2022-2023

[Course Number] – Applications of machine learning in biology

Semester B

Time: [Day] [From-Until], Room [Room Number]

Instructor: Dr. David Deutsch, **Email:** ddeutsch1@univ.haifa.ac.il

Office Hours: Tuesday, 9-10 (Dr. Deutsch's office). Outside of these hours – by appointment only. Please do not arrive outside the office hours without coordinating an appointment.

Course Level: MA+PhD

Course Type & Format: Elective.
Attendance is not mandatory. However, students are strongly encouradged to attend all classes.

Number of Hours/Credits: 2

Prerequisites: Basic programming skills in Python or Matlab and basic knowledge in statistics. If you have programming experience in another language and like to take the course, please contact the instructor.

Course Overview (Short Abstract):

This course gives an overview on machine learning, focusing on some specific methods that are commonly used by biologists. No previous knowledge in machine learning is assumed, but students are expected to have basic programming skills to be able to complete the home assignments. Basic knowledge in statistics is also assumed.

We will start by refreshing some basic concepts in probability and statistics. We will then talk about regression models, dimensionality reduction (linear and non-linear) and learning (supervised and unsupervised). Last, we will focus on neural networks, reviewing the basic math and practical uses.

In each topic we cover we will first introduce the basic concepts, and then will look into specific applications in biology. While the examples are given mostly from datasets that come from the field of neurobiology, applying the same techniques to other types of datasets should be straight forward.

Learning Outcomes (What are the skills, abilities, or major concepts a student is expected to acquire in this course?) – At the end of the course students will be able to:

1. Understand basic concepts in machine learning.
2. Be familiar with specific methods for regression, dimensionality reduction, learning and neural networks.
3. Be able to apply these methods, using programming, to scientific problems.

Assessment (Assessment Method and Grade Composition):

Homework assignments - 50%

Final project – 50%

Week-by-Week Content and Assignments:

Week # Date	Topic	Assignment
1 2.3	Probability and statistics I	Home assignment 1: statistics, regression
2 9.3	Probability and statistics II, intro to ML	
3 16.3	Linear regression, multiple regression	
4 23.3	Generalized linear models (GLM) - I	Logistic regression ?
5 30.3	GLM - II	Poisson regression
6.4	PASSOVER	No class

6 13.4	Unsupervised learning - Dimentionality reduction – I (PCA)	Home assignment 3: dimentionality reduction
7 20.4	Unsupervised learning – non-linear dimentionality reduction – II (t-SNE/UMAP?)	Alon Rubin (Weizmann)
8 27.4	Unsupervised learning - clustering (k-means ?)	Oren Forkosh (HUJI)
9 4.5	Supervised learning- classification (KNN)	Home assignment 3: classification
10 11.5	Supervised learning - classification (Support vector machine)	
11 18.5	supervised learning - classification (decision trees)	Final project - Due 25.6
25.5	SHAVUOT	No class
12 1.6	Neural networks - introduction	Ogen Drukerman (U Haifa)
13 8.6	Recurrent neural networks	Omri Barak (Technion)
14 15.6	Deep neural network applications : text processing, image processing	Ogen/Dudi
Week of 25.6	Defending final projects	

Course book: “Hands-on Machine Learning with Scikit-Learn, Kares & TensorFlow” by Aurelien Geron (3rd edition, 2022).