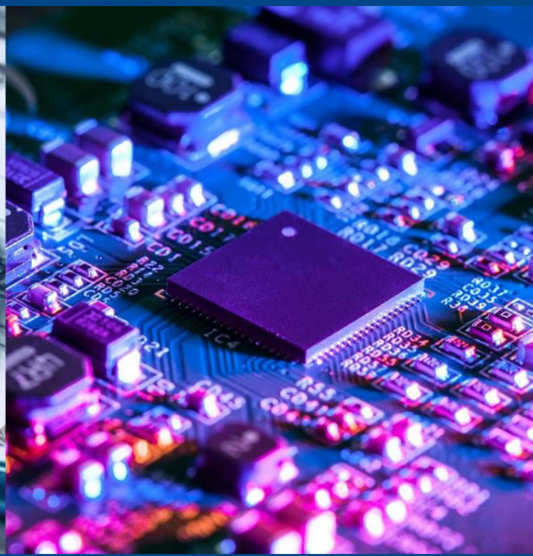
ISSN: 2582-7219

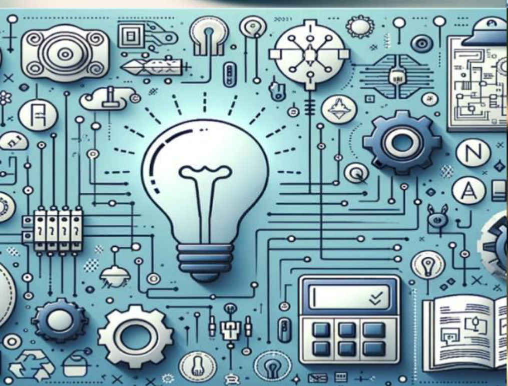




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 4, April 2025

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Abalone Age Estimation Using K-NN Regression

Shravani Kailas Raykar, S.P.Sarode

Dept. of Computer, JSPM Polytechnic, Pune, Maharashtra, India

Guide, Dept. of Computer, JSPM Polytechnic, Pune, Maharashtra, India

ABSTRACT: The research establishes and assesses a machine learning age estimation technique which utilizes abalone physical traits to determine their age. Laborious manual shell ring counting serves as the traditional method to determine abalone age while becoming both time-consuming and demanding human effort. The purpose of this study involves the utilization of K-Nearest Neighbors (K-NN) regression to determine abalone age through available physical measurements. The UCI Machine Learning Repository provides the dataset with shell length, diameter, height and sex attributes serving as input data for the K-NN model. The study reveals that machine learning technology provides companies with a quicker and more scalable replacement for existing conventional manual techniques. The performance evaluation of the model includes various metrics and demonstrates its integration into a friendly Streamlit web application which makes marine resource management data accessible to stakeholders

KEYWORDS: Abalone, K-Nearest Neighbors, Age Estimation, Regression, Machine Learning, Streamlit, UCI Repository, Physical Features, Marine Resource Management

I. INTRODUCTION

Marine mollusks belonging to Haliotis species maintain crucial biological and ecological research importance due to their commercial importance. The determination of abalone ages enables researchers to understand both their population dynamics and their health indicators along with their growth trends. Historically researchers have utilized shell examination to determine abalone age by measuring growth rings in the same way as tree ring analysis.

This technique delivers exact findings yet it needs extensive manual work and skilled personnel and specialized instruments. Machine learning regression algorithms have progressed enough to streamline age estimation processes through an efficient system that provides faster solutions. A research investigates how the K-Nearest Neighbors (K-NN) regression algorithm utilizes abalone physical measurements such as length, diameter, height, and sex to produce age estimations of these sea snails.

K-NN lending its strength to existing datasets provides approximate age forecasts that operate more quickly than standard manual counting methods thus serving as an attractive instrument for marine resource research. The research describes the method for K-NN training which entails data cleaning and selecting features along with evaluation measures. The web application based on Streamlit integrates the model to provide users with an interface for submitting abalone measurement data to get automatic age estimates.

II. LITERATURE SURVEY

The correct estimation of animal age plays an essential role in sustaining healthy abalone populations. The shell ring counting method for determining age requires each ring to represent yearly growth because this technique assumes one ring for one year. The counting of annual rings using shell analysis produces unreliable data because environmental factors alter growth while rings 1 become harder to distinguish from each other in aged abalones. Researcher have investigated new physical measurement methods that use shell length together with height and diameter measurements.
The application of machine learning approaches to ecological data analysis has become prevalent for conducting classification and regression activities. The estimation of abalone age through support vector machines (SVM) and random forests as well as artificial neural networks (ANNs) has resulted in divergent levels of success. The ability of these models to detect non-linear patterns between abalone characteristics and age is promising yet they require additional complexity while being relatively expensive to compute compared to K-NN.

K-NN stands as an excellent non-parametric algorithm suitable for regression applications since it functions effectively when the underlying relations between predictor variables and the target variable remain unknown. The method functions by combining the outcome values of k near neighbors to each data point to deliver predictions for continuous values such as abalone age.

III. THIS APP UTILIZES

Users can access the " Abalone Age Predictor" web application through which the K-Nearest Neighbors (K-NN) regression algorithm predicts the age of abalone by analyzing physical measurements. Streamlit enables the creation of user-friendly interactive web applications to handle the predictive model through its Python library interface. Through the web application users can submit their abalone features including shell length and diameter and height together with sex information to receive an estimated age in terms of growth ring predictions.

The application enables users to v"ew v'sual age-related abalone characteristics through useful graphical representations in addition to receiving performance statistics that detail the underlying model's operations. The application combines educational purposes with functional capabilities which meet the needs of marine biologists and fisheries managers.

IV. SYSTEM ANALYSIS

A framework includes two major elements which encompass the backend machine learning model together with the front-end Streamlit application. The K-NN backend model receives training using the UCI Abalone data while maintaining transformed data through normalization steps alongside categorical feature label encoding. Using cross-validation the training process chooses the most appropriate number of neighbors (k) in order to prevent both underfitting and overfitting.

Non-technical users can operate the Streamlit application successfully because it was developed with an intuitive user-friendly interface that accepts required physical measurements from end-users. The app gives users on-screen predictions about age with additional charts that display test results. The combination of machine learning technology with web application functionality creates better technical opportunities for marine biological stakeholders and resource management professionals.

V. IDEAL OVERVIEW

The exact application for the " Abalone Age Predictor" application should focus on marine biology research through abalone population management and research. Scientists need abalone age estimation as a basis to study their lifecycle patterns together with reproductive behavior and general health. The app enables marine scientists together with fisheries managers to perform speedy and precise abalone age estimations which boosts their efficiency when conducting population assessments and conservation work.

Installing this tool into marine research activities supports the development of data-oriented population management strategies that protect sustainable abalone populations. The app provides convenient accessibility and user-friendly experience which transforms it into an essential educational tool that enables students and enthusiasts to interact with authentic biological data.

VI. EASE OF USE

User experience plays a central role in the development of the Streamlit-based application interface. An easy-to-use interface presents form fields where users enter physical measurements of abalone such as shell length and diameter and height and sex. Following data entry the application uses processing power to yield an age prediction in ring counts.

The app presents two graphical displays including scatter plots and residual plots to help users view the relationship between feature inputs and age predictions. The application interfaces result data in a manner that makes them accessible to people with basic computational literacy and straightforward understanding abilities.

VII. CONCLUSION

The K-Nearest Neighbors (K-NN) regression method allows effective estimation of abalone age through physical measurement analysis which replaces conventional ring counting approaches. Through a developed web application users receive fast age predictions which require basic inputs from the model embedded into a user-friendly platform.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

More precise modeling can be achieved by utilizing advanced prediction models in combination with supplementary information sources although the K-NN solution provides satisfactory delivery of results.

The "Abalone Age Predictor" application serves as a practical instrument for marine biologists together with researchers and fisheries managers to assess abalone age through this faster and scalable process. The subsequent work should focus on improving the model through continual updates and broadening its features and investigating fresh methods to enhance the predictive capabilities by resolving minority class restrictions and developing advanced feature engineering methods.

ACKNOWLEDGMENTS

The Abalone dataset from the UCI Machine Learning Repository enabled our research through its provision. The developers of Streamlit framework supported the creation of user-friendly interactive web applications and we extend our appreciation. Our appreciation goes to the colleagues and mentors from marine biology and machine learning who offered indispensable feedback during this research.

REFERENCES

1. Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 46(3), 175-185.
2. Clark, D., Schreter, Z., & Adams, A. (1996). A Quantitative Comparison of Dystal and Backpropagation. Submitted to the Australian Conference on Neural Networks (ACNN'96).
3. Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.
4. Nash, W. J., Sellers, T. L., Talbot, S. R., Cawthorn, A. J., & Ford, W. B. (1994). The Population Biology of Abalone (Haliotis species) in Tasmania. I. Blacklip Abalone (H. rubra) from the North Coast and Islands of Bass Strait. Sea Fisheries Division, Technical Report No. 48 (ISSN 1034-3288).
5. Waugh, S. (1995). Extending and benchmarking Cascade-Correlation. PhD thesis, Computer Science Department, University of Tasmania.

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com