POLITECHNIKA POZNAŃSKA

INSTYTUT INFORMATYKI
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
ul. Piotrowo 2, 60-965 Poznań
tel.: +48 (61) 665 2997, +48 (61) 665 2999
e-mail: office_cs@put.poznan.pl
www.cs.put.poznan.pl

dr hab. inż. Mirosław Ochodek, prof. uczelni
Instytut Informatyki,
Wydział Informatyki i Telekomunikacji
Politechnika Poznańska,
tel. 61 665 2944
e-mail: miroslaw.ochodek@put.poznan.pl

Poznań, 28.02.2026

RECENZJA ROZPRAWY DOKTORSKIEJ

Tytuł rozprawy: **An Interactive Machine Learning-Based Multi-objective Optimization Framework for Software Overtime Planning**
Autor rozprawy: **mgr inż. Hammed Adeleye Mojeed**
Promotor rozprawy: **dr hab. inż. Rafał Szłapczyński**
Dyscyplina: **Informatyka techniczna i telekomunikacja**

Cel i zakres tematyczny rozprawy

Celem badań przedstawionych w rozprawie doktorskiej było opracowanie nowego, interaktywnego podejścia/algorytmu (ang. *framework*) optymalizacji wielokryterialnej opartego na uczeniu maszynowym dla problemu planowania nadgodzin w projektach IT. Podejście ma uwzględniać preferencje kierowników projektów poprzez wykorzystanie modeli uczenia maszynowego do predykcji tych preferencji, co pozwala na automatyzację procesu optymalizacji przy jednoczesnym zachowaniu zgodności z rzeczywistymi potrzebami decydentów. Zdefiniowany cel badawczy łączy ze sobą aspekty istotne z perspektywy zarządzania projektami informatycznymi z zaawansowanymi technikami optymalizacji wielokryterialnej oraz uczenia maszynowego, stanowiąc odpowiedź na praktyczne wyzwania związane z opóźnieniami projektów.

W rozprawie nie sformułowano klasycznej tezy badawczej, określając w zastępstwie cztery centralne pytania badawcze, które można przetłumaczyć w następujący sposób:

- **RQ1:** Jak skutecznie model predykcyjny uczenia maszynowego może uchwycić i odwzorować subiektywne preferencje kierowników projektów w planowaniu nadgodzin?
- **RQ2:** W jakim stopniu integracja modelu uczenia maszynowego z interaktywnym algorytmem optymalizacji wielokryterialnej może zastąpić ciągłą interakcję z człowiekiem w pętli?
- **RQ3:** Jak rozwiązania planowania nadgodzin wygenerowane przez interaktywne podejście oparte na uczeniu maszynowym mają się do rozwiązań otrzymywanych przy użyciu tradycyjnych metod uwzględniających człowieka w pętli w zakresie nadgodzin, kosztów i jakości projektu?
- **RQ4:** Jak skalowalne i adaptowalne jest zaproponowane w rozprawie podejście do optymalizacji oparte na uczeniu maszynowym przy zastosowaniu go w różnorodnych środowiskach projektów IT?

POLITECHNIKA POZNAŃSKA

INSTYTUT INFORMATYKI
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
ul. Piotrowo 2, 60-965 Poznań
tel.: +48 (61) 665 2997, +48 (61) 665 2999
e-mail: office_cs@put.poznan.pl
www.cs.put.poznan.pl

Zakres rozprawy obejmuje logicznie powiązany ze sobą cykl badań rozpoczynający się od przeglądu literatury w postaci tzw. systematycznego mapowania literatury (ang. *literature mapping study*) identyfikującego luki badawcze dotyczące integracji technik optymalizacji i uczenia maszynowego w obszarze planowania projektów IT, poprzez opracowanie i walidację modelu uczenia maszynowego do predykcji preferencji kierowników projektów, zaprojektowanie i implementację interaktywnego podejścia optymalizacyjnego ML-iMOSFLA, aż po empiryczną walidację skuteczności tego podejścia na zbiorze sześciu projektach IT.

Biorąc pod uwagę wcześniejsze badania dotyczące planowania projektów IT oraz badania dotyczące zastosowań uczenia maszynowego i optymalizacji wielokryterialnej w inżynierii oprogramowania należy podkreślić, że zaproponowany cel badawczy jest interesujący i nowatorski. Szczególnie wartościowym aspektem jest integracja trzech rzadko łączonych obszarów w obszarze planowania projektów IT: optymalizacji wielokryterialnej, uczenia maszynowego oraz interaktywnej optymalizacji uwzględniającej preferencje decydenta. Wkład wynikający z przeprowadzonych badań empirycznych jest cenny dla rozwoju dyscypliny, zarówno w aspekcie teoretycznym (nowe podejście do optymalizacyjny, systematyczne mapowanie literatury), jak i praktycznym (podejście wspomagające podejmowanie decyzji w zarządzaniu projektami, walidacja na rzeczywistych projektach).

Zawartość i układ pracy

Recenzowana rozprawa doktorska napisana jest w języku angielskim, obejmuje 145 stron i podzielona jest na 6 rozdziałów merytorycznych, odnośniki literaturowe oraz trzy dodatki. Rozprawa zachowuje jednorodny, spójny i kompletny charakter pozycji książkowej.

Pierwsze dwa rozdziały rozprawy stanowią wprowadzenie. Rozdział 1 zarysowuje problem badawczy, motywację oraz cele i główne pytania badawcze rozprawy, omawiając także jej strukturę. Rozdział 2 zawiera obszerny przegląd literatury, w tym systematyczne mapowanie **71 artykułów z lat 2012-2022** dotyczących zastosowań optymalizacji wielokryterialnej i uczenia maszynowego w planowaniu projektów IT. W ramach tego rozdziału doktorant identyfikuje kluczową lukę badawczą – tylko **4 prace (5.63%)** synergistycznie integrują optymalizację wielokryterialną i uczenie maszynowe, przy czym **wszystkie w obszarze szacowania pracochłonności (SEE), a żadne w kontekście planowania harmonogramu (SPS) czy nadgodzin (SOP)**. Dodatkowo metody interaktywne uwzględniające preferencje kierowników projektów w planowaniu nadgodzin są całkowicie nieobecne w literaturze.

W rozdziale 3 przedstawiono metodologię badań. Doktorant opisuje zaproponowane podejście (framework) **ML-iMOSFLA** składające się z trzech głównych komponentów: modelu uczenia maszynowego, interaktywnego algorytmu optymalizacji wielokryterialnej oraz modułu integrującego oba komponenty. Zaprezentowano matematyczne sformułowanie problemu planowania nadgodzin jako zadanie optymalizacji 4-kryterialnej. Opisano również dane projektowe wykorzystane w badaniach (6 projektów IT o różnej wielkości: od 185 do 635 punktów funkcyjnych, z łączną liczbą 102 pakietów roboczych) oraz metryki oceny skuteczności algorytmu.

W rozdziale 4 przedstawiono szczegółowe wyniki badań dotyczących opracowania i walidacji **modelu uczenia maszynowego** do predykcji preferencji kierowników projektów dla planów nadgodzin. Opisano proces przygotowania zbioru danych obejmujący wygenerowanie rozwiązań za pomocą algorytmu MOSFLA, ich

POLITECHNIKA POZNAŃSKA

INSTYTUT INFORMATYKI
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
ul. Piotrowo 2, 60-965 Poznań
tel.: +48 (61) 665 2997, +48 (61) 665 2999
e-mail: office_cs@put.poznan.pl
www.cs.put.poznan.pl

ocenę przez **20 kierowników projektów** (łącznie 1622 rozwiązania po usunięciu wartości odstających) oraz porównanie **8 modeli uczenia maszynowego** spośród których algorytm Random Forest Regression (RFR) wykazał najlepszą skuteczność. Następnie przeprowadzono optymalizację hiperparametrów tego modelu za pomocą metody **Greedy Halving Grid Search**. Doktorant przeprowadził także analizę interpretowalności modelu za pomocą metody SHAP (SHapley Additive exPlanations), analizę skalowalności modelu względem złożoności projektów oraz walidację krzyżową między projektami.

Rozdział 5 koncentruje się na **empirycznej walidacji skuteczności ML-iMOSFLA**. Przedstawiono wyniki 30 niezależnych uruchomień algorytmu dla każdego z 6 projektów oraz przeprowadzono statystyczne porównanie z algorytmem bazowym MOSFLA oraz z algorytmem interaktywnym wymagającym ciągłego udziału człowieka. Wyniki wykazały znaczącą przewagę ML-iMOSFLA nad MOSFLA w 16 z 18 porównań wskaźników jakości. W porównaniu z algorytmem wymagającym ciągłego udziału człowieka zaobserwowano, że ML-iMOSFLA osiąga statystyczny parytet przy 150-200 iteracjach, jednocześnie eliminując potrzebę ciągłej interakcji z ekspertem.

Rozdział 6 zamyka rozprawę podsumowując odpowiedzi na pytania badawcze, osiągnięte cele badawcze, wkład naukowy (teoretyczny i praktyczny), ograniczenia przeprowadzonych badań oraz propozycje przyszłych kierunków rozwoju.

Ocena zastosowanego piśmiennictwa w ramach rozprawy doktorskiej

Dobór literatury w recenzowanej rozprawie doktorskiej jest kompleksowy i obejmuje 223 odniesienia literaturowe. Bibliografia obejmuje zarówno prace dotyczące optymalizacji wielokryterialnej, uczenia maszynowego, jak i planowania projektów IT. Szczególnie wartościowe jest włączenie wyników systematycznego mapowania literatury obejmującego 71 artykułów z lat 2012-2022, co świadczy o gruntownym rozpoznaniu aktualnego stanu wiedzy w badanym obszarze. Literatura została poprawnie wyselekcjonowana, aby wspierać poszczególne etapy badań – od uzasadnienia problemu badawczego, poprzez wybór odpowiednich metod, aż po określenie metryk oceny skuteczności algorytmu. Bibliografia jest aktualna i obejmuje zarówno klasyczne pozycje z zakresu optymalizacji wielokryterialnej i uczenia maszynowego, jak i nowsze publikacje dotyczące zastosowań tych technik w inżynierii oprogramowania. Można zatem stwierdzić, że bogaty zakres prezentowanej bibliografii świadczy o gruntownym przygotowaniu teoretycznym doktoranta.

Ocena doboru i poprawności zastosowania metod badawczych

Recenzowana rozprawa oparta jest o empiryczne metody badawcze. Badanie przedstawione w rozdziale 2 przeprowadzono zgodnie z metodą systematycznego mapowania literatury według wytycznych zaproponowanych przez Petersena, Vakkalanke oraz Kuzniarza. Proces ten został przeprowadzony systematycznie, chociaż należy zaznaczyć pewne braki w raportowaniu szczegółów metodologicznych (omówione dalej w rozdziale poświęconym uwagom do pracy).

W rozdziale 4, dotyczącym opracowania modelu uczenia maszynowego, zastosowano klasyczne metody empiryczne z obszaru uczenia maszynowego, w tym walidację krzyżową (10-fold cross-validation), porównanie wielu alternatywnych modeli predykcyjnych oraz optymalizację hiperparametrów. Doktorant

POLITECHNIKA POZNAŃSKA

INSTYTUT INFORMATYKI
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
ul. Piotrowo 2, 60-965 Poznań
tel.: +48 (61) 665 2997, +48 (61) 665 2999
e-mail: office_cs@put.poznan.pl
www.cs.put.poznan.pl

przeprowadził także walidację modelu obejmującą analizę interpretowalności za pomocą metody SHAP oraz analizę skalowalności względem złożoności projektów.

Jeśli chodzi o ewaluację empiryczną opisaną w rozdziale 5, doktorant zastosował schemat badania z wielokrotnymi powtórzeniami (30 niezależnych uruchomień per algorytm per projekt) oraz statystyczne testowanie hipotez przy użyciu testów statystycznych.

Uzyskane wyniki i ich znaczenie dla rozwoju dyscypliny naukowej oraz praktyki

W pracy przedstawiono szereg wartościowych wyników badawczych. Do najbardziej interesujących w moim odczuciu należy systematyczne mapowanie literatury identyfikujące lukę w integracji technik optymalizacji i uczenia maszynowego w obszarze planowania projektów IT, opracowanie i walidacja modelu osiągającego wysoką skuteczność predykcji preferencji kierowników projektów oraz empiryczna demonstracja, że zaproponowane podejście ML-iMOSFLA stanowi perspektywiczną alternatywę dla podobnych metod uwzględniających człowieka w pętli decyzyjnej.

Uwagi krytyczne i pytania

W moim przekonaniu doktorant wykazał się dbałością o szczegóły w trakcie realizacji prowadzonych badań, niemniej jednak lektura rozprawy skłania mnie do zadania kilku pytań oraz sformułowania kilku ogólnych uwag co do jej treści.

1. **Reproduktywne badania.** Istotną cechą prawidłowo wykonanych badań empirycznych jest ich przejrzystość oraz reprodukowalność. W związku z tym intensywnie promowaną praktyką badawczą w obszarze empirycznej inżynierii oprogramowania jest uzupełnianie badań o tzw. pakiety reprodukcyjne, pozwalające zweryfikować narracje autorów prac naukowych korzystając bezpośrednio z materiałów źródłowych. Myślę, że doktorant mógł się osobiście przekonać jak ważne jest takie podejście – ponieważ w swoich badaniach bazuje na zbiorze danych projektowych udostępnionym przez autorów wcześniejszych prac właśnie w duchu otwartej nauki. W rozprawie brak jest niestety odnośników do repozytoriów z danymi, kodem źródłowym algorytmów czy szczegółowymi protokołami badań będących rezultatem prac badawczych nad rozprawą doktorską. W rozdziale 5.1 doktorant wspomina o implementacji narzędzia w języku Java z integracją biblioteki Weka, jednak brak jest odnośnika do repozytorium kodu tego projektu. W moim odczuciu uzupełnienie tych materiałów byłoby niezwykle wartościowe i znacząco zwiększyłoby przejrzystość oraz możliwość weryfikacji przedstawionych wyników.
2. **Rozbieżności z deklarowanymi wytycznymi w systematycznym mapowaniu literatury.** Doktorant jawnie deklaruje stosowanie wytycznych do tego typu przeglądów sformułowanych przez Petersena, Vakkalanke oraz Kuzniarza (odnośnik [101]). W przeprowadzonym systematycznym mapowaniu literatury brakuje kilku elementów metodologicznych rekomendowanych lub wręcz oczekiwanych przez te (oraz podobne) wytyczne. (1) Brak informacji czy proces selekcji i oceny artykułów prowadził doktorant samodzielnie czy z drugim recenzentem. (2) Opis procedury wyszukiwania metodą kuli śnieżnej (ang. *snowballing*) jest niejasny. Wprawdzie w rozprawie jest opisane czym jest snowballing, ale brakuje informacji o tym jak został on zaaplikowany (czy był to tzw. forward

POLITECHNIKA POZNAŃSKA

INSTYTUT INFORMATYKI
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
ul. Piotrowo 2, 60-965 Poznań
tel.: +48 (61) 665 2997, +48 (61) 665 2999
e-mail: office_cs@put.poznan.pl
www.cs.put.poznan.pl

snowballing, backward snowballing, a może oba?). Brakuje także opisu kryteriów zatrzymania dla snowballingu. (3) Brak również informacji czy strategia wyszukiwania była testowana na zbiorze znanych istotnych prac (tzw. złotym zbiorze). (4) W pracy brakuje pełnych tabel ekstrakcji danych, np. w formie załącznika (co jest zalecane przez wspomniane wytyczne dla zachowania pełnej transparentności). (5) Nie jest też jasne czy kryteria włączenia/wyłączenia były łączone operatorem AND czy OR. (6) Wreszcie, co szczególnie istotne z mojej perspektywy – brak systematycznej dyskusji zagrożeń dla wiarygodności badania (ang. *validity threats*), co jest kluczowe dla komunikacji wyników badań prowadzonych metodami empirycznymi.

3. **Problem zmiennej liczby cech w modelu uczenia maszynowego.** Analizowane projekty mają różną liczbę pakietów roboczych (między 10 a 27 jak przedstawiono w tabeli 4.6), co przekłada się na różną liczbę cech reprezentujących dane wejściowe w modelu uczenia maszynowego. Zgodnie z tabelą 4.5 cechy te reprezentują alokację nadgodzin per pakiet roboczy. Klasyczne modele uczenia maszynowego rozważane w rozprawie wymagają stałej liczby cech wejściowych. Zatem model predykcyjny dla projektu ACAD oczekiwałby 10 cech, a model dla projektu PARM oczekuje 27 cech, a zatem nie można użyć jednego modelu dla wszystkich projektów bez transformacji danych. W rozprawie brak jakiejkolwiek informacji jak ten problem został rozwiązany. Z mojej perspektywy jest to niezwykle istotna kwestia, ponieważ jeśli model musiałby być trenowany zawsze od zera w ramach konkretnego projektu, aby zamodelować preferencje kierownika projektu to korzyści z jego użycia są przynajmniej mocno wątpliwe.
4. **Proces walidacji modeli uczenia maszynowego – niejasności wokół zbioru danych i procedury.** W rozdziale 4.3 (str. 77) doktorant deklaruje: "*The dataset for each project was split into an 80% training set and a 20% test set*". Jednak tabele 4.7-4.12 pokazują tylko wyniki walidacji krzyżowej na zbiorze treningowym. Pojawia się zatem pytanie do czego został użyty zbiór 20% przypadków testowych? Jedno ze standardowych podejść do oceny jakości predykcji modeli uczenia maszynowego zakłada wybór modelu (i jego parametrów) na podstawie walidacji krzyżowej oraz jego ostateczną ocenę na zbiorze testowym (hold-out) co pozwala lepiej ocenić zdolności generalizacji modelu. Niestety w pracy nie znalazłem ponownego nawiązania do zbioru testowego. Dodatkowo w kontekście zastosowania metody walidacji krzyżowej na zbiorze treningowych brak jest także informacji o sposobie podziału zbioru treningowego oraz czy testowano różne ziarna losowości dla oceny odporności wyników.
5. **Brak informacji o hiperparametrach modeli bazowych.** W rozdziale 4.3 porównano 8 modeli uczenia maszynowego (tabele 4.7-4.12). Niestety brakuje informacji o konfiguracji hiperparametrów poszczególnych z nich. Dopiero w rozdziale 4.5, już po przeprowadzonym porównaniu, doktorant ujawnia, że model RFR używany wcześniej miał ustawione „domyślne" parametry ("*default RFR configurations*" - str. 90) i był testowany bez ich strojenia ("*without tuning*" - str. 89). Zatem najprawdopodobniej wszystkie 8 modeli w tym porównaniu używało domyślnych parametrów (czyli jakich?). Co istotne, doktorant sam podkreśla kluczowe znaczenie strojenia hiperparametrów: "*These improvements underscore the crucial role of advanced hyperparameter optimization in overcoming the limitations of default RFR configurations, which frequently suffer from suboptimal parameter selection*" (str. 90). Pojawia się zatem istotne pytanie, czy jeśli pozostałe modele zostałyby poddane strojeniu wynik tej analizy byłby taki sam (zwłaszcza biorąc pod uwagę, że wrażliwość na

POLITECHNIKA POZNAŃSKA

INSTYTUT INFORMATYKI
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
ul. Piotrowo 2, 60-965 Poznań
tel.: +48 (61) 665 2997, +48 (61) 665 2999
e-mail: office_cs@put.poznan.pl
www.cs.put.poznan.pl

dobór hiperparametrów może być bardzo różna w zależności od algorytmu uczenia i architektury modelu)?

6. **Potencjalne zanieczyszczenie zbioru w badaniu oceniającym ML-iMOSFLA.** W rozdziale 5.2.1 doktorant pisze o tym, że rozwiązania wygenerowane przez algorytm zostały poddane ocenie eksperckiej: "*re-evaluating all solutions produced by the algorithm using PM assessments*". Pojawia się w tym miejscu pytanie: kim byli kierownicy projektów, którzy wzięli udział w tym badaniu? Fragment rozprawy na str. 106 cyt. „*To compute MoSD, a target solution has to be established. To do this, the standard MOSFLA algorithm was applied, and* ***the 20 participant PMs*** *ranked the non-dominated solutions generated*" wydaje się sugerować, że mogły to być te same osoby, które brały udział we wcześniejszym badaniu związanym z trenowaniem i oceną modeli uczenia maszynowego (rozdział 4.2.3). Jeśli faktycznie tak było to preferencje oraz oceny tych osób posłużyłyby zarówno do trenowania modeli oraz do oceny porównawczej między MOSFLA a ML-iMOSFLA, a to stanowiłoby istotne zagrożenie do poprawności tego badania. Preferencje tych kierowników byłyby niejako „zaszyte" w modelach co zapewne przełożyłoby się na ocenę jakości rozwiązań. Niestety w rozdziale 5.2.1 brakuje szczegółowych informacji pozwalających rozwiać te wątpliwości.
7. **Niekompletne raportowanie statystyczne w porównaniach algorytmów.** W rozdziale 5.2.2 przeprowadzono statystyczne porównania ML-iMOSFLA vs MOSFLA używając testu Mann-Whitney U (α=0.05), jednak raportowanie wyników jest niekompletne. Brak informacji: (1) czy test był jednostronny (hipoteza alternatywna: ML-iMOSFLA > MOSFLA) czy dwustronny (hipoteza alternatywna: ML-iMOSFLA ≠ MOSFLA). (2) Brak również miar wielkości efektu. (3) Dodatkowo przeprowadzono 18 porównań statystycznych (tabela 5.3) na podstawie, których wyciągnięto wniosek o wyższości metody ML-iMOSLFA nad MOSFLA bez zastosowania korekcji z uwagi na wielokrotne testowanie hipotez. Przy poziomie istotności α=0.05 i 18 niezależnych testach prawdopodobieństwo uzyskania przynajmniej jednego fałszywie pozytywnego wyniku wynosi około 60%. Zatem pojawia się ryzyko, że niektóre wyniki uznane za "statystycznie istotne" mogą być artefaktem wielokrotnego testowania. Stąd zasadne wydaje się pytanie - czy wyniki pozostają nadal istotne statystycznie po zastosowaniu odpowiedniej korekcji?
8. **Niewystarczające uzasadnienie wyboru punktu odniesienia do oceny ML-iMOSFLA.** Doktorant porównuje zaproponowane podejście ML-iMOSFLA z dwoma algorytmami: bazowym MOSFLA oraz interaktywnym HIL-iMOSFLA. Wybór MOSFLA jako punktu odniesienia jest uzasadniony odesłaniem do pracy [30], której doktorant jest także pierwszym autorem: "*MOSFLA is chosen... because it has been shown in a previous study [30] to outperform NSGA-II*" (str. 61-66). Jednak badania w pracy [30] prowadzone były na tym samym zbiorze 6 projektów, co stawia pod znakiem zapytania uniwersalność tego wniosku. Czy zatem faktycznie zasadne było odrzucenie porównania z NSGA-II, skoro także wedle dokonanego przeglądu literatury jest to algorytm bardziej popularny niż algorytm MOSFLA?
9. W pracy można znaleźć także drobne niejasności oraz problemy metodologiczne, które nie mają jednak istotnego wpływu na zawartość i ostateczną jakość merytoryczną samej rozprawy, np.: (1) w opisie podejścia doktorant używa sformułowania "*instances of initial optimal solutions*" – co w kontekście optymalizacji wielokryterialnej, jeśli pozostaje niewyjaśnione budzi pewne wątpliwości, m.in., jeśli już „optymalne" to po co nad nim dalej pracować. (2) Często używany jest przymiotnik

POLITECHNIKA POZNAŃSKA

INSTYTUT INFORMATYKI
WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
ul. Piotrowo 2, 60-965 Poznań
tel.: +48 (61) 665 2997, +48 (61) 665 2999
e-mail: office_cs@put.poznan.pl
www.cs.put.poznan.pl

"*significant*" ("*significantly lower*", "*significantly improving*") w sytuacjach, kiedy nie są stosowane metody wnioskowania statystyczne co może rodzić wątpliwości czytelnika - czy mowa jest o istotności statystycznej czy o subiektywnej ocenie doktoranta. (3) Na str. 31 pojawia się także błędne odniesienie do załącznik A jako listy publikacji wybranych w ramach przeglądu literatury.

Podsumowanie

Recenzowana rozprawa stanowi interesujący wkład w obszar zarządzania projektami IT oraz zastosowań uczenia maszynowego i optymalizacji wielokryterialnej w inżynierii oprogramowania. Autor wykazał umiejętność prowadzenia badań naukowych obejmujących różne etapy procesu badawczego: od systematycznego przeglądu literatury i formułowania problemu badawczego, poprzez projektowanie i implementację rozwiązania, aż po przeprowadzenie walidacji empirycznej i interpretację wyników.

Należy docenić wysiłek organizacyjny związany z przeprowadzeniem badań empirycznych wymagających zaangażowania 20 kierowników projektów do oceny rozwiązań oraz koordynacji eksperymentów. Wyniki badań prezentowanych w rozprawie były publikowane w czasopismach i prezentowane na konferencjach naukowych co dodatkowo wzmacnia recenzowaną rozprawę.

Zaproponowany „framework" ML-iMOSFLA przedstawia potencjalnie użyteczne podejście do wspomagania decyzji w planowaniu nadgodzin w projektach IT, oferując alternatywę dla metod wymagających ciągłej interakcji z decydentem. Praktyczna wartość rozwiązania wymaga jednak dalszej weryfikacji na szerszym zbiorze projektów oraz w zróżnicowanych kontekstach organizacyjnych.

Biorąc pod uwagę powyższe, stwierdzam, iż przedstawiona do recenzji rozprawa doktorska autorstwa mgr inż. Hammeda Adeleye Mojeeda pt. "An Interactive Machine Learning-Based Multi-objective Optimization Framework for Software Overtime Planning" spełnia wymogi formalne stawiane rozprawom doktorskimi i wnoszę o jej dopuszczenie do publicznej obrony.

Dr hab. inż. Mirosław Ochodek, prof. uczelni

Politechnika Poznańska