Annexe A: New/Revised Course Content in OBTL+ Format

Course Overview

The sections shown on this interface are based on the templates UG OBTL+ or PG OBTL+
If you are revising/duplicating an existing course and do not see the pre-filled contents you expect in the subsequent sections e.g. Course Aims, Intended Learning Outcomes etc. please refer to Data Transformation Status for more information.

Expected Implementation in Academic Year	AY2019-2020
Semester/Trimester/Others (specify approx. Start/End date)	Semester 1
Course Author * Faculty proposing/revising the course	Lee-Chua Lee Hong
Course Author Email	clhlee@ntu.edu.sg
Course Title	Introduction to Data Science and Artificial Intelligence
Course Code	CV0003
Academic Units	3
Contact Hours	39
Research Experience Components	Not Applicable

Course Requisites (if applicable)

Pre-requisites	CV1014 Introduction to Computation Thinking
Co-requisites	
Pre-requisite to	
Mutually exclusive to	
Replacement course to	
Remarks (if any)	

Course Aims

In today's era of Information, 'Data' is the new driving force, provided we know how to extract relevant 'Intelligence'. This course will start with the core principles of Data Science, and will equip you with the basic tool and techniques of data handling, exploratory data analysis, data visualization, data-based inference, and data-focussed communication. The course will also introduce you to the fundamentals of Artificial Intelligence – state space representation, uninformed search, and reinforcement learning.
The course will motivate you to work closely with data and make data-driven decisions in your field of study. The course will also touch upon ethical issues in Data Science and Artificial Intelligence, and motivate you to explore the cutting-edge applications related to Big Data, Neural Networks and Deep Learning. Python will be the language of choice to introduce hands-on computational techniques.

Course's Intended Learning Outcomes (ILOs)

Upon the successful completion of this course, you (student) would be able to:

ILO 1	Identify and define data-oriented problems and data-driven decisions in real life
ILO 2	Illustrate the problems in terms of data exploration and visualization
ILO 3	Apply basic machine learning tools to extract inferential information from the data
ILO 4	Compose an engaging “data-story” to communicate the problem and the inference
ILO 5	Outline the roles and requirements of artificial intelligence in practical application
ILO 6	Explain and discuss fundamentals of state space search and reinforcement learning

Course Content

	Topics	LAMS/TEL (Hours)	Example Classes (2-Hour Sessions)
1	**Data-Analytic Thinking** What is Data Science? – The core problems and solutions. Extracting Intelligence from Data – formulating problems.	1	Problem Formulation, Data Wrangling, Cleaning and Preparation (2 weeks)
2	**The Data Pipeline** Types of Data in various practical Data Science scenarios. Data Wrangling, Cleaning and Preparation using Python.	1	
3	**Data Presentation** Basic concepts in Statistics and Exploratory Data Analysis. Data Exploration and Data Visualization using Python. Case Studies involving Structured and Unstructured Data	2	Basic Statistics, Data Exploration and Visualization (2 weeks)
4	**Data-driven Inference** Basics of Machine Learning : Prediction and Classification. Prediction and Classification techniques using Scikit-Learn.	2	Prediction and Classification (2 weeks)

5	**Data-driven Identification** Basics of Machine Learning : Clustering and Anomalies. Clustering and Anomaly Detection using Scikit-Learn.	1	Clustering and Anomaly Detection (1 week)
6	**Digital Storytelling** Data-driven Dashboards, Websites and Presentations. Data Presentation using Python Notebooks and Plotly.	1	Data Presentation and Dashboards (1 week)
7	**Artificial Intelligence** What is Artificial Intelligence? – History and State-of-Art. Principles of problem solving and the State Space Search. Case Studies for State Space Search and Search Algorithms	2	State Space Search and misc. Search Algorithms (2 weeks)
8	**Reinforcement Learning and AI** Introduction to Reinforcement Learning in context of AI. Fundamentals of Markov Processes and Q-Learning.	2	Markov Processes and Q-Learning (2 weeks)
9	**Ethics in DS&AI** Ethical considerations and the idea of responsible DS&AI.	0.5	Ethical Data Science and AI (1 week)
10	**State-of-the-Art in DS&AI** Progress in Big Data, Neural Networks and Deep Learning.	0.5	
	Check for Hours	= 13	= 26

Reading and

References (if applicable)

There is no single textbook for the course. The following books and resources will be used as references.

1. Python Data Science Handbook : Jake VanderPlas : O'Reilly (1st edition)
2. An Introduction to Statistical Learning : James, Witten, Hastie, Tibshirani
3. Artificial Intelligence: A Modern Approach : Russell and Norvig (3rd edition)

Additional resources, if required, will be shared with you in the LAMS/TEL videos and Example Classes.

Planned Schedule

Week or Session	Topics or Themes	ILO	Readings	Delivery Mode	Activities
1	Data-Analytic Thinking What is Data Science? – The core problems and solutions. Extracting Intelligence – formulating problems.	1, 2	Online Video (LAMS/TEL)	In-person	Defining a Data Science Problem in real-life. Familiarization with Python tools for DS.
2	The Data Pipeline Types of Data in various practical Data Science scenarios. Data Wrangling, Cleaning, Preparation.	1, 2	Online Video (LAMS/TEL)	In-person	Extraction, Wrangling, Cleaning, Preparation of Data using Pandas.
3	Data Exploration Basic concepts in Statistics and Exploratory Data Analysis.	1, 2	Online Video (LAMS/TEL)	In-person	EDA using Case Studies involving Structured and Unstructured Data
4	Data Presentation Data Exploration and Data Visualization using Python.	2, 4	Online Video (LAMS/TEL)	In-person	Visualization tools in Python and the basics of Data Visualization
5	Data-driven Predictions Prediction using techniques of Regression and Time Series	2, 3	Online Video (LAMS/TEL)	In-person	Using Prediction tools from Scikit-Learn.

Week or Session	Topics or Themes	ILO	Readings	Delivery Mode	Activities
6	Data-driven Classification Classification using techniques of Decision Trees and Support Vectors	2, 3	Online Video (LAMS/TEL)	In-person	Using Classification tools from Scikit-Learn.
7	Data-driven Identification Clustering and Anomaly Detection.	2, 3	Online Video (LAMS/TEL)	In-person	Using Clustering tools from Scikit-Learn.
8	Digital Storytelling Data-driven Dashboards, Websites and Presentations.	2, 4	Online Video (LAMS/TEL)	In-person	Data Presentation using Notebooks and Plotly.
9	Artificial Intelligence What is Artificial Intelligence? – History and State-of-Art. Principles of problem solving and State Space.	5, 6	Online Video (LAMS/TEL)	In-person	Case Studies for State Space Search and Search Algorithms
10	Uninformed Search Search Algorithms : breadth-first, depth-first, IDA, uniform-cost.	5, 6	Online Video (LAMS/TEL)	In-person	Case Studies for State Space Search and Search Algorithms

Week or Session	Topics or Themes	ILO	Readings	Delivery Mode	Activities
11	Reinforcement Learning Introduction to Reinforcement Learning in context of AI. Basics of Markov Processes and Q-Learning.	5, 6	Online Video (LAMS/TEL)	In-person	Case Studies for Reinforcement Learning
12	Reinforcement Learning Introduction to Reinforcement Learning in context of AI. Basics of Markov Processes and Q-Learning.	5, 6	Online Video (LAMS/TEL)	In-person	Case Studies for Reinforcement Learning
13	Ethics and State-of-the-Art Ethical considerations and the idea of responsible DS&AI. Progress in Big Data, Neural Net, Deep Learning.	1, 5	Online Video (LAMS/TEL)	In-person	Ethical considerations and the idea of responsible DS&AI.

Learning and Teaching Approach

Approach	How does this approach support you in achieving the learning outcomes?
LAMS/ TEL (Online Video)	Topics will be delivered as a series of online videos lectures, and you will also be provided reference materials for self-study to achieve the ILOs.
Example Class (Face-to-Face)	Example Classes will be used for seminar sessions for students to discuss, debate and clarify the contents of the online LAMS/TEL contents, as well as hands-on sessions to equip students with practical knowledge on data science, machine learning and artificial intelligence, and to guide in terms of the design and implementation of a mini project, to achieve the ILOs.

Assessment Structure

Assessment Components (includes both continuous and summative assessment)

No.	Component	ILO	Related PLO or Accreditation	Weightage	Team/Individual	Rubrics	Level of Understanding
1	Continuous Assessment (CA): Others(TEL participation and TEL MCQs)	1,5,6	a,b,h,l	10	Individual	Analytic	Multistructural
2	Continuous Assessment (CA): Test/Quiz(Quiz)	2,5,6	a,b,h	40	Individual	Holistic	Relational
3	Continuous Assessment (CA): Test/Quiz(Coding Assignment in Class)	3,4,6	a,b,c,d,e,h	20	Individual	Analytic	Multistructural
4	Continuous Assessment (CA): Project(Mini Project with Presentation)	1,2,3,4,5,6	a,b,c,d,e,f,i,j	30	Team	Analytic	Relational

Description of Assessment Components (if applicable)

For CA 4
- A Modification Factor (MF) will be applied to the 30% weightage. This is to account for individual contribution to the project work. The MF is derived from panel judges and peer assessment. For more details on the MF calculation, please see item 4 of the Rubric file attached

IMPORTANT NOTE :
** To pass this course, students are required to take the compulsory 40% Quiz Session listed in the Continuous Assessment 2. Details of the quiz date will be announced in the first teaching week.

Formative Feedback

TEL participation and TEL MCQs : This is an online exercise. You will see you scores, your answers, the correct answers, feedback on your incorrect answers, and explanations for the correct answers, immediately after you have submitted your answers online.

Coding Assignment : These are partially based on online exercises based on MCQs, and partially on classwork submissions. For the MCQs, you will see you scores, your answers, the correct answers, feedback on your incorrect answers, and explanations for the correct answers, immediately after you have submitted your answers online. For the classwork submissions, Individual feedback will be provided to you after proper evaluation of your submissions. The answers will be discussed in the class, and you will also get to know the basic score statistics of the other students in the same cohort.

Mini Project in Example Class : You will be guided in choosing the topic, and the instructor will also help you

during the course of the project, as and when required. Regular interactions with the instructor will be arranged to monitor your progress, and to provide you with constructive criticism.

NTU Graduate Attributes/Competency Mapping

This course intends to develop the following graduate attributes and competencies (maximum 5 most relevant)

Attributes/Competency	Level
Problem Solving	Advanced
Critical Thinking	Advanced
Embrace Challenge	Advanced

Course Policy

Policy (Academic Integrity)

Good academic work depends on honesty and ethical behaviour. The quality of your work as a student relies on adhering to the principles of academic integrity and to the NTU Honour Code, a set of values shared by the whole university community. Truth, Trust and Justice are at the core of NTU's shared values. As a student, it is important that you recognize your responsibilities in understanding and applying the principles of academic integrity in all the work you do at NTU. Not knowing what is involved in maintaining academic integrity does not excuse academic dishonesty. You need to actively equip yourself with strategies to avoid all forms of academic dishonesty, including plagiarism, academic fraud, collusion and cheating. If you are uncertain of the definitions of any of these terms, you should go to the academic integrity website for more information. On the use of technological tools (such as Generative AI tools), different courses / assignments have different intended learning outcomes. Students should refer to the specific assignment instructions on their use and requirements and/or consult your instructors on how you can use these tools to help your learning. Consult your instructor(s) if you need any clarification about the requirements of academic integrity in the course.

Policy (General)

Students are expected to take responsibility to follow up with course notes, assignments and course related announcements. You are also expected to participate in class discussion and submit the project before the stipulated deadline.

Policy (Absenteeism)

Valid reasons include falling sick supported by a medical certificate and participation in NTU's approved activities supported by an excuse letter from the relevant bodies. There will be no make-up opportunities for in-class activities.

Policy (Others, if applicable)

As a student of the course, you are required to abide by both the University Code of Conduct and the Student Code of Conduct. The Codes provide information on the responsibilities of all NTU students, as well as examples of misconduct and details about how students can report suspected misconduct. The University also has the Student Mental Health Policy. The Policy states the University's commitment to providing a supportive environment for the holistic development of students, including the improvement of mental health and wellbeing. These policies and codes concerning students can be found in the following link: http://www.ntu.edu.sg/SAO/Pages/Policies-concerning-students.aspx

Last Updated Date: 29-04-2025 09:21:00

Last Updated By: Yang, En-Hua

Rubrics for CV0003 Introduction to Data Science and Artificial

1. Assessment Criteria for CA1:TEL MCQs

You will complete 12 online LAMS/TEL sessions, including embedded MCQs. The maximum score is 10% of your total marks.

2. Assessment Criteria for CA2: QUIZ

You will take 1 quiz based on MCQs during the semester. The maximum score is 40% of your total marks.

3. Assessment Criteria for Coding Assignment in Class

You will take 1 hands-on Lab Quiz during the semester, based on the material covered during the Labs or the Example Classes. You will need to code for this quiz (at least major part of it), and the maximum score for the Lab Quiz is 20% of your total marks.

4. Assessment Criteria for Mini Project

Criteria	**Standards**		
	Fail standard (0-40 %)	**Pass standard (41-74 %)**	**High standard (75-100 %)**
Identify the core definition of the problem and plan the data-driven solution. (LO 1, 3, 5)	Identifying completely wrong definitions of the problems, and planning solutions that are somewhat related but are not the actual solutions expected for the problems.	Identifying the correct and relevant definitions of the problems in line with the course materials, planning solutions reasonably in line with solutions expected for the problems, and trying to relate the course materials to the planned solutions. Accuracy and clarity can be further improved.	Identifying the correct and relevant definitions of the problems in line with the course materials, planning technically accurate steps for the solutions that are expected for the problems, and clearly connecting the course materials to the planned solutions.
Explore the data effectively and devise required models to solve the problems. (LO 2, 3, 6)	Ad hoc analysis of the data and arbitrary steps in building the model without properly connecting the concepts with relevant concepts from the course. No or little evidence of critical	Logical exploration of the data that demonstrates a good understanding of the concepts from the course, and building models with reasonable accuracy to solve the problems. Reasonable evidence of critical thinking related to the proposed solution, and producing solutions with	Clear logical flow of data exploration of that demonstrates a good understanding of the concepts from the course (and beyond), and building models with high accuracy to solve the problems. Extensive evidence of critical

	evaluation of the proposed solution.	some degree of intuition and justification (rigorous steps for model-building or validation of models and results may be missing).	thinking related to the proposed solution, and producing solutions with clear intuition and proper justification, including rigorous steps for model-building and validation of the models and results.
Overall Editorial Standard of the Solution and the Final Report. (LO 4)	Disorganised format and arrangement of the code and report, without any comment or little/no mention of references /resources.	Clear logical flow and well-formatted arrangement of the code and report, with all essential components. Reasonable comments and reasonable documentation of references /resources.	Clear logical flow and well-formatted arrangement of the code and report, with all essential components. Detailed set of technical comments to illustrate the choices made towards the solution, and to highlight the inferences. Proper documentation of references /resources.

You will submit the code(s) for data analysis, the visualization dashboard, and a final report to illustrate the Mini Project – both the problem and the solution. Mini-Project will be graded out of 100 points, with 80 points for the Team Exercise (code, presentation, report) and 20 points for Individual contribution. The Individual contribution will be judged based on an Oral Evaluation after project presentation. The score for the Mini-Project, graded out of 100, will then be scaled down to 30% of your total marks.

Your Individual contribution (20 points out of 100) towards the Mini-Project will be judged based on an Oral Evaluation, as per the following rubrics.

Criteria	**Standards**		
	Fail standard (0-40 %)	**Pass standard (41-74 %)**	**High standard (75-100 %)**
Understanding of the Project and Individual Contribution. (LO 1, 2, 3)	Little understanding of problem definition, solution techniques, data exploration and machine learning tools used in the project. Individual contribution is too	Decent understanding of problem definition, solution techniques, data exploration and machine learning tools used in the project. Individual contribution to the	Clear understanding of problem definition, solution techniques, data exploration and machine learning tools used in the project. Individual contribution to the project is

	low compared to the team-mates.	project is proportional to the team size and project difficulty.	significantly high compared to team-mates.

5. Peer Evaluation of Contributions to the Team Project

	Performance Level/Criteria			
Performance Indicators	**Outstanding: 4**	**Good: 3**	**Average, meet expectation: 2**	**Below expectations: 1**
Collaborative behaviour	Cooperative and always delivered assigned tasks on time. Take initiative to help other to ensure success of team project.	Cooperative and always delivered assigned tasks on time. Willing to assist others upon request.	Stop short at delivering assigned tasks, sometimes after reminder(s).	Uncooperative, non-committed, always miss deadlines.
Quality of works	Quality of works higher than overall group quality, or go extra miles to assist teammate to enhance the quality of group works.	Good quality of deliverables under individual responsibility.	Acceptable quality of deliverables under individual responsibility.	Quality of works not acceptable.
Ideas & participations	Active participation and initiatives, good ideas & suggestions in enhancing the quality of group works.	Contributed suggestions and ideas to enhance the quality of group works.	Somewhat contributed in enhancing the quality of group works.	Did not participate in group works.

Average Peer Assessment Score	**MF**
3.51 to 4.00	1.05
2.76 to 3.50	1.00
2.51 to 2.75	0.95
2.00 – 2.50	0.9
Below 2.0	Separate Assessment

Peer assessment exercise will be anonymous and done towards the end of the semester.

For student who has average peer assessment score below 2.0, the coordinator might contact them, and/or contact any other group member(s) to further assess the appropriate MF.