**Time Series Forecasting in Chaotic Dynamical Systems [ORIGINAL]**

*Lucas Wilson*

**Interest**

*Explain interest in the subject*

I have an interest in chaotic systems, specifically systems of emergence such as those in ecosystems. By nature, they are very difficult to predict, so I am interested in using modern machine learning algorithms and measure their effectiveness in solving chaotic time series forecasting problems compared to traditional models. Examples of chaotic systems include weather patterns, ecosystems in nature, and stock markets, so research in this area would be valuable to a variety of fields.

**Project Description**

*Central themes, research questions, design project, and/or creative activity*

What are the creative activities, research questions, and/or design features that are foundational to your thesis?

I will be using recurrent neural nets (RNNs) to perform time series forecasting on chaotic problems: such as the double pendulum, Lorenz system, and emergent systems found in nature. The Lorenz system is a set of differential equations modeling atmospheric convection, so it could have useful application in forecasting weather. Emergent systems are systems defined by simple rules, but the results can’t easily be described by a combination of those simple rules: for example, the crystallization of water in a snow flake or the movement of a school of fish. Other simpler chaotic systems such as Conway’s Game of Life can be built for a baseline understanding of chaos and testing of my models. Conway’s Game of Life is a step-time series system (as opposed to a continuous one, such as the Lorenz problem) with simple rules, so it would be a good initial test during the developmental phase of building a recurrent neural network. The creative aspect lies in developing this model. I need to consider using autoencoders which will reduce the dimensionality of a set of data points. They do this by mapping the high dimensional space to a lower one, and then from the lower dimensional space, recreating the higher dimensions. It results in two functions: an encoder and a decoder. The reduced dimensionality can increase accuracy and reduce training time for the RNN. I can research how well the autoencoders reduce dimensionality and how well the RNN performs afterwards. Another creative aspect lies in understanding the nature of chaotic data. This involves generating or collecting chaotic data. In order to understand the data, I will investigate how well visualization techniques such as t-sne perform. Like autoencoders t-sne can also be used to reduce dimensionality. I can research how well t-sne can group chaotic data. T-sne takes high dimensional data and reduces it to a lower dimension while preserving proximity in the space. Due to the sensitivity of chaotic systems, points which are close in space are likely not close after a few time steps. T-sne will be most interesting in seemingly periodic systems like the Lorenz system. In each of these models--t-sne, RRN, autoencoders--, I can explore different configurations for each of the models individually which will affect how well they perform.

**Major Tasks**

How will you discover answers to research questions, create artistry, and/or develop design projects?

First, I will need to understand the current research on the topic of recurrent neural networks, time series forecasting, t-sne, and chaotic systems. I will do this by reading available information explaining the topics and by reading papers written on the topics. Once I understand how the topics work, I can begin to design software to implement RNNs to solve time series forecasting problems. Then, I will need to collect data. To generate chaotic data, I would have to simulate a chaotic system and collect data from it. To collect data, I would have to search through available databases and convert it from its given format into a form which my RNN model can use. The research I do in the first step will help me understand what format the data needs to be in in order to work most effectively. Then, I will investigate the data with autoencoders and t-sne. With autoencoders, I will be looking at how much they can reduce the dimensionality of the collected data while still being able to preserve the data’s original form. With t-sne, I will investigate how effective it is in mapping the data to a low dimensional space. This will involve looking at graphs to see if clusters are preserved, and it will involve identifying what a cluster is. Finally, I will build and train the RNN with the output of the autoencoders. I will also test whether it performs better without the autoencoders, but as is the nature of autoencoders, this will likely not be the case. Lastly, I will analyze how well the RNN was able to forecast the chaotic systems specified earlier, and I will conclude with how useful they might be on chaotic systems in general.

**Key Resources**

*6-12 book, material, people, journals*

**Timeline of Completion**

*Penultimate draft to advisor and committee members*

*Presentation date*

**Final Product Description**

*Writing component is one of the parts*

Describe the creative activity and the expected research result, final product, and/or artistic expression that is the core of your thesis.

The creative activity is building and designing the model. I will experiment with different techniques until I am able to find one that works very well specifically with chaotic systems. The final result will be an understanding of how well RNNs can solve chaotic time series forecasting problems such as the Lorenz system, double pendulum, and emergent systems related to nature.