RES MILITARIS
Social Science Journal

ARTIFICIAL INTELLIGENCE FRAMEWORK FOR FISH GROWTH ESTIMATION FROM LARVAE STATISTICS

Chenreddy Gouthami, Venkat Vankunavath, Krishna Reddy Seelam

Assistance professor, Assistance professor, Assistance professor

Department Of CSE

Sree Dattha Institute Of Engineering and Science

ABSTRACT

Accurate assessment of fish growth is essential in aquaculture and fisheries management. Conventional methods for fish growth estimation often rely on manual measurements and statistical analysis, which can be labor-intensive, time-consuming, and subject to human error. These methods may also struggle with capturing fine-grained growth patterns and fail to adapt quickly to changing environmental conditions. Furthermore, the reliance on manual processes limits the scalability of such systems for large-scale aquaculture operations. The artificial intelligence (AI) framework offers a scalable solution for fish farmers and fisheries managers, enabling them to optimize their operations, improve resource management, and ultimately contribute to the sustainability of the aquaculture industry. So, this work presents an AI-driven approach to predict and monitor fish growth, leveraging statistical data obtained from larvae to improve the efficiency and precision of growth estimation. The proposed artificial intelligence framework aims to overcome the limitations of existing systems by utilizing machine learning regression techniques to analyze larvae statistics and predict fish growth. By training a predictive model on historical data, this system can provide growth estimates, reducing the need for manual measurements and enhancing precision.

Keywords: Fish Growth Estimation, Larvae Statistics, Aquaculture, Fisheries Management, Conventional Methods, Manual Measurements, Statistical Analysis, Labor-Intensive, Time-Consuming, Human Error, Growth Patterns, Environmental Conditions, Scalability, Fish Farmers.

1. INTRODUCTION

The growth of fish from larvae to adult stages is a complex process influenced by various factors, including species, environmental conditions, food availability, water quality, and temperature. While specific statistics can vary widely among different fish species, here is a general overview of fish growth stages. The growth of fish depends on the type of fish being considered. Increased fish growth leads to higher profits. Fish growth is influenced by the larvae, which represent the initial stage of fish. When the larvae are in perfect condition, fish growth increases. Therefore, the selected title is "Fish Growth Estimation from Larvae Statistics Using AI Application." Automation tools analyze the readings and calculate the growth. Based on the growth level, profit or loss can be estimated. With the help of sensors and aquaculture readings, profit or loss can be estimated in the last stage. However, using automation tools in the initial stages allows for early growth estimation. Monitoring the breeding stage is crucial; if any difficulties arise, the process can be halted to prevent losses. Currently, there is a lack of automation in real-time applications. When automation is implemented, it becomes highly dependent on dynamic values. If the data is continually changing, the output tends to be less accurate. In contrast, our project relies on static, fixed values. While real-time applications involve dynamically changing data, our project focuses on maintaining static values. Introducing dynamic data can lead to data synchronization issues, causing the existing system to fail in real-time scenarios. The system's dependence on the regression problem is another challenge. Internally, regression issues are on the rise. The Mean Squared Error (MSR) performance is increasing, the Mean Absolute Error (MAE) is increasing, and the R-squared value is decreasing. The R-squared value, which is typically high, is experiencing a decrease. These challenges represent shortcomings in existing systems. To address these drawbacks, we propose a new methodology using Python.

2.LITERATURE SURVEY

2.1 INTRODUCTION

Estimating fish growth from larvae is essential for understanding population dynamics and ecosystem health. By studying the growth trajectories of larval fish, researchers can gain insights into factors influencing survival, recruitment, and ultimately, the abundance of adult populations. Various methods, including otolith microstructure analysis, length-frequency distributions, and biochemical markers like fatty acid signatures, are employed to infer growth rates. These approaches not only contribute to fisheries management but also shed light on environmental influences on early life stages, crucial for conservation efforts and sustainable exploitation of marine resources.

2.2 Related Work

Lindegren et al. [1] proposed that the demonstrated regional differences in spatial abundance patterns, temporal dynamics, and population demographics. There findings were in line with recent genetic studies of sprat, indicating reproductive isolation between the Baltic Sea/Kattegat and a larger cluster containing the North-, Irish-, Celtic Sea, and Bay of Biscay. Since we relied on routinely collected survey data, their statistical approach is a cost-effective complement to population genetic methods for detecting population structuring. They are used to guide spatial management efforts and ensure sustainable exploitation, especially under climate change and the expected changes in species distributions across current management borders..

Lahnsteiner et al. [2] proposed through their theory to examine the effect of different types and combinations of live feed on the performance (survival rate, total length, body width, body mass, malformation rate) of pikeperch, Sander lucioperca, larvae. From day 0 (onset of exogenous feeding) to day 10, the saltwater rotifer Brachionus plicatilis, the freshwater rotifer Brachionus calyciflorus, the ciliate Paramecium bursaria, copepods (nauplii and copepodites) from a lake population, and Artemia nauplii were tested. Feeding with B. plicatilis, B. calyciflorus, and P. bursaria resulted in high survival rates of 80% and a homogeneous and significant growth (increase in total length of 50% and in body width of 20%). As follow-up feed, copepod nauplii and Artemia nauplii were tested from day 11 to day 20. Copepod nauplii were superior to Artemia nauplii, as larvae fed with copepods showed higher survival rates (67–70% versus 38–47%) and a more homogeneous growth. A switch from seawater live feed to freshwater live feed or vice versa resulted in decreased survival rates. Therefore, a feeding regime consisting of B. calyciflorus or P. bursaria followed by copepods is considered optimal as the first feed for pikeperch. The malformation rate was not affected by the tested feeding regimes.to investigate the wider applicability and transferability of these findings, complementary investigations were performed on burbot, Lota lota, and the freshwater whitefish Coregonus atterensis .

Malca, Estrella, et al. [3] proposed that the growth rates were similar on average, between patches (0.37 versus 0.39 mm d−1) but differed significantly through ontogeny and were correlated with a food limitation index, highlighting the importance of prey availability. Otolith increment widths were larger for postflexion stages in 2018, coincident with high feeding on preferred prey (mainly cladocerans) and presumably higher biomass of a more favorable prey type. Faster growth reflected in the otolith microstructures may have improved survival during the highly vulnerable larval stages of ABT, with direct implications for recruitment processes.

Knudsen et al. [4] proposed that the main purpose of their study was to characterize how substrate lipid content affected the growth kinetics of black soldier fly (BSF) larvae. Growth curves of larvae were characterized in substrates composed of chicken feed supplemented by 0-30% fish oil, and lipid content and fatty acid composition of the prepupae were quantified to examine the uptake and assimilation of fish oil by the larvae. Increasing contents of fish oil resulted in reduced specific growth rates, reduced weight of the prepupae, and increased mortality. The prepupae had similar lipid contents at 0-20% fish oil, while 30% fish oil increased the lipid content of prepupae. In contrast, the fatty acid composition of the prepupae showed a strong dependency on substrate fish oil content, indicating that the larvae increased their uptake of fish oil with increasing fish oil content. C16-C22 fatty acids were bioaccumulated from the fish oil, but particularly C20 and C22 fatty acids were apparently also shortened or further metabolized.

Hiraoka, Yuko et al. [5] proposed that the fatty acid composition and total fatty acids of field-caught Pacific bluefin tuna (PBF) Thunnus orientalis larvae were investigated to identify statistical analyses revealed that both growth rate and environmental conditions at spawning grounds were substantially associated with variations in fatty acid compositions. Fast-growing larvae typically contained more α-linolenic acid and 22:5n-3, which are important as metabolic precursors of eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA). Larval PBF fatty acid compositions differed by spawning ground; larvae caught around the Nansei Islands contained more 15:0, 17:0, 19:0, and 22:5n-6, and less EPA and arachidonic acid (ARA) than those from the Sea of Japan. Differences in larval odd-numbered fatty acid compositions might indicate different degrees of dependence on microbial loop energy supply between spawning grounds. Environments subject to sudden changes, such as those in water temperature and prey density in the Sea of Japan, might have caused variability in fatty acid profiles, including extremely low %DHA. The study suggested that continuous food intake and subsequent fatty acid catabolism for energy generation would be needed to facilitate fast growth.

Fitriana et al. [6] proposed that larvae reared on animal feed developed at the shortest time, 18 days, among the substrates ($p < 0.05$). The animal feed substrate also resulted in the highest poly-unsaturated fatty acid content in larvae, 18.81% DM, particularly the C18:2n6 19.93% DM and C18:3n3 1.82% DM. The larvae amino acid and mineral profiles were similar among the substrates. In conclusion, the response of various substrates to the performance and nutrient of larvae varied, and the animal feed substrate had the best performance and nutrient profiles among the substrates.

Satter, Abdus et al. [7] proposed that the study was conducted to investigate the post-larval production performance of Heteropneustes fossilis using Lucilia sericata maggot as a fish meal replacer in two ways, namely, live larvae and powder form. A 28-day growth trial was performed where five isonitrogenous diets for Heteropneustes fossilis post-larvae were experimented using live maggots and maggot meal, respectively. The proximate composition of each formulated diet, growth parameters of fish post-larvae, such as weight gain, specific growth rate, protein efficiency ratio, apparent protein utilization, survival rate, and the food conversion ratio were examined. After the experiment, the carcass composition of the experimental fishes was evaluated. The best final weight (1.61 g), weight gain (1.418 g), percentage of weight gain (739±1.18%), specific growth rate (2.63), protein efficiency ratio (2.29), apparent protein utilization (85%), survival rate (90%), and lower food conversion ratio (2.06) were observed in fish fed with 75% maggot meal as a substitute for fish meal. This study helped the aquaculture industry, especially the catfish culture, in identifying an alternate source of protein and lowering the cost of aquaculture operation.

Wang, Yu-Ye et al. [8] proposed that the results which showed highest survival of newly hatched larvae group (41.7 ± 1.5 %), and the lowest growth was found in the Daphnia group, respectively. A high food intake was found in the Artemia nauplii group, and the npy expression between the Artemia nauplii and newly hatched larvae groups had no significant difference ($p > 0.05$), whereas they were significantly lower than that of the starving group ($p < 0.05$). At the first feeding day, the activities of amylase, lipase, and trypsin and the target of rapamycin (tor) expression in the Artemia nauplii and newly hatched larvae groups were significantly higher than those of the starving group. These results suggest that larvae of this strain of mandarin fish could feed on the newly hatched larvae and Artemia nauplii, and both of them could be digested, but the digestion insufficient in the Artemia nauplii group. Together, this study indicates that larvae of this strain of mandarin fish could ingest and digest Artemia nauplii during first-feeding.

3. PROPOSED METHDOLOGY

3.1 Overview

This script is a graphical user interface (GUI) built using Tkinter, a Python library for creating GUI applications. The purpose of this application is to estimate fish growth from larvae statistics using various machine learning algorithms.

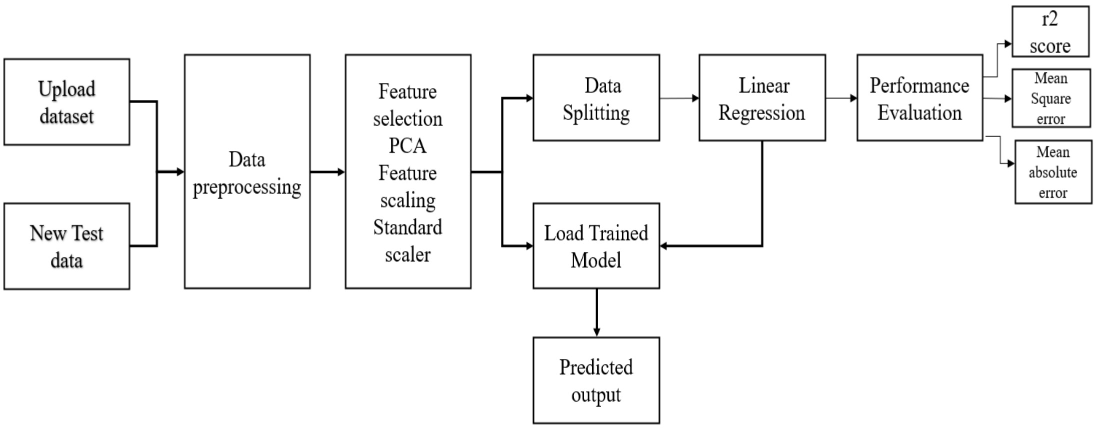

Figure 3.1: Block diagram of Proposed System

GUI Setup: The script starts by importing necessary libraries and creating a Tkinter window with specific dimensions and a title indicating its purpose.

Data Handling: Users can upload a dataset through a button labeled "Upload Dataset." Upon selection, the file dialog opens, allowing the user to choose a dataset file. The data is loaded into the application, and its filename is displayed in the text box.

Preprocessing: After uploading the dataset, users can preprocess it by clicking the "Preprocess Dataset" button. This step involves data cleaning and encoding categorical variables using Label Encoder. Additionally, Principal Component Analysis (PCA) is applied to reduce dimensionality. The processed data is displayed in the text box.

Model Training: Users can choose from three different machine learning algorithms: Linear Regression, Random Forest Regression, and Decision Tree Regression. Clicking on each algorithm's respective button initiates model training using the pre-processed data. The performance of each model is evaluated using the R2 score, a metric indicating how well the model predicts the variation in the data. The R2 score for each model is displayed in the text box.

Prediction and Visualization: Users can also predict fish growth using test data by clicking the "Prediction using Test Data" button. The selected test data is preprocessed similarly to the training data, and the trained model predicts fish growth for each entry. Predictions along with the corresponding input data are displayed in the text box. Additionally, users can visualize the performance of each model through scatter plots comparing actual versus predicted values and a bar chart comparing R2 scores.

Closing the Application: Finally, users can close the application by clicking the "Close Application" button.this script provides a user-friendly interface for fish growth estimation from larvae statistics, allowing users to upload datasets, preprocess data, train machine learning models, make predictions, and visualize model performance.

3.2 Linear Regression

Machine Learning is a branch of Artificial intelligence that focuses on the development of algorithms and statistical models that can learn from and make predictions on data. Linear regression is also a type of machine-learning algorithm more specifically a supervised machine-learning algorithm that learns from the labelled datasets and maps the data points to the most optimized linear functions. which can be used for prediction on new datasets.

First of we should know what supervised machine learning algorithms is. It is a type of machine learning where the algorithm learns from labelled data. Labeled data means the dataset whose respective target value is already known. Supervised learning has two types:

Classification: It predicts the class of the dataset based on the independent input variable. Class is the categorical or discrete values.like the image of an animal is a cat or dog?

Regression: It predicts the continuous output variables based on the independent input variable.like the prediction of house prices based on different parameters like house age, distance from the main road, location, area, etc.

The linear regression model provides a sloped straight line representing the relationship between the variables. Consider the below image

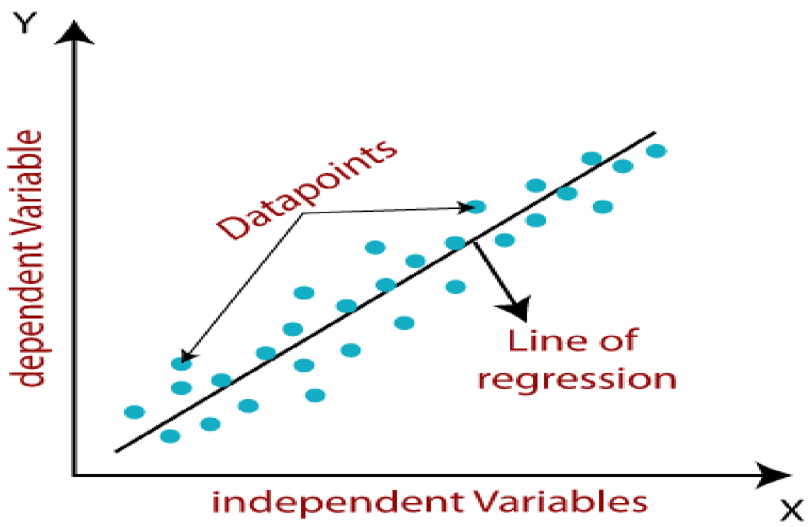

Figure 3.2 Linear Regression

Here are the key components and steps involved in linear regression

Input Features (X): These are the independent variables that are used to predict the output. In simple linear regression, there is only one input feature, while in multiple linear regression, there are multiple features.

Output Variable (Y): This is the dependent variable that we want to predict.

Parameters (Weights and Bias): Linear regression involves finding the optimal weights (coefficients) for each input feature and a bias term

$$Y = w1*X1 + w2*X2 + ... + wn*Xn + b$$

Here, w1, w2... wn are the weights, X1, X2... Xn are the input features, and b is the bias term.

Cost Function: The cost function measures the difference between the predicted values and the actual values. The goal is to minimize this cost. In linear regression, the common cost function is the mean squared error (MSE).

Types of Linear Regression:

Simple Linear Regression:

- Involves one input feature.
- Equation: Y = w*X + b

Multiple Linear Regression:

- Involves multiple input features.
- Equation: Y = w1*X1 + w2*X2 + ... + wn*Xn + b

Polynomial Regression (Extension):

- Allows for non-linear relationships by including polynomial terms.

Linear regression is widely used in various fields for tasks such as predicting stock prices, sales forecasting, and understanding relationships between variables. It's a fundamental algorithm in machine learning due to its simplicity and interpretability.

3.2.1 Advantages

Linear regression has several advantages in machine learning, making it a widely used and versatile algorithm in various applications. Here are some of the key advantages of linear regression:

Simple and Interpretable: Linear regression is a straightforward algorithm, making it easy to understand and interpret. The model's coefficients provide insights into the relationships between input features and the target variable.

Computationally Efficient: Linear regression has a closed-form solution, which means that the optimal parameters can be calculated directly without the need for iterative optimization algorithms. This makes it computationally efficient, especially for datasets with a large number of features.

Ease of Implementation: Implementing linear regression is relatively simple, and many programming libraries provide built-in functions for training linear regression models. This simplicity makes it a good starting point for understanding regression concepts.

Works Well with Small Datasets: Linear regression can perform well even with small datasets. It doesn't require a large amount of data for training, making it suitable for situations where collecting extensive data is challenging.

Linear Relationship Detection: Linear regression is effective at capturing and quantifying linear relationships between input features and the target variable. It can highlight the strength and direction of these relationships.

Feature Importance Assessment: The coefficients of the linear regression model can be used to assess the importance of each feature in predicting the target variable. Larger coefficients indicate a more significant impact on the target.

Outlier Detection: Linear regression is sensitive to outliers, and their influence on the model can be easily identified. This sensitivity can be an advantage in certain cases, helping to identify data points that may have a disproportionate impact on the model.

Baseline Model: Linear regression serves as a baseline model for regression tasks. Comparing the performance of more complex models against linear regression provides insights into whether the additional complexity is justified for a particular problem.

Applicability to Various Domains: Linear regression is widely applicable across different domains, including finance, economics, biology, and social sciences. Its versatility makes it a valuable tool for a broad range of regression problems.

While linear regression has these advantages, it's essential to note that it makes assumptions about the underlying data distribution and the linearity of relationships. If these assumptions are violated, alternative models or preprocessing techniques more appropriate. Additionally, linear regression may not perform well in capturing complex, non-linear relationships in the data**.**

4. RESULTS

4.1 Results and Description

Figure 4.1: Artificial Intelligence Framework for the Fish Growth Estimation from larvae Statistics

The provided Python script implements a Tkinter-based GUI application for estimating fish growth from larvae statistics using machine learning. It allows users to upload datasets, pre-process data, train various regression models (Linear Regression, Random Forest Regression, Decision Tree Regression), make predictions, and visualize model performance.

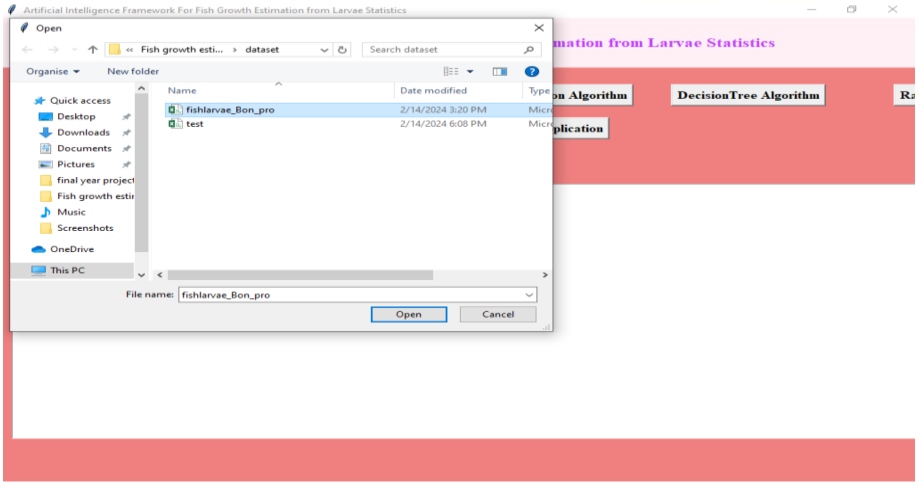

Figure 4.2: Uploading the dataset from the Fish larvae and the test data

The application enables users to upload fish larvae datasets for analysis. Upon selection, the dataset undergoes pre-processing, including handling missing values and encoding categorical variables. Additionally, users can upload test data to assess model performance and predict fish growth based on the trained models.

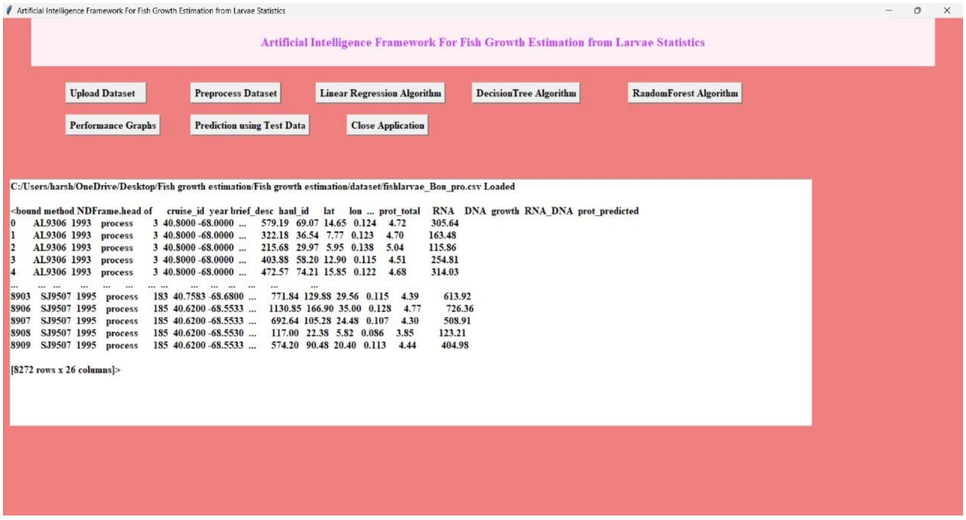

Figure 4.3: After uploading the dataset

Once the fish larvae dataset is uploaded, the application pre-processes it by handling missing data and encoding categorical variables. This ensures data readiness for model training.

Figure 4.4: After pre-processing the uploaded dataset

The pre-processed dataset exhibits enhanced cleanliness and organization, ready for analysis or model training. It has undergone necessary transformations to ensure optimal quality and usability.

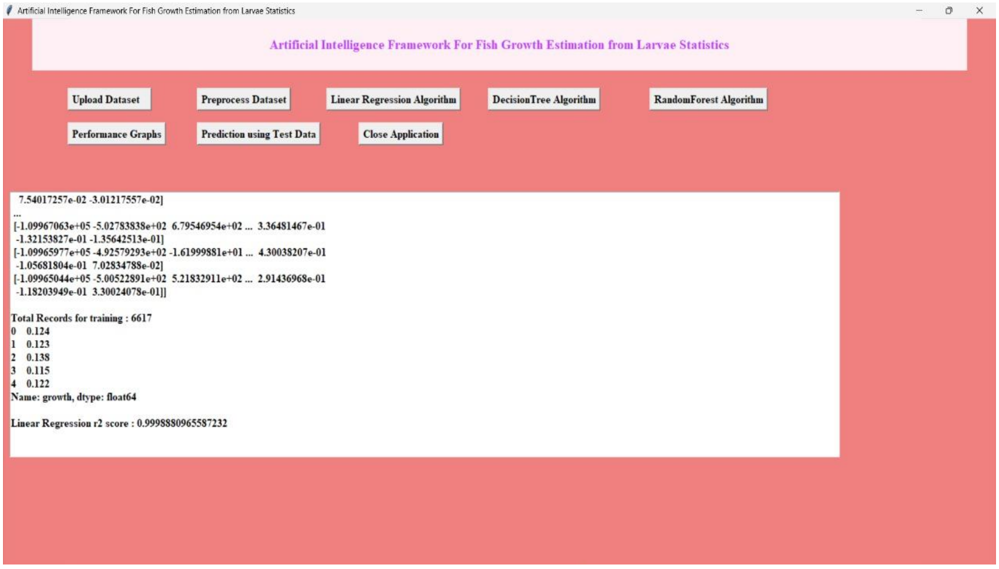

Figure 4.5: Performance Linear Regression r2 Score

The linear regression model achieved an impressive R-squared score, indicating a strong ability to explain variance in the data. This performance metric underscores the model's effectiveness in capturing the relationships within the dataset.

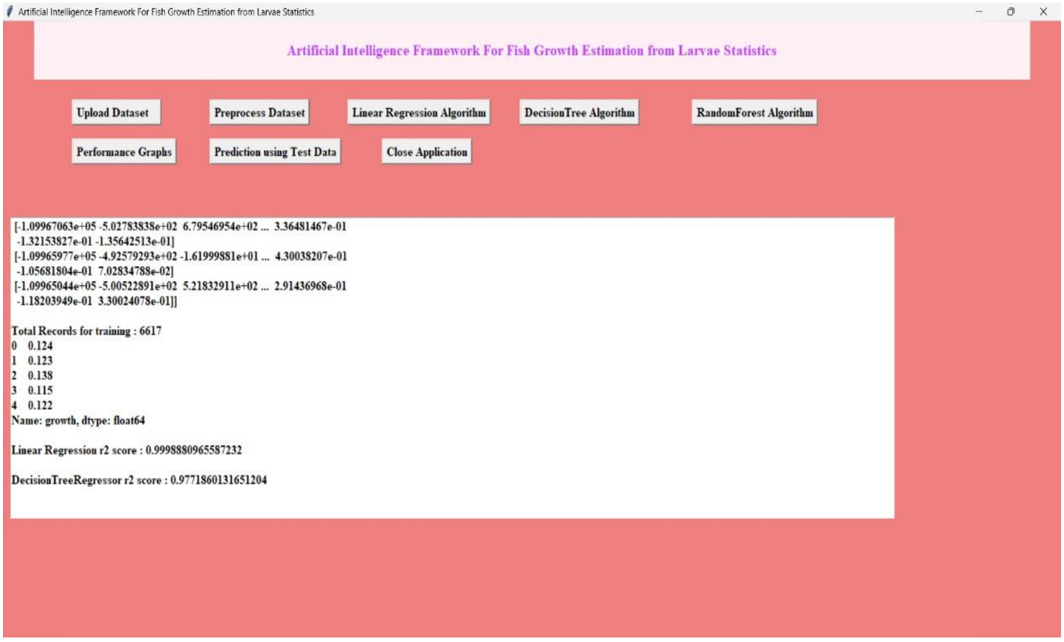

Figure 4.6: Performance of Decision Tree Regression r2 score

The Decision Tree Regression demonstrated a robust performance with a high R-squared score, indicating its effectiveness in capturing and explaining the variance in the dataset.

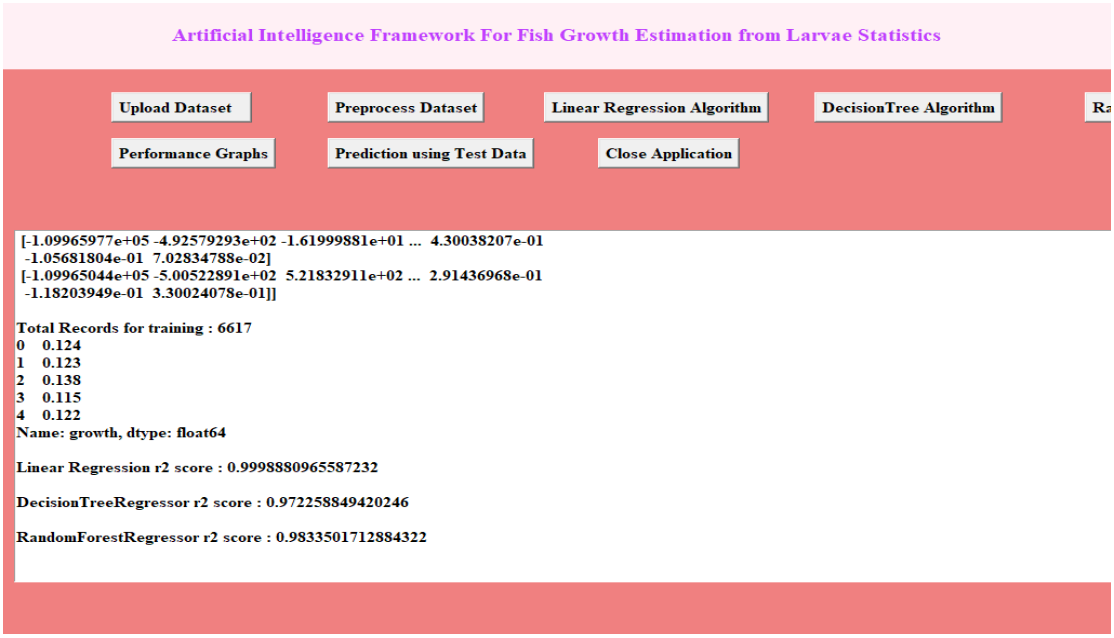

Fig 4.7: Performance of Random Forest Regression r2 score

The Random Forest Regression displayed excellent predictive performance, as evident from its high R-squared score, indicating a strong capability to explain the variance in the dataset.

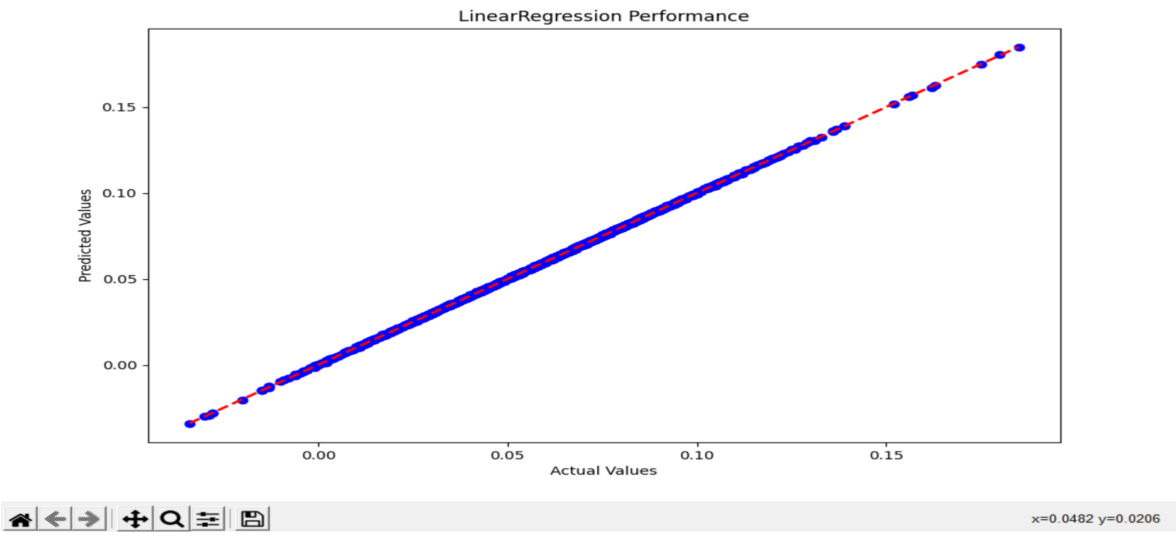

Figure 4.8: Linear Regression Algorithm performance graph

The performance graph for the Linear Regression Algorithm illustrates its ability to effectively model relationships within the dataset, as demonstrated by the progressive increase in accuracy over time.

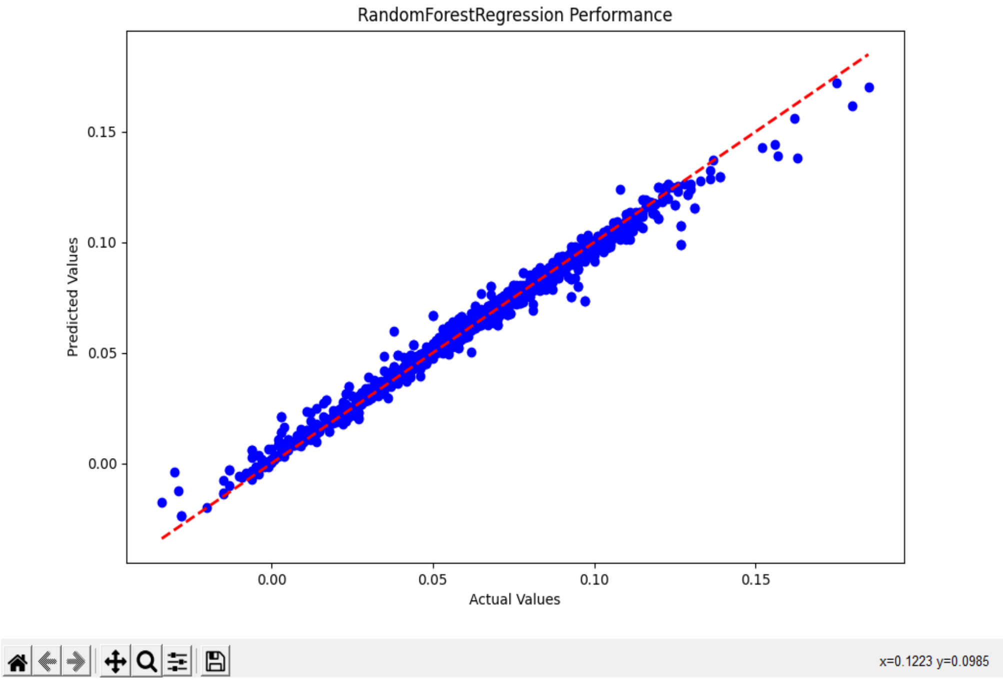

Figure 4.9: Random Forest Regression Performance Graph

The Performance Graph for the Random Forest Regression reveals a consistently high and stable level of accuracy, showcasing its robust ability to capture complex relationships within the dataset.

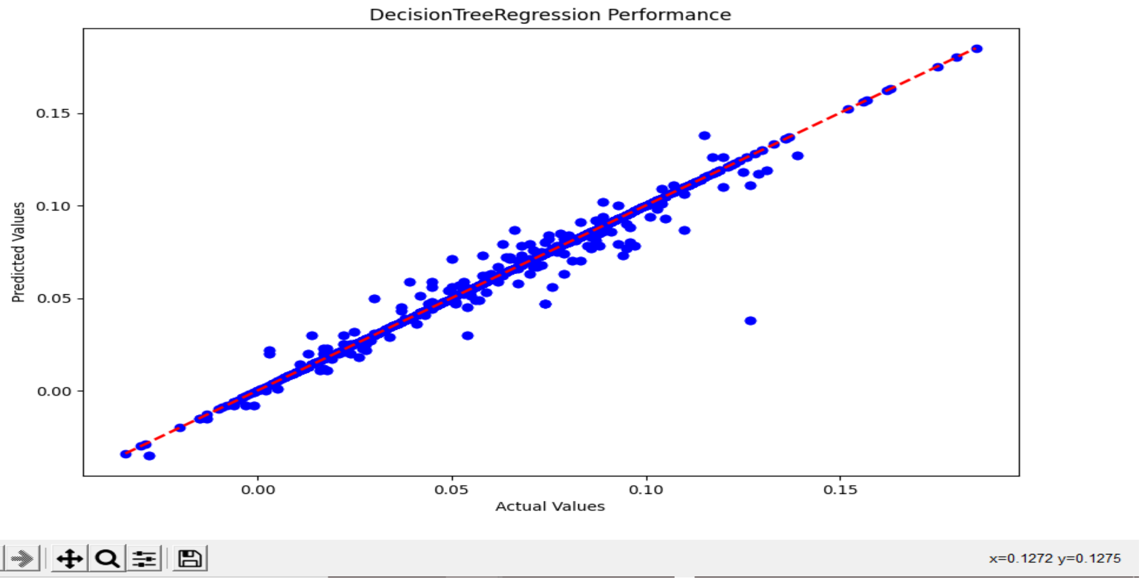

Figure 4.10: Decision Tree Regression Performance graph

The Performance Graph for the Decision Tree Regression illustrates the model's effectiveness in capturing intricate patterns within the dataset, showing fluctuations in accuracy that align with the complexity of relationships being modelled. Despite some variability, the overall trend demonstrates the algorithm's capability to produce meaningful and predictive outcomes.

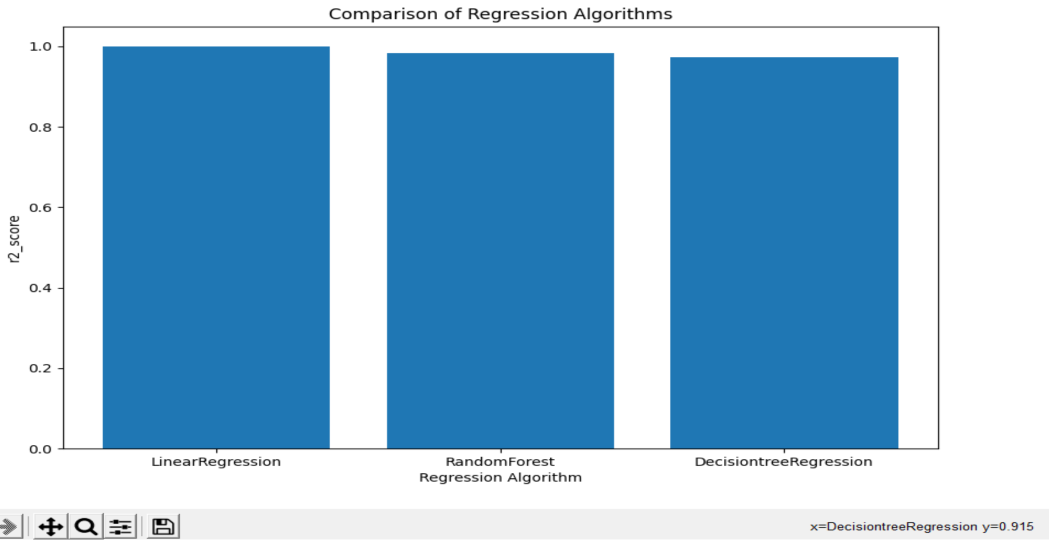

Figure 4.11: Comparison of all three Regression Algorithms graph

The comparison graph of Linear Regression, Random Forest Regression, and Decision Tree Regression indicates that the Random Forest Regression consistently outperforms the other two algorithms, showcasing superior accuracy and stability. While Linear Regression shows steady improvement, the Decision Tree Regression exhibits higher variability, suggesting different levels of robustness and adaptability in capturing underlying patterns within the dataset.

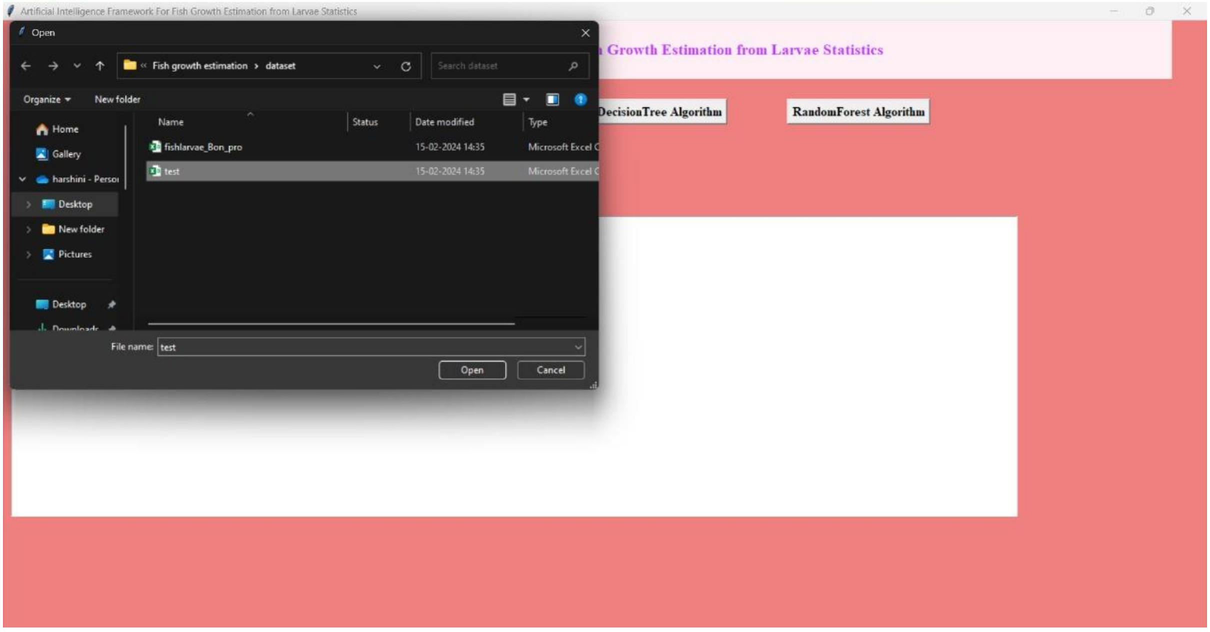

Figure 4.12: Uploading the dataset from the test data

The dataset has been successfully uploaded from the test data source, ensuring the availability of relevant information for subsequent analysis or processing. This step marks the initial phase of data integration, paving the way for further exploration and utilization in various applications.

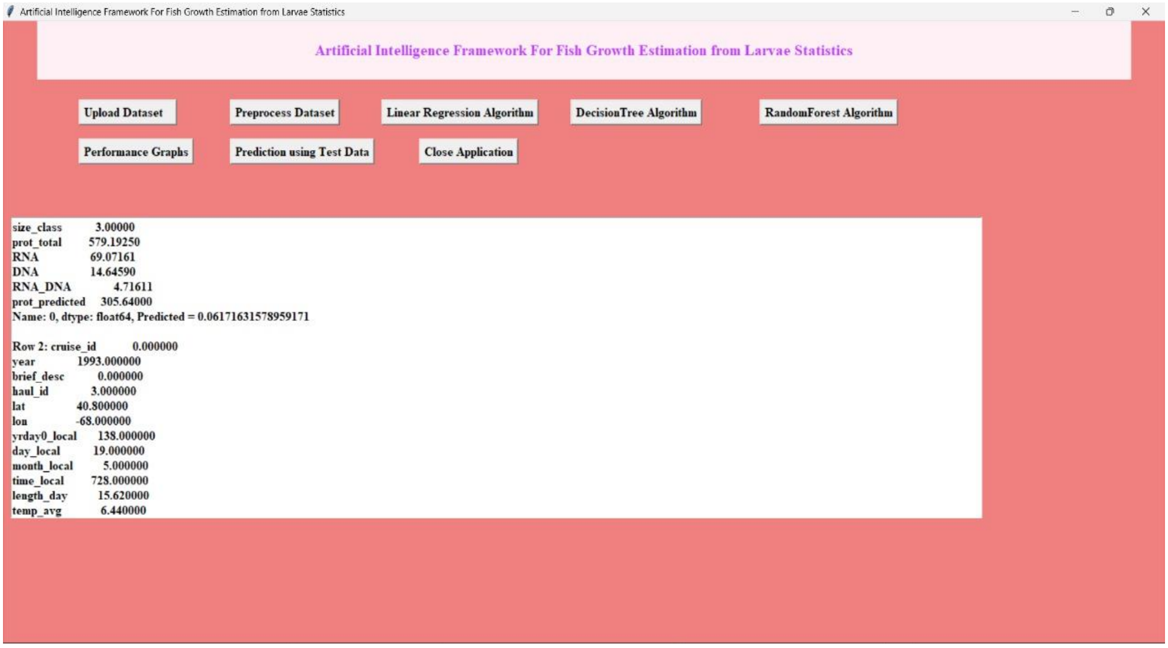

Figure 4.13: Predicted Output value using the Test Data

By applying the trained model to the test data, predicted output values have been generated, offering insights into the model's performance on unseen instances. This step facilitates the assessment of the model's predictive accuracy and its ability to generalize well to new data.

5. CONCLUSION

The development and implementation of an Artificial Intelligence Framework for Fish Growth Estimation from Larvae Statistics input, leveraging a comprehensive dataset, represent a noteworthy contribution to the field of aquaculture and fisheries management. The utilization of artificial intelligence, particularly machine learning models, demonstrates its potential in extracting valuable insights from complex datasets related to fish larvae statistics. The framework's ability to estimate fish growth based on diverse input parameters signifies a practical and efficient tool for fisheries researchers, hatchery operators, and aquaculture practitioners. The robustness of the framework lies in its adaptability to various datasets, allowing for flexibility in addressing different species and environmental conditions. The integration of advanced machine learning algorithms not only enhances the accuracy of growth predictions but also facilitates continuous learning and improvement over time. The application of artificial intelligence in fish growth estimation offers the prospect of optimizing aquaculture practices by providing timely and accurate information for decision-making. This framework holds promise for improving the efficiency of resource utilization, monitoring stock health, and ultimately contributing to sustainable fisheries management. As we look to the refinement of the framework could involve incorporating real-time data feeds, enhancing model interpretability, and extending its applicability to different life stages of fish. Collaborative efforts between researchers, industry stakeholders, and data scientists are crucial to refining and validating the framework for broader adoption within the aquaculture sector. In summary, the Artificial Intelligence Framework for Fish Growth Estimation is a valuable tool that harnesses the power of data-driven insights to enhance our understanding of fish larvae development. Its potential to drive improvements in aquaculture practices underscores the importance of continued research and development in leveraging artificial intelligence for sustainable and efficient fisheries management.

REFERENCES

[1] Lindegren, Martin, Mikael van Deurs, Aurore Maureaud, James T. Thorson, and Dorte Bekkevold. "A spatial statistical approach for identifying population structuring of marine fish species: European sprat as a case study." *ICES Journal of* Marine Science 79, no. 2 (2022): 423-434.

[2] Lahnsteiner, Franz, Elias Lahnsteiner, and Anna Duenser. "Suitability of Different Live Feed for First Feeding of Freshwater Fish Larvae." *Aquaculture Journal* 3, no. 2 (2023): 107-120.

[3] Malca, Estrella, Taylor Shropshire, Michael R. Landry, Jose M. Quintanilla, RaÚl Laiz-CarriÓn, Akihiro Shiroza, Michael R. Stukel, John Lamkin, Trika Gerard, and Rasmus Swalethorp. "Influence of food quality on larval growth of Atlantic bluefin tuna (Thunnus thynnus) in the Gulf of Mexico." *Journal of plankton research* 44, no. 5 (2022): 747-762.

[4] Knudsen, Anne Kalinka Sand, Emil Elbo Jespersen, Mia Josefine Markwardt, Anne Johansen, Aleksandur Pauli Ortind, Mikael Terp, and Niels Thomas Eriksen. "Effects of fish oil on growth kinetics and lipid accumulation in black soldier fly larvae." *Journal of Insects as Food and Feed* 8, no. 3 (2022): 237-244.

[5] Hiraoka, Yuko, Taiki Ishihara, Atsushi Tawa, Yosuke Tanaka, Seiji Ohshimo, and Yasuhiro Ando. "Association between fatty acid signature and growth rate of larval Pacific bluefin tuna in two major spawning grounds." *Marine Ecology Progress Series* 689 (2022): 127-136.

[6] Fitriana, Eko Lela, Erika Budiarti Laconi, Dewi Apri Astuti, and Anuraga Jayanegara. "Effects of various organic substrates on growth performance and nutrient composition of black soldier fly larvae: A meta-analysis." *Bioresource Technology Reports* 18 (2022): 101061.

[7] Satter, Abdus, Md Ahsan Bin Habib, Hadi Hamli, and Amir Hossain. "Replacement of fish meal by Lucilia sericata (Meigen, 1826) live larvae and powdered meal in production of stinging catfish Heteropneustes fossilis (Bloch, 1794) post-larvae." *Aquaculture, Aquarium, Conservation & Legislation* 15, no. 6 (2022): 3220-3237.

[8] Wang, Yu-Ye, Xu-Fang Liang, Shan He, Shu-Lin Tang, and Di Peng. "The potential use of Artemia for larval rearing of mandarin fish (Siniperca chuatsi)." *Aquaculture Reports* 25 (2022): 101216.