**Lecture Series: "Bayesian statistics, Data Mining and Machine Learning”**

Lecture series composed of two “mini courses” in May '23, and Dec ’23 to Jan ’24,

at the Rudjer Boskovic Institute (Zagreb, Croatia).

Aim

For whom

Where

Modality

Prerequisites

**Mini-course 1: “Practical introduction to Bayesian statistics”**

6 lectures: **May** 4, 8, 11, 15, 22, 29,**2023**

**Topics:**

a) **Introduction to Maximum Likelihood method**

         - parameter estimation for normal distribution

         - parameter estimation for Cauchy (Lorentz) distribution

         - beating sqrt(N) improvement for uniform distribution

         - straight line fit with errors on both axes

         - polynomial fit: which order to take?

         - penalized likelihood regression

         - Gaussian Process regression

         - over-fitting, under-fitting and cross-validation

b) **Introduction to Bayesian statistics**

         - basic concepts, priors, credible regions

         - simple examples of parameter estimation

         - simple examples of model selection

         - hierarchical Bayes: see

               https://github.com/astroML/astroML-workshop\_AAS235/blob/master/hierarchical\_bayes/AAS2019\_HBexample.ipynb

         - perhaps ABC? see

               https://github.com/astroML/astroML-workshop\_AAS235/blob/master/approximate\_bayesian\_computation/AAS2019\_ABCexample.ipynb

c) **MCMC and its applications in Bayesian statistics**

         - non-linear function regression

         - model selection (examples: polynomials, burst shapes in time series)

Definitely watch this before teaching Bayes

https://www.youtube.com/playlist?list=PLKW2Azk23ZtQSHmwOpObPEr58Pe1rpIdB

**Mini-course 2: “Introduction to Data Mining and Machine Learning”**

6 lectures: 3 in December 2023, and 3 in January 2024

**Topics:**

a) **Density estimation and clustering**

         - non-parametric vs. parametric estimation

         - Knuth histograms

         - Bayesian Blocks algorithm

         - K-means clustering

         - Gaussian Mixtures model

b) **Unsupervised and supervised classification**

         - generative vs. discriminative classification

         - evaluating success: ROC curves

         - K-Nearest Neighbor Classifier

         - decision trees

         - neural networks

c) **Dimensionality reduction**

         - the curse of dimensionality

         - Principal Component Analysis

         - Non-negative Matrix Factorization

         - manifold learning

d) **Deep Learning (Convolutional Neural Networks)**

         - introduction to neural networks

         - using CNNs to classify images