USE OF LOGISTIC REGRESSION MODELS AS A SUPERVISED LEARNING ALGORITHM TO IDENTIFY IMPULSIVE SOUNDS IN MONITORED SOUND DATA

Harry Ao Cai *[1]
[1]HGC Engineering, Mississauga, Ontario, Canada

1 Introduction

Impulsive sounds, characterized by their transient nature, often pose challenges in sound monitoring applications such as environmental noise assessments. Noise emission regulations, such as those given in NPC-300 of the Ontario Ministry of Environment, Conservation and Parks[1], and specific project needs can require impulsive sounds to be identified and processed separately from other impulsive and non-impulsive sounds. This paper investigates the use of logistic regression models, implemented as a supervised learning algorithm, to identify impulsive sounds from monitored sound data and to calculate their logarithmic mean impulse sound level (LLM).

2 Background and Existing Methods

Several methods exist to identify impulsive sounds, such as by listening to recorded audio or manual examination of the logged frequency-spectral data. These methods can be time-consuming for long-term monitoring projects. Methods that automatically tag or identify segments of the data based on a trigger criterion still requires manual a review of each segment.

Recent advances in data science have also introduced various audio classification machine learning algorithms. These methods can involve large volumes of audio recording files and require extensive computing power. The presented method aims to automate some of the analysis procedure required for handling large volumes of spectral data involving impulsive sounds without processing audio files.

3 Methodology

3.1 Data Source

Logged spectral sound pressure data, in the form of raw third-octave impulsive sound pressure levels (LISPL) and equivalent sound pressure level (LEQ), measured in 1 second windows from 6.3 Hz to 20 kHz, from sound level meters were used as input.

The data source for this paper came from sound pressure level data measured at four locations around a facility containing a pulse jet dust collector system. In this case, the targeted impulsive sound are the short bursts of pressurized air from the dust collectors, which were observed to emit an impulsive sound with similar amplitude. Four datasets were assessed, each containing around 40 minutes of data at each monitoring location.

3.2 Logistic Regression Model

A logistic regression model was constructed as a supervised machine learning algorithm to classify each 1-second window of frequency-spectral data, with each second being a data point. Logistic regression is a statistic model, which is similar to linear regression such that it predicts the probability of an event based on several independent variables, used to predict discrete categorical values.

The model presented herein outputs a binary prediction; an output of “1” if a data point is predicted to be an instance of the targeted impulsive sound and a “0” if otherwise. Raw LISPL and LEQ data were processed to create two sets of custom variables that mirror the manual method of identifying impulsive sounds; the arithmetic difference in LISPL between two data points, and the arithmetic difference between LISPL and LEQ.

3.3 Training and Evaluation

The model was constructed using Python, using pandas libraries to store and manipulate data and SciPy libraries to construct the logistic regression model. Approximately 20% of each dataset was labelled manually for the targeted impulsive sound and was used train the model. The logistic regression then assigns parameters of different weights to the custom variables, which ideally would capture the determinant characteristics of the targeted impulsive sound.

A logistic regression model was created for each dataset, since the frequency content of the impulsive sound can vary from measurement location to location. The same model, constructed for the first dataset, was also used to predict impulsive sounds in the other three datasets to investigate the feasibility of reusing the same model for different measurement locations.

4 Results

4.1 Sound Identification

The model was validated by splitting the training dataset into a “learning” and “validation” subset, such that the parameters of the regression model were developed on the “learning” subset and tested on the “validation” subset. The confusion matrix below, which is a measurement of machine learning classification, shows the performance of the first dataset (for the first measurement location), yielding a total accuracy of 97 %. It is also noted that the performance of the model was noticeably improved with custom features instead of relying on raw LFISPL and LEQ data. It is noted that the models are more prone to making false-positive predictions than false-

* hcai@hgcengineering.com

negatives. This makes intuitive sense as the learning data contains far more negatives than positives.

Figure 2 below is a time-history plot of A-weighted LISPL values of dataset 1, with the impulsive sounds of interest marked in orange, shown for datapoints around the cut-off between labelled datapoints and the unlabeled data.

4.2 Determining LLM

After the model has predicted the impulsive sounds, LLMs were calculated for the manually labelled impulsive sounds and for the predicted impulsive sounds. Table 1 shows the labelled data LLM and predicted LLM for each dataset. Ideally, the predicted and the labelled LLM would be similar, assuming that the targeted impulsive sound and the soundscape at the measurement location remained do not significantly change throughout the duration.

5 Discussion

The results show that for monitoring locations where the general soundscape remained the same, the impulsive sounds identified by the model matches very closely to the manually labelled data, implying that the model was able to identify impulsive sounds accurately. However, in cases where the soundscape varied throughout the monitoring duration, such as in the case of dataset 4 with numerous helicopter pass-bys, the model can over-predict and result in a much higher LLM than the labelled data. Since the helicopter pass-bys were not present in the training data, the frequency-spectra data of helicopter sounds can resemble the characteristics of the targeted impulsive sound, triggering false-positives.

The use of one prediction model for multiple datasets was investigated since the targeted sound is expected to have similar spectral characteristics at different locations. This approach would only require one training dataset, reducing the amount of work required to manually label data. The results also show that the model tends to over-predict and when presented with new transient and intrusive, more so than using dataset-specific models. This suggests that dataset-specific models are required for each monitoring location, to account for different soundscapes.

6 Conclusion

The supervised learning model can identify targeted impulsive sounds in scenarios where the training dataset can sufficiently represent the soundscape at the measurement location. However, variations in their soundscape can greatly reduce the accuracy of the learning model. Changes in background levels or the presence of sounds that were not described in the training dataset can trigger false-positive predictions, which can over-predict the LLM. Further work should be conducted to isolate the background sound and the non-impulsive sounds from the targeted impulsive sound. Another limitation of this model is the inability to identifying impulsive sounds when the targeted impulsive sound levels are lower than the background sound level, which could introduce a bias such that only the louder instances of the targeted sound are identified.

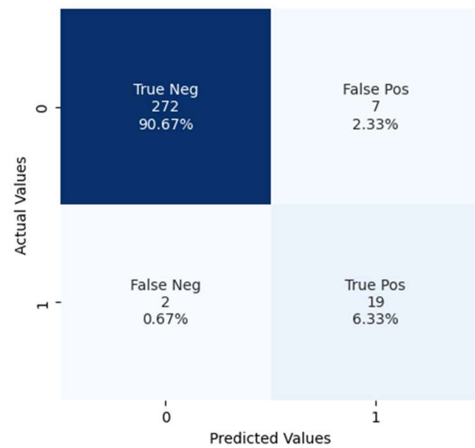

Figure 1: Confusion Table of Dataset 1 Model

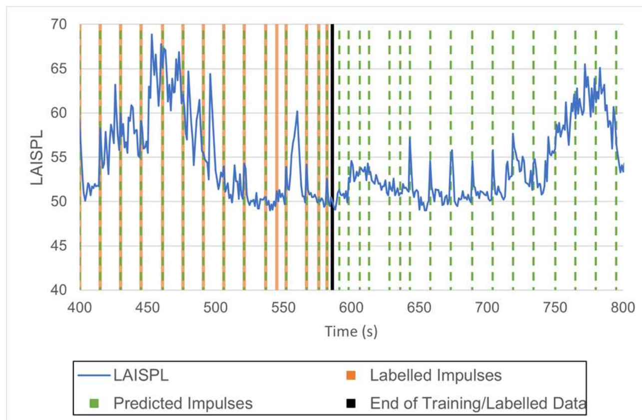

Figure 2: Identified Impulses over Time

Table 1: Model Results

		LLM [dBAI]			
Dataset	Acc.	Manually Labelled Data	Predicted using dataset-Specific Model	Predicted using Dataset 1 Model	Qualitative Notes Over Duration of Dataset
1	97%	58.2	58.2	58.2	No significant variations
2	98%	55.5	57.4	61.5	Increasing intrusive sounds
3	94%	57.0	57.1	57.6	No significant variations
4	96%	58.4	64.0	65.1	Helicopter pass-bys & intrusive sounds

References

[1] Ontario Ministry of the Environment. (August 2013). *NPC-300, Environmental Noise Guideline Stationary and Transportation Sources - Approval and Planning.* Retrieved from https://www.ontario.ca/page/environmental-noise-guideline-stationary-and-transportation-sources-approval-and-planning