

National Aeronautics and Space Administration

EXPLORE FLIGHT
WE'RE WITH YOU WHEN YOU FLY



Document Classification Techniques for Aviation Letters of Agreement

Anna Batra, Swetha Rajkumar, David Nielsen (KBR, Inc.), Stephen S.B. Clarke (FRA), and Krishna Kalyanam

NASA Ames Research Center, Moffett Field, CA, 94035, USA

Kanvasi Tejasen, Melissa Ohsfeldt (Silver Maple LLC), Mike Copp (LS Technologies)

Federal Aviation Administration, Washington D.C., USA

Overview

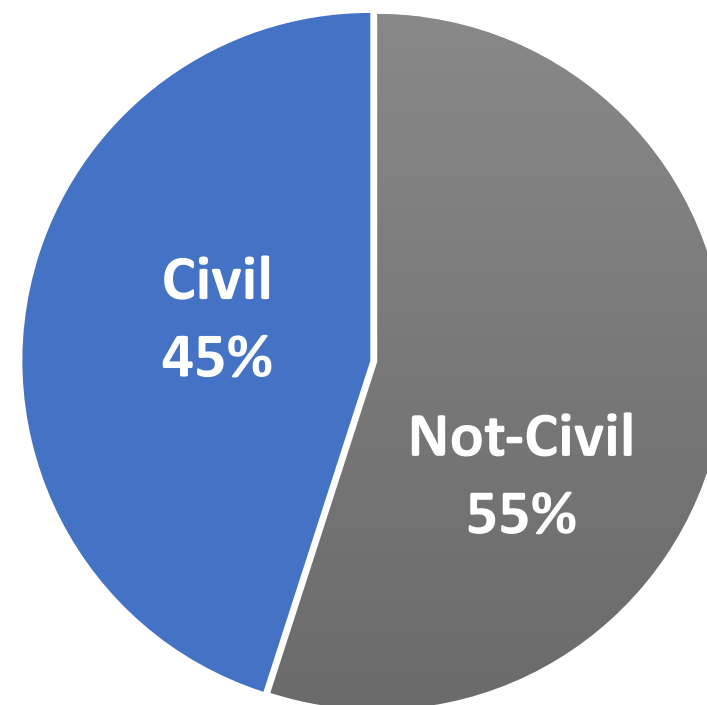


Letters of Agreement

- Letters of Agreement (LOA) formalizes airspace operations
 - Always two or more parties
 - Can be between FAA entities or FAA and private companies or the military
- Objective: represent the constraints contained in the LOAs in machine-interpretable digital format.
 - Currently stored as PDF documents available only to controllers
 - Digital LOA constraints could be made available to airspace users for tasks such as trajectory planning
- Natural language processing (NLP) is crucial to the digitization of LOAs.

- We want to digitize only **Civil** LOAs:
 - Civil LOAs have only FAA parties
 - No military or private company signatories
 - Focus of digitization
- Parties are *not* stored in LOA metadata
- This is a **two-class** document **classification** problem
 - Civil/not-civil class labels
 - Classification is the machine learning (ML) process of labeling documents by their class

493 labeled LOAs



Methodology



Preparing the LOA data

6

- Convert PDF to text
- Filter document to relevant sections
 - When possible, used section titles ***Document, Title, Facility, Effective (date), Subject, Purpose, and Scope***
 - Otherwise, used first page
- Clean and tokenize the text

(Name) Center/Approach Control and (Name) FSS

LETTER OF AGREEMENT

EFFECTIVE: _____

SUBJECT: Special VFR Operations within (Name) Airport Surface Area

1. **PURPOSE:** To provide operating procedures for Special VFR flight handling in the (name) surface area without individual coordination.

2. **SCOPE:** The procedures outlined herein are for use in the conduct of Special VFR operations within the (name) Airport surface area at or below _____ feet. These procedures are applicable only to aircraft equipped with functioning 2-way radio in order to effect a recall when required by traffic or weather conditions.

3. **RESPONSIBILITIES:** Upon request by the (name) FSS, the Center/Approach Control Facility may authorize Special VFR operations in the (name) Airport surface area for specific periods of time. The Center/Approach Control Facility must retain the authority to withdraw the provisions of this agreement at any time.

4. **PROCEDURES:**

a. Local Special VFR operations. The (name) FSS must not authorize more than one aircraft to operate simultaneously in the surface area unless pilots agree to maintain visual separation with other aircraft operating in the surface area.

b. IFR Arrivals and Departures. IFR arrivals and departures must be controlled by the (name) Center/Approach Control during the following periods:

(1) From 10 minutes prior to the estimated time of arrival of an IFR aircraft until it is over the approach fix until it is on the ground (IFR arrivals must not be interrupted by Special VFR operations.)

(2) From 10 minutes prior to the estimated time of departure of an IFR aircraft until it starts the surface area.

Air Traffic Manager, (Name) FSS

Air Traffic Manager, (Name) ARTCC/Approach Control



Embedding & Classification Pipeline

7

LOA Text

Embedding
Method

Embedding
(Numeric
Vector)

Classification
Model

Document
Label
(Civil or
Not-civil)

TF-IDF
GLoVe
Word2Vec
Doc2Vec
RoBERTa

KNN
Random Forest
Logistic Regression
SVM
Naïve Bayes
FNN
MLP
CNN
LSTM



Classification Models Comparison

Frequency-based embeddings

Term Frequency-Inverse Document Frequency (TF-IDF) Unigram	Document based method based on how frequently a term appears divided by how many documents it appears in
TF-IDF Bigram	Document based method based on how frequently pairs of terms appears divided by how many documents the pair appears in

Contextual Embeddings

GloVe	Pretrained method that creates word-embeddings based on co-occurrence matrix
Word2Vec	Document based method that trains a neural network on the continuous bag of words (CBOW) which is then used to create word embeddings
Doc2Vec	As Word2Vec but a document embedding is trained as well
RoBERTa	Large language model (LLM) pretrained on 160GB of English text which can produce word and document embeddings
Transfer learning (TL)	Fine-tuning a pretrained model on the 493 LOA documents



Classification Models Comparison

9

Classic machine learning (ML) methods

K-nearest neighbors (KNN)	Labels from closest embeddings
Random forests (RF)	Random ensemble of decision trees that average results
Logistic regression (LR)	Calculates the probability of an embedding's class
Support vector machines (SVM)	Calculates decision boundary that maximizes distance from each class
Naïve Bayes (NB)	Calculates class label from conditional probability

Neural ML methods

Feedforward neural network (FNN)	Single layer network with no backpropagation
Multilayer perceptron (MLP)	Multiple layers of FNN
Convolutional neural network (CNN)	Using convolution to extract features from the embeddings
Long short-term memory (LSTM)	Sequential focused model

Results



Results & Conclusion

11

Embeddings	Classification Models								
	KNN	RF	SVM	LR	NB	MLP	FNN	CNN	LSTM
TFIDF Unigram	0.919	0.750	0.877	0.880	0.860	0.854	0.714	0.840	0.788
TFIDF Unigram + Bigram	0.859	0.878	0.859	0.879	0.879	0.828	0.853	0.853	0.833
TFIDF Bigram	0.797	0.853	0.878	0.879	0.859	0.807	0.830	0.819	0.807
Averaged GloVe	0.818	0.814	0.816	0.778	0.740	0.814	0.407	0.359	0.569
TFIDF-Averaged GloVe	0.819	0.855	0.816	0.739	0.714	0.801	0.760	0.643	0.661
Averaged Word2Vec	0.838	0.816	0.877	0.800	0.797	0.845	0.699	0.754	0.756
TFIDF-Averaged Word2Vec	0.840	0.898	0.877	0.799	0.731	0.836	0.760	0.758	0.814
Doc2Vec	0.853	0.758	0.820	0.836	0.407	0.792	0.644	0.407	0.706
RoBERTa	0.756	0.740	0.838	0.820	0.720	0.359	0.359	0.359	0.359
TL Averaged GloVe	0.816	0.773	0.731	0.754	0.660	0.812	0.714	0.714	0.696
TL TFIDF-Averaged GloVe	0.699	0.773	0.811	0.731	0.700	0.803	0.773	0.660	0.760
TL RoBERTa	0.836	0.811	0.795	0.838	0.760	0.685	0.359	0.359	0.359



Results & Conclusion

Embeddings	Classification Models								
	KNN	RF	SVM	LR	NB	MLP	FNN	CNN	LSTM
TFIDF Unigram	0.919	0.750	0.877	0.880	0.860	0.854	0.714	0.840	0.788
TFIDF Unigram + Bigram	0.859	0.878	0.859	0.879	0.879	0.828	0.853	0.853	0.833
TFIDF Bigram	0.797	0.853	0.878	0.879	0.859	0.807	0.830	0.819	0.807
Averaged GloVe	0.818	0.814	0.816	0.778	0.740	0.814	0.407	0.450	0.569
TL Averaged GloVe	0.819	0.855	0.816	0.774	0.704	0.810	0.750	0.445	0.691
Doc2Vec	0.853	0.758	0.820	0.836	0.407	0.792	0.644	0.407	0.706
RoBERTa	0.756	0.740	0.838	0.820	0.720	0.359	0.359	0.359	0.359
TL Averaged GloVe	0.816	0.773	0.731	0.754	0.660	0.812	0.714	0.714	0.696
TL TFIDF-Averaged GloVe	0.699	0.773	0.811	0.731	0.700	0.803	0.773	0.660	0.760
TL RoBERTa	0.836	0.811	0.795	0.838	0.760	0.685	0.359	0.359	0.359

Use TF-IDF and SVM or LR to set a baseline before examining more complex methods for datasets of this size.



Future work

- Use TF-IDF and SVM to perform constrain classification, named entity recognition, and other NLP tasks for LOA digitization.
- Could more LOA and aviation documents improve the transfer learning performance?
- LLMs end-to-end: can we use transfer learning adapted LLM as an end-to-end classifier?

Questions?

David Nielsen: david.l.nielsen@nasa.gov