data Mining, Visualization and Decision Making**
Course Unit Code	BI410
Type of Unit	Elective
Level of Course Unit	First cycle
Year of Study	First
Number of ECTS Credits	6.0 ECTS
Class Contact Hours	28
Minimum Learning Effort (In Hours)	150
Course Unit Objectives	Data is considered as the oil of the 21st century and precisely for this reason many companies and organizations are investing a lot of resources on data mining/analytics techniques in order to discover non-trivial patterns or correlations hidden in the data they collect. These findings could be used to enhance their decision making capabilities and improve their business operations. In this course, students will be taught the state-of-the-art techniques applied in data science for mining, analysis, visualization and interpretation of data. Both statistical and machine-learning based techniques will be taught with emphasis on the application of programmable solutions, visualization, interpretation and communication of the results obtained from the application of such techniques. In addition, the students will understand the uncertainty hidden in their results due to the probabilistic nature of the statistical and machine-learning techniques.
Learning Outcomes	The students completing the course should be able to

Learning Outcomes		
	CILO 1	Understand and appreciate the value of data and data-driven decision making via data mining/analytics and how this improves business decisions.
	CILO 2	Demonstrate ability in data mining techniques in order to individuate regularities, discover anomalies, correlations and patters in complex datasets that suit business applications.
	CILO 3	Understand fundamental machine-learning techniques such as supervise, unsupervised and semi-supervised learning and demonstrate ability to apply techniques such as clustering and classification that suit business applications.
	CILO 4	Visualize effectively the results obtained from data analysis using a scripting programming language such as Python.
	CILO 5	Understand the uncertainty hidden behind the results of a statistical analysis by understanding the associated reliability metrics; accuracy, false positive/negative rates, true positive/negative rates, recall, precision.
	CILO 6	Increase their capabilities as managers to think in a more statistical and data-driven way and acquire skills to provide leadership in statistical methods for the staff in their area of responsibility.
Name of Lecturer(s)		

Mode of delivery	Face to Face	
Prerequisites or corequisites	BI 420 Programming for Business Analytics	
	1. Introduction to data science and to data lifecycle: From collection to decision making through analysis, visualization and interpretation.	CILO 1,6
	2. Introduction to data mining techniques and the importance of data quality.	CILO 1,2
	3. Machine Learning Algorithms using Python (scikit-learn): supervised, semi-supervised and unsupervised techniques.	CILO 2,3
	4. Clustering algorithms using Python (scikit- learn): k-means	CILO 3
	5. Classification techniques using Python (scikit- learn): binary classification, logistic regression, decision trees, random forests and neural networks	CILO 3
	6. Interpretation and Visualization Techniques in Data Science using Python libraries (matplotlib, ggplot, pygal): boxplots, histograms, scatter plots, network plots, treemap, streamgraph, heatmap.	CILO 4,5,6
	7. Decision making and communicating with uncertainty: Understanding of probabilistic models and reliability metrics false positives, false negatives, precision, recall, ROC curve.	CILO 5,6
Recommended or required reading	**Required Reading:** 1. Sebastian Raschka. *Python Machine Learning*. Packt Publishing, 2015. 2. Simon Rogers and Mark Girolami. *A first course in machine learning*. CRC Press, 2011. **Recommended Reading:** Textbooks 3. Jared Dean. *Big Data, Data Mining: Value Creation for Business Leaders and Practitioners*. Wiley, 2014. 4. Wayne Winston, Christian Albright and Christopher Zappe. *Data Analysis and Decision Making*. Cengage Learning (4th Edition), 2009. Research Articles 5. Jeffrey Steinhoff and Terry Carnahan. *Smart Use of Data Mining is Good Business and Good Government*.	

	Journal of Governmental Financial Management, Spring 2012, Vol. 61 Issue 1, p 16-22, 2012.	
	6. Peter Mouncey. *Creating Value with Big Data analytics: making smarter marketing decisions.* International Journal of Market Research, Vol 58(5), 2016.	
	7. Carl Carande, Paul Lipinski and Traci Gusher. *How to Integrate Data and Analytics into Every Part of your organization.* Harvard Business Review, 2017.	
	8. Steve Lavalle, Eric Lesser, Rebecca Shockley, Michael Hopkins and Nina Kruschwitz. *Big data, analytics and the path from insights to value.* MIT Sloan Management Review 52(2), 2011.	
Planned learning activities and teaching methods	lectures, group work, lab work, role playing, project-based learning, homework	
Assessment methods and criteria	Class participation: 10% Group Assignments & Class Participation: 40% In-class examination: 50%	
Language of Instruction	English	
Work Placement(s)	Not applicable	