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

   A lecture series at the Ruđer Bošković Institute (IRB) composed of two “mini-courses”.

By Zeljko Ivezic, University of Washington, a Fulbright Visiting Scholar to IRB.

**Aim:**to discuss modern methods in Bayesian statistics, Data Mining and Machine Learning

      with researchers at IRB and students from the University of Zagreb.

**For whom:**graduate students and young researchers, in exceptional cases undergraduate

           students and senior researchers, focused on quantitative data analysis in physics,

           chemistry, biology and related fields.

When: May 2023 and January/February 2024

**Where:** at IRB (exact location TBD).

**Modality:**lectures will be interactive, and based on jupyter python notebooks.

**Prerequisites:** at least basic undergraduate background in mathematics and statistics, and at least

a rudimentary experience with python. Attending Mini-course 1 is not a prerequisite for attending Mini-course 2, though it is highly recommended.

**The class size is limited to 25 attendees.**Exact lecture times (to be chosen from 10:30-12:00, 13:30-15:00 and 14:30-16:00) will be decided by polling those who sign up.**To sign up, please send email to ivezic@uw.edu by Mar 17, 2023**, with a Subject line: “IRB Lecture Series”.

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

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

**Topics:**

**1)  Introduction to Maximum Likelihood Estimation method**

   - robust statistics

         - parameter estimation with MLE

         - linear regression with heteroscedastic uncertainties

         - cost functions and penalized likelihood

         - non-gaussian likelihoods

         - conceptual difficulties with MLE

**2) Introduction to Bayesian statistics and inference**

         - basic concepts, priors, nuisance parameters, credible regions

         - simple examples of parameter estimation

         - simple examples of model selection

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

         - hierarchical Bayes

         - ABC (approximate Bayesian computation)

**3) Markov Chain Monte Carlo and its applications in Bayesian statistics**

         - non-linear regression

         - model selection

         - Bayesian Blocks algorithm

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

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

**Topics:**

**1) Density estimation and clustering**

         - non-parametric vs. parametric estimation

         - Knuth histograms

         - Bayesian Blocks algorithm

         - K-means clustering

         - Gaussian Mixtures model

**2) Unsupervised and supervised classification**

         - generative vs. discriminative classification

         - evaluating success: ROC curves

         - K-nearest neighbor classifier

         - decision trees

         - neural networks

**3) Dimensionality reduction**

         - the curse of dimensionality

         - Principal Component Analysis

         - Non-negative Matrix Factorization

         - manifold learning

**4) Deep Learning**

         - introduction to neural networks

         - using Convolutional Neural Networks to classify images