Bayesian ML: project topics

Julyan Arbel, Rémi Bardenet

February 7, 2020

1 Nature of the project

Students should form groups of two to three, each group undertaking one project. We suggest in Section 4 a few scientific papers that can each lead to a project, but you can choose another paper, subject to our approval.

For the paper your group will have chosen, you should: (1) explain the theoretical, computational and/or empirical methods, (2) emphasize the main points of the paper, and (3) apply it to real data of your choice when applicable. Bonus points will be considered if you are creative and add something insightful that is not in the original paper: this can be a theoretical point, an illustrative experiment, etc. The whole point is to read the paper with a critical mind.

2 Assignment of papers

As a first step, we ask each group to fill the spreadsheet at

https://lite.framacalc.org/9eu9-mva-bml

with the title of the paper, a link to it (if available), and the composition of the group. We ask that you fill in the form **before February 13**.

3 Format of the deliverable

You can use either Python or R for the programming part. Please have each group send

- one report as a pdf (≤ 10 pages) in the NeurIPS template,
- the link to a GitHub or GitLab repository containing your code and a detailed readme file with instructions to (compile/install and) run the code.

to both teachers[1] **no later than March 11**. There will be no deadline extension.

[1] if the above link is broken, this means: julyan.arbel@inria.fr, and remi.bardenet@gmail.com

4 Proposed papers

Lecture 1

[A1] P. Germain, F. Bach, A. Lacoste, and S. Lacoste-Julien. Pac-Bayesian theory meets Bayesian inference. In *Advances in Neural Information Processing Systems*, pages 1884–1892, 2016.

[A2] P. Grünwald and T. Van Ommen. Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Analysis*, 12(4):1069–1103, 2017.

[A3] P. Rigollet and A. Tsybakov. Exponential screening and optimal rates of sparse estimation. *The Annals of Statistics*, 39(2):731–771, 2011.

Lecture 2

[B1] M. Rabinovich, E. Angelino, and M. Jordan. Variational consensus Monte Carlo. In *Advances in Neural Information Processing Systems*, pages 1207–1215, 2015.

[B2] Y. W. Teh, D. Newman, and M. Welling. A collapsed variational Bayesian inference algorithm for latent Dirichlet allocation. In *Advances in neural information processing systems*, pages 1353–1360, 2007.

[B3] M. Welling and Y. W. Teh. Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 681–688, 2011.

Lecture 3

[C1] W. Chu and Z. Ghahramani. Preference learning with Gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*, pages 137–144. ACM, 2005.

[C2] Ferenc Huszár and David Duvenaud. Optimally-weighted herding is Bayesian quadrature. In *Proceedings of the conference on Uncertainty in Artificial Intelligence (UAI)*, 2012.

[C3] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951–2959, 2012.

[C4] Michalis K Titsias. Variational learning of inducing variables in sparse gaussian processes. In *International Conference on Artificial Intelligence and Statistics*, pages 567–574, 2009.

Lecture 4

[D1] Federico Camerlenghi, Antonio Lijoi, Peter Orbanz, and Igor Prünster. Distribution theory for hierarchical processes. *The Annals of Statistics*, 2018.

[D2] F. Caron and Y. W. Teh. Bayesian nonparametric models for ranked data. In *Advances in Neural Information Processing Systems*, pages 1520–1528, 2012.

[D3] Michael Lavine. Some aspects of Polya tree distributions for statistical modelling. *The Annals of Statistics*, pages 1222–1235, 1992.

[D4] Jim Pitman and Marc Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *The Annals of Probability*, 25(2):855–900, 1997.

[D5] Yee W Teh, Michael I Jordan, Matthew J Beal, and David M Blei. Sharing clusters among related groups: Hierarchical Dirichlet processes. In *Advances in neural information processing systems*, pages 1385–1392, 2005.

[D6] R. Thibaux and M. I. Jordan. Hierarchical beta processes and the Indian buffet process. In *International conference on artificial intelligence and statistics*, pages 564–571, 2007.

Lecture 5

[E1] Y. Gal and Z. Ghahramani. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *International conference on machine learning*, pages 1050–1059, 2016.

[E2] Soufiane Hayou, Arnaud Doucet, and Judith Rousseau. On the impact of the activation function on deep neural networks training. In *International Conference on Machine Learning*, 2019.

[E3] Mohammad Emtiyaz E Khan, Alexander Immer, Ehsan Abedi, and Maciej Korzepa. Approximate Inference Turns Deep Networks into Gaussian Processes. In *Advances in Neural Information Processing Systems*, pages 3088–3098, 2019.

[E4] J. Lee, J. Sohl-Dickstein, Jeffrey Pennington, Roman Novak, Sam Schoenholz, and Yasaman Bahri. Deep neural networks as Gaussian processes. In *International Conference on Machine Learning*, 2018.

[E5] Wesley J Maddox, Pavel Izmailov, Timur Garipov, Dmitry P Vetrov, and Andrew Gordon Wilson. A simple baseline for bayesian uncertainty in deep learning. In *Advances in Neural Information Processing Systems*, pages 13132–13143, 2019.