

**Course Number: 692X**

**Course Name: Seminar: machine learning on biological sequence data**

**Meeting Days** Monday, 2:30-3:30pm

**Classroom: CS 140**

Capacity: 12

**Credits: 1**

**Instructor:**

Anna Green

[annagreen@umass.edu](mailto:annagreen@umass.edu)

Office: CS 348

Pronouns: she/her

**TA:** None

**Office Hours:** by appointment

**Course Description:** A seminar in which students will read, present, and discuss research papers on recent and advanced topics in computational biology, specifically related to machine learning models fit to biological sequence data (proteins and DNA). This semester, the seminar will primarily cover the following topics: foundation models of DNA and protein sequences (including transformer-based models), predicting the effects of biological mutations, predicting the structure of proteins (including AlphaFold), and supervised vs. unsupervised learning on sequences. Students are expected to read up to two papers per week. For one or more sessions in the semester, students are expected to make summary presentations and lead discussion of the papers. Students should have taken COMPSCI 690U, Computational Biology and Bioinformatics, or have comparable background. 1 credit.

**Justification:** There has been a recent explosion in application of deep learning models to computational biology problems. This course will discuss the foundational and recent papers in the field, positioning students for research projects in the area, as well as providing valuable experience discussing and presenting research. As the course covers the particular challenges of working with biological sequence data, it will broaden the perspective of core ML students, who are likely more familiar with either image or natural language data.

**Course URL:** TBD

**Grading basis:** Optional Grading

**Eligibility:** CS Grads only, will accept PhDs from related disciplines with override

**Prerequisites:** No formal prerequisites

**Class Notes:** A background in computational biology and bioinformatics is required, such as found in COMPSCI 690U. Students needing special permission must request overrides via the on-line form: <https://cics.umass.edu/overrides>

**Textbook:** None (Required readings will be free electronic materials provided by the instructor, including pdfs of journal and conference articles and pdfs of free textbooks)

