

A Time Series Classification Dataset Based on the Average Price of Concrete in major Cities in China

Qing Liu¹, Woon-Seek Lee², and Min Du¹

¹Huainan Normal University

²Pukyong National University

March 17, 2022

Abstract

Time series classification (TSC) is an important and challenging problem in data mining. Time series data sets are an important basis for this research and are widely used in baseline verification of various algorithm models. Aiming at the problem that there are few domestic data sets and the current TSC data set is relatively old, a new data set for TSC task is established based on the average price data of concrete in major cities in China, which provides new data support for the research of TSC algorithm. We made use of the data center of Oriental Fortune to disclose the sample data of the average price of concrete from 2013-10-23 to 2021-01-20, created 730 autoregression-based series data sets by using sliding windows of different lengths, and then selected the appropriate sliding window length through machine learning model verification, finally, convolutional neural network (CNN) and long and short memory (LSTM) network, which are good at processing temporal features, are used to verify the data and prove the validity of the dataset. The dataset is freely available at <https://gitee.com/lq2012/tsc-dataset>.

Hosted file

A Time Series Classification Dataset Based on the Average Price of Concrete in major Cities in China - available at <https://authorea.com/users/465724/articles/560218-a-time-series-classification-dataset-based-on-the-average-price-of-concrete-in-major-cities-in-china>