Small Sample Learning on Small Datasets

Quing Key

June 18, 2022

We explore the potential for encoding graphs into unitary sequences to enable learning of long and multidimensional embeddings over them. The task is relevant to both learning of traffic networks and predicting trajectories of pedestrians. We specify graph generation as a part of the problem and solve it using a graph-theoretic encoder-decoder architecture. We also propose a new measure of embedding length called the L^1 norm and show that it is able to capture a large class of embeddings. We evaluate the model on a range of datasets and show that it generalizes well across datasets.

The information bottleneck principle is a convergence-guaranteed objective for information-theoretically motivated in machine learning. Similar to information bottleneck theory, this principle analyzes the loss function of a model over a subset of data even if there are no assumptions about the full data distribution. This allows us to analyze the mutual information of models and to visualize the information-theoretic bound of these objectives. Based on information bottleneck theory, we provide two key results: (1) the mutual information term for the information-theoretic objective can be explicitly expressed in terms of a class of information-theoretic limiters (that is, by Kernel Entropies). (2) the mutual information term under varying conditions between the data and training models can be explicitly characterized. The proposed objective functions are equipped with minimum mean discrepancy (MMD) analyses, where an aleatoric and the negative log likelihood are used to determine the set of essential parameters. We demonstrate the generality of the results through experimental evaluations.

We introduce the problem of density estimation over multinomial base onsities with incomplete data. To resolve this new problem, we propose an algorithm, called D Dimple (DP), which smoothly interpolates between adding a doubly-robust OOD detector and the healthy distribution, without solving an exponential number of OOD cases. DP finds a OOD population of a random OOD sample, after regularizing with a population dependent barrier. This instance-independent algorithm alleviates OOD situations when DP is not solved in a straightforward fashion. Experiments on various benchmark datasets demonstrate that DP outperforms DP and achieves the state-of-the-art results. Our source code is available at https://github.com/jw97/DP.

The recent progress of artificial intelligence (AI), especially the technological advances of deep neural networks (DNNs), has led to vast availability of high-quality data (e.g., images and language data). However, this high quality has some important implications on several applications such as disease diagnosis and healthcare research. To demonstrate the effectiveness of DNNs for the development of health research tools needed in developing countries, we conducted a comprehensive experimental study to evaluate the performance of various AI algorithms to diagnose onset of AD in diabetic patients. We trained four different DNN models (four following–Cheng, HDP-II, ResNet-18, ResNet-34) in tensor methods (PGVAE, ResNet-50, ResNet-56) based on different datasets. The four models were trained on different datasets (MNIST, CIFAR-10, and CIFAR-100) and their performance was compared based on various relevant performance metrics. We observed that the most robust of the four models, ResNet-34 with greater than 9535.0%, and an average precision of 2.75×10^3, for the prediction of AD1 for digits 4 and 7. We further observed that HDP-II model outperforms several algorithms including Random Forest, and that it outperforms both word and parameter prediction. We further observed that in some datasets, the WordNet model outperforms a Random Forest model. The obtained result suggests that DNNs, based on simple and efficient architectures, have the potential to improve the performance of many other research applications.