1. Course Code

2262

2. Course Title

G52e: Data Science Exercises

3. Teacher

HAMIDULLAH, Sokout

4. Term

Spring 2

5. Course Requirements (Courses / Knowledge for this course) and Important Information

Data Science Course (2261)

6. Course Overview and Objectives

The major goal of Data Science is to solve companies (institutions, government, private sectors, .. etc) problems using available data. Data Science entails everything that has something to do with data such as: Collecting the data, Cleaning up the data, Visualizing data, Analyzing it, and Creating predictive models based on the data. Playing with data will pave the way to gain insights that could help us with good decision making.

In this course we will start from the very basics (pre-processing data), we will build up your skills and soon you will be able to solve advanced statistics tasks using machine learning algorithms in Python such as CLUSTERING, REGRESSION, CLASSIFICATION. As well as you learn with integration and visualization with Python. For this year, we also added the part of Deep Learning for this course.

7. Course Outline

1	Course Orientation and Data Visualization with Python
2	Data Visualization with Python Exercises (Basics and Interactive)
3	Data Preprocessing + Regression with Python
4	Data Preprocessing Exercises with Machine Learning (Regression)
5	Machine Learning Algorithmes (Classification)
6	Classification Exercises
7	Machine Learning Algorithms (Clustering)
8	Clustering Exercises
9	Model Selection
10	Exercises
11	Integration Concept
12	Integration Exercises (In-Class Practice)
13	Deep Learning Concepts
14	Exercises with Deep Learning Algorithem
15	Students Presentation
16	Students Presentation

8. Textbooks (Required Books for this course)

9. Reference Books (optional books for further study)

(1) Machine Learning with Python Cookbook, ISBN: 9781491989388.
(2) Advanced Data Analytics using Python, ISBN: 978-1-4842-3449-5.
(3) Available resources on Internet

10. Course Goals (Attainment Targets)

(1)	Understanding and applying Visulization with Python
(2)	Able to preprocess data using Python
(3)	Gain in-depth familiarity with various Machine Learning algorithms (supervised learning algorithms and unsupervised learning algorithms), as well as Deep Learning
(4)	Able to Implement machine learning algorithms to real-world problems, and rigorously evaluate their performance using different methods.
(5)	Understand the concpet of Integration Programming and implement in real-world problems.
(6)	
(7)	
(8)	

11. Correspondence relationship between Educational goals and Course goals

Educational goals of the school			Course Goals
High level ICT skills	Basic academic skills		
	Specialized knowledge and literacy		(1) (2) (3) (4) (5)
Human skill (Tankyu skill)	Ability to continually improve own strengths		(1) (2) (3) (4) (5)
	Ability to discover and resolve the problem in society	Problem setting	
		Hypothesis planning	
		Hypothesis testing	
		Practice	
	Fundamental Competencies for Working Persons	Ability to step forward	(3) (4) (5)
		Ability to think through	(3) (4) (5)
		Ability to work in a team	(3) (4) (5)
Professional ethics			

12. Evaluation

Goals	Evaluation method & point allocation					
	Examination	Quiz	Reports	Presentation	Deliverables	Other
(1)		○		○	○	
(2)		○		○	○	
(3)	○			○	○	
(4)	○	○		○		
(5)					○	
(6)						
(7)						
(8)						
Allocation	20	20		20	40	

13. Evaluation Criteria

Examination	This facilitates students growth and improvement by allowing students to ask for assistance and teachers to focus on areas that might require additional attention. As well as, to evaluate the understanding of students and motivate them for further learning.
Quiz	Every week multiple choice quizzes are used to evaluate the understanding of students and motivate them for further learning.
Reports	
Presentation	In the final presentation, students will be asked to work on a real-world problem using Machine Learning algorithm. The evaluation will be based on participation in the group presentation, the structure of presentation, the relavance of the argument, time managment and relation to the course contents.
Deliverables	Individual and group assignment will be assigned for the students with focus on learning goals (1,3,4, and 5). The evaluation will be based on how the students understand the exercises and participation.
Other	

14. Active Learning

Hourly percentage of active learning within the whole class time	80%
1 Active learning such as problem solving assignment using the knowledge and skills acquired in class.	All the time
2 Active learning such as group works and discussions.	All the time
3 Outcome presentations and feedbacks.	Sometimes
4 Students actively make decisions on how the class should be conducted.	Not at all

15. Notes

Please bring your computers in the class.

16. Course plan

(Notice) This plan is tentative and might be changed at the time of delivery

Lessen 1: (Course Orientation)	Lecture + Exercises
1. Course Orientation and Introduction to Data Visualization 2. Data Visualization Exercises with Pyton	
Lessen 2: (Data Visualization with Python)	Lecture + Exercises
1. Bascis of Visualizations 2. Interactive way of Visualization with Python	
Lessen 3: Data Preprocessing + Regression with Python	Lecture + Exercises
1. Machine Learning Algorithmes (Regression) 2. Exercises	
Lessen 4: Regression with Python Exercises	Lecture + Exercises
1. Exercises and Practices	
Lessen 5: Machine Learning Algorithmes (Classification Concept)	Lecture + Exercises
1. Machine Learning Algorithmes (Classification Concept) 2. Exercises	

Lessen 6: Machine Learning Algorithmes (Classification Exercises)	Lecture + Exercises
1. Exercises and Practices	
Lessen 7: Machine Learning Algorithmes (Unsupervised Concept)	Lecture + Exercises
1. Machine Learning Algorithmes (Unsupervised Concept) 2. K-Means	
Lessen 8: Machine Learning Algorithmes (Unsupervised Concept)	Lecture + Exercises
1. Exercises and Practices	
Lessen 9: Model Selection	Lecture + Exercises

1. Model Selection
    a. K-fold cross validation
    b. Grid Search
    c. XG-Boost
    d. Cat-Boost	

Lessen 10: Model Selection Exercises	Lecture + Exercises
1. Integration Concpet 2. Integration Exercises and Practice	
Lessen 11: Integration Concept and Exercises	Lecture + Exercises

1. Integration Concept
    a. Python
    b. My-SQL
    c. Tableau	

Lessen 12: Integration Concept and Exercises	Lecture + Exercises
1. Devlopment Model with Integration Concept	
Lessen 13: Deep Learning Algorithms Concept and Exercises	Exercises
1. Deep Learning Concpet 2. Exercises with Deep Learning Algorithm	

Lessen 14: Deep Learning Algorithms Concept and Exercises	Exercises
1. Exercises with Deep Learning Algorithm	
Lessen 15: (Student Presentation)	Presentation(90 min)
Students Presentation	
Lessen 16 : Final Examination of the Course	Presentation (90 min)
Examination	