

HSE-Yandex autumn school on generative models

Tuesday, 26 November 2019 - Friday, 29 November 2019

Yandex, Moscow

Scientific Programme

Generative models

Denis Derkach (HSE University)

Simple generative models. Generative adversarial networks.

Advanced generative models. Introduction to normalizing flows.

Applications of deep generative models can be found in a variety of domains nowadays. These lectures consider modern architectures of generative models and corresponding training algorithms. Practice part covers industrial applications of generative adversarial networks (GAN) and normalising flows. After lectures you will understand pros and cons of those methods and will be able to choose and apply the best-fitting type of model.

Optimization methods for optimal transport

Pavel Dvurechensky (WIAS Berlin)

In this lecture I will give a short introduction to optimal transport problem with some motivating examples from modern machine learning, including image retrieval and classification. The effectiveness of optimal transport in applications comes with the price of heavy computations and I will discuss two types of methods which allow to efficiently compute optimal transport distance (Wasserstein distance). The first method is Sinkhorn's algorithm and the second one is accelerated gradient method. If time allows I will also discuss a next level problem of finding Wasserstein barycenter of a set of measures, which works quite well in image analysis. Numerical methods for this problem will be discussed.

Introduction to scalable Bayesian methods

Dmitry Vertrov and Ekaterina Lobacheva (HSE University) and Nadia Chirkova (HSE University)

In this mini-course, we will discuss the advantages of using Bayesian methods in machine learning and particularly in deep learning. The attendees will learn how to use probabilistic modeling to construct neural generative and discriminative models, how to train these models using approximate Bayesian inference and what tricks are needed to make this technique scalable. Theoretical and practical assignments will follow the lectures.

Towards Photorealistic Neural Avatars

Victor Lempitsky (Samsung AI Center, Skolkovo Institute of Science and Technology)

I will present the results on modeling the appearance of humans using generative convolutional

neural networks that have been obtained recently in Samsung AI Center. This includes full body avatars as well as "few-shot" head avatars, i.e. neural head avatars that can be learned from few photographs. Time permitting I will also present results for 3D human body pose tracking using deep convolutional networks.

Wasserstein-2 Generative Networks

Evgeny Burnaev (Skoltech)
(joint work with Alexander Korotin, Vage Egiazarian, Arip Asadulaev)

Abstract: Modern generative learning is mainly associated with Generative Adversarial Networks (GANs). Training such networks is always hard due to the minimax nature of the optimization objective. In this paper we propose a novel algorithm for training generative models, which gets rid of mini-max GAN objective, thus significantly simplified model training. The proposed algorithm uses the variational approximation of Wasserstein-2 distances by Input Convex Neural Networks. We also provide the results of computational experiments, which confirms the efficiency of our algorithm in application to latent spaces optimal transport and image-to-image style transfer.

RPGAN: random paths as a latent space for GAN interpretability

Andrey Voynov

Sequence modeling with unconstrained generation order

Dmitry Emelianenko

Bayesian inference vs stochastic optimization

Vladimir Spokoiny (WIAS, HSE University)

The talk discusses a new approach to Bayesian inference which allows to study a finite sample performance of Bayesian credible sets. The main results claim near normality of the posterior distribution with the mean at the penalized maximum likelihood estimation. We also discuss the relation between Bayesian inference and stochastic optimization and applications to nonlinear inverse problems.

Uncertainty estimation: can your neural network provide confidence for its predictions?

Maxim Panov (Skoltech, HSE University)

Neural networks paved their way as a state-of-the-art approach in almost any machine learning application. However, neural networks often make very confident predictions for the out-of-sample data or the data in-between classes. In many applications, this is unacceptable, and the ability to provide confidence for prediction is thus crucial. However, an uncertainty estimation for neural networks remains a non-trivial problem, and the existing approaches still demonstrate moderate performance in terms of either accuracy (dropout-based approaches) or computational resources (ensembles and variational inference).

In this talk, we discuss the existing approaches for uncertainty estimation and focus on the improvement of dropout-based methods by introducing an additional diversity among the dropout masks. We consider several approaches for diversification, including the sampling based on determinantal point processes. Numerical experiments demonstrate that a single model with an enhanced diversity is comparable or even better than an ensemble of neural networks. The resulting approach may be applied to the existing models and architectures to enable accurate uncertainty estimation.

Approximation of multivariate functions using deep learning with applications

Ivan Oseledets (Skoltech)

TBA

Structure-adaptive manifold estimation

Nikita Puchkin (HSE University)

We consider a problem of manifold estimation from noisy observations. Many manifold learning procedures locally approximate a manifold by a weighted average over a small neighborhood. However, in the presence of large noise, the assigned weights become so corrupted that the averaged estimate shows very poor performance. We suggest a novel computationally efficient structure-adaptive procedure, which simultaneously reconstructs a smooth manifold and estimates projections of the point cloud onto this manifold. The proposed approach iteratively refines the weights on each step, using the structural information obtained at previous steps. After several iterations, we obtain nearly "oracle" weights, so that the final estimates are nearly efficient even in the presence of relatively large noise. In our theoretical study we establish tight lower and upper bounds proving asymptotic optimality of the method for manifold estimation under the Hausdorff loss. Our finite sample study confirms a very reasonable performance of the procedure in comparison with the other methods of manifold estimation

Fast Simulation Using Generative Adversarial Networks in LHCb

Artem Maevskiy (HSE University)

Differentiating the Black-Box: Optimization with Local Generative Surrogates

Vladislav Belavin (HSE University)

We propose a novel method for gradient-based optimization of black-box simulators using differentiable local surrogate models. Many processes are modelled with non-differentiable simulators with intractable likelihoods. In these cases, the optimization of this forward process is particularly challenging, especially when the process is stochastic. Problems of this nature are common in applied fields such as experimental physics, for instance in the design of magnets, detection apparatus, accelerators, etc. To address the problem of optimizing processes represented by stochastic, non-differentiable simulators, our approach uses deep generative models to approximate the simulator in successive local neighbourhoods of the parameter space. These local surrogates are capable of approximating the gradients of the simulator, thus enabling gradient-based optimization of the simulator's parameters. This technique is further generalized to perform gradient-based optimization in the global parameter space. We show that our method can attain the same minima faster than existing approaches, including recently proposed REINFORCE-based strategies.

Adaptive Divergence for Rapid Adversarial Optimization

Maxim Borisyak (HSE University)

Adversarial Optimization (AO) provides a reliable, practical way to match two implicitly defined distributions, one of which is usually represented by a sample of real data, and the other is defined by a generator. Typically, AO involves training of a high-capacity model on each step of the optimization. In this work, we consider computationally heavy generators, for which training of high-capacity models is associated with substantial computational costs. To address this problem, we introduce a novel family of divergences, which varies the capacity of the underlying model, and allows for a significant acceleration with respect to the number of samples drawn from the generator.

We demonstrate the performance of the proposed divergences on several tasks involving tuning parameters of Pythia event generator.

$(1+\epsilon)$ -class Classification: an Anomaly Detection Method for Highly Imbalanced or Incomplete Data Sets

Andrey Ustyuzhanin, (HSE University)

Anomaly detection is not an easy problem since distribution of anomalous samples is unknown a priori. We explore a novel method that gives a trade-off possibility between one-class and two-class approaches, and leads to a better performance on anomaly detection problems with small or non-representative anomalous samples. The method is evaluated using several data sets and compared to a set of conventional one-class and two-class approaches.

