

Course Title	:	Business Analytics for e-Business
Course Code	:	CDS3008
Recommended Study Year	:	3
No. of Credits/Term	:	3
Mode of Tuition	:	Sectional Approach
Class Contact Hours	:	3 hours per week
Category in Major Prog.	:	Required
Prerequisite(s)	:	BUS1102 Statistics for Business
Co-requisite	:	Nil
Exclusion	:	Nil
Exemption Requirement	:	Nil

Brief Course Description:

E-Business is the conduct of business processes on the Internet. E-Business organizations need a huge amount of information and knowledge from all data sources to give one holistic view of business partners, customers, products, services, and processes. This allows managers to make informed decisions to achieve effective cooperation among partners, customer acquisition and retention, cross selling, up selling, customer lifetime value maximization, appropriate marketing strategies, products and services optimization, innovations, and business process reengineering. The rapid accumulation of various kinds of data has prompted managers to use business analytics to extract useful hidden knowledge to assist decision making and gain competitive advantage.

Aims:

This course provides the fundamental concepts of business analytics and its applications for e-Business. Students will participate in the workflow of the business analytics process and understand different business analytics methods. Topics will include problem understanding, data understanding, data curation, data preprocessing, clustering, classification, model evaluation, visualization tools, and model deployment. In particular, case studies will be given to show how to extract and present implicit, previously unknown, and potentially useful information from the results of the business analytics methods, such as k-means clustering, decision trees, and random forest. Software packages such as IBM SPSS Modeler, Microsoft Azure ML, or SAS Enterprise Miner, will be the major tools for the whole process.

Learning Outcomes (LOs):

Upon completion, students should be able to:

1. Recognize and describe the basic concepts of business analytics and its applications in e-Business
2. Explain the business analytics process
3. Describe and compare the characteristics of different business analytics methods
4. Apply the business analytics process and suitable business analytics methods for different applications
5. Evaluate and validate the discovered knowledge

Indicative Contents:

Introduction to business analytics and its applications in e-Business

- Overview of business analytics
- Overview of some applications in e-Business
- Motivation of using business analytics for e-Business

Business analytics Process

- Defining a study
- Data preparation
- Selecting methods and building models
- Evaluating and validating models
- Understanding the discovered models
- Applying the models

Business Analytics Methods and Their Characteristics

- Data preprocessing methods
 - Data visualization methods
 - Statistical methods
 - Decision trees
 - Random Forest
 - XGBoost
 - Clustering
 - Market basket analysis
-

Business Analytics Methods for E-Business

- Identify appropriate business analytics methods for different e-Business applications

Model Evaluation and Validation Procedure for E-Business Applications

- Introduction to evaluation metrics, cross-validation, nested cross-validation

Teaching Method:

The concepts and principles of business analytics and its applications in e-Business are covered in lectures. Business analytics software packages such as IBM SPSS Modeler, Microsoft Azure ML, or SAS Enterprise Miner, will be described during the laboratories. A case study will be carried out throughout the course to demonstrate the process of applying business analytics to process data and discover knowledge from a huge e-Business database with more than 100,000 records and hundreds of variables. Students are required to perform a group project or Service-Learning group project* to apply the concepts and principles covered in this course to induce useful and valid models from another e-Business database with many cases and variables. They are required to present their findings and the business implications obtained in carrying out the project.

Depending on the availability of a service-learning community partner, Service-Learning will be implemented in the course. When implemented, all students will undertake the service-learning approach.

Assessment:

Class Attendance and Participation	5%
Assignments	30%
Group Project* or Service-Learning (S-L) Group Project**	35%
<u>Examination</u>	<u>30%</u>
Total	100%

*** [Non-S-L mode] Group Project**

The class will be divided into groups. Each group is given a dataset for a business case study. Students are required to apply appropriate data analysis techniques to analyse different data types and provide relevant business recommendations. Students are also encouraged to collect additional data from external sources (e.g., the Internet) to supplement their analysis.

*** [S-L mode] S-L Group Project**

The class will be divided into groups. Each group is required to analyse the data provided by the assigned NGO or Social Enterprise in conjunction with the Office of Service-Learning (OSL). Through the S-L group project, students will have opportunities to learn and apply an analytical and practical approach to enhance the collaborating partners' business performance and produce constructive recommendations for the community. Consultants from the agency partner and OSL will be invited to attend the training workshops and consultation sessions to give feedback to student groups to enhance their learning.

Measurement of Learning Outcomes:

	Class Attendance and Participation	Assignments	Group Project / S-L Group Project	Examination
Recognize and describe the basic concepts of business analytics and its applications in e-Business	x			x
Explain the business analytics process	x			x
Describe and compare the characteristics of different business analytics methods		x	x	x
Apply the business analytics process and suitable business analytics methods for different applications		x	x	
Evaluate and validate the discovered knowledge		x	x	x

1. There are a number of classroom activities to evaluate if the students can recognize different practical applications and the whole business analytics process (LO1, LO2).
2. Assignments require students to demonstrate their understanding of different business analytics methods, apply the different methods, and evaluate and validate the discovered knowledge (LO3-L05).
3. The group project / S-L group project requires students to demonstrate their understanding of different business analytics methods and apply the methods provided in different software packages. The appropriate procedures and metrics should be used to evaluate and validate the knowledge induced by different algorithms (LO3-5).

4. The examination can evaluate if the students can recognize the basic concepts, different practical applications, and the whole business analytics process. Students are required to demonstrate their understanding of different business analytics techniques, evaluation methods, and validation procedure (LO1-3, LO5).

Required/Essential Readings:

1. Provost, F. and Fawcett, T., *Data Science for Business: What you need to know about data mining and data-analytic thinking*, O'Reilly Media, 2013.

Recommended/Supplementary Readings:

1. Baesens, B., *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*, Wiley, 2014.
2. Bartlett, R., *A Practitioner's Guide to Business Analytics: Using Data Analysis Tools to Improve Your Organization's Decision Making and Strategy*, McGraw-Hill, 2013.
3. Foreman, J. W., *Data Smart: Using Data Science to Transform Information into Insight*, Wiley, 2013.
4. Han, J., Kamber, M., and Pei, J., *Data Mining: Concepts and Techniques, 3rd Edition*, Morgan Kauffmann, 2011.
5. Hansen, D., Shneiderman, B., and Smith, M. A., *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*, Morgan Kauffmann, 2010.
6. Howson, C., *Successful Business Intelligence: Unlock the Value of BI & Big Data*, McGraw-Hill, 2013.
7. Linoff, G. S. and Berry, M., *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, 3rd Edition*, Wiley, 2011.
8. Maisel, L. and Cokins, G., *Predictive Business Analytics: Forward Looking Capabilities to Improve Business Performance*, Wiley, 2014.
9. Siegel, E. and Davenport, T. H., *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die, 2nd Edition*, Wiley, 2016.
10. Tan, Pang-Ning, Steinbach, Michael, and Kumar, Vipin. *Introduction to Data Mining*, Pearson Education Dorling Kindersley, 2016.
11. Witten, I. H. and Frank, E., *Data Mining: Practical Machine Learning Tools and Techniques, 4th Edition*, Morgan Kauffmann, 2016.

Important Notes:

1. Students are expected to spend a total of 9 hours (i.e. 3 hours of class contact and 6 hours of personal study) per week to achieve the course learning outcomes.
2. Students shall be aware of the University regulations about dishonest practice in course work, tests and examinations, and the possible consequences as stipulated in the Regulations Governing University Examinations. In particular, plagiarism, being a kind of dishonest practice, is "the presentation of another person's work without proper

acknowledgement of the source, including exact phrases, or summarised ideas, or even footnotes/citations, whether protected by copyright or not, as the student's own work". Students are required to strictly follow university regulations governing academic integrity and honesty.

3. Students are required to submit writing assignment(s) using Turnitin.
4. To enhance students' understanding of plagiarism, a mini-course "Online Tutorial on Plagiarism Awareness" is available on <https://pla.ln.edu.hk/>.

Prepared by Prof. Wendy Wan Yee HUI

Rubric for Final Examination of CDS3008 - Business Analytics for e-Business

(Passing mark = 3)

Criteria	Very good (5-6)	Satisfactory (3-4)	Unsatisfactory (0-2)
Recognize and describe the basic concepts of business analytics and its applications in e-Business	Student demonstrates a clear understanding of the basic concepts and its practical applications in e-Business. Student can elaborate nearly all concepts and practical applications. The elaborations are completely / nearly completely correct and precise.	Student demonstrates a reasonable understanding of the basic concepts and its practical applications in e-Business. Student can elaborate most concepts and practical applications. Some elaborations are not completely correct and precise.	Student demonstrates limited understanding of the basic concepts and its practical applications in e-Business. Student can only elaborate a limited number of practical applications.
Describe the whole business analytics process	Student can describe all/most procedures in the whole process correctly.	Student can describe many procedures in the whole process correctly.	Student can only describe a limited number of procedures in the whole process correctly.
Recognize different business analytics techniques and their characteristics	Student gives correct descriptions of most business analytics techniques.	Student gives correct descriptions of many business analytics techniques.	Student gives correct descriptions of few business analytics techniques.
Evaluate and validate the discovered knowledge	<p>All of the following points are achieved:</p> <ul style="list-style-type: none"> ▪ Employ the right validation procedure(s). ▪ Use the right evaluation metrics. ▪ Calculate all measurement values correctly. 	<p>Two of the following points are achieved:</p> <ul style="list-style-type: none"> ▪ Employ the right validation procedure(s). ▪ Use the right evaluation metrics. ▪ Calculate all measurement values correctly. 	<p>None or One of the following tasks are achieved</p> <ul style="list-style-type: none"> ▪ Employ the right validation procedure(s). ▪ Use the right evaluation metrics. ▪ Calculate all measurement values correctly.
Presentation	Content of submission is well presented with very few grammatical mistakes.	Content of submission is satisfactorily presented with no more than a few grammatical mistakes.	Content of submission is not properly presented and/or there are more than a few grammatical mistakes.

**Rubric for Individual Assignments of CDS3008 – Business Analytics for e-Business
(Passing mark = 3)**

Criteria	Very good (5-6)	Satisfactory (3-4)	Unsatisfactory (0-2)
Able to load the datasets, assign feature characteristics, handle missing and noisy data	<p>All of the following points are achieved:</p> <ul style="list-style-type: none"> ▪ Load large datasets. ▪ Assign the right characteristics to all features. ▪ Handle all missing and noisy data correctly. 	<p>Two of the following points are achieved:</p> <ul style="list-style-type: none"> ▪ Load large datasets. ▪ Assign the right characteristics to all features. ▪ Handle all missing and noisy data correctly. 	<p>None or One of the following points are achieved:</p> <ul style="list-style-type: none"> ▪ Load large datasets. ▪ Assign the right characteristics to all features. ▪ Handle all missing and noisy data correctly.
Able to apply business analytics techniques to select the features	Can use the appropriate feature selection methods to select all/ almost all of the relevant features and remove most of the irrelevant.	A small number of the relevant features have not been selected or a small number of the irrelevant features have not been removed.	Quite a few of the relevant features have not been selected or quite a few of the irrelevant features have not been removed.
Can employ business analytics techniques to perform advanced feature transformations	Appropriate feature transformation methods have been performed. Suitable new features have been generated. The settings of the methods are completely correct.	Appropriate feature transformation methods have been performed, but the transformed features have some minor issues, because the settings of the methods are not completely correct.	Feature transformations have not been performed or most feature transformations have been done incorrectly.
Can perform data visualization	Appropriate data visualization methods have been performed. Suitable visualization results have been obtained. The settings of the methods are completely correct.	Appropriate data visualization methods have been performed, but the results have some minor issues, because the settings of the methods are not completely correct.	Data visualization has not been performed or most data visualizations have been done incorrectly.
Recognize different techniques for various data	Appropriate business analytics methods have been used for all datasets.	Appropriate business analytics methods have been used for most datasets.	Inappropriate business analytics methods have been used for most datasets.
Able to use business analytics techniques to learn models from data	The procedure of using the business analytics methods and the settings of these methods are completely correct. The learnt models are accurate.	The procedure of using the business analytics methods and the settings of these methods are not completely correct. Thus the learnt models have some issues.	The learnt models are incorrect.
Able to use the right techniques to evaluate the performance of the learnt models	Appropriate methods (e.g. cross-validation, nested cross-validation) have been applied to divide the dataset into training, testing and/or validation data sets. Learn and validate the models using the training and/or the validation datasets. Apply the suitable methods to evaluate the performance of the model on the testing dataset.	The dataset has been divided into training, testing and/or validation data sets, but the splitting ratio or methods used may not be appropriate. Methods used to learn and validate the models are generally correct. Performance evaluation strategy is generally correct. But the evaluation results are somewhat biased or unstable.	The dataset has not been correctly divided into training, testing and/or validation data sets. Methods used to learn and validate the models are incorrect. Performance evaluation has not been done or the evaluation strategy is incorrect.

Rubric for Group Project or S-L Group Project Presentation of CDS3008 - Business

Analytics for e-Business

(Passing mark = 3)

Criteria	Very good (5-6)	Satisfactory (3-4)	Unsatisfactory (0-2)
Appropriate time allocation and pace	Allocate time appropriately, and managed time effectively, with smooth progression. Appropriate pace. Start presentation punctually.	Marginally long or marginally short but uses time reasonably effectively. Reasonable pace. Start presentation punctually.	Significantly too short or too long and does not use time effectively. Pace is significantly too fast or too slow. Does not start presentation punctually.
Clear, logically organized and relevant content	Information included is always relevant. Clearly stated and developed points. Material flows extremely well and is well-organized. No ambiguities are left unexplained.	Information included is generally relevant. Key points are relatively clear. Most information presented in logical sequence; sufficiently well-organized with generally satisfactory flow. Some ambiguities are left unexplained.	Much of the information included is not relevant and even key points are not clear. Presentation is choppy or disjointed, does not flow well, or has no apparent logical order.
Effective use of presentation tools	Balanced and proper use of presentation tools with little or no distraction (e.g., appropriate animation/pictures, appropriate information on each slide, good color combination, clear titles, etc.)	Generally good use of presentation tools. Some distractions but they are not overwhelming (e.g., reasonable animation/pictures, fair information on each slide, fair color combination, fair titles, etc.)	Poor use of presentation tools and/or many distractions (e.g., too much animation/pictures, too much information on a slide, poor color combination, absence of titles, etc.)
Uses good body language, eye contact, appropriate voice tone	Shows poise and composure; makes good eye contact with audience; balanced posture; shows enthusiasm and confidence; uses voice tone effectively.	Fairly poised and composed; makes fairly good eye contact with audience; balanced posture; shows some enthusiasm and confidence; uses voice tone relatively effectively.	Little poise and composure; makes little or no eye contact with audience; poor posture; shows little or no enthusiasm and confidence; uses voice tone ineffectively or too monotone.
Gains/holds attention	Provides good motivation to engage the audience's interest. Presents the content in a manner that captivates the audience's attention.	Provides reasonable motivation to engage the audience's interest. Audience is reasonably engaged but there are instances where the presentation is otherwise dull.	Provides insufficient motivation to engage the audience's interest. Dull presentation of content that does not engage the audience.
Uses instructor defined role appropriate dress	Professionally dressed as expected by the instructor.	Minor deviations from instructor's expectations.	Do not dress in a manner expected by the instructor.
Clarity of speech/Accuracy of grammar & pronunciation	Voice is consistently comprehensible; grammar and pronunciation are accurate.	Voice is generally comprehensible; grammar and pronunciation are adequate but with some mistakes.	Voice is incomprehensible on several occasions; many mistakes in terms of grammar and pronunciation.

Rubric for Group Project or S-L Group Project of CDS3008 - Business Analytics for e-Business

(Passing mark = 3)

Criteria	Very good (5-6)	Satisfactory (3-4)	Unsatisfactory (0-2)
Problem definition (demonstrate the understanding of the problem and consider alternative solutions)	Clearly state the problem, list out related constraints, and be able to identify alternative solutions.	The problem is stated but related constraints and alternative solutions are not considered thoroughly.	The problem is marginally defined and with no consideration of constraints and alternative solutions.
Solution design	The design of the solution is strongly related to the problem. Clearly explains the approaches and techniques involved. The design of the solution is very innovative.	The design of the solution is related to the problem. Does not clearly explain the approach and techniques involved. The design of the solution is not very innovative.	The design of the solution is weakly related to the problem and no explanation of the approach and technique provided. The design of the solution is not innovative.
Able to collect datasets, load the datasets, assign feature characteristics, handle missing and noisy data	All of the following points are achieved: <ul style="list-style-type: none"> ▪ Collect datasets. ▪ Load large datasets. ▪ Assign the right characteristics to all features. ▪ Handle all missing and noisy data correctly. 	Two of the following points are achieved: <ul style="list-style-type: none"> ▪ Collect datasets. ▪ Load large datasets, ▪ Assign the right characteristics to all features, ▪ Handle all missing and noisy data correctly. 	None or One of the following points are achieved <ul style="list-style-type: none"> ▪ Collect datasets. ▪ Load large datasets. ▪ Assign the right characteristics to all features. ▪ Handle all missing and noisy data correctly.
Able to apply business analytics techniques to select the features	Can use the appropriate feature selection methods to select all/ almost all of the relevant features and remove most of the irrelevant.	A small number of the relevant features have not been selected or a small number of the irrelevant features have not been removed.	Quite a few of the relevant features have not been selected or quite a few of the irrelevant features have not been removed.
Can employ business analytics techniques to perform advanced feature transformations	Appropriate feature transformation methods have been performed. Suitable new features have been generated. The settings of the methods are completely correct.	Appropriate feature transformation methods have been performed, but the transformed features have some minor issues, because the settings of the methods are not completely correct.	Feature transformations have not been performed or most feature transformations have been done incorrectly.
Can perform data visualization	Appropriate data visualization methods have been performed. Suitable visualization results have been obtained. The settings of the methods are completely correct.	Appropriate data visualization methods have been performed, but the results have some minor issues, because the settings of the methods are not completely correct.	Data visualization has not been performed or most data visualizations have been done incorrectly.
Apply different business analytics	Appropriate business analytics methods have been used for all datasets.	Appropriate business analytics methods have been used for most datasets.	Inappropriate business analytics methods have been used for most datasets.

Criteria	Very good (5-6)	Satisfactory (3-4)	Unsatisfactory (0-2)
techniques for various data			
Able to use business analytics techniques to learn models from data	The procedure of using the business analytics methods and the settings of these methods are completely correct. The learnt models are accurate.	The procedure of using the business analytics methods and the settings of these methods are not completely correct. Thus the learnt models have some issues.	The learnt models are incorrect.
Able to use the right techniques to evaluate the performance of the learnt models	Appropriate methods (e.g. cross-validation, nested cross-validation) have been applied to divide the dataset into training, testing and/or validation data sets. Learn and validate the models using the training and/or the validation datasets. Apply the suitable methods to evaluate the performance of the model on the testing dataset.	The dataset has been divided into training, testing and/or validation data sets, but the splitting ratio or methods used may not be appropriate. Methods used to learn and validate the models are generally correct. Performance evaluation strategy is generally correct. But the evaluation results are somewhat biased or unstable.	The dataset has not been correctly divided into training, testing and/or validation data sets. Methods used to learn and validate the models are incorrect. Performance evaluation has not been done or the evaluation strategy is incorrect.
Able to interpret the results	The interpretation is completely correct and insightful. Critical evaluation of the limitations is correct.	There are some issues in the interpretation of the results. Critical evaluation of the limitations is not completely correct.	The interpretation of the results is incorrect. Critical evaluation of the limitations is not provided.

Rubric for Class Attendance and Participation of CDS3008 - Business Analytics for e-Business

(Passing mark = 3)

Criteria	Very good (5-6)	Satisfactory (3-4)	Unsatisfactory (0-2)
Attendance	Student attends almost all of the classes.	Student attends many of the classes.	Student does not attend most of the classes.
Punctuality	Student is always or almost always punctual.	Student is usually punctual.	Student is usually not punctual.
Recognize the practical applications of business analytics in e-Business	Student can describe nearly all practical applications.	Student can describe most practical applications. Some descriptions are not completely correct and precise.	Student can only describe a few practical applications.
Describe the whole business analytics process	Student can describe all procedures in the whole process correctly.	Student can describe most procedures in the whole process correctly.	Student can only describe a few procedures in the whole process correctly.