

# Course Syllabus Part II

## DSC550 – Data Mining

3 Credit Hours

### Course Resources

#### Course Texts:

*Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning (1st Edition).*  
 Benjamin Bengfort, Rebecca Bilbro, & Tony Ojeda. O'Reilly Media ISBN-13: 978-1491963043 ISBN-10: 1491963042

*Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning (1st Edition).*  
 Chris Albon O'Reilly Media ISBN-13: 978-1491989388 ISBN-10: 1491989386

#### Required Resources

In this course, you will need to be able to:

- Access the Internet.
- Access Cyberactive.
- Access to Github.
- Collaborate Online via Video and Voice.
- Collaborate while writing a single document.
- Submit a Word Document.
- Access to GitHub account.
- Python programming environment using PyCharm, Anaconda, and Jupyter Notebook.

### Course Schedule

Week	Topic	Assigned Reading
1	Text preprocessing, transformation, vectorization, and feature extraction	Machine Learning with Python Chapters 1-4 Applied Text Analysis Chapter 1
2	Handling Categorical Data, Text, Dates & Times	Machine Learning with Python Chapters 16-18 Applied Text Analysis with Python Chapters 2-5
3	Graphical Displays	Machine Learning with Python Chapters 5-7
4	Topic modeling, document similarity, and Context aware text analysis	Applied Text Analysis with Python Chapters 6-7
5	Graph analysis and social networks	Machine Learning with Python Chapters 8-10 Applied Text Analysis with Python Chapters 8-9
6	Unsupervised learning Part 1: Collaborative filtering and frequent pattern mining	Machine Learning with Python Chapters 11-12 Applied Text Analysis with Python Chapter 10
7	Unsupervised learning Part 2: Hierarchical clustering and dimensionality reduction	Machine Learning with Python Chapters 13-15
8	Model Evaluation and Selection	Machine Learning with Python Chapters 19-20
9&10	Practical considerations including project organization, infrastructure, scaling, and ethics and advanced topics in data mining	Machine Learning with Python Chapters 8 and 20- 21 Applied Text Analysis with Python Chapters 11-12

11&12	Final project	
-------	---------------	--

## Course Activities

In this section of the syllabus, I will describe what we will be doing in each of the activities for each week. Specifically, I will be describing your deliverables – those items you need to submit at or before the deadline. You can find more detail on grading criteria for each category by viewing its detailed rubric.

### Discussion/Participation

Every week you will be required to make 10 posts via an online platform. The goal is to simulate real world discussion and participation – there will not be formal posts required or required topics to discuss. There may be optional topics provided to start discussion, however, sharing information, troubleshooting, asking questions/feedback, etc. will be the primary focus for discussion/participation. Discussion/Participation will be graded as follows:

0	0%
5	50%
10	100%

### Exercises

Each week, you will be assigned an exercise or series of exercises based on the weekly topic to complete and submit to the assignment link. These are no group assignments to complete and should be done on your own. However, if you have questions about a specific exercise, you are encouraged to discuss with your classmates, without completing the assignment together.

### Term Project

This course has one major project, that has five total milestones (including the final submission). In this project, students will create a fully functional Python program that applies knowledge and skills from previous lessons.

### Grade and Point Breakdown

Component	Percentage	Point Value	Number of Times	Total
Discussion	20%	20	10 Times per Week for 12 Weeks	240
Exercises	45%	60	9 Per Term	540
Term Project	35%	Varies	4 Milestones worth 55pts each, Final Project Submission worth 200pts	420
				1200

### Late Work

Late work is not accepted unless arrangements are made with the instructor for very special, unavoidable circumstances. If you do not alert the professor before or shortly after something that will make you late, the chances of special arrangements are much lower. If in doubt, please email as soon as possible.

### Participation

Students are required to login often and contribute to the class on a regular basis, including posting in the online platform, submitting assignments, and participating in group activities as required. If you have specific participation requirements related to your educational funding or student status, you are expected to monitor your own participation to ensure you are in compliance with those requirements.

**Expectations for Students:**

- Students should expect to spend approximately 10-15 hours per week to complete the activities and assignments in this course.
- Students will log in as often as needed to complete their assignments and progress through the course.
- Students will treat their classmates and the instructor with respect and courtesy.
- Students are responsible for keeping current with the reading assignments and coming to class prepared to discuss the work assigned.
- Students are responsible for knowing what assignments are due and when.
- Students will submit only their own work and will not commit plagiarism or other acts of academic dishonesty.
- Students will contact the instructor as soon as personal problems arise that may affect the student's ability to complete assignments on time.

**Expectations for Faculty:**

- The instructor will treat all students with respect and courtesy.
- The instructor will make grading criteria clear and follow the criteria scrupulously in evaluating student work.
- The instructor will provide feedback about student work within 6 days of due dates (or 24 hours prior to the next due date)—feedback that helps the student learn and improve.
- The instructor will respond to all student messages within 48 hours.