



XXXIV KONFERENCJA NAUKOWA

SEKCJI KLASYFIKACJI I ANALIZY DANYCH
POLSKIEGO TOWARZYSTWA STATYSTYCZNEGO

XXXIX KONFERENCJA

TAKSONOMICZNA

KLASYFIKACJA I ANALIZA DANYCH

– TEORIA I ZASTOSOWANIA

PROGRAM I KSIĄŻKA STRESZCZEŃ

23-24 września 2025

Uniwersytet Ekonomiczny we Wrocławiu

Wydział Ekonomii i Finansów

ul. Komandorska 118/120, 53-345 Wrocław

Program konferencji

Dzień 1 – 23 września 2025 (wtorek)

9:00 – Otwarcie konferencji

Sala główna (4W)

9:15 – 10:00 – Sesja keynote

Sala główna (4W)

Prowadzący: Józef Pociecha

Mark de Rooij (Leiden University)

Reduced Rank Regression for Mixed Predictor and Response Variables

10:00 – 10:20 – Przerwa kawowa

10:20 – 11:00 – Sesja keynote GUS

Sala główna (4W)

Prowadzący: Eugeniusz Gatnar

Marek Cierpiął-Wolan (Prezes GUS, Statistics Poland, University of Rzeszów)

Statistics in Hyper-Turbulent Era

11:00 – 11:30 – SKAD – 35 lat (referat historyczny)

Sala główna (4W)

Prowadzący: Paweł Lula

Józef Pociecha, Marek Walesiak, Krzysztof Jajuga

35 lat Sekcji Klasyfikacji i Analizy Danych PTS

11:30 – 12:45 – Sesje równoległe

Sesja równoległa ANG – Sala W4

Prowadzący: Krzysztof Najman

1. Andrzej Dudek

Data Preprocessing Step in Forecasting with Deep Learning Models

2. Paweł Rokita, Radosław Pietrzyk, Piotr Grzybowski

Toxic Order Flow, VPIN Spikes, and Price Changes: Implications for Market Makers in Unregulated Crypto-Token Market

Sesja równoległa PL1 – Sala W3**Prowadzący: Marek Walesiak****1. Beata Bieszk-Stolorz***Entropia czasu przeżycia w ocenie szansy podjęcia zatrudnienia***2. Krzysztof Dmytrów***Wielowymiarowa ocena metod porządkowania liniowego w wyborze lokalizacji podczas kompletacji towarów***3. Joanna Michalak, Małgorzata Szczepaniak***Profilowanie postaw wobec nierówności dochodowych w Polsce przy użyciu analizy skupień***12:45 – 13:30 – Obiad****13:30 – 14:30 – Sesja plenarna 1***Sala główna (4W)***Prowadzący: Andrzej Dudek****1. Volodymyr Melnykov***Conditional mixture modeling with applications***2. Paweł Lula***Metody klasyfikacji dokumentów tekstowych***14:30 – 15:45 – Sesje równoległe****Sesja równoległa PL2 – Sala W4****Prowadzący: Marcin Salamaga****1. Krzysztof Najman, Kamila Migdał-Najman***Wyjaśnialna Sztuczna Inteligencja – krok w stronę zwiększania zaufania do modeli AI***2. Joanna Trzęsiok***O dyskryminowaniu bez stronniczości***3. Katarzyna Raca, Alisson Soares***Dyskusje wokół teorii "Wielkiego Resetu" na Reddicie: analiza treści i dynamiki***Sesja równoległa PL3 – Sala W3****Prowadzący: Grażyna Dehnel****1. Janusz Wątroba***Uwzględnienie dodatkowych źródeł zmienności wyników badań przy ocenie istotności statystycznej efektów***2. Marta Dziechciarz-Duda, Marcin Pełka***Porządkowanie liniowe danych na podstawie danych symbolicznych histogramowych***3. Monika Świąciochowska, Łukasz Smaga***Metoda losowych projekcji w wieloczynnikowej analizie wariancji dla danych funkcjonalnych*

15:45 – 16:05 – Przerwa kawowa

15:45 – 16:30 – Sesja plakatowa

Hol główny

1. Andrzej Geise

Digitalization and innovation impact on export of high-tech products in European countries – a panel data approach

2. Dominik Krężolek

Nieparametryczne metody klasyfikacji ryzyka dla wybranych spółek GPW w Warszawie

3. Norbert Jaworski

A Novel Density-Based Metric for Measuring Spatial Competition in Urban Agglomerations

4. Michał Boda, Marta Karaś, Michał Stachura

Beyond Traditional Risk Models: A Robust Interaction-Based Framework for Systemic Risk Assessment

5. Elżbieta Antczak, Agnieszka Sobol

Ubóstwo energetyczne w Polsce – ocena stanu i trendów przestrzennych w oparciu o agregatowy miernik rozwoju

6. Paweł Kaczmarczyk

VAR models application in approximation and prediction of the telecommunication market. The empirical results in terms of effectiveness

16:15 – 17:00 – Posiedzenie Sekcji SKAD PTS

Sala główna (4W)

19:00 – Uroczysta kolacja

Dzień 2 – 24 września 2025 (środa)

9:30 – 11:10 – Sesje równoległe

Sesja równoległa PL4 – Sala W4

Prowadząca: Kamila Migdał-Najman

1. Joanna Landmesser-Rusek

Sieci neuronowe GAT w modelowaniu kursów walutowych

2. Marcin Salamaga

Badanie wpływu internacjonalizacji przedsiębiorstw na ich innowacyjność cyfrową z wykorzystaniem modelowania równań strukturalnych

3. Dorota Rozmus

Wpływ inflacji bieżącej i oczekiwanej na aktywa przedsiębiorstw przetwórstwa przemysłowego sektora publicznego

Sesja równoległa PL5 – Sala W3**Prowadząca: Katarzyna Kuziak**

1. **Tomasz Klimanek, Sylwia Filas-Przybył**
Próba oceny wpływu zmian dostępności czasowej na migracje na pobyt stały
2. **Kamil Zientarski**
O poprawności doboru i ważenia zmiennych diagnostycznych w procedurach taksonometrycznych
3. **Marcin Pełka, Aneta Rybicka**
Badanie preferencji konsumentów piwa z wykorzystaniem metody profilów ukrytych
4. **Aleksander Mercik, Krzysztof Piontek**
Efekt przenoszenia zmienności między giełdami na rynkach cyfrowych aktywów

11:10 – 11:30 – Przerwa kawowa**11:30 – 13:00 – Sesja plenarna 2***Sala główna (4W)***Prowadząca: Joanna Landmesser-Rusek**

1. **Marek Walesiak, Grażyna Dehnel**
Realizacja celu SDG 4 w państwach UE w perspektywie roku docelowego 2030: wielowymiarowa analiza wskaźnikowa z wykorzystaniem podejścia hybrydowego
2. **Katarzyna Kopczewska**
Przestrzenne reguły asocjacyjne w odkrywaniu relacji urbanizacja - zazielenienie: zmiany użytkowania gruntów pod wpływem zmian populacyjnych w Europie w latach 2006-2018
3. **Jacek Batóg, Barbara Batóg**
Wielowymiarowa analiza zadłużenia publicznego największych gospodarek świata

13:00 – 13:15 – Zamknięcie konferencji*Sala główna (4W)***13:15 – 14:00 – Obiad**

Miejsce konferencji:

Budynek W (Biblioteka)

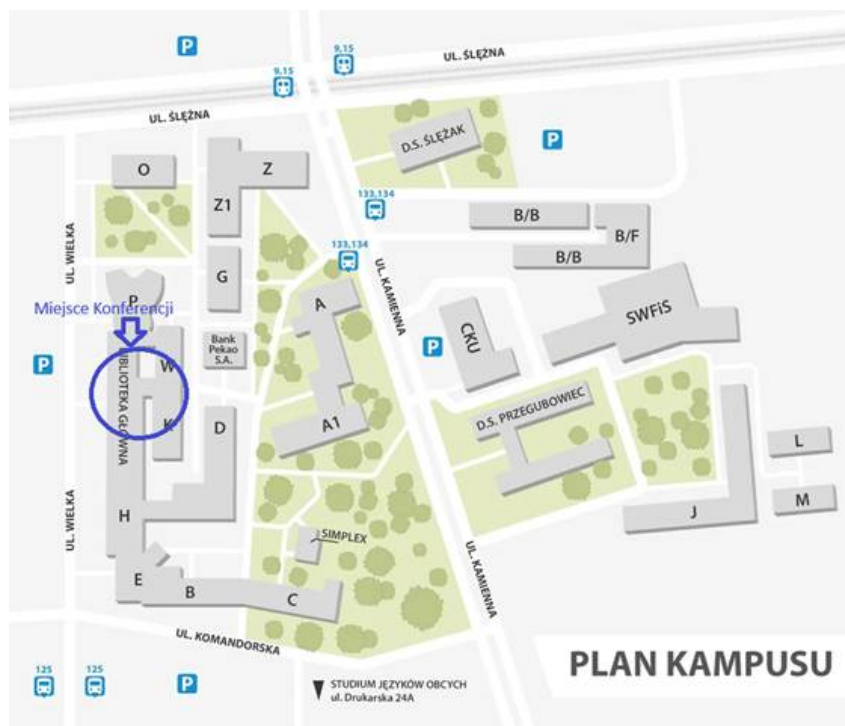
Sala główna: 4W

Sala dodatkowa: 3W

Uniwersytet Ekonomiczny we Wrocławiu

Wydział Ekonomii i Finansów

ul. Komandorska 118/120, 53-345 Wrocław



Streszczenia

Streszczenia uporządkowane alfabetycznie według nazwiska pierwszego autora

Ubóstwo energetyczne w Polsce – ocena stanu i trendów przestrzennych w oparciu o agregatowy miernik rozwoju	11
Wielowymiarowa analiza zadłużenia publicznego największych gospodarek świata . . .	11
Entropia czasu przeżycia w ocenie szansy podjęcia zatrudnienia	12
Beyond Traditional Risk Models: A Robust Interaction-Based Framework for Systemic Risk Assessment	12
Wielowymiarowa ocena metod porządkowania liniowego w wyborze lokalizacji podczas kompletacji towarów	13
Data Preprocessing Step in Forecasting with Deep Learning Models	13
Porządkowanie liniowe danych na podstawie danych symbolicznych histogramowych . .	14
Digitalization and innovation impact on export of high-tech products in European countries – a panel data approach	14
A Novel Density-Based Metric for Measuring Spatial Competition in Urban Agglomerations	15
VAR models application in approximation and prediction of the telecommunication market. The empirical results in terms of effectiveness	16
Próba oceny wpływu zmian dostępności czasowej na migracje na pobyt stały	16
Przestrzenne reguły asocjacyjne w odkrywaniu relacji urbanizacja - zazielenienie: zmiany użytkowania gruntów pod wpływem zmian populacyjnych w Europie w latach 2006-2018	17
Nieparametryczne metody klasyfikacji ryzyka dla wybranych spółek GPW w Warszawie	17
Sieci neuronowe GAT w modelowaniu kursów walutowych	18
Metody klasyfikacji dokumentów tekstowych	18
Conditional mixture modeling with applications	19
Profilowanie postaw wobec nierówności dochodowych w Polsce przy użyciu analizy skupień	19
Wyjaśnialna Sztuczna Inteligencja – krok w stronę zwiększania zaufania do modeli AI	20
Badanie preferencji konsumentów piwa z wykorzystaniem metody profilów ukrytych . .	20
Dyskusje wokół teorii "Wielkiego Resetu" na Reddicie: analiza treści i dynamiki	21
Toxic Order Flow, VPIN Spikes, and Price Changes: Implications for Market Makers in Unregulated Crypto-Token Market	21
Wpływ inflacji bieżącej i oczekiwanej na aktywa przedsiębiorstw przetwórstwa przemysłowego sektora publicznego	22
Badanie wpływu internacjonalizacji przedsiębiorstw na ich innowacyjność cyfrową z wykorzystaniem modelowania równań strukturalnych	22
Metoda losowych projekcji w wieloczynnikowej analizie wariancji dla danych funkcjonalnych	23
O dyskryminowaniu bez stronniczości	23
Realizacja celu SDG 4 w państwach UE w perspektywie roku docelowego 2030: wielowymiarowa analiza wskaźnikowa z wykorzystaniem podejścia hybrydowego	24

Uwzględnienie dodatkowych źródeł zmienności wyników badań przy ocenie istotności statystycznej efektów	25
O poprawności doboru i ważenia zmiennych diagnostycznych w procedurach taksono- metrycznych	25

Ubóstwo energetyczne w Polsce – ocena stanu i trendów przestrzennych w oparciu o agregatowy miernik rozwoju

Elżbieta Antczak¹, Agnieszka Sobol²

¹Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny

²Instytut Ochrony Środowiska

Streszczenie

W artykule przeprowadzono badanie stanu i analizę trendów przestrzennych ubóstwa energetycznego gospodarstw domowych w 380 powiatach w Polsce. W tym celu skonstruowano agregatowy miernik rozwoju na podstawie bezwzorcowej metody porządkowania liniowego. Wartości cech zostały uprzednio uśrednione i znormalizowane z wykorzystaniem metody unitaryzacji zerowanej. Takie podejście pozwoliło na kompleksowe ujęcie komponentu ubóstwa energetycznego oraz ocenę możliwości pomiaru efektów realizacji Społecznego Planu Klimatycznego (SPK) jako ważnego narzędzia polskiej polityki publicznej. Rozpoznane zostało również tło złożonych wyzwań i uwarunkowań realizacji tego dokumentu oraz powiązanego z nim Społecznego Funduszu Klimatycznego. Wyniki badań wskazały na nierówności powiatowe w poziomie ubóstwa oraz tendencję do koncentracji przestrzennej tego zjawiska. Zwrócono też szczególną uwagę na potrzebę systemowego ujęcia ubóstwa energetycznego i propozycję jego ograniczania w długookresowej strategii rozwoju kraju i samorządów. Przedstawione zostały także kluczowe wnioski i rekomendacje dla polityki publicznej.

Słowa kluczowe: Społeczny Plan Klimatyczny, ubóstwo energetyczne gospodarstw domowych, agregatowy miernik rozwoju, unitaryzacja zerowana, powiaty

Wielowymiarowa analiza zadłużenia publicznego największych gospodarek świata

Jacek Batóg¹, Barbara Batóg¹

¹Uniwersytet Szczeciński

Streszczenie

Spowolnienia gospodarcze ostatnich dwóch dekad doprowadziły do poważnego kryzysu zadłużenia publicznego. Potrzeba wsparcia systemów finansowych spowodowała, że wiele krajów przekroczyło dopuszczalne poziomy długu publicznego, co ograniczyło ich możliwości rozwojowe. Nadmierny deficyt budżetowy i zbyt wysoki poziom zadłużenia powodują również spadek konsumpcji, a także spadek inwestycji prywatnych i produkcji przemysłowej. Sformułowanie realnych i efektywnych strategii redukcji nadmiernego zadłużenia publicznego wymaga przede wszystkim określenia jego długoterminowego poziomu. Podstawowym celem badania będzie projekcja długookresowych poziomów relatywnego długu publicznego największych światowych gospodarek oraz określenie roli jaką odegrały w tym kontekście dwa ostatnie kryzysy gospodarcze. Efektem przeprowadzonych analiz będzie również klasyfikacja gospodarek według skali zadłużenia krótkoterminowego, mierzonego wartościami deficytów budżetowych oraz zadłużenia długoterminowego, z uwzględnieniem relacji tych zadłużeń do wzrostu gospodarczego.

Słowa kluczowe: dług publiczny, klasyfikacja obiektów, kryzysy gospodarcze

Entropia czasu przeżycia w ocenie szansy podjęcia zatrudnienia

Beata Bieszk-Stolorz

Uniwersytet Szczeciński

Streszczenie

Pojęcie entropii w swojej wersji pierwotnej, jako miara nieporządku, zostało opracowane w dziedzinie termodynamiki, a następnie rozszerzono je na mechanikę statystyczną. Obecnie pojęcie to ma fundamentalne znaczenie w innych dziedzinach nauki, takich jak rachunek prawdopodobieństwa i statystyka, teoria komunikacji i ekonomia. Jest to ogólna miara informacji ujawnionej przez zdarzenie. W analizie trwania, entropia czasu trwania może być uważana za informacyjny odpowiednik oczekiwanej dalszej długości trwania życia. W takim ujęciu mierzy ona matematyczną oczekiwaną ilość informacji ujawnionych przez określoną długość życia i pozwala na pomiar stopnia ryzyka związanego z wiekiem w chwili śmierci. Jeśli zmienna losowa X oznacza czas trwania jednostki w określonym stanie, wówczas entropia może być użyta jako miara związanej z tym niepewności. Celem badania jest próba oceny szansy podjęcia pracy przez bezrobotnych przy wykorzystaniu entropii resztowej i entropii przeszłej dla wykładniczego rozkładu czasu trwania w bezrobociu. W badaniu zastosowano miarę entropii czasu przeżycia skonstruowaną w oparciu o entropię Shannona. Analizowano jednostkowe dane o osobach bezrobotnych zarejestrowanych w PUP w Szczecinie.

Słowa kluczowe: analiza przeżycia, entropia, bezrobocie

Beyond Traditional Risk Models: A Robust Interaction-Based Framework for Systemic Risk Assessment

Michał Boda¹, Marta Karaś², Michał Stachura³

¹*Uniwersytet Ekonomiczny w Krakowie*

²*Uniwersytet Ekonomiczny we Wrocławiu*

³*Uniwersytet Jana Kochanowskiego*

Abstract

This paper introduces a novel methodological framework for analyzing systemic risk and its impact on banks' capital adequacy, combining interaction panel models with asymmetric ΔCoVaR . Our approach is specifically designed to handle complex interdependencies among 104 endogenous and exogenous variables, enabling a nuanced assessment of dynamic risk transmission mechanisms. To develop, validate, and apply this framework, we estimated over 400 model specifications, through which we established a comprehensive and rigorous robustness testing strategy. This extensive validation process demonstrates that both our methodological innovations and empirical findings are highly robust across a wide range of scenarios. We illustrate the utility of the approach through an empirical application to 120 systemically important European banks across major financial and geopolitical events—including the global financial crisis, the European debt crisis, Brexit, the COVID-19 pandemic, and the war in Ukraine. The findings—such as the reversal of conventional macroeconomic effects on total capital ratios (TCRs) post-crisis, the reduced role of housing market dynamics in capital resilience, and evidence of persistent transatlantic spillovers—highlight the power of our methodology in uncovering structural changes in systemic

risk dynamics that traditional models often miss.

Keywords: Systemic Risk, Interaction Panel Models, ΔCoVaR , Robustness Testing, Capital Adequacy

Wielowymiarowa ocena metod porządkowania liniowego w wyborze lokalizacji podczas kompletacji towarów

Krzysztof Dmytrów
Uniwersytet Szczeciński

Streszczenie

W składowaniu współdzielonym każdy indeks towarowy może być składowany w wielu lokalizacjach w magazynie. Jeżeli stosowane jest składowanie współdzielone, podczas kompletacji, należy wybrać lokalizacje do odwiedzenia przez magazyniera (który to problem nie występuje przy składowaniu dedykowanym, w którym każdy indeks towarowy jest składowany w jednej lokalizacji). Do wyboru lokalizacji w składowaniu współdzielonym można zastosować metody porządkowania liniowego. Za ich pomocą można zapewnić realizację różnych strategii wyboru lokalizacji. W badaniu wykorzystano cztery metody porządkowania liniowego: COPRAS (COMplex PROportional ASsessment), TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) oraz TMAL (Taksonomiczna Miara Atrakcyjności Lokalizacji), w dwóch wersjach: TMAL(+), oparta na odległości od wzorca oraz TMAL(−), oparta na odległości od antywzorca. Jest ona oparta na syntetycznym mierniku rozwoju Hellwiga. Celem badania jest ocena tych metod porządkowania liniowego pod względem zapewnienia realizacji różnych strategii wyboru lokalizacji. Ocena została dokonana za pomocą jednoczynnikowej analizy wariancji pod względem najważniejszych z punktu widzenia realizacji każdej strategii wyboru zmiennych/charakterystyk.

Słowa kluczowe: kompletacja zamówień, składowanie współdzielone, porządkowanie liniowe

Data Preprocessing Step in Forecasting with Deep Learning Models

Andrzej Dudek
Uniwersytet Ekonomiczny we Wrocławiu

Abstract

Data preprocessing plays a crucial role in the accuracy and reliability of forecasting tasks, especially when using deep learning models. One of the essential steps is data transformation, which ensures that the input time series is appropriately prepared for modeling. Direct comparisons of forecasting methods using identical raw datasets can lead to misleading conclusions, as each model type has different requirements for input structure. Thus, comparing model outputs without accounting for proper reverse transformation of the data can distort benchmark results. In particular, raw data is typically unsuitable for deep learning models without preprocessing, as these models require normalized or integrated inputs to learn effectively.

Statistical forecasting models such as ARIMA and GARCH, on the other hand, inherently assume certain properties like stationarity, and thus necessitate integration in cases of non-stationarity—an almost ubiquitous condition in real-world time series like stock market prices

or energy consumption. Consequently, the need for an appropriate transformation of input data is not only critical but also method-dependent. Therefore, the first and foundational step in any meaningful benchmark analysis should be the tailored preprocessing of data aligned with the requirements of the respective modeling approach.

In the presentation five preprocessing strategies for time series forecasting: raw data, max-min scaled data, integrated data (using return ratios), integrated and scaled to percentage form, and finally, integrated followed by max-min scaling will be compared on real stock data timesets. The study demonstrates that careful consideration of preprocessing is essential to unlock the full potential of deep learning models in forecasting applications and to ensure fair comparisons across methods.

Keywords: time series, forecasting, deep learning, deep neural networks

Porządkowanie liniowe danych na podstawie danych symbolicznych histogramowych

Marta Dziechciarz-Duda¹, Marcin Pełka¹

¹*Uniwersytet Ekonomiczny we Wrocławiu*

Streszczenie

W artykule zaprezentowano rozwiązania metodyczne pozwalające na przeprowadzenie porządkowania liniowego dla danych symbolicznych histogramowych. Podstawą do zastosowania porządkowania dla tego typu danych jest miara odległości Wassersteina oraz skalowanie wielowymiarowe dla danych symbolicznych histogramowych, które pozwoli na graficzną reprezentację wyników oraz wyznaczenie izokwant rozwoju. W części empirycznej na podstawie danych regionalnych OECD przygotowano dane opisujące kraje Unii Europejskiej z wykorzystaniem danych symbolicznych histogramowych, które poddano dwustopniowej procedurze porządkowania liniowego.

Słowa kluczowe: porządkowanie liniowe, dane symboliczne histogramowe, miara odległości Wassersteina

Digitalization and innovation impact on export of high-tech products in European countries – a panel data approach

Andrzej Geise

Uniwersytet Mikołaja Kopernika w Toruniu

Abstract

In the era of artificial intelligence (AI) and accelerating technological change, high-tech export is becoming crucial factor for economic development and growth, as well for building the competitiveness of the economy. High-tech export in economy provides a valuable information about the position of country in the global technology market, as well it helps to understand comparative advantages created by innovation and high-tech products across countries. Also the most important determinant of development is knowledge accumulation, which helps to achieve economic and social progress by concentrating on exporting products with high added value, based on technology, digitalization and artificial intelligence, rather than exporting products with lower added value.

The analysis of connectedness among digitalization, innovation, export of high-tech products and economic growth of EU countries is motivated by several economic policy concerns, where the most important problems are related to building the strategic autonomy of EU based on innovation, knowledge commercialization and rule of law. Other concerns are related to building resilience of EU member economies to geopolitical tensions, digital transformation policy and innovation policy effectiveness.

Based on this relationships the study aims to analyze the impact of the digitalization and innovation of the economy on the high-tech exports and economic growth in European Union countries. Linear and power dynamic panel models built in a sample of 27 economies over the period 2017-2022 were used to meet the objective. The GMM estimator was used to determine the parameters.

Keywords: digitalization, innovation, high-tech, panel model, GMM

A Novel Density-Based Metric for Measuring Spatial Competition in Urban Agglomerations

Norbert Jaworski

University of Warsaw, Faculty of Economic Sciences

Abstract

Traditional methods of analyzing firm competition often overlook the spatial dimension of accessibility and consumer distribution, particularly in dense urban settings. There is a growing need for spatially sensitive approaches that integrate geographic and demographic data to assess market dynamics more accurately.

Understanding spatial competition at this granular level is increasingly important for several reasons. As Kopczewska (2020, 2021) highlights, spatial interactions between consumers and firms, and the structure of urban agglomerations, are crucial for understanding location behavior of firms. In urban planning, fine-grained spatial analyses can inform zoning policies, infrastructure investment, and service provision, helping to address spatial inequalities in access to amenities. Unlike traditional models of spatial competition that rely on stylized geometries, such as the linear city or circular city, our approach employs Voronoi tessellation to partition the urban landscape into mutually exclusive zones of influence, each anchored on a retail facility. Drawing inspiration from ecological models of inter- and intra-species competition and adapting classical spatial econometric frameworks that incorporate heterogeneous neighborhood densities, we translate the concept of territorial influence into an urban retail context.

By integrating ecological concepts of territorial competition with contemporary spatial econometric techniques, our metric provides urban economists, city planners, and retail strategists with a rigorous, demographically grounded tool. This facilitates a deeper understanding of competitive dynamics, enables optimization of location strategies, and supports informed urban planning and antitrust policy decisions.

Keywords: competition, spatial, urban economics, supermarket, voronoi

VAR models application in approximation and prediction of the telecommunication market. The empirical results in terms of effectiveness

Paweł Kaczmarczyk

Akademia Mazowiecka w Płocku

Abstract

The aim of the described research is to assess the effectiveness of VAR models in approximating and predicting the seasonality of the demand for telecommunications services for hourly data - times of calls outgoing from the selected operator's network in seconds summed up every hour.

The approximation and the prediction efficiency of the models estimated on the basis of data from a two-year period was examined. The data made it possible to take into account daily, weekly, monthly and annual seasonality for a given type of connection and a specific subscriber group.

VAR models were shown to be more efficient than selected single-equation models.

Keywords: VAR models, demand for telecommunication services, seasonality, approximation, prediction

Próba oceny wpływu zmian dostępności czasowej na migracje na pobyt stały

Tomasz Klimanek^{1,2}, Sylwia Filas-Przybył²

¹*Uniwersytet Ekonomiczny w Poznaniu*

²*Urząd Statystyczny w Poznaniu*

Streszczenie

Po 2004 r. (przystąpienie do UE) Polska przeszła znaczące zmiany w migracjach wewnętrznych na pobyt stały, w szczególności w przemieszczaniu się mieszkańców z dużych miast na okoliczne tereny podmiejskie. W zaprezentowanym wystąpieniu podjęto próbę oceny wpływu lepszej dostępności, a mianowicie skrócenia czasu podróży do ośrodków miejskich, na wzorce migracji, wykorzystując jako studium przypadku Poznań.

Proces suburbanizacji spowodował wzrost populacji na obszarach peryferyjnych, co wpływa na planowanie urbanistyczne, rozwój infrastruktury i świadczenie usług. Zrozumienie czynników determinujących tę migrację jest niezbędne do stworzenia spójnych i trwałych interakcji miejsko-podmiejskich.

W niniejszym badaniu wykorzystano analizę sieciową do mapowania trendów migracyjnych między Poznaniem a sąsiednimi gminami, korelując te ruchy ze zmianami w dostępności w badanym okresie. Badanie umożliwiły macierze dostępności opublikowane przez Ośrodek Statystyki Miast Urzędu Statystycznego w Poznaniu, wykorzystujące dane OpenStreetMap.

Wyniki dostarczają istotnych spostrzeżeń na temat relacji między poprawą dostępności a migracjami wewnętrznymi na pobyt stały. Badania podkreślają wpływ infrastruktury transportowej na wzorce migracji na pobyt stały.

Słowa kluczowe: migracje na pobyt stały, dostępność, suburbanizacja, analiza sieciowa

Przestrzenne reguły asocjacyjne w odkrywaniu relacji urbanizacja - zazielenienie: zmiany użytkowania gruntów pod wpływem zmian populacyjnych w Europie w latach 2006-2018

Katarzyna Kopczewska

Uniwersytet Warszawski

Streszczenie

Ważnym pytaniem w zakresie zrównoważonego rozwoju jest w jaki sposób zmiany w populacji wpływają na zmiany w użytkowaniu gruntów. W przypadku wzrostu populacji można oczekiwać zagęszczania osadnictwa, które nie wiąże się z zajmowaniem gruntów lub rozlewaniem się w przestrzeni, które przekształca grunty w tereny zabudowane (urbanizacja). W przypadku spadku populacji, grunty mogą pozostać terenami zabudowanymi lub zostać ponownie wykorzystane do innych celów (zazielenianie). Przedstawiona analiza dotyczy relacji dla Europy w latach 2006-2018, w oparciu o dane LUCAS dotyczące użytkowania gruntów i spis powszechny ludności. Jest to analiza o niskiej granulacji w grid 2 km x 2 km dla ok. 1 miliona komórek siatki. Wykorzystuje metody uczenia bez nadzoru – geo-lokalizowane reguły asocjacyjne w celu wykrycia głównych trendów oraz ekonometryczne modele probitowe w celu znalezienia czynników, które określają prawdopodobieństwo zmian w użytkowaniu gruntów. Wyniki pokazują, że nie ma silnej presji populacyjnej. W Europie obserwowane są wszystkie możliwe wzory zmian, jednak obszary zabudowane pozostają takie, jakie były lub nawet znikają przy wzroście populacji. Wyjaśnienie tego zjawiska leży w zazielenianiu miast i polityce zrównoważonego rozwoju. Artykuł powstał w ramach projektu LAMASUS finansowanego z programu Horyzont 2020.

Słowa kluczowe: presja populacyjna, reguły asocjacyjne, dane o niskiej granulacji, baza danych LUCAS, gridowa baza danych, zazielenianie miast

Nieparametryczne metody klasyfikacji ryzyka dla wybranych spółek GPW w Warszawie

Dominik Krężolek

Uniwersytet Ekonomiczny w Katowicach

Streszczenie

Celem artykułu jest zastosowanie wybranych nieparametrycznych metod klasyfikacji do oceny poziomu ryzyka inwestycyjnego spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie w roku 2024. Analiza obejmuje wybrane spółki, dla których zgromadzono dane o dziennych stopach zwrotu oraz wskaźnikach fundamentalnych (takich jak P/E, ROE, zadłużenie). Stopy zwrotu zostały uśrednione do poziomu rocznego, natomiast zmienna zależna opisująca poziom ryzyka została zdefiniowana na podstawie liczby dni, w których roczna seria stóp zwrotu spółki przekraczała wartość VaR na poziomie 0,05 dla indeksu WIG20. Na tej podstawie utworzono trzystopniową klasyfikację ryzyka: ryzykowne, bardzo ryzykowne i ekstremalnie ryzykowne.

W badaniu wykorzystano cztery nowoczesne, nieparametryczne algorytmy klasyfikacyjne: Random Forest, Gradient Boosting Machines (XGBoost), Maszynę wektorów nośnych (SVM) oraz k-Nearest Neighbors (k-NN). Modele oceniano za pomocą walidacji krzyżowej i miar klasyfikacyjnych (accuracy, AUC, macierz pomyłek). Najlepsze rezultaty uzyskano dla metod zespołowych (Random Forest, XGBoost), co potwierdza ich wysoką skuteczność w analizie ryzyka

spółek giełdowych oraz elastyczność nieparametrycznego podejścia w modelowaniu złożonych danych finansowych.

Słowa kluczowe: klasyfikacja ryzyka, metody nieparametryczne, GPW w Warszawie, ekstremalne stopy zwrotu

Sieci neuronowe GAT w modelowaniu kursów walutowych

Joanna Landmesser-Rusek

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Streszczenie

Modelowanie kursów walutowych jest ważnym zagadnieniem w analizach rynków finansowych. Kursy walut tworzą dynamiczną sieć powiązań, którą można skutecznie modelować przy użyciu grafowych sieci neuronowych (graph neural networks, GNN). Kluczowym mechanizmem GNN jest propagacja informacji między węzłami grafu, umożliwiającą lepsze modelowanie interakcji walutowych. Każdy węzeł aktualizuje swoją reprezentację, agregując cechy od sąsiadów i łącząc je ze swoimi własnymi. W popularnych grafowych sieciach konwolucyjnych (convolutional graph neural networks, GCN) wszystkie sąsiednie węzły są traktowane w ten sam sposób, jednak w rzeczywistości niektóre z nich mogą mieć większy wpływ niż inne. Aby uwzględnić takie zmienne znaczenie sąsiadów, wprowadzono sieci GAT (graph attention networks), które umożliwiają przypisywanie różnych wag węzłom w tym samym sąsiedztwie.

Celem badania jest ocena skuteczności sieci GAT w modelowaniu kursów walutowych. Analiza obejmie szeregi czasowe dla głównych walut światowych w ciągu ostatnich trzech lat. Wyniki modelowania uzyskane przy użyciu GAT zostaną porównane z wynikami z modeli benchmarkowych, takich jak ARIMA, GARCH, MLP, GCN, LSTM-GCN. Badanie wskaże, czy sieci GAT przewyższają klasyczne metody statystyczne. Wyniki mogą mieć praktyczne zastosowanie, wspierając inwestorów i analityków w podejmowaniu decyzji i zarządzaniu ryzykiem.

Słowa kluczowe: GAT, grafowe sieci neuronowe, rynek walutowy

Metody klasyfikacji dokumentów tekstowych

Paweł Lula

Uniwersytet Ekonomiczny w Krakowie

Streszczenie

W pracy zostaną zaprezentowane badania dotyczące metod klasyfikacji tekstów.

Rozważane będą modele: a) Pozwalające na przeprowadzenie klasyfikacji płaskiej i hierarchicznej, b) Pozwalające na klasyfikację dwu- lub wielostanową, c) Przypisujące dokumenty do jednej lub do wielu klas jednocześnie, d) Pracujące na poziomie wyrazów, zdań i pełnych dokumentów, e) Oparte na regułach lub wykorzystujące klasyfikatory powstałe w procesie uczenia maszynowego, f) Wykorzystujące reprezentację tekstu w postaci sekwencji znaków, macierzy częstości i osadzeń, g) Wykorzystujące podejście jedno- i wielomodelowe.

Przedstawione zostaną etapy realizacji procesu klasyfikacji oraz możliwości ich wsparcia za pomocą dostępnych pakietów obliczeniowych. Zaprezentowane zostaną miary pozwalające na ocenę jakości klasyfikacji. We wnioskach przedstawiona zostanie ocena prezentowanych podejść

oraz uwagi dotyczące wyboru najlepszej metody w zależności od charakterystyki zadania.

Słowa kluczowe: przetwarzanie języka naturalnego, klasyfikacja tekstów

Conditional mixture modeling with applications

Volodymyr Melnykov

ISM Department

Abstract

Finite mixture models are often at the risk of overparameterization due to a large number of parameters even in those cases when the number of components is low. In practice, it can cause overfitting and mixture order underestimating. One traditional approach to relax this issue is to reduce the number of parameters by considering parsimonious models. As the number of parameters is especially high in covariance matrices, the majority of mixture modeling methods consider their reparameterizations. We propose an alternative procedure that shows remarkable modeling flexibility and demonstrate its utility on simulated as well as well-known classification data sets.

Keywords: finite mixture model, model-based clustering, parsimonious models

Profilowanie postaw wobec nierówności dochodowych w Polsce przy użyciu analizy skupień

Joanna Michalak¹, Małgorzata Szczepaniak¹

¹*Uniwersytet Mikołaja Kopernika w Toruniu*

Streszczenie

Dotychczasowe badania nad nierównościami dochodowymi koncentrują się głównie na pomiarze ich poziomu, identyfikacji źródeł oraz analizie społeczno-ekonomicznych konsekwencji. W niniejszym badaniu uwaga skierowana została jednak na subiektywną percepcję nierówności, która nie zawsze koreluje z ich obiektywnym poziomem. Postawy wobec nierówności kształtowane są przez indywidualne doświadczenia, porównania społeczne oraz dominujące narracje, co może prowadzić do wzrostu poczucia niesprawiedliwości, obniżenia satysfakcji z życia oraz większego poparcia dla polityk redystrybucyjnych. Choć większość osób postrzega nierówności jako zbyt duże i deklaruje preferencję dla większego poziomu równości, postawy te są silnie zróżnicowane i zależą m.in. od poziomu dochodu oraz pozycji jednostki w strukturze społeczno-ekonomicznej.

Głównym celem badania była identyfikacja oraz charakterystyka zróżnicowanych postaw wobec nierówności dochodowych w Polsce za pomocą metod analizy skupień (profilowania). Celem wspierającym było porównanie skuteczności wybranych algorytmów klasteryzacyjnych w kontekście badania postaw społecznych.

W analizie zastosowano szereg zaawansowanych metod grupowania: K-means, MiniBatchK-Means, DBSCAN, aglomeracyjną hierarchiczną analizę skupień, Gaussian Mixture Model (GMM) oraz Spectral Clustering. Najlepsze rezultaty uzyskano dla algorytmu MiniBatchKMeans, który osiągnął najwyższy indeks Calinskiego-Harabasz, najwyższy silhouette score oraz najniższy wskaźnik Daviesa-Bouldina. Analiza wykazała, że kluczowym czynnikiem różnicującym wyodrębnione klastry postaw wobec nierówności dochodowych były subiektywne oceny własnej pozycji

dochodowej.

Słowa kluczowe: profilowanie postaw, nierówności, analiza skupień

Wyjaśnialna Sztuczna Inteligencja – krok w stronę zwiększania zaufania do modeli AI

Krzysztof Najman¹, Kamila Migdał-Najman¹

¹Uniwersytet Gdański

Streszczenie

Zdecydowana większość modeli sztucznej inteligencji to modele tzw. czarnej skrzynki. Wiemy jakie dane są argumentami wejściowymi i możemy zaobserwować wartości wyjściowe modelu. Bardzo trudno jest jednak wyjaśnić w jaki sposób wartości te uzyskano. Jest to w wielu zastosowaniach empirycznych poważny problem.

Wyjaśnialna sztuczna inteligencja (Explainable Artificial Intelligence, XAI) to stosunkowo nowa dziedzina badań nad modelami sztucznej inteligencji, która koncentruje się na uczynieniu modeli AI bardziej przejrzystymi i zrozumiałymi dla ludzi. Jej celem jest dostarczanie jasnych, zrozumiałych dla ludzi wyjaśnień decyzji generowanych przez modele AI, zwiększając zaufanie użytkowników do uzyskanych wyników. XAI ma kluczowe znaczenie dla odpowiedzialnego wdrażania AI, umożliwiając interesariuszom zrozumienie, w jaki sposób działają systemy AI i dlaczego formułują one takie a nie inne przewidywania.

W referacie dokonano identyfikacji i oceny kluczowych elementów XAI z punktu widzenia interesariuszy różnych modeli AI.

Słowa kluczowe: sztuczna inteligencja, modele AI, Explainable Artificial Intelligence

Badanie preferencji konsumentów piwa z wykorzystaniem metody profilów ukrytych

Marcin Pełka¹, Aneta Rybicka¹

¹Uniwersytet Ekonomiczny we Wrocławiu

Streszczenie

Badanie preferencji konsumentów jest jednym z kluczowych zagadnień w ramach badań rynkowych i ekonomicznych. W teorii ekonomii wyróżnia się preferencje ujawnione (dokonane zakupy, dane historyczne) oraz wyrażone (intencje zakupu). W pomiarze preferencji wykorzystać można m.in. metody profilów ukrytych (latent profile analysis), które nie wymagają ustalenia liczby badanych profilów, a jedynie ich atrybutów lub poziomów. W prezentacji przedstawiono kompozycyjne wykorzystanie metody profilów ukrytych w badaniu preferencji konsumentów piwa. Otrzymane wyniki wskazują, że cechy takie jak zawartość alkoholu, wielkość oraz typ opakowania są najważniejszymi cechami w ramach dwóch wyodrębnionych profilów.

Słowa kluczowe: analiza profilów ukrytych, badanie preferencji, tidyLPA

Dyskusje wokół teorii "Wielkiego Resetu" na Reddicie: analiza treści i dynamiki

Katarzyna Raca¹, Alisson Soares²

¹*Uniwersytet Gdański*

²*Digital Humanities Lab - Federal University of Bahia - Brazil*

Streszczenie

Teorie spiskowe są obecne w wielu dziedzinach życia, zwłaszcza w obszarach zdrowia publicznego, polityki, zmian klimatycznych oraz zjawisk pozaziemskich. Od lat niektóre grupy społeczne formułują alternatywne narracje, dotyczące m.in. istnienia UFO, płaskiej Ziemi czy szkodliwości szczepień. Powstawanie i utrwalanie tego typu przekonań wiąże się z występującymi w mediach zniekształceniami informacji, jak i z podatnością społeczeństwa do przyjmowania uproszczonych wyjaśnień.

Rozwój Internetu i mediów społecznościowych znacznie ułatwił rozpowszechnianie się tego typu narracji, a ich autorzy coraz częściej czerpią z tego korzyści finansowe. Jedną z nowszych teorii spiskowych jest "Wielki Reset", zniekształcona interpretacja idei promowanych przez Światowe Forum Ekonomiczne. Teoria ta przyczynia się do wzrostu nieufności wobec polityków i instytucji, pogłębiając społeczną polaryzację, zwłaszcza w państwach demokratycznych.

Celem niniejszego badania jest analiza treści i dynamiki dyskusji wokół teorii "Wielkiego Resetu" na platformie Reddit. W badaniu wykorzystano metody analizy sieci społecznościowych, pozwalające na wyodrębnienie wpływowych grup użytkowników, oraz narzędzia z zakresu text miningu, które umożliwiają identyfikację dominujących tematów i emocji pojawiających się w dyskusjach. Wyniki badania mogą stanowić istotny wkład w zrozumienie mechanizmów stojących za rosnącą popularnością teorii spiskowych.

Słowa kluczowe: wielki reset, teorie spiskowe, text mining, analiza sieci społecznościowych

Toxic Order Flow, VPIN Spikes, and Price Changes: Implications for Market Makers in Unregulated Crypto-Token Market

Paweł Rokita^{1,2}, Radosław Pietrzyk^{1,2}, Piotr Grzybowski²

¹*Uniwersytet Ekonomiczny we Wrocławiu*

²*EMPIRICA Sp. z o.o.*

Abstract

This study refers to and builds on our recent findings regarding the co-occurrence of Volume-Synchronized Probability of Informed Trading (VPIN) spikes, specifically those exceeding the threshold value of 0.7, and anomalous short-term price movements that deviate from established longer-term market trends. This time, we are going further, to provide an in-depth analysis of the VPIN-price relationship. Furthermore, we outline actionable recommendations for market makers, assisting them in mitigating risks associated with informed trading and manipulative practices. This research contributes to a larger, ongoing project aimed at developing a robust modeling framework capable of evaluating and scoring financial exchanges based on vulnerability to manipulative behaviors, including toxic order flow and pump-and-dump schemes.

Keywords: VPIN, Toxic Flow, Crypto-Exchanges, Market Manipulations, Market Makers

Wpływ inflacji bieżącej i oczekiwanej na aktywa przedsiębiorstw przetwórstwa przemysłowego sektora publicznego

Dorota Rozmus

Uniwersytet Ekonomiczny w Katowicach

Streszczenie

Celem referatu jest porównanie wpływu inflacji mierzonej wskaźnikiem CPI r/r oraz oczekiwań inflacyjnych na aktywa przedsiębiorstw przetwórstwa przemysłowego należących do sektora publicznego w Polsce. Badanie odpowiada na pytanie, która z tych zmiennych silniej oddziałuje na wartość aktywów – bieżąca dynamika cen czy oczekiwania inflacyjne.

W badaniu zastosowano modele wektorowej autoregresji (VAR) oraz test przyczynowości Grangera analizując dane kwartalne z lat 2014–2024. Źródłem danych były kwartalne sprawozdania o aktywach i pasywach finansowych (RF-01) publikowane przez Główny Urząd Statystyczny (GUS) oraz dane Narodowego Banku Polskiego (NBP).

W referacie zostaną przedstawione wyniki badania, w tym wnioski dotyczące wpływu inflacji na aktywa przedsiębiorstw.

Słowa kluczowe: inflacja, oczekiwania inflacyjne, aktywa przedsiębiorstw, modele VAR, test przyczynowości Grangera

Badanie wpływu internacjonalizacji przedsiębiorstw na ich innowacyjność cyfrową z wykorzystaniem modelowania równań strukturalnych

Marcin Salamaga

Uniwersytet Ekonomiczny w Krakowie

Streszczenie

Technologie cyfrowe są kluczowym znakiem rozpoznawczym współczesnego społeczeństwa i gospodarki. Bez nich skuteczne konkutowanie przedsiębiorstw na globalnym rynku jest trudne lub niemożliwe. Inwestowanie w technologie cyfrowe i aktywność innowacyjna przedsiębiorstwa w tym zakresie staje się często koniecznością. Związek pomiędzy umiędzynarodowieniem przedsiębiorstw i ich innowacyjnością ma swoje podstawy w teoriach z zakresu ekonomii i nauk o zarządzaniu. Jak do tej pory nie podjęto jednak szerszych badań nad wpływem internacjonalizacji firm na ich aktywność w zakresie cyfrowych innowacji. Niniejszy referat próbuje wypełnić tę lukę. W referacie zaproponowano wykorzystanie modelowania równań strukturalnych opartego na metodzie cząstkowych najmniejszych kwadratów (ang. Partial Least Squares – PLS) do badania wpływu różnych form internacjonalizacji firm na ich innowacyjność cyfrową. Analizie poddano losową próbę firm z woj. małopolskiego i śląskiego. Celem referatu jest wskazanie tych metod internacjonalizacji przedsiębiorstw, które w największym stopniu wpływają na cyfrową innowacyjność produktową, procesową, marketingową i organizacyjną. Wyniki badań mogą być przydatne dla menadżerów firm rozważających ekspansję na rynki zagraniczne.

Słowa kluczowe: równania strukturalne, internacjonalizacja, innowacyjność

Metoda losowych projekcji w wieloczynnikowej analizie wariancji dla danych funkcjonalnych

Monika Świąciochowska¹, Łukasz Smaga¹

¹Uniwersytet im. Adama Mickiewicza w Poznaniu

Streszczenie

Dane funkcjonalne stały się w ciągu ostatnich dekad istotnym obszarem badań współczesnej statystyki matematycznej. Ich powstanie związane jest z wykonywaniem bardzo częstych pomiarów zmiennych ilościowych ciągłych w czasie lub przestrzeni, co zostało umożliwione przez postęp technologiczny. Ich przykładem może być mierzenie wilgotności powietrza codziennie przez rok lub notowanie cen akcji na giełdzie każdego dnia przez pewien okres czasu. Do istotnych problemów analizy danych funkcjonalnych należy problem badania różnorodności danych funkcjonalnych w zależności od warunków (grup), w jakich zostały otrzymane. W celu uzyskania odpowiedzi na tego typu pytania można stosować funkcjonalną analizę wariancji (FANOVA). W pracy rozważamy taką analizę w wieloczynnikowym układzie doświadczalnym z powtarzanymi pomiarami. W literaturze dla tego problemu rozważono metody bazujące na rozwinięciu bazowym. Metody te mają jednak pewne restrykcje - wymagają normalności danych oraz homogeniczność funkcji kowariancji, co w praktyce może okazać się trudne do spełnienia. Aby uniknąć takich problemów zastosowaliśmy metodę losowych projekcji wykorzystującą nieparametryczne testy statystyczne. Teoretyczne uzasadnienie poprawności tej metody można udowodnić korzystając z wcześniejszych wyników probabilistycznych. Własności skończenie próbkowe zostaną zbadane za pomocą badań symulacyjnych metodą Monte-Carlo, które wykazują lepsze własności testów projekcyjnych nad testami bazującymi na rozwinięciu bazowym. W szczególności porównane zostaną rozmiary i moce testów statystycznych. Ponadto, testy te zostaną zastosowane do rzeczywistego zbioru danych.

Słowa kluczowe: analiza danych funkcjonalnych, analiza wariancji, metoda losowych projekcji, powtarzane pomiary, weryfikacja hipotez statystycznych

O dyskryminowaniu bez stronniczości

Joanna Trzęsiok

Uniwersytet Ekonomiczny w Katowicach

Streszczenie

Przy dużym wzroście zastosowań metod eksploracji danych (Machine Learning) oraz szeroko rozumianej sztucznej inteligencji (AI), wzmógł się problem zaufania do modeli predykcyjnych. Zagadnienie to jest opisywane w różnych kontekstach, ponieważ możemy je definiować jako: (1) kontrolowanie jakości danych, (2) zapewnianie wyjaśnialności decyzji podejmowanych na podstawie tych modeli, ale też (3) zagwarantowanie spójności, transparentności i zgodności z odpowiednimi regulacjami w całym cyklu życia modelu. W obszarach zastosowań metod ML związanych np. z edukacją, procesami rekrutacyjnymi, podstawowymi usługami publicznymi i prywatnymi, jak choćby w ocenie zdolności kredytowej, kwestie te są również ściśle regulowane prawem Unii Europejskiej (EU Artificial Intelligence Act).

Jak więc sprawdzać i testować modele, czy spełniają wymóg zaufania nakładany na narzędzia analizy danych? Powszechna jest już świadomość, że nie wystarczy zbudować model o dobrych

zdolnościach predykcyjnych. Konieczne jest również zapewnienie, wspomnianej już wyjaśnialności wskazań tego modelu, ale też kontrolowanie jego odporności, transparentności oraz stronniczości. Stawia to przed badaczami wyzwania konstruowania nowych mierników oraz procedur testowania modeli predykcyjnych.

Referat będzie miał na celu przybliżenie pojęcia stronniczości modeli, rozumianego jako faworyzowanie lub odwrotnie – dyskryminowanie grupy obiektów. Przedstawione oraz porównane zostaną mierniki, nie opisywane szerzej w literaturze polskiej, pozwalające ocenić na ile modele dyskryminacyjne są stronnicze. Wyniki zostaną pokazane na rzeczywistych zbiorach danych metodami symulacyjnymi.

Słowa kluczowe: eksploracja danych, dyskryminacja, mierniki stronniczości

Realizacja celu SDG 4 w państwach UE w perspektywie roku docelowego 2030: wielowymiarowa analiza wskaźnikowa z wykorzystaniem podejścia hybrydowego

Marek Walesiak¹, Grażyna Dehnel²

¹*Uniwersytet Ekonomiczny we Wrocławiu*

²*Uniwersytet Ekonomiczny w Poznaniu*

Streszczenie

Celem referatu jest uszeregowanie 27 krajów UE w 2023 roku według ich postępów w realizacji celu SDG 4 „Zapewnić wszystkim inkluzywną i wysokiej jakości edukację oraz promować uczenie się przez całe życie” oraz wyznaczenie ich dystansu w stosunku do celu określonego na rok 2030. Eurostat postępy w realizacji celu SDG 4 monitoruje i ocenia za pomocą sześciu wskaźników. Na bazie tych wskaźników w artykule skonstruowano miarę agregatową według założeń podejścia hybrydowego łączącego skalowanie wielowymiarowe z porządkowaniem liniowym. Najmniejszy dystans w stosunku do celu wyznaczonego dla SDG 4 można zaobserwować w przypadku Finlandii, Niderlandów, Danii, Szwecji i Estonii. Najgorszą sytuację w aspekcie realizacji celu SDG 4 zaobserwować można dla Rumunii, Bułgarii, Cypru i Grecji.

Słowa kluczowe: jakość edukacji, zrównoważony rozwój, miary agregatowe, metoda hybrydowa, skalowanie wielowymiarowe

Uwzględnienie dodatkowych źródeł zmienności wyników badań przy ocenie istotności statystycznej efektów

Janusz Wątroba
StatSoft Polska

Streszczenie

Na ocenę wyników badań empirycznych ma zazwyczaj wpływ wiele różnorodnych czynników. Są one źródłem zmienności wyników. Umiejętne uwzględnienie potencjalnych źródeł zmienności umożliwia wykazanie postulowanych przez badacza efektów. Wymaga to poprawnego zaplanowania badania oraz przyjęcia odpowiedniego modelu analizy jego wyników. W referacie zostaną omówione i zilustrowane przykładami dwa modele: model analizy wariancji w układzie blokowym oraz analiza kowariancji. Dla uwidocznienia korzyści wynikających z zastosowania powyższych modeli każdy z prezentowanych przykładów zostanie w dwóch wariantach: bez uwzględnienia dodatkowych źródeł zmienności wyników oraz z ich uwzględnieniem.

Słowa kluczowe: źródła zmienności wyników, analiza wariancji w blokach, analiza kowariancji

O poprawności doboru i ważenia zmiennych diagnostycznych w procedurach taksonometrycznych

Kamil Zientarski
Uniwersytet Marii Curie-Skłodowskiej (UMCS)

Streszczenie

Celem referatu jest identyfikacja popularnych, lecz dyskusyjnych pod względem poprawności sposobów prowadzenia badań w ramach procedur taksonometrycznych oraz poddanie ich krytycznej analizie. Szczególną uwagę zwrócono na ich empiryczną weryfikację z wykorzystaniem jak najogólniejszych przykładów. Przegląd literatury fachowej wykazuje, iż problematyka metodyki badań taksonometrycznych nie jest częstym przedmiotem badań, pomimo bardzo dużego rozpowszechnienia samych metod; szczególnie w zakresie weryfikacji procedur często wykorzystywanych, a nie nowych propozycji metodycznych.

W niniejszym referacie analizowane będą dwa najczęściej wykorzystywane sposoby wyboru i ważenia zmiennych diagnostycznych: za pomocą współczynnika zmienności, a także współczynnika korelacji liniowej Pearsona. Weryfikowana będzie hipoteza badawcza brzmiąca – Współczynnik zmienności, a także współczynnik korelacji liniowej Pearsona jako wyłączne sposoby wyboru i ważenia zmiennych diagnostycznych nie są optymalne, skutkują bowiem utratą znaczącej części ładunku informacyjnego względem procedury bardziej holistycznej oraz zniekształcają wyniki agregacyjnych mierników.

Metodyka badań prezentowanych w referacie opierać się będzie na obliczeniu za pomocą metod taksonometrycznych syntetycznych mierników w czterech wariantach – w pierwszym z nich wyznaczanie oraz ważenie zmiennych diagnostycznych odbywać się będzie wyłącznie za pomocą współczynnika zmienności, w drugim wyłącznie za pomocą współczynnika korelacji liniowej Pearsona, w trzecim – za pomocą współczynnika zmienności i korelacji liniowej Pearsona oraz w czwartym – za pomocą autorskiej metody, uwzględniającej współczynnik zmienności, względnej amplitudy zmiennej, asymetrii oraz współczynnika korelacji liniowej Pearsona lub korelacji rang

Spearmana w zależności od wyników weryfikowanej hipotezy o rozkładzie normalnym (z uwzględnieniem statystycznej istotności korelacji).

Słowa kluczowe: metody taksonometryczne, dobór zmiennych, ważenie zmiennych, współczynnik zmienności, współczynnik korelacji liniowej Pearsona