Ph. D.

IT-770

09. Neural networks and deep learning for biological data:

Pre-requisites: Students opting for this course must be familiar with basics of machine learning and artificial intelligence.

Objective: This course is aimed at introducing cutting edge artificial neural network techniques a applied to biological problems. Different types of neural networks, their training and validatio strategies and limitations will be discussed. Specific examples from published literature such a sequence specificity predictions will be taken up.

Syllabus:

Background of machine learning methods: Regression and classification (2 hrs)

Overview of discretization and data reprsentation techniques (3 hrs)

Introduction to neural networks architecture and training methods: Mulilayer perceptron backpropagation, quickpropagation and other learning methods. Batch and online training. (5 hrs)

Neural network applications in Bioinformatics: Secondary structure, solvent accessibility, side chain orientation predictions. Gene ontology predictions. Functional annotation prediction using neural networks and other machine learning methods. (5 hrs)

Introduction to Time delay neural networks, RBF neural networks. Self organizing maps and Kohonen learning. (3 hrs)

Fundamentals of Deep learning strategies. Boltzman machine and auto-encoders. Denoising autoencoders. Amino acid autoencoder and dimensionality assessment. (4 hrs)

Inferring nucleic acid sequence specificity and other biological applications of deep learning (2 hrs)

Suggested readings:

1. Deep Learning Made Easy with R: A Gentle Introduction for Data Science CreateSpace Independent Publishing Platform (ISBN-13: 978-1519514219)
2. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning, Alipanahi B, Delong A, Weirauch MT, Frey BJ, Nature Biotechnol. 2015 Aug;33(8):831-8. doi: 10.1038/nbt.3300.
3. Deep learning for regulatory genomics, Park Y, Kellis M, Nature Biotechnol. 2015 Aug;33(8):825-826. doi: 10.1038/nbt.3313.
4. Reversible auto-encoding of amino-acid residues in reduced space: an application to predicting DNA-binding proteins, Shandar Ahmad, Third IAPR International Conference

on Pattern Recognition in Bioinformatics (PRIB) (2008, Melbourne, Australia) Monash University Research Repositories (http://arrow.monash.edu.au/hdl/1959.1/63725)

5. Dimensionality of amino acid space and solvent accessibility prediction with neural networks,Marcos J. Araúzo-Bravo , Shandar Ahmad, , Akinori Sarai, Computational Biology and Chemistry, 2006, 30, 160–168