

Automatic SARIMA Order Identification Convolutional Neural Network

Paisit Khanarsa and Krung Sinapiromsaran

Abstract—Deep learning concept is popularly used to solve a classification problem including the model identification in time series analysis. For the time series models such as the autoregressive integrated moving average (ARIMA) model and the seasonal autoregressive integrated moving average (SARIMA) model, statisticians mostly identify the ARIMA/SARIMA orders before building the forecasting model. To identify them, they investigate sample ACF and PACF to extract p , d , q while most researchers automate this step using the likelihood based-method via AIC ignoring the residual diagnostic. This paper uses ACF, PACF and differencing time series images as inputs to the convolutional neural network architecture that automatically identifies the ARIMA/SARIMA orders, called the automatic ARIMA order identification convolutional neural network (AOC). The performance of AOC outperforms the likelihood based-method in terms of identifying ARIMA order via precision, recall and f1-score. Moreover, AOC is extended to identify the SARIMA order, called the automatic SARIMA order identification convolutional neural network (ASOC). The performance of the ASOC model provides better performance than the likelihood method via precision, recall and f1-score.

Index Terms—Convolutional neural network, ARIMA, SARIMA, ACF, PACF.

I. INTRODUCTION

In recent years, deep learning has received much more attention for solving several machine learning problems [1], especially classification. One of the deep learning architectures is a convolutional neural network (CNN) having success in training images from ImageNet [2]. For the traditional CNN architecture, LeNet was proposed by LeCun *et al.* in 1998 [3] for classifying digits. Later, CNN has been used in many fields, in particular, AlexNet was proposed by Krizhevsky *et al.* in 2012 [4]. Later, there were many CNN architectures such as VGGNet by Simonyan and Zisserman in 2014 [5], GoogLeNet by Google in 2014 [6] and ResNet by Kaiming He *et al.* in 2015 [7]. Baloglu *et al.*, in 2019 [8] used deep CNN for classification of myocardial infarction with multi-lead ECG signals. Wang *et al.*, in 2019 [9] proposed a new CNN for applying on point clouds, including classification and segmentation. Hekler *et al.*, in 2019 [10]

applied deep learning to 11 pathologists in the classification of histopathological melanoma images. Fewaz *et al.*, in 2016 [11] proposed a novel CNN to data augmentation for time series classification. Brunel *et al.*, in 2019 [12] adapted A CNN to time series for the classification of Supernovae. Moreover, there are some other researches about classification problems, such as Mohammadreza *et al.* in 2012 [13], Liu *et al.* in 2018 [14] and Abdulsalam *et al.* in 2019 [15]

Time series analysis is a branch of statistics that studies phenomena of time series data. It employs different models which may include the autoregressive integrated moving average model (ARIMA) and the seasonal autoregressive integrated moving average model (SARIMA) by Box and Jenkins in 1970 [16], the self-exciting threshold autoregressive model (SETAR) by Tsay in 1989 [17], the autoregressive conditional heteroskedasticity model (ARCH) by Engle in 1983 [18] and the generalized autoregressive conditional heteroskedasticity model (GARCH) by Bollerslev in 1990 [19].

Two broad models for univariate applied time series models are ARIMA and SARIMA which the future value derives linearly from previous values. To build these models, the Box-Jenkins method and the likelihood-based method are used. At present, they contribute to solve many time series domains such as annual runoff series by Wang *et al.* in 2015[20], monthly and weekly stock prices by Sima and Naminf in 2018 [21], wheat production in India by Bhola Nath *et al.* in 2019 [22]. For SARIMA, Valipour in 2015 [23] forecasted long-term runoff in the United States using SARIMA. Fang and Lahdelma in 2016 [24] evaluated a multiple linear regression and the SARIMA model in forecasting heat demand for district heating systems. After that, Lang *et al.* in 2018 [25] proposed a new model combining the SARIMA model and the ordinary kriging method for forecasting concentrations of PM 2.5 in the urban area of Hangzhou.

To achieve the best forecasting model, the ARIMA order or the SARIMA order must be identified before applying the Box-Jenkins parameter estimation. The sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) are plotted for identifying orders called the Box-Jenkins identification [26]. Alternatively, the likelihood-based method automatically generates the SARIMA models of all possible orders based on AIC [27] or BIC [28] and it reports the SARIMA model with the minimum AIC or BIC. Currently, various researchers attempted to improve the method for identifying these orders such as Lee and Oh in 1996 [29]. Chenoweth *et*

Manuscript received February 12, 2020; revised April 7, 2020.

The authors are with the Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand (e-mail: paisitkhanarsa@gmail.com, krung.s@chula.ac.th).

al. in 2000 [30], Al-Qawasmi *et al.* in 2010 [31]. Newly, Tang and Röllin in 2018 [32] used the deep learning model, called ResNet50, to identify the ARMA order and compare the performance of the deep learning model with the likelihood-based method based on AIC.

The issue of identifying the ARIMA order by the Box-Jenkins identification uses sample ACF and sample PACF via human which may provide different orders from different viewpoints. Even though the likelihood-based method automatically identifies order without human interaction, the results of identifying the ARIMA order are still disappointing due to improper residual behavior of the resulting time series models. Consequently, this paper resolves this issue using ACF, PACF and differencing time series data images as inputs to the deep learning model, called the automatic ARIMA order identification convolutional neural network (AOC). Subsequently, AOC is extended to identify the SARIMA order, called the automatic SARIMA order identification convolutional neural network (ASOC). This model identifies the SARIMA order by approximating the seasonal length using the spectral analysis [33]. Then the aggregate time series data by the seasonal length as input to AOC to identify the order for the seasonal autoregressive integrated moving average model.

This paper consists of seven sections. The first section is the introduction. The second section is the review of the convolution neural network. The third section covers ARMA, ARIMA and SARIMA models. The fourth and fifth sections describe the AOC model and the ASOC model. The sixth section covers the results of these models. The last section is the conclusion and discussion.

II. REVIEW OF CONVOLUTIONAL NEURAL NETWORK FOR TIME SERIES

In deep learning concept, the convolutional neural network (CNN) is one of deep neural networks, commonly applied to analyzing visual images [2]. The concept of CNN is to extract the feature of an image by applying the filter matrix and followed by additional convolutions such as pooling layers, fully connected layers and normalization layers. Presently, CNN is applied popularly to various fields including time series data. Selvin *et al.* in 2017[34] used LSTM, RNN and CNN-sliding window model to predict the Stock price. Li *et al.* in 2017[35] used a deep learning model for anomaly detection in hyperspectral imagery. Lu *et al.* in 2018 [36] combined three-dimensional CNN and LSTM for applying with video kinematic data. Ji *et al.* in 2019 [37] applied carbon futures price forecasting to the ARIMA-CNN-LSTM model. Bogaerts *et al.* in 2020 [38] used trajectory data of traffic as inputs to the CNN-LSTM neural network for predicting short and long-term future values. Kim and Cho in 2019 [39] used CNN-LSTM neural networks to predict residential energy consumption. Clearly, CNN can effectively apply to a wide range of problems. Thus, this research will take advantage of CNN to identify the ARIMA order by training stationary time series data using ACF and PACF images.

III. TIME SERIES MODELS

In this section, the formula of the autoregressive moving average (ARMA) model, the autoregressive integrated moving average (ARIMA) model, and the seasonal autoregressive integrated moving average (SARIMA) model are shown.

The ARMA model of time series analysis of order (p, q) can be expressed

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

or

$$\phi_p(L) z_t = \theta_q(L) a_t$$

where

$$\phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

and

$$\theta_q(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$$

where $\phi_1, \phi_2, \dots, \phi_p$ represent the coefficients of the AR component; $\theta_1, \theta_2, \dots, \theta_q$ represent the coefficients of the MA component. L represents the backward shift operator $z_{t-1} = Lz_t$.

The general ARMA model having the differencing order is the ARIMA model of order (p, d, q) which can be written as

$$\phi_p(L) \nabla^d z_t = \theta_q(L) a_t$$

where d is the number of differencing and ∇ is the differencing operator computing by $\nabla z_t = z_t - z_{t-1}$.

More generally, seasonal effect is applied to the ARIMA model defining the seasonal period to be s . The model is called SARIMA of order $(p, d, q) \times (P, D, Q)$.

$$\Phi_p(L^s) \phi_p(L) (1-L)^d (1-L^s)^D z_t = \theta_q(L) \Theta_Q(L^s) a_t$$

where

$$\phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p,$$

$$\theta_q(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q,$$

$$\Phi_p(L^s) = 1 - \phi_1 L^s - \phi_2 L^{2s} - \dots - \phi_p L^{Ps},$$

$$\Theta_Q(L^s) = 1 - \theta_1 L^s - \theta_2 L^{2s} - \dots - \theta_Q L^{Qs}.$$

P is the order of seasonal autoregressive term; Q is the order of moving average term and D is the number of seasonal differencing which $\phi_1, \phi_2, \dots, \phi_p$ represent the coefficients of the seasonal AR and $\theta_1, \theta_2, \dots, \theta_Q$ represent the coefficients of the seasonal MA.

To identify the ARIMA order, the plots of ACF and PACF are acquired [23]. The formula for calculating sample ACF and sample PACF are described as follows:

Give an observed time series data z_1, z_2, \dots, z_n , sample ACF ($\hat{\rho}_k$) at lag k is calculated by

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2}, \quad k = 0, 1, 2, \dots$$

$$\text{where } \bar{z} = \frac{\sum_{t=1}^n z_t}{n}.$$

Moreover, sample PACF ($\hat{\varphi}_{kk}$) at lag k is calculated by

$$\hat{\varphi}_{kk} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^k \hat{\varphi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\varphi}_{kj} \hat{\rho}_j}$$

where $\hat{\rho}_k$ is sample ACF at lag k and

$$\hat{\varphi}_{k+1,j} = \hat{\varphi}_{kj} - \hat{\varphi}_{k+1,k+1} \hat{\varphi}_{k,k+1-j}, j = 0, 1, 2, \dots, k$$

IV. THE AUTOMATIC ARIMA ORDER IDENTIFICATION CONVOLUTIONAL NEURAL NETWORK

This section demonstrates the method for synthesizing time series data and the CNN model for building the automatic ARIMA order identification convolutional neural network (AOC). The AOC model is constructed by training ACF and PACF images computing from the time series data varying p, q from 0-5. In the experiments, the simulating process for the ARMA order randomly generates coefficients from the standard normal distribution. The last p^{th} coefficient of AR coefficients is set as $\max\{\phi_1, \phi_2, \dots, \phi_p\} + 0.1$ and the last q^{th} coefficient of MA coefficients is set as $\max\{\phi_1, \phi_2, \dots, \phi_q\} + 0.1$.

Simulated time series data are generated by varying p and q where the number of training time series data, validating time series data and testing time series data are 7200, 3600 and 360, respectively. In case of identifying d order, the time series data from simulating p and q order is accumulated by varying d from 0 to 3 having 28800 for training, 14400 for validating and 1440 for testing. Then, sample ACF and PACF are calculated and plotted to black and white images as Fig. 1.

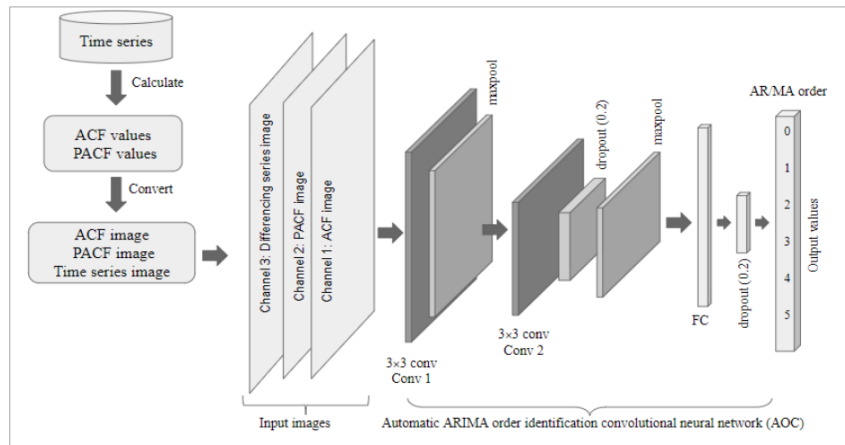


Fig. 2. The architecture of the AOC model for training the AR/MA order.

V. THE AUTOMATIC SARIMA ORDER IDENTIFICATION CONVOLUTIONAL NEURAL NETWORK

After obtaining p, d, q from AOC, it is extended to identify the SARIMA order (P, D, Q). The whole process is called the automatic SARIMA order identification convolutional neural network (ASOC). The first step, the seasonal length is calculated by spectral analysis of the time series [33]. In the concept of spectral analysis, a linear trend is removed, then the spectral density function is approximated from the best

The architecture of the AOC model composes of the input layer, two convolutional layers, one fully connected layer and softmax layer. The input layer of the AOC model comprises of 1, 2 or 3 channels of images. The input image is sent through a convolutional layer of size 3×3 having the stride and the spatial padding are set to 1. Then, max-pooling layers are applied. The window size of max-pooling is set to 2×2 having the stride = 2. The last part of the architecture is followed by fully connected layers which include 512 nodes and the activation function is a linear function. Finally, the final layer is the softmax layer. The AOC model is divided into three types consisting of the p -AOC model for training the AR order, the q -AOC model for training the MA order and the d -AOC model for training the differencing order. The architecture of these AOC models are shown in Fig. 2 and Fig. 3.

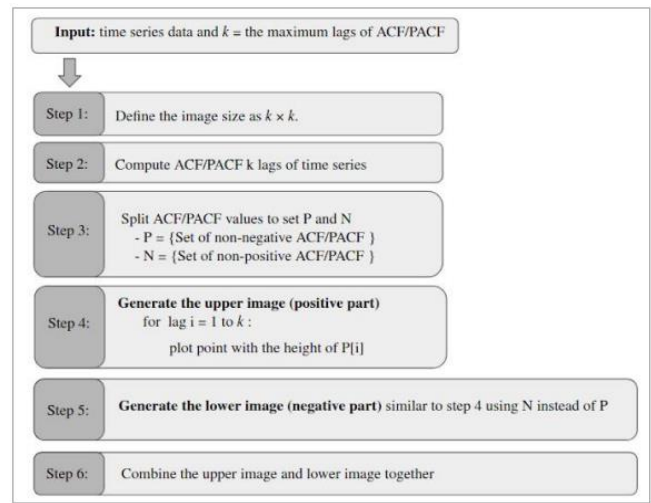


Fig. 1. The method for generating images of ACF/PACF.

fitting autoregressive model at frequency f . If there is a maximum in spectral density function at f , the seasonal length will be $1/f$. After obtaining the seasonal length, the time series data will be aggregated by the seasonal length. Finally, the aggregate time series data will be used to identify seasonal order P, D and Q by AOC. The whole process of ASOC is shown in Fig. 4 by defining s as the seasonal length, x_1, x_2, x_3, \dots are the original time series data and y_1, y_2, y_3, \dots are the aggregate time series data.

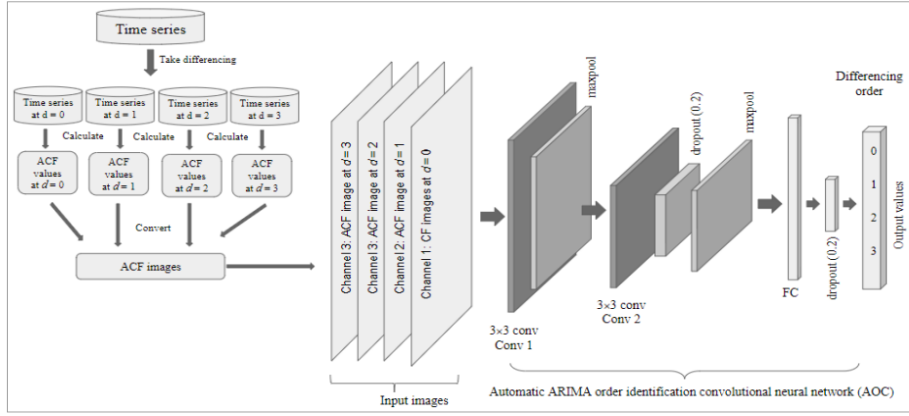


Fig. 3. The architecture of the AOC model for training differencing order.

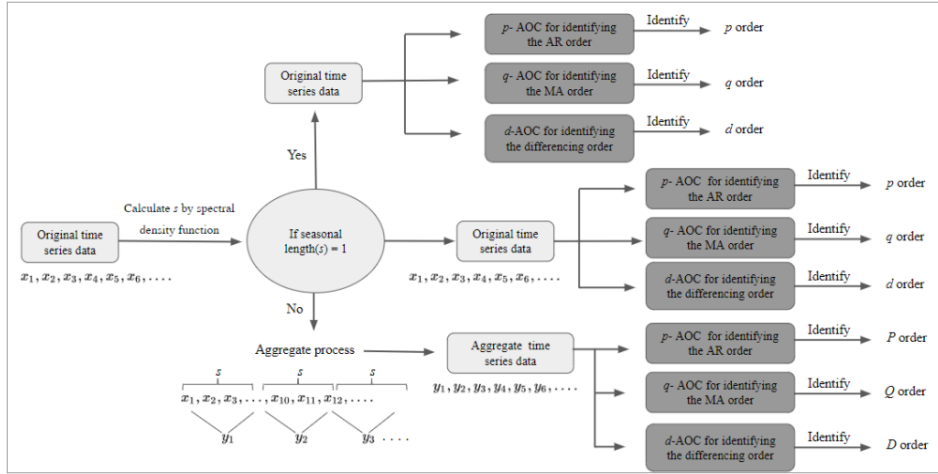


Fig. 4. The automatic SARIMA order identification convolutional neural network (ASOC).

TABLE I: DESCRIPTION OF EACH CHANNEL IN THE MODELS

Model	Model	Case: p varying q	Case: q varying p	Case: d varying p and q
		Channel(s)	Channel(s)	Channel(s)
Model A1	AOC	1 channel: PACF images	1 channel: ACF images	3 channels: ACF images with $d = 0, 1, 2$ and 3
Model A3	AOC	2 channels: PACF and ACF images	2 channels: ACF and PACF images	None
Model A2	AOC	3 channels: PACF, ACF and differencing series images	3 channels: PACF, ACF and differencing series images	None
Model B	ResNet50	Series	Series	None
Model C	Likeli hood method	Series	Series	Series

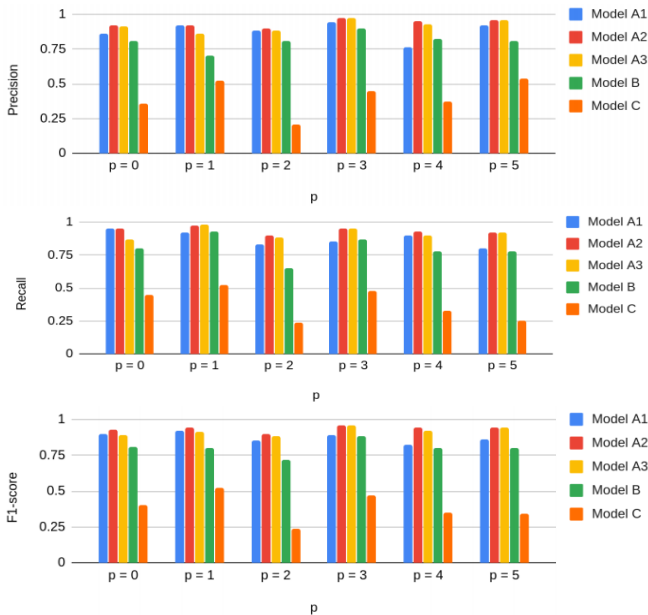


Fig. 5. Precision, recall and f1-score of the models by considering p varying q .

VI. EXPERIMENTAL RESULTS

The experiments are divided into three cases where the first case is called p varying q fixing the AR order p in $\{0, 1, 2, 3, 4, 5\}$ while varying the MA order from 0 to 5 repeating 200 times for training, 100 times for validating and 10 times for testing. The second case is called q varying p fixing the MA order q in $\{0, 1, 2, 3, 4, 5\}$ while varying the AR order from 0 to 5 repeating 200 times for training, 100 times for validating and 10 times for testing. The third case is called d varying p and q fixing the differencing order d in $\{0, 1, 2, 3\}$ while varying both the AR order and the MA order from 0 to 5 repeating 200 times for training, 100 times for validating and 10 times for testing. Table I characterizes different models having different channels of inputs.

A. Results of the AOC Model in the Case of p Varying q

Fig. 5 summarize the scores of all models corresponding to precision, recall and f1-score, respectively. From this figure,

model A1, A2 and A3 are the top-3 best scores providing scores close to 1. The performance of model A1, A2 and A3 is quite similar however model A2 using 2 channels of PACF and ACF is the best while both model B and model C demonstrates low performance.

B. Results of the AOC Model in the Case of q Varying p

From Fig. 6, model A2 and A3 provides the best precision, recall and f1-score. Nevertheless, the precision and f1-score of A3 are highest with respect to other models. Model B provides the lowest scores of precision, recall, f1-score predicting q . Model C is better than model B, yet it is inferior to model A1, model A2, model A3 for all measures.

C. Results of the AOC Model in the Case of d Varying p and q

Fig. 7 demonstrates precision, recall and f1-score between model A1 and model C which it is clear that model A1 gives the best precision, recall and f1-score.

TABLE II: DESCRIPTION OF EACH CHANNEL IN THE MODELS

Model	Case: P varying Q Channel(s)	Case: Q varying P Channel(s)	
			Model A1
Model C	Likeli hood method	Series	Series

D. Results of the ASOC Model in the Case of Identifying P and Q

In the case of testing performance of the ASOC model, precision, recall and f1-scores are also applied for evaluating performance of identifying seasonal order P and Q . Table II

demonstrates different models for cases having different inputs. The synthesized datasets are divided into two cases where the first case is called P varying Q fixing the seasonal AR order P in $\{0, 1, 2, 3, 4, 5\}$ while varying the seasonal MA order Q from 0 to 5. Another case Q varying P is done in the same way.

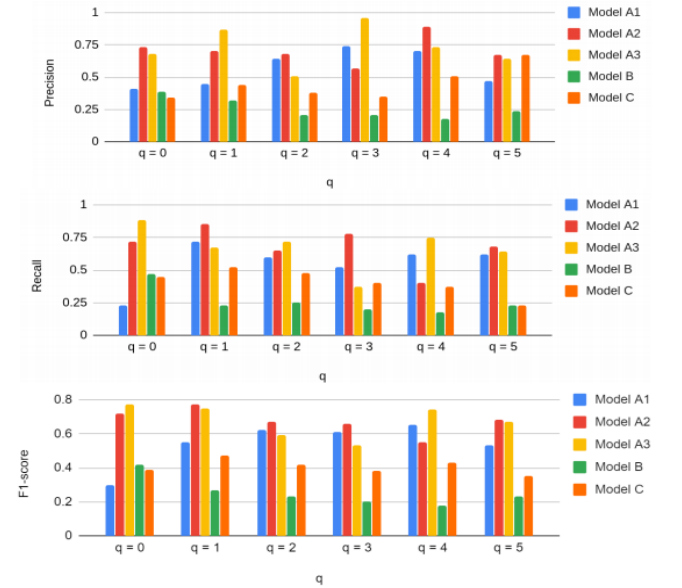


Fig. 6. Precision, recall and f1-score of the models by considering q varying p .

From Fig. 8, model A1 gives the best precision and f1-score in all cases of identifying order P except the cases of identifying $P = 1$ and 2 and model C provides high recall. However model A1 provide the highest recall in case of identifying $P = 0$. It is clear that model A1 can identify the seasonal order better than model C.

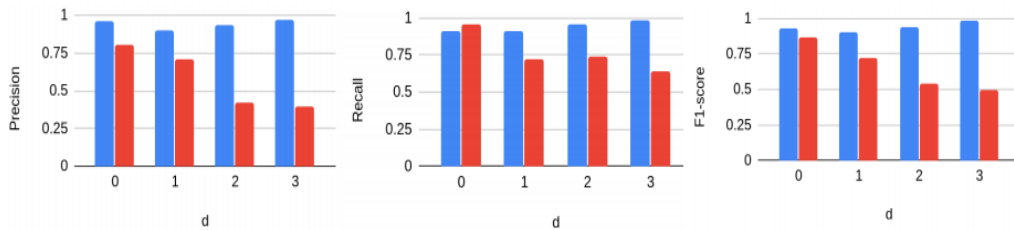


Fig. 7. Precision, recall and f1-score of the models by considering d varying p and q .

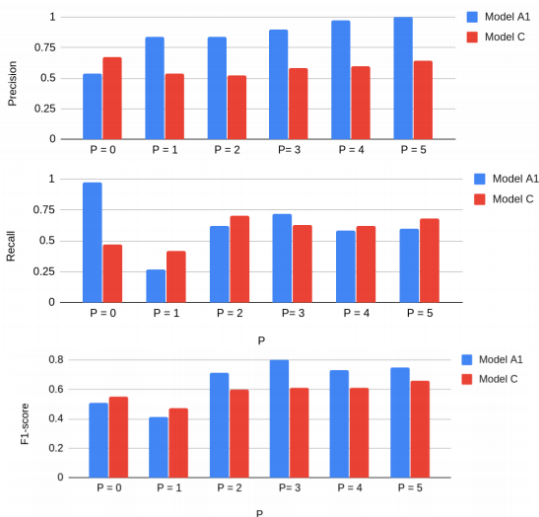


Fig. 8. Precision, recall and f1-score of the models by considering P varying Q .

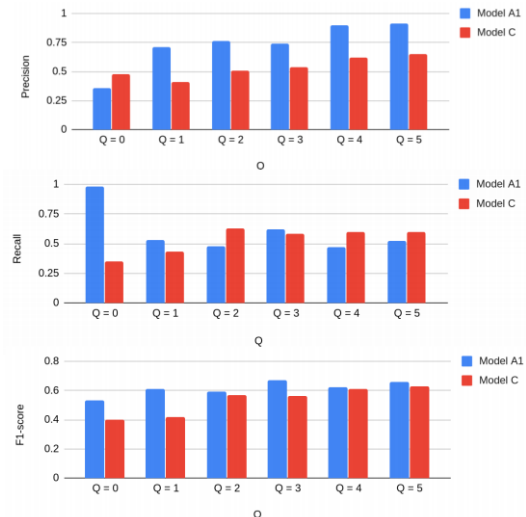


Fig. 9. Precision, recall and f1-score of the models by considering Q varying P .

From Fig. 9, model A1 outperforms model C via precision and f1-score except the case of identifying $Q = 0$.

Nevertheless, Model C provides the highest recall except the case of identifying $Q = 0$ which model A1 shows high recall. These results can be concluded that model A1 can identify the seasonal order better than model C.

VII. CONCLUSION AND DISCUSSION

In this research, the automatic ARIMA order identification convolutional neural network (AOC) model is created by training ACF and PACF images as the input data to identify the ARIMA order. The results of the AOC model provide the suitable ARIMA order better than the likelihood-based method.

Moreover, the AOC model is improved for identifying the SARIMA order, called the automatic SARIMA order identification convolutional neural network (ASOC) model.

From the experiments, the ASOC model can identify the SARIMA order better than the likelihood-based method when considering precision, recall and f1-score. The outcome from this research affirms that the use of sample ACF and sample PACF is still suitable for model identification of the SARIMA order since ACF and PACF can demonstrate the actual relation that occurs in time series data, whereas the likelihood-based method does not identify order from the actual relation, it just picks the order according to the lowest residual error.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Author Paisit Khanarsa designed the framework of the deep learning problem and defined the process of building the AOC model and ASOC model to solve the identifying problem of the ARIMA and SARIMA order. Author Krung Sinapiromsaran helped to analyze and criticize the empirical results and helped editing the manuscript. Both authors wrote the paper and approved the final version.

ACKNOWLEDGMENT

I would like to express appreciation to my advisor, Assistant Professor Krung Sinapiromsaran for his invaluable help and constant encouragement of this research. In addition, I most gratefully acknowledge Science Achievement Scholarship of Thailand (SAST) for funding this research and the Applied Mathematics and Computational Science, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University for supporting this work.

REFERENCES

- [1] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational Intelligence and Neuroscience*, 2018.
- [2] J. Deng, W. Dong, R. Socher *et al.*, "Imagenet: A large-scale hierarchical image database," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248-255.
- [3] Y. LeCun. (2015). LeNet-5, convolutional neural networks. [Online]. Available: <http://yann.lecun.com/exdb/lenet>
- [4] N. Aloysius and M. Geetha, "A review on deep convolutional neural networks," in *Proc. 2017 International Conference on Communication and Signal Processing (ICCSP)*, 2017, pp. 0588-0592.
- [5] A. S. Modi, "Review article on deep learning approaches," in *Proc. 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2018, pp. 1635-1639.
- [6] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27-48, 2016.
- [7] S. Targ, D. Almeida, and K. Lyman, "Resnet in resnet: Generalizing residual architectures," arXiv: 1603.08029, 2016.
- [8] U. B. Baloglu, M. Talo, O. Yildirim, R. S. Tan, and U. R. Acharya, "Classification of myocardial infarction with multi-lead ECG signals and deep CNN," *Pattern Recognition Letters*, vol. 122, pp. 23-30, 2019.
- [9] Y. Wang, Y. Sun, Z. Liu *et al.*, "Dynamic graph cnn for learning on point clouds," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 5, pp. 1-12, 2019.
- [10] A. Hekler, J. S. Utikal, A. H. Enk *et al.*, "Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images," *European Journal of Cancer*, vol. 118, pp. 91-96, 2019.
- [11] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P. A. Muller, "Data augmentation using synthetic data for time series classification with deep residual networks," arXiv:1808.02455, 2018.
- [12] A. Brunel, J. Pasquet, J. Pasquet, N. Rodriguez, F. Comby, D. Fouchez, and M. Chaumont, "A CNN adapted to time series for the classification of Supernovae," *Electronic Imaging*, 2019, vol. 14, pp. 90-91, 2019.
- [13] M. Keyvanpour, N. Izadpanah, and H. Karbasforoushan, "Classification and evaluation of high-dimensional image indexing structures," *International Journal of Machine Learning and Computing*, vol. 2, no. 3, pp. 252-256, 2012.
- [14] Q. Liu, Z. Zhou, S. R. Shakya, P. Uduthalappally, M. Qiao, and A. H. Sung, "Smartphone sensor-based activity recognition by using machine learning and deep learning algorithms," *International Journal of Machine Learning and Computing*, vol. 8, no. 2, pp. 121-126, 2018.
- [15] W. H. Abdulsalam, R. S. Alhamdani, and M. N. Abdullah, "Facial emotion recognition from videos using deep convolutional neural networks," *International Journal of Machine Learning and Computing*, vol. 9, no. 1, p. 6, 2019.
- [16] G. E. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control Holden-Day*, San Francisco, p. 498, 1970.
- [17] R. S. Tsay, "Testing and modeling threshold autoregressive processes," *Journal of the American Statistical Association*, vol. 84, no. 405, pp. 231-240, 1989.
- [18] R. F. Engle, "Estimates of the variance of US inflation based upon the ARCH model," *Journal of Money, Credit and Banking*, vol. 15, no. 3, pp. 286-301, 1983.
- [19] T. Bollerslev, "Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model," *The Review of Economics and Statistics*, pp. 498-505, 1990.
- [20] W. C. Wang, K. W. Chau, D. M. Xu, and X. Y. Chen, "Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition," *Water Resources Management*, vol. 29, no. 8, pp. 2655-2675, 2015.
- [21] S. Siami-Namini and A. S. Namin. (2018). Forecasting economics and financial time series: ARIMA vs. LSTM. [Online]. Available: <https://arxiv.org/abs/1803.06386v1>, vol. abs/1803.06386
- [22] B. Nath, D. S. Dhakre, and D. Bhattacharya, "Forecasting wheat production in India: an ARIMA modelling approach," *Journal of Pharmacognosy and Phytochemistry*, vol. 8, no. 1, pp. 2158-2165, 2019.
- [23] M. Valipour, "Long-term runoff study using SARIMA and ARIMA models in the United States," *Meteorological Applications*, vol. 22, no. 3, pp. 592-598, 2015.
- [24] T. Fang and R. Lahdelma, "Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system," *Applied Energy*, vol. 179, pp. 544-552, 2016.
- [25] Y. Lang, L. Xiao, and C. George, "Forecasting concentrations of PM 2.5 in main urban area of Hangzhou and mapping using SARIMA model and ordinary Kriging method," *Acta Sci. Circum*, vol. 1, pp. 62-70, 2018.
- [26] W. W. Wei, "Time series analysis," *The Oxford Handbook of Quantitative Methods in Psychology*, vol. 2, 2006.
- [27] H. Akaike, "Factor analysis and AIC," *Selected papers of Hirotugu Akaike*, Springer, New York, NY, pp. 371-386, 1987.
- [28] A.A. Neath and J. E. Cavanaugh, "The Bayesian information criterion: background, derivation, and applications," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 199-203, 2012.
- [29] K. C. Lee and S. B. Oh, "An intelligent approach to time series identification by a neural network-driven decision tree classifier," *Decision Support Systems*, vol. 17, no. 3, pp. 183-197, 1996.

- [30] T. Chenoweth, R. Hubata, and R. D. S. Louis, "Automatic ARMA identification using neural networks and the extended sample autocorrelation function: A reevaluation," *Decision Support Systems*, vol. 29, no. 1, pp. 21-30, 2000.
- [31] K. E. Al-Qawasmī, A. M. Al-Smadi, and A. Al-Hamami, 2010, "Artificial neural network-based algorithm for ARMA model order estimation," in *Proc. International Conference on Networked Digital Technologies*, Springer, Berlin, Heidelberg, pp. 184-192.
- [32] W. H. Tang and A. R. ċlin, "Model identification for ARMA time series through convolutional neural networks," arXiv: 1804.04299, 2018.
- [33] R. J. Hyndman, 2019, *CRAN Task View: Time Series Analysis*.
- [34] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in *Proc. 2017 International Conference on Advances in Computing, Communications and Informatics*, pp. 1643-1647.
- [35] W. Li, G. Wu, and Q. Du, "Transferred deep learning for anomaly detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 597-601, 2017.
- [36] N. Lu, Y. Wu, L. Feng, and J. Song, "Deep learning for fall detection: Three-dimensional CNN combined with LSTM on video kinematic data," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 314-323, 2018.
- [37] L. Ji, Y. Zou, K. He, and B. Zhu, "Carbon futures price forecasting based with ARIMA-CNN-LSTM model," *Procedia Computer Science*, vol. 162, pp. 33-38, 2019.
- [38] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx, "A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 62-77, 2020.
- [39] T. Y. Kim and S. B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72-81, 2019.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



Paisit Khanarsa received the B.Sc. degrees (first-class honours) in mathematics from Kasetsart University, Bangkok, Thailand, in 2014 and the M.S. degree in applied mathematics and computational science from Chulalongkorn University, Bangkok, Thailand, in 2016. Since 2017, he has been a Ph.D. candidate in applied mathematics and computational science program at Chulalongkorn University. His research interests include time series and deep learning.



Krung Sinapiromsaran received his B.S. in mathematics from Chulalongkorn University, his M.S. and Ph.D. in computer science from the University of Wisconsin-Madison. He is currently an assistant professor in the Department of Mathematics, Chulalongkorn University. His ongoing research work are related to deep learning, machine learning, artificial intelligence, data mining, knowledge discovery, and optimization.