

RADAR 2022 – Artificial Intelligence - Month 9 – Deep Learning Project

Predicting the Age of Abalone

Hello Radar-ees!

This is our second ML project for the Radar Artificial Intelligence stream.

It will conclude our semester 2 work on *Introduction to Deep Learning* and next month we will move on to *Computer Vision*!

Project definition

Build a neural network model that predicts Abalone age (number of rings).

(<https://en.wikipedia.org/wiki/Abalone>)

Make this model as accurate as possible and submit your code and results!

Data

The dataset can be found on the UCL ML Repository:

<https://archive.ics.uci.edu/ml/datasets/Abalone>

Can be downloaded from here:

<https://archive.ics.uci.edu/ml/machine-learning-databases/abalone/>

```
[1]: import pandas as pd

[3]: header_list = ["sex", "length", "diameter", "height", "whole weight", "shucked weight", "viscera weight", "shell weight", "rings"]

[4]: pd.read_csv("abalone.data", names = header_list )

[4]:
```

	sex	length	diameter	height	whole weight	shucked weight	viscera weight	shell weight	rings
0	M	0.455	0.365	0.095	0.5140	0.2245	0.1010	0.1500	15
1	M	0.350	0.265	0.090	0.2255	0.0995	0.0485	0.0700	7
2	F	0.530	0.420	0.135	0.6770	0.2565	0.1415	0.2100	9
3	M	0.440	0.365	0.125	0.5160	0.2155	0.1140	0.1550	10
4	I	0.330	0.255	0.080	0.2050	0.0895	0.0395	0.0550	7
...
4172	F	0.565	0.450	0.165	0.8870	0.3700	0.2390	0.2490	11
4173	M	0.590	0.440	0.135	0.9660	0.4390	0.2145	0.2605	10
4174	M	0.600	0.475	0.205	1.1760	0.5255	0.2875	0.3080	9
4175	F	0.625	0.485	0.150	1.0945	0.5310	0.2610	0.2960	10
4176	M	0.710	0.555	0.195	1.9485	0.9455	0.3765	0.4950	12

4177 rows x 9 columns

The *abalone.data* file can be loaded into Pandas with the *read_csv()* method.

The *abalone.names* file can be opened in a text editor and is a description of the data.

As we can see, the features are:

- Sex
- Length
- Diameter
- Height
- Whole weight
- Shucked weight
- Viscera weight
- Shell weight

And we will be using these to try and predict the label:

- Rings

Pre-requisites

- LiL courses: [Radar AI Tech Stream - Semester 2](#)

How will this be judged?

We will select the best entry by overall quality of submission and quality of results – we will evaluate of the following criteria:

1. Good quality code
2. Comments briefly describing sections of code
3. Notes explaining reasoning behind decisions made
4. Good quality data exploration and visualisation
5. Correct feature engineering and feature selection
6. Understanding of how to build a Multi-layer Perceptron (MLP) in Keras
7. Trial of different MLP architectures and regularisation
8. Analysis of results

Extension

Finished? Too easy? Try:

- Building an MLP that can predict the quality of wine from the data features using this data set: <https://archive.ics.uci.edu/ml/datasets/Wine+Quality>
- Let me know if you try this, the results are interesting!

Conclusion

There are many good solutions to this problem scattered around the internet. Try to solve it yourself first before using them for guidance.

Enjoy and best of luck!

Submission deadline Friday 25th November 2022.