

Received February 3, 2022, accepted February 20, 2022, date of publication February 25, 2022, date of current version March 8, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3154812

AIS-Based Intelligent Vessel Trajectory Prediction Using Bi-LSTM

CHENG-HONG YANG^{1,2,3}, (Senior Member, IEEE), CHIH-HSIEN WU¹, JEN-CHUNG SHAO¹,
YI-CHUAN WANG⁴, AND CHIH-MIN HSIEH⁵

¹Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 80778, Taiwan

²Institute of Biomedical Engineering, Kaohsiung Medical University, Kaohsiung 80708, Taiwan

³School of Dentistry, Kaohsiung Medical University, Kaohsiung 80708, Taiwan

⁴Department of Business Administration, CTBC Business School, Tainan 709237, Taiwan

⁵Department of Maritime Information and Technology, National Kaohsiung University of Science and Technology, Kaohsiung 80778, Taiwan

Corresponding author: Cheng-Hong Yang (chyang@nkust.edu.tw)

This work was supported by CECI Engineering Consultants, Inc., under Grant 10944.

ABSTRACT Accurate vessel trajectory prediction is essential for maritime traffic control and management. In addition to collision avoidance, accurate vessel trajectory prediction can help in planning navigation routes, shortening the sailing distance, and increasing navigation efficiency. Vessel trajectory prediction with automatic identification system (AIS) data has thus attracted considerable attention in the maritime industry. Original AIS data may contain noise, which limits their application in real-world maritime traffic management. To overcome this problem, this study proposes a vessel trajectory prediction method that combines data denoising and a deep learning prediction model. In this method, data denoising is realized in three steps: trajectory separation, data denoising, and standardization. First, outliers from the original AIS data samples are removed, after which the moving average model is employed to further clean up the data; finally the denoised data are standardized into uniformly distributed time-series data. Bidirectional long short-term memory (Bi-LSTM) is then applied for vessel trajectory prediction. The performance of the proposed prediction model was verified using data on the trajectories of ten vessels and comparing the results obtained with those obtained using other prediction models (exponential smoothing, autoregressive integrated moving average, support vector regression, recurrent neural network, and LSTM models); the trajectory data were downloaded from a public AIS database. The experimental results revealed that model prediction accuracy increased after the data denoising process. Specifically, the Bi-LSTM model had the lowest mean absolute error, mean absolute percentage error, and root-mean-square error, demonstrating that the proposed method is highly efficient for trajectory prediction and can help vessel traffic controllers predict accurate vessel tracks; this would enable them to take early preventive measures to avoid collisions and thus improve the efficiency and safety of maritime traffic.

INDEX TERMS Deep learning, bidirectional long short-term memory, trajectory prediction, collision avoidance.

I. INTRODUCTION

Approximately 71% of the Earth's surface is covered by water, and only approximately 21% of the surface is land. Taiwan is an island country located at the intersection of Northeast Asia and Southeast Asia. The eastern half of Taiwan is close to the largest ocean in the world. Oil tankers, freighters, merchant vessels, and other vessels often travel across the surrounding Pacific Ocean. However, because vessels frequently traverse the western Pacific Ocean

international trade routes and because of frequent and prosperous fishing activities, the marine traffic flow around Taiwan is of medium to high complexity. According to data from the Lloyd's Register Casualty Returns, the Taiwan sea area is classified as a moderate-risk environment [1]. Moreover, according to relevant statistical data from the Ministry of Transportation and Communications R.O.C. and the Coast Guard Administration, Ocean Affairs Council, in recent years, the average number of marine vessel disasters that occur in the waters around Taiwan has reached 100 [2]. Maritime transportation constitutes more than 90% of the global cargo trade. Because maritime traffic accidents are

The associate editor coordinating the review of this manuscript and approving it for publication was Sudipta Roy¹.

likely to cause considerable loss of life and environmental damage, improving maritime traffic safety has become a priority. The safety of ships sailing at sea is a key problem in maritime areas or ports with high traffic density and complicated conditions. A vessel traffic service (VTS), whose purpose is to accurately and effectively monitor and predict vessel trajectories (including in real time), provides valuable technical support for the early warning of marine traffic accidents [3]. To improve the safety of ships sailing in an environment with complex and ever-changing sea conditions, it is necessary to provide trajectory prediction and danger warning functions to a ship's intelligent navigation system. However, the maritime navigation environment is prone to many accidents, especially in crowded port waters, and it is not easy to predict moving targets.

The safety and integrity of marine vessels against hazards should be maintained when planning the ship's course. Considering this, the International Maritime Organization (IMO) proposes the application of various navigation systems, such as a bridge navigational watch alarm (BNWAS), automatic identification system (AIS), and electronic chart display and information (ECDIS), as tools to support the officers on watch (OOWs). Information collected by an automatic identification system (AIS) can help prevent maritime accidents and improve maritime situational awareness. An AIS collects all types of maritime surveillance data to provide accurate early warning information, including maritime traffic spatial information, to maritime traffic operators and provide support for various navigation operation decisions. Thus, an AIS can obtain information in a timely manner along the vessel. Studies have demonstrated that AIS data have a major impact on maritime traffic safety analysis. Therefore, improving the quality of AIS data has become an active topic in maritime informatics research [4]–[6]. The popularity of AISs has led to the recording and storage of a large quantity of vessel navigation data; nearly a trillion pieces of information are stored every day. However, this large quantity of data includes abnormal data caused by factors such as signal interference and mechanical failure [7]. Methods for processing abnormal data from AISs involve the use of statistics [8], unsupervised machine learning, and artificial neural networks [9], [10]. In addition, many methods, theories, and technologies involving the use of massive data, data exploration, and machine learning have been proposed. Developing an approach for efficiently using the information obtained to increase maritime transportation intelligence has become critical [11], [12]. However, the AIS onboard a vessel usually adds one piece of information every 2 s to 3 min, which results in an extremely large AIS dataset; this also engenders the data redundancy problem, which makes it difficult to use the data for research and practical applications. Data sparsity and erroneous information automatically extracted from unstructured or semistructured sources may result in the creation of the same product two or more times (i.e., duplicates). Product duplicates are not only one of the main sources of bad product experiences but also make

product matching harder as multiple entities of the same product exist with overlapping information. The duplicates can have many fatal effects, including preventing machine learning algorithms from discovering important consistencies in product representations [13].

If repeated data is only generated 2–3 times, it will make the model more significant because the variance would be reduced. However, we should ideally remove repeated data. Duplicates are an extreme case of nonrandom sampling, and they introduce bias in the fitted model.

Therefore, data compression technology is highly crucial in an AIS [14]–[16]. Automation is more challenging at sea than it is at land. Unlike cars on the road, sea vessels can travel in all directions. In addition, compared with car traffic rules, rules stipulated by the Convention on the International Regulations for Preventing Collisions at Sea are less quantifiable and their implementation relies more on experience; this thus increases the difficulty of predicting the navigation behavior of vessels. Numerous studies have presented models and procedures for track prediction using AIS data; examples of such models and procedures include a Kalman filter model [17], Markov model [18], optimal route estimation based on clustering, and an ant colony algorithm [19]. Some researchers also used neural networks to predict vessel trajectories without prior knowledge [20]. The Multihypothesis of ship navigation addresses the occasion where ships may have multiple shipping lanes. Multihypothesis means a ship may choose to follow one of two or more shipping lanes, which may lead to different predictions. Ship track prediction may also classify the initial ships by clustering and predict a similar navigation trajectory to achieve a more accurate trajectory prediction [21].

Recent developments in deep learning methods, such as deep neural networks [22] (including convolutional neural networks [23] and a recurrent neural networks (RNNs) [24]), have had a considerable impact in the fields of computer vision, natural language processing, and speech recognition [25]. Long short-term memory (LSTM) can be used to model a complex function and extract various features from a large dataset [26]. Bidirectional LSTM (Bi-LSTM) used to predict DNA-Protein binding [27]. However, studies on the use of deep learning methods for vessel trajectory prediction are scarce; only a few recent studies have explored the use of deep learning methods for vessel trajectory prediction [28]–[31].

Inspired by the successful use of deep learning methods in sequence prediction, we investigated whether an RNN model can be employed for vessel track prediction. In this study, a method for vessel trajectory denoising and vessel trajectory prediction was proposed for determining future vessel trajectories according to the given AIS observation sequence. The proposed method is based on the bidirectional LSTM (Bi-LSTM) structure, which has become an effective and scalable structure for sequence prediction. Considering the lag between the input sample and the output to be predicted, the ability of an LSTM network to learn data

with long-term time dependence renders it ideal for use in the trajectory prediction task [26]. The main contributions of this study are as follows: (1) cleaning of the original AIS data through trajectory separation, outlier deletion, and data standardization; (2) development of a Bi-LSTM model for predicting the trajectory of a vessel using the denoised AIS data; (3) verification of the performance of the proposed model using the trajectory of ten vessels. This study can help vessel traffic controllers accurately predict the trajectory of vessels, which would allow them to take preventive measures to avoid collisions and improve the efficiency and safety of maritime traffic.

II. RELATED WORK

Trajectory prediction methods based on statistical methods are commonly used in the maritime industry, among which methods based on Gaussian process regression are the most common. Anderson *et al.* [32] used time as the independent variable, obtained the measured value in discrete time, and regarded the trajectory as a one-dimensional Gaussian process.

Neural networks are computing systems with interconnected nodes that operate akin to neurons in the human brain. By using algorithms, they can recognize hidden patterns and correlations in raw data, cluster and classify them, and continually learn and improve over time [33]. Neural networks are ideally suited to solving complex problems in real-life situations. They can learn and model the relationships between inputs and outputs that are nonlinear and complex; make generalizations and inferences; reveal hidden relationships, patterns, and predictions; and model highly volatile data (such as time series data) [34]. Many types of neural networks like Artificial Neural Network (ANN) [35]–[37], Back-Propagation (BP) network [38], Multiple Layer Neural Network (MLNN) [30], Convolutional Auto-Encoder Neural Network (CAENN) [31] have been considered for the prediction task. With the popularization of artificial intelligence, neural networks have also been gradually applied to the field of maritime navigation [39], [40]. Historical vessel trajectory data and trajectory characteristics are used as an input for the neural network to output the predicted vessel trajectory data [40]. The clusters achieved by the first step were used to train the artificial neural network (ANN) to predict the vessels' trajectories [41]. The results showed a 70% prediction accuracy. Tang *et al.* constructed a neural network with two long short-term memory (LSTM) layers, which can observe the first 10 min of the vessel's state to predict the location of the vessel after 20 min [28]. Tu *et al.* proposed a deep learning method that integrates multiple ship movements, which can be adapted to predict various categories of vessel trajectories after training the neural network appropriately [42]. Overall, the resulting accuracy varies as a function of ship categories, which entails a need to improve the modeling approach.

A so-called sequence-to-sequence recurrent neural network model has been developed to mesh and serialize a vessel trajectory into a neural network model to predict the main

trajectory and arrival time [43]. An LSTM model was introduced to predict the ship's position by evaluating the probability distribution, and it provides relatively accurate results [28]. To improve the accuracy of the prediction mechanisms, a multiple azimuth autonomous device sensor has been used as an additional data input; however, this approach requires a large amount of AIS data and is thus computationally expensive [44]. Although large AIS historical data sets can be used for a reference to predict maritime trajectories, good data quality is often not guaranteed and data redundancy is a major problem. At present, there are approximately 1600 AIS receivers on the coastline of more than 150 countries and 65,000 ships sailing at a time [45]. Furthermore, abnormal data due to either environmental conditions or technical problems are likely to generate significant trajectory prediction errors [46], [9]. The most appropriate balance between the use of large data sets (to train prediction mechanisms well) with the minimization of redundant and noisy data is still a crucial issue in practice [47].

The advantages of common statistical methods are that they use data that occupy less storage space during the calculation process, they can be used to realize short-term trajectory prediction, and their calculation method is relatively lightweight; these allow statistical methods to perform comparably with their deep learning counterparts. The disadvantage is that the initial state of the model and violations of the assumptions of ideal conditions greatly affect the prediction results. However, unlike deep learning, statistical methods cannot learn the effects of shallow reefs, islands, and other spatial factors on the trajectory of ships. This feature makes deep learning more practical in trajectory prediction. This motivates our search for a trajectory prediction approach that accounts for the impact of redundant and noisy data on neural network training and that optimizes the input trajectory dataset to improve the final quality of the modeling approach. This led us to combine a neural network with a Bi-LSTM framework, which is described as follows.

Murray and Perera proposed a novel dual linear autoencoder approach for predicting the trajectory of a selected vessel [29]. Forti *et al.* explored neural sequence-to-sequence models based on the LSTM encoder–decoder architecture to effectively capture long-term temporal dependencies of sequential AIS data and increase the overall predictive power [48]. Although scholars have focused on long-term track prediction (within 5–30 min) to completely prevent the occurrence of short-range collisions, vessel trajectories are affected by factors such as terrain features, ocean currents, and wind direction; therefore, making accurate predictions is difficult. Advances in artificial intelligence have facilitated the development of intelligent transportation systems; the main aim of these systems is to improve the safety and efficiency of maritime traffic. Vessel collision avoidance is one of the most important issues in marine safety. Vessel collision avoidance involves controlling the direction in which the vessel moves and obtaining reliable track prediction. Therefore, track prediction should be accurate and instantaneous.

AIS data are used to understand the historical behavior of a vessel and thus predict its trajectory. An AIS stores numerous vessel parameters in a database. The sailing mode of a vessel is easier to determine through the AIS; therefore, a prediction model using AIS data can be optimized [19], [49]. In this study, by denoising and analyzing original AIS data, we could reduce the complexity of the input prediction model data and the calculation time, thereby increasing the prediction accuracy; thus, fast and accurate vessel trajectory prediction could be achieved and the efficiency and safety of maritime traffic could be increased. The prediction performance of the proposed model was compared with that of the ETS, ARIMA, SIR, RNN, and LSTM models; the advantages and disadvantages of each model were analyzed as well.

ETS is a data-averaging method that considers three factors: the error, trend, and season [50]. Moreover, the weight of ETS-weighted data decays exponentially. The weight of the latest data is the highest, with the weights decreasing with the age of the data. However, because of the lack of calculation, a considerable gap exists between the predicted value and the observation value in ETS. The ARIMA model predicts values by examining the differences between time-series values. The ARIMA model comprises three components: AR, integrated (I), and MA components. It also includes a total of three parameters: p , d , and q . To achieve accurate prediction results, the ARIMA model must refer to a large quantity of historical data to determine its optimal parameter combination and must determine the AR (p) and MA (q) parameters through the Akaike information criterion and Bayesian information criterion. In general, a statistical model cannot be used to solve nonlinear problems easily. As displayed in Fig. 9, the ETS and ARIMA models always exhibited the highest error values in this study.

SVR is a classic machine learning method that has been successfully used for bus passenger flow prediction [51], Covid-19 case prediction in India [52], and vessel trajectory analysis [53]. SVR has three hyperparameters: the regularization parameter (C), kernel function bandwidth (σ), and ϵ -insensitive loss function (ϵ). Any changes to these parameters would considerably affect the SVR prediction accuracy. However, the automatic adjustment of the three hyperparameters in SVR remains a challenge [54], [55]. The experimental results reveal that the MAE, RMSE, and MAPE of the SVR model were lower than those of the ETS and ARIMA models when the default hyperparameters were used.

Owing to the transient memory, an RNN is suitable for modeling time-series data. An RNN maintains an excitation parameter vector for each time step, especially when short-term correlations are included in the input data. However, if gradient descent is used to train an RNN, it becomes difficult for the RNN to learn the long-term dependence in the input sequence because of the gradient vanishing problem [51], [52]. Hochreiter and Schmidhuber developed the LSTM architecture [26]; in an LSTM network, a special structure and memory unit are adopted to maintain the forward and backward transmission between layers within

a stable signal range in order to solve the problem of gradient vanishing and compensate for the inadequate long-term dependence of the RNN on vessel trajectory data feature extraction [24], [53].

An RNN has a fewer number of computations, fewer neurons, and hidden layers when compared with an LSTM network; however, an RNN has a higher error rate. Although LSTM has an outstanding feature extraction ability, its unidirectional memory propagation is not sufficient to achieve the accuracy required for trajectory prediction tasks; therefore, LSTM was not the most ideal method for this study, and its error rate was similar to that of an RNN.

The relationship between past and future time points is crucial for solving the time-series problem. The Bi-LSTM model uses its special valve structure (gate) to control memory access; this gate allows the network to remember the characteristics of long series data and obtain a model for the relationship between future data points and past data points through the bidirectional design, strengthen the original time series, and make similar predictions for continuous data. Therefore, the Bi-LSTM model can outperform an LSTM network in solving nonlinear problems and can more effectively fit a dataset; hence, it can be used in sequence analysis and provide more accurate predictions.

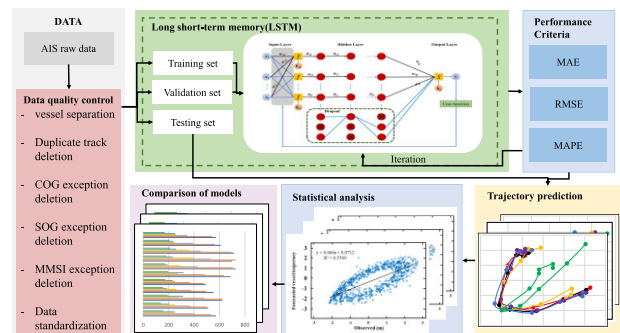


FIGURE 1. Flowchart of study method.

III. METHOD

In this study, we applied data denoising along with a Bi-LSTM model to predict vessel trajectories. The flowchart of the study method is illustrated in Fig. 1. First, data on the vessel trajectory, speed, heading, and other features were collected. Subsequently, the data were cleaned using trajectory separation, outlier deletion, and data standardization, after which the Bi-LSTM model was employed to eliminate noise in AIS trajectory prediction. The vessel trajectory predicted by the Bi-LSTM model was then evaluated, and the results were compared with those of other algorithms. Finally, the predicted trajectory was compared with the original trajectory data to verify the prediction efficiency of the algorithm.

A. COLLECTION OF DATA ON VESSEL TRAJECTORY, SPEED, COURSE, AND OTHER FEATURES

Vessel track prediction was the main objective of this study. The AIS dataset was downloaded from a public database; the

data included the MMSI, speed over ground (SOG), course over ground (COG), record time, and boat length. Finally, the difference of vessel trajectory prediction after data denoising is discussed.

B. DATA CLEANING THROUGH TRACK SEPARATION, OUTLIER DELETION, AND DATA STANDARDIZATION

Vessel trajectory data (AIS data) are typically stored in a database through data transmission and reception. Therefore, abnormal AIS data must be removed before vessel track analysis. This study enhanced the quality of the collected vessel trajectory data through data cleaning, data standardization, and deduplication.

1) TRACK SEPARATION

The raw AIS dataset contained numerous ship features along with the data of hundreds of ships. An MMSI can uniquely identify ships; therefore, it can be used to separate the AIS data samples of different ships. This helps us to improve the accuracy of neural network prediction and to establish different prediction models for different vessel trajectories. We can predict the trajectories more accurately under similar navigation modes [56].

The trajectory data of the same vessel are separated. In navigable waters, owing to the large number of ships and the constraints of the AIS working mechanism, network communication is blocked. For example, the data that should have been received 1 s ago is received after a delay of 1 s by the network. This 1 s is the time interval. Thus, the AIS cannot reserve or listen to idle time slots, causing the AIS information to be delayed and the trajectory data of a given ship to appear at larger intervals. Overall, a continuous vessel trajectory is separated based on the timestamp information of the AIS data.

2) DATA CLEANING

This study was conducted off the island of Taiwan, at a latitude between 20° and 25° N and at a longitude between 120° and 123° E. Raw AIS data collected from a single vessel were cleaned through the deletion of duplicate tracks, deletion of data abnormal COG and SOG data, deletion of abnormal MMSI number data, and standardization of the remaining data. As presented in Fig. 2, several factors can engender erroneous or noisy data. Changes in direction (latitude) are common among vessels; hence, to ensure the accuracy of captured data, the moving average (MA) method should be used for location data standardization during the cleaning process. Moreover, manual inspection of the original AIS data revealed abnormalities in the data regarding the speed of the vessels, with some of the speeds exceeding >30 knots. Such pieces of abnormal data were also deleted. To ensure vessel safety, vessel speeds should be standardized for specific navigation conditions. On the basis of the average navigation direction of vessels, the average navigation direction of vessels traveling in a region can be predicted.

3) DATA STANDARDIZATION

During data preprocessing, wandering or anchoring trajectories in the original dataset were eliminated. We set the minimum time interval of the trajectory to 1200 s because the AIS information receiving interval is generally specified to be 5–10 min and because an information interval higher than 20 min is used in the next stage of the navigation status [28]. Many noisy data points in the original data were eliminated. The original ship dataset contained data on many floating and anchored ships. The speed of these ships is affected by wind and ocean currents and is often less than 1 knot, which is an abnormal navigation state. The data required for the experiment were for sailing ships; accordingly, a smooth route was obtained along the motion trajectory, making each trajectory easier to analyze. Because the time interval of the ship track point acquisition depends on the ship, the ship track can be regarded as a continuous time series. When encountering outliers, the MA model was used to standardize the data series. Taking Figure 2.(a) as an example, when we find that an abnormal SOG value suddenly appears, MA processing is performed on two adjacent data, and the obtained value is used to replace the original abnormal value [47].

C. APPLICATION OF BI-LSTM TO AIS-BASED TRAJECTORY PREDICTION WITH NOISE ELIMINATION

1) LSTM NETWORK

The Bi-LSTM model was applied to model individual navigation features by using the denoised dataset. Consequently, the accuracy of vessel trajectory prediction could be increased. An LSTM network is a special RNN that is suitable for analyzing sequential data [57], [58]. In an LSTM network, a memory cell replaces the hidden layer function that is present in a traditional RNN. An LSTM unit usually comprises a memory cell, forget gate, input gate, and output gate. The LSTM unit is used to enhance the long-term memory ability of an LSTM model and to resolve the long-term dependence problem. Generally, an LSTM network consists of at least one unidirectional LSTM layer and one LSTM unit. This study combines two bidirectional LSTM layers and thus has four LSTM units. The formula for an LSTM network can be expressed as follows:

At time t , x_t is the input data of the LSTM unit, h_t is the output of the LSTM unit, h_{t-1} is the output of the LSTM unit at $t-1$, and C_t is the value of the memory unit. The operating procedures of an LSTM network are outlined as follows:

- 1) The value of the forget gate F_t is calculated. The forget gate controls the update of the historical data to the state value of the memory unit, where W_f represents the weight matrix and b_f represents the bias.

$$F_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

- 2) The value of the candidate memory unit \tilde{C}_t is calculated, where W_c represents the weight matrix and b_c represents the bias.

$$\tilde{C}_t = \sigma(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2)$$

MMSI	SOG	COG	Longitude	Latitude	Record_Time
133506759	8.3	264.8	121.436613	27.567739	2019-06-03 10:07:33.56
133506759	8.1	258.5	121.437027	27.560658	2019-06-03 10:09:17.83
133506759	7.9	258.3	121.434955	27.554104	2019-06-03 10:09:26.41
133506759	8.2	253.2	121.431071	27.556089	2019-06-03 10:10:54.62
133506759	8.2	259.5	121.436559	27.558083	2019-06-03 10:11:51.05
133506759	2.3	313.9	121.312346	27.694966	2019-06-03 10:12:43.46
133506759	8.5	261.1	121.433695	27.557933	2019-06-03 10:14:05.73
133506759	8.6	258.8	121.428682	27.551047	2019-06-03 10:14:16.35
133506759	8.6	255.7	121.423309	27.554338	2019-06-03 10:15:38.16
133506759	8.3	259.5	121.422358	27.558220	2019-06-03 10:17:19.81
133506759	1.0	318.2	121.313826	27.701323	2019-06-03 10:18:13.45
133506759	8.4	257.3	121.421157	27.556812	2019-06-03 10:18:24.33

(a)

MMSI	SOG	COG	Longitude	Latitude	Record_Time
416671530	6.3	84.9	119.733412	23.516139	2019-06-03 16:30:08.57
416671530	6.3	87.1	119.733625	23.516742	2019-06-03 16:33:24.85
416671530	6.5	88.5	119.734055	23.517332	2019-06-03 16:37:31.49
416671530	6.6	84.6	119.734571	23.517917	2019-06-03 16:41:55.64
416671530	6.3	381.2	119.735159	23.517993	2019-06-03 16:44:27.07
416671530	6.2	85.4	119.736045	23.518626	2019-06-03 16:46:06.48
416671530	6.2	84.6	119.737295	23.519440	2019-06-03 16:49:15.79
416671530	6.4	85.1	119.737823	23.519732	2019-06-03 16:52:53.34
416671530	6.5	88.2	119.738309	23.520309	2019-06-03 16:55:42.13
416671530	6.2	86.5	119.739654	23.520658	2019-06-03 16:56:38.84

(b)

MMSI	SOG	COG	Longitude	Latitude	Record_Time
373698000	12.7	193.5	118.372753	24.382658	2019-06-22 09:42:31.52
373698000	12.3	193.7	118.372722	24.421853	2019-06-22 09:44:26.81
373698000	11.8	197.6	118.370671	24.475931	2019-06-22 09:45:17.64
373698000	11.8	199.4	118.369780	24.515264	2019-06-22 09:47:35.25
373698000	11.8	199.4	118.369780	24.515264	2019-06-22 09:48:12.95
373698000	11.8	199.4	118.369780	24.515264	2019-06-22 09:52:51.11
373698000	12.4	188.5	118.349183	25.622649	2019-06-22 09:53:08.05
373698000	12.1	174.3	118.347157	25.650326	2019-06-22 09:55:16.34
373698000	11.9	185.3	118.344215	25.698513	2019-06-22 09:59:33.68
373698000	12.3	191.1	118.341642	25.715005	2019-06-22 10:02:42.79
373698000	12.3	191.1	118.341642	25.715005	2019-06-22 10:07:13.15
373698000	11.7	193.7	118.335018	25.795624	2019-06-22 10:08:24.83

(c)

MMSI	SOG	COG	Longitude	Latitude	Record_Time
2722	6.5	211.2	120.165783	28.342658	2019-05-16 14:20:26.07
2722	6.5	217.3	120.162795	28.341868	2019-05-16 14:26:44.52
2722	6.7	227.4	120.175638	28.375961	2019-05-16 14:30:10.49
2722	6.8	219.4	120.175729	28.395282	2019-05-16 14:34:08.35
2722	7.1	224.6	120.164328	28.392661	2019-05-16 14:40:52.44
2722	7.1	205.8	120.169539	28.397765	2019-05-16 14:46:41.82
2722	7.3	212.6	120.143389	28.372635	2019-05-16 14:51:57.76
2722	7.1	211.4	120.151105	28.370334	2019-05-16 14:55:06.46
2722	6.9	209.9	120.156244	28.378598	2019-05-16 14:59:31.25
2722	6.6	206.6	120.146659	28.365067	2019-05-16 15:05:28.13

(d)

FIGURE 2. Four types of erroneous and noisy effects. (a) Abnormalities exist in the data in the SOG column. Generally, it is difficult for a vessel to accelerate or decelerate very quickly over a short period (it would require 52 s to slow down from 8.2 to 2.3 knots); therefore, the values highlighted are deemed abnormal. (b) Abnormalities exist in the data in the COG column. COG values range from 0 to 360. (c) Duplicate data, which are not conducive to neural network training and may cause overfitting, are removed. (d) MMSI column is abnormal. The standard MMSI length should be 9.

- 3) The value of the input gate I_t is then calculated. The input gate controls the update of the current input data to the state value of the memory unit, where σ represents the sigmoid function, W_i represents the weight matrix, and b_i represents the bias.

$$I_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

- 4) The value of the current memory unit C_t is calculated; C_{t-1} represents the state value of the previous LSTM unit.

$$C_t = F_t \times C_{t-1} + I_t \times \tilde{C}_t \quad (4)$$

- 5) The value of the output gate O_t is calculated. The output gate controls the output of the state value of the memory unit, where W_o represents the weight matrix and b_o represents the bias.

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

- 6) The output of the LSTM unit h_t is calculated, where \tanh is a nonlinear activation function. It converges the permissible amplitude range of the output signal to a finite value. The function is expressed as follows:

$$h_t = O_t \times \tanh(C_t) \quad (6)$$

The three control gates and memory cell of an LSTM unit facilitate the process of maintaining, resetting, and updating

long-term information. Because of the weight-sharing mechanism in LSTM, the number of dimensions can be controlled by setting the weight matrix. In an LSTM unit, a long delay exists between forward and back propagation because the internal state of the memory cell in the LSTM structure maintains a constant data size, reducing the probability of gradient explosion and gradient vanishing.

2) BI-LSTM NETWORK

In this study, we designed a Bi-LSTM model to predict the trajectory of vessels; this network model computes input vectors containing information about the past and future within a specific time range (Fig. 3). In the proposed method, a regular LSTM state neuron is divided into two parts: one part is responsible for the positive time direction (forward state) and the other is responsible for the opposite time direction (backward state). Our aim was to develop a network that can outperform an RNN. We experimentally tested this network model by dividing the training set into several batches, with each batch containing continuous vessel trajectories; the number of trajectories was determined by the batch size. Studies have shown that a neural network with a batch size of 64 achieves the best prediction results. Orthogonal initialization was employed to initialize the weight and deviation settings, and each model was fitted to training data by 2000 iterations. The learning rate was set to 0.001, dropout was set to [0.2, 0.2], and number of Bi-LSTM neurons was

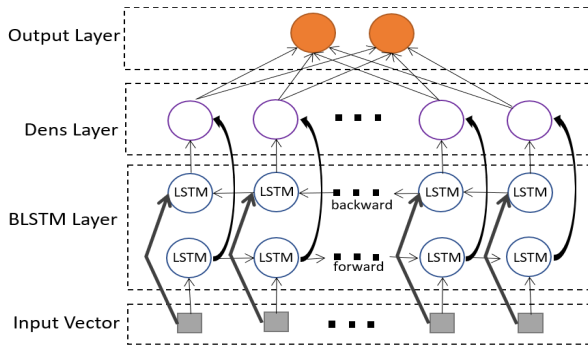


FIGURE 3. Network structure of Bi-LSTM.

set to [200, 100]. The processes involved in the prediction of vessel trajectory are described subsequently.

An AIS sequence after noise elimination is represented as $M_N = \{x, y, lat, lon, s, c\}$, where N is the total length of the sequence. M_N is divided into a training set and a test set, with the length of the training set being t and the length of test set being $N - t$; (lat, lon, s, c) is the input vector to the Bi-LSTM model, and (x, y) is the output vector. The Bi-LSTM model includes three layers: an input layer, a hidden layer, and an output layer. Trajectory data (lat, lon, s, c) are input to the input layer, and the output layer outputs the vector (x, y) . The hidden layer consists of an LSTM layer and a fully connected layer. The model training process includes forward propagation and backpropagation. In forward propagation, the prediction model between the input vector (lat, lon, s, c) and output vector (x, y) is established. First, the training samples (lat, lon, s, c) are divided into several groups according to the batch size and sent to the input layer. Subsequently, they are propagated through the Bi-LSTM layer according to (1)–(6) to obtain the output h_t . Finally, the prediction offset (x_p, y_p) is derived using the fully connected layer. During backpropagation, the prediction model is optimized by adjusting the weight parameters. The error between the actual vectors $(x$ and $y)$ and the predicted offsets $(x_p$ and $y_p)$ is calculated using the loss function (7). The weight parameters are adjusted by minimizing the loss value to improve the prediction accuracy of the model. The loss function is expressed as follows:

$$Loss = \frac{\sum_{i=1}^n |x_i^o - x_i^p|}{n} \quad (7)$$

where t denotes the time step, n denotes the length of the test set, x_i^o denotes the observation value, and x_i^p denotes the predicted value.

3) DROPOUT

Overfitting affects the statistical noise in the training data for model learning, resulting in the poor timeliness of the evaluation model on the new data (test set). Due to overfitting, generalization error also increases. Dropout is a regularization method that is similar to training a large number of neural networks with different structures in parallel. In the training process, some neuron outputs are ignored or

discarded randomly, making the hidden layer appear similar to a new network structure with a different number of neurons to reduce overfitting.

During neural network training, the prevention of overfitting is crucial. Srivastava *et al.* proposed the dropout method to prevent overfitting in neural network training [59]. During the training process between layers in a neural network, the dropout method randomly drops some neurons with a certain probability, as displayed in Fig. 4.

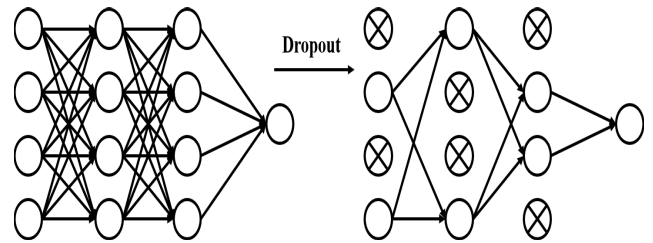


FIGURE 4. Flowchart of dropout.

D. EVALUATION OF THE ALGORITHM AND COMPARISON WITH OTHER MODELS

To increase the accuracy of vessel trajectory prediction and verify the performance of the data denoising process and the Bi-LSTM model, we compared the error rate of our vessel trajectory prediction algorithm with those of common time-series prediction methods in the literature, including exponential smoothing (ETS) [50], autoregressive (AR) integrated MA (ARIMA) [60], support vector regression (SVR) [61], RNN [62], and LSTM [58] algorithms.

1) ETS ALGORITHM

The ETS algorithm proposed by Brown in 1956 is a classic time-series prediction method [50]. ETS predicts the future value of the time series by using the weighted average of past observations of the time series. This method gives decreasing weights to past observations and higher weights to more recent observations. This framework enables reliable estimates to be produced quickly in most applications. The simple exponential smoothing model was first extended to seasonality components by Winters in 1960 [63], after which Holt used it to determine trends [64]. In these models, the trend components can be multiplicative (m), multiplicative damped (MD), additive (a), additive damped (AD), or absent (n), and the seasonality components can be multiplicative (m), additive (a), or absent (n). The simple ETS formula is as follows:

$$S_t = Wa \cdot y_t + (1 - a)S_{t-1} \quad (8)$$

where S_t denotes the smoothing value at time t , y_t denotes the observation value at time t , S_{t-1} denotes the smoothing value at time $t - 1$, a denotes the smoothing constant in the range [0,1], and S_t denotes the weighted arithmetic mean of y_t and S_{t-1} . The influence of y and S_{t-1} on S_t can be determined by changing a . When $a = 1$, $S_t = y_t$. When $a = 0$, $S_t = S_{t-1}$.

2) ARIMA MODEL

Autoregression is a statistical method used to analyze the relationship between a single variable and a group of independent variables. It is usually employed for time-series forecasting. However, the autoregression can only be used in models that can be fitted with extracted time-series data. Therefore, to obtain favorable experimental results, a large quantity of data must be collected. The ARIMA model was proposed by Box and Jenkins in 1976; this model is also known as the Box–Jenkins model [60]. The ARIMA model can be fitted to a time series to better understand the future value in the series; it is a simple linear method. The ARIMA model consists of two parts: AR and MA components.

3) SVR MODEL

On the basis of the structural risk minimization principle proposed by Vapnik, a loss function ε was derived; subsequently, an SVR model was developed to solve nonlinear problems, especially time-series prediction problems. SVR was proposed by Vapnik *et al.* in 1997 [61] and has been used in many forecasting tasks such as short-term load forecasting [65] and monthly rainfall forecasting [66]. To obtain good forecasting performance, all three hyperparameters (C , ε , and σ , a kernel parameter) of the SVR model must be determined. These hyperparameters are usually determined through data resampling, which is computationally time consuming. Thus, an efficient approach to simultaneously determine all parameters is necessary.

4) RNN MODEL

Based on Rumelhart's work in 1986, the aim of which was to extract long-term dependency in sequential data [67], [68], an RNN has a unique memory unit that allows it to be used in short sequence prediction; however, in practical applications, the length of the problem sequence is not known, which may lead to gradient vanishing or gradient explosions during the learning process. Therefore, the practicability of an RNN is limited. However, variants of RNNs exist, such as LSTM [26] and gate recurrent unit [69] models. A simple RNN has only one internal memory unit h_t , which is represented by (16):

$$h_t = f(Wx_t + U_f h_{t-1} + b) \quad (9)$$

where f is the activation function, U and W are the weight matrices of the hidden layer, b is the bias, and x_t is the input vector at time t [70].

E. CRITERIA

The mean absolute error (MAE), root-mean-square error (RMSE), and mean absolute percentage error (MAPE) are typically used to determine the deviations of observed values from predicted values for evaluating the prediction performance of models. These metrics were used in this study to evaluate the performance of the proposed method. Specifically, assuming y_i is the ground truth and \hat{y}_i^j is the predicted

value at time t , they are defined as follows:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N \left| \hat{y}_t^i - y_t^i \right| \quad (10)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_t^i - y_t^i)^2} \quad (11)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{y}_t^i - y_t^i}{y_t^i} \right| \times 100\% \quad (12)$$

IV. RESULTS AND DISCUSSION

A. DATA COLLECTION

To verify the efficiency and effectiveness of the Bi-LSTM model, we recorded data from more than 300 vessels in the waters around Taiwan over a specific period (Fig. 5). The vessel trajectory was randomly selected from the AIS dataset between 00:00:00 on July 6, 2019, and 03:00:00 on July 6, 2019. The MMSI numbers of the vessels were 249558000, 229069000, 357402000, 305097000, 249290000, 271045019, 351296000, 671115100, 667001698, and 636014592. Because of the equipment and signal drift, information on the longitude, latitude, direction, and vessel speed in AIS dataset changes considerably.

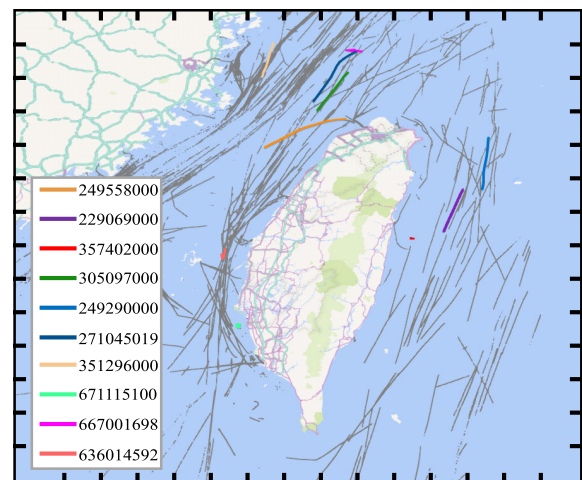


FIGURE 5. UTM zone map and data source location.

To ensure data availability, the data had to be preprocessed in this study. The transcoded and preprocessed data are presented in Fig. 6. For trajectory training, the data of 1,364 vessels with high activity frequency were selected from the dataset, and 1-month navigation track data were obtained; a total of 1,048,575 track points were obtained. In Fig. 6, Lon represents longitude, Lat represents latitude, and Record_time represents the track point recording time. Table 1 lists the azimuths corresponding to the COG value in the AIS data, and Table 2 lists the speed status corresponding to the SOG value in the AIS data.

B. NETWORK PARAMETER SETTINGS

In this study, Keras was used to help develop the Bi-LSTM model, and the Adaptive Moment Estimation (Adam) algorithm that optimizes random objective functions was used to

MMSI	Lon	Lat	COG	SOG	Record_Time
249558000	121.14959330	25.25619833	263.2	14.2	2019/7/6 00:00
249558000	121.14138000	25.25523500	262.6	14.2	2019/7/6 00:01
249558000	121.14138000	25.25523500	262.6	14.2	2019/7/6 00:02
249558000	121.13284170	25.25422333	262.5	14.1	2019/7/6 00:03
249558000	121.13284170	25.25422333	262.5	14.1	2019/7/6 00:04
249558000	121.12478830	25.25334500	263.5	14.1	2019/7/6 00:05
249558000	121.12002000	25.25288333	264.5	14.1	2019/7/6 00:06
249558000	121.12002000	25.25288333	264.5	14.1	2019/7/6 00:07
249558000	121.10721330	25.25180667	264.2	14.0	2019/7/6 00:09
249558000	121.10721330	25.25180667	264.2	14.0	2019/7/6 00:10
249558000	121.09864830	25.25105000	264.6	14.0	2019/7/6 00:11
249558000	121.09864830	25.25105000	264.6	14.0	2019/7/6 00:12
249558000	121.09004500	25.25030833	264.5	14.0	2019/7/6 00:13
249558000	121.09004500	25.25030833	264.5	14.0	2019/7/6 00:14
249558000	121.08147170	25.24961167	265.3	14.1	2019/7/6 00:15
249558000	121.08147170	25.24961167	265.3	14.1	2019/7/6 00:16
249558000	121.07291830	25.24898833	264.1	14.1	2019/7/6 00:17
249558000	121.07291830	25.24898833	264.1	14.1	2019/7/6 00:18
249558000	121.06446500	25.24819833	263.2	14.0	2019/7/6 00:19
249558000	121.064465	25.24819833	263.2	14	2019/7/6 00:20

FIGURE 6. Transcoded AIS Data.

TABLE 1. List of COG statuses.

Course Over Ground (COG)	Statuses
337.5~360 & 0~22.5	North
22.5~67.5	Northeast
67.5~112.5	East
112.5~157.5	Southeast
157.5~202.5	South
202.5~247.5	Southwest
247.5~292.5	West
292.5~337.5	Northwest

TABLE 2. List of SOG statuses.

Speed Over Ground (SOG)	Statuses
0~3	Slow
3~14	Medium
14~23	High
23~99	Very High
Over 99	Exception

update the parameters in the training process. Adam dynamically adjusts the learning rate of each parameter according to the loss function (MAE, RMSE, and (7)) and the first-order matrix estimation and second-order moment estimation of the gradient of each parameter. Moreover, this algorithm is based on the gradient descent algorithm [71], but a specific range is specified for the learning step of each parameter; additionally, a larger gradient would not lead to an excessively large learning step, and this ensures a relatively stable parameter value. The learning rate was set to 0.0001, and the decay_rate was set to 0.5, which indicates the decay_rate in the learning rate after each update of the parameters. In addition, to prevent model overfitting, the dropout algorithm was employed. During the training process, the weight of neurons was randomly reset to 0, rendering the neurons ineffective. Setting the weight to 0.1 would signify a 10% dropout probability for each neuron. Because the training dataset was relatively small, we employed a batch size of 1000 in this experiment.

We used 80% of the dataset as the training data and used the remaining 20% as the test data.

C. RESULTS OF THE PREDICTION OF MULTIPLE VESSEL TRAJECTORIES

Table 3 presents a comparison of the vessel trajectory prediction results of the proposed Bi-LSTM model and those of five other models. The five models were the ETS, ARIMA, SVR, RNN, and LSTM models.

In this study, the trajectories of ten vessels (vessels 1–10) randomly selected from the AIS data were predicted. The MAE, MAPE, and RMSE were used as the criteria for model validation. The experimental results revealed that among the models, the Bi-LSTM model had the lowest prediction error for vessels 1–10. The MAPE of the Bi-LSTM model was lower than those of the ETS, ARIMA, SVR, RNN, and LSTM models by 75.8%, 74.5%, 66.9%, 46.0%, and 44.8%, respectively. The Wilcoxon-signed rank test was used to test for statistical significance. The results also indicated that the prediction performance of the Bi-LSTM model was significantly higher than that of the other models.

To better understand the effectiveness of track prediction, we can consider the results for vessel 1 as an example. As illustrated in Fig. 7, the predicted values (red dots) of ETS and ARIMA deviated considerably from the observed values (blue dots); this can be attributed to the smoothing coefficient of ETS and the problem in optimizing the ARIMA hyperparameters p , d , and q . Moreover, the statistical model requires considerable preprocessing before the data are input into the model to ensure that the input sequence is stationary and to increase the prediction accuracy considerably. The prediction performance of the SVR model was superior to that of the statistical model; however, the hyperparameter optimization process remains a challenge. LSTM is a network structure optimized by an RNN. The prediction accuracy of the LSTM network was expected to be higher than that of the RNN. However, as presented in Table 3, we did not observe considerable difference between the prediction accuracies of the LSTM and RNN models. The proposed Bi-LSTM model exhibited a considerable improvement in prediction performance. The superior prediction performance of the Bi-LSTM model can be attributed to the bidirectional structure, which strengthens the feature extraction process in a sequence.

In the present study, Vessels 1 to 6 were randomly selected for trajectory prediction. Because of the inertia of vessels, finding the complex trajectories of heavy vessels in the ocean is a difficult task. To verify the effectiveness of the proposed method, Vessels 7 to 10, which had a turning nature, were added. The experimental results revealed that accurate trajectory predictions were more difficult to achieve with the other methods (MAPE > 10). The proposed Bi-LSTM yielded favorable prediction results for the ten vessels (including those with complex trajectories), and a p-value test revealed a significant difference between the results of the proposed method and those of other methods (Fig. 8).

TABLE 3. Vessel trajectory prediction errors for different models for ten vessels.

VESSEL	MMSI	Criteria	ETS	ARIMA	SVR	RNN	LSTM	Bi-LSTM
1	249558000	MAE	552.11	517.85	389.21	272.81	258.37	162.7
		MAPE(%)	37.07	34.60	26.42	17.22	16.04	10.18
		RMSE(m)	571.00	533.00	406.00	255.19	247.00	157.00
2	229069000	MAE	503.18	52.03	364.00	237.00	228.00	151.29
		MAPE(%)	33.24	32.39	24.11	16.19	13.31	9.04
		RMSE(m)	512.00	498.00	371.00	249.81	205.00	139.00
3	357402000	MAE	609.00	622.00	504.00	269.00	313.00	190.00
		MAPE(%)	40.32	40.58	31.42	18.85	20.32	9.57
		RMSE(m)	621.00	625.00	483.00	278.13	313.00	194.00
4	305097000	MAE	545.00	524.00	378.00	296.00	224.00	188.00
		MAPE(%)	35.90	32.04	24.55	15.63	12.21	9.13
		RMSE(m)	553.00	493.00	378.00	261.04	188.00	156.00
5	249290000	MAE	518.27	492.00	353.72	261.43	225.03	186.94
		MAPE(%)	31.88	30.73	23.84	15.94	15.06	8.29
		RMSE(m)	491.00	473.00	367.00	249.18	232.00	174.00
6	271045019	MAE	703.26	627.90	471.07	287.00	294.00	177.68
		MAPE(%)	44.67	41.22	32.49	16.07	20.65	7.65
		RMSE(m)	688.00	635.00	482.00	273.48	318.00	195.00
7	351296000	MAE	731.84	712.47	463.05	326.81	206.87	175.42
		MAPE(%)	40.56	38.77	30.62	20.19	12.04	8.15
		RMSE(m)	726.51	704.48	429.79	288.97	191.68	170.53
8	671115100	MAE	723.79	734.58	557.29	426.04	236.58	190.65
		MAPE(%)	45.36	47.11	35.26	25.13	14.34	8.77
		RMSE(m)	699.30	726.52	535.41	394.81	231.17	183.60
9	667001698	MAE	723.16	699.07	465.52	362.3	268.49	222.03
		MAPE(%)	46.21	45.20	30.51	23.75	16.09	9.23
		RMSE(m)	712.5	694.83	453.69	359.7	259.83	211.79
10	636014592	MAE	616.79	576.49	424.27	255.83	251.96	218.47
		MAPE(%)	39.89	36.58	26.43	14.77	14.02	9.64
		RMSE(m)	610.22	563.49	393.24	247.81	235.67	206.81
Average	Average	MAE	571.80*	472.63	410.00*	270.54*	257.07*	176.10
		MAPE(%)	37.18*	35.26*	27.13*	16.65*	16.27*	8.98
		RMSE(m)	573.00*	543.00*	347.00*	261.14*	250.50*	169.17

MAE, mean absolute error; MAPE, mean absolute percentage error; RMSE, root mean square error; Boldface, the lowest values in each row.

ETS, exponential smoothing; ARIMA, autoregressive integrated moving average; SVR, support vector regression; RNN, recurrent neural network; LSTM, long short-term memory. Bi-LSTM, bidirectional long short-term memory. * $p < .05$, ** $p < .01$, *** $p < .001$.

D. STATISTICAL DESCRIPTION

To verify the advantages of the proposed Bi-LSTM model, a Wilcoxon signed-rank test was used to compare the MAPE and RMSE values obtained for the ETS, ARIMA, SVR, RNN, LSTM models with those obtained for the Bi-LSTM model. The Wilcoxon signed-rank test is a nonparametric statistical hypothesis test used to compare two related samples, matched samples, or repeated measurements on a single sample to assess whether their population mean ranks differ (i.e., it is a paired difference test). All statistical analyses were conducted using Python. As presented in Table 3, the prediction accuracy of the Bi-LSTM model improved by 75.8%, 74.5%, 66.9%, 46.0%, and 44.8% when compared with that of the

ETS, ARIMA, SVR, RNN and LSTM models, respectively, and the differences were statistically significant (p value less than .05).

Fig. 9 shows the predicted track values generated by the Bi-LSTM, LSTM, RNN, SVR, ARIMA and ETS models during the prediction period. The results indicated that the average MAPE of the Bi-LSTM model was lower than those of the LSTM, RNN, SVR, ARIMA, and ETS models by 44.8%, 46.0%, 66.9%, 74.5%, and 75.8%, respectively; the average RMSE of the Bi-LSTM model was lower than those of the LSTM, RNN, SVR, ARIMA, and ETS models by 32.4%, 35.2%, 51.2%, 68.8%, and 70.4%, respectively; and the average MAE of the Bi-LSTM model was lower than

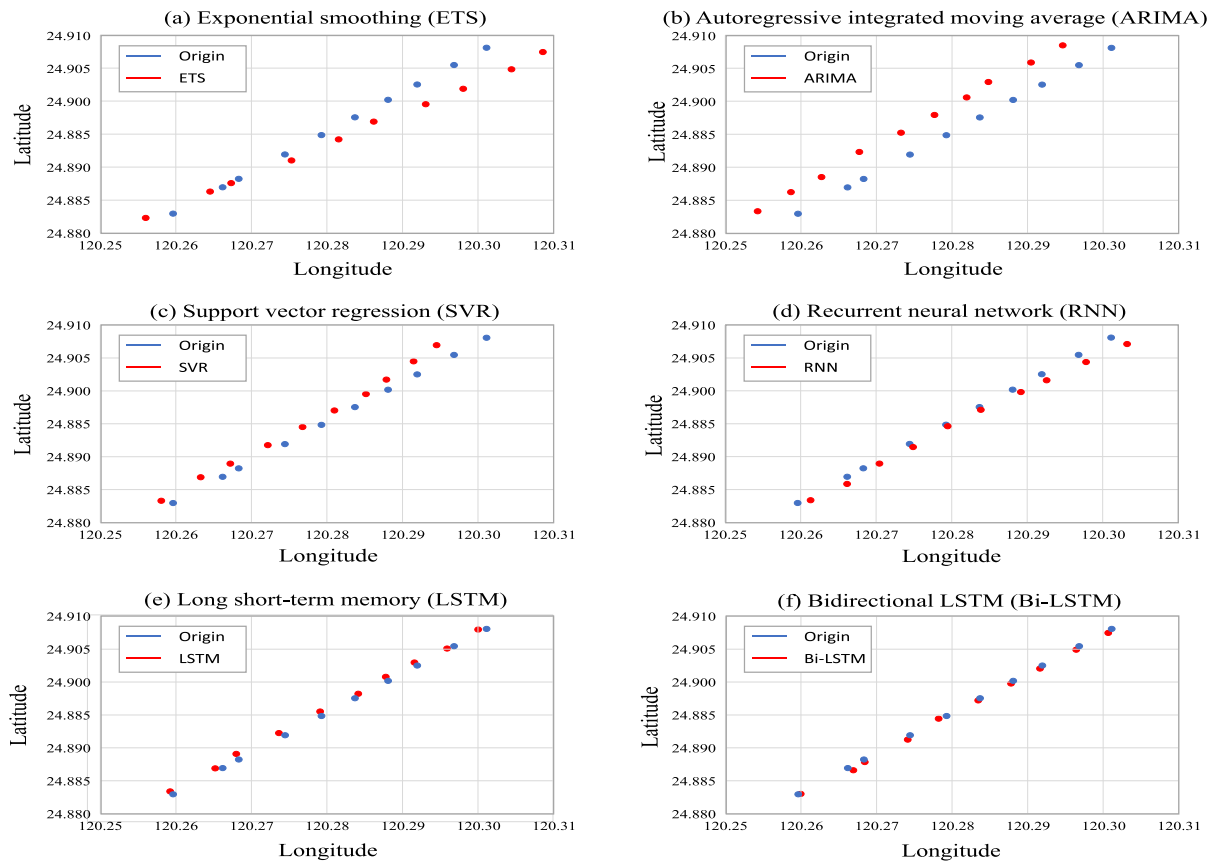


FIGURE 7. Vessel trajectory prediction result for vessel 1 using six models: (a) ETS, (b) ARIMA, (c) SVR, (d) RNN, (e) LSTM, and (f) Bi-LSTM.

those of the LSTM, RNN, SVR, ARIMA, and ETS models by 31%, 34%, 57%, 62%, and 69%, respectively. Accordingly, the Bi-LSTM model was superior to the ETS, ARIMA, SVR, RNN, and LSTM models in all aspects of vessel trajectory prediction. The Bi-LSTM model exhibited the best prediction accuracy; the experimental results indicate that the Bi-LSTM model can provide more robust predictions with lower error rates than can other models.

E. STUDY OBJECTIVE

Accurate vessel route prediction is crucial for maritime traffic control and management. In addition to collision avoidance, vessel route prediction can help in planning routes in advance, shortening the sailing distance, and increasing navigation efficiency. In this study, we proposed a new AIS data denoising and prediction method. The main contributions of this study are outlined as follows: (1) cleaning of the original AIS data through track separation, outlier deletion, and data standardization; (2) development of a Bi-LSTM model for accurately predicting vessel trajectories by using denoised AIS data; and (3) verification of the performance of the Bi-LSTM model by using the trajectory data of ten vessels and by comparing its performance with those of other common prediction models in the literature. Moreover, the study

findings can help vessel traffic controllers to predict the accurate track of vessels, which can enable them to take preventive measures to avoid collisions and improve the efficiency and safety of maritime traffic.

Notably, the method proposed in this study focuses on the trajectory of large ships moving across the ocean. Because of its weight and inertia, a ship's trajectory is straightforward and easy to predict. When the method is applied to small- and medium-sized ships, the high mobility and high speed of small ships must be considered. During the training of the model, we must focus more on the effects of COG and SOG changes on trajectory prediction. The in-depth learning prediction model is also a key method for achieving improvements. It improves not only the prediction accuracy but also the operation speed, both of which contribute considerably to reducing maritime collisions. Moreover, the proposed models exhibit applicability not just for vessel trajectory prediction but also for other applications (e.g., prediction of hurricane, human, and animal movement trajectories). Future studies should include contextual information regarding similarity measures and location predictions of trajectories. Environmental factors such as wind, wave, and weather conditions have considerable effects on the movement of vessels and may cause them to deviate from their normal paths.

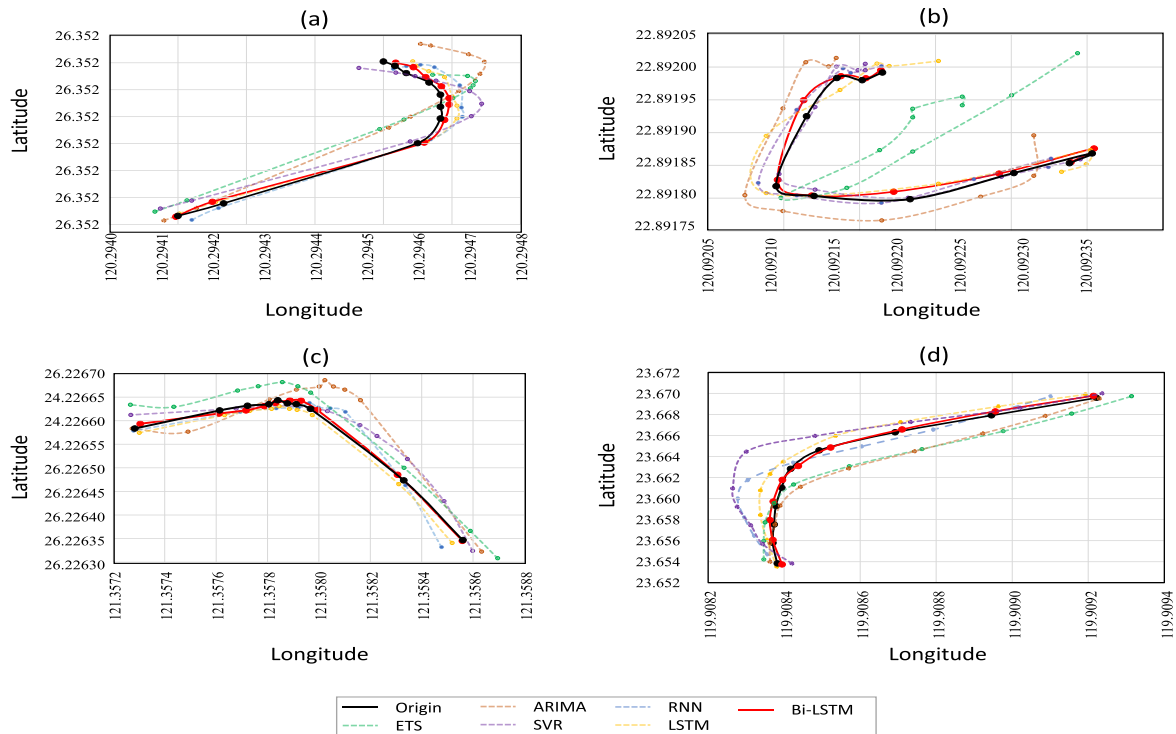


FIGURE 8. Vessel trajectory prediction result for vessel 7-10. (a) vessel 7. (b) vessel 8. (c) vessel 9. (d) vessel 10.

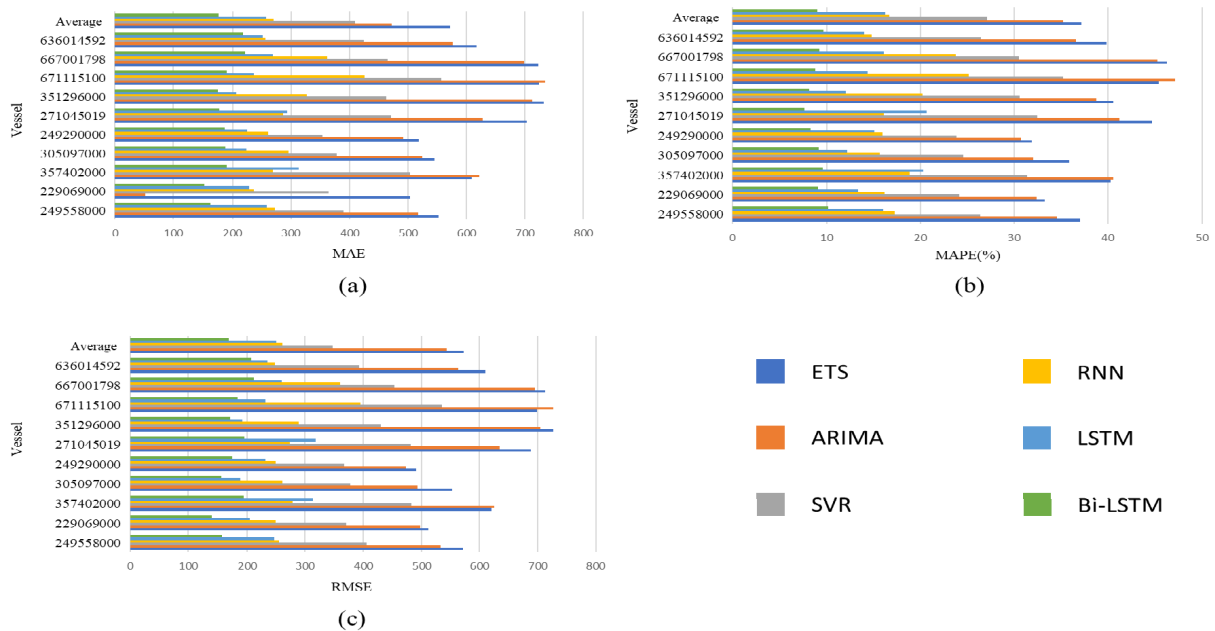


FIGURE 9. Vessel trajectory prediction error via AIS data using ETS, ARIMA, SVR, RNN, LSTM, and Bi-LSTM. (a) MAE; (b) MAPE; and (c) RMSE.

V. CONCLUSION

In this study, a method for predicting vessel trajectory using AIS data denoising and a Bi-LSTM model was proposed. The model can accurately predict the trajectory sequence and be used for short-term prediction. We used data on the trajectory

of ten vessels to evaluate our model. Our model was found to outperform the ETS, ARIMA, SVR, RNN, and LSTM models, demonstrating its suitability for vessel trajectory prediction. The proposed model can encode, delete outliers, standardize data, and effectively extract features of historical

vessel tracks through data denoising. Furthermore, the proposed model can extract features and apply them to vessel track point prediction, which helps considerably reduce errors in vessel trajectory prediction. Thus, the study findings can help maritime traffic users predict the accurate trajectory of vessels, which can enable them to take preventive measures to avoid collisions and improve the efficiency and safety of maritime traffic.

REFERENCES

- [1] *Casualty Returns*, LLloyd's Register of Shipping, London, U.K., 2019.
- [2] O. A. C. Coast Guard Administration. (2019). *2019 Coast Guard Annual Statistical Report*. [Online]. Available: https://www.cga.gov.tw/GipOpen/wSite/lp?ctNode=11351&mp=999&idPath=600_604_8755_8794_11351
- [3] G. Praetorius, "Vessel traffic service (VTS): A maritime information service or traffic control system?: Understanding everyday performance and resilience in a socio-technical system under change," Ph.D. dissertation, Chalmers Tekniska Högskola, Gothenburg, Sweden, 2014.
- [4] A. Bakdi, I. K. Glad, E. Vanem, and Ø. Engelhardt, "AIS-based multiple vessel collision and grounding risk identification based on adaptive safety domain," *J. Mar. Sci. Eng.*, vol. 8, no. 1, p. 5, Dec. 2019.
- [5] J. M. Mou, C. van der Tak, and H. Ligteringen, "Study on collision avoidance in busy waterways by using AIS data," *Ocean Eng.*, vol. 37, nos. 5–6, pp. 483–490, 2010.
- [6] X. Chen, J. Ling, Y. Yang, H. Zheng, P. Xiong, O. Postolache, and Y. Xiong, "Ship trajectory reconstruction from AIS sensory data via data quality control and prediction," *Math. Problems Eng.*, vol. 2020, Aug. 2020, Art. no. 7191296.
- [7] A. Norris, "AIS implementation—success or failure?" *J. Navigat.*, vol. 60, no. 1, p. 1, 2007.
- [8] L. Etienne, T. Devoegele, and A. Bouju, "Spatio-temporal trajectory analysis of mobile objects following the same itinerary," *Adv. Geo-Spatial Inf. Sci.*, vol. 10, pp. 47–57, Jul. 2012.
- [9] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 6, pp. 2218–2245, 2013.
- [10] H. Rong, A. P. Teixeira, and C. Guedes Soares, "Data mining approach to shipping route characterization and anomaly detection based on AIS data," *Ocean Eng.*, vol. 198, Feb. 2020, Art. no. 106936.
- [11] P.-L. Sanchez-Gonzalez, D. Díaz-Gutiérrez, T. Leo, and L. Núñez-Rivas, "Toward digitalization of maritime transport?" *Sensors*, vol. 19, no. 4, p. 926, Feb. 2019.
- [12] I. Jurdana, A. Krylov, and J. Yamnenko, "Use of artificial intelligence as a problem solution for maritime transport," *J. Mar. Sci. Eng.*, vol. 8, no. 3, p. 201, Mar. 2020.
- [13] K. Shah, S. Koprü, and J. D. Ruvini, "Neural network based extreme classification and similarity models for product matching," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics: Hum. Lang. Technol.*, vol. 3, 2018, pp. 8–15.
- [14] L. P. Perera and B. Mo, "Ship performance and navigation data compression and communication under autoencoder system architecture," *J. Ocean Eng. Sci.*, vol. 3, no. 2, pp. 133–143, Jun. 2018.
- [15] S. Sun, Y. Chen, Z. Piao, and J. Zhang, "Vessel AIS trajectory online compression based on scan-pick-move algorithm added sliding window," *IEEE Access*, vol. 8, pp. 109350–109359, 2020.
- [16] J. Liu, H. Li, Z. Yang, K. Wu, Y. Liu, and R. W. Liu, "Adaptive douglas-peucker algorithm with automatic thresholding for AIS-based vessel trajectory compression," *IEEE Access*, vol. 7, pp. 150677–150692, 2019.
- [17] L. Qi and Z. Zheng, "Trajectory prediction of vessels based on data mining and machine learning," *J. Digit. Inf. Manage.*, vol. 14, no. 1, pp. 33–40, 2016.
- [18] T. Xiaopeng, C. Xu, S. Lingzhi, M. Zhe, and W. Qing, "Vessel trajectory prediction in curving channel of inland river," in *Proc. Int. Conf. Transp. Inf. Saf. (ICTIS)*, Jun. 2015, pp. 706–714.
- [19] L. Zhang and Q. Meng, "Probabilistic ship domain with applications to ship collision risk assessment," *Ocean Eng.*, vol. 186, Aug. 2019, Art. no. 106130.
- [20] C. Wang, H. Ren, and H. Li, "Vessel trajectory prediction based on AIS data and bidirectional GRU," in *Proc. Int. Conf. Comput. Vis., Image Deep Learn. (CVIDL)*, Jul. 2020, pp. 260–264.
- [21] B. R. Dalsnes, S. Hexeberg, A. L. Flaten, B.-O.-H. Eriksen, and E. F. Brekke, "The neighbor course distribution method with Gaussian mixture models for AIS-based vessel trajectory prediction," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Jul. 2018, pp. 580–587.
- [22] T. Wang, Q. Wu, J. Zhang, B. Wu, and Y. Wang, "Autonomous decision-making scheme for multi-ship collision avoidance with iterative observation and inference," *Ocean Eng.*, vol. 197, Feb. 2020, Art. no. 106873.
- [23] F. Sharifzadeh, G. Akbarizadeh, and YS. Kaviani, "Ship classification in SAR images using a new hybrid CNN-MLP classifier," *J. Indian Soc. Remote Sens.*, vol. 47, no. 4, pp. 551–562, Apr. 2019.
- [24] S. Yang, P. Xinya, D. Zexuan, and Z. Jiansen, "An approach to ship behavior prediction based on AIS and RNN optimization model," *Int. J. Transp. Eng. Technol.*, vol. 6, pp. 16–21, Mar. 2020.
- [25] F. Cheraghchi, I. Abualhaol, R. Falcon, R. Abielmona, B. Raaheemi, and E. Petriu, "Modeling the speed-based vessel schedule recovery problem using evolutionary multiobjective optimization," *Inf. Sci.*, vols. 448–449, pp. 53–74, Jun. 2018.
- [26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [27] Y. Zhang, S. Qiao, S. Ji, and Y. Li, "DeepSite: Bidirectional LSTM and CNN models for predicting DNA–protein binding," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 4, pp. 841–851, 2020.
- [28] H. Tang, Y. Yin, and H. Shen, "A model for vessel trajectory prediction based on long short-term memory neural network," *J. Mar. Eng. Technol.*, pp. 1–10, doi: 10.1080/20464177.2019.1665258.
- [29] B. Murray and L. P. Perera, "A dual linear autoencoder approach for vessel trajectory prediction using historical AIS data," *Ocean Eng.*, vol. 209, Aug. 2020, Art. no. 107478.
- [30] R. W. Liu, J. Nie, S. Garg, Z. Xiong, Y. Zhang, and M. S. Hossain, "Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6G-enabled maritime IoT systems," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5374–5385, Apr. 2021.
- [31] M. Liang, R. W. Liu, S. Li, Z. Xiao, X. Liu, and F. Lu, "An unsupervised learning method with convolutional auto-encoder for vessel trajectory similarity computation," *Ocean Eng.*, vol. 225, Apr. 2021, Art. no. 108803.
- [32] S. Anderson, T. D. Barfoot, C. H. Tong, and S. Särkkä, "Batch nonlinear continuous-time trajectory estimation as exactly sparse Gaussian process regression," *Auto. Robots*, vol. 39, no. 3, pp. 221–238, Oct. 2015.
- [33] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, Nov. 2018, Art. no. e00938.
- [34] W. Zaremba, I. Sutskever, and O. Vinyals, "Recurrent neural network regularization," 2014, *arXiv:1409.2329*.
- [35] B. L. Young, *Predicting Vessel Trajectories From AIS Data Using R*. Monterey, CA, USA: Naval Postgraduate School, 2017.
- [36] A. Daranda, "Neural network approach to predict marine traffic," *Baltic J. Mod. Comput.*, vol. 4, no. 3, pp. 483–495, 2016.
- [37] M. Lacki, "Intelligent prediction of ship maneuvering," *TransNav, Int. J. Mar. Navigat. Saf. Sea Transp.*, vol. 10, no. 3, pp. 511–516, 2016.
- [38] T. Xu, X. Liu, and X. Yang, "A novel approach for ship trajectory online prediction using BP neural network algorithm," *Int. J. Adv. Inf. Sci. Service Sci.*, vol. 4, no. 11, pp. 271–277, 2012.
- [39] G. K. D. de Vries and M. van Someren, "Machine learning for vessel trajectories using compression, alignments and domain knowledge," *Expert Syst. Appl.*, vol. 39, no. 18, pp. 13426–13439, Dec. 2012.
- [40] C. Zhong, Z. Jiang, X. Chu, and L. Liu, "Inland ship trajectory restoration by recurrent neural network," *J. Navigat.*, vol. 72, no. 6, pp. 1359–1377, Nov. 2019.
- [41] S. Gan, S. Liang, K. Li, J. Deng, and T. Cheng, "Ship trajectory prediction for intelligent traffic management using clustering and ANN," in *Proc. UKACC 11th Int. Conf. Control (CONTROL)*, Aug. 2016, pp. 1–6.
- [42] E. Tu, G. Zhang, S. Mao, L. Rachmawati, and G.-B. Huang, "Modeling historical AIS data for vessel path prediction: A comprehensive treatment," 2020, *arXiv:2001.01592*.
- [43] D.-D. Nguyen, C. Le Van, and M. I. Ali, "Vessel trajectory prediction using sequence-to-sequence models over spatial grid," in *Proc. 12th ACM Int. Conf. Distrib. Event-Based Syst.*, Jun. 2018, pp. 258–261.
- [44] P. Borkowski, "The ship movement trajectory prediction algorithm using navigational data fusion," *Sensors*, vol. 17, no. 6, p. 1432, 2017.
- [45] D. Zissis, E. K. Xidias, and D. Lekkas, "Real-time vessel behavior prediction," *Evolving Syst.*, vol. 7, no. 1, pp. 29–40, Mar. 2016.
- [46] P. R. Lei, J. Su, W.-C. Peng, W. Y. Han, and S. P. Chang, "A framework of moving behavior modeling in the maritime surveillance," *J. Chung Cheng Inst. Technol.*, vol. 40, no. 2, pp. 33–44, 2011.

- [47] A. Valsamis, K. Tserpes, D. Zissis, D. Anagnostopoulos, and T. Varvarigou, "Employing traditional machine learning algorithms for big data streams analysis: The case of object trajectory prediction," *J. Syst. Softw.*, vol. 127, pp. 249–257, May 2017.
- [48] N. Forti, L. M. Millefiori, P. Braca, and P. Willett, "Prediction of vessel trajectories from AIS data via sequence-to-sequence recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 8936–8940.
- [49] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, and G.-B. Huang, "Exploiting AIS data for intelligent maritime navigation: A comprehensive survey from data to methodology," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1559–1582, May 2018.
- [50] R. G. Brown, "Exponential smoothing for predicting demand," *Oper. Res.*, vol. 5, no. 1, p. 145, 1957.
- [51] C. Li, X. Wang, Z. Cheng, and Y. Bai, "Forecasting bus passenger flows by using a clustering-based support vector regression approach," *IEEE Access*, vol. 8, pp. 19717–19725, 2020.
- [52] D. Parbat and M. Chakraborty, "A Python based support vector regression model for prediction of COVID19 cases in India," *Chaos, Solitons Fractals*, vol. 138, Sep. 2020, Art. no. 109942.
- [53] H. Namgung and J.-S. Kim, "Vessel trajectory analysis in designated harbor route considering the influence of external forces," *J. Mar. Sci. Eng.*, vol. 8, no. 11, p. 860, Oct. 2020.
- [54] H.-H. Liu, L.-C. Chang, C.-W. Li, and C.-H. Yang, "Particle swarm optimization-based support vector regression for tourist arrivals forecasting," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–13, Sep. 2018.
- [55] S. Yang, H.-C. Chen, W.-C. Chen, and C.-H. Yang, "Forecasting outbound Student mobility: A machine learning approach," *PLoS ONE*, vol. 15, no. 9, Sep. 2020, Art. no. e0238129.
- [56] L. You, S. Xiao, Q. Peng, C. Claramunt, X. Han, Z. Guan, and J. Zhang, "ST-Seq2Seq: A spatio-temporal feature-optimized Seq2Seq model for short-term vessel trajectory prediction," *IEEE Access*, vol. 8, pp. 218565–218574, 2020.
- [57] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.
- [58] C.-H. Yang, C.-H. Wu, and C.-M. Hsieh, "Long short-term memory recurrent neural network for tidal level forecasting," *IEEE Access*, vol. 8, pp. 159389–159401, 2020.
- [59] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, Jun. 2014.
- [60] G. E. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control San Francisco*, 5th ed. Hoboken, NJ, USA: Wiley, 2015.
- [61] V. Vapnik, S. E. Golowich, and A. J. Smola, "Support vector method for function approximation, regression estimation and signal processing," in *Proc. Adv. Neural Inf. Process. Syst.*, 1997, pp. 281–287.
- [62] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.
- [63] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Manage. Sci.*, vol. 6, no. 3, pp. 324–342, 1960.
- [64] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *Int. J. Forecasting*, vol. 20, no. 1, pp. 5–10, 2004.
- [65] E. Ceperic, V. Ceperic, and A. Baric, "A strategy for short-term load forecasting by support vector regression machines," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4356–4364, Nov. 2013.
- [66] A. D. Mehr, V. Nourani, V. K. Khosrowshahi, and M. A. Ghorbani, "A hybrid support vector regression-firefly model for monthly rainfall forecasting," *Int. J. Environ. Sci. Technol.*, vol. 16, no. 1, pp. 335–346, 2019.
- [67] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.
- [68] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. 30th Int. Conf. Mach. Learn.*, 2013, pp. 1310–1318.
- [69] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," 2014, *arXiv:1409.1259*.
- [70] F. Kratzert, D. Klotz, C. Brenner, and K. Schulz, "Rainfall-runoff modelling using long short-term memory (LSTM) networks," *Hydrol. Earth Syst. Sci.*, vol. 22, no. 11, pp. 6005–6022, 2018.
- [71] S. Ruder, "An overview of gradient descent optimization algorithms," 2016, *arXiv:1609.04747*.



CHENG-HONG YANG (Senior Member, IEEE) is currently the Chair Professor with the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan. He served as the President for the National Kaohsiung University of Applied Science, Taiwan, from 2012 to 2016. He has authored/coauthored over 380 refereed publications and a number of book chapters. His main research interests include fuzzy control, evolutionary computation, bioinformatics, and deep learning. He is also an editorial board member of multiple international journals.



CHI-HSIEN WU received the B.S. degree from the Department of Information Management, Shu-Te University, in 2019. He is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan. He has rich experience in machine learning, computer programming, database design and management, and systems programming and design. His main research interests include time series, data mining, machine learning, and deep learning.



JEN-CHUNG SHAO received the B.S. degree from the Department of Maritime Information and Technology, National Kaohsiung Marine University, in 2015. He is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan. His research interests include the machine learning, deep learning, and time series.



YI-CHUAN WANG received the Ph.D. degree in civil engineering from National Taiwan University, in 2003. He is currently an Assistant Professor with the Department of Business Administration, CTBC Business School. His research interests include the public transportation planning and operation, transit-oriented development, transportation economics, and traffic engineering.



CHI-MIN HSIEH was born in Taiwan, in 1973. He is currently pursuing the master's and Ph.D. degrees in ocean engineering with the Department of Hydraulic and Ocean Engineering, National Cheng Kung University. From 2006 to 2007, he was a Postdoctoral Fellow with the Institute of Physics, Academia Sinica, Taipei, Taiwan. He is also a Professor with the Department of Maritime Information and Technology, National Kaohsiung University of Science and Technology. His research interests include the signal analysis, and computed fluid dynamics for wave dynamics, sediment transport, and coastal engineering.