Bagadi, R. (2017). Finding The Next Term Of Any Time Series Type Sequence Using Total Similarity & Dissimilarity {Version 5}. *PHILICA.COM Article number 1172*. http://philica.com/display_article.php?article_id=1172

Finding The Next Term Of Any Time Series Type Sequence Using Total Similarity & Dissimilarity {Version 4}

ISSN 1751-3030

Authored By
Ramesh Chandra Bagadi
Affiliation 1:
Data Scientist
INSOFE (International School Of Engineering)
Gachibowli, Hyderabad, Telengana State, India
Affiliation 2:
Founder, Owner, Director & Advising Scientist In Principal
Ramesh Bagadi Consulting LLC (R420752), Madison, Wisconsin 53726
United States Of America

Email: rameshcbagadi@uwalumni.com
Telephone: +91 9440032711

Abstract

In this research investigation, the author has detailed a novel scheme of finding the next term of any given time series type sequence.

Theory

Given any Sequence of the Time Series kind, $S=\{y_1,y_2,y_{3,}.......,y_{n-1},y_n\}$ which represent some Time Series data of concern, we write a Truth Statement Equation as follows:

$$y_{n+1}=\frac{\left\{\overbrace{\sum_{i=1}^{n}\{Smaller(y_i,y_{n+1})\}}^{DirectSimilarity}\right\}+\left\{\overbrace{\sum_{i=1}^{n}\{L\arg er(y_i,y_{n+1})-Smaller(y_i,y_{n+1})\}}^{DirectDissimilarity}\right\}}{n}\qquad\text{Equation 1}$$

The above Equation cannot be solved for y_{n+1} but can be used to find y_{n+1} by guessing its value. For the correct guess, i.e., the true value of y_{n+1}, i.e., the next Term of the Sequence, the above Equation is satisfied, i.e., LHS=RHS.

One can note that this Grand Equation can be used to find the Next Prime as well, given a sequence of Primes from the beginning, while considering 1 as Prime as well, i.e., the beginning or first Prime. One can note the concepts of Similarity & Dissimilarity from author's [1]. The author calls $\sum_{i=1}^{n}\{Smaller(y_i,y_{n+1})\}$ as Direct Dissimilarity and $\sum_{i=1}^{n}\{L\arg er(y_i,y_{n+1})-Smaller(y_i,y_{n+1})\}$ as Direct Dissimilarity.

Bagadi, R. (2017). Finding The Next Term Of Any Time Series Type Sequence Using Total Similarity & Dissimilarity {Version 5}. *PHILICA.COM Article number 1172*. http://philica.com/display_article.php?article_id=1172

For Guessing, we can usually start with a Guess value much smaller than the smallest data value of the dataset and and keep increasing its value by very small increments till the value of the δ_j tends to zero within the limits of our computational ability to guess. The δ_j is given by

$$\delta_j = y_{n+1\,Gues\,j} - \frac{\left\{\overbrace{\sum_{i=1}^{n}\left\{Smaller\left(y_i, y_{n+1\,Gues\,j}\right)\right\}}^{Similarity}\right\} + \left\{\overbrace{\sum_{i=1}^{n}\left\{L\arg er\left(y_i, y_{n+1Gues\,j}\right) - Smaller\left(y_i, y_{n+1Gues\,j}\right)\right\}}^{Dissimilarity}\right\}}{n}$$

Equation 2

where y_{n+1} is the j^{th} Guess for y_{n+1}

Example

For the data given below

Annual GNP deflator, U.S., 1889 to 1970

Exported from datamarket.com	
Date exported	2017-11-22 07:10
View online	https://datamarket.com/data/set/22wo/annual-gnp-deflator-us-1889-to-1970#!ds=22wo&display=line
License	Unknown; please assume a restricted license (all rights reserved); contact DataMarket if you need different licensing
Provider	Time Series Data Library
Source URL	
Units	

	Annual GNP deflator, U.S., 1889 to 1970		Annual GNP deflator, U.S., 1889 to 1970		Annual GNP deflator, U.S., 1889 to 1970		Annual GNP deflator, U.S., 1889 to 1970
Year		Year		Year		Year	
1889	25.9	1913	31.1	1937	44.5	1961	104.6
1890	25.4	1914	31.4	1938	43.9	1962	105.8
1891	24.9	1915	32.5	1939	43.2	1963	107.2
1892	24	1916	36.5	1940	43.9	1964	108.8
1893	24.5	1917	45	1941	47.2	1965	110.9
1894	23	1918	52.6	1942	53	1966	113.9
1895	22.7	1919	53.8	1943	56.8	1967	117.6
1896	22.1	1920	61.3	1944	58.2	1968	122.3
1897	22.2	1921	52.2	1945	59.7	1969	128.2
1898	22.9	1922	49.5	1946	66.7	1970	135.3
1899	23.6	1923	50.7	1947	74.6		
1900	24.7	1924	50.1	1948	79.6		
1901	24.5	1925	51	1949	79.1		
1902	25.4	1926	51.2	1950	80.2		
1903	25.7	1927	50	1951	85.6		
1904	26	1928	50.4	1952	87.5		
1905	26.5	1929	50.6	1953	88.3		
1906	27.2	1930	49.3	1954	89.6		
1907	28.3	1931	44.8	1955	90.9		
1908	28.1	1932	40.2	1956	94		
1909	29.1	1933	39.3	1957	97.5		
1910	29.9	1934	42.2	1958	100		
1911	29.7	1935	42.6	1959	101.6		
1912	30.9	1936	42.7	1960	103.3		

The above stated authors algorithm predicted the 83rd data element (corresponding to the year 1970) correctly as 135.3 when the first 82 data elements (corresponding to the years 1889-1917) were used to predict the 83rd data element.

Furthermore, when the author used the first 74 data elements (corresponding to the years 1889-1961) to predict the 75th data element, the Prediction Error was zero. Similarly, this Accumulated Progressive Error of Prediction (for the next 10 steps) was Zero for the next 10 steps, i.e., until we predicted the last 83rd data element. By Accumulated Progressive Error (for One Step), we mean the Prediction Error obtained using the last Predicted data element to Predict the next data element. If we get a non-Zero Prediction Error in the beginning, the Accumulated Progressive Error of Prediction keeps increasing.

Bagadi, R. (2017). Finding The Next Term Of Any Time Series Type Sequence Using Total Similarity & Dissimilarity {Version 5}. *PHILICA.COM Article number 1172*. http://philica.com/display_article.php?article_id=1172

References

Bagadi, R. (2017). Total Intra Similarity And Dissimilarity Measure For The Values Taken By A Parameter Of Concern. {Version 2}. ISSN 1751-3030. *PHILICA.COM Article number 1153*.
http://www.philica.com/display_article.php?article_id=1153

http://vixra.org/author/ramesh_chandra_bagadi

http://philica.com/advancedsearch.php?author=12897