PhD Project in Active Deep Learning for Nano Sensor Systems

At DTU Compute, Sections for Cognitive Systems and IDUN Sensor.

Project Description

This 3-year PhD project starting August 1st 2017 at DTU Compute's Sections for Cognitive Systems is a part of IDUN center of excellence, which is a multi-disciplinary project lead by DTU Nanotech with participation of DTU Compute and The University of Copenhagen, and funded by the Danish National Research Foundation and the Villum Foundation. The center is divided into IDUN Drug focusing on drug delivery and IDUN Sensor focusing on nano-mechanical sensors.

The PhD project is associated with IDUN Sensor research that focuses on development and exploration of nano-mechanical biosensors. Data processing and modelling are indispensable multipurpose tools for sensor development and evaluation and analysis of results in demonstration activities. In sensor development, data modelling tools provide other views and insight into the physical and chemical properties of the sensor as well as sensing principle; hence, improving sensor development in terms of time-use, but also the ability to robustly confirm hypotheses about the sensor's functionality. In relation to sensor demonstration activities, data modelling is important for obtaining robust sensor performance by suppression of noise caused by undesired physical and chemical properties of the sensor as well as uncontrollable experimental factors.

We are living in the information age, where data is the raw material. But access to data can be sparse and expensive, especially when human annotation and expert supervision is needed in the data collection process. Therefore optimising data efficiency in machine learning and developing optimal experimental designs is the two methodological objectives in this PhD project.

The main hypothesis driving the project is that generalisation performance and data effiency of machine learning can be enhanced increasingly with more intelligent data collection and interaction with human annotators through the combination of better active learning techniques in deep neural networks.
This will have a great impact on the information society as a whole and affect how researchers interact with and set up their experiments and how humans interact with their intelligent devices and tools. The teacher-student interaction will be more friction- and effortless, so as to implement it in everyday life and lower the expenses on data collection.

Recent advances within the deep learning field has shown remarkable performance in a great variety of data processing tasks. The PhD project will focus on developing new active learning methods for deep neural network models. Such methods can provide optimal experimental design and hypothesis testing for sensor development, and further reduce the need for user labels in connection with demonstration of detection and predictive sensing capabilities. The methodological objectives and research relates to, and will leverage from, current advances in Bayesian optimization; one-shot-learning; unsupervised and semi-supervised generative models like generative adversarial networks and variational auto-encoders; users-in-the-loop models, where the user is the sensor developer and/or an end-user providing labeled information.

The project will also be about showing that active deep learning models are generic and will perform well across several kinds of data sets as well as humans reuse pattern recognition and memory across several tasks. This can be done using knowledge transfer and one-shot-learning, reusing previously trained parameters and prior knowledge about the sensor's environment, for converting the model from one problem to another. Learning more global representations of the world is the goal, so patterns learned from some objects can be recognised in others. By building very generic multipurpose models, we can start to solve problems across different research fields.

Generative models as the variational auto-encoder (VAE) can work in the semi-supervised regime, by encoding very crucial information and global structure in both labeled and unlabeled data in latent representations using probabilistic inference. From the latent representation it is then possible to cluster, identify, classify and generate the excessive amount of unlabeled data given the more rare and expensive labeled data. The VAE model is superior at learning the global structure in the features and smoothing out local variances and noise, which will be highly useful for generating feature representations with low signal-to-noise ratios.

Generative adversarial networks (GAN) are optimised in a mini-max game played by a generator network generating false data close to the real data to fool the discriminator network in it's objective of distinguishing between these. By inferring the local structure in the features the generator can generate new very natural looking data points. By using the more flexible VAE model as generator the forces can be combined and features can be generated holding both the global and local structure.

In active learning, the learning network may place a limited number of queries for the expert, e.g. sensor or human annotator, to get labels. The objective compared to passive learning is now to learn a query strategy optimising both the efficiency, lowering the time and resources spend on fulfilling queries, and efficacy, obtaining the best generalisation performance. Efficacy is the key objective of machine learning models. Adding the cost of data collection in the models puts a new efficiency constraint on the objective. By jointly optimising these can lead to even better experimental designs and users-in-the-loop systems run and monitored by the models. When solving harder problems has a greater learning potential the model's ability to query and select new data can be turned into an advantage rather than a hindrance.

The first part of the project will mainly have a methodological focus on developing novel techniques in active learning for deep neural networks, through reproducing and comparing state-of-the-art results of active deep learning on benchmarking and synthetic datasets and leveraging these into new hypotheses on improvements of the methods. Using completely labeled datasets and making a fraction of the labels available to the model, we can develop pool-based active learning methods to query for the unlabeled data.

The second part of the project will be setting up optimal experimental designs and users-in-the-loop systems, related to problems addressed in the IDUN project, where the methods can be proven.

The third part of the project will be applying the optimal active learning and deep neural network models in a nano-mechanical biosensor and biological case.