← Go Back to Practical Data Science

Course Content

White Noise and Random Walk

A brief overview of White Noise and Random Walk

No matter how powerful they get, machine learning algorithms cannot predict everything. A well-known area where ML algorithms can become pretty helpless is time series forecasting. Despite the availability of a large suite of autoregressive models and many other algorithms for time series forecasting, we cannot predict the target distribution if it is White Noise or follows a Random Walk. So, we must first detect such unpredictable distributions before making any further efforts.

White Noise

The concept of White Noise is essential for time series analysis and forecasting. White Noise tells you if you should further optimize the model or not. White Noise is a series that's not predictable, as it's a sequence of random numbers, i.e, the next value does not rely on the sequential past values, instead, the values occur randomly. If you build a model and its residual (the difference between predicted and actual) values look like White Noise, then you know you've done everything to make the model as good as possible. However, if there are visible patterns in the residuals, there's potentially a better model for your dataset.

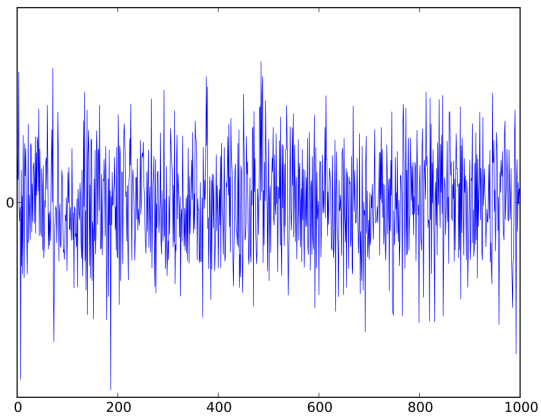
For a time series to be categorized as White Noise, the following conditions must be met:

. The mean value should be zero
. The standard deviation should be constant, it should not change over time
. There must be zero autocorrelation at all lags

Methods for determining if a time series resembles White Noise:

. By comparing the mean and standard deviation over time
. By plotting the time series
. By examining autocorrelation plots

The below plot is an example of what would be considered a White Noise Time Series:



Even though there are occasional spikes, there are no identifiable patterns visible, i.e., the distribution looks completely random.

Image Source

Random Walk

Random Walk time series, like White Noise, are unpredictable. However, in Random Walk, the values aren't simply a list of random numbers with no relation to their previous sequential data. In a Random Walk, the current value is dependent on the previous value along with some added noise.

In Random Walks, we can use the previous value for forecasting but this does not ensure results, because only one previous value does not guarantee the desired outcome, and the addition of white noise to the previous values makes it even more difficult. So, it is difficult to forecast Random Walks too, especially for a long period of time.

To make a dummy Random Walk series, follow the below steps:

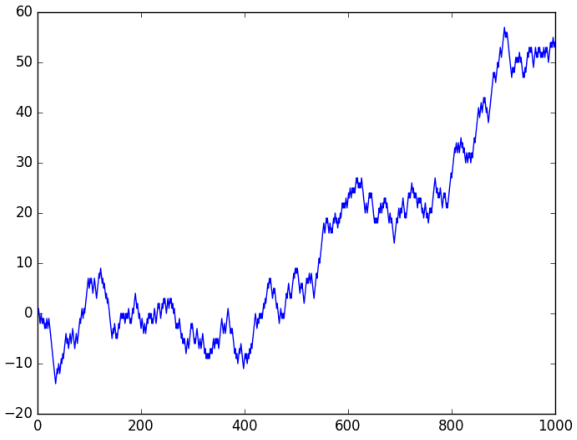
. Begin with an arbitrary value, such as zero.
. The next value is the previous value plus some added random fluctuation. You can go through the procedure of adding more values as many times as you like.

A Random Walk is mathematically defined as:

$$X_t = X_{t-1} + W_t$$

Where, X_t is the present value, X_{t-1} is the previous value, and W_t is the white noise addition

An example of a Random Walk plot is shown below:



In the above graph, the series seems to somehow maintain a trend but due to the addition of white noise at each stage, the trend is not very clean and appears to have spikes and troughs.

Image Source

Happy Learning!

Previous

Next