GATS: A Time-Series Dataset for Addressing General Aviation Flight Safety

Aidan LaBella [1 *] **Aditya Iyer** [1 *] **Charles Duong** [1 *] **Nathan DePiero** [1 *] **Pak Iong Long** [1 *] **Elise Carman** [1 *]
Randall Balestriero [1] **Travis Desell** [2]

Abstract

Assessing general aviation operations has become critical for improving the safety of airspace systems worldwide. Yet, machine learning research in the domain is nearly nonexistent due to the extremely limited amount of publicly available flight data. To encourage research in airspace safety and general aviation as a whole, we release GATS, a dataset comprising more than 7,000 flights anonymously sampled with permission from the privately held US National General Aviation Flight Information Database (NGAFID), corresponding to 10,641 total hours of data recordings. This dataset sets itself apart from previous works with its inclusion of 2 new aircraft types and 76 different flight data sensor parameters, including navigational information and aircraft orientation. We benchmark this dataset on 2 aviation-domain tasks. The first is aircraft classification, a proof-of-concept problem to establish that advanced machine learning methods can be applied effectively on time-series flight data. The second is missing data reconstruction, a more rigorous safety-critical task necessary in real-world environments where sensors can fail and information must be restored for flight analysis purposes. We achieved near-perfect accuracy on the aircraft classification task, but failed to generate meaningful reconstructions on the missing data task. The poor performance on the second task with the chosen models indicates the opportunity for future research into better techniques for understanding and improving flight safety using this dataset. [1]

1. Introduction

General aviation safety remains a critical concern as the National Airspace System (NAS) continues to grow in size and complexity. Between 2002 and 2022, the United States recorded 1,205 general aviation accidents, of which 214 were fatal (National Transportation Safety Board, 2024). These accidents result from a complex interaction of factors such as pilot error, mechanical failures, adverse weather, and aging aircraft. Improving safety in this domain requires more than basic sensor flags and post-flight debriefings. It calls for data-driven methods that can learn from extensive flight data and anticipate risks before they materialize.

Beyond accident prevention, modern machine learning has the potential to transform aviation more broadly. Detailed flight representations can enable predictive modeling for enhanced warnings during flights, detailed and tailored analytics for post-flight debriefing, custom and automated planning for air traffic control, and general improvement in the efficiency and safety of aviation operations.

However, there is a substantial lack of publicly available data sources for general aviation. Most flight data is collected by manufacturers, governments, or flight schools and is typically unavailable because it is proprietary or subject to pilot privacy concerns. For educational training flights, federal regulations such as the Family Educational Rights and Privacy Act (FERPA, 20 U.S.C. 1232g) requires that the identification information (e.g., operator name, location and aircraft tail number) is anonymized before any student-related data can be shared. This legal constraint, while essential for protecting privacy, further limits the availability of open datasets. The data that is available tends to be task or aircraft-specific, limiting generalizability. This scarcity has inhibited the development of machine learning for the domain.

To address this data shortage problem, we introduce **GATS**, an anonymous, FERPA-compliant **G**eneral **A**viation **T**ime-**S**eries dataset of 10,641 flight hours from 7,679 student pilot training flights across 97 propeller-driven aircraft on 3 different airframe types. Each flight in GATS includes various features that describe key aspects such as engine performance, aircraft orientation, environmental conditions, navigation information, and flight instrument readings. Together, these categories provide a holistic view of each flight and enable more precise analysis of aircraft behavior and conditions that lead to incidents.

In our study, we provide 2 aviation-relevant benchmark tasks that help establish the validity and complexity of this dataset:

[1] **Huggingface Dataset (Anonymous Author Version Link):** https://huggingface.co/datasets/NGAFID2025ICML/NGAFID-LOCI-GATS-Anonymous

aircraft classification and missing data reconstruction. Each task serves a different purpose. Aircraft classification is to help the audience understand how machine learning is applied on aviation data. Missing data reconstruction is a critical task that addresses sensor failures, which is a real-world condition that can negatively affect an aircraft's flight and accident analysis.

These experiments evaluated both self-supervised learning (SSL) and supervised learning techniques. The SSL models employed masked autoencoders and contrastive learning, both of which are well-suited to learning robust feature representations of a flight that can be used in flight analysis and other downstream tasks that do not pertain to specific labels. The supervised model utilized a deep neural network architecture (Yang et al., 2022; Xiao et al., 2020), which is effective for task-specific applications when labeled data are available. Together, these approaches offer a comprehensive assessment of how the dataset can be applied across different machine learning paradigms.

Our contributions can be summarized accordingly:

- We introduce GATS, a novel general aviation dataset consisting of thousands of hours of flight data across 3 aircraft platforms.
- We provide an anonymized, FERPA-compliant dataset in both raw and preprocessed versions, with documentation and implementation details to ensure interpretability and reproducibility.
- We build 2 SSL models based on contrastive learning and masked autoencoder techniques.
- We evaluate the dataset on 2 domain-relevant tasks using both the constructed SSL models and pre-existing supervised learning frameworks.

This dataset addresses the shortage of public data in the aviation domain, encouraging advancement in the aviation field and enabling a new domain to evaluate cutting-edge machine learning techniques on time-series data.

2. Related Works

2.1. National General Aviation Flight Information Database (NGAFID)

The NGAFID (LaBella et al., 2022) is a FAA-funded project where any aviator can upload and analyze their own flight data.

Previous works with NGAFID data mainly focus on identifying anomalous flight and aircraft performance conditions, as well as tasks like phase of flight and approach type classification. (LaBella et al., 2022) implemented a novel method for identifying dangerous flight conditions, such as Loss of Control In-Flight (LOC-I) (Balogh, 2017) and aerodynamic stalls (Administration, 2009). Using subject matter expert validation and a controlled test flight, this work showed that it was possible to calculate probabilistic metrics indicative of the risk that a flight will experience one of these dangerous conditions at any given timestep.

(Yang et al., 2022) leveraged mining aircraft logbook records (Akhbardeh et al., 2021) to create training data for aircraft predictive maintenance, namely the Cessna 172S airframe. This work utilized a convolutional multi-headed self-attention model (ConvMHSA) (Xiao et al., 2020) to output the probability that a flight occurred before or after maintenance. This work also produced an associated dataset (Yang & Desell, 2022) with 28 sensor parameters, unlike GATS, which contains 76 parameters. Moreover, this dataset only contains flight data from one aircraft type (the Cessna 172S), unlike GATS, which contains 3 different aircraft types, including the PA-28-181 and PA-44-180. Additionally, this dataset contains only aircraft engine parameters, which do not allow for the analysis of flight safety with regard to pilot inputs.

There have also been works that utilize NGAFID data for phase of flight and approach type identification. (Karboviak et al., 2018) developed a method to classify the type and quality of the approach of a flight in the NGAFID. (Lyu et al., 2024) utilized minimally-supervised learning and self-organized maps (MS-SOMs)(Kohonen, 1990) to predict the phase of flight with limited labeled data. Work by (Clachar, 2015) looked into using supervised and unsupervised approaches for finding atypical flight patterns in NGAFID data, i.e. anomaly detection.

2.2. Time-Series Data Problems

Classification Time-series classification (Ismail Fawaz et al., 2019) is niche in the machine learning community, especially when compared to the vast amount of literature on vision and language tasks. Current deep learning approaches include LSTM-RNNs (Karim et al., 2017; Kong et al., 2025), CNNs (Wang et al., 2017; Ismail Fawaz et al., 2020), and convolutional transformers (Yang et al., 2022). Many time-series classification tasks come from the medical community (Huang et al., 2024), for tasks such as ECG data classification (Gupta et al., 2024; Sakib et al., 2023). Another common example of time-series data found in the literature is related to stock market data, such as stock price classification (Kumari et al., 2024), risk analysis (Petchpol & Boongasame, 2025), and trading behavior analysis (Kong et al., 2020).

Forecasting Time-series forecasting (Lim & Zohren, 2021) has been a popular task for many applications, such

as coal power plant parameter prediction (Lyu et al., 2021), air quality forecasting (Anggraini et al., 2024), and weather dynamics (Zhang et al., 2024). Forecasting has the potential to be applied to NGAFID for tasks such as engine performance indicators, weather metrics (such as wind speeds), and pilot input readings. Such tasks have been accomplished before with evolutionary neural networks using NGAFID Cessna 172S engine data (Desell et al., 2014; 2020; ElSaid et al., 2021), including methods to increase explainability (Murphy et al., 2024).

Regression Regression tasks, such as missing column reconstruction, can be a very useful task for many downstream applications where missing and/or noisy sensor readings occur frequently. Past works have utilized electrical signals to reconstruct data for arc-fault detection (Jiang & Zheng, 2023), remote sensing (Zhou et al., 2023), and molecular biology (Laporte et al., 2024). Parameter regression has the potential to be used in NGAFID data because many parameters have missing readings. Other potential applications could include regressing parameters that are not recorded by the aircraft's flight data recorder, such as flap settings or rudder inputs.

Anomaly Detection Previous works have looked at transformer models with data from the physical sciences (Xu et al., 2021), graph-attention networks (Zhao et al., 2020), as well as autoencoders and RNN-based models (Delibasoglu & Heintz, 2024). Anomaly detection could be a useful tool for NGAFID data as it can identify potentially unsafe flying patterns that are not already found by classification or forecasting models.

While there are many existing tasks that utilize time-series data for classification, forecasting, regression, and anomaly detection, the NGAFID data stands out because it is subject to the external physical environment, as well as human inputs and mechanical processes. There exist very few other datasets that are dependent on these factors, making the NGAFID data unique.

3. A Novel General Aviation Dataset

GATS was extracted directly from the NGAFID, particularly for two US flight schools over a two-month period and consists of 10,641 flight hours from 7,679 flights spanning 3 airframe types. Only NGAFID developers and administrators are able to access all of its data, as individual users can only see their own data. In cooperation with the NGAFID and its stakeholders, we obtained permission to organize a subset of the NGAFID's data, with operator-identifying attributes removed for privacy.

The NGAFID's main feature, event detection, provides a set of 20 dangerous event definitions. The occurrences of these events are logged on a per flight basis.

3.1. Data Collection Methodology

The data collected for this work mostly consists of student pilot training and evaluation flights. Due to FERPA, any identifying information such as operator name, location and/or tail number is anonymized. Additionally, this data was pulled directly from the NGAFID resulting in extra columns, such as `Stall Index` and `Density Ratio`, which are calculated based on raw sensor readings. These extra columns are used by the LOC-I and Stall Index event calculations (LaBella et al., 2022; Balogh, 2017) to derive probabilistic metrics that indicate the likelihood that the event will occur.

3.2. Dataset Statistics and Analysis

Aircraft Types As shown in Appendix B.2, this dataset is comprised of three aircraft types: Cessna 172S, PA-28-181, and PA-44-180, depicted in Appendix B.1. While the Cessna 172S is high wing, both PA-28-181 and PA-44-180 are low wing. Additionally, Cessna 172S and PA-28-181 have a single engine while PA-44-180 has two. Eleven of the columns in the raw column set correspond to the PA-44-180's second engine.

Potentially Dangerous In-Flight Events The LOC-I events in this dataset correspond to special incidents within a flight. Our dataset comprises 8,802 events occurrences of the 20 event types listed in Appendix G.1. Additionally, Appendix G.4 shows the distribution of the number of events per flight. Of the 7,679 flights, only 2,616 experience one or more events and only 1,525 experience two or more.

Appendix G.6 shows the most prevalent event types across the dataset. The foremost being Low Airspeed on Climbout, which occurs in just over 1,250 of the 2,616 eventful flights. This event describes aircraft takeoffs with unexpectedly low airspeed. Another frequent event is High Altitude Stalls, which describe the occurrence of an aircraft stall above a threshold altitude. These occur in around 1,100 flights. Overall, these events are more likely to occur in GATS due to the dataset comprising of only training flights. Specifically, in training, student pilots practice many techniques including High Altitude Stalls, thus leading to higher frequency counts. The remaining event types are half as common, occurring in 500 flights or less.

Appendix G.5 examines the number of occurrences of high altitude stall events in flights. Unlike most other events which occur at most 6 times throughout a flight, high altitude stalls can occur at most 47 times.

Flight lengths The flights in the dataset have a variety of durations, ranging from 3 minutes to 9 hours. The distribution of these flight lengths is pictured, in Figure 3.

NaN Values The columns in the dataset represent different sensors on the aircraft and their readings. These may contain NaN values because of sensor failures at certain timesteps, or from missing data fields for specific aircraft types. Overall, the dataset contains 2.49% percent NaN values. These NaN values are concentrated in a subset of the columns, as seen in Appendix H.

Flight Count per Column Because different aircraft are equipped with different sensor configurations, the set of recorded columns varies across flights and is not standardized. The occurrences of columns by number of flights is pictured in Appendix F.2.

3.3. Preprocessing

Because of the previously described aspects of the dataset, we provide both a preprocessing script and preprocessed version of the dataset. This allows for a set of flights with uniform columns and length without any NaN values. A more detailed preprocessing methodology can be found in Appendix C.

3.4. Models and Benchmarks

To demonstrate the practical value and versatility of this dataset, we establish two critical benchmark tasks in aviation: airframe classification and masked column regression. The first benchmark assesses whether the dataset supports simple yet meaningful classification problems. Airframe classification is divided into two subtasks—airframe type classification, which identifies the specific aircraft platform among three possible types, and airframe class classification, which distinguishes between single-engine and multi-engine aircraft. These tasks reflect core capabilities needed for aircraft identification in real-world applications, such as maintenance scheduling, training simulations, and automated air traffic control systems. The second benchmark, masked column regression, focuses on reconstructing missing sensor data—a frequent and high-stakes problem in aviation where sensor failures can impair both post-flight analysis and in-flight decision-making. Accurate reconstruction of missing values supports flight understanding and enhances the reliability of downstream safety systems. Details about setup, models, and results can be found in Appendix D.

4. Limitations

This dataset provides a valuable interface for machine learning researchers to study general aviation flight data and have a direct impact on flight safety analysis. However, the dataset itself includes limitations that future users must be aware of. The most prominent are the derived columns. As stated in Section D.2, several feature columns in this dataset are calculated from other feature columns. For instance, simple angle of attack (`AOASimple`) is directly calculated from columns such as pitch (`pitch`). Thus, machine learning models could shortcut learning the underlying representation of the data and can instead learn the simple correspondences. A full table of the correspondences can be viewed in Appendix E. A second limitation of this data is the large number of NaN values in the feature columns. Because this is real-world flight data, sensors may not always be operational, and so we invite future researchers to use their own preprocessing methods to fill these values in more advanced or appropriate ways in their own works. Another limitation of the dataset was that certain sensor fields are inherently different between aircraft. Lastly, a limitation is this was general aviation flight training data and not commercial aviation data. This means that this dataset applies to researchers studying general aviation problems, with the caveat that this data applies more directly to student pilots, rather than trained aviators.

5. Conclusions and Future Work

This paper addresses the lack of publicly available aviation data by introducing GATS, a novel and expansive time-series dataset of FERPA-compliant real-world flight recordings. We evaluate the data on two benchmark tasks, showcasing that machine learning is a valid method for analyzing this data, and providing an example task that would have a direct impact on improving flight safety.

We identify the aircraft classification tasks to be a simple problem, and the performances for both supervised and self-supervised models tested were very high for aircraft type classification, with accuracy divergence only happening in aircraft class classification. These results indicate the compatibility of the GATS dataset in machine learning frameworks. On the other hand, we observe that the missing data reconstruction task is challenging for both the self-supervised learning models tested, regardless of the training parameters (varying mask lengths and ratios). While the masked autoencoder method performed better than the contrastive technique, they both ultimately failed to accurately restore the data. Thus, it remains unclear what the optimal model solution is for this regression task, and future research efforts should be made to solve this problem with alternative architectures. A few potential methods that researchers could experiment with include attention (Vaswani et al., 2023), joint-embedding based architectures (Assran et al., 2023; Bardes et al., 2024), and non-contrastive SSL techniques (Oquab et al., 2024).

We encourage researchers to explore the GATS dataset as a benchmark for developing and evaluating machine learning methods on real-world, safety-critical time-series data.

References

Administration, U. S. F. A. *Pilot's Handbook of Aeronautical Knowledge*. Skyhorse Publishing Inc., 2009.

Ahunt. 1957 cessna 172 skyhawk photograph. https://commons.wikimedia.org/wiki/File:Cessna172Skyhawk1957model01.jpg, 2005. Public domain.

Akhbardeh, F., Alm, C. O., Zampieri, M., and Desell, T. Handling extreme class imbalance in technical logbook datasets. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 4034–4045, 2021.

Anggraini, L., Noersasongko, E., Purwanto, and Marjuni, A. Evaluation of time series forecasting techniques for air quality prediction: Case study of no2 levels. *2024 International Seminar on Application for Technology of Information and Communication (iSemantic)*, pp. 522–527, 2024. URL https://api.semanticscholar.org/CorpusID:274372458.

Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabbat, M., LeCun, Y., and Ballas, N. Self-supervised learning from images with a joint-embedding predictive architecture, 2023. URL https://arxiv.org/abs/2301.08243.

Balogh, S. Use of foqa data to estimate the probability of vehicle upset or loss of control in-flight. Master's thesis, University of North Dakota, Grand Forks, ND, USA, 2017.

Bardes, A., Garrido, Q., Ponce, J., Chen, X., Rabbat, M., LeCun, Y., Assran, M., and Ballas, N. Revisiting feature prediction for learning visual representations from video, 2024. URL https://arxiv.org/abs/2404.08471.

Bordes, F., Balestriero, R., Garrido, Q., Bardes, A., and Vincent, P. Guillotine regularization: Why removing layers is needed to improve generalization in self-supervised learning. *arXiv preprint arXiv:2206.13378*, 2022.

Burdett, M. Piper pa-28-180 cherokee c over northrepps. https://commons.wikimedia.org/wiki/File:G-AVRZ_Piper_PA-28-180_Cherokee_C_at_Northrepps.jpg, 2021. Creative Commons Attribution–ShareAlike 2.0 Generic (CC-BY-SA 2.0).

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. A simple framework for contrastive learning of visual representations. In III, H. D. and Singh, A. (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 1597–1607. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/chen20j.html.

Clachar, S. A. Identifying and analyzing atypical flights by using supervised and unsupervised approaches. *Transportation research record*, 2471(1):10–18, 2015.

Delibasoglu, I. and Heintz, F. Time series anomaly detection leveraging mse feedback with autoencoder and rnn. In *Time*, 2024. URL https://api.semanticscholar.org/CorpusID:273506497.

Desell, T., Clachar, S., Higgins, J., and Wild, B. Evolving neural network weights for time-series prediction of general aviation flight data. In *Parallel Problem Solving from Nature–PPSN XIII: 13th International Conference, Ljubljana, Slovenia, September 13-17, 2014. Proceedings 13*, pp. 771–781. Springer, 2014.

Desell, T., ElSaid, A., and Ororbia, A. G. An empirical exploration of deep recurrent connections using neuroevolution. In *Applications of Evolutionary Computation: 23rd European Conference, EvoApplications 2020, Held as Part of EvoStar 2020, Seville, Spain, April 15–17, 2020, Proceedings 23*, pp. 546–561. Springer, 2020.

ElSaid, A., Karns, J., Lyu, Z., Ororbia, A. G., and Desell, T. Continuous ant-based neural topology search. In *International Conference on the Applications of Evolutionary Computation (Part of EvoStar)*, pp. 291–306. Springer, 2021.

Feichtenhofer, C., Fan, H., Li, Y., and He, K. Masked autoencoders as spatiotemporal learners, 2022. URL https://arxiv.org/abs/2205.09113.

Gupta, P., Murugan, A. A., Chordia, D., Yannam, P. K. R., and Gupta, M. A comparative study of few-shot learning methods for 1-d ecg time-series classification. In *2024 IEEE 12th International Conference on Healthcare Informatics (ICHI)*, pp. 604–611. IEEE, 2024.

He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition, 2015. URL https://arxiv.org/abs/1512.03385.

Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Huang, N., Wang, H., He, Z., Zitnik, M., and Zhang, X. Repurposing Foundation Model for Generalizable Medical Time Series Classification, October 2024. URL http://arxiv.org/abs/2410.03794. arXiv:2410.03794 [cs].

Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., and Muller, P.-A. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, 33(4):917–963, March 2019. ISSN 1573-756X. doi: 10.1007/s10618-019-00619-1. URL http://dx.doi.org/10.1007/s10618-019-00619-1.

Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., Webb, G. I., Idoumghar, L., Muller, P.-A., and Petitjean, F. Inceptiontime: Finding alexnet for time series classification. *Data Mining and Knowledge Discovery*, 34(6):1936–1962, 2020.

Jiang, R. and Zheng, Y. Series arc fault detection using regular signals and time-series reconstruction. *IEEE Transactions on Industrial Electronics*, 70:2026–2036, 2023. URL https://api.semanticscholar.org/CorpusID:248142366.

Karboviak, K., Clachar, S., Desell, T., Dusenbury, M., Hedrick, W., Higgins, J., Walberg, J., and Wild, B. Classifying aircraft approach type in the national general aviation flight information database. In *Computational Science–ICCS 2018: 18th International Conference, Wuxi, China, June 11–13, 2018, Proceedings, Part I 18*, pp. 456–469. Springer, 2018.

Karim, F., Majumdar, S., Darabi, H., and Chen, S. Lstm fully convolutional networks for time series classification. *IEEE access*, 6:1662–1669, 2017.

KGG1951. Piper PA–44 seminole (lz-fto) over bulgaria. https://commons.wikimedia.org/wiki/File:Piper-pa-44.jpg, 2010. Image licensed under CC BY-SA 3.0.

Kohonen, T. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990.

Kong, A., Azencott, R., and Zhu, H. Pattern recognition in trading behaviors before stock price jumps: new method based on multivariate time series classification. *arXiv: Statistical Finance*, 2020. URL https://api.semanticscholar.org/CorpusID:226290158.

Kong, Y., Wang, Z., Nie, Y., Zhou, T., Zohren, S., Liang, Y., Sun, P., and Wen, Q. Unlocking the power of lstm for long term time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 11968–11976, 2025.

Kumari, B. A., Manju, K., Amrutha, M., Amrutha, M., and Neeraja, C. Time series data classification for precise stock market price prediction using ml. *2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)*, pp. 1–6, 2024. URL https://api.semanticscholar.org/CorpusID:269241138.

LaBella, A. P., Karns, J. A., Akhbardeh, F., Desell, T., Walton, A. J., Morgan, Z., Wild, B., and Dusenbury, M. Optimized flight safety event detection in the national general aviation flight information database. In *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*, pp. 1570–1579, 2022.

Laporte, M. H., Gambarotto, D., Bertiaux, E., Bournonville, L., Louvel, V., Nunes, J. M., Borgers, S., Hamel, V., and Guichard, P. Time-series reconstruction of the molecular architecture of human centriole assembly. *Cell*, 187:2158 – 2174.e19, 2024. URL https://api.semanticscholar.org/CorpusID:269034316.

Lim, B. and Zohren, S. Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A*, 379(2194):20200209, 2021.

Lyu, Z., Patwardhan, S., Stadem, D., Langfeld, J., Benson, S., Thoelke, S., and Desell, T. Neuroevolution of recurrent neural networks for time series forecasting of coal-fired power plant operating parameters. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 1735–1743, 2021.

Lyu, Z., Thapa, P., and Desell, T. Minimally supervised topological projections of self-organizing maps for phase of flight identification. In *2024 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2024.

Maaten, L. v. d. and Hinton, G. Visualizing data using t-SNE. 9(86):2579–2605. URL http://jmlr.org/papers/v9/vandermaaten08a.html.

Murphy, J., Kar, D., Karns, J., and Desell, T. Exa-gp: Unifying graph-based genetic programming and neuroevolution for explainable time series forecasting. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 523–526, 2024.

National Transportation Safety Board. Civil aviation dashboard, 2024. URL https://www.ntsb.gov/safety/StatisticalReviews/Pages/CivilAviationDashboard.aspx. Accessed: 2024-07-02.

Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Assran, M., Ballas, N., Galuba, W., Howes, R., Huang, P.-Y., Li, S.-W., Misra, I., Rabbat, M., Sharma,

V., Synnaeve, G., Xu, H., Jegou, H., Mairal, J., Labatut, P., Joulin, A., and Bojanowski, P. Dinov2: Learning robust visual features without supervision, 2024. URL https://arxiv.org/abs/2304.07193.

Petchpol, K. and Boongasame, L. Enhancing predictive capabilities for identifying at-risk stocks using multivariate time-series classification: A case study of the thai stock market. *Applied Computational Intelligence and Soft Computing*, 2025. URL https://api.semanticscholar.org/CorpusID:278320898.

Robertson, S. Understanding inverse document frequency: on theoretical arguments for idf. *Journal of documentation*, 60(5):503–520, 2004.

Sakib, M., Mustajab, S., and Siddiqui, T. Deep learning-based heartbeat classification of 12-lead ecg time series signal. In *2023 4th International Conference on Data Analytics for Business and Industry (ICDABI)*, pp. 273–278. IEEE, 2023.

Silva, T. S. Exploring simclr: A simple framework for contrastive learning of visual representations. *https://sthalles.github.io*, 2020. URL https://sthalles.github.io/simple-self-supervised-learning/.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. Attention is all you need, 2023. URL https://arxiv.org/abs/1706.03762.

Wang, Z., Yan, W., and Oates, T. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pp. 1578–1585. IEEE, 2017.

Xiao, X., Zhang, D., Hu, G., Jiang, Y., and Xia, S. Cnn–mhsa: A convolutional neural network and multi-head self-attention combined approach for detecting phishing websites. *Neural Networks*, 125:303–312, 2020.

Xu, J., Wu, H., Wang, J., and Long, M. Anomaly transformer: Time series anomaly detection with association discrepancy. *ArXiv*, abs/2110.02642, 2021. URL https://api.semanticscholar.org/CorpusID:238408395.

Yang, H. and Desell, T. A large-scale annotated multivariate time series aviation maintenance dataset from the ngafid. *arXiv preprint arXiv:2210.07317*, 2022.

Yang, H., LaBella, A., and Desell, T. Predictive maintenance for general aviation using convolutional transformers. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 12636–12642, 2022.

Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., and Eickhoff, C. A transformer-based framework for multivariate time series representation learning, 2020. URL https://arxiv.org/abs/2010.02803.

Zhang, W., Huang, J., Wang, R., Wei, C., Huang, W., and Qiao, Y. Integration of mamba and transformer - mat for long-short range time series forecasting with application to weather dynamics. *2024 International Conference on Electrical, Communication and Computer Engineering (ICECCE)*, pp. 1–6, 2024. URL https://api.semanticscholar.org/CorpusID:272653578.

Zhao, H., Wang, Y., Duan, J., Huang, C., Cao, D., Tong, Y., Xu, B., Bai, J., Tong, J., and Zhang, Q. Multivariate time-series anomaly detection via graph attention network. *2020 IEEE International Conference on Data Mining (ICDM)*, pp. 841–850, 2020. URL https://api.semanticscholar.org/CorpusID:221507538.

Zhou, J., Menenti, M., Jia, L., Gao, B., Zhao, F., Cui, Y., Xiong, X., Liu, X., and Li, D. A scalable software package for time series reconstruction of remote sensing datasets on the google earth engine platform. *International Journal of Digital Earth*, 16:988 – 1007, 2023. URL https://api.semanticscholar.org/CorpusID:257712743.