

Prof. dr hab inż Maciej Kuboń

Kraków 12.01.2023 r.

Katedra Inżynierii Produkcji, Logistyki

i Informatyki Stosowanej

Wydział Inżynierii Produkcji i Energetyki

Uniwersytet Rolniczy im. Hugona Kołłątaja w Krakowie

30-149 Kraków, ul. Balicka 116B

Tel. (012) 662-46-99, e-mail: maciej.kubon@urk.edu.pl

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr inż. Dominiki Sierackiej

pt.:

**„Identyfikacja parametrów charakterystycznych w plonowaniu nasion
odmiany konopi przemysłowych (*Cannabis sativa* L.)
z użyciem metod sztucznej inteligencji”**

1. Podstawa opracowania recenzji

Recenzję opracowano na podstawie zlecenia Przewodniczącego Rady Naukowej dyscypliny „Inżynieria Mechaniczna” Uniwersytetu Przyrodniczego w Poznaniu, prof. UPP dr hab. inż. Macieja Zaborowicza, z dnia 15 listopada 2022 roku (nr pisma WI-4000-61/2022). Podstawę formalno-prawną opracowania recenzji rozprawy doktorskiej mgr inż. Dominiki Sierackiej na temat: „Identyfikacja parametrów charakterystycznych w plonowaniu nasion odmiany konopi przemysłowych (*Cannabis sativa* L.) z użyciem metod sztucznej inteligencji” stanowi Uchwała Rady Naukowej Dyscypliny inżynieria mechaniczna Uniwersytetu Przyrodniczego w Poznaniu z dnia 15 listopada 2022 roku.

2. Ogólna charakterystyka i ocena formalna rozprawy doktorskiej

Przedstawiona do recenzji rozprawa doktorska została napisana w Katedrze Inżynierii Biosystemów na Wydziale Inżynierii Środowiska i Inżynierii Mechanicznej Uniwersytetu Przyrodniczego w Poznaniu przez mgr inż. Dominikę Sieracką pod opieką naukową promotora prof. UPP dra hab. inż. Macieja Zaborowicza oraz promotora pomocniczego dra inż. Jakuba Frankowskiego.

Praca liczy 82 strony i podzielona jest na 9 rozdziałów głównych wraz z podrozdziałami. Układ pracy prawidłowy, z zachowaniem właściwych proporcji poszczególnych rozdziałów, zgodny z ogólnie przyjętymi zasadami w tego typu pracach naukowych. Można w nich dostrzec logikę badań, jak również logikę w interpretacji osiągniętych efektów badań.

Pierwszy z rozdziałów zatytułowany „Wstęp”, przedstawia podstawowe informacje na temat roli i znaczenia konopi przemysłowych oraz badań z wykorzystaniem sztucznej inteligencji.

W rozdziale 2 Doktorantka dokonała przeglądu literatury w aspekcie prawnym i technologicznym uprawy konopi przemysłowych oraz ich wykorzystania, a także dokonała oceny metod modelowania procesów zarówno w rolnictwie, jak i w branży konopnej.

W rozdziale 3 w sposób poprawny przedstawiono cel i zakres pracy, problem naukowy wraz z problemami cząstkowymi oraz hipotezę roboczą, czyli wszystko to co powinna zawierać praca naukowa. O ile cel i zakres pracy nie budzą żadnych zastrzeżeń, prosiłbym doktorantkę o bliższe wyjaśnienie w jaki sposób można dokonać „...**efektywnej** predykcji plonowania...” o której wspomina w problemie naukowym oraz na **jakiej podstawie można odrzucić** tak postawioną hipotezę roboczą. Uważam również, że pierwszy problem czątkowy powinien brzmieć w sposób następujący: Które cechy procesu uprawy konopi przemysłowych są **istotne** w procesie predykcji ich plonu.

Rozdział 4 „Materiał i metody” zawiera charakterystykę materiału badawczego oraz opis metodyki badań, ze szczególnym uwzględnieniem metodyki pozyskiwania materiału badawczego, przygotowania danych do modelowania oraz charakterystyki uczenia modeli neuronowych.

W rozdziale 5 przedstawiono wyniki uzyskane z badań, a dyskusja wyników została zawarta w rozdziale 6. Praca zakończona została podsumowaniem i wnioskami, w których Doktorantka w sposób syntetyczny przedstawiła zakres wykonanych prac oraz wnioski z przeprowadzonych badań. Na końcu znajduje się wykaz literatury zawierający 96 pozycji publikacji naukowych oraz wykaz 9 stron internetowych. Spośród wspomnianych 96 pozycji literatury, 64 stanowią opracowania w języku angielskim, a pozostałe w języku polskim. Należy także zaznaczyć, że do wszystkich zamieszczonych pozycji bibliograficznych zostały zastosowane w tekście stosowne odsyłacze. W zebranej literaturze dominują publikacje z ostatniej dekady, w tym 26% prac opublikowanych było po roku 2020, a do rzadkości należy zaliczyć te, które pochodzą sprzed 2010 roku. Wykorzystanie przez Doktorantkę tak ogromnej ilości źródeł informacji, zwłaszcza obcojęzycznych, wymagało dużego nakładu pracy. Zatem dobór literatury i jej opracowanie oceniam pozytywnie. Do pracy dołączono streszczenie w języku polskim i angielskim.

Stwierdzam, że przedstawiona do oceny praca zawiera wszystkie wymagane elementy rozprawy doktorskiej oraz ma charakter naukowo-badawczy.

3. Ocena merytoryczna rozprawy

Tematyka rozprawy mieści się w dyscyplinie inżynieria mechaniczna, w obszarze inżynieria rolnicza i stanowi istotne zagadnienie w punktu widzenia gospodarczego.

Konopie przemysłowe (*Cannabis sativa* L.) to jedne z bardziej wszechstronnych roślin uprawianych w Polsce. W zależności od odmiany, sposobu zasiewu czy terminu zbioru pozyskiwane z nich surowce rolne – kwiatostany, nasiona i łodygi – mogą być wykorzystywane w różnych gałęziach przemysłu: spożywczym, kosmetycznym, włókienniczym,

motoryzacyjnym i papierniczym, a nawet budowlanym. Z kolei dzięki niskim wymaganiom glebowym i hydrologicznym roślina ta może być uprawiana praktycznie na każdym gruncie, bez stosowania nawozów i to niezależnie od warunków klimatycznych.

Rynek produktów wytworzonych na bazie konopi przemysłowych rozwija się w Polsce w tempie dwucyfrowym. Sprzyja temu duża otwartość konsumentów na nowe trendy, historyczne przywiązanie do medycyny naturalnej i otwarcie rolników na rozpoczęcie upraw konopi. Obszar zasiewów zgłoszonych przez gospodarstwa rolne do dopłat bezpośrednich przekroczył w 2019 roku 3 tys. hektarów – to wzrost o 78 proc. w ujęciu rocznym. Konopie przemysłowe stanowią 0,3 proc. całego areалу upraw roślin przemysłowych w Polsce.

Szybkiemu wzrostowi sprzedaży produktów konopnych nie towarzyszy jednak wystarczająco szybki wzrost areалу upraw *Cannabis sativa*, co prowadzi do wzrostu importu surowców rolnych pozyskiwanych z konopi i do słabego rozwoju firm przetwórczych zależnych od dostępu do lokalnie pozyskiwanych surowców (np. biokompozyty, domy z konopi, włókiennictwo).

Największą częścią polskiego przetwórstwa konopnego jest wytwarzanie produktów zawierających kannabidiol (CBD), w tym żywności, suplementów diety i kosmetyków. Według szacunków wartość przychodów firm zajmujących się przetwórstwem CBD przekroczyła w 2019 r. 210 mln zł, z czego za około dwie trzecie odpowiadały rodzime mikroprzedsiębiorstwa.

Doktorantka w swojej rozprawie podjęła się bardzo istotnego zadania, jakim była identyfikacja parametrów charakterystycznych w plonowaniu wybranych odmian konopi przemysłowych. Rozwiązanie tego zadania wymagało przeprowadzenia szeroko zakrojonych badań i analiz, dających odpowiedź na tak postawiony cel rozprawy. Aby zrealizować cel pracy konieczne było:

- zebranie i odpowiednie oznaczenie danych wejściowych do modelowania,
- wygenerowanie modeli SNN,
- scharakteryzowanie jakościowe i dokonanie oceny wrażliwości wygenerowanych modeli sieci neuronowych przy użyciu funkcji *Automatycznego Projektanta* oraz *Projektanta Sieci Użytkownika*,
- ustalenie rangi zmiennych wraz z nadaniem im odpowiedniej hierarchii.

Postawiony cel pracy uważam za poprawnie sformułowany, a postawioną hipotezę za prawidłowo sformułowaną (z drobną uwagą o której wspominałem w pkt. 2) dla podjętego tematu pracy doktorskiej i realizacji jej celu.

Autorka rozprawy trafnie uzasadnia podjęcie badań w tym temacie, wynikające z braku dostatecznie pogłębionych rozważań oraz dużego rozproszenia problematyki modelowania w literaturze przedmiotu. Istnieje wiele opracowań, jak podaje Doktorantka z tego tematu, ale nie ma praktycznych rozwiązań w zakresie omawianego tematu. Stąd też wybór tematu i zakresu jest w pełni uzasadniony.

Metodyka pracy jest przedstawiona w sposób jasny i czytelny, zgodnie z obowiązującymi zasadami metodologii badań naukowych. W sposób prawidłowy Doktorantka scharakteryzowała materiał badawczy, a następnie metodę badawczą w oparciu o zastosowane narzędzie w postaci sztucznych sieci neuronowych (Pakiet Statistica).

W pierwszej kolejności Doktorantka dokonała charakterystyki jakościowej i dokonała oceny wrażliwości modeli sieci neuronowych wytworzonych przy użyciu funkcji *Automatycznego Projektanta* oraz *Projektanta sieci Użytkownika*. W rezultacie na podstawie przeprowadzonych badań i opracowanego zbioru uczącego wygenerowano po 10 sieci dla każdego z modeli, a następnie, cytując „...wybrano **po 3 najlepsze sieci** dla każdego z modeli wygenerowanych przy użyciu każdej z dwóch powyższych funkcji.” Tutaj rodzi się pytanie jakie było kryterium wyboru najlepszej sieci?

Przyjęte modele są jak najbardziej prawidłowe w kontekście realizacji założonego celu pracy. Wyniki oceny jakościowej modeli wygenerowanych przy użyciu funkcji *Automatycznego Projektanta* przedstawiono w tabelach 9, 11 i 13, a przy użyciu funkcji *Projektanta Sieci Użytkownika* w tabelach 15, 17 i 18. Oceny dokonano w oparciu o następujące parametry: jakość uczenia, walidacyjna testowa oraz błąd uczenia, walidacyjny, testowy oraz parametr nigdzie nie wyjaśniony w pracy: **uczenie**, w postaci oznaczeń – np. **KM, KN**. Proszę o wyjaśnienie tego parametru i ww. oznaczeń.

Na podstawie danych zamieszczonych w tabelach 9, 11 i 13 można zauważyć, że parametr jakość uczenia mieścił się w granicach od 0,9626 dla modelu z dwoma zmiennymi wyjściowymi: kiełkowaniem i plonem nasion do 0,7659 dla modelu z jedną zmienną wyjściową: plon. Natomiast odwrotną sytuację można zaobserwować w przypadku jakości walidacji, gdzie przyjmuje ona wartość od 0,9756 do 1,2494. Dodatkowo wartość **1,2494** została w tabeli specjalnie oznaczona.

Biorąc pod uwagę sieci wygenerowane przy użyciu funkcji *Projektanta Sieci Użytkownika* stwierdzono, że parametr jakość uczenia mieścił się w granicach od 0,9847 dla modelu z dwoma zmiennymi wyjściowymi: kiełkowaniem i plonem nasion do 0,9898 dla modelu z jedną zmienną wyjściową: plon, a jakość walidacji kształtowała się w przedziale od 0,9934 do 0,9905.

Brakuje mi tutaj analizy uzyskanych wyników badań, lub też wyjaśnienia istotnych parametrów oceny modeli. Proszę o wyjaśnienie co oznacza jakość walidacyjna i dlaczego została ona specjalnie oznaczona w modelu wygenerowanym przy użyciu funkcji *Automatycznego Projektanta*.

Analiza wrażliwości zmiennych sieci niezależnie od użytej funkcji została przedstawiona w ujęciu tabelarycznym, w sposób prawidłowy i czytelny.

Na stronie 59, w akapicie poniżej tabeli 13 Autorka pracy pisze: „... W przypadku modeli z dwoma zmiennymi wyjściowymi: plon i kiełkowanie (RBF 15:41-3-2:2) oraz z jedną zmienną wyjściową: kiełkowanie (RBF 15:41-6-1:1), jakości sieci **uznano za dobre, a błędy za niskie**. Jednak w przypadku modelu z jedną zmienną wyjściową: plon, wygenerowane modele **wykazywały cechy przeuczenia**, w związku z czym postanowiono kontynuować badania przy użyciu funkcji *Projektanta Sieci Użytkownika*”. W pracy nie doszukałem się informacji na podstawie czego Autorka uznała jakość sieci za dobrą, jak też jakie parametry sieci wskazywały na przuczenie sieci? Proszę również o odpowiedź na to pytanie.

Kolejnym krokiem w celu rozwiązania problemu naukowego była ocena i analiza wrażliwości wytworzonych modeli neuronowych. W tabeli 20 Doktorantka przedstawiła w sposób syntetyczny wszystkie najistotniejsze parametry charakteryzujące wygenerowane sieci, a w tabeli 21 wyniki analizy wrażliwości zmiennych poszczególnych zbiorów uczących, które

brały udział w procesie modelowania neuronowego. Szkoda tylko, że Autorka ograniczyła się tylko do przepisania danych z tabeli w tekst opisujący poszczególne modele a nie wskazała i uzasadniła wyboru najlepszych i najgorszych rozwiązań modelowych lub różnicy w uzyskanych parametrach. Podobnie jest w przypadku tabeli 21, gdzie Autorka tylko sklasyfikowała określone zmienne dla poszczególnych modeli bez żadnego komentarza.

Przedstawiona Dyskusja w mojej opinii jest kontynuacją przeglądu literatury w zakresie zastosowania SNN w procesach produkcyjnych a nie dyskusją uzyskanych wyników z pracami o podobnej tematyce i zakresie. Jedyne ostatni akapit odnosi się do uzyskanych wyników gdzie Autorka pisze: „...Modele typu RBF wygenerowane w drugim etapie badań, z wykorzystaniem funkcji *Projektanta Sieci Użytkownika* charakteryzują się wysoką jakością rzędu: 97%–98%. Jest to jakość wyższa od jakości sieci wygenerowanych przez badaczy prowadzących badania nad plonowaniem np. trzciny cukrowej [Meder i in.2019] - 83,49% i porównywalna z jakością sieci wytworzonej przez zespół Ghandiego [Ghandi i in. 2016] badający plonowanie ryżu – model tych badaczy charakteryzował się jakością 97,5%”. Tutaj rodzi się kolejne pytanie: czy możemy porównywać ze sobą obiekty, które charakteryzują się innymi cechami?

Pracę kończy rozdział Podsumowanie oraz Wnioski, w których Doktorantka w sposób syntetyczny odnosi się do uzyskanych wyników badań, odpowiadając zarazem na postawiony cel pracy i problem badawczy. O ile Doktorantka odnosi się we wnioskach do rozwiązania problemu naukowego to nie zauważyłem wniosku odnoszącego się do weryfikacji postawionej hipotezy. Czy została ona pozytywnie czy negatywnie zweryfikowana.

Pomimo wymienionych uwag, uzyskane wyniki badań, ich opracowanie analityczne oraz wnioskowanie pozwalają na stwierdzenie, że Autorka zrealizowała postawiony w rozprawie cel naukowy, a wykorzystana metoda badawcza była właściwa i odpowiada aktualnemu stanowi wiedzy naukowej. Przedstawione w zakończeniu rozprawy wnioski dają odpowiedź na postawiony cel pracy. Oceniana praca posiada duże walory naukowe i praktyczne.

4. Uwagi szczegółowe

- s.23. „Suma ważona „ Σ ” jest algebraiczna sumą” – powinno być Suma ważona „ Σ ” jest **algebraiczną** sumą
- s. 23 „, Sieci są z zbudowane” – powinno być Sieci są **z** zbudowane
- s. 23. „, Warstwy mogą mieć dowolną ilość neuronów, ale należy zwrócić uwagę, iż przy zbyt dużej ich ilości sieć...” – powinno być Warstwy mogą mieć dowolną **liczbę** neuronów, ale należy zwrócić uwagę, iż przy zbyt dużej ich **liczbie** sieć...
- s. 25 Zebrane dane, aby mogły być wykorzystane w tworzeniu modeli SSN, zapisuje się w formacie CSV – czy zawsze należy zapisywać dane do format CSV?
- s. 25. „... z powodzeniem zostać użyta do w problemach” – powinno być z powodzeniem zostać użyta **do** w problemach...
- s. 34 „...Które cechy procesu uprawy konopi przemysłowych są **reprezentatywne...**” czy istotne?

- s. 49 Tabela 10 – kolumna przedplon – brak wyjaśnienia oznaczeń AB i AC
- s. 50 „...Z tak przygotowanego zbioru uczącego wygenerowano 3 modele SSN (Tab. 8)” – modele generuje się ze zbioru uczącego czy zbioru danych?
- s. 50 „... Przy użyciu Projektanta sieci użytkownika możliwa jest kontrola procesu modelowania neuronowego poprzez wybór rodzaju sieci, określenie wejść, warstw, **ilości** neuronów w warstwach, metody uczenia sieci i **warunków wyboru optymalnego modelu**” – powinno być **liczby**. Proszę o wyjaśnienie ostatniego z kryteriów: warunków wyboru optymalnego modelu.
- s. 55 „...Podczas tego etapu modelowania wygenerowano 3 modele SSN zawierających po 10 sieci. Wygenerowane SSN poddano analizie, a kryterium wyboru stanowiły: jakość uczenia, jakość walidacyjna, jakość testowa, błąd uczenia, błąd uczenia, błąd walidacyjny i błąd testowy” – powinno być Podczas tego etapu modelowania wygenerowano 3 modele SSN **zawierające** po 10 sieci. Wygenerowane SSN poddano analizie, a kryterium **wyboru** stanowiły: jakość uczenia, jakość walidacyjna, jakość testowa, ~~błąd uczenia~~, błąd uczenia, błąd walidacyjny i błąd testowy.
- s. 56 „... **Wybrano po 3 najlepsze sieci** dla każdego z modeli wygenerowanych przy użyciu każdej z dwóch powyższych funkcji” – proszę wymienić kryteria wyboru najlepszych sieci
- s. 64 „...wartością błędu RMS...” – powinno być wartością błędu **RMSE**
- s. 65 „...Do procesu uczenia sieci użyto algorytmów KM, KN i PI”. – proszę wyjaśnić skróty KM i KN
- s. 66 „...Ważnym punktem w procesie modelowania neuronowego jest przeprowadzenie analizy wrażliwości, podczas którego dla poszczególnych...” – powinno być „Ważnym punktem w procesie modelowania neuronowego jest przeprowadzenie analizy wrażliwości, podczas **której** dla poszczególnych”
- s. 66 „... Uważa się, że w przypadkach, w których iloraz błędu jest mniejszy lub równy ~~od~~ jedności...”

Proszę, aby Autorka pracy w trakcie obrony ustosunkowywała się do uwag zawartych w recenzji, a także uwag szczegółowych. Omówione niedociągnięcia, drobne potknięcia edycyjne oraz uwagi nie umniejszają merytorycznej wartości pracy, którą oceniam wysoko. Uwzględnienie przez Doktorantkę wymienionych w niniejszej recenzji uwag pozwoli na udoskonalenie warsztatu pisarskiego oraz pozwoli na uniknięcie różnych uchybień i niedociągnięć na etapie przygotowania publikacji, bądź referatów konferencyjnych.

5. Ocena końcowa

Recenzowana rozprawa doktorska dotyczy identyfikacji parametrów charakterystycznych w plonowaniu nasion odmiany konopi przemysłowych (*Cannabis sativa* L.) z użyciem metod sztucznej inteligencji. Na podkreślenie zasługuje tutaj połączenie wiedzy i doświadczenia Doktorantki z zakresu agronomii (technologia uprawy konopi przemysłowych) i inżynierii rolniczej (modelowania procesów z wykorzystaniem SNN), co w efekcie pozwoliło

na zidentyfikowanie podstawowych parametrów wpływających na plonowanie nasion konopi przemysłowych.

Mgr inż. Dominika Sieracka wykazała się odpowiednią wiedzą teoretyczną i praktyczną w dyscyplinie naukowej inżynieria mechaniczna (w obszarze inżynieria rolnicza) oraz wysokimi umiejętnościami samodzielnego planowania oraz realizacji badań naukowych.

Biorąc pod uwagę walory naukowe, poznawcze oraz aplikacyjne recenzowanej rozprawy doktorskiej pt. „Identyfikacja parametrów charakterystycznych w plonowaniu nasion odmiany konopi przemysłowych (*Cannabis sativa L.*) z użyciem metod sztucznej inteligencji” stwierdzam, że spełnia ona wszystkie wymagania określone w art. 13, ust. 1 stawiane rozprawom doktorskim zawarte w Ustawie z dnia 14 marca 2003 roku o stopniach i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2017 r. poz. 1789), art. 179 Ustawy z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce (Dz.U. poz. 1669) oraz wnoszę o dopuszczenie mgr inż. Dominiki Sierackiej do publicznej obrony przed Radą dyscypliny inżynieria mechaniczna Uniwersytetu Przyrodniczego w Poznaniu.

Kraków, 12.01.2023 r



.....
Prof dr hab. inż. Maciej Kuboń