

Efficient Graph-Based Active Learning with Probit Likelihood via Gaussian Approximations

Kevin Miller, Hao Li, & Andrea Bertozzi

University of California Los Angeles

July 18, 2020



Graph-Based SSL Objective:

$$\mathbf{u}^* = \arg \min_{\mathbf{u} \in \mathbb{R}^N} \frac{1}{2} \langle \mathbf{u}, L_\tau \mathbf{u} \rangle + \sum_{j \in \mathcal{L}} \ell(u_j, y_j) =: \arg \min_{\mathbf{u} \in \mathbb{R}^N} J_\ell(\mathbf{u}; \mathbf{y}), \quad (1)$$

for different loss functions ℓ .

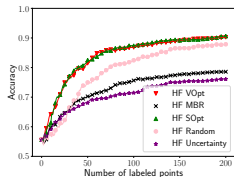
Bayesian Probabilistic Perspective:

- $\mathbb{P}(\mathbf{u}|\mathbf{y}) \propto \exp(-J_\ell(\mathbf{u}; \mathbf{y}))$
- Most choices of ℓ lead to non-Gaussian posterior, $\mathbb{P}(\mathbf{u}|\mathbf{y})$

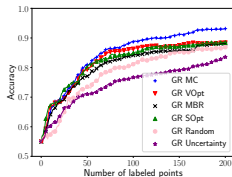
Main Idea: Use Gaussian approximations of non-Gaussian posterior distributions to allow for more general uses of Gaussian-based acquisition functions in active learning.

- “Model Change” acquisition function

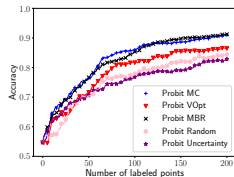
Checkerboard Results:



(a) HF

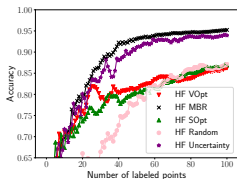


(b) GR

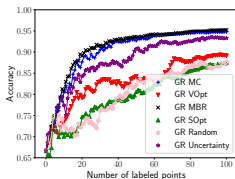


(c) Probit

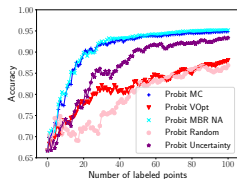
MNIST Results:



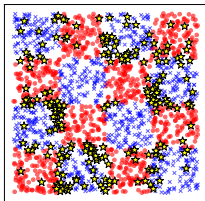
(a) HF



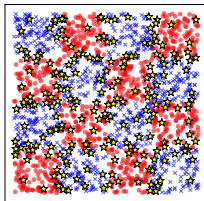
(b) GR



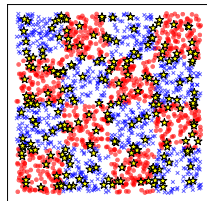
(c) Probit



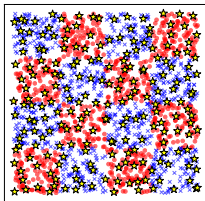
(a) HF-MBR



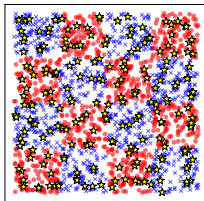
(b) GR-MC



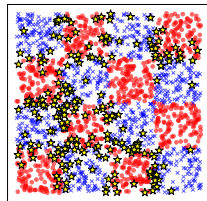
(c) Probit-MC



(d) HF-Vopt



(e) GR-Vopt



(f) Probit-Uncertainty



Ji, Ming and Jiawei Han. "A Variance Minimization Criterion to Active Learning on Graphs". en. In: *Artificial Intelligence and Statistics*. ISSN: 1938-7228 Section: Machine Learning. Mar. 2012, pp. 556–564. URL: <http://proceedings.mlr.press/v22/ji12.html> (visited on 06/11/2020).



Ma, Yifei, Roman Garnett, and Jeff Schneider. " Σ -Optimality for Active Learning on Gaussian Random Fields". In: *Advances in Neural Information Processing Systems 26*. Ed. by C. J. C. Burges et al. Curran Associates, Inc., 2013, pp. 2751–2759. URL: <http://papers.nips.cc/paper/4951-optimality-for-active-learning-on-gaussian-random-fields.pdf> (visited on 06/11/2020).



Zhu, Xiaojin, Zoubin Ghahramani, and John Lafferty. "Semi-supervised learning using Gaussian fields and harmonic functions". In: *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*. ICML'03. Washington, DC, USA: AAAI Press, Aug. 2003, pp. 912–919. ISBN: 978-1-57735-189-4. (Visited on 06/11/2020).



Zhu, Xiaojin, John Lafferty, and Zoubin Ghahramani. "Combining Active Learning and Semi-Supervised Learning Using Gaussian Fields and Harmonic Functions". In: *ICML 2003 workshop on The Continuum from Labeled to Unlabeled Data in Machine Learning and Data Mining*. 2003, pp. 58–65.