한국컴퓨터정보학회 동계학술대회 논문집 제31권 제1호 (2023. 1)

# 인공지능 기반(ML) 양식장 관리시스템 개발을 위한 수질 데이터 분석

심 현*, 심흥섭O
*순천대학교 산학협력단,
O동양대학교 컴퓨터군사학과

# Water quality data analysis for development of artificial intelligence-based fish farm management system

Hyun Sim*, Heung Sup SimO
*Industry-Academic Cooperation Foundation, Suncheon National University,
ODept. of Computer and Military Affairs, Dongyang University

● 요 약 ●

양식장에서 최적의 생육환경을 유지할 수 있는 제어시스템 개발을 위해 수질에 영향을 미치는 요인들의 상관관계 분석을 위한 머신러닝 모델을 개발하고자 한다. 데이터간의 상관관계 분석 및 예측모델 생성을 위해 알고리즘의 결정계수와 MSE, RMSE 등의 수치를 통하여 데이터의 적합성을 검증하고자 한다.

**키워드:** 수질 데이터분석, 정확도, 데이터셋, 결측값

## I. 서론

최근 양식업은 청년층 감소와 어촌의 고령화 등으로 인한 인력난 문제와 경제성 저하, 불법 어선의 침입, 수산 재해 등으로 재산 피해가 문제가 대두되고 있다[1]. 가장 많이 보급된 새우양식은 과거에는 주기적인 환수를 통해 수질을 유지하는 방식이 대부분으로 배출수로 인해 연안 환경오염과 수평감염의 주된 경로로 문제점이 제기되었다. 국내 새우의 생산량은 대부분 양식을 통한 생산이며 국내 새우 소비량에 비해 생산량은 10% 밖에 되지 않으며 대부분 해외에서 수입하고 있다. 이러한 양식업의 문제점인 환경오염을 해결하기 위해서 바이오폴락 양식 기술을 활용하고 있으나 줄어드는 어업 인구수로 인하여 새우 생산량은 줄어들고 있으며 어업인의 연령이 고령화 되어 실질적으로 양식업의 유지가 위기에 처해있다[2-3]. 이러한 문제점을 해결하기 위해 본 연구에서는 AI 기반의 스마트 바이오폴락 양식 기술을 활용하여 양식장을 자동으로 제어 관리하는 시스템을 개발하는 것이 목적이며 이를 위해서 양식장 수질 환경에 대한 분석과 예측 모델이 필수적이다. 최근 다양한 인공지능 기술을 활용하여 데이터 기반의 최적화를 통한 양식 자동화 기술이 활발히 연구되어 적용되고 있다[8-11]. 이러한 예측 모델을 설계하는데 추가적으로 고려해야 할 사항들은 많이 있지만 특히 비용적인 문제를 고려하지 않을 수 없다. 육상해수양식장의 다양한 전기장비 및 수질유지 약품 등 운영비용을 절감하기 위해, 수질데이터 분석을 통한 자율제어시스템이 필요하다[4-7]. 바이오폴락의 경우 미생물을 이용하기 위해서 지속적인 산소공급을 하게 되고 이로 인한 전력 소모가 심하기 때문에 이러한 문제해결을 위해서는 최적의 생육환경을 찾아내어 최소한의 전력을 사용할 필요성이 제기된다. 최적의 생육환경 제어시스템 개발을 위한 수질 데이터간의 상관관계 분석 및 예측모델 생성을 위한 절차는 데이터 수집 및 처리, R을 활용한 관계분석 및 회귀모델 평가, 도구를 활용한 회귀모델 평가 및 데이터 예측, 분석결과 및 예측에 대한 결과를 도출한다. 추가적으로 본 연구에서 제시하는 양식장은 밀폐된 환경에서 운영되는 양식장 구축을 연구하기 때문에 추후에 공기질의 상태와 수질과의 연관관계를 분석할 필요가 있다. 이러한 다양한 문제점을 해결하기 위해서 수질의 다양한 요인들의 데이터를 집적하고 이를 상관관계를 분석하고, 또한 공기질의 데이터를 집적하여 수질의 상관관계를 분석 실시간 예측함으로써 효율적으로 양식장을 관리할 수 있는 자동제어 시스템을 개발하고자 한다. 이를 위해 본 연구에서는 새우 양식을 위한 실제 수조 사이트를 구축하고 수질과 관련된 요인들의 상관관계 분석을 위한 머신러닝 모델을 개발하고자 한다. 최종적으로 AI기반의 머신러닝기술의 적용을 통해 효율적으로 양식장 생장환경정보, 생육환경 정보, 기후 변화에 따른 대응 시스템 등을 구축하여 양식에 필요한 수질 기준에 맞춰 수질 제어를 위한 장비 관제 및 사료, 약품 투여 등을 제어할 수 있는 AI 기반 수질 자율제어 시스템을 개발한다.

본 논문의 구성은 다음과 같다. 2장에서는 바이오폴락, 의사결정트리, 시계열분석에 대해서 알아보고, 3장에서는 AI 기반 양식장 수질 예측 기술에 대해서 알아보고, 4장에서는 결론 및 향후 연구과제를

제시하면 끝을 맺는다.

## II. 관련 연구

### 2.1 바이오폴락

바이오 폴락기술이란 광합성 및 타가영양 세균이 유기탄소를 이용하여 독성의 암모니아를 세균 단백질로 직접 동화시키며 증식된 세균은 미세조류, 원생동물 등 다른 미생 동물 및 미세 유기물 등과 결합하여 폴락(무생물층)을 형성하며 이것이 사육생물의 먹이가 된다. 물을 교환하지 않거나 최소화하기 때문에 배출 수에 의한 환경오염을 줄이고 전염병을 차단한다. 또한 사육밀도를 높이고 사육생물에 의한 폴락 재섭식을 통해 사료효율을 높일 수 있고, 항생제 및 인체에 해로운 약품을 쓰지 않아 친환경 새우 양식이 가능하다.

### 2.2 의사결정트리

의사결정나무(decision tree)는 어떤 항목에 대한 관측값과 목표값을 연결시켜주는 예측 모델이다. 트리 모델 중 목표변수가 유한한 수의 값을 가지는 것을 분류 트리라 한다. 이 트리구조에서 잎은 클래스 라벨을 나타내고 가지는 클래스라벨과 관련 있는 특징들의 논리곱을 나타내며, 결정 트리 중 목표 변수가 연속하는 값, 일반적으로 실수를 가지는 것은 회귀 트리라고 한다.

### 2.2 인공지능 모델(시계열분석)

시계열 데이터는 머신러닝 분야에서 관심이 많은 데이터 영역이다. 실제 세상에서는 시간적 요소가 중요한 데이터들이 많은 케이스가 많기 때문이다. 시계열 데이터는 시간 변수와 한 개 혹은 여러 개의 변수들로 구성되어 있으며, 시간과 한 개의 변수로 데이터가 구성되어 있을 땐 분석이 비교적 쉽다. 하지만, 시간과 여러 개의 변수로 데이터가 구성되어 있으면 분석이 어렵다. 시계열 데이터가 아닌 데이터에 대한 다변수 분석과 비교했을 때도 다변수 시계열 데이터는 분석이 더 어렵다. 단변량 시계열 데이터를 예측하는 방법으로는 머신 러닝이나 딥러닝 방법보다는 전통적인 ARIMA, ETS 방법이 효과적이나, 탐색적 데이터 마이닝을 통한 회귀분석 중심의 모델링 알고리즘을 연구하고자 한다.

## III. AI 기반 양식장 수질 예측 기술

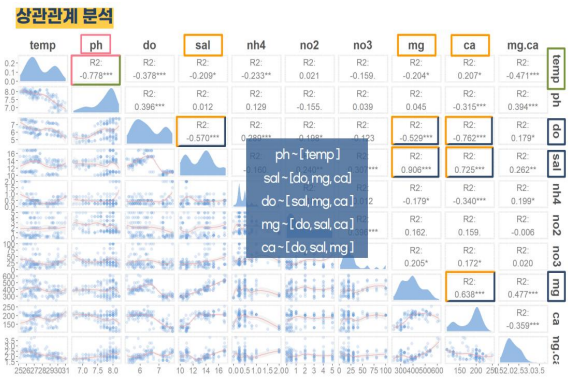
본 장에서는 AI 기반 양식장 수질 예측 기술에 대해 설명한다.

본 연구는 7월 ~ 12월까지 수집한 데이터로, 수질 센서를 통해서 취득한 값을 측정하는 방식이다. 양식수조에 수질센서를 부착하여 데이터를 수집하고, 수집된 센서 데이터를 실시간으로 모니터링하며 시계열 데이터 기반 머신러닝을 사용하여 수질을 예측하여 수조에 발생하는 문제 상황을 대비할 수 있도록 수조환경을 제어하는 시스템을 구현한다. 수질 센서를 통해서 염도, 탁도, pH, Do, 수온 등 정보를 측정하고 해당 측정 값들을 모니터로 실시간을 확인할 수 있도록 한다. Fig. 1은 실제 구축 운영하고 있는 수조 사이트 사진이다. Fig. 2는 수질 센서로 염도, 탁도, pH, Do, 온도 등의 정보를 측정한다.

Fig. 1. Fish farm site

Fig. 2. Water quality sensor

### 3.1 수질 요인간의 상관관계 분석

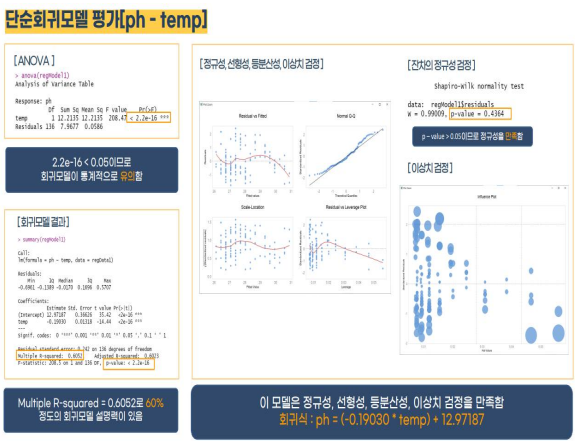



-데이터의 종류 : 엑셀, csv 데이터, 10개 정도의 관측항목 구성
-데이터 프레임 형태 : 정형데이터

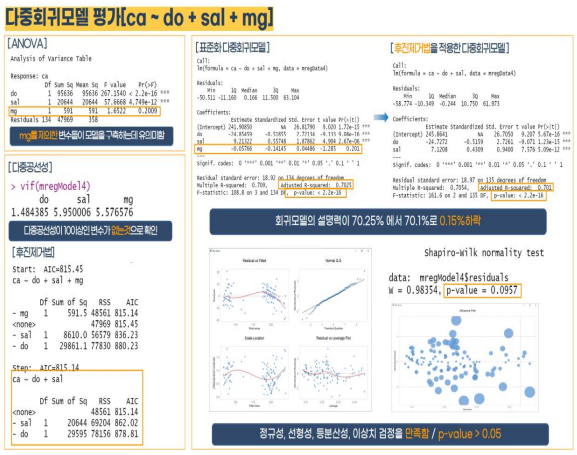
### 3.2 회귀모델 평가

관측항목의 독립변수와 종속변수의 지정에 따른 데이터의 결정계수가 통계적으로 유의하게 도출되고, 정규성, 정상성, 등분산성, 선형성 등의 잔차분석 모형의 가정이 유의하다고 판단되므로 오차에 대한 가정들의 성립 관련하여 검증할수 있다.

-단순회귀 분석 모델링.(일부)

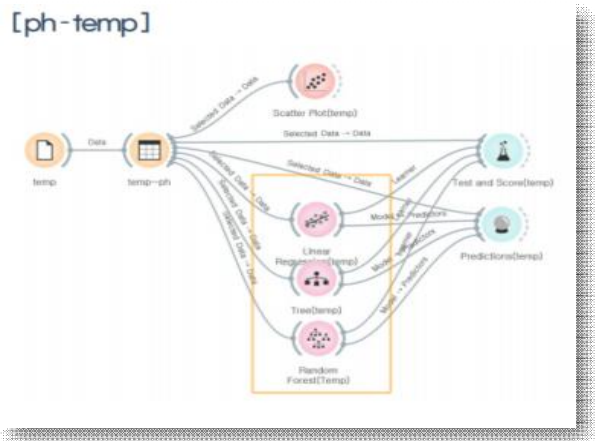


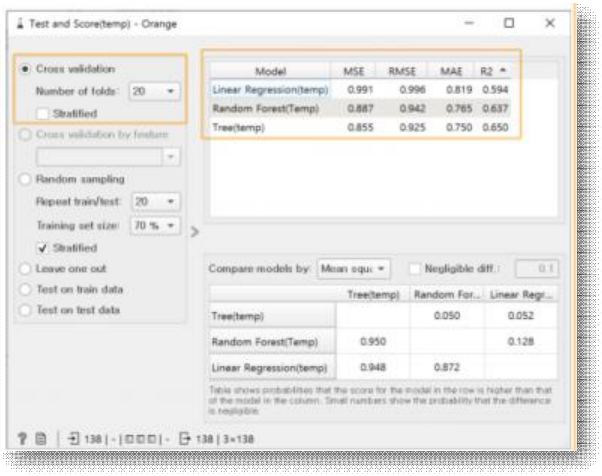

-다중회귀분석 모델링(일부)




## 3.3 Orange를 이용한 실시간 수질 예측 인공지능 모델

R과 함께, ML분석 도구로 오렌지 데이터마이닝(3.34.0)를 사용하여, 회귀분석 ML모델(수치 예측), 의사결정트리와 랜덤 포레스트의 3가지 ML모델을 통하여 검증하였다.

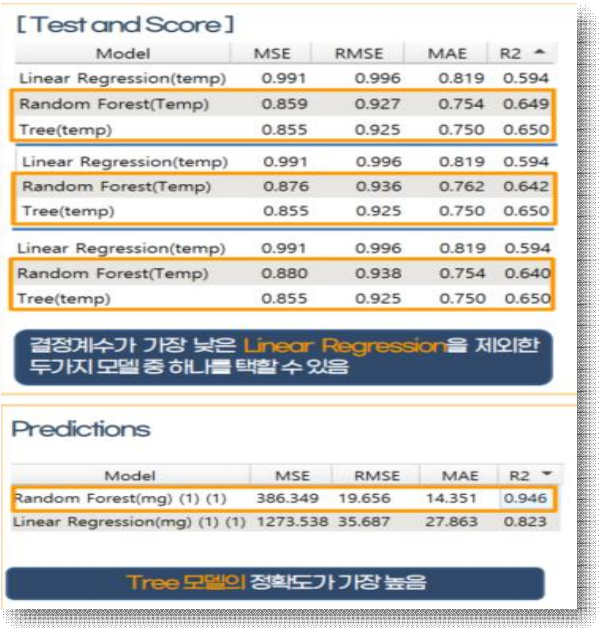



## 3.4 모델 평가 검증 방안

데이터의 셋의 환경에 따라서, 회귀분석(0.823)보다는 랜덤포레스트(0.946)의 예측 정확도가 훨씬 높게 나왔다.

따라서, 온도, 수온, pH, Do, 탁도의 기본조건과 이 양식장의 구조에서는 제한적으로 랜덤 포레스트의 머신러닝 모델을통하여 수치를 예측하는 것이 회귀분석의 예측과 의사 결정 트리의 예측치보다 정확하다라고 볼수 있다.(양식환경에 따른 오차를 배제시)




# IV. 결론 및 향후 연구

본 연구에서는 양식장 수질을 예측하기 위한 기존 연구들이 직접 사람들이 수동으로 수질 데이터 값을 측정하는 전통적 양식장 관리 방식으로 인해 다수의 노동력과 많은 노동시간이 요구되는 문제점을 개선한다. 또한 과도한 개체 및 사료의 투입으로 수질 환경을 악화시키고 질병에도 취약하여 폐사율이 증가하고 경제성이 떨어지는 등

전주기에서 악순환을 유발하는 문제점이 있다. 이러한 문제를 해결하기 위해서 본 연구에서는 AI기반의 스마트 양식장을 구현하였다.

이를 위해 본 연구에서는 머신 러닝 기법을 중심으로 수치 예측의 방식의 랜덤포레스트, 의사결정트리, 회귀분석의 중에 효과적인 랜덤 포래스트 분석을 이용하여 온도, 수온, pH, Do, 탁도 등의 기본 수질 데이터를 통해 시간 단위의 수질을 높은 정확도록 예측하는 기법을 제안하고 이를 통한 최적의 수질 관리를 위한 제어값을 제공하는 양식장 수질 관리 시스템으로 개선시킬 예정이다. 본 연구를 바탕으로 밀폐형 양식장에서 고려해야 하는 공기질과 수질의 상호관계를 분석하여 신뢰도 높은 예측 Modeling 구축을 위하여 추가 연구가 진행중이다.

## REFERENCES

[1] Song JH. “A study on development process of enterprise-type business in fish aquaculture: case by yellowtail aquaculture in Japan”, KSFBA, 36, 1, 139~153. 2005.

[2] FAO. “The state world fisheries and aquaculture 2018”. FAO, 1~227. 2018.

[3] Korean Statistical Information Service (KOSIS). “Agriculture, forestry and fisheries survey.Census of agriculture, forestry and fisheries”.*Retrieved from https://kosis.kr on May 12. 2019.

[4] N.K.Asmel, R.R.Al-Nima, F.I.Mohammed, A.M. Saadi, and A.A. Ganiyu, "Forecasting Effluent Turbidity and pH In Jar Test Using Radial Basis Neural Network," Towards a Sustainable Water Future, pp. 361-370, Jan. 2021.

[5] D.H.Ryu, T.W.Choi, "Development of the Smart Device for Real time Water Quality Monitoring,"The Journal of the Korea institute of electronic communication sciences, Vol. 14, No. 4, pp. 723-728,Aug. 2019.

[6] S.J.Lee, J.W.Kim, "Enhancement of Water Quality Prediction System for Scientific Management of Water Quality," The Korean Society of Agricultural Engineer, Vol. 2019, No. 0, pp. 220-220, Oct. 2019.

[7] K.P.Singh, A.Basant, A.Malik, G.Jain, "Artificial neural network modeling of the river water quality .A case study," Ecological Modelling, Vol. 220, No. 6, pp. 888-895, Mar. 2009.

[8] AKVA Hompage, https://www.akvagroup.com/

[9] Aqua Manager Hompage, https://www.aquamanager.com/

[10] Innovasea Hompage, https://rtaqua.com/

[11] Aquasend Hompage, https://www.aquasend.com/