

Understanding Data Pre-processing using Sliding Window Method in Forecasting Tasks

Annie Roser*

Department of Biostatistics, Science and Technology of New York, New York, USA

Introduction

Normalization is the process of creating shifted and scaled versions of statistics with the goal of eliminating the effects of certain gross influences, like in an anomaly time series, by comparing the corresponding normalized values of different datasets (heterogenic data). As a result, the procedure for heterogeneous data transformation brings all attributes to the same scale. Indeed, the decimal scaling method is one of the quantitative data normalization techniques that moves the decimal point of the data's values. We divide each data value by the maximum absolute to normalize the data using this method. The original data are subjected to linear transformation in the minimum-maximum (Min-Max) data normalization method, whereas in the z-score data normalization procedure, values are normalized using the mean and standard deviation parameters. On the basis of these evidences, quantitative data standardization and normalization procedures may have distinct parametric distribution, such as the normal distribution, and data variability reduction capabilities [1].

Description

Quantitative data used for the present study were drawn from previous experiments as described. Briefly, collected data included four growth parameters (diameter, plant height, leaf length and leaf number) of two maize varieties, treated by both rhizobacteria and foliar bio-fertilizing. Further, collected data for each treatment were summarized in a matrix including four columns describing variables parameters (two maize varieties growth parameters) and ninety-six rows corresponding to the observation number. Next, we submitted the above-mentioned data matrix to Box-Cox, Logarithm, Square Root, Inverse and Z-score, Minimum, Exponential and Minimum-Maximum quantitative data standardization as well as normalization (data transformation) procedures. Biometric verification is a method for checking a person's personality by using a piece of their identity, like their finger impression, facial features, or iris design. These features contain unique information that can't be duplicated. Despite their numerous benefits, certain biometrics, particularly facial recognition, have recently come under fire for being an infringement on privacy. Considering everything, your "face print" is your information, and many people don't like the idea that their face prints could be used or shared without their consent. This may eliminate the obscurity that many people anticipate in open areas, such as online. Even the idea of "connecting" a person's face to yet another source of personal data has been floated [2].

As the COVID-19 pandemic evolves, many pediatricians and infants' parents have been left with renewed questions about the consequences of infection on children and steps to be taken if their child has symptoms of, or tests positive for, COVID-19. Methods & Results: Literature reviews and recent studies revealed that children are better than adults at combating SARS – CoV 2. There was conflicting evidence on age-related differences in ACE2 expression in the nose and lungs. However, measurements of SARS-CoV-2 viral load have shown no

*Address for Correspondence: Annie Roser, Department of Biostatistics, Science and Technology of New York, New York, USA, E-mail: roser89@edu.in

Copyright: © 2023 Roser A. This is an open-access article distributed under the terms of the creative commons attribution license which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Received: 02 January, 2023, Manuscript No. Vcrh-23-94787; Editor assigned: 03 January, 2023, Pre QC No. P-94787; Reviewed: 16 January, 2023, QC No. Q-94787; Revised: 21 January, 2023, Manuscript No. R-94787; Published: 28 January, 2023, DOI: 10.37421/2736-657X.2023.7.173

clear difference. Strikingly, cross-reactive antibodies from previous exposure to coronavirus common cold do not offer any special protection in both children and adults. The kid's immune response against SARS CoV-2 infection is initiated with low immunological tone to prevent overactive immunity and is characterized by rapid lung damage repair in contrast to stormy waves in adults. Omicron variant surge threatened to change the Kid's immune story, as more children were hospitalized, however, even with more Omicron cases, severe illness among infected kids remains extremely rare. Conclusion: One of the few silver linings of the COVID-19 pandemic is that children are relatively spared and winning with their immunity [3-5].

Conclusion

We focused on eight quantitative data transformation systems in the present comparative study. Processed quantitative data standardization and/or normalization procedures are as following Box-Cox (Box), Exponential (Expo), Inverse, Logarithmic normalization, Maximum, Minimum-Maximum, Square Root and Z-score. Above-mentioned data transformation systems was applied to the same data matrix (collected data) generating a new data set for each standardization and/or normalization methods. The present study provided a systematic comparative study that highlighted difference as well as similitude between eight quantitative data standardization methodologies providing useful tool to researchers, in choosing adequately data transformation methodologies that well fitting for their investigations. The same survey displayed smaller bias transformation by using the Box-Cox transformation as opposite to logarithm transformation. The same study revealed that the mean squared error of estimation is smaller with the Box-Cox transformation; and as well, the Box-Cox transformation leads to systematically higher estimated values than Logarithmic transformation. Hence, the Box-Cox transformation should be considered as a viable alternative in statistical modelling if the transformation of variables is required. Low aptitude with regard Exponential and Inverse data transformation in reducing data variability as well as in adjusting data normality could be due to processed positive value of analysed data. Indeed, our analysis suspected Exponential data transformation as a potential source of transformed data variability.

Acknowledgement

We thank the anonymous reviewers for their constructive criticisms of the manuscript. The support from ROMA (Research Optimization and recovery in the Manufacturing industry), of the Research Council of Norway is highly appreciated by the authors.

Conflict of Interest

The authors declare that there was no conflict of interest in the present study.

References

1. Chan, Chi N., Benjamin Trinité and David N. Levy. "Potent inhibition of HIV-1 replication in resting CD4 T cells by resveratrol and pterostilbene." *Antimicrob Agents Chemother* 61 (2017): e00408-17.
2. Gonzalez, Hugo, Catharina Hagerling and Zena Werb. "Roles of the immune system in cancer: From tumor initiation to metastatic progression." *Genes Dev* 32 (2018): 1267-1284.

3. Woller, Norman, Engin Gurlevik, Bettina Fleischmann-Mundt and Anja Schumacher. "Viral infection of tumors overcomes resistance to PD-1-immunotherapy by broadening neoantigenome-directed T-cell responses." *Mol Ther* 23 (2015): 1630-1640.
4. Robert, Caroline. "A decade of immune-checkpoint inhibitors in cancer therapy." *Nat Commun* 11 (2020): 1-3.
5. Mondal, Moumita, Jingao Guo, Ping He and Dongming Zhou. "Recent advances of oncolytic virus in cancer therapy." *Hum Vaccin Immunother* 16 (2020): 2389-2402.

How to cite this article: Roser, Annie. "Understanding Data Pre-processing using Sliding Window Method in Forecasting Tasks." *Virol Curr Res* 7 (2023): 173.