JOIV : Int. J. Inform. Visualization, 6(4) - December 2022 821-828

INTERNATIONAL JOURNAL
ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Aircraft Flight Movement Anomaly Detection using Automatic Dependent Surveillance-Broadcast

Abdul Azzam Ajhari [a], I Gede Putra Kusuma Negara [a,*]

[a] *Computer Science Department, BINUS Graduate Program - Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia*

*Corresponding author: *inegara@binus.edu*

Abstract*—Automatic Dependent Surveillance-Broadcast (ADS-B) is an aircraft backup radar device that transmits aircraft sensor information via radio frequency. This data can be used to detect aircraft changes that occur significantly or abnormally (anomaly). Anomaly detection in this study aims to reduce and prevent flight accidents by analyzing abnormal data on aircraft flights using the Deep Learning (DL) model. Bidirectional LSTM (Bi-LSTM) and Bidirectional GRU (Bi-GRU) models are proposed as DL models which are implemented on ADS-B data using data mining methods. The data is generated from the ADS-B device that records the plane crash incident and is stored on the Flightradar24 community server. Data containing sensor changes from anomalous aircraft movements are studied for predictability on other flight data. The class breakdown is divided into two, anomaly and normal, based on information on the time span of anomaly occurrences in the accident investigation report of each aircraft using the sliding window technique in the data mining methodology. In evaluation, the confusion matrix measurement method is used to predict predictive analysis of the tested data. The results of the model evaluation performance show that the Bi-LSTM proposed in this study has the best accuracy of 99.44% and the f1-score of 99.51% is slightly superior to the Bi-GRU model. The model in this study can be applied in the ADS-B device to detect aircraft movement anomalies and as material for reviewing technicians in periodic maintenance and measuring the accuracy of the ADS-B device used as a backup radar.

***Keywords*— Aircraft flight movement; anomaly detection; aircraft ADS-B device; flight anomalies; data mining.**

Manuscript received 4 Jun. 2022; revised 21 Jul. 2022; accepted 5 Aug. 2022. Date of publication 31 Dec. 2022.

I. INTRODUCTION

The absence of a decrease in aviation accidents shows that flight supervision in Indonesia is still less than optimal since article 91 of the safety regulations on general operations and flight rules was made in 2017 [1]. For this reason, applying a technique to reduce and prevent aviation accidents is necessary for analyzing abnormal data on aircraft flight movements. Anomalies often form outliers, abnormalities, rare occurrences, or deviations of data points or styles that do not conform to expected behavior [2]. Some anomalies were found in ADS-B by transmitting manipulated data [3] which forms the pattern of his attack behavior [4] through ADS-B protocol security [5] and categorized as cyber-attacks [6]. Another anomaly occurs due to a mismatch of the aircraft's distance with a predetermined trajectory [7], [8]. So that, cauterization is widely used to predict the farthest distance, which is considered an anomaly [9], [10]. Significant changes in ADS-B data information that occur frequently and can affect aircraft eligibility are known as sensor anomalies [11]. Usually, sensor anomaly detection is used to determine the feasibility of an aircraft by testing it on an aircraft prototype [12], [13]. Other research on detecting aircraft flight movement anomaly sensors has been carried out on the ADS-B X-plane game data simulation [14].

This study focuses on aircraft movement anomalies based on the analysis of significant changes in the ADS-B sensor data using DL. The limitations of previous research are only detecting movement anomalies based on ADS-B sensor data in prototypes and simulation games. This study implemented anomaly detection on real ADS-B data using a community server from Flightradar24. Flightradar24 is a global aviation digital map service that provides real-time aircraft sensor data worldwide [15]. In addition, in understanding the sequence of anomaly data that occurs in airplane accidents and minimizing bias, labeling is carried out using the sliding windows technique from the sequence of data on accident investigation reports for each aircraft. The model used is the [14] model, namely LSTM and GRU, which is compared with the proposed model, namely Bi-LSTM and Bi-GRU. All models are evaluated with a confusion matrix to determine the best and most optimal model. Thus, this research can contribute to the aviation sector's feasibility of ADS-B devices monitoring air traffic. In addition, the detected anomaly data can be used as a review recommendation for flight technicians and officers in making decisions to reduce and prevent aviation accidents.

II. MATERIALS AND METHOD

Based on the type of anomaly detection performed using ADS-B data, the literature review in this study is divided into three parts: cyber-attack anomaly detection, aircraft trajectory anomaly detection, and sensor anomaly detection.

In the detection of cyber-attack anomalies, research [3] used the LSTM encoder to predict the number of anomalies manipulated by injecting ADS-B data which resulted in a false alarm warning error of about 4.5%. Anomaly detection is carried out based on the Generative Adversarial Networks (DAD-GAN) model to classify attack behavior and distinguish standard data from anomalous data [4]. Constant deviation injection attacks were detected with 98.2% accuracy, 99.3% random deviation injection attacks were detected, and deviation injection attacks increased by 90.6%. Abnormal classifications that have the potential to compromise the security of the ADS-B protocol have been detected, such as interference, modification, and injection [5] using the LSTM model with an accuracy of f1-score random noise detection at 0.9548, fixed offset+ at 0.9830, fixed offset - at 0.9834, route offset at 0.9885, altitude offset+ at 0.8928, altitude offset- at 0.9200, speed offset+ at 0.8853, speed offset- at 0.6470 at 0.5052, and climb rate at 0.2969. Cyberattacks can be adapted in [6] to create a fake message anomaly detection system on the injected ADS-B data. The superior LSTM model has a constant deviation attack percentage of 93.96%, flight replacement attacks of 99.71%, data replay attacks of 96.52%, and DoS attacks of 99.0%.

In detecting aircraft trajectory anomalies, Pusadan et al. [7] used the Agglomerative Hierarchical Clustering (AHC) model to detect anomalies in the plane trajectory where the farthest distance from the prediction is referred to as an anomaly. Their study showed a cophenetic correlation coefficient (c) of 0.691 c 0.974. Pusadan et al. [8] continued previous research to detect anomalies based on cluster segment predictions using the DBSCAN and K-Means models with a Dunn index value of 0.645 and a Silhouette index value of 0.89. Based on previous research, Pusadan et al. [9] carried out deeper anomaly detection based on segment formation and testing process from cluster distance, resulting in 96% K-NN and K-Means accuracy and 93% SVM. The detection of large and scattered aircraft trajectories in the ADS-B flight data is presented in the study [10] with the DBSCAN model to check for odd outliers resulting in 95% detection accuracy.

Furthermore, the anomaly detection sensor classifies the data sent by the ADS-B transponder using phase signal patterns and Neural Networks with a message detection accuracy of 64% and a secret field of 69% [11]. ADS-B is also used in testing the COMAC C919 aircraft on 19 sensor attributes using the LSTM model, which predicts a Root Mean Square Error (RMSE) value of 0.003 [12]. Then, research [13] used an optimized LSTM model to detect anomalies that occurred in the COMAC airplane test flight, which increased the accuracy of the prediction model up to 38% of the RMSE calculation analysis. Nanduri & Sherry [14] used 478 training data and 22 test data obtained from X-plane simulation data. 22 test data is divided into 11 normal data and 11 anomalous data. Eleven anomalous data consists of 11 different anomalies, so there is only one data for a certain type of anomaly. Next, they tried several LSTM, GRU, MKAD, and Autoencoder models to test for 11 anomalies. Although the LSTM and GRU models are relatively acceptable, they experiment with minimal anomalous data from the simulation data.

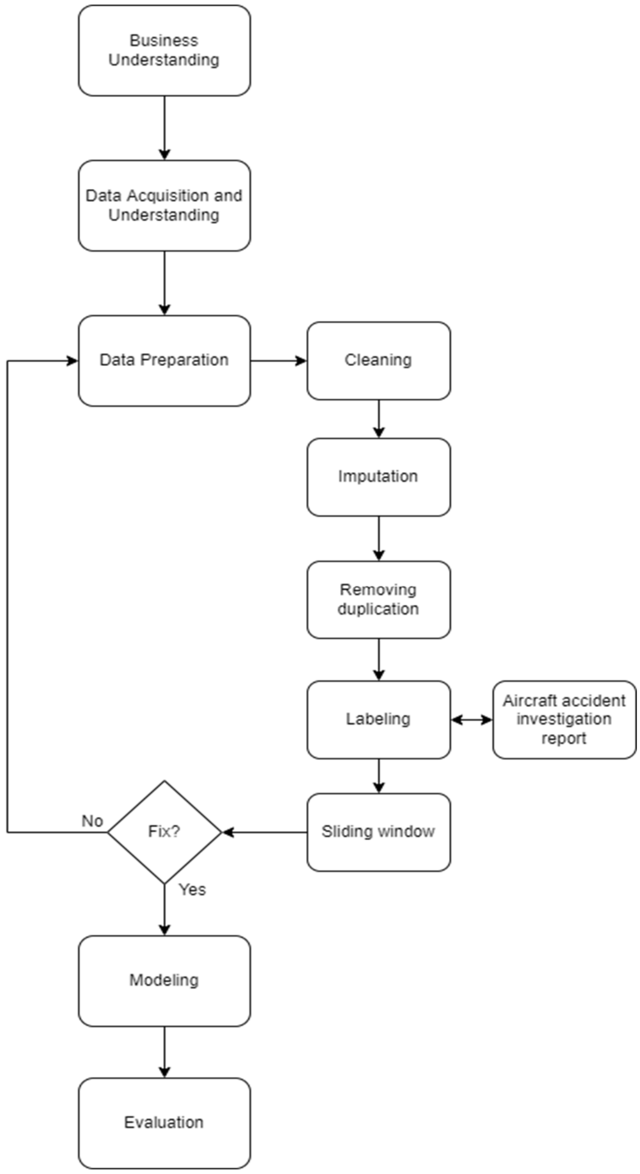

Fig. 1 Data mining research method process

This study proposes several LSTM and GRU models compared to the Bi-LSTM and Bi-GRU models to detect aircraft movement anomalies. Previous research has limitations, namely only detecting movement anomalies based on ADS-B sensor data on prototypes and simulation games. This study implemented anomaly detection on real ADS-B data using a community server from Flightradar24 from real plane crash data. Fig. 1 shows the main stages in the process of data mining research methodology.

Fig. 1 indicates 5-stages Cross-Industry Standard Process for Data Mining cycle (CRISP-DM) [16], starting from business understanding to reporting the performance evaluation test of the DL model used. This study aims to analyze and validate the method that has provided the best accurate results to be applied and carried out predictive testing on real ADS-B data with the following stages.

A. Business Understanding

The flight accident of the Sriwijaya Air SJ 182 aircraft, which occurred in early 2021 on January 9, was a sign of an emergency for aircraft maintenance. The US Federal Aviation Administration (FAA), on July 24, 2020, issued an emergency statement to airlines that own Boeing 737 aircraft to inspect their planes that have been parked for more than seven days. The reason is that it has not been used during the COVID-19 pandemic, which can cause corrosion. It is understood that the SJ 182 had been parked for nine months and was already operating before the crash. Therefore, we need a model that can analyze and predict the presence of anomalies in the ADS-B device as information to be validated.

B. Data Acquisition and Understanding

The data was collected from aircraft accident investigation reports from 2016 to 2021. The data is captured by various ADS-B who are scattered and joined in the Flightradar24 community, where the information is stored on the community server. In this study, 14 aircraft crash data which can be seen in Table I were analyzed based on aircraft accident investigation reports determining anomalies and regular classes as ground truth using the sliding window technique.

TABLE I
FLIGHTRADAR24 DATASETS

No	Accident Date	Aircraft Name
1	January 9, 2021	Sriwijaya Air SJ182
2	May 22, 2020	Pakistan International Airlines PK8303
3	January 26, 2020	S-76 Kobe Bryant Helicopter
4	May 5, 2019	Aeroflot 1492
5	March 10, 2019	Ethiopian Airlines 302
6	February 23, 2019	Atlas Air 5Y3591
7	October 29, 2018	Lion Air JT610
8	April 17, 2018	Southwest Airlines 1380
9	February 18, 2018	Iran Aseman Airlines 3704
10	February 11, 2018	Saratov Airlines 703
11	September 30, 2017	Air France AF66
12	January 15, 2017	MyCargo ACT Airlines TK6491
13	August 3, 2016	Emirates 521
14	March 19, 2016	FlyDubai 981

There are seven features or attributes used: latitude, longitude, altitude, ground speed (speed), heading, vertical speed (vspeed), and milliseconds. The millisecond attribute is obtained from the conversion of the timestamp data. While one feature, namely ICAO/hex is not included because the feature is only an aircraft identification code. The description and explanation of the attributes of the research dataset can be seen in Table II.

TABLE II
FLIGHTRADAR24 DATASETS ATTRIBUTES

No	Attribute Name	Attribute Type	Description
1	latitude	Numeric	Position of aircraft latitude in decimal degrees with a range of -90 to 90.
2	longitude	Numeric	Position of aircraft longitude in decimal degrees with a range of -180 to 180.
3	altitude	Numeric	The aircraft's altitude is in feet, with a range of 0 to 43100 feet for commercial aircraft.
4	speed	Numeric	Aircraft ground speed in knots ranges from 0 to 400 knots for commercial aircraft.
5	heading	Numeric	The direction of the plane's longitudinal motion is seen from the nose of the aircraft in degrees from the north.
6	vspeed	Numeric	The speed of an airplane moving vertically uphill or downhill in feet per minute.
7	millisecond	Ordinal	Time/duration of the aircraft made the flight converted to milliseconds.

C. Data Preparation

Five subprocesses are carried out at the data preparation stage with the following explanation.

- Data cleaning is the stage of removing attributes that have empty values and NaN.
- Due to a large amount of missing data and having no value, an imputation technique is needed. The linear interpolation imputation technique is used to fill in missing data values between rows or data series [17].
- Remove duplicate data with the same value on each attribute.
- As seen in Table III, data are labeled as an anomaly or normal according to the time series of each aircraft's crash investigation report.

TABLE III
DATASETS LABELING BASED ON AIRCRAFT ACCIDENT INVESTIGATION REPORT

No	Aircraft Name	Callsign	Label Data	
			Normal	Anomaly
1	Sriwijaya Air SJ182	SJ182	1,125	68
2	Pakistan International Airlines PK8303	PK8303	0	1,097
3	S-76 Kobe Bryant Helicopter	N72EX	4,632	136
4	Aeroflot 1492	SU1492	1,355	3,529
5	Ethiopian Airlines 302	ET302	0	101
6	Atlas Air 5Y3591	5Y3591	13,435	189
7	Lion Air JT610	JT610	0	680
8	Southwest Airlines 1380	WN1380	267	217
9	Iran Aseman Airlines 3704	EP3704	0	3,328
10	Saratov Airlines 703	SOV703	0	1,520
11	Air France AF66	AFR066	261	176
12	MyCargo ACT Airlines TK6491	TK6491	0	232
13	Emirates 521	UAE521	626	3,421
14	FlyDubai 981	FDB981	3,822	20,382
		Total Data	**25,523**	**35,076**

- The sliding window in this process, as shown in Fig. 2, is described as an algorithm that controls data based on a sequence to fit various data into a single standard to avoid duplication and data loss [18].

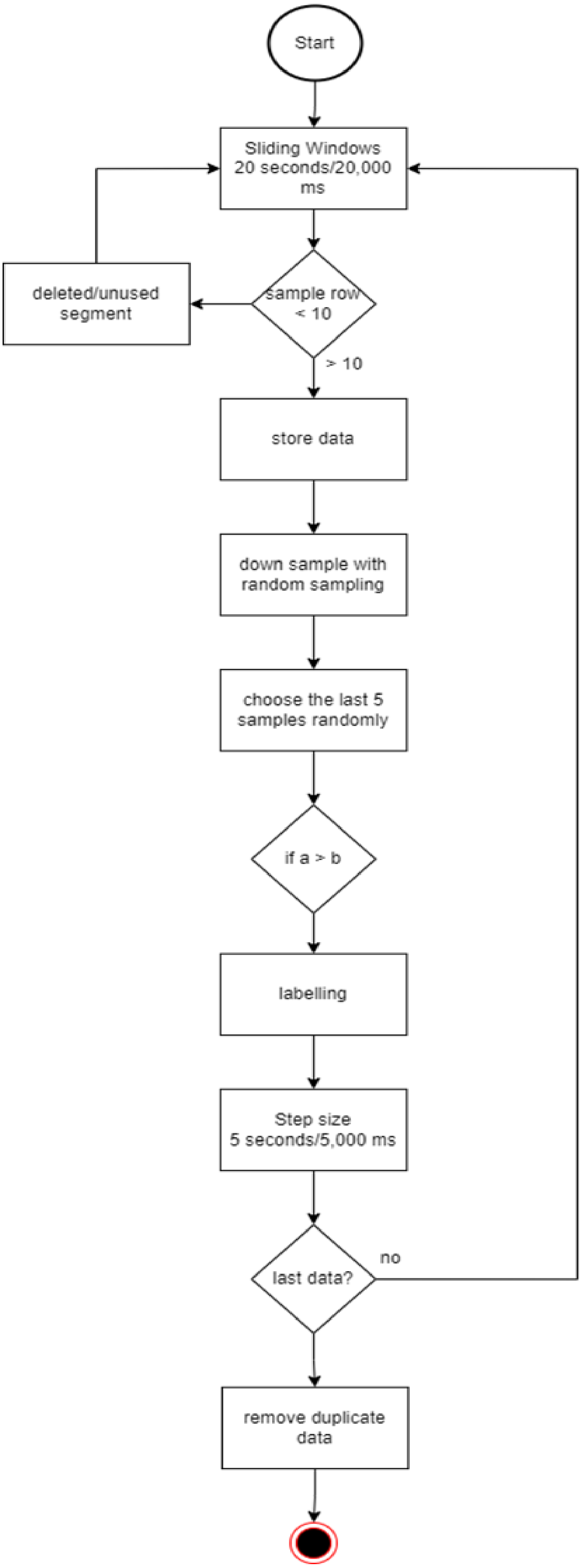

Fig. 2 Sliding window process

The meaning of the sliding window process in Fig. 2 is the sliding window parameter set to 20,000 ms (20 sec) with a step size of 5,000 ms (5 sec). Next, threshold parameters are installed in each segment. Ten data is minimum to be one segment. If there are less than 10 data in a segment, the data and segment are deleted. After that, the data from each stored segment will be carried out, and the sample is taken by random sampling (taken at random and adjusted in one segment with ten sample data). The last five sample data are taken in each segment, and a vote is taken to label all the segment data. Finally, data on all segments are combined. For each duplicated sample, data adjustments are made by selecting the label that appears most often and will be the primary data. From this algorithm sliding window process, it is expected that every two seconds (2,000 ms), there is at least one sample of data.

So that from the total anomaly data as much as 35,076 and normal data as much as 25,523 in Table III there was a change in the decrease in data after using the sliding window as shown in Table IV. Data with a dash symbol (-) is defined as insufficient data during the sliding window algorithm process so that the three datasets on the callsign aircraft ET302, WN1380, and AFR066 are not used in the next process.

TABLE IV
DATASETS LABELING BASED ON AIRCRAFT ACCIDENT INVESTIGATION REPORT AND IMPLEMENTATION SLIDING WINDOW

No	Aircraft Name	Callsign	Label Data	
			Normal	**Anomaly**
1	Sriwijaya Air SJ182	SJ182	1,049	74
2	Pakistan International Airlines PK8303	PK8303	0	761
3	S-76 Kobe Bryant Helicopter	N72EX	4,702	55
4	Aeroflot 1492	SU1492	1,236	3,558
5	Ethiopian Airlines 302	ET302	-	-
6	Atlas Air 5Y3591	5Y3591	13,239	80
7	Lion Air JT610	JT610	0	679
8	Southwest Airlines 1380	WN1380	-	-
9	Iran Aseman Airlines 3704	EP3704	0	3,281
10	Saratov Airlines 703	SOV703	0	1,310
11	Air France AF66	AFR066	-	-
12	MyCargo ACT Airlines TK6491	TK6491	0	213
13	Emirates 521	UAE521	590	3,305
14	FlyDubai 981	FDB981	3,821	20,355
		Total Data	**24,637**	**33,671**

D. Modeling

All the best models in previous studies, namely LSTM and GRU [12], [14] as well as the proposed model Bi-LSTM and Bi-GRU were used with the architectural design as follows.

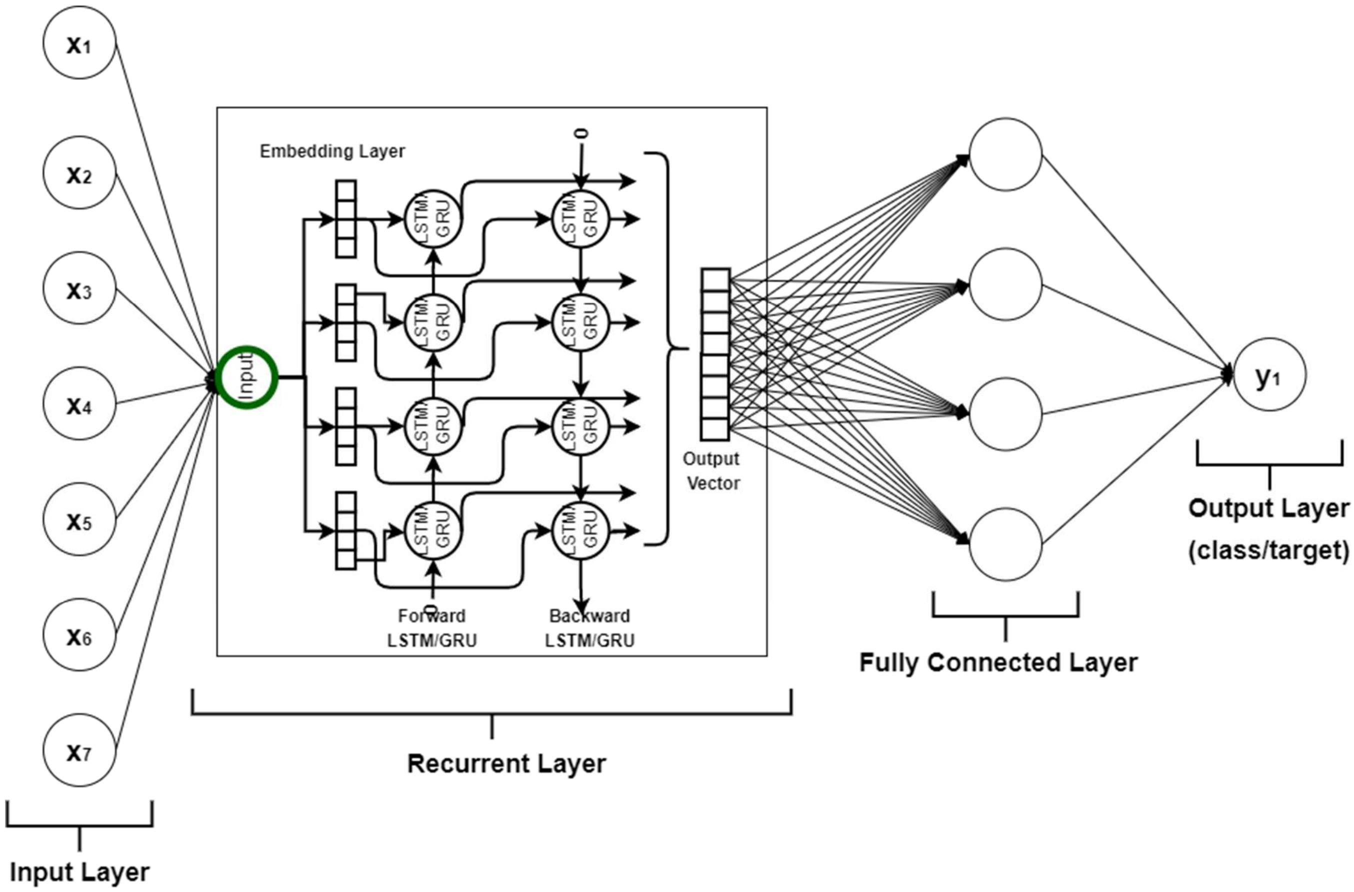

Fig. 3 Proposed deep learning model architecture

The DL architecture model proposed in this study, as shown in Figure 3 uses seven attributes at the input layer in vector order. First, the length of the vector is adjusted to the size of the attribute per row. The features used are altitude, heading, latitude, longitude, vertical speed, duration, and movement speed. Then it is determined that the sequence length or timestep is four neurons/units that are part of the hyperparameter in the number of rows. After the attributes are entered into the input layer, the data will be forwarded to a hidden layer (recurrent layer) using several models as experiments. The Bidirectional model has a different layer from the LSTM and GRU models because the recurrent layer in Bidirectional works in both forward and backward directions [19]. Every model determines seven neurons/unit dimensions as the number of hidden units. The hidden layer contains data that is processed using a fully connected layer with several hyperparameters, four neurons/unit. After that, the classification output in the output layer (y) is adjusted to the results with one neuron/unit or class, namely anomaly and normal.

E. Evaluation

At the evaluation stage, there are 58,308 data lines with 24,637 normal data and 33,671 anomaly data from 11 aircraft datasets used. Then, the dataset is broken down into training, validation, and test data. The training data uses a percentage of 60%, the percentage of validation data is 20%, and the test data is 20%. Next, the training data is used to train the model. Data validation is then used for hyperparameter adjustment according to Table V. Finally, test data is used to measure performance. The results of the training and validation phases are reported graphically as training loss, validation loss, training accuracy, and validation accuracy.

TABLE V
HYPERPARAMETER SUMMARY

No	Hyperparameter	Alternative Value
1	Input layer	Seven attributes in the form of a sequence vector
2	Recurrent layer	Hidden unit with seven dimensions
3	Fully connected layer	Neuron/unit (4)
4	Output layer	Neuron/unit (1)
5	Sequence/timesteps	4
6	Epochs	300 with EarlyStops function
7	Batch size	32 and 64
8	Loss function	binary_crossentropy
9	Optimizer	Adam and SGD
10	Learning rate	0.001 and 0.0001

After the data is trained, the new data is entered into the validation data for the evaluation process using the hyperparameters set in Table V. For example, in the early stop or early stop function, validation accuracy training is monitored using patience, set to a value of 10 (patience = 10). The patience parameter will stop training when there is no increase in the validation accuracy value in the iteration epoch. Finally, the data is tested to evaluate the performance of the model from the output layer. In determining its performance, it is necessary to compare the classification of test data with its basic truth. Therefore, the comparison of the results of the classification and testing of ground truth data is tabulated into a confusion matrix, illustrated in Fig. 4 [20].

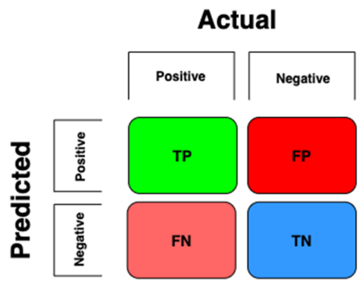

Fig. 4 Confusion matrix

The results are analyzed between certain numbers and basic truth numbers using a confusion matrix [21]. The confusion matrix calculation is divided into four stages: accuracy, precision, recall, and f1-score, which can be explained as the following equation.

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{1}$$

Equation 1 is used to consider the classification accuracy in measuring the prediction performance of the data.

$$precision = \frac{TP}{TP+FP} \tag{2}$$

In addition to accuracy, the results need to be analyzed further to strengthen the performance measure of data using equation 2. Precision is a technique to measure the level of the positive observation ratio that is predicted correctly.

$$recall = \frac{TP}{TP+FN} \tag{3}$$

The next step is to measure the recall value using equation 3, where the ratio of positive observations is correctly predicted to all statements in the actual class.

$$f1 - score = 2 \bullet \frac{precision \bullet recall}{precision+recall} \tag{4}$$

From all the equation techniques in the confusion matrix, measurements using the f1-score in equation 4 will produce a balanced value performance between calculating the average precision and recall.

III. RESULTS AND DISCUSSION

The four models namely LSTM, GRU, Bi-LSTM, and Bi-GRU were evaluated for train and validation using the hyperparameters tuning shown in Table V. Each model was carried out several experiments to find the best model from the results of the training performance evaluation and validation as well as a comparison of the results obtained to be carried out in the evaluation of the test. The total data for the 11 ADSB datasets is 58,308 rows of data divided into training, validation, and test data. The percentages used are 60% for training data of 34,981 data, 20% for validation data of 11,661 data, and 20% for test data of 11,661 data. Training data is used to train the model. Then the validation data is used to optimize the hyperparameters. Finally, test data are used to measure performance. The training and validation phases are reported in the form of graphs of training loss, validation loss, training accuracy, and validation accuracy.

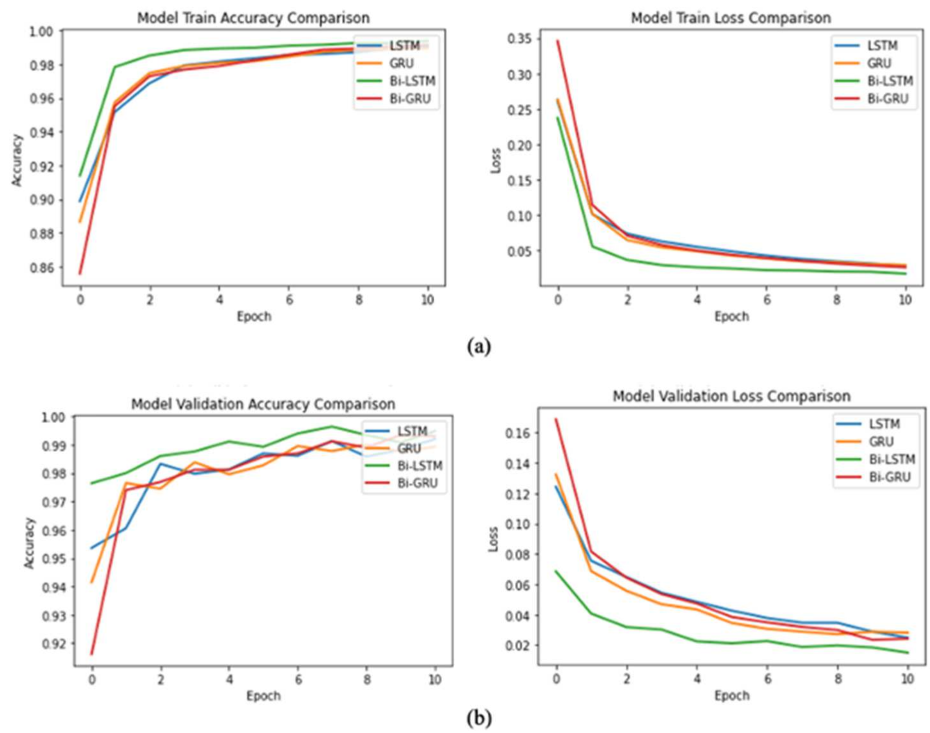

Fig. 5 The best evaluation model train (a) and validation (b) comparison

The training and validation performance results in a comparison visualization can be seen in Fig. 5, describing each model's uses early stopping, which looks at changes in validation performance. This function will stop the training of each model when there is no performance improvement and loss reduction [22]. The best results of the train and validation of the complete model visualization evaluation are presented in Table VI.

TABLE VI
THE BEST TRAIN AND VALIDATION RESULT

Model	Train		Validation	
	Loss	Accuracy	Val Loss	Val accuracy
LSTM	2.85%	99.06%	2.48%	99.20%
GRU	2.97%	98.95%	2.81%	98.93%
Bi-LSTM	**1.76%**	**99.38%**	**1.50%**	**99.49%**
Bi-GRU	2.64%	99.15%	2.41%	99.31%

The best training and validation place the Bi-LSTM model in the first position, followed by Bi-GRU in the second position, LSTM in the third position, and the GRU in the last position. Then performed a performance evaluation with the confusion matrix presented in Table VII.

TABLE VII
THE BEST TESTING RESULT

Model	Testing (Confusion Matrix)			
	Accuracy	Precision	Recall	F1-score
LSTM	99.29%	99.68%	99.09%	99.38%
GRU	98.79%	98.41%	99.5%	98.95%
Bi-LSTM	**99.44%**	**99.48%**	**99.55%**	**99.51%**
Bi-GRU	99.23%	99.39%	99.28%	99.33%

Table VII describes the best performance evaluation obtained by the proposed model: Bi-LSTM, with an overall accuracy of 99.44% and an f1-score of 99.51%. Second place was acquired by LSTM model, which rises to second place with an overall accuracy of 99.29% and an f1-score of 99.38%. The difference between the Bi-GRU model and the LSTM has slight difference, where the overall accuracy obtained is 99.23%, and the f1-score is 99.33% placing the Bi-GRU model in third place. GRU occupies the fourth position with an overall accuracy of 98.79% and an f1-score of 98.95%.

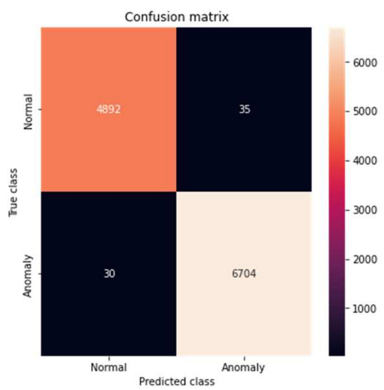

Fig. 6 Predicted class result Bi-LSTM model

In Figure 6, the Bi-LSTM model can predict optimally, where 11,661 test data evaluated resulted in a misclassification of 35 data from normal data to anomaly data and 30 from anomaly data to normal data.

IV. CONCLUSION

This study proposed the DL models to detect aircraft flight movement anomalies using ADS-B data. Various models have been trained, evaluated, and tested using ADS-B data.

The Bi-LSTM model has the best performance evaluation results, with an overall accuracy of 99.44% and an f1-score of 99.51%. Which of these models performs better than the results of each of the previous models, where the dataset used is simpler than the data used in this study. The results of this study are promising to be applied to the aviation industry because the ADS-B device can be used as a backup radar in monitoring and detecting aircraft movement anomalies. In addition, for future research, the model can be implemented on ADS-B monitoring server to generate reports as material for aircraft technician studies to make decisions about the feasibility of the aircraft on the next flight in preventing and reducing the rate of aircraft accidents. The dataset in this study can be accessed for future comparison studies on the Flightradar24 community server.

ACKNOWLEDGMENT

We thank Computer Science Department, BINUS Graduate Program - Master of Computer Science, Bina Nusantara University for funding this research.

REFERENCES

[1] A. A. Ajhari, R. Ibrahim, A. Pramodana, J. S. Pramudito, J. R. Tasyam, and W. Hilmy, "ADS-B Mobile Ground Station Receiver Flight Surveillance Architecture," in *2021 International Conference on Artificial Intelligence and Computer Science Technology (ICAICST)*, 2021, pp. 174–178.

[2] V. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artif. Intell. Rev.*, vol. 22, no. 2, pp. 85–126, 2004.

[3] E. Habler and A. Shabtai, "Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages," *Comput. Secur.*, vol. 78, pp. 155–173, 2018.

[4] T. Li, B. Wang, F. Shang, J. Tian, and K. Cao, "ADS-B Data Attack Detection Based on Generative Adversarial Networks," in *International Symposium on Cyberspace Safety and Security*, 2019, pp. 323–336.

[5] J. Wang, Y. Zou, and J. Ding, "ADS-B spoofing attack detection method based on LSTM," *EURASIP J. Wirel. Commun. Netw.*, vol. 2020, no. 1, pp. 1–12, 2020.

[6] T. Li, B. Wang, F. Shang, J. Tian, and K. Cao, "Dynamic temporal ADS-B data attack detection based on sHDP-HMM," *Comput. Secur.*, vol. 93, p. 101789, 2020.

[7] M. Y. Pusadan, J. L. Buliali, and R. V. H. Ginardi, "Anomaly detection of flight routes through optimal waypoint," in *Journal of Physics: Conference Series*, 2017, vol. 801, no. 1, p. 12041.

[8] M. Y. Pusadan, J. L. Buliali, and R. V. H. Ginardi, "Anomaly detection on flight route using similarity and grouping approach based-on automatic dependent surveillance-broadcast," *Int. J. Adv. Intell. Informatics*, vol. 5, no. 3, pp. 285–296, 2019.

[9] M. Y. Pusadan, J. L. Buliali, and R. V. H. Ginardi, "Cluster Phenomenon to Determine Anomaly Detection of Flight Route," *Procedia Comput. Sci.*, vol. 161, pp. 516–526, 2019.

[10] J. Sun, J. Ellerbroek, and J. Hoekstra, "Flight extraction and phase identification for large automatic dependent surveillance–broadcast datasets," *J. Aerosp. Inf. Syst.*, vol. 14, no. 10, pp. 566–572, 2017.

[11] M. Leonardi, L. Di Gregorio, and D. Di Fausto, "Air traffic security: Aircraft classification using ADS-B message's phase-pattern," *Aerospace*, vol. 4, no. 4, p. 51, 2017.

[12] D. Wu *et al.*, "Custom machine learning architectures: towards realtime anomaly detection for flight testing," in *2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2018, pp. 1323–1330.

[13] Z. Cao *et al.*, "Improving prediction accuracy in LSTM network model for aircraft testing flight data," in *2018 IEEE International Conference on Smart Cloud (SmartCloud)*, 2018, pp. 7–12.

[14] A. Nanduri and L. Sherry, "Anomaly detection in aircraft data using Recurrent Neural Networks (RNN)," in *2016 Integrated Communications Navigation and Surveillance (ICNS)*, 2016, pp. 5C2-1.

[15] A. Y. Nuryantini and B. W. Nuryadin, "Learning vector of motion using FlightRadar24 and Tracker motion analysis," *Phys. Educ.*, vol. 55, no. 1, p. 15019, 2019.

[16] G. Nguyen *et al.*, "Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 77–124, 2019.

[17] S. Daberdaku, E. Tavazzi, and B. Di Camillo, "A combined interpolation and weighted K-nearest neighbours approach for the imputation of longitudinal ICU laboratory data," *J. Healthc. Informatics Res.*, vol. 4, no. 2, pp. 174–188, 2020.

[18] D. M. Burns and C. M. Whyne, "Seglearn: A python package for learning sequences and time series," *J. Mach. Learn. Res.*, vol. 19, no. 1, pp. 3238–3244, 2018.

[19] Y. Cao, J. Cao, Z. Zhou, and Z. Liu, "Aircraft Track Anomaly Detection Based on MOD-Bi-LSTM," *Electronics*, vol. 10, no. 9, p. 1007, 2021.

[20] K. M. Ting, "Confusion matrix.," *Encycl. Mach. Learn. data Min.*, vol. 260, 2017.

[21] P. Bintoro and A. Harjoko, "Lampung Script Recognition Using Convolutional Neural Network," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 16, no. 1, pp. 23–34, 2022, [Online]. Available: https://jurnal.ugm.ac.id/ijccs/article/view/70041/33160

[22] Y. Bai *et al.*, "Understanding and Improving Early Stopping for Learning with Noisy Labels," *Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, [Online]. Available: https://proceedings.neurips.cc/paper/2021/file/cc7e2b878868cbae992d1fb743995d8f-Paper.pdf.