International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

# Prospective Applications of Data Mining in Indian Fisheries

**J. Gladju, Ph. D Scholar, Department of Computer Science, Nallamuthu Gounder Mahalingam College, Pollachi, Tamil Nadu, India, gladju@gmail.com**

**Dr. A. Kanagaraj, Assistant Professor, Department of Computer Science, Nallamuthu Gounder Mahalingam College, Pollachi, Tamil Nadu, India, a.kanagaraj@gmail.com**

**Abstract: In India, capture fisheries and aquaculture produced 12.6 million tonnes of different aquatic organisms in 2018 to feed millions of people. Further, as an important economic activity, it provided livelihood to more than 14.5 million people and earned substantial foreign exchange from fish and fishery products. Corresponding to the rapid growth and technological progress witnessed in this sector, the volume of information and scientific data available in fisheries, aquaculture and allied activities is also increasing day by day. Few attempts have been made to capture the essence of information from the available data to develop intelligent decision support systems for fisheries management, but there is a need for more comprehensive efforts. On the other hand, the adoption of artificial intelligence tools based on data mining and machine learning algorithms is already making significant impact on aquaculture production systems, particularly in the optimisation of feed use, disease prevention, biomass monitoring and market intelligence. In this milieu, we illustrate and briefly discuss the data mining process and techniques and their scope of application in capture fisheries and aquaculture.**

***Keywords* — *Aquaculture, Artificial Intelligence, Data mining, Decision support, Fish, Fisheries***

## I. INTRODUCTION

India is naturally bestowed with diverse inland and marine water resources which harbour a rich aquatic biodiversity and valuable fisheries resources. Utilising this, the country has made rapid strides in fisheries and aquaculture sector, and at present, India is world's second largest producer of fish and other aquatic products. According to official statistics, the total fish production in 2018 was nearly 12.6 million metric tonnes, with 3.7 million metric tonnes coming from the marine sector and 8.9 million metric tonnes coming from the inland sector. Out of this, 1.4 million tonnes of fish and fish products were exported earning valuable foreign exchange to the tune of 45,000 crore rupees [3]. Besides contributing more than 5% to the agriculture GDP, fisheries sector also provides direct employment to over 14 million people and nutritional security to many more. Through commendable scientific innovations and infrastructure development, the capture fisheries sector in India is evolving from artisanal to technologically advanced commercial fishing. Likewise, new technologies, better management and diversification of culture species have made aquaculture more efficient and specialized. Among the technological advancement, the application of artificial intelligence is contributing to huge improvements in the efficiency of global aquaculture. Particularly, data mining and machine learning tools are used in decision making processes, which for example helps in the reduction of feed wastage (Observe Technologies, eFishery and Umitron Cell); prevention of fish diseases (AquaCloud and Aquaconnect); and growth prediction to price tracking (XpertSea) [1]. Similarly, there is a good scope for developing smart decision support systems by mining data and discovering useful information from available Indian data sets on 1) species-wise, year-wise and area-wise aquaculture and capture fish production; 2) seasonal and annual fluctuations in major aquatic environmental factors such as temperature and primary production; 3) farm management data related to water quality changes, feed usage, fish / shrimp growth patterns and behavioural indices; 4) fish trade data such as species-wise export and import volumes; and 5) nutritional composition of fishes with relevance to human dietary intake recommendations in terms of protein, lipid, amino acid, fatty acid, mineral and vitamin content (Fig. 1).

 DOI : 10.35291/2454-9150.2021.0139 

ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

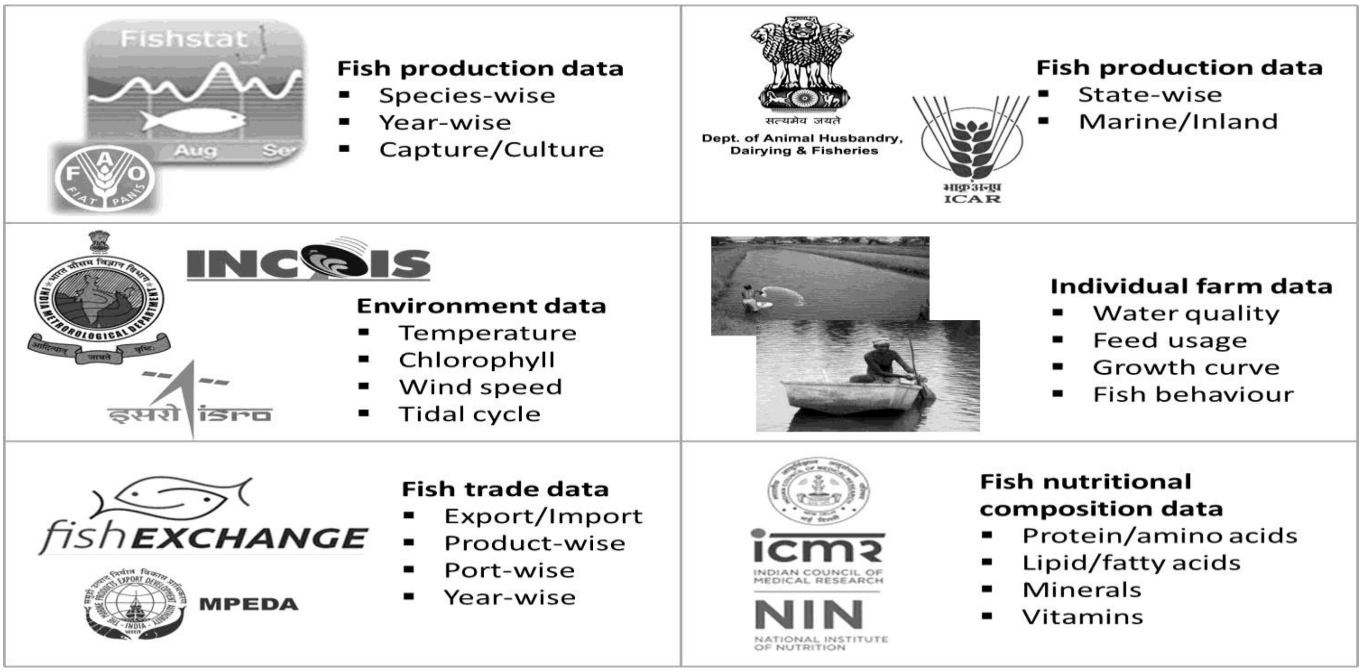


**Fig. 1. Sources of Data on Indian Fisheries**

## II. A SIMPLE OVERVIEW OF DATA MINING TECHNIQUES

Data mining is the computing process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics and database system [11]. The process of extracting important and useful information from large sets of data (i.e., data mining) and their applications in different research areas are a fertile research field. Data mining techniques can be mainly divided into supervised or predictive (classification, regression and time series analysis) and unsupervised or descriptive (clustering, association rules, summarization and sequence discovery) techniques [9], [10]. For example, k-means is a data mining technique for clustering, where the aim is to group data based on measure of similarities between data samples in an unclassified data set according to some distance measure. The k-means belongs to the category of expectation maximization algorithms, which are powerful for finding maximum likelihood solutions for models with latent variables. On the other hand, k-nearest neighbour is a common classification technique where a training set is known and it is used to classify samples of unknown classification, with the basic assumption that similar samples should have similar classification. Advanced to this, is the artificial neural networks data classification system inspired by the research on human brain. In this network, each node represents a neuron (to perform simple tasks) and each link represents the way two neurons interact (to perform complex tasks). The structure of the network determines the ability of a neural network to perform a given task and the most used type of artificial neural network is the multilayer perceptron where neurons are organized in layers. Further, to classify data samples in two disjoint or linearly separable classes, binary classifiers like support vector machines are used [13]. We have provided a schematic representation of the data mining process in Fig.2. Also, we have provided the categories and organisation of data mining techniques in Fig.3. Some of these data mining techniques are known to have varied applications in the field of agriculture, fisheries management and aquaculture, as briefed below.

 DOI : 10.35291/2454-9150.2021.0139 

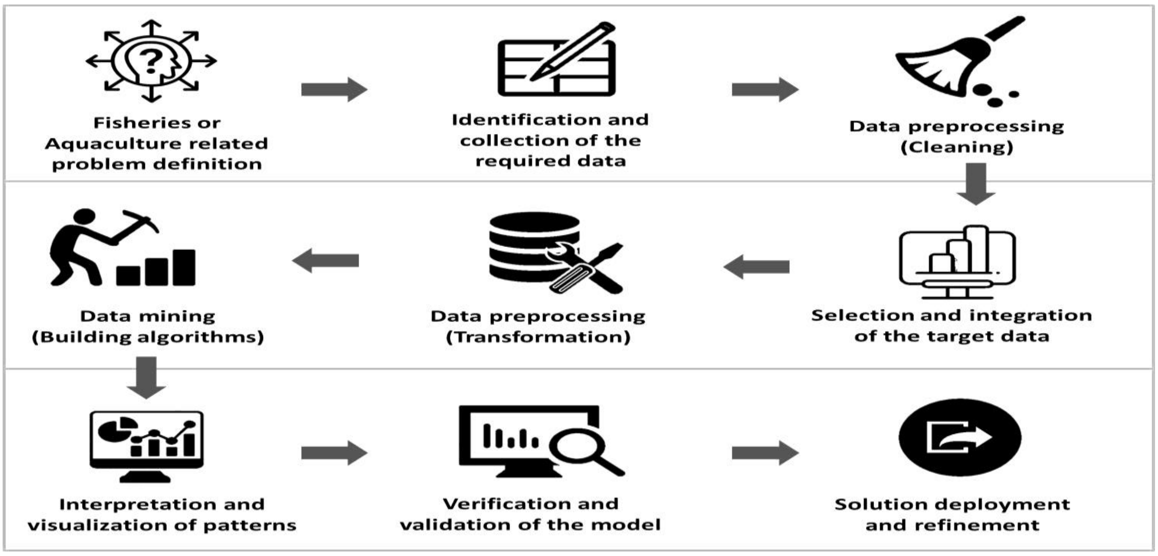


Fig. 2. Schematic Representation of Data Mining

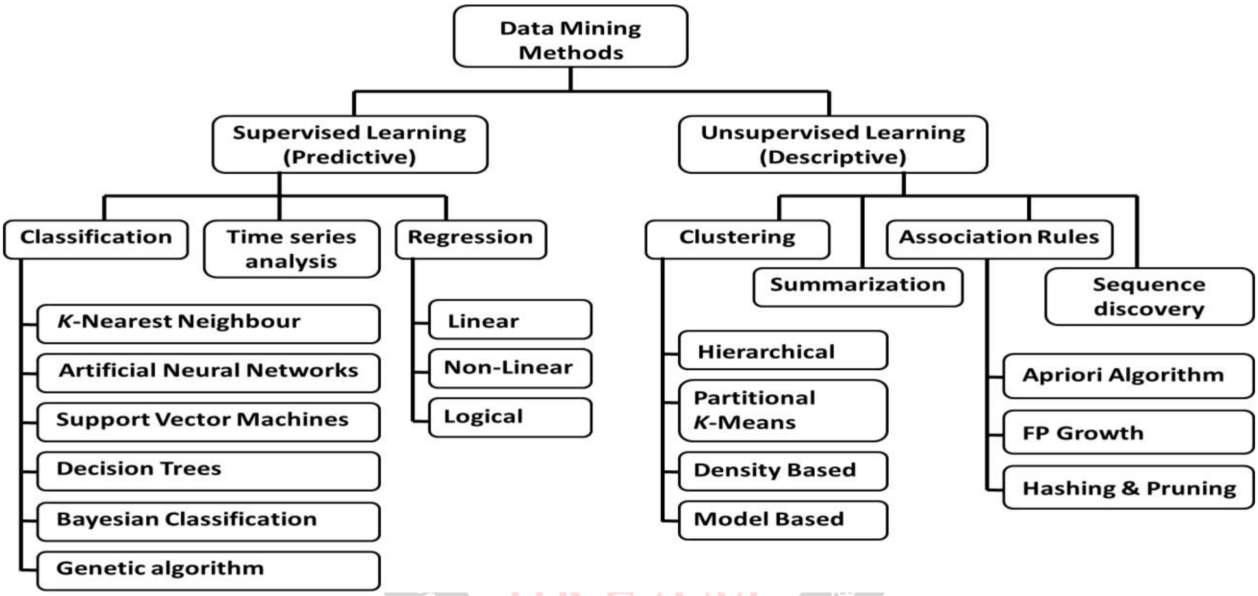


Fig. 3. Organisation of Data Mining Techniques

## III. APPLICATION OF DATA MINING IN AGRICULTURE AND LIVESTOCK FARMING

Agriculture is a sector where farmers and enterprises have to make several decisions every day, considering all the intricate complexities involved and the various factors influencing the yield estimates. Data mining techniques can be an effective approach for accomplishing practical and effective solutions for several of these problems [11]. At present, the data mining techniques used in agriculture are related to spatio-temporal weather conditions (e.g. k-nearest neighbour for classifying daily precipitation), soil characteristics (e.g. k-means for clustering soil types), accurate yield estimation / prediction for crops, quality of the agricultural commodity (e.g. artificial neural networks for classifying fruits based on colour and juiciness) and post-harvest process monitoring (e.g. support vector machine classification based on smell or taste sensors) [13], [17]. Similarly in livestock, data mining techniques can be used to study sound recognitions problems i.e., support vector machine can classify the sound of diseases and healthy animals. Artificial neural networks can classify eggs as fertile and non-fertile using computer vision. Data mining techniques are also useful in monitoring and controlling the housing environment of livestock and in classifying meat based on specific characteristics during post-harvest processing. Other specialized applications include the detection of prohibited ingredients in animal feeds using a combination of near-infrared spectroscopy and support vector machines; prediction of changes in agricultural productivity due to climate vagaries [13], [17]. Data availability in agriculture and livestock is ever increasing from many sources and correspondingly the opportunity and utility for data mining applications is growing. In the near future, more sophisticated and tailor-made data mining techniques will be required to address complex problems and provide better solutions in agriculture-related fields. This includes the use of complex techniques such as biclustering and high-performance parallel computing systems [13]. Further, as data mining tools predict future trends and behaviours, it can assist agriculture and livestock researchers / policy makers to take proactive and knowledge-driven decisions [17]. On a cautionary note,

 DOI : 10.35291/2454-9150.2021.0139 

ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

many agricultural models have suffered from fragmentation in implementation, redundancy and difficulty to apply and generate real solutions for the agricultural sector. This situation can be improved only if an open, self-sustained, and committed community starts developing agricultural models, associated data and tools as a common resource [8]. The opportunities and challenges for data mining applications in fisheries are very similar to the agriculture and livestock sector.

## IV. APPLICATION OF DATA MINING IN FISHERIES

In general, fisheries sector includes the industry that harvest fish from the oceans and inland water bodies, the farming system which produces fish and other aquatic organisms from water and the food processing and value addition industry which develops various products for local consumption and export [18]. As the fisheries sector is primarily driven by the demand and profit generated, there are many unforeseen trends at different phases of its development, necessitating the use of data mining applications. For instance, the spatiotemporal dynamics of a living organism and its life cycle processes are closely related with environmental factors. To study this ecological aspect, a specific data-mining approach i.e., the spatiotemporal assignment mining model (STAMM) has been used to extract the spatiotemporal pattern and assignment of environmental factors (e.g. temperature) which controls fish distribution in yellow sea, China. Neighbourhood rules and relationships were used to construct a decision table, with recursively processed indices that expresses the probabilities of the ecological association [14]. Another data mining approach to identify hydrologic indicators related to fish community was based on genetic programming using the Shannon index and the number of individuals of a fish community [19]. Similarly, artificial neural network based self-organising feature maps was found to be useful for exploring the relationship between fish community and water quality by clustering, analysing and visualizing heterogeneous datasets [16]. In another angle, data mining approach has been used for stock assessment of fisheries resources based on catch per unit effort analysis and species-wise fisheries statistics [5]. Further, data mining algorithms such as AGRID+ (a grid density based clustering for high dimensional data) was found to be useful in the determination of potential fishing zones based on temporal and daily aggregates of fishes in different spatial clusters [6]. Data mining techniques such as multiple regression models and k-means clustering have also been used to ascertain and forecast the economic welfare of fishermen and fish farmers, using microeconomic and macroeconomic indicators [15]. In the Indian fisheries context, hitherto, no substantial effort has been made to elicit useful data patterns and information from the existing databases, but the opportunities are plenty. Here below, in fig. 4, 5, 6 and 7, we have illustrated and schematically represented the utility of data mining techniques for different fisheries and aquaculture applications.

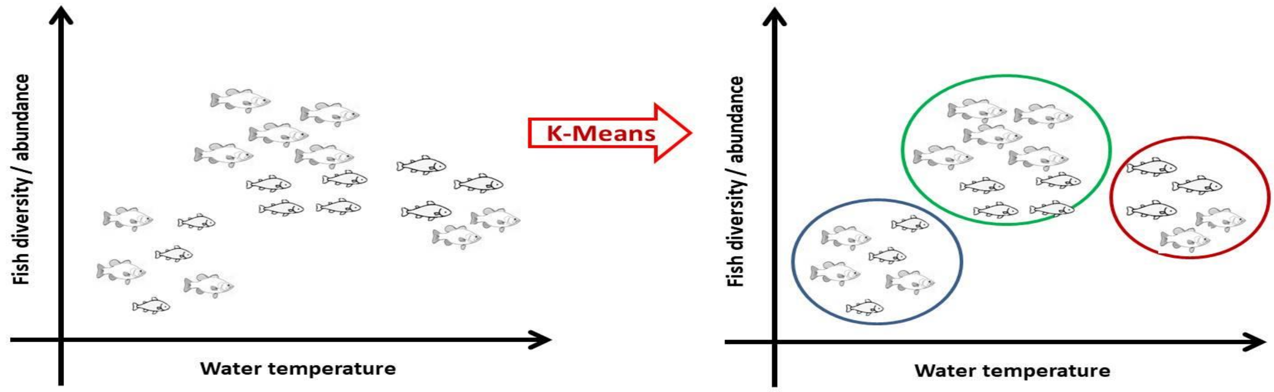


**Fig. 4. K-means Clustering for Fish Stock**

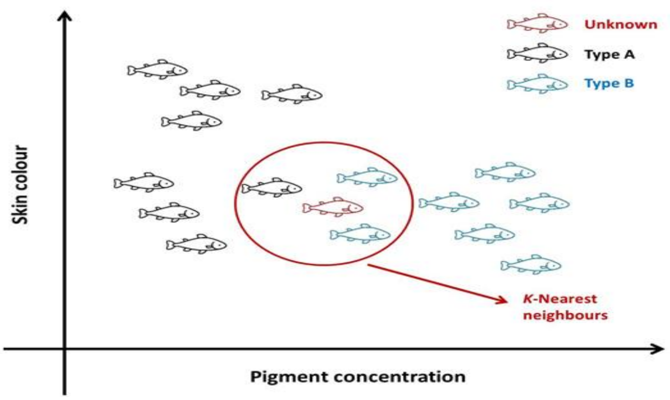


**Fig. 5. *K*-Nearest Neighbour Classification for Fish Colour Analysis**

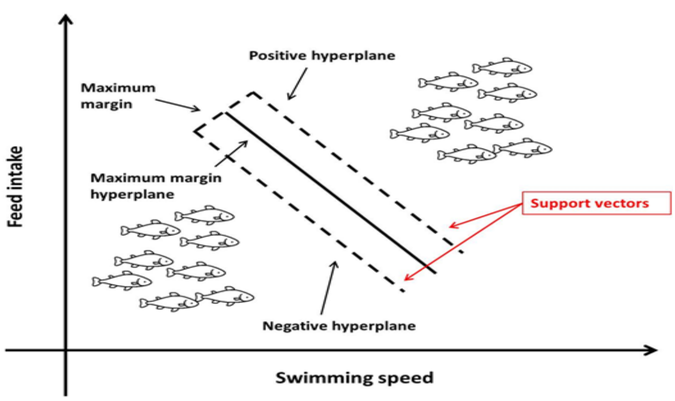


**Fig. 6. Support Vector Machine for Monitoring Fish Behaviour**


DOI : 10.35291/2454-9150.2021.0139

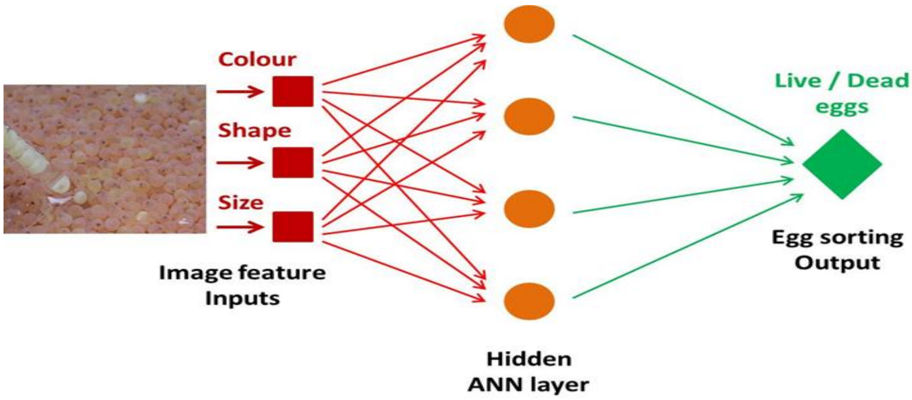


**Fig. 7. Artificial Neural Network for Sorting Live and Dead Eggs**

As illustrated in the above simple examples, data mining algorithms can be used to process and analyse huge volumes of data related to water quality, visual images of fishes, acoustic features of fish distribution or video recordings of fish behaviour in fish farms or natural aquatic habitats, as soon as it is collected by technologically advanced sensor networks [2], [7], [4]. Based on the patterns, sequences and knowledge which is discovered through the data mining process, we can develop intelligent decision support systems which can help in several aspects of fisheries and aquaculture such as 1) fish stock assessment with respect to environmental changes; 2) monitoring and regulation of fishing activities; 3) identification of fishing grounds or suitable marine areas for protection; 4) surveillance of aquatic ecosystem changes with respect to natural (climate change) dynamics and anthropogenic activities; 5) monitoring and control of aquaculture production systems; 6) optimization of resource (water) and input (feed) use; 7) monitoring of fish biomass and behavioural changes for predicting growth curves, stress status and disease incidences; 8) fish processing and value addition systems; 9) market intelligence such as price projections; and 10) socio-economics of fishers and fish farmers in India and elsewhere [2], [12], [20].

## V. Conclusion

With technological developments, it is becoming easier to obtain huge volumes of data in fisheries. But, rational analysis of these data is getting more complicated and difficult. Development and deployment of appropriate data mining algorithms could be the solution for unravelling and using the hidden knowledge in the amassed data. However, to generate useful and accurate information, data collection methods should be standardized; data sharing and security protocols should be well defined; speed and efficiency of data processing should be continuously upgraded; and data mining tools should be user-friendly and cost-effective. In near future, it is definitely possible to address the multiple challenges that limits fish production systems by effectively using data mining and artificial intelligence for accurate assessment and smart use of the available fisheries and aquatic resources.

## Acknowledgment

The academic support of Dr. Antony Selvadoss Thanamani, Head, Department of Computer Science, in the PhD research of the first author is gratefully acknowledged. The authors also gratefully acknowledge the guidance of Dr. Biju Sam Kamalam, Scientist, ICAR-Directorate of Coldwater Fisheries Research, in the preparation of this manuscript.

## References

[1] J.V. Beijen and G. Yan, "A practical guide to using AI in aquaculture", Retrieved from https://thefishsite.com/articles/a-practical-guide-to-using-ai-in-aquaculture, 2020.

[2] D. Bradley, M. Merrifield, K.M. Miller, S. Lomonico, J.R. Wilson and M. G. Gleason, "Opportunities to improve fisheries management through innovative technology and advanced data systems. Fish and fisheries", 20(3), pp. 564-583, 2019.

[3] Department of Animal Husbandry, Dairying and Fisheries (2018), "Annual report 2017-18", Ministry of Agriculture and Farmers Welfare, Government of India

[4] O. F. El-Gayar, "The use of information technology in aquaculture management. Aquaculture Economics & Management", 1(1-2), pp. 109-128, 1997.

[5] K. Enomoto, S. Ishikawa, M. Hori, H. Sitha, S. L. Song, N. Thuok, and H. Kurokura, "Data mining and stock assessment of fisheries resources in Tonle Sap Lake, Cambodia", Fisheries Science, 77(5), pp. 713-722, 2011.

[6] D. Fitrianah, H. Fahmi, A.N. Hidayanto, and A. M. Arymurthy, "A Data Mining Based Approach for Determining the Potential Fishing Zones", International Journal of Information and Education Technology, 6(3), pp. 187, 2016.

[7] M. Føre, K. Frank, T. Norton, E. Svendsen, J. A. Alfredsen, T. Dempster, H. Eguiraun, W. Watson, A. Stahl, L.M. Sunde, C. Schellewald, K. R. Skøien, M. O. Alver and D. Berckmans, "Precision fish farming: A

 DOI : 10.35291/2454-9150.2021.0139 

new framework to improve production in aquaculture. Biosystems Engineering", 173, pp. 176-193, 2018.

[8] S. J. Janssen, C. H. Porter, A. D. Moore, I. N. Athanasiadis, I. Foster, J. W. Jones, and J. M. Antle, J. M. "Towards a new generation of agricultural system data, models and knowledge products: Information and communication technology", Agricultural systems, 155, 200-212, 2017.

[9] S. H. Liao, P. H. Chu, and P. Y. Hsiao, "Data mining techniques and applications-A decade review from 2000 to 2011", Expert systems with applications, 39(12), 11303-11311. 2012.

[10] H. A. Madni, Z. Anwar, and M. A. Shah, "Data mining techniques and applications-a decade review", In: 23rd International Conference on Automation and Computing, 1-7. 2017.

[11] J. Majumdar, S. Naraseeyappa, and S. Ankalaki, S, "Analysis of agriculture data using data mining techniques: Application of big data", Journal of Big data, 4(1), pp. 20, 2017.

[12] J. R. Mathiassen, E. Misimi, M. Bondø, E. Veliyulin and S. O. Østvik, " Trends in application of imaging technologies to inspection of fish and fish products", Trends in Food Science & Technology, 22(6), pp. 257-275, 2011.

[13] A. Mucherino, P. Papajorgji and P. M. Pardalos, P. M. "A survey of data mining techniques applied to agriculture". Operational Research, 9(2), pp. 121-140, 2009.

[14] F. Su, C. Zhou, V. Lyne, Y. Du, and W. Shi, "A data-mining approach to determine the spatio-temporal relationship between environmental factors and fish distribution", Ecological Modelling, 174(4), pp. 421-431, 2004.

[15] W. A. Teniwut, F. Pentury, and Y. A. Ngamel, "Forecasting the welfare of fishermen and aquaculture farmers in Indonesia: Data Mining Approach", Journal of Physics: Conference Series, Vol. 1175, No. 1, pp. 012066. IOP Publishing, 2019.

[16] W. P. Tsai, S. P. Huang, S. T. Cheng, K. T. Shao, and F. J. Chang, "A data-mining framework for exploring the multi-relation between fish species and water quality through self-organizing map", Science of the Total Environment, 579, pp. 474-483, 2017.

[17] S. Veenadhari, B. Misra, B and C. D. Singh, "Data mining techniques for predicting crop productivity - A review article", International Journal of Computer Science and Technology, 2(1), pp. 90-100, 2011.

[18] C. L. Vincent, V. Singh, V, K. Chakraborty and A. Gopalakrishnan, "Patent data mining in fisheries sector: An analysis using Questel-Orbit and Espacenet", World Patent Information, 51, pp. 22-30, 2017.

[19] Y. C. E. Yang, X. Cai, and E. E. Herricks, "Identification of hydrologic indicators related to fish diversity and abundance: A data mining approach for fish community analysis", Water Resources Research, 44(4), 2008.

[20] B. Zion, "The use of computer vision technologies in aquaculture - a review", Computers and Electronics in Agriculture, 88, pp. 125-132, 2012.

 DOI : 10.35291/2454-9150.2021.0139 