Big Data & Extreme Scale Computing, 2nd Series, (BDEC2)
Workshop 1: Bloomington, Indiana, November 28-30, 2018

BDEC-2 WHITE PAPER:
Category/Focus Area:

Novel models of integrated inquiry: Unprecedented new methods in high-end data analysis (HDA) that are being pioneered in Big Data communities, such as Deep Learning, will increasingly be combined and integrated with the simulation-centric approaches of traditional high performance computing (HPC). In other words, the impact of the digital revolution on scientific methodologies has entered a dramatic new phase.

Author: William Tang (Princeton University/PPPL)

Background & Challenge: Accelerating delivery of accurate predictions in key scientific domains of current interest can best be accomplished by engaging modern big-data-driven statistical methods featuring machine learning/deep learning/artificial intelligence (ML/DL/AI). Associated techniques have been formulated and adapted to enable new avenues of data-driven discovery in prominent scientific applications areas such as the quest to deliver Fusion Energy -- identified by the 2015 CNN "Moonshots for the 21st Century" series as one of 5 exciting grand challenges. An especially time-urgent and very challenging problem facing the development of a fusion energy reactor is the need to reliably predict and avoid large-scale major disruptions in magnetically-confined tokamak systems such as the EUROfusion Joint European Torus (JET) today and the burning plasma ITER device in the near future. Significantly improved methods of prediction with better than 95% predictive capability are required to provide sufficient advanced warning for disruption avoidance or mitigation strategies to be effectively applied before critical damage can be done to ITER -- a ground-breaking $25B international burning plasma experiment with the potential capability to exceed "breakeven" fusion power by a factor of 10 or more. This truly formidable task demands accuracy beyond the near-term reach of hypothesis-driven /"first-principles" extreme-scale computing (HPC) simulations that dominate current research and development in the field.

Approach & Advances: Recent HPC- relevant advances in the deployment of deep learning convolutional and recurrent neural nets have been demonstrated in exciting scaling studies of Princeton's Deep Learning Code -- "FRNN (Fusion Recurrent Neural Net) Code -- on modern GPU systems. This is clearly a "big-data" project in that it has direct access to the huge EUROFUSION/JET disruption data base of over a half-petabyte to drive these supervised machine-learning studies. FRNN implements a distributed data parallel synchronous stochastic gradient approach with "Tensorflow" libraries at the backend and MPI for communication. This deep learning software has demonstrated excellent scaling up to 6000 GPU's on Titan that has enabled clear progress toward the goal of establishing the practical feasibility of using leadership class supercomputers to greatly enhance training of neural nets that can enable transformational impact on key discovery science application domains such as Fusion Energy Science. In addition to (1) Titan at the OLCF, powerful systems currently deployed by the Princeton U/PPPL deep learning software include: (2) Japan's "Tsubame 3" system with 3000 P-100 GPU's; and (3) OLCF'S "Summit" system during it's Early Access Phase. Achieving accelerated progress in statistical Deep Learning/AI software trained on very large data sets hold exciting promise for delivering much-needed predictive tools capable of accelerating scientific knowledge discovery in HPC. The associated creative methods being developed also has significant potential for cross-cutting benefit to a number of important application areas in science and industry. This work was recently awarded the 2018

NVIDIA Global Achievement Award – (https://insidehpc.com/2018/03/princeton-team-using-deep-learning-develop-fusion-energy/; and https://www.princeton.edu/news/2018/04/02/william-tang-wins-2018-global-impact-award-advance- development-ai-software-help.) Moreover, the project on "Accelerated Deep Learning Discovery in Fusion Energy Science" has been selected as one of the DOE-ALCF-21 Early Science Projects that will feature advanced INTEL architectures:
https://www.alcf.anl.gov/articles/alcf-selects-data-and-learning-projects-aurora-early-science-program; https://www.hpcwire.com/off-the-wire/deep-learning-to-predict-fusion-disruptions-picked-for-first-us-exascale-system/

Relevance AI/Deep Learning Co-Design: The applied math and computer science algorithms being developed in our exemplar Tokamak fusion DL/AI project have significant potential for cross-cutting benefit to a number of important application areas in science and industry. For example, this work is featured as the Plasma Fusion "exemplar" in the Big Data & Extreme Computing (BDEC) Report (J. Dongarra, et al. 2018, https://www.exascale.org/bdec/). In addition, as evident from examining some of the hyper-parameter tuning workflows being developed in the fusion energy application studies, there are some similarities with those displayed in deep-learning projects such as "Candle" – the DOE/NIH exascale deep learning and simulation studies of precision medicine for Cancer. Accordingly, algorithmic formulations with results (including ROC curves) from current work provide opportunities for productive cross-disciplinary comparisons of methodologies with associated "lessons learned" for co-design. This would also provide a desirable step forward to the delivery of modern DL/AI software capabilities that can be largely "machine and hardware agnostic" with inherent adaptability to expected improvements in architectural designs. As highlighted in IBM's Power-9 rollout, it was noted that "whether its GPU accelerators or FPGA's, our aim is to provide the links and hooks to give all an equal footing in the new server."

Accelerated progress in the further development and deployment of advanced DL/AI R&D for the prediction and mitigation of disruptions in burning plasma fusion Tokamak systems will be significantly enhanced by cross-cutting engagement with leading DOE laboratory scientists. As noted earlier, exciting progress can be realistically anticipated from the collaboration between the clean fusion energy "FRNN" deep learning /AI project at Princeton U/PPPL and the cancer precision medicine "Candle" project at ANL. For example, mutually addressing generically similar but highly challenging hyper-parameter tuning workflow complexities holds great promise for benefiting both application domains. In addition, the new DL/AI capability to include for the first time multi-dimensional signals into the pre-disruption classifiers opens the door for delivery of such classifiers via path-to-exascale HPC simulations – a major step forward in illustrating possible practical "convergence" between big-data-driven AI/DL with exascale HPC simulations.

It is important to highlight at this point another hot topical area of strong interest -- the *demonstration of the actual connection of DL/AI prediction to control in real-world situations.* The fusion energy DL/AI exemplar provides a natural pathway to do so – i.e. moving from accurate prediction to actual control in an active tokamak laboratory environment. This is a key goal which will require collaborative development -- with diagnostics experts -- of actuators capable of showing how reinforcement learning, inference, etc. can positively impact real-time control in a realistic laboratory environment. For example, the leadership of the DIII-D tokamak experiment at General Atomics in San Diego, CA. has already expressed enthusiasm to deploy the Princeton FRNN "predictors" on their plasma control system (PCS). While we continue work on further improving our DL/AI disruption predictor, it's first deployment on an actual PCS promises to be quite exciting. Realistic control theory capabilities connected to DL/AI predictors will in general be a huge area of growth.in many practical application domains.