



FLORIDA STATE UNIVERSITY
COLLEGE OF COMMUNICATION & INFORMATION
School of Information

LIS 5765 DATA MINING AND ANALYTICS

TERM 20XX | TIME: XXXX | LOCATION: XXX

MODE OF INSTRUCTION: Online

Students all meet with instructor for class weekly, synchronously, using a multi-media conferencing system such as Collaborate, at a specific class time indicated in the University's course schedule for each semester. Additional asynchronous interactions (e.g., discussion forums) among students and with instructor may also be required to complete the course.

Instructor: Dr. Zhe He

Office:

Office Hours:

Email:

COURSE PREREQUISITES:

Completion of at least one database course in the undergraduate or graduate level.

LIS2780, LIS3781, LIS3784, OR LIS5782

COURSE DESCRIPTION:

This course provides an introduction to data mining methods and applications. In this course, students will learn basic concepts and tools for data mining, including data sources, data cleaning tools and methods, mainstream algorithms for data mining, statistical modeling, popular tools for mining structured data and unstructured data. Students will also learn how data mining can be effectively used in various application areas, with the focus on in healthcare, to drive decisions and actions. The students will conduct a project of data analytics, and use the tools introduced in the course to tackle the problem. This course is appropriate for students with basic knowledge and skills in database management systems. Prior programming skills are helpful but not required.

COURSE OBJECTIVES:

At the end of the course, the student will be able to:

- Describe and define basic concepts in data analytics.
- Discuss prominent algorithms in data analytics (e.g., K-means clustering, association rule mining).
- Explain the process of inferring knowledge and insights from heterogeneous data sources.
- Use data analytics tools with graphical user interface (e.g., Weka and Orange Data Mining) to perform basic data analytics tasks such as clustering, association rule mining, classification, visualization, and text mining.
- Conduct a problem-solving task and use appropriate tools and algorithms to solve the problem.

COURSE MATERIALS:

Required:

All required reading material such as articles, whitepapers, case studies, etc. will be posted on the course site.

Optional Texts:

Han, J., Kamber, M, and Pei, J. (2012) *Data Mining: Concepts and Techniques*, 3rd Edition. San Francisco, CA: Morgan Kaufmann. ISBN-10: 0123814790

Downey, A.B. (2015) *Think Python: How to Think Like a Computer Scientist*, Second Edition. O'Reilly Media. ISBN-10: 1491939362

Required Instructional Technologies:

(1) Hardware and software requirements for people enrolled in iSchool courses

<http://ischool.cci.fsu.edu/academics/online/requirements/>

(2) Python, Jupyter Notebook, MySQL, Weka, Orange Data Mining

(3) Each student should apply for a license of the Unified Medical Language System at <https://uts.nlm.nih.gov//license.html>

COURSE ASSIGNMENTS:

Attendance (5%)

Attendance will be taken in each online session.

Show and Tell (5%)

This assignment is aimed to allow students to explore major and trendy issues related to Data Science and share their thoughts and opinions with their classmates. Students are required to select one news article related to data science/big data and give a brief presentation on his/her choice of article during class. One student will present during every session. A brief Q&A or discussion session on the issue will follow.

Term Project (50%)

There will be a project for this course. You will be asked to choose a problem out of the given two problems that requires a data analytics solution. You will be asked to present a final presentation and submit a project report.

Content and Format

Project Weekly Reports (10%)

In each week, students should report their progress with the project.

Project Final Presentation (15%)

The project presentation will include an introduction to the problem, the methodology framework to solve the problem, the experimental results, and the interpretation of the output.

Project Report (25%)

All material that is not your own work must be cited (APA or Vancouver Style) or will be considered an instance of plagiarism (see Academic Honor Policy, page 7). The project report will be submitted to the course website.

Problem Sets (20%)

You will have open-book quizzes to test your understanding and mastery of class concepts. The quizzes will be offered online in the course site. You will be given two weeks to work on each quiz.

Class activities (20%)

- In every class session, students will participate in an activity.
- Student will receive 5 points for each activity through their participation.
- These class activities are designed to help you gain a better grasp of the class content that is essential for your project.
- No make-up will be available for the activities. But students can be waived for missing activities up to 2 times during the semester.

GRADE CALCULATION:

The course grade will be calculated by converting the total points earned to a percentage and assigning the corresponding letter grade.

Item	Point Value
Attendance	5
Show and Tell	5
Class Activities	20
Project Weekly Reports	10
Project Final Presentation	15
Project Report	25
Problem Sets	20
Total number of points	100/100

GRADING SCALE:

The course will use the standard grading scale for courses taught at FSU:

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

Note: As a general policy, no INC. grades will be given for this course except under legitimate extenuating circumstances and only if the student requesting the incomplete has attended at least 50% of the class meetings and can present valid medical documentation for inability to comply with the course requirements.

COURSE SCHEDULE (SUBJECT TO CHANGE):

Week	Course Topics
1	Course overview
2	Introduction to data analytics
3	Getting to know your data
4	Collecting & cleaning data
5	Data warehousing & online analytical processing (OLAP)
6	Association rule mining
7	Predictive analysis & basic classification (Part A)
8	Predictive analysis & basic classification (Part B)
9	Clustering analytics
10	Text mining and analysis (Part A)
11	Text mining and analysis (Part B)
12	Health data analytics
13	Statistical tests
14	Deep learning
15	Trends in Big Data
16	Project Final presentations and feedback

UNIVERSITY ACADEMIC POLICIES

UNIVERSITY ATTENDANCE POLICY:

University-wide policy requires all students to attend the first day of class meeting of all classes for which they are registered. Students who do not attend the first class meeting of a course for which they are registered will be dropped from the course by the academic department that offers the course. This policy applies to all

levels of courses and to all campuses and study centers. **It remains the student's responsibility to verify course drops and check that fees are adjusted.**

Excused absences include documented illness, deaths in the family and other documented crises, call to active military duty or jury duty, religious holy days, and official University activities. These absences will be accommodated in a way that does not arbitrarily penalize students who have a valid excuse. Consideration will also be given to students whose dependent children experience serious illness.

ACADEMIC HONOR POLICY:

The Florida State University Academic Honor Policy outlines the University's expectations for the integrity of students' academic work, the procedures for resolving alleged violations of those expectations, and the rights and responsibilities of students and faculty members throughout the process. Students are responsible for reading the Academic Honor Policy and for living up to their pledge to "...be honest and truthful and...[to] strive for personal and institutional integrity at Florida State University." (Florida State University Academic Honor Policy, found at <http://fda.fsu.edu/Academics/Academic-Honor-Policy>)

ADA STATEMENT:

Students with disabilities needing academic accommodation should:

- (1) register with and provide documentation to the Student Disability Resource Center; and
- (2) bring a letter to the instructor indicating the need for accommodation and what type.

Please note that instructors are not allowed to provide classroom accommodation to a student until appropriate verification from the Student Disability Resource Center has been provided.

This syllabus and other class materials are available in alternative format upon request.

For more information about services available to FSU students with disabilities, contact the:

Student Disability Resource Center
874 Traditions Way
108 Student Services Building
Florida State University
Tallahassee, FL 32306-4167
(850) 644-9566 (voice)
(850) 644-8504 (TDD)
sdrc@admin.fsu.edu
<http://www.disabilitycenter.fsu.edu/>

SYLLABUS CHANGE POLICY:

Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice.

SCHOOL OR DISCIPLINARY POLICIES:

COPYRIGHT STATEMENT:

Some of the materials in this course are possibly copyrighted. They are intended for use only by students registered and enrolled in this course and only for instructional activities associated with, and for the duration of, the course. They may not be retained in another medium or disseminated further. They are provided in compliance with the provisions of the *Technology, Education, And Copyright Harmonization* (TEACH) Act (refer to the 3/7/2001 TEACH Act at www.copyright.gov/legislation/archive/).

SEXUAL HARASSMENT POLICY:

It is the policy of the University that its employees and students neither commit nor condone sexual harassment in any form. http://registrar.fsu.edu/bulletin/grad/info/university_notices.htm

SCHOOL OF INFORMATION HARDWARE AND SOFTWARE REQUIREMENTS:

A list of all hardware and software requirements for students participating in courses offered by the School of Information (iSchool) can be found at the following location:

<http://ischool.cci.fsu.edu/academics/online/requirements/>.

STUDENT ELIGIBILITY FOR AN INCOMPLETE GRADE:

Incomplete ("I") grades will not be assigned, except in the case of exceptional unforeseen circumstances that occur within the last three weeks of the semester and your work has otherwise been satisfactory (C average).