**Course Outline - DEEP LEARNING METHODS IN POPULATION GENOMICS AND PHYLOGEOGRAPHY**

**Daily schedule (2-8PM Berlin time)**

2 - 3.15 Session 1

3.15-3.30 BREAK

3.30 -4.45 Session 2

4.45 - 5.15 BREAK

5.15 - 6.30 Session 3

6.30-6.45 BREAK

6.45 - 8 Session 4

**Day 1 - Monday (February 20th, 2023)**

**Part 1 (Manolo)**

- Intro to course.

- Basic concepts on simulation-based likelihood free methods.

**Part 2 (Manolo)**

- Modelling demographic history with genetic data simulators.

**Part 3 (Manolo)**

- Practical: conceiving demographic scenarios.

**Part 4 (Manolo)**

- Practical: building a script to simulate genetic data under competing demographic scenarios.

*Useful resources:*

Hahn, M. W. (2018). *Molecular population genetics*. Sunderland, MA: Sinauer Associates. ISBN 978-0878939657.

ms software and manual: <http://home.uchicago.edu/~rhudson1/source/mksamples.html>

stdpopsim: <https://popsim-consortium.github.io/stdpopsim-docs/stable/index.html>

**Day 2 - Tuesday (February 21st, 2023)**

**Part 1 (Manolo)**

- A gentle introduction to supervised deep learning.

**Part 2 (Manolo)**

- Understanding the basic CNN architecture for image recognition.

**Part 3 (Manolo)**

- Practical: Image classification with CNN.

**Part 4 (Manolo)**

- Using CNN to learn directly from genetic data.

*Useful resources:*

- 3Blue1Brown (videos for understanding how Deep Learning works): <https://www.youtube.com/watch?v=aircAruvnKk&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=1>

Fastai course: <https://course.fast.ai/>

**Day 3 - Wednesday (February 22nd, 2023)**

**Part 1 (Matteo)**

- Introduction to approaches for detecting selection.

**Part 2 (Matteo)**

- Recognizing signatures of selection with deep learning.

**Part 3 (Matteo)**

- Practical: simulating genetic data and using CNN to predict whether a given locus is under selection.

*Useful resources:*

Sheehan S, Song YS. Deep learning for population genetic inference. PLoS computational biology. 2016 Mar 28;12(3):e1004845. <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004845>

Torada L, Lorenzon L, Beddis A, Isildak U, Pattini L, Mathieson S, Fumagalli M. ImaGene: a convolutional neural network to quantify natural selection from genomic data. BMC bioinformatics. 2019 Nov;20:1-2. <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-019-2927-x>

**Day 4 - Thursday (February 23rd, 2023)**

**Part 1 (Manolo)**

- The building blocks of a CNN script.

**Part 2 (Manolo)**

- Practical: Comparing demographic scenarios with deep learning.

**Part 3 (Manolo)**

- Quick overview of other applications .

- Future perspectives.

- Course wrapup.

*Useful resources:*

Kirschner P, Perez MF, Záveská E, Sanmartín I, Marquer L, Schlick-Steiner BC, Alvarez N, STEPPE Consortium, Steiner FM, Schönswetter P. Congruent evolutionary responses of European steppe biota to late Quaternary climate change. Nature Communications. 2022 Apr 8;13(1):1921. <https://www.nature.com/articles/s41467-022-29267-8>

Korfmann K, Gaggiotti OE, Fumagalli M. Deep learning in population genetics. Genome Biology and Evolution. 2023 Jan 23:evad008. <https://academic.oup.com/gbe/article/15/2/evad008/6997869>

For more advanced courses:

<https://www.coursera.org/specializations/deep-learning>. This specialization goes more deep into the math of Deep Learning.

Fastai course. This is a more practical course and is accompanied by a book: <https://course.fast.ai/>

The EvoGenomics network. Where you can get updates on interesting Deep Learning applications and has periodical talks. <https://www.evogenomics.ai/>

Stanford course on machine learning with Graphs. <https://www.youtube.com/watch?v=JAB_plj2rbA&list=PLoROMvodv4rPLKxIpqhjhPgdQy7imNkDn>