CSCI 3360 Data Science I  
Spring 2019

Instructor  
Dr. Sheng Li  
Department of Computer Science  
University of Georgia  
549 Boyd GSRC, Athens, GA 30602  
E-mail: sheng.li@uga.edu  

Time and Location of the Lectures:  
TR: 12:30 pm - 1:45 pm Chemistry 551  
W: 12:20 am - 1:10 pm Dawson Hall 208

Instructor Office Hours and Location  
Thursday: 1:10 pm - 2:30 pm or by an email appointment.  
Location: Boyd GSRC 549

TA Office Hours and Location  
TBA

Course Description  
This course presents a rigorous overview of methods for data mining, image processing, natural language processing, and scientific computing. Core concepts in supervised and unsupervised analytics, dimensionality reduction, deep learning, and data visualization will be explored in depth.

Recommended Prerequisites  
Programming; Data Structures; Linear Algebra; Probability

Credit Hours  
4
Text(s)

The main textbook for this course is: “An Introduction to Statistical Learning with Applications in R” by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. Springer. The PDF version of this book is available on the author’s homepage.

Course Topics

1. Review of data science
2. Data collection and data preprocessing
3. Data visualization
4. Regression
5. Classification
6. Clustering
7. Dimensionality reduction
8. Feature selection
10. Fundamental of natural language processing
11. Fundamental of recommender system
12. Overview of visual analytics, causal inference, deep reinforcement learning, etc.

Grade Distributions

- Class Participation  5%
- Homework  35%
  - 5 individual assignments involving problem solving and programming
  - Homework should be submitted to the eLC by due date (11:59pm).
- Team Project  25%
  - Project Proposal (5%); Progress Review (5%); Final Presentation and Report (15%)
- Midterm Exam  15%
- Final Exam  20%
  - Both exams are closed-books/notes.

Grade Conversion

- A:  [93,100]
- A-:  [90,92]
- B+:  [87,89]
- B:  [83,86]
- B-:  [80,82]
- C+:  [77,79]
- C:  [73,76]
- C: [70,72]
- D+: [67,69]
- D: [63,66]
- D-: [60,62]
- F: [0,59]

**Grading Policy**
- **Late Submission Policy**: Late submissions will be penalized by deducting 10% of the score for each day beyond due time.
- **Regrade Request**: The regrade request must be submitted by email (to TA and Instructor) within one week of distribution of your grade. Any regrade request after one week will NOT be considered.

**Team Project**
Students are required to work on a team (no more than five students) data science project over the semester. Team project will be evaluated based on the novelty, efforts, technical soundness, presentations, and the quality of final report.

**Academic Integrity and Ethics**
We will strictly follow UGA’s Academic Honesty Policy. Dishonest behavior will not be tolerated and may result into failing the course. Please contact the instructor if you have any concerns regarding this issue.