

Uczenie maszynowe w analizie danych finansowych: przegląd literatury, metody i wyzwania

projekt realizowany w ramach przedmiotu *Metodologia badań naukowych*

1. Obszar badawczy i cel badania

Obszar i problem badawczy:

W dobie cyfryzacji i rosnącej złożoności danych, współczesne rynki finansowe generują ogromne wolumeny informacji, których skuteczna analiza wymaga zastosowania zaawansowanych metod obliczeniowych. Coraz większym zainteresowaniem środowisk akademickich oraz sektora finansowego cieszą się techniki uczenia maszynowego (machine learning), które znajdują zastosowanie m.in. w prognozowaniu cen aktywów, ocenie ryzyka kredytowego, wykrywaniu oszustw, analizie sentymentu rynkowego oraz automatyzacji procesów decyzyjnych. Problem badawczy niniejszej pracy koncentruje się na identyfikacji konkretnych metod uczenia maszynowego wykorzystywanych w analizie danych finansowych, określeniu kontekstów, w których przynoszą one najlepsze rezultaty, oraz wskazaniu głównych wyzwań towarzyszących ich implementacji – zarówno technicznych, jak i organizacyjnych czy etycznych.

Cel badania:

Głównym celem badania jest dokonanie przeglądu aktualnej literatury naukowej dotyczącej zastosowania metod uczenia maszynowego w analizie danych finansowych. W szczególności analizie poddane zostaną obszary zastosowań, najczęściej wykorzystywane algorytmy oraz pojawiające się w badaniach wyzwania technologiczne i normatywne. Celem uzupełniającym jest identyfikacja dominujących trendów badawczych oraz charakterystyka kierunków rozwoju tej dynamicznie rozwijającej się dziedziny.

2. Pytania badawcze

- Jakie metody uczenia maszynowego są najczęściej wykorzystywane w analizie danych finansowych według przeglądu literatury naukowej?
- W jakich obszarach finansów (np. detekcja oszustw, ocena ryzyka, predykcja cen) zastosowanie uczenia maszynowego jest najbardziej intensywne i skuteczne?
- Jakie wyzwania techniczne, etyczne lub interpretacyjne są identyfikowane w literaturze dotyczącej wykorzystania uczenia maszynowego w finansach?

3. Źródła danych do analizy literatury

- **Google Scholar**
Ogólnodostępna wyszukiwarka literatury naukowej obejmująca artykuły, książki i

preprinty. Pozwala na szybkie dotarcie do szerokiego zakresu publikacji, również spoza najważniejszych baz.

- **Scopus**

Multidyscyplinarna baza danych zawierająca recenzowane publikacje naukowe z zakresu m.in. finansów i sztucznej inteligencji. Umożliwia filtrowanie według cytowań, autorów i lat publikacji oraz oferuje narzędzia bibliometryczne.

4. Słowa kluczowe i szablon przeszukiwania

Wybrane pytanie badawcze:

Jakie metody uczenia maszynowego są najczęściej wykorzystywane w analizie danych finansowych według przeglądu literatury naukowej?

Słowa kluczowe:

- *machine learning*
- *financial data*
- *deep learning*
- *neural networks*
- *anomaly detection*
- *fraud detection*
- *financial forecasting*

Szablon wyszukiwania:

(machine learning OR deep learning OR neural networks) AND

(financial data OR finance) AND

(anomaly detection OR fraud detection OR forecasting)

5. Przeszukiwanie literatury w bazach danych

Google Scholar

Tabela 1. Zestawienie pięciu publikacji w Google Scholar wyszukanych w dniu 04.05.2025, wybór z ~19700 publikacji na podstawie dopasowania do tematu

L. p.	Opis bibliograficzny zgodnie ze stylem APA	Słowa kluczowe (z publikacji)	Uzyskane rezultaty lub wkład artykułu oraz cel artykułu i/lub zastosowane metody
1	Huang, W., Lai, K. K., Nakamori, Y., Wang, S., & Yu, L. (2007). Neural networks in finance and economics forecasting. <i>International Journal of Information Technology &</i>	<ul style="list-style-type: none">• Artificial neural networks• finance forecasting• economic forecasting• input variables selection	Celem pracy jest przegląd zastosowań sztucznych sieci neuronowych (ANN) w prognozowaniu kursów walut, indeksów giełdowych oraz wzrostu gospodarczego. Autorzy analizują rodzaje wejść danych (jedno- i wielowymiarowe), typy modeli (MLP, RNN, GRNN, PNN) oraz porównują ich skuteczność z metodami klasycznymi.

	<i>Decision Making</i> , 6(01), 113-140.	<ul style="list-style-type: none"> • performance comparisons 	Artykuł wskazuje, że sieci neuronowe często przewyższają modele liniowe, zwłaszcza przy danych nieliniowych lub hałaśliwych. Opisano także integracje z innymi technologiami (algorytmy genetyczne, analiza falkowa, logika rozmyta) oraz hybrydowe systemy decyzyjne, podkreślając skuteczność modeli kombinowanych w prognozowaniu finansowym.
2	Ashtiani, M. N., & Raahemi, B. (2021). Intelligent fraud detection in financial statements using machine learning and data mining: a systematic literature review. <i>Ieee Access</i> , 10, 72504-72525.	<ul style="list-style-type: none"> • Fraud detection • financial statement • machine learning • data mining • outlier detection • systematic literature review 	<p>Celem artykułu jest systematyczny przegląd 47 badań nad wykrywaniem oszustw w sprawozdaniach finansowych przy użyciu metod uczenia maszynowego i eksploracji danych. Autorzy klasyfikują techniki (m.in. SVM, regresja logistyczna, sieci neuronowe, metody zespołowe) oraz analizują wykorzystywane typy danych (wskaźniki finansowe, dane tekstowe, dane głosowe).</p> <p>Artykuł identyfikuje dominację podejść nadzorowanych i podkreśla rosnące znaczenie danych nienumerycznych. Wskazuje luki badawcze: niski udział metod nienadzorowanych i semi-nadzorowanych oraz niewielkie wykorzystanie danych tekstowych i głosowych, sugerując dalsze badania z użyciem danych nieustrukturyzowanych i algorytmów heurystycznych.</p>
3	Parimi, S. S. (2017). Leveraging Deep Learning for Anomaly Detection in SAP Financial Transactions. <i>Available at SSRN</i> 4934907.	<ul style="list-style-type: none"> • Anomaly detection • machine learning algorithms • deep learning integration • SAP Financial transactions 	<p>Celem pracy jest ocena zastosowania głębokiego uczenia (autoenkodery, RNN, CNN) do wykrywania anomalii w transakcjach finansowych systemów SAP. Autor porównuje te podejścia z metodami tradycyjnymi (reguły, statystyka), wskazując ich lepszą zdolność do adaptacji, wykrywania subtelnych wzorców i pracy na danych sekwencyjnych.</p> <p>Artykuł systematyzuje aktualne podejścia deep learningowe do detekcji anomalii w środowiskach SAP i wskazuje wyzwania wdrożeniowe: interpretowalność modeli,</p>

			skalowalność, jakość danych oraz adaptację do zmiennych warunków. Zidentyfikowano trendy takie jak federated learning, analiza danych multimodalnych oraz wdrożenia w czasie rzeczywistym z użyciem edge computingu.
4	Singla, A., & Jangir, H. (2020, February). A comparative approach to predictive analytics with machine learning for fraud detection of realtime financial data. <i>In 2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3)</i> (pp. 1-4). IEEE.	<ul style="list-style-type: none"> • Predictive Analytics • Artificial Intelligence • Machine learning • streaming data • real-time applications • deep learning 	<p>Celem artykułu jest porównanie różnych metod uczenia maszynowego i analityki predykcyjnej w kontekście wykrywania oszustw finansowych w czasie rzeczywistym. Autorzy analizują m.in. techniki głębokiego uczenia, drzewa decyzyjne oraz algorytmy wykrywania anomalii, zwracając uwagę na znaczenie inżynierii cech i doboru parametrów.</p> <p>Praca identyfikuje najbardziej efektywne podejścia do detekcji oszustw na podstawie danych strumieniowych, podkreśla rolę automatyzacji i niskich wskaźników fałszywych alarmów oraz proponuje wykorzystanie metod ML jako wsparcie w podejmowaniu decyzji biznesowych w sektorze finansowym.</p>
5	Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., ... & Saif, A. (2022). Financial fraud detection based on machine learning: a systematic literature review. <i>Applied Sciences</i> , 12(19), 9637.	<ul style="list-style-type: none"> • financial fraud • fraud detection • machine learning • data mining • systematic literature review • Kitchenham approach 	<p>Celem pracy jest systematyczny przegląd 93 badań nad wykrywaniem oszustw finansowych z wykorzystaniem metod uczenia maszynowego (ML). Autorzy stosują metodologię Kitchenham, analizując techniki ML (SVM, ANN, Naive Bayes, HMM, drzewa decyzyjne, regresję, algorytmy genetyczne, ensemble) oraz klasyfikują typy oszustw (karty kredytowe, sprawozdania finansowe, cyberoszustwa, ubezpieczenia).</p> <p>Artykuł pokazuje, że dominującymi technikami są metody nadzorowane (szczególnie Naive Bayes, SVM, ANN), a najczęściej badanym typem oszustwa jest fraud kartowy. Zidentyfikowano luki w zastosowaniu metod nienadzorowanych, danych nieustrukturyzowanych i w</p>

			uwzględnianiu konceptu <i>concept drift</i> . Autorzy sugerują kierunki przyszłych badań: eksploracja tekstów finansowych (np. MD&A), uczenie transferowe, aktywne i przyrostowe oraz techniki wykrywania anomalii.
--	--	--	---

Źródło: opracowanie własne

Tabela 2. Zestawienie pięciu publikacji w Google Scholar wyszukanych w dniu 04.05.2025, wybór z ~17400 publikacji z lat 2019-2025 na podstawie dopasowania do tematu

L. p.	Opis bibliograficzny zgodnie ze stylem APA	Słowa kluczowe (z publikacji)	Uzyskane rezultaty lub wkład artykułu oraz cel artykułu i/lub zastosowane metody
1	Olubusola, O., Mhlongo, N. Z., Daraojimba, D. O., Ajayi-Nifise, A. O., & Falaiye, T. (2024). Machine learning in financial forecasting: A US review: Exploring the advancements, challenges, and implications of AI-driven predictions in financial markets. <i>World Journal of Advanced Research and Reviews</i> , 21(2), 1969-1984.	<ul style="list-style-type: none"> Artificial Intelligence Finance Machine Learning Financial Forecasting Financial Market 	<p>Techniki zastosowane w artykule: Linear Regression, Random Forest Regression, Support Vector Regression, Deep Learning, LSTM (Long Short-Term Memory), Transformers, Uczenie nadzorowane i nienadzorowane (k-NN, Auto-ARIMA), Explainable AI (LIME, SHAP), Automatyczne modelowanie (AutoML) i ensemble models (np. XGBoost + Random Forest).</p> <p>Wyzwania wymienione przez autorów: Jakość danych, Interpretowalność modeli (black-box), Przeuczenie (overfitting), Wymogi regulacyjne i etyczne, Złożoność integracji z klasycznymi modelami finansowymi</p> <p>Techniki uczenia maszynowego opisane w artykule dają najlepsze rezultaty w kontekstach, gdzie dostępne są duże zbiory danych, wyraźne wzorce historyczne i gdzie szybkość analizy ma kluczowe znaczenie</p>
2	Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. <i>Applied soft</i>	brak w publikacji	<p>Techniki zastosowane w artykule: LSTM, RNN, GRU, CNN, DMPLP, autoenkodery, DBN, RBM oraz uczenie przez wzmocnienie</p> <p>Wyzwania wymienione przez</p>

	<i>computing</i> , 93, 106384.		<p>autorów: Brak interpretowalności modeli, nadmierne dopasowanie (overfitting), ograniczona dostępność i jakość danych, wysokie koszty obliczeniowe oraz trudność w zastosowaniu modeli w rzeczywistych warunkach rynkowych</p> <p>Techniki opisane w artykule dają najlepsze rezultaty w kontekstach takich jak: prognozowanie cen aktywów i indeksów giełdowych (szeregi czasowe), wykrywanie oszustw finansowych, scoring kredytowy, zarządzanie portfelem inwestycyjnym oraz analiza sentymentu na podstawie danych tekstowych i wiadomości rynkowych.</p>
3	Olowe, K. J., Edoh, N. L., Zouo, S. J. C., & Olamijuwon, J. Review of predictive modeling and machine learning applications in financial service analysis.	<ul style="list-style-type: none"> • Predictive Modeling • Machine Learning • Financial Service Analysis • Review 	<p>Techniki zastosowane w artykule: regresja logistyczna, drzewa decyzyjne, las losowy (Random Forest), SVM, sieci neuronowe, uczenie głębokie, uczenie przez wzmacnianie, klasteryzacja (k-means) oraz analiza szeregów czasowych</p> <p>Wyzwania wymienione przez autorów: niska jakość danych, trudności z integracją danych z różnych źródeł, brak interpretowalności modeli, ryzyko uprzedzeń algorytmicznych, problemy z regulacjami (np. RODO) oraz trudności w skalowaniu modeli ML w instytucjach finansowych.</p> <p>Techniki i metody opisane w artykule dają najlepsze rezultaty w kontekstach takich jak: ocena ryzyka kredytowego, wykrywanie oszustw finansowych, segmentacja i personalizacja klientów, prognozowanie rynków finansowych, zarządzanie portfelem inwestycyjnym oraz</p>

			algorytmiczny handel w czasie rzeczywistym.
4	Cao, L. (2020). AI in finance: A review. <i>Available at SSRN, 3647625, 1.</i>	<ul style="list-style-type: none"> • Artificial intelligence • AI • data science • data analytics • advanced analytics • machine learning • FinTech • Finance • Economics • AI in FinTech • AI in finance 	<p>Techniki zastosowane w artykule: Uczenie głębokie (deep learning), sieci neuronowe (RNN, CNN, LSTM), regresja, drzewa decyzyjne, reinforcement learning, analiza klastrow i modelowanie agentowe.</p> <p>Wyzwania wymienione przez autorów: Trudność modelowania złożonych zjawisk finansowych, niska przejrzystość modeli AI (black-box), problemy z jakością i dynamiką danych, ograniczona interpretowalność wyników oraz wyzwania etyczne i regulacyjne.</p> <p>Techniki opisane w artykule dają najlepsze rezultaty w kontekstach takich jak: prognozowanie rynków finansowych, zarządzanie portfelem, scoring kredytowy, wykrywanie oszustw, analiza zachowań klientów, robo-doradztwo, oraz analiza sentymentu i danych alternatywnych (np. z mediów społecznościowych)</p>
5	Maple, C., Szpruch, L., Epiphaniou, G., Staykova, K., Singh, S., Penwarden, W., ... & Avramovic, P. (2023). The ai revolution: opportunities and challenges for the finance sector. <i>arXiv preprint arXiv:2308.16538.</i>	brak w publikacji	<p>Techniki zastosowane w artykule: uczenie nadzorowane, uczenie nienadzorowane, uczenie głębokie (deep learning), reinforcement learning, sieci neuronowe, a także metody takie jak NLP i analiza szeregów czasowych.</p> <p>Wyzwania wymienione przez autorów: brak przejrzystości modeli AI (black-box), niskiej jakości lub stronnicze dane, ryzyko systemowe i etyczne, brak regulacji lub ich niedopasowanie, problemy z integracją AI w istniejącej infrastrukturze, a także zagrożenie wykluczeniem zawodowym i społecznym.</p> <p>Techniki i metody opisane w</p>

			artykule dają najlepsze rezultaty w kontekstach takich jak: algorytmiczny handel, zarządzanie ryzykiem finansowym, wykrywanie oszustw, automatyzacja procesów decyzyjnych, ocena zdolności kredytowej, analiza sentymentu rynkowego oraz personalizacja usług finansowych.
--	--	--	---

Źródło: opracowanie własne

Scopus

Tabela 3. Zestawienie pięciu wybranych publikacji w Scopus wyszukanych w dniu 05.05.2025, wybór z 68035 publikacji na podstawie liczby cytowań.

L. p.	Opis bibliograficzny zgodnie ze stylem APA	Słowa kluczowe (z publikacji)	Uzyskane rezultaty lub wkład artykułu oraz cel artykułu i/lub zastosowane metody
1	He, H., & Garcia, E. A. (2009). Learning from imbalanced data. <i>IEEE Transactions on knowledge and data engineering</i> , 21(9), 1263-1284.	<ul style="list-style-type: none"> • Active learning • Assessment metrics • Classification; • Cost-sensitive learning • Imbalanced learning • Kernel-based learning • Sampling methods 	<p>Celem pracy jest przegląd metod radzenia sobie z problemem nie zrównoważonych danych w uczeniu maszynowym. Omawiane techniki to m.in. oversampling (np. SMOTE), uczenie kosztowo-wrażliwe, metody oparte na jądrach oraz uczenie aktywne. Poruszono także temat doboru odpowiednich metryk, takich jak AUC czy krzywa precyzja-recall.</p> <p>Autorzy systematyzują istniejące podejścia do uczenia z nie zrównoważonych danych, identyfikują wyzwania oraz wskazują kierunki dalszych badań, szczególnie w kontekście zastosowań w finansach i dużych systemach danych.</p>
2	Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and	<ul style="list-style-type: none"> • Accountability • Comprehensibility • Data Fusion • Deep Learning • Explainable Artificial Intelligence • Fairness • Interpretability • Machine Learning 	<p>Celem pracy jest przedstawienie definicji, klasyfikacji i przeglądu metod wyjaśnialnej sztucznej inteligencji (XAI), ze szczególnym uwzględnieniem technik post-hoc dla złożonych modeli takich jak sieci neuronowe. Autorzy proponują taksonomie XAI oraz omawiają poziomy przejrzystości modeli ML, w tym techniki takie jak</p>

	<p>challenges toward responsible AI. <i>Information fusion</i>, 58, 82-115.</p>	<ul style="list-style-type: none"> • Privacy • Responsible Artificial Intelligence • Transparency 	<p>LIME, SHAP, reguły decyzyjne i wizualizacje.</p> <p>Artykuł systematyzuje pojęcia XAI, wskazuje na kluczową rolę wyjaśnialności w kontekście AI odpowiedzialnej (Responsible AI) i identyfikuje wyzwania badawcze, m.in. brak jednolitych metryk i problem zależności od odbiorcy (audience-dependence) przy interpretacji modeli.</p>
3	<p>Kuhn, M., & Johnson, K. (2013). <i>Applied predictive modeling</i> (Vol. 26, p. 13). New York: Springer.</p>	<p>(ze strony internetowej wydawnictwa)</p> <ul style="list-style-type: none"> • Model • Non-Linear • Predictive Models • R • Regression Models • Regression Trees 	<p>Publikacja przedstawia kompletny, praktyczny przewodnik po procesie budowy modeli predykcyjnych, ze szczególnym uwzględnieniem przetwarzania danych, doboru i strojenia modeli, oceny wydajności oraz przeciwdziałania przeuczeniu. Autorzy wykorzystują szeroki zakres metod: regresję liniową, modele penalizowane, drzewa decyzyjne, sieci neuronowe, SVM, boosting, random forests itp., wraz z ich implementacją w języku R.</p> <p>Książka systematyzuje cały proces modelowania predykcyjnego i stanowi praktyczny podręcznik dla analityków danych. Zawiera przykłady z obszarów takich jak finanse, farmacja czy przetwarzanie sygnałów biologicznych.</p>
4	<p>Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. <i>International journal of forecasting</i>, 14(1), 35-62.</p>	<p>(ze strony internetowej wydawnictwa)</p> <ul style="list-style-type: none"> • Neural Networks • Forecasting 	<p>Celem pracy jest przegląd zastosowań sztucznych sieci neuronowych (ANN) w prognozowaniu. Autorzy omawiają różne architektury sieci (głównie MLP), metody treningu (backpropagation, BFGS, Levenberg–Marquardt), normalizację danych oraz krytyczne kwestie modelowania, takie jak dobór liczby wejść, warstw ukrytych i funkcji aktywacji.</p>

			Artykuł systematyzuje dotychczasowe badania nad wykorzystaniem ANN w prognozowaniu, porównuje ich skuteczność z metodami statystycznymi (np. ARIMA), i identyfikuje kluczowe wyzwania — m.in. brak standardów w projektowaniu sieci oraz nadmierne dopasowanie modeli. Zaproponowano również kierunki dalszych badań, w tym łączenie ANN z klasycznymi modelami.
5	Lütkepohl, H. (2005). New introduction to multiple time series analysis. <i>Springer Science & Business Media</i> .	(ze strony internetowej wydawnictwa) <ul style="list-style-type: none"> • Analysis • Dynamic Econometric Modeling • Forecasting • Multiple Time Series • Multiple Time Series Analysis • Regression • Time Series Analysis • model • simulation • statistics 	<p>Celem publikacji jest przedstawienie nowoczesnego i kompletnego ujęcia analizy wielowymiarowych szeregów czasowych (VAR, VECM, SVAR, VARMA). Autor omawia metody estymacji, testowania, modelowania współzależności, analizy przyczynowości Grangera, dekompozycji błędów prognozy oraz analizy impulsowo-odpowiedziowej.</p> <p>Książka stanowi rozszerzoną i zaktualizowaną wersję wcześniejszej pracy autora, wprowadzając m.in. pełną analizę kointegracji w modelach VAR/VECM, estymację bayesowską oraz modele MGARCH. Zawiera także liczne przykłady empiryczne, procedury diagnostyczne i materiał do zastosowania w ekonomii i finansach.</p>

Źródło: opracowanie własne

Tabela 4. Zestawienie pięciu wybranych publikacji w Scopus wyszukanych w dniu 05.05.2025, wybór z 53891 publikacji z lat 2019-2025 na podstawie liczby cytowań.

L. p.	Opis bibliograficzny zgodnie ze stylem APA	Słowa kluczowe (z publikacji)	Uzyskane rezultaty lub wkład artykułu oraz cel artykułu i/lub zastosowane metody
-------	--	-------------------------------	--

1	<p>Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. <i>International journal of information management</i>, 57, 101994.</p>	<ul style="list-style-type: none"> • AI • Artificial intelligence • Cognitive computing • Expert systems • Machine learning • Research agenda 	<p>Celem artykułu jest przedstawienie przekrojowego przeglądu zastosowań, wyzwań i szans związanych z wdrażaniem AI z perspektywy różnych dyscyplin (biznes, sektor publiczny, prawo, technologia, edukacja). Autorzy analizują wpływ AI na decyzje, zatrudnienie, gospodarkę i etykę, proponując agendę badawczą uwzględniającą aspekty społeczne, ekonomiczne, organizacyjne, technologiczne i legislacyjne.</p> <p>Artykuł integruje opinie ponad 40 ekspertów i wskazuje główne wyzwania (m.in. brak przejrzystości algorytmów, zagrożenia etyczne, luki legislacyjne) oraz możliwości (automatyzacja, analiza big data, decyzje wspomagane AI) i podkreśla konieczność zrównoważonego i odpowiedzialnego rozwoju sztucznej inteligencji.</p>
2	<p>Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. <i>Telematics and informatics</i>, 36, 55-81.</p>	<ul style="list-style-type: none"> • Applications • Blockchain • Classification 	<p>Celem pracy jest systematyczny przegląd zastosowań technologii blockchain w różnych sektorach (m.in. finanse, zdrowie, IoT, administracja publiczna). Autorzy klasyfikują aplikacje blockchainowe według obszarów zastosowań oraz analizują ograniczenia i wyzwania związane z ich wdrożeniem. Przegląd opiera się na analizie 260 publikacji naukowych i raportów.</p> <p>Artykuł dostarcza obszernej klasyfikacji zastosowań blockchain, wskazuje na dynamiczny wzrost zainteresowania tą technologią oraz identyfikuje luki badawcze, takie jak brak standaryzacji, ograniczenia skalowalności i trudności w doborze odpowiednich mechanizmów konsensusu dla różnych domen.</p>
3	<p>Brunton, S. L., & Kutz, J. N. (2022). Data-driven science and engineering: Machine</p>	<p>(ze strony internetowej wydawnictwa)</p> <ul style="list-style-type: none"> • Mathematics 	<p>Celem książki jest połączenie metod uczenia maszynowego z klasycznymi technikami analizy układów dynamicznych i</p>

	learning, dynamical systems, and control. <i>Cambridge University Press.</i>	<ul style="list-style-type: none"> • Computational Science • Engineering • Control Systems • Optimisation 	<p>sterowania. Autorzy prezentują nowoczesne podejście do analizy danych naukowych i inżynierskich, oparte na takich metodach jak PCA/SVD, regresja, sieci neuronowe, dynamiczna dekompozycja modalna (DMD) czy systemy rzadkie (sparse identification of dynamics - SINDy).</p> <p>Publikacja dostarcza spójnego frameworku do analizy i modelowania systemów dynamicznych z danych pomiarowych. Łączy teorię z licznymi przykładami praktycznymi (w tym kodem w MATLAB/Pythonie), co czyni ją wartościowym źródłem wiedzy zarówno dla naukowców, jak i praktyków pracujących z danymi czasowymi, predykcyjnymi i systemami sterowania.</p>
4	Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020, August). Connecting the dots: Multivariate time series forecasting with graph neural networks. <i>In Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining</i> (pp. 753-763).	<ul style="list-style-type: none"> • graph neural networks • graph structure learning • multivariate time series forecasting • spatial-temporal graphs 	<p>Celem pracy jest opracowanie modelu MTGNN - frameworku do prognozowania wielowymiarowych szeregów czasowych z wykorzystaniem grafowych sieci neuronowych (Graph Neural Networks), który automatycznie uczy się struktury grafu na podstawie danych. Model integruje warstwę uczenia grafu, propagację wieloskokową (mix-hop), konwolucje czasowe i uczy się w sposób end-to-end.</p> <p>MTGNN osiąga lepsze lub porównywalne wyniki względem istniejących metod na 6 benchmarkowych zbiorach danych bez potrzeby znajomości wcześniej zdefiniowanej struktury grafu. Autorzy wprowadzają też strategię uczenia progresywnego (curriculum learning) i pokazują, że model dobrze radzi sobie z dużą liczbą zmiennych i długimi sekwencjami.</p>
5	Van Houdt, G., Mosquera, C., &	<ul style="list-style-type: none"> • Deep learning 	Celem pracy jest kompleksowy przegląd modelu LSTM - jego

	Nápoles, G. (2020). A review on the long short-term memory model. <i>Artificial Intelligence Review</i> , 53(8), 5929-5955.	<ul style="list-style-type: none"> • Long short-term memory • Recurrent neural networks • Vanishing/exploding gradient 	<p>architektury, treningu, wariantów oraz szerokiego spektrum zastosowań, takich jak prognozowanie szeregów czasowych, NLP, analiza obrazów i rozpoznawanie mowy. Artykuł zawiera także praktyczny przykład implementacji LSTM w TensorFlow.</p> <p>Artykuł systematyzuje wiedzę o LSTM jako jednej z najskuteczniejszych architektur do modelowania danych sekwencyjnych. Pokazuje, że w wielu zastosowaniach (np. finanse, zdrowie, transport) model ten przewyższa klasyczne podejścia, oraz podkreśla jego elastyczność i skuteczność w hybrydowych architekturach (np. CNN-LSTM, BiLSTM).</p>
--	---	---	--

Źródło: opracowanie własne

6. Analiza literatury cytowanej

Do dalszej analizy literatury zostały wykorzystane artykuły z Tabeli 2, bibliografia tych artykułów prezentuje szeroki zakres źródeł w temacie uczenia maszynowego, liczba źródeł w bibliografiach wynosi od 38 do 202.

Tabela 5. Wybrane źródła zawarte w bibliografiach artykułów z Tabeli 2.

L. p.	Nr publ., z której pochodzi pozycja	Opis bibliograficzny publikacji zgodnie ze stylem APA	Słowa kluczowe (z publikacji)	Uzyskane rezultaty lub wkład artykułu oraz cel artykułu i/lub zastosowane metody
1	2	Sezer, O. B., Ozbayoglu, M., & Dogdu, E. (2017). A deep neural-network based stock trading system based on evolutionary optimized technical analysis parameters. <i>Procedia computer science</i> , 114, 473-480.	<ul style="list-style-type: none"> • Stock Trading • Stock Market • Deep Neural-Network • Evolutionary Algorithms • Technical Analysis 	Celem pracy było zaprojektowanie zautomatyzowanego systemu handlu akcjami, który łączy głębokie sieci neuronowe (MLP) z algorytmami genetycznymi (GA) w celu optymalizacji parametrów analizy technicznej. Artykuł

				wnosi istotny wkład w dziedzinę zautomatyzowanego handlu i inteligentnej optymalizacji wskaźników technicznych, pokazując, że hybrydowe podejście może prowadzić do skutecznych strategii inwestycyjnych.
2	1	Ahuja, R., Kumar, Y., Goyal, S., Kaur, S., Sachdeva, R. K., & Solanki, V. (2023, March). Stock Price Prediction By Applying Machine Learning Techniques. <i>In 2023 International Conference on Emerging Smart Computing and Informatics (ESCI)</i> (pp. 1-5). IEEE.	<ul style="list-style-type: none"> • Linear Regression • Support Vector Regressor • Random Forest Regressor • Mean Absolute error • Stock Market • Root Mean Squared Error • Knowledge Engineering 	Celem publikacji było opracowanie systemu predykcji cen akcji z wykorzystaniem modelu LSTM (Long Short-Term Memory) oraz regresji liniowej, aby umożliwić inwestorom podejmowanie trafniejszych decyzji inwestycyjnych. Publikacja prezentuje skuteczne połączenie modelu LSTM i regresji liniowej do predykcji cen akcji, pokazując, że taka hybrydowa metoda może zwiększyć dokładność prognoz i stanowić solidną podstawę dla inteligentnych systemów wspomagania decyzji inwestycyjnych.
3	1	Cohen, G. (2022). Algorithmic trading and financial forecasting using advanced artificial intelligence methodologies. <i>Mathematics</i> , 10(18), 3302.	<ul style="list-style-type: none"> • algorithmic trading • fuzzy systems • LSTM • Forecasting • Artificial Intelligence 	Artykuł ma na celu przedstawienie nowoczesnych metod sztucznej inteligencji wykorzystywanych w algorytmicznym handlu i prognozowaniu finansowym, takich jak sieci neuronowe (ANN, LSTM), SVM, systemy rozmyte czy analiza sentymentu. Autor

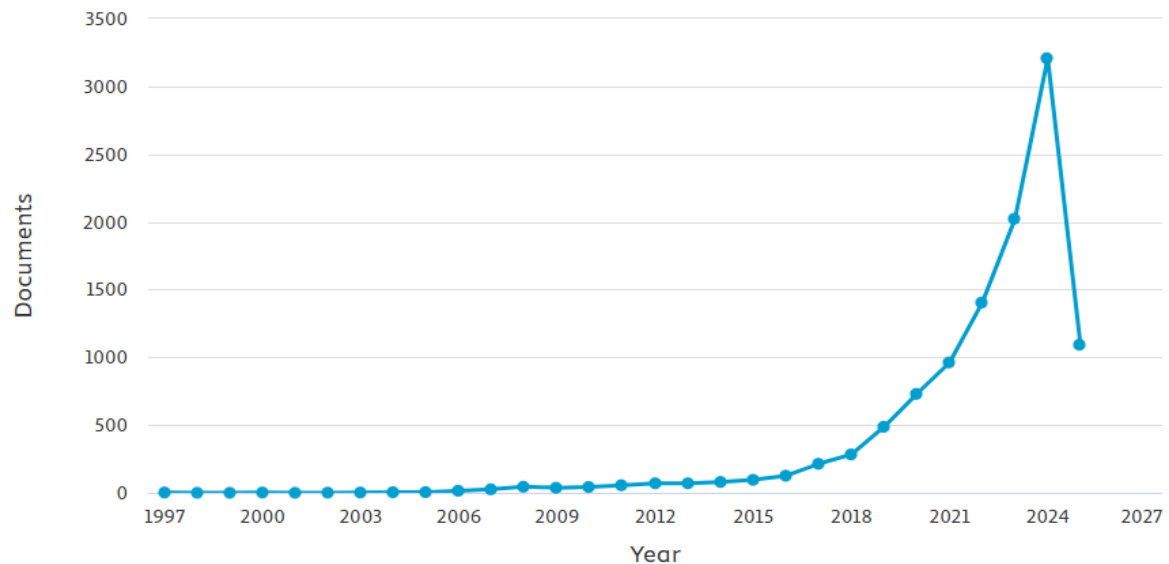
				<p>pokazuje, że integracja danych rynkowych, fundamentalnych i społecznych pozwala tworzyć bardziej trafne modele predykcyjne, które mogą przewyższać tradycyjne strategie inwestycyjne, co stanowi istotny wkład w rozwój inteligentnych systemów wspomagania decyzji finansowych.</p>
4	4	<p>Albashrawi, M. (2016). Detecting financial fraud using data mining techniques: A decade review from 2004 to 2015. <i>Journal of Data Science</i>, 14(3), 553-569.</p>	<ul style="list-style-type: none"> • Financial fraud • fraud detection • data mining techniques • literature review 	<p>Artykuł przegląda techniki data mining stosowane do wykrywania oszustw finansowych w latach 2004–2015, klasyfikując je według typu oszustwa i skuteczności. Celem było ocenienie 41 metod (np. regresji, SVM, drzew decyzyjnych, sieci neuronowych), a wkładem – stworzenie systematycznego przeglądu i zaleceń, które techniki najlepiej sprawdzają się w danym kontekście.</p>
5	4	<p>Baldominos, A., Blanco, I., Moreno, A. J., Iturrarte, R., Bernárdez, Ó., & Afonso, C. (2018). Identifying real estate opportunities using machine learning. <i>Applied sciences</i>, 8(11), 2321.</p>	<ul style="list-style-type: none"> • real estate • Appraisal • Investment • machine learning • artificial intelligence 	<p>Artykuł prezentuje porównanie 4 metody ML w celu identyfikacji nieruchomości poniżej wartości rynkowej, wskazując, że najskuteczniejsze są zespoły drzew regresyjnych. Wkładem pracy jest pokazanie, że ML może wspierać inwestorów w ocenie okazji inwestycyjnych na rynku nieruchomości.</p>

Źródło: opracowanie własne

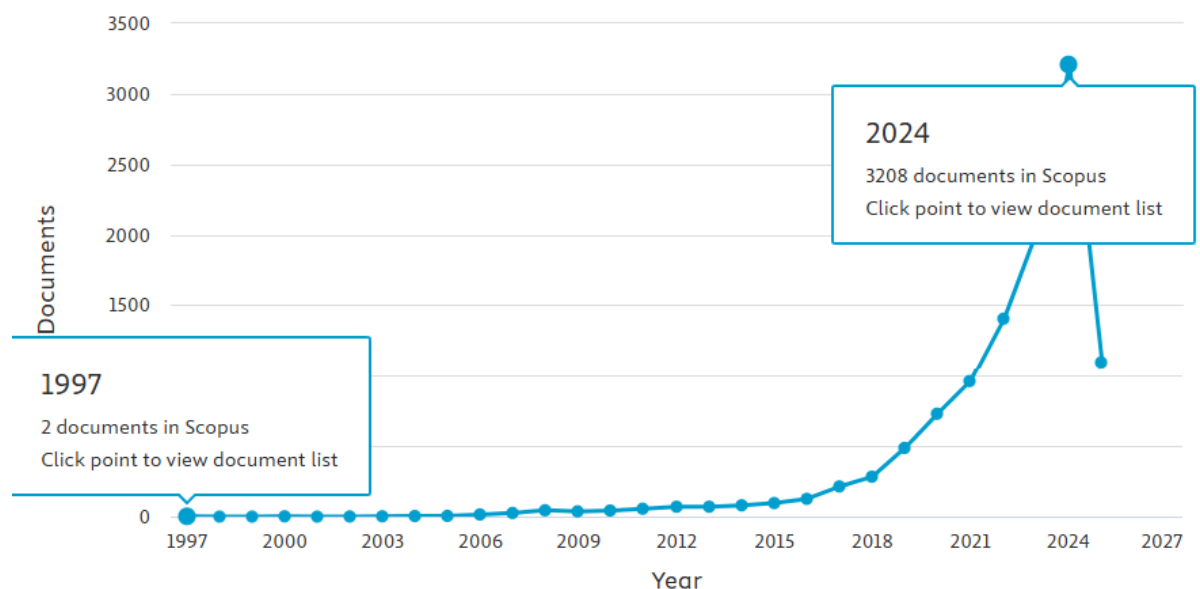
7. Wizualizacja w bazie Scopus

Wykres 1: Wizualizacja lat publikacji z szczegółowym przedstawieniem roku pierwszej publikacji i ilości publikacji w tym roku oraz roku z największą ilością publikacji

Documents by year



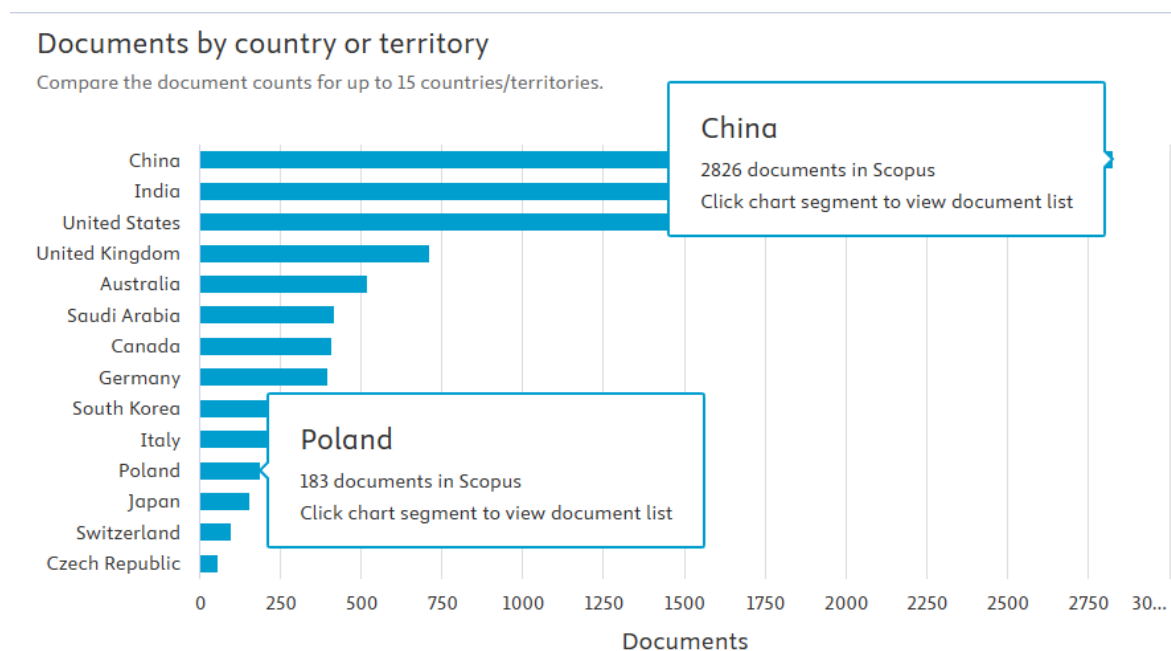
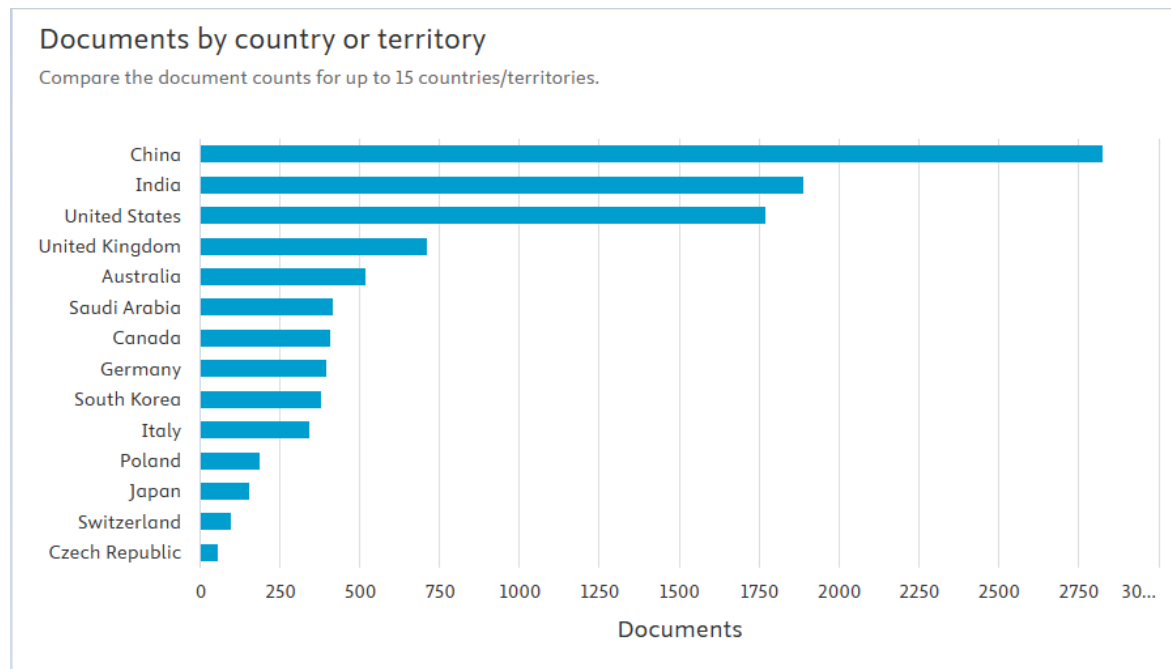
Documents by year



Analiza liczby publikacji w latach 1997–2025 wskazuje na wyraźny wzrost zainteresowania tematyką wykrywania anomalii w danych finansowych z wykorzystaniem uczenia maszynowego i głębokiego. Pierwsze publikacje pojawiły się w 1997 roku. Początkowo, w latach 1997–2017, liczba publikacji była stosunkowo niska, co może świadczyć o wczesnym etapie rozwoju tej dziedziny. Od 2018 roku obserwuje się stopniowy wzrost liczby publikacji, osiągając szczyt w latach 2022–2025, a najwięcej pojawiło się w 2024 roku, bo aż 3208 publikacji. Ten trend może

być związany z rosnącą dostępnością danych finansowych, postępem w technologiach uczenia maszynowego oraz zwiększonym zapotrzebowaniem na skuteczne metody wykrywania oszustw finansowych.

Wykres 2. Wizualizacja ilości publikacji ze względu na kraj

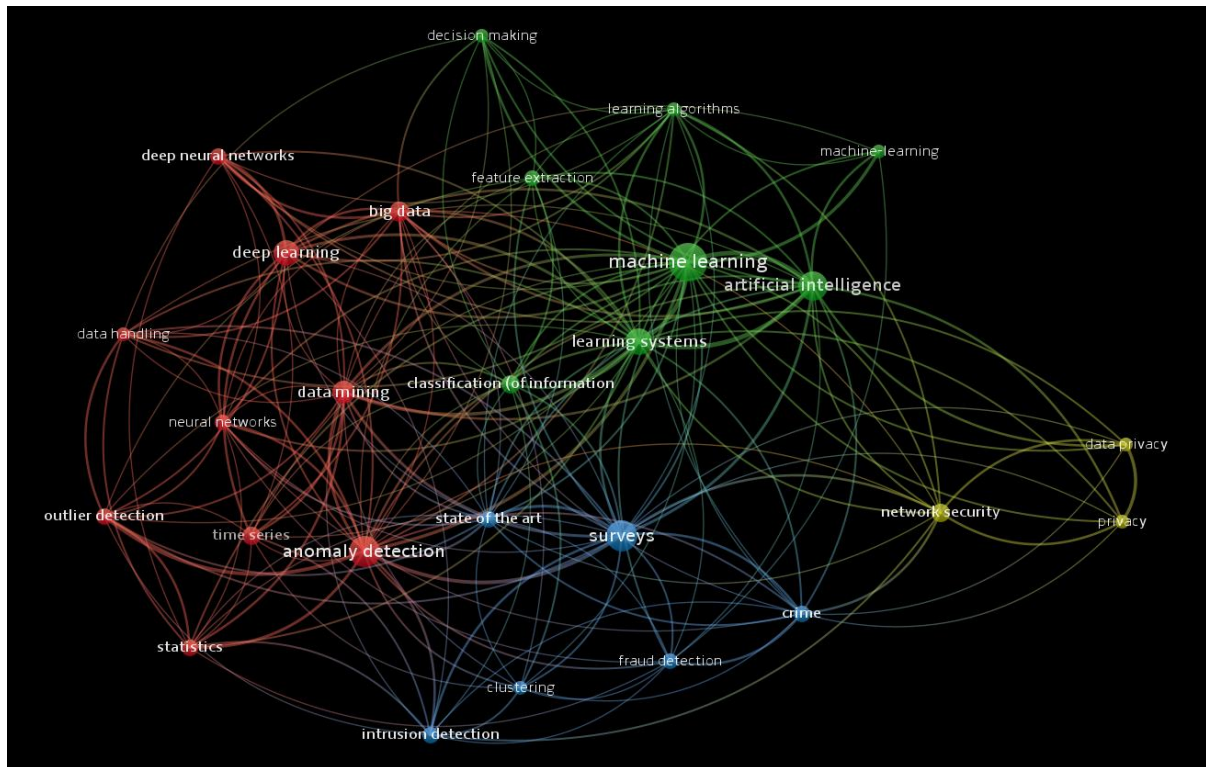


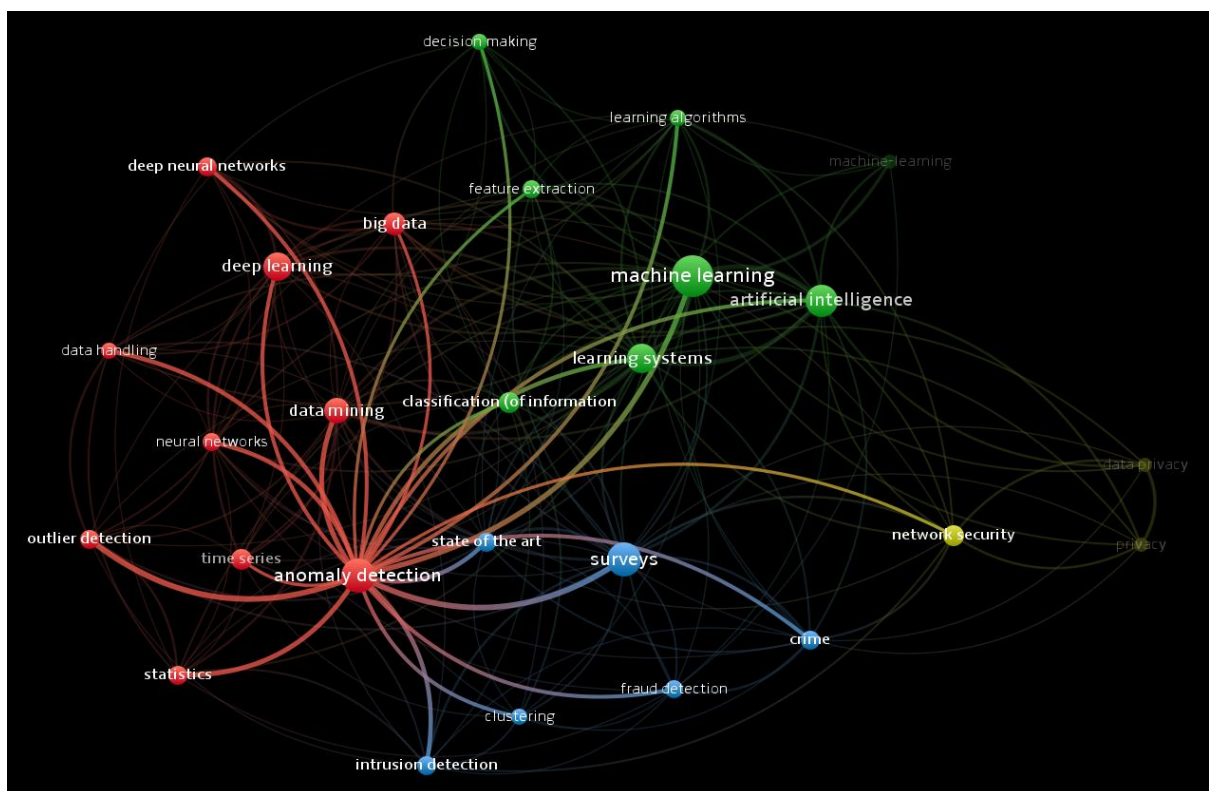
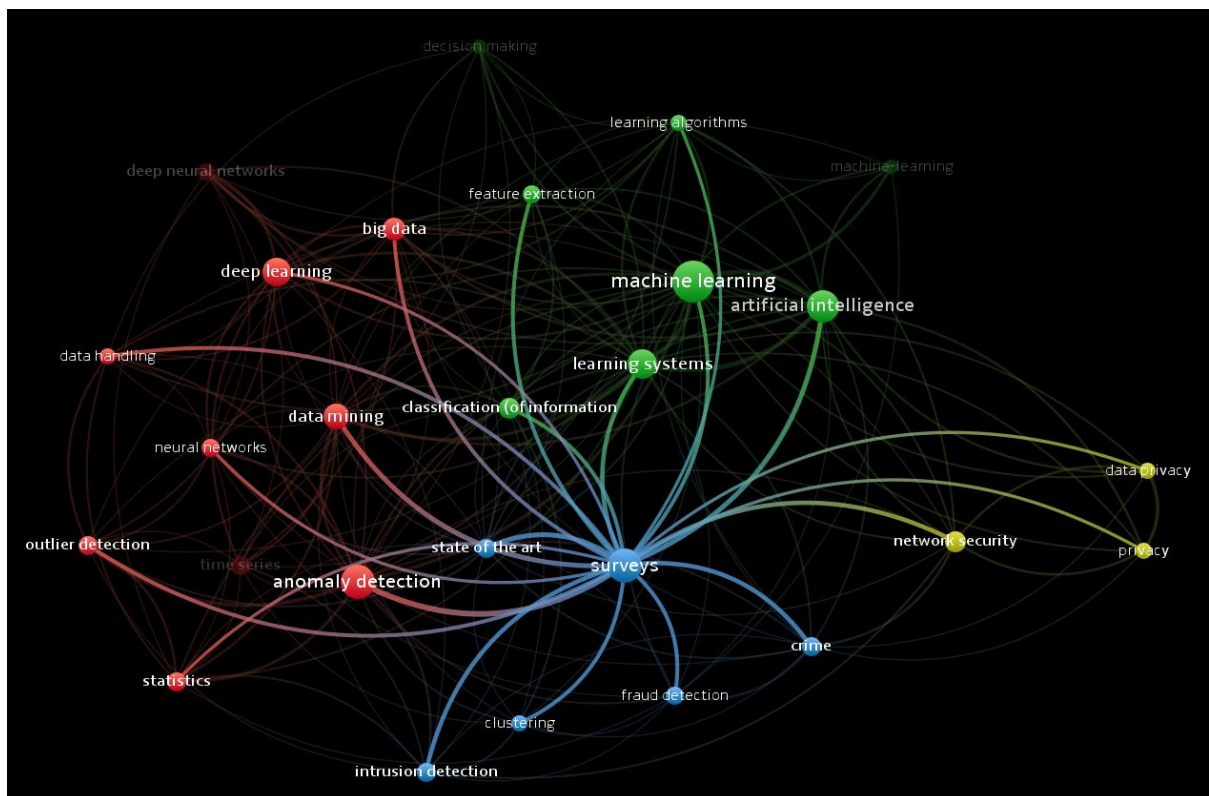
Analiza geograficznego pochodzenia autorów publikacji w zakresie wykrywania anomalii w danych finansowych przy użyciu metod uczenia maszynowego pokazuje wyraźną dominację Chin (2826 dokumentów) i Indii. Stany Zjednoczone również zajmują wysoką pozycję, co potwierdza ich silną pozycję w badaniach nad sztuczną inteligencją. Polska znajduje się niżej w rankingu z 183 publikacjami, co jednak wskazuje na rosnące zainteresowanie tym tematem w

regionie Europy Środkowo-Wschodniej. Widoczna jest globalizacja badań, choć największy wkład pochodzi z Azji i Ameryki Północnej.

8. Wizualizacja z innego narzędzia

Wykres 3. Przedstawienie graficzne najczęściej występujących słów kluczowych w 100 najczęściej cytowanych pracach naukowych przy pomocy narzędzia VOSviewer.





W 100 najczęściej cytowanych artykułach naukowych przeanalizowane zostały słowa kluczowe. Wyselekcjonowane zostały te, które pojawiały się więcej niż 5 razy. Najczęściej występuje “Machine learning” spośród 100 publikacji pojawił się w 24 z nich. Kolejnymi najczęściej występującymi słowami kluczowymi są “Surveys” oraz “Anomaly detection”

występują one 18 razy. Co klasyfikuje wyżej wymienione czynniki jako najczęściej występujące spośród najczęściej cytowanych publikacji.

9. Podsumowanie wyników

Analiza literatury wykazała, że dominującymi słowami kluczowymi w badaniach nad zastosowaniem uczenia maszynowego w finansach są: *machine learning*, *deep learning*, *financial forecasting*, *fraud detection*, *anomaly detection*, *neural networks*, a także *predictive analytics* i *credit scoring*. Wskazuje to na silną koncentrację badań wokół zastosowań predykcyjnych, wykrywania nieprawidłowości oraz wspomagania decyzji finansowych.

Przeglądane publikacje najczęściej dotyczą takich zagadnień jak: prognozowanie cen aktywów i indeksów giełdowych, wykrywanie oszustw w danych transakcyjnych i sprawozdaniach finansowych, ocena ryzyka kredytowego, algorytmiczny handel, analiza sentymentu rynkowego oraz personalizacja usług finansowych. Często poruszane są również problemy interpretowalności modeli (Explainable AI), jakości danych, aspektów regulacyjnych i etycznych, a także integracji modeli z rzeczywistymi systemami finansowymi.

W badaniach dominują metody nadzorowane, w tym sieci neuronowe (MLP, LSTM, RNN, CNN), SVM, regresja logistyczna i lasy losowe. W części prac pojawiają się również metody nienadzorowane (np. analiza skupień, autoenkodery) i podejścia hybrydowe, które łączą różne techniki. Zwraca się uwagę na wyzwania związane z dostępnością danych, ich nieustrukturyzowanym charakterem oraz potrzebą skalowalnych i adaptacyjnych rozwiązań.

Z przeglądu wynika, że badania prowadzone są głównie w celu poprawy trafności prognoz, automatyzacji decyzji i wykrywania zagrożeń finansowych. Mają charakter stosowany, ukierunkowany na praktyczne wdrożenia w dynamicznych warunkach rynkowych.

Wykorzystanie AI:

W procesie przygotowania niniejszej pracy wykorzystano narzędzie sztucznej inteligencji *Open AI Chat GPT 4o*. Wsparcie obejmowało przetwarzanie treści artykułów i identyfikację głównych metod badawczych w Tabelach 1-5 oraz wstępne przygotowanie wybranych tekstów analitycznych. Treści otrzymane w wyniku użycia tych narzędzi zostały przez zespół projektowy zweryfikowane oraz zredagowane w stopniu pozwalającym na stwierdzenie, że jest to opracowanie autorskie.