

Sprawozdanie z XXVIII Konferencji Naukowej „Klasyfikacja i analiza danych – teoria i zastosowania”

Jacek BATÓG¹

Krzysztof JAJUGA²

Marek WALESIAK³

XXVIII Konferencja Naukowa Sekcji Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego (SKAD PTS) (XXXIII Konferencja Taksonomiczna) „Klasyfikacja i analiza danych – teoria i zastosowania” odbyła się 18–20 września 2019 r. w Szczecinie. Jej organizatorami były Sekcja Klasyfikacji i Analizy Danych PTS oraz Wydział Ekonomii, Finansów i Zarządzania Uniwersytetu Szczecińskiego (US).

Komitetowi Organizacyjnemu konferencji przewodniczył dr hab. Jacek Batóg, prof. US, a funkcje sekretarzy naukowych pełnili dr hab. Iwona Forys, prof. US, dr Barbara Batóg, dr Monika Rozkruit i dr Paweł Baran.

Głównymi celami konferencji były: zaprezentowanie osiągnięć i wymiana doświadczeń z zakresu teoretycznych i stosowanych zagadnień klasyfikacji i analizy danych. Konferencja stanowi coroczne forum umożliwiające badaczom podsumowanie aktualnego stanu wiedzy, przedstawienie i upowszechnianie nowatorskich dokonań oraz wskazanie kierunków dalszych prac i badań.

Tematyka konferencji obejmowała problematykę teoretyczną i praktyczną. Do tej pierwszej należały kwestie taksonomii, analizy dyskryminacyjnej, metod porządkowania liniowego, metod statystycznej analizy wielowymiarowej, metod analizy zmiennych ciągłych, metod analizy zmiennych dyskretnych, metod analizy danych symbolicznych oraz metod graficznych. Wśród dyskutowanych zastosowań znalazły się: analizy danych finansowych, marketingowych, przestrzennych, a także wykorzystanie analizy danych w medycynie, psychologii, archeologii oraz aplikacje komputerowe metod statystycznych.

¹ Uniwersytet Szczeciński, Wydział Ekonomii, Finansów i Zarządzania, Instytut Ekonomii i Finansów, Katedra Ekonometrii i Statystyki, ul. Mickiewicza 64, 71-101 Szczecin, e-mail: jacek.batog@usz.edu.pl, ORCID: <https://orcid.org/0000-0003-1413-7692>.

² Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii i Finansów, Katedra Inwestycji Finansowych i Zarządzania Ryzykiem, ul. Komandorska 118-120, 53-345 Wrocław, e-mail: krzysztof.jajuga@ue.wroc.pl, ORCID: <https://orcid.org/0000-0002-5624-6929>.

³ Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii i Finansów, Katedra Ekonometrii i Informatyki, ul. Nowowiejska 3, 58-500 Jelenia Góra, e-mail: marek.walesiak@ue.wroc.pl, ORCID: <https://orcid.org/0000-0003-0922-2323>.

W konferencji wzięło udział 85 pracowników oraz doktorantów reprezentujących: Akademię Górniczo-Hutniczą w Krakowie, Akademię im. Jakuba z Paradyża w Gorzowie Wlkp., Akademię Pomorską w Słupsku, Politechnikę Opolską, Państwową Wyższą Szkołę Zawodową w Skierniewicach, Szkołę Główną Gospodarstwa Wiejskiego w Warszawie, Szkołę Główną Handlową w Warszawie, Uniwersytet Ekonomiczny w Katowicach, Uniwersytet Ekonomiczny w Krakowie, Uniwersytet Ekonomiczny w Poznaniu, Uniwersytet Ekonomiczny we Wrocławiu, Uniwersytet Gdański, Uniwersytet Łódzki, Uniwersytet im. Adama Mickiewicza w Poznaniu, Uniwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu, Uniwersytet Mikołaja Kopernika w Toruniu, Uniwersytet Przyrodniczy w Poznaniu, Uniwersytet Szczeciński, Uniwersytet Rolniczy w Krakowie, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, a także Główny Urząd Statystyczny i urzędy statystyczne w Szczecinie, Poznaniu i we Wrocławiu.

W trakcie dwóch sesji plenarnych oraz dwunastu sesji równoległych wygłoszono 44 referaty. Odbyla się również sesja plakatowa, licząca 32 wystąpienia. Poszczególnym sesjom przewodniczyli: Andrzej Sokołowski, Krzysztof Jajuga, Eugeniusz Gatnar, Iwona Konarzewska, Marek Walesiak, Mirosław Krzyśko, Grażyna Dehnel, Barbara Pawełek, Józef Pocięcha, Andrzej Bąk, Grażyna Trzpiot, Feliks Wysocki, Andrzej Dudek, Krzysztof Najman.

Wygłoszono następujące referaty:

- Krzysztof Jajuga, *Data Science a analiza danych – nowa jakość czy nowe szaty?*;
- Eugeniusz Gatnar, *Modele pomiaru oczekiwań inflacyjnych w Polsce*;
- Barbara Pawełek i Józef Pocięcha, *Problem wartości odstających w prognozowaniu upadłości przedsiębiorstw z wykorzystaniem metody Logit leaf model*;
- Paweł Ulman i Małgorzata Ćwiek, *Ubóstwo mieszkaniowe polskich gospodarstw domowych oraz jego zróżnicowanie*;
- Iwona Bąk i Katarzyna Wawrzyniak, *Wykorzystanie metod wielowymiarowej analizy statystycznej w badaniu wypalenia zawodowego nauczycieli i wykładowców wyższych uczelni*;
- Joanna Landmesser, *Różnice w rozkładach dochodów kobiet i mężczyzn w krajach Unii Europejskiej*;
- Agnieszka Wałęga i Grzegorz Wałęga, *Psychologiczne konsekwencje nadmiernego zadłużenia gospodarstw domowych w Polsce*;
- Marek Walesiak i Grażyna Dehnel, *Zastosowanie skalowania wielowymiarowego w ocenie zmian w procesie starzenia się ludności województw Polski dla lat 2002, 2010 i 2017*;
- Jacek i Barbara Batógowie, *Analiza wpływu selekcji zmiennych na zgodność i stabilność klasyfikacji*;
- Kamila Migdał-Najman i Krzysztof Najman, *Algorytmy przyrostowe w analizie skupień*;

- Mariusz Kubus, *Wpływ niezbilansowania próby uczącej na dobór zmiennych i jakość klasyfikatorów*;
- Romana Głowicka-Wołoszyn i Feliks Wysocki, *Problem identyfikacji stopnia samodzielności finansowej gmin ze względu na występowanie asymetrii rozkładów cech prostych*;
- Przemysław Szczuciński, *Badanie atrakcyjności turystycznej gmin województwa lubuskiego z wykorzystaniem miary agregatowej*;
- Romana Głowicka-Wołoszyn, *Koncepcja przestrzennego taksonomicznego miernika potencjału. Przykład zastosowania w ocenie samodzielności finansowej gmin województwa wielkopolskiego*;
- Adam Sagan i Mariusz Grabowski, *Analiza wrażliwości w identyfikacji przyczynowych efektów mediacji w modelu TAM*;
- Justyna Majewska i Grażyna Trzpiot, *Modelowanie wskaźnika umieralności dla jednej i dwóch populacji dla wybranych krajów Europy Środkowej i Wschodniej*;
- Marcin Pełka i Aneta Rybicka, *Identyfikacja czynników decydujących o odejściu klienta z zastosowaniem hybrydowego modelu dla danych symbolicznych*;
- Paweł Baranowski, Karol Korczak i Jarosław Zajac, *Prognozowanie frekwencji na seansach kinowych w Polsce*;
- Maria Kaźmierska-Zatoń i Wojciech Zatoń, *Rola barier działalności przedsiębiorstw przemysłowych w ocenie koniunktury gospodarczej w Polsce*;
- Barbara Batóg i Katarzyna Wawrzyniak, *Porównanie propozycji transformacji nominant na stymulanty na przykładzie wskaźników finansowych dla spółek notowanych na GPW*;
- Dominik Krężolek, *Zastosowanie estymatora Hilla do oceny ryzyka ekstremalnego na rynku metali*;
- Katarzyna Cheb i Iwona Bąk, *Wykorzystanie wielowymiarowej analizy statystycznej w identyfikacji powiązań pomiędzy zrównoważonym rozwojem i zieloną gospodarką w krajach Unii Europejskiej*;
- Mariusz Doszyń, *Ekonometryczny model cen nieruchomości z restrykcjami w postaci nierówności*;
- Katarzyna Frodyma, Monika Papież i Sławomir Śmiech, *Oddzielenie wzrostu gospodarczego od emisji dwutlenku węgla w krajach Unii Europejskiej*;
- Dorota Rozmus, *Miejsce Polski w UE pod względem zrównoważonego rozwoju w świetle różnych miar stabilności grupowania*;
- Iwona Markowicz i Paweł Baran, *Przegląd wskaźników jakości danych dotyczących wewnątrzspółnotowych dostaw towarów. Wybór wskaźnika a uprządkowanie krajów*;
- Marcin Salamaga, *Propozycja badania zależności pomiędzy konkurencyjnością technologiczną i eksportową krajów Europy Środkowo-Wschodniej*;

- Iwona Konarzewska, *Wpływ siły i struktury współzależności liniowej kryteriów na wyniki wielokryteriowej analizy porównawczej*;
- Andrzej Dudek, *Indeks sylwetkowy jako narzędzie oceny jakości klasyfikacji*;
- Beata Bieszk-Stolorz, *Modele Prentice'a-Williamsa-Petersona w ocenie wpływu cech osób bezrobotnych na intensywność kolejnych zarejestrowań w urzędzie pracy*;
- Wojciech Roszka i Marcin Szymkowiak, *Estymacja wybranych charakterystyk rynku pracy w ujęciu przestrzennym z wykorzystaniem techniki symulowanego wyżarzania*;
- Andrzej Sokołowski i Małgorzata Markowska, *Elastyczne porządkowanie liniowe obiektów*;
- Jacek Białek, *Pomiar inflacji przy zastosowaniu danych skanowanych*;
- Bartłomiej Jefmański, *Syntetyczna miara rozwoju Hellwiga dla intuicjonistycznych zbiorów rozmytych*;
- Grzegorz Szczerbak, *Analiza techniczna z wykorzystaniem algorytmów rozpoznawania obrazów*;
- Grażyna Trzpiot i Justyna Majewska, *Wpływ długowieczności na wycenę zwrotów z inwestycji długoterminowych: przypadek wybranych krajów europejskich*;
- Andrzej Bąk, *Pomiar gotowości do zapłaty za dobra nierynkowe z wykorzystaniem analizy conjoint*;
- Adam Idczak, *Porównanie wybranych metod analizy text miningowej do klasyfikacji tekstów marketingowych*;
- Paweł Baran, *Wykorzystanie technik word embedding do poprawy jakości baz danych handlu zagranicznego – studium empiryczne*;
- Tomasz Górecki, Mirosław Krzyśko i Waldemar Wołyński, *Mierzenie i testowanie wzajemnej zależności dla danych funkcjonalnych*;
- Artur Zaborski, *Zastosowanie niepełnej metody tetrad do pomiaru podobieństw w niemetrycznym skalowaniu wielowymiarowym*;
- Marcin Pełka, *Wzmocnienie zdolności predykcyjnych podejścia wielomodelowego danych symbolicznych z zastosowaniem wyodrębniania zmiennych objaśniających z sieci neuronowych*;
- Wioletta Grzenda, *Dyskretyzacja zmiennych ciągłych a moc predykcyjna klasyfikatorów na przykładzie modeli regresji logistycznej oraz sztucznych sieci neuronowych*;
- Dominik Rozkrut, *AI i ML w statystyce publicznej: stan i perspektywy rozwoju*.

Pierwszego dnia konferencji odbyło się posiedzenie członków Sekcji Klasyfikacji i Analizy Danych PTS pod przewodnictwem Krzysztofa Jajugi. Plan zebrania obejmował następujące punkty:

- sprawozdanie z działalności SKAD PTS;
- informacje dotyczące planowanych konferencji krajowych i zagranicznych;

- organizacja konferencji SKAD PTS w następnych latach;
- wybór przedstawiciela SKAD PTS do International Federation of Classification Societies (IFCS) Council na kadencję 2020–2023.

Sprawozdanie z działalności SKAD PTS przedstawiła Barbara Pawełek, sekretarz naukowy sekcji. Poinformowała, że sekcja liczy obecnie 236 członków, a także przypomniała, że na stronie internetowej sekcji znajduje się regulamin oraz deklaracja członkowska.

Kolejna informacja dotyczyła artykułów po konferencji SKAD 2018 (10–12 września 2018 r. w Ciechocinku), które można było składać do czasopism: „Statistics in Transition new series” (w języku angielskim), „Przegląd Statystyczny” (w języku polskim lub angielskim), „Econometrics. Advances in Applied Data Analysis” (w języku angielskim), „Dynamic Econometric Models” (w języku angielskim) oraz „Folia Oeconomica Stetinensia” (również po angielsku). W zeszytcie 1 „Przeglądu Statystycznego” z 2019 r. ukazało się sprawozdanie z konferencji autorstwa Krzysztofa Jajugi, Tadeusza Kufła i Marka Walesiaka (s. 94–99).

Następnie Barbara Pawełek przedstawiła informacje dotyczące działalności międzynarodowej oraz udziału członków SKAD w ważnych konferencjach:

- GPSDAA 2019, Bayreuth, Niemcy, 17 marca 2019 r. (Komitet Naukowy: Józef Pociecha, przewodniczenie sesji: Józef Pociecha, Józef Dziechciarz, wystąpienia z Polski: 7);
- ECDA 2019, Bayreuth, Niemcy, 18–20 marca 2019 r. (Komitet Naukowy: Krzysztof Jajuga, Józef Pociecha, Adam Sagan, przewodniczenie sesji: Krzysztof Jajuga (dwukrotnie), Józef Pociecha, Adam Sagan, Józef Dziechciarz, wystąpienia z Polski: 13);
- IFCS 2019, Saloniki, Grecja, 26–29 sierpnia 2019 r. (Komitet Naukowy: Krzysztof Jajuga, przewodniczenie sesji: Krzysztof Jajuga, Józef Pociecha, Tomasz Górecki, wystąpienia z Polski: 10);
- CLADAG 2019, Cassino, Włochy, 11–13 września 2019 r. (Komitet Naukowy: Józef Pociecha, przewodniczenie sesji: Józef Pociecha, wystąpienia z Polski: 2).

Kolejny punkt posiedzenia sekcji obejmował zapowiedzi pokrewnych tematycznie konferencji krajowych: XXXVIII Konferencji Naukowej „Multivariate Statistical Analysis MSA 2019” (Łódź, 4–6 listopada 2019 r.), XIV Międzynarodowej Konferencji Naukowej im. Profesora Aleksandra Zeliasia „Modelowanie i prognozowanie zjawisk społeczno-gospodarczych” (Zakopane, 12–14 maja 2020 r.) oraz II Konferencji Naukowej z cyklu Metodologia Badań Statystycznych (MET2020, Warszawa, 1–3 lipca 2020 r.), jak również konferencji zagranicznych: ECDA 2020 (Neapol, 16–18 września 2020 r.), IFCS 2021, CLADAG 2021 i GPSDAA 2021. W trakcie posiedzenia podjęto też kwestię organizacji kolejnych konferencji SKAD. We wrześniu 2020 r. konferencję zorganizuje w Sopocie Uniwersytet Gdański.

W dalszej części zebrania dokonano wyboru przedstawiciela SKAD do IFCS Council. Powołano komisję skrutacyjną w składzie: Iwona Konarzewska – przewodnicząca, Andrzej Bąk – członek i Jacek Batóg – członek. Krzysztof Jajuga zgłosił kandydaturę Andrzeja Sokołowskiego, który wyraził zgodę na kandydowanie. Następnie zgłoszony i jednogłośnie poparty został wniosek o zamknięcie listy kandydatów. W tajnym głosowaniu uczestniczyło 27 członków sekcji, jednogłośnie popierając kandydaturę Andrzeja Sokołowskiego.

Na zakończenie konferencji ogłoszono przyznanie nagrody firmy StatSoft na najlepszy plakat. Jury w składzie Józef Pociecha (przewodniczący), Mirosław Krzyśko i Iwona Konarzewska zdecydowało o przyznaniu nagrody Małgorzacie Miształ i Aleksandrze Kupis-Fijałkowskiej za plakat *Rozwój technologii informacyjno-telekomunikacyjnych w Polsce na tle Unii Europejskiej – wielowymiarowa analiza statystyczna*.

Teksty referatów konferencyjnych, które zostaną przygotowane w formie recenzowanych artykułów naukowych, ukażą się w monografii *Classification and Data Analysis – Theory and Applications, Springer: Studies in Classification, Data Analysis, and Knowledge Organization* pod red. Jajugi, Batóga i Walesiaka, która ukaże się nakładem wydawnictwa Springer. Ponadto artykuły pokonferencyjne publikowane będą w czasopismach „Folia Oeconomica Stetinensia”, „Econometrics. Advances in Applied Data Analysis” i „Przegląd Statystyczny”.

Informacje o konferencji dostępne są na stronie www.skad2019.usz.edu.pl.