Risk Radar

System wczesnego ostrzegania przed reżimami podwyższonej zmienności akcji spółek z giełdy NASDAQ

Mateusz Weryński · Informatyka Stosowana, Uniwersytet Ekonomiczny w Krakowie, 2026
Podsumowanie projektu na podstawie pracy magisterskiej „Zastosowanie sztucznej inteligencji w analizie i ocenie ryzyka inwestycji w akcje spółek notowanych na rynku NASDAQ”
Promotor: dr hab. Marcin Jędrzejczyk, prof. UEK

> **Nota o wersji.** Niniejsze podsumowanie opisuje system w postaci zbadanej w pracy magisterskiej. Aktualnie dostarczona wersja Risk Radar została od tego czasu poprawiona i ulepszona, dlatego jej działanie oraz wyniki mogą się różnić od zaprezentowanych poniżej.

W skrócie

Risk Radar to zbudowany w Pythonie klasyfikator uczenia maszynowego, który z wyprzedzeniem szacuje prawdopodobieństwo wejścia spółki w okres gwałtownej zmienności kursu. Model oparty na lasach losowych, douczany co roku w schemacie Walk-Forward, utrzymał przewagę nad losowością w 9-letnim teście (2017–2026) we wszystkich badanych koszykach spółek.

0,61–0,67	**3 × 3**	**10**	**7**
AUC w teście 9-letnim	koszyki × horyzonty testu	scenariuszy kryzysowych	cech wejściowych + SHAP

Risk Radar nie jest sygnałem kupna ani sprzedaży. Mierzy ryzyko silnych wahań, aby inwestor mógł świadomie dostosować wielkość pozycji i zabezpieczenie portfela.

1. Problem i hipotezy badawcze

Spółki technologiczne z NASDAQ są skrajnie wrażliwe na stopy procentowe i sentyment, co pokazały przeceny z lat 2000, 2008, 2020 i 2022. Projekt odpowiada na trzy słabości klasycznych narzędzi:

- **symetryczne miary ryzyka**, w których wzrost o 10% waży tyle samo co spadek o 10%,
- **starzenie się modeli** (*concept drift*) w niestacjonarnym środowisku rynkowym,
- **brak interpretowalności** decyzji modeli typu „czarna skrzynka”.

Zweryfikowano cztery hipotezy:

H1. Random Forest wzbogacony o indeks VIX osiąga skuteczność powyżej losowości: AUC $> 0{,}5$.

H2. Walidacja Walk-Forward z rozszerzającym się oknem ogranicza dryf koncepcji w długim horyzoncie.

H3. Model oparty wyłącznie na danych rynkowych nie przewiduje „Czarnych Łabędzi” (granica użyteczności).

H4. Metoda SHAP pozwala lokalnie wyjaśnić czynniki stojące za oceną ryzyka w konkretnym dniu.

Wszystkie cztery hipotezy zostały potwierdzone, w tym H3, która opisuje ograniczenie modelu.

2. Dane i inżynieria cech

System pobiera z Yahoo Finance (biblioteka `yfinance`) 15 lat dziennych notowań spółki (cena zamknięcia P_t, wolumen V_t) oraz indeksu strachu VIX. Szeregi są łączone po dacie sesji, braki uzupełniane ostatnią znaną wartością (*forward fill*), a rekordy z wartościami nieskończonymi

usuwane. Podstawą obliczeń są logarytmiczne stopy zwrotu:

$$r_t = \ln \frac{P_t}{P_{t-1}}. \tag{1}$$

Surowe ceny są niestacjonarne, dlatego model otrzymuje wyłącznie wskaźniki względne. Zmienność krocząca w oknie $n \in \{10, 30\}$ sesji, w ujęciu rocznym (252 sesje):

$$\sigma_t^{(n)} = \sqrt{252} \cdot \sqrt{\frac{1}{n-1} \sum_{i=0}^{n-1} \left(r_{t-i} - \bar{r}_t^{(n)}\right)^2}, \qquad \bar{r}_t^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} r_{t-i}. \tag{2}$$

Impet (momentum), dynamika wolumenu i dzienna zmiana VIX:

$$\mathrm{MOM}_t^{(10)} = \frac{P_t}{P_{t-10}} - 1, \qquad \Delta V_t = \frac{V_t}{V_{t-1}} - 1, \qquad \Delta \mathrm{VIX}_t = \frac{\mathrm{VIX}_t}{\mathrm{VIX}_{t-1}} - 1. \tag{3}$$

Odpowiedzią na problem symetrii odchylenia standardowego jest asymetryczna miara ryzyka, kroczące maksymalne obsunięcie kapitału względem szczytu z 30 sesji:

$$\mathrm{MDD}_t^{(30)} = \frac{P_t}{\max\limits_{0 \le k \le 29} P_{t-k}} - 1 \ \in \ [-1, 0]. \tag{4}$$

Wartość 0 oznacza cenę na lokalnym szczycie, a wartości ujemne mierzą głębokość trwającej wyprzedaży. Wektor wejściowy modelu ma postać:

$$\mathbf{x}_t = \left[\, \sigma_t^{(10)},\ \sigma_t^{(30)},\ \mathrm{MOM}_t^{(10)},\ \Delta V_t,\ \mathrm{MDD}_t^{(30)},\ \mathrm{VIX}_t,\ \Delta \mathrm{VIX}_t \,\right]^\top \in \mathbb{R}^7. \tag{5}$$

3. Zmienna celu bez wycieku danych

Dla każdego dnia t wyznaczana jest zmienność, która faktycznie wystąpi w kolejnych $H = 20$ sesjach:

$$\sigma_t^{\mathrm{fwd}} = \sqrt{252} \cdot \mathrm{std}(r_{t+1}, r_{t+2}, \dots, r_{t+H}). \tag{6}$$

Próg anomalii to 75. percentyl rozkładu tej zmienności, liczony wyłącznie z obserwacji historycznych, już zrealizowanych w dniu t:

$$\tau_t = Q_{0,75}\Big(\{\sigma_s^{\mathrm{fwd}} : s \le t - H\}\Big), \tag{7}$$

a etykieta binarna przyjmuje postać:

$$y_t = \mathbb{1}[\sigma_t^{\mathrm{fwd}} > \tau_t] = \begin{cases} 1 & \text{reżim podwyższonej zmienności (wysokie ryzyko),} \\ 0 & \text{rynek stabilny.} \end{cases} \tag{8}$$

Dynamiczny próg zapobiega wyciekowi informacji z przyszłości (*target leakage*) i dopasowuje się do charakteru spółki: w badanych scenariuszach wynosił około 25% dla Microsoftu i około 63% dla Tesli. Ostatnie H obserwacji, dla których σ_t^{fwd} nie jest jeszcze znane, są usuwane ze zbioru uczącego.

4. Model: lasy losowe i wskaźnik ryzyka

Cechy są standaryzowane parametrami wyznaczonymi *wyłącznie* na zbiorze treningowym:

$$z_{t,j} = \frac{x_{t,j} - \mu_j^{\mathrm{train}}}{s_j^{\mathrm{train}}}. \tag{9}$$

Okresy wysokiej zmienności są z definicji rzadsze (około 25% obserwacji), dlatego klasy ważone są odwrotnie do liczności (`class_weight='balanced'`), gdzie N to liczba obserwacji, $K = 2$ liczba klas:

$$w_c = \frac{N}{K \cdot N_c}, \qquad c \in \{0, 1\}. \tag{10}$$

Każde z B drzew lasu losowego dzieli przestrzeń cech, minimalizując ważoną nieczystość Giniego w węzłach:

$$G(m) = 1 - \sum_{c\in\{0,1\}} \hat{p}_{m,c}^2, \qquad \hat{p}_{m,c} = \frac{\sum_{i\in m} w_{y_i}\,\mathbb{1}[y_i = c]}{\sum_{i\in m} w_{y_i}}, \tag{11}$$

a prognoza lasu jest średnią prawdopodobieństw z poszczególnych drzew. Na jej podstawie powstaje wskaźnik ryzyka w skali 0–100, przypisywany następnie do pięciostopniowej drabiny (od „bardzo niskiego" do „ekstremalnego" ryzyka):

$$\hat{p}(\mathbf{x}_t) = \frac{1}{B}\sum_{b=1}^{B} \hat{p}_b(y_t = 1 \mid \mathbf{z}_t), \qquad S_t = \mathrm{round}(100 \cdot \hat{p}(\mathbf{x}_t)). \tag{12}$$

5. Walidacja Walk-Forward

Losowy podział danych (*k-fold*) pozwala modelowi „zajrzeć" w przyszłość, a stałe okno uczące szybko się starzeje. Zastosowano więc walidację postępującą z rozszerzającym się oknem. Dla kolejnych dat granicznych $T_0 < T_1 < \cdots < T_K$, oddalonych o rok, k-ty krok wykorzystuje zbiory:

$$\mathcal{D}_k^{\text{train}} = \{(\mathbf{x}_t, y_t) : t < T_k\}, \qquad \mathcal{D}_k^{\text{test}} = \{(\mathbf{x}_t, y_t) : T_k \le t < T_{k+1}\}. \tag{13}$$

Algorytm. Uczenie postępujące (expanding window)

1. Dla $k = 0, 1, \ldots, K-1$:
   a) wyznacz $\mu^{\text{train}}, s^{\text{train}}$ na $\mathcal{D}_k^{\text{train}}$ i przekształć oba zbiory,
   b) wytrenuj nowy las losowy z wagami w_c na $\mathcal{D}_k^{\text{train}}$,
   c) wykonaj „ślepą" predykcję $\hat{p}(\mathbf{x}_t)$ dla $t \in \mathcal{D}_k^{\text{test}}$ i zapisz wyniki,
   d) dołącz $\mathcal{D}_k^{\text{test}}$ do historii: $\mathcal{D}_{k+1}^{\text{train}} = \mathcal{D}_k^{\text{train}} \cup \mathcal{D}_k^{\text{test}}$.
2. Policz metryki na połączonych predykcjach ze wszystkich kroków.

6. Wyjaśnialna AI: wartości Shapleya

Każda predykcja jest rozkładana na wkłady cech metodą SHAP (`TreeExplainer`). Wkład cechy j to jej średni marginalny wpływ na wynik modelu po wszystkich podzbiorach $\mathcal{S}$ pozostałych cech $\mathcal{F}$, gdzie $|\mathcal{F}| = 7$:

$$\phi_j = \sum_{\mathcal{S}\subseteq\mathcal{F}\setminus\{j\}} \frac{|\mathcal{S}|!\,(|\mathcal{F}| - |\mathcal{S}| - 1)!}{|\mathcal{F}|!}\Big[f_{\mathcal{S}\cup\{j\}}(\mathbf{x}) - f_{\mathcal{S}}(\mathbf{x})\Big]. \tag{14}$$

Wkłady sumują się dokładnie do predykcji (własność addytywności), więc $\phi_j > 0$ oznacza cechę podnoszącą ryzyko, a $\phi_j < 0$ cechę je obniżającą:

$$\hat{p}(\mathbf{x}) = \phi_0 + \sum_{j=1}^{7} \phi_j, \qquad \phi_0 = \mathbb{E}[\hat{p}(\mathbf{x})]. \tag{15}$$

Przykład: dla Alphabet 15.05.2017 wskaźnik ryzyka wyniósł 0/100, a najsilniej obniżał go niski VIX (10,42 pkt, $\phi = -0{,}27$). Globalnie uzupełnieniem jest ważność cech (*Feature Importance*) lasu losowego.

7. Architektura aplikacji

- **Moduł 1 – Zintegrowany Model Scoringowy.** Trening na całej historii do wczoraj, wskaźnik S_t na bieżącą sesję, pięciostopniowa klasyfikacja i lista czynników SHAP.
- **Moduł 2 – Wehikuł Czasu.** Odtworzenie predykcji z dowolnego „Dnia Zero" bez dostępu do kolejnych 20 sesji, zestawienie z rzeczywistością i werdykt TAK/NIE z wykresem.
- **Moduł 3 – Globalny backtest.** Walk-Forward dla całego koszyka spółek: skuteczność i czułość per spółka, zagregowana macierz błędów, krzywa ROC z AUC oraz ważność cech.

8. Metryki i wyniki testów portfelowych

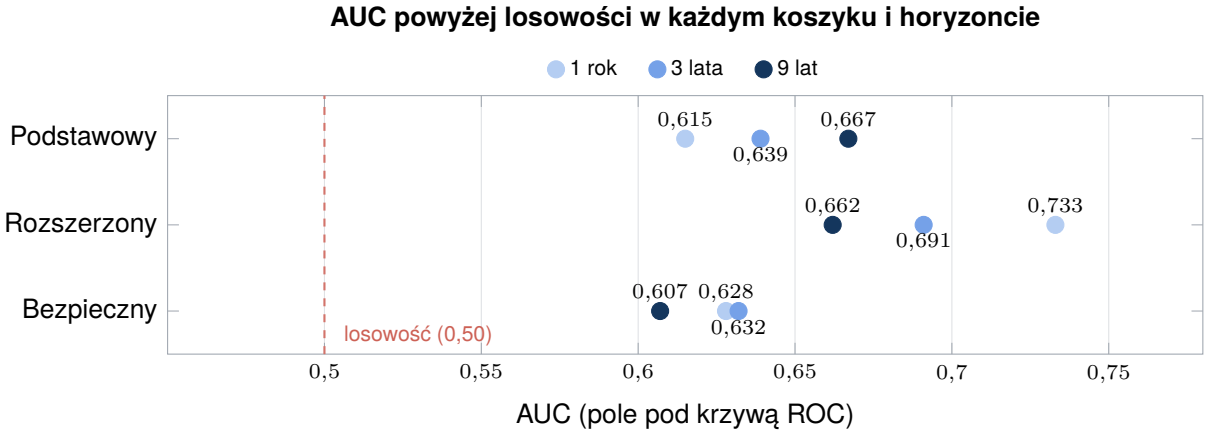
Przy przewadze dni spokojnych sama skuteczność (*accuracy*) jest myląca, dlatego kluczowe są czułość (jaka część rzeczywistych okresów ryzyka została oflagowana) oraz AUC:

$$\mathrm{Acc} = \frac{TP + TN}{TP + TN + FP + FN}, \qquad \mathrm{Recall} = \mathrm{TPR} = \frac{TP}{TP + FN}, \qquad \mathrm{FPR} = \frac{FP}{FP + TN}, \tag{16}$$

$$\mathrm{AUC} = \int_0^1 \mathrm{TPR}\,\mathrm{d}(\mathrm{FPR}) = \Pr\big(\hat{p}(\mathbf{x}^+) > \hat{p}(\mathbf{x}^-)\big). \tag{17}$$

AUC to szansa, że dzień przed okresem wysokiej zmienności dostanie wyższą ocenę niż dzień spokojny: 0,50 to losowość, 0,60–0,70 dobry model dla rynku akcji, a powyżej 0,85 sygnał błędu (np. wycieku danych).

Testy objęły trzy koszyki: **Podstawowy** (AAPL, MSFT, AMD, TSLA, NVDA), **Rozszerzony** (podstawowy + PLTR, NFLX, INTC, AMZN) oraz **Bezpieczny** (JNJ, PG, KO, WMT, MCD), każdy w horyzontach 1, 3 i 9 lat.



Skuteczność / czułość (%) w podziale na koszyki

Koszyk	1 rok	3 lata	9 lat
Podstawowy	69,6 / 21,9	71,6 / 34,1	67,0 / 49,5
Rozszerzony	71,1 / 55,7	72,4 / 48,6	66,9 / 51,6
Bezpieczny	59,0 / 38,4	67,2 / 33,5	63,8 / 42,1

Macierz błędów: koszyk bezpieczny, 9 lat

	Prog.: 0	Prog.: 1
Fakt.: 0	TN = 5703	FP = 2019
Fakt.: 1	FN = 2072	TP = 1506

0 – stabilnie, 1 – ryzyko

Najważniejszy wynik dotyczy dryfu koncepcji. W horyzoncie 9 lat model wytrenowany jednorazowo spadał do poziomu rzutu monetą (AUC $\approx$ 0,50). Po wdrożeniu corocznego douczania ten sam algorytm utrzymał AUC = 0,667 dla koszyka Big Tech, a jego czułość wzrosła do około 50%.

9. Dziesięć scenariuszy rynkowych

W Wehikule Czasu model sprawdzono w 10 celowo trudnych momentach. Trafił w 3 z nich, a scenariusze wyznaczyły granice jego użyteczności.

Scenariusz	Spółka	Dzień Zero	S_t	MDD 20 sesji	$\sigma^{\mathrm{fwd}}/\tau$	Trafność
Wybuch pandemii COVID-19	MSFT	19.02.2020	2	−27,9%	107,6% / 25,0%	NIE
Volmageddon (eksplozja VIX)	AAPL	30.01.2018	6	−9,7%	35,2% / 27,3%	NIE
Szok stóp procentowych Fed	TSLA	15.09.2022	7	−32,1%	53,9% / 62,5%	TAK
Pęknięcie bańki „work from home”	ZM	15.10.2020	4	−31,7%	100,9% / 76,0%	NIE
Przesilenie hossy AI	NVDA	10.04.2024	0	−13,1%	63,9% / 49,8%	NIE
Katastrofa raportu kwartalnego	META	01.02.2022	7	−40,4%	115,1% / 38,0%	NIE
Załamanie popytu na sprzęt	AMD	01.10.2018	11	−48,2%	96,5% / 64,6%	NIE
Korelacja z rynkiem kryptowalut	MSTR	08.11.2021	11	−34,2%	57,4% / 42,2%	NIE
Stabilna hossa (fałszywy alarm?)	GOOGL	15.05.2017	0	−2,3%	19,0% / 24,5%	TAK
Ocena na bieżących danych	ASTS	07.07.2026	1	−28,9%	110,8% / 111,3%	TAK

- **Model nie widzi „Czarnych Łabędzi”.** 19.02.2020 cena Microsoftu była na szczycie ($\mathrm{MDD}^{(30)} = 0$), momentum było silne, a VIX spokojny. Informacja, która nie trafiła jeszcze do notowań (pandemia, raport Meta, krach kryptowalut), jest dla modelu niewidoczna – to potwierdzenie H3 i rozróżnienia Knighta między ryzykiem a niepewnością.
- **Zmienność to nie kierunek.** Tesla spadła o 32%, ale $\sigma^{\mathrm{fwd}} < \tau$, więc ocena „niskie ryzyko” była formalnie poprawna: model mierzy siłę wahań, nie znak stopy zwrotu.
- **Brak fałszywych alarmów w spokojnym rynku.** Alphabet 2017 został poprawnie oceniony jako bezpieczny, a w czasie Volmageddonu „niskie ryzyko” dla Apple uchroniło przed zamknięciem pozycji, która po 20 sesjach była na plusie 7,1%.

10. Ograniczenia i kierunki rozwoju

- **Brak backtestu strategii.** AUC pokazuje przewagę statystyczną, ale nie krzywą kapitału. Pełna walidacja wymaga reguł wejścia i wyjścia, kosztów transakcyjnych, poślizgu oraz porównania z „kup i trzymaj” na NASDAQ-100 przy użyciu wskaźników Sharpe'a i Sortino:

$$\mathrm{SR} = \frac{\bar{R}_p - R_f}{\sigma_p}, \qquad \mathrm{SoR} = \frac{\bar{R}_p - R_f}{\sigma_d}, \qquad \sigma_d = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \min(0,\, R_{p,t} - R_f)^2}. \tag{18}$$

- **Brak modeli bazowych.** Kolejnym krokiem jest porównanie z regresją logistyczną, regułą progową na VIX oraz gradient boostingiem (np. XGBoost).
- **Tylko dane ustrukturyzowane.** Rozbudowa o modele językowe (LLM), NLP i podejście agentowe pozwoliłaby analizować komunikaty spółek, wiadomości i nastroje w mediach w czasie rzeczywistym.
- **SHAP wyjaśnia model, nie rynek.** Wartości ϕ_j opisują logikę estymatora, a nie obiektywne mechanizmy giełdy.

11. Weryfikacja hipotez i wnioski

Hipoteza	Wynik	Dowód
H1	potwierdzona	AUC $\in$ [0,607; 0,733] we wszystkich 9 testach
H2	potwierdzona	9 lat: AUC = 0,61–0,67 zamiast $\approx$ 0,50 dla modelu statycznego
H3	potwierdzona	przeoczone szoki zewnętrzne: COVID-19, raport Meta, krach kryptowalut
H4	potwierdzona	lokalne wkłady ϕ_j dla każdej predykcji (np. VIX dla Alphabet, MDD dla ASTS)

Risk Radar dostarcza stabilnej, mierzalnej przewagi statystycznej w wykrywaniu reżimów wysokiej zmienności, zachowując pełną przejrzystość decyzji. Jednocześnie projekt jasno pokazuje, gdzie kończą się możliwości modelu opartego wyłącznie na danych cenowych – i właśnie tę granicę mają przesunąć kolejne wersje systemu.

12. Stack technologiczny i kompetencje

Technologia	Zastosowanie
Python 3.10+, Jupyter	logika aplikacji, eksperymenty i dokumentacja badań
pandas, NumPy	przetwarzanie szeregów czasowych, wektoryzacja obliczeń
yfinance	automatyczne pobieranie notowań spółek i VIX
scikit-learn	`RandomForestClassifier`, `StandardScaler`, metryki ROC AUC i recall
SHAP	lokalne wyjaśnianie predykcji (`TreeExplainer`)
Matplotlib, Seaborn	krzywe ROC, macierze błędów, wykresy analityczne

Projekt łączy uczenie maszynowe na szeregach czasowych przy niezrównoważonych klasach, rygorystyczną walidację bez wycieku danych, inżynierię cech finansowych, wyjaśnialną AI oraz wiedzę domenową o ryzyku rynkowym – wraz z krytyczną oceną granic własnego modelu.