



CSC425: PRINCIPLES OF MACHINE LEARNING

Credit Hours: 3

Contact Hours: This is a 3-credit course, offered in accelerated format. This means that 16 weeks of material is covered in 8 weeks. The exact number of hours per week that you can expect to spend on each course will vary based upon the weekly coursework, as well as your study style and preferences. You should plan to spend 14-20 hours per week in each course reading material, interacting on the discussion boards, writing papers, completing projects, and doing research.

Faculty Information: Faculty contact information and office hours can be found on the faculty profile page.

COURSE DESCRIPTION AND OUTCOMES

Course Description:

This course provides an introduction to concepts associated with machine learning and pattern recognition theory. Students will be introduced to topics associated with supervised learning, unsupervised learning, learning theory, reinforcement learning and adaptive control. Students will gain an understanding of machine learning in regards to applications in speech recognition, data processing, data mining, and robotic control.

Course Overview:

Have you ever wondered how Amazon knows what you want to buy before you make a purchase? In this class, you will learn a lot about machine learning and how this technology analyzes data and looks for patterns. This class will help provide a deeper look into identifying structures in data and discuss applications of the technology in different fields of computing.

Course Learning Outcomes:

1. Identify machine learning techniques for a given application.
2. Discuss the characteristics of a given dataset for solving a machine learning problem.
3. Formulate machine learning techniques for a given application.
4. Apply techniques associated with supervised learning for a given scenario.
5. Solve an adaptive control problem for a given scenario.
6. Discuss applications of machine learning to various applied computing fields.
7. Determine techniques to use in an unsupervised learning problem.

PARTICIPATION & ATTENDANCE

Prompt and consistent attendance in your online courses is essential for your success at CSU-Global Campus. Failure to verify your attendance within the first 7 days of this course may result in your withdrawal. If for some reason you would like to drop a course, please contact your advisor.

Online classes have deadlines, assignments, and participation requirements just like on-campus classes. Budget your time carefully and keep an open line of communication with your instructor. If you are having technical problems, problems with your assignments, or other problems that are impeding your progress, let your instructor know as soon as possible.

COURSE MATERIALS

Required:

Kelleher, J. D., Namee, B. M., & DArcy, A. (2015). *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*. Cambridge, MA: The MIT Press.

- Print ISBN: 9780262029445
- eText ISBN: 9780262331746

Suggested:

Students will need to install the latest version of Python 3 on their computers. The Python 3 distribution (Python 3.6 or later) may be downloaded from <https://www.python.org/downloads/release/python-362/>

NOTE: All non-textbook required readings and materials necessary to complete assignments, discussions, and/or supplemental or required exercises are provided within the course itself. Please read through each course module carefully.

COURSE SCHEDULE

Due Dates

The Academic Week at CSU-Global begins on Monday and ends the following Sunday.

- **Discussion Boards:** The original post must be completed by Thursday at 11:59 p.m. MT and peer responses posted by Sunday at 11:59 p.m. MT. Late posts may not be awarded points.
- **Opening Exercises:** Take the Opening Exercise before reading each week's content to see which areas you will need to focus on. You may take these exercises as many times as you need. The Opening Exercises will not affect your final grade.
- **Mastery Exercises:** Students may access and retake Mastery Exercises through the last day of class until they achieve the scores they desire.
- **Critical Thinking:** Assignments are due Sunday at 11:59 p.m. MT.

WEEKLY READING AND ASSIGNMENT DETAILS

Module 1

Readings

- Chapter 1 in *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*

Opening Exercise (0 points)

Discussion (25 points)

Critical Thinking (65 points)

OPTION #1 & #2: Machine Learning Scenarios

Part 1:

Identify at least two to three machine learning scenarios and distinguish how they differ. Describe how each scenario uses test data to evaluate the learning algorithm.

Part 2:

Throughout the course we will utilize python to complete various exercises. Please ensure that you have installed python.

1. Install Python3 on your computer if you do not have it already installed it. You can download it from <https://www.python.org/downloads/release/python-362/>.
2. Make sure you check the box to include the Python executable in your environment path.
3. Edit your Python program using your choice of editor such as Notepad, Notepad++, or Idle. Idle is a simple Python interactive development environment that installed with your Python package.
4. Save your Python code using the file name ***CSC425_CTA1_Option1.py***.
5. Execute your Python code in command prompt as ***python CSC425_CTA1_Option1.py***.

Submission Requirements

- Your well-written paper should be four to five pages in length not including the required cover and references pages.
- Support your paper with at least four to six scholarly references (of which the textbook may be one). The [CSU-Global Library](#) is a good place to find these references.
- Format your paper according to the [CSU-Global Guide to Writing and APA](#).
- Submit the text file named ***CSC425_CTA1_Option1.py*** that contains your Python code into the Module 1 drop box.

Mastery Exercise (10 points)

Portfolio Reminder

Take some time to review the Portfolio Project options, assignment requirements, and rubric, which can be found in Module 8. Choose one of the two options. Then, review Week 3 and Week 4 Portfolio Project Milestone assignments, which are both worth 25 points. The final Portfolio Project will be due at the end of the course before midnight of Week 8.

You are encouraged to visit the [CSU-Global Library](#) for additional resources or contact the CSU-Global Librarian at Library@CSUGlobal.edu. The [Purdue Online Writing Lab's APA Formatting and Style Guide](#) is also a great resource for additional information.

If you have questions, contact your instructor sooner rather than later so that you have ample time to discuss the assignment.

Module 2

Readings

- Chapter 2 in *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*

Opening Exercise (0 points)

Discussion (25 points)

Critical Thinking (60 points)

Choose one of the following two assignments to complete this week. Do *not* complete both assignments. Identify your assignment choice in the title of your submission.

OPTION #1: Predictive Model

Consider the scenario that Net-tastic, an online gaming company, is having a problem losing gaming customers each month. Net-tastic would like to increase their subscribers and reduce the customers loss each month. What are some of the predictive data analytics tools that can be used to address this business problem? Please be sure to describe the predictive data models that you have selected and identify the benefit that each model would post to solving this problem.

Submission Requirements

- Your well-written paper should be 4-5 pages in length not including the required cover and references pages.
- Support your paper with at least three scholarly references (of which the textbook may be one). The [CSU-Global Library](#) is a good place to find these references.
- Format your paper according to the [CSU-Global Guide to Writing and APA](#).

OPTION #2: Predictive Model Presentation

Consider the scenario that Net-tastic, an online gaming company, is having a problem losing gaming customers each month. Net-tastic would like to increase their subscribers and reduce the customers loss each month. What are some of the predictive data analytics tools that can be used to address this business problem? Please be sure to describe the predictive data models that you have selected and identify the benefit that each model would post to solving this problem.

Submission Requirements

Your well-written presentation should meet the following requirements:

- Be 10-12 slides in length not including the required title and references pages.
- Use any presentation software you prefer (PowerPoint, Prezi, Google Slides, or Apache Open Office, for example).
- Be formatted for a formal presentation including headings per slide and major points per slide.
- Include presentation speaker notes to support the slide content, as if you were doing the actual presentation. If you are using PowerPoint, please place them in the speaker notes. If your software choice doesn't have a place for speaker notes, provide them in a separate Word document with a link to the presentation.
- Depending on the type of presentation software chosen, upload either a PowerPoint file, or a single Word document with a link to your presentation, instructions for viewing, and speaker notes. Do not submit a PDF file.

NOTE: Since you can to submit only one document, if you need to include speaker notes, be sure your link AND speaker notes are in the same document.

- Support for your presentation should include at least three scholarly references (of which the textbook may be one) with associated in-text citations. The [CSU-Global Library](#) is a good place to find these references.

Mastery Exercise (10 points)

Module 3

Readings

- Chapter 3 in Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies
- Gallagher, R. (2018, September 13). Senior google scientist resigns over “forfeiture of our values” in china. The Intercept. Retrieved from <https://theintercept.com/2018/09/13/google-china-search-engine-employee-resigns/>

Opening Exercise (0 points)

Discussion (25 points)

Critical Thinking (60 points)

Choose one of the following two assignments to complete this week. Do *not* complete both assignments. Identify your assignment choice in the title of your submission.

OPTION #1: Machine Learning to Drive Customer Retention

Using the available datasets at <https://skymind.ai/wiki/open-datasets>, pick an appropriate dataset and construct a table that identifies the descriptive features of the dataset. You should identify at least four descriptive features and determine the average of these descriptive features. To complete this task, you may utilize Excel or python. Document your findings for this analysis in a two-page paper. Include your python script or excel file that was utilized for analyzing the data.

Submission Requirements

- Your well-written paper should be two pages in length not including the required cover and references pages.
- Support your paper with at least three scholarly references (of which the textbook may be one). The CSU-Global Library is a good place to find these references.
- Include your python script or Excel file.
- Format your paper according to the [CSU-Global Guide to Writing and APA](#).

OPTION #2: Machine Learning to Drive Customer Retention

Using the available datasets at <https://skymind.ai/wiki/open-datasets>, pick an appropriate dataset and construct a table that identifies the descriptive features of the dataset. You should identify at least four descriptive features and determine the average of these descriptive features. To complete this task, you may utilize excel or python. Document your findings for this analysis in a two-page paper. Include your python script or excel file that was utilized for analyzing the data.

Submission Requirements

Your well-written presentation should meet the following requirements:

- Be 10-12 slides in length not including the required title and references pages.

- Use any presentation software you prefer (PowerPoint, Prezi, Google Slides, or Apache Open Office, for example).
- Be formatted for a formal presentation including headings per slide and major points per slide.
- Include presentation speaker notes to support the slide content, as if you were doing the actual presentation. If you are using PowerPoint, please place them in the speaker notes. If your software choice doesn't have a place for speaker notes, provide them in a separate Word document with a link to the presentation.
- Depending on the type of presentation software chosen, upload either a PowerPoint file, or a single Word document with a link to your presentation, instructions for viewing, and speaker notes. Do not submit a PDF file.

NOTE: Since you can submit only one document, if you need to include speaker notes, be sure your link AND speaker notes are in the same document.

- Support for your presentation should include at least three scholarly references (of which the textbook may be one) with associated in-text citations. The [CSU-Global Library](#) is a good place to find these references.
- Format your presentation according to the [CSU-Global Guide to Writing and APA](#).

Mastery Exercise (10 points)

Portfolio Milestone (25 points)

Options #1 and #2

You have a choice between two Portfolio Projects in this course. Go to the Week 8 Assignments page to read both Portfolio Project descriptions. Select only one project--do not do both projects.

This week, identify your Portfolio Project choice in the title of a Word document and then briefly explain why you selected that project. Review the Portfolio Project Grading Rubric on the Course Information page to understand how you will be graded on the project.

- Format your paper according to the [CSU-Global Guide to Writing and APA Requirements](#).
- Your description should be between one-half to one page in length.

Module 4

Readings

- Chapter 4 in *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*
- Mumtaz, P., Steven, A., Wilkinson, J., Patel, L., & Baker, P. (2016). Value of supervised learning events in predicting doctors in difficulty. *Medical Education*, 50(7), 746-756.

Opening Exercise (0 points)

Discussion (25 points)

Critical Thinking (60 points)

Choose one of the following two assignments to complete this week. Do *not* complete both assignments. Identify your assignment choice in the title of your submission.

OPTION #1: ID3 Decision Tree Induction Algorithm

Using the available datasets at <https://skymind.ai/wiki/open-datasets>, pick an appropriate dataset and construct an appropriate decision tree using the ID3 decision tree induction algorithm. Discuss the steps

taken to construct the decision tree and identify findings that can be obtained from the construction. Document your findings for this analysis in a two-page paper.

Submission Requirements

- Your well-written paper should be two to three pages in length not including the required cover and references pages.
- Support your paper with at least three scholarly references (of which the textbook may be one). The [CSU-Global Library](#) is a good place to find these references.
- Format your paper according to the [CSU-Global Guide to Writing and APA](#).

OPTION #2: ID3 Decision Tree Induction Algorithm

Using the available datasets at <https://skymind.ai/wiki/open-datasets>, pick an appropriate dataset and construct an appropriate decision tree using the ID3 decision tree induction algorithm. Discuss the steps taken to construct the decision tree and identify findings that can be obtained from the construction. Document your findings for this analysis in a 10-12 slide presentation.

Submission Requirements

Your well-written presentation should meet the following requirements:

- Be 10-12 slides in length not including the required title and references pages.
- Use any presentation software you prefer (PowerPoint, Prezi, Google Slides, or Apache Open Office, for example).
- Be formatted for a formal presentation including headings per slide and major points per slide.
- Include presentation speaker notes to support the slide content, as if you were doing the actual presentation. If you are using PowerPoint, please place them in the speaker notes. If your software choice doesn't have a place for speaker notes, provide them in a separate Word document with a link to the presentation.
- Depending on the type of presentation software chosen, upload either a PowerPoint file, or a single Word document with a link to your presentation, instructions for viewing, and speaker notes. Do not submit a PDF file.

NOTE: Since you can to submit only one document, if you need to include speaker notes, be sure your link AND speaker notes are in the same document.

- Support for your presentation should include at least three scholarly references (of which the textbook may be one) with associated in-text citations. The [CSU-Global Library](#) is a good place to find these references.
- Format your presentation according to the [CSU-Global Guide to Writing and APA](#).

Mastery Exercise (10 points)

Portfolio Milestone Options #1 and #2: (25 points)

Submit an outline of your Portfolio Project

- State your paper topic.
- Provide an outline of your paper.
 - Your outline should include headers (the major topics), major resources, and the intended flow of the information in the project.
 - Under each header, write a few sentences on what you think you might cover in that section.

- Provide three articles that you might consider using for your final Portfolio Project. These sources cannot include the textbook or other course materials.
- Give a short reason why you feel each source would be pertinent to your project. This is not expected to be a final list. The goal here is to motivate you to begin examining research that might help you in your final Portfolio Project.

Module 5

Readings

- Chapter 5 in *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*
- Hübner, D., Verhoeven, T., Schmid, K., & Tangermann, M. (2017). Learning from label proportions in brain-computer interfaces: Online unsupervised learning with guarantees. *PLoS One*, 12(4), E0175856.

Opening Exercise (0 points)

Discussion (25 points)

Critical Thinking (65 points)

Choose one of the following two assignments to complete this week. Do *not* complete both assignments. Identify your assignment choice in the title of your submission.

OPTION #1: Implementing Forward Sequential Selection

In this module, we have learned about similarity-based learning techniques. For this assignment, you are to provide an implementation of the forward sequential selection search in python. Upon completing your model, test it on the dataset of your choice. Provide a two-page summary that highlights your findings.

Submission Requirements

- Your well-written paper should be two pages in length not including the required cover and references pages.
- Support your paper with at least four to six scholarly references (of which the textbook may be one). The [CSU-Global Library](#) is a good place to find these references.
- Provide python source code (.py)
- Format your paper according to the [CSU-Global Guide to Writing and APA](#).

OPTION #2: Implementing Forward Sequential Selection

In this module, we have learned about similarity-based learning techniques. For this assignment, you are to provide an implementation of the forward sequential selection search in python. Upon completing your model, test it on the dataset of your choice. Support your assertions with a 10-slide presentation.

Submission Requirements

Your well-written presentation should meet the following requirements:

- Be 10 slides in length not including the required title and references pages.
- Use any presentation software you prefer (PowerPoint, Prezi, Google Slides, or Apache Open Office, for example).
- Be formatted for a formal presentation including headings per slide and major points per slide.

- Include presentation speaker notes to support the slide content, as if you were doing the actual presentation. If you are using PowerPoint, please place them in the speaker notes. If your software choice doesn't have a place for speaker notes, provide them in a separate Word document with a link to the presentation.
- Depending on the type of presentation software chosen, upload either a PowerPoint file, or a single Word document with a link to your presentation, instructions for viewing, and speaker notes. Do not submit a PDF file.

NOTE: Since you can to submit only one document, if you need to include speaker notes, be sure your link AND speaker notes are in the same document.

- Support for your presentation should include at least three scholarly references (of which the textbook may be one) with associated in-text citations. The [CSU-Global Library](#) is a good place to find these references.
- Format your presentation according to the [CSU-Global Guide to Writing and APA](#).

Mastery Exercise (10 points)

Module 6

Readings

- Chapter 6 in *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*
- Zhou, Y., Duval, B., & Hao, J. (2018). Improving probability learning based local search for graph coloring. *Applied Soft Computing*, 65: 542-53.

Opening Exercise (0 points)

Discussion (25 points)

Critical Thinking (60 points)

Choose one of the following two assignments to complete this week. Do *not* complete both assignments. Identify your assignment choice in the title of your submission.

OPTION #1: Naïve Bayes Model Using Python

Using the available datasets at <https://skymind.ai/wiki/open-datasets>, pick an appropriate dataset and create a Naïve Bayes model using python. Discuss the steps taken to construct the model and identify findings that can be obtained from the construction. Be sure to identify the probability density function used to model specific descriptive features in the dataset. Document your findings for this analysis in a two-page paper.

Submission Requirements

- Your well-written paper should be three pages in length not including the required cover and references pages.
- Support your paper with at least three scholarly references (of which the textbook may be one). The [CSU-Global Library](#) is a good place to find these references.
- Provide python source code (.py)
- Format your paper according to the [CSU-Global Guide to Writing and APA](#).

OPTION #2: Naïve Bayes Model Using Python

Using the available datasets at <https://skymind.ai/wiki/open-datasets>, pick an appropriate dataset and create a naïve bayes model using python. Discuss the steps taken to construct the model and identify findings that can be obtained from the construction. Be sure to identify the probability density function used to model specific descriptive features in the dataset. Document your findings for this analysis in a two-page paper.

Submission Requirements

Your well-written presentation should meet the following requirements:

- Be 10-12 slides in length not including the required title and references pages.
- Use any presentation software you prefer (PowerPoint, Prezi, Google Slides, or Apache Open Office, for example).
- Be formatted for a formal presentation including headings per slide and major points per slide.
- Include presentation speaker notes to support the slide content, as if you were doing the actual presentation. If you are using PowerPoint, please place them in the speaker notes. If your software choice doesn't have a place for speaker notes, provide them in a separate Word document with a link to the presentation.
- Depending on the type of presentation software chosen, upload either a PowerPoint file, or a single Word document with a link to your presentation, instructions for viewing, and speaker notes. Do not submit a PDF file.

NOTE: Since you can to submit only one document, if you need to include speaker notes, be sure your link AND speaker notes are in the same document.

- Support for your presentation should include at least three scholarly references (of which the textbook may be one) with associated in-text citations. The [CSU-Global Library](#) is a good place to find these references.
- Format your presentation according to the [CSU-Global Guide to Writing and APA](#).

Mastery Exercise (10 points)

Portfolio Reminder

Continue working on your Portfolio Project, referring to the Portfolio Project assignment and corresponding Portfolio Project grading rubric in Module 8. The Portfolio Project will be due at the end of the course before midnight of Week 8.

Module 7

Readings

- Chapter 7 in *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*
- deLong, J. B. (2017). Artificial intelligence and artificial problems. *Project Syndicate*. Retrieved from <https://csuglobal.idm.oclc.org/login?url=https://search-proquest-com.csuglobal.idm.oclc.org/docview/1883347593?accountid=38569>
- Wong, A. L., & Shelhamer, M. (2011). [Exploring the fundamental dynamics of error-based motor learning using a stationary predictive-saccade task](#). *PLoS ONE*, (6)9, 1-13.

Opening Exercise (0 points)

Discussion (25 points)**Mastery Exercise (10 points)****Portfolio Reminder**

Continue working on your Portfolio Project, referring to the Portfolio Project assignment and corresponding Portfolio Project grading rubric in Module 8. The Portfolio Project will be due at the end of the course before midnight of Week 8.

Module 8**Readings**

- Chapter 8 in Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies

Opening Exercise (0 points)**Discussion (25 points)****Mastery Exercise (10 points)****Portfolio Project Option #1 and #2 Detailed Data Analysis Project (300 points)**

Using openly available data, you are to perform a detailed analysis using appropriate machine learning techniques. Specifically, you will create a program that is able to employ machine learning for data analysis and prediction. Your python program should accomplish the following:

- Import data into the program
- Provide for analysis of descriptive features
- Utilize data exploration techniques
- Utilize information-based learning models and techniques
- Utilize similarity-based learning models and techniques
- Utilize probability-based learning models and techniques

Upon completion, you will provide an eight-page paper that summarizes your findings using these machine learning techniques.

Submission Requirements:

- Your well-written plan should be eight pages in length, not including the required title and references pages. Headings and sub-headings may help you organize your work.
- Include at least eight academic or other legitimate sources (not including your textbook) to support your findings. The [CSU-Global Library](#) can help you with finding quality resources.
- Review the Portfolio Project grading rubric, which can be accessed from the Module 8 folder.
- Include all appropriate python scripts (.py).
- Format your paper according to the CSU-Global Guide to Writing & APA.
- Please be sure to reach out to your instructor at any point in the course if you have questions about the assignment.

COURSE POLICIES

Grading Scale	
A	95.0 – 100
A-	90.0 – 94.9
B+	86.7 – 89.9
B	83.3 – 86.6
B-	80.0 – 83.2
C+	75.0 – 79.9
C	70.0 – 74.9
D	60.0 – 69.9
F	59.9 or below

Course Grading

20% Discussion Participation
0% Opening Exercises
8% Mastery Exercises
37% Critical Thinking Assignments
35% Final Portfolio Project

IN-CLASSROOM POLICIES

For information on late work and incomplete grade policies, please refer to our [In-Classroom Student Policies and Guidelines](#) or the Academic Catalog for comprehensive documentation of CSU-Global institutional policies.

Academic Integrity

Students must assume responsibility for maintaining honesty in all work submitted for credit and in any other work designated by the instructor of the course. Academic dishonesty includes cheating, fabrication, facilitating academic dishonesty, plagiarism, reusing/repurposing your own work (see CSU-Global Guide to Writing & APA for percentage of repurposed work that can be used in an assignment), unauthorized possession of academic materials, and unauthorized collaboration. The CSU-Global Library provides information on how students can avoid plagiarism by understanding what it is and how to use the Library and internet resources.

Citing Sources with APA Style

All students are expected to follow the CSU-Global Guide to Writing & APA when citing in APA (based on the most recent APA style manual) for all assignments. A link to this guide should also be provided within most assignment descriptions in your course.

Disability Services Statement

CSU-Global is committed to providing reasonable accommodations for all persons with disabilities. Any student with a documented disability requesting academic accommodations should contact the Disability Resource Coordinator at 720-279-0650 and/or email ada@CSUGlobal.edu for additional information to coordinate reasonable accommodations for students with documented disabilities.

Netiquette

Respect the diversity of opinions among the instructor and classmates and engage with them in a courteous, respectful, and professional manner. All posts and classroom communication must be conducted in accordance with the student code of conduct. Think before you push the Send button. Did you say just what you meant? How will the person on the other end read the words?

Maintain an environment free of harassment, stalking, threats, abuse, insults, or humiliation toward the instructor and classmates. This includes, but is not limited to, demeaning written or oral comments of an ethnic, religious, age, disability, sexist (or sexual orientation), or racist nature; and the unwanted sexual advances or intimidations by email, or on discussion boards and other postings within or connected to the online classroom. If you have concerns about something that has been said, please let your instructor know.