CIS DL TECHNICAL TALK: Anthropomorphic Machine Learning:How to get Fast,Interpretable Deep Learning"

Share

Background:

Machine Learning (ML) and AI justifiably attract the attention and interest not only of the wider scientific community and industry, but also society and policy makers. Recent developments in this area range from accurately recognizing images and speech to beating the best players in games like Chess, Go and Jeopardy. In such well-structured problems, the ML and AI algorithms were able to surpass the human performance, acting autonomously. These breakthroughs in performance were made possible due to the dramatic increase of computational power and the amount and ubiquity of the data available. This data-rich environment, however, led to the temptation to shortcut from raw data to the solutions without getting a deep insight and understanding of the underlying dependencies and causalities between the factors and the internal model structure. Even the most powerful (in terms of accuracy) algorithms such as deep learning (DL) can give a wrong output, which may be fatal. Recently, a crash by a driverless Uber car was reported raising issues such as responsibility and the lack of transparency, which could help analyze the cause and prevent future crashes. Due to the opaque and cumbersome model structure used by DL, some authors started to talk about a dystopian “black box” society. Having the true potential to revolutionize industries and the way we live, the recent breakthroughs in ML and AI also raised many new questions and issues. These are related primarily to their transparency, explainability, fairness, bias and their heavy dependence on large quantities of labeled training data.

Despite the success in this area, the way computers learn is still principally different from the way people acquire new knowledge, recognize objects and make decisions. Children during their sensory-motor development stage (first two years of a child’s life) imitate observed activities and are able to learn from one or few examples in “one-shot learning”. People do not need a huge amount of annotated data. They learn by example, using similarities to previously acquired prototypes, not by using parametric analytical models. They can explain and pass aggregated knowledge to other humans. They predict based on rules they formulate using prototypes and examples.

Current ML approaches are focused primarily on accuracy and overlook explainability, the semantic meaning of the internal model representation and its link with the problem domain. They also overlook the efforts to collect and label training data and rely on assumptions about the data distribution that are often not satisfied. For example, the widely used assumption that the validation data has the same distribution as that of the training data is usually not satisfied in reality and is the main reason for poor performance. The typical assumption for classification, that all validation data come from the same classes as the training data, may also be incorrect. It does not consider scenarios in which new classes appear. For example, if a driverless car is confronted with a scene that was never used in the training data or if a new type of malware or attack appears in a cybersecurity domain. In such scenarios, the best existing approach of transfer learning will require a heavy and long process of training with huge amounts of labeled data. While driving in real time, the car will be helpless. In the cybersecurity area it is not possible to pre-train for all possible attacks and viruses. Therefore, the ability to detect the unseen and unexpected and start learning this new class/es in real time with no or very little supervision is critically important and is something that no currently existing classifier can offer. Another big problem with the currently existing ML algorithms is that they ignore the semantic meaning, explainability and reasoning aspects of the solutions they propose. The challenge is to fill this gap between high level of accuracy and the semantically meaningful solutions.

The most efficient algorithms that have fueled interest towards ML and AI recently are also computationally very hungry – they require specific hardware accelerators such as GPU, huge amounts of labeled data and time. They produce parameterized models with hundreds of millions of coefficients, which are also impossible to interpret or be manipulated by a human. Once trained, such models are inflexible to new knowledge. They cannot dynamically evolve their internal structure to start recognizing new classes. They are good only for what they were originally trained for. They also lack robustness, formal guarantees about their behavior and explanatory and normative transparency. This makes problematic use of such algorithms in high stake complex problems such as aviation, health, bailing from jail, etc. where the clear rationale for a particular decision is very important and the errors are very costly.

All these challenges and identified gaps require a dramatic paradigm shift and a radical new approach. In this talk I will present such a new approach towards the next generation of computationally lean ML and AI algorithms that can learn in real-time using normal CPUs on computers, laptops, smartphones or even be implemented on chip that will change dramatically the way these new technologies are being applied. It will open a huge market of truly intelligent devices that can learn lifelong, improve their performance and adapt them to the user’s demands. I call this approach anthropomorphic, because it shares similar characteristics to the way people learn, aggregate, articulate and exchange knowledge. It focuses on addressing the open research challenge of developing highly efficient, accurate ML algorithms and AI models that are transparent, interpretable, explainable and fair by design. Such systems are able to self-learn lifelong, and continuously improve without the need for complete re-training, can start learning from few training data samples, explore the data space, detect and learn from unseen data patterns, collaborate with humans or other such algorithms seamlessly.

Speaker Bio -

Prof. Angelov (MEng 1989, PhD 1993, DSc 2015) is a Fellow of the IEEE, of the IET and of the HEA. His PhD supervisor, Dr. Dimitar P. Filev is now Member of the National Academy of Engineering, USA. Professor Angelov is Vice President of the International Neural Networks Society (INNS) for Conference and Governor of the Systems, Man and Cybernetics Society of the IEEE.

Date and Time

Location

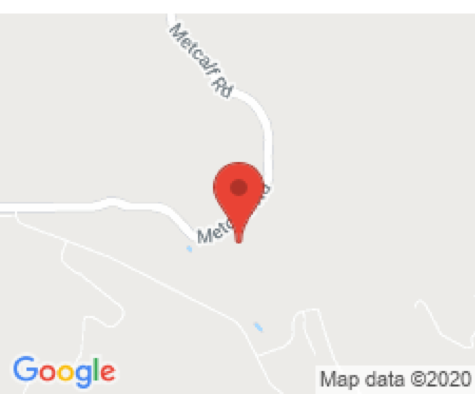
Hosts

Registration

Date: **11 Dec 2019**
Time: **06:30 PM to 08:30 PM**
All times are US/Pacific
Add Event to Calendar
- Outlook (vCal)
- iCal
- Google Calendar

NA

Santa Clara, California
United States



Santa Clara Valley Section Chapter,CIS11

Agenda

About this Event

Title: Towards Anthropomorphic Machine Learning: How to get Fast, Interpretable Deep Learning

Speaker: Prof **Plamen Angelov**, PhD, DSc, FIEEE from School of Computing and Communications, Lancaster University, UK and Vice President-International Neural Network Society.

Date: Wednesday, December 11, 2019. 5:30pm to 8:30pm Pacific Time.

Venue: Intel Santa Clara, SC9 auditorium (2200 Mission College Blvd, Santa Clara, CA 95054)

PROGRAM

- 6:30 - 7:00 PM Networking & Refreshments
- 7:00 - 8:00 PM Talk
- 8:00 - 8:30 PM Q&A

Organized by: IEEE Computational Intelligence Society

Event **co-sponsored** by : IEEE Computer Society (CS) chapter of Silicon Valley, IEEE Solid-State Circuits Society chapter of Silicon Valley and ValleyML.ai

Admission Fee:

Open to all to attend (**Online registration is needed**. If you did not register, seating is not guaranteed.)

- IEEE members - free
- non-IEEE members - free. You do not need to be an IEEE member to attend!