Open-Source machine learning BANTER acoustic classification of a cryptic echolocating species

Shannon Rankin[a,1,*], Taiki Sakai[a,2], Frederick Archer[a,3], Jay Barlow, Danielle Cholewiak, Annamaria DeAngelis, Jennifer L. K. McCullough, Erin Oleson, Melissa Soldevilla

[a] *Southwest Fisheries Science Center, Marine Mammal & Turtle Division, 8901 La Jolla Shores Dr., La Jolla, 92037*

Abstract

Passive acoustic monitoring is increasingly used for assessing populations of marine mammals; however, analysis of large datasets is limited by our ability to easily classify sounds detected. Classification of beaked whale acoustic events, in particular, require evaluation of multiple lines of evidence by expert analysts. Here we present a highly automated approach to acoustic detection and classification using supervised machine learning and open source software methods. Data from four large scale surveys of beaked whales (North Atlantic, South Atlantic, Hawaii, and Eastern Pacific) were analyzed PAMGuard (acoustic detection), PAMpal (acoustic analysis) and BANTER (hierarchical random forest classifier). Overall correct classification results ranged from 86% for the South Atlantic data to 99% for the North Atlantic. Results for many species could likely be improved with increased sample sizes and consideration of alternative automated detectors. These methods provide a highly automated approach to acoustic detection and classification using open source methods that can be readily adopted for species and geographic regions.

Keywords: bioacoustics, machine learning, random forest, species classification, marine mammals, passive acoustic monitoring, beaked whale

Please make sure that your manuscript follows the guidelines in the Guide for Authors of the relevant journal. It is not necessary to typeset your manuscript in exactly the same way as an article, unless you are submitting to a camera-ready copy (CRC) journal.

For detailed instructions regarding the elsevier article class, see https://www.elsevier.com/authors/policies-and-guidelines/latex-instructions

1. Bibliography styles

Here are two sample references: Feynman and Vernon Jr. (1963) Dirac (1953).

By default, natbib will be used with the `authoryear` style, set in `classoption` variable in YAML. You can sets extra options with `natbiboptions` variable in YAML header. Example

*Corresponding author

Email addresses: `shannon.rankin@noaa.gov` (Shannon Rankin), `taiki.sakai@noaa.gov` (Taiki Sakai), `eric.archer@noaa.gov` (Frederick Archer), `jaybarlow33@yahoo.com` (Jay Barlow), `danielle.cholewiak@noaa.gov` (Danielle Cholewiak), `annamaria.deangelis@noaa.gov` (Annamaria DeAngelis), `jennifer.mccullough@noaa.gov` (Jennifer L. K. McCullough), `erin.oleson@noaa.gov` (Erin Oleson), `melissa.soldevilla@noaa.gov` (Melissa Soldevilla)

[1] This is the first author footnote.

[2] Another author footnote, this is a very long footnote and it should be a really long footnote. But this footnote is not yet sufficiently long enough to make two lines of footnote text.

[3] Yet another author footnote.

`natbiboptions: longnamesfirst,angle,semicolon`

There are various more specific bibliography styles available at https://support.stmdocs.in/wiki/index.php?title=Model-wise_bibliographic_style_files. To use one of these, add it in the header using, for example, `biblio-style: model1-num-names`.

1.1. *Using CSL*

If `cite-method` is set to `citeproc` in `elsevier_article()`, then pandoc is used for citations instead of `natbib`. In this case, the `csl` option is used to format the references. By default, this template will provide an appropriate style, but alternative `csl` files are available from https://www.zotero.org/styles?q=elsevier. These can be downloaded and stored locally, or the url can be used as in the example header.

2. Equations

Here is an equation:

$$f_X(x) = \left(\frac{\alpha}{\beta}\right)\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}; \alpha, \beta, x > 0.$$

Inline equations work as well: $\sum_{i=2}^{\infty}\{\alpha_i^{\beta}\}$

3. Figures and tables

Figure 1 is generated using an R chunk.

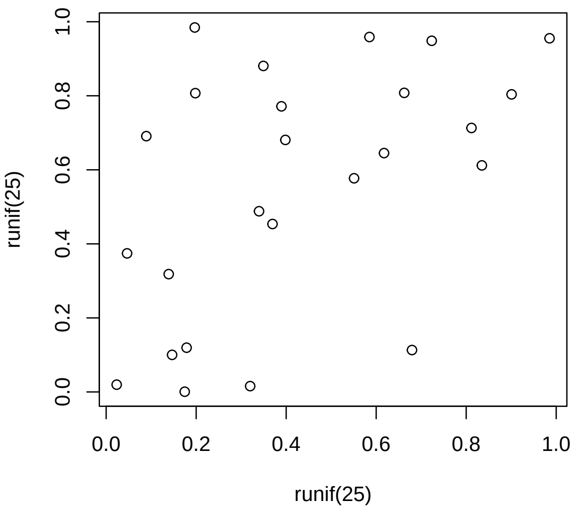

Figure 1: A meaningless scatterplot

4. Tables coming from R

Tables can also be generated using R chunks, as shown in Table 1 example.

```r
knitr::kable(head(mtcars)[,1:4])
```

Table 1: Caption centered above table

	mpg	cyl	disp	hp
Mazda RX4	21.0	6	160	110
Mazda RX4 Wag	21.0	6	160	110
Datsun 710	22.8	4	108	93
Hornet 4 Drive	21.4	6	258	110
Hornet Sportabout	18.7	8	360	175
Valiant	18.1	6	225	105

References

Dirac, P.A.M., 1953. The Lorentz transformation and absolute time. Physica 19, 888–896. doi:10.1016/S0031-8914(53)80099-6.

Feynman, R.P., Vernon Jr., F.L., 1963. The theory of a general quantum system interacting with a linear dissipative system. Annals of Physics 24, 118–173. doi:10.1016/0003-4916(63)90068-X.