Training Session 8: Hoo hoo hoo are you?

Before the session:
1. Watch Larissa's An Introduction to Machine Learning 2. Watch Laurel's Creating Training Datasets for Machine Learning Models

Goals for today's lesson:

1. Understand how to apply machine learning for automated bioacoustics analysis
2. Familiarize yourself with key machine learning terms and concepts
3. Create and annotate a training dataset for bioacoustic analysis
4. Implement and evaluate a machine learning model on annotated data

Carmelo Lopez Abad

Philippe Verbelen

Breakout Room: Room XX

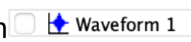
Participants: XX, XX, XX, XX
Note-taker: XX
Screen-sharer: XX

Part 1: Group warm-up

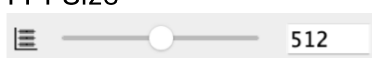
Question 1: What products or software that you have used use machine learning? What are they using for their training data? What features are they using to train their models?

Part 2: Annotating sound files in Raven

1. Open all sound files in Raven 1.6 together using the “Page sound” option

2. Deselect the wave form 

3. Before annotating, record your chosen spectrogram settings here:

Color palette 	Brightness 	Contrast 	FFT Size 	Low freq zoom (Hz)	High freq zoom (Hz)

4. Open the selection table entitled “TS8_Owl_Selection_Table_Participant_Copy.selections.txt” by clicking “File” → . This selection table contains 67 annotations.

5. Add the measurements “Begin File”, “Begin Path”, and “File Offset (s)”.

6. Review and correct the annotations. Most are well-done, but there are some mistakes. 5 of them have typos in the annotation labels, 6 of the annotation boxes are not tightly bound, and 7 of the annotations were wrongly given to the other species. (Hint: Use the Begin File names to see which were incorrectly labelled)

7. Make additional annotations. Aim for a total of 100 total annotations. For the Rinjani scops owl (*Otus jolandae*) use the abbreviation **RISO**. For the Sunda scops owl, use the abbreviation **SUSO**. Importantly, try to annotate across a range of recordings to best capture the variability present.

If your group wants a challenge, you have the option to annotate for a third owl species, the Moluccan Scops-owl. Use **MOSO** for the class label. Now, you will be building a machine learning classifier for three owl species.

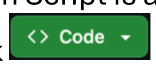
8. Add acoustic measurements. These are the **features** that will be used to train the machine learning model. Add at least 5 acoustic measurements and no more than 10 acoustic measurements. An explanation of the various acoustic measurements is provided below in **Table 1**.

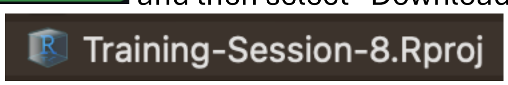
9. Question: What are the main sources of variability associated with the calls in these recordings? List at least four types of variation present that would be helpful to account for in the training dataset:

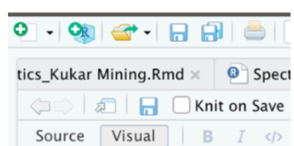
    1. ________________
    2. ________________
    3. ________________
    4. ________________

10. Re-save the selection table with your changes.

Part 3: Machine Learning in R

11. The R Markdown Script is available on the BEAT Github Page. To download the materials, click 

and then select "Download ZIP".
12. Double-click the 

to open this R Project in R Studio.
13. Now, click the "TS8_RMarkdown_Code.Rmd" script to open up the R Markdown script.
14. Make sure to switch it to "Visual" mode 

15. Follow the directions in the R Markdown Script

Part 4: Model Results

After finishing the R exercise, document the settings and results of your final model below.

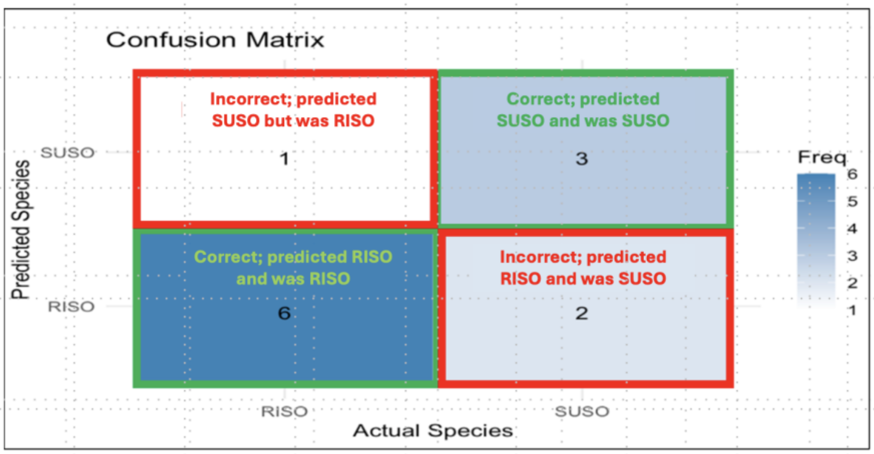
Training / testing split: _____ %
Number of examples in training dataset: ________ RISO, ________ SUSO, ________ MOSO
Number of examples in testing dataset: ________ RISO, ________ SUSO ________ MOSO

List the features included in the model:

1. ________________________
2. ________________________
3. ________________________
4. ________________________
5. ________________________
(Continue as needed)

Confusion matrix:
*** Paste your confusion matrix here***

Example confusion matrix:



Model performance:
There are many ways to evaluate the performance of a model. We will focus more on evaluating model performance in Training Session 10.

For now, just calculate the following:

How many examples did the model correctly identify as RISO? ________ (6 in the above example).
How many examples did the model incorrectly identify as RISO? _______ (2 in the above example).
How many examples did the model correctly identify as SUSO? _______ (3 in the above example).
How many examples did the model incorrectly identify as SUSO? ______ (1 in the above example).

Discussion questions

Q1: Would you say this is a good model? Why or why not?

Q2: Which features did you find to be the most important in predicting between owl species? Why? Would you say their songs differed more in frequency or temporal elements?

Q3: How can you apply what you’ve learned today to your own bioacoustics projects?

Table 1. Acoustic measurements available in Raven 1.65 (modified from Batist et al. 2024)

Measurement	Units	Definition of measure
		1. Frequency Measurements
Delta frequency	Hz	Frequency range between the high frequency and low frequency bounds of the selection box
Freq 5% (Hz)	Hz	Frequency at which signal is divided in two frequency intervals of 5% and 95% energy respectively
Freq 25% (Hz)	Hz	Frequency at which signal is divided in two frequency intervals of 25% and 75% energy respectively
Center frequency	Hz	Frequency at which signal is divided in two frequency intervals of equal energy (50% and 50%)
Freq 75% (Hz)	Hz	Frequency at which signal is divided in two frequency intervals of 75% and 25% energy respectively
Freq 95% (Hz)	Hz	Frequency at which signal is divided in two frequency intervals of 95% and 5% energy respectively
Peak frequency	Hz	Frequency at which the highest amplitude occurs across a signal
BW 50%	Hz	Freq 75% (Hz) – Freq 25% (Hz)
BW 90%	Hz	Freq 95% (Hz) – Freq 5% (Hz)
		2. Time Measurements
Delta Time (s)	sec	End time of selection box – start time of selection box
Duration 90%	sec	Shortest duration that contains 90% of the energy
Duration 50%	sec	Shortest duration that contains 50% of the energy
Time 25% Relative	sec	Time at which signal is divided in two time intervals of 25% and 75% energy respectively
Center time Relative	sec	Time at which signal is divided in two time intervals of equal energy (50% and 50%)
Time 75% Relative	sec	Time at which signal is divided in two time intervals of 75% and 25% energy respectively
Interquartile range of time	sec	Time range between Time 75% Rel. and Time 25% Rel.
		3. Tonality Measurements
PFC Number of Inflection Points	none	The "Peak Frequency Contour Number of Inflection Points" refers to the number of times the peak frequency contour changes direction within a given selection. The peak frequency

		contour represents the highest amplitude frequency at each time slice of the signal.
Avg. Entropy (bits)	0-1 scale	Spectral entropy is a measure of the randomness or disorder within the frequency distribution of a sound signal. It is calculated based on the distribution of power across different frequency bands. 0=pure tone, 1=noisy tone

Key terms:

Modified from Google's Introduction to Machine Learning Course

Machine learning

Machine learning is the process of training a model to make predictions or find patterns on never-before-seen data.

Labels

A label is what we are trying to predict or identify in our analysis. In bioacoustics, labels can include animal species, call type, the identity of an individual, the size of an animal, and much more. For this exercise, our labels are owl species (SUSO, RISO, MOSO).

Features

A feature is an input variable used to train a model. Simple machine learning models might use just one feature, while more complex models can use millions. In this exercise, our features include acoustic measurements such as "Center Frequency," "Duration 90%," and "Peak Frequency."

Examples

An example is a specific instance of data. In a Raven Selection Table, each row represents one example. Examples can be either labeled or unlabeled. Labeled examples include both the input features and the correct output label, which we use to train our model. Once the model is trained with these labeled examples, it can predict the labels for new, unlabeled examples.

Models

A model defines the relationship between features and label. For example, a gibbon model might learn to associate certain acoustic features strongly with gibbon sounds. Let's highlight two phases of a model's life:

- **Training** means creating the model by showing the model labeled examples. Thus, the model gradually learns the relationships between features and label.
- **Inference** means applying the trained model to unlabeled examples. That is, you use the trained model to make useful predictions. For example, during inference, you can predict "species" for new unlabeled examples.

Helpful resources on machine learning:

Tutorials:
Google's Introduction to Machine Learning Course (available in English and Indonesian)

Articles:
Visualize introduction to machine learning (available in English and Indonesian):

Videos:
BioacousTalks: Machine Learning for Tropical Acoustic Monitoring