

## **MSE 477/542: Data Science and Materials Informatics**

### **Course Description: Official Version for UW course catalog.**

Introduction to data science approaches and their applications to materials science and engineering. Basic skills in data processing, data visualization, statistical inference, and machine learning for materials research topics taught through case studies and other methodologies.

### **Course Content and Purpose**

The purpose of this course is to introduce how data science approaches can be applied to materials science and engineering. Students will learn basic skills of data processing, data visualization, statistical inference, and machine learning through case studies and exercises that cover the routine procedure of utilizing data science approaches on materials research topics. Examples and exercises will start with the scraping and processing of data and building a data set for input, and end with using machine learning methods for data analysis and visualization. The class will include online contents to help students refresh their knowledge of python. In-class lectures will focus on case studies and materials informatics.

### **Course Prerequisites**

CSE 160 or MSE 546 or MSE 547, or equivalent courses, or proof of proficiency in Python or another programming language.

### **Student Performance Evaluation**

Each students' final grade in the course will be based on 2 elements: Homework assignments that focus on application of the data science methods, and a light-weight final course project

- Homework assignments: 70%  
Assignments will be submitted online. Discussions among peers are encouraged while “copy-paste” is forbidden.
- Final project:30%  
Grades will be assigned based on a short report.

### **Learning objectives**

Upon successful completion of the class, students will demonstrate the ability to

- Identify different data types used in materials science, and the data structure for each type of data.

- Use basic data processing techniques to manipulate and clean experimental or simulated materials data to prepare a dataset that can be further studied by using machine learning methods.
- Choose and apply appropriate supervised and unsupervised learning models such as linear regression, logistic regression, clustering, dimensionality reduction, K-NN, pipeline, etc., to different types of materials data.
- Perform scientific and technical analyses to the machine learning results and use data visualization tools on those results.
- Use critical thinking and knowledge learned from this class to complete a data science project related to materials science and engineering.

### **Course format and tentative schedule**

This will be a 3-credit course with meetings 3 times per week, 2 lectures and 1 hands-on section. Topics that will be covered by the course including:

- provide basic python language overview, introduce basic packages such as Numpy, SciPy, Matplotlib, Pandas, Pickle and joblib, and computational environment such as Jupyter notebook with a focus of data science application using course provided online resources.
- Data challenges and opportunity of materials informatics. Understanding the different data types in materials research.
- Introduction of main data types of materials research and possible data methods associate to each type
- Introduction Existing common data bases for materials informatics.
- Introduction to supervised learning such as Neural networks, Tree, KNN or SVM,
- Introduction to unsupervised learning such as linear regression, PCA, K-means clustering, SOM
- Data scraping and processing. Introduce to natural language process machine learning tools, such as SciBert, and its' application in building a dataset
- Case study of using data science methods to get insights for materials research, reveal materials information using clustering, and visualization of data

### **Requesting Religious Accommodation**

Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW's policy, including more information about how to request an accommodation, is available at [Religious Accommodations Policy \(https://registrar.washington.edu/staffandfaculty/religious-accommodations-policy/\)](https://registrar.washington.edu/staffandfaculty/religious-accommodations-policy/). Accommodations must be requested within the first two weeks of this course using the [Religious Accommodations Request form \(https://registrar.washington.edu/students/religious-accommodations-request/\)](https://registrar.washington.edu/students/religious-accommodations-request/).”