2EL6190 – Bayesian methods for machine learning

Instructors: Simon Leglaive
Department: CAMPUS DE RENNES
Language of instruction: ANGLAIS
Campus: CAMPUS DE RENNES
Workload (HEE): 60
On-site hours (HPE): 35,00
Elective Category : Fundamental Sciences
Advanced level : Yes

Description

Bayesian modeling, inference and prediction techniques have become commonplace in machine learning. Bayesian models are used in data analysis to describe, through latent factors, the generative process of complex data (medical images, audio, documents, etc.) The discovery of these latent or hidden variables from observations is based on the notion of posterior probability distribution, the calculation of which corresponds to the Bayesian inference step.

Let's take the example of a technique called "Latent Dirichlet Allocation" or LDA. This is a Bayesian method, which in particular is used to discover hidden topics in a set of observed documents. If we apply this technique to analyze a set of 1102 abstracts of scientific articles on Bayesian machine learning published in *Journal of Machine Learning Research (JMLR)*, the following topics emerge:

Topic #1: model models data process latent bayesian dirichlet hierarchical nonparametric inference
Topic #2: features learn problem different knowledge learning image object example examples
Topic #3: method neural bayesian using linear state based kernel approach model
Topic #4: belief propagation nodes local tree posterior node nbsp given algorithm
Topic #5: learning data bayesian model training classification performance selection prediction sets
Topic #6: inference monte carlo markov sampling variational time algorithm mcmc approximate
Topic #7: function optimization algorithm optimal learning problem gradient methods bounds state
Topic #8: learning networks variables structure network bayesian em paper distribution algorithm
Topic #9: bayesian gaussian prior regression non estimation likelihood sparse parameters matrix

Topic #10: model information bayesian human visual task probability sensory prior concept
(credits: Rémi Bardenet, https://github.com/rbardenet/bml-course/blob/master/notebooks/00_topic_modelling_for_Bayesian_ML_papers.ipynb)

Recognizable topics stand out, such as *Topic #6* on approximate Bayesian inference methods or *Topic #8* on learning in Bayesian networks.

The Bayesian machine learning approach has the advantage of being interpretable, and it makes it easy to include expert knowledge through the definition of priors on the latent variables of interest. In addition, it naturally offers uncertainty information about the prediction, which can be particularly important in certain application contexts, such as medical diagnosis or autonomous driving for example.

This course is built as a "journey towards variational autoencoders (VAEs)". Introduced in 2014, VAEs lie at the intersection of Bayesian modeling and inference techniques and deep learning with artificial neural networks. This type of model is at the heart of many current challenges in artificial intelligence (weakly-supervised learning, causality, etc.). The different sessions of this module aim at introducing fundamental notions allowing at the end an in-depth understanding of VAEs, while remaining generalizable to many other application contexts. After two sessions on the fundamentals of the Bayesian methodology and machine learning, we will study Bayesian networks and exact inference techniques for latent variable models. As exact inference is not always possible, we will move on to approximate techniques based on variational and Markov chain Monte Carlo (MCMC) methods. We will end with deep-learning-based generative models, exploiting recent variational inference methods that scale for large datasets or for high-dimensional data

Theoretical concepts will be applied on concrete data, in particular during lab sessions in Python. Different supervised and unsupervised Bayesian learning models will be implemented (Gaussian mixture model, Bayesian and sparse linear regression, variational autoencoder). These examples will allow the students to study the influence of the prior on the parameters of the model and on the obtained prediction compared to a non-Bayesian approach.

Quarter number
SG6

Prerequisites (in terms of CS courses)
Basics of statistics and probabilities. Fundamentals of machine learning: empirical risk minimization, maximum likelihood approach, supervised

learning (linear models for regression and classification), unsupervised learning (dimensionality reduction, clustering). The 1st-year course "statistics and learning" provides all these requirements.

Syllabus

Lectures:

- Fundamentals of Bayesian modeling and inference
- Fundamentals of machine learning
- Bayesian networks and inference in latent variable models
- Variational inference
- Markov Chain Monte Carlo
- Deep generative models

Lab sessions:

- Gaussian mixture model
- Bayesian linear regression

Class components (lecture, labs, etc.)

The course is organized in 7 lectures of 3 hours, 3 lab sessions of 3 hours on Python, and 1 revision session. Most of the lectures also include a short practical session (in Python) to apply the theoretical concepts. Students will be asked to do theoretical preparatory work before the lab sessions.

Grading

Students will be evaluated through lab-session reports in the form of Jupyter notebooks, including the answers to the theoretical exercises and the implementation of the algorithms. The evaluation of these reports represents 30% of the final grade, the remaining 70% correspond to a final exam of 2 hours.

Course support, bibliography

Course materials (slides, Jupyter notebooks, Python code and teaching activities) will be made available on Edunao.

References:

- Christopher M. Bishop, « *Pattern Recognition and Machine Learning* », Springer, 2006 (freely available online)

- Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong, « *Mathematics for Machine Learning* » Cambridge University Press, 2020 (freely available online)
- Kevin P. Murphy, « *Machine Learning, A Probabilistic Perspective* », MIT Press, 2012 (available at the library)

Resources
Teaching team: Simon Leglaive
Software tools: Anaconda (Python package manager).

Learning outcomes covered on the course
At the end of the course, students are expected to:

- know when it is useful or necessary to use a Bayesian machine learning approach;
- have a view of the main approaches in Bayesian modeling and inference;
- know how to identify and derive a Bayesian inference algorithm from the definition of a model;
- be able to implement standard supervised or unsupervised Bayesian learning methods.

Description of the skills acquired at the end of the course
C1. Analyze, design, and build complex systems with scientific, technological, human, and economic components
C6. Be operational, responsible, and innovative in the digital world