

Time Series Regression with Meta-Clusters

Authors : Monika Chuchro

Abstract : This paper presents a preliminary attempt to apply classification of time series using meta-clusters in order to improve the quality of regression models. In this case, clustering was performed as a method to obtain a subgroups of time series data with normal distribution from inflow into waste water treatment plant data which Composed of several groups differing by mean value. Two simple algorithms: K-mean and EM were chosen as a clustering method. The rand index was used to measure the similarity. After simple meta-clustering, regression model was performed for each subgroups. The final model was a sum of subgroups models. The quality of obtained model was compared with the regression model made using the same explanatory variables but with no clustering of data. Results were compared by determination coefficient (R^2), measure of prediction accuracy mean absolute percentage error (MAPE) and comparison on linear chart. Preliminary results allows to foresee the potential of the presented technique.

Keywords : clustering, data analysis, data mining, predictive models

Conference Title : ICMSCS 2014 : International Conference on Mathematical, Statistical and Computational Sciences

Conference Location : Paris, France

Conference Dates : December 30-31, 2014