**Overview and Situation**

The impact of aircraft noise on a “community” is well regulated by the Federal Aviation Administration (FAA). Aircraft manufacturers such as Boeing must adhere to these standards. Data retrieval and analysis for aircraft community noise testing requires extensive equipment and human labor. Moreover, such data, primarily in the form of acoustic signals, often becomes “contaminated” by the environment. Sources of contamination include, but not limited to, bird chirps, widelife/livestock vocalizations, insect noises, traffic noises, and aircraft noises borne from aircraft other than the designated airplane of interest. Too much contamination in the acoustic signals will have corrupted the data, rendering it unusable.

Our project is an environmental community noise detector. The goal is to monitor, detect, and classify the presence of environmental noise contamination in real-time. The scope of the classification should be sufficient to guide testing engineers on necessary recourse: whether the flight test need be redone, what type(s) of contaminant noise sources need to be removed the testing site, etc. Human labor designated for these tasks, while reasonably effective, is taxing and not cost-efficient. We propose a signal processing and machine learning-based system for this problem.

**Target Audience**

Ideally, our project would remove the need for multiple on-site work stations and lead to significant cost reductions for Boeing. The performance of our system will have increased accuracy and classification, further streamlining regulatory noise analysis at the Boeing Test and Evaluation unit which oversees this work.

**Constraints and Requirements**

The project, its software code, and data, must not limit the commercial use by The Boeing Company of such code or data, or models depending on such code or data. Furthremore, all software code must be executable in MATLAB, the preferred language for Boeing Test & Evaluation noise testing division.

**Deliverables:**

* Trained algorithm, explanation of its selection, and theoretical background, for detecting audio signals contaminated by environmental noise sources.
* Measures of algorithm performance when deployed in various “environments” – i.e. for a wide spectrum of noisy audio signals
* Algorithm pros and cons, with a listing of its benefits over alternative methods utilized in literature, as well as its limitations
* Well-documented process for algorithm retraining and overall project reproducibility

**Information Required From the Sponsor**

We require that the example training data (acoustic signals with/without contamination) must be classified.

**Proposal**

As aforementioned, we seek to research, develop and implement a signal processing into machine-learning pipeline for automation of environmental noise contamination associated acoustic measurements. A proposed methodology for achieving such results could be:

1. The data from the sponsor (Boeing BT & E) will be of the .wav file format (audio signal). Therefore the first step will be to extract features from these signals. This portion of the project will require research and implementation of novel audio signal processing techniques in MATLAB. Signal filtering, Fourier analysis, signal-to-noise ratio/signal energy spectrum methods, wavelet transforms and other time-series transient signal classification methods are potential methods for this purpose.
2. Given a set of features/covariates from (1) above, we develop classification algorithms using the training data. These machine-learning algorithms will range from basic to complex: i.e. from softmax multi-classification and multi-layer perceptrons, to recurrent neural networks. Because we plan on delivering an automated system *for real-time* analysis by on-site engineers, the algorithm must be efficient and not require heavy computational resources. Furthermore the algorithm(s) must be deployable in MATLAB
3. Steps (1) and (2) will be continuously re-evaluated according to model performance(s). Perhaps the features engineered in (1) are not sufficient to result in accurate classification, and different signal processing techniques must be researched and deployed. Likewise for the ML portion in (2)
4. To assist in the prior steps, we may augment the data set from the sponsor with other examples of environmental noise, such as data from the United States National Park Service.
5. Once satisfies with performance of the pipeline, we will provide robust and sufficient documentation to retrain the algorithm(s) and reproduce and/or adapt our project for potentially differing situations for BT&E.