

Fur Seal Data README

Audrey Der

Updated: August 7, 2018

Here's a link to the repository of the original data. There are 72 files, each over 7000 datapoints long, recorded at 25 Hz. There was no documentation available (that I could find) but all the data should have converted properly. Please let me know if you run into any issues; I can be reached at ader003@ucr.edu.

1 Characteristics

- **depth**
- **x, y, z**
- **behaviour**: 26 behaviors, converted from strings to numbers as follows:

| | | | | |
|----------------|------------------|----------------|---------------|-------------------|
| 'other': 0 | 'moving': 1 | 'in': 2 | 'foraging': 3 | 'manipulation': 4 |
| 'chewing': 5 | 'thrash': 6 | 'slow': 7 | 'swimming': 8 | 'holdntear': 9 |
| 'still': 10 | 'out': 11 | 'sitting': 12 | 'sailing': 13 | 'rubbing': 14 |
| 'feeding': 15 | 'scratch': 16 | 'lying': 17 | 'jump': 18 | 'fast': 19 |
| 'facerub': 20 | 'bowing': 21 | 'juggling': 22 | 'playing': 23 | 'shake': 24 |
| 'grooming': 25 | 'stationary': 26 | | | |

- **type**: types of behaviors, converted from strings to numbers as follows:
 - Other': 0
 - 'Travelling': 1
 - 'Foraging': 2
 - 'Resting': 3
 - 'Grooming': 4
- **place**: the seal's location, converted from strings to numbers as follows:
 - 'land': 1
 - 'surface': 0
 - 'underwater': -1

2 What was Removed

- **number, value**: no obvious significance or relationship with any other columns or data
- **behaviour_event**: information from 'number' and 'behaviour' concatenated
- **type_event**: information from 'value' and 'type' concatenated
- **date, time**: the date of presumed recording and time

- **doe**: presumably 'date of entry'; matches the original filename on the github repository
- **raw**: a column with much missing data
- **fps**: iterating numbers
- **location**: redundant information

3 Using the Jupyter Notebook

Note: this notebook is hardcoded in many places, and as a result, is specific to this particular dataset of cleaned fur seal data. Edit at your own risk.

3.1 What You'll Need

Aside from jupyter notebook, you will also need:

- Python 3.6
- SQLite
- Apache Spark

Ideally, the notebook (*findbehaviours.ipynb*) will be **in the same directory as the folder of data you'd like to consider** (not *in* the folder of data, just the same directory as the folder).

Aside from the database file, given a specified behavior to generate graphs for, the script will generate one .png for each range classified as the behavior. That is, considering any one behavior b over the 72 files (f_1, f_2, \dots, f_{72}) in the dataset, there will be r_1, r_2, \dots, r_n number of ranges classified as b , and $\sum_{n=1}^{72} r_i$ number of .png files will be generated. The figures are set to be transparent.

3.2 Making Changes to the Notebook

Looking inside the notebook, I've denoted particular important notes and parameters that need to be (mandatory changes are denoted by (*)) or can be changed with comments in capital letters. I've also compiled a list of places below:

1. Cell 3, Line 7: Comment/uncomment as desired to drop the table
2. Cell 3, Line 8: Change the name of the database created (so multiple .db files can be saved)
3. Cell 4, Line 2: Change the database connection to your preferred database
4. *Cell 4, Line 9; Cell 5, Line 2: **Change the behavior** you'd like graphed
5. *Cell 5, Line 13: Change data directory name
6. Cell 5, Lines 20-25; Cell 5, Lines 31-36: Change the features graphed (by default, depth, x, y, and z accelerometer data is set)
7. Cell 5, Line 46: figure dpi and figure dimensions can be changed
8. Cell 5, Line 58: file format and status of transparency can be changed