

SafeAeroBERT: Towards a Safety-Informed Aerospace-Specific Language Model

Sequoia Andrade^{1, 2}

Hannah Walsh²

HX5 LLC.¹; NASA Ames Research Center ²

AIAA Aviation, June 12th-16th 2023

This material is a work of the U.S. Government and is not subject to copyright protection in the United States
Published by the American Institute of Aeronautics and Astronautics, Inc., with permission.

Motivation

- Aviation generates many documents, such as compliance documents, mishap safety reports, and requirements
- The jargon in aviation documents limits the effectiveness of state-of-the-art machine learning model
- BERT models [1]:
 - Have been successfully trained for various domain-specific use cases (e.g., biomedical [2], legal [2], finance [3])
 - Show superior performance on nearly all natural language processing (NLP) tasks

DO NOT REPORT AIRCRAFT ACCIDENTS AND CRIMINAL ACTIVITIES ON THIS FORM.

ACCIDENTS AND CRIMINAL ACTIVITIES ARE RECORDED IN THE AARS AND SHOULD NOT BE SUBMITTED TO NASA.

ALL ITEMS CONTAINED IN THIS REPORT MUST BE RETURNED TO ISSUING COMPLETE OFFICE AFTER ANALYSIS.

(OFFICIAL USE ONLY - RESERVED FOR NASA QUALITY STRATEGIC)

DEFINITION STATE: Please fill all blank fields to ensure return of ID strip to you.
NO RECORD WILL BE KEPT OF YOUR DEFINITION. This section will be returned to you.

TELEPHONE NUMBERS who may reach you for further details at this occurrence:

HOME Area No. _____ Hours _____	WORK Area No. _____ Hours _____
---------------------------------	---------------------------------

NAME _____
ADDRESS/PO BOX _____
CITY _____ STATE _____ ZIP _____

DATE OF OCCURRENCE _____
LOCAL TIME (At clock) _____

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA has established an Aviation Safety Reporting System (ASRS) to identify errors in the aviation system which need to be addressed. The program of which this form is a part is described in Detail in FAA Advisory Circular 00-46F. Your assistance in referring us about such issues is essential to the success of the program. Please fill out this form as completely as possible, enclose in an sealed envelope, affix proper postage, and send it directly to us.

The information you provide on this identity strip will be used by NASA to determine if it is necessary to contact you for further information. This identity strip WILL BE RETURNED DIRECTLY TO YOU. The return of the identity strip assures your anonymity.

AVIATION SAFETY REPORTING SYSTEM

Section 91.25 of the Federal Aviation Regulations (14 CFR § 91.25) prohibits reports filed with NASA from being used for FAA enforcement purposes. This report will not be made available to the FAA for civil penalty or certification actions for violations of Federal Aviation Regulations. Your identity strip, stamped by NASA, is proof that you have submitted a report to the Aviation Safety Reporting System. You can only return the strips to you if they are provided a mailing address. Equally important, we can often obtain additional useful information if our safety analysts can talk with you directly by telephone. For this reason, we have provided telephone numbers where we may reach you.

Thank you for your contribution to aviation safety.

Note: Aircraft Accidents and Incidents Should Not Be Reported On This Form. Such events should be filed with the National Transportation Safety Board as required by NTSB Regulation 830.6 (49 CFR 830.6).

Paperwork Reduction Act Statement: This information collection meets the requirements of 44 U.S.C. 3507, as amended by section 2 of the Paperwork Reduction Act of 1995. You do not need to answer those questions outside what you desire a valid Office of Management and Budget control number. The OMB control number for this information collection is 2700-0176. We estimate that it will take about 30 minutes to read the instructions, gather the facts, and answer the questions. You may send comments on our time estimates above to: PO Box 199 Moffett Field, CA 94035-0199.

If you want to mail this form please place pages, enclose in a sealed, stamped envelope, and mail to:



NASA AVIATION SAFETY REPORTING SYSTEM
POST OFFICE BOX 199
MOFFETT FIELD, CA 94035-0199

DESCRIPTION EVENT/SITUATION

Keeping in mind the topics shown below, discuss those which you find most relevant and anything else you think is important. Indicate what you believe caused the problem, and what can be done to prevent a recurrence, or correct the situation. (USE ADDITIONAL PAPER IF NEEDED.)

REPORTER	FLYING TIME (in hrs)	CERTIFICATES & RATINGS	AIR EXPERIENCE
☐ Captain	Total Time _____ in hrs	☐ Student	☐ Pilot
☐ First Officer	☐ Instructor	☐ Sport/flight	☐ Multi-engine
☐ Jetting flying	Last 90 Days Total _____ in hrs	☐ Private	☐ Instrument
☐ Last not flying	☐ Dispatch _____ in hrs	☐ Commercial	☐ Non-superior
☐ Relief pilot	☐ Other _____ in hrs	☐ ATP	☐ Military
☐ Check airman			

BASETYPE	CONDITIONS / WEATHER ELEMENTS	LIGHT / VISIBILITY	RAMP ADVISORY SCUE
☐ Class A	☐ VMC	☐ Daylight	☐ ATC
☐ Class B	☐ IMC	☐ Night	☐ Rotor
☐ Class C	☐ MC	☐ Cloud	☐ Ground
☐ Class D	☐ Chemo/mile	☐ Ceiling	☐ UNCOM
☐ Class E	☐ Mixed	☐ Visibility _____ in mi	☐ TRACON
☐ Class F	☐ Xing	☐ Other	☐ CTAF
☐ Class G	☐ Marginal	☐ Other	☐

Other Aircraft Type	Operating FAR Part	Other Aircraft Type	Operating FAR Part
☐ Air carrier	☐ military	☐ Air carrier	☐ military
☐ Taxi	☐ general	☐ Taxi	☐ general
☐ Corporate	☐ other	☐ Corporate	☐ other

Mission

☐ passenger	☐ cargo/freight	☐ ferry
☐ commercial	☐ training	☐ other

Flight Plan

☐ VFR	☐ SVFR	☐ none
☐ IFR	☐ DVFR	☐ none

Flight Phase

☐ taxi	☐ climb	☐ final approach
☐ missed/GAR	☐ level	☐ missed/GAR
☐ descent	☐ landing	☐ landing
☐ initial climb	☐ initial descent	☐ initial approach

Route In Use

☐ direct (D)	☐ STAR (D)	☐ visual approach
☐ indirect	☐ oceanic	☐ none
☐ vector	☐ vectors	☐ none

if more than two aircraft were involved, please describe the additional aircraft in the "Describe Event/Situation" section.

LOCATION	DIRECTION	Estimate min distance in feet	horz	vert
Altitude _____ (single value) MSL _____ AGL _____	from _____	_____ ft	O/T	O/H
Distance _____ and/or Radar (bearing) _____	from _____	_____ ft	O/T	O/H

What was TCAS-A or TA doing? Yes ☐ No ☐
Contributing factors _____

CHART OF EVENTS Yes ☐ No ☐
How the problem arose _____ How it was discovered _____
Corrective actions _____

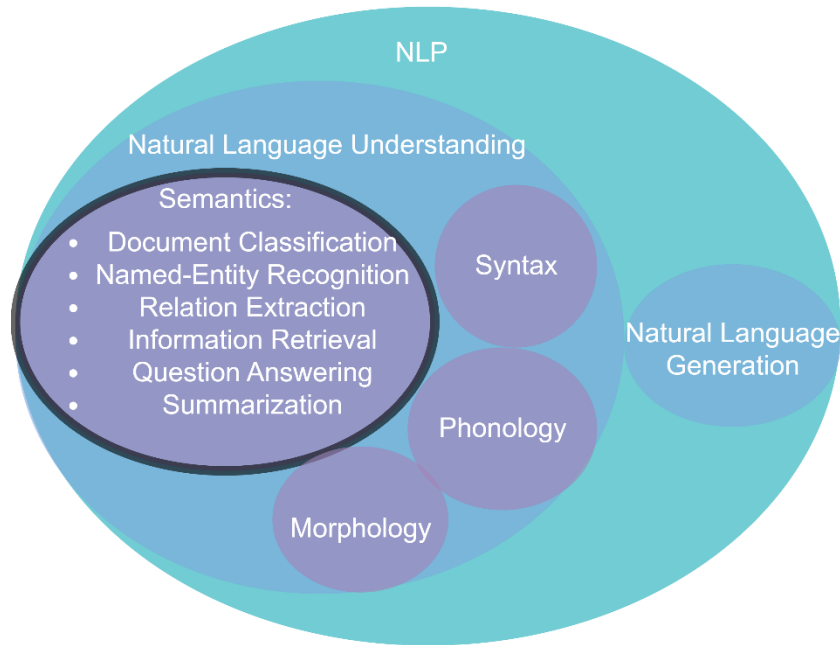
HUMAN PERFORMANCE CONSIDERATIONS
+ Perceptions, judgments, decisions + Actions or reactions
+ Factors affecting the quality of human performance

ASRS Form 7778 (May 2009)
GENERAL

OMB No. 2700-0176

Figure 1: Example ASRS [5] worksheet

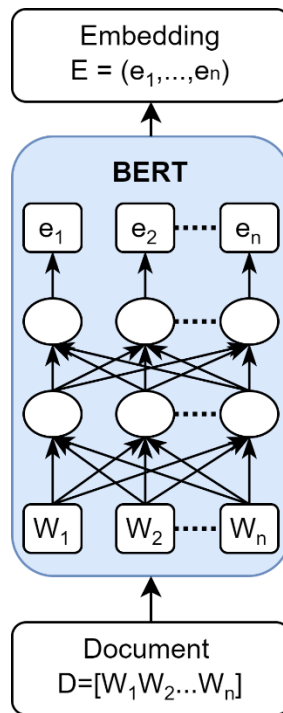
Natural Language Processing



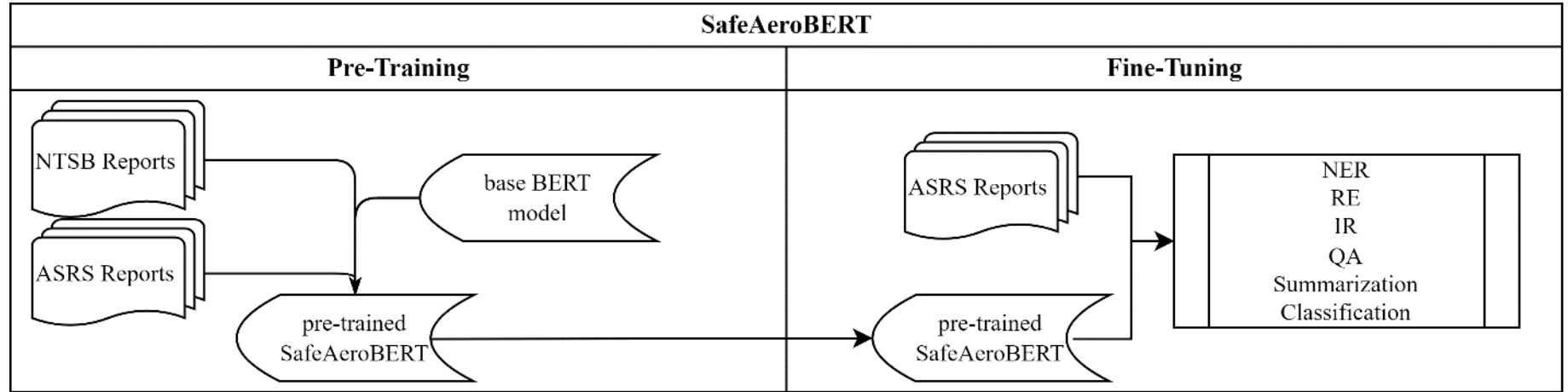
- In Aviation
 - Classification for cause identification and severity categorization
 - NER for NOTAM digitalization and FMEA extraction
 - QA and Summarization
- The proposed SafeAeroBERT model could improve aviation NLP results by learning aviation jargon

BERT Model

- Bi-directional masked language model
 - Provides contextual meaning of words within a sentence
- Next sentence prediction
 - Provides contextual meaning of sentences in a document
- Pre-training: upfront training on millions of documents to learn word context
- Fine-Tuning: additional training on fewer documents for a specific task, such as classification

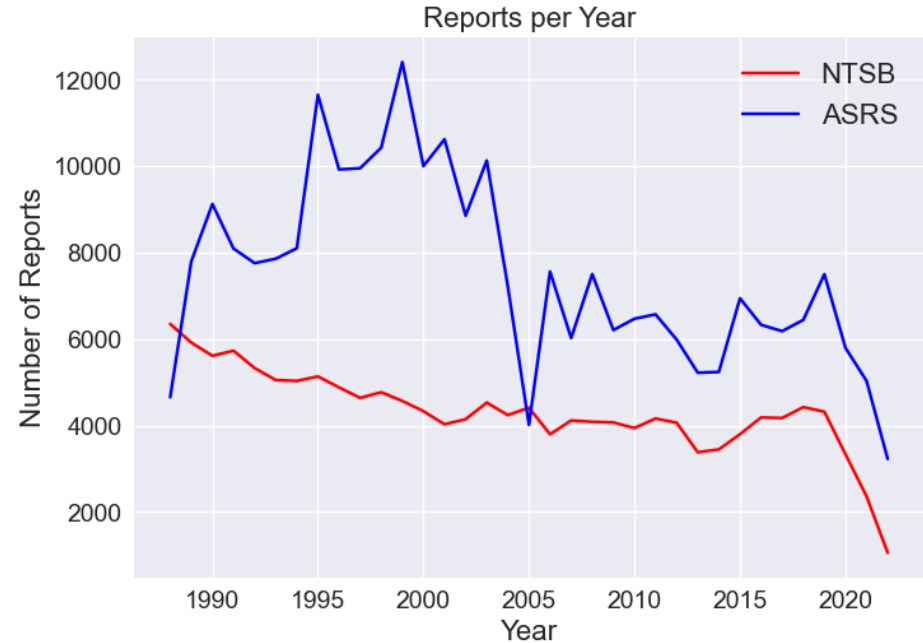


Method



Pre-Training: Data

- ASRS [5]
 - 262,935 documents
 - 5 possible narrative text sections
- NTSB
 - 161,460 documents
 - 6 possible narrative text sections
- Total of 2,283,435 narrative text sections



Pre-Training

- Training:
 - Trained for > 2 epochs
 - Batch size = 8
 - Gradient accumulation = 16
 - Total batch = 64
 - 13,000 steps
 - ~ 2 months run time on 3 GPUs

Fine-Tuning

- Classification
 - Compared SciBERT, base BERT, and SafeAeroBERT
 - 5 epochs
 - Batch size of 32
 - 3 GPUs
 - ~1.2 hour to train each model for each classification task

Contributing Factor		Train	Validation	Test
Aircraft	Present	2083	260	260
	Not Present	5917	740	740
	Total	8000	1000	1000
Human Factors	Present	3610	451	451
	Not Present	4390	549	549
	Total	8000	1000	1000
Procedure	Present	1172	146	146
	Not Present	6828	854	854
	Total	8000	1000	1000
Weather	Present	1035	129	129
	Not Present	6965	871	871
	Total	8000	1000	1000

Classification Results

Contributing Factor	Metric	BERT	SciBERT	SafeAeroBERT
Aircraft	Accuracy	0.747	0.726	0.740
	Precision	0.716	0.691	0.548
	Recall	0.747	0.726	0.740
	F-1	0.719	0.699	0.629
Human Factors	Accuracy	0.608	0.557	0.549
	Precision	0.618	0.586	0.527
	Recall	0.608	0.557	0.549
	F-1	0.572	0.426	0.400
Procedure	Accuracy	0.766	0.755	0.845
	Precision	0.766	0.762	0.742
	Recall	0.766	0.755	0.845
	F-1	0.766	0.758	0.784
Weather	Accuracy	0.807	0.808	0.871
	Precision	0.803	0.769	0.759
	Recall	0.807	0.808	0.871
	F-1	0.805	0.788	0.811

- SafeAeroBERT outperforms BERT and SciBERT on procedure and weather document classification
- BERT outperforms SciBERT and SafeAeroBERT on aircraft and human factors
- SafeAeroBERT tends to have a higher accuracy and recall, indicating that it has better success at identifying all of the positive cases, potentially at the cost of a lower precision and higher false positive rate

Discussion

- SafeAeroBERT performs better than both SciBERT and BERT on some downstream tasks
- SafeAeroBERT can be used as a backbone in a knowledge management toolkit
 - Knowledge Discovery
 - Information Retrieval
- Entity Relation, Information retrieval, and Summarization still need to be tested
- Ideally, SafeAeroBERT would also include NER in pre-training
 - Future aviation specific BERT models must identify the set of relevant entities and build a training set
- Intended for safety applications

Conclusion

- In this research, we pre-trained a custom safety-informed aviation-specific large language model on NTSB and ASRS reports
- The resulting SafeAeroBERT model shows promise for improving NLP applications in aviation when used in place of standard BERT or SciBERT
- Will be made publicly available in the Future via NASA GitHub

Contact Information

➤ Emails:

- Sequoia.r.Andrade@nasa.gov
- Hannah.s.Walsh@nasa.gov

➤ Additional Resources:

- Robust Software Engineering:
<https://ti.arc.nasa.gov/tech/rse/>

References

- [1] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," arXiv, 2018.
- [2] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So and J. Kang, "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," Bioinformatics, 2019.
- [3] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras and I. Androutsopoulos, "LEGAL-BERT: The Muppets straight out of Law School," in Findings of the Association for Computational Linguistics: EMNLP 2020, 2020
- [4] D. T. Araci, "FinBERT: Financial Sentiment Analysis with Pre-trained," arXiv, 2019.
- [5] B. L. Hooey, "ASRS Program Briefing," NASA, 2021.



AMERICAN INSTITUTE OF
AERONAUTICS AND ASTRONAUTICS