ADAPT in SC: DNN infused Causal Bayesian Networks - A Beneficial Combination to Balance Explainability and Tractability

***Ivan Dungan**

Assistant Professor, Francis Marion University

Keywords - Explainability, Deep Learning, Causality

Abstract - We will introduce projects regarding our partner's multi-dimensional data in which inference and prediction are required; further, explainability and more generally, trustworthiness are crucial to successfully counseling stakeholders. Causal Bayesian Networks (CBN), a type of Structural Causal Network, are graphical models that are of particular interest in the AI market for encoding and discovering causal information which meets our needs. However, the computational requirements, particularly, the memory usage of CBNs is a shortcoming of these models due to the possibility of a large number of states. Target variables for prediction are especially prone to this difficulty. Even with the availability of memory, a large number of states can subdue the information in the data. We will introduce CBNs and review their strengths and weaknesses in our projects, and introduce an approach, which infuses Deep Neural Networks (DNN) into local components of CBNs to compensate for these weaknesses without losing their strengths. We will compare the original model with the infused one via memory usage and k-fold validation on simulated data, and discuss the implications to explainability of the model.

A Sustainable Way for Minimizing Environmental Impact of Large/Deep Neural Network Implementations

***G Kumar Venayagamoorthy, PhD, MBA, FIEEE**
Duke Energy Distinguished Professor of Power
Engineering Real-Time Power and Intelligent Systems
Laboratory Department of Electrical and Computer
Engineering Clemson University, Clemson, SC 29634

Keywords— Climate change, complex systems, computational energy, modern power systems, sustainable AI

Abstract

Complex networks continue to generate quintillion bytes of data periodically, leading to the pressing needs for new efforts in dealing with the grand challenges brought by Big Data. There is consensus in computational intelligence (CI) communities that data volume presents a pertaining challenge to the issue of scalability. This is only one side of the volume study. The other side of the volume is the dimensionality. Complex networks have several outputs, and this makes it challenging for CI methods to accurately model such systems, The aim of multi-output learning is to simultaneously predict multiple outputs given an input. It is an important learning problem for decision-making, since making decisions in the real world often involves multiple complex factors and criteria. Many of the solutions have been shown to involve complex AI algorithms such as deep neural networks and deep learning systems.

Optimal operation and management of modern power systems are key to clean energy sustainability. Power systems are complex systems consisting of components and subsystem exhibiting nonlinear phenomena and changing dynamics with different operating conditions. Components such as power generators (including wind and solar power plants) and network dynamics including power flows can be represented by differential algebraic equations (DAEs). Large-scale neural networks can be effectively used to learn the solutions of such DAEs, and thus, they can be suitable tools for modeling and monitoring the behavior of power system components and network phenomena.

Mainstream artificial intelligence (AI) implementation overheads are computationally expensive, requiring long training time, huge energy requirements, and are not scalable and sustainable. Development of large-scale neural networks and learning systems such as deep neural networks consume mega-watts of power and burn a lot of energy. This is environmentally detrimental, with high carbon emissions, and will contribute to worsening climate conditions. The efficiency and scalability of today’s AI and machine learning approaches for mission-critical applications is questionable. An integration of dynamic systems theory, computational neuroscience, and neural networks are necessary to learn the spatial and temporal dynamics of complex nonlinear systems in real life. ‘Sustainable AI’ is defined as a movement to foster change in the entire lifecycle of AI products (i.e. idea generation, training, re-tuning, implementation, governance) towards greater ecological integrity and social justice. Sustainable AI has two branches; AI for sustainability and sustainability of AI (e.g. reduction of carbon emissions and computing power).

In the context of the two branches of sustainable AI, cellular computational network (CCN) is proposed as a much efficient approach to implement scalable AI for complex systems including modern power systems. Cellular computational networks, a class of sparsely connected dynamic recurrent networks, are known to be scalable and have the potential to reduce the AI implementation computational energy and time overheads without sacrificing AI’s performance and accuracy. CCN (Figs. 1 and 2) is a directed graph, distributed, and scalable framework for learning of learning systems (LOLS) and predicting/estimating the dynamics of complex network systems, utilizing asynchronous and synchronous mechanisms. The CCN

This research is funded by US NSF grants: CNS 2131070, CNS 2318612 and ECCS 2234032 and the Duke Energy Distinguished Professor Endowment Fund. Any opinions, findings and conclusions or recommendations expressed in this material are those of author and do not necessarily reflect the views of NSF and Duke Energy.

suggests a composable modularity that can divide a large system into small subsystems with corresponding computational cells. Each cell can be implemented with a simple shallow network, and such a framework design enables multi-resolution for both global and local information. The computational energy and time overhead for learning and operation is highly reduced in the utilization of a CCN compared to a single large neural network. It is interesting to note that deep neural networks (DNNs) can be directly realized via CCNs, thus simplifying the development overheads needed by DNNs.

This paper makes the case for CCNs for sustainable AI implementation beyond its known capability of CCN as a scalable AI approach. The computational energy and time overhead of a CCN during its learning and operational phases are lower than traditional neural networks and much lower than deep neural networks. Some challenges for CCNs to be successful is deriving the topology for computing. However, our recent work is focusing on data-driven and graph theory approaches to deriving the topology of complex system/network for CCNs to be deployed easily. The value of CCNs will be illustrated with case studies addressing challenges in modern power systems operations and management, especially providing situational intelligence to system operators in power system control centers.

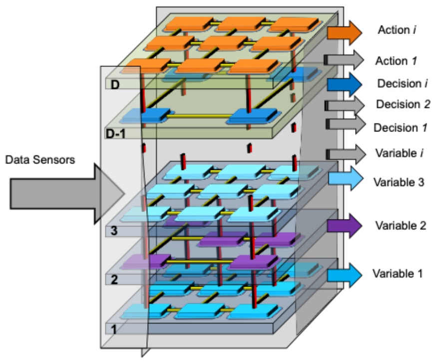

Fig. 1 *N*-diminesional, *D*-layer CCN.

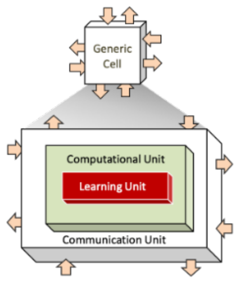

Fig. 2 A generic cell of a CCN consisting of three units: communication, computational and learning.

ADAPT in SC: Personalized Eating Detection by Monitoring Wrist Motion Using Individually Trained Convolution Neural Networks

***Yu Xuan and Adam Hoover**
Doctoral Student, Clemson University and Professor, Clemson University

Presentation Keywords:
personalized health; automated dietary monitoring; eating detection; mobile health; obesity

Abstract:
Our research considers the problem of automatically detecting when a person is eating, and for how long, based on smartwatch sensor data. Eating behaviors play an important role in the etiology, prevention, and treatment of many highly prevalent and debilitating chronic diseases, including overweight/obesity and type 2 diabetes. To monitor the eating behaviors automatically, we utilize wrist-mounted motion sensors contained in a smartwatch to detect when a person is eating. This reduces the burden and inaccuracy of participants to self-report their eating states in a traditional questionnaire. Previous works have demonstrated the viability of this method by training wrist motion classifiers on datasets collected by groups of people. However, we hypothesize that the motion features or patterns of each individual may vary due to distinct habits, preferred food types, eating utensils, intake speed, and other individual-specific characteristics. To test this hypothesis, we collected new wrist motion data from 9 individuals which we call the Clemson Multi-day Dataset (CMD). Then, we trained a personalized eating classifier for each individual separately.

For baseline, we used the Clemson All-day dataset (CAD). CAD contains 354 day records from 351 participants. CMD collected 119 day records (ranging from 9 days to 21 days for each participant). We are currently building a new smartwatch app to allow us to collect more individual data (N=30).

Our group previously proposed a top-down classifier for detecting eating episodes from wrist motion data using a convolution neural network (CNN). The top-down classifier slides a large window (6 minutes) along the 6-axis IMU data of wrist motions (acceleration and angular velocity) and trains the CNN on each window. While other methods learn eating detection by coupling hands-to-mouth gestures or single bites, this top-down classifier only requires an individual to label each meal's start and end time. As a result, an individual could easily input their eating data in a free-living environment to personalize an eating detector by simple operations such as clicking the button on a smartwatch twice per meal.

To investigate the personalized classifier for eating detection, we trained the CNN architecture once on CAD (group model), and once for each individual in CMD (individual models). We used

early stopping to train the group model and 5-fold cross-validation (splitting by days) to train the individual model and to evaluate all models, averaging performance across folds. We selected 1:20 weighted accuracy (wAcc) as a time metric to measure performance as eating typically occurs in only 5% of the time in a day. We also selected episode true positive rate (TPR) and episode false positive over true positive ratio (FP/TP) to measure the model's ability to detect each meal episode.

On average, the individual models outperform the group model with significant improvement in wAcc and TPR, though the FP/TP ratios of individual models are higher. For cross-validation, one participant with 9-day records was excluded to make sure that each fold contained at least 2-day records. For the other 8 participants, the average wAcc, TPR, and FP/TP of the group model is 0.724, 0.770, and 2.08, respectively, and the average wAcc, TPR, and FP/TP of the individual model is 0.792, 0.881, and 2.89. Moreover, the improvement of individual models over the group model can vary per individual. For instance, the improvement of wAcc ranges from -0.63% to 16.7% where 7 are positive.

This work demonstrates the individually trained wrist motion models provide greater accuracy in recognizing when a person is eating compared to a group model. However, the range of improvement varied, with 2 out of 9 participants actually showing a decrease in accuracy.
It may be that more data is needed for these individuals. In future work we plan to collect more data to allow for a deeper exploration. We also plan to explore hybrid methods that include both group and individual model training.