Department of Health Outcomes and Biomedical Informatics, University of Florida Gainesville, FL USA

INTRODUCTION

The Short Course of Data Science will last for **fifteen days**, from **Jan. 18th** to **Feb. 1st**. This workshop provides students with an introduction to a wide range of concepts, theories, and practical techniques in data science. Data science or data-driven science is an interdisciplinary field about scientific methods, processes, and systems to extract knowledge or insights from data in various forms, either structured or unstructured. Data and data analysis are essential components in a wide range of fields. It is critical for students to understand and gain practical experience with the entire life cycle of data, from data collection to data analysis to the dissemination and archiving of valuable results. In this workshop, students will be introduced to the broad landscape of data science. Students will learn to design and implement computerized databases for data collection, perform basic queries and prepare the data for analytical tasks, and understand basic data analytical skills. This workshop will be a foundation for students who are interested in becoming data scientists.

Importantly, the program provides the opportunity for students to visit UF research laboratories and work with the UF course staff. Students may discuss future MS and PhD opportunities with the department and university. Our program aims to help international students to learn theoretical knowledge, develop research ideas and skills, and improve their opportunities for successful application to US Graduate Schools. While on campus, students may meet with other faculty and visit ongoing classes and research labs at UF.

WORKSHOP INFORMATION

- **Contents:**

The workshop consists of 2 weeks of lectures. Teaching methods include lecture, discussion, and hands-on data assessment, analysis, and presentation. The goals of the course are:

- To provide an overview of the data, questions, and tools that data analysts and data scientists work with.
- To provide students with basic principles of data management.
- To introduce the concepts of big data, and the associated tools, methods and approaches.
- To give students access to exploratory data analysis techniques and tools.
- To give students an opportunity to conduct their own data science project and apply the skills learned in this workshop.

- **Schedule and Topic:**

The workshop will last 2 weeks and be hosted at the University of Florida Gainesville campus in January and February 2018. There will be 6 lectures per week (12 lectures total), and each lecture will be 2 hours in the morning. In the afternoon, a 3-hour session will be dedicated to either (1) office hours for Q&A sessions or (2) labs that help students gain hands-on experience in data science, with the help of a teaching assistant (i.e., a graduate-level student with either computer science or biomedical informatics background). Table 1 below shows a tentative schedule and associated lecture topics.

Table 1. A tentative schedule and associated lecture topics (January 18, 2019 to February 1, 2019).

Date	Topic
Day 1 (Lecture 1-2)	Introduction and course overview: data science articulated, history and context, examples, technology landscape
Day 2 (Lecture 3-4)	Data science toolbox: ideas and tools • Ideas behind turning data into actionable knowledge • Tools that are commonly used in the realm of data science
Day 3 (Lecture 5-6)	Python tutorial 1: Intro to Python
Day 4 (Lecture 7-8)	Python tutorial 2: Data cleaning, management, and visualization in Python
Day 5 (Lecture 9-10)	Machine Learning: lecture on supervised learning
Day 6 (Lecture 11-12)	Supervised learning in Weka/Python
Day 7 (Lecture 13-14)	Machine Learning: lecture on unsupervised learning
Day 8 (Lecture 15-16)	Unsupervised learning in Weka/Python
Day 9 (Lecture 17-18)	Case Study 1
Day 10 (Lecture 19-20)	Case Study 2
Day 11 (Lecture 21-22)	Case Study 3
Day 12	Final Presentation of Course Project

- Students will have the unique opportunity to participate in:

    - Intensive instruction by University of Florida professors in Engineering on Health Outcomes and Biomedical Informatics.
    - Customized English language program to improve skills in listening and speaking, grammar, reading and writing
    - Instruction on technical report writing (including grammar, style, and content)
    - Instruction on technical presentation delivery (including preparation skills)
    - Cultural interaction with UF faculty and students inside and outside of the classroom
    - Professional development lectures conducted by workshop staff

INSTRUCTORS

- Instructor 1--Jiang Bian, PhD, Principal Investigator.

Dr. Bian is an Assistant Professor of Biomedical Informatics in Health Outcomes & Biomedical Informatics (HOBI) at the University of Florida (UF). He has strong expertise in developing informatics tools and systems as well as extensive experience in machine learning methods for the interpretation of biomedical and textural data.

- Instructor 2--Yi Guo, PhD, MSPH, Principal Investigator

Dr. Guo is an Assistant Professor of Biostatistics in HOBI at UF. His research areas include bio-statistical methodology, cancer prevention and control, and cancer epidemiology. His expertise includes power and sample size analysis, analysis of multi-level and longitudinal designs, and psychometric analysis.

- Instructor 3-- François Modave, PhD, Principal Investigator

Dr. Modave is an Associated Professor of Biomedical Informatics in HOBI at UF. Dr. Modave is trained as a computer scientist with a specific focus on decision theory, machine learning, and more generally, intelligent systems.

BENEFITS

Specific benefits of the short course include the following:

- The international students will benefit from exposure to a high-quality and comprehensive educational experience under instruction from UF professors, along with relevant industrial and cultural experiences.
- Students will have the opportunity to interact with UF students serving as mentors in the classroom and on trips to foster cultural exchange.

HOUSING

During the program, student will be housed in the 3-star hotel assigned by the University of Florida. High-speed Internet access is available in each room and the breakfast is included for each student. Students will have access to on-campus dining options and 24/7 on-call staff.

TRANSPORTATION

University of Florida will be responsible for picking up students to campus after their arrival at Gainesville Regional Airport (GNV) and send them back to the Gainesville Regional Airport (GNV) after the program is finished. At orientation, students will be given a 1-month bus pass that gives them unlimited access to all Gainesville Regional Transit System (RTS) buses for the duration of the program. The hotel will provide shuttle to the campus on weekdays.

INSURANCE

During the program, Liberty Insurance covers students' health insurance.

ADMISSION REQUIREMENTS

Students must meet the following requirements:

- Minimum GPA 3.0
- CET-6 500 or TOEFL 80
- Major in Computer Science, Computer Engineering, Software, (Bio-)Medical Science or related field.
- Junior or senior level standing.
- Students who do not meet the requirements will need an interview (contact Sam Zheng).

PROGRAM ADMISSION

The logistical administration of this program will be carried out by Sam Zheng. His email address is dzheng1@ufl.edu(link sends e-mail).

TUITION FEE

$3,500 per student, including accommodation and breakfasts during three weeks, insurance, transportation on campus, education and teaching resource charge; but round-trip flight, meal cost (except breakfast) and personal cost are not included.

APPLICATION DETAILS

Please contact Sam Zheng for application process, costs and other details. The deadline for application is **Sep 30th, 2018**. Please contact Sam Zheng if you have any questions.