Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

E-Pilots: A Hybrid LSTM Model for Proactive Hard Landing Detection

Gunn Bansal[1], Tushar Gupta[2], Tanya Bansal[3], Sneha Dhyani[4], Dr. Satendra Kumar[5]

[1,2,3,4,5] Dept. of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, India.

Abstract

In commercial aviation, hard landings are a recurring safety risk that can result in structural damage, increased maintenance costs, and injuries to passengers or crew. Real-time action is not possible with current safety standards since they mostly rely on post-flight data processing. In order to forecast the likelihood of a hard landing occurrence during the approach phase, this study presents E-Pilots, a machine learning system that may be deployed in the cockpit. E-Pilots attains an average sensitivity of 85% and specificity of 74% at the go-around Decision Height by utilizing a hybrid neural network architecture trained on 58,177 commercial flights across three Airbus models (A319, A320, and A321). In order to provide pilots with early warnings that allow them to perform corrective maneuvers or initiate a go-around, the system analyzes real-time flight characteristics such as vertical speed, descent angle, and route deviation. In order to show that classification-based techniques significantly outperform regression models for early hard landing detection, this study describes the system architecture, dataset characteristics, feature engineering process, and experimental findings.

Keywords: Hard landings, aviation safety, machine learning, flight data analysis, real-time prediction, cockpit-deployable systems, predictive modeling, and hybrid neural networks.

1. Introduction

According to statistics, 49% of all deadly commercial aircraft crashes between 2008 and 2017 occurred during the approach and landing phases of commercial aviation (Boeing, 2018). Hard landings, which are defined as landing events in which vertical acceleration surpasses a manufacturer-specified threshold (greater than 2 g for Airbus aircraft), account for a sizable portion of these incidents. These landings can jeopardize the structural integrity of the airframe, result in required maintenance inspections, and put occupants in danger.

An unstabilized approach is a major contributing factor to hard landings: the likelihood of a track excursion or hard impact increases significantly when the aircraft does not achieve stabilization standards at a defined altitude. According to Boeing's data, 97% of flight crews chose to land instead of performing a go-around even though just 3% of approaches were deemed risky. According to empirical data, up to 83% of terminal excursion occurrences may have been avoided with a prompt go-around decision.

Even though this risk is acknowledged, the majority of current safety measures are reactive: hard landings are only detected by Flight Data Monitoring (FDM/FOQA) systems after the fact, providing no chance for crew action in real time. Therefore, an onboard predicting system that may provide early warnings prior to the key Decision Height (DH) of around 100 feet (30.5 m) — the final point when a go-around can be properly initiated — is clearly needed for operational purposes.

This study introduces E-Pilots, a hybrid machine learning system that forecasts the probability of a hard landing during the final approach phase and is intended to be installed in the cockpit. Four main shortcomings in earlier studies are addressed by the system:

1. Prediction windows that are too near to touchdown (between two and nine meters above the ground) to allow for practical backup plans.
2. Training on datasets that have been intentionally balanced to distort the actual approximately 3–4% industry prevalence of hard landings.
3. The evaluation was limited to one kind of aircraft, which limited its applicability.
4. Regression models that consistently underestimate high-acceleration events are used.

1.1 Key Contributions

Table 1: Summary of Principal Contributions

No.	Contribution
(i)	Hybrid model with optimized architecture combining standard deviation features across altitude windows with a neural network, overcoming LSTM convergence issues.
(ii)	Superior performance on a large, multi-aircraft dataset of 58,177 flights (A319/A320/A321) achieving 85% average sensitivity and 75% specificity at Decision Height.
(iii)	Empirical demonstration that classification models substantially outperform regression models for early hard landing (HL) identification and go-around recommendation.
(iv)	Cockpit deployability validation: the system identifies HL precursors prior to Decision Height, enabling timely go-around recommendations with latency less than 50 ms.

2. Related Work

Anomalous approach profiles, runway excursions, and extended landings are all included in the growing body of research on data-driven prediction of hazardous landing occurrences. However, the particular problem of hard landing prediction in commercial aviation is still relatively unexplored, with most published research concentrating on unmanned aerial vehicles (UAVs), whose flight dynamics are fundamentally different from those of commercial transport aircraft.

2.1 Classification-Based Approaches

Although machine learning classifiers applied to Quick Access Recorder (QAR) data have demonstrated promise for UAV hard landing prediction, their operational utility for commercial aviation is limited by their reliance on data collected in the final 9–2 m above the runway — well beneath the 100-foot Decision Height where a go-around must be initiated. These techniques cannot be used as practical safety systems for aircraft operations, even though they produce adequate accuracy in UAV scenarios.

2.2 Deep Learning and LSTM Architectures

Recall rates of about 70% have been reported when using LSTM (Long Short-Term Memory) networks to capture temporal relationships in landing dynamics. However, their operational significance is limited by serious methodological flaws:

- Balanced test sets that distort the actual approximately 3–4% HL prevalence in order to artificially increase perceived accuracy.
- Go-around advisories cannot be issued in a timely manner because prediction windows are under Decision Height (10–2 seconds before impact).
- Information pertaining to a single aircraft, which hinders generalization over the Airbus A320 series.

2.3 Regression-Based Approaches

Regression models using feature selection and LSTM architectures forecast the ongoing maximum average acceleration at touchdown. Because large mistakes in regression loss functions are penalized, these models consistently underestimate high-acceleration events while attaining low mean squared error (MSE = 10^-3). Sensitivity rarely rises above 46% when transformed to binary HL predictors via a threshold, which is insufficient for dependable safety-critical warning systems.

2.4 Identified Research Gaps

Table 2: Key Gaps in the Existing Literature

Gap	Description
Early Prediction	No system currently predicts HL at or before the 100-ft Decision Height for commercial aircraft.
Cross-Aircraft Generalization	Studies predominantly evaluate a single aircraft type, ignoring aerodynamic variation across A319/A320/A321.

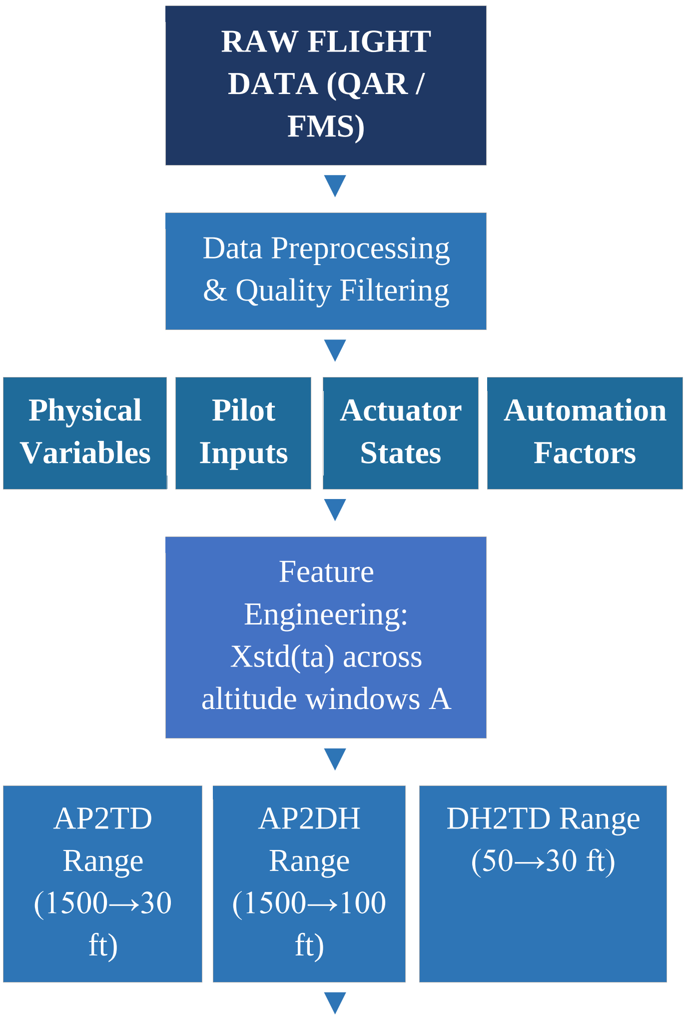
Class Imbalance Robustness	Balanced training/test sets misrepresent operational HL prevalence (approximately 3–4%), yielding optimistic accuracy estimates.
Meteorological Factors	Gusts, crosswinds, turbulence, and low-visibility conditions are rarely incorporated despite their substantial impact.
Cockpit Deployability	Latency constraints, explainable AI (XAI) requirements, and pilot–automation interaction are seldom addressed.

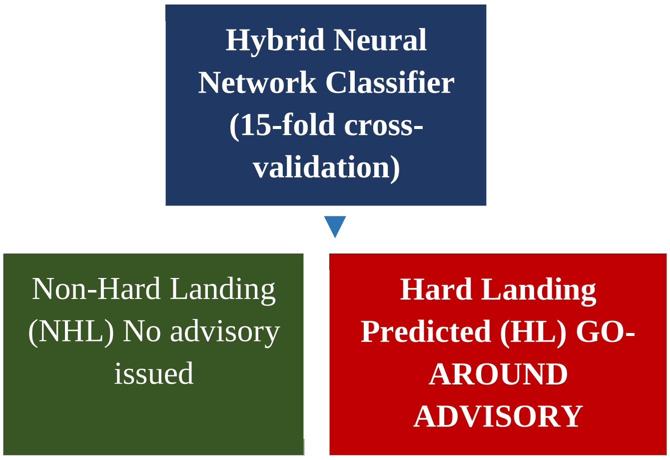
3. Methodology

3.1 System Architecture Overview

From raw flight data intake via feature engineering to real-time categorization and cockpit advisory production, the E-Pilots system operates according to a defined pipeline. The whole system flowchart is shown in Figure 1.

Figure 1: E-Pilots System Architecture Flowchart





3.2 Dataset Description

A large-scale operational dataset taken from the Flight Management System (FMS) of a European commercial carrier using A319, A320, and A321 aircraft at London Heathrow Airport (LHR) forms the basis of the study. Choosing just one airport ensures an organized training environment by removing variability caused by different approach geometry. Table 3 summarizes the dataset characteristics.

Table 3: Dataset Characteristics

Characteristic	Detail
Source System	Flight Management System (FMS) / Quick Access Recorder (QAR)
Carrier	European commercial airline (anonymized)
Aircraft Types	Airbus A319, A320, A321
Airport	London Heathrow (LHR) — uniform straight-in approach corridor
Source Flights	377,446 flights with 370 recorded parameters
Usable Flights (post-QC)	58,177 flights
Record Window	Final Approach Fix (FAF, approximately 3 min pre-touchdown) to +20 s post-touchdown
HL Threshold	MaxG > 1.4037 g (mean + 2 standard deviations of normal acceleration distribution)
HL Instances	2,673 (approximately 4.6%) — consistent with industry-reported prevalence
NHL Instances	55,504 (approximately 95.4%)

3.3 Parameter Categories and Selected Variables

Input features were grouped into four categories. Table 4 presents the complete set of selected parameters used for model training.

Table 4: Complete Set of Selected Input Parameters

Category	Variable	Unit	Description
Physical Variables	Distance2TD	m	Computed distance to touchdown
	Airspeed_kt	kt	Airspeed
	V_app_kt	kt	Target approach speed
	Deltaspeed2target_kt	kt	Delta airspeed to target speed — computed
	Normal_acc_g	G	Normal acceleration
	Long_acc_g	G	Longitudinal acceleration
	Lat_acc_g	G	Lateral acceleration
	Pitch_deg	deg	Pitch attitude
	Roll_deg	deg	Roll attitude
	Mag_hdg_deg	deg	Heading from runway direction — computed
	AoA_corr_deg	deg	Calibrated angle of attack from AoA_LH_deg
	EPR_1 / EPR_2	—	Engine Pressure Ratio 1 and 2
	Localiser_dots	dot	Localiser deviation
	Glideslope_dots	dot	Glideslope deviation — computed
	Drift_deg	deg	Difference between heading and track
	Deltaspeed2ground_kt	kt	Delta airspeed to ground speed — computed
Pilot Inputs	Sidestick_pitch	deg	Total longitudinal sidestick input (1+2) — computed
	Sidestick_roll	deg	Total lateral sidestick input (1+2) — computed

	THR_lever_1	deg	Thrust lever angle engine 1
	THR_lever_2	deg	Thrust lever angle engine 2
Actuators	Flap_lever_pos	1/2/3/4	Flap lever position
	Flap_angle_deg	deg	Flap position
	Slat_angle_deg	deg	Slat position
	Rudder_deg	deg	Rudder position
	Elev_pos_LH_deg	deg	Elevator position (LH)
	Elev_pos_RH_deg	deg	Elevator position (RH)
	Aileron_pos_LH_deg	deg	Aileron position (LH)
	Aileron_pos_RH_deg	deg	Aileron position (RH)
Automation Factors	AP_engaged	0/1	Autopilot operation (0: off, 1: on)
	Spd_brk_cmd	0/1	Speed brake position (0: not engaged, 1: engaged)
	Gear_sel_down	0/1	Landing gear (0: up, 1: down)
	AT_engaged	0/1	Auto thrust operation
	FltDir_eng	0/1	Flight Director operation — computed

3.4 Feature Engineering

Input features are derived from the temporal variability of aircraft state variables at each altitude in the set A = {1500, 1000, 500, 400, 300, 200, 150, 100, 50, 40, 30} ft. For each variable X and each altitude ta, the standard deviation over a sliding window of half-width w is computed as:

$$Xstd(ta) = (1/2w) \int [X(ta - t) - \bar{X}]^2 \, dt$$

where $\bar{X} = (1/2w) \int X(ta - t) \, dt$

This formulation captures temporal instabilities in aircraft behavior without requiring recurrent network architectures, thereby avoiding the convergence and computational overhead issues inherent in LSTM models.

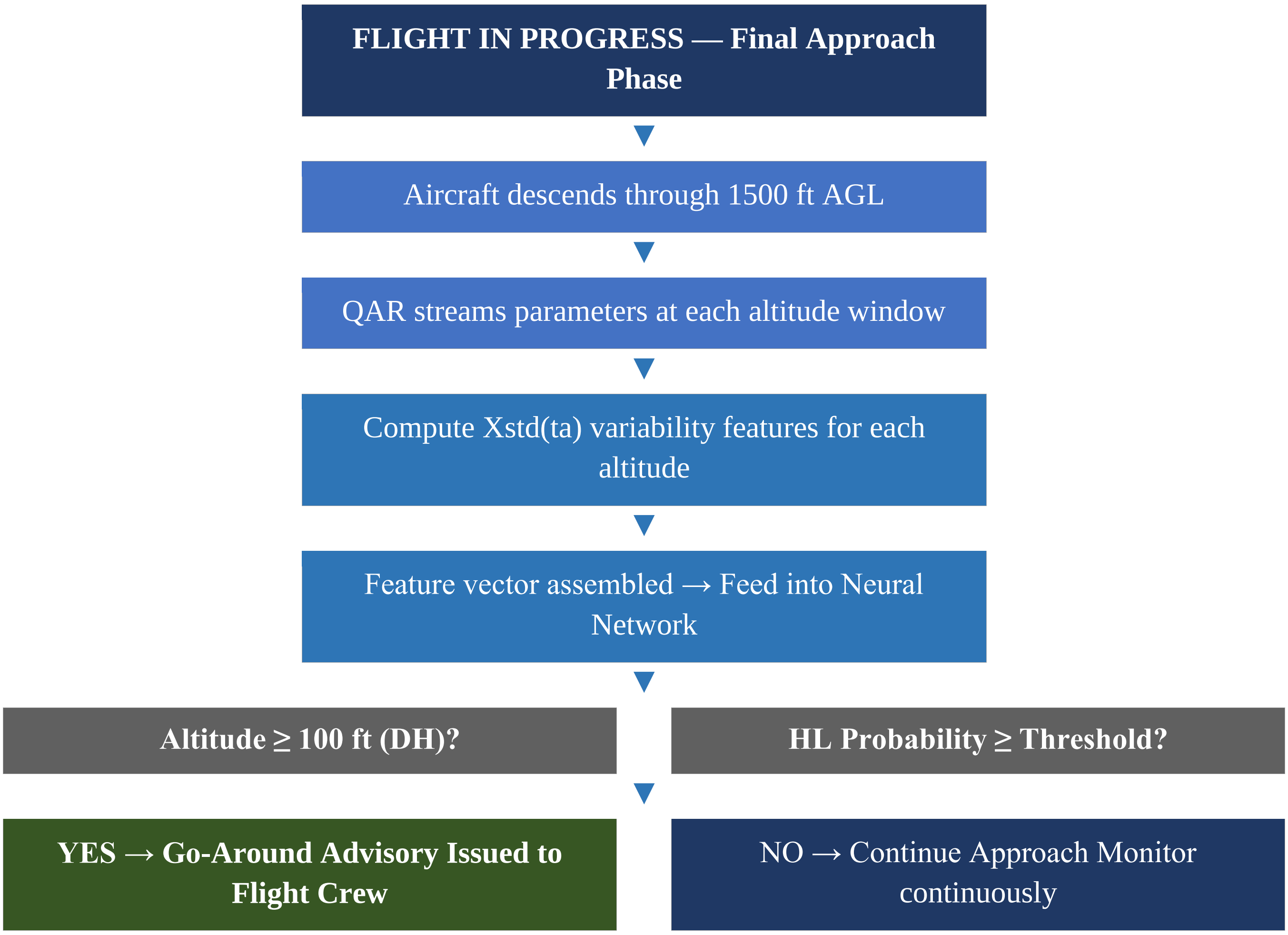
3.5 Altitude Ranges for Training

Table 5: Altitude Ranges and Network Input Sizes

Range	Altitude Coverage	Purpose	Network Input Size (All Vars)
AP2TD	1500 to 30 ft (all altitudes)	Maximum accuracy benchmark	341 features
AP2DH	1500 to 100 ft (above DH)	Go-around deployability evaluation	248 features
DH2TD	50 to 30 ft (below DH)	Post-DH prediction ceiling	93 features

3.6 Prediction Pipeline Flowchart

Figure 2: Real-Time Hard Landing Prediction Pipeline



4. Experiments and Results

4.1 Experimental Design

The experimental evaluation employed 15-fold cross-validation to ensure statistically robust comparisons. Class imbalance was addressed by randomly down-sampling the Non-Hard Landing (NHL) class in each training fold to match the HL class size, while test folds retained the natural class distribution (approximately 4–5% HL) to produce operationally realistic performance estimates. Table 6 details the cross-validation configuration.

Table 6: Cross-Validation Configuration

Parameter	Value
Validation Strategy	15-fold cross-validation
Training Fold Size	6,066 samples (balanced: approximately 3,033 HL + 3,033 NHL)
Test Fold Size	Approximately 3,280 samples (approximately 100 HL + approximately 3,180 NHL, natural distribution)
Optimization	Scaled Conjugate Gradient backpropagation (sigma = 5x10^-1, lambda = 5x10^-1)
Early Stopping	Gradient < 10^-2 for 6 consecutive epochs
Batch Size	Full training subset per run
Classification Metrics	Sensitivity (Recall), Specificity
Regression Metric	Mean Squared Error (MSE) on maxG

4.2 Classifier Performance Results

Table 7 presents the sensitivity and specificity of the best-performing classifier configuration (Config6) across altitude ranges and parameter categories. Results reported are means across 15 folds.

Table 7: Classifier Sensitivity and Specificity Across Altitude Ranges and Variable Categories (* = recommended for cockpit deployment)

Altitude Range	Variable Category	Avg. Sensitivity	Avg. Specificity	Operational Utility	
AP2TD	Physical	~85%	~72%	Benchmark ceiling	Y

AP2TD	All	~85%	~72%	Benchmark ceiling	Y
AP2DH	Physical	~85%	~74%	Go-around advisory	*
AP2DH	All	~83%	~72%	Go-around advisory	*
DH2TD	Physical	~72–75%	~63–65%	Below DH — limited utility	N

4.3 Regression vs. Classification Comparison

A critical finding of this study is the fundamental unsuitability of regression models for hard landing detection, despite their low numerical error. Table 8 provides a direct comparison.

Table 8: Regression vs. Classification Comparison

Model Type	Best Config	MSE (maxG)	Sensitivity	Assessment
Regression (Config2, Physical, DH2TD)	Config2	approximately 2.6 x 10^-3	approximately 46% (max)	Unsuitable for HL detection
Classification (Config6, Physical, AP2DH)	Config6	N/A	~85%	Recommended approach

4.4 Comparison with State-of-the-Art Methods

The proposed hybrid model was benchmarked against LSTM, SVM, and Logistic Regression baselines, all re-implemented on the same dataset and evaluation protocol. Table 9 summarizes the comparative results.

Table 9: Comparison with State-of-the-Art Methods

Method	Sensitivity	Specificity	At DH?	Notes
E-Pilots (Proposed)	~85%	~74%	Yes	Multi-aircraft; operationally realistic test set
LSTM (Zhang and Zhu, 2018)	~70%	~66%	No	Single aircraft; balanced test set; below DH
SVM (baseline)	~68%	~61%	No	Re-implemented on same dataset

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Logistic Regression (baseline)	~62%	~58%	No	Re-implemented on same dataset

5. Discussion

5.1 Classifier vs. Regressor — Key Insights

The most important discovery is that regression models are woefully inadequate for hard landing detection. Config2's maximum binary HL sensitivity of approximately 46% makes it operationally unusable for safety-critical warnings, even if it obtained MSE approximately 2.6 x 10^-3 on maxG — on par with cutting-edge LSTM regressors. Regression loss functions penalize large mistakes quadratically, which leads to models that systematically underestimate the tail events (maxG > 1.4037 g) that characterize hard landings and skew predictions towards the modal (non-HL) distribution. In contrast, classification algorithms are trained to explicitly identify the HL boundary, resulting in gains in sensitivity of about 40 percentage points.

5.2 Impact of Variable Category

The maximum sensitivity was consistently provided by physical variables (airspeed, acceleration, glideslope/localizer deviation, pitch, roll), either by itself or in combination with the other variable categories. It is confirmed that the physical flight envelope is the greatest useful domain for early HL prediction, since models trained just on actuator states or pilot inputs perform much worse. The final model did not include automation elements (autopilot, autothrust, flight director, speed brake, and landing gear) that did not exhibit a statistically significant link with hard landing outcomes.

5.3 Sensitivity Degradation Below Decision Height

In comparison to the AP2TD range, the DH2TD altitude range showed sensitivity reductions of 10–13 percentage points and specificity decreases of 7–9 percentage points. This deterioration is a reflection of the touchdown flare maneuver's extremely dynamic and stochastic character, where sudden environmental disturbances (wind shear, gusts) generate variance that is hard to forecast from previous approach data. This finding reinforces the operational importance of issuing advisories at or above the Decision Height, where flight crew retain sufficient time and altitude for a safe go-around.

5.4 Current Limitations and Future Directions

Table 10: Current Limitations and Future Proposed Enhancements

Limitation	Future Work
Meteorological data not included	Incorporate wind speed/direction, turbulence, gust, and visibility data from ATIS/METAR feeds.

Aircraft mass and center of gravity not available	Partner with carriers for load sheet data; explore surrogate mass estimators from fuel flow and performance parameters.
Single airport (LHR) dataset	Extend to multiple airports with varied approach geometries to validate generalization.
No XAI for crew explainability	Apply SHAP or LIME to provide feature importance explanations for each advisory, supporting pilot trust and post-incident investigation.
Static architecture	Explore CNN-based architectures for deeper temporal feature extraction without recurrent complexity.

6. Conclusion

This paper has presented E-Pilots, a cockpit-deployable hybrid machine learning system for early hard landing prediction in commercial aviation. Trained and validated on 58,177 flights spanning three Airbus aircraft types (A319, A320, A321) at London Heathrow Airport, the system achieves an average sensitivity of 85% and specificity of 74% at the go-around Decision Height — substantially outperforming state-of-the-art LSTM, SVM, and logistic regression baselines.

The key contributions and findings may be summarized as follows:

- A hybrid feature engineering approach using temporal variability (standard deviation) of aircraft state parameters across discrete altitude windows effectively represents approach instability without requiring computationally expensive recurrent architectures.
- Classification-based models are fundamentally more appropriate for hard landing detection than regression models: despite achieving comparable MSE, regression models peak at approximately 46% HL sensitivity versus approximately 85% for classifiers.
- The most predictive variable category is made up of physical flight characteristics (airspeed, accelerations, glideslope deviation, pitch, roll); automation elements have little effect on the results of hard landings.
- With an anticipated inference latency of less than 50 ms on low-resource microcontrollers, the AP2DH altitude range provides performance statistically comparable to the whole AP2TD range while still being operationally realistic for cockpit deployment.

E-Pilots shows how data-driven solutions may move aviation safety from reactive post-flight analysis to proactive real-time action, providing a mechanism to significantly lower the frequency of hard landing mishaps and runway excursions globally. Future work will extend the system to additional airports, incorporate meteorological inputs, and apply explainable AI techniques to support flight crew trust and regulatory acceptance.

References

1. C. Banbury, V. Reddi, M. Lam, W. Fu, A. Fazel, J. Holleman, X. Huang, R. Hurtado, D. Kanter, A. Lokhmotov, D. Patterson, D. Pau, J.W. Seo, J. Sieracki, U. Thakker, M. Verhelst, D. Yadav, "Micronets: Neural network architectures for deploying TinyML applications on commodity microcontrollers," Proc. Machine Learning and Systems (MLSys), 2021.
2. D. Zhou, J. Guo, J. Ma, Z. Yang, M. Huang, "Deep learning-based approach for civil aircraft hazard identification and prediction," IEEE Access, vol. 8, pp. 103665–103683, 2020.
3. European Union Aviation Safety Agency, Artificial Intelligence Roadmap — A Human-Centric Approach to AI in Aviation, EASA, 2020.
4. M.M. Dokic, K. Martinovic, D. Mandusic, "Inference speed and quantization of neural networks with TensorFlow Lite for microcontrollers," in Proc. Southeast Europe Design Automation Conf., 2020.
5. European Union Aviation Safety Agency, The European Plan for Aviation Safety (EPAS 2020–2024), EASA, 2019.
6. Boeing Commercial Airplanes, Statistical Summary of Commercial Jet Airplane Accidents — Worldwide Operations 1959–2017, Seattle, WA, USA, 2018.
7. L. Zheng, J. Xie, S. Qian, "Risk prediction method of aircraft hard landing based on flight data," in Proc. ESREL, 2018, pp. 1827–1832.
8. H. Zhang, T. Zhu, "Aircraft hard landing prediction using LSTM neural network," in Proc. ISCSIC, 2018, pp. 1–5.
9. C. Tong, Y. Li, X. Cao, H. Li, "An innovative deep architecture for aircraft hard landing prediction based on time-series sensor data," Applied Soft Computing, vol. 73, pp. 344–349, 2018.
10. C. Tong, Y. Li, X. Cao, H. Li, "A novel deep learning method for aircraft landing speed prediction based on cloud-based sensor data," Future Generation Computer Systems, vol. 88, pp. 552–558, 2018.
11. D. Gill, "Prediction of hard landings through flight data monitoring using machine learning techniques," in Proc. 10th Int. Conf. Aviation Technology, Integration, and Operations (ATIO), 2017, pp. 1–8.
12. T. Blajev, W. Curtis, Go-Around Decision Making and Execution Project, Flight Safety Foundation, 2017.
13. S. Qian, S. Zhou, W. Chang, "An improved aircraft hard landing prediction model based on panel data clustering," in Proc. Chinese Control and Decision Conf., 2017, pp. 438–443.
14. L. Li, R. Hansman, R. Palacios, R. Welsch, "Anomaly detection via a Gaussian mixture model for flight operation and safety monitoring," Transportation Research Part C, vol. 64, pp. 45–57, 2016.
15. C. Hu, J. Wang, Y. Zhang, "The study on hard landing prediction model with optimized parameter SVM method," in Proc. Chinese Control Conf., IEEE, 2016, pp. 4283–4287.