

Politechnika Wroclawska
Wybrzeże Wyspiańskiego 27
50-370 Wrocław
Dyscyplina naukowa: **Informatyka Techniczna
i Telekomunikacja**
za pośrednictwem:
Rady Doskonałości Naukowej
pl. Defilad 1
00-901 Warszawa
(Pałac Kultury i Nauki, p. XXIV, pok. 2401)

Mariusz Topolski
Politechnika Wroclawska
Wydział Informatyki i Telekomunikacji
Katedra Systemów i Sieci Komputerowych
Wybrzeże Wyspiańskiego 27
50-370 Wrocław

Wniosek

z dnia 28.08.2023

o przeprowadzenie postępowania w sprawie nadania stopnia doktora habilitowanego w dziedzinie **Nauk inżyniersko-technicznych** w dyscyplinie¹ **Informatyka techniczna i telekomunikacja**.

Określenie osiągnięcia naukowego będącego podstawą ubiegania się o nadanie stopnia doktora habilitowanego: monografia naukowa zatytułowana „**Metody ekstrakcji cech w uczeniu maszynowym. Nowe trendy inżynierii cech**”.

Wnoszę – na podstawie art. 221 ust. 10 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 zm.) – aby komisja habilitacyjna podejmowała uchwałę w sprawie nadania stopnia doktora habilitowanego w głosowaniu **jawnym**².

Zostałem poinformowany, że:

Administratorem w odniesieniu do danych osobowych pozyskanych w ramach postępowania w sprawie nadania stopnia doktora habilitowanego jest Przewodniczący Rady Doskonałości Naukowej z siedzibą w Warszawie (pl. Defilad 1, XXIV piętro, 00-901 Warszawa).

Kontakt za pośrednictwem e-mail: kancelaria@rdn.gov.pl, tel. 22 656 60 98 lub w siedzibie organu.

Dane osobowe będą przetwarzane w oparciu o przesłankę wskazaną w art. 6 ust. 1 lit. c) Rozporządzenia UE 2016/679 z dnia z dnia 27 kwietnia 2016 r. w związku z art. 220 - 221 oraz art. 232 – 240 ustawy z dnia 20 lipca 2018 roku - Prawo o szkolnictwie wyższym i nauce, w celu przeprowadzenia postępowania o nadanie stopnia doktora habilitowanego oraz realizacji praw i obowiązków oraz środków odwoławczych przewidzianych w tym postępowaniu.

Szczegółowa informacja na temat przetwarzania danych osobowych w postępowaniu dostępna jest na stronie www.rdn.gov.pl/klauzula-informacyjna-rodo.html



.....
(podpis wnioskodawcy)

¹ Klasyfikacja dziedzin i dyscyplin wg. rozporządzenia Ministra Nauki i Szkolnictwa Wyższego z dnia 20 września 2018 r. w sprawie dziedzin nauki i dyscyplin naukowych oraz dyscyplin w zakresie sztuki (Dz. U. z 2018 r. poz. 1818).

² Niepotrzebne skreślić.

Załączniki:

1. Dane wnioskodawcy
2. Kopia dokumentu potwierdzającego posiadanie stopnia doktora
3. Autoreferat wnioskodawcy
4. Wykaz osiągnięć naukowych
5. Deklaracje współautorów dotyczące wspólnych artykułów
6. Monografia wchodząca w skład osiągnięcia naukowego „Metody ekstrakcji cech w uczeniu maszynowym. Nowe trendy inżynierii cech”. Monografia jest podzielona na dwie części 6a i 6b.

Autoreferat wnioskodawcy

1 Imię i nazwisko

Mariusz Topolski

2 Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania oraz tytułu rozprawy doktorskiej

2008 Stopień doktora nauk technicznych w dyscyplinie Informatyka.

Nadany uchwałą Rady Wydziału Elektroniki Politechniki Wrocławskiej.

Tytuł w języku polskim: *Komputerowe algorytmy rozpoznawania sekwencyjnego z modelem łączącym teorię ewidencji matematycznej z teorią zbiorów rozmytych*

Promotor: *prof. dr hab. inż. Marek Kurzyński*

2003 Tytuł zawodowy magistra inżyniera na kierunku: Fizyka techniczna, w zakresie specjalizacji Inżynieria biomedyczna - Zastosowanie komputerów w medycynie.

Wydział Podstawowych Problemów Techniki Politechniki Wrocławskiej.

Tytuł pracy magisterskiej: *Algorytmy generowania reguł rozmytych z danych empirycznych - zastosowanie w komputerowo wspomaganiej diagnostyce medycznej*

Promotor: *prof. dr hab. inż. Marek Kurzyński*

3 Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych

03.2019 – obecnie **Adiunkt**

Katedra Systemów i Sieci Komputerowych

Wydział Informatyki i Telekomunikacji

Politechnika Wroclawska

(do roku 2021 w Katedrze Systemów i Sieci Komputerowych Wydziału Elektroniki PWr)

10.2014 – 02.2019 **Adiunkt**

Wyższa Szkoła Bankowa we Wrocławiu

10.2013 – 09.2014 **Adiunkt**

Wrocławska Wyższa Szkoła Informatyki Stosowanej
(w trakcie zatrudnienia pełniłem funkcje prorektora i rektora)

08.2004 – 09.2013 **Adiunkt**

Międzynarodowa Wyższa Szkoła Logistyki i Transportu
(do 2008 roku asystent, a od 2008 adiunkt)

4 *Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.)*

4.1 *Tytuł osiągnięcia naukowego*

W ramach niniejszego wniosku habilitacyjnego prezentowane jest osiągnięcie w formie monografii naukowej wydanej przez wydawnictwo *Akademicka Oficyna Wydawnicza Exit*, które w roku opublikowania monografii w ostatecznej formie jest ujęte w wykazie sporządzonym przez Ministerstwo Nauki i Szkolnictwa Wyższego i posiada Unikatowy Identyfikator Wydawnictwa 14300.

Metody ekstrakcji cech w uczeniu maszynowym. Nowe trendy inżynierii cech

4.2 *Wykaz zawartości monografii naukowej*

Zamieszczony spis głównych rozdziałów stanowi zakres monografii naukowej.

Wykaz ważniejszych symboli i skrótów

1. Wstęp
2. Wprowadzenie

Część I Metody ekstrakcji bazujące na składowych głównych

3. Modyfikacja metody składowych głównych uwzględniająca rotację czynników wg centroidów klas
4. Zastosowanie metody gradientów stochastycznych do szacowania składowych głównych
5. Zastosowanie metody składowych głównych w zadaniu imputacji danych

Część II Metody ekstrakcji cech sygnałów o charakterystyce czasowej oraz dwuwymiarowych

6. Metoda szacowania liczby cech i komponentów w analizie składowych głównych
7. Model fuzji gradientów stochastycznych i składowych głównych w zadaniu ekstrakcji cech obrazu
8. Zastosowanie transformaty falkowej oraz analizy składowych niezależnych w zadaniu ekstrakcji cech sygnałów

Część III Zastosowanie metod inżynierii cech w zadaniu optymalizacji wielokryterialnej

9. Statystyczne metody zwiększania mocy dyskryminacyjnej w zadaniu ekstrakcji cech w ujęciu metaheurystyk
10. Ekstrakcja cech z wykorzystaniem zmodyfikowanego wielokryterialnego algorytmu genetycznego
11. Zastosowanie metod optymalizacji wielokryterialnej do budowy zespołów klasyfikatorów

Zakończenie

Bibliografia

5 Omówienie celu naukowego wyżej wymienionych prac w monografii, osiągniętych wyników oraz ich ewentualnego wykorzystania

W pierwszym rozdziale monografii¹ zaproponowałem autorską metodę ekstrakcji cech opierającą się na składowych głównych z rotacją czynników wg centroidów klas (ang. Centroid Class Principal Component Analysis, CCPCA). Jest ona modyfikacją klasycznej metody składowych głównych (ang. Principal Component Analysis, PCA). Celem badawczym było uzyskanie takiej przestrzeni cech, która będzie wyjaśniać porównywalny lub większy procent całkowitej wariancji, w porównaniu z innymi znanymi z literatury metodami ekstrakcji, co przełoży się na zwiększenie jakości poprawnych klasyfikacji.

W metodzie przyjąłem, że punkt C będzie określał środek ciężkości danej klasy i jest oddalony od początku układu współrzędnych o wektor $\vec{u} = [-C_x, -C_y]$. Istotą CCPCA jest obrót współrzędnych osi, aby zwiększyć zawartość informacji o kowariancji względem klas. W pierwszym kroku dokonujemy normalizacji każdej cechy, wykorzystując skalę σ dla każdej i -tej cechy

$$z_i = \frac{x_i - \mu_i}{\sigma_i} \quad (3.1)$$

gdzie z_i jest i -tą nową znormalizowaną zmienną, x_i jest i -tą pierwotną cechą, μ jest wartością średnią ze wszystkich wzorców oraz σ jest odchyleniem standardowym.

Proces konwersji wartości x_i na skalę z_i powoduje przesunięcie bezstratne całej przestrzeni cech do środka układu współrzędnych. W metodzie CCPCA po normalizacji (3.1) wg centroidów każdej i -tej klasy oś obrotu należy przemieścić o wektor $\vec{u}_j = [-C_{j,x_1} - C_{j,x_2}, \dots, -C_{j,x_n}]$, gdzie n oznacza n -tą cechę.

W przypadku rzeczywistych zbiorów danych najczęściej dysponujemy wektorami składającymi się z wielu cech $X_j = [x_{j,1}, x_{j,2}, \dots, x_{j,n}]$. W zadaniu ekstrakcji dążymy do maksymalizacji całkowitej wariancji. W tym celu obliczamy j macierzy kowariancji będącą liczbową reprezentacją ilości informacji zawartych pomiędzy n cechami.

$$A_j A_j^T = \begin{bmatrix} E[(x_1 - C_1)(x_1 - C_1)] & \cdots & E[(x_1 - C_1)(x_n - C_n)] \\ E[(x_2 - C_2)(x_1 - C_1)] & \cdots & E[(x_2 - C_2)(x_n - C_n)] \\ \vdots & \ddots & \vdots \\ E[(x_n - C_n)(x_1 - C_1)] & \cdots & E[(x_n - C_n)(x_n - C_n)] \end{bmatrix} \quad (3.2)$$

Następnie na podstawie macierzy (3.2) wyznaczamy $B_j B_j^T$

$$B_j B_j^T = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}, \quad (3.3)$$

Wartości na przekątnej reprezentują wariancję między tymi samymi

¹ Część I. Metody ekstrakcji bazujące na składowych głównych "Modyfikacja metody składowych głównych uwzględniająca rotację czynników wg centroidów klas"

oraz z pozostałymi cechami. Proces dekompozycji wartości własnych przedstawic możemy następująco

$$A_j A_j^T = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \times \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}, \quad (3.4)$$

gdzie m jest liczbą wzorców, n jest liczbą cech.

Podczas tego procesu otrzymujemy wektory własne, które stanowią nowe wymiary CCPCA w przestrzeni cech oraz wartości własne, które zawierają informację o wartościach wektorów własnych. Metoda CCPCA dokonuje dekompozycji większej przestrzeni cech na mniejszą, uwzględniając informacje o klasach za pomocą wartości osobliwych (ang. Singular Value Decomposition, SVD). W ten sposób, dowolną macierz A o n cechach i m wzorcach możemy zapisać w postaci

$$A_{j(m \times n)} = U_{j(m \times n)} S_{j(m \times n)} V_{j(n \times n)}^T, \quad (3.5)$$

gdzie U i V są macierzami ortogonalnymi z ortonormalnymi własnymi wartościami z macierzy AA^T i $A^T A$. Natomiast S jest macierzą diagonalną o dodatnich wartościach.

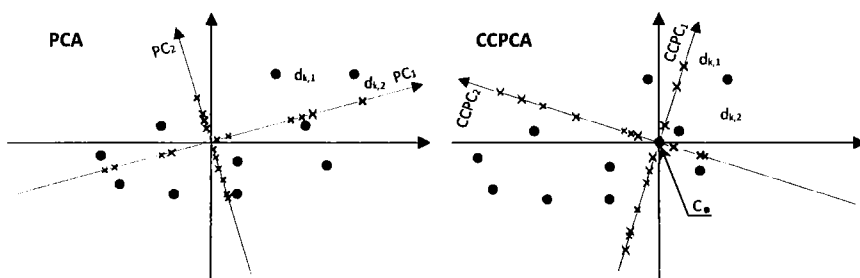
Macierze U i V są ortogonalne co oznacza, że ich iloczyn krzyżowy wynosi 0. Natomiast macierz ortonormalna oznacza, że (a) macierze są ortogonalne, (b) wyznacznik wynosi 1.

W nowej przestrzeni każda składowa główna, jak to zostało już wspomniane tworzy wektor własny, z wartościami własnymi dla każdej cechy:

$$CCPC_{j,k} = (\omega(x_1), \omega(x_2), \dots, \omega(x_n)), \quad (3.6)$$

gdzie: k oznacza k -ty komponent dla j -tej klasy.

Na rysunku 5 przedstawiono zasadę działania metody CCPCA w porównaniu z PCA.



Rysunek 3.1: Przykładowy schemat obrazujący wyznaczenia komponentów CCPCA wg. centroidów dystanse od nich.

Zauważmy, że rotując wg. centroidów klas z jednej strony otrzymujemy więcej komponentów ale również to na czym nam zależy lepszą ich dyskryminację. W celu doboru odpowiedniej liczby cech do danego komponentu w pierwszej kolejności należy posortować malejąco wartości własne z równania (3.6). Następnie zaczynając od dwóch cech o największych wartościach własnych z macierzy $A_j A_j^T$ (3.2) obliczamy miarę spójności skali

$$\alpha = \frac{n}{n+1} \left(1 - \frac{\sum_{i=1}^n S_i^2}{S_{sc}^2} \right), \quad (3.7)$$

gdzie: n - jest liczbą cech, S_i^2 jest wartością pozycji, a S_{sc}^2 jest wartością skali.

Wartość pozycji S_i^2 możemy zapisać wzorem

$$S_i^2 = \sum_{i=1}^n E[(x_i - C_i)(x_i - C_i)] \quad (3.8)$$

Natomiast wartość skali S_{sc}^2 obliczamy

$$S_{sc}^2 = \sum_{k=1}^n \sum_{i=1}^n E[(x_k - C_k)(x_i - C_i)], \quad (3.9)$$

gdzie i to poszczególne kolumny macierzy $A_j A_j^T$ (3.2), natomiast k określa jej wiersze.

Dla każdej składowej z macierzy (3.2) wyznaczamy spójność (3.7) dodając kolejno cechy, aż do momentu uzyskania maksymalnej wartości. Jeżeli wartość maksymalna $\alpha < 0.7$ wówczas cały komponent usuwamy. Oznacza to, że cechy w tej składowej nie wnoszą istotnej informacji do dyskryminacji klas przez rozpatrywany w tym komponencie wektor cech. Ponadto jeżeli dwa wyodrębnione ww. metodą komponenty zawierają będą tą samą podprzestrzeń cech, należy usunąć wszystkie spójne składowe pozostawiając tylko jedną.

W części eksperymentalnej w pierwszej kolejności wykorzystałem dane syntetyczne wygenerowane za pomocą metody *make_classification* pochodzącej z biblioteki *scikit-learn*². Przeprowadzenie eksperymentów na danych syntetycznych (przypadek binarny) miało na celu dokonanie oceny jak kształtuje się jakość klasyfikacji w zależności od: (a) liczby cech, (b) stopnia niezbalansowania oraz (c) wskaźnika odsetkowego cech informatywnych (w tym przypadku sprawdzałem dwie wartości 5% oraz 10%). W tym zadaniu do oceny na ile klasyfikator zwraca poprawną decyzję, a nie losową zastosowałem wskaźnik *Kappa Cohana*. Testy wykonałem dla różnych metod ekstrakcji tj.: (a) składowych głównych PCA, (b) nieliniowej analizy składowych głównych (ang. Kernel Principal Component Analysis, KPCA), (d) składowych niezależnych ICA (ang. Independent Component Analysis, ICA) oraz proponowanej metody CCPCA. Dla kontrastu wyników jako jeden z wariantów przyjąłem brak ekstrakcji. W badaniach zastosowałem pięć bazowych klasyfikatorów tj.: (a) k najbliższych sąsiadów (ang. K-Nearest Neighbors, KNN), (b) maszynę wektorów nośnych (ang. Support Vector Classifier, SVC), (c) drzewo klasyfikacyjne (ang. Classification And Regression Tree, CART), (d) naiwny klasyfikator Bayesa (ang. Gaussian Naive Bayes, GNB) oraz (e) sieć neuronową (ang. Multilayer Perceptron, MLP).

Uzyskane wyniki pozwoliły na wysnucie następujących wniosków:

- Wraz ze wzrostem liczby cech 30-800 zwiększa się jakość klasyfikacji. Ponadto im bardziej dane są zbalansowane tym lepsza jest uzyskiwana jakość klasyfikacji dla wszystkich klasyfikatorów. Zauważyłem, że wraz ze wzrostem wolumenu danych rośnie skuteczność ekstrakcji i

² F. Pedregosa i in. "Scikit-learn: Machine Learning in Python". W: *Journal of Machine Learning Research* 12 (2011), s. 2825-2830

jakość klasyfikacji. Dla danych gdzie wystąpiło 10% cech informatywnych uzyskałem lepsze jakości klasyfikacji niż dla przypadku 5% takich danych.

- Przeprowadzone eksperymenty na danych syntetycznych wskazują, że po ekstrakcji dokonanej za pomocą opracowanej metody CCPCA uzyskano dla większych wolumenów danych $n > 50$ lepsze jakości klasyfikacji niż w przypadku innych metod ekstrakcji.
- W przypadku danych niezbalansowanych uzyskane wyniki jakości klasyfikacji dla różnej liczności cech informatywnych uzyskałem co najmniej tak dobre, a nawet lepsze jakości klasyfikacji po zastosowaniu innych metod.

W celu szerszego spojrzenia na skuteczność opracowanej metody CCPCA w kolejnym eksperymencie wykorzystałem 41 rzeczywistych zbiorów danych, z których 10 pochodzi z prywatnego repozytorium, a pozostałe 31 z ogólnie dostępnej bazy KEEL (ang. Knowledge Extraction based on Evolutionary Learning)³. W monografii zamieściłem tylko wyniki dla sieci neuronowej MLP dla której uzyskano najlepsze jakości klasyfikacji. Ze względu na różny stopień niezbalansowania zastosowałem protokół eksperymentalny w postaci walidacji krzyżowej stratyfikowanej pięciokrotnej powtórzonej dwa razy. Dla każdego zbioru jakość uzyskana przez sieć MLP dla różnych metod ekstrakcji była porównywana testem Wilcoxon na poziomie istotności statystycznej $p=0,05$. Dla przypadków binarnych zastosowałem metrykę Gmean, a dla wieloklasowych zbalansowaną dokładność BAC. Uzyskane wyniki pozwoliły mi na wysnucie następujących wniosków:

³ www.keel.es

- W przypadku proponowanej metody CCPCA uzyskano dla każdego zbioru rzeczywistego istotnie wyższe jakości poprawnych klasyfikacji w porównaniu do przypadku kiedy nie zastosowano żadnej metody ekstrakcji.
- Dla 18 zbiorów danych dla przypadków binarnych dla proponowanej metody CCPCA uzyskałem lepsze jakości klasyfikacji w 4 przypadkach od metody PCA, w 7 przypadkach od metody KPCA i w 7 przypadkach od metody ICA. Dla pozostałych przypadków wyniki były porównywalne statystycznie do ww. metod.
- Dla przypadków wieloklasowych zaobserwowałem interesujące wyniki. Mianowicie w zdecydowanej większości zbiorów danych po rotacji wg. centroidów klas CCPCA uzyskałem lepsze jakości klasyfikacji.
- Zaproponowane przeze mnie podejście pozwoliło na uzyskanie na wszystkich zbiorach średnio wyższej całkowitej wyjaśnianej wariancji o 6,3% niż w PCA, o 7,1% niż w KPCA i o 7,3% niż w ICA.
- Po bardziej szczegółowej analizie statystycznej danych zauważyłem, że zaproponowana metoda CCPCA jak i inne metody ekstrakcji opierające się na składowych głównych tym gorzej sobie radzą z zadaniem ekstrakcji im większe jest odchylenie standardowe cech oraz kiedy występują przypadki odstające tzw. outliery.

Opracowana metoda CCPCA znalazła praktyczne zastosowanie w projektach badawczych:

- **Model klasyfikacji nasilenia objawów lęku i depresji u dzieci z wcześniej rozpoznanymi zaburzeniami psychicznymi w czasie pandemii COVID-19 w Polsce.** Na podstawie zbioru danych z badań prowadzonych przez pracowników naukowych Kliniki Psychiatrii Dziecięcej Warszawskiego Uniwersytetu Medycznego. Rezultatem jest ekstrakcja cech i stworzenie modelu drzewa decyzyjnego typu CART oraz wspólny artykuł naukowy ⁴.
- **Model klasyfikacji ryzyka depresji u pacjentów po operacjach kardiochirurgicznych.** Na podstawie zbioru danych z badań prowadzonych przez pracowników naukowych: (a) Katedry Medycyny Rodzinnej, Wydziału Lekarskiego, Uniwersytetu Medycznego w Gdańsku, (b) Oddział Kardiochirurgii, Kaszubskiego Centrum Chorób Serca i Naczyń, Szpitala Specjalistycznego Ceynowa, (c) Kliniki Otolaryngologii Uniwersyteckiego Centrum Klinicznego Gdańskiego Uniwersytetu Medycznego, (d) Zakładu Socjologii Medycyny i Patologii Społecznej, Wydziału Nauk o Zdrowiu, Gdańskiego Uniwersytetu Medycznego oraz (e) Biura Ochrony Radiologicznej Uniwersyteckiego Centrum Klinicznego Gdańskiego Uniwersytetu Medycznego. Rezultatem jest ekstrakcja cech i stworzenie modelu drzewa decyzyjnego typu CART oraz wspólny artykuł naukowy ⁵
- **Zastosowanie metod uczenia maszynowego w predykcji poziomu wypalenia zawodowego, prężności oraz poczucia własnej skuteczności u psychologów klinicznych i lekarzy psychiatrów.** Projekt jest realizowany pod patronatem Rzecznika Praw Pacjenta ⁶. Rezultatem jest ekstrakcja cech metodą CCPCA i stworzenie modelu sieci neuronowej typu MLP. Aktualnie powstaje wspólny artykuł, który planujemy opublikować w czasopiśmie *Pattern Recognition*.

Możliwe obszary zastosowania proponowanej metody CCPCA:

- W zadaniu ekstrakcji cech dla różnych zbiorów, szczególnie dla zadana klasyfikacji problemów wieloklasowych.
- Na potrzeby redukcji cech przy danych niebalansowanych.

Wnioski do jakich doszedłem podczas badań nad opracowaną metodą CCPCA nakłoniły mnie do dalszego poszukiwania rozwiązań które pozwoliły by na lepszą jakość ekstrakcji w sytuacjach: (a) kiedy cechy mają wysokie wartości odchylenia standardowego, (b) występują przypadki odstające oraz (c) rozkłady cech w klasach mogą mieć dowolne funkcje gęstości prawdopodobieństwa. W związku z powyższym w kolejnym rozdziale ⁷ zaproponowałem metodę maksymalizacji całkowitej wariancji z zastosowaniem gradientów stochastycznych GPCA.

W przypadku GPCA wektory własne szukamy dla macierzy K takiej, że

$$K_{i,j} = L(Z_i, Z_j), \quad (3.10)$$

⁴ Anna Maria Kalenik i in. "The impact of the COVID-19 pandemic on the mental health of children with psychiatric diagnoses—multidimensional CCPCA Model". W: *BMC psychiatry* 22.1 (2022)

⁵ Katarzyna Nowicka-Sauer i in. "Applying Machine Learning to Construct a Model of Risk of Depression in Patients Following Cardiac Surgery with the Use of the SF-12 Survey". W: *nt. J. Environ. Res. Public Health* 20.6 (2023)

⁶ <https://www.gov.pl>

⁷ Część I. Metody ekstrakcji bazujące na składowych głównych "Zastosowanie metody gradientów stochastycznych do szacowania składowych głównych"

gdzie L jest funkcją celu, Z jest zmienną standaryzowaną, $L(Z_i, Z_j)$ jest jądrem

$$L(Z_i, Z_j) = \sum_{i=1}^n (x_i - \varrho^T Z_j)^2, \quad (3.11)$$

gdzie ϱ^T jest gradientem.

W przeciwieństwie do metody KPCA w tym punkcie możemy zdefiniować inne np. nieliniowe funkcje celu.

Minimalizację błędu $e(L(Z_i, Z_j))$ funkcji celu rozpoczynamy od wybranego rozwiązania startowego $\varrho_0 = 0$. Następnie wyznaczany jest gradient w punkcie $\varrho_{k-1}, \alpha_k \nabla_L(\varrho_{k-1})$. Kolejno wyznaczany jest krok wzdłuż negatywnego gradientu

$$\varrho_k = \varrho_{k-1} - \alpha_k \nabla_L(\varrho_{k-1}), \quad (3.12)$$

gdzie α_k jest długością kroku ustaloną przed poszukiwanie liniowe.

Gradient ∇_L obliczamy za pomocą różniczki

$$\frac{\partial (Z_i - \varrho^T Z_j)^2}{\partial \varrho_j} = -2 (Z_i - \varrho^T Z_j) Z_{ij}. \quad (3.13)$$

Ostatecznie

$$\nabla_L(\varrho) = -2 (x_i - \varrho^T Z_j) Z_j. \quad (3.14)$$

Składowe główne można przedstawić teraz jako liniową kombinację pierwotnych zmiennych Z

$$G_{kij} = \sum_{i=1}^k \sum_{j=1}^m a_{kij} Z_j, \quad (3.15)$$

gdzie m jest liczbą zmiennych pierwotnych w zbiorze uczącym, k to liczba składowych głównych, Z_j jest zmienną standaryzowaną, G_{kij} jest składową główną, a_{kij} to ładunki czynnikowe.

Ostatecznie należy dokonać agregacji cech w odpowiednich komponentach. W tym celu dla każdej klasy j i komponentu k obliczyć należy przedziały ufności dla ładunków czynnikowych $a_{k,j}$

$$\forall k, j \left(a_{k,j} - u_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}, a_{k,j} + u_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right), \quad (3.16)$$

gdzie $a_{k,j}$ jest to średnia wartość ładunku czynnikowego, s jest odchyleniem standardowym, α jest poziomem istotności statystycznej, $u_{1-\frac{\alpha}{2}}$ jest to dystrybuanta rozkładu normalnego oraz n jest liczbą cech.

Otrzymano w ten sposób przedziały ufności dla każdej klasy w każdym komponentcie. Następnie dla każdej składowej przyjmujemy tylko te cechy które należą do części wspólnej wyznaczonych przedziałów ufności. Jeżeli dla kilku składowych wystąpiła ta sama cecha, o jej przynależności do jednego z kilku składowych decyduje większa wartość ładunku czynnikowego.

Zaproponowana metoda GPCA może znaleźć zastosowanie w przestrzeniach nieliniowych cech. Można zaproponować inne funkcje jądra w zależności od klasy problemu. Ponieważ bazuje ona podobnie jak CCPCA na centroidach klas może okazać się lepsza pod względem ekstrakcji od

ww. metody w przypadku pewnych charakterystycznych zbiorów danych, dla których cechy mają wyższe odchylenia standardowe, występują przypadki odstające, czy istnieje nieliniowy związek między cechami w klasach. Rozwiązując zadanie optymalizacyjne dla każdej klasy osobno możemy uzyskać separację optymalnych składowych głównych dla każdej z nich.

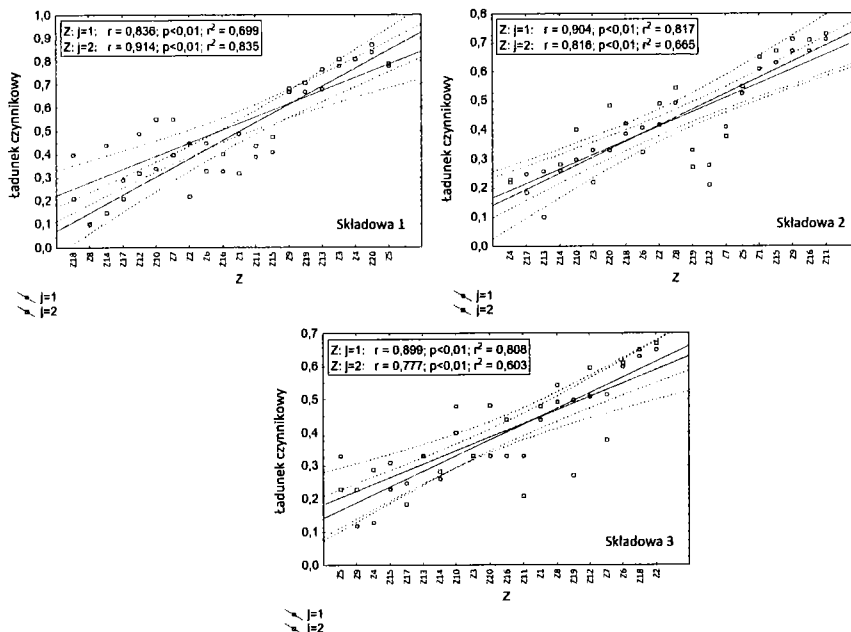
Otrzymujemy w ten sposób przedziały ufności dla każdej klasy w każdym komponencie. Następnie dla każdej składowej przyjmujemy tylko te cechy które należą do części wspólnej wyznaczonych przedziałów ufności. Jeżeli dla kilku składowych wystąpiła ta sama cecha, o jej przynależności do jednego z kilku składowych decyduje większa wartość ładunku czynnikowego.

W części metodologicznej przeprowadziłem trzy eksperymenty. W każdym zastosowałem protokół eksperymentalny tj. walidację krzyżową pięć-krotną, powtórzoną dwa razy. W celu porównania jakości klasyfikacji po przeprowadzonym procesie ekstrakcji, zastosowałem test Wilcoxon. Istotność różnic między różnymi metodami ekstrakcji weryfikowałem na poziomie istotności $p < 0,05$. W eksperymencie pierwszym testy przeprowadziłem dla 41 rzeczywistych zbiorów danych. Ocenę jakości klasyfikacji przeprowadziłem dla sieci neuronowej MLP oraz metod ekstrakcji: (a) PCA, (b) KPCA, (c) CCPCA oraz dla proponowanej (c) GPCA. Analizując otrzymane wyniki doszedłem do następujących wniosków:

- Na 41 zbiorów danych rzeczywistych w trzech przypadkach po ekstrakcji metodą GPCA uzyskano istotnie statystycznie lepszą jakość poprawnych klasyfikacji niż w metodzie CCPCA. W pozostałych przypadkach wyniki obu metod są porównywalne statystycznie. Mimo braku istotności zauważyłem, że istnieją pewne tendencje, a mianowicie jakość klasyfikacji po zastosowaniu metody GPCA nieznacznie jest wyższa niż uzyskana za pomocą metody CCPCA. Ten przypadek występuje w znacznej części zbiorów. Szczególnie ma to znaczenie w zbiorach w których współczynnik zmienności pewnych cech jest większy niż 20%.
- W przypadku wieloklasowym podobnie jak w pierwszej części monografii uzyskałem wyniki jakości klasyfikacji lepsze dla metod GPCA i CCPCA niż dla przypadków klasyfikacji binarnej. Zatem rotacja wg centroidów klas może mieć częściej lepsze zastosowanie w danych wieloklasowych niż binarnych.

W tej części badań szczególne uwagę skupiłem na jednym zbiorze danych rzeczywistych klasyfikacji białaczki limfocytowej u dzieci. Jest to podyktowane współpracą z lekarzami onkologii. Bazując tylko na tym zbiorze danych przeprowadziłem dwa kolejne eksperymenty. W pierwszym z tych dwóch eksperymentów dla różnych klasyfikatorów bazowych i metod ekstrakcji. Weryfikowałem doświadczalnie przy jakim progu całkowitej wyjaśnianej wariancji jakość klasyfikacji będzie największa. W tym eksperymencie zastosowałem trzy metody ekstrakcji PCA, CCPCA oraz GPCA. Przyjąłem progi wyjaśnianej całkowitej wariancji w przedziale 0 do 100 %. Eksperymentalnie uzyskałem po ekstrakcji różnymi metodami cechy wyjaśniające 68-71% całkowitej wariancji. Badania przeprowadziłem z wykorzystaniem różnych klasyfikatorów bazowych. Stosując

metodę GPCA uzyskałem trzy składowe spełniające warunek całkowitej wyjaśnianej wariancji 72%. Przykład metody agregacji cech po zadaniu ekstrakcji dla dwóch klas obrazuje rysunek 3.2



Rysunek 3.2: Wykres zależności doboru cech obiektu do każdej z trzech składowych głównych od wartości ładunków czynnikowych

W procesie agregacji po zastosowaniu części wspólnej przedziałów ufności ładunków czynnikowych i kryterium rozstrzygającego maksymalnej wartości ładunku czynnikowego w sytuacji przynależności cechy do kilku kryteriów uzyskałem wektor cech, który pozwolił mi uzyskać lepsze jakości klasyfikacji w porównaniu z innymi metodami ekstrakcji, co potwierdzono w trzecim eksperymencie. Najlepsze wyniki jakości klasyfikacji uzyskałem po zastosowaniu metody GPCA dla drzewa decyzyjnego CART (83,1%), następnie dla CCPCA (81%), gdzie dla metody PCA wynik jakości klasyfikacji był istotnie niższy (77%). Zatem w przypadku tego zbioru udało mi się potwierdzić skuteczność proponowanych metod w porównaniu z innymi metodami ekstrakcji.

Możliwe obszary zastosowania proponowanej metody GPCA:

- W zadaniu ekstrakcji cech dla zadania klasyfikacji białaczki limfocytowej u dzieci.
- Dla zbiorów danych dla których: (a) istnieją cechy z dużymi wartościami odchylenia standardowego, (b) istnieją cechy z wartościami odstającymi oraz (c) związki między cechami mają charakter nieliniowy.

Współpracując z różnymi zespołami naukowymi, zwłaszcza z dziedziny medycyny, często napotykałem trudności związane z pozyskiwaniem danych. Problemy te wynikały z braków w dokumentacji medycznej, kosztownych badań oraz małej liczby pacjentów z określonym problemem medycznym. W przypadku braków danych, różne modele pomijają całe

wzorce, co jest problematyczne, gdyż pozyskane dane są istotne i wartościowe z punktu widzenia wiedzy o klasach problemu. Dlatego też, omawiając temat imputacji danych, poświęciłem mu kolejny rozdział ⁸.

W zadaniu imputacji danych w zaproponowałem podejście, zakładające, że dana cecha jest przypisana tylko do jednej składowej głównej. Zapewnia to lepszą moc dyskryminacyjną wg klas ale również silniejsze powiązanie różnych cech. Ponieważ w klasycznym podejściu każda cecha przynależy do każdego komponentu k z pewną wartością własną ω , przeprowadziłem normalizację do przedziału [0-1] tak aby ich łączny rozkład prawdopodobieństwa każdej cechy we wszystkich komponentach był równy 1. Zauważyłem, że występują sytuacje, kiedy istotnie statystycznie (test dla wielu wskaźników odsetkowych) zmienna może przynależeć do kilku komponentów jednocześnie. Dla takich przypadków wstępnie założyłem, że dana cecha może przynależeć do więcej niż jednego komponentu. W celu rozwiązania problemu przypisania prawdopodobieństwa do kilku zdarzeń zastosowałem teorię ewidencji matematyczną. Jest to podyktowane tym, że w probabilistyce prawdopodobieństwo przypisujemy do pojedynczych zdarzeń. Zbiór K_1 określa wszystkie kombinacje składowych głównych

$$\binom{k}{p} = \frac{k!}{p!(k-p)!}, \quad (3.17)$$

gdzie k jest liczbą komponentów, a p jest wyborem każdego z p elementów ze zbioru k , bez zwracania uwagi na kolejność, czyli jako kombinacje bez powtórzeń.

Zgodnie z teorią ewidencji matematycznej ⁹ przyjąłem funkcją masy, która podobnie jak w rachunku prawdopodobieństwa wyznaczana jest na zasadzie wystąpień danej kombinacji komponentów w stosunku do wszystkich wzorców

$$\sum_{K_1 \in 2^\Omega} m(K_1) = 1. \quad (3.18)$$

Kolejnym kryterium dyskryminacyjnym jest tzw. współczynnik Lambda Wilksa

$$\lambda = \frac{\det(G)}{\det(G+H)}, \quad (3.19)$$

gdzie $\det(G)$ jest wyznacznikiem macierzy wewnątrzgrupowej kowariancji, $\det(H)$ jest wyznacznikiem macierzy międzygrupowej kowariancji.

Dla każdej cechy obliczana jest wartość cząstkowa mocy dyskryminacyjnej którą zmodyfikowano aby była zgodna korelacyjnie

$$\forall x_i \in X \quad \lambda_i = 1 - \frac{\lambda_{i,next}}{\lambda_{i,previous}}, \quad (3.20)$$

gdzie $\lambda_{i,next}$ jest wartością współczynnika Lambda Wilksa dla modelu po wprowadzeniu do niego i -tej zmiennej, $\lambda_{i,previous}$ jest wartością współczynnika Lambda Wilksa dla modelu przed wprowadzeniem do niego i -tej cechy. Miara rzetelności komponentów jest wyznaczana

$$\alpha = \frac{n}{n+1} \left(1 - \frac{\sum_{i=1}^n S_i^2}{S_{sc}^2} \right), \quad (3.21)$$

⁸ Część I. Metody ekstrakcji bazujące na składowych głównych "Zastosowanie metody składowych głównych w zadaniu imputacji danych"

⁹ M.A. Kłopotek. "Metody identyfikacji i interpretacje struktur rozkładów przekonań w teorii Dempstera-Shafera". W: Instytut Podstaw Informatyki Polskiej Akademii Nauk (1998)

gdzie S_i^2 jest wartością pozycji, a S_{sc}^2 jest wartością skali, n jest liczbą cech.

Następnie na podstawie kryteriów (3.20) oraz (3.21) dla każdej składowej głównej $i = 1, 2, \dots, k$ oraz każdej cechy $j = 1, 2, \dots, n$ wyliczany jest miara dyskryminacyjna

$$\mu_{i,j} = \alpha_{i,j} \lambda_{i,j}. \quad (3.22)$$

W ten sposób otrzymałem drugi bazowy rozkład oceny ważności cech

$$\sum_{K_2 \in 2^\Omega} m(K_2) = 1. \quad (3.23)$$

Metoda szacowania funkcji masy jest taka sama jak w przypadku zbioru K_1 , z tą różnicą, że dla K_2 normalizujemy wg komponentów wartość μ . W celu złożenia dwóch bazowych rozkładów $m(K_1)$ oraz $m(K_2)$ zastosowałem regułę kombinacji wywodzącą się z teorii ewidencji matematycznej

$$m(K) = m(K_1) \oplus m(K_2) = \frac{\sum_{K_1 \cap K_2 = K} m(K_1) m(K_2)}{1 - \sum_{K_1 \cap K_2 = \emptyset} m(K_1) m(K_2)}. \quad (3.24)$$

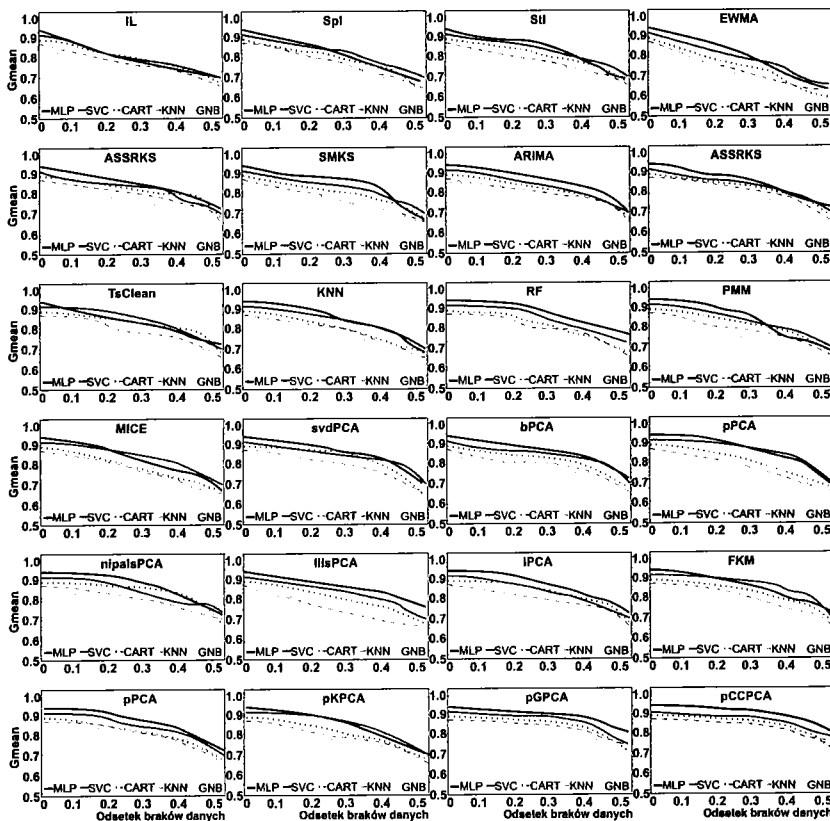
W wyniku złożenia (3.24) nadal mogą występować przypisane prawdopodobieństwa do kilku komponentów jednocześnie. Aby obliczyć rozkłady bazowe dla pojedynczych składowych zastosowałem tzw. funkcje przekonania

$$Bel(K') = \sum_{K \in K'} m(K). \quad (3.25)$$

Komponenty składające się wyłącznie z pustych zbiorów zmiennych są usuwane. Po tej operacji otrzymałem k' komponentów, z których każdy ma unikalny zestaw cech. Zadanie imputacji polega na wypełnieniu brakujących wartości cech w wektorze. Aby to zrobić, należy najpierw określić, do jakiej klasy należy adresowany wektor oraz do którego komponentu należy brakująca wartość cechy. Dla tak wyznaczonej podprzestrzeni cech należy je uśrednić wartościami rozpatrywanego wektora.

W badaniach wykorzystałem protokół eksperymentalny oparty na walidacji krzyżowej 5-krotnej powtórzonej dwukrotnie. W celu weryfikacji różnic statystycznie istotnych między różnymi metodami imputacji brakujących danych zastosowałem test Wilcozona. W celu urealnienia wyników eksperymenty przeprowadziłem dla 41 zbiorów danych rzeczywistych. W pierwszej kolejności usuwałem losowo z każdego zbioru 5%, 15% oraz 25% danych, a następnie uzupełniłem ww. braki za pomocą opracowanej metody. Podczas eksploracji różnych modeli imputacji do eksperymentów wybrałem te, dla których jakość klasyfikacji była największa tj. (a) algorytm z estymacją bayesowską uwzględniający niepewność wyników (ang. Bayesian Principal Component Analysis, *BPCA*), (b) z estymacją składowych głównych poprzez iteracyjne rozwiązywanie równania PCA wykorzystujące podejście Martensaza (ang. Nonlinear Iterative Partial Least Squares, *NIPALSPCA*), (c) wykorzystuje analizę wartości osobliwych SVD do dekompozycji macierzy danych na nadrzędne składowe (ang. Least Squares Principal Component Analysis, *LLSPCA*) oraz (d) uwzględniający rozmyty opis skupień k średnich (ang. Fuzzy K-Means, *FKM*)

Analizując otrzymane wyniki zauważyłem, że dla 5%, 15% oraz 25% braków danych metoda PCCPCA w zadaniu imputacji często prowadzi do uzyskania istotnie statystycznie wyższych wyników jakości poprawnych klasyfikacji w porównaniu z algorytmami takimi jak: (a) NIPALSPCA, (b) LLSPCA, (c) FKM oraz (d) BPCA. Ponadto, odsetek utraconych danych jest ujemnie skorelowany z jakością klasyfikacji. Szczególnie widoczne są te różnice dla zbiorów w wieloma klasami. Dla 5% brakach danych dla 37, 15% dla 34 i 25% dla 31 zbiorów danych uzyskałem porównywalne istotnie statystycznie ($p \geq 0,05$) jakości klasyfikacji do przypadku kompletnych danych. Bardziej szczegółowe badania przeprowadziłem na danych z biopsji, których wyniki znajdują się na rysunku 3.3.



Rysunek 3.3: Wykres przedstawiający wyniki eksperymentu dla różnych metod imputacji i klasyfikacji

Kolorem niebieskim oraz litrą p oznaczyłem metody składowych głównych w których zastosowałem opracowane przeze mnie rozwiązanie imputacji danych, takich jak PPCA, PKPCA, PGPCA oraz PCCPCA. Bardzo interesującym wnioskiem jest to, że rotując czynniki według centroidów, zaobserwowałem bardziej spłaszczony przebieg wykresu jakości klasyfikacji. Przy dużych stratach danych, dochodzących do 50%, opracowana przeze mnie metoda pozwoliła na odtworzenie danych i uzyskanie przewagi w granicach 5% - 9% w jakości klasyfikacji w porównaniu do innych metod. Przyczyną takiej różnicy jest wyodrębniona większa liczba istotnych składowych, co pozwoliło na uzyskanie bardziej dyskryminujących cech dla zadania imputacji oraz klasyfikacji.

Możliwe obszary zastosowania proponowanej metody imputacji:

- Dla różnych zbiorów danych których cechy mają charakter ilościowy lub porządkowy.
- Dla zbiorów danych charakteryzujących się wartościami odstającymi, które zaburzają klasyfikację.
- Dla danych, których pozyskanie jest kosztowne, a usuwanie wzorców zmniejsza zawartość informacyjną.

W kolejnej części monografii¹⁰ zaproponowałem autorską metodę szacowania liczby cech oraz składowych głównych. Analizując literaturę oraz stosowane praktyki doboru liczby komponentów, najczęściej stosuje się dwa kryteria. Pierwsze z nich to reguła *Kaisera*, która zakłada, że dana składowa jest istotna, gdy jej wartość własna ω jest większa od 1. Natomiast drugim stosowanym kryterium jest tzw. *osypisko czynników*. Jest to metoda subiektywnej oceny, polegająca na tym, że jeżeli wartości własne kolejnych komponentów nie różnią się istotnie między sobą, reszta składowych jest usuwana. Nie rzadko stosuje się podejście, w którym przyjmuje się wszystkie składowe główne, które łącznie wyjaśniają 80% całkowitej wyjaśnianej wariancji. Podejścia subiektywne często wiążą się z doświadczeniem eksperta, co przekłada się na jakość klasyfikacji. Celem badawczym było opracowanie metody szacowania liczby cech i składowych głównych dopasowanych do danego zbioru danych, których zastosowanie zwiększy jakość poprawnych klasyfikacji w porównaniu z ww. stosowanymi powszechnie metodami. Cel badań został przeze mnie osiągnięty. W modelu zgodnie z całą strukturą monografii przyjąłem, że k jest liczbą składowych głównych, a n liczbą wszystkich cech. Dla każdej składowej głównej otrzymałem następujące wektory własne

$$\forall k \quad \Omega_k = [w_{k,1}, w_{k,2}, \dots, w_{k,n}]. \quad (3.26)$$

Niech $\Phi(\Omega_k)$ będzie standardową funkcją rozkładu normalnego $N(0, 1)$. Oryginalne podejście do zadania określenia liczby cech i elementów polega na wykorzystaniu kryterium niezawodności określonego wzorem

$$\beta = \frac{n\bar{r}\bar{r}}{1 + (n-1)\bar{r}\bar{r}}, \quad (3.27)$$

gdzie \bar{r} jest średnią wartością współczynnika korelacji między wszystkimi parami pozycji, \bar{r} współczynnik determinacji określa siłę, z jaką dana cecha wyjaśnia przynależność do składowej głównej oraz n jest liczbą cech.

Ostatecznie o tym, że dana cecha należy do k -tej składowej, decyduje spełnienie zależności

$$\Phi(w_{k,i}) \geq \beta, \quad (3.28)$$

gdzie $i = 1, 2, \dots, n$.

W sytuacji, gdy pojedyncza cecha została zaklasyfikowana do więcej niż jednej składowej głównej, przynależność do jednego z k komponentów wyznaczana jest za pomocą kryterium $\max[|w_{k,i}|]$. Dla składowych wyjaśniających coraz mniejszy procent wariancji całkowitej liczba cech należących do tych składowych maleje i w przypadku niespełnienia kryterium (3.28) składowe te są usuwane.

¹⁰ Część II. Metody ekstrakcji cech sygnałów o charakterystyce czasowej oraz dwuwymiarowych "Metoda szacowania liczby cech i komponentów w analizie składowych głównych"

Rozważyłem również przykład nieliniowej separacji głównych składowych. Dobra separacja związana jest z oszacowanym marginesem $[\beta_-; \beta_+]$. W tym celu można zastosować nieliniowy model SVM. W przestrzeni danych Ω zdefiniowane są wektory β_i , które tworzą próbkę uczącą D należąca do dwóch klas

$$D = \{(\beta_i, c_i) \mid \beta_i \in [0; 1], c_i = \{-1; 1\}\}_{i=1}^N. \quad (3.29)$$

Zadanie polegało na wyznaczeniu klasyfikatora, który pozwoli na podzielenie całej przestrzeni cech na dwa odrębne obszary odpowiadające klasom $(-1, 1)$ co pozwala na lepsze ich dopasowanie do składowych głównych. W tym celu, zastosowałem podejście nieliniowej maszyny wektorów nośnych. Jeżeli żadne $K(x, \beta)$ nie oznacza funkcji symetrycznej dwóch wektorów jądra, takich że $x \in X$, $\beta \in [0, 1]$ i $X \in R$, to transformację można określić jako $\phi : X \mapsto \Theta$, gdzie

$$K(x, \beta) = \phi(x) \cdot \phi(\beta). \quad (3.30)$$

Przestrzeń Θ , do której dokonuje się odwzorowania, nazywana jest przestrzenią przekształconych zmiennych.

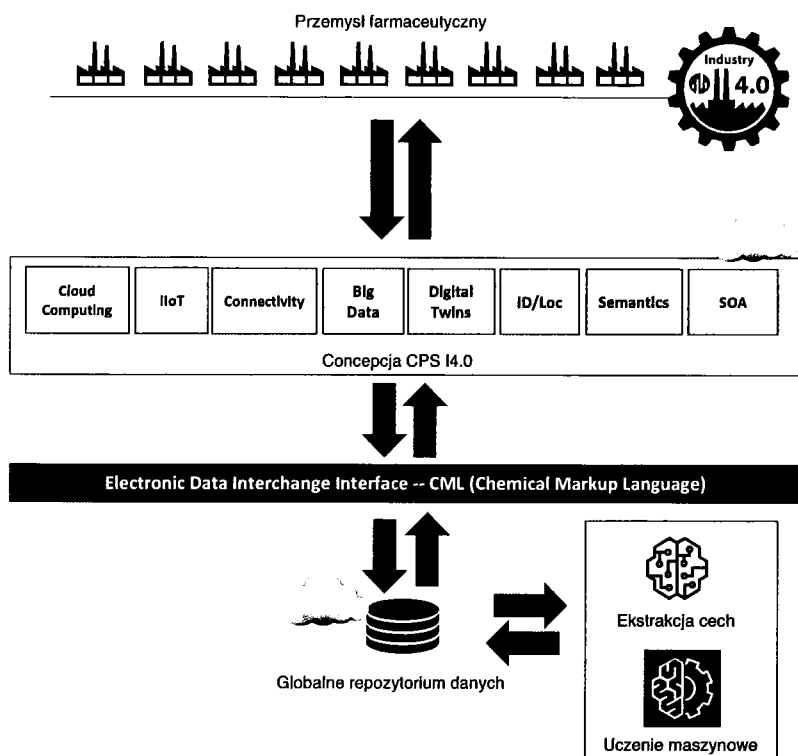
W eksperymentach wykorzystałem łącznie 41 zbiorów danych rzeczywistych. Wszystkie testy przeprowadziłem dla różnych klasyfikatorów bazowych tj. MLP, SVC, CART, GNB oraz KNN. Po analizie różnych hiperparametrów modeli ostatecznie przyjąłem do dalszych badań sieć neuronową MLP. Dla metod składowych głównych PCA oraz CCPCA porównałem jakość klasyfikacji wg różnych kryteriów szacowania liczby składowych głównych i cech tj. (a) *Kaisera*, (b) *osypiska czynników* oraz (c) łącznie *Kaisera* i *osypiska czynników* oraz (d) autorskiego rozwiązania. Eksperymentalnie udało mi się udowodnić, że w przypadku 10 na 41 zbiorów danych dla PCA i 17 na 41 dla CCPCA jakość klasyfikacji była istotnie lepsza po zastosowaniu autorskiego rozwiązania w porównaniu z podejściem przyjęcia jednocześnie kryterium *Kaisera* oraz *osypiska czynników*. W każdej z ww. metod składowych głównych jakość klasyfikacji jest istotnie wyższa w porównaniu z ww. kryteriami stosowanymi oddzielnie. Potwierdziłem również, że jakość klasyfikacji po ekstrakcji metodą CCPCA jest wyższa niż po PCA.

Przeprowadziłem również bardziej szczegółowe eksperymenty na danych rzeczywistych ryzyka chemicznego. Opracowana przeze mnie metoda pozwala uzyskać wyższe jakości klasyfikacji po zastosowaniu jej w metodach PCA, CCPCA oraz GPCA. W przypadku analizowanego zbioru danych nie uzyskałem istotnej różnicy dla metody KPCA.

Wykazałem również, że dla różnych klasyfikatorów bazowych i metod ekstrakcji opartych na składowych głównych najlepsze jakości klasyfikacji uzyskać można dla przyjętych składowych głównych które wyjaśniają około 80% całkowitej wariancji.

Ponieważ moje badania poruszają praktyczne aspekty, w ostatniej części zaproponowałem architekturę zintegrowanego systemu zarządzania ryzykiem, który przedstawiłem na rysunku 3.4.

Zgodnie z architekturą przeprowadziłem eksperymenty których celem było zmierzenie czasów odpowiedzi modułu ekstrakcji i klasyfikacji na



Rysunek 3.4: Model klasyfikacji ryzyka chemicznego w przemyśle farmaceutycznym

podstawie odebranych danych w czasie rzeczywistym. Założyłem różne rozwiązania architektur będących integralną częścią Przemysłu 4.0 tj.: (a) Cloud Computing, (b) IIoT, (c) Connectivity, (d) Big Data, (e) Digital Twins, (f) ID/Loc, (g) Semantics oraz (h) SOA. Modelowanie przeprowadziłem w środowisku Anylogic i Python. Na podstawie przeprowadzonych eksperymentów uzyskałem bardzo interesujące wyniki. W przypadku Big Data i Digital Twins szybkość przetwarzania danych jest mniejsza niż dla innych rozważanych architektur. Ze względu na czas odpowiedzi maszyna wektorów nośnych SVC wydaje się lepszym rozwiązaniem niż MLP. Dotychczasowe eksperymenty wykazały, że jakość klasyfikacji jest podobna dla wyżej wymienionych klasyfikatorów. Ekstrakcja cech pozwala na szybsze przeprowadzenie procesu klasyfikacji i uzyskanie odpowiedzi. Analizując otrzymane wyniki można stwierdzić, że czasy odpowiedzi są stabilne i mieszczą się w przedziale od 16 ms do 56 ms dla MLP oraz od 18 ms do 59 ms dla SVC. Dlatego w najlepszym przypadku, aby zapewnić dobrą wydajność systemu, możemy wysłać od 17 (56 ms) do 62 zapytań (16 ms) na sekundę.

Możliwe obszary zastosowania proponowanej metody szacowania liczby składowych głównych i cech:

- We wszystkich zastosowaniach ekstrakcji cech opierających się na składowych głównych i niezależnych.
- Do szacowania liczby skupień w metodach k-średnich, oraz hierarchicznej Warda.

W kolejnym rozdziale monografii ¹¹ zaproponowałem autorską metodę ekstrakcji cech bazującą na składowych głównych oraz gradientach stochastycznych. Głównym celem było zbudowanie modelu ekstrakcji cech obrazów, który pozwoli na uzyskanie co najmniej tak dobrych i lepszych jakości poprawnych klasyfikacji w porównaniu z innymi znanymi z literatury metodami ekstrakcji, w tym głębokich sieci neuronowych. Motywacja do badań wynikała z poruszanego z ekspertami medycyny problemu mało licznych zbiorów obrazów diagnostycznych. Takie zbiory często utrudniają, a nawet uniemożliwiają ekstrakcję cech np. z wykorzystaniem metod głębokich uczenia. Przyczyn mniejszych zbiorów może być wiele, natomiast w dziedzinie medycyny najczęściej są one związane albo z kosztem badań, albo z rzadkimi chorobami tj.: (a) zespół Alporta, (b) rzadkie choroby serca u dzieci, (c) mukowiscydoza, (d) zespół Marfana i inne.

¹¹ Część II. Metody ekstrakcji cech sygnałów o charakterystyce czasowej oraz dwuwymiarowych
"Model fuzji gradientów stochastycznych i składowych głównych w zadaniu ekstrakcji cech obrazu"

Istotą proponowanej przeze mnie metody sgPCA jest wykorzystanie gradientów stochastycznych w zadaniu ekstrakcji cech. W ogólnym przypadku punkt początkowy na obrazie wyznaczany jest w sposób losowy. W zaproponowanej metodzie punkty te wprawdzie są losowe lecz pula z jakiej są wybierane jest szacowana z wykorzystaniem analizy składowych głównych. W pierwszej kolejności dla obrazów typu RGB zdefiniowałem jasność piksela jako zbiór jasności z trzech kanałów kolorów: czerwonego (R), zielonego (G) i niebieskiego (B) w postaci $r_i = \{r_{Ri}, r_{Gi}, r_{Bi}\}$. W następnej kolejności przeprowadzana jest wstępna ekstrakcja cech z wykorzystaniem analizy składowych głównych. Cechy te nazywane są kluczowymi, a ich liczba jest zredukowana zgodnie z zaproponowanym kryterium 3.28 do liczby t . W analizowanym przypadku losowość determinująca stochastyczność gradientów polega na losowym wyborze składowej głównej i cechy. Zbiór wszystkich pikseli wyłonionych za pomocą metody PCA możemy zapisać w postaci: $r_{0i}, r_i \in [1, 2, \dots, n]$. W celu uniknięcia obliczania całkowitego gradientu, poszukiwanie lokalnego minimum jasności zastosowałem aktualizowanie gradientu stochastycznego

$$r_{i+1} = r_i - \eta_i \nabla f_{ij}(r_i), \quad (3.31)$$

gdzie η jest wektorem wag modelu.

W kolejnych krokach następuje aktualizacja gradientu $\nabla f(r_i)$. Ważnym elementem poszukiwania jest ustalenie odpowiedniego kroku. Przyjąłem, że wielkość kroku jest związana z kryterium minimalizacji wariancji gradientów, której celem jest minimalizacja fluktuacji przy optimum lokalnym. Aby zminimalizować wariancję gradientów w pierwszej kolejności przyjąłem, że φ jest średnią kroczącą (ruchomą). Natomiast φ_t jest adaptacyjną średnią kroczącą która będzie przyjmować wartości minimalizujące szacowaną wariancję dla każdej iteracji t . Zerowy moment gradientu zapisujemy w postaci

$$m_t^{(0)}(\varphi_t) = \varphi_t m_{t-1}^{(0)} \varphi_{t-1} + (1 - \varphi_t). \quad (3.32)$$

Pierwszy moment będzie miał postać

$$m_t^{(1)}(\varphi_t) = \varphi_t m_{t-1}^{(1)} \varphi_{t-1} + (1 - \varphi_t) g_t, \quad (3.33)$$

gdzie $g_t = 1 - \varphi_t$ jest rozmiarem kroku (współczynnikiem adaptacyjnym uczenia). Natomiast drugi moment gradientu zapiszemy w postaci

$$m_t^{(2)}(\varphi_t) = \varphi_t m_{t-1}^{(2)} \varphi_{t-1} + (1 - \varphi_t) g_t^2. \quad (3.34)$$

Ostatecznie szukamy takiego φ_t które będzie minimalizowało wariancję gradientów stochastycznych

$$\varphi_{ti} = \varphi \arg[\min(\sigma_{ti}^2)] = \varphi \arg[\min m_{ti}(\varphi)] - [m_{ti}(\varphi)]^2 \quad (3.35)$$

oraz

$$m_t(\varphi) = m_t^1(\varphi) / m_t^0(\varphi). \quad (3.36)$$

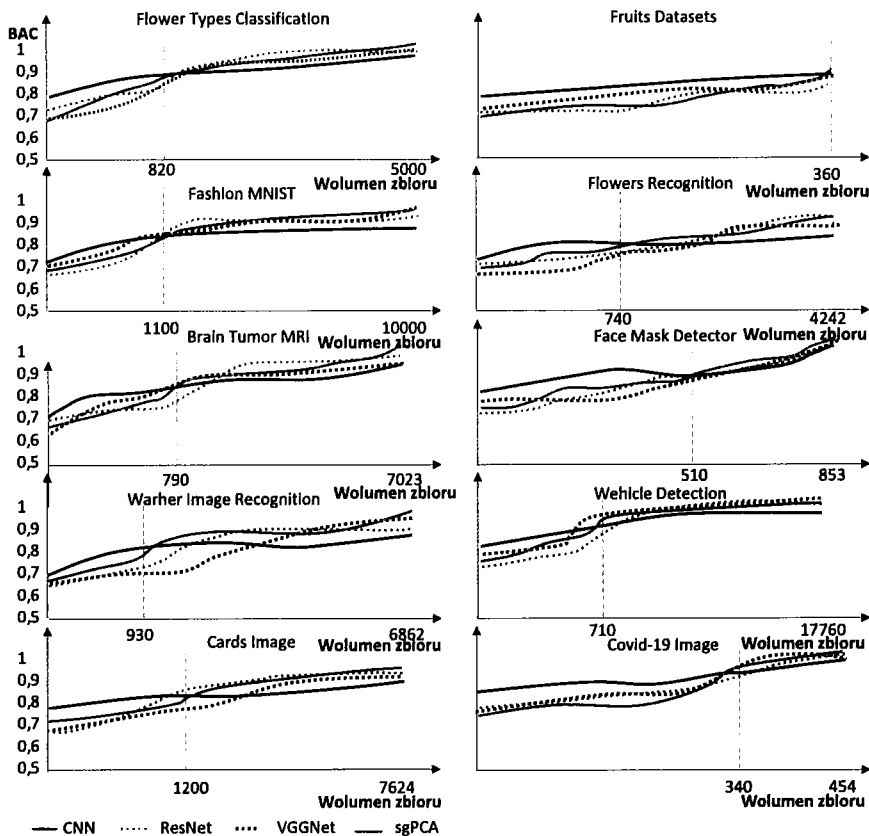
W ostatnim kroku, r_i zostaje zaktualizowana na wartość znaną w danym regionie obrazu zgodnie z metodą gradientów stochastycznych. W różnych fragmentach obrazu mogą istnieć różne lokalne minima, które, w zależności od punktu inicjującego, dostarczają nowych informacji o cechach obrazu. Całą procedurę powtarzamy dla każdej wyodrębnionej cechy. W ten sposób tworzy się nowa przestrzeń cech, które są następnie zapisywane w postaci tabelarycznej w celu klasyfikacji.

W części eksperymentalnej istotne było dla mnie porównanie opracowanej metody z innymi znanymi z literatury rozwiązaniami do wyodrębniania cech kluczowych tj.: SURF, BRIEF, ORB, BRISK oraz ekstraktorami obiektów tj.: HOG, LBP, FFT oraz PCA. Ponieważ wg różnych badań jednymi z najlepszych ekstraktorów cech z obrazów są głębokie sieci neuronowe w związku z czym zasadnym było dla mnie porównaniu jakości klasyfikacji również w stosunku do sieci tj.: CNN, RESNET oraz SGPCA. W eksperymentach zastosowałem protokół w postaci walidacji pięć-krotnej powtórzonej dwukrotnie. Aby ocenić powtarzalność wyników wykorzystałem 10 zbiorów z obrazami z ogólnodostępnego repozytorium KEEL (<https://keel.es>). Ponadto badania przeprowadziłem dla różnych klasyfikatorów bazowych tj.: MLP, SVC, CART, KNN oraz GNB.

Przeprowadzone przeze mnie eksperymenty wykazały, że jakość klasyfikacji po ekstrakcji opracowaną przeze mnie metodą jest istotnie statystycznie porównywalna i częściej lepsza niż po zastosowaniu metod: SURF, BRIEF, ORB, BRISK, HOG, LBP, FFT oraz PCA. W przypadku HOG dla różnych klasyfikatorów bazowych uzyskałem najbardziej porównywalne do autorskiej metody wyniki jakości klasyfikacji. Średnio dla 2-4 zbiorów obrazów opracowana przeze mnie metoda pozwoliła na uzyskanie lepszych wyników jakości klasyfikacji. W przypadku zbiorów z obiektami o wyraźniejszych krawędziach wyniki były porównywalne w metodzie SGPCA i HOG. Zauważyłem, że po ekstrakcji metodą HOG uzyskuje się duże wektory cech. Po zastosowaniu na tych wynikach powtórnej ekstrakcji np. metodą PCA można uzyskać porównywalne jakości klasyfikacji.

Istotnie ważne wyniki uzyskałem porównując opracowaną przeze mnie metodę SGPCA z różnymi architekturami głębokich sieci neuronowych. Uzyskane rezultaty przedstawiam na rysunku 3.5.

Analizując otrzymane rezultaty uzyskałem założony na wstępie cel. Oznacza to, że dla małych wolumenów zbiorów obrazów opracowana przeze mnie metoda pozwala na bardziej skuteczną ekstrakcję cech co



Rysunek 3.5: Wyniki jakości klasyfikacji BAC opracowanej metody ekstrakcji cech obrazu sgPCA oraz różnych architektur głębokich sieci neuronowych wg liczebności zbioru danych.

przekłada się na większą jakość klasyfikacji w porównaniu z analizowanymi głębokimi sieciami neuronowymi. Progi te oznaczyłem na rysunku 3.5 pionową przerywaną linią. W związku z powyższym można nakreślić następujące możliwe zastosowania opracowanej przeze mnie metody:

- dla małych zbiorów danych których pozyskanie wiąże się z kosztami badań i anotacji,
- w wyszukiwaniu kluczowych regionów na obrazie, które różnicują te obszary względem innych. Może to mieć szczególne zastosowanie w obrazach RTG złamań, gdzie wyłoniła się składowa może opisywać kość, a kolejne komponenty dalsze zmiany tj. pęknięcia, czy złamania,
- wykrywanie fragmentów uszkodzeń aorty serca u małych dzieci. Na temat tego rozwiązania aktualnie prowadzę rozmowy z ekspertami nauk medycznych ze Specjalistycznego Szpitala Zachodniego im. Św. Jana Pawła II,
- klasyfikacji udarów mózgu (na przykładzie obrazów z rezonansu magnetycznego) i jej wpływu na decyzje terapeutyczne odnośnie wyboru leku przeciwkrzepliwego.

W kolejnym rozdziale książki¹² zaproponowałem autorską metodę łą-

¹² Część II. Metody ekstrakcji cech sygnałów o charakterystyce czasowej oraz dwuwymiarowych "Zastosowanie transformaty falkowej oraz analizy składowych niezależnych w zadaniu ekstrakcji cech sygnałów"

czącą składowe niezależne z transformatą falkową w zadaniu ekstrakcji cech sygnału. Metoda może być zastosowana do różnych sygnałów. Jako przykład podałem przebieg EKG. Motywacją do wyboru takiego zbioru danych była potrzeba stworzenia narzędzia, które może być zastosowane do urządzeń wbudowanych i nie będzie wymagało dużego obciążenia obliczeniowego, jakie często jest związane z metodami uczenia głębokiego. W literaturze występują różne podejścia do zadania ekstrakcji sygnałów. Często wykorzystywane są pojedyncze modele. Można napotkać również prace w które przedstawiają różne algorytmy, jednak działają one niezależnie. Stosowanie pojedynczych metod tj. składowych głównych i niezależnych natrafia na trudności tj. wrażliwość wyników ekstrakcji na początkowe warunki, czyli początkowe estymaty składników sygnału. Różne początkowe estymaty mogą prowadzić do różnych wyników, co utrudnia osiągnięcie stabilnych i powtarzalnych rezultatów. Aby rozwiązać ten problem i poprawić jakość klasyfikacji, zastosowałem dodatkowo transformatę falkową, która modyfikuje macierz mieszania w metodzie ICA. Dzięki takiemu podejściu uzyskałem większą stabilność w analizie różnych składników pracy serca, takich jak skurcze komór, skurcze przedsionków czy zakłócenia elektryczne. Celem moich badań było stworzenie modelu ekstrakcji cech sygnałów, który łączyłby metodę ICA z transformatą falkową, a co najważniejsze, osiągałby co najmniej tak dobre wyniki jakości klasyfikacji jak inne znane metody ekstrakcji opisane w literaturze. Dodatkowo, porównałem opracowaną metodę z wynikami uzyskanymi za pomocą głębokich sieci neuronowych. Pełny model wstępnego przetwarzania, ekstrakcji i klasyfikacji przedstawiłem na rysunku 3.6.

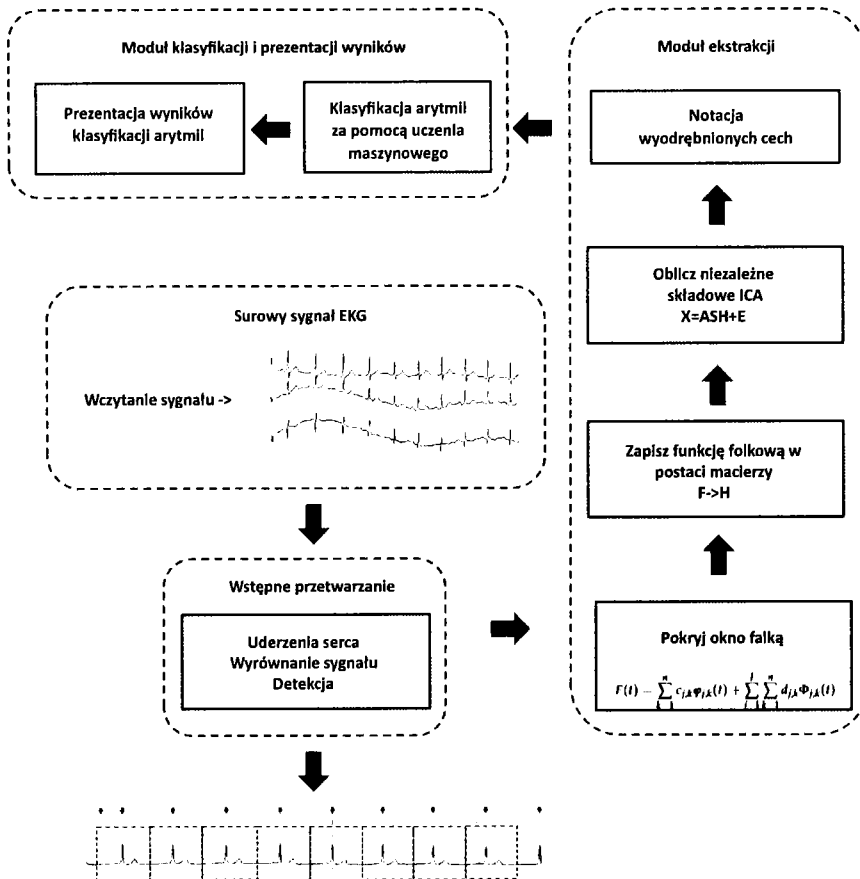
Najpierw przetwarzany jest surowy sygnał EKG. Wykonywana jest segmentacja pojedynczych taktów serca. Przed przystąpieniem do ekstrakcji należy wykonać wyrównanie załamka QRS do ustalonej szerokości okna d . Zaproponowałem aby okna były dobierane w taki sposób, aby dopasować ich środek do amplitudy R. Powoduje to wyrównanie fazowe załamka R w sygnale EKG ze środkiem funkcji falkowej. Niewspółosiowość może wprowadzać szum do wyodrębnionych elementów i wpływać na wydajność klasyfikacji. W celu ustalenia wpływu wyrównania sygnału na jakość ekstrakcji cech wprowadziłem współczynnik przesunięcia fazowego α . Wartość t_1 jest indeksem czasu rozpoczęcia okna, natomiast t_d jest indeksem czasu końca okna, a $d = t_d - t_1$ jest szerokością okna. Środek okna określa równanie

$$t_{\frac{d}{2}} = t_1 + \left\lfloor \frac{d}{2} \right\rfloor. \quad (3.37)$$

Niech t_R będzie indeksem czasowym pojawienia się fali R. Przesunięcie fazowe między $\frac{d}{2}$ a t_R jest wtedy zdefiniowane jako

$$\alpha = \frac{\left| t_R - t_{\frac{d}{2}} \right|}{\frac{d}{2}}. \quad (3.38)$$

W celu wykrycia załamka R zaproponowałem zastosowanie algorytmu Pan-Tompkinsa. Każdy segment jest wyrównany do ustalonego rozmiaru okna d i zawiera załamek R w środku. Drugim etapem jest ekstrakcja cech.



Rysunek 3.6: Diagram koncepcyjny proponowanej metody.

Rozpoczyna się od pokrycia segmentowanego sygnału funkcją falkową

$$F(t) = \sum_{k=1}^l c_{j,k} \varphi_{j,k}(t) + \sum_{j=1}^l \sum_{k=1}^l d_{j,k} \Phi_{j,k}(t), \quad (3.39)$$

gdzie $c_{j,k}$ jest to współczynnik aproksymacji dla skali j i lokalizacji k , $d_{j,k}$ jest to współczynnik szczegółowości, l jest wartością dekompozycji, l to długość oryginalnego sygnału, Φ to funkcja falkowa, $\varphi_{j,k}(t)$ jest filtrem dolnoprzepustowym i/lub górnoprzepustowym.

Ze względu na wykorzystanie notacji macierzowej w ICA, obliczona falka jest reprezentowana jako macierz współczynników L będących reprezentacją pojedynczego rytmu serca. Konwersję macierzy L przeprowadziłem z wykorzystaniem algorytmu Stranga. Wyjście algorytmu Stranga dla każdego sygnału EKG jest łączone w macierz H , którą wykorzystuje się w modelu matematycznym ICA.

W zaproponowanej przeze mnie metodzie założyłem, że zmienne są liniowymi kombinacjami niezależnych składowych

$$X = ASH + E, \quad (3.40)$$

gdzie $X_{|D| \times n}$ jest macierzą zawierającą przykłady aktualnie przetwarzanych przykładów uczących D , $A = [a_{pj}]_{n \times m}$ jest macierzą współczynników składowych głównych, $S = [S_1, S_2, \dots, S_m]$ jest macierzą składowych głównych, H to macierz współczynników wyodrębniona z sygnału źródła

dłowego za pomocą transformaty falkowej, $E = [E_1, E_2, \dots, E_m]$ jest macierzą szumów.

W kolejnym kroku obliczane są składowe niezależne. Funkcje wyodrębnione ze zmodyfikowanymi niezależnymi komponentami są przechowywane dla zadania uczenia. Transformata falkowa reprezentuje dodatkowe źródło informacji o sygnale EKG. Stosując ICA z transformatą falkową szukałem takiej transformacji, która maksymalizuje negentropię sygnałów wyjściowych, poszukując takich składników sygnału, które mają największą niegaussowskość.

W części eksperymentalnej wykorzystałem bazę danych PhysioNet MIT-BIH Arrhythmia Database <https://www.physionet.org/content/mitdb/1.0.0/>. Analizowałem przypadek pięcio klasowy. Ze względu na stosowanie powyższego zbioru do badań przez różnych autorów mogłem swoje wyniki skonfrontować z innymi znanymi z literatury. W części eksperymentalnej zastosowałem różne klasyfikatory bazowe tj.: MLP, CART, SVC, GNB oraz KNN, z których najlepszym okazała się sieć neuronowa MLP.

W eksperymencie pierwszym weryfikowałem, która z metod ekstrakcji tj. transformata falkowa i ICA pozwoli wyodrębnić cechy z sygnału EKG na podstawie których jakość klasyfikacji będzie lepsza. W tym celu do badań zastosowałem 10 różnych funkcji falkowych. Najlepsze jakości klasyfikacji 92,23% uzyskałem dla sieci neuronowej MLP i metody ICA. W przypadku transformaty falkowej najwyższą jakość poprawnych klasyfikacji BAC miały miejsce dla funkcji falkowej typu *db6*. Stosując autorską metodę i falkę *db6* udało mi się zwiększyć jakość klasyfikacji do 95,97% dla MLP. Taka kombinacja metody ICA i transformaty falkowej z falką *db6* były wariantem dającym najlepsze wyniki jakości klasyfikacji, które wykorzystano do porównań z innymi metodami występującymi w literaturze.

W kolejnym eksperymencie weryfikowałem, czy zastosowanie innych funkcji zamiast falkowych może przyczynić się do zwiększenia jakości klasyfikacji. Zaproponowałem rozwiązania w postaci rozkładów tj.: (a) Poissona, (b) wykładniczy, (c) t-Studenta, (d) Gaussa, (e) logarytmiczny oraz (f) logistyczny. W każdym eksperymencie wyniki dla tych rozkładów okazały się istotnie statystycznie gorsze o około 8%-10% od zaproponowanej przeze mnie metody.

W kolejnej części porównałem wyniki jakości klasyfikacji uzyskane za pomocą autorskiej metody z rezultatami innych zastosowanych metod prezentowanych w literaturze. W porównaniu z różnymi metodami bazującymi na składowych głównych, niezależnych oraz SVM uzyskałem istotnie lepsze wyniki dla zaproponowanej autorskiej metody. W przypadku głębokich sieci neuronowych otrzymałem gorsze wyniki jakości klasyfikacji.

Następnie przeprowadziłem eksperymenty dla 7 różnych rodzajów arytmii. Skutkowało to zmniejszeniem wolumenu danych. W rezultacie uzyskane jakości klasyfikacji z zastosowaniem fuzji transformaty falkowej oraz ICA były istotnie wyższe od sieci ALEXNET oraz VGGNET dla arytmii typu: (a) przedwczesnych kompleksów przedsionkowych APC, (b) nagłego stanu migotania komór VFW oraz przedwczesnych skróczów komór serca VEB.

Na podstawie przeprowadzonych badań można wskazać możliwe obszary zastosowania tj.:

- proponowana metoda wymaga mniejszej liczby obliczeń w czasie testu w porównaniu z głębokimi sieciami neuronowymi, co umożliwia szersze zastosowanie w przetwarzaniu brzegowym lub urządzeniach wbudowanych,
- do klasyfikacji sygnałów których koszt pozyskania jest duży i/lub związane są z rzadko występującymi przypadkami.

W trzeciej części monografii ¹³ zaproponowałem autorską metodę szacowania rozwiązań przestrzeni cech, z wykorzystaniem metaheurystyk. Metoda ta polega na szacowaniu rozwiązań wektorów cech przy użyciu funkcji: (a) dyskryminacji Lambda Wilksa, (b) odległości Mahalanobisa, (c) wariancji całkowitej oraz (d) V Rao. Celem było opracowanie metody ekstrakcji cech która umożliwi uzyskanie wyższej całkowitej wyjaśnianej wariancji w stosunku do rozwiązań binarnych. Finalnym celem było uzyskanie poprawy jakości klasyfikacji dla różnych metod szacowania rozwiązań przestrzeni cech. Zdecydowałem się na zastosowanie metaheurystyk z kilku powodów. Pierwszym była ich odporność na szum w danych, ponieważ mogą uwzględniać różne scenariusze i niezależnie oceniać wpływ cech na wyniki. Drugim powodem jest możliwość zastosowania wielu kryteriów optymalizacyjnych, co ma znaczenie np. w przypadku uwzględnienia jednocześnie kosztu pozyskania cech i jakości klasyfikacji.

Rozwiązanie R w zaproponowanej przeze mnie metodzie dla n - wymiarowego wektora cech ma postać $R = [\varphi_1 r_1, \varphi_2 r_2, \dots, \varphi_n r_n]$, gdzie $r_i \in \{0, 1\}$, natomiast $\varphi_i r_i \in [0, 1]$. W przypadku binarnym r_i przyjmuje wartość 1 w sytuacji kiedy dana i -ta cecha została wybrana do rozwiązania, natomiast wartość 0 w sytuacji przeciwnej. Dodatkowo wprowadziłem współczynnik $\varphi_i = \lambda_i$, który jest mocą dyskryminacyjną klas dla każdej cechy wyrażoną wzorem

$$\lambda = 1 - \prod_{i=1}^q \frac{1}{1 + \omega_i}, \quad (3.41)$$

gdzie ω_i jest i -tą wartością własną, q jest natomiast maksymalną liczbą funkcji dyskryminacyjnych.

Dla $r_i = 0$ dana cecha jest odrzucona. Natomiast kiedy $r_i = 1$ to o ostatecznym przyjęciu lub odrzuceniu cechy decyduje współczynnik φ_i . Wraz ze wzrostem φ_i rośnie moc dyskryminacyjna. Cecha będzie mimo $r_i = 1$ odrzucona kiedy $\varphi_i < \beta$. Natomiast współczynnik β jest hiperparametrem modelu i jest wyznaczany eksperymentalnie.

Takie podejście ma sens w sytuacji kiedy mimo przyjęcia danej cechy do rozwiązania jest siła dyskryminacyjna jest na słabym poziomie. Pierwszą przyjętą przeze mnie funkcją dyskryminacyjną była λ_i Wilksa.

Drugą moją propozycją jest wartość całkowitej wyjaśnianej wariancji wg klas $\varphi = F$

¹³ Część III. Zastosowanie metod inżynierii cech w zadaniu optymalizacji wielokryterialnej "Statystyczne metody zwiększania mocy dyskryminacyjnej w zadaniu ekstrakcji cech w ujęciu metaheurystyk"

$$F = \frac{1}{N} \sum n_i s_i^2, \quad (3.42)$$

gdzie s_i^2 jest wariancją cech w i -tej klasie, N jest to sumaryczna liczebność wszystkich wzorców, n_i jest liczebnością i -tej klasy, .

Kolejnym rozwiązaniem jest odległość Mahalanobisa $\varphi = D$

$$D = [X - \mu]^T K^{-1} [X - \mu]^{0,5}, \quad (3.43)$$

gdzie X jest wektorem cech, μ jest wektorem wartości średnich cech, K jest macierzą kowariancji cech, T - oznacza transpozycję macierzy.

Odległość Mahalanobisa uwzględnia zależności między cechami i pozwala na porównywanie obiektów o różnych wartościach cech.

Innym przyjętym przeze mnie kryterium jest V Rao, który jest estymatorem wariancji dla sondażu wielostopniowego. Można go obliczyć za pomocą wzoru

$$V = \frac{\sum w_i (\bar{x}_i - \bar{x})^2}{\sum w_i^2}, \quad (3.44)$$

gdzie $\sum w_i$ to suma wag i -tej klasy, \bar{x}_i jest wartością średnią cech w danej klasie, \bar{x} jest średnią wartością wszystkich cech.

Kluczowym elementem dowolnej metaheurystyki jest funkcja celu, która przypisuje każdemu rozwiązaniu R wartość określającą jego jakość. Umożliwia to porównywanie rozwiązań oraz wyznaczanie kierunku poszukiwań. Zaproponowana przeze mnie funkcja celu to F

$$F(R) = \frac{BAC(R)}{1 + \gamma \frac{n(R)}{n}}, \quad (3.45)$$

gdzie $BAC(R)$ to zbalansowana jakość klasyfikacji z wykorzystaniem podzbioru cech R , $n(R)$ jest liczbą cech w danym rozwiązaniu, a γ jest parametrem metody.

Zadanie optymalizacji dla różnych zastosowanych przeze mnie metaheurystyk polegało na maksymalizacji przyjętej funkcji celu. Parametrem, który określa, jak bardzo istotna jest ekstrakcja danych, jest γ . Jeżeli $\gamma = 0$ poszukiwane były rozwiązania R o najlepszej jakości klasyfikacji BAC bez względu na wielowymiarowość cech. W sytuacji kiedy $\gamma > 0$ akceptowane były gorsze rozwiązania przy mniejszej liczbie cech.

W części eksperymentalnej przeprowadziłem porównania różnych zaproponowanych przez siebie kryteriów dyskryminacyjnych tj.: (a) Lambda Wilska, (b) odległości Mahalanobisa, (c) wariancji całkowitej oraz (d) V Rao, dla różnych metaheurystyk tj:

- iteracyjnego przeszukiwania przestrzeni rozwiązań (ang. Hill Climbing, HC),
- symulowanego wyżarzania (ang. Simulated Annealing , SW),
- przeszukiwania tabu (ang. Tabu Search, TS),
- optymalizacji rojem cząstek (ang. Optimization Swarm, PSO),
- algorytmu mrówkowego (ang. Ant Colony Optimization, ACO),

- algorytmu genetycznego (ang. Genetic Algorithms, GA).

W badaniach zastosowałem protokół eksperymentalny w postaci walidacji pięciokrotnej powtórzonej dwa razy. Wyniki jakości klasyfikacji jakie uzyskałem dla różnych kryteriów dyskryminacyjnych w obrębie każdej zastosowanej metaheurystyki porównywałem testem statystycznym Wilcoxon na poziomie istotności $p < 0,05$.

Ważnym wnioskiem płynącym z przeprowadzonych badań jest to, że zastosowanie funkcji dyskryminacyjnej Lambdy Wilksa, zwłaszcza w przypadkach wieloklasowych, prowadzi do uzyskania lepszych wyników jakości klasyfikacji dla większości zbiorów danych. Dla każdej porównywanej metaheurystyki zaobserwowałem podobieństwo w kształtowaniu się różnic dla różnych kryteriów rozwiązań. Zastosowanie kryterium odległości Mahalanobisa, całkowitej wyjaśnianej wariancji oraz V Rao pozwoliły na uzyskanie w większości zbiorów danych porównywalnych rezultatów jak dla klasycznego rozwiązania binarnego. Najlepsze wyniki jakości klasyfikacji otrzymałem po przeprowadzeniu ekstrakcji cech algorytmem genetycznym GA, a następnie algorytmem mrówkowym PSO. Natomiast najniższe wyniki jakości klasyfikacji uzyskano dla HILL CLIMBING.

Możliwe zastosowania opracowanej metody obejmują zadania:

- ekstrakcji cech dla strumieni danych gdzie można samodzielnie odkrywać i ekstrahować cechy na podstawie zdefiniowanych kryteriów optymalizacyjnych,
- ekstrakcji danych które zawierają szum,
- ekstrakcji cech w problemach klasyfikacji wielokryterialnej,
- ekstrakcji cech z danych charakteryzujących się wysoką nieliniowością

W kolejnym rozdziale ¹⁴ zaproponowałem autorską metodę modyfikacji algorytmu MOGA (ang. Multiple Objective Genetic Algorithm), której celem jest optymalizacja wielokryterialna. Metoda ta wykorzystuje techniki niszowania oraz rankingowania rozwiązań niezdominowanych. Takie podejście umożliwia odnalezienie wyników znajdujących się na froncie Pareto rozwiązań niezdominowanych oraz zapewnia różnorodność w populacji. Celem jaki postawiłem jest modyfikacja algorytmu MOGA poprzez nową propozycję wyznaczania rangi dla danego rozwiązania. W proponowanej modyfikacji algorytmu mMOGA, zaproponowałem aby wartość *fitness* była wyznaczana na podstawie stosunku wariancji między niszami δ_{ii} do wariancji wewnątrz niższy δ_{iw} . W danej niszy znajdują się osobnicy o tych samych wartościach przedziału rang. Liczba takich przedziałów wynosi \sqrt{N} . Długość każdego przedziału wynosi $(r_{max} - r_{min}) / \sqrt{N}$. W rezultacie chciałem uzyskać algorytm, którego jakość klasyfikacji będzie tak dobra, a nawet lepsza w stosunku do innych algorytmów genetycznych stosowanych w optymalizacji wielokryterialnej.

W analizie decyzyjnej, identyfikacja i eliminacja rozwiązań zdominowanych stanowi ważny etap, ponieważ pozwala na skupienie się na rozwiązaniach optymalnych. Dzięki temu można zoptymalizować proces po-

¹⁴ Część III. Zastosowanie metod inżynierii cech w zadaniu optymalizacji wielokryterialnej "Ekstrakcja cech z wykorzystaniem zmodyfikowanego wielokryterialnego algorytmu genetycznego"

dejmowania decyzji i uzyskać najlepsze wyniki. W pierwszej fazie wszystkie rozwiązania traktuje się jako zdominowane. W pracy zaproponowałem inny sposób wyznaczania rangi dla danego rozwiązania i , który wyznaczamy w następujący sposób

$$r_i = 1 + h_i \left(1 - \prod_{i=1}^q \frac{1}{1 + \omega_i} \right), \quad (3.46)$$

gdzie ω_i jest i -tą wartością własną macierzy składowych głównych, q - jest natomiast maksymalną liczbą funkcji dyskryminacyjnych danego rozwiązania, r_i oznacza przypisaną rangę, natomiast h_i - liczbę osobników, które dominują nad rozwiązaniem i .

Minimalna ranga wynosząca 1 może występować w dwóch przypadkach. Pierwszy zakłada, że nie ma osobników, którzy dominują nad rozwiązaniem (czyli $h_i = 0$). Wariant drugi zakłada, że moc dyskryminacyjna klas danego rozwiązania wynosi zero. Ranga zatem wzrasta w przedziale od 1 do N w zależności od liczby rozwiązań które dominują nad danym rozwiązaniem oraz od ich mocy dyskryminacyjnej.

Jednym z ważnych mechanizmów jest tzw. *nisza* która zapewnia, że w populacji zachowana jest różnorodność i uniknięte zostają lokalne optima. Nisza to nic innego jak przestrzeń która jest charakterystyczna dla pewnych rozwiązań. Dzięki temu algorytm nie wykazuje zbieżności tylko do jednego rozwiązania ale wykazuje zdolność do odnajdywania rozwiązań w różnych częściach rozpatrywanej przestrzeni. Umieszczenie rozwiązań w niszy nie powoduje jednak zwiększenia istotnie ich różnorodności. W takim rozwiązaniu kondycja każdego osobnika znajdującego się w niszy jest zależna od jego sąsiedztwa.

Istotą badań była weryfikacja czy wprowadzenie dodatkowej informacji do rangowania rozwiązań w postaci kryterium mocy dyskryminacyjnej oraz uwzględniania stosunku wariancji między niszami do wewnątrz nisz we współczynniku oddziaływania między rozwiązaniami wpływa na poprawę klasyfikacji przy zastosowaniu metody optymalizacji wielokryterialnej MOGA.

W proponowanym podejściu oprócz rozwiązań $R_{(mMOGA)}$ wyekstrahowanego wektora cech metodą mMOGA szacowane jest również rozwiązanie $R_{(CCPCA)}^{(k)}$ dla każdej składowej k w wykorzystaniem CCPCA, która jest opisana w pierwszej części książki. Ostateczne bazowe rozwiązanie jest wyznaczane

$$\forall k \quad R^{(k)} = R_{mMOGA} \cap R_{CCPCA}^{(k)}. \quad (3.47)$$

W metodologii badawczej wykorzystałem protokół eksperymentalny walidację krzyżową 5-cio krotną powtórzoną dwa razy. Eksperymenty przeprowadziłem na 41 zbiorach danych rzeczywistych. W badaniach zastosowałem algorytmy: (a) TOPSIS b) WSM (c) NSGA-II (d) MOGA oraz (e) mMOGA będąca modyfikacją MOGA która została opisana w sekcji z metodą. Analizowałem różne klasyfikatory bazowe tj.: (a) KNN, (b) (c) MLP, (d) CART oraz (e) GNB. Najlepsze jakości poprawnych klasyfikacji uzyskałem w większości zbiorów danych dla sieci neuronowej MLP, której wyniki zaprezentowałem w monografii. W eksperymentach przyjąłem

kryteria optymalizacyjne w postaci: (a) maksymalizacji czynników ortogonalnych i zbalansowanej dokładności BAC, (b) maksymalizacji funkcji logarytmu wiarygodności i zbalansowanej dokładności BAC.

Na podstawie uzyskanych wyników możemy wnioskować, że opracowana metoda mMOGA jest co najmniej porównywalna do metod tj.: TOPSIS, WSM, NSGA-II oraz MOGA. Częściej jest ona lepsza pod względem zbalansowanej dokładności od metody MOGA. W przypadku 5 z 41 zbiorów danych dla czynników ukośnych oraz ortogonalnych i 2 na 41 zbiorów dla maksymalizacji funkcji logarytmu wiarygodności mMOGA uzyskała lepsze jakości klasyfikacji w porównaniu z innymi metodami. Zauważyłem również, że metody tj. NSGA-II, MOGA i mMOGA uzyskują często lepsze wyniki jakości klasyfikacji w porównaniu z metodami TOPSIS oraz WSB. Analizując złożoność czasową, to metody TOPSIS oraz WSM w krótszym czasie dokonują optymalizacji wielokryterialnej niż NSGA-II, MOGA i mMOGA.

Możliwe zastosowania praktyczne opracowanej metody:

- klasyfikacja różnych problemów uwzględniająca różne kryteria tj. jakość i koszt, co ma szczególne znaczenie w obszarach w których cechy wiążą się z dużym kosztem pozyskania,
- w zadaniach klasyfikacji leczenia nowotworów za pomocą radioterapii, gdzie konieczne jest uwzględnienie dawkowania promieni, uwzględniając różne czynniki, takie jak: (a) minimalizacja uszkodzenia zdrowych tkanek, (b) maksymalizacja dawki dla guza i (c) minimalizacja czasu trwania leczenia,
- w zadaniu wyboru najlepszego wariantu leczenia farmakologicznego uwzględniając różne czynniki, takie jak skuteczność leków, efekty uboczne i koszty,
- w zadaniu klasyfikacji rodzaju operacji, gdzie należy uwzględnić kryteria tj.: (a) minimalizacja czasu trwania operacji, (b) minimalizacja utraty krwi, (c) minimalizacja ryzyka powikłań i (d) maksymalizacja szans na powodzenie operacji.

W ostatniej części monografii ¹⁵ zaproponowałem metodę ekstrakcji cech w zadaniu optymalizacji wielokryterialnej dla zespołów klasyfikatorów.

W proponowanej metodzie, w pierwszej kolejności należy przeprowadzić analizę składowych głównych ze względu na centroidy klas, co zostało opisane w pierwszej części pierwszego rozdziału niniejszej monografii. Aby zapewnić różnorodność rozwiązań, nie agreguje się komponentów dla różnych klas, lecz pozostawia w wyłonionej surowej postaci. Każda wyodrębniona składowa posiada wartość własną ω_k .

$$CCPCA_{j,k} = [\omega(1), \omega(2), \dots, \omega(k)]. \quad (3.48)$$

Dla wyodrębnionych podprzestrzeni powstał zbiór cech na podstawie których będzie dokonywana klasyfikacja, przez jeden z m klasyfikatorów bazowych będących w zespole. Ostateczne dla wyłonionych k komponentów możemy zapisać zespół klasyfikatorów w postaci

¹⁵ Część III. Zastosowanie metod inżynierii cech w zadaniu optymalizacji wielokryterialnej "Zastosowanie optymalizacji wielokryterialnej do budowy zespołów klasyfikatorów"

$$\Psi = [\Psi_1, \Psi_1, \dots, \Psi_k]. \quad (3.49)$$

W zadaniu optymalizacji wielokryterialnej zastosowano dwa kryteria. Pierwszym z nich jest zbalansowana dokładność BAC, natomiast drugim jest InfoGain (ang. Information gain), który mierzy zysk informacyjny dla każdej cechy niezależnie. W tym celu zostało zastosowane kryterium

$$InfoGain(X, Y) = Entr(Y) - Entr(X), \quad (3.50)$$

gdzie $Entr(Y)$ jest entropią zmiennej objaśnianej tj. klasy, natomiast $Entr(X)$ jest entropią warunkową.

Entropię $Entr(Y)$ możemy zapisać jako

$$Entr(Y) = - \sum_{y \in Y} p(y) \log p(y). \quad (3.51)$$

Natomiast entropię $Entr(X)$ zapisujemy jako:

$$Entr(X) = - \sum_{x \in X} p(x) Entr(x). \quad (3.52)$$

Każde z tak utworzonych k rozwiązań, co jest równie liczbie składowych głównych stanowi przestrzeń cech na których jest przeprowadzana klasyfikacja przez jeden algorytm będący częścią zespołu klasyfikatorów.

W badaniach zastosowałem 41 rzeczywistych zbiorów danych oraz protokół eksperymentalny z walidacją 5x2cv.

Wykorzystałem zespoły klasyfikatorów homogenicznych tj.: KNN, CART, GNB, SVC, MLP. W prezentacji wyników przedstawiłem jedynie rezultaty uzyskane za pomocą sieci neuronowej MLP, która okazała się najlepszym klasyfikatorem dla rozpatrywanego zadania. Decyzja z zespołu klasyfikatorów była podejmowana na podstawie reguły agregacji tzw. głosowania większościowego. Porównałem również różne metody optymalizacji wielokryterialnej tj.: (a) TOPSIS, (b) WMS, (c) NSGS-II, (d) MOGA oraz (e) MMOGA.

Na początku przeprowadziłem optymalizację wielokryterialną dla różnych ww. metod. W ten sposób otrzymałem rozwiązania R z postaci przestrzeni cech. Następnie dla trzech rodzajów metod wyodrębniania w metodzie CCPCA tj.: (a) składowych ukośnych, (b) składowych ortogonalnych oraz (c) funkcji logarytmu wiarygodności otrzymałem dla każdej składowej rozwiązania $R_{CCPCA}^{(k)}$. Finalnie były one agregowane

$$\forall k \quad R^{(k)} = R \cap R_{CCPCA}^{(k)}. \quad (3.53)$$

Każda z tak wyodrębnionych składowych głównych stanowiła wektor cech dla pojedynczego klasyfikatora z zespołu.

Dodatkowym kryterium optymalizacyjnym dla MMOGA przy wyborze najbardziej reprezentatywnych zespołów były zysk informacyjny oraz zbalansowana dokładność BAC. Test Wilcoxon na poziomie istotności $p < 0,05$ wykazał, że zaproponowana modyfikacja metody MOGA z CCPCA dla czynników ortogonalnych, ukośnych oraz metody funkcji logarytmu wiarygodności pozwoliło osiągnąć najlepsze wyniki poprawnych klasyfikacji BAC.

W każdym z eksperymentów wykazałem, że zaproponowana metoda budowy zespołów klasyfikatorów w przypadku MOGA z CCPCA dla części zbiorów ma porównywalną jakość klasyfikacji do MOGA. Dla przypadków wieloklasowych jakości zbalansowanej dokładności dla MOGA z CCPCA częściej jest istotnie wyższa w porównaniu z: (a) TOPSIS, (b) WSM, (c) NSGA-II oraz (d) MOGA. Wysoka różnorodność kombinacji cech w komponentach jest przyczyną poprawy wyników klasyfikacji dla przypadków wieloklasowych.

Przykłady zastosowania praktycznej opracowanej metody:

- zadanie do rozpoznawania różnych rodzajów chorób które są silnie determinowane przez pewne kluczowe cechy przy jednoczesnym wyborze najlepszego leczenia dla pacjenta, uwzględniając różne czynniki, takie jak skuteczność, koszty i skutki uboczne.
- do klasyfikacji i predykcji rynków finansowych charakteryzujących się różnymi konstruktami cech z optymalizacją wielokryterialną uwzględniającą różne cele inwestycyjne, ryzyko i zwrot z inwestycji.

5.1 Podsumowanie osiągnięcia naukowego i kierunki dalszych badań

Prace opisane w ramach zaprezentowanej monografii stanowią podsumowanie kolekcji metod ekstrakcji cech pozwalających modelom klasyfikacji rozwiązywać szerokie spektrum problemów decyzyjnych.

Koncentracja trudności, przykładowo, w wielokryterialnych metodach analizy danych, czy dla mało licznych zbiorów zawierających dodatkowo wartości odstające wymaga metod specyficznych, radzących sobie z wieloma wyzwaniami równocześnie.

Zaproponowany przeze mnie zbiór metod stara się wychodzić tym trudnościom naprzeciw, wykazując w większości przypadków statystycznie istotnie lepsze wyniki nad rozwiązaniami znanymi z literatury, zalecanymi często do typowych problemów, dostosowanymi do ogólnych przypadków klasyfikacji danych.

Znacząca część opracowanych przeze mnie metod opiera się na składowych głównych i niezależnych, metaheurystykach i optymalizacji wielokryterialnej. Na szczególną uwagę zasługuje opracowana metoda ekstrakcji cech która poprzez rotację wg centroidó klas wydobywa cechy najlepiej je dyskryminujące. Metoda ta ma praktyczne zastosowania w zadaniu klasyfikacji i została wykorzystana w psychiatrii^{16,17} przy projektowaniu aparatury do psychotestów z wykorzystaniem metod uczenia maszynowego wspomagającej klasyfikację.

W badaniach rozwiązałem problem ekstrakcji cech w warunkach dużej zmienności wartości cech, szumów oraz wyników odstających tzw. *outlier*. W tym celu zaproponowałem w metodzie składowych głównych zamiast podejścia regresyjnego gradienty stochastyczne. Zostały one również wykorzystane w metodzie ekstrakcji cech obrazów oraz szukania kluczowych cech.

Innym ważnym osiągnięciem naukowym jest opracowana przeze mnie metoda imputacji brakujących danych oraz szacowania liczby cech i skła-

¹⁶ Anna Maria Kalenik i in. "The impact of the COVID-19 pandemic on the mental health of children with psychiatric diagnoses—multidimensional CCPCA Model". W: *BMC psychiatry* 22.1 (2022)

¹⁷ Katarzyna Nowicka-Sauer i in. "Applying Machine Learning to Construct a Model of Risk of Depression in Patients Following Cardiac Surgery with the Use of the SF-12 Survey". W: *nt. J. Environ. Res. Public Health* 20.6 (2023)

dowych głównych, które wyznaczone są analitycznie. Jest to ważne osiągnięcie, gdyż dotychczas wybór cechy i komponentów był dokonywany arbitralnie bądź na podstawie kryteriów *Kaisera* oraz *osypiska czynników*. Często powodowało to niepełne dopasowanie przestrzeni cech do rozpatrywanego zadania klasyfikacji.

Wychodząc na przód oczekiwaniom kardiologii, współpracując z ekspertami naukowymi w dziedziny medycyny opracowałem metodę ekstrakcji cech sygnału EKG będącej fuzją metody składowych niezależnych i transformaty falkowej. Rezultatem było osiągnięcie lepszych wyników klasyfikacji specyficznych typów arytmii serca od głębokich sieci neuronowych. Istotnym atrybutem tej metody jest to, że może ona być zastosowana w urządzeniach wbudowanych cechujących się mniejszą mocą obliczeniową i dla małych zbiorów danych.

Opracowałem również modyfikacje metod optymalizacji wielokryterialnej w zadaniu ekstrakcji cech dzięki którym dla różnych kryteriów optymalizacyjnych uzyskałem przestrzenie cech charakteryzujące się większym procentem wyjaśnianej całkowitej wariancji, co przełożyło się na wyższe istotnie jakości klasyfikacji w porównaniu z innymi metodami znanymi z literatury.

Do najważniejszych osiągnięć opisanych w monografii zaliczyć można:

- **Opracowanie metody ekstrakcji cech bazującą na składowych głównych oraz na rotacji czynników wg centroidów klas**, która pozwala na poprawienie jakości klasyfikacji poprzez rotację czynników według centroidów klas. Znalazła ona zastosowanie praktyczne w medycynie oraz przy projektowaniu aparatury do psychotestów.
- **Zaproponowanie algorytmu ekstrakcji wykorzystującą gradienty stochastyczne przy wyodrębnianiu składowych głównych**, który pozwala na ekstrakcję cech charakteryzujących się dużą zmiennością, nielinowością oraz wartościami odstającymi.
- **Opracowanie metody imputacji brakujących danych** pozwalającej na uzupełnienie brakujących danych z wykorzystaniem składowych głównych, co umożliwiło dokładniejszą imputację w przypadku większego odsetka brakujących danych (20-50%) w porównania ze znanymi z literatury metodami.
- **Sformułowanie algorytmu szacowania liczby cech i składowych głównych** pozwalającego na analityczne wyznaczanie liczby składowych głównych oraz cech w ich obrębie zwiększającą całkowitą wyjaśnianą wariancję cech.
- **Zaproponowanie metody ekstrakcji cech obrazu**, w której do wyznaczenia kluczowych cech obrazu oraz ekstrakcji cech oparto się na składowych głównych oraz gradientach stochastycznych.

- **Opracowanie algorytmu ekstrakcji cech sygnałów o przebiegach czasowych** wykorzystującego fuzję analizy składowych niezależnych oraz transformaty falkowej. Eksperymenty bazowano na przebiegu sygnału EKG.
- **Sformułowano metodę ekstrakcji cech sygnałów w oparciu o metaheurystyki** uwzględniającą zastosowanie wskaźników do rozwiązań wyekstrahowanej przestrzeni cech, tj.: mocy dyskryminacyjnej Lambda Wilksa, odsetku całkowitej wyjaśnianej wariancji, odległości Mahalanobisa i współczynnika V Rao.
- **Zaproponowanie algorytmu optymalizacji wielokryterialnej** wykorzystującego metodę ekstrakcji składowych ortogonalnych, głównych i logarytmu funkcji wiarygodności w szacowaniu rozwiązań przestrzeni cech najsilniej dyskryminującej klasy.
- **Opracowanie rozwiązania optymalizacji wielokryterialnej dla zespołów klasyfikatorów** wykorzystującego autorską modyfikację algorytmu genetycznego MOGA oraz metodę ekstrakcji składowych ortogonalnych, głównych i logarytmu funkcji wiarygodności w szacowaniu rozwiązań przestrzeni cech najsilniej dyskryminującej klasy. Jako kryteria optymalizacyjne przyjęto zysk informacyjny i zbalansowaną dokładność.

Część uzyskanych rezultatów została wcześniej opublikowana w formie artykułów w czasopiśmie i referatów na konferencjach międzynarodowych, gdzie była wykorzystywana w zadaniu uczenia maszynowego do budowy modelu ryzyka depresji u pacjentów po operacjach kardiologicznych¹⁸ oraz do oceny wpływu pandemii COVID-19 na zdrowie psychiczne dzieci z diagnozami psychiatrycznymi¹⁹, a także w budowie narzędzi diagnostycznych dla psychologów.

Zastosowanie zaproponowanych metod ekstrakcji w przypadku mniejszych wolumenów danych okazuje się mieć kluczowe znaczenie w sytuacjach, kiedy koszty pozyskania badań czy ich etykietyzacji są wysokie. Opracowane metody ekstrakcji są adresowane nie tylko do osób zajmujących się omawianą tematyką, ale również do studentów oraz naukowców pracujących nad zgłębianiem i rozwijaniem dziedziny.

Opracowane metody mogą mieć zastosowanie w zbiorach:

- zawierających trudne dane,
- cechujących się wielowymiarowością,
- niejednorodnych, z brakami danych,
- dla których wymagana jest wysoka dokładność,
- dla których ważna jest interpretowalność uzyskanych cech,
- wymagających specjalistycznej wiedzy.

¹⁸ Katarzyna Nowicka-Sauer i in. "Applying Machine Learning to Construct a Model of Risk of Depression in Patients Following Cardiac Surgery with the Use of the SF-12 Survey". W: *nt. J. Environ. Res. Public Health* 20.6 (2023)

¹⁹ Anna Maria Kalenik i in. "The impact of the COVID-19 pandemic on the mental health of children with psychiatric diagnoses—multidimensional CCPCA Model". W: *BMC psychiatry* 22.1 (2022)

Zagadnienia przedstawione w książce nie obejmują całej tematyki związanej z ekstrakcją cech. Dziedzina ta nadal dynamicznie się rozwija i wymaga kolejnych usprawnień. Mając to na uwadze, nakreślić można dalsze kierunki badań:

- opracowanie metod ekstrakcji cech dla zadania uczenia nadzorowanego, w tym uczenia aktywnego;
- opracowanie metod ekstrakcji cech dla detekcji z obrazów przesyłanych w czasie rzeczywistym, co może mieć szerokie zastosowanie w problemach medycznych;
- dalsze rozwijanie metod ekstrakcji w optymalizacji wielokryterialnej redukującej złożoność obliczeniową dla wielu kryteriów, co w efekcie ma zmniejszyć czas uczenia; jest to szczególnie istotne w problemach, dla których złożoność obliczeniowa ma kluczowe znaczenie.

Ze względu na charakter pracy pominięto wiele aspektów dynamicznej ekstrakcji cech, niezbędnych w przetwarzaniu strumieni danych czy sekwencji wideo. Niemniej, wyniki otrzymane w badaniach eksperymentalnych wskazują na poprawę jakości klasyfikacji z wykorzystaniem przedstawionych algorytmów w porównaniu z innymi metodami znanymi z literatury.

Oprócz tematyki zawartej w prezentowanej monografii moje zainteresowania naukowe dotyczą również innych tematów związanych ze sztuczną inteligencją oraz wielowymiarowym przetwarzaniem danych na potrzeby uczenia maszynowego. W przypadku metod sztucznej inteligencji prowadzą badania związane z głębokimi sieciami neuronowymi i ich zastosowaniem w przetwarzaniu i detekcji obrazów, szczególnie na potrzeby obrazowania z rezonansu magnetycznego i wykrywania anomalii płodu we wczesnym stadium. Wykorzystuje również metody głębokiego uczenia do ekstrakcji cech pracy serca u osób dużego ryzyka niedokrwienia mięśnia sercowego.

Prowadzę również prace badawcze związane z uczeniem aktywnym. Rozwijam metody uczenia aktywnego do wykrywania anomalii w systemach finansowych i cyberbezpieczeństwie. Pracuję nad algorytmami które aktywnie eksplorują różne wzorce i anomalie w danych, aby zidentyfikować potencjalne przypadki nadużyć finansowych.

W przypadku technik statystycznych rozwijam metody tj.: modele ekstrakcji i selekcji, modele przyczynowe szeregów czasowych, analiza spektralna oraz różne metodyki data mining.

Opracowana przeze mnie metoda CCPCA została praktycznie wdrożona w komputerowo wspomaganych testach psychologicznych, które są wykorzystywane w praktyce psychologicznej i placówkach medycznych. Potwierdzenie w postaci rekomendacji zamieściłem na końcu niniejszego załącznika.

6 *Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej*

W okresie od obrony pracy doktorskiej, aż do dzisiaj uczestniczyłem w wielu różnych projektach naukowo-badawczych które były realizowane przez różne jednostki naukowe z Polski i zagranicy. Opracowane przeze mnie metody ekstrakcji cech w zadaniu uczenia maszynowego były i są popularyzowane w różnych obszarach naukowych tj. informatyka, medycyna, budownictwo, zarządzania etc. W ramach aktywności naukowej przeprowadziłem wiele szkoleń, wykładów, konsultacji z zakresu wielowymiarowej analizy danych, ekstrakcji cech i metod uczenia maszynowego. Łącznie współpracowałem z ponad 15 jednostkami naukowymi Uczelni w Polsce. Najważniejsze z nich przedstawiam poniżej.

W roku 2021 współpracowałem naukowo z zespołem Uniwersytetu Jagiellońskiego Collegium Medicum Katedry Mikrobiologii w zakresie analizy wielowymiarowej danych w tym data mining. Współpraca dotyczyła opracowania modeli klasyfikacji danych na potrzeby mikrobiologii. W efekcie powstał wspólny artykuł naukowy ²⁰.

W roku 2022 podjąłem współpracę naukową z zespołem naukowym lekarzy psychiatrów Warszawskiego Uniwersytetu Medycznego, Kliniki Psychiatrii Wieku Rozwojowego. Celem podjętych wspólnie konsultacji było wypracowanie modeli wspomagających klasyfikacje dzieci z diagnozami psychiatrycznymi w dobie COVID-19. W ramach współpracy powstały różne modele ekstrakcji cech i klasyfikacji z zastosowaniem drzewa klasyfikacyjnego CART. Owocem wspólnych prac jest opublikowany artykuł naukowy ²¹.

W latach 2019-2022 prowadziłem również współpracę naukową z zespołem naukowo-badawczym ówczesnego Instytutu Informatyki i Zarządzania Politechniki Wrocławskiej. Moim wkładem było tworzenie modeli ekstrakcji cech oraz wielowymiarowej analizy danych, w tym tzw. data mining. W latach 2020 i aktualnie współpracuję z zespołem Wydziału Budownictwa Lądowego i Wodnego, Katedry Inżynierii Materiałów i Procesów Budowlanych Politechniki Wrocławskiej, gdzie we współpracy z Panią Profesor Bożeną Hołą wdrażam w projektach różne metody uczenia maszynowego do klasyfikacji i predykcji wypadków w budownictwie ²². Taka współpraca pozwala mi rozwijać moje modele dla różnych obszarów nauki. Owocem współpracy jest wspólny artykuł naukowy poświęcony predykcji wypadków w budownictwie. Aktualnie jestem promotorem pomocniczym pracy doktorskiej i wspomagam merytorycznie doktorantkę w obszarze uczenia maszynowego i ekstrakcji cech.

W latach od 2020 r. do dzisiaj współpracuję z zespołem Wydziału Filologicznego Uniwersytetu Gdańskiego w zakresie prac naukowych związanych z zastosowaniem metod uczenia maszynowego i wielowymiarowej analizy danych w zagadnieniach organizacji i zarządzania bibliotekami, marketingu bibliotecznego, promocji czytelnictwa oraz zarządzania wiedzą w instytucjach komercyjnych i non profit. W ramach współpracy powstał artykuł naukowy który aktualnie jest w recenzji wydawniczej "*Prediction of social capital levels among library employees using machine learning*

²⁰ Wojkowska-Mach J i in. "Antibiotic consumption in long-term care facilities in Poland and other European countries in 2017". W: *Antimicrobial Resistance Infection Control* 10.1 (2021), s. 1-7

²¹ Anna Maria Kalenik i in. "The impact of the COVID-19 pandemic on the mental health of children with psychiatric diagnoses—multidimensional CCPCA Model". W: *BMC psychiatry* 22.1 (2022)

²² Hoła Bożena i in. "Prediction model of seasonality in the construction industry based on the accidentality phenomenon". W: *Archives of Civil and Mechanical Engineering* 22 (2022), s. 1-13

methods".

Wymieniam się doświadczeniami z zakresu projektowania eksperymentów rozpoznawania wzorców – wspólnie z zespołami *Uniwersytetu Gdańskiego* i *Gdańskiego Uniwersytetu Medycznego* które przyczyniły się do realizacji wspólnych projektów naukowo-badawczych, których owocem jest wspólny artykuł²³ dotyczący zastosowania metod uczenia maszynowego do budowy modelu i klasyfikacji ryzyka depresji u pacjentów po operacjach kardiochirurgicznych z wykorzystaniem kwestionariusza SF-12. Podczas blisko sześcioletniej współpracy z zespołami naukowo-badawczymi uczestniczyłem we wspólnej realizacji dziewięciu projektów popularyzując zastosowanie różnych metod ekstrakcji cech oraz uczenia maszynowego. W ramach współpracy w roku 2021 została podpisana umowa o współpracę między *Uniwersytetem Gdańskim*, *Gdańskim Uniwersytetem Medycznym* oraz *Politechniką Wrocławską*, gdzie jestem w roli osoby kierującej projektem którego tematem jest "Zastosowanie metod uczenia maszynowego w predykcji poziomu wypalenia zawodowego, prężności oraz poczucia własnej skuteczności u psychologów klinicznych i lekarzy psychiatrów". Badania są realizowane pod patronatem Rzecznika Praw Pacjenta. Aktualnie opracowywany jest wspólny artykuł naukowy. Potwierdzenie współpracy zamieściłem na końcu niniejszego załącznika. W toku swojej pracy naukowej odbyłem także dwa krótko i długotrzmnowe staże naukowe:

- Pierwszy staż naukowy zrealizowałem w ramach współpracy z *National Mining University of Ukraine - Dnepropetrowsk*²⁴, który zakończył się nawiązaniem trwałej współpracy badawczej z zespołem *The Institute of Economics at the National Mining University in Dnietropietrowsk* kierowanym przez profesora Alexandra Ivanovicha Sharova. Współpraca ta wyraża się zarówno we wspólnych pracach badawczych, jak i w wymianie doświadczeń projektowych, której podsumowaniem może być wygłoszony wykład dla pracowników naukowych i doktorantów ww uczelni z zakresu zastosowania metod uczenia maszynowego w naukach ekonomicznych. W ramach stażu prowadziłem badania związane zastosowaniem metod uczenia maszynowego w zadaniach wykrywania nadużyć finansowych. Ponadto brałem udział w sympojach na których wymieniałem się doświadczeniami naukowo-badawczymi oraz metodami kształcenia studentów jednocześnie popularyzując zastosowanie sztucznej inteligencji w różnych problemach decyzyjnych i predykcyjnych. Przeprowadziłem również ponad 90 godzin zajęć dydaktycznych ze studentami z Informatyki oraz warsztaty z zakresu sztucznej inteligencji dla ówczesnych kół naukowych. Opracowałem również ekspertyzy z zakresu wyposażenia laboratoriów w sprzęt umożliwiający prowadzenie badań naukowych oraz dydaktyki. Staż odbyłem w okresie: 1.04.2013 – 15.11.2013.

²³ Katarzyna Nowicka-Sauer i in. "Applying Machine Learning to Construct a Model of Risk of Depression in Patients Following Cardiac Surgery with the Use of the SF-12 Survey". W: *Int. J. Environ. Res. Public Health* 20.6 (2023)

²⁴ *National Mining University in Dnietropietrowsk* Odesa, Ukraine

- Drugi staż naukowy zrealizowałem w Instytucji Naukowej *Fundación ICIL*²⁵, co pozwoliło mi na rozwinięcie badań z zakresu klasyfikacji różnych scenariuszy łańcuchów dostaw. W ramach rozwijania zakresu współpracy analizowana była tematyka *sztucznej inteligencji* dla zagadnień planowania i klasyfikacji łańcuchów dostaw w krajach Unii Europejskiej. W trakcie stażu przeprowadziłem warsztaty i wykład na temat zastosowania metaheurystyk oraz metod uczenia maszynowego w zadaniu klasyfikacji łańcuchów dostaw. Staż odbyłem w okresie: 3.09.2014 – 17.09.2014.

²⁵
Fundación ICIL, Barcelona
Hiszpania

Potwierdzenie odbytych staży zamieściłem na końcu niniejszego załącznika.

Oprócz staży zagranicznych odbyłem również dwa staże krajowe w przedsiębiorstwie VISTEL (09.04.2014 - 09.08.2014) oraz GOLF TELECOM (15.11.2012 - 15.03.2013). Brałem udział w badaniach dotyczących wdrażania metod sztucznej inteligencji dla potrzeb przemysłu. Ponadto pełniłem funkcję eksperta z zakresu algorytmów metod optymalizacji. Potwierdzenie odbytych staży oraz ich zakres zamieściłem na końcu niniejszego załącznika.

Również aktywnie udzielam się w przemyśle zwłaszcza w obszarze transportu dla którego opracowuję różne metody optymalizacji oraz ekstrakcji dla zadania planowania tras w złożonych łańcuchach dostaw. Potwierdzenie w postaci zaświadczeń od przedsiębiorstw: Demotrans, FTS oraz Tom Trans zamieściłem na końcu niniejszego załącznika.

7 Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę

Podczas blisko dwudziestoletniej pracy naukowo-dydaktycznej prowadziłem trzynaście kursów dla studentów kierunków *Informatyka, Informatyka Techniczna* i *Teleinformatyka*, w siedmiu (pierwszych wymienionych) będąc głównym wykładowcą i autorem materiałów dydaktycznych:

1. Głębokie sieci neuronowe.
2. Miękkie metody obliczeniowe.
3. Metody techniki systemów w medycynie.
4. Zastosowania informatyki w medycynie.
5. Projektowanie telemedycznych systemów internetowych i mobilnych.
6. Aplikacje mobilne.
7. Współczesne trendy w Informatyce.
8. Uczenie maszyn.
9. Projekt naukowo – wdrożeniowy.
10. Podstawy obliczeń neuronowych.
11. Projektowanie gier komputerowych.

12. Metody sztucznej inteligencji.

13. Sieci komputerowe.

Angażuję się również w opracowywanie materiałów dydaktycznych dla studentów, uczniów oraz nauczycieli:

- Przygotowałem recenzowaną monografię jako materiał dydaktyczny dla studentów, z zakresu systemów komputerowych w logistyce: **Mariusz Topolski, *Systemy komputerowe w logistyce*, Oficyna Wydawnicza NDiO, ISBN 978-83-89908-13-1, (2011).**
- Przygotowałem recenzowaną monografię jako materiał dydaktyczny dla studentów, z zakresu utrzymania i bezpieczeństwa systemów informacyjnych: **Mariusz Topolski, *Utrzymanie i bezpieczeństwo systemów informacyjnych*, Oficyna Wydawnicza NDiO, ISBN 978-83-89908-06-3, (2010).**
- W 2021 roku opracowałem recenzowany skrypt dla studentów, z tematu: **Algorytmy optymalizacji w VBA (ang. Visual Basic for Applications)**
- W 2021 roku opracowałem recenzowany skrypt dla studentów, z tematu: **Business Intelligence w łańcuchu dostaw**
- W 2016 roku opracowałem recenzowany skrypt dla studentów z tematu: **Metody optymalizacyjne w łańcuchu dostaw**

Angażuję się w prowadzenie szkoleń i konsultacji dla pracowników naukowych z zakresu zastosowania metaheurystyk oraz zaawansowanych metod symulacyjnych w narzędziu Anylogic.

Aktywnie uczestniczę także w opracowywaniu kart przedmiotów, będąc opiekunem trzech kursów:

- Głębokie sieci neuronowe,
- Metody techniki systemów w medycynie,
- Projektowanie telemedycznych systemów internetowych i mobilnych,

Ponadto, od trzech lat pełnię rolę sekretarza komisji dyplomowej specjalności *Systemy i Sieci Komputerowe*. W okresie od kiedy posiadam stopień doktora nauk technicznych byłem także promotorem ponad 140 prac magisterskich i ponad 160 prac inżynierskich.

Na Politechnice Wrocławskiej, na kierunku Informatyka Techniczna i specjalizacji Systemy informatyki w medycynie pełnię funkcje doradcy toku studiów. Do moich obowiązków należy:

- opiniowanie Indywidualnych Programów Studiów,
- określanie Indywidualnych Programów Studiów uczestnikom wymiany międzynarodowej (program Erasmus),

- wskazywanie zamienników w programie studiów dla studentów danego kierunku studiów (w związku z planowanym pobytem na uczelni zagranicznej, wniosków studentów związanych z przeniesieniem, wznowieniem studiów, po powrocie z urlopu).

W roku 2021 otrzymałem wyróżnienie i nagrodę Rektora Politechniki Wrocławskiej za działalność na rzecz Uczelni. W roku 2022 otrzymałem nagrodę od Rektora Politechniki Wrocławskiej za wybitne osiągnięcia naukowe.

Aktywnie udzielam się również w organizacji konferencji i warsztatów naukowych, wśród których chciałbym wymienić:

- Organizację konferencji *The 10th IEEE International Conference on Data Science and Advanced Analytics (DSAA) DisA 2023*, 9-13.10.2023, Thessaloniki, Grecja. Zasięg międzynarodowy <http://http://disa.kssk.pwr.edu.pl/>
- Organizację konferencji *The 13 International Conference on Computer Recognition Systems CORES*, 29.6.2023 r., Wrocław - zdalnie, Polska. Zasięg międzynarodowy <http://cores.pwr.wroc.pl>
- Organizację konferencji *Polskie Porozumienie na rzecz Rozwoju Sztucznej Inteligencji PP-RAI'2019*, 16-18.10.2019., Wrocław, Polska. Zasięg krajowy <http://pp-rai.pwr.edu.pl/>
- Organizację konferencji *The 12 International Conference on Computer Recognition Systems CORES*, 28-30.6.2021 r., Polanica-Zdrój, Polska. Zasięg międzynarodowy <http://cores.pwr.wroc.pl>
- Organizację konferencji *The 11 International Conference on Computer Recognition Systems CORES*, 20-22.5.2019 r., Polanica-Zdrój, Polska. Zasięg międzynarodowy <http://cores.pwr.wroc.pl>
- Organizację konferencji *The 11 International Conference on Computer Recognition Systems CORES*, 22-24.5.2017 r., Polanica-Zdrój, Polska. Zasięg międzynarodowy <http://cores.pwr.wroc.pl>
- Komitet naukowy konferencji *I Konferencja Naukowa Wrocławskiej Wyższej Szkoły Informatyki Stosowanej we Wrocławiu, "Współczesne zastosowania metod informatycznych" WWSIS*, 3-4.3.2012 r., Polanica, Polska. Zasięg krajowy <http://cores.pwr.wroc.pl>
- Organizację konferencji *IV KONFERENCJA NAUKOWA LOGISTYKA 2018 „Logistyka w naukach o zarządzaniu”*, 1-2.12.2022 r., Opole, Polska. Zasięg krajowy. <https://www.konferencjalogistyka.pl/rada-naukowa/>
- Komitet naukowy konferencji *Konferencja Naukowo-Techniczna, „Rola transportu regionalnego w funkcjonowaniu III Paneuropejskiego Korytarza Transportowego”*, 26-27.11.2009 r., Wrocław, Polska. Zasięg krajowy. <https://www.konferencjalogistyka.pl/rada-naukowa/>
- Pomoc przy konferencji *The 6 International Conference on Computer Recognition Systems CORES*, 25-28.5.2009 r., Jelenia Góra, Polska. Zasięg międzynarodowy <http://cores.pwr.wroc.pl>

Jestem recenzentem dwóch rozpraw doktorskich:

- Raquel Redondo Guevara, "Hybridization of machine learning for advanced manufacturing", Supervised by: Emilio Santiago Corchado Rodríguez (Director), Álvaro Herrero Cosío (Director), Javier Sedano Franco (Director), Universidad de Salamanca 2020
- Magda Friedjungová, "Advanced methods of asymmetric heterogeneous transfer learning", Supervised by: doc. RNDr. Ing. Marcel Jiřina, Ph.D., Czech Technical University in Prague, Faculty of Information Technology 2020

Aktualnie jestem w trakcie recenzowania trzeciej rozprawy doktorskiej "Machine learning based passive image forgery classification using hybrid-handcrafted and CNN-based features", której promotorem jest prof. Vineet Khandelwal, z uczelni Jaypee Institute of Information Technology, Dept. of Electronics and Communication Engineering, Noida(U.P), India.

8 Dane bibliometryczne

8.1 Prace autorstwa Mariusza Topolskiego spoza monografii

Zbiór prac opublikowanych po uzyskaniu stopnia naukowego doktora liczy 76 pozycji z czasopism krajowych i zagranicznych oraz materiałów konferencyjnych.

- 1.) [IF: 4,614, PKT: 140] Nowicka-Sauer Katarzyna, Jarmoszewicz Krzysztof, Molisz Andrzej, Sobczak Krzysztof, Sauer Marta, Topolski Mariusz, *Applying Machine Learning to Construct a Model of Risk of Depression in Patients Following Cardiac Surgery with the Use of the SF-12 Survey*, International Journal of Environmental Research and Public Health, vol. 20, number 6, 2023.
- 2.) [IF: 4,042, PKT: 140] Hoła Bożena, Topolski Mariusz, Szer Iwona, Szer Jacek, Blazik-Borowa Ewa, *Prediction model of seasonality in the construction industry based on the accidentality phenomenon*, Archives of Civil and Mechanical Engineering, vol. 22, p. 1-13, 2022.
- 3.) [IF: 1,056, PKT: 40] Topolski Mariusz, Beza Marcin, *Modification of the Principal Component Analysis Method Based on Feature Rotation by Class Centroids*, JOURNAL OF UNIVERSAL COMPUTER SCIENCE, vol. 3, number 28, p. 227-248, 2022.
- 4.) [IF: 2,838, PKT: 100] Białko Marta, Hoła Bożena, Topolski Mariusz, Woźniak Zuzanna, *Waste Reduction Methods Used in Construction Companies with Regards to Selected Building Products*, Applied Sciences, vol. 12, number 11, 2022.
- 5.) [IF: 3,282, PKT: 100] Jarmoszewicz Krzysztof, Topolski Mariusz, Hajduk Adam, Banaszek Dorota, Nowicka-Sauer Katarzyna, *Prevalence and Predictors of Suicidal Ideation in Patients Following Cardiac Surgery*, World Journal of Surgery, vol. 46, number 8, p. 1997-2004, 2022.

- 6.) [IF: 4,144, PKT: 100] Kalenik Anna Maria, Topolski Mariusz, Górnik Justyna, Wolańczyk Tomasz, *The impact of the COVID-19 pandemic on the mental health of children with psychiatric diagnoses—multidimensional CCPCA Model*, BMC psychiatry, vol. 22, number 1, 2022.
- 7.) [IF: –, PKT: 140] Grzyb Joanna, Topolski Mariusz, Woźniak Michał, *Application of multi-objective optimization to feature selection for a difficult data classification task*, International Conference on Computational Science, p. 81-94, 2021.
- 8.) [IF: 3,847, PKT: 100] Topolski Mariusz, *Application of Feature Extraction Methods for Chemical Risk Classification in the Pharmaceutical Industry*, Sensors, vol. 21, number 17, 2021.
- 9.) [IF: 6,456, PKT: 100] Wojkowska-Mach Jadwiga, Brudło Michał, Topolski Mariusz, Bochenek Tomasz, Jachowicz Estera, Siewierska Małgorzata, Różańska Anna, *Antibiotic consumption in long-term care facilities in Poland and other European countries in 2017*, Antimicrobial Resistance Infection Control, vol. 10, number 1, p. 1-7, 2021.
- 10.) [IF: 3,752, PKT: 100] Topolski Mariusz, Kozal Jędrzej, *Novel feature extraction method for signal analysis based on independent component analysis and wavelet transform*, Plos one, vol. 16, number 12, 2021.
- 11.) [IF: –, PKT: 140] Topolski Mariusz, *Application of the stochastic gradient method in the construction of the main components of PCA in the task diagnosis of multiple sclerosis in children*, Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part IV 20, p. 35-44, 2020.
- 12.) [IF: 1,139, PKT: 40] Topolski Mariusz, *The Modified Principal Component Analysis Feature Extraction Method for the Task of Diagnosing Chronic Lymphocytic Leukemia Type B-CLL.*, J. Univers. Comput. Sci., vol. 26, number 6, p. 734-746, 2020.
- 13.) [IF: –, PKT: 20] Topolski Mariusz, *Algorithm of multidimensional analysis of main features of PCA with blurry observation of facility features detection of carcinoma cells multiple myeloma*, Progress in Computer Recognition Systems 11, p. 286-294, 2020.
- 14.) [IF: 4,639, PKT: 70] Jachowicz Estera, Różańska Anna, Pobiega Monika, Topolski Mariusz, Wójkowska-Mach Jadwiga, *Consumption of Antibiotics and Epidemiology of Clostridioides difficile in the European Union in 2016—Opportunity for Practical Application of Aggregate ECDC Data*, AntibiotICS, vol. 9, number 3, 2020.
- 15.) [IF: –, PKT: 10] Topolski Mariusz, Topolska Katarzyna, *Algorithm for constructing a classifier team using a modified PCA (Principal Component Analysis) in the task of diagnosis of acute lymphocytic leukaemia type B-CLL*, Hybrid Artificial Intelligent Systems: 14th International Conference, HAIS 2019, León, Spain, September 4–6, Proceedings 14, p. 614-624, 2019.

- 16.) Topolska Katarzyna, Topolski Mariusz, *Wspomaganie zarządzania inteligentnym transportem wewnętrznym za pomocą wielowymiarowego modelu korespondencji MCA*, *Przedsiębiorczość i Zarządzanie*, vol. 20, number 7, cz. 2 *Logistyka w naukach o zarządzaniu*, p. 379-389, 2019.
- 17.) Topolska Katarzyna, Topolski Mariusz, *Zastosowanie inteligentnego sterowania ruchem drogowym*, *Przedsiębiorczość i Zarządzanie*, vol. 20, number 7, cz. 2 *Logistyka w naukach o zarządzaniu*, p. 369-378, 2019.
- 18.) Topolski Mariusz, *Application of smart road intelligent traffic control as a determinant of reduction of Co2 to the air*, *Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach*, number 10, p. 105-110, 2018.
- 19.) Topolski Mariusz, *Multidimensional MCA correspondence model supporting intelligent transport management*, *Archives of Transport System Telematics*, vol. 11, 2018.
- 20.) Topolski Mariusz, *Architektura systemu w planowaniu transportu w aspekcie błędu prognoz*, *Przedsiębiorczość i Zarządzanie*, vol. 19, number 12, cz. 2 *Współczesne problemy zarządzania publicznego*, p. 331-338, 2018.
- 21.) Topolski Mariusz, *A model for controlling the flow of materials in the production process*, *AUTOBUSY–Technika, Eksploatacja, Systemy Transportowe*, vol. 19, number 6, p. 1141-1144, 2018.
- 22.) Topolski Mariusz, *Zarządzanie inteligentnym transportem wewnętrznym poprzez komputerowe algorytmy probabilistyczne*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 19, 2018.
- 23.) Topolska Katarzyna, Topolski Mariusz, *Rola instrumentów zarządzania przewozami pasażerskimi w transporcie miejskim*, *Przedsiębiorczość i Zarządzanie*, vol. 19, number 11, cz. 3 *Logistyka w naukach o zarządzaniu. Część 1*, p. 363-374, 2018.
- 24.) Topolski Mariusz, *Management of intelligent internal transport through computer-based probabilistic algorithms*, *AUTOBUSY–Technika, Eksploatacja, Systemy Transportowe*, vol. 19, number 6, p. 968-972, 2018.
- 25.) Topolski Mariusz, *Wielowymiarowy model korespondencji wspomagający zarządzanie inteligentnym transportem w służbie zdrowia*, *Przedsiębiorczość i Zarządzanie*, vol. 19, number 12, cz. 2 *Współczesne problemy zarządzania publicznego*, p. 115-128, 2018.
- 26.) Tomala Borys, Topolski Mariusz, *Wirtualna elektrownia. Przyszłość zarządzania popytem i podażą energii*, *Energetyka Ciepła i Zawodowa*, 2018.
- 27.) Topolski Mariusz, *Model sterowania przepływem materiałów w procesie produkcyjnym*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 19, number 6, 2018.
- 28.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *Planning model of production process in a manufacturing enterprise*, *Research in Logistics Production*, vol. 7, number 5, p. 461-475, 2017.

- 29.) Topolski Mariusz, *Komputerowo wspomagany łańcuch logistyczny z zastosowaniem miękkich metod matematycznych*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, 2017.
- 30.) Topolski Mariusz, *Randomized algorithms supporting management of an intelligent internal transport system*, Archives of Transport System Telematics, vol. 10, number 3, p. 46-50, 2017.
- 31.) Topolski Mariusz, *Zastosowanie miękkich metod obliczeniowych w komputerowo wspomaganym zarządzaniu magazynem. Cz. 1*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number 6, 2017.
- 32.) Topolski Mariusz, Sitna Klaudia, *Analiza perspektyw rozwoju CSR w Polsce*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number 12, 2017.
- 33.) Mariusz Topolski, Katarzyna Topolska, *Computer recognition of data structures using cluster analysis and the theory of mathematical records*, Journal of KONBiN, vol. 41, number 1, p. 5-20, 2017.
- 34.) Topolski Mariusz, *Zastosowanie metod heurystycznych do celu identyfikacji ograniczeń procesów logistycznych*, Przedsiębiorczość i Zarządzanie, vol. 18, number 8, p. 249-258, 2017.
- 35.) Topolski Mariusz, *Zastosowanie metod heurystycznych w zadaniu optymalizacji procesów magazynowych. Cz. 3*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, 2017.
- 36.) Topolski M, *The use of Markov chains of kth row in computer management of transport in a supply chain*, Archives of Transport System Telematics, vol. 10, 2017.
- 37.) [IF: -, PKT: 20] Topolski M, *Computer algorithms probabilistic supporting intelligent transportation management of internal*, Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach, number 9, p. 93-100, 2017.
- 38.) Topolski Mariusz, *Zastosowanie zbiorów rozmytych z analizą skupień w zadaniu klasyfikacji towarów. Cz. 2*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number 6, 2017.
- 39.) Topolski Mariusz, *Komputerowe wspomaganie zarządzanie jakością transportu*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number, 2017.
- 40.) Topolska Katarzyna, Kierzek Diana, Topolski Mariusz, *Negatywny wpływ subkultur na zarządzanie organizacją*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 17, number 12, p. 1876-1879, 2016.
- 41.) Katarzyna Topolska, Mariusz Topolski, Michał Janicki, Czesław Kolanek *Using Computer Methods to Identify the Factors Affecting the Management of an Urban Parking Lot*, Business Challenges in the Changing Economic Landscape-Vol. 2: Proceedings of the 14th Eurasia Business and Economics Society Conference, p. 165-174, 2016.

- 42.) Topolski Mariusz, *Planowanie optymalnej trasy przejazdu transportu samochodowego z wykorzystaniem miękkich metod obliczeniowych*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, number 6, p. 1174-1179, 2016.
- 43.) Topolski Mariusz, Bujak Andrzej, *The use of telematics tools in controlling the flow of materials in a production process*, *Archives of Transport System Telematics*, vol. 9, number 4, p. 43-46, 2016.
- 44.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *The use of cluster analysis and the theory of mathematical records in the process of planning the production-warehouse flow*, *Research in Logistics Production*, vol. 6, number 3, p. 259-267, 2016.
- 45.) Topolska Katarzyna, Topolski Mariusz, *Model systemu wspomagania decyzji z wykorzystaniem koncepcji zarządzania wiedzą w przedsiębiorstwie produkcyjnym*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, 2016.
- 46.) Topolski Mariusz, Topolska Katarzyna, Janicki Michał, Kolanek Czesław, *The management of urban parking LOTS*, *Archives of Transport System Telematics*, vol. 9, 2016.
- 47.) Topolska Katarzyna, Topolski Mariusz, *Model systemu informacyjnego w procesie weryfikacji lojalności klientów przedsiębiorstwa produkcyjnego*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, 2016.
- 48.) Topolski Mariusz, *Application of a telematics system to the improvement of transport processes in intelligent high bay warehouses*, *Archives of Transport System Telematics*, vol. 9, 2016.
- 49.) Topolski Mariusz, *Komputerowy model planowania zapasów bezpieczeństwa z wykorzystaniem miękkich metod obliczeniowych*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, number 6, p. 1802-1806, 2016.
- 50.) Topolska Katarzyna, Topolski Mariusz, *Telematics system architecture to manage the internal transport*, *Archives of Transport System Telematics*, vol. 9, number 3, p. 46-49, 2016.
- 51.) Topolska Katarzyna, Topolski Mariusz, *Rozwój przedsiębiorstw transportowo-spedycyjnych w warunkach międzynarodowych*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, 2016.
- 52.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *Managing light-signalling devices in view of city eco-logistics*, *Research in Logistics Production*, vol. 5, 2015.
- 53.) Bujak Andrzej, Topolski Mariusz, Miler Ryszard, *Telematics as a tool of efficient supply chain management*, *Archives of Transport System Telematics*, vol. 8, 2015.
- 54.) Topolska Katarzyna, Topolski Mariusz, *Gospodarka odpadami medycznymi w funkcjonowaniu placówki zdrowia*, *TTS Technika Transportu Szybowego*, vol. 22, 2015.

- 55.) Bujak Andrzej, Topolski Mariusz, *Concept of a Telematics System Model in Crisis Management*, Tools of Transport Telematics: 15th International Conference on Transport Systems Telematics, TST 2015, Wrocław, Poland, April 15-17, 2015. Selected Papers 15, p. 180-187, 2015.
- 56.) Bujak Andrzej, Topolski Mariusz, Topolska Katarzyna, *Rola informatycznych systemów telematycznych w zarządzaniu współczesnym łańcuchem dostaw*, Logistyka, number 4, p. 2681-2688, 2015.
- 57.) Topolski Mariusz, Topolska Katarzyna, Kobyłt Agata, *Sterowanie przepływem materiałów w procesie produkcyjnym*, Logistyka, number 2, p. 765-772, 2015.
- 58.) Topolska Katarzyna, Topolski Mariusz, *Capabilities for development an ITS system in the city of Wrocław*, Archives of Transport System Telematics, vol. 8, number 4, p. 29-32, 2015.
- 59.) Topolski Mariusz, Topolska Katarzyna, *Komputerowe algorytmy wspomagające planowanie sprzedaży w przedsiębiorstwach produkcyjnych*, Logistyka, 2014.
- 60.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *Zastosowanie inteligentnych systemów do analizy zagrożeń komunikacyjnych*, Logistyka, number 5, p. 1760-1765, 2014.
- 61.) Topolska Katarzyna, Bujak Andrzej, Topolski Mariusz, *Badanie zapotrzebowania na usługi kolejowego transportu pasażerskiego na przykładzie miasta Wrocławia*, Studia Miejskie, vol. 15, p. 63-78, 2014.
- 62.) Topolski Mariusz, , *Komputerowa analiza błędów prognoz w transporcie*, Logistyka, vol. , number 6,, 2014.
- 63.) Topolska Katarzyna, Topolski Mariusz, *Use of mergers and genetic fuzzy classifiers in the task of controlling traffic lights*, Logistyka, number 6, 2014.
- 64.) Topolska Katarzyna, Topolski Mariusz, *Optymalizacja logistycznej obsługi klienta z zastosowaniem systemu symulacyjnego Anylogic*, Logistyka, 2014.
- 65.) Topolski Mariusz, *Komputerowy model rozmytego klasyfikatora Bayesa w ocenie rentowności produkcji*, Logistyka, number 6, 2014.
- 66.) Bujak Andrzej, Kobyłt Agata, Topolska Katarzyna, Topolski Mariusz, *Inteligentne rozwiązanie sterowania ruchem drogowym w logistyce miasta*, Logistyka, number 5, p. 1755-1759, 2014.
- 67.) Janicki Michał, Kolanek Czesław, Topolska Katarzyna, Topolski Mariusz, *Logistics of urban car parks*, Logistyka, number 6, 2014.
- 68.) Andrzej Bujak, Kobyłt Agata, Topolska Katarzyna, Topolski Mariusz, *Intelligent traffic control solutions for logistics center*, Logistyka, vol. 5, p. 1755-1759, 2014.

- 69.) Topolska Katarzyna, Topolski Mariusz, Błachut Bartłomiej, Haber Marcin, Piekarczyk Aleksander, *Zastosowanie fuzji klasyfikatorów rozmytych i genetycznych w zadaniu sterowania sygnalizacją świetlną*, Biuletyn Naukowy Wrocławskiej Wyższej Szkoły Informatyki Stosowanej. Informatyka, vol. 2, p. 30-33, 2012.
- 70.) Katarzyna Topolska, Topolski Mariusz, , *Optymalizacja i niezawodność inteligentnych systemów sterowania ruchem drogowym*, Problemy Eksploatacji Instytut Technologii Eksploatacji. Journal of KONBiN, vol. 20, numer 4, p. 133-140, 2011.
- 71.) Topolski Topolska Katarzyna Mariusz, *Inteligentne systemy wieloagentowe w analizie zagrożeń wypadków komunikacyjnych*, Logistyka Współczesne wyzwania. Wydawnictwo PWSZ, p. 175-181, 2011.
- 72.) [IF: -, PKT: 20] Topolska Katarzyna, Topolski Mariusz, *Zastosowanie zbiorów rozmytych w planowaniu zapasów w łańcuchu dostaw*, Logistics and Transport, vol. 8, number 1, p. 169-178, 2009.
- 73.) Kurzynski Marek, Wolczowski Andrzej, Topolski Mariusz, *Control of Hand Bioprosthesis Via Sequential Recognition of Patient's Intent Using Combination of Fuzzy Sets and Dempster-Shafer Theory*, Information Technologies in Biomedicine, p. 421-428, 2008.
- 74.) [IF: -, PKT: 20] Topolski Mariusz, Topolska Katarzyna, *Fuzja systemu ekspertowego z technologią RFID na przykładzie wybranego fragmentu łańcucha logistycznego*, Logistics and Transport, vol. 6, p. 69-78, 2008.
- 75.) [IF: -, PKT: 20] Topolski Mariusz, Topolska Katarzyna, *Warunki eksploatacji pojazdów samochodowych-analiza czynności wykonywanych podczas przeglądów samochodowych*, Logistics and Transport, vol. 7, p. 99-106, 2008.
- 76.) [IF: -, PKT: 20] Topolski Mariusz, Topolska Katarzyna, *Inteligentne systemy logistyczne w produkcji samochodów i części zamiennych*, Logistics and Transport, vol. 7, number 2, p. 91-98, 2008.

8.2 Prace autorstwa Mariusza Topolskiego przed uzyskaniem stopnia naukowego doktora inżyniera

Zbiór prac opublikowanych przed uzyskanie stopnia naukowego doktora liczy 13 pozycji z czasopism krajowych i zagranicznych oraz materiałów konferencyjnych.

- 1.) Topolska Katarzyna, Walkowiak Wojciech, Topolski Mariusz, *Method of analysis of measurement result cohesion with the theory of mathematical record using and concentrations analysis*, Journal of KONES, vol. 14, number 2, p. 521-528, 2007.
- 2.) Katarzyna Topolska, Walkowiak Wojciech, Mariusz Topolski, *Metoda analizy spójności wyników pomiarów z wykorzystaniem teorii ewidencji matematycznej i analizy skupień*, Journal of KONES, vol. 14, number 2, p. 521-528, 2007.

- 3.) [IF: -, PKT: 20] Topolski Mariusz, *Rozmyte sterowanie sygnalizacją świetlną w newralgicznych punktach miejskich*, Logistics and Transport, number 2, p. 101-110, 2006.
- 4.) [IF: -, PKT: 20] Topolska Katarzyna, Topolski Mariusz, *Struktura organizacyjna i system informatyczny w łańcuchu logistycznym sektora motoryzacyjnego*, Logistics and Transport, vol. 3, p. 75-82, 2006.
- 5.) [IF: -, PKT: 20] Topolska Katarzyna, Topolski Mariusz, *System logistyczny przedsiębiorstwa i jego struktura przestrzenna*, Logistics and Transport, vol. 3, p. 67-73, 2006.
- 6.) Topolska Katarzyna, Walkowiak Wojciech, Topolski Mariusz, *The modeling of stability of mechanical component based on Dempster-Shafer theory*, Journal of KONES Powertrain and Transport, vol. 13, number 4, 2006.
- 7.) Topolski Mariusz, , *Rola logiki rozmytej i teorii ewidencji matematycznej w systemach logistycznych*, Zeszyty Naukowe. Logistyka i Transport/Międzynarodowa Wyższa Szkoła Logistyki i Transportu we Wrocławiu, vol. 2, number 1, p. 87-99, 2006.
- 8.) Topolska Katarzyna, Walkowiak Wojciech, Topolski Mariusz, *The validation methods of durability of ceramic coating under laboratory conditions*, Journal of KONES Powertrain and Transport, vol. 13, number 4, 2006.
- 9.) Mariusz Topolski, *Dynamiczna kontrola trwałości rozrusznika serca*, W: Niekonwencjonalne metody oceny trwałości i niezawodności. XXXIV Zimowa Szkoła Niezawodności, Szczyrk,[8-14 stycznia], p. 342-351, 2006.
- 10.) Topolski Mariusz, *The Part of Fuzzy Systems Assisting the Decision in Diagnostics of Fuel Engine Subassemblies Defects*, Journal of KONES Internal Combustion Engines, vol. 12, p. 3-4, 2005.
- 11.) Topolska Katarzyna, Topolski Mariusz, *The Dempster-Shafer model of mechanical object durability in laboratory conditions*, Journal of KONES Internal Combustion Engines, vol. 12, p. 3-4, 2005.
- 12.) Topolski Mariusz, *The fuzzy-probabilistic sequent system for controlling the spark ignition in fuel engine*, Journal of KONES Internal Combustion Engines, vol. 12, p. 3-4, 2005.
- 13.) [IF: -, PKT: 40] Topolska Katarzyna, Topolski Mariusz, , *Optimization and reliability of intelligent traffic lights control systems/Optymalizacja i niezawodność inteligentnych systemów sterowania ruchem drogowym*, Journal of KONBiN, vol. 20, number 1, p. 133-140, 2005.



PODPIS ZAUFANY

MARIUSZ
TOPOLSKI
28.08.2023 06:21:01 | GMT+2
Dokument podpisany elektronicznie
podpisem zaufanym

.....
(podpis wnioskodawcy)



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«НАЦІОНАЛЬНИЙ ГІРНИЧИЙ УНІВЕРСИТЕТ»

просп. К. Маркса, 19, Дніпропетровськ, 49005, Україна
Телефон: +38 (056) 744-62-19, (0562) 46-40-62; факс: +38 (056) 744-62-11;
e-mail: rector@nmu.org.ua, nmu@nmu.org.ua; http://www.nmu.org.ua

№ _____

на № _____

Mariusz Topolski
ul. Kasztanowa 22
55-002 Dobrzykowice

15th November 2013

Certificate of completing an internship

Mr. Dr. Eng. Mariusz Topolski took an internship in the National Mining University in Dnipropietrowsk.

Period of internship placement:

from 1.04.2013 to 15.11.2013

including participation in the International Summer School for Leaders in Crimea.

The place of internship:

The Institute of Economics at the National Mining University in Dnipropietrowsk.

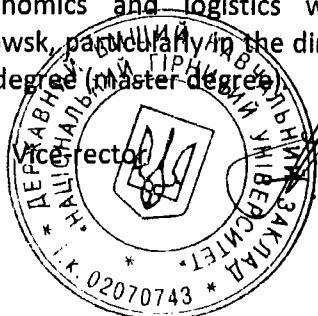
Internship coordinator:

Director of the Institute of Economics professor Alexander Ivanovich Sharov.

Tasks performed by the person, who took internship:

1. Experience exchange in students education.
2. Participation in the design of students training program.
3. Analysis methods of lecture, exercises and laboratories.
4. Observation and analysis of laboratory equipment.
5. Leading didactic lessons.
6. Cooperation with students scientific club.
7. Participation in the substantive proposals for scientific research.
8. Establishing cooperation under the international researches and projects.
9. Exchange of scientific-professional experiences.
10. Conduct activities for the International Summer School Leaders.
11. Analysis of the possibilities of computer algorithms in the task of planning and forecasting logistic systems.
12. Exchange of experience in the scope of computer recognition methods and decision support of integrated management system.

As part of the internship was carried out more than 90 hours of lectures on Polish computer science, economics and logistics with students of National Mining University in Dnipropietrowsk, particularly in the direction of the classroom management software, line: logistics at II degree (master degree).



O.M.Shashenko

Barcelona el día 17. de septiembre de 2014

(lugar, fecha)

Documento de confirmación de participación de las personas:

dr inż. Mariusz Topolski

(nombre y apellido de la persona/personas)

en una formación al extranjero, del personal didáctico, en la entidad:

FUNDACIÓN ICIL

(nombre de la entidad, dirección)

C/Conxita Supervía, 11, bajos

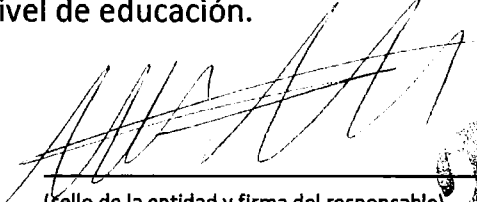
08028 Barcelona

En el periodo del 3.09.2014_ hasta _17.09.2014_

Campo de formación:

1. Intercambio de las experiencias en la educación de los estudiantes, de los métodos didácticos aplicados durante cursos, ejercicios y laboratorios
2. Visitas y análisis del equipamiento de los laboratorios
3. Discusión sobre cooperación potencial en los proyectos de investigación / proyectos internacionales.

El objetivo de la formación fue profundizar conocimientos y intercambio de las experiencias sobre los métodos aplicados en las escuelas superiores para garantizar un más alto nivel de educación.


(Sello de la entidad y firma del responsable)

(*) el documento tiene que ser presentado en el idioma oficial de la entidad

Desde 1950 impulsando la Logística
C/Conxita Supervía, 11, bajos
08028 Barcelona
Tel: +34 93 12 006 100
Fax: +34 93 12 006 101
www.icil.org

Gdańsk, 4.11.2022

ZAŚWIADCZENIE

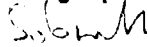
Pan dr inż. Mariusz Topolski, adiunkt w Katedrze Systemów i Sieci Komputerowych Wydziału Informatyki i Telekomunikacji Politechniki Wrocławskiej w ramach współpracy z zespołami psychologów, socjologów medycyny i lekarzy współuczestniczył w badaniach prowadzonych w Zakładzie Medycyny Rodzinnej oraz Zakładzie Socjologii Medycyny i Patologii Społecznej, Gdańskiego Uniwersytetu Medycznego stosując m.in. różne techniki wielowymiarowe i autorskie modele ekstrakcji cech. Pozwoliło to uzyskać różne ważne przestrzenie cech dla rzeczywistych problemów medycznych.

Dr inż. Mariusz Topolski uczestniczył lub wspomagał prace statystyczne w ramach takich projektów jak:

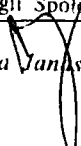
1. *Zastosowanie metod uczenia maszynowego w predykcji poziomu wypalenia zawodowego, prężności oraz poczucia własnej skuteczności u psychologów klinicznych i lekarzy psychiatrów (Patronat Rzecznika Praw Pacjenta).*
2. *Samoocena przygotowania studentów ostatniego roku medycyny do przekazywania niekorzystnych wiadomości medycznych (opublikowano 2 prace w International Journal of Environmental Research and Public Health, łączny IF = 9,228).*
3. *Factors associated with high preoperative anxiety: results from cluster analysis. Publikacja w czasopiśmie World Journal of Surgery, 2020; IF=3.352.*
4. *Predictors of patient-reported health following cardiac surgery. Publikacja w czasopiśmie Journal of Cardiovascular Surgery, 2021; IF=1.415.*
5. *Prevalence and predictors of suicidal ideation in patients following cardiac surgery. Publikacja w czasopiśmie World Journal of Surgery, 2022; IF=3.352.*
6. *Applying machine learning to identify risk factors for depression in patients following cardiac surgery. Artykuł przygotowany do publikacji.*
7. *Determinants of anxiety related to anesthesia among patients scheduled for cardiac surgery. Artykuł przygotowany do publikacji.*
8. *Illness perception and perceived benefits of illness among persons with type 1 diabetes. Artykuł przyjęty do publikacji w czasopiśmie Health Psychology Report.*
9. *Badanie wpływu powikłań w gojeniu się ran na jakość życia u chorych po operacjach serca. S. Batkiewicz, rozprawa na stopień doktora nauk medycznych, GUMed, 2019.*

Pan dr inż. Topolski współpracując przy wymienionych wyżej projektach wykazał się wysokim zaangażowaniem, gotowością do działania, a Jego merytoryczny wkład w powstałe prace niewątpliwie podniósł ich wartość przyczynił się do ich opublikowania.

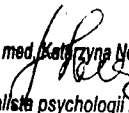
Gdański Uniwersytet Medyczny
Zakład Socjologii Medycyny
i Patologii Społecznej


dr n. med. Krzysztof Sobczak

p.o. KIEROWNIKA
Zakładu Socjologii Medycyny
i Patologii Społecznej


dr Agata Vanaszczyk

Gdański Uniwersytet Medyczny
Zakład Medycyny Rodzinnej
Katedra Medycyny Rodzinnej
ul. Dębinki 2
80-211 Gdańsk
tel. 58 349 15 75


dr hab. n. med. Katarzyna Nowicka-Sauer
specjalista psychologii klinicznej

NAZWA JEDNOSTKI WIELKIMI LITERAMI

ul. Adres Jednostki, 80-123 Gdańsk | 58 122 45 67 | nazwajednostki@gumed.edu.pl

REFERENCJE

Pan dr inż. Mariusz Topolski współpracował z Wrocławską Wyższą Szkołą Informatyki Stosowanej od 2009 roku. W latach 2013/2014 został zatrudniony na umowę o pracę i pełnił funkcje Prorektora i Rektora.

Do obowiązków Pana Rektora dr inż. Mariusza Topolskiego należało m.in. :

- reprezentowanie uczelni w Konferencji Rektorów Akademickich Szkół Polskich oraz innych gremiach, których uczestnictwo jest dla spraw Uczelni pożądane,
- kształtowanie polityki kadrowej oraz nadzór nad sposobami realizacji obowiązków pracowniczych, tak nauczycieli akademickich, jak i pracowników niebędących nauczycielami akademickimi,
- zawieranie i rozwiązywanie umów o pracę, powoływanie i odwoływanie osób na kierownicze stanowiska w jednostkach organizacyjnych Uczelni (katedry, zakłady, studia, biblioteka, administracja) oraz ustalanie ich obowiązków, kompetencji, zasad i zakresu odpowiedzialności, a także tworzenie i znoszenie jednostek administracyjnych w Uczelni na zasadach określonych w statucie
- kierowanie działalnością naukowo – badawczą
- sprawowanie nadzoru nad działalnością dydaktyczną, naukową, wychowawczą oraz administracyjną Uczelni
- nadzór nad międzyuczelnianą współpracą krajową i zagraniczną
- podejmowanie decyzji w formie zarządzeń, ustalania regulaminów, rozstrzygnięć w sprawach uregulowanych w Statucie, należących do kompetencji Rektora i niezastrzeżonych do kompetencji innych organów
- nadzór nad przestrzeganiem obowiązującego w uczelni Prawa o szkolnictwie wyższym oraz innych przepisów, które mają zastosowanie w każdej dziedzinie działalności Uczelni.
- nadzór nad przebiegiem procesu dydaktycznego
- kierowanie rekrutacją na studia,
- opieka merytoryczna nad projektami dydaktycznymi,
- współpraca z samorządem studentów,
- koordynowanie współpracy z przedsiębiorstwami w zakresie tworzenia nowoczesnych laboratoriów oraz programów studiów odpowiadających potrzebom współczesnej gospodarki,
- koordynowanie prac związanych z tworzeniem Akademickiego Inkubatora Przedsiębiorczości,
- tworzenie kół naukowych,

Pan dr inż. Mariusz Topolski wykonywał swoje zadania systematycznie i obowiązkowo. Jest osobą samodzielną i zaangażowaną zarówno w pracę jak i w życie Uczelni. W czasie pracy wykazał się ogromnym doświadczeniem i głęboką wiedzą w dziedzinie zarządzania Uczelnią wyższą.

Pan dr inż. Mariusz Topolski jest osobą ambitną, ciągle podnoszącą swoje kwalifikacje. W czasie pracy we Wrocławskiej Wyższej Szkole Informatyki Stosowanej wykorzystał wszelkie możliwe szanse na zdobycie dodatkowych kwalifikacji.

Znakomicie sprawdzał się zarówno w sytuacjach wymagających indywidualnego wysiłku, jak i w pracy zespołowej. Ma wysokie umiejętności interpersonalne - łatwo nawiązuje kontakty, a w swoim zespole wytworzył atmosferę wzajemnej pomocy i zrozumienia.

Jestem przekonany, że Pan dr inż. Mariusz Topolski jest osobą, która sprosta wielu ambitnym zadaniom i spełni oczekiwania najbardziej wymagającego pracodawcy.

Pan dr inż. Mariusz Topolski jest bardzo dobrym dydaktykiem. Od początku swojej działalności był jednym z najlepszych, a w latach 2013/2014 został najlepszym z wykładowców w ocenie studentów.

W razie potrzeby z przyjemnością udzielę dokładniejszych informacji na jego temat.



PROREKTOR
Wrocławská Wyższa Szkoła
Informatyki Stosowanej
Hubert Zarzycki
dr inż. Hubert Zarzycki

Krzysztof Horoszkiewicz Pracownia Psychologiczna
62-100 Wągrowiec, Powstańców Wlkp. 20b/2
NIP 766-143-06-98, REGON 634240621
Dystrybutor Systemu Diagnostyki Psychofizjologicznej
SDP System - Współautor Narzędzi Psychologicznych

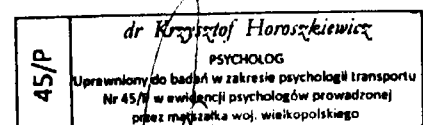
Wągrowiec, dnia 03.12.2022

Rekomendacja

Niniejszym zaświadczam, że opracowana przez dr inż. Mariusza Topolskiego metoda składowych głównych CCPCA została zastosowana do opracowania autorskich kwestionariuszy dla wielu zmiennych.

Została ona wykorzystana w licznych testach i ostatecznie za jej pomocą zostały zbudowane końcowe narzędzia psychologiczne wykorzystywane w praktyce psychologicznej i placówkach medycznych.

Z poważaniem



DEMOTRANS

Dobroszyce, dnia 28 marca 2022

Demotrans Spedycja & Transport
Wrocławska 56, 56-410 Dobroszyce

ZAŚWIADCZENIE

Zaświadczam, że Pan dr inż. Mariusz Topolski w dniach od 9 do 28 marca 2022 roku współpracował z przedsiębiorstwem Demotrans Spedycja & Transport w roli eksperta w zakresie wykorzystania metod uczenia maszynowego i metaheurystyk w planowaniu transportu krajowego i międzynarodowego. W tym okresie przeprowadził szkolenia dla developerów z zakresu integracji opracowanych algorytmów z aktualnie stosowanymi rozwiązaniami w systemie przedsiębiorstwa. Zaproponowane rozwiązania są bardzo istotne w szczególności dla zadania wielu zleceń transportowych, gdzie optymalizacja odgrywa kluczową rolę.

DEMOTRANS

SPEDYCJA & TRANSPORT
Paweł Kaleta, 55-106 Zawonia, ul. Oleśnicka 18
ODDZIAŁ DOBROSZYCE
56-410 Dobroszyce, ul. Wrocławska 56
NIP 915-128-44-48

DEMOTRANS
DYREKTOR GENERALNY
[Signature]
Anatolij Bruch



FTS Artur Grzyb
ul. Południowa 1, 56-400 Oleśnica
NIP: 911-100-21-07
REGON: 930589850



TOM TRANS Tomasz Walczak
Energetyczna 2/2a, 56-400 Oleśnica
NIP 911-126-76-27
REGON 931125933

Oświadczenie wydane na podstawie współpracy w ramach wdrożenia metod optymalizacji planowania harmonogramu transportu, prac w betoniarni oraz audytu stron internetowych.

W dniach 17-21 lipca 2021 roku dr inż. Mariusz Topolski przeprowadził konsultacje z zakresu wdrożenia systemów informatycznych i metod optymalizacyjnych wspomagających planowanie harmonogramowania transportu w zakresie dostarczania materiałów budowlanych do budowy dróg i mostów kołowych oraz integracji tych rozwiązań z pracą betoniarni. W ramach współpracy została przeprowadzona ekspertyza stosowanych rozwiązań IT oraz audyt stron internetowych. Opracowane koncepcje informatyczne przyczyniły się do zwiększenia jakości planowania prac poprzez redukcję niepotrzebnych kosztów. Jednocześnie po przeprowadzonym audycie stron internetowych zostały wdrożone nowe zabezpieczenia danych oraz metody pozycjonowania. Składamy podziękowania Panu Doktorowi za pomoc w opracowanych metodach i ich wdrożeniu.

FIRMA TRANSPORTOWO-SPRZĘTOWA
Grzyb Artur
56-400 Oleśnica, ul. Południowa 1
NIP 911-100-21-07, Regon 930589850
Konto: 85 Oleśnica 69 9584 0008 2001 0008 7027 0001
tel./fax 71-399 32, tel./fax 71-399 10 52
Komy 605-421-668

"TOM TRANS"
Tomasz Walczak
56-400 Oleśnica, ul. Energetyczna 2-2A
NIP: 911-126-76-27, Regon: 931125933

Wrocław, dnia 09.08.2014

Imię i nazwisko: Mariusz Topolski
Stopień naukowy: dr inż.
Adres zamieszkania: ul. Kasztanowa 22
55-002 Dobrzykowice

Zaświadczenie o odbyciu stażu naukowego w przedsiębiorstwie

Zaświadczam, że Pan dr inż. Mariusz Topolski odbył staż naukowy w przedsiębiorstwie

.... **VISTEL Dariusz Ptak**.....

Okres odbywania stażu: od 09.04.2014, do dnia 09.08.2014

Opiekun osoby odbywającej staż:mgr inż. Paweł Grzeszczuk.....

Zadania realizowane przez osobę odbywającą staż:

1. Wymiana doświadczeń dotyczących oprogramowania oraz sprzętu informatycznego
2. Prowadzenie badań naukowych w ramach nowoczesnych technologii informatycznych.
3. Proponowanie innowacyjnych rozwiązań informatycznych
4. Wspomaganie zarządzania projektami informatycznymi w przedsiębiorstwie
5. Konsultowanie programów nauczania w celu ich dostosowania do aktualnych potrzeb rynku pracy
6. Analiza proponowanych rozwiązań informatycznych w oparciu o SWOT i alternatywne narzędzia analityczne
7. Tworzenie opisów proponowanych rozwiązań innowacyjnych na podstawie przeprowadzonych badań i analiz oraz odpowiedzi na zidentyfikowane problemy badawcze.
8. Wspomaganie zarządzania logistycznego przedsiębiorstwem.
9. Wymiana doświadczeń naukowo-badawczych.
10. Współpraca przy propozycjach wdrożeniowych systemów i sieci informatycznych.

.....
Paweł Grzeszczuk
.....
miejsowość, data, podpis opiekuna stażysty

.....
Dariusz Ptak
.....
(podpis osoby reprezentującej firmę)

VISTEL Dariusz Ptak
55-200 Olawa, ul. F. Chopina 34/3
REGON 021152624, NIP 9121576293
www.vistel.pl e-mail: biuro@vistel.pl

.....
(pieczęćka firmy)

Wrocław, dnia: 15.03.2013

Zaświadczenie o odbytym stażu krajowym w przedsiębiorstwie

Niniejszym zaświadczamy, że Pan dr inż. Mariusz Topolski

odbył staż w przedsiębiorstwie/ instytucji

w okresie od 15.11.2012 do 15.03.2013

w wymiarze 450 godzin

W ramach stażu zostały zrealizowane następujące czynności:

1. Badania naukowo – techniczne
2. Konsultacje z zakresu dostosowania programów nauczania do aktualnych potrzeb przedsiębiorstw
3. Udział przy wdrażaniu, tworzeniu i utrzymaniu rozwiązań informatycznych
4. Tworzenie rozwiązań algorytmów sztucznej inteligencji dla potrzeb urzędów przemysłowych
5. Proponowanie i wdrażanie metod optymalizacji rozwiązań sprzętowych i programistycznych dla potrzeb przemysłowych.
6. Analiza błędów rozwiązań informatycznych.
7. Analizy biznesowe i logistyczne.

Ocena merytoryczna stażysty

Pan dr inż. Mariusz Topolski trakcie odbywania stażu wykazał się bardzo dogłębną wiedzę z zakresu wdrażanie, tworzenia i utrzymania systemów informatycznych. Jego wiedza z zakresu metod sztucznej inteligencji pozwoliła nakreślić nowe kierunki programowe dla urzędów technicznych. Wspólnie ze specjalistami działu IT Pan doktor opracował różnice programowe dla studentów studiów informatycznych niezbędne do dostosowania efektów kształcenia do potrzeb rynku pracy. Pan Doktor Mariusz Topolski dał się poznać jako człowiek przedsiębiorczy oraz bardzo dobry metodyk i badacz naukowy.

Artur Pietrzyk
.....
miejsowość, data, podpis opiekuna stażysty

GOLEF TELECOM
SPÓŁKA Z OGRANICZONĄ ODPOWIEDZIALNOŚCIĄ
SPÓŁKA KOMITETOWA
ul. Rakowa 5, Krynitzko, 55-110 Szlania Wola
NIP 9151781454, REGON 021846145

PREZES ZARZĄDU
Robert Borys
.....
(podpis osoby reprezentującej firmę)
Robert Borys

Wykaz osiągnięć naukowych albo artystycznych, stanowiących znaczny wkład w rozwój określonej dyscypliny

1 INFORMACJA O OSIĄGNIĘCIACH NAUKOWYCH ALBO ARTYSTYCZNYCH O KTÓRYCH MOWA W ART. 219 UST. 1. PKT 2 USTAWY

1.1 Monografia naukowa wydana przez wydawnictwo, które w roku opublikowania monografii w ostatecznej formie było ujęte w wykazie sporządzonym zgodnie z przepisami wydanymi na podstawie art. 267 ust. 1 pkt 2 lit. a

W ramach niniejszego wniosku habilitacyjnego prezentowane jest osiągnięcie w formie monografii naukowej wydanej przez wydawnictwo *Akademicka Oficyna Wydawnicza Exit*, które w roku opublikowania monografii w ostatecznej formie jest ujęte w wykazie sporządzonym przez Ministerstwo Nauki i Szkolnictwa Wyższego i posiada Unikatowy Identyfikator Wydawnictwa 14300. Numer monografii ISBN 978-83-7837-139-7

Metody ekstrakcji cech w uczeniu maszynowym. Nowe trendy inżynierii cech

Wszystkie badania zawarte w monografii pochodzą z okresu po uzyskaniu stopnia doktora.

1.2 Cykl powiązanych tematycznie artykułów naukowych

2 INFORMACJA O AKTYWNOŚCI NAUKOWEJ ALBO ARTYSTYCZNEJ

2.1 Wykaz opublikowanych monografii naukowych (z zaznaczeniem pozycji niewymienionych w pkt 1.1).

Jestem autorem dwóch monografii naukowych.

- **Mariusz Topolski**
Utrzymanie i bezpieczeństwo systemów informacyjnych
Oficyna Wydawnicza NDiO, ISBN 978-83-89908-06-3, (2010).
- **Mariusz Topolski**
Systemy komputerowe w logistyce
Oficyna Wydawnicza NDiO, ISBN 978-83-89908-13-1, (2011).

Jestem autorem trzech skryptów dla studentów.

- **Mariusz Topolski**
Algorytmy optymalizacji w VBA (ang. Visual Basic for Applications) - (2021).
- **Mariusz Topolski**
Business Intelligence w łańcuchu dostaw - (2021).
- **Mariusz Topolski**
Metody optymalizacyjne w łańcuchu dostaw - (2016).

2.2 Wykaz opublikowanych rozdziałów w monografiach naukowych.

Jestem autorem pięciu rozdziałów w monografiach:

- **Mariusz Topolski**
Algorithm of multidimensional analysis of main features of PCA with blurry observation of facility features detection of carcinoma cells multiple myeloma.
W: Progress in Computer Recognition Systems / Robert Burduk, Marek Kurzyński, Michał Woźniak (Eds.). Switzerland : Springer, cop. 2020. s. 286-294.
Advances in Intelligent Systems and Computing, ISSN 2194-5357; vol. 977; 2020 r.
- **Katarzyna Topolska, Mariusz Topolski**
Zastosowanie inteligentnego sterowania ruchem drogowym.
W. Przedsiębiorczość i zarządzanie, Logistyka w naukach o zarządzaniu. Andrzej Bujak (red.), Alicja Gębczyńska (red.), Marta Brzozowska (red.). Wydawnictwo SAN, ISSN 2543-8190; 369 s., 2019 r.
- **Mariusz Topolski Katarzyna Topolska**
Wspomaganie zarządzania inteligentnym transportem wewnętrznym za pomocą wielowymiarowego modelu korespondencji MCA.
W. Przedsiębiorczość i zarządzanie, Logistyka w naukach o zarządzaniu. Andrzej Bujak (red.), Alicja Gębczyńska (red.), Marta Brzozowska (red.). Wydawnictwo SAN, ISSN 2543-8190; 379 s., 2019 r.
- **Mariusz Topolski Katarzyna Topolska**
Modelowanie problemów decyzyjnych systemów transportowych.
W. Przedsiębiorczość i zarządzanie, Logistyka w naukach o zarządzaniu. Andrzej Bujak (red.), Katarzyna Topolska (red.), Katarzyna Kolasińska-Morawska (red.). Wydawnictwo SAN, ISSN 2543-8190; 175 s., 2018 r.

• **Mariusz Topolski,**

Zastosowanie metod ekspertowych do optymalizacji procesów magazynowych.

W. Integracja-Dezintegracja-Entropia, Wyzwania dla zarządzania przedsiębiorstwem. Aneta Pachura (red.). Wydawnictwo Politechniki Częstochowskiej, ISBN 978-83-65179-75-3; 177 s., 2016 r.

2.3 *Informacja o członkowskie w redakcjach naukowych monografii*

2.4 *Wykaz opublikowanych artykułów w czasopismach naukowych (z zaznaczeniem pozycji niewymienionych w pkt I.2)*

**Informacje dot. liczby punktów MEiN oraz współczynnika IF
podane na podstawie wskaźników z dnia 28 sierpnia 2023.**

2.4a *Artykuły opublikowane po uzyskaniu stopnia doktora, zgłoszone w punkcie I.2:* Ponieważ, osiągnięciem naukowym jest monografia to w punkcie I.2 nie wykazano osiągnięć naukowych w postaci cyklu.

2.4b *Artykuły opublikowane po uzyskaniu stopnia doktora, niezgłoszone w punkcie I.2.:*

- 1.) [IF: 4,614, PKT: 140] Nowicka-Sauer Katarzyna, Jarmoszewicz Krzysztof, Molisz Andrzej, Sobczak Krzysztof, Sauer Marta, Topolski Mariusz, *Applying Machine Learning to Construct a Model of Risk of Depression in Patients Following Cardiac Surgery with the Use of the SF-12 Survey*, International Journal of Environmental Research and Public Health, vol. 20, number 6, 2023.
- 2.) [IF: 4,042, PKT: 140] Hoła Bożena, Topolski Mariusz, Szer Iwona, Szer Jacek, Blazik-Borowa Ewa, *Prediction model of seasonality in the construction industry based on the accidentality phenomenon*, Archives of Civil and Mechanical Engineering, vol. 22, p. 1-13, 2022.
- 3.) [IF: 1,056, PKT: 40] Topolski Mariusz, Beza Marcin, *Modification of the Principal Component Analysis Method Based on Feature Rotation by Class Centroids*, JOURNAL OF UNIVERSAL COMPUTER SCIENCE, vol. 3, number 28, p. 227-248, 2022.
- 4.) [IF: 2,838, PKT: 100] Białko Marta, Hoła Bożena, Topolski Mariusz, Woźniak Zuzanna, *Waste Reduction Methods Used in Construction Companies with Regards to Selected Building Products*, Applied Sciences, vol. 12, number 11, 2022.
- 5.) [IF: 3,282, PKT: 100] Jarmoszewicz Krzysztof, Topolski Mariusz, Hajduk Adam, Banaszkiewicz Dorota, Nowicka-Sauer Katarzyna, *Prevalence and Predictors of Suicidal Ideation in Patients Following Cardiac Surgery*, World Journal of Surgery, vol. 46, number 8, p. 1997-2004, 2022.
- 6.) [IF: 4,144, PKT: 100] Kalenik Anna Maria, Topolski Mariusz, Górnik Justyna, Wolańczyk Tomasz, *The impact of the COVID-19 pandemic on the mental health of children with psychiatric diagnoses—multidimensional CCPCA Model*, BMC psychiatry, vol. 22, number 1, 2022.
- 7.) [IF: -, PKT: 140] Grzyb Joanna, Topolski Mariusz, Woźniak Michał, *Application of multi-objective optimization to feature selection for a difficult data classification task*, International Conference on Computational Science, p. 81-94, 2021.
- 8.) [IF: 3,847, PKT: 100] Topolski Mariusz, *Application of Feature Extraction Methods for Chemical Risk Classification in the Pharmaceutical Industry*, Sensors, vol. 21, number 17, 2021.

- 9.) [IF: 6,456, PKT: 100] Wojkowska-Mach Jadwiga, Brudło Michał, Topolski Mariusz, Bochenek Tomasz, Jachowicz Estera, Siewierska Małgorzata, Róžańska Anna, *Antibiotic consumption in long-term care facilities in Poland and other European countries in 2017*, *Antimicrobial Resistance Infection Control*, vol. 10, number 1, p. 1-7, 2021.
- 10.) [IF: 3,752, PKT: 100] Topolski Mariusz, Kozal Jędrzej, *Novel feature extraction method for signal analysis based on independent component analysis and wavelet transform*, *Plos one*, vol. 16, number 12, 2021.
- 11.) [IF: –, PKT: 140] Topolski Mariusz, *Application of the stochastic gradient method in the construction of the main components of PCA in the task diagnosis of multiple sclerosis in children*, *Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part IV* 20, p. 35-44, 2020.
- 12.) [IF: 1,139, PKT: 40] Topolski Mariusz, *The Modified Principal Component Analysis Feature Extraction Method for the Task of Diagnosing Chronic Lymphocytic Leukemia Type B-CLL.*, *J. Univers. Comput. Sci.*, vol. 26, number 6, p. 734-746, 2020.
- 13.) [IF: –, PKT: 20] Topolski Mariusz, *Algorithm of multidimensional analysis of main features of PCA with blurry observation of facility features detection of carcinoma cells multiple myeloma*, *Progress in Computer Recognition Systems* 11, p. 286-294, 2020.
- 14.) [IF: 4,639, PKT: 70] Jachowicz Estera, Róžańska Anna, Pobiega Monika, Topolski Mariusz, Wójkowska-Mach Jadwiga, *Consumption of Antibiotics and Epidemiology of Clostridioides difficile in the European Union in 2016—Opportunity for Practical Application of Aggregate ECDC Data*, *AntibioticCS*, vol. 9, number 3, 2020.
- 15.) [IF: –, PKT: 10] Topolski Mariusz, Topolska Katarzyna, *Algorithm for constructing a classifier team using a modified PCA (Principal Component Analysis) in the task of diagnosis of acute lymphocytic leukaemia type B-CLL*, *Hybrid Artificial Intelligent Systems: 14th International Conference, HAIS 2019, León, Spain, September 4–6, Proceedings* 14, p. 614-624, 2019.
- 16.) Topolska Katarzyna, Topolski Mariusz, *Wspomaganie zarządzania inteligentnym transportem wewnętrznym za pomocą wielowymiarowego modelu korespondencji MCA*, *Przedsiębiorczość i Zarządzanie*, vol. 20, number 7, cz. 2 *Logistyka w naukach o zarządzaniu*, p. 379-389, 2019.
- 17.) Topolska Katarzyna, Topolski Mariusz, *Zastosowanie inteligentnego sterowania ruchem drogowym*, *Przedsiębiorczość i Zarządzanie*, vol. 20, number 7, cz. 2 *Logistyka w naukach o zarządzaniu*, p. 369-378, 2019.
- 18.) Topolski Mariusz, *Application of smart road intelligent traffic control as a determinant of reduction of Co2 to the air*, *Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach*, number 10, p. 105-110, 2018.
- 19.) Topolski Mariusz, *Multidimensional MCA correspondence model supporting intelligent transport management*, *Archives of Transport System Telematics*, vol. 11, 2018.
- 20.) Topolski Mariusz, *Architektura systemu w planowaniu transportu w aspekcie błędu prognoz*, *Przedsiębiorczość i Zarządzanie*, vol. 19, number 12, cz. 2 *Współczesne problemy zarządzania publicznego*, p. 331-338, 2018.
- 21.) Topolski Mariusz, *A model for controlling the flow of materials in the production process*, *AUTOBUSY–Technika, Eksploatacja, Systemy Transportowe*, vol. 19, number 6, p. 1141-1144, 2018.
- 22.) Topolski Mariusz, *Zarządzanie inteligentnym transportem wewnętrznym poprzez komputerowe algorytmy probabilistyczne*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 19, 2018.
- 23.) Topolska Katarzyna, Topolski Mariusz, *Rola instrumentów zarządzania przewozami pasażerskimi w transporcie miejskim*, *Przedsiębiorczość i Zarządzanie*, vol. 19, number 11, cz. 3 *Logistyka w naukach o zarządzaniu. Część 1*, p. 363-374, 2018.

- 24.) Topolski Mariusz, *Management of intelligent internal transport through computer-based probabilistic algorithms*, AUTOBUSY–Technika, Eksploatacja, Systemy Transportowe, vol. 19, number 6, p. 968-972, 2018.
- 25.) Topolski Mariusz, *Wielowymiarowy model korespondencji wspomagający zarządzanie inteligentnym transportem w służbie zdrowia*, Przedsiębiorczość i Zarządzanie, vol. 19, number 12, cz. 2 Współczesne problemy zarządzania publicznego, p. 115-128, 2018.
- 26.) Tomala Borys, Topolski Mariusz, *Wirtualna elektrownia. Przyszłość zarządzania popytem i podażą energii*, Energetyka Ciepła i Zawodowa, 2018.
- 27.) Topolski Mariusz, *Model sterowania przepływem materiałów w procesie produkcyjnym*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 19, number 6, 2018.
- 28.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *Planning model of production process in a manufacturing enterprise*, Research in Logistics Production, vol. 7, number 5, p. 461-475, 2017.
- 29.) Topolski Mariusz, *Komputerowo wspomagany łańcuch logistyczny z zastosowaniem miękkich metod matematycznych*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, 2017.
- 30.) Topolski Mariusz, *Randomized algorithms supporting management of an intelligent internal transport system*, Archives of Transport System Telematics, vol. 10, number 3, p. 46-50, 2017.
- 31.) Topolski Mariusz, *Zastosowanie miękkich metod obliczeniowych w komputerowo wspomaganym zarządzaniu magazynem. Cz. 1*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number 6, 2017.
- 32.) Topolski Mariusz, Sitna Klaudia, *Analiza perspektyw rozwoju CSR w Polsce*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number 12, 2017.
- 33.) Mariusz Topolski, Katarzyna Topolska, *Computer recognition of data structures using cluster analysis and the theory of mathematical records*, Journal of KONBiN, vol. 41, number 1, p. 5-20, 2017.
- 34.) Topolski Mariusz, *Zastosowanie metod heurystycznych do celu identyfikacji ograniczeń procesów logistycznych*, Przedsiębiorczość i Zarządzanie, vol. 18, number 8, p. 249-258, 2017.
- 35.) Topolski Mariusz, *Zastosowanie metod heurystycznych w zadaniu optymalizacji procesów magazynowych. Cz. 3*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, 2017.
- 36.) Topolski M, *The use of Markov chains of kth row in computer management of transport in a supply chain*, Archives of Transport System Telematics, vol. 10, 2017.
- 37.) [IF: -, PKT: 20] Topolski M, *Computer algorithms probabilistic supporting intelligent transportation management of internal*, Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach, number 9, p. 93-100, 2017.
- 38.) Topolski Mariusz, *Zastosowanie zbiorów rozmytych z analizą skupień w zadaniu klasyfikacji towarów. Cz. 2*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number 6, 2017.
- 39.) Topolski Mariusz, *Komputerowe wspomaganie zarządzanie jakością transportu*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 18, number, 2017.
- 40.) Topolska Katarzyna, Kierzek Diana, Topolski Mariusz, *Negatywny wpływ subkultur na zarządzanie organizacją*, Autobusy: technika, eksploatacja, systemy transportowe, vol. 17, number 12, p. 1876-1879, 2016.
- 41.) Katarzyna Topolska, Mariusz Topolski, Michał Janicki, Czesław Kolanek *Using Computer Methods to Identify the Factors Affecting the Management of an Urban Parking Lot*, Business Challenges in the Changing Economic Landscape-Vol. 2: Proceedings of the 14th Eurasia Business and Economics Society Conference, p. 165-174, 2016.

- 42.) Topolski Mariusz, *Planowanie optymalnej trasy przejazdu transportu samochodowego z wykorzystaniem miękkich metod obliczeniowych*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, number 6, p. 1174-1179, 2016.
- 43.) Topolski Mariusz, Bujak Andrzej, *The use of telematics tools in controlling the flow of materials in a production process*, *Archives of Transport System Telematics*, vol. 9, number 4, p. 43-46, 2016.
- 44.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *The use of cluster analysis and the theory of mathematical records in the process of planning the production-warehouse flow*, *Research in Logistics Production*, vol. 6, number 3, p. 259-267, 2016.
- 45.) Topolska Katarzyna, Topolski Mariusz, *Model systemu wspomagania decyzji z wykorzystaniem koncepcji zarządzania wiedzą w przedsiębiorstwie produkcyjnym*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, 2016.
- 46.) Topolski Mariusz, Topolska Katarzyna, Janicki Michał, Kolanek Czesław, *The management of urban parking LOTS*, *Archives of Transport System Telematics*, vol. 9, 2016.
- 47.) Topolska Katarzyna, Topolski Mariusz, *Model systemu informacyjnego w procesie weryfikacji lojalności klientów przedsiębiorstwa produkcyjnego*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, 2016.
- 48.) Topolski Mariusz, *Application of a telematics system to the improvement of transport processes in intelligent high bay warehouses*, *Archives of Transport System Telematics*, vol. 9, 2016.
- 49.) Topolski Mariusz, *Komputerowy model planowania zapasów bezpieczeństwa z wykorzystaniem miękkich metod obliczeniowych*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, number 6, p. 1802-1806, 2016.
- 50.) Topolska Katarzyna, Topolski Mariusz, *Telematics system architecture to manage the internal transport*, *Archives of Transport System Telematics*, vol. 9, number 3, p. 46-49, 2016.
- 51.) Topolska Katarzyna, Topolski Mariusz, *Rozwój przedsiębiorstw transportowo-spedycyjnych w warunkach międzynarodowych*, *Autobusy: technika, eksploatacja, systemy transportowe*, vol. 17, 2016.
- 52.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *Managing light-signalling devices in view of city eco-logistics*, *Research in Logistics Production*, vol. 5, 2015.
- 53.) Bujak Andrzej, Topolski Mariusz, Miler Ryszard, *Telematics as a tool of efficient supply chain management*, *Archives of Transport System Telematics*, vol. 8, 2015.
- 54.) Topolska Katarzyna, Topolski Mariusz, *Gospodarka odpadami medycznymi w funkcjonowaniu placówki zdrowia*, *TTS Technika Transportu Szynowego*, vol. 22, 2015.
- 55.) Bujak Andrzej, Topolski Mariusz, *Concept of a Telematics System Model in Crisis Management*, *Tools of Transport Telematics: 15th International Conference on Transport Systems Telematics, TST 2015, Wrocław, Poland, April 15-17, 2015. Selected Papers 15*, p. 180-187, 2015.
- 56.) Bujak Andrzej, Topolski Mariusz, Topolska Katarzyna, *Rola informatycznych systemów telematycznych w zarządzaniu współczesnym łańcuchem dostaw*, *Logistyka*, number 4, p. 2681-2688, 2015.
- 57.) Topolski Mariusz, Topolska Katarzyna, Kobyłt Agata, *Sterowanie przepływem materiałów w procesie produkcyjnym*, *Logistyka*, number 2, p. 765-772, 2015.
- 58.) Topolska Katarzyna, Topolski Mariusz, *Capabilities for development an ITS system in the city of Wrocław*, *Archives of Transport System Telematics*, vol. 8, number 4, p. 29-32, 2015.
- 59.) Topolski Mariusz, Topolska Katarzyna, *Komputerowe algorytmy wspomagające planowanie sprzedaży w przedsiębiorstwach produkcyjnych*, *Logistyka*, 2014.
- 60.) Bujak Andrzej, Topolska Katarzyna, Topolski Mariusz, *Zastosowanie inteligentnych systemów do analizy zagrożeń komunikacyjnych*, *Logistyka*, number 5, p. 1760-1765, 2014.

- 61.) Topolska Katarzyna, Bujak Andrzej, Topolski Mariusz, *Badanie zapotrzebowania na usługi kolejowego transportu pasażerskiego na przykładzie miasta Wrocławia*, Studia Miejskie, vol. 15, p. 63-78, 2014.
- 62.) Topolski Mariusz, , *Komputerowa analiza błędów prognoz w transporcie*, Logistyka, vol. , number 6,, 2014.
- 63.) Topolska Katarzyna, Topolski Mariusz, *Use of mergers and genetic fuzzy classifiers in the task of controlling traffic lights*, Logistyka, number 6, 2014.
- 64.) Topolska Katarzyna, Topolski Mariusz, *Optymalizacja logistycznej obsługi klienta z zastosowaniem systemu symulacyjnego Anylogic*, Logistyka, 2014.
- 65.) Topolski Mariusz, *Komputerowy model rozmytego klasyfikatora Bayesa w ocenie rentowności produkcji*, Logistyka, number 6, 2014.
- 66.) Bujak Andrzej, Kobyłt Agata, Topolska Katarzyna, Topolski Mariusz, *Inteligentne rozwiązanie sterowania ruchem drogowym w logistyce miasta*, Logistyka, number 5, p. 1755-1759, 2014.
- 67.) Janicki Michał, Kolanek Czesław, Topolska Katarzyna, Topolski Mariusz, *Logistics of urban car parks*, Logistyka, number 6, 2014.
- 68.) Andrzej Bujak, Kobyłt Agata, Topolska Katarzyna, Topolski Mariusz, *Intelligent traffic control solutions for logistics center*, Logistyka, vol. 5, p. 1755-1759, 2014.
- 69.) Topolska Katarzyna, Topolski Mariusz, Błachut Bartłomiej, Haber Marcin, Piekarczyk Aleksander, *Zastosowanie fuzji klasyfikatorów rozmytych i genetycznych w zadaniu sterowania sygnalizacją świetlną*, Biuletyn Naukowy Wrocławskiej Wyższej Szkoły Informatyki Stosowanej. Informatyka, vol. 2, p. 30-33, 2012.
- 70.) Katarzyna Topolska, Topolski Mariusz, , *Optymalizacja i niezawodność inteligentnych systemów sterowania ruchem drogowym*, Problemy Eksploatacji Instytut Technologii Eksploatacji. Journal of KONBiN, vol. 20, numer 4, p. 133-140, 2011.
- 71.) Topolski Topolska Katarzyna Mariusz, *Inteligentne systemy wieloagentowe w analizie zagrożeń wypadków komunikacyjnych*, Logistyka Współczesne wyzwania. Wydawnictwo PWSZ, p. 175-181, 2011.
- 72.) [IF: -, PKT: 20] Topolska Katarzyna, Topolski Mariusz, *Zastosowanie zbiorów rozmytych w planowaniu zapasów w łańcuchu dostaw*, Logistics and Transport, vol. 8, number 1, p. 169-178, 2009.
- 73.) Kurzynski Marek, Wolczowski Andrzej, Topolski Mariusz, *Control of Hand Bioprosthesis Via Sequential Recognition of Patient's Intent Using Combination of Fuzzy Sets and Dempster-Shafer Theory*, Information Technologies in Biomedicine, p. 421-428, 2008.
- 74.) [IF: -, PKT: 20] Topolski Mariusz, Topolska Katarzyna, *Fuzja systemu ekspertowego z technologią RFID na przykładzie wybranego fragmentu łańcucha logistycznego*, Logistics and Transport, vol. 6, p. 69-78, 2008.
- 75.) [IF: -, PKT: 20] Topolski Mariusz, Topolska Katarzyna, *Warunki eksploatacji pojazdów samochodowych- analiza czynności wykonywanych podczas przeglądów samochodowych*, Logistics and Transport, vol. 7, p. 99-106, 2008.
- 76.) [IF: -, PKT: 20] Topolski Mariusz, Topolska Katarzyna, *Inteligentne systemy logistyczne w produkcji samochodów i części zamiennych*, Logistics and Transport, vol. 7, number 2, p. 91-98, 2008.

2.4c Artykuły opublikowane przed uzyskaniem stopnia doktora :

- 1.) Topolska Katarzyna, Wałkowiak Wojciech, Topolski Mariusz, *Method of analysis of measurement result cohesion with the theory of mathematical record using and concentrations analysis*, Journal of KONES, vol. 14, number 2, p. 521-528, 2007.

- 2.) Katarzyna Topolska, Walkowiak Wojciech, Mariusz Topolski, *Metoda analizy spójności wyników pomiarów z wykorzystaniem teorii ewidencji matematycznej i analizy skupień*, Journal of KONES, vol. 14, number 2, p. 521-528, 2007.
- 3.) [IF: -, PKT: 20] Topolski Mariusz, *Rozmyte sterowanie sygnalizacją świetlną w newralgicznych punktach miejskich*, Logistics and Transport, number 2, p. 101-110, 2006.
- 4.) [IF: -, PKT: 20] Topolska Katarzyna, Topolski Mariusz, *Struktura organizacyjna i system informatyczny w łańcuchu logistycznym sektora motoryzacyjnego*, Logistics and Transport, vol. 3, p. 75-82, 2006.
- 5.) [IF: -, PKT: 20] Topolska Katarzyna, Topolski Mariusz, *System logistyczny przedsiębiorstwa i jego struktura przestrzenna*, Logistics and Transport, vol. 3, p. 67-73, 2006.
- 6.) Topolska Katarzyna, Walkowiak Wojciech, Topolski Mariusz, *The modeling of stability of mechanical component based on Dempster-Shafer theory*, Journal of KONES Powertrain and Transport, vol. 13, number 4, 2006.
- 7.) Topolski Mariusz, *Rola logiki rozmytej i teorii ewidencji matematycznej w systemach logistycznych*, Zeszyty Naukowe. Logistyka i Transport/Międzynarodowa Wyższa Szkoła Logistyki i Transportu we Wrocławiu, vol. 2, number 1, p. 87-99, 2006.
- 8.) Topolska Katarzyna, Walkowiak Wojciech, Topolski Mariusz, *The validation methods of durability of ceramic coating under laboratory conditions*, Journal of KONES Powertrain and Transport, vol. 13, number 4, 2006.
- 9.) Mariusz Topolski, *Dynamiczna kontrola trwałości rozrusznika serca*, W: Niekonwencjonalne metody oceny trwałości i niezawodności. XXXIV Zimowa Szkoła Niezawodności, Szczyrk, [8-14 stycznia], p. 342-351, 2006.
- 10.) Topolski Mariusz, *The Part of Fuzzy Systems Assisting the Decision in Diagnostics of Fuel Engine Subassemblies Defects*, Journal of KONES Internal Combustion Engines, vol. 12, p. 3-4, 2005.
- 11.) Topolska Katarzyna, Topolski Mariusz, *The Dempster-Shafer model of mechanical object durability in laboratory conditions*, Journal of KONES Internal Combustion Engines, vol. 12, p. 3-4, 2005.
- 12.) Topolski Mariusz, *The fuzzy-probabilistic sequent system for control-ling the spark ignition in fuel engine*, Journal of KONES Internal Combustion Engines, vol. 12, p. 3-4, 2005.
- 13.) [IF: -, PKT: 40] Topolska Katarzyna, Topolski Mariusz, *Optimization and reliability of intelligent traffic lights control systems/Optymalizacja i niezawodność inteligentnych systemów sterowania ruchem drogowym*, Journal of KONBiN, vol. 20, number 1, p. 133-140, 2005.

2.5 Wykaz osiągnięć projektowych, konstrukcyjnych, technologicznych (z zaznaczeniem pozycji niewymienionych w pkt I.3)

—

2.6 Wykaz publicznych realizacji dzieł artystycznych (z zaznaczeniem pozycji niewymienionych w pkt I.3)

—

2.7 Informacja o wystąpieniach na krajowych lub międzynarodowych konferencjach naukowych lub artystycznych, z wyszczególnieniem przedstawionych wykładów na zaproszenie i wykładów plenarnych.

- 1.)

Konferencja: PP-RAI 2021 : 3rd Polisch Conference on Artificial Intelligence
Miejsce: Gdynia, Polska, 25-27 Kwiecień
Referat: *Feature Space Discrimination model for Machine Learning Methods with Applied of the Principal Component Analysis Method Based on Feature Rotation by Class Centroids.*
- 2.)

Konferencja: ICCS 2021 : 21th International Conference on Computational Science
Miejsce: Kraków, Polska, 16-18 czerwca
Referat: *Application of multi-objective optimization to feature selection for a difficult data classification task.*
- 3.)

Konferencja: ICCS 2020 : 20th International Conference of Computational Science
Miejsce: Amsterdam, Holandia, 3-5 czerwca 2020
Referat: *Application of the stochastic gradient method in the construction of the main components of PCA in the task diagnosis of multiple sclerosis in children.*
- 4.)

Konferencja: Hybrid Artificial Intelligent Systems : 14th International Conference, HAIS 2019
Miejsce: León, Spain, 4-6 września 2019
Referat: *Algorithm for constructing a classifier team using a modified PCA (Principal Component Analysis) in the task of diagnosis of acute lymphocytic leukaemia type B-CLL.*
- 5.)

Konferencja: 18th International Conference on Transport System Telematics, TST 2018
Miejsce: Kraków, Polska, 20-23 marca 2020
Referat: *Multidimensional MCA correspondence model supporting intelligent transport management.*
- 6.)

Konferencja: IV Konferencja Naukowa Logistyka 2018
Miejsce: Kraków, Polska, 8-9 marca 2018
Referat: *A Multidimensional Correspondence Model Supporting Intelligent Transport Management in Healthcare.*
- 7.)

Konferencja: Smart solutions in today's transport: 17th International Conference on Transport Systems Telematics, TST 2017
Miejsce: Katowice-Ustroń, Polska, 5-8 czerwca 2017
Referat: *Randomized algorithms supporting management of an intelligent internal transport system.*
- 8.)

Konferencja: Challenge of Transport Telematics: 16th International Conference on Transport Systems Telematics, TST 2016
Miejsce: Katowice-Ustroń, Polska, 16-19 marca 2016
Referat: *The use of telematics tools in controlling the flow of materials in a production process.*
- 9.)

Konferencja: Challenge of Transport Telematics: 16th International Conference on Transport Systems Telematics, TST 2016
Miejsce: Katowice-Ustroń, Polska, 16-19 marca 2016
Referat: *Application of a telematics system to the improvement of transport processes in intelligent high bay warehouses.*

10.)

Konferencja: Tools of Transport Telematics: 15th International Conference on Transport Systems Telematics, TST 2015
Miejsce: Wrocław, Polska, 15-17 kwietnia 2015
Referat: *Concept of a Telematics System Model in Crisis Management.*

11.)

Konferencja: Tools of Transport Telematics: 15th International Conference on Transport Systems Telematics, TST 2015
Miejsce: Wrocław, Polska, 15-17 kwietnia 2015
Referat: *Capabilities for development an ITS system in the city of Wrocław.*

2.7b Aktywny udział (prezentacja pracy) podczas międzynarodowych konferencji naukowych przed uzyskaniem stopnia doktora:

1.)

Konferencja: Tools of Transport Telematics: 15th International Conference on Transport Systems Telematics, TST 2015
Miejsce: Wrocław, Polska, 15-17 kwietnia 2015
Referat: *Capabilities for development an ITS system in the city of Wrocław.*

2.7c Wykłady (prezentacje) wygłoszone dla zagranicznych zespołów badawczych:

1.) Wykład w ramach stażu naukowego

Application of machine learning methods in economic sciences

7 października 2013, *National Mining University of Ukraine - Dnepropetrovsk, Odesa, Ukraine*

2.) Wykład podczas stażu naukowego

The use of metaheuristics and machine learning methods in the task of classifying supply chains

8 września 2014, *Fundación ICIL, Barcelona, Hiszpania*

2.8 Informacja o udziale w komitetach organizacyjnych i naukowych konferencji krajowych lub międzynarodowych, z podaniem pełnionej funkcji.

8a.) Członek komitetu technicznego podczas międzynarodowych konferencji naukowych po uzyskaniu stopnia doktora:

- The 12 International Conference on Computer Recognition Systems, CORES 2021,
- International Conference on Computational Science 2021, ICCS 2021
- International Conference on Computational Science 2020, ICCS 2020
- The 11 International Conference on Computer Recognition Systems, CORES 2019,
- CORES 2009: 6th International Conference on Computer Recognition Systems,

8b.) Członek komitetu technicznego podczas międzynarodowych konferencji naukowych przed uzyskaniem stopnia doktora

- CORES 2007: 5th International Conference on Computer Recognition Systems,
- CORES 2005: 4th International Conference on Computer Recognition Systems,
- Międzynarodowa Konferencja Naukowa KONES 2007
- Międzynarodowa Konferencja Naukowa KONES 2005

8c.) Członek komitetu organizacyjnego specjalnej sesji naukowej po uzyskaniu stopnia doktora

- Organizacja konferencji Polskie Porozumienie na rzecz Rozwoju Sztucznej Inteligencji, 16- 18 października 2019, Wrocław, Polska. Zasięg krajowy.

8d.) Członek komitetu organizacyjnego specjalnej sesji naukowej przed uzyskaniem stopnia doktora

- Organizacja konferencji The 5 International Conference on Computer Recognition Systems, CORES 2007, Wrocław, Polska. Zasięg międzynarodowy

2.9 Informacja o uczestnictwie w pracach zespołów badawczych realizujących projekty finansowane w drodze konkursów krajowych lub zagranicznych, z podziałem na projekty zrealizowane i będące w toku realizacji, oraz z uwzględnieniem informacji o pełnionej funkcji w ramach prac zespołów.

9a.) Projekty w toku, rozpoczęte po uzyskaniu stopnia doktora:

1.)

Tytuł	System Wykrywania Dezinformacji Metodami Sztucznej Inteligencji
Źródło finansowania	NCBiR
Budżet	8 657 006,00 zł
Okres realizacji	2021-12-02 – 2024-04-01
Partnerzy	Matic S.A., Politechnika Bydgoska
Rola w projekcie	Projekt i dokumentacja protokołu eksperymentalnego przetwarzania kaskad informacyjnych i procesu eksperymentowania.

2.)

Tytuł	Długotrwałe uczenie maszynowe na podstawie danych strumieniowych
Źródło finansowania	NCN
Budżet	596 700 zł
Okres realizacji	2021-02-01 – 2024-02-16
Partnerzy	Politechnika Wroclawska, Technical University of Ostrava
Rola w projekcie	Główny wykonawca

3.)

Tytuł	Math Solution innowacyjna platforma wspomagająca uczniów i korepetytorów w procesie nauczania indywidualnego lub wspólnego w oparciu o zaawansowane algorytmy przetwarzania obrazu i uczenia maszynowego w zakresie matematyki i innych przedmiotów ścisłych
Źródło finansowania	NCBiR
Budżet	9 162 812,50 zł
Okres realizacji	2021-11-01 – 2023-10-01
Partnerzy	Sirius Education Spółka z ograniczoną odpowiedzialnością
Rola w projekcie	Ekspert AI

4.)

Tytuł	SPIN - system spiralnej produkcji i nadzorowania danych przemysłowych do eksperymentów w procesach badawczo-rozwojowych
Źródło finansowania	NCBiR
Budżet	7 641 322,30 zł
Okres realizacji	2021-11-01 – 2023-10-31
Partnerzy	OKE Poland Spółka z ograniczoną odpowiedzialnością
Rola w projekcie	Ekspert ds. uczenia maszynowego / Ekspert AI w zakresie algorytmów klasyfikacji

9b.) Projekty zakończone, realizowane po uzyskaniu stopnia doktora:

1.)

Tytuł	Incat FaaS AI - Opracowanie platformy bezpieczeństwa operacyjnego podmiotów finansowych w oparciu o zaawansowane mechanizmy uczenia maszynowego
Źródło finansowania	NCBiR
Budżet	8 267 875 zł
Okres realizacji	2020-04-01 – 2023-03-31
Partnerzy	INCAT Spółka z ograniczoną odpowiedzialnością
Rola w projekcie	Ekspert AI

2.)

Tytuł	Opracowanie i wdrożenie systemu pomiarowo-diagnostycznego do ewidencji, monitorowania oraz automatycznej oceny terenów zadarnionych
Źródło finansowania	NCBiR
Budżet	6 662 000 zł
Okres realizacji	2021-06 – 2023-03
Partnerzy	OPTIDATA Spółka z ograniczoną odpowiedzialnością
Rola w projekcie	Architekt systemu AI

9c.) Projekty zakończone, realizowane przed uzyskaniem stopnia doktora:

—

2.10 Członkostwo w międzynarodowych lub krajowych organizacjach i towarzystwach naukowych wraz z informacją o pełnionych funkcjach

—

2.11 *Informacja o odbytych stażach w instytucjach naukowych lub artystycznych, w tym zagranicznych, z podaniem miejsca, terminu, czasu trwania stażu i jego charakteru.*

1.)

Jednostka	National Mining University in Dnietropietrowsk Odesa, Ukraine
Termin stażu	1-04-2013 – 15-11-2013
Tematyka	<i>Zastosowanie metod uczenia maszynowego w klasyfikacji nadużyć finansowych oraz popularyzacja metod sztucznej inteligencji w kształceniu studentów</i>
Charakter stażu	Celem stażu była współpraca zarówno we wspólnych pracach badawczych, jak i w wymianie doświadczeń projektowych. W ramach stażu prowadziłem badania związane zastosowaniem metod uczenia maszynowego w zadaniach wykrywania nadużyć finansowych. Ponadto uczestniczyłem w sympozjach na których wymieniałem się doświadczeniami naukowo-badawczymi oraz metodami w kształceniu studentów popularyzując zastosowanie sztucznej inteligencji w różnych problemach decyzyjnych i predykcyjnych. Celem stażu było również przeprowadzenie 90 godzin zajęć dydaktycznych dla studentów z Informatyki oraz warsztaty z zakresu sztucznej inteligencji dla ówczesnych kół naukowych. Współpraca nie zakończyła się jeszcze wspólnym artykułem, a jej owocem są raporty z badań.

2.)

Jednostka	Fundación ICIL Barcelona, Hiszpania
Termin stażu	3-09-2014 – 17-09-2014
Tematyka	<i>Zastosowanie metod sztucznej inteligencji w planowaniu i klasyfikacji łańcuchów dostaw w krajach Unii Europejskiej</i>
Charakter stażu	Celem stażu naukowego było prowadzenie badań z wykorzystaniem metod sztucznej inteligencji w klasyfikacji różnych scenariuszy łańcuchów dostaw w krajach Unii Europejskiej. Celem było również prowadzenie warsztatów oraz wykładu na temat zastosowania metaheurystyk oraz metod uczenia maszynowego w zadaniu klasyfikacji łańcuchów dostaw. W ramach tej współpracy powstały raporty wewnętrzne.

11a.) Staże zrealizowane przed uzyskaniem stopnia doktora: —

2.12 *Członkostwo w komitetach redakcyjnych i radach naukowych czasopism wraz z informacją o pełnionych funkcjach*

Jestem redaktorem Special Issue "Advances in Computer Recognition, Image Processing and Communications, Selected Papers from CORES 2023 and IPC 2023" czasopisma Entropy MDPI
<https://www.mdpi.com/journal/entropy/specialissues/Q8W90223KC>

2.13 *Informacja o recenzowanych pracach naukowych lub artystycznych, w szczególności publikowanych w czasopismach międzynarodowych.*

Wykonywałem recenzje prac dla następujących czasopism naukowych:

CZASOPISMO	IMPACT FACTOR
Machine Learning	2.940
Applied Soft Computing Journal	8.263
Journal of Computational Science	3.817
Plos	3.700
Entropy	2.700
Algorithms	2.300
Journal of Universal Computer Science	1.497

Wykonywałem recenzje prac dla następujących konferencji naukowych:

LP	KONFERENCJA
1.	10th IEEE International Conference on Data Science and Advance Analytics (DSAA-2023)
2.	WSL FORUM 2021
3.	Advances and Practical Applications of Deep and Shallow Machine Learning (APADeMaL22)
4.	3rd International Workshop on Learning with Imbalanced Domains: Theory and Applications (LIDTA2021)
5.	International Conference on Computational Science (ICCS 2021)
6.	The 12th International Conference on Computer Recognition Systems (CORES'21)
7.	International Conference on Computational Science 2020 (ICCS 2020)
8.	J.UCS Special Issue Pattern Recognition and Artificial Intelligence – current challenges, novel solutions and emerging applications (PRAISI2020)
9.	20th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL 2019)
10.	The 11th International Conference on Computer Recognition Systems CORES (CORES-19)

2.14 Informacja o uczestnictwie w programach europejskich lub innych programach międzynarodowych

Tytuł	LOGICAL - innowacyjne technologie informacyjne w transporcie i logistyce (realizowany w ramach Programu dla Europy Środkowej).
Źródło finansowania	UE
Partnerzy	Uniwersytet Ekonomiczny we Wrocławiu, Uniwersytet w Lipsku, Logistics Network w Lipsku i Halle, KIUT Regional Development Association w Zahony, Interporto Bologna S.p.a. w Bolonii, Regional Development Agency of Usti Region, PLC w Usti nad Łabą, Luka Koper Port, C.L. Consulting i Logistyka Sp. z o.o. z Wrocławia
Okres realizacji	2011-2014
Rola w projekcie	Specjalista

2.15 *Informacja o udziale w zespołach badawczych, realizujących projekty inne niż określone w pkt. II.9.*

Tytuł	Zastosowanie metod uczenia maszynowego w predykcji poziomu wypalenia zawodowego, prężności oraz poczucia własnej skuteczności u psychologów klinicznych i lekarzy psychiatrów
Okres realizacji	2021 – 2023
Patronat	Rzecznik Praw Pacjenta
Partnerzy	Politechnika Wrocławska, Gdański Uniwersytet Medyczny, Uniwersytet Gdański
Rola w projekcie	Kierowanie projektem ze strony Politechniki Wrocławskiej oraz prowadzenie badań z zastosowaniem różnych metod ekstrakcji cech oraz predykcji z zastosowaniem metod uczenia maszynowego.

2.16 *Informacja o uczestnictwie w zespołach oceniających wnioski o finansowanie badań, wnioski o przyznanie nagród naukowych, wnioski w innych konkursach mających charakter naukowy lub dydaktyczny*

3 **INFORMACJA O WSPÓŁPRACY Z OTOCZENIEM SPOŁECZNYM I GOSPODARCZYM**

3.1 *Wykaz dorobku technologicznego*

3.2 *Informacja o współpracy z sektorem gospodarczym*

- W roku 2021 w miesiącu wrześniu dla pracowników naukowych oraz developerów pracujących w przedsiębiorstwie OPTIDATA Sp. z o.o. przeprowadziłem 20 godzin szkoleń z zakresu architektury rozwiązań: (a) ekstrakcji z systemów produkcyjnych TMS klientów Optidata, (b) Ekstrakcji z systemów zewnętrznych np. GUS, IMGW, OpenWeather, (c) centralnej hurtowni danych przechowującej dane z systemów, (d) elementów predykcyjnych i sposobu ich połączenia z hurtownią danych wewnętrznych oraz systemu TMS Optidata. Druga część szkoleń dotyczyła zastosowania metod uczenia maszynowego w zagadnieniach budowy modeli uwzględniających: (a) przewidywanie czasów załadunku w zależności od kategorii kierowcy, dnia tygodnia/święta, magazynu oraz odbiorcy (sklepu), (b) wyliczenie wskaźnika spowolnienia/przyspieszenia dla kategorii kierowcy na podstawie historii realizowanych tras. Owocem szkoleń było późniejsze zatrudnienie mnie w roli architekta systemu AI w projekcie finansowanym przez NCBiR.
- Współpraca z Pracownią Psychologiczną Krzysztof Horoszkiewicz zajmującą się projektowaniem testów psychologicznych oraz dystrybucją systemu diagnostyki psychofizjologicznej SDP System. Współpraca polegała na przeprowadzeniu analiz statystycznych na danych pilotażowych oraz ekstrakcji i selekcji cech najlepiej dyskryminujących klasy diagnozowanych problemów psychologicznych (01-24.11.2022r. - potwierdzenie w postaci rekomendacji w pliku z autoreferatem).
- Dla przedsiębiorstwa Demotrans Spedycja oraz Transport opracowałem algorytmy planowania transportu oraz przeprowadziłem szkolenia dla developerów firmy z zakresu możliwości ich inte-

gracji z aktualnie stosowanym systemem (9-28.03.2022r. - potwierdzenie w postaci zaświadczenia w pliku z autoreferatem).

- Dla przedsiębiorstw FTS Artur Grzyb oraz TOM TRANS Tomasz Walczak przeprowadziłem konsultacje z zakresu wdrożenia systemów informatycznych i metod optymalizacyjnych wspomagających planowanie harmonogramowania transportu w zakresie dostarczania materiałów budowlanych do budowy dróg i mostów kołowych oraz integracji tych rozwiązań z pracą betoniarni (17-21.07.2021r. - potwierdzenie w postaci oświadczenia w pliku z autoreferatem).
- W roku 2014 będąc zatrudnionym na stanowisku Rektora Wrocławskiej Wyższej Szkoły Informatyki Stosowanej koordynowałem współpracę z różnymi przedsiębiorstwami – szczególnie z obszaru IT w zakresie tworzenia nowoczesnych laboratoriów oraz programów studiów odpowiadających potrzebą ówczesnej gospodarki. Jak również koordynowałem prace związane z tworzeniem Akademickiego Inkubatora Przedsiębiorczości (01.10.2013r. - 31.09.2014r. - potwierdzenie w postaci rekomendacji w pliku z autoreferatem).
- W 2014 roku odbyłem staż naukowy krajowy w przedsiębiorstwie *VISTEL Dariusz Ptak*. Celem była wymiana doświadczeń z zakresu nowoczesnych metod uczenia studentów Informatyki, prowadzenia badań z zastosowaniem nowoczesnych technologii informatycznych oraz konsultacji z zakresu wdrażania i utrzymania różnych rozwiązań teleinformatycznych (08.04.2014r. – 09.08.2014r - potwierdzenie w postaci zaświadczenia w pliku z autoreferatem).
- W 2013 roku odbyłem staż naukowy krajowy w przedsiębiorstwie *GOLF TELECOM*. Celem była wymiana doświadczeń z zakresu nowoczesnych metod uczenia studentów Informatyki, opracowanie metod optymalizacyjnych oraz sztucznej inteligencji dla rozwiązań sprzętowych i programistycznych dla potrzeb przemysłu (15.11.2012r. – 15.03.2013r - potwierdzenie w postaci zaświadczenia w pliku z autoreferatem).

3.3 Uzyskane prawa własności przemysłowej, w tym uzyskane patenty krajowe lub międzynarodowe

—

3.4 Informacja o wdrożonych technologiach

Po uzyskaniu stopnia doktora wdrożyłem następujące technologie:

- Modele oparte o zaawansowane mechanizmy uczenia maszynowego dla innowacyjnej platformy bezpieczeństwa operacyjnego podmiotów finansowych w oparciu o model SaaS (Software as a Service), dedykowanej w pierwszej kolejności dla organizacji FinTech i Challenger Banków. Model powstał dla firmy Incat spółki z ograniczoną odpowiedzialnością w ramach projektu NCBiR (2023 rok).
- System do zarządzania Ogólnopolską Olimpiadą Informatyczną dla Wrocławskiej Wyższej Szkoły Informatyki Stosowanej (2013 rok).
- System do zarządzania Ogólnopolską Olimpiadą Informatyczną dla Wrocławskiej Wyższej Szkoły Informatyki Stosowanej (2013 rok).
- System zarządzania biblioteką oparty na protokole Z39.50 wyszukiwania i pobierania informacji z różnych baz danych dla Międzynarodowej Wyższej Szkoły Logistyki we Wrocławiu (2012 rok).
- System do zarządzania Ogólnopolską Olimpiadą Logistyczną dla Międzynarodowej Wyższej Szkoły Logistyki we Wrocławiu (2011 rok).
- System do nauki zdalnej języków obcych dla Międzynarodowej Wyższej Szkoły Logistyki we Wrocławiu (2011 rok).

3.5 Informacja o wykonanych ekspertyzach lub innych opracowaniach wykonanych na zamówienie instytucji publicznych lub przedsiębiorców

Dla dwóch przedsiębiorstw transportowych FTS Artur Grzyb oraz TOM TRANS Tomasz Walczak przeprowadziłem w roku 2021 przeprowadziłem ekspertyzę stosowanych rozwiązań IT oraz audyt stron internetowych. W wyniku prac opracowałem zabezpieczenia systemów IT oraz stron internetowych ww. przedsiębiorstw.

3.6 Informacja o udziale w zespołach eksperckich lub konkursowych

W roku 2011 zaprojektowałem system do przeprowadzania Ogólnopolskiej Olimpiady Logistycznej. W latach 2011 - 2013 w zespole konkursowym byłem głównym organizatorem oraz sprawowałem nadzór merytoryczny.

<https://news.edubaza.pl/s/1777/42203-NEWSY-edukacja/...>

W roku 2013 zaprojektowałem system do przeprowadzania Ogólnopolskiej Olimpiady Informatycznej. W ramach zatrudnienia we Wrocławskiej Wyższej Szkole Informatyki Stosowanej byłem organizatorem I oraz II edycji ww. konkursu. Moją rolą było kierowanie zespołami informatycznym i merytorycznym.

<https://1liceum.pl/archiwum/297>

3.7 Informacja o projektach artystycznych realizowanych ze środowiskami pozaartystycznymi.

—

4 INFORMACJE NAUKOMETRYCZNE

Informacje dot. liczby punktów MEiN, współczynnika IF oraz liczby cytowań podane na podstawie wskaźników z dnia 28 sierpnia 2023.

1. Informacja o punktacji Impact Factor (w dziedzinach i dyscyplinach w których parametr ten jest powszechnie używany jako wskaźnik naukometryczny).

	Liczba prac z IF	Suma IF
Ogółem	11	39,81
Po uzyskaniu stopnia doktora	11	39,81
Przed uzyskaniem stopnia doktora	0	0

2. Informacja o liczbie cytowań publikacji wnioskodawcy, z oddzielnym uwzględnieniem autocytowań.

	Liczba wszystkich cytowań	Liczba cytowań bez autocytowań
Google Scholar	218	Brak danych
Web of Science	28	11
Scopus	48	32

3. Informacja o posiadanym indeksie Hirscha.

	h-indeks
Google Scholar	7
Web of Science	2
Scopus	4

4. Informacja o liczbie punktów MEiN.

	Liczba prac z listy MEiN	Suma punktów MEiN
Ogółem	22	1140
Po uzyskaniu stopnia doktora	12	1050
Przed uzyskaniem stopnia doktora	10	90



PODPIS ZAUFANY

MARIUSZ
TOPOLSKI
28.08.2023 06:23:38 [GMT+2]
Dokument podpisany elektronicznie
podpisem zaufanym

.....
(podpis wnioskodawcy)