Learning Theory and Graphical Models(B659) Project Proposal:

Due on Thursday, March 7 , 2019

Dr.Roni Khardon

Arnav Arnav (aarnav)

March 7, 2019

Contents

Representation Learning Using Gaussian Processes

- **The Machine Learning Problem :** Representation learning is the problem of finding a latent representation from the data that is compact and can still encode important information about the data. If such a representation can be obtained, it can then be used for subsequent tasks such as regression or clustering. Obtaining such representations can be difficult and there has been much research in this field.

  The focus of this project is to understand and implement how Gaussian Process (GP) based models can be used to learn such representations.

- **The Probablistic Model :** This project explores the use of Bayesian Gaussian Process Latent Variable Models (GP-LVM) that assume that the data is generated from a Gaussian Process and use variational inference (with mean field approximation) to estimate the distribution over the latent input variables [1]. The model is robust to overfitting and can learn the dimensionality of the latent space.

  Damianou and Lawrence in [2] extend the bayesian GP-LVM model and allow a hierarchical structure of Gaussian processes that generate the data. The authors show that variational inference can be used to train the deep model and use the Variational Lower Bound (VLB) for model selection between models with different hierarchies. The authors show that the deep hierarchy can be used to learn meaningful latent representations even on small datasets, and the model can learn the dimensionality of the latent space for each layer in the deep hierarchy (by using RBF covariance kernel for the GPs).

  This project aims to understand and implement the two models and compare the performance on various datasets. the models can be evaluated based on the VLB as suggested on the paper, and the quality of the representation can be compared by using the nearest neighbor error(when true labels are known) and using reconstruction error given partially observed inputs.

- **What needs to be done :**

  1. Understand in depth the models described in the papers, and prove by hand the update rules derived in the papers.
  2. Implement the models in python, either from scratch or by extending existing implementations of Gaussian Processes available in libraries like PyMC3, if that is not too complicated.
  3. Replicate the results achieved in the papers. There are freely available datasets such as the MNIST dataset, Frey Faces dataset and the oil flow dataset that have been used in the papers to evaluate the models. In particular this step involves comparing the quality of learned representations with different number of layers in the deep model. The authors show in [2] that the deeper models learn better representations.
  4. Evaluate the effect of dataset size on training time of the deep GP models, and compare the effect of using sparse Gaussian processes at each layer in the deep GP model.

References

[1] Michalis Titsias and Neil D Lawrence. Bayesian gaussian process latent variable model. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pages 844–851, 2010.

[2] Andreas Damianou and Neil Lawrence. Deep gaussian processes. In *Artificial Intelligence and Statistics*, pages 207–215, 2013.