Joint Conference on

2nd International Conference of

AI and Data Science

October 26-27, 2022, Dubai, UAE

UNITED RESEARCH FORUM
(CONNECT WITH RESEARCH WORLD)

Deep Learning and Fairmindedness towards Missing values

Viji Venugopal Kochath

Data Scientist/ Researcher R. K. University, India.

Abstract

Machine Learning (ML) and Deep Learning (DL) based predictions mainly depend on the data collected and fed to the algorithms. Missing data can happen due to several reasons such as incomplete responses, ambiguous response choices, dropouts in longitudinal data, etc. which affect the development of robust ML algorithms and sometimes contribute to a black box in DL algorithms. This paper discusses the impact of missing values and demonstrates a novel imputation technique implemented on Google's state-of-the-art transformer named Temporal Fusion Transformer (TFT). This paper compares the output of the original TFT algorithm where missing data is imputed using LOCF (Last Observation Carried Forward) imputation with novel imputation techniques where data is imputed based on its type such as dependent or independent variables, inter-related patterns among the missing values compared to non-missing values based on Selected Mean Value Imputation technique. This paper also showcases the increase in the learning capacity of the network by using Variable Selection Networks (VSN) by stacking multiple GRN layers on top of the VSN layer. The result of imitated dataset shows that the proposed imputation methods with the VSN algorithm have a noticeably better performance than conventional methods. Besides, they can also handle extreme column-wise missing data cases which are not addressed in TFT.

Biography

Viji is a data science enthusiast having 11 years of experience in Data Analysis, worked across all aspects of data from engineering to building sophisticated visualizations, machine learning models and experiments. Currently working as a faculty with GEMS education and as a researcher at Rise Hydroponics India to optimize crop growth using Machine Learning and Deep Learning algorithms to build sustainable food security for the future. She has also proved her skills as a data engineer at Union Bank of India and was appointed as a data analyst at Central Salt and Marine Central Research Institute (CSMCRI), one of the top performing national R&D laboratories in India which are based on the salt and marine chemicals analysis, water desalination and purification, renewable energy, salinity tolerance and waste management. While working with an educational institution in UAE she led a Machine learning project about "21st Century Transportation" for Think Science UAE (2018) event. She is currently pursuing PhD in Deep Learning and previously completed her degree & master's degree in Computer Applications (2012) and has done a diploma in Machine Learning Specialization and Deep Learning from IIT Roorkee (2020). Her current research interest is Artificial neural networks, backpropagation algorithms in deep learning, and exploratory data analysis. Etc.