JOURNAL OF MARITIME RESEARCH

Vol XXII. No. I (2025) pp 75–81

ISSN: 1697-4840, www.jmr.unican.es

Automatic Identification System (AIS) - Based Ship Trajectory Modelling for Indonesian Sea Transportation Monitoring

Natalia Damastuti[1,2], Aulia Siti Aisjah[1,*], Agoes Masroeri[3]

ARTICLE INFO

Article history:
Received 16 Jun 2024;
in revised from 26 Jun 2024;
accepted 23 Jul 2024.

Keywords:
AIS, Vessel Navigation, Prediction, Trajectory, Data Mining.

ABSTRACT

The aim of this research is to build a ship trajectory model and display ship predictions for the next period, using 10 minutes of preliminary data on the trajectory of ships sailing in the Java Sea, and then predict the trajectory of the next ship in the next 20 minutes and the next 40 minutes. This study uses data mining steps in AIS data mining which is then processed using the LSTM learning algorithm with sequence prediction. By using some static data on AIS, that is SOG, Latitude, Longitude and MMSI from the ship during November 2018 in the Java Sea, Indonesia. Results of the study indicated that the RMSE resulted in a value of 0.13. This research is original in its approach to utilizing AIS data with the LSTM algorithm for ship trajectory prediction in the Java Sea. The integration of sequence prediction with specific ship data parameters provides a novel methodology for maritime navigation and safety management. This study contributes valuable insights into the predictive modelling of ship movements, which can enhance the efficiency and safety of maritime operations. The impact of this research is significant for maritime navigation and safety. By accurately predicting ship trajectories, it can help in avoiding potential collisions and improving route planning. This can lead to better fuel efficiency and reduced operational costs for shipping companies. Additionally, the methodology can be adapted to other regions and types of vessels, providing a broader application for global maritime safety and logistics.

1. Introduction.

AIS technology has been long introduced, yet only certain ships have used it. Since 2019, the Regulation of the Minister of Transportation Number PM 7/2019 obliged the installation of AIS system for every ship sailing in Indonesian waters. AIS system is functioned not only as a means of communication between ships or transmitter stations but also as a vessel traffic service, search and rescue and monitoring of shipping in the Indonesian Sea (Aisjah et al. n.d.).

AIS data is also very useful for analyzing vessel movements and evaluating the risk of collision. In its media release, the Ministry of Transportation informed that 54 marine transportation accidents occurred in Indonesia from 2010-2016 with 32% of the accidents were ships collisions while the other was weather and other factors. AIS is an electronic device that functions as a navigation system for marine transportation. AIS is able to identify the location where the ship is sailing, exchange data electronically including the identification of the position, activity or state of the ship, and speed, with other nearby ships and Vessel Traffic Services (VTS) stations (Aisjah 2018)(Silveira, Teixeira, and Soares 2013).

AIS data can be used for maritime purposes, especially to understand the characteristics of navigation through analysis of

[1]Department of Physics Engineering, Faculty of Technology Industry and Engineering System, Institut Teknologi Sepuluh Nopember Surabaya. ITS Campus, Keputih, Sukolilo 61111, Surabaya, Indonesia .

[2]Departement of Computer System, Faculty of Computer Science, Universitas Narotama Surabaya Jl. Arief Rahman Hakim No. 51, Surabaya, Indonesia.

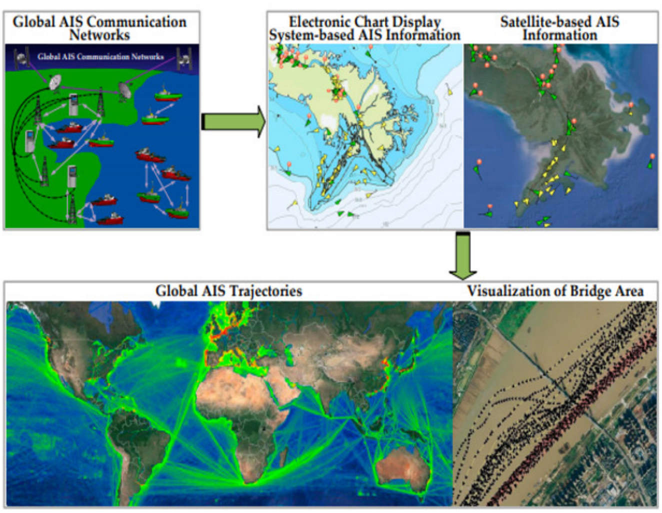
[3]Departement of Marine, Faculty of Marine, Institut Teknologi Sepuluh Nopember Surabaya. ITS Campus, Keputih, Sukolilo 61111, Surabaya, Indonesia.

*Corresponding author: Aulia Siti Aisjah. E-mail Address: auliasa20@gmail.com.

traffic patterns in the waters. Visually, AIS technology is very helpful in real-time ship surveillance. (Li et al. 2017)(Parry and Sampath 2013). The AIS system is used in almost all ships sailing in Indonesian waters, the information display on AIS has several visualizations as shown in Figure 1. There are several providers of AIS data such as marine traffic, FletMon, MarineCadastre, Aprs and VT Explorer. There are several providers providing data access online and some of them are free (Mao et al. 2017).

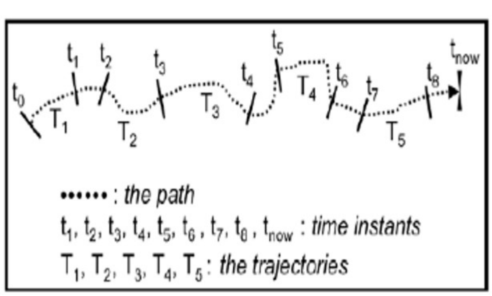
In the research, AIS data were obtained from the online website www.marinetraffic.com. Trajectory is related to the movement of an object in space and time. Data can be recorded in the form of geographic location in a certain time in sequence. The sequence of geographic location and time sequence will form a path where the path represents the paths that are passed by the object. Figure 2 is a description of the movement of an object in a certain time and location. The point where the movement passes is called trajectory.

Figure 1: AIS information in several forms of visualization.



Source: Authors.

Figure 2: The movement of objects in a certain period of time.



Source: Authors.

Spatio-temporal trajectory model is mostly used to understand the behavior of object movements. The use of large, complex, heterogeneous and exponential AIS can be presented in a trajectory data mining model [5]. However, storing and collecting large trajectory data such as AIS data is not easy as it requires complex data processing stages. Machine learning technique is used in this study to proceed the data. In maritime sector, tracking the ship's position can also use the AIS system. Furthermore, AIS data can also be used to predict ship trajectories using machine learning methods (Mao et al. 2017), (Fernández Calvo et al. 2017) AIS data is so susceptible to noise that it is necessary to remove the noise as performed by (Zhang et al. 2018),(Kim and Lee 2018).

However, the noise removal process is still carried out based on the ship's destination, while in this study the noise removal was carried out in more than one field (SOG, Lat, Long, travel time). Predicting a ship's trajectory is different from that of airplane. Currently, there have been new breakthroughs in predicting ship trajectories using traditional methods (Valsamis et al. 2017)(Tu et al. n.d.)(Vries and Someren 2010), yet future behavior has not been demonstrated. Traditional model can work well, but it is a model-driven, meaning that the model can only be applied in that environment until finally other studies offer pattern recognition and image processing based research, where they claimed that the accuracy is better (Tang, Yin, and Shen 2019). However, the model offers few variables regardless of external factors such as weather and ocean waves. In this study, a Long Short-Term Memory (LSTM)- based ship trajectory model in Java-Indonesia region is proposed by considering the external factor of wave height.

2. Materials and Methods.

2.1. Data Collection.

The present study used data from marrinetrafic.com and NASDEC-ITS while data of wave height was obtained from BMKG Surabaya with the period November 2018 with latitude of -9.395191-1.043314 and a longitude of 107.362342-117.931183. As shown in Figure 3, the research object is focused in the Java Sea where the area is an area with heavy ship traffic and 160,000 vessels were recorded during that time period.

Figure 3: Visualization.

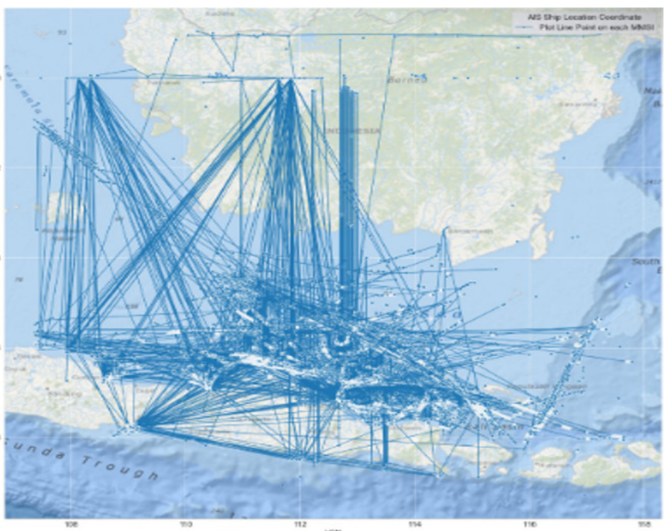
Source: marinetraffic.com.

The raw data obtained in the scope of AIS is a very large data stream that some fields are filtered prior to processing.

2.2. PreProcessing Stage.

Figure 4 is a display of raw data. Extracting raw data is difficult, thus, preprocessing needs to be carried out to select the desired fields, eliminate noise and normalize the data.

Figure 4: Display of raw data (the x axis is the longitude and the y axis is the latitude).



Source: Authors.

Data cleaning and normalization were carried out to proceed raw data. The data attribute values with different ranges needed to be normalized since the difference in the range of values causes the malfunction of the attribute that has a much smaller value compared to other attributes. Therefore, data transformation with normalization to equalize the range of values is needed (Nasution, Khotimah, and Chamidah 2019). Normalization is carried out in a small range of values [0,1] or [-1,1]. Thus, the attributes have the same weight. In this study, normalization used min-max normalization as formulated in formula 1.

$$normalized(x) = \frac{minRange + (x - minValue)(maxRange - minRange}{maxValue - minValue} \quad (1)$$

Where x 'is the result of normalization of the data set range [min value, max value].

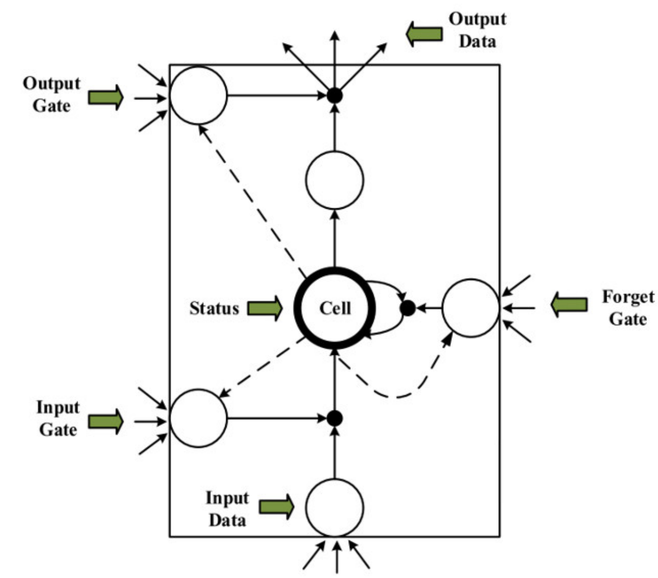
2.3. Training Process and Testing Process.

Figure In the present study, LSTM model was used. The data WERE divided into training and testing data. Training data was used to train algorithm networks while testing data was used to determine the performance of previously trained algorithms. The result of the training is called a model. Separating the data into training data and testing data was carried out so that the model obtained have good generalizability in making data predictions. The amount of training data was greater than that of the testing data. In this study, 80% of the training data and 20% of the testing data were used by using the train_test_split () function in the scikit-learn library.

2.4. Prediction Process.

After training, track prediction was carried out. In the present study, LSTM method is used. This is an implementation of an artificial neural network with ability of storing previous memory inputs and predicting time series efficiently. This method can also be applied to predictions on land such as pedestrians and vehicles (Altche and De La Fortelle 2018).

Figure 5: General structure of LSTM network.



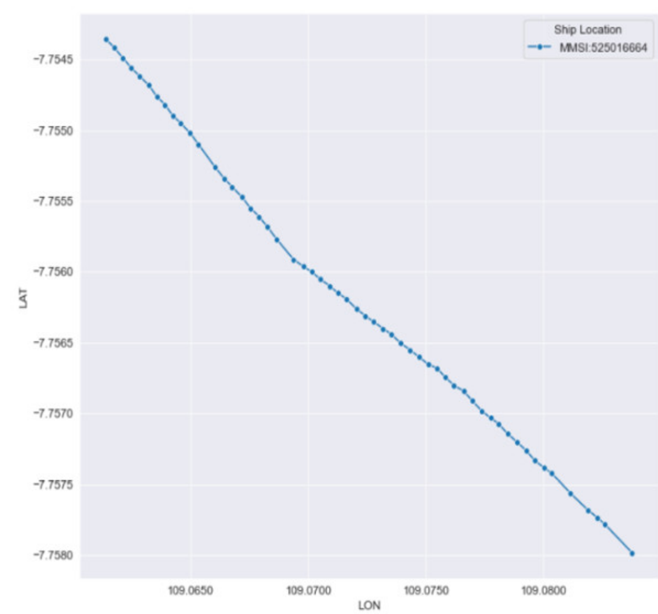
Source: Authors.

Algorithm was built using python software and the library in it to handle very large data problems. Unlike plane trajectories that can be 3-dimensional modeled, ship trajectories follow patterns of geographic location (x, y), SOG, Course, and other information. The path is highly influenced by time that can be defined $S[x_j^t, y_j^t, v_j^t, c_j^t]$ with vessel j = 1,2... N and time t = k-n,.... , k. path prediction $\left(x_j^{t_{k+m}}, y_j^{t_{k+m}}\right)$as presented in [13]. The LSTM network structure is shown in Figure 5. LSTM is proposed as a solution of prediction where a memory cell is added to store information for a relatively long time using a 'forgot gate'. If i is 'forgot gate input' in time t, then it can be formulated as follows: $i_t = \sigma(W_{hi}h_{t-1} + W_{xi}x_t)$ Where σ is the activation function, W is the weight, h is the hidden layer.

3. Result and Discussion

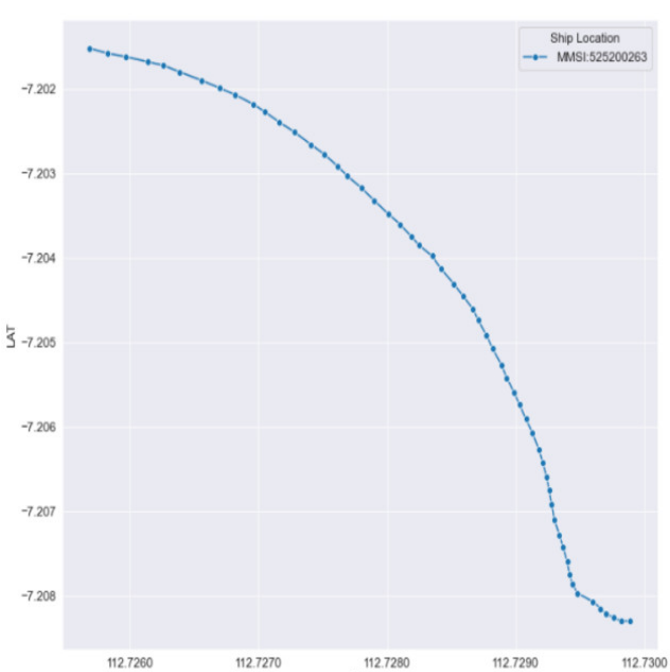
Big AIS data requires special handling so that the data can be processed. The data obtained is still raw data. During 3 months there were 2.165.612 data that could be recorded. In this case, we use one month's data, that is November 2018, where the recorded data is 12.942 and there are 64 types of ships that were sailing in that period including Tug, Container, Passenger, Patrol Vessel, Crane Ship, Ro-Ro, Tanker. Static and dynamic information as well as the destinations of these vessels are also recorded in AIS. Figure 5-8 is a sampling of several ship trajectories from different ship types. The track uses a time-stamp for 10 minutes. Figure 6 is a sampling of several ship trajectories from different ship types. The trajectories are time-stamped at 10 minutes.

Figure 6: Trajectory (1) based on defferende vessel types.



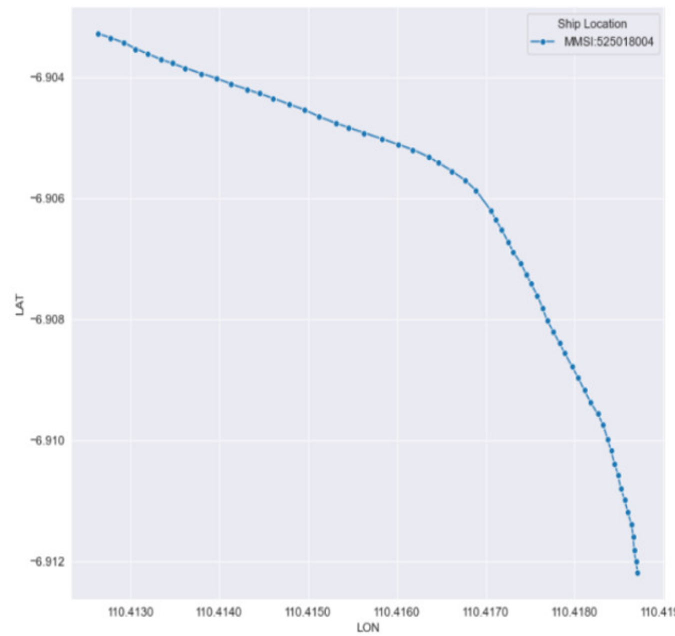
Source: Authors.

Figure 7: Trajectory (2) based on defferende vessel types.



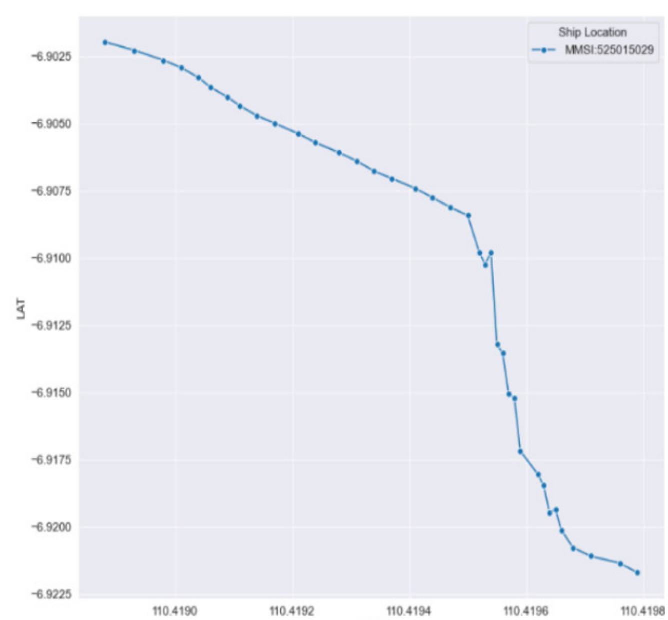
Source: Authors.

Figure 8: Trajectory (3) based on defferende vessel types.



Source: Authors.

Figure 9: Trajectory (4) based on defferende vessel types.



Source: Authors.

The table 1 is the result of preprocessing. While the table 2 is the result of normalization of preprocessing data.

The training process in data processing uses the LSTM model which has several gates, that is as input 'forget gates', 'input gates', 'cell gates', and 'output gates'. The input through the 'forget gates' will be selected which will be stored in the 'memory cell' using 'sigmoid' activation with the output binary value 0 and 1, where if the output is 1 it will be stored, and if the out-

Table 1

	MMSI	LAT	LON	SPEED	COURSE	DATE	DATE_FIRST	DATE_PREV	TIMEDIFF	MINUTES	DISTANCE
0	244370000	1.727872	-1.305227	0.147783	0.163814	13	13	NaN	0.005441	0.000000	0.000000
1	244370000	1.773303	-1.270349	0.142857	0.163814	13	13	13.0	0.005731	9.300000	0.005594
2	244370000	1.772440	-1.271030	0.142857	0.163814	13	13	13.0	0.005441	9.300000	0.000109
3	244370000	2.150964	-0.906572	0.142857	0.212714	13	13	13.0	0.008448	105.683333	0.056833
4	244370000	2.151367	-0.904101	0.142857	0.212714	13	13	13.0	0.005467	106.500000	0.000366
...	...	...	...	...	...	...	...	...	...	...	...
12937	525900057	-0.712770	0.563095	0.029557	0.471883	23	23	23.0	0.005470	116.116667	0.000050
12938	525900057	-0.715014	0.562359	0.059113	0.608802	23	23	23.0	0.005464	116.866667	0.000153
12939	525900057	-0.715791	0.561454	0.088670	0.625917	23	23	23.0	0.005455	117.300000	0.000139
12940	525900057	-0.717143	0.559048	0.098522	0.628362	23	23	23.0	0.005469	118.200000	0.000361
12941	525900057	-0.718668	0.557519	0.068966	0.603912	23	23	23.0	0.005462	118.866667	0.000237

Source: Authors.

Table 2

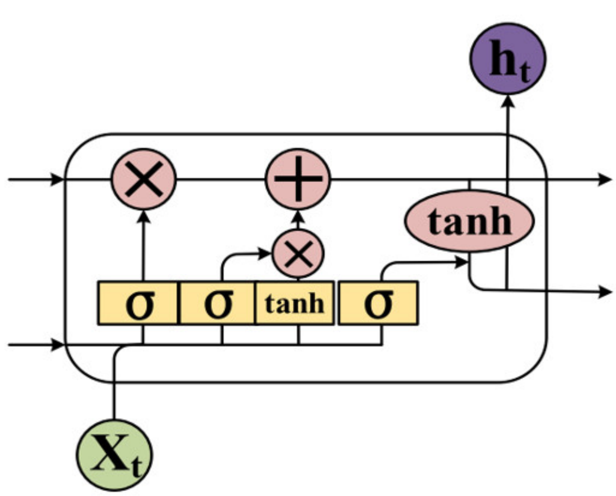
	SHIP TYPE	WIDTH	DWT	LENGTH
count	76225	76225	76225	76225
mean	6.269177	21.259508	24789.5856	120.01224
std	1.887615	10.618657	45990.85	67.095835
min	0	2	7	2
25%	6	14	3288	68
50%	7	20	8500	108
75%	7	27	24789.5856	166
max	9	70	308491	433

Source: Authors.

put is 0 it will be discarded. It follows the formula (Gao, Shi, and Li 2018):

$$f_t = sigmoid(W_f\,[h_{t-1},\ x_t] + b_f)$$

Figure 10: Memory Cell LSTM.



Source: Authors.

The next process is through 'input gates' which will decide the value to be updated using sigmoid activation, while tanh activation will produce a new value that will be stored in the memory cell using the formula:

$$C_t = tanh(W_c\,[h_{t-1},\ x_t] + b_c)$$

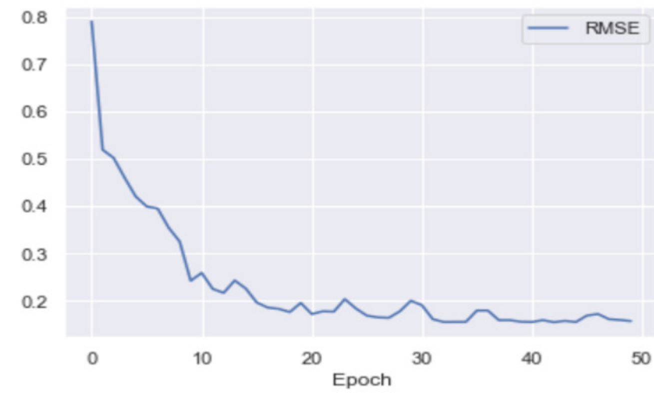
The output value of the previous process becomes input in this 'output gate' which will then complete the training process. The formula uses something like this:

$$o_t = sigmoid(W_o\,[h_{t-1},\ x_t] + b_o)$$

$$h_t = \tanh(C_t)$$

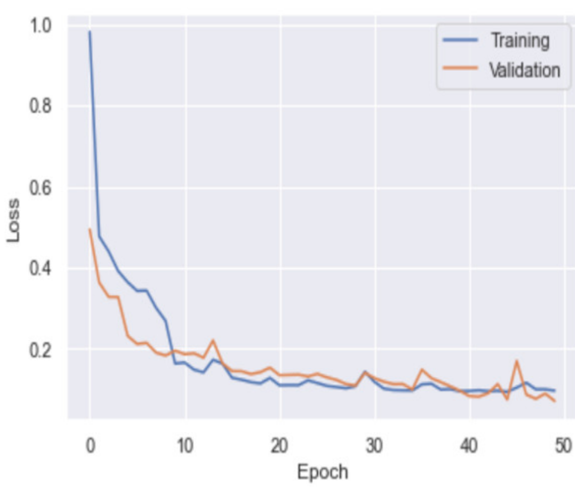
This research uses the python programming language and as initialization the basic parameters required are the learning rate, number of hidden layers, number of neurons, target error, and epoch.

Figure 11: Regression Analytics.



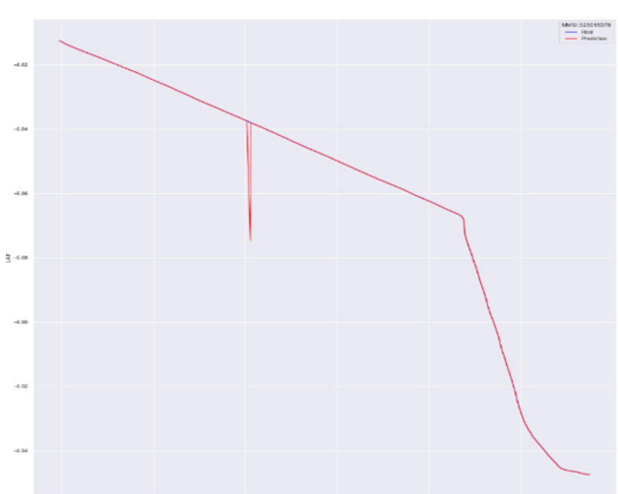
Source: Authors.

Figure 12: Model Loss.



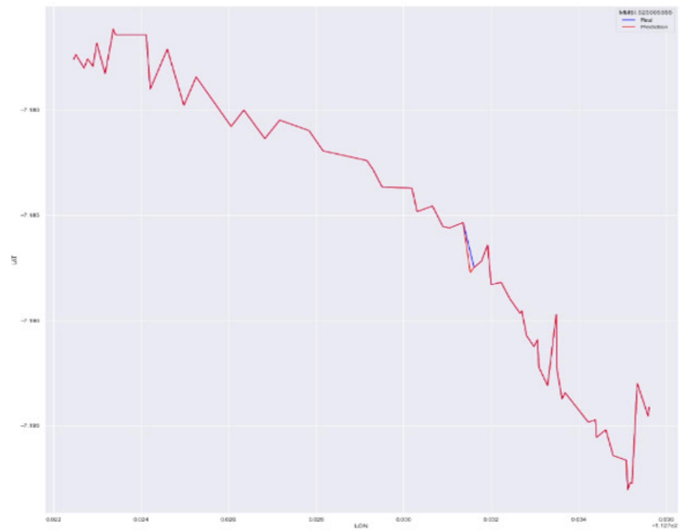
Source: Authors.

Figure 13: Result of prediction trajectory from different MMSI (1).



Source: Authors.

Figure 14: Result of prediction trajectory from different MMSI (2).



Source: Authors.

Conclusions.

According to the discussion, the model that we can build with the LSTM algoritms to predict the location of vessel at the 20th minute and 40th minute. a sequence LSTM layers is delivered and the network is trained by prediction. The historical AIS data of Java Sea, Indonesia are used to train and testing the LSTM model. The model perform is better RMSE, its about 0.130.

Acknowledgements.

The authors gratefully thank to the Department of Physics Engineering of the Institute of Technology Sepuluh Nopember (ITS) and LPPM for all the support of this research.

References.

Aisjah, Aulia Siti. 2018. Pemetaan Pola Gerak Ilegal Fishing Dan Ilegal Transhipment Pada Vessel Monitoring System Berbasis Data AIS. Surabaya.

Aisjah, Aulia Siti, A A Masroeri, Moch Yusuf Santoso, and Syamsul Arifin. "A FUZZY DECISION SUPPORT SYTEM OF IUU – TRANSHIPMENT IN INDONESIA Department of Engineering Physics , Faculty of Industrial Technology Department of Marine Engineering, Faculty of Maritime Technology Institut Teknologi Sepuluh Nopember (ITS) Shipbuidin." : 53–60.

Altche, Florent, and Arnaud De La Fortelle. 2018. "An LSTM Network for Highway Trajectory Prediction." IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC 2018-March(January): 353–59.

Fernández Calvo, Esther et al. 2017. "DART: A Machine-Learning Approach to Trajectory Prediction and Demand - Capacity Balancing." Seventh SESAR Innovation Days (November).

Gao, Miao, Guoyou Shi, and Shuang Li. 2018. "Online Prediction of Ship Behavior with Automatic Identification System Sensor Data Using Bidirectional Long Short-Term Memory Recurrent Neural Network." Sensors (Switzerland) 18(12).

Kim, Kwang Il, and Keon Myung Lee. 2018. "Mining of Missing Ship Trajectory Pattern in Automatic Identification System." International Journal of Engineering and Technology(UAE) 7(2): 167–70.

Li, Huanhuan et al. 2017. "A Dimensionality Reduction-Based Multi-Step Clustering Method for Robust Vessel Trajectory Analysis." Sensors (Switzerland) 17(8).

Mao, Shangbo et al. 2017. "An Automatic Identification System (AIS) Database for Maritime Trajectory Prediction and Data Mining."

Nasution, Darnisa Azzahra, Hidayah Husnul Khotimah, and Nurul Chamidah. 2019. "Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-NN." Computer Engineering, Science and System Journal 4(1): 78.

Parry, David, and Premalatha Sampath. 2013. "Trajectory Analysis Using Automatic Identification System (AIS) in New Zealand Waters." International Journal of Computer and Information Technology 02(01): 132–36. http://aut.researchgateway-.ac.nz/handle/10292/5333.

Silveira, P. A.M., A. P. Teixeira, and C. Guedes Soares. 2013. "Use of AIS Data to Characterise Marine Traffic Patterns and Ship Collision Risk off the Coast of Portugal." Journal of Navigation 66(6): 879–98.

Tang, Huang, Yong Yin, and Helong Shen. 2019. "A Model for Vessel Trajectory Prediction Based on Long Short - Term Memory Neural Network." Journal of Marine Engineering and Technology 0(0): 1–10. https://doi.org/20464177.2019.1665258.

Tu, Enmei et al. "Exploiting AIS Data for Intelligent Maritime Navigation : A Comprehensive Survey." : 1–24.

Valsamis, Angelos et al. 2017. "Employing Traditional Machine Learning Algorithms for Big Data Streams Analysis: The Case of Object Trajectory Prediction." Journal of Systems and Software 127(September): 249–57.

Vries, Gerben De, and Maarten Van Someren. 2010. "Clustering Vessel Trajectories with Alignment." : 296–311.

Zhang, Liye, Qiang Meng, Zhe Xiao, and Xiuju Fu. 2018. "A Novel Ship Trajectory Reconstruction Approach Using AIS Data." Ocean Engineering 159(February): 165–74. https://doi.-org/10.1016/j.oceaneng.2018.03.085.