

Modelowanie i prognozowanie struktury i dynamiki kursów walutowych przy użyciu grafowych sieci neuronowych



Joanna Landmesser-Rusek

*Szkoła Główna Gospodarstwa Wiejskiego
w Warszawie*



V Kongres Statystyki Polskiej
1-3 lipca 2025, Warszawa

CEL BADANIA

Celem badania jest ocena skuteczności **grafowych sieci neuronowych** (*graph neural networks*, GNN) w prognozowaniu kursów walutowych ze szczególnym uwzględnieniem zależności strukturalnych między walutami.

Analiza dotyczy szeregów czasowych dla stóp zwrotu 15 głównych walut światowych w okresie 2022-2024.

Cele szczegółowe:

1. Określenie **struktury sieci walutowej** na podstawie korelacji między zwrotami przy użyciu **minimalnego drzewa rozpinającego** (*minimum spanning tree*, MST), w którym węzły reprezentują poszczególne waluty, a krawędzie współzależności między nimi.
2. Zastosowanie **grafowych sieci neuronowych**, w których kluczowym mechanizmem jest propagacja wiadomości (*message passing*) pomiędzy węzłami, do **prognozowania kursów walutowych**.
3. Porównanie uzyskanych za pomocą GNN wyników prognozowania i klasyfikacji z **benchmarkami**.
4. Analiza **osadzenia węzłów** (*embeddings*) celem identyfikacji walut wpływowych.

METODY – Grafy topologiczne

Grafem nazywamy parę $G = (V, E)$, gdzie V jest zbiorem węzłów ($|V| = n$) i $E \subset V \times V$ zbiorem krawędzi.

Krawędź $e_{ij} = (v_i, v_j)$ łączy węzeł v_i z węzłem v_j .

Sąsiedztwo węzła $v \in V$ definiuje się jako $N(v) = \{u \in V | (u, v) \in E\}$.

Macierz sąsiedztwa $A = (a_{ij})$ jest macierzą wymiaru $n \times n$ pokazującą połączenia pomiędzy węzłami w grafie, przy czym $a_{ij} = 1$ jeśli $e_{ij} \in E$ oraz $a_{ij} = 0$ jeśli $e_{ij} \notin E$.

Wektor x_v jest wektorem cech węzła v .

Jeśli każdej krawędzi przypisana jest waga reprezentująca siłę połączenia między węzłami, to jest to graf ważony.

W przeprowadzonym badaniu **węzły grafów** prezentują **waluty**.

Krawędzie odzwierciedlają **współzależności** pomiędzy walutami, a ich siła jest mierzona współczynnikami korelacji pomiędzy szeregami czasowymi stóp zwrotu.

Do opisu struktury zależności na rynku walutowym wykorzystano **minimalne drzewa rozpinające** (*minimum spanning trees*, MSTs).

MST to podzbiór krawędzi w połączonym, ważonym grafie, który łączy wszystkie wierzchołki bez żadnych cykli i przy minimalnej możliwej całkowitej wadze krawędzi.

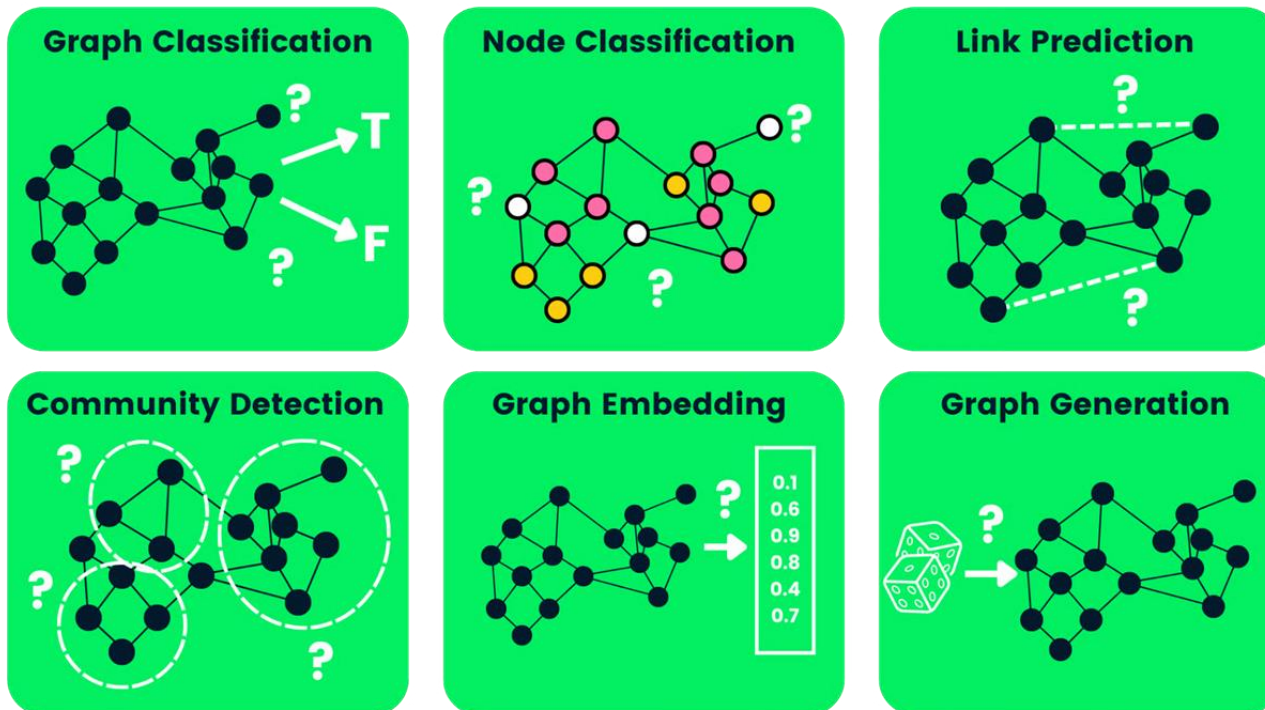
MST stanowiły punkt wyjścia do tworzenia macierzy sąsiedztwa dla GNN konstruowanych w następnym kroku.

METODY – Graflowe sieci neuronowe

Graflowe sieci neuronowe są narzędziem do uczenia się na podstawie danych o strukturze graflowej. Są one w stanie uchwycić złożone relacje w strukturach grafów, co pozwala na ich wykorzystanie w różnych dziedzinach.

GNN są używane głównie do przewidywania:

- **węzłów** (→ tu główny cel to klasyfikacja węzłów),
- **krawędzi** (→ przewidywanie brakujących połączeń między węzłami),
- **całych grafów** (→ klasyfikacja grafów + generowanie grafów).



METODY – Grafowe sieci neuronowe

Pierwsze zastosowanie sieci neuronowych do skierowanych grafów acyklicznych:

Sperduti, A., & Starita, A. (1997). Supervised neural networks for the classification of structures. *IEEE Transactions on Neural Networks*, 8(3), 714-735.

Sformułowanie koncepcji grafowych sieci neuronowych:

Gori, M., Monfardini, G., & Scarselli, F. (2005). A new model for learning in graph domains. *Proceedings of IEEE International Joint Conference on Neural Networks*, 2, 729-734.

Przykłady zastosowań GNN w finansach:

- sieć transakcji -> analiza recenzji user-item,
- dane transakcyjne -> wykrywanie oszustw i przeciwdziałanie praniu pieniędzy,
- scoring kredytowy, analiza ryzyka -> ocena zdolności kredytowej kredytobiorców,
- grafy giełdowe -> przewidywanie cen akcji i ich trendów, klasyfikacja akcji,
- prognozowanie cen kryptowalut, kursów walutowych Romanova (2024), Zhang (2025).

Typologia GNN:

- grafowe sieci rekurencyjne (RGNN),
- grafowe sieci konwolucyjne (GCN),
- grafowe autoenkodery (GAE),
- grafowe sieci przeciwstawne (GAN),
- grafowe sieci przestrzenno-czasowe (STGNN),
- grafowe sieci uczone metodami ze wzmocnieniem (GRL).

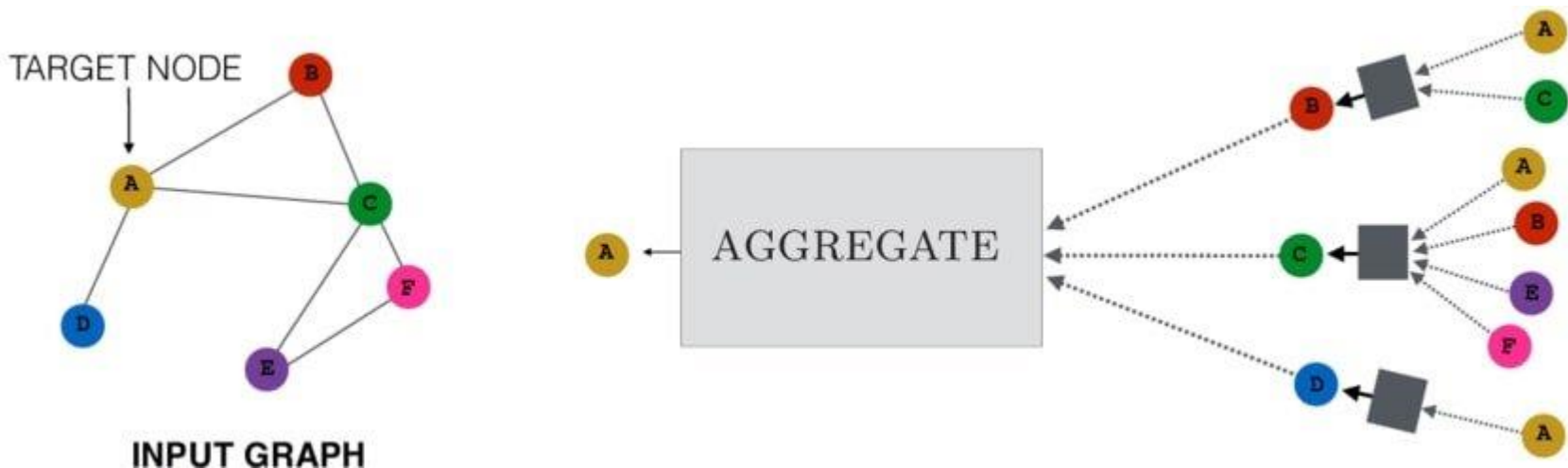
Graflowe sieci neuronowe

Uogólniona formuła dla aktualizacji stanu węzła v w modelach GNN w kroku t :

$$h_v^{(t+1)} = \text{UPDATE} \left(h_v^{(t)}, \text{AGGREGATE} \left(\{h_u^{(t)} : u \in N(v)\} \right) \right)$$

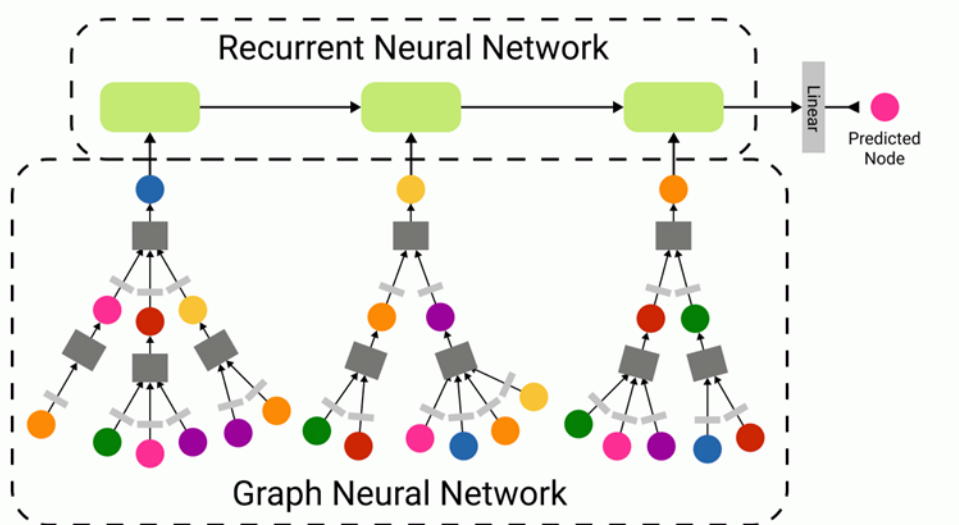
AGGREGATE to funkcja, która integruje informacje z sąsiednich węzłów, co można zrobić za pomocą metod takich jak sumowanie, uśrednianie, przyjmowanie maksimum lub stosowanie mechanizmów uwagi.

UPDATE to funkcja, która aktualizuje reprezentację węzła na podstawie jego poprzedniej reprezentacji i zagregowanych informacji od jego sąsiadów.



Graflowe sieci rekurencyjne

Graflowe sieci rekurencyjne (*recurrent graph neural networks*, RGNNs, 2005) uczą się reprezentacji węzła poprzez iteracyjne propagowanie informacji o sąsiadach, aż osiągną stabilną równowagę.



W grafowych sieciach rekurencyjnych (RGNN) ukryty stan węzła jest stale aktualizowany przez

$$h_v^{(t+1)} = f\left(x_v, \{h_u^{(t)} : u \in N(v)\}\right)$$

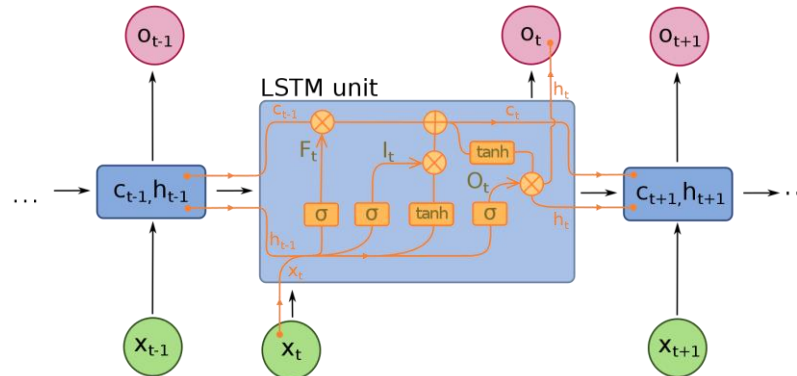
gdzie f jest funkcją parametryczną.

Ukryta reprezentacja węzła jest funkcją lokalnej struktury grafu i cech jego sąsiadów, umożliwiając modelowi przechwytywanie informacji strukturalnych i semantycznych.

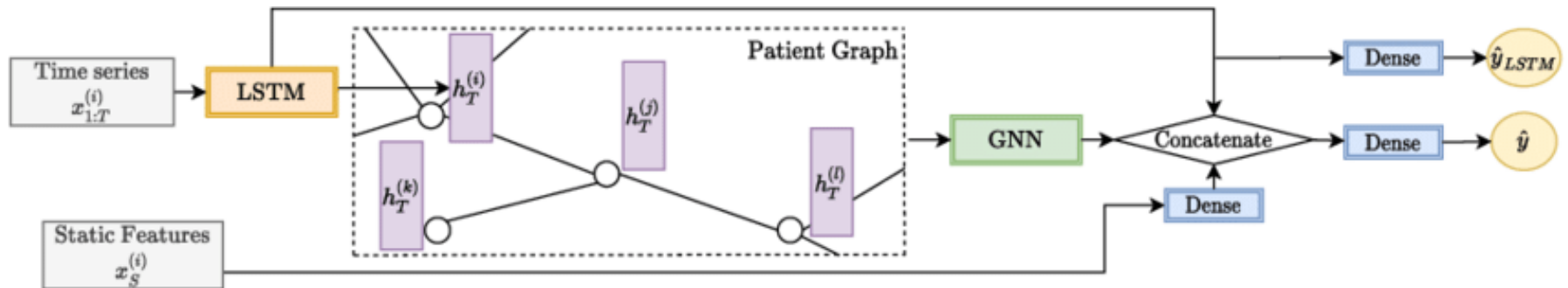
Graflowe sieci rekurencyjne

Sieć LSTM (*long short-term memory*) wykorzystuje trzy bramki (*input gate*, *forget gate*, *output gate*) oraz stan komórki $c_v^{(t)}$ (*cell state*), który umożliwia przechowywanie informacji przez długie okresy czasu.

$$h_v^{(t+1)} = LSTM(x_v^{(t)}, h_v^{(t)}, c_v^{(t)})$$

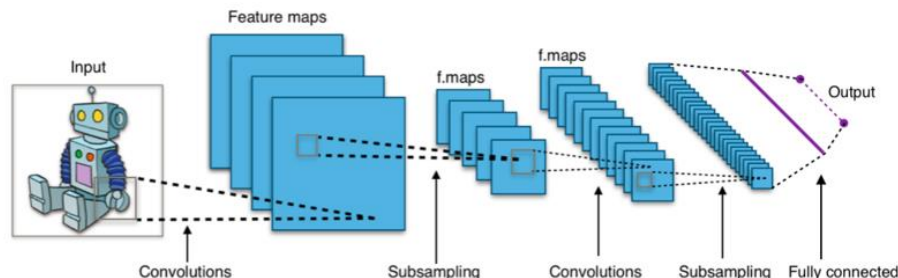


W architekturze **LSTM-GNN**, GNN agreguje informacje od sąsiadów węzła, podczas gdy komórka LSTM jest używana do aktualizacji ukrytego stanu węzła w czasie lub w krokach propagacji.

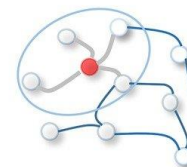
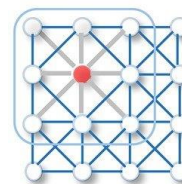


Grafowe sieci konwolucyjne

Sieci konwolucyjne (*convolutional neural networks*, CNNs) - typ sieci przystosowany do klasyfikacji obrazów.

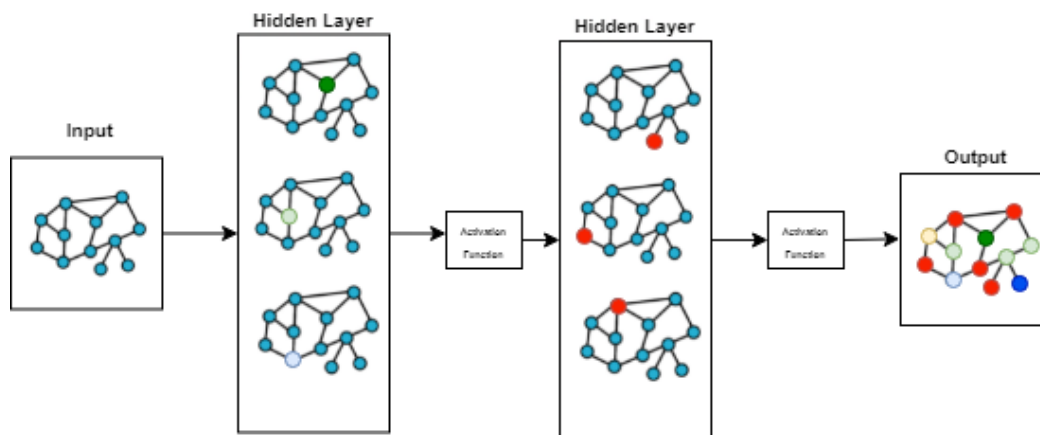


Grafowe sieci konwolucyjne (*graph convolutional neural networks*) rozszerzają ideę konwolucji z obrazów na grafy.



GCNs działają na węzłach i ich sąsiadach w strukturze grafu.

Każdy węzeł aktualizuje swoją reprezentację poprzez agregowanie obrazów od swoich sąsiadów i łączenie ich z własnymi. Pozwala to sieci na przechwytywanie lokalnej struktury grafu i informacji o atrybutach. Sieci GCN łączą wiele warstw konwolucyjnych w celu wyodrębnienia reprezentacji węzłów.



Graflowe sieci konwolucyjne

GCN dziedziczą po RGNN ideę przekazywania wiadomości.

GCN wykorzystują kilka warstw konwolucyjnych do stopniowego gromadzenia i przekształcania informacji z sąsiedztwa węzła, umożliwiając modelowi uczenie się znaczących cech węzła.

W tym przypadku macierz reprezentacji węzła w warstwie l ma postać:

$$H^{(l+1)} = \sigma(\widehat{D}^{-1/2} \widehat{A} \widehat{D}^{-1/2} H^{(l)} W^{(l)})$$

gdzie $\widehat{A}$ jest znormalizowaną macierzą sąsiedztwa z dodanymi pętlami własnymi, $\widehat{D}$ to macierz diagonalna, która mówi, ile krawędzi ma każdy wierzchołek w grafie, $W^{(l)}$ jest macierzą wag dla warstwy l , która jest uczona podczas treningu, a σ jest funkcją aktywacji (taką jak ReLU).

Osadzenie węzła v w $(l+1)$ -ej warstwie jest reprezentowane na podstawie osadzenia jego sąsiadów w l -tej warstwie w następujący sposób:

$$h_v^{(l+1)} = \sigma \left(\sum_{u \in N(v) \cup \{v\}} \frac{1}{\sqrt{d_v d_u}} W^{(l)} h_u^{(l)} \right)$$

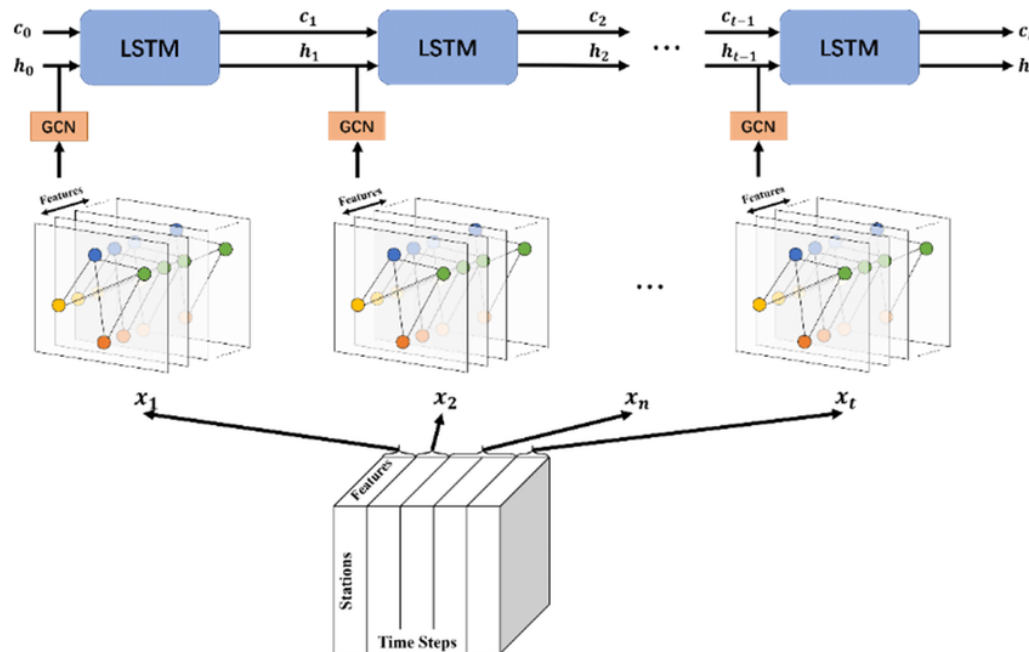
gdzie d_v - stopień wierzchołka v (liczba krawędzi połączonych z v).

LSTM-GCN

LSTM-GCN łączy warstwy konwolucyjne grafu z jednostkami LSTM w celu przetwarzania informacji sekwencyjnych.

W każdej warstwie l model agreguje informacje z sąsiednich węzłów za pomocą operacji konwolucji i aktualizuje stany węzłów, biorąc pod uwagę zarówno strukturę grafu, jak i relacje czasowe poprzez przetwarzanie LSTM:

$$H^{(l+1)} = \sigma(\hat{D}^{-1/2} \hat{A} \hat{D}^{-1/2} LSTM(H^{(l)}) W^{(l)}).$$



ESTYMACJA I EWALUACJA

Możliwe **funkcje straty**:

- w prognozowaniu - średni błąd kwadratowy (MSE), który mierzy różnicę między wartościami przewidywanymi a rzeczywistymi,
- w klasyfikacji binarnej - Binary Cross-Entropy (BCE).

Osadzenia (**embeddings**) węzłów $h_v^{(t)}$ są używane do tworzenia prognoz:

$$\hat{y}_v = f(W h_v^{(L)} + b),$$

gdzie f jest funkcją liniową (w prognozowaniu) lub sigmoidalną (w klasyfikacji), $h_v^{(L)}$ - ostateczne osadzenie węzła v , W - macierz wag, b - odchylenie.

Ocena prognoz:

- **RMSE** (*root mean squared error*) - bardziej karze większe błędy,
- **MAE** (*mean absolute error*) - traktuje wszystkie błędy jednakowo.

Ocena klasyfikacji:

- **Acc** - dokładność (*accuracy*) - udział procentowy poprawnych prognoz w całkowitej liczbie prognoz.

Współczynnik wpływu dla węzła - odzwierciedla jak „centralny” lub „ważny” jest węzeł pod względem jego wyuczonej reprezentacji:

$$Influence\ Score\ (v) = \|h_v\| = \sum_{i=1}^d h_{v,i}^2$$

gdzie h_v jest wektorem osadzenia węzła v , a d jest jego wymiarem.

DANE WYKORZYSTANE W BADANIU EMPIRYCZNYM

Celem badania była ocena skuteczności GNN w prognozowaniu kursów wymiany. Oczekiwano, że GNN zapewnią dokładniejsze prognozy w porównaniu z innymi metodami poprzez uchwycenie strukturalnych relacji między walutami.

Na potrzeby analizy pozyskano dzienne kursy wymiany **15 walut** w stosunku do dolara nowozelandzkiego (X/NZD) z okresu **od 3.01.2022 do 31.12.2024**.

Źródło: <https://stooq.com> (dostęp 1.01.2025).

Analizowano kursy dla:

Abbreviation	Currency name	Abbreviation	Currency name
CAD	Canadian Dollar	KRW	South Korean Won
CHF	Swiss Franc	NOK	Norwegian Krone
CNY	Chinese Yuan	PLN	Polish Zloty
EUR	Euro	RUB	Russian Ruble
GBP	Pound Sterling	SEK	Swedish Krona
HKD	Hong Kong Dollar	SGD	Singapore Dollar
ILS	New Israeli Shekel	USD	US Dollar
JPY	Japanese Yen		

NZD, waluta o niskim znaczeniu globalnym, została potraktowana jako **numéraire** w celu utrzymania w sieci węzłów dla wszystkich innych ważnych walut.

Dla walut obliczono **logarytmiczne stopy zwrotu**: $r^{(t)} = \log p^{(t+1)} - \log p^{(t)}$.

Badanie empiryczne

Analizę przeprowadzono w **3 podokresach**:

- 3.01.-31.12.2022 (258 days),
- 1.01.-31.12.2023 (258 days),
- 1.01.-31.12.2024 (259 days).

W każdym podokresie:

- **prognozowano** wartości zwrotów,
- **klasyfikowano** zwroty na dodatnie i ujemne,
- **porównywano** poszczególne waluty między sobą.

Wykorzystane narzędzia:

- **Model GCN** zbudowany z czterech warstw konwolucyjnych, na koniec których miała miejsce aktywacja ReLU i dropout w celu regularizacji.
- **Model LSTM-GCN** łączący warstwę LSTM do sekwencyjnego przetwarzania danych z dwiema warstwami konwolucyjnymi i dropoutem.
- Benchmark 1: **Model MLP** składający się z dwóch warstw ukrytych, po których następowała aktywacja ReLU i dropout.
- Benchmark 2: **Model ARIMA**, w którym rzędy p, q, d ustalono w procedurze automatycznej.

WYNIKI EMPIRYCZNE

Część I: Prognozy dla jednego okresu (ostatniego w szeregu czasowym)

800 epok;

optimizer Adam ze współczynnikiem uczenia 0,005;

dropout ze współczynnikiem 0,2;

funkcja straty MSE.

Tabela 1. Porównanie skuteczności różnych metod prognozowania (najlepsze wyniki pogrubiono)

Methods	Forecast Errors					
	2022		2023		2024	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
LSTM-GCN	0.585	0.435	0.663	0.467	0.447	0.352
GCN	0.773	0.549	0.525	0.372	0.738	0.703
MLP	0.606	0.490	0.890	0.540	1.300	0.864
ARIMA	0.442	0.291	0.543	0.379	0.538	0.449
Naïve	0.607	0.483	1.056	0.645	1.480	0.923

WYNIKI EMPIRYCZNE

Część II: Klasyfikacja stóp zwrotu (2 klasy: wzrostów i spadków)

dane treningowe – 90%, testowe – 10%.

200 epok;

32 batche na 1 epokę;

optymalizator AdamW ze współczynnikiem uczenia 0,001;

dropout ze współczynnikiem 0,2;

funkcja straty BCE.

Tabela 2. Wyniki klasyfikacji stóp zwrotu dla walut - porównanie wyników (najlepsze wyniki pogrubiono)

Methods	Accuracy		
	2022	2023	2024
LSTM-GCN	71.54 %	71.79 %	72.05 %
GCN	70.51 %	67.18 %	70.77 %
MLP	63.33 %	65.64 %	63.08 %

WYNIKI EMPIRYCZNE

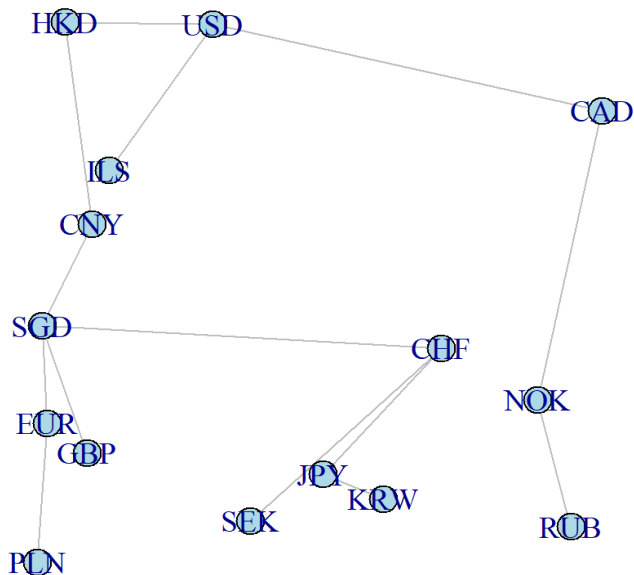
Część III: Analiza współczynników wpływu

- przykładowa wizualizacja MST oraz GNN

Embedding - wektor liczb, który reprezentuje walutę w przestrzeni cech:

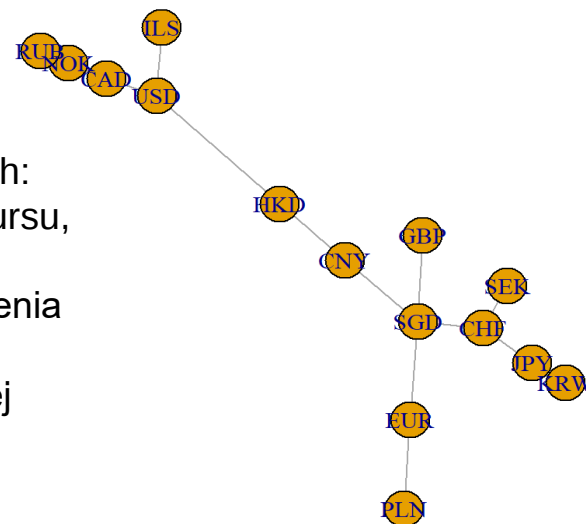
- „streszcza” najważniejsze informacje o danej walucie (historię jej kursu, powiązania z innymi, czyli informacje od sąsiadów w grafie)
- zawiera „siłę wpływu” widoczną w tym, jak bardzo dana waluta zmienia embeddingi sąsiadów
- model przypisuje większą wagę tym walutom, które pozwalają lepiej przewidywać zmiany u innych

Graph Neural Network - Embeddings (Currencies)



v	h_v,1	h_v,2	Influence score_v
CAD	19,13	38,35	1836,40
CHF	13,57	19,76	574,62
CNY	1,46	29,46	870,28
EUR	-0,10	13,92	193,70
GBP	1,27	11,58	135,66
HKD	0,51	45,29	2051,73
ILS	2,05	33,67	1137,66
JPY	9,46	9,89	187,33
KRW	11,55	7,96	196,64
NOK	16,90	15,70	532,22
PLN	-0,45	2,99	9,14
RUB	18,04	5,81	359,32
SEK	6,94	6,25	87,20
SGD	-0,29	21,50	462,41
USD	5,64	45,11	2066,91

Graph MST



WYNIKI EMPIRYCZNE

Część III: Analiza współczynników wpływu walut

Tabela 3. Współczynniki wpływu obliczone na podstawie osadzeń i dokładności klasyfikacji dla walut w oparciu o LSTM-GCN (najwyższe wartości pogrubiono, drugie i trzecie najwyższe podkreślono, najniższe zaznaczono kursywą)

Currency	Influence Score			Accuracy		
	2022	2023	2024	2022	2023	2024
CAD	4.9	<u>1769.7</u>	159.9	<u>80.8</u>	<u>73.1</u>	<u>73.1</u>
CHF	4.9	34.0	286.3	<u>76.9</u>	80.8	<u>76.9</u>
CNY	5.3	25.4	1078.1	<u>50.0</u>	<u>76.9</u>	<u>76.9</u>
EUR	<u>206.9</u>	548.3	<u>410.4</u>	<u>80.8</u>	80.8	<u>73.1</u>
GBP	187.8	34.0	26.2	<u>80.8</u>	<u>73.1</u>	61.5
HKD	3.9	<u>837.9</u>	237.7	73.1	<u>76.9</u>	<u>73.1</u>
ILS	4.9	3.4	159.9	<u>80.8</u>	69.2	61.5
JPY	4.9	3.4	<u>429.1</u>	53.8	57.7	<u>76.9</u>
KRW	3.6	9.2	297.0	69.2	<u>73.1</u>	65.4
NOK	114.9	109.4	67.8	69.2	65.4	69.2
PLN	<u>394.4</u>	34.0	26.2	65.4	<u>73.1</u>	<u>73.1</u>
RUB	8.3	3.4	26.2	61.5	69.2	65.4
SEK	436.2	81.4	39.0	84.6	<u>73.1</u>	88.5
SGD	163.2	638.0	60.0	69.2	69.2	<u>76.9</u>
USD	3.7	1802.4	156.3	<u>76.9</u>	65.4	69.2

WNIOSKI

- ❑ GCN i LSTM-GCN osiągnęły najniższe wartości RMSE i MAE spośród wszystkich modeli w latach 2023 i 2024. W 2022 r. przewyższył je ARIMA. Wniosek: na bardzo zmiennym rynku FX mogą zdarzać się okresy, w których GNN te nie są w stanie uchwycić szybko zmieniających się wzorców.
- ❑ Jeśli chodzi o dokładność klasyfikacji, model LSTM-GCN wyraźnie przewyższył inne metody, konsekwentnie osiągając dokładność powyżej 71%.
- ❑ Sieci GNN oferują również cenne informacje na temat zachowania poszczególnych walut:
 - Odnotowano wzrost wpływu USD, CAD, HKD w 2023 r. oraz CNY i CHF w 2024 r.
 - EUR utrzymywało wysoki i stabilny wpływ przez cały okres badania.
 - GBP wykazał zauważalny spadek wpływu.
 - Wysoka dokładność klasyfikacji dla CAD, CHF, EUR, HKD i SEK, niższa dla CNY, JPY i RUB.
- ❑ Ograniczenia GNN:
 - kosztowne obliczeniowo (złożoność obliczeniowa może ograniczać skalowalność GNN w przypadku większych zestawów danych lub danych o wyższej częstotliwości),
 - skomplikowane dostrajanie hiperparametrów.

Dziękuję za uwagę!