CANDLE: As part of the Precision Medicine and the Cancer Moonshot initiatives, the Department of Energy (DOE) has entered into a partnership with the National Cancer Institute (NCI) of the National Institutes of Health (NIH). This partnership has identified three key challenges that the combined resources of DOE and NCI can accelerate. Four DOE national laboratories are collaborating with the NCI and the NCI-supported Frederick National Laboratory for Cancer Research to advance these challenges. Within each of these three key challenges, we have focused on the machine learning aspect of the three problems and, in particular, we are focused on building single scalable deep neural network code we call CANDLE (CANcer Distributed Learning Environment) that can be used to address all three challenges. ~~In the RAS pathway problem, we guide multi-scale molecular dynamics (MD) runs through a large-scale state-space search, using unsupervised learning to determine the scope and scale of the next series of simulations based on the history of previous simulations. In the drug response problem, we use supervised machine learning methods to capture the complex, non-linear relationships between the properties of drugs and the properties of the tumors to predict response to treatment and therefore develop a model that can provide treatment recommendations for a given tumor. In the treatment strategy problem, we use semi-supervised machine learning to automatically read and encode millions of clinical reports into a form that can be computed upon. These encoded reports will be used by the national cancer surveillance program to understand the broad impact of cancer treatment practices and drive simulations of entire cancer populations to determine optimal treatment strategies for patient cohorts.~~

As part of CANDLE’s development, we have created a set of benchmark programs, all of which highlight the computational needs of diverse deep learning (including unsupervised, semi-supervised and supervised) algorithms. [Note: “we” means Oak Ridge National Laboratory, ORNL]

Upulee, the strike-through text above can be sent to NSF in the proposal, but for now, it should not otherwise be shared.

From the CANDLE ‘owner’:

* Candle is written using Python and c/c++.
* It will be in the public domain. (when?? TBD) So, how can students access the code to do preliminary testing? [LP – ask]
* Hardware requirements are mainly for gpu based clusters.
* We will welcome testing and evaluation of the code.