**DATA MINING: With Applications in the Life Sciences**

Computer Science 4170/5170 (class number: 2138/2139, section: 100/100)

Monday and Wednesday, 9:40 AM to 10:35 AM, [Academic & Research Center 321](http://www.ohio.edu/registrar/info/buildingmaps/ARC.jpg)

Dr. Welch, [welch@ohio.edu](mailto:welch@ohio.edu); office hours: Monday and Wednesday, 10:40 AM to 11:30 AM (via Teams)

**Introduction and Learning Outcomes[[1]](#footnote-1)**:

This course will prepare students to construct sophisticated data mining software for gaining important insights from data sets in a variety of contexts, including genomics. Students will be able to implement the processes and methods of data mining in order to discover new knowledge from available data sets. Students will be able to use knowledge from the domain of genomics throughout all phases of the data mining process.

**Assessment[[2]](#footnote-2),[[3]](#footnote-3)**:

* 15% - Class participation (attendance is mandatory)
* 25% - Quizzes (20% for CS 5170)
* 40% - Data mining activities
* 20% - Wikipedia article editing (25% for CS 5170)
* Deadline extensions will not be granted, except for legitimate reasons (see the student handbook).
* Grades will be assigned as follows:

|  |  |
| --- | --- |
| A-, A | 90-93, 94-100% |
| B-, B, B+ | 80-83, 84-86, 87-89% |
| C-, C, C+ | 70-73, 74-76, 77-79% |
| D-, D, D+ | 60-63, 64-66, 67-69% |

**Textbook (online):**

Introduction to data mining, Discovering Knowledge in Data. D.R. Larose and C.D. Larose. Wiley. 2014. <https://alice.library.ohio.edu/record=b5187242?> (click on the link “Connect to resource OhioLink”)

**Course Format and Organization**:

Students will learn course concepts by implementing data mining algorithms, resulting in a progressively sophisticated software pipeline. The sequence of major topics is shown in the table below. *Each week* you will be introduced to new concepts via readings and/or videos, and a pre-recorded introductory lecture by Professor Welch. The materials, a list of specific concepts to learn, and a pre-recorded introductory lecture will be provided every *Friday*. After you have learned the concepts, you will take an online quiz (to be completed no later than the start of the class session on *Monday*). Interactive lecture sessions will be held in the classroom at the scheduled class times, on *Mondays* and *Wednesdays*. You will employ the concepts by performing data mining tasks on an actual genomics data set from a cancer research project of the National Institutes of Health. The results of your data mining activities will be submitted by email each *Thursday*. The Blackboard system[[4]](#footnote-4) will hold all items related to the course: links to readings and videos, lists of specific learning objectives, introductory lectures, quizzes, data mining project assignments, data sets, class grades and syllabus.

|  |  |
| --- | --- |
| **Weeks** | **Topics** |
| **1-3** | exploratory data analysis |
| **4-6** | classifiers |
| **7-9** | classifierevaluation |
| **10-12** | feature selection and feature association |
| **13-15** | final project |

### Classroom COVID-19 policies

Ohio University is committed to maintaining a safe learning environment for all members of its community – students, faculty, and staff – regardless of whether they have been vaccinated against COVID-19.

Because the university is not requiring vaccination nor are faculty or staff permitted to ask students about their vaccination status, all members of the OHIO community [are expected to take certain precautions](https://www.ohio.edu/coronavirus/presidential-health-directive) in classrooms and elsewhere on campus to keep each other safe:

* Select a testing pathway as part of the [COVID-19 Testing Pathway Program](https://www.ohio.edu/coronavirus/testing-pathway-program). If you have not already made your pathway selection, please do so immediately. If you select the Weekly Testing Pathway, I expect you to adhere to your required testing schedule.
* Maintain physical distance as directed on signage in classrooms, vaccinated or not.
* All individuals are required to wear masks in indoor public spaces on campus, including classrooms.
* Complete a daily [COVID-19 symptom assessment](https://ohio.qualtrics.com/jfe/form/SV_3CWnZed0IFjRzx3) and do not come to class if you feel sick, even if you are vaccinated against COVID-19. If you need a thermometer, you can request one at the Guest Services desk on the fourth floor of Baker Center.
* If you test positive, feel sick or suspect exposure to COVID-19, it is your responsibility to follow the [OHIO COVID-19 Protocol](https://www.ohio.edu/coronavirus/protocol).

If you do test positive or need to isolate or quarantine this semester, after you have taken care of yourself and followed all the steps in the [OHIO COVID-19 Protocol](https://www.ohio.edu/coronavirus/protocol), please email me so that we can develop a plan for you to receive necessary course content. COVID-related illness, quarantine, isolation, and remain-in-room directives are legitimate university absences, and I will work with you to manage your academic requirements and connect you to resources.

Approaching your professor with one-on-one questions, to avoid crowding:

* approach one at a time,
* maintain physical distance if a line forms, and
* use digital means for one-on-one conversations (virtual office hours via Teams, email).

If you feel that your class performance is being impacted by COVID-19, please talk with me and/or contact COVID Operations by phone (740.566.8445) or email ([COVIDoperations@ohio.edu](mailto:COVIDoperations@ohio.edu)) Monday through Friday, 8 a.m. to 8 p.m., and noon to 5 p.m. on weekends. The University has resources available to help with quarantine and isolation support, as well as access to COVID-19 testing, counseling services, food assistance, and more.

1. Implements the following ABET outcomes: (b) An ability to analyze a problem, and identify and define the computing requirements appropriate to its solution. (c) An ability to design, implement, and evaluate a computer-based system, process, component, or program to meet desired needs. [↑](#footnote-ref-1)
2. Accommodations for Students with Disabilities: Any student who suspects s/he may need an accommodation based on the impact of a disability should contact the class instructor privately to discuss the student’s specific needs and provide written documentation from the Office of Student Accessibility Services (OSAS). If the student is not yet registered as a student with a disability, s/he should contact the OSAS. [↑](#footnote-ref-2)
3. Individuals performing plagiarism, copying (e.g., using software code written by another student) and other forms of *academic misconduct* (see student handbook for further details) will receive an ‘F’ in the course and referral to Ohio University Judiciaries. [↑](#footnote-ref-3)
4. Access the Blackboard system at [https://blackboard.ohio.edu](https://blackboard.ohio.edu/) and look under “My Courses” for this course. [↑](#footnote-ref-4)