

<b>Code</b>	<b>M17VFB05E E</b>	<b>ECTS Credit</b>	<b>6</b>	<b>HUN Credit</b>	<b>6</b>
<b>Module Title:</b>	<b>BUSINESS INTELLIGENCE</b>				
<b>Term</b>	spring				
<b>Module Leader:</b>	<b>Dr. Ferenc Kruzslicz, PhD.</b>	<b>Sessions</b>			
<b>Telephone:</b>	+36-72/501-599, 23113	<b>E-mail:</b>	kruzslicz@tkk.pte.hu		
<b>Short Description:</b>	<p>Business Intelligence generally is about how decision making processes can be improved through various interaction with data. The course focuses on data analytics application of internal and external enterprise data sources. The course structure follows the steps of the industrial standard CRISP-DM methodology. After a concise introduction to the data warehouse and the multidimensional data model, the typical data mining tasks are detailed. The semester is dedicated to predictive modeling tools, followed by frequently used descriptive models. These issues are illustrated by real world examples with special emphasis given to Machine Learning. Related case studies are discussed and solved by recent data mining softwares as hands-on experimentation with algorithms.</p>				
<b>Sessions (weeks): 14 (7<sup>th</sup> week is break) Schedule is tentative and subject to change.</b>					
<b>1. week</b>	<i>RA: 1<sup>st</sup> part of BInDM chapter 1 &amp; BInDM chapter 12</i>				
	Introduction (BI Terminology, Big Data)				
<b>2. week</b>	<i>RA: 1<sup>st</sup> part of BInDM chapter 4 &amp; BInDM chapter 13</i>				
	Data (Data Quality, Data Modeling Primer, CRISP-DM)				
<b>3. week</b>	<i>RA: 2<sup>nd</sup> part of BInDM chapter 1 &amp; BInDM chapter 3</i>				
	Exploring Data (Visualization, Data Warehouses, OLAP Operations)				
<b>4. week</b>	<i>RA: 2<sup>nd</sup> part of BInDM chapter 4</i>				
	Data Analytics (Data Mining, Data Preprocessing, Data Cleaning)				
<b>5. week</b>	<i>RA: BInDM chapter 5</i>				
	Classification Model (Nearest-Neighbor Classifier, Decision Tree Classifiers, Decision Borders)				

<b>6. week</b>	<i>RA: 1<sup>st</sup> part of PAnDM chapter 4</i>
	Classification Evaluation (Validation, Performance Metrics, Overfitting)
<b>7. week</b>	<i>Break</i>
	No class this week
<b>8. week</b>	<i>RA: 2<sup>nd</sup> part of PAnDM chapter 4</i>
	Statistical Classification Bayesian Classifier, Rules based Classifiers
<b>9. week</b>	<i>RA: BInDM chapter 7</i>
	Machine Learning Artificial Neural Networks and Applications
<b>10. week</b>	<i>RA: BInDM chapter 9</i>
	Association analysis (Frequent Itemset, Association Rule Generation, Performance Measures)
<b>11. week</b>	<i>RA: BInDM chapter 8</i>
	Segmentation (Prototype, Hierarchical Clustering Methods)
<b>12. week</b>	<i>RA: PAnDM chapter 7</i>
	Segmentation Evaluation (Model evaluation, Density Based Clustering, DBSCAN and SOM)
<b>13. week</b>	<i>RA: BInDM chapter 10 &amp; BInDM chapter 11</i>
	Unstructured Data (Text Mining, Web Mining, Sentiment Analysis)
<b>14. week</b>	<i>Project presentations</i>
	Midterm test
<b>Rationale Including Aims:</b>	Achieving the course students will recognize data analytics aspects and requirements of business intelligence problems. They will be able to actively manage and efficiently participate in data mining projects. Additionally to understanding of popular data mining techniques students have knowledge to identify advantages and limitations of these tools. Based on their solid practical skills using data mining softwares students can perform powerful data analysis, or build and implement automated applications.

<p><b>Learning Outcomes:</b></p> <p><b>Knowledge</b></p>	<p>Differentiate traditional and data mining tasks.</p> <p>Recognize data warehouse and data mining functionalities.</p> <p>Able to follow knowledge discovery methodologies.</p> <p>Describe data mining primitives, languages, and system architectures</p> <p>Define current state and requirements in data mining research.</p> <p>Sense on security and social issues of data mining.</p> <p>Understand the foundation of machine learning and its various learning strategies.</p>
<p><b>Learning Outcomes:</b></p> <p><b>Skills</b></p>	<p>Apply data preprocessing techniques to improve data quality.</p> <p>Use data mining software and understand its process flow.</p> <p>Able to scale the known techniques on large databases.</p> <p>Perform standard analytics (classification, rule generation, anomaly detection).</p> <p>Match business problems to appropriate analytic tasks.</p>
<p><b>Teaching and Learning Strategies:</b></p>	<p>Theoretical foundations strictly follow the text book. Reading assessments are short quizzes used for measuring the preparation level. Presentations and other supplementary course materials are provided. In class the selected topic is discussed and small scale illustrative examples and exercises are solved. The corresponding computer lab exercises are simulated data mining projects of different fields and real data sets. Students have to read the case study and do the basic data exploratory steps in advance, and group work is allowed. Along with the course progresses less detailed instructions are given and the data analytic goals are getting more ambiguous. Midterm exams based on similar exam cases, where students have to understand and build a working model on computer, evaluate it, and interpret the results.</p>

<b>Assessment Scheme:</b>	<p>Reading assessment: 15%</p> <p>Group problem solving: 15%</p> <p>Midterm test: 35%</p> <p>Final exam: 35%</p> <p><b>Reading assessment (RA):</b> Each quiz is based on the weekly topic and available online. Before class reading tests contain 5 questions about the current book chapter. (Not available in exam course.)</p> <p><b>Group problem solving (GPS):</b> Groups of 3-4 students work on small case studies, and submit their solutions to 5 questions on a weekly basis. (Not available in exam course.)</p> <p><b>Midterm tests:</b> A practice focused exam case must be solved using the computer. Any kind of additional resources are allowed to use, except communication. Midterm resit opportunity is scheduled at the first week of the exam period.</p> <p><b>Final exam:</b> Students have to answer questions in written form from selected parts of the core materials of the course. The exam may contain end of chapter exercises.</p>
<b>Core Learning Materials:</b>	<p><b>[BInDM]</b> Anil Maheshwari: Business Intelligence and Data Mining (Big Data and Business Analytics) Business Expert Press 2014, (1<sup>st</sup> Edition) ISBN: 978-1631571206, pp. 180</p> <p><b>[PAnDM]</b> Bala Deshpande, Vijay Kotu: Predictive Analytics and Data Mining (Concepts and Practice with RapidMiner) Morgan Kaufmann Publishing 2014, (1<sup>st</sup> Edition) ISBN: 978-0128014608, pp. 446</p>
<b>Optional Learning Material:</b>	<p>Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Anuj Karpatne: Introduction to Data Mining, Pearson 2012, (2<sup>nd</sup> Edition) ISBN: 978-0-13-312890-1, pp. 792</p> <p>Markus Hofmann, Ralf Klinkenberg: RapidMiner: Data Mining Use Cases and Business Analytics Applications, Chapman &amp; Hall/CRC, 2013 ISBN: 1482205491, pp. 525</p>