Indian Institute of Management, Udaipur
Post-Graduate Programme in Management
Year 2016-17

Course Outline for Predictive Analytics I

Credit : 4 **Term: IV**

Instructor : Professor V. Nagadevara

Email :

Course Objectives: This course aims to introduce participants with some of the modern methods for data mining, with their applications in text mining, social media and network analysis, and web mining.

Textbooks:

1. "Data Mining for Business Intelligence" by G. Shmueli, N. Patel and P. Bruce, Wiley. (Referenced as SPB)
2. "Web Data Mining: Exploring Hyperlinks, Contents and Usage data (Second Edition)", by Bing Liu, Springer. (Referenced as BL)
3. A number of other books/online notes/videos may be referenced throughout the course.

Course Structure: Presentation will be case based, and learning will be hands on.

Class coverage: Given below is a tentative schedule of topics which will be introduced with practical applications.

Date	Topics to be covered	Reading
Session 1	Introduction to data mining (discussions on structured and unstructured data, supervised and unsupervised learning methods)	SPB: Chapter 2 BL: Chapter 6
Session 2	Introduction to softwares to be	

	used in the subsequent sessions: for example, R, XLMiner, Weka (for text mining)	
Session 3	Data visualization	SPB: Chapter 3
Session 4	Introduction to text mining; Overview of different techniques used in text mining – Decision trees, Naïve Bayes classifier, k-nearest neighbor, Support vector machines, Clustering; Classifier evaluation	BL: Chapter 6
Session 5	Naïve Bayes method for classification	BL: Chapters 3.6 and 3.7
Session 6	k-nearest neighbor algorithm	SPB: Chapter 7
Session 7	Support vector machines	BL: Chapter 3.8
Session 8	Ensemble of classifiers: Bagging and Boosting	
Session 9	Cluster analysis – Data standardization, Hierarchical clustering	SPB: Chapter 12 BL: Chapters 4.1 to 4.4
Session 10	K-means clustering; Cluster evaluation	SPB: Chapter 12 BL: Chapter 4.1 to 4.4
Session 11	Text mining: text and web-page preprocessing	BL: Chapter 6.5
Session 12	Latent semantic indexing	BL: Chapter 6.7
Session 13	Opinion mining and sentiment analysis: Document sentiment classification, Aspect-based opinion mining	BL: Chapter 11.1 to 11.5

Session 14	Opinion mining and sentiment analysis: Mining comparative opinions; Further topics	BL: Chapter 11.6
Session 15	Social network analysis: Co-citation and bibliographic coupling, PageRank	BL: Chapter 7.1 to 7.3
Session 16	Social media and network analysis: HITS algorithm, Community discovery	BL: Chapter 7.4 and 7.5
Session 17	Introduction to web-mining, Web structure mining	Hand-out, BL: Chapter 8
Session 18	Web content mining	BL: Chapters 9 and 10
Session 19	Web usage mining	BL: Chapter 12
Session 20	Project presentation	

Evaluation: The following will be the evaluation scheme