General Course Information

Course Title: **DATA 5000 Data Science Seminar**

Credit Weight: 0.5 credits

Prerequisites: None

Course Description

Cloud-based distributed systems, statistics, machine learning, use of complex ecosystems of tools and platforms, data ethics, and communication skills to explain advanced analytics. Students choose a project in Big Data management and/or analysis, deliver a paper and give a class presentation on their findings.

Course Rationale

Students will examine the role of data science in a modern organization and will be able to apply the best practices-based methodology and techniques to solve real-world data science problems.

Course Learning Objectives

After successfully completing this course, you will be able to:

Explain the role of data science and data scientists in a modern organization.

Apply best practices-based methodology to complete data science projects.

Apply data pre-processing techniques.

Apply data visualization and exploration techniques to explore and analyze data.

Apply machine learning algorithms to solve data science problems.

Course Schedule

Module	Topics	Materials	Deliverable
Getting started	Getting started. The course plan.	N/A	Introduce yourself discussion due on Day 7 of Week 1.
Setting the Stage	Current industry trends. Data science and AI pitfalls. Data science life cycle. The role of the data scientist.	N/A	Paper selection due on Day 7 of Week 1.
It's All About the Data	The evolution of data analytics. Accelerator technologies for analytics.	N/A	Business vs. operational intelligence discussion due on Day 7 of Week 2. Traditional vs. cloud computing discussion due on Day 7 of Week 2.
Working with Data	Accessing data. Data pre-processing. Visualizing data.	N/A	Data exploration and visualization discussion due on Day 7 of Week 3. Paper review presentation due on Day 7 of Week 3.
Introduction to Machine Learning	Introduction to machine learning. Evaluating machine learning model performance.	N/A	Evaluating machine learning performance discussion due on Day 7 of Week 4. Types of machine learning discussion due on Day 7 of Week 4.
Machine Learning Algorithms 1	Machine learning algorithms 1.	N/A	Best algorithm for the job 1discussion due on Day 7 of Week 5.

Module	Topics	Materials	Deliverable
Machine Learning Algorithms 2	Machine learning algorithms 2.	N/A	Best algorithm for the job 2 discussion due on Day 7 of Week 6.
Course Project	Your course project.	Computational resources as needed.	Project ideas discussion due on Day 7 of Week 3. Project proposal due on Day 7 of Week 3. Project presentation due on Day 7 of Week 6. Project report due on Day 7 of Week 7.

Learning Materials

Textbook

N/A

Other Resources

Several of the course modules include sub-modules containing links to extra self-help learning resources. These resources are proven to help students build extra skills relevant to their projects. The use of these resources is completely optional.

If, while planning the course project, it becomes clear that additional computational resources are required, the student is encouraged to reach out to the instructor to see if access to additional resources can be arranged.

Late Assignments

Late assignments are not accepted for any reason and will earn a grade of 0 on that assignment. Requests for extension will be considered in cases of illness, family emergency, or other exceptional circumstances.