Received: 14 March 2022 | Revised: 21 April 2022 | Accepted: 8 May 2022

DOI: 10.1002/eng2.12541

SHORT COMMUNICATION

Engineering Reports Open Access WILEY

A set of time-series classification datasets based on the average price of concrete in major Chinese cities

Qing Liu[1,2] | Minghao Huang[2] | Woon-Seek Lee[2] | Yamin Du[1]

[1] School of Economics and Management, Huainan Normal University, Huainan City, Anhui Province, China

[2] Graduate School of Management of Technology, Pukyong National University, Busan, Republic of korea

Correspondence
Woon-Seek Lee and Minghao Huang, 365, Sinseon_ro, Nam-gu, Busan, 48547, Republic of korea.
Email: iewslee@pknu.ac.kr and mhuang@pknu.ac.kr

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 set of time-series classification datasets are 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 October 23, 2013 to January 20, 2021, created 1093 autoregression-based series data sets by using sliding windows of different lengths, and then through the experimental verification of linear regression model, decision tree model, and random forest model, a set of data sets with lengths of 170, 842, and 1052 were selected. We use three convolutional neural network (CNN) models and three long short-term memory (LSTM) network models to verify the validity of the data, the CNN's obtained the best accuracy rate of 93.20%, and the LSTM networks obtained the best accuracy rate of 92.99%. The establishment of the new data set has certain significance for the research of TSC and provides some references for other researchers to create datasets. The datasets are freely available at https://gitee.com/lq2012/tsc-dataset.

KEYWORDS

average value of concrete, dataset, time series classification

JEL CLASSIFICATION

Computer and software engineering

1 | INTRODUCTION

Any type of data acquisition with a certain ordering regularity will generate time series, which makes this type of data very common in data mining problems.[1] There are many ways to apply time-series data. Reference 2 reviewed in-depth three important applications of learning to time series data:

Classification: Commonly used in supervised learning, Reference 3 studies a time series classification algorithm based on fully convolutional networks.

Prediction: Based on the fluctuation rule of historical time series data, the trend of future fluctuation is predicted. Reference 4 considers the correlation between space and time and proposes a new traffic flow prediction method based on deep learning.

Anomaly detection: Reference 5 applies the unsupervised feature learning architecture of deep belief network (DBN) to anomaly detection of sleep data.

The three types of applications summarized in Reference 6 often overlap and have similarities. For example, anomaly detection can be converted into dichotomies through supervised learning to mark whether the heart is abnormal or not, while the stock prediction problem is also a three-classification problem with price rising, falling, or unchanged in many cases. Therefore, time series classification (TSC) is an important and challenging problem in data mining,[7] hundreds of TSC algorithms have been proposed since 2015,[8] and all of these algorithms need to use data sets for validity verification. Time series classification data sets are an important basis to support TSC research.

Reference 3 proposes a robust time series classification baseline model, which is validated on 44 UCR (UCR time series classification archive) datasets. Reference 9 proposed a TSC network model based on LSTM and FCNN and verified it on almost all UCR datasets. Reference 10 studied transfer learning for TSC and verified the performance of the research results on 85 datasets of UCR. Reference 11 studied a binary distribution tree-based method for time series classification, using the UCR dataset to compare the performance with 5 other powerful baseline methods. Reference 12 provides a systematic review of a large number of new time series classification algorithms proposed in the literature over the past 5 years, all of which use the UCR archived dataset for performance evaluation. It is not difficult to see that the time series data set is the cornerstone of the research on the TSC algorithm, and the continuous update and enrichment of the data set is of great significance to the research of time series classification.

The last major update of the world's largest open-source time series classification archive website (UCR time series classification archive) was completed in the fall of 2018, after which the data set was stopped due to a lack of funding support, which makes the current datasets used for TSC algorithm benchmarking relatively old, it is extremely unfavorable for TSC research. This article collects the average price data of concrete (C20) in China from October 23, 2013 to January 20, 2021 from Oriental Fortune.com and is used to construct 1093 time series classification datasets with different feature lengths, and then a set of datasets with excellent performance was selected by the machine learning algorithm. We use convolutional neural network (CNN), which is good at mining data spatial features, and long short-term memory (LSTM) network, which is good at extracting data time series features, to verify the effectiveness of the dataset. The results show that the emergence of new datasets has implications for the research of TSC algorithm.

Section 2 of this article introduces the process of creating time series classification datasets, Section 3 introduces the selection of time series feature lengths, Section 4 uses CNN and LSTM networks to verify the validity of the dataset, and Section 5 is the conclusion.

2 | DATASET CREATION

The research purpose of this article is to use the average price data of concrete in major cities in China to create a set of univariate autoregressive series for the baseline test of the TSC task model and algorithm. This section mainly introduces the construction process of the dataset.

2.1 | Related concepts

References 13,14 divide time series into univariate time series (UTS) and multivariate time series (MTS), referring to the description of time series in References 13,15, UTS and MTS are defined as follows

Definition 1. Univariate time series (UTS) $\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_T}]$ is a group of ordered actual observed values, the length of X is T, and $\mathbf{x_i} \in \mathrm{R}$.

Definition 2. Multivariable time series (MTS) $\mathbf{X} = [\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_T}]$ contains M different univariate time series, and $\boldsymbol{X^i} \in \mathrm{R}^T$.

Definition 3. Dataset $\mathbf{D} = \{(\mathbf{X_1Y_1}), (\mathbf{X_2Y_2}), \ldots, (\mathbf{X_NY_N})\}$ is the set of time series $(\boldsymbol{X_i}, \boldsymbol{Y_i})$, $\boldsymbol{X_i}$ is the feature data of time series, and $\boldsymbol{Y_i}$ is the one-hot label vector corresponding to $\boldsymbol{X_i}$ or the natural number label starting from 0.

Autoregressive sequence (ARS) is a special kind of time series, as the name suggests, it does not use feature data X to predict classification label y, but uses X to predict X itself, which is a regression of the variable itself, the basic logic of autoregressive sequence can be described by the following formula:

$$\boldsymbol{X_{t+d}} = \boldsymbol{f}(\boldsymbol{X_t}, \boldsymbol{X_{t-1}}, \cdots)\ \boldsymbol{d} \in (\mathbf{1{,}2{,}3}, \cdots). \tag{1}$$

Reference 16 believes that in many cases, the relationship between past and future observations is not deterministic, which is equivalent to expressing conditional probability distribution as a function of past observations, and the formula is described as follows

$$\boldsymbol{p}\,(\boldsymbol{X_{t+d}} | \boldsymbol{X_t}, \boldsymbol{X_{t-1}}, \cdots) = \boldsymbol{f}(\boldsymbol{X_t}, \boldsymbol{X_{t-1}}, \cdots)\ \boldsymbol{d} \in (\mathbf{1{,}2{,}3}, \cdots). \tag{2}$$

2.2 | Dataset Description

To facilitate the construction of the dataset, it is necessary to systematically describe the construction process of the concrete autoregressive sequence.

Let P_d^T be the price of concrete on the $\boldsymbol{d}$ day in the time period of consecutive T days, and $\boldsymbol{x_T^d}$ be the price change value of the concrete on the d day in the time period of consecutive T days, that is

$$\boldsymbol{x_d^T} = \boldsymbol{P_d^T} - \boldsymbol{P_{d-1}^T}, \boldsymbol{P_d^T} = \boldsymbol{x_d^T} + \boldsymbol{P_{d-1}^T}, \boldsymbol{d} = \mathbf{1{,}2{,}3}, \cdots, \boldsymbol{T}, \tag{3}$$

then

$$\boldsymbol{x_0^T} = \mathbf{0}, \boldsymbol{X^T} := \left\{ \boldsymbol{x_d^T} \,\middle|\, \boldsymbol{d} = \mathbf{0{,}1{,}2}, \cdots, \boldsymbol{T} - \mathbf{1} \right\} \tag{4}$$

is a sequence of concrete price changes for T consecutive days.

Let P_d^T be the price of concrete on the $\boldsymbol{d}$ day in the time period of consecutive T days, and $\Delta 10 x_d^T$ be is the change value of the concrete price on the d day relative to the $\boldsymbol{d} - \mathbf{10}$ day in the period of T consecutive days, then there are:

$$\Delta 10 x_T^T = \boldsymbol{P_T^T} - \boldsymbol{P_{T-10}^T}, \boldsymbol{P_T^T} = \Delta 10 x_T^T + \boldsymbol{P_{T-10}^T}, \tag{5}$$

at this time, the label data y^T indicating the direction of the price change of the concrete price on the T-day relative to the price before $\boldsymbol{T} - \mathbf{10}$ day is:

$$\mathbf{y}^{\boldsymbol{T}} = \begin{cases} \mathbf{0}\ \left(\Delta 10 x_T^T = \mathbf{0}\right), \\ \mathbf{1}\ \left(\Delta 10 x_T^T > \mathbf{0}\right), \\ \mathbf{2}\ \left(\Delta 10 x_T^T < \mathbf{0}\right), \end{cases} \tag{6}$$

and the supervised learning sample for concrete price forecast 10 days ago can be described as $(\boldsymbol{X^T}, \mathbf{y}^{\boldsymbol{T}})$,The concrete price change dataset with feature length T and sample number N can be described as:

$$\boldsymbol{D_N^T} := \left\{ \left(\boldsymbol{X_n^T}, \boldsymbol{y_n^T}\right) \,\middle|\, \boldsymbol{n} = \mathbf{0{,}1}, \cdots \boldsymbol{N} - \mathbf{1}, \boldsymbol{y_n^T} \in \boldsymbol{C} := \{\mathbf{0{,}1{,}2}\} \right\}, \tag{7}$$

where $\mathbf{y}^{\boldsymbol{T}} = 0$ means the price remains unchanged, $\mathbf{y}^{\boldsymbol{T}} = 1$ means the price has risen, and $\mathbf{y}^{\boldsymbol{T}} = 2$ means the price has fallen.

2.3 | Data acquisition and processing

The original data set is collected from the average price data of concrete (C20) in major Chinese cities from October 23, 2013 to January 20, 2021 collected by Oriental Fortune network (https://www.eastmoney.com), and the site did not disclose a specific list of the "major Chinese cities" mentioned in the data's title. We use crawler technology to obtain the original data from the website. The valid fields only include the date and price fields, and there are 1789 original data samples in total. The original data style is shown in Table 1.

Table 2 summarizes the monthly distribution of the 1789 original sample data between October 2013 and January 2021. It is not difficult to see that the number of samples in each month is missing, and the original data is not continuous, to make the original data complete and continuous, it is necessary to fill in the vacant data.

Assuming that in a continuous period of T days, if the sample with the statistical dated has a sample with the statistical dated-1, the sample on the date d is called a continuous sample, otherwise, the sample d is not continuous, to choose a suitable blank data filling method, we made basic statistics on the price changes of continuous data in the original sample, and the price changes were defined by formula (3),

$$\boldsymbol{x_d^T = P_d^T - P_{d-1}^T, P_d^T = x_d^T + P_{d-1}^T, d = 1, 2, 3, \cdots, T}.$$

The price trend statistics of the original data samples are shown in Table 3

The data with constant prices in the continuous sample accounts for 65% of all continuous samples, so we have reason to believe that the main reason for the vacancy in data is that the price has not changed. Therefore, if the data on date d

TABLE 1 The raw data of the website

Date	Price
October 25, 2013	294.1
October 28, 2013	294.3

TABLE 2 Statistical table of sample size distribution

	01	02	03	04	05	06	07	08	09	10	11	12
2013										7	21	22
2014	22	17	21	21	21	20	23	21	22	19	20	22
2015	23	17	22	21	10	21	23	21	21	18	21	23
2016	19	18	23	20	21	21	21	23	21	18	22	21
2017	19	19	23	19	21	22	21	23	22	17	22	21
2018	22	17	22	20	22	20	18	21	21	18	22	20
2019	22	15	21	22	21	19	23	22	21	19	21	22
2020	16	20	22	22	19	21	23	21	23	17	21	23
2021	13											

TABLE 3 Statistics of price trends on continuous samples

$\boldsymbol{x_d^T}$	Count	Rate
$x_d^T = 0$	923	65%
$x_d^T > 0$	281	20%
$x_d^T < 0$	210	15%
Total	1414	100%

date	price
2013-11-02	295.9
2013-11-04	296.5
2013-11-05	298.1
2013-11-06	298.0

date	price
2013-11-02	295.9
2013-11-03	**295.9**
2013-11-04	296.5
2013-11-05	298.1
2013-11-06	298.0

FIGURE 1 Schematic diagram of discontinuous sample completion

is vacant, the price data on day d − 1 is used to complete it, the process is shown in Figure 1. After completing the data, there are 2647 data samples between October 23, 2013 and January 20, 2021.

On the basis of complementing continuous samples, the time series classification dataset can be constructed according to formula (7)

$$\boldsymbol{D}_{\boldsymbol{N}}^{\boldsymbol{T}} := \left\{ \left(\boldsymbol{X}_{\boldsymbol{n}}^{\boldsymbol{T}}, \boldsymbol{y}_{\boldsymbol{n}}^{\boldsymbol{T}} \right) \middle| \ \boldsymbol{n} = \boldsymbol{0}, \boldsymbol{1}, \cdots \boldsymbol{N} - \boldsymbol{1}, \boldsymbol{y}_{\boldsymbol{n}}^{\boldsymbol{T}} \in \boldsymbol{C} := \{\boldsymbol{0}, \boldsymbol{1}, \boldsymbol{2}\} \right\}.$$

For a specific data set $\boldsymbol{D}_{\boldsymbol{N}}^{\boldsymbol{T}}$, the number of time series samples N constructed will change with the change of the time series length T, the variable N_s is used to represent the number of data samples after completion, and the variable *Ahead* is used to represent the predicted days in advance, then N can be calculated as follows:

$$N = \boldsymbol{N_s} - \boldsymbol{T} - Ahead + 1, \boldsymbol{T} = \boldsymbol{2}, \boldsymbol{3}, \cdots \boldsymbol{N_s} - \boldsymbol{1}.$$

3 | SELECTION OF FEATURE LENGTH

For many mature application scenarios, the choice of feature-length is a question with mature answers, such as a heartbeat fluctuation dataset that detects whether the heartbeat is abnormal, while for a brand-new application scenario when creating a time series dataset, the selection of feature-length is an issue to be investigated.

Considering the general market law, we believe that the price fluctuation data 3 years ago has no reference value for the current price forecast. Therefore, we created 1093 datasets with a minimum feature-length of 2 and a maximum feature-length of 1094 (365 × 3 − 1) in a traversal manner as the candidate datasets, the collection of these datasets can be described as

$$\boldsymbol{D} := \left\{ \boldsymbol{D}^{\boldsymbol{T}} \middle| \ \boldsymbol{T} = \boldsymbol{2,3,4}, \cdots \boldsymbol{1094} \right\}.$$

For the TSC task, accuracy is undoubtedly the most important statistical indicator. The definition of accuracy is the ratio of the number of correctly predicted samples to all predicted samples, and the formula is described as follows:

$$\textbf{Accuracy} = \frac{\boldsymbol{I}(\boldsymbol{f}(\boldsymbol{X}) = \boldsymbol{Y})}{\boldsymbol{I}(\boldsymbol{f}(\boldsymbol{X}) = \boldsymbol{Y}) + \boldsymbol{I}(\boldsymbol{f}(\boldsymbol{X}) \neq \boldsymbol{Y})}.$$

In the formula, $I(\cdot)$ is the indicator function. Therefore, we use the accuracy as the evaluation index, and use the linear regression model, the decision tree model and the random forest model with the best performance in the classification field to test on 1093 candidate data sets. The function $\boldsymbol{f_M}(\boldsymbol{T})$ represents the prediction accuracy of the model M with the given parameters as a function of the time series length T, Figure 2 shows the fitted curve of $\boldsymbol{f_M}(\boldsymbol{T})$ when $T \in I_n := \{2, 3, \cdots 1094\}$, the horizontal axis of the curve is the feature-length, and the vertical axis is the accuracy.

As shown in Figure 2, the highest classification accuracy of the three machine learning models in 730 data is 0.912, random forest model performs the best, decision tree model is the second, and linear regression model is the worst. For decision tree model and random forest model, when the long march length is greater than 100, the $\boldsymbol{f_M}(\boldsymbol{T})$ function curve tends to converge, so if a decision tree model or a random forest model is used to make predictions on this dataset, any

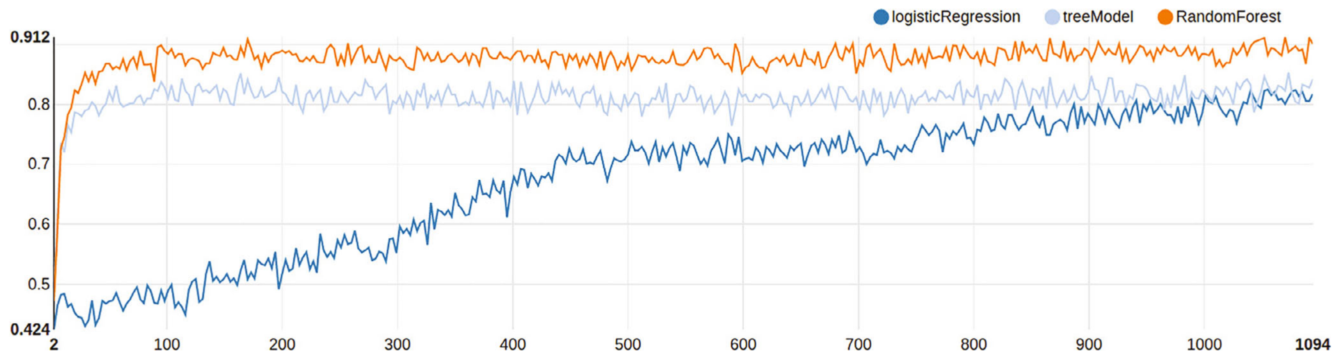

FIGURE 2 Machine learning classification accuracy curve under different length time series

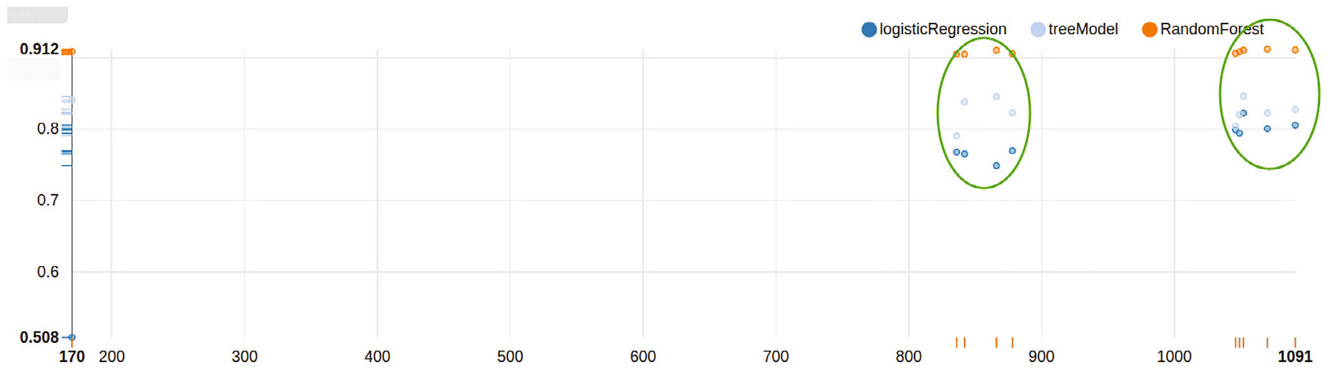

FIGURE 3 Top10 feature-length scatter plot

length of the dataset feature greater than 100 is appropriate, choosing a larger feature length may give the model more room to play, and it will also increase the consumption of computing power. For the linear regression model, when feature length is greater than 500, the $\boldsymbol{f_M(T)}$ function tends to be stable, but it is still slowly improving, so a larger feature length is needed to improve the prediction performance, but this is obviously not a good choice.

To further observe the performance of the three models on these datasets, we screened out the top 10 data with the highest prediction accuracy and presented them in the form of a scatter graph to observe the distribution of data. As shown in Figure 3, since [170] is free, the optimal feature lengths are concentrated in the two intervals of [800–900] and [1000–1091].

The following work is to choose one or more "appropriate" feature lengths between [170], [800–900], and [1000–1091], we use "appropriate" rather than optimal because, for different models and different model parameters, this optimal value is different. Table 4 presents the above top10 data in tabular form, it is not difficult to see that there is only a slight difference between the prediction results of the [800–900] and [1000–1091] intervals, so we tend to choose a dataset in each interval to provide researchers with more choices. We selected three data sets of different lengths, $\boldsymbol{D^{170}}$, $\boldsymbol{D^{842}}$, and $\boldsymbol{D^{1052}}$ based on the more comprehensive evaluation data of average accuracy. At this time, on the three algorithm models, they obtained average accuracies of 75.3%, 83.6%, and 86.6%, respectively.

4 | VALIDITY VERIFICATION

CNN is good at mining the spatial features of data, and LSTM network, as the representative of recurrent neural networks (RNN), is very good at mining the time-series features of data, which makes CNN and LSTM widely used in all kinds of TSC tasks. Currently, almost all research on TSC algorithms is based on extensions of these 2 nets,[3,9,10,12] which is why we use them to verify the validity of the dataset. Since this is a confirmatory experiment, only a very simple network structure is used in this article.

TABLE 4 Best accuracy top 10 statistics

Feature length	Logistic regression	Decision tree	Random forest	Average
170	0.508	0.841	0.909	0.753
836	0.768	0.791	0.905	0.821
842	0.765	0.838	0.905	0.836
866	0.749	0.846	0.911	0.835
878	0.770	0.823	0.906	0.833
1046	0.799	0.804	0.906	0.836
1049	0.794	0.820	0.909	0.841
1052	0.822	0.846	0.911	0.860
1070	0.801	0.823	0.912	0.845
1091	0.806	0.828	0.911	0.848

4.1 | Validation with convolutional neural networks

We designed three simple convolutional neural networks for the usability verification of the datasets, the specific parameters of the network are such as shown in Table 5.

Using the above networks, 20 rounds of training were performed on the training set of the datasets

$$D := \left\{D^{170}, D^{842}, D^{1052}\right\}$$

TABLE 5 Architectures of CNN

CNN1	Layer	Operate	Input	Step	Padding	Output
	1	Conv 1D_RELU_1	(None, T,1)	1	Same	(None, T,32)
	2	Conv 1D_RELU_2	(None, T,32)	1	Same	(None, T,32)
	3	MaxPooling1D	(None, T,32)			(None,13,32)
	4	Conv 1D_RELU_3	(None,13,32)	1	Same	(None,13,32)
	5	GlobalAveragePooling1D	(None,13,32)			(None,32)
	6	Dropout	(None,32)			(None,32)
	7	Dense	(None,32)			3
CNN2	**Layer**	**Operate**	**input**	**Step**	**Padding**	**Output**
	1	Conv 1D_RELU_1	(None, T,1)	1	Same	(None, T,64)
	2	MaxPooling1D	(None, T,64)			(None,13,64)
	3	Conv 1D_RELU_2	(None,13,64)	2	Same	(None,13,64)
	4	GlobalAveragePooling1D	(None,13,64)			(None,64)
	5	Dropout	(None,64)			(None,64)
	6	Dense	(None,64)			3
CNN3	**Layer**	**Operate**	**Input**	**step**	**Padding**	**Output**
	1	Conv 1D_RELU_1	(None, T,1)	2	Same	(None, T,32)
	2	Conv 1D_RELU_2	(None, T,32)	1	Same	(None, T,32)
	3	MaxPooling1D	(None, T,32)			(None,13,32)
	4	Dense	(None,13,32)			(None,100)
	5	Dense	(None,100)			3

and then the model was tested on the test set. The prediction results are shown in Table 6.

Among the three models, CNN2 obtained the best prediction accuracy rate of 93.29% in the dataset 93.29%, and the lowest accuracy was obtained by the CNN2 model on the dataset $\boldsymbol{D}^{842}$. It is not difficult to see that three datasets can be used to evaluate the performance of the CNN model on the TSC task.

4.2 | Verification of long short-term memory networks

As with the validation of the datasets by the convolutional neural networks, this article designs three simple long-short memory networks for the validation of the datasets. The network architectures of LSTM are shown in Figure 4.

Using the above networks, 20 rounds of training were performed on the training set of the datasets

$$\boldsymbol{D} := \left\{\boldsymbol{D}^{170}, \boldsymbol{D}^{842}, \boldsymbol{D}^{1052}\right\}$$

and then the model was tested on the test set. The prediction results are shown in Table 7.

TABLE 6 Prediction results of CNN

T	CNN1				CNN2				CNN3			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score
170	0.8729	0.8760	0.8733	0.8736	0.9116	0.8989	0.9336	0.9126	0.9116	0.9054	0.9100	0.9073
842	0.8871	0.8970	0.8878	0.8892	0.8598	0.8502	0.9038	0.8641	0.8780	0.8936	0.8558	0.8663
1052	0.9076	0.9177	0.9076	0.9103	0.9329	0.9195	0.9490	0.9323	0.8659	0.8482	0.8779	0.8601

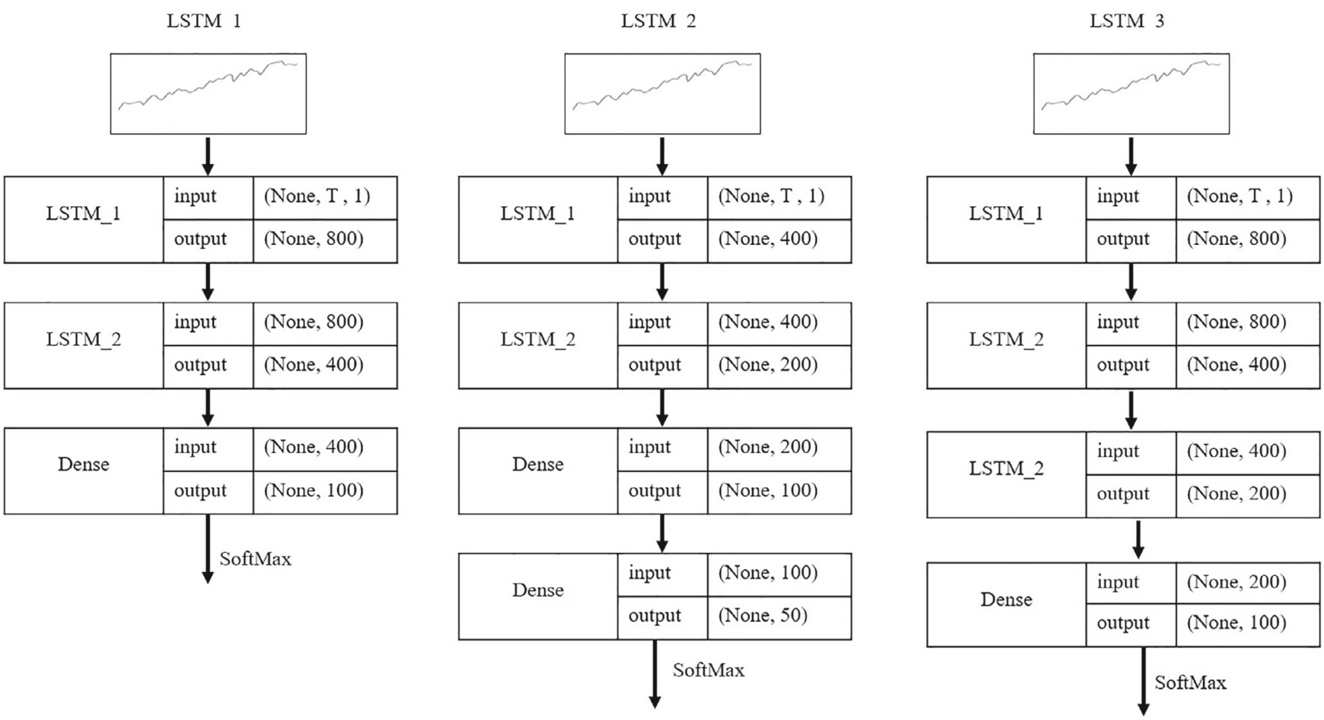

FIGURE 4 LSTM network architecture

TABLE 7 Prediction results of LSTM

T	LSTM1					LSTM2				LSTM3			
	Accuracy	Precision		Recall	F1 score	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score
170	0.8171	0.8977		0.7481	0.7897	0.8567	0.9051	0.8093	0.8427	0.9085	0.9202	0.8899	0.9034
842	0.9085	0.8897		0.9339	0.9071	0.8750	0.9054	0.8465	0.8639	0.9299	0.9265	0.9315	0.9284
1052	0.7866	0.8376		0.7171	0.7525	0.9177	0.9112	0.9182	0.9127	0.9177	0.9143	0.9131	0.9122

Among the three models, LSMT3 obtained the best prediction accuracy of 92.99% in the dataset $\boldsymbol{D}^{\mathbf{842}}$, the lowest accuracy was obtained for the LSTM1 model at D^{1052}. The experimental results show that there is a close correlation between the feature vector and the label vector of our time-series datasets.

The time series classification datasets we created based on the average price of concrete in China can be used for baseline testing of TSC models and algorithms.

5 | CONCLUSION

In this article, 1093 time series datasets with different feature lengths are constructed by using the average price data of concrete in major cities in China, and then we selected a set of "appropriate" data sets from 1093 data sets through prediction experiments of multiple machine learning algorithms: $\boldsymbol{D} := \left\{\boldsymbol{D}^{\mathbf{170}}, \boldsymbol{D}^{\mathbf{842}}, \boldsymbol{D}^{\mathbf{1052}}\right\}$, and verify the usability of this dataset using CNN and LSTM networks widely adopted in time series classification tasks. Experiments show that our time series classification dataset can be used for baseline testing of TSC models and algorithms, which is beneficial to make up for the outdated problem of current TSC datasets.

During the research, we found that the collection and open-source of datasets is a meaningful thing. In the future, we will publish more datasets to provide a research basis for algorithm researchers.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/eng2.12541.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in gitee at https://gitee.com/lq2012/tsc-dataset.

CONFLICT OF INTEREST

Authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS

Qing Liu: Data curation (equal); investigation (equal); writing – original draft (equal). **Minghao Huang:** Project administration (equal); resources (equal); validation (equal). **Woon-Seek Lee:** Project administration (equal); resources (equal); supervision (equal); validation (equal). **Yamin Du:** Project administration (equal); visualization (equal).

ORCID

Qing Liu https://orcid.org/0000-0001-5398-8006

REFERENCES

1. Längkvist M, Karlsson L, Loutfi A. A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognit Lett*. 2014;42:11-24.
2. Gamboa JC. Deep learning for time-series analysis. ArXiv Abs/170101887, 2017.

3. Wang Z, Yan W, Oates T. Time series classification from scratch with deep neural networks: a strong baseline. Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN); 2017:1578-1585.
4. Lv Y, Duan Y, Kang W, Li Z, Wang FY. Traffic flow prediction with big data: a deep learning approach. *IEEE Trans Intell Transp Syst*. 2015;16:865-873.
5. Längkvist M, Karlsson L, Loutfi A. Sleep stage classification using unsupervised feature learning. *Adv Artif Neural Syst*. 2012;2012:107046:1-107046:9.
6. Manjunath YS, Zhao S, Zhang XP. Time-distributed feature learning in network traffic classification for internet of things. ArXiv abs/210914696, 2021.
7. Esling P, Agón C. Time-series data mining. *ACM Comput. Surv*. 2012;45:12:1-12:34.
8. Bagnall A, Lines J, Bostrom A, Large J, Keogh E. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining Knowl Discov*. 2016;31:606-660.
9. Karim F, Majumdar S, Darabi H, Chen S. LSTM fully convolutional networks for time series classification. *IEEE Access*. 2018;6:1662-1669.
10. Ismail Fawaz H, Forestier G, Weber J, Idoumghar L, Muller P. Transfer learning for time series classification. Proceedings of the 2018 IEEE International Conference on Big Data (Big Data); 2018:1367-1376.
11. Ma C, Shi X, Zhu W, Li W, Cui X, Gui H. An approach to time series classification using binary distribution tree. Proceedings of the 2019 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN); 2019:399-404.
12. Bagnall A, Bostrom AG, Large J, Lines J. The great time series classification bake off: an experimental evaluation of recently proposed algorithms. Extended Version. ArXiv abs/1602.01711, 2016.
13. Ismail Fawaz H, Forestier G, Weber J, Idoumghar L, Muller PA. Deep learning for time series classification: a review. *Data Mining Knowl Discov*. 2019;33:917-963.
14. Mengchen Y, Xudong C, Peng C, Hum N. A review of research on early time series classification methods. *J East China Normal Univ (Nat Sci Ed)*. 2021;05:115-133.
15. Zhang W, Wang ZH, Yuan JD, Hao SHL. A time series discriminative feature dictionary construction algorithm. *J Softw*. 2020;31(10):3216-3237. doi:10.13328/j.cnki.jos.005852
16. Binkowski M, Marti G, Donnat P. Autoregressive convolutional neural networks for asynchronous time series. ArXiv abs/1703.04122, 2018.

How to cite this article: Liu Q, Huang M, Lee W-S, Du Y. A set of time-series classification datasets based on the average price of concrete in major Chinese cities. *Engineering Reports*. 2022;4(11):e12541. doi: 10.1002/eng2.12541