Institute of Technology of Cambodia

Department of Applied Mathematics and Statistics

Syllabus of Course

Programming for Data Science

Year 2022-2023

Course: Programming for Data Science   Course's Code: AMSI51PDS

Semester: 1, Year: 5   Number of Credit: 3

Instructor's name: Mr. CHAN Sophal   Latest Degree Obtained: MSIT

Pre-requisites   Programing for DS (semester 2, year 3)

Tel: +855 95 201 434   Email: Sophal.chan@itc.edu.kh

I. Course Description

Programming for data science introduces student to techniques of science and machine learning with a focus on data analysis by using Python or R (optional) programming language. One new technique is covered every week, including: NumPy, pandas, data visualization, scikit learn, principal component analysis, hierarchical clustering, decision tree learning, neural networks, network science, agent-based modeling, text mining and image classification. The main assessment is a final paper where the students are asked to pick any data set (preferably from their own research) and apply one or multiple techniques from the course. No programming experience is required, but the course includes Python coding.

II. Course Learning Outcomes (CLOs)

Upon completing this course, students will be able to

No	Course Learning Outputs (CLOs)	PLOs	Bloom's taxonomy
CLO1	Understand how to apply programming (python) for data science field.	PLO1, PLO11, PLO16	Analyzing, Applying, Remembering
CLO2	Defined how to apply data science library such as: NumPy, pandas, scikit learn and etc. to data science project	PLO12, PLO15, PLO16	Analyzing, Applying, Remembering
CLO3	Understand how to collection in the data via API and use EDA for data collection and ML algorithm	PLO1, PLO11, PLO12PLO17	Analyzing, Applying, Remembering
CLO4	Apply rigorously one or multiple of these techniques learned in their own research	PLO1, PLO6, PLO15, PLO17	Analyzing, Applying, Remembering

III. Teaching Approaches

- Give Lecture
- Practice tutorial/programming
- Make group discussion and Presentation
- Give homework
- Give project/Assignment
- Have Quiz

IV. Assessment Policy

No	Assessment Task	Weighting (%)	Responded Course Learning Outcomes
1	Attendance	10	
2	Class Activities and Quiz	10	CLO1, CL02, CLO3, CLO4
3	Assignment/Report and Presentation	10	CLO1, CL02, CLO3, CLO4
4	Midterm	30	CLO1, CL02
5	Final	40	CL03, CLO4

V. Grading and Evaluation Criterion

In order to pass this subject, student need to

- Get total score of at least 30 (in 100) if the average score more than 50%
- Get total score of at least 50 (in 100) if the average score lower than 50%
- Otherwise will be judge by jury of the department.

VI. Details of Contents, CLOs, LLOs, Teaching and Learning Activities, Assessment, and Supported Materials

Weeks	Sessions	CLOs	LLOs	Content	Teaching Activities	Learning Activities	Assessment	Materials
1-2	1-8	CLO1	• Ability to remember the basic programing. • Ability to understand data structure. • Ability to write OOP programing • Ability to install python environment for data science. • Ability to use some useful library for data science	Chapter 1: basic python and recommended library	• Lecture Presentation • Tutorials Practice • Lab • Case Study • Q&A	• Listen and Note Taking • Indiviudal Homework • Group Discussion and Presentation • Coding/Project Presentation • Q&A	• Attendance • Class Activities • Formative Assement • Quiz and Q&A	• Slide Presentation • LCD Projector • Laptop • Slide Pointer • Lab
3-4	9-16	CLO2 CLO3	• Ability to best data science tool. • Ability to select the best IDE and website for data science process • Ability to use NumPy library • Ability to write complex problem with NumPy • Ability to analyze structure data • Ability to use data frame • Ability to write code with pandas for series, data frame and panel.	Chapter 2: Data science tool (jupyter) and NumPy, Pandas	• Lecture Presentation • Tutorials Practice • Lab • Case Study • Q&A	• Listen and Note Taking • Indiviudal Homework • Group Discussion and Presentation • Coding/Project Presentation • Q&A	• Attendance • Class Activities • Formative Assement • Quiz and Q&A	• Slide Presentation • LCD Projector • Laptop • Slide Pointer • Lab

4-6	17-24	CLO2 CLO3	• Ability to transform the data into visualization • Ability to illustrate the structure data into graph, chart and etc. • Ability to use power BI • Ability to use and write programing with matplotlib and etc. •	Chapter 3: Visualization	• Lecture Presentation • Tutorials Practice • Lab • Case Study • Q&A	• Listen and Note Taking • Indiviudal Homework • Group Discussion and Presentation • Coding/Project Presentation • Q&A	• Attendance • Class Activities • Formative Assement • Quiz and Q&A	• Slide Presentation • LCD Projector • Laptop • Slide Pointer • Lab
7-9	25-36	CLO2 CLO3	• Ability to transform the data into visualization • Ability to illustrate the structure data into graph, chart and etc. • Ability to use power BI • Ability to use and write programing with matplotlib and etc. • Can understand and programing the EDA task	Chapter 4: Data collection and Exploratory data analysis	• Lecture Presentation • Tutorials Practice • Lab • Case Study • Q&A •	• Listen and Note Taking • Indiviudal Homework • Group Discussion and Presentation • Coding/Project Presentation • Q&A •	• Attendance • Class Activities • Formative Assement • Quiz and Q&A	• Slide Presentation • LCD Projector • Laptop • Slide Pointer • Lab
10	37-40		Lecture Review Midterm Exam					
11-12	31-48	CLO3 CLO4	• Ability to analyze the feature of the data • Ability to write code for feature scoring • Ability to analyze and select the important feature to put in ML model	Chapter 5: Feature Engineering	• Lecture Presentation • Tutorials Practice • Class Facilitating • Case Study Q&A	• Listen and Note Taking • Indiviudal Homework • Group Discussion and Presentation • Report/Project Presentation Q&A	• Attendance • Class Activities • Formative Assement • Quiz and Q&A	• Slide Presentation • LCD Projector • Laptop • Slide Pointer • Writting Board

13-14	49-56	CLO3 CLO4	• Ability to understand the machine learn in term of unsupervised • Ability to identify the problem in clustering • Ability to write programing for clustering problem with some advance algorithm.	Chapter 6: Machine learning (Advance Unsupervised)	• Lecture Presentation • Tutorials Practice • Class Facilitating • Case Study Q&A	• Listen and Note Taking • Indiviudal Homework • Group Discussion and Presentation • Report/Project Presentation Q&A	• Attendance • Class Activities • Formative Assement • Quiz and Q&A	• Slide Presentation • LCD Projector • Laptop • Slide Pointer • Writting Board
15-16	57-64	CLO3 CLO4	• Ability to understand the machine learn in term of supervised • Ability to identify the problem in regression and classification • Ability to write programing for regression and classification	Chapter 7: Machine learning (Advance supervised and deep learning algorithm)	• Lecture Presentation • Tutorials Practice • Class Facilitating • Case Study Q&A	• Listen and Note Taking • Indiviudal Homework • Group Discussion and Presentation • Report/Project Presentation Q&A	• Attendance • Class Activities • Formative Assement • Quiz and Q&A	• Slide Presentation • LCD Projector • Laptop • Slide Pointer • Writting Board
17			Prepare for Final Exam					
18			Final Exam					

VII. Internal Regulation Related to Students' Learning and Assessment

To preserve the learning good environment in our classrooms, students are expected to adhere to the following rules:

- Student are expected to come to class punctually and regularly. Punctuality reflects that you are ready and willing to undertake the task at hand and are respectful of others involved. If you are late, please come in quietly and take a seat in the back of the room.
- Treat everyone in the classroom with respect and be tolerant of questions asked by fellow classmates. This is a diverse community and we need to respect each other's differences.
- Be respectful when engaging in online discourse.
- Pay attention and participate actively in the classroom conversation. Participation in class discussion is highly encouraged.
- Refrain from talking to other students during class or interrupting others. No "sidebars."
- Come to class prepared: always have your assignments, textbook, notebook, and pen.
- Turn off all cell-phones and other electronic devices not used for educational purposes during class.
- If you must leave during class, exit and re-enter as quietly as possible.
- Do not leave class during exam sessions.
- Do not litter in the classroom. Clean up around your desk before you leave.
- Wait until class has ended before you pack up your bags

VIII. References

Main Books

- Jake VanderPlas. *Python Data Science Handbook: Essential Tools for Working with Data,* 1st Edition, O'Reilly Media. 2016
- Andreas C. Müller and Sarah Guido*, Introduction to Machine Learning with Python: A Guide for Data Scientists*, 1st Edition, O'Reilly Media. 2017
- Stephen Klosterman, *Data Science Projects with Python: A case study approach to successful data science projects using Python, pandas, and scikit-learn*, 1st Edition, Packt Publishing. 2019

Others related course materials

- https://ieeexplore.ieee.org/document/7557668
- https://www.coursera.org/specializations/jhu-data-science
- https://www.coursera.org/professional-certificates/ibm-data-science
- https://ieeexplore.ieee.org/document/9141187
- https://ieeexplore.ieee.org/document/8701415

Remark: This syllabus is intended to provide guidance as to students 'and the instructor's obligations for this course, and to provide an outline of topics to be covered during the semester. However, the instructor reserves the right to modify syllabus items as needs arise. Students will be notified in advance of any modifications.