Luke Godfrey
Statement of Research Philosophy

Machine learning and artificial intelligence have exploded in popularity in the past several years. Thanks to advancements in hardware and a few important discoveries, machine learning has become not only feasible but unexpectedly effective at solving problems, often surpassing even human ability. Tasks that were once thought to be beyond the capabilities of computers are now being performed at human or super-human levels by artificially intelligent agents: self-driving cars, handwriting recognition, medical diagnosis, translation between languages, and even video gaming (which has implications for artificially intelligent agents in the real world). Artificial neural networks, inspired by the brain, are among the most successful machine learning models and are the primary subject of my research.

Much of my work, both in academics and in industry, has been on the application of artificial neural networks to the task of time-series analysis and multi-step forecasting. This type of problem poses unique challenges, but is particularly useful in many domains. My work in this area builds on key insights from traditional statistical time-series approaches, utilizing non-traditional neural network components and implicitly learning seasonalities and trends. This is not only useful for typical time-series problems such as sales forecasting or market predictions, but also for modeling dynamical systems and learning to anticipate in such an environment, which is especially important for automated systems like robots or self-driving vehicles.

I have also researched the use of artificial neural networks for various classification tasks, including handwriting recognition and image labeling. The goal is to take high-dimensional data and encode it in a meaningful way; in other words, to understand the data and to be able to map raw inputs to meaningful concepts. Much of the research in this area uses huge neural networks that take a network of high-end computers months to train. My work seeks to train smaller networks in a strategic, targeted way such that, with less time and fewer computational resources, the smaller network can reach the same level of accuracy as a much larger network, speeding up training and enabling greater scalability. One method for accomplishing this is allowing the network to learn things that are typically hard-coded, such as the activation function. The demand for faster-training neural networks will continue to rise as machine learning applications become more and more prevalent and are used with larger and increasingly complex tasks.

Currently, I am researching methods of multivariate time-series forecasting. While many approaches to forecasting train a new model for every task, some problems lack the volume of historical data required to sufficiently capture seasonalities but are composed of many similar tasks. By applying multi-task learning to time-series analysis, a model can leverage information from one time-series when making predictions for another. This kind of approach can be used to synthesize many kinds of data, such as the images and sound in a video. It can also be used to transfer learning from one domain to another, such as training an artificially intelligent agent inside a simulation and using that trained model to power a machine in the real world.

One of the most important problems to be solved in artificial intelligence is that of scalability. Although neural networks have surpassed human ability in a number of noteworthy applications, trained models rarely generalize outside their intended tasks. Although theoretically possible, transferring learning from a simulation to the real world is particularly difficult in practice due to an explosive increase in the number of variables and amount of noise. I believe that this barrier can be overcome using some combination of 1) developing techniques for training more robust models, 2) constructing more realistic simulations, and 3) using augmented reality as a bridge from the simulation to the target application. As virtual and augmented reality is becomes more widespread, this third approach is especially interesting, as it potentially combines computer vision, reasoning, and human-AI interaction.

A related problem is that of real-time adaptation in uncertain environments, a topic which the NSF has recently expressed interest in funding (https://www.nsf.gov/cise/iis/ri_pgm12.jsp). Humans are exceptionally adept at making quick, almost instinctive decisions in real-time when faced with a new problem; artificial intelligence, in its current state, simply lacks this kind of flexibility. For example, autonomous vehicles are good at driving when everything is predictable, such as clear lane markings, well-lit street signs, and no surprises. These systems have difficulty, however, in other circumstances, such as construction zones or when an animal crosses the road unexpectedly. The key to solving this problem is the application of dimensionality reduction techniques that promote meaningful low-dimensional encodings of raw sensor data, enabling an autonomous system to relate events based on appropriate reactions rather than on raw correlation.