Subject Description Form

Subject Code	COMP5152
Subject Title	Advanced Data Analytics
Credit Value	3
Level	5
Pre-requisite/ Co-requisite/ Exclusion	Nil (but some knowledge in machine learning, data analytics, and programming language is preferable)
Objectives	The objectives of this course are to: 1. get familiar with data analytic frameworks; 2. learn existing machine learning/data mining/data analysis algorithms and techniques; 3. differentiate various data types and build suitable models to perform analytics; 4. utilize existing software or write programs to explore various data analytics problems; 5. conduct a group project based on real-world data.
Intended Learning Outcomes	Upon completion of the subject, students will be able to: a. conduct an end-to-end data analytics project in various domains following general steps/framework for data analytics b. deeply understand various algorithms and techniques for data analytics and how they should be used to analyze data that they encounter in their research and/or job; c. apply advanced data analytical techniques to various applications such as social media analytics, financial analytics, news analytics, video and image analytics, data analytics, etc.; d. carry out in-depth analysis of the data encountered in sophisticated applications and/or research.
Subject Synopsis/ Indicative Syllabus	1. Introduction to data analytics, background, definitions, and terminologies. 2. General framework/steps to perform data analytics. Possible data types and characteristics of the data. 3. Statistical methods for data analysis which include linear and logistic regression, principal component analysis, independent component analysis, hypothesis testing, ANOVA, ARMA and ARIMA. 4. Data analytics involving special data types such as texts, sequential and genomic data, temporal, spatial, temporal-spatial data. 5. Advanced analytics techniques and how to apply them in data analytics: Convolutional networks, autoencoders, deep belief networks, recurrent neural networks, long short term memory, deep reinforcement learning, deep and restricted Boltzmann machines, generative adversarial network. 6. Programming languages and tools for data analytics. 7. Useful applications in financial analytics, news analytics, and social media analytics.

Teaching/Learning Methodology

Lectures teach students on the main concepts and methods of the course, together with comprehensive examples, and class questions/answers/discussions for easy understanding.

Tutorials and lab sessions offer the opportunity for students to review and consolidate the lecture and reference materials through exercises and also software tools.

Project assignments give students the opportunity to solve practical data analysis problems.

Written assignments help students to develop a solid foundation of data analytics.

Assessment Methods in Alignment with Intended Learning Outcomes

Specific assessment methods/tasks	% weighting	Intended subject learning outcomes to be assessed			
		a	b	c	d
1. Assignments	10		✓	✓	✓
2. Projects	25	✓			✓
3. Examination	65		✓	✓	✓
Total	100				

Assignment(s): assessment of the theoretic studies with respect to the understanding of the relevant subject matters including new concepts, algorithms and techniques by proving answers to the assignment questions.

Project: assessment of the ability for problem solving through real case studies and implementation of a prototype system for demonstration .

Examination: assessment of the overall performance by written report , oral presentation and exam or quiz.

Student Study Effort Expected

Class contact:	
▪ Lecture/Tutorial/Lab	39 Hrs.
Other student study effort:	
▪ Reading	40 Hrs.
▪ Prepare written and project and assignments, quizzes/tests, examination	26 Hrs.
Total student study effort	**105 Hrs.**

Reading List and References

Pandas: https://pandas.pydata.org/docs/getting_started/index.html#getting-started.

WesMcKinney Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython 2nd Edition.

Kaggle: https://www.kaggle.com/.

Aurelien Geron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems.