Richard Ma

First paper: Support vector machine active learning with application to text classification

Tong, S., & Koller, D. (2001). Support vector machine active learning with applications to text classification. *Journal of machine learning research*, *2*(Nov), 45-66.

This paper has a broad goal of developing a new algorithm by which a support vector machine can request new data to be labeled by a human in a large pool of unlabeled data, in a context in which manually labeling large amounts of data is resource-intensive. In the context of this paper, the use of support vector machines for classification has already been well-established. The novelty/contributions of the paper lies in the algorithm developed to request new labels.

The proposed methodology is as follows (most of these steps are also the paper's novel contributions): The paper makes use of the existing concept of a 'version space', which is the set of all possible hyperplanes that linearly (usually in a high dimensional) separate the classified data, considering that the goal of a SVM is to establish a hyperplane with the greatest separation between the data points. A small batch of the training pool contains a larger version space than that of the entire pool, with the former enclosing the latter. The paper proposes a greedy approach in which new requests for points seek to minimize the version space as rapidly as possible. The paper proposes doing so by seeking to halve the size of the version space with every iteration, and so to favor choosing points in the 'middle' of the version space so that version space is maximally decreased no matter what label the human ultimately chooses for it.

The paper then proposes novel heuristics for assessing which data most effectively halves the size of the version space.

Potential downsides: A potential downside of the work is its relatively limited scope.

Second paper: Active learning with statistical models

Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). Active learning with statistical models. *Journal of artificial intelligence research*, *4*, 129-145.

The motivation of this paper is to introduce a broadly applicable framework for implementing an active learning cycle into various machine learning methodologies for regression analysis, and to evaluate their performance and suitability in different scenarios. For example, the paper gives as an recurring example industrial settings in which physical experiments must be performed to acquire new data, and so the use of heavy computational time to produce an optimal labeling request is justified. On the other hand, in a situation in which acquiring new labeled data is cheap, the computational cost of calculating optimal labeling requests must be balanced.

The paper works towards this motivation by proposing a generally applicable statistical model for prediction variance, and by requesting new labeled data to minimize prediction variance. The machine learning technique is assumed to have some way of approximating these output distributions. The technique is applied to neural networks, mixture of Gaussians and locally weighted regression machine learning techniques. The performance of the proposed statistical active learning method using these machine learning models is assessed on a test function, and the relative strengths and weaknesses are reported.

Novelties/contributions:

- General proposal of the statistical variance minimization method for selecting the next data request point
- Evaluation of the computational efficiency of neural networks in estimating output variance
- Analytical derivations for output variance using the Gaussian models
- Evaluation of the performance of this framework on a test dataset.

Potential downside of the work: The evaluation of the active learning techniques described were done on analytical test functions (predicting the position of a robot arm with two degrees of freedom, as a function of the input angles that the arm was bent to, a geometrically solvable problem). Evaluation potentially could have been performed on non-analytically solvable problems.

Third paper: Deep Bayesian Active Learning with Image data

Gal, Y., Islam, R., & Ghahramani, Z. (2017, July). Deep bayesian active learning with image data. In *International conference on machine learning* (pp. 1183-1192). PMLR.

The paper's stated motivation/gap is that deep learning models potentially have difficulties in incorporating the active learning process: They often require large amounts of data to learn such that the incremental increasing of acquired data entailed by active learning may be inadequate, and the estimations for parameter space uncertainty/heuristics for choosing new data to acquire may be difficult to make for deep learning models.

The paper proposes using Bayesian convolutional neural networks to solve this problem. It sets up a model architecture with parameters chosen to optimize for performance in an active learning context.

Novelties/contributions:

- In general, a large number of existing algorithms are united and adapted for use for this particular purpose
- New acquisition functions acquired, and their performance evaluated in the specific context of Bayesian convolutional neural network active learning.

Potential downsides: Evaluations could potentially take place over a larger variety and larger quantity of test data images.