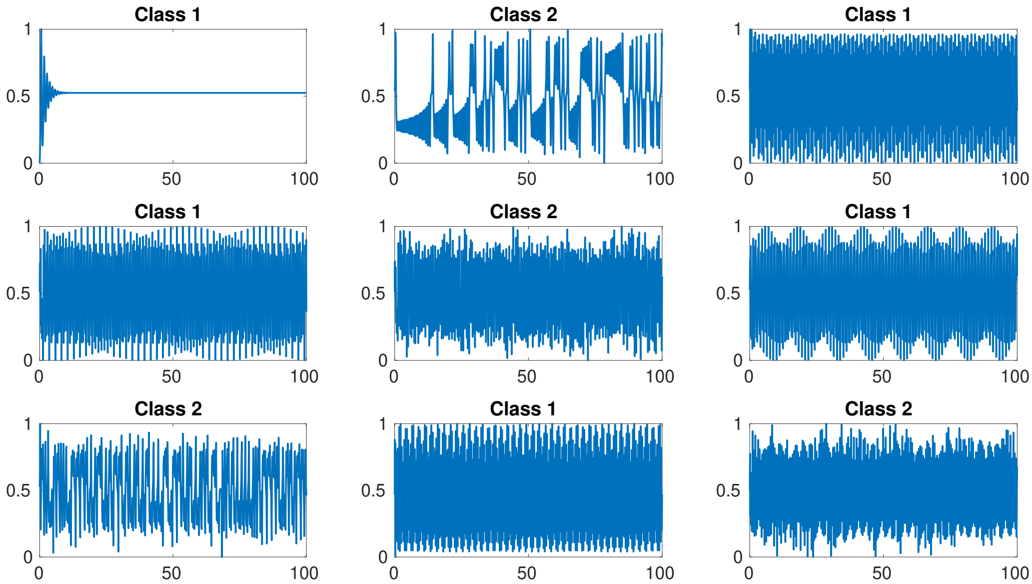
EPSRC
Engineering and Physical Sciences Research Council

InFoMM
Industrially Focused Mathematical Modelling

EPSRC Centre for Doctoral Training in Industrially Focused Mathematical Modelling



Classification of chaotic time series with deep learning

Nicolas Boullé

UNIVERSITY OF OXFORD

CCFE
CULHAM CENTRE FOR FUSION ENERGY

Contents

1 Introduction **1**

Dynamical systems and chaos 1

Deep learning 2

Glossary of terms 3

2 Chaotic time series classification **3**

Discrete dynamical systems 3

Continuous dynamical systems 5

3 Discussion, conclusions, & recommendations **6**

4 Potential impact **6**

References **6**

1 Introduction

The primary aim of scientific research at the Culham Center for Fusion Energy (CCFE) is to enable nuclear fusion as a viable source of energy. CCFE is a key player in the experimental ITER reactor, which is an international nuclear fusion research project under construction in the south of France. Nuclear fusion involves controlling plasmas at temperatures at 100 Million Kelvin, which is ten times the temperature of the sun. Unwanted turbulences can occur in the tokamak due to the huge temperature and density gradients. One of the challenges for CCFE is to rapidly identify such chaotic scenarios when they occur in order to stabilize the plasma.

Traditional time series classification approaches are computationally expensive.

Time series are generated from many experiments across different scientific fields such as nuclear fusion, quantum mechanics, biology, or medical observations. The size of the data sets is often large and analysing these time series represents a huge computational challenge. However, the spectacular success of using machine learning and deep learning techniques for image classification, which have recently surpassed human-level performance on the imagenet data set, has inspired the development of neural network techniques for time series forecasting and classification. Neural networks are inspired by networks in the brain and can learn to perform tasks by considering examples. Machine learning has also been applied to the field of dynamical systems in order to perform model-free predictions of chaotic dynamical systems. However, deep learning requires a consequent data set to train the neural network, which might not be available due to experimental constraints.

Our aim is to use a deep-learning approach to automatically classify time series generated by discrete and continuous dynamical systems according to their potential chaotic behaviour. Contrary to standard machine learning techniques, our focus will be on training the neural network on a different and simpler set than the *testing set* of interest in order to simulate the lack of experimental training data. The main challenge is to learn the chaotic features of the *training set*, without overfitting, and generalise on a second data set, which behaves differently.

Dynamical systems and chaos

A *discrete* dynamical system is defined by the iterations of a function f. Iterating nonlinear but simple functions sometimes leads to complex and unpredictable results.

Unlike discrete dynamical systems, *continuous* dynamical systems depend on a continuous parameter t, which we call time. Continuous dynamical systems often arise in applications in biology or physics. One of the most famous examples is the Lorenz system, originally introduced by Lorenz as a model for atmospheric convection. A solution to the Lorenz system is shown in Figure 1, where the blue line demonstrates the trajectory passing through a particular initial condition.

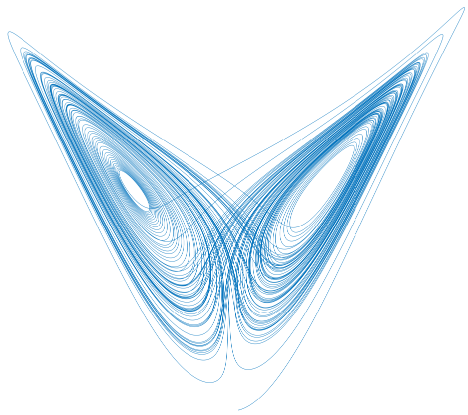


Figure 1 – A solution to the Lorenz system.

These discrete and continuous nonlinear systems can have different and complex behaviours, which can lead to periodic or chaotic patterns. Chaos theory has been extensively studied during the twentieth century and appears in many applied fields such as climate prediction, biology, or road traffic. Although there is not a universal definition of this phenomenon, Hilborn wrote [1]:

> *The noun chaos and the adjective chaotic are used to describe the time behaviour of a system when that behaviour is aperiodic (it never exactly repeats) and is apparently random or "noisy".*
>
> Robert C. Hilborn

In Figure 2, we illustrate a chaotic behaviour, and more particularly the sensitivity to initial conditions, of a dynamical system defined by repeatedly applying the function $f(x) = 3.9x(1-x)$. We observe that two trajectories, starting from $x_0 = 0.5$ (blue) and $x_0 = 0.499$ (red), rapidly diverge from each other, which is one characteristic of chaos.

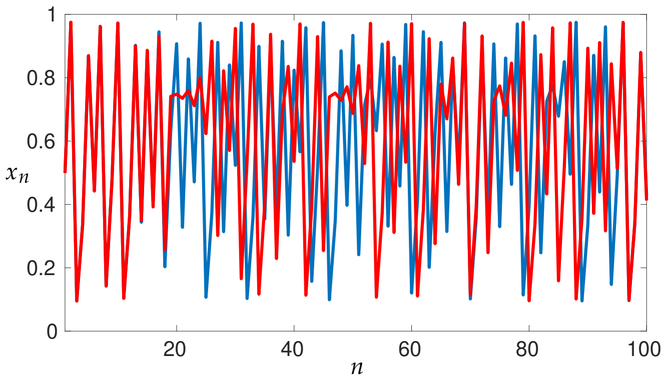


Figure 2 – The chaotic map $f(x) = 3.9x(1-x)$, iterated from $x_0 = 0.5$ (blue) and $x_0 = 0.499$ (red).

> The Lyapunov exponent of a simple dynamical system is easy to compute, and quantify sensitivity to changes.

Given a dynamical system or time series from experimental data, we are interested in methods for quantifying or estimating the underlying chaotic behaviour. These quantifiers are particularly interesting to discriminate chaotic from noisy behaviour or sort systems into classes. We use the Lyapunov exponent, which is a measure of how fast two close trajectories diverge from each other in time, to quantify the sensitivity of the system to changes in the initial condition. A positive Lyapunov exponent corresponds to a chaotic behaviour.

Deep learning

Artificial intelligence has became a major field with applications in many research and industrial areas such as image recognition, natural language processing, speech recognition, or drug discovery. Artificial intelligence techniques are used to find structures in a data set in order to perform prediction or classification. The artificial intelligence evolution has been made possible due to the development of computational power contained in graphic processing units (GPUs) or multi-core processors. Another factor is the increasing access to large data sets that can be used to train the artificial intelligence. Moreover, the development of high level Python libraries such as Tensorflow or Pytorch have helped make the field more accessible.

One of the major successes of artificial intelligence goes back to the 90s with the defeat of the Chess world champion Garry Kasparov by Deep Blue, an artificial intelligence created by IBM, while a more recent breakthrough is the first victory of a Go playing system, AlphaGo developed by Google Deepmind, against a Go world champion. This was mainly achieved thanks to the development of a technique called deep learning.

In standard machine learning, the features used to represent the data have to be provided by the user. The aim of deep learning is to automatically construct features to describe the data using a neural network. The depth of the network determines the range of features available, with a deep neural network able to identify complex patterns in a given set.

Glossary of terms

- **Artificial neural network:** A trainable computational model for forecasting or classification.
- **Bifurcation parameter:** A parameter of a dynamical system which leads to different behaviours.
- **Chaotic behaviour:** The behaviour of a system which is characterised by a small change in a critical parameter giving a large change in the solution.
- **Convolution:** A mathematical operation which highlights features in a dataset.
- **Partial differential equation:** A mathematical equation depending on spatial and temporal variations used to describe physical phenomena.
- **Testing set:** A dataset used to evaluate a model given by a neural network.
- **Time series:** A sequence of points in time.
- **Training set:** Set of examples used to determine the parameters of a neural network.

2 Chaotic time series classification

Time series classification is one of the most challenging problems in machine learning with a wide range of applications in human activity recognition, acoustic scene classification, and cybersecurity. Our aim is to perform chaotic time series classification using deep learning.

The problem can be stated as follows: given a time series generated by a dynamical system (discrete or continuous), can we determine whether the time series has a chaotic or non-chaotic behaviour ?

Moreover, our aim is to be able to classify a given time series while training the network on a time series generating by a different dynamical system whose chaotic behaviour can be determined *a priori* using the its Lyapunov exponent.

> Convolutional neural networks can identify the main characteristics of a dataset without prior knowledge.

Convolutional neural networks (CNNs) were first introduced to perform handwritten digit recognition and have been successfully applied to images and time series. This type of neural networks consist of a succession of operations called *convolutions* on the input data set and is able to efficiently capture complex features.

We will use the neural network depicted in Figure 3 to perform time series classification. The input data passes through two convolutional layers, followed by two fully connected layers from which we obtain the classification.

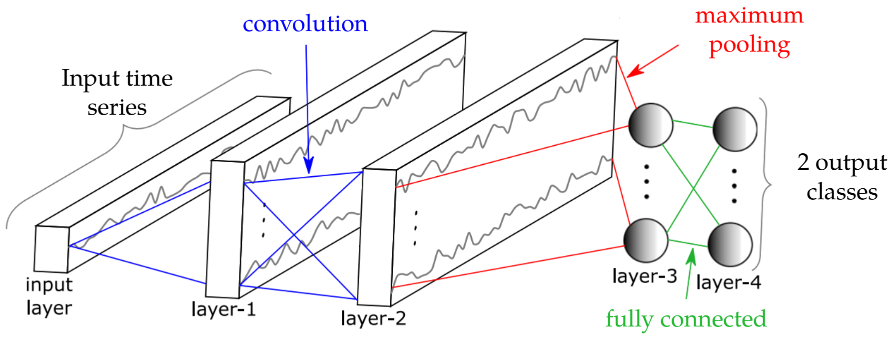


Figure 3 – A convolutional neural network architecture for time series classification. Figure adapted from [2].

Discrete dynamical systems

We consider two discrete dynamical systems called the logistic map and the sine-circle map. The logistic map is defined by repeatedly applying the function $f(x) = \mu x(1 - x)$, where μ is a *bifurcation parameter* varying between zero and four. This system exhibits periodic or chaotic behaviour depending on the value of μ. In Figure 4, we see that with two different values of μ we find two different kind of solutions of the logistic map, namely periodic and chaotic sequences.

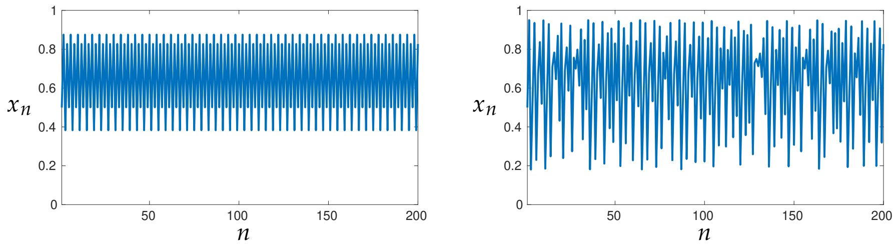


Figure 4 – A periodic (left) and a chaotic (right) sequences of length two hundred of the logistic map with respective parameters $\mu = 3.5$ and $\mu = 3.8$.

The second dynamical system considered in this section is the sine-circle map, which is sometimes referred to as the circle map. It is defined by iterating a nonlinear map, which depends on two parameters Ω and μ, where μ is the relevant parameter, which measures the strength of the nonlinearity.

Similarly to the logistic map, iterating the sine-circle map leads to periodic or chaotic signals depending on the value of the bifurcation parameter μ. We illustrate two signals with different behaviours, generated using a bifurcation parameter of $\mu = 2.1$ (left) and $\mu = 2.3$ (right), in Figure 5.

Sine-circle signals look significantly different from logistic signals.

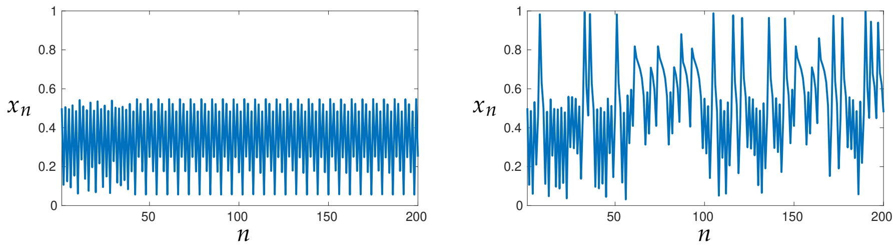


Figure 5 – A periodic (left) and a chaotic (right) sequences of length two hundred of the sine-circle map with respective parameters $\mu = 2.1$ and $\mu = 2.3$.

We now want to classify signals generated by the logistic and sine-circle maps according to their chaotic and non-chaotic behaviour. Our main goal, and challenge, is to find a neural network that is able to learn the features characterising chaotic sequences using data from the logistic map and generalise them to sequences generated by the sine-circle map.

To do this, we generate two data sets, one for the logistic map and one for the sine-circle map, by computing sequences of length one thousand for five thousand different values of the parameter μ. We randomize the logistic map data set across the bifurcation parameter and we choose two thirds of the data to be the training set.

We compute a classification of the training and testing signals in order to provide the expected classification to the neural network during the training phase and validate the output of the neural network on the testing set. This classification of chaotic time series is done using the Lyapunov exponent.

In real applications, computing the Lyapunov exponent of a time series without knowing the expression of the underlying dynamical system can be infeasible or computationally expensive, which justifies the approach of using a machine learning algorithm to perform the classification automatically.

We found that state-of-the-art neural networks for time series classification classify sequences from the sine-circle map with an accuracy less than 65%. Our convolutional network, however, seems to override overfitting issues on the training set by capturing the main features of chaotic and periodic sequences and gets and achieves an average classification score of 83.5%.

Continuous dynamical systems

We now consider continuous dynamical systems generated by partial differential equations that exhibit temporal or spatiotemporal chaos. The aim is to determine whether a neural network trained on a specific continuous dynamical system is able to generalise and classify time series generated by another system.

We first consider the Lorenz system since it is one of the simplest continuous dynamical systems able to generate chaotic time series. The Lorenz system is defined by three partial differential equations and depends on three parameters: σ, β, and ρ. We fix the first two and vary the last variable to generate non-chaotic and chaotic time series (as shown in Figure 6). Moreover, these time series are classified according to their Lyapunov exponent and constitute our training set for the neural network.

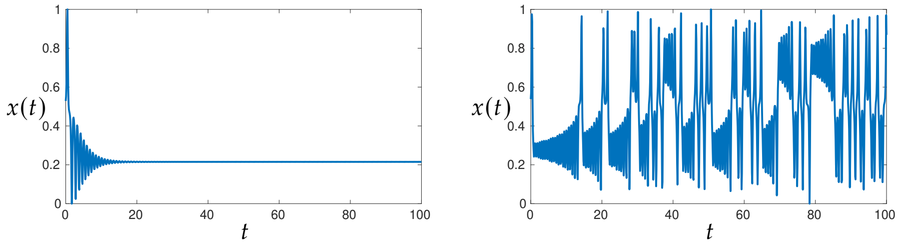


Figure 6 – A non-chaotic (left) and a chaotic (right) time series generated by the Lorenz system.

We then consider the Kuramoto–Sivashinsky (KS) equation, which is an example of partial differential equation which exhibits spatiotemporal chaos depending on its birfurcation parameter α. This equation was originally derived to model instabilities in laminar flame fonts and arises in a wide range of physical problems such as plasma physics, flame propagation, or free surface film flows.

Solving this equation for different parameters α yields a wide range of solutions as illustrated in Figure 7.

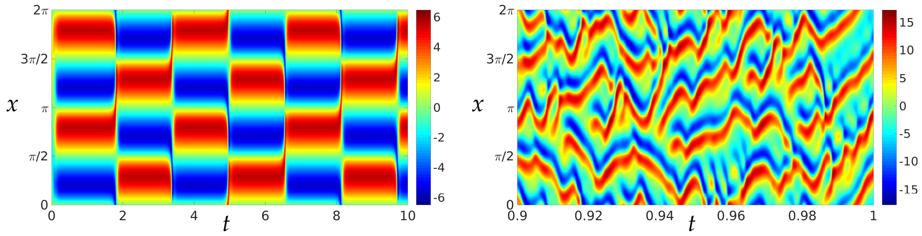


Figure 7 – A non-chaotic (left) and a chaotic (right) solutions to the Kuramoto–Sivashinsky equation.

The time series describe the main time variations of the solutions to the KS equation.

We solve the KS equation numerically for various α and we generate a testing set of one thousand time series, equally divided between chaotic and non-chaotic behaviours.

We observe a global accuracy of 94.4% in classifying the KS time series between chaotic and non-chaotic after training the convolution neural network on the Lorenz training set. In particular, convergent and chaotic time series are classified correctly with a score above 95% while low frequency time series are mostly misclassified by the neural network. We expect this misclassification to be due to qualitative differences between the corresponding time series of the KS equation and the periodic time series of the Lorenz system. In particular, the KS data set contains periodic time series with low frequency oscillations in this regime, while the Lorenz system generates periodic time series with high frequency oscillations. The neural network is then unable to classify features that are not present in the training set and hence fails to extrapolate to the low frequency periodic time series of the Kuramoto–Sivashinsky equation.

3 Discussion, conclusions, & recommendations

Many challenging real-life applications are so complicated that they cannot be precisely modelled, which makes the identification of different dynamical behaviours unfeasible or in the best case scenario computationally expensive. For this reason, we have introduced a deep learning approach for classifying time series generated by discrete and continuous dynamical systems. Our approach is to train our neural network on time series obtained by solving a basic dynamical system and then to use this network to classify time series of a more intricate dynamical system.

We developed a convolutional neural network which is able to learn the chaotic features of these systems and classify with high accuracy. We applied our approach to classify time series generated by the sine-circle map and the Kuramoto–Sivashinsky equation, using the logistic map and the Lorenz system as training data sets, respectively. We observed a classification accuracy greater than 80% on the sine-cicle map and more than 90% on the KS equation. Our method outperforms state-of-the-art neural networks, which tend to overfit the training data set.

Further developments require the study of more complicated dynamical systems.

There are many directions in which our results can be extended. In particular, attempting to classify time series obtained from real-life applications is crucial. In that respect, the effect of noise in the training and testing data sets is an important aspect to be studied in order to determined the influence of the noise on the accuracy of the networks to classify the time series.

4 Potential impact

The development of deep learning techniques, which can generalise the knowledge acquired from a training set to give meaningful information about a different testing set, is likely to be valuable for classifying time series obtained from real-life applications.

Dr Debasmita Samaddar, Computational Plasma Physicist at CCFE, said: "*Harnessing a turbulent plasma at very high temperatures is one of the biggest challenges in the path towards achievement of commercial fusion reactors. Being able to predict turbulent behavior is a major outcome of Nicolas' work. Although he has demonstrated his case with problems simpler than a fully developed turbulent scenario as encountered within a fusion device, the work paves the way to a breakthrough in tackling turbulence.*"

References

[1] R. C. Hilborn. *Chaos and nonlinear dynamics: an introduction for scientists and engineers*. Oxford University Press on Demand, 2000.

[2] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller. Deep learning for time series classification: a review. *Data Min. Knowl. Disc.*, pages 1–47, 2019.