Applying Natural Language Processing Tools to Occurrence Reports

Loren Groff, PhD
Chief Data Scientist
U.S. NTSB
ICAO December 2018

Overview

- Motivation for automated tools
- Brief explanation of methods
- Examples
- Applications and limitations

Volume of Safety Data

- US NTSB: 1,800 accident, serious incident, and other occurrence cases per year
- US FAA: additional 2500+ mandatory incident reports per year, and more than 670 voluntary reporting programs
- European Central Repository: 250,000 mandatory and voluntary reports per year

Challenge

- Large amount of structured and unstructured data that must be reviewed and validated to evaluate safety concerns
- An expectation of results for all this effort
- Time and resource pressures

Natural language processing (NLP)

- Automated classification, keyword extraction, and identification of similar records
- Triage incoming reports, generate or validate structured data, and improve quality control of large occurrence databases

NLP and Machine Learning Tools

- Open source packages can provide exceptional performance
  - *spaCy* – text preprocessing, tagging, parsing and entity recognition
  - *scikit-learn* – modeling, classification
  - *gensim* – word embeddings and semantic text analyses
- Source code and trained models can be shared and adapted

Example 1 – Classifying Event Notifications

- Input: accident and incident notifications

N11810 CRASHED WHILE ATTEMPTING A RETURN TO KEWK AFTER REPORTING AN ENGINE OUT.

- Output: CAST/ICAO occurrence category
- Training data: Analysis narratives and occurrence classifications from 10,000 NTSB database records

Classifying Event Notifications - Methods

- Use *spaCy* to parse text, remove stopwords, and combine different word forms

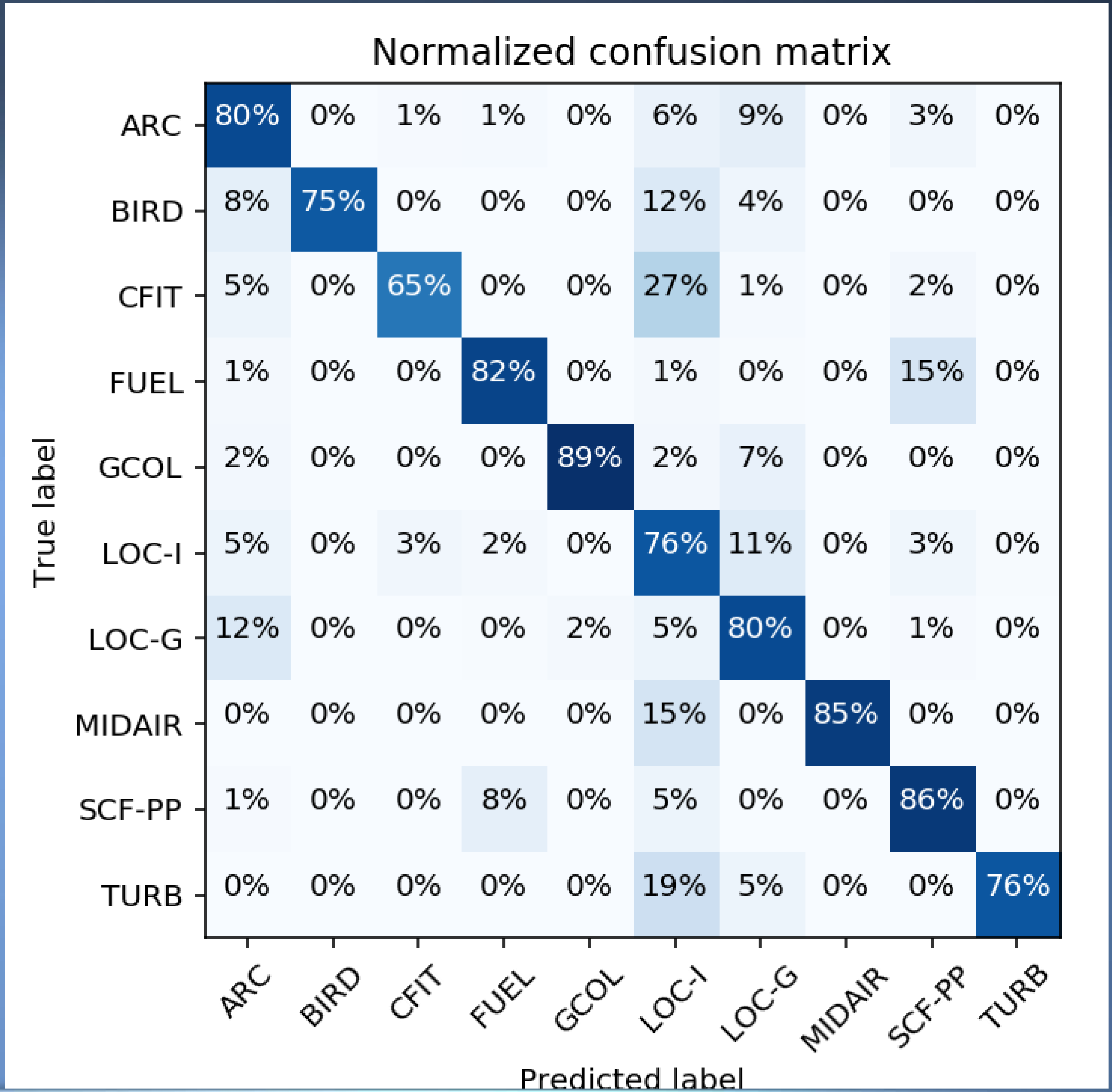
The flightcrew commenced the takeoff roll after lining up with the runway centerline

['flightcrew','commence','takeoff','roll','line','runway',’centerline’]

- Use *scikit-learn* to split data into training and test datasets, generate word vectors (tf-idf), and train model (SVC)

Results

- 80% accuracy with minimal model tuning and data cleaning
- Classification results used to improve model



Applying Classification Model

- Use model to predict occurrence categories from text notifications of new events
- Can also be used to validate new or existing occurrence category data

ERA13CA392 predicted = Abnormal Runway Contact | actual = Loss of Control In-Flight
ERA11LA461 predicted = Fuel Related | actual = System/Component Failure - Powerplant
ERA13LA053 predicted = Fuel Related | actual = System/Component Failure - Powerplant
NYC08CA170 predicted = Abnormal Runway Contact | actual = Other
ERA12LA584 predicted = Fuel Related | actual = System/Component Failure - Powerplant
ANC13CA109 predicted = Loss of Control on Ground | actual = Loss of Control In-Flight

Applying Classification Model

For added value

- Parsed input text can be compared with civil registry to identify aircraft
- Outputs can be compared to existing records to return attributes of recently investigated similar occurrences

Demo

Notification text: N11810 CRASHED WHILE ATTEMPTING RETURN TO RUNWAY AFTER REPORTING AN ENGINE OUT.

****

N11810 is a CESSNA 150

notification text is similar to System/Component Failure - Powerplant (67.8%)

****

Recent 150 case(s) citing System/Component Failure - Powerplant:

CEN16LA306

A partial loss of engine power due to a fatigue crack of the No. 2 cylinder cooling fin, which resulted in failure of the No. 2 cylinder

Example 2 – Semantic Analysis

- Input: NTSB investigations
- Output: A model specifically trained to identify occurrence report narratives describing similar issues
- Training data: NTSB database, reports, safety recommendations, and technical reports

Semantic Analysis - Methods

- *gensim* used to train word context from aviation occurrence reports
- Calculates word vectors from a window of words appearing before and after

- Trained word vectors loaded in *spaCy* to enable NLP tailored to aviation occurrences

Applying Semantic Model

- Use trained model to compare new investigation narratives with existing reports and identify events with similar safety issues
- Semantic model can also be used to identify keywords and related terms
- Works best with longer narratives

Demo

Selected event: SEA86LA050 (100.0% similar): DURING LANDING ROLL ON THE 1800 FOOT LONG 30 FOOT WIDE WET RUNWAY, THE AIRCRAFT VEERED RIGHT OFF THE RUNWAY AND THEN NOSED OVER. THE PILOT APPLIED BRAKES TWICE WITH NO EFFECT AND THE THIRD TIME THEY GRABBED. NO MALFUNCTIONS WERE FOUND DURING INSPECTION OR EXAMINATION. HYDROPLANING PROBABLY OCCURRED.

Most similar narrative: WPR10CA176 (85.9% similar): The pilot reported that he felt a slight wind gust on touchdown. The landing was smooth and occurred in the first 1/3 of the runway. When the nosewheel made contact with the runway, the airplane veered left. The pilot applied right rudder and braking in an attempt to correct the airplane, but before it stopped the airplane veered off the runway, went down an embankment, and nosed over. No mechanical malfunctions or failures were found during the postaccident examination of the airplane.

Second most similar narrative: SEA99LA154 (85.8% similar): The pilot reported that during the landing roll on a wet runway surface, he applied braking action and the airplane skidded to the right and veered off the runway. The aircraft ground-looped and came to rest in the mud next to the runway. The pilot reported that there were no mechanical failures or malfunctions with the airplane at the time of the accident.

Applications

- Similar machine learning techniques can be applied to automate transcription, classify images, and detect anomalies or unique occurrences
- Performance can be further improved using an ensemble of multiple models

…and Limitations

- Proper data cleaning and preprocessing can require significant effort
- Requires care to avoid bias and to properly train and maintain models
- Only as good as the training data, and can't create knowledge that isn't present in the source reports

Cited Software Packages

spaCy - spacy.io

gensim - radimrehurek.com/gensim/

scikit-learn - scikit-learn.org

National Transportation Safety Board