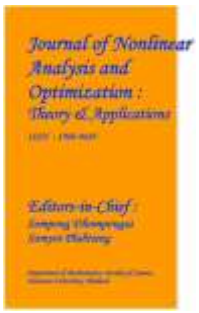
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, Issue. 2 : 2023
ISSN : **1906-9685**



AGE PREDICTION OF ABALONE USING CREATED NEURAL NETWORKS

[1]Dr. E.Nagarjuna, [2]Ch. Padma, [3]D.Rammohanreddy, [4]Baireddy Bhargav Reddy
[1]Professor, [2,3]Assistant Professor, [4]Student, Dept. of Computer Science Engineering, Newton's Institute of Engineering, Macherla, Andhra Pradesh, India.

ABSTRACT:
Predicting the age of an abalones,which are very popular in foreign countries, is considered as utmost important, in order to eliminate the cost of traditional cutting methods, which are very expensive. In this paper, artificial neural networks concept is used to predict the age of abalone based on their physical characteristics. Multi-Layer perceptron classifier (MLPClassifier) is used to implement and train the neural network. To increase the accuracy of our neural network, training a neural network is needed which involves adjusting the sizes of the hidden layers, maximum iterations and setting up a suitable activation function. A webpage is designed in order to display the predicted value i.e., the age of an abalone.

Keywords: Abalone, Activation Function, Artificial Neural Network, Hidden Layers, MLP Classifier.

1. INTRODUCTION:

Abalone are a very common type of shellfish. Their flesh is considered to be a delicacy and their shells are popular in jewellery. They can live up to 50 years, depending on a species. The speed of their growth is primarily determined by environment factors related to water flow and wave activity. Due to differences in food availability, those from sheltered water typically grow more slowly than those from exposed reef areas. Estimating the age of abalone is therefore difficult mainly because their size depends not only on their age, but on the availability of food as well.

The cost of abalone depends on its age. However, the process of finding the age of an abalone is a complicated one.The rings in the inner shell of the abalone grows at the rate of one ring per year. Counting the number of rings involves cutting of shell. The lab technician counts the number of rings using the microscope, after polishing and staining.This heavy process increases the price of an abalone and its popularity.

Machine learning (ML) algorithms and Neural Networks which is a subset of ML became popular tools for making predictions using previous data. A dataset provided by the University of California Irvine Machine Learning Repository consists of physical characteristics of abalones and their ages. This study is a classification problem which aims to predict the age class of an abalones by considering their physical attributes.

1.1 Introduction to ANN:

http://doi.org/10.36893/JNAO.2023.V14I2.349-353

Artificial Neural Networks (ANN), which is a subset of Machine Learning, is a collection of artificial neurons which performs various computations and arranged in mainly three layers. They are one input layer which is used for data entry, one output layer which provides information to the user and the hidden layer which has the processed data. ANN gathers knowledge by identifying the patterns and relationships between the attributes and trained through experience, rather than from programming. Once our neural network is trained, it can be used to predict the outcome of a new set of input data. Larger training datasets results in better predictions.

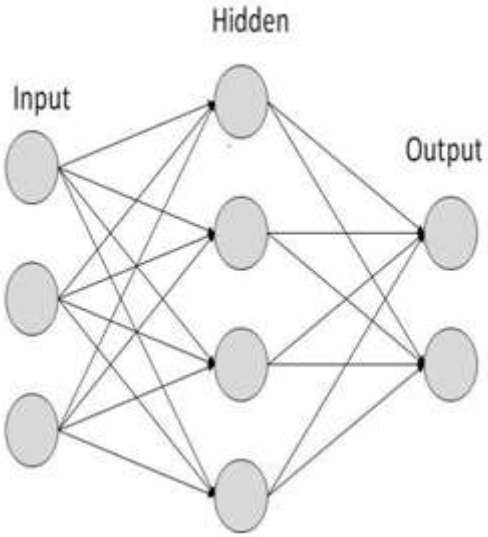


Fig. -1. Architecture of Artificial Neural Networks

The four advantages of artificial neural networks are organic learning, storing information, fault tolerance and self-repair.

- Organic Learning: The output of artificial neural networks is not only generated from the inputs and the results given, but they can also learn by themselves and then make the predictions. Organic learning involves generalizing the inputs.
- Storing Information: ANN stores the input data in its own networks instead of using external devices like database.
- Fault Tolerance: The corruption of neurons or nodes of ANN does not affect the network formed and it cannot prevent it from generating end result i.e., output. Hence ANN has high potential for fault tolerance.
- Self-Repair: Once the ANN is trained completely, it has an ability to work with the incomplete data i.e., it can generate output with missing information. Until and unless the missing of important data occurs, the network is not affected.

1. Methodology:

2.1 Data pre-processing:

The dataset is pre-processed to transform the problem of predicting the age to a classification problem.In this project, the considered dataset contains 8 physical characteristics of an abalone. They are sex, length, diameter, height, whole weight, shucked weight, viscera weight, shell weight and no. of rings. Among all the attributes, the target variable is 'Rings' of abalone, which is a discrete integer from 1 to 29. A new column ‘Age’ is added to the dataset. Age of an abalone is classified based on its rings count. Age column is given three values, they are 1, 2 and 3.

- 1 is for the rings in the range of 1 to 8
- 2 is for the rings in the range of 9 to 10

http://doi.org/10.36893/JNAO.2023.V14I2.349-353

- 3 is for the rings in the range of 11 to 29

Algorithm for adding new column Age:

Input: Rings column

Output: Returns Age class labels as 1, 2 and 3 based on the constraints.

Steps:

1. Start
2. Define a function for classifying labels for age column
2.1. Read the rings column
2.2. If rings are greater than or equal to 1 and less than or equal to 8 then
Return age class label as 1
2.3. If rings are greater than or equal to 9 and less than or equal to 10 then
Return age class label as 2
2.4. If rings are greater than or equal to 11 and less than or equal to 29 then
Return age class label as 3
2.5. Otherwise return 0
3. Apply above defined function to all the rows of a dataset using lambda function.
4. Stop

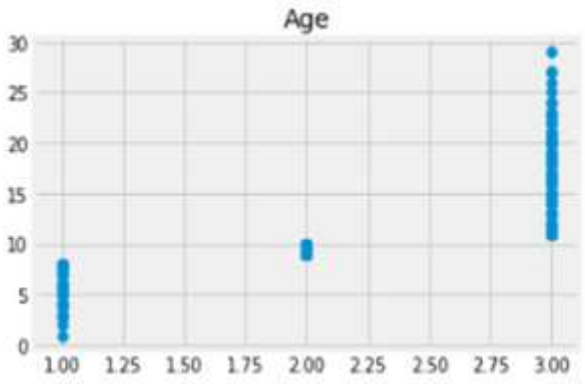


Fig. -2. Scatter Plot between Age and Rings

(x-axis indicates Age, y-axis indicates Rings)

2.2 Model building:

Multi-Layer Perceptron (MLP), a widely-used ANN architecture, is composed of a group of non-linear neurons (perceptron). MLP architecture is the best algorithm for classifying and predicting new observations. The neurons present in MLP are interlinked to each other in a multilayer structure. The input layer neurons divide the input signal among the neurons present in the hidden layer. Then, every neuron in the hidden layer iterates over multiple hidden layers and gathers knowledge. Finally, after all connections and iterations end output is generated and classified.

MLPClassifier algorithm contains many attributes like hidden layer sizes, max iterations, activation function and an optimizer.

- Hidden layer sizes: The hidden layer sizes attribute takes tuple of three numeric values which indicates the number of nodes present in input layer, hidden layer and an output layer. Among the tuple of three values, it is necessary to mention at least one value in the tuple, remaining two values are optional.

http://doi.org/10.36893/JNAO.2023.V14I2.349-353

Basically, the first two values in the tuple are the powers of 2 and the third value will be the size of output class.

hidden_larer_sizes = (64, 64, 4)

- Max iterations: Max iterations represents number of epochs. The number of epochs defines the number of times that the algorithm should work through the entire training dataset. For a good performance, algorithm should pass over the training data multiple times in order to get better accuracy. Thus, the value of this attribute is numeric.

max_iter = 98

- Activation function: Activation function decides, whether a neuron should be activated or not. Without activation function, the neural network acts as a simple linear regression model. Activation function performs non-linear transformations to the input and makes the network to perform complex tasks. There are many activation functions like linear, sigmoid, tanh, relu and softmax. Relu and softmax are used when the class labels are more than 2.

activation = 'relu'

- Optimizer: Optimizers are responsible for reducing the losses and generates accurate results. There are many optimizers like gradient descent, momentum, adadelta, adam. Adam is the best optimizers which trains the neural network in less time and more efficiently.

optimizer = 'adam'

Algorithm for model building:

1. Start
2. Read the data set
3. Pre-process the data and make age class labels
4. Split into training and testing data
5. Transform the training data using standard scalers.
6. Fit the MLPClassifier with suitable attributes values to the training data
7. Generate predictions for new data
8. Stop

2. Experiments &Results:

After splitting the dataset into training and testing data at the rate of 90 percent and 10 percent and applying the MLPCLassifier algorithm with hidden layer sizes as 64 and 98 epochsand with relu activation function, the maximum accuracy reached is 69.1387

```
#accuracy
print("Accuracy of Neural networks is: ")
print(accuracy_score(test_y, y_pred)*100)

Accuracy of Neural networks is:
69.13875598086125
```

Fig. -3. Accuracy of the model

Flask framework is used to deploy implemented algorithm and integrate it with the designed webpage. Hence the output i.e., the age class label either 1 or 2 or 3 is seen through the webpage. Webpage contains the eight input labels which takes the values of eight physical characteristics of abalone and a predict button which displays the age label.

http://doi.org/10.36893/JNAO.2023.V14I2.349-353

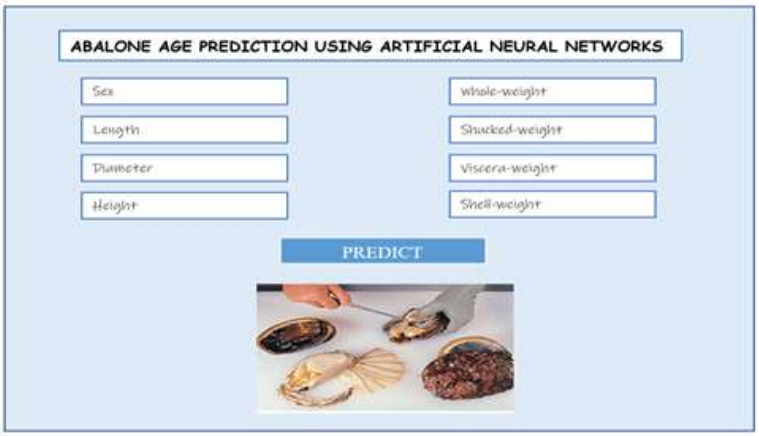


Fig. -4. Snippet of user interface

3. Conclusion:

Abalone age can be predicted based on its physical characteristics in the form of rings by applying regression algorithms and also in the form of age class labels by applying classifier algorithms. In this paper, classifier algorithm is used to predict the age of the abalone in the form of age class labels. This model generates more accuracy score compared to the regression models. Applying classifier algorithms involves building up artificial neurons and a dedicated neural networked which is trained with the dataset.

References:

[1] Khadija, Abalone Age Prediction using Artificial Neural Network, IOSR Journal of Computer Engineering, Vol. 18, pp. 34-38, September, 2016.
[2] Muhammad FaizMisman, Prediction of Abalone Age Using Regression-Based Neural Network, International Conference on Artificial Intelligence, 2019.
[3] Zhengjie Wang, Abalone Age Prediction Employing A Cascade Network Algorithm and Conditional Generative Adversarial Networks, Research School of Computer Science, Australian National University.
[4] Runze Guo, A New Method of Measuring the Age of Abalone Based on Data Visualization Analysis, Journal of Physics: Conference Series, 2020.

http://doi.org/10.36893/JNAO.2023.V14I2.349-353