JURNAL
PAMATOR

Vol. 16, No. 2, 2023, p. 269-279
https://journal.trunojoyo.ac.id/pamator
ISSN: 2654-7856 (Online)
ISSN : 1829 -7935 (Print)

SUCCESSFUL VANAME SHRIMP HARVEST: PREDICTION AND PROCESS IMPROVEMENT

Iryanti Djaja[1], A. Arviansyah[2]

[1,2] Master of Management, Faculty of Economics and Business, Universitas Indonesia, Program Studi Magister Manajemen, Gedung Prof. Wahjudi Prakarsa, Jl. Salemba Raya 4 Jakarta, Indonesia, iryanti.djaja@ui.ac.id

DOI : https://doi.org/10.21107/pamator.v16i2.19794

Manuscript received 6th April 2023, Revised 30th April 2023, Published 27th May 2023

Abstract

The demand for shrimp from Indonesia continues to increase every year, thus creating more significant interest in the shrimp farming industry. Although shrimp is relatively easy to farm, many variables affect the harvest's success. The harvest in shrimp farming is calculated using % SR (Survival Rate). Our research uses machine learning approaches: logistic regression (LR), decision tree (DT) and k-Nearest Neighbour (KNN). LR, DT and KNN will be used to predict whether we will have a successful harvest. From these predictions, we also provide suggestions for business improvements to utilize data. The expected result of such advice is that the business can improve its performance and get more consistent results.

Keywords: Shrimp Farming, Aquaculture, Machine Learning, Logistic Regression, Decision Tree, K-Nearest Neighbour, BPMN

INTRODUCTION

The development of shrimp farming increased, especially after the issuance of the Minister of Marine Affairs and Fisheries Decree of the Republic of Indonesia number 41/2001. The government took this step as an anticipatory step after the decline in tiger shrimp *(Penaeus monodon)* production since 1996. The development of Vaname shrimp is so rapid that in 2004 Vaname shrimp production could beat Tiger shrimp production, with total global production reaching 1.1 million tons[1]. The market for Vaname shrimp was kept strong, although small or mid-scale industries suffered during the Covid-19 [2]

[1] BPS, *BPS 2020*, *Statistik Indonesia 2020*, 2020, ; BPS, *BPS 2019*, *Statistik Indonesia 2019*, 2019.

[2] Yenik Pujowati, Putri Ari Saruhun Hasibuan, and Sucahyo Tri Budiono, 'Analisis Dampak Covid-19 Terhadap Pendapatan UMKM (Usaha Mikro, Kecil Menengah) Di Kabupaten Nganjuk', *Jurnal*

Bangka Belitung Province is a suitable area for the development of Vaname shrimp. The suitability is proven by the continued increase in farmers on Bangka Island since 2015, reaching 1 million hectares [3]. Contrary, the biggest problem for shrimp farmers is harvest failure. Many things can cause the failure, starting from the quality of seed, feed, water conditions and development, plankton conditions, weather, and so on[4]

This research was based from PT FEI. PT FEI is a traditional shrimp farm located on Bangka Island. PT FEI's pond has been operating since 2020. This company is a small and medium-sized company whose all operating processes are still carried out traditionally. PT FEI had experienced harvest failure in the last 3-4 harvests. The biggest problem is that they do not know what caused the successful and failed harvest.

From the results of interviews in the field, they confirmed that everything they did was the same even though the results were different. We realize that the economic conditions of workers were increasingly tricky due to continuously failed harvests without knowing the cause. From the field visit, we could observe that PT FEI has made various efforts and tests so that yields improved. It is exciting to investigate by linking several scientific principles of operations management to increase productivity characterized by successful harvests.

In general, one of the challenges in operations management is how to improve successful harvest that PT FEI also experienced. With various crop failures that have been experienced, PT FEI needs to improve its operation management steps. The Operational Management[5] reference provides an example of how process improvements in operating processes can increase productivity. An article[6] explains how the use of data can help to improve results.

This study discusses the process of shrimp cultivation, which consists of pond preparation, hatchery, and growth. The growing process includes selecting and spreading seeds, maintaining water and feed quality, and disease control to harvest. Harvest success is calculated using Survival Rate, which compares the number of shrimps at harvest time with the number of seeds speeded. Data-driven analysis can be done using machine learning to increase agriculture and manufacturing productivity. This study aims to predict harvest success factors using machine learning, design improvements to the "Data Utilization" business process to produce more optimal results, and design data

Pamator : Jurnal Ilmiah Universitas Trunojoyo, 15.1 (2022), 100–112 <https://doi.org/10.21107/pamator.v15i1.13922>.

[3] Arya Bima Mahendra, 'Jadi Daerah Dengan Tambak Udang Vaname Terluas Di Bangka Belitung, Sayang Belum Bisa Sumbangkan PAD - Bangkapos.Com', *BangkaPos*, 2022.

[4] Nono Hartanto, 'SOP Pembesaran Udang Vaname Di Tambak Milenial', 2020; Atikah Indri Prastianti, 'Faktor-Faktor Yang Mempengaruhi Produksi Udang Vannamei (Litopeneaus Vannamei) Di Desa Pantai Bahagia, Kecamatan Muara Gembong' (Universitas Islam Negeri Syarif Hidayatullah, 2021).

[5] Barry Heizer, Jay; Render, *Operation Management Sustainability and Supply Chain Management*, *Pearson*, 2017, XII.

[6] Tim Lawrence, 'Exploiting Big Data and Analytics to Improve Productivity in Manufacturing', *Manufaturing Digital*, May 2020.

visualizations to support further actions.[7] This study also aims to provide strategies for utilizing daily data to support productivity.[8] This paper uses data from PT FEI (aka).

RESEARCH METHODS

This research combines qualitative and quantitative approaches with business strategies to achieve optimal production results[9]. The qualitative approach is carried out by field observation and interviews, while the quantitative approach is carried out by data analysis using machine learning. This research includes classifying predictive analysis with KNN, LR, and DT approaches[10][11]. The data used came from PT FEI, with 930 data analyzed. The operational variables used are pH, salinity, crop yield, NH_4^+, NOx, and ratio of adverse bacteria. Data is processed using Python[12]. The model refers to previous research using machine learning frameworks and methods[13]. This study aims to increase productivity results in the shrimp fishing industry by predicting harvest success[14]. The machine learning approach in aquaculture has been used in several industries like

[7] Heizer, Jay; Render, XII.

[8] Lawrence.

[9] Lars Buitinck and others, 'API Design for Machine Learning Software: Experiences from the Scikit-Learn Project', 2013, 1–15.

[10] Buitinck and others.

[11] IBM, 'What Is Logistic Regression?', *IBM.Com*, 2022, p. 1 <https://www.ibm.com/topics/logistic-regression#:~:text=Logistic regression estimates the probability,bounded between 0 and 1.>.

[12] Dewi Alima Nostalia Suseno and others, 'Analisis Faktor Produksi Budidaya Udang Vannamei (Litopenaeus Vannamei) Di Tambak HDPE (High Density Polyethilene) Pulokerto Pasuruan', *Balitbang Kkp*, 19 (2021), 99–104.

[13] Delbert Buentello, Alejandro J; Gatlin, 'Effects of Water Temperature and Dissolved Oxygen on Daily Feed', *Aquaculture*, 182.2000 (2000), 339–52; Huan Gao and others, 'Quantitative Analysis of Temperature, Salinity and PH on WSSV Proliferation in Chinese Shrimp Fenneropenaeus Chinensis by Real-Time PCR', *Aquaculture*, 312.1–4 (2011), 26–31 <https://doi.org/10.1016/j.aquaculture.2010.12.022>; J. Gladju, Biju Sam Kamalam, and A. Kanagaraj, 'Applications of Data Mining and Machine Learning Framework in Aquaculture and Fisheries: A Review', *Smart Agricultural Technology*, 2.April (2022), 100061 <https://doi.org/10.1016/j.atech.2022.100061>.

[14] Joel Grus, *Data Science from Scratch*, *O'Reilly*, 1st edn (O'Reilly, 2015), MDXLII; Thomas C Kenett, Ron S; Redman, 'The Real Work of Data Science', *Forbes*, 204.1 (2021), 52–54 <https://doi.org/10.1002/9781119570790.ch1>; Olshen R, 'Sckitlearn Decision Trees', *Sckit-Learn Developers*, 2009 <https://scikit-learn.org/stable/modules/tree.html>; Adam Shafi, 'K-Nearest Neighbors (KNN) Classification with An Overview of K-Nearest Neighbors', February 2023; Ming Sun, Xiaofen Yang, and Yinggang Xie, 'Deep Learning in Aquaculture: A Review', *Journal of Computers*, 31.1 (2020), 294–319 <https://doi.org/10.3966/199115992020023101028>; David G Whiting, H Dennis Tolley, and Gilbert W Fellingham, *An Empirical Bayes Procedure for Adaptive Forecasting of Shrimp Yield*, *Aquaculture*, 2000, CLXXXII; Jinyong Zhu, 'Big Data : How Much Data Is Big Data ? What Is Big Data ?', 2022.

shellfish[15], oyster[16], fish[17]

RESULT AND DISCUSSION

Data Analysis with Machine Learning

The selected machine learning model is based on several reference journals[18,19]: k-Nearest Neighbor, Decision Tree, and Logistic Regression. Good predictions are indicated by higher accuracy.

The variables used are data on PT FEI: Salinity, pH, NH_4^+, NOx, and the ratio of adverse bacteria. Category for successful harvest if SR ≥ 75%. The results of previous studies show that NH4+, salinity, pH, and NOx can determine the success of shrimp harvest.

Table 1. Variables from previous research

No.	Variables Determining Crop Success	Information From Previous Research
1.	NH_4^+ and salinity	20
2.	pH and salinity	21
3.	pH, Salinity, NH_4^+, NOx	22

15 Md Sumon Shahriar, Ashfaqur Rahman, and John McCulloch, 'Predicting Shellfish Farm Closures Using Time Series Classification for Aquaculture Decision Support', *Computers and Electronics in Agriculture*, 102 (2014), 85–97 <https://doi.org/10.1016/j.compag.2014.01.011>.

16 Erick Bataller and Andrew D. Boghen, 'Elimination of the Gill Worm Urastoma Cyprinae (Graff) from the Eastern Oyster Crassostrea Virginica (Gmelin) Using Different Salinity-Temperature Combinations', *Aquaculture*, 182.3–4 (2000), 199–208 <https://doi.org/10.1016/S0044-8486(99)00281-1>.

17 Shili Zhao and others, 'Application of Machine Learning in Intelligent Fish Aquaculture: A Review', *Aquaculture*, 540.March (2021), 736724 <https://doi.org/10.1016/j.aquaculture.2021.736724>.

18 Ashfaqur Rahman, Stuart Arnold, and Joel Janek Dabrowski, 'Identification of Variables Affecting Production Outcome in Prawn Ponds: A Machine Learning Approach', *Computers and Electronics in Agriculture*, 156 (2019), 618–26 <https://doi.org/10.1016/j.compag.2018.12.015>.

19 Ashfaqur Rahman and others, 'An Integrated Framework of Sensing, Machine Learning, and Augmented Reality for Aquaculture Prawn Farm Management', *Aquacultural Engineering*, 95.July (2021), 102192 <https://doi.org/10.1016/j.aquaeng.2021.102192>.

20 Whiting, Tolley, and Fellingham, CLXXXII; Rahman, Arnold, and Dabrowski.

21 Rahman and others.

22 Muhammad Faiz Fuady, - Haeruddin, and Mustofa Nitisupardjo, 'Pengaruh Pengeololaan Kualitas Air Terhadap Tingkat Kelulushidupan Dan Laju Pertumbuhan Udang Vaname Di PT Indokor

Vol. 16, No. 2, 2023
ISSN: 2654-7856 (Online)
ISSN: 1829-7935 (Print)

Here is a prediction of a successful harvest using Python

k Nearest Neighbor

```
classifier = KNeighborsClassifier(n_neighbors = 2)
classifier.fit(X_train, y_train)
y_pred = classifier.predict(X_test) acy_score
accuracy = accuracy_score(y_test, y_pred)
```

Decision Tree

```
classifier = DecisionTreeClassifier()
classifier = classifier.fit(X_train,y_train)
y_pred = classifier.predict(X_test)
print('Accuracy Score:', accuracy_score(y_test,y_pred))
```

Logistic Regression

```
model=LogisticRegression()
model.fit(X_train,y_train)
y_pred=model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
```

We do several variations of model analysis as follows:

1. Variables according to references

The accuracy of predicting crop success with machine learning according to reference variables results in different accuracy. The accuracy with the highest value comes from the Decision Tree model. The kNN and Logistic Regression’s accuracy result almost the same range.

Bangun Desa, Yogyakarta’, *Management of Aquatic Resources Journal*, 2.4 (2013), 155–62 <https://doi.org/10.14710/marj.v2i4.4279>.

Table 2. % Acuracy using variables that influenced the success of shrimp harvest from previous studies

No.	Variables Determining Crop Success	kNN	DT	LR
1.	NH_4^+ and salinity	77	86	75
2.	pH and salinity	70	73	75
3.	pH, salinity, NH_4^+, NOx	71	88	74

The accuracy of the Decision Tree gives the best results. The variables with the lowest accuracy are pH and salinity; it's quite difference from the references that will be analyzed further.

2. Are pH and salinity variables that can predict crop yield?

Based on references above, pH and salinity are critical factors determining the harvest's success. However, judging by the model, the accuracy is between 70-88%. After iterating again, accuracy will be better without using pH and salinity. The result was different than the reference that mentioned pH and salinity are critical factor to determine harvest's success[23]

Table 3. Accuracy Prediction With Variables From References

No.	Variables Determining Crop Success	kNN	DT	LR
From Previous Study	pH, Salinity, NH_4^+, NOx	71	88	74
2.	pH, NH_4^+, NOx	80	88	75
3.	NH_4^+, NOx	87	92	76

3. How does the ratio of adverse bacteria affect the accuracy of the model?

[23] Rahman and others.

Vol. 16, No. 2, 2023
ISSN: 2654-7856 (Online)
ISSN: 1829-7935 (Print)

The ratio of adverse bacteria is one of the measurement results in PT FEI. We also investigate how the adverse bacteria ratio affects the harvest's success.

Table 4. Accuracy Prediction With Ratio of Adverse Bacteria Variables

No.	Variables Determining Crop Success	kNN	DT	LR
From Previous Study	pH, Salinity, NH_4^+, NOx	71	88	74
2.	The ratio of adverse bacteria, pH, NH_4^+, NOx	75	90	75
3.	The ratio of adverse bacteria, NH_4^+, NOx	89	95	76
4.	The ratio of adverse bacteria, NH_4^+	91	96	75
5.	The ratio of adverse bacteria, NOx^+	89	94	75

The analysis results show that the adverse bacteria ratio affects the model's accuracy, especially for kNN and DT. The combination that provides the best accuracy is the ratio of adverse bacteria, and NH_4^+. This result is somewhat different from the reference, which states that salinity and pH are variables that affect successful harvest.

From the model analysis results, pH and salinity provide less accuracy, while other variables, such as the ratio of adverse bacteria, NH_4^+, and NO_x, have a positive influence. We conducted a fishbone analysis to look for potential root causes,

a. Salinity and pH are the results of internal measurements. External labs made other measurements, such as NOx, NH_4^+, and adverse bacteria ratios. Although internal and external measurement are derived from the same sample.
b. Internal logging is still not tidy, with many input errors and inconsistent recording. PT FEI have not utilized the data yet. Data from internal and external are only for record.

Analysis of process improvements for data utilization

The potential root causes of data utilization is in-line with the observation. From the observation and interview at PT FEI as well as looking at the results of data analysis they need process improvement so that data can be utilized as well as possible, especially for performance improvement and predict the successful harvest. This condition is because there is no awareness of using existing data. From the dmachine learning approach results, we can know that crop yields can be used to predict crop success.

Vol. 16, No. 2, 2023
ISSN: 2654-7856 (Online)
ISSN: 1829-7935 (Print)

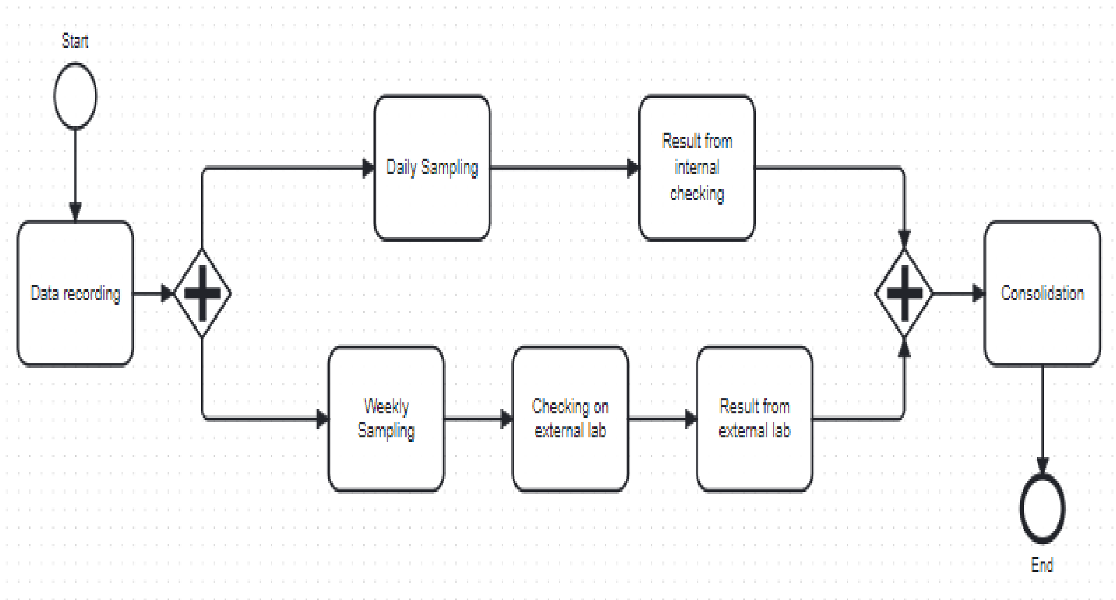

Figure 1. Current data utilization

Using the fishbones analysis approach, the factor that causes the data not to be utilized properly is that there is no process for utilizing the data to help the decision-making process. Apart from the absence of a process, PT FEI does not have tools to facilitate work. The following is an improvement process for data utilization so that it can help the overall final result.

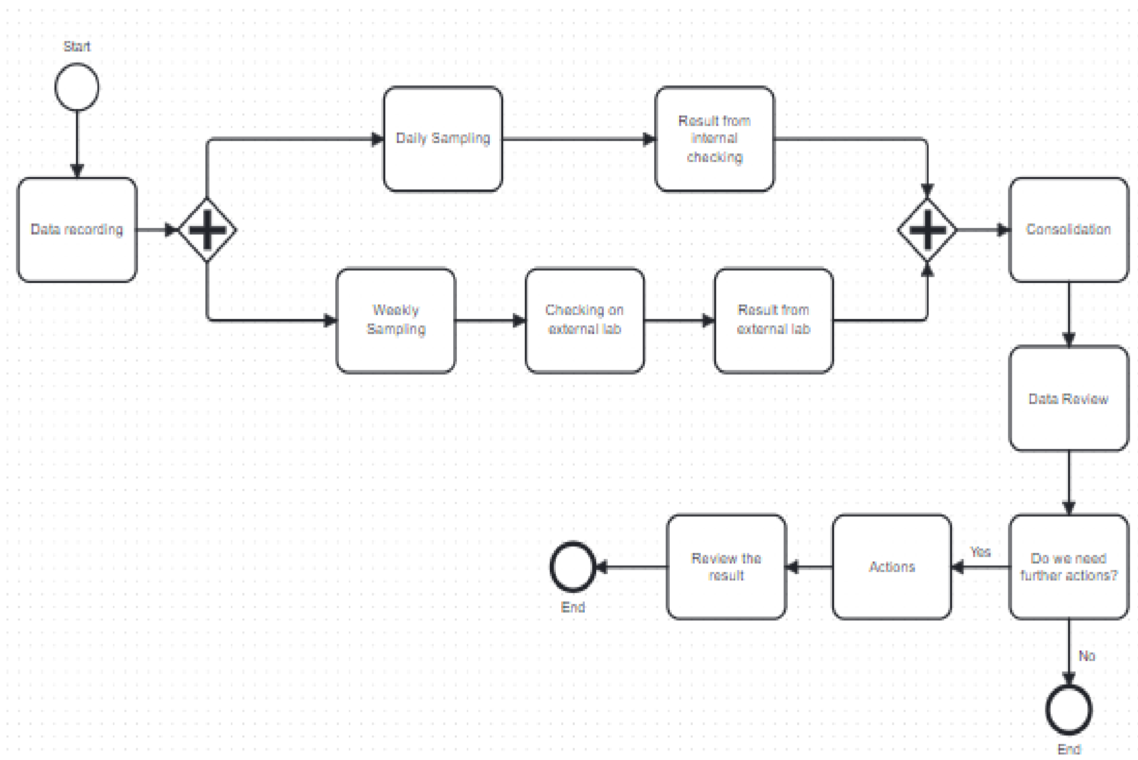

Figure 2. Process of improving data utilization

CONCLUSION

Machine Learning with k-Nearest Neighbour models, Decision Trees, and Logistic Regression can be used to predict shrimp harvest success. The Decision Tree model provides the best accuracy results, and the accuracy results of Regression Logistics are the lowest. The critical thing in machine learning is the data analysed. The quality of the data used affects the accuracy of the results. Productivity can be increased by utilized data optimally as the guidance for the decision.

BIBLIOGRAPHY

Alima Nostalia Suseno, Dewi, Buyung Purnomo Waluyo, Sugeng Rahardjo, Djoko Surahmat, Bambang Supriyadi, Dan Bowo Priono, and others, ‘Analisis Faktor Produksi Budidaya Udang Vannamei (Litopenaeus Vannamei) Di Tambak HDPE (High Density Polyethilene) Pulokerto Pasuruan’, *Balitbang Kkp*, 19 (2021), 99–104

Bataller, Erick, and Andrew D. Boghen, ‘Elimination of the Gill Worm Urastoma Cyprinae (Graff) from the Eastern Oyster Crassostrea Virginica (Gmelin) Using Different Salinity-Temperature Combinations’, *Aquaculture*, 182.3–4 (2000), 199–208 <https://doi.org/10.1016/S0044-8486(99)00281-1>

BPS, *BPS 2019, Statistik Indonesia 2019*, 2019

———, *BPS 2020, Statistik Indonesia 2020*, 2020,

Buentello, Alejandro J; Gatlin, Delbert, ‘Effects of Water Temperature and Dissolved Oxygen on Daily Feed’, *Aquaculture*, 182.2000 (2000), 339–52

Buitinck, Lars, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, and others, ‘API Design for Machine Learning Software: Experiences from the Scikit-Learn Project’, 2013, 1–15

Fuady, Muhammad Faiz, - Haeruddin, and Mustofa Nitisupardjo, ‘Pengaruh Pengeololaan Kualitas Air Terhadap Tingkat Kelulushidupan Dan Laju Pertumbuhan Udang Vaname Di PT Indokor Bangun Desa, Yogyakarta’, *Management of Aquatic Resources Journal*, 2.4 (2013), 155–62 <https://doi.org/10.14710/marj.v2i4.4279>

Gao, Huan, Jie Kong, Zhanjun Li, Guangxia Xiao, and Xianhong Meng, ‘Quantitative Analysis of Temperature, Salinity and PH on WSSV Proliferation in Chinese Shrimp Fenneropenaeus Chinensis by Real-Time PCR’, *Aquaculture*, 312.1–4 (2011), 26–31 <https://doi.org/10.1016/j.aquaculture.2010.12.022>

Gladju, J., Biju Sam Kamalam, and A. Kanagaraj, ‘Applications of Data Mining and Machine Learning Framework in Aquaculture and Fisheries: A Review’, *Smart Agricultural Technology*, 2.April (2022), 100061 <https://doi.org/10.1016/j.atech.2022.100061>

Grus, Joel, *Data Science from Scratch*, *O’Reilly*, 1st edn (O’Reilly, 2015), MDXLII

Hartanto, Nono, ‘SOP Pembesaran Udang Vaname Di Tambak Milenial’, 2020

Heizer, Jay; Render, Barry, *Operation Management Sustainability and Supply Chain*

Management, *Pearson*, 2017, XII

IBM, 'What Is Logistic Regression?', *IBM.Com*, 2022, p. 1 <https://www.ibm.com/topics/logistic-regression#:~:text=Logistic regression estimates the probability,bounded between 0 and 1.>

Kenett, Ron S; Redman, Thomas C, 'The Real Work of Data Science', *Forbes*, 204.1 (2021), 52–54 <https://doi.org/10.1002/9781119570790.ch1>

Lawrence, Tim, 'Exploiting Big Data and Analytics to Improve Productivity in Manufacturing', *Manufaturing Digital*, May 2020

Mahendra, Arya Bima, 'Jadi Daerah Dengan Tambak Udang Vaname Terluas Di Bangka Belitung, Sayang Belum Bisa Sumbangkan PAD - Bangkapos.Com', *BangkaPos*, 2022

Prastianti, Atikah Indri, 'Faktor-Faktor Yang Mempengaruhi Produksi Udang Vannamei (Litopeneaus Vannamei) Di Desa Pantai Bahagia, Kecamatan Muara Gembong' (Universitas Islam Negeri Syarif Hidayatullah, 2021)

Pujowati, Yenik, Putri Ari Saruhun Hasibuan, and Sucahyo Tri Budiono, 'Analisis Dampak Covid-19 Terhadap Pendapatan UMKM (Usaha Mikro, Kecil Menengah) Di Kabupaten Nganjuk', *Jurnal Pamator : Jurnal Ilmiah Universitas Trunojoyo*, 15.1 (2022), 100–112 <https://doi.org/10.21107/pamator.v15i1.13922>

R, Olshen, 'Sckitlearn Decision Trees', *Sckit-Learn Developers*, 2009 <https://scikit-learn.org/stable/modules/tree.html>

Rahman, Ashfaqur, Stuart Arnold, and Joel Janek Dabrowski, 'Identification of Variables Affecting Production Outcome in Prawn Ponds: A Machine Learning Approach', *Computers and Electronics in Agriculture*, 156 (2019), 618–26 <https://doi.org/10.1016/j.compag.2018.12.015>

Rahman, Ashfaqur, Mingze Xi, Joel Janek Dabrowski, John McCulloch, Stuart Arnold, Mashud Rana, and others, 'An Integrated Framework of Sensing, Machine Learning, and Augmented Reality for Aquaculture Prawn Farm Management', *Aquacultural Engineering*, 95.July (2021), 102192 <https://doi.org/10.1016/j.aquaeng.2021.102192>

Shafi, Adam, 'K-Nearest Neighbors (KNN) Classification with An Overview of K-Nearest Neighbors', February 2023

Shahriar, Md Sumon, Ashfaqur Rahman, and John McCulloch, 'Predicting Shellfish Farm Closures Using Time Series Classification for Aquaculture Decision Support', *Computers and Electronics in Agriculture*, 102 (2014), 85–97 <https://doi.org/10.1016/j.compag.2014.01.011>

Sun, Ming, Xiaofen Yang, and Yinggang Xie, 'Deep Learning in Aquaculture: A Review', *Journal of Computers*, 31.1 (2020), 294–319 <https://doi.org/10.3966/199115992020023101028>

Whiting, David G, H Dennis Tolley, and Gilbert W Fellingham, *An Empirical Bayes*

Vol. 16, No. 2, 2023
ISSN: 2654-7856 (Online)
ISSN: 1829-7935 (Print)

Procedure for Adaptive Forecasting of Shrimp Yield, *Aquaculture*, 2000, CLXXXII

Zhao, Shili, Song Zhang, Jincun Liu, He Wang, Jia Zhu, Daoliang Li, and others, 'Application of Machine Learning in Intelligent Fish Aquaculture: A Review', *Aquaculture*, 540.March (2021), 736724 <https://doi.org/10.1016/j.aquaculture.2021.736724>

Zhu, Jinyong, 'Big Data : How Much Data Is Big Data ? What Is Big Data ?', 2022