



Department of Mathematical, Information, and Computer Sciences

## MTH4142 Data Science Project I

2 Units

*Fall 2024*

**Meeting Location/Time:** RS365 on Tuesdays, 4:30pm-5:30pm

**Final Exam:** The final project presentation is the final exam for this class.

<b>Instructor title and name:</b>	Dr. José Manjarrés
<b>Phone:</b>	(619) 849-2451
<b>Email:</b>	josemanjarres@pointloma.edu
<b>Office location and hours:</b>	RS 276   <a href="#">Schedule office hours here</a>

### PLNU Mission

#### To Teach ~ To Shape ~ To Send

Point Loma Nazarene University exists to provide higher education in a vital Christian community where minds are engaged and challenged, character is modeled and formed, and service is an expression of faith. Being of Wesleyan heritage, we strive to be a learning community where grace is foundational, truth is pursued, and holiness is a way of life.

### Department Mission

The Mathematical, Information, and Computer Sciences department at Point Loma Nazarene University is committed to maintaining a curriculum that provides its students with the tools to be productive, the

passion to continue learning, and Christian perspectives to provide a basis for making sound value judgments.

## **Course Description**

This course follows the complete data science process. Students will work in teams to scope a real-world problem, gather data to answer the question, wrangle the data, model it, validate the models, draw conclusions, and communicate results. The course includes study of the principles of data science and technical communication.

Letter grade.

**Prerequisite(s):** MTH 3083, Junior Standing, and consent of instructor.

## **Program and Course Learning Outcomes**

Student Outcomes:

1. Understand the fundamental concepts of Python programming language.
2. Manipulate and analyze data using computational tools such as NumPy, Pyplot, and Pandas.
3. Identify the stages of a machine learning project.
4. Implement linear and polynomial regression models and evaluate them using appropriate metrics.
5. Implement classification models and evaluate them using appropriate metrics.
6. Explain and apply the concepts of ensemble learning.
7. Select the appropriate clustering method for a dataset based on problem-specific metrics.
8. Create a comprehensive project integrating various machine learning techniques covered throughout the course.

Program Outcomes:

- An ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics. (LO1)
- An ability to communicate effectively with a range of audiences. (LO3)
- An ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions. (LO6)
- An ability to acquire and apply new knowledge as needed, using appropriate learning strategies. (LO7)

## Required Texts and Recommended Study Resources

- A. Géron. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 3rd edition. O'Reilly. 2022. ISBN: 978-1098125974.

## Course Credit Hour Information

In the interest of providing sufficient time to accomplish the stated Course Learning Outcomes, this class meets the PLNU credit hour policy for a 2-unit class delivered over 14 weeks. It is anticipated that students will spend a minimum of 37.5 participation hours per credit hour on their coursework. For this course, students will spend an estimated 75 total hours meeting the course learning outcomes.

## Assessment and Grading

This course includes three types of activities: entry tickets, homework assignments, and project presentations. Entry tickets are short summaries of key ideas that students highlight from their independent study, followed by lingering questions they have for the in-person portion of the course. Homework assignments are small practice problems aimed at reinforcing key concepts introduced early in each unit. Project presentations are real-life scenarios where the students employ an end-to-end machine learning project framework to solve a problem. Presentations are video recordings where students use client-friendly language to explain their findings about the data and proposed solutions.

The table below outlines the assessment criteria for this course.

Activity	Points Per Activity	Quantity	Total Points
Homework Assignments	50	4	200
Project Presentations	100	4	400
Entry Tickets	20	10	200
			800

Grades will be based on the following:

### Sample Standard Grade Scale Based on Percentages

<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>F</b>
A 93-100	B+ 87-89	C+ 77-79	D+ 67-69	F Less than 59

A- 90-92	B 83-86	C 73-76	D 63-66	
	B- 80-82	C- 70-72	D- 60-62	

### **Incompletes and Late Assignments**

All assignments are to be submitted/turned in by when they are due. Late assignments are deducted 20% of its grade. Incompletes will only be assigned in extremely unusual circumstances.

### **Missed Exams**

No examination shall be missed without prior consent or a well-documented emergency beyond the student’s control. A score of zero will be assigned for an examination that is missed without prior consent or a well-documented emergency beyond the student’s control. If a student misses an online test, any attempt to complete it outside of the classroom will be considered an act of academic dishonesty and will nullify the test score as well as disciplinary actions.

### **PLNU Attendance and Participation Policy**

Regular and punctual attendance at all class sessions is considered essential to optimum academic achievement. If the student is absent for more than 10 percent of class sessions, the faculty member will issue a written warning of de-enrollment. If the absences exceed 20 percent, the student may be de-enrolled without notice until the university withdrawal date or, after that date, receive an “F” grade.

### **Artificial Intelligence (AI) Policy**

Use of Artificial Intelligence (AI) tools (e.g, ChatGPT, iA Writer, Marmot, Botowski) is not permitted, and use of these tools will be treated as plagiarism.

### **PLNU Copyright Policy**

Point Loma Nazarene University, as a non-profit educational institution, is entitled by law to use materials protected by the US Copyright Act for classroom education. Any use of those materials outside the class may violate the law.

### **PLNU Academic Honesty Policy**

Students should demonstrate academic honesty by doing original work and by giving appropriate credit to the ideas of others. Academic dishonesty is the act of presenting information, ideas, and/or concepts as one’s own when in reality they are the results of another person’s creativity and effort. A faculty

member who believes a situation involving academic dishonesty has been detected may assign a failing grade for that assignment or examination, or, depending on the seriousness of the offense, for the course. Faculty should follow and students may appeal using the procedure in the university Catalog. See [Academic Policies](#) for definitions of kinds of academic dishonesty and for further policy information.

## **PLNU Academic Accommodations Policy**

PLNU is committed to providing equal opportunity for participation in all its programs, services, and activities in accordance with the Americans with Disabilities Act (ADA). Students with disabilities may request course-related accommodations by contacting the Educational Access Center (EAC), located in the Bond Academic Center (EAC@pointloma.edu or 619-849-2486). Once a student's eligibility for an accommodation has been determined, the EAC will work with the student to create an Accommodation Plan (AP) that outlines allowed accommodations. The EAC makes accommodations available to professors at the student's request.

PLNU highly recommends that students speak with their professors during the first two weeks of each semester/term about the implementation of their AP in that particular course. Accommodations are not retroactive so clarifying with the professor at the outset is one of the best ways to promote positive academic outcomes.

Students who need accommodations for a disability should contact the EAC as early as possible (i.e., ideally before the beginning of the semester) to assure appropriate accommodations can be provided. It is the student's responsibility to make the first contact with the EAC. Students cannot assume that because they had accommodations in the past, their eligibility at PLNU is automatic. All determinations at PLNU must go through the EAC process. This is to protect the privacy of students with disabilities who may not want to disclose this information and are not asking for any special accommodations.

## **State Authorization**

State authorization is a formal determination by a state that Point Loma Nazarene University is approved to conduct activities regulated by that state. In certain states outside California, Point Loma Nazarene University is not authorized to enroll online (distance education) students. If a student moves to another state after admission to the program and/or enrollment in an online course, continuation within the program and/or course will depend on whether Point Loma Nazarene University is authorized to offer distance education courses in that state. It is the student's responsibility to notify the institution of any change in his or her physical location. Refer to the map on [State Authorization](#) to view which states allow online (distance education) outside of California.

## **Use of Technology**

In order to be successful in the online or hybrid environment, you'll need to meet the minimum technology and system requirements; please refer to the *Technology and System Requirements*

information. Additionally, students are required to have headphone speakers, microphone, or webcams compatible with their computer available to use. Please note that any course with online proctored exams requires a computer with a camera (tablets are not compatible nor allowable) to complete exams online.

Problems with technology do not relieve you of the responsibility of participating, turning in your assignments, or completing your class work.

#### Semester Schedule Outline

<b>Date</b>	<b>Topic</b>
<b>3-Sep</b>	Welcome
<b>5-Sep</b>	Introduction to Python and Jupyter
<b>10-Sep</b>	Basic Python Review
<b>12-Sep</b>	Numpy and Pyplot
<b>17-Sep</b>	Numpy and Pyplot Review
<b>19-Sep</b>	Pandas
<b>24-Sep</b>	Pandas Review
<b>26-Sep</b>	Project 1
<b>1-Oct</b>	Introduction to Machine Learning
<b>3-Oct</b>	Linear Regression
<b>8-Oct</b>	Linear Regression Review
<b>10-Oct</b>	Polynomial and Regularized Regression
<b>15-Oct</b>	Polynomial and Regularized Regression Review
<b>17-Oct</b>	Machine Learning Projects
<b>22-Oct</b>	Project 2
<b>24-Oct</b>	No class - Fall break
<b>29-Oct</b>	Classification
<b>31-Oct</b>	Logistic and Softmax Regression
<b>5-Nov</b>	Logistic and Softmax Regression Review
<b>7-Nov</b>	Support Vector Machines
<b>12-Nov</b>	Support Vector Machines Review
<b>14-Nov</b>	Decision Trees
<b>19-Nov</b>	Decision Trees Review

<b>21-Nov</b>	Ensemble Learning, Boosting, and Stacking
<b>26-Nov</b>	Project 3
<b>28-Nov</b>	No class - Thanksgiving
<b>3-Dec</b>	PCA
<b>5-Dec</b>	Clustering
<b>10-Dec</b>	Clustering Review
<b>12-Dec</b>	Project 4
<b>17-Dec</b>	Comprehensive Project Presentations