1ST PAN AFRICAN CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SMART SYSTEMS
Windhoek, Namibia – OCTOBER 7-9, 2021

Workshop/Mini Course/Extended Tutorial

Title

On the Ubiquity of the Bayesian Paradigm in Artificial Intelligence and Statistical Machine Learning

Author

Ernest Fokoué
School of Mathematical Sciences
Rochester Institute of Technology
Rochester, New York,
USA

Duration and Preferred Time of Day

4 hours, preferably in the morning

Prerequisites

Calculus 2
Vector and matrix algebra
Basic Probability
Basic Statistics
Basic Algorithmics
Basic Knowledge of R and/or Python
Eagerness and earnestness to learn
Familiarity with data manipulation

Abstract

The Bayesian paradigm directly or indirectly permeates almost every single aspect of the building blocks of statistical machine learning, providing a natural and almost quintessential framework for reasoning about and solving all kinds of artificial intelligence and data science problems. Mindful of this all-pervading influence and natural appeal, this four hour workshop/tutorial aimed at both practitioners and methodologists, will travel the length and breadth of the Bayesian school of thought as pertaining to its use in artificial intelligence, with the finality of highlighting and evidencing its ubiquity, but even more crucially its appealing and compelling usefulness and practicality in applications, computation, methodology and theory. This extended tutorial will unfold at the rhythm of the four aspects of statistical machine learning, namely:

- (A) Applications [1st hour]
- (M) Methodology [2nd hour]
- (T) Theory [3rd hour]
- (C) Computation [4th hour]

Objectives

Upon completing this extended tutorial, the participant shall not only appreciate the ubiquity and practical usefulness of the Bayesian paradigm in statistical machine learning and artificial intelligence, but also build/construct a formidable arsenal replete with powerful tools for tackling and solving a wide variety of impactful problems.

Applications

Focus: Infusion and cultivation of the awareness of the importance of clear formulation of problems What class of problems are typically tackled and solved in artificial intelligence? Weak AI? Strong AI? Learnability?

Motivating applications of a compellingly appealing type will be used to set things in motion, with a primary focus on the critical importance of good formulation, along the role of the Bayesian thought as an aid in the process of clarifying the key aspects of the task at hand.

Computer Vision
Image Analysis/Processing
Automatic Speech Recognition
Speech Synthesis
Automatic Language Translation
Text Analysis
Video Analysis
Autonomous navigation

Methodology

Focus: Identification of existing or emerging or even totally novel paradigms consistent or compatible with the formulation created or designed.

- Model class or Function Space Identification or Specification
- Model Estimation and Learning Algorithms and Learning Machines Model selection and model refinement
- Model aggregation

An excursion into methodology will invite an exploration of a class of function spaces along with the vast array of potential algorithms, methods and techniques most compatible and most suitable to the formulation arrived at, with once again a harnessing of the power of the Bayesian machinery in the context of model building and function estimation. This second module will put the spotlight on elements of estimation, inference, prediction, forecasting and simulation, all of which are central to state of the art artificial intelligence. Algorithmic and Methodological derivations with allusions to well-known methods like:

Gaussian Mixture Models
Hidden Markov Models
Markov Random Fields
Gaussian Processes
Nearest Neighbor Methods
Kernel Methods
Neural Networks
Support Vector Machines
Random Forests
Generalized Linear Models
Generative Models
Discriminative Models
Prior specification and understanding of the function space of interest

Theory

Focus: Exploration or at least awareness of the theoretical foundations or underpinnings of the frameworks under consideration. Including amongst other topics:

Limit theorems
Inequalities
Convergence and converge rates
Bounds on the generalization error
Confidence and credible sets
Inference and computational complexity.
Relationship to the Bayes Risk
Sample Complexity
VC Dimension
Rademacher Complexity
Etc …

Computation

Focus: Thorough investigation of aspects like algorithmic stability, computational scalability and wellposedness in Hadamard's Sense

Computational aspects and consideration
Gradient descent
Newton Raphson
Stochastic gradient descent
Maximum Likelihood Estimation
Expectation-Maximization Algorithm
Markov Chain Monte Carlo
Variational Inference

Gibbs Sampler
Cross Validation
Bootstrap
Stochastic Hold Out
Etc …

-- Notes

To achieve the highest impact, it is ideal to advertise all tutorials far ahead of time and limit the number of places. For this tutorial specifically, I anticipate engaging the audience in a hands on experience with their computer, specifically encouraging them to run some of the functions along with me. To this end, it is vital to limit the size of the tutorial audience.