AI SYSTEM FOR DETECTING OPTIMAL CONDITIONS FOR AQUACULTURE

Penka Georgieva, Burgas Free University, penka.georgieva@bfu.bg
Maria Miroslavova Daskalova, maria.mdaskalova@gmail.com

Abstract: This study presents the development of an artificial intelligence–based system for the analysis and classification of marine environmental. The test dataset includes key physicochemical parameters such as water temperature, salinity, dissolved oxygen, chlorophyll-a concentration, and pH. Data preprocessing involved interpolation of short missing values, removal of longer gaps, and aggregation to daily temporal resolution. A one-dimensional convolutional neural network (1D-CNN) was applied to analyze temporal patterns and perform binary classification of marine conditions as favorable or unfavorable. The results demonstrate that the proposed approach can effectively extract patterns from marine time series and identify periods with favorable environmental conditions.

Keywords: *artificial intelligence, machine learning, time series analysis, Black Sea, aquaculture, 1D-CNN.*

AI СИСТЕМА ЗА ОТКРИВАНЕ НА ОПТИМАЛНИ УСЛОВИЯ ЗА АКВАКУЛТУРИ

Пенка Георгиева, Бургаски Свободен Университет, penka.georgieva@bfu.bg
Мария Мирославова Даскалова, maria.mdaskalova@gmail.com

Абстракт: Това изследване представя разработка на система, базирана на изкуствен интелект за анализ и класификация на морската среда. Тестовият набор от данни включва ключови физикохимични параметри като температура на водата, соленост, разтворен кислород, концентрация на хлорофил-а и pH. Предварителната обработка на данните включва интерполация на кратки липсващи стойности, премахване на по-дълги пропуски и агрегиране до дневна времева резолюция. За анализ на времеви модели и извършване на двоична класификация на морските условия като благоприятни или неблагоприятни е приложена едномерна конволюционна невронна мрежа (1D-CNN). Резултатите показват, че предложеният подход може ефективно да извлича модели от времеви редове и да идентифицира периоди с благоприятни условия на околната среда.

Ключови думи: *изкуствен интелект, машинно обучение, времеви серии, Черно море, аквакултури, 1D-CNN.*

1. Въведение

През последните години наблюдението и анализът на морската среда придобиват все по-голямо значение във връзка с климатичните промени, устойчивото управление на морските ресурси и развитието на синята икономика. Морските екосистеми, и по-специално крайбрежните зони, са чувствителни към промени във физико-химичните параметри на водата, което налага използването на съвременни подходи за обработка и интерпретация на големи обеми от сензорни данни.

Традиционните методи за анализ на морските условия често разчитат на статични прагове и експертни оценки, които трудно улавят сложната динамика и нелинейните зависимости във времевите редове. В този контекст методите на изкуствения интелект и машинното обучение предоставят възможност за по-прецизно моделиране на морските процеси, както и за автоматично откриване на подходящи или рискови условия въз основа на реални наблюдения [4][6].

В рамките на тази статия се представя разработката и прилагането на модулна AI система за анализ и класификация на морските условия, базирана на сензорни данни от българското Черноморие. Системата използва предварителна обработка на времеви серии, включително интерполация на липсващи стойности и агрегация, последвана от обучение на едномерна конволюционна невронна мрежа за бинарна класификация на условията. Резултатите се съхраняват в база данни и се визуализират с топлинни карти с цел улесняване на анализа и вземането на решения.

Основната цел е да се оцени приложимостта на изкуствения интелект за автоматично откриване на благоприятни условия въз основа на реални сензорни измервания, както и да се демонстрира потенциалът на подобни системи като инструмент за подпомагане на устойчивото управление на морските ресурси и аквакултурите в региона на Черно море.

2. Методология

Разработената система е фокусирана върху анализ на морските условия в крайбрежната зона на българското Черноморие чрез прилагане на методи на машинното обучение върху многопараметрични времеви серии. Данните са получени от Copernicus Marine Service[3] и DOORS[10] и съдържат дългосрочни наблюдения с дневна времева резолюция. Анализирани са физико-химични параметри, включително температура на морската вода (°C), соленост, разтворен кислород, концентрация на хлорофил-а и pH. За всяка изследвана локация данните са структурирани по параметри и обработени независимо преди обединяване във времевия модел.

Извличането на данните от файловете е реализирано чрез библиотеката xarray в Python среда. След зареждане на времевите серии, данните са конвертирани в pandas DataFrame формат с цел улесняване на последващата обработка. Първоначалният етап включва проверка за липсващи стойности и аномалии, като прекъсвания във времевите редове.

Обработката на липсващите стойности е извършена в два последователни етапа с цел запазване на статистическата структура на данните. Кратки пропуски с продължителност до три последователни дни са запълнени чрез линейна интерполация [1]. При по-дълги липсващи сегменти (над три дни) съответните времеви интервали са изключени от анализа, за да се избегне въвеждането на изкуствени зависимости. Този подход осигурява баланс между запазване на информация и ограничаване на систематични грешки.

За целите на класификационния модел е използвана бинарна схема за етикетиране на данните. Всеки времеви сегмент е класифициран като „благоприятен“ или „неблагоприятен“ въз основа на предварително дефинирани диапазони за оптимални условия за отглеждане на аквакултури. Класификацията се извършва чрез комбинирано условие, отчитащо едновременно всички включени параметри, което позволява по-реалистично отразяване на екологичната среда.

Разгледани са няколко подходящи модела за класификация на времеви данни: CNN, Random Forest, SVM, XGBoost, LSTM и CNN-LSTM. Random Forest и XGBoost са добре приложими при таблични данни, но не моделират директно времевата последователност на наблюденията. SVM е ефективен при класификационни задачи, но е по-ограничен при работа с многомерни времеви серии. LSTM е предназначен за анализ на времеви зависимости, но има по-сложна архитектура и по-висока изчислителна цена. CNN-LSTM съчетава двата подхода, но въвежда допълнителна сложност, която не е необходима за настоящата задача.

Класификацията е реализирана чрез едномерна конволюционна невронна мрежа (1D CNN) [9], имплементирана с TensorFlow/Keras. Този модел е избран поради способността му да извлича локални зависимости и характерни модели от многомерни времеви серии при сравнително ниска изчислителна сложност. Това го прави подходящ за анализ на последователните екологични измервания, използвани в настоящото изследване.

Архитектурата на невронната мрежа включва:

- Конволюционни слоеве (Conv1D) – извличат локални характеристики от входните последователности;
- ReLU активация – осигурява нелинейност и стабилност на обучението;
- MaxPooling слоеве – запазват размерността на изхода и акцентират върху най-значимите признаци;
- Dropout – добавя регуляризация чрез случайно изключване на неврони по време на обучението;
- Flatten слой – трансформира многомерния тензор в едномерна структура за плътните слоеве;
- Плътно-свързан изходен слой – отговаря за класификацията на прозореца.

Обучението на модела е извършено с функция на загуба Binary Cross-Entropy и оптимизатор Adam с начална learning rate стойност 0.001. Използван е batch size от 32 и 50

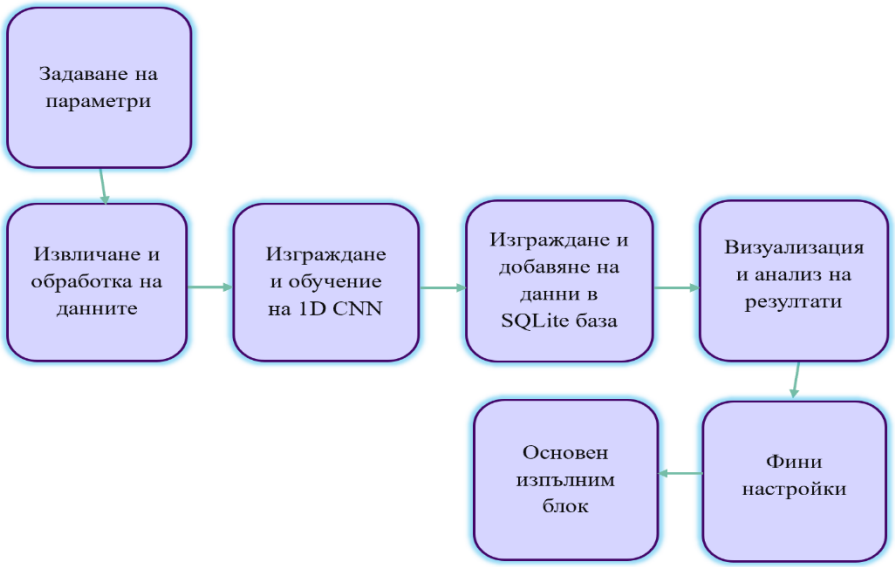
епохи. Данните са разделени на обучаващ, валидиращ и тестови набор, като 20% от общия обем е заделен за тестова оценка. За ограничаване на пренастройването (overfitting) е приложен Early Stopping механизъм с параметър patience = 10, базиран на валидиращата загуба.

Оценката се извършва чрез стандартни метрики за бинарна класификация: точност, прецизност, чувствителност и F1-мярка [2]. Този набор от метрики позволява комплексна оценка на способността на модела да разграничава благоприятни от неблагоприятни условия, особено в случаи на потенциален дисбаланс между класовете.

Данните се съхраняват в SQLite база данни като тя се допълва автоматично при всяко ново стартиране на процеса за обработка или моделиране, като се избягва презаписване чрез проверка за вече налични записи (по комбинация от параметър, дата и локация).

След първоначалното обучение се предвижда възможност за прилагане на фина настройка с новопостъпили данни. Тази процедура представлява целенасочено оптимизиране на вече обучената невронна мрежа чрез допълнителни цикли на обучение върху нова информация, без да се губи придобитото знание от първоначалната фаза.

Цялата система е реализирана в Python среда с използване на библиотеките NumPy, Pandas, Xarray и TensorFlow/Keras. На фигура 1 е представена блокова схема на работа на софтуерната система.



Фигура 1.Блокова схема - архитектура на софтуерната система.

3. Проведени тестове и резултати

В този раздел се представят резултатите от обучението и тестването на разработения модел, насочен към оценка на условията за отглеждане на черна морска мида (Mytilus

galloprovincialis) [7][8]. Избрани са пет основни екологични параметъра: температура, соленост, pH, хлорофил-a и разтворен кислород [5]. Изборът им се основава на биологичното им значение за физиологичното състояние, растежа и жизнеспособността на вида, както и на насоките на FAO[8] за оценка на условията при този вид аквакултура.

Приетите прагови стойности за дадената аквакултура са:

Параметър	**Минимална стойност**	**Максимална стойност**	**Единица**
Температура	10.0	20.0	°C
Соленост	15.0	19.0	PSU
pH	8.0	8.3	pH
Хлорофил-a	0.5	8.0	mg/m³
Разтворен кислород	166.0	320.0	mmol/m³

Таблица 1. Прагови стойности

Температурата влияе върху метаболизма и растежа, солеността – върху осморегулацията и физиологичното състояние, а pH – върху физиологичните процеси и формирането на черупката. Хлорофил-a се използва като индикатор за наличието на фитопланктон и хранителна база, докато разтвореният кислород е необходим за нормалните метаболитни процеси.

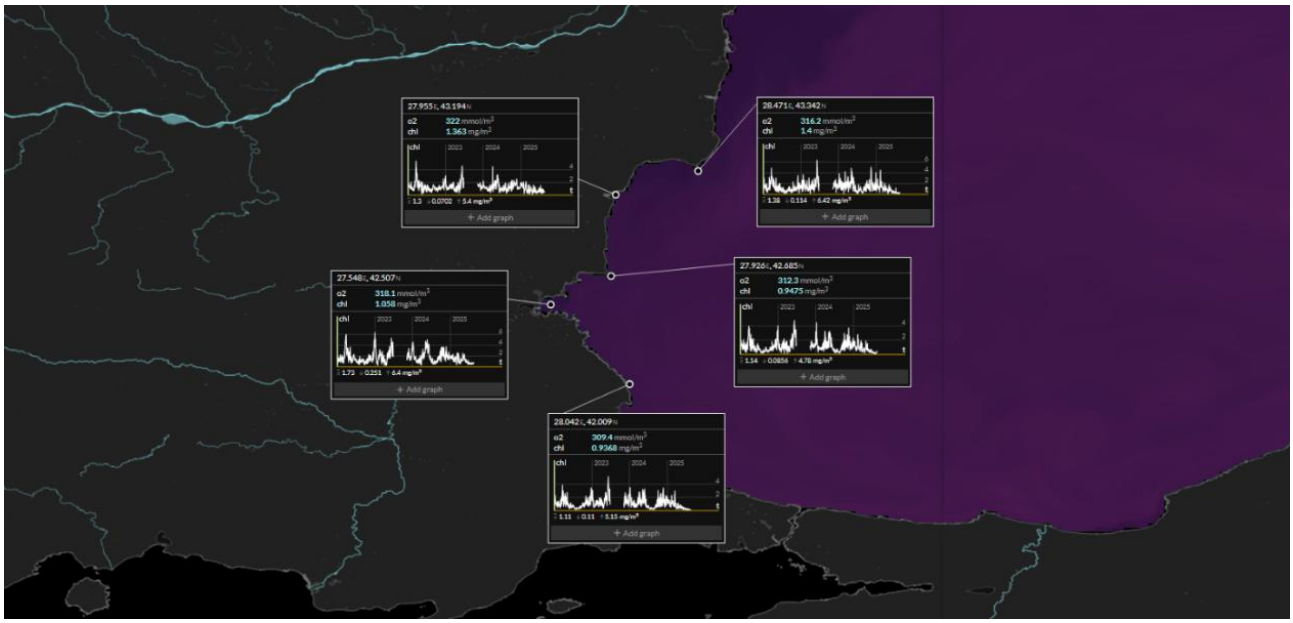
Въз основа на тези прагове условията се разделят на два класа. Благоприятни условия (клас 1) са налице, когато всички пет параметъра попадат в съответните допустими диапазони. Неблагоприятни условия (клас 0) се определят, когато поне един от параметрите е извън определения диапазон. По този начин етикетирането се извършва чрез обективен и еднозначен критерий, който впоследствие се използва за обучение на 1D CNN модела.

Черната морска мида (Mytilus galloprovincialis) е избрана като обект на изследване поради естественото ѝ разпространение в Черно море, адаптацията ѝ към местните условия и потенциала ѝ за устойчиво развитие на морските аквакултури.

Данните са събрани от 5 локации по българското Черноморие – Бургас, Варна, н. Емине, Резово и Шабла (фигура 2). Обхващат 365 последователни дневни наблюдения. Използван е времеви прозорец WINDOW = 1, поради което всяка входна извадка съдържа стойностите на петте параметъра за един ден.

Етикетите са генерирани автоматично чрез сравняване на стойностите с предварително определените прагове. Получените класове представляват целевата променлива за обучение и оценка на разработения модел.

С цел оптимизиране на обучението на модела е проведен експеримент за настройка на основните хиперпараметри. Анализирано е влиянието на различни стойности на параметрите като брой епохи, размер на партидата, скорост на обучение и брой филтри. Тестваните конфигурации за представени в Таблица 2.



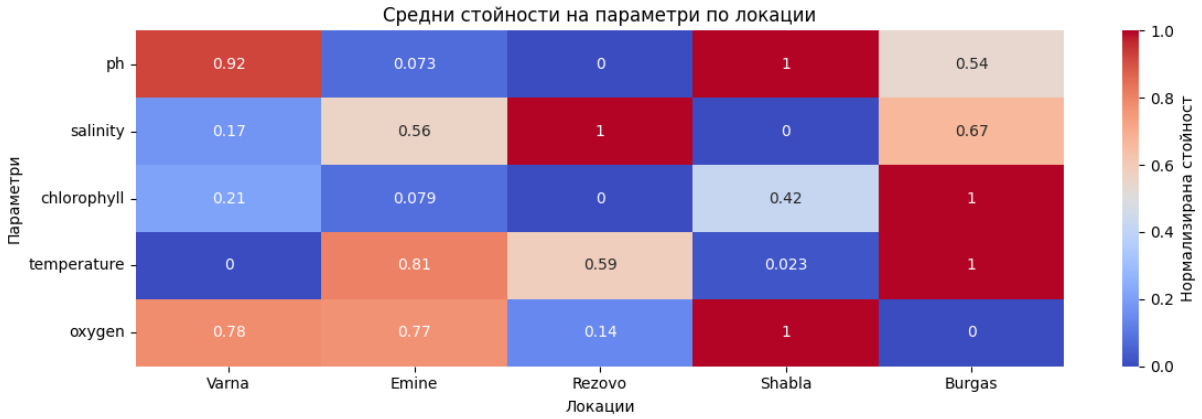
Фигура 2. Визуализация на локациите и част от параметрите в платформата Copernicus [3].

Параметър	**Тествани стойности**	**Най-добра стойност**	**Описание**
Брой епохи (Epochs)	25, 50, 75, 100	50 – 75	Брой итерации на обучение върху обучаващия набор
Размер на партидата (Batch size)	8, 16, 32, 64	8	Брой проби, обработвани едновременно при една стъпка на обучение
Скорост на обучение (Learning rate)	0.01, 0.001, 0.0001, 0.005	0.01	Стъпка на обновяване на теглата при оптимизация
Брой филтри (Conv1D filters)	16, 32, 64, 128	32	Брой филтри в конволюционните слоеве за извличане на характеристики

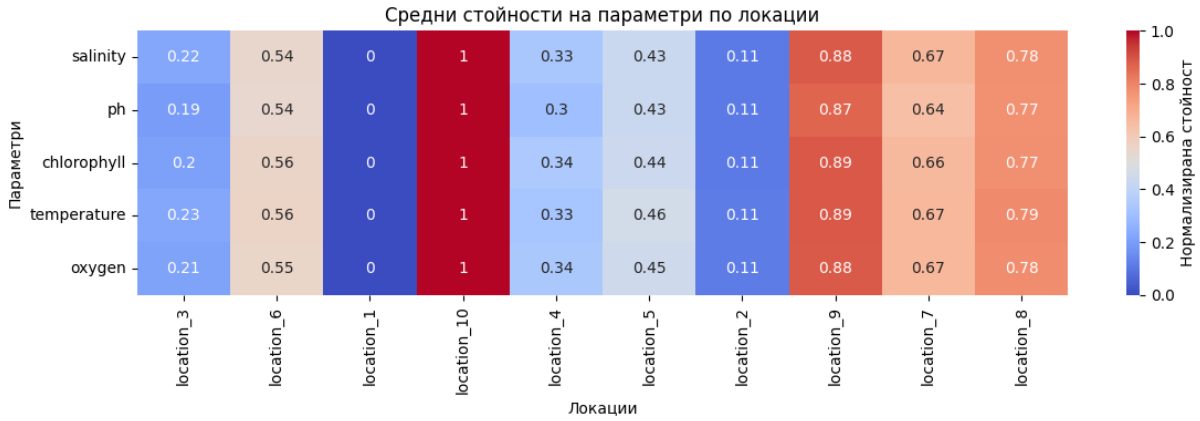
Таблица 2. Резултати от проведени тестове с различни настройки на хиперпараметрите.

Топлинните карти, даващи оценка на условията по локации на база средни стойности, изобразени след нормализация са показани на фигури 3 и 4, като на фигура 3 са визуализирани реални данни от измерени параметри по локации, докато фигура 4 илюстрира резултати, базирани на изкуствено генерирани данни и пространствени координати.

При прилагане на MinMax мащабиране е възможно отделни параметри да получат стойности, равни на 0 или 1, без действителните им измерени стойности да съвпадат с абсолютния минимум или максимум във физически смисъл. Това се обяснява с факта, че алгоритъмът нормализира спрямо най-ниската и най-високата налична стойност, в целия използван набор от данни, като се вземат предвид всички налични измервания във времето и по локации. Ако тези екстремни стойности са резултат от единични или редки наблюдения, те определят границите на скалата, което води до компресиране на останалите стойности и изместването им към краищата на интервала [0, 1], дори когато реално не достигат тези граници.



Фигура 3.Топлинна карта с резултати от реални локации и данни [1].



Фигура 4. Топлинна карта с резултати след тест с изкуствено генерирани данни [1].

Моделът показва надеждна способност за разграничаване на благоприятни и неблагоприятни морски условия. Получените резултати потвърждават, че предложената система може да извлича съществени зависимости от сложни многомерни времеви серии и да идентифицира периоди характеризиращи се с благоприятни екологични условия.

Получените данни се съхраняват в SQLite база данни, като преди тяхното записване се извършва проверка за наличие на дублирани записи (Таблица 3).

id	location	time	prediction
26	Varna	2024-01-27	0
27	Varna	2024-01-28	0
28	Varna	2024-01-29	0
29	Varna	2024-01-30	0
30	Varna	2024-01-31	1
31	Varna	2024-02-01	0
32	Varna	2024-02-02	0
33	Varna	2024-02-03	1
..	…	…	…

Таблица 3. Визуализация на част от базата данни.

4. Анализ на резултатите

Получените резултати показват, че използването на 1D CNN за анализ на времеви серии позволява ефективно откриване на зависимости между екологичните параметри и общото състояние на морската среда. Моделът демонстрира добра способност за извличане на локални времеви зависимости от последователните измервания и за използване на информацията от няколко параметъра при определяне на условията за отглеждане на черна морска мида.

В Таблица 4 са представени основните показатели за ефективността на модела, изчислени върху тестовия набор. Получената точност от 0.85 показва, че приблизително 85% от тестовите наблюдения са класифицирани правилно. При клас 0 (неблагоприятни условия) са постигнати прецизност 0.89, чувствителност 0.91 и F1-мярка 0.90, което показва много добро разпознаване на този клас. При клас 1 (благоприятни условия) стойностите са по-ниски – прецизност 0.67, чувствителност 0.62 и F1-мярка 0.64. Тази разлика следва да се разглежда във връзка с неравномерното разпределение на класовете в тестовия набор, съдържащ 1127 наблюдения от клас 0 и 320 от клас 1. Въпреки това претеглен F1-резултат от 0.85 показва добро общо представяне на модела.

	precision	recall	f1-score	support
0.0	0.89	0.91	0.90	1127
1.0	0.67	0.62	0.64	320
accuracy			0.85	1447
macro avg	0.78	0.77	0.77	1447
weighted avg	0.84	0.85	0.85	1447

Таблица 4. Общи метрики за всички локации.

Сравнено с подходи, основани единствено на фиксирани прагови стойности или анализ на отделни параметри, разработената система позволява едновременно разглеждане на множество екологични фактори. Това е съществено при анализа на морската среда, тъй като нейното състояние се определя от взаимодействието между различни физични, химични и биологични характеристики. Използването на 1D CNN позволява автоматично извличане на значими характеристики от входните времеви серии, без необходимост от ръчно дефиниране на всички зависимости между параметрите.

Използването на времеви прозорци позволява данните да бъдат разглеждани като последователности, а не като напълно независими измервания. Това създава възможност моделът да отчита краткосрочни изменения и зависимости в динамиката на морските параметри. При настоящата конфигурация обаче WINDOW = 1, поради което тази възможност е ограничена и моделът на практика използва информацията от еднодневно наблюдение. Увеличаването на времевия прозорец в бъдещи изследвания би позволило по-пълно изследване на времевата динамика.

Въпреки добрите резултати, системата има и определени ограничения. Наличието на липсващи стойности налага използването на интерполация и изключване на определени сегменти, което може да повлияе върху пълнотата на данните. Освен това моделът е обучен върху ограничен брой екологични параметри и географски локации, поради което прилагането му към други морски региони би изисквало допълнително обучение и валидиране. Бъдещи изследвания могат да включат допълнителни параметри, метеорологични данни и информация от дистанционни наблюдения, както и да изследват по-сложни архитектури като CNN-LSTM и трансформър-базирани модели.

Разработената система демонстрира потенциал за приложение при мониторинг и оценка на морската среда, както и като инструмент за подпомагане на решенията при планиране и управление на аквакултурни дейности.

Приносът на разработката се изразява в прилагането на 1D CNN за оценка на пригодността на морската среда за отглеждане на черна морска мида, чрез едновременно използване на множество океанографски параметри. Модулната архитектура позволява системата да бъде разширявана с допълнителни източници на данни, модели и видове аквакултури, създавайки основа за бъдещи AI-базирани решения в областта на устойчивите морски аквакултури.

5. Заключение

Настоящото изследване представя разработването на система за анализ и класификация на морски условия чрез използване на методи на изкуствения интелект и многопараметрични времеви серии. Използвайки данни от сензорни наблюдения в българското Черноморие, е създаден модел на база едномерна конволюционна невронна мрежа способен да идентифицира благоприятни и неблагоприятни екологични условия.

Резултатите показват, че предложената методология позволява ефективно извличане на зависимости между различни физико-химични параметри на морската среда и осигурява надеждна класификация на условията. Комбинацията от подходяща предварителна обработка на данните, времеви прозорци и дълбоко обучение се оказва ефективна стратегия за анализ на сложни екологични времеви серии.

Получените резултати потвърждават потенциала на методите на изкуствения интелект като инструмент за подпомагане на мониторинга и управлението на морските екосистеми. Подобни системи могат да подпомогнат развитието на устойчиви практики в аквакултурите и да допринесат за по-добро разбиране на динамиката на морската среда в условията на климатични промени.

6. Литература

[1] Даскалова М, AI Система за откриване на оптимални условия за аквакултури, Дипломна работа за степен Магистър по Изкуствен интелект, защитена на 11.09.2025г.
[2] Google Developers. Classification: Accuracy, recall, precision, and related metrics, 2024. https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall
[3] Copernicus Marine MyOcean Viewer . Copernicus.eu. 2024. https://data.marine.copernicus.eu/viewer/
[4] Yang X, Song Z, Liu J, Zhou C. Deep learning for smart fish farming: applications, opportunities and challenges. ResearchGate . 2020 Apr 6 https://www.researchgate.net/publication/340939333_Deep_learning_for_smart_fish_farming_applications_opportunities_and_challenges
[5] Christenser J, Galeazzi R, Mogensen L, Andersen J. Detection, Localization and Classification of Fish and Fish Species in Poor Conditions using Convolutional Neural Networks. https://www.researchgate.net/publication/329244872_Detection_Localization_and_Classification_of_Fish_and_Fish_Species_in_Poor_Conditions_using_Convolutional_Neural_Networks. 2018.
[6] Rather MA, Ahmad I, Shah A, Younis Ahmad Hajam, Amin A, Saba Khursheed, Exploring opportunities of Artificial Intelligence in aquaculture to meet increasing food demand. Food chemistry X. 2024 Jun 1;
[7] Verma DK, Singh S, Narendra Kumar Maurya, Kumar P, Rajesh Jayaswal. Important Water Quality Parameters in Aquaculture: An Overview. Agriculture and Environment . 2022 Mar 6;

https://www.researchgate.net/publication/362667844_Important_Water_Quality_Parameters_in_Aquaculture_An_Overview
[8] FAO - Mytilus galloprovincialis . Fao.org. 2024.
https://www.fao.org/fishery/docs/CDrom/aquaculture/I1129m/file/en/en_mediterraneanmussel.htm
[9] Jain A. Understanding the 1D Convolutional Layer in Deep Learning . Medium. 2025.
https://medium.com/@abhishekjainindore24/understanding-the-1d-convolutional-layer-in-deep-learning-7a4cb994c981
[10] xcube Viewer . Brockmann-consult.de. 2025.
https://doors.viewer.brockmann-consult.de/