Using Machine Learning to Improve Numerical Weather Prediction

Conor Casey
Project Proposal for the BTYSTE 2021

September 29, 2020

This project is a continuation of my SciFest project and provides major enhancements over the original project (it must be noted that the original project won the Science Foundation Ireland Best Project Award in IT Tralee). Weather forecasting has traditionally been done by physical models of the atmosphere, which are unstable to perturbations, and thus are inaccurate for large periods of time. Building a simulator of a complex dynamical system, such as the atmosphere, can also involve years of work, and even the best simulators are often inaccurate due to insufficient knowledge of, or difficulty in approximating, the underlying physics and parameters. An attractive alternative to traditional simulators is to use neural networks to train simulators directly from observed data. As such, the purpose of this project is to investigate the utilisation of neural networks in combination with existing physical atmospheric models to determine if such solutions are viable and more accurate than using physical models alone.

Hypothesis

The hypothesis of this project is that, it is possible to train a recurrent neural network on an atmospheric reanalysis dataset based on data from the 15 years, that such a neural network captures crucial weather patterns and can predict the future evolution of atmosphere, and that such a machine learning model will ultimately improve numerical weather prediction in comparison to established physics-based models.

Benchmarking Method

Benchmarks are necessary to contextualise the performance of a weather forecast generated by any given scheme. Five select metrics, normalised root mean squared error being an example of such a metric, will be used to examine the performance of the software in comparison with climatology, a persistence forecast, and a traditional deterministic physics-based model operated by the National Oceanic and Atmospheric Administration. The execution time required to generate a weather forecast by the software will also be measured and recorded. This parameter will provide insight into how the software could potentially be incorporated into a meteorological setting, and it will also test the assumption that utilising a machine learning model will lead to a significant decrease in the computational resources required to generate a weather forecast.

Time Frame

In regards to time frame, a variation of this particular project has actively worked on for the last two years. The focus in recent few months, however, has shifted to the application of machine learning in numerical weather prediction. Due to the time constraints imposed by the Leaving Certificate, it was necessary to complete the vast majority of the project work during the summer period. As such, the machine learning model and the accompanying software has been programmed and tested extensively, with the majority of the benchmarking similarly being complete. Hence, I do not foresee any obstacle that could obstruct the completion of this project in time for the exhibition in January.