

ARTYKUŁ PRZEGLĄDOWY

Uczenie maszynowe jako narzędzie wspierające rozwój produkcji rzepaku (*Brassica napus* L.)Machine learning as a tool supporting the development of rapeseed (*Brassica napus* L.) productionMarek Wójtowicz¹, Andrzej Wójtowicz^{2*}, Ilona Świerczyńska²**Streszczenie**

Rzepak (*Brassica napus* L.) jest jedną z najważniejszych roślin oleistych na świecie. Coraz większe znaczenie w jego produkcji zyskuje uczenie maszynowe, które umożliwia prognozowanie plonów, ocenę stanu roślin oraz wspomaganie decyzji agrotechnicznych. W pracy przedstawiono przegląd badań dotyczących wykorzystania różnych algorytmów – od klasycznych modeli, takich jak regresja liniowa, drzewa decyzyjne, las losowy czy maszyna wektorów nośnych, po bardziej zaawansowane, w tym sieci neuronowe. Omówiono zastosowania danych z teledetekcji, obrazowania zarejestrowane z pokładu bezzałogowych statków powietrznych i satelitów oraz czujników naziemnych do monitorowania chorób, stresów środowiskowych i potrzeb nawozowych. Podkreślono potencjał uczenia maszynowego w zwiększaniu efektywności i zrównoważeniu produkcji rzepaku, a także wyzwania związane z jakością danych, złożonością modeli i ich interpretacją.

Słowa kluczowe: uczenie maszynowe, rzepak (*Brassica napus* L.), prognozowanie plonów, diagnostyka chorób roślin, rolnictwo precyzyjne

Abstract

Oilseed rape (*Brassica napus* L.) is one of the most important oil crops worldwide. Machine learning methods are gaining increasing importance in oilseed rape production, enabling yield prediction, crop condition assessment, and support for agronomic decision-making. This paper presents a review of studies on the use of various algorithms – from classical models such as linear regression, decision trees, random forest, and support vector machine to more advanced ones, including neural networks. Applications of remote sensing data, unmanned aerial vehicles and satellite imagery, as well as ground-based sensors for monitoring diseases, environmental stresses, and nutrient requirements are discussed. The potential of machine learning to enhance efficiency and sustainability in rapeseed production is highlighted, along with challenges related to data quality, model complexity, and interpretability.

Keywords: machine learning, oilseed rape (*Brassica napus* L.), yield prediction, plant disease diagnostics, precision agriculture

¹Instytut Hodowli i Aklimatyzacji Roślin – Państwowy Instytut Badawczy
Radzików, 05-870 Błonie

²Instytut Ochrony Roślin – Państwowy Instytut Badawczy
ul. Władysława Węgorka 20, 60-318 Poznań

*corresponding author: a.wojtowicz@iorpib.poznan.pl

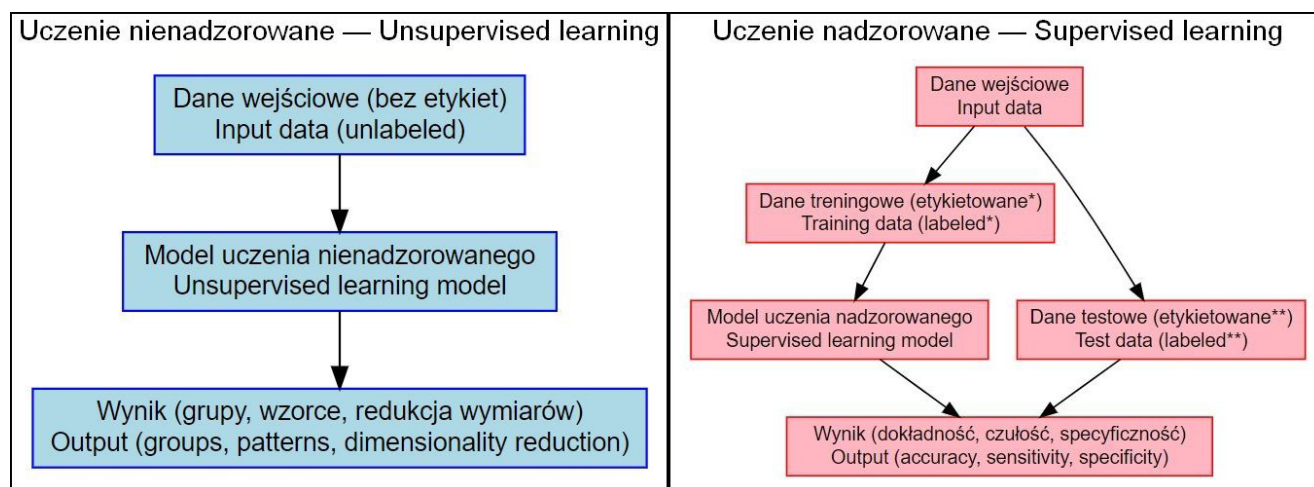
Wstęp / Introduction

Rzepak (*Brassica napus* L.) jest jedną z kluczowych roślin oleistych uprawianych w Europie i na świecie. Jest istotnym surowcem do produkcji oleju roślinnego, biopaliw oraz pasz dla zwierząt. Charakteryzuje się stałym trendem wzrostu produkcji i powierzchni uprawy. W roku 2023 powierzchnia uprawy rzepaku na świecie wynosiła około 43,5 miliona hektarów, a produkcja osiągnęła 91,9 miliona ton (FAOSTAT 2023). Wzrost zapotrzebowania na produkty pochodzące z rzepaku oraz zmieniające się warunki klimatyczne stawiają przed producentami nowe wyzwania związane z maksymalizacją plonów przy zachowaniu jakości nasion i zrównoważonego wykorzystania zasobów. W tym kontekście rośnie zainteresowanie nowoczesnymi narzędziami analitycznymi, w tym metodami uczenia maszynowego, które umożliwiają integrację i przetwarzanie dużych zbiorów danych rejestrowanych w warunkach laboratoryjnych i polowych.

Uczenie maszynowe polega na tym, że algorytmy realizują określone zadania na podstawie danych wejściowych bez konieczności wcześniejszego programowania reguł. Modele uczenia maszynowego można podzielić na dwa główne typy: nienadzorowane i nadzorowane. Modele uczenia nienadzorowanego nie korzystają z etykietowanych danych. Ich celem jest odkrywanie w danych ukrytych struktur, zależności lub grup. Algorytmy nienadzorowane przetwarzają dane w całości i samodzielnie identyfikują wzorce, takie jak grupowanie podobnych obserwacji (klasteryzacja) czy redukcja wymiarów, co pozwala lepiej zrozumieć charakter danych i znaleźć w nich ukryte informa-

cje. W przypadku modeli nadzorowanych proces uczenia polega na tym, że najpierw zbierane są dane etykietowane, które następnie dzieli się na zbiór treningowy i testowy. W procesie uczenia nadzorowanego etykiety pełnią dwie różne funkcje: podczas treningu są wykorzystywane do kierowania uczeniem modelu, umożliwiając mu dopasowanie przewidywań do rzeczywistych wyników, natomiast w fazie ewaluacji służą wyłącznie do oceny jakości predykcji, pozwalając zweryfikować, na ile model poprawnie przewiduje wyniki dla nowych danych. Model uczy się na danych treningowych, dopasowując swoje parametry, aby jak najlepiej przewidywać wyniki. Po wytrenowaniu model jest gotowy do użycia, a jego skuteczność jest sprawdzana na danych testowych, oceniając, jak dobrze radzi sobie z nowymi, wcześniej niewidzianymi danymi. Na rysunku 1. przedstawiono schemat działania uczenia maszynowego nienadzorowanego i nadzorowanego, a w tabeli 1. krótkie charakterystyki najbardziej popularnych algorytmów.

Najważniejsze zalety algorytmów uczenia maszynowego to: zastępowanie powtarzalnej i żmudnej pracy człowieka oraz zdolność do odkrywania złożonych i subtelnych wzorców, niedostępnych dla przeciętnego obserwatora. Uczenie maszynowe znalazło z powodzeniem zastosowanie w różnych dziedzinach od inżynierii kosmicznej, finansów i rozrywki, po biologię, medycynę i rolnictwo. W rolnictwie wykorzystywane jest w wielu obszarach, w tym w prognozowaniu plonów, wykrywaniu chorób i szkodników, fenotypowaniu odmian, a także w ocenie jakości nasion i optymalizacji praktyk agrotechnicznych. Integracja uczenia maszynowego z danymi pochodzącymi z dronów, satelitów, sensorów polowych oraz laboratoriów analitycz-



*podczas treningu, etykiety są wykorzystywane do nadzorowania uczenia modelu – during training, labels are used to guide the model's learning

**w fazie ewaluacji etykiety służą wyłącznie do oceny jakości predykcji – in the evaluation phase, labels are used solely to assess the quality of predictions

Rys. 1. Schemat działania uczenia maszynowego nienadzorowanego i nadzorowanego

Fig. 1. Scheme of unsupervised and supervised machine learning

Tabela 1. Charakterystyki wybranych modeli uczenia maszynowego

Table 1. Characteristics of selected machine learning models

Nazwa algorytmu Algorithm name	Opis algorytmu Algorithm description
Uczenie nienadzorowane – Unsupervised learning	
K-means	Popularny algorytm klastrowania dzielący dane na k klastrów na podstawie podobieństwa punktów. Wymaga podania liczby klastrów z góry. Zastosowanie: segmentacja klientów, analiza obrazów. A popular clustering algorithm that divides data into k clusters based on the similarity of points. It requires specifying the number of clusters in advance. Applications: customer segmentation, image analysis.
HCA (Hierarchical Cluster Analysis)	Służy do grupowania danych w hierarchiczną strukturę klastrów. Wynikiem jest dendrogram pokazujący podobieństwa między obiektami. Zastosowanie: biologia, analiza rynku. It is used to group data into a hierarchical cluster structure. The result is a dendrogram showing the similarities between objects. Applications: biology, market analysis.
PCA (Principal Component Analysis)	Wykorzystywany jest do redukcji wymiarowości poprzez przekształcenie danych w nowe, niepowiązane zmienne. Zachowuje najważniejszą zmienność danych. Zastosowanie: wizualizacja danych, redukcja redundancji. It is used for dimensionality reduction by transforming data into new, uncorrelated variables. It preserves the most important variability in the data. Applications: data visualization, redundancy reduction.
Uczenie nadzorowane – Supervised learning	
RF (Random Forest)	Model zespołowy łączący wiele drzew decyzyjnych; klasa obiektu ustalana większością głosów. Zastosowanie: klasyfikacja obrazów, prognozowanie plonów. Model combines multiple decision trees; the class of an object is determined by majority voting. Applications: image classification, crop yield prediction.
SVM (Support Vector Machine)	Tworzy granice między punktami danych w przestrzeni cech. Zastosowanie: rozpoznawanie obrazów, analiza tekstu. It creates boundaries between data points in the feature space. Applications: image recognition, text analysis.
ELM (Extreme Learning Machine)	Szybka, jednokierunkowa sieć neuronowa do regresji i klasyfikacji w czasie rzeczywistym. Zastosowanie: rozpoznawanie obrazów. A fast, feedforward neural network for real-time regression and classification. Applications: image recognition.
CNN (Convolutional Neural Network)	Głęboka sieć neuronowa z warstwami konwolucyjnymi, skuteczna w analizie danych przestrzennych i obrazów. Zastosowanie: rozpoznawanie obrazów, detekcja obiektów. A deep neural network with convolutional layers, effective in analyzing spatial data and images. Applications: image recognition, object detection.

nych pozwala nie tylko na automatyzację procesów, ale również na podejmowanie decyzji na poziomie precyzyjnego rolnictwa (Xie i Yang 2020).

Celem niniejszego artykułu jest przegląd aktualnych zastosowań metod uczenia maszynowego w produkcji rzepaku oraz omówienie możliwości, jakie te technologie oferują w kontekście zwiększenia efektywności i zrównoważonego rozwoju upraw. Artykuł obejmuje zagadnienia związane z hodowlą i fenotypowaniem, diagnostyką chorób i stresów środowiskowych, prognozowaniem plonów, oceną jakości nasion oraz optymalizacją praktyk agrotechnicznych (tab. 2). Ponadto wskazane są wyzwania, ograniczenia oraz perspektywy dalszego rozwoju tej dziedziny.

Przegląd literatury został przeprowadzony z wykorzystaniem wyszukiwarki Google Scholar. Kryteriami doboru publikacji były: zakres czasowy obejmujący lata 2017–2024, język publikacji (artykuły w języku angielskim) oraz dostępność pełnego tekstu lub streszczenia umożliwiającego ocenę metod i wyników. Wyszukiwanie opierało się na kombinacjach słów kluczowych, takich jak: „machine learning” and „rapeseed”, „yield prediction” and „Brassica napus”, a także wariantach związanych z diagnostyką chorób roślin i rolnictwem precyzyjnym. Proces selekcji obejmował wstępne przeglądanie tytułów i abstraktów, a następnie pełnych tekstów artykułów spełniających kryteria tematyczne i jakościowe. Łącznie przeanalizowano

Tabela 2. Przykłady zastosowania metod uczenia maszynowego w badaniach nad rzepakiem
Table 2. Examples of the application of machine learning methods in rapeseed research

Literatura References	Zastosowanie Application	Rodzaj danych Data type	Źródło danych Data source	Wybrany model Selected model	Metryki statystyczne Statistical metrics
Shahsavari i wsp. (2023)	plon nasion seed yield	liczba dni do dojrzałości fizjologicznej, liczba łuszczyń na roślinie, wysokość roślin number of days to physiological maturity, number of pods per plant, plant height	laboratorium laboratory	multilayer perceptron neural network (MPNN)	$R^2 = 0,84$ RMSE = 0,283 MAE = 0,224
Xie i wsp. (2022)	dynamika kwitnienia flowering dynamics	zobrazowania RGB RGB imagery	bezzałogowy statek powietrzny unmanned aerial vehicle	decision tree-based segmentation model (DTSM)	$R^2 = 0,97$ RMSE = 0,051
Qadri i wsp. (2021)	klasyfikacja odmian rzepaku rapeseed variety classification	zobrazowania RGB RGB imagery	laboratorium laboratory	artificial neural network (ANN)	dokładność – accuracy = 0,95–0,98%
Konga i wsp. (2018)	identyfikacja zgnilizny twardzikowej identification of Sclerotinia rot	zobrazowania hiperspektralne hyperspectral imagery	laboratorium laboratory	partial least squares- discriminant analysis (PLS-DA)	dokładność – accuracy = 0,90; 0,97; 1,00
Tang i wsp. (2023)	monitoring wilgotności gleby soil moisture monitoring	pomiary hiperspektralne hyperspectral measurements	pole doświadczalne experimental field	random forest (RF)	$R^2 = 0,94$ RMSE = 0,005 MRE = 0,031
Zhang i wsp. (2023)	monitoring zawartości wody w liściach leaf water content monitoring	zobrazowania hiperspektralne hyperspectral imagery	laboratorium laboratory	convolutional neural network regression with attention mechanism and long short-term memory (CNN-ATT-LSTM-R)	$R^2 = 0,81$ RMSEP = 0,005
Xia i wsp. (2019)	stres zalania roślin plant flooding stress	zobrazowania hiperspektralne hyperspectral imagery	laboratorium laboratory	quadratic discriminant analysis (QDA)	dokładność – accuracy = 0,96
Li i wsp. (2022)	stres wymarzania frost stress	zobrazowania RGB RGB imagery	bezzałogowy statek powietrzny unmanned aerial vehicle	convolutional neural network (CNN)	dokładność – accuracy = 0,93 F-score = 0,80 czułość – sensitivity = 0,80
Zhou i wsp. (2023)	wykrywanie metali ciężkich heavy metal detection	zobrazowania hiperspektralne hyperspectral imagery	laboratorium laboratory	transfer stacked auto- encoder (T-SAE)	dokładność – accuracy = 0,99 $R^2 = 0,92$ –0,93
Hu i wsp. (2024)	predykcja plonu nasion seed yield prediction	zobrazowania RGB i hiperspektralne RGB and hyperspectral imagery	bezzałogowy statek powietrzny unmanned aerial vehicle	random forest regression (RFR)	$R^2 = 0,93$ RRMSE = 0,059
Shi i wsp. (2024)	predykcja plonu nasion seed yield prediction	zobrazowania multispektralne multispectral imagery	bezzałogowy statek powietrzny unmanned aerial vehicle	extreme gradient boosting regression (XGBR)	$R^2 = 0,58$ –0,77 RMSE = 0,005–0,011
Rajković i wsp. (2022)	predykcja plonu nasion seed yield prediction	genotyp, rok genotype, year	pole doświadczalne experimental field	random forest regression (RFR)	$R^2 = 0,91$ RMSE = 227,082
Zhang i wsp. (2023)	indeks powierzchni liścia leaf area index	pomiary hiperspektralne hyperspectral measurements	pole doświadczalne experimental field	random forest regression (RFR)	$R^2 = 0,81$ RMSE = 0,455

Tabela 2. Przykłady zastosowania metod uczenia maszynowego w badaniach nad rzepakiem – cd.**Table 2.** Examples of the application of machine learning methods in rapeseed research – continued

Literatura References	Zastosowanie Application	Rodzaj danych Data type	Źródło danych Data source	Wybrany model Selected model	Metryki statystyczne Statistical metrics
Wei i wsp. (2017)	indeks powierzchni liścia leaf area index	zobrazowania multispektralne multispectral imagery	satelita satellite	random forest regression (RFR)	$R^2 = 0,52$ RMSE = 0,923 RRMSE = 0,937
Rajković i wsp. (2022)	predykcja masy 1000 nasion, zawartości oleju i białka prediction of thousand- seed weight, oil content, and protein content	genotyp, rok genotype, year	laboratorium laboratory	random forest regression (RFR)	$R^2 = 0,89; 0,94;$ 0,94 RMSE = 0,179; 0,581, 0,501
Yu i wsp. (2018)	zawartość azotu w liściach leaf nitrogen content	zobrazowania hiperspektralne hyperspectral imagery	laboratorium laboratory	fully-connected neural network (FNN)	$R^2 = 0,90$ RMSEP = 0,003
Abdalla i wsp. (2020)	zawartość azotu, fosforu i potasu w roślinie nitrogen, phosphorus, and potassium content in the plant	zobrazowania RGB RGB imagery	pole doświadczalne experimental field	long-short term memory (LSTM)	dokładność – accuracy = 0,95
Feng i wsp. (2024)	termin zbioru rzepaku rapeseed harvest date	zobrazowania hiperspektralne nasion rzepaku hyperspectral imagery of rapeseed seeds	laboratorium laboratory	support vector machine (SVM)	dokładność – accuracy = 0,98

20 publikacji, koncentrując się na badaniach empirycznych istotnych dla zastosowania uczenia maszynowego w produkcji rzepaku.

Wykorzystanie uczenia maszynowego do optymalizacji hodowli i uprawy rzepaku / Application of machine learning for optimizing rapeseed breeding and cultivation

Hodowla odmian / Variety breeding

Hodowla rzepaku wymaga dokładnej oceny cech fenotypowych, takich jak: wysokość roślin, liczba i długość rozgałęzień, liczba kwiatostanów i łuszczyń czy masa tysiąca nasion. Tradycyjne metody pomiaru są czasochłonne i podatne na subiektywizm, co utrudnia szybki postęp w programach hodowlanych (Li i wsp. 2022). W ostatnich latach rosnące możliwości pozyskiwania danych z dronów (UAV), zdjęć multispektralnych i hiperspektralnych, a także systemów wysokoprzepustowego fenotypowania przyczyniają się do rozwoju nowoczesnych metod wspierających selekcję odmian. Uczenie maszynowe pozwala na automatyczne przetwarzanie dużych zbiorów danych fenotypowych i efektywną selekcję cech w procesie hodowlanym, ukierunkowaną na wybór tych najsilniej związanych z plonem roślin. Takie badania prowadzili Shahsavari i wsp. (2023), którzy porównali skuteczność pięciu metod regresyjnych

w prognozowaniu plonu nasion rzepaku z wykorzystaniem piętnastu cech morfologiczno-rozwojowych (wysokość rośliny, liczba łuszczyń na pędzie głównym, liczba łuszczyń na pędach bocznych, liczba łuszczyń na roślinie, liczba pędów na roślinie, długość pędu głównego, wysokość osadzenia pierwszej łuszczyzny od powierzchni gleby, długość łuszczyzny, liczba dni do początku kwitnienia, liczba dni do końca kwitnienia, liczba dni do fizjologicznej dojrzałości, okres kwitnienia, masa tysiąca nasion, liczba nasion w łuszczyźnie, średnica łodygi). Najwyższą dokładność prognozy uzyskano przy zastosowaniu modelu nu-support vector regression (NuSVR) ($R^2 = 0,86$; RMSE = 0,266; MAE = 0,210), a najważniejszymi cechami istotnie wpływającymi na plon nasion okazały się: liczba dni do dojrzałości fizjologicznej, liczba łuszczyń na roślinie oraz wysokość roślin. Wymienione cechy uwzględniono przy opracowaniu modelu multilayer perceptron neural network (MLPNN), który charakteryzował się następującymi parametrami: $R^2 = 0,84$; RMSE = 0,283; MAE = 0,224. Wyniki badań wykazały, że zastosowanie metod uczenia maszynowego umożliwia dokładne przewidywanie plonu rzepaku na podstawie mniejszej liczby cech, a tym samym pomaga w optymalizacji i przyspieszeniu programów hodowlanych rzepaku.

Z kolei Xie i wsp. (2022) zajmowali się oceną dynamiki kwitnienia różnych genotypów rzepaku z zastosowaniem

zdjęć RGB rejestrowanych z pokładu bezzałogowego statku powietrznego. Autorzy wykazali, że dokładność wyodrębniania obszarów kwitnienia przy użyciu modelu decision tree-based segmentation model (DTSM) była wyższa niż przy zastosowaniu: naive Bayes (NB), k-nearest neighbours (KNN), random forest (RF) i support vector machine (SVM) dla wszystkich odmian i terminów kwitnienia, osiągając $R^2 = 0,97$ i $RMSE = 0,051$. W badaniu wykazano zróżnicowanie dynamiki kwitnienia analizowanych odmian. Wszystkie odmiany zostały sklasyfikowane w cztery klastry na podstawie analizy skupień metodą k-średnich (k-means), które różniły się przede wszystkim terminem osiągnięcia pełni kwitnienia oraz terminem rozpoczęcia kwitnienia. Wyniki tego badania mogą stanowić podstawę do prowadzenia hodowli rzepaku opartej na analizie dynamiki kwitnienia.

Uczenie maszynowe znalazło również zastosowanie w klasyfikacji odmian rzepaku na podstawie zdjęć RGB (Qadri i wsp. 2021). Autorzy wykazali, że zastosowanie modelu artificial neural network (ANN) pozwalało skutecznie rozróżniać osiem odmian rzepaku uzyskując bardzo wysoką dokładność na poziomie 0,95–0,98.

Diagnostyka chorób i stresów środowiskowych / Diagnosis of diseases and environmental stresses

Choroby oraz stresy środowiskowe, takie jak susza, mróz czy niedobory składników odżywczych, znacząco ograniczają plon i jakość nasion rzepaku. Tradycyjne metody diagnostyki, oparte na wizualnej ocenie pola lub laboratoryjnych testach, są czasochłonne i nie zawsze pozwalają na szybkie wykrycie zagrożeń.

Przykładem ilustrującym efektywne wykorzystanie uczenia maszynowego do wykrywania chorób roślin są badania Konga i wsp. (2018), którzy zastosowali cztery modele dyskryminacyjne, w tym partial least squares-discriminant analysis (PLS-DA), support vector machine (SVM), soft independent modeling of class analogies (SIMCA) i k-nearest neighbors (KNN) do identyfikacji objawów zgnilizny twardzikowej na liściach rzepaku ozimego zarejestrowanych z wykorzystaniem kamery hiperspektralnej (ImSpector V10E, Spectral Imaging Ltd., Oulu, Finland). Za pomocą wymienionych modeli autorzy analizowali zmienność jasności pikseli obrazu, odpowiadających obszarom liści zdrowych oraz porażonych w różnym stopniu. Przeprowadzone na trzech niezależnych zestawach danych badania wykazały, że najefektywniejszym modelem był PLS-DA, którego dokładność wykrywania analizowanej choroby wynosiła w zależności od zestawu użytych danych 0,90; 0,97 i 1.

Uczenie maszynowe znalazło również zastosowanie w identyfikacji reakcji roślin na stresy abiotyczne. Tang i wsp. (2024) potwierdzili możliwość dokładnego monitorowania wilgotności gleby w strefie korzeniowej rzepaku ozimego przy użyciu danych hiperspektralnych zarejestrowanych z użyciem spektrometru polowego (ASD Field-

-Spec 3). Do estymacji wilgotności gleby zastosowano cztery modele: support vector machine (SVM), random forest (RF), back propagation neural network (BPNN) i extreme learning machine (ELM). Najlepsze wyniki osiągnął model random forest (RF) dla głębokości gleby 0–20 cm, uzyskując wysoki współczynnik determinacji $R^2 = 0,94$, niski błąd średniokwadratowy $RMSE = 0,005$ oraz średni względny błąd $MRE = 0,031$.

Oprócz badań poświęconych zawartości wody w glebie prowadzone były prace nad zawartością wody w liściach rzepaku. Przykładowo Zhang i wsp. (2023a) z wykorzystaniem zobrażeń zarejestrowanych za pomocą kamery hiperspektralnej FX17 (Spectral Imaging Ltd., Oulu, Finland) ocenili przydatność do realizacji tego celu następujących modeli: support vector regression (SVR), partial least squares regression (PLSR), least absolute shrinkage i selection operator (LASSO), convolutional neural network regression with attention mechanism and long short-term memory (CNN-ATT-LSTM-R). Największą dokładność wykazał model CNN-ATT-LSTM-R. Współczynnik determinacji (R^2) i błąd średniokwadratowy dla zbioru testowego (RMSEP) tego modelu wynosiły odpowiednio 0,81 i 0,005.

Z kolei Xia i wsp. (2019) podjęli próbę identyfikacji stresu wynikającego z zalania roślin rzepaku. Do budowy modeli klasyfikacyjnych porównujących spektra liści poddanych różnym poziomom zalania wodnego w trzech zbiorach danych zgromadzonych z użyciem kamery hiperspektralnej Pika XC, zastosowali Quadratic discriminant analysis (QDA), k-nearest neighbor (KNN) i support vector machine (SVM). Najlepszą dokładność klasyfikacji wynoszącą 0,96 uzyskano z zastosowaniem modelu QDA.

Ciekawe badania przedstawili Li i wsp. (2022), którzy zastosowali convolutional neural network (CNN) do rozpoznawania materiałów hodowlanych rzepaku odpornych na przemarzanie, osiągając następujące wartości kluczowych parametrów statystycznych: dokładność = 0,93; F-score = 0,80 oraz czułość = 0,80. Badanie wykazało, że wykorzystanie zobrażeń RGB zarejestrowanych z pokładu bezzałogowego statku powietrznego (UAV) w połączeniu z głębokim uczeniem pozwala automatycznie i efektywnie identyfikować materiały hodowlane odporne na mróz w dużych populacjach roślin, oferując ekonomiczną, wygodną i precyzyjną metodę wspierającą hodowców w selekcji i zarządzaniu materiałami roślinnymi.

Kolejnym zagadnieniem podejmowanym z zastosowaniem uczenia maszynowego w zakresie detekcji stresów środowiskowych są badania zawartości metali ciężkich w roślinach rzepaku. Zhou i wsp. (2023) opracowali model do przewidywania zawartości ołowiu (Pb) w roślinach rzepaku, wykorzystujący obrazowanie hiperspektralne. Dane spektralne liści i korzeni rzepaku pozyskano przy różnych wartościach Pb. Opracowany model – transfer stacked auto-encoder (T-SAE) osiągnął 0,99 dokładności klasyfikacji gradientu stresu Pb, a współczynniki determinacji dla za-

wartości Pb w liściach i korzeniach wyniosły odpowiednio 0,92 i 0,93.

Prognozowanie plonów oraz parametrów liści i nasion / Prediction of yields, as well as leaf and seed traits

Prognozowanie plonów jest kluczowe zarówno dla planowania produkcji, jak i zarządzania ryzykiem w rolnictwie. Tradycyjne metody oparte na analizie danych historycznych, obserwacjach polowych oraz modelach empirycznych, są czasochłonne i podatne na błędy. Uczenie maszynowe umożliwia wykorzystanie dużych zbiorów danych pochodzących z dronów, satelitów, sensorów polowych oraz laboratoriów analitycznych do tworzenia precyzyjnych modeli predykcyjnych (Omia i wsp. 2023).

Przykładem ilustrującym zastosowanie dronów są badania Hu i wsp. (2024), którzy analizowali zdjęcia RGB i hyperspektralne rzepaku w fazie siewki, pąkowania, kwitnienia oraz tworzenia łuszczyń. Z obrazów hyperspektralnych wyodrębniono i obliczono refleksję pokrywy roślinnej oraz wskaźniki spektralne. Po przeprowadzeniu analizy korelacji i analizy głównych składowych z zastosowaniem modelu principal component analysis (PCA) wybrano wskaźniki spektralne, które posłużyły do budowy modeli predykcji plonów przy użyciu random forest regression (RF), multiple linear regression (MLR) i support vector machine regression (SVM). Wyniki wykazały, że podejścia oparte na uczeniu maszynowym, mogą poprawić dokładność predykcji plonów w porównaniu z metodami tradycyjnymi, takimi jak regresja wielokrotna. Model oparty na RF uzyskał najwyższy współczynnik determinacji ($R^2 = 0,93$) oraz najniższy względny błąd pierwiastkowy średniokwadratowy (RRMSE = 0,059). Natomiast model regresji wielokrotnej osiągnął $R^2 = 0,73$ oraz RRMSE = 0,113. Ponadto, stwierdzono że etap pąkowania był najlepszy do prognozowania plonów rzepaku w porównaniu z innymi fazami wzrostu.

Z kolei Shi i wsp. (2024) w prognozowaniu plonu rzepaku na podstawie obrazowań multispektralnych zarejestrowanych z pokładu bezzałogowego statku powietrznego, oprócz tradycyjnej regresji liniowej, zastosowali trzy modele uczenia maszynowego: random forest regression (RFR), support vector regression (SVR) oraz extreme gradient boosting regression (XGBR). Autorzy wykazali, że wszystkie trzy algorytmy przewyższają regresję liniową pod względem dokładności prognoz, a najdokładniejszym modelem okazał się XGBR.

Przykładem zastosowania uczenia maszynowego do prognozowania plonu nasion rzepaku są również badania Rajkovića i wsp. (2022). Na podstawie informacji o genotypie, odzwierciedlającym zmienność biologiczną oraz roku badań, odzwierciedlającym zmienność środowiskową, opracowali oni dwa algorytmy uczenia maszynowego: artificial neural networks (ANN) i random forest regression (RFR). Na podstawie uzyskanych wyników wykazali prze-

wagę modelu RFR ($R^2 = 0,91$; RMSE = 227,082) nad ANN ($R^2 = 0,83$; RMSE = 303,249).

Oprócz prognozowania plonu, uczenie maszynowe znalazło zastosowanie w szacowaniu indeksu powierzchni liści (LAI), kluczowego wskaźnika dla oceny wzrostu i rozwoju roślin uprawnych. Zhang i wsp. (2023b) na podstawie danych zarejestrowanych z użyciem polowego spektrometru ASD sprawdzili przydatność do tego celu trzech algorytmów uczenia maszynowego: back propagation neural network (BPNN), support vector machine (SVM) i random forest (RF). Spośród testowanych algorytmów najwyższą dokładność estymacji indeksu LAI uzyskano dla modelu RF ($R^2 = 0,81$; RMSE = 0,455), co potwierdza jego przydatność w monitorowaniu wzrostu rzepaku.

Z kolei Wei i wsp. (2017) podjęli się opracowania modelu do szacowania LAI na podstawie zdjęć satelitarnych. Do mapowania LAI rzepaku ozimego wykorzystano random forest regression (RFR) w ramach hybrydowej metody inwersji, łączącej optymalne wskaźniki roślinności (NDVI, MSR, ARVI) z odbiciem satelitarnym wysokiej rozdzielczości, co pozwoliło na przestrzenne odwzorowanie LAI na poziomie polowym. Zgodność estymowanych wartości LAI z pomiarami terenowymi wyniosła $R^2 = 0,52$.

Zakres zastosowań uczenia maszynowego w badaniach nad rzepakiem obejmuje również ocenę jakości nasion. Do najważniejszych cech jakościowych nasion rzepaku należą zawartość tłuszczu i białka oraz profil kwasów tłuszczowych, determinujący wartość odżywczą i technologiczną nasion. Rajković i wsp. (2022) wykazali, że wykorzystanie algorytmu random forest regression (RFR) umożliwia skuteczną predykcję takich cech, jak masa 1000 nasion ($R^2 = 0,89$; RMSE = 0,179), a także zawartość oleju ($R^2 = 0,94$; RMSE = 0,581) i białka ($R^2 = 0,94$; RMSE = 0,501). Do opracowania modelu wykorzystano informacje o genotypie, odzwierciedlającym zmienność biologiczną oraz roku badań, odzwierciedlającym zmienność środowiskową.

Optymalizacja praktyk agrotechnicznych / Optimization of agronomic practices

Optymalizacja praktyk agrotechnicznych jest kluczowa dla zwiększenia wydajności rzepaku, poprawy jakości nasion oraz ograniczenia negatywnego wpływu na środowisko. Tradycyjne podejścia oparte na doświadczeniu rolników lub ogólnych zaleceniach agronomicznych nie zawsze uwzględniają zmienność warunków lokalnych, co może prowadzić do suboptymalnych wyników. Metody uczenia maszynowego pozwalają na tworzenie modeli wspomagających decyzje dotyczące nawożenia oraz terminów zbiorów.

Yu i wsp. (2018) potwierdzili możliwość szybkiego i nieinwazyjnego oznaczania zawartości azotu w liściach rzepaku przy użyciu metod uczenia maszynowego połączonych z technikami obrazowania hyperspektralnego. Spośród testowanych modeli (PLS – partial least square, SVM

– support vector machine, FNN – fully-connected neural network) najwyższą skutecznością charakteryzował się model FNN, dla którego wartości R^2 wyniosły 0,90; a RMSEP 0,003.

Większy zakres badań obejmujący oprócz azotu również szacowanie zawartości fosforu i potasu zaproponowali Abdalla i wsp. (2020), którzy z zastosowaniem modelu long-short term memory (LSTM), przeprowadzili klasyfikację stanu odżywienia rzepaku na podstawie sekwencji zobrażeń RGB. Opracowana metoda pozwala z dokładnością równą 0,95 na automatyczną ocenę stanu odżywienia w całym cyklu życia roślin i ma potencjał w zastosowaniach rolnictwa precyzyjnego.

Przykładem badań ukierunkowanych na ustalenie optymalnego terminu zbioru rzepaku prowadzonych z wykorzystaniem metod uczenia maszynowego jest praca Fenga i wsp. (2024), którzy zastosowali kilka algorytmów uczenia maszynowego [extreme learning machine (ELM), k-nearest neighbor (KNN), random forest (RF), partial least-squares discriminant analysis (PLSDA), support vector machine (SVM)] w połączeniu z obrazowaniem hiperspektralnym nasion rzepaku. Najlepszy wynik uzyskano z zastosowaniem modelu SVM, osiągając dokładność 0,98. Wyniki pokazują, że dojrzałość rzepaku może być szybko i nieinwazyjnie oceniana za pomocą obrazowania hiperspektralnego, co ułatwia wybór optymalnego terminu zbioru.

Wyzwania i perspektywy rozwoju / Challenges and development prospects

Pomimo rosnącego zainteresowania uczeniem maszynowym w produkcji rzepaku ozimego oraz licznych sukcesów, nadal istnieje wiele wyzwań, które ograniczają pełne wykorzystanie tych technologii. Jednym z głównych problemów jest zbieranie i jakość danych, ponieważ dane rolnicze są często heterogeniczne, niestrukturalne i niekompletne, co utrudnia ich gromadzenie i analizę, a dodatkowo na ich wiarygodność wpływają takie czynniki, jak warunki pogodowe, obecność szkodników czy występowanie chorób. Kolejną trudnością jest brak standaryzacji, gdyż rolnictwo charakteryzuje się dużym zróżnicowaniem upraw, rodzajów gleb i systemów gospodarowania, co utrudnia ujednolicenie procesów zbierania danych i tym samym ogranicza możliwość tworzenia dokładnych i skalowalnych modeli uczenia maszynowego. Istotnym wyzwaniem pozostaje również ograniczona dostępność danych treningowych, niezbędnych do prawidłowego działania algorytmów, które wymagają dużych zbiorów informacji, a w rolnictwie często ich brakuje. Dodatkową barierę stanowi integracja uczenia maszynowego z istniejącymi praktykami rolniczymi, ponieważ wielu rolników nie zna zasad jego działania lub nie posiada odpowiednich narzędzi i zasobów, aby skutecznie je wdrożyć. Ważnym ograniczeniem jest także infrastruktura, gdyż algorytmy uczenia maszynowego wymagają znacznej mocy obliczeniowej i dużej pojemności magazynowania

danych, co bywa problematyczne w regionach o słabo rozwiniętej infrastrukturze, typowych dla obszarów wiejskich. Nie bez znaczenia są również koszty, ponieważ wdrażanie rozwiązań opartych na uczeniu maszynowym w rolnictwie jest kosztowne, a zwrot z inwestycji często nie jest natychmiastowy ani oczywisty, co szczególnie utrudnia korzystanie z tych technologii małym gospodarstwom i lokalnym społecznościom.

Perspektywy rozwoju uczenia maszynowego w produkcji rzepaku są jednak obiecujące. Integracja uczenia maszynowego z sensorami polowymi oraz systemami satelitarnymi pozwala na tworzenie modeli predykcyjnych działających w czasie rzeczywistym. Połączenie danych rejestrowanych z użyciem urządzeń pomiarowych zamontowanych na maszynach rolniczych, dronach czy satelitach zagwarantuje precyzyjne zarządzanie procesem produkcji i wczesne reagowanie na zagrożenia. Dalszy rozwój uczenia maszynowego w rolnictwie wymaga także szerszej współpracy między hodowcami, agronomami i specjalistami od danych, aby tworzyć bazy danych i modele dostosowane do potrzeb praktyki polowej. Standaryzacja procedur zbierania danych, opracowanie platform analitycznych oraz popularyzacja technologii wśród rolników będą kluczowe dla efektywnego wdrożenia tych narzędzi w codziennej produkcji rzepaku.

Podsumowanie / Summary

Uczenie maszynowe odgrywa coraz większą rolę w produkcji rzepaku ozimego, wspierając rozwój nowoczesnych metod hodowli, fenotypowania, diagnostyki chorób, prognozowania plonów oraz optymalizacji praktyk agrotechnicznych. Metody uczenia maszynowego pozwalają na przetwarzanie i analizę dużych zbiorów danych pochodzących z dronów, satelitów, sensorów polowych i laboratoriów analitycznych, co znacząco przyspiesza procesy decyzyjne i zwiększa dokładność przewidywań. W hodowli i fenotypowaniu uczenie maszynowe umożliwia automatyczną identyfikację cech morfologicznych, liczenie roślin, segmentację kwiatostanów oraz selekcję najbardziej progностycznych cech, wspierając wybór odmian o wysokim potencjale plonotwórczym i odpornościowym. W diagnostyce chorób i stresów środowiskowych algorytmy uczenia maszynowego pozwalają na wczesne wykrywanie infekcji, monitorowanie uszkodzeń spowodowanych suszą czy mrozem oraz wspierają selekcję odmian odpornych na niekorzystne warunki. W prognozowaniu plonów i ocenie jakości nasion uczenie maszynowe umożliwia tworzenie modeli predykcyjnych z dużą dokładnością, wykorzystując dane rejestrowane z poziomu pola, bezzałogowego statku powietrznego i satelity. Modele te wspierają planowanie produkcji oraz kontrolę jakości nasion, co jest istotne zarówno dla hodowców, jak i przemysłu olejarskiego. W optymalizacji praktyk agro-

technicznych uczenie maszynowe wspiera precyzyjne dawowanie nawozów i wyznaczanie terminów zbiorów. Mimo licznych sukcesów, pełne wykorzystanie uczenia maszynowego w produkcji rzepaku napotyka na wyzwania związane z dostępnością i standaryzacją danych oraz interpretowalnością wyników. Dodatkowo istotnym ograniczeniem jest ryzyko nadmiernego dopasowania (overfitting), zwłaszcza w przypadku niewielkich lub mało zróżnicowanych zbiorów danych, a także brak pełnej walidacji krzyżowej. Przenoszenie modeli na inne warunki środowiskowe, odmiany roślin czy różne sezony może prowadzić do spadku dokład-

ności prognoz, co podkreśla potrzebę ostrożnej interpretacji wyników oraz dalszych badań nad generalizacją modeli.

Współpraca między hodowcami, agronomami i specjalistami od danych stwarza jednak ogromny potencjał dla zwiększenia efektywności, zrównoważenia produkcji i lepszego zarządzania ryzykiem w uprawie rzepaku. Podsumowując, uczenie maszynowe staje się nieocenionym narzędziem wspierającym rozwój produkcji rzepaku, integrując dane z różnych źródeł, zwiększając precyzję decyzji oraz umożliwiając szybszy postęp w hodowli i praktyce rolniczej.

Literatura / References

- Abdalla A., Cen H., Wan L., Mehmood K., He Y. 2020. Nutrient status diagnosis of infield oilseed rape via deep learning-enabled dynamic model. *IEEE Transactions on Industrial Informatics* 17 (6): 4379–4389. DOI: 10.1109/TII.2020.3009736
- Feng H., Chen Y., Song J., Lu B., Shu C., Qiao J., Liao Y., Yang W. 2024. Maturity classification of rapeseed using hyperspectral image combined with machine learning. *Plant Phenomics* 6: 0139. DOI: 10.34133/plantphenomics.0139
- Food and Agriculture Organization of the United Nations (FAO). FAOSTAT Statistical Database 2023. <https://www.fao.org/faostat/en> [dostęp: 03.09.2025].
- Hu H., Ren Y., Zhou H., Lou W., Hao P., Lin B., Zhang G., Gu Q., Hua S. 2024. Oilseed rape yield prediction from UAVs using vegetation index and machine learning: a case study in East China. *Agriculture* 14 (8): 1–15. DOI: 10.3390/agriculture14081317
- Kong W., Zhang C., Cao F., Liu F., Luo S., Tang Y., He Y. 2018. Detection of sclerotinia stem rot on oilseed rape (*Brassica napus* L.) leaves using hyperspectral imaging. *Sensors* 18 (6): 1–15. DOI: 10.3390/s18061764
- Li L., Qiao J., Yao J., Li J., Li L. 2022. Automatic freezing-tolerant rapeseed material recognition using UAV images and deep learning. *Plant Methods* 18 (1): 1–13. DOI: 10.1186/s13007-022-00838-6
- Omia E., Bae H., Park E., Kim M.S., Baek I., Kabenge I., Cho B.K. 2023. Remote sensing in field crop monitoring: a comprehensive review of sensor systems, data analyses and recent advances. *Remote Sensing* 15 (2): 1–46. DOI: 10.3390/rs15020354
- Qadri S., Qadri S.F., Razzaq A., Rehman M.U., Nazir A., Nawaz S.A., Saher N., Akhta N., Khan D.M. 2021. Classification of canola seed varieties based on multi-feature analysis using computer vision approach. *International Journal of Food Properties* 24 (1): 493–504. DOI: 10.1080/10942912.2021.1900235
- Rajković D., Marjanovic J.A., Pezo L., Loncar B., Zanetti F., Monti A., Kondic Špika A. 2022. Yield and quality prediction of winter rapeseed – artificial neural network and random forest models. *Agronomy* 12 (1): 58. DOI: 10.3390/agronomy12010058
- Shahsavari M., Mohammadi V., Alizadeh B., Alizadeh H. 2023. Application of machine learning algorithms and feature selection in rapeseed (*Brassica napus* L.) breeding for seed yield. *Plant Methods* 19 (1): 1–22. DOI: 10.1186/s13007-023-01035-9
- Shi H., Li Z., Xiang Y., Tang Z., Sun T., Du R., Li W., Liu X., Huang X., Liu Y., Zhang F. 2024. Integrating multi-source remote sensing and machine learning for root-zone soil moisture and yield prediction of winter oilseed rape (*Brassica napus* L.): A new perspective from the temperature-vegetation index feature space. *Agricultural Water Management* 305: 109129. DOI: 10.1016/j.agwat.2024.109129
- Tang Z., Zhang W., Xiang Y., Liu X., Wang X., Shi H., Li Z., Zhang F. 2024. Monitoring of soil moisture content of winter oilseed rape (*Brassica napus* L.) based on hyperspectral and machine learning models. *Journal of Soil Science and Plant Nutrition* 24 (1): 1250–1260. DOI: 10.1007/s42729-024-01626-y
- Wei C., Huang J., Mansaray L.R., Li Z., Liu W., Han J. 2017. Estimation and mapping of winter oilseed rape LAI from high spatial resolution satellite data based on a hybrid method. *Remote Sensing* 9 (5): 1–16. DOI: 10.3390/rs9050488
- Xia J.A., Cao H., Yang Y., Zhang W., Wan Q., Xu L., Ge D., Zhang W., Ke Y., Huang B. 2019. Detection of waterlogging stress based on hyperspectral images of oilseed rape leaves (*Brassica napus* L.). *Computers and Electronics in Agriculture* 159: 59–68. DOI: 10.1016/j.compag.2019.02.022
- Xie Z., Chen S., Gao G., Li H., Wu X., Meng L., Ma Y. 2022. Evaluation of rapeseed flowering dynamics for different genotypes with UAV platform and machine learning algorithm. *Precision Agriculture* 23 (5): 1688–1706. DOI: 10.1007/s11119-022-09904-4
- Xie C., Yang C. 2020. A review on plant high-throughput phenotyping traits using UAV-based sensors. *Computers and Electronics in Agriculture* 178: 105731. DOI: 10.1016/j.compag.2020.105731
- Yu X., Lu H., Liu Q. 2018. Deep-learning-based regression model and hyperspectral imaging for rapid detection of nitrogen concentration in oilseed rape (*Brassica napus* L.) leaf. *Chemometrics and Intelligent Laboratory Systems* 172: 188–193. DOI: 10.1016/j.chemolab.2017.12.010
- Zhang C., Li C., He M., Cai Z., Feng Z., Qi H., Zhou L. 2023a. Leaf water content determination of oilseed rape using near-infrared hyperspectral imaging with deep learning regression methods. *Infrared Physics & Technology* 134: 104921. DOI: 10.1016/j.infrared.2023.104921
- Zhang W., Li Z., Pu Y., Zhang Y., Tang Z., Fu J., Xu W., Xiang Y., Zhang F. 2023b. Estimation of the leaf area index of winter rapeseed based on hyperspectral and machine learning. *Sustainability* 15 (17): 1–13. DOI: 10.3390/su151712930
- Zhou X., Zhao C., Sun J., Yao K., Xu M. 2023. Detection of lead content in oilseed rape leaves and roots based on deep transfer learning and hyperspectral imaging technology. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 290: 122288. DOI: 10.1016/j.saa.2022.122288