Artificial Neural Network and Time Series Analysis and Forecasting

1.1.1 Artificial Neural Network

Artificial neural networks (**ANNs**), usually simply called **neural networks** (**NNs**), are computing systems vaguely inspired by the biological neural networks that constitute animal brains.

An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons. An artificial neuron that receives a signal then processes it and can signal neurons connected to it. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called *edges*. Neurons and edges typically have a *weight* that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that a signal is sent only if the aggregate signal crosses that threshold. Typically, neurons are aggregated into layers. Different layers may perform different transformations on their inputs. Signals travel from the first layer (the input layer), to the last layer (the output layer), possibly after traversing the layers multiple times.

1.2.1 Describing vs Predicting

We have different goals depending on whether we are interested in understanding a dataset or making predictions.

Understanding a dataset, called *time series analysis*, can help to make better predictions, but is not required and can result in a large technical investment in time and expertise not directly aligned with the desired outcome, which is forecasting the future.

> *In descriptive modeling, or time series analysis, a time series is modeled to determine its components in terms of seasonal patterns, trends, relation to external factors, and the like. … In contrast, time series forecasting uses the information in a time series (perhaps with additional information) to forecast future values of that series*

1.2.2 Time Series Analysis

When using classical statistics, the primary concern is the analysis of time series.

Time series analysis involves developing models that best capture or describe an observed time series in order to understand the underlying causes. This field of study seeks the "*why*" behind a time series dataset.
This often involves making assumptions about the form of the data and decomposing the time series into constitution components.

The quality of a descriptive model is determined by how well it describes all available data and the interpretation it provides to better inform the problem domain.

The primary objective of time series analysis is to develop mathematical models that provide plausible descriptions from sample data

1.2.3 Time Series Forecasting

Making predictions about the future is called extrapolation in the classical statistical handling of time series data.

More modern fields focus on the topic and refer to it as time series forecasting.

Forecasting involves taking models fit on historical data and using them to predict future observations.

Descriptive models can borrow for the future (i.e. to smooth or remove noise), they only seek to best describe the data.

An important distinction in forecasting is that the future is completely unavailable and must only be estimated from what has already happened.

The purpose of time series analysis is generally twofold: to understand or model the stochastic mechanisms that gives rise to an observed series and to predict or forecast the future values of a series based on the history of that series

The skill of a time series forecasting model is determined by its performance at predicting the future. This is often at the expense of being able to explain why a specific prediction was made, confidence intervals and even better understanding the underlying causes behind the problem.

1.3.1 The Model

The model can be written as

yi=f(yt−i)+ϵtyi=f(yt−i)+ϵt

where **yt−1=(yt−1,yt−2,⋯,yt−8)′yt−1=(yt−1,yt−2,⋯,yt−8)′** is a vector containing lagged values of the series, and ff is a neural network with 4 hidden nodes in a single layer.

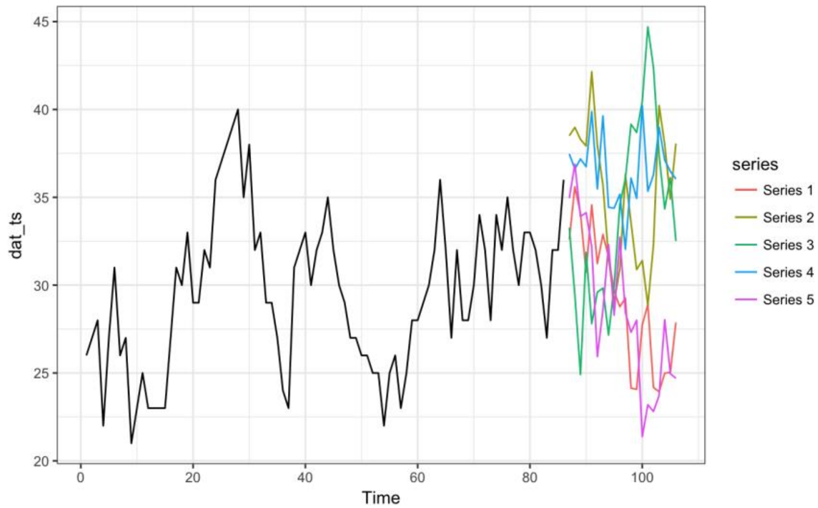
The error series **{ϵt}{ϵt}** is assumed to be homoscedastic (and possibly also normally distributed).
We can simulate future sample paths of this model iteratively, by randomly generating a value for ϵtϵt, either from a normal distribution, or by resampling from the historical values.

So if **{ϵ∗T+1}{ϵT+1∗}** is a random draw from the distribution of errors at time **T+1T+1** , then
y∗T+1=f(yT+1)+ϵ∗T+1yT+1∗=f(yT+1)+ϵT+1∗

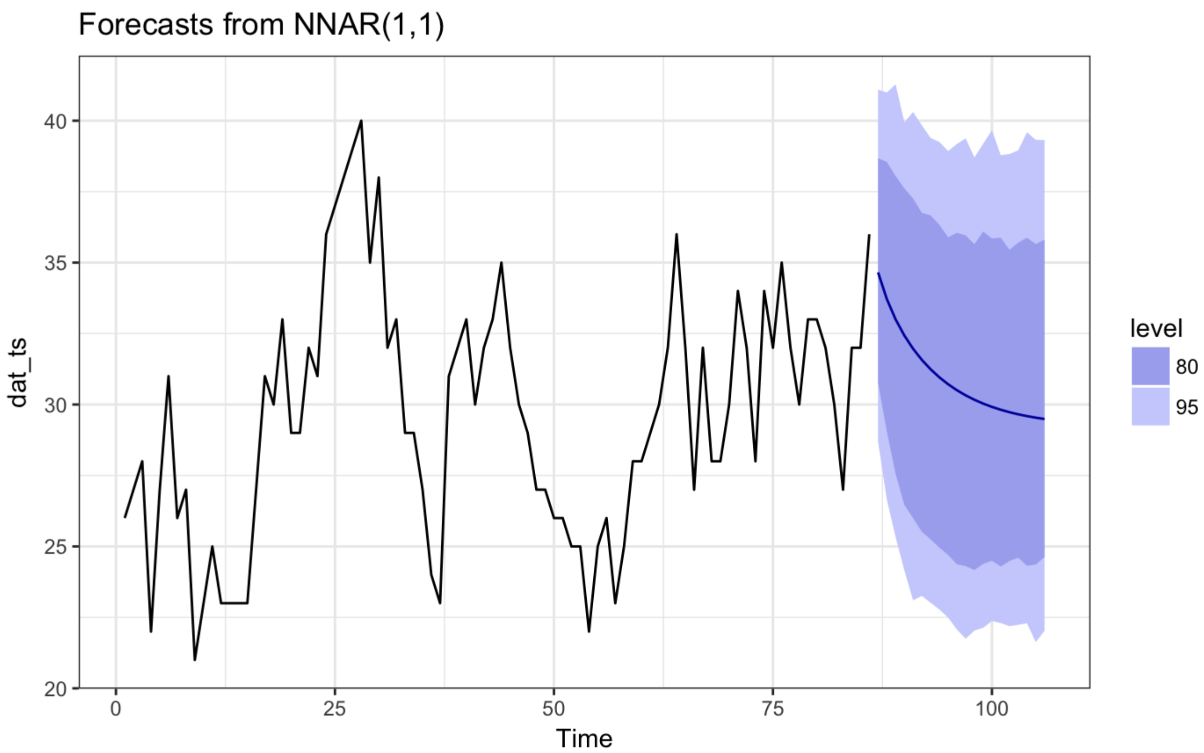
is one possible draw from the forecast distribution for **yT+1yT+1**. Setting **y∗T+1=(y∗T+1,yT,⋯,yT−6)′yT+1∗=(yT+1∗,yT,⋯,yT−6)′** we can then repeat the process to get

yT+2=f(y∗T+1)+ϵ∗T+1yT+2=f(yT+1∗)+ϵT+1∗

In this way, we can iteratively simulate a future sample path. By repeatedly simulating sample paths, we build up knowledge of the distribution for all future values based on the fitted neural network. Here is a simulation of 9 possible future sample paths for the lynx data. Each sample path covers the next 20 years after the observed data.



Now If we do this a few hundred or thousand times, we can get a very good picture of the forecast distributions.



Artificial Neural Network and Time Series Analysis and Forecasting

2.1.1 neuralquotes.com

Now let us take what we know about Artificial Neural Networks and incorporate into Time Series Analysis and Time Series Forecasting to predict how Ubers Stock will do over the next 30 days.

Data Provided is for Actuals is based on Yahoo Finance and reflects closing price of the NASDAQ composite avg as Market Close for each day since inception.

Artificial Neural Network and Time Series Analysis and Forecasting