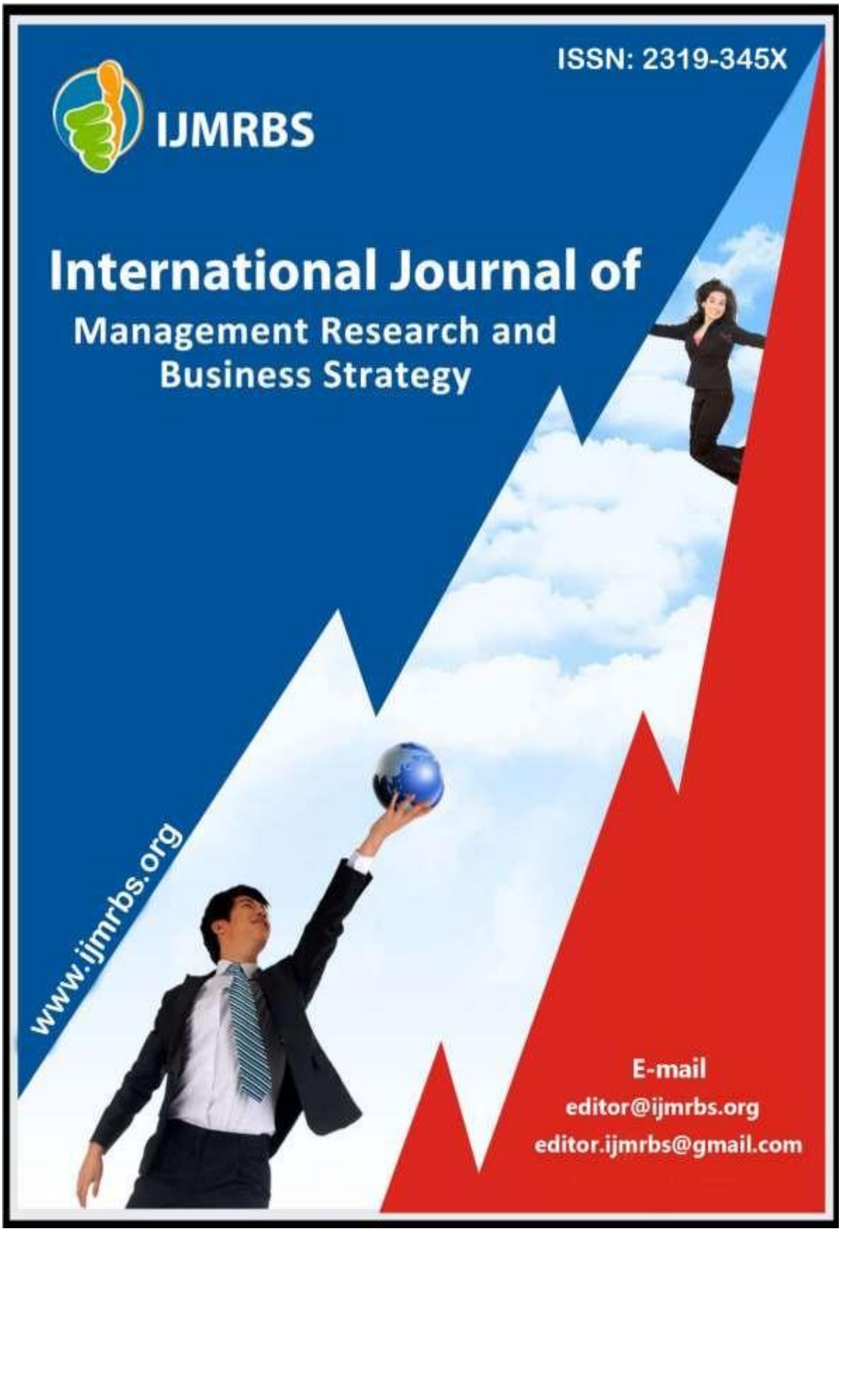

ISSN: 2319-345X
IJMRBS
International Journal of
Management Research and
Business Strategy
www.ijmrbs.org
E-mail
editor@ijmrbs.org
editor.ijmrbs@gmail.com

International Journal of Management Research and Business Strategy

ISSN 2319-345X

www.ijmrbs.com

Vol. 13, Issuse. 1, Mar 2024

EARLY DETECTION OF FISH DISEASES BY ANALYSING QUALITY OF WATER USING MACHINE LEARNING ALGORITHM

MOTAPOTULA NAGA SAI LIKHITHA*1, K. NAVEED KUMAR REDDY*2, YERUVA SIREESHA*3, PALANI DINESH*4, *5

* 1,3,4,5 B. Tech Students, *2 Assistant Professor
Dept. of Computer Science and Engineering,
RISE Krishna Sai Prakasam Group of Institutions

ABSTRACT:

Diseases that affect fish in aquaculture pose a substantial threat to the industry's ability to provide enough nutrition. The lack of sufficient infrastructure makes it difficult to identify diseased fish in aquaculture at an early stage, making it difficult to figure out whether or not fish have been infected. A necessary step in the fight against the further spread of illness is the prompt and accurate identification of diseased fish. Because salmon aquaculture is the world's fastest-growing system for the production of food, and because salmon aquaculture accounts for 70 percent of the market (2.5 million tonnes), the goal of this investigation is to discover the illness that affects salmon fish in aquaculture. We are able to detect the diseased fishes caused by a variety of pathogens thanks to a synergistic partnership between faultless image processing and machine learning method. This work is separated into two distinct parts. In the preliminary stage, picture pre-processing and segmentation have been used to, respectively, lessen the amount of noise and amplify the appearance of the image. In the second part of our analysis, we use an approach for machine learning called Support Vector Machine (SVM) with a kernel function to categorise illnesses and then extract the characteristics that are involved in each disease. The processed pictures from the first section were run through this support vector machine (SVM) model. After that, we harmonise a full experiment using the suggested combination of approaches on the salmon fish picture dataset that was used to investigate the fish illness. This work has been presented on a new dataset that includes both cases of image augmentation and cases where it was not used. The findings have led us to the conclusion that the SVM we have been using operates well, achieving an accuracy of 91.42 and 94.12 percent, respectively, with and without augmentation.

I. INTRODUCTION

1.1 BRIEF INFORMATION

A multi-sensor water quality monitoring system was used in order to carry out the monitoring of the Chemical Oxygen Demand (COD), Total Suspended Solids (TSS), and Oil & Grease (O&G). The use of reduced weights and the construction of models based on weighted variables were utilised in an enhanced boosting strategy for the purpose of suppressing quality-relevant factors. Predictions of wastewater quality were made using these models. The observation framework was tested out in the field with successful results, which showed the feasibility of this approach to assess the water quality online. In 2014, a second research was conducted utilising methods of machine learning to determine the level of pollution in the ocean off the coast of South Korea. Between 2011 and 2012, 63 samples were gathered by them. The analysis of the Geostationary Ocean Colour Imager (GOCI) photos was used in the research, and it was effective in discussing the distribution of water quality metrics. In 2012, a research project led by S. Shah was carried out in twenty distinct sites throughout the Indian state of Kerala. According to the findings of the research, the majority of the water samples collected in the area were appropriate for agricultural use. In order to make the water suitable for the marine life that lived in it, a fairly simple pretreatment was required. N. Karlar conducted research on the water samples he obtained from ten different communities. He examined physicochemical factors such as temperature, conductivity, pH, ionic strength (Cl-), total dissolving solids (TDS), alkalinity, Ca^{2+}, and Mg^{2+}. Because the WQI of these samples varied from 40.67 to 69.59, it was clear that the water taken from the block surface needed to be treated before it could be used. Usha conducted study in the year 2013 to determine the water quality rating and health of urban water bodies in the town of Bilari. During the first three months of 2011, water samples were collected from 10 different places around the country. Several physico-synthetic characteristics of surface water were investigated over

International Journal of Management Research and Business Strategy

ISSN 2319-345X

www.ijmrbs.com

Vol. 13, Issuse. 1, Mar 2024

the course of this study. According to the WQI, the level of pollution in the water was becoming worse every day. A. B. Frontier engaged in the process of doing his study. In order to conduct an investigation on the physicochemical characteristics of the water, samples were collected from a variety of communities located within the Nasik district of the Kalwan Tahsil. A random sample of water was taken from each of the five different places.

Techniques for machine learning, such as Gradient boosting, use decision-tree based in-built hierarchical structures that are based on regression algorithms to categorise difficult situations and assist in the automation of decision-making [10,11]. In this work, we used the approach to forecast fish infections by solving classification problems based on actual datasets that were comprised of required criteria of water purity. These achievements of such a technique in a variety of difficult conditions inspired us to use the technique, and we were motivated to do so by the successes of such a technique in those situations. Our strategy for developing an automated fish illness detection algorithm and using it in the decision-making process may be summed up as the following series of phases, in the order in which they are presented:

Step 1: Take a sample of water to identify the water quality.

Step 2: Predicting the water quality using a machine learning algorithm. We have already collected and prepared a dataset and trained our algorithm to predict probable fish diseases based on the water quality parameters.

Step 3: Analyzing the disease and identifying.

Step 4: Making wise decisions to minimize harm to the fish farm and ensure healthy habitat.

1.2 PURPOSE

Trends Sci. of 11 Biochemical Oxygen Demand (BOD) measures the amount of dissolved oxygen used by aerobic microorganisms when the organic matter is decomposed in water. It offers an index for evaluating the impact of wastewater discharged on the recipient area. A high BOD value indicates many available organic compounds for bacteria that consume oxygen . Nitrate is formed by combining oxygen or ozone with nitrogen. Nitrogen is helpful for all living bodies. But a higher level of nitrate is harmful to all the live bodies underwater. The term Coliform indicates bacteria that are present in animals, including humans. Coliform does not cause diseases, but some coliform like E. coli can cause serious harm to the living body.

1.3 SCOPE

Feature extraction refers to the use of a computer to extract image information from an image, where the features are the same type of measurement data in the image. Different features could be used for the automatic and accurate detection of fish diseases. Common feature extraction can be divided into a color feature, texture feature, shape feature, and spatial feature. The more prevalent ones are the Haar feature, LBP feature, HOG feature, Shif feature, etc. Different features have to be selected when studying different targets, while the accuracy of single-feature extraction and the accuracy of manual feature screening by artificial neural networks are no longer sufficient for fish disease detection in aquaculture. Table 2 below describes the application of feature extraction in aquaculture fish detection. With the development of computer vision, deep learning has been greatly applied in the direction of automatic feature extraction, which has greatly improved accuracy.

1.4 MOTIVATION

There are mainly two types of aquaculture. The first one is Mariculture which is the farming of marine organisms for food and other products such as pharmaceuticals, food additives, jewelry (e.g., cultured pearls), nutraceuticals, and cosmetics. Marine organisms are farmed either in the natural marine environment or in the land- or sea-based enclosures, such as cages, ponds, or raceways. Seaweeds, mollusks, shrimps, marine fish, and a wide range of other minor species such as sea cucumbers and sea horses are among the wide range of organisms presently farmed around the world's coastlines. It contributes to sustainable food production and the economic development of local communities. However, sometimes at a large scale of marine firming become a threat to marine and coastal environments like degradation of natural habitats, nutrients, and waste discharge, accidental release of alien organisms, the transmission of diseases to wild stocks, and displacement of local and indigenous communities.

II. LITERATURE SURVEY

Literature survey is the most important step in software development process. Before developing the tool, it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, and then next steps are to determine which operating system and language used for developing the tool. Once the programmers start building the tool, the programmers need lot of external support. This support obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into account for developing the proposed system.

1) **Autonomous underwater vehicles (AUVs): their past, present and future contributions to the advancement of marine geoscience**

Author: R. B. Wynn, V. A. I. Huvenne, T. P. Le Bas et al.,

Autonomous Underwater Vehicles (AUVs) have a wide range of applications in marine geoscience, and are increasingly being used in the scientific, military, commercial, and policy sectors.

International Journal of Management Research and Business Strategy

ISSN 2319-345X

www.ijmrbs.com

Vol. 13, Issuse. 1, Mar 2024

Their ability to operate autonomously of a host vessel makes them well suited to exploration of extreme environments, from the world's deepest hydrothermal vents to beneath polar ice sheets. They have revolutionised our ability to image the seafloor, providing higher resolution seafloor mapping data than can be achieved from surface vessels, particularly in deep water. This contribution focuses on the major advances in marine geoscience that have resulted from AUV data. The primary applications are i) submarine volcanism and hydrothermal vent studies, ii) mapping and monitoring of low-temperature fluid escape features and chemosynthetic ecosystems, iii) benthic habitat mapping in shallow- and deep-water environments, and iv) mapping of seafloor morphological features (e.g. bedforms generated beneath ice or sediment-gravity flows.

2) Integration of navigation systems for autonomous underwater vehicles

Author: M. Dinc and C. Hajiyev

An autonomous underwater vehicle (AUV) requires a precise navigational system for localization, positioning, path tracking, guidance and control. The main navigational device for an AUV is an inertial navigation system (INS) because high-precision navigational devices such as the Global Positioning System have a limited usage in the underwater environment. In this study, based on the dynamic mathematical model of AUV, we develop two types of low-cost integrated navigational system for AUVs based on error models of INS and its aiding devices such as Doppler velocity logs, compasses and pressure depth sensors. Nine- and 15-state INS error models and corresponding measurement models of aiding devices are derived for the Kalman filter (KF). We compare the perform ance of those two integrated navigation systems. The simulation results confirm that low-cost inertial measurement unit sensors produce a notable amount of noisy measurements, but KF-based integrated navigation system models for AUV can effectively mitigate those drawbacks.

3) Design of autonomous underwater vehicle (AUV) control unit

Author: Y. Zhou, S. Cui, Y. Wang, and C. Ai

This paper outlines the conception, modeling and control of a rover type modular mini submarine. The development and implementation of the mechanical structure as well as the embedded electronics is described. The onboard instrumentation and sensors required to collect data on the environment and on its own position and orientation are also described. The mathematical representation to describe the movement of an underwater vehicle is analyzed considering the characteristics and limitations of the underwater robot. Furthermore, a control algorithm is implemented based on Lyapunov theory and Backstepping Integral Adaptive (BIA). This control strategy stabilizes the vehicle in position and orientation. The proposed control algorithm is validated in numerical simulations as well as in experimental tests which confirm the good performance of the prototype and the controller.

III. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

The CNN system with perfect images without noise. Next step is to build our own dataset of fish in the ocean. Because it is difficult to obtain images from other kinds of objects such as the sea turtle, coral and so on. This part of the research, fish is the only object to be detected. For the collection of 410 images, many of them have multiple fish in one image, so the detection is challenging. The same method was chosen to create ground truth image. And all the parameters introduced before remain the same, only the class information is made up of a 1 × 1 × 18 vector instead of 1 × 1 × 30 because of reduce in the classes. one labelled image example. It is obvious that this data set is totally different from the ideal images from ImageNET.

Disadvantages of Existing System

- Less accuracy
- It apply on perfect images with no noise

3.2 PROPOSED SYSTEM

The paper author is analyzing water quality by applying machine learning algorithm to predict fish diseases. If water contains high toxins or viruses then it will affect fish lungs which causes disease inside fish or fish may die. So we can collect water samples and then apply on machine learning model to predict water quality and if quality is not good then we can say fish is not healthy. In propose paper author has used water quality dataset from KAGGLE website and then trained with Gradient Boosting Algorithm. This algorithm giving more than 95% accurate prediction accuracy on test data.

Advantages of Proposed System

- High accuracy
- Any type of Images is consider

3.5 HARDWARE REQUIREMENTS

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. The following sub-sections discuss the various aspects of hardware requirements.

- System : CORE i3 Processor.
- Hard Disk : 100 GB.
- RAM : 4 GB.

3.6 SOFTWARE REQUIREMENTS

Software Requirements deal with defining software resource requirements and pre-requisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or pre-requisites are generally not included in the

International Journal of Management Research and Business Strategy
ISSN 2319-345X
www.ijmrbs.com
Vol. 13, Issuse. 1, Mar 2024

software installation package and need to be installed separately before the software is installed.

- Operating system : Windows 7 Ultimate(min)
- Coding Language : Python
- Front-End : Python
- Designing : HTML,CSS.
- Dataset : Kaggle

IV. SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

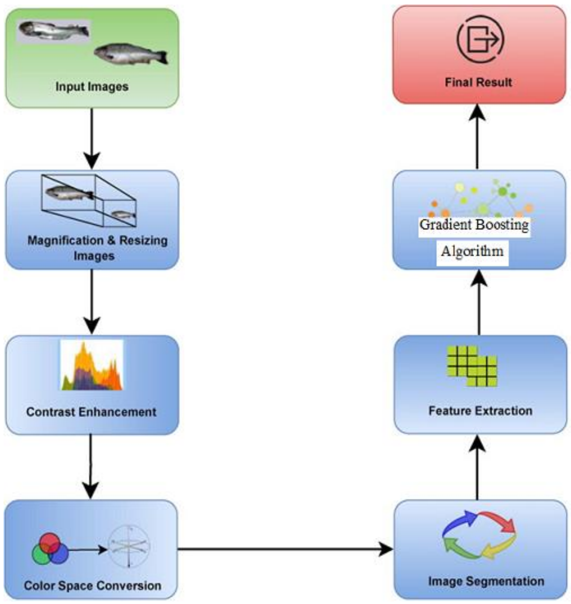


Fig:1. System Architecture

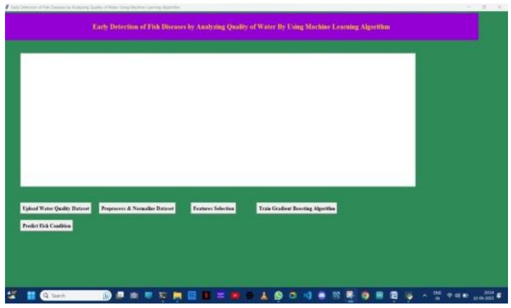
4.2 MODULES

Implementation is the stage where the theoretical design is converted into programmatically manner. In this stage we will divide the application into a number of modules and then coded for deployment. The front end of the application takes Python, and as a Back-End Data base we took data from Kaggle. The application is divided mainly into following modules. They are as follows:

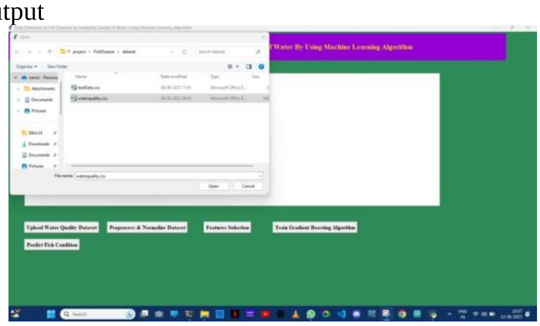
- **Upload Water Quality Dataset:** using this module we will upload dataset to application
- **Preprocess & Normalize Dataset:** using this module will convert all non-numeric data to numeric data and then normalize all values
- **Features Selection:** using this module application will select X training features and Y class label and then split dataset into train and test where application using 80% dataset for training and 20% for testing
- **Train Gradient Boosting Algorithm:** 80% training data will be input to Gradient Boosting algorithm to train a model and this model will be applied on test data to calculate prediction accuracy
- **Predict Fish Condition:** using this module we will upload test data and then algorithm will predict weather fish is healthy or disease affected.

V. OUTPUT SCREENS

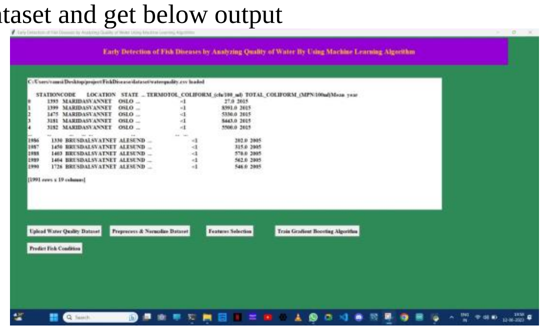
To run project double click on 'run.bat' file to get below screen



In above screen click on 'Upload Water quality Dataset' button to upload dataset and get below output



In above screen click on 'Open' button to load dataset and get below output



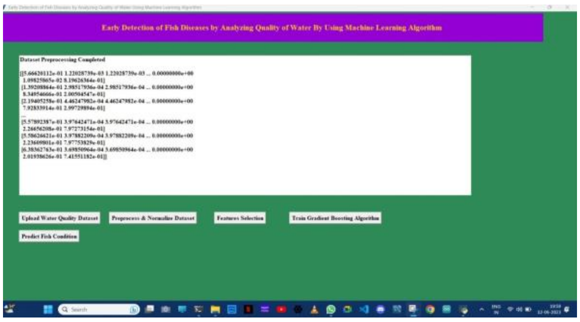
In above screen dataset loaded and we can see dataset contains numeric and non-numeric values and machine learning algorithms accept only numeric dataset so by applying label encoder class we can convert non-numeric data to numeric values so click on 'Preprocess & Normalize Dataset' button to get below output

International Journal of Management Research and Business Strategy

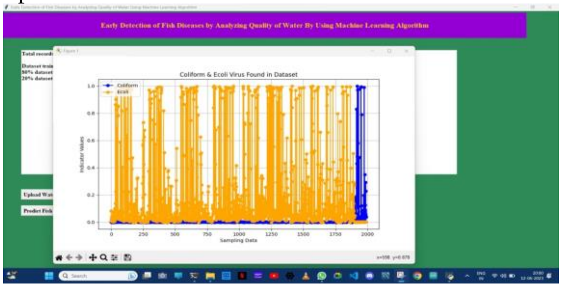
ISSN 2319-345X

www.ijmrbs.com

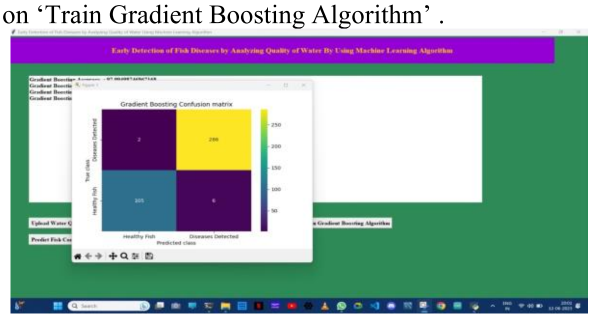
Vol. 13, Issuse. 1, Mar 2024



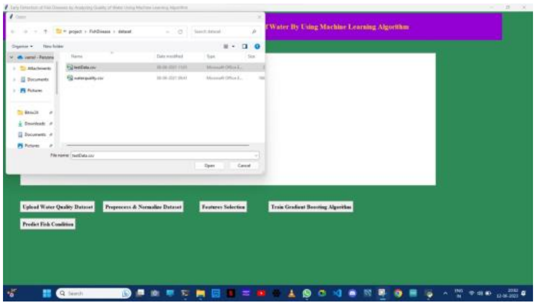
In above screen all values are converted to numeric format and now click on 'Features Selection' button to extract X and Y features from dataset and then split into train and test values and get below output



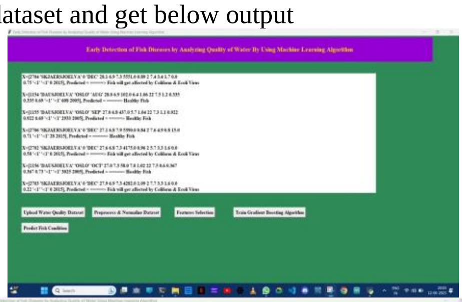
In above screen we can see dataset total values and then we can see training and testing dataset size and in graph x-axis represents number of records and y-axis represents presence quality of 'Coliform and Ecoli' virus where yellow line is for Ecoli and blue line for Coliform and now close above graph and then click on 'Train Gradient Boosting Algorithm' .

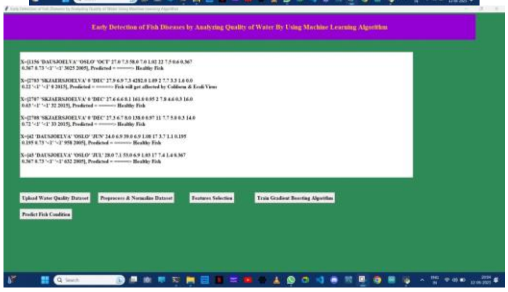


In above screen with Gradient Boosting we got 97% accuracy and in confusion matrix graph x-axis represents Predicted Labels and y-axis represents True Labels and blue colur boxes represents Incorrect prediction count which is 2 only and different colour boxes contains correct prediction count. Now close above graph and then click on 'Predict Water Quality & Risk' button to get below output



In above screen selecting and uploading 'testData.csv' file and then click on 'Open' button to load dataset and get below output





In above screen in square bracket we can see test data values and after arrow symbol we can see predicted values as healthy or disease affected fish

VI. CONCLUSION AND FUTURE SCOPE

The authors of this study created a neural network model to identify fish. The data augmentation strategy was used to provide adequate dataset to help the training process. The overfitting issue was resolved by using the dropout method. Additionally, the loss function was improved to update the network's parameters. These methods resulted in significant reductions in training time and training loss. In order to explore a practical solution for fish detection, this article revised loss function and other CNN parameters, established the data set to include real blur ocean water conditions, and targeted the system at an embedded system for AUV design with all feasible optimisations.

FUTURE SCOPE

We want to use different Convolutional Neural Networks (CNN) architecture in the future to recognise fish sickness more accurately and thoroughly. Additionally, we will concentrate on putting the suggested solution into practise using a real-world IoT device. By doing this, aquaculture producers may be able to recognise diseased salmon fish and take the necessary action to prevent any

International Journal of Management Research and Business Strategy
ISSN 2319-345X
www.ijmrbs.com
Vol. 13, Issuse. 1, Mar 2024

unanticipated losses to their operations. We'll experiment with other fish datasets to improve our system's applicability to other aquaculture fields. As salmon fish is one of the most in-demand commodities globally, we will also focus on expanding our current information.

REFERENCES

1) R. B. Wynn, V. A. I. Huvenne, T. P. Le Bas et al., “Autonomous underwater vehicles (AUVs): their past, present and future contributions to the advancement of marine geoscience,” *Marine Geology*, vol. 352, pp. 451–468, 2014.
2) M. Dinc and C. Hajiyev, “Integration of navigation systems for autonomous underwater vehicles,” *Journal of Marine Engineering & Technology,* vol. 14, no. 1, pp. 32–43, 2015.
3) Y. Zhou, S. Cui, Y. Wang, and C. Ai, “Design of autonomous underwater vehicle (AUV) control unit,” in *2015 ASEE Gulf-Southwest Annual Conference*, pp. 25–27, ASEE Gulf-South, San Antonio, TX, 2015.
4) Y. Zhou, S. Cui, Y. Wang, and L. Zhai, “A refined attitude algorithm for AUV based on IMU,” in *15th International Conference on Scientific Computing (CSC’17)*, pp. 16–22, CSREA Press ©, Las Vegas, NV, 2017.
5) A. Cadena, P. Teran, G. Reyes, J. Lino, V. Yaselga, and S. Vera, “Development of a hybrid autonomous underwater vehicle for benthis monitoring,” in *Proceedings of 2018 4th International Conference on Control, Automation and Robotics (ICCAR)*, pp. 20–23, IEEE, Auckland, New Zealand, 2018.
6) M. Eichhorn, H. C. Woithe, and U. Kremer, “Parallel of path planning algorithms for auvs concepts, opportunities, and program – technical implementation,” in *2012 Oceans - Yeosu*, MTS/IEEE, Yeosu,South Korea, 2012.
7) Greenfacts. Fisheries latest data, Available at: https://www.greenfacts.org/en/fisheries/l-2/06-fishconsumption.htm, accessed May 2020.
8) Department of Fisheries Bangladesh. Yearbook of fisheries statistics of Bangladesh, 2017-18. Vol 35. Fisheries Resources Survey System (FRSS), Department of Fisheries, Ministry of Fisheries, Bangladesh, 2018, p. 129.
9) N Bagum, MS Monir and MH Khan. Present status of fish diseases and economic losses due to incidence of disease in rural freshwater aquaculture of Bangladesh. J. Innov. Dev. Strat. 2013; 7, 48- 53.
10) M Chaurasia, I Gupta, R Kumar, P Prasad and RA Sohony. Water quality analysis of surface water: A web approach. Environ. Monit. Assess. 2013, 185, 5987-992.
11) N Rahmanian, SHB Ali, M Homayoonfard, NJ Ali, M Rehan, Y Sadef and AS Nizami. Analysis of physiochemical parameters to evaluate the drinking water quality in the State of Perak, Malaysia. J. Chem. 2015; 2015, 716125.
12) C Nagamani, CS Devi, and A Shalini. Physico-chemical analysis of water samples. Int. J. Sci. Eng. Res. 2015; 6, 2149-55.