**Summary**

1. Overview

The search space for optimal parameters in a typical complex deep learning (DL) system is vast even by today’s computing standards. Conventional DL algorithms are complex, are extremely computationally expensive, and often result in suboptimal solutions, negatively impacting performance and generalization. Quantum computing (QC) offers the potential for rapid training of such models and the ability to find better solutions quickly in these large search spaces. We plan to demonstrate that our QC-based training can find better parameters than conventional parameter optimization approaches, as well as prior approaches utilizing QC and overcome the deficiencies of current DL technology on state-of-the-art challenges such as automatic interpretation of electroencephalography (EEG) signals.

Generative models based on undirected probabilistic graphical models, such as Markov Random Fields (MRFs), paved the road for the deep learning revolution. Unfortunately, classical training algorithms for undirected models, which include Markov Chain Monte Carlo\ sampling, have taken a back seat to other machine learning (ML) models utilizing backpropagation, due to their computational complexity and lack of scalability on popular high performance computing architectures.

In addition to the lack of scalability of the learning approaches used for undirected models, there still remain concerns about how closely the model distribution of an MRF can approach the target distribution. Despite those difficulties, MRFs remain very promising, in part due to their ability to learn and model very complex probability distributions, as well as their suitability for unsupervised DL of training data deprived of classification labels. However, previous attempts of training MRFs with QC resulted in only modest improvements.

We propose a novel way of applying Quantum Annealing Computers (QACs) to training and sampling from Deep Boltzmann Machine (DBM) model distributions. Our preliminary results revealed the ability of QACs to sample states (finding local valleys in the energy function) that do not suffer from many limitations of classical sampling techniques. We propose a method that combines QAC and classical sampling that will achieve unprecedented improvements in trainability of undirected probabilistic models on many real-world complex datasets. Our preproposal results indicated that the use of QAC supplements the sample with states that have high probability and represent important regions of the probability distribution but are often missed by classical search/sampling algorithms because they have narrow basins of attraction. Identification of such states is critical for enabling the sample to reflect the effects of rare training events that often arise in unbalanced data sets. Accurate classification of rare events is extremely important in a wide range of DL applications where these rare events contain important information (e.g., drug discovery).

*Keywords: quantum computing, quantum annealing, machine learning, deep Boltzmann machines*

1. Intellectual Merit

The primary intellectual merit of this proposal is to advance learning algorithms for complex DL systems by accelerating the training process and improving accuracy. By applying QAC to DBMs, we will demonstrate the efficacy of QC-based methods on two extremely challenging ML “needle in the haystack” tasks: (1) seizure detection on EEG signals, which involves learning effective temporal and spatial dependencies, and (2) automatic interpretation of digital pathology images, which involves analyzing extremely large high-resolution images (50K x 50K pixels) by integrating local and global dependencies.

1. Broader Impacts

The primary immediate broader impact of this proposal is that it will make training of larger, more complex deep learning systems more tractable. Demonstrating performance improvements on two important bioengineering applications will attract attention from researchers across a wide range of fields who face applications where data is limited. This will allow DL technology to be applied to a much broader range of bioengineering applications that can significantly impact healthcare. This will be one of the first significant applications of QC to mainstream ML problems where technological progress is well calibrated.

 Xx xx xx Xx xx xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx Xx xx xx Xx