One-Shot Transfer Learning of Non-linear Ordinary and Partial Differential Equations

Hari Raval
Harvard University
hraval@g.harvard.edu

Pavlos Protopapas
Harvard University
pprotopapas@g.harvard.edu

Abstract

Efficiently and accurately solving differential equations is essential across many fields of research, including biology and physics. Traditional methods of solving these differential equations make use of iterative numerical methods such as Runge-Kutta. Recently, the use of Physics Informed Neural Networks has gained popularity due to the numerous benefits they provide over numerical approaches. One area of research that remains under-explored is the application of transfer learning to solving these problems. Applying transfer learning to solving differential equations wold mean that highly accurate solutions to unknown differential equations can be obtained instantly, without having to train complex networks from scratch. In this work, we aim to build a framework that applies transfer learning to Physics Informed Neural Networks to solve non-linear ordinary and partial differential equations. We will use a feature representation to linerize the neural network model for special cases and then generalize it to solving these more complex problems.

1 Motivation and Background

The primary motivation for choosing this project is to fulfill my research interests in machine learning applied to new, real-world applications that have a broader impact. Transfer learning is a revolutionary area of research that has illustrated significance in fields such as computer vision and natural language processing. Applying this technique to solving differential equations is largely unexplored, yet has been shown to work for specific linear cases [1]. In this regard, this project has the unique characteristic of being novel, yet showing proof-of-work via prior research conducted in Dr. Protopapas' research group.

The background that motivates my choice of this project can be summarized in terms of my academic background, current course work, and and future plans. First, my undergraduate degree was in computer science with a minor in machine learning. With my studies at Princeton University, I was able to learn the fundamentals of machine learning and apply it to a smaller semester-long research project in natural language processing. Hence, by working on this current project with Dr. Protopapas, I will be able to build upon and expand my knowledge of machine learning through an application that I have never worked on before. Moreover, this project combines my interests in mathematics, computation, and data science. I am pursuing a Computational Science and Engineering master's degree at Harvard, yet I am equally interested in data science and its various applications. In this regard, this project aligns well with my curiosity for both computational science as well as data science.

Ultimately, my future plans revolve around working as a data scientist in industry. The current project is largely computation-based and will provide me with a year-long exposure into applying cutting-edge tools to a novel problem, which can more generally prepare me for work after the thesis.

I plan on working on application-driven work in data science, as opposed to strictly theoretically-oriented problems.

2 Goals

At a high-level, the first goal of this project is to produce a generalization of the work done on one-shot transfer learning for linear ordinary and partial differential equations [1]. In other words, we want to use a feature representation to linearize a model for special cases and then see if it can generalize to non-linear cases. Moreover, we want to validate the effectiveness of this approach by comparing it against the true solutions of specific non-linear differential equations. The second goal of this project is to produce a framework that can work well for multiple equations, enabling usability across a wider set of non-linear equations. Finally, the third goal is to produce a publication of this work, assuming it yields reasonable success in it generalization.

3 Scientific Justification

Differential equations are essential in a broad variety of applications, having notable impact in fields such as biology in modeling epidemics and physics in modeling motion. Due to this breadth of relevance across many domains, efficiently solving these equations is important. Traditional methods of solving these equations include using iterative numerical methods such as Runge-Kutta, however Physics-Informed Neural Networks have also appeared as a viable alternative to solving these equations [4]. Physics-Informed Neural Networks leverage backpropagation to compute partial derivatives and so can enforce a known differential equation via a loss function. One benefit of using such neural networks to solve differential equations is the ability to exploit transfer learning. Transfer learning enables the storing of knowledge gained in one problem and then applying it to a different, but related problem. Currently, applying transfer learning to solve differential equations remains under-explored, despite the benefits it can provide in terms of avoiding the training of the networks from scratch.

With this in mind, it becomes evident that transfer learning is well-motivated for application in solving differential equations. Prior work mentioned in Section 4 below illustrate that transfer learning has been successfully applied to linear ordinary and partial differential equations. My project is innovative in that it can build off of this prior research by generalizing the work to non-linear ordinary and partial differential equations. These non-linear equations are significant in modeling complex, real-world systems that exhibit non-linear properties.

4 Literature Review

Prior work has been done in both creating the networks to solve ordinary and partial differential equations and in solving the equations themselves. Namely, researchers have shown that Physics Informed Neural Networks can be used to solve supervised learning tasks while satisfying given physics laws described by general nonlinear partial differential equations [4]. This work is significant in that it serves as an alternate to traditional numerical methods in solving these equations. Next, work has been done which creates a general framework for performing transfer learning on Physics Informed Neural Networks that results in one-shot inference for linear systems of both ordinary and partial differential equations [1]. The work by these researchers is able to instantly generate accurate solutions to a set of unknown differential equations, such as the Poisson equation, without retraining a network from scratch.

In addition to Physics Informed Neural networks, relevant prior work has been done in developing a method for solving differential equations with unsupervised neural networks that apply Generative Adversarial Networks to learn the loss function for optimizing the neural network [5]. This research has yielded empirical results that illustrate significantly lower mean squared errors than an alternative unsupervised neural network method.

Finally, application-driven work also exists in multiple domains. One recent project studied the spread of COVID-19 through the use of a semi-supervised neural network. The researchers started with an unsupervised neural network that learned solutions of differential equations for different modeling parameters and initial conditions. After this, their supervised method solved the inverse problem by determining the optimal conditions that generate functions which fit the data for those infected by, recovered from, and deceased due to COVID-19 [3]. Another project focused on creating a network for discovering eigen functions and eigenvalues for differential eigenvalue problems. This project utilized a network optimization that was data-free and depended only on the predictions. The unsupervised method is used to solve the quantum infinite well and quantum oscillator eigenvalue problems [2].

5 Methodology

At a high-level,the methodology for my project will involve three key components: 1.) deriving the mathematical formulation of the problem 2.) creating a network to represent the formulation 3.) evaluating the network's results. With regards to software, the modeling code will all be written in Python, relying on the PyTorch library for machine learning. The primary techniques used for the computational portion, as mentioned earlier, are building, training, and evaluating neural networks and applying transfer learning to them.

References

[1] Shaan Desai, Marios Mattheakis, Hayden Joy, Pavlos Protopapas, and Stephen J. Roberts. One-shot transfer learning of physics-informed neural networks. *CoRR*, abs/2110.11286, 2021.

[2] Henry Jin, Marios Mattheakis, and Pavlos Protopapas. Unsupervised neural networks for quantum eigenvalue problems, 2020.

[3] Alessandro Paticchio, Tommaso Scarlatti, Marios Mattheakis, Pavlos Protopapas, and Marco Brambilla. Semi-supervised neural networks solve an inverse problem for modeling covid-19 spread. *CoRR*, abs/2010.05074, 2020.

[4] M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[5] Dylan Randle, Pavlos Protopapas, and David Sondak. Unsupervised learning of solutions to differential equations with generative adversarial networks. *CoRR*, abs/2007.11133, 2020.