

Credit Scorecard

Model Oceny Ryzyka Kredytowego

Agenda

- Wprowadzenie do projektu
- Metodologia i architektura modelu
- Preprocessing i feature engineering
- Wyniki modelu Basic Scorecard
- Analiza kalibracji i generalizacji
- Interpretacja zmiennych (WoE)
- Rating Scorecard (AAA-D)
- Podsumowanie i wnioski
- Advanced Scorecard (preview)

Cel projektu i zakres analizy

- **Problem biznesowy:**
 - Automatyzacja oceny ryzyka kredytowego firm.
 - Wsparcie decyzji komitetu kredytowego.
- **Zakres projektu:**
 - Model Basic Scorecard: Regresja logistyczna + WoE encoding.
 - Cel: Predykcja default (0/1) i mapping na rating AAA-D.
- **Korzyści biznesowe:**
 - Szybsza ocena, transparentność, zgodność z regulacjami (Basel).

Architektura modelu Basic Scorecard

[Dane surowe] → [Preprocessing & Scaling]



[WoE Encoding] → [Feature Selection (Optuna)]



[Logistic Regression] → [Kalibracja (Beta)]



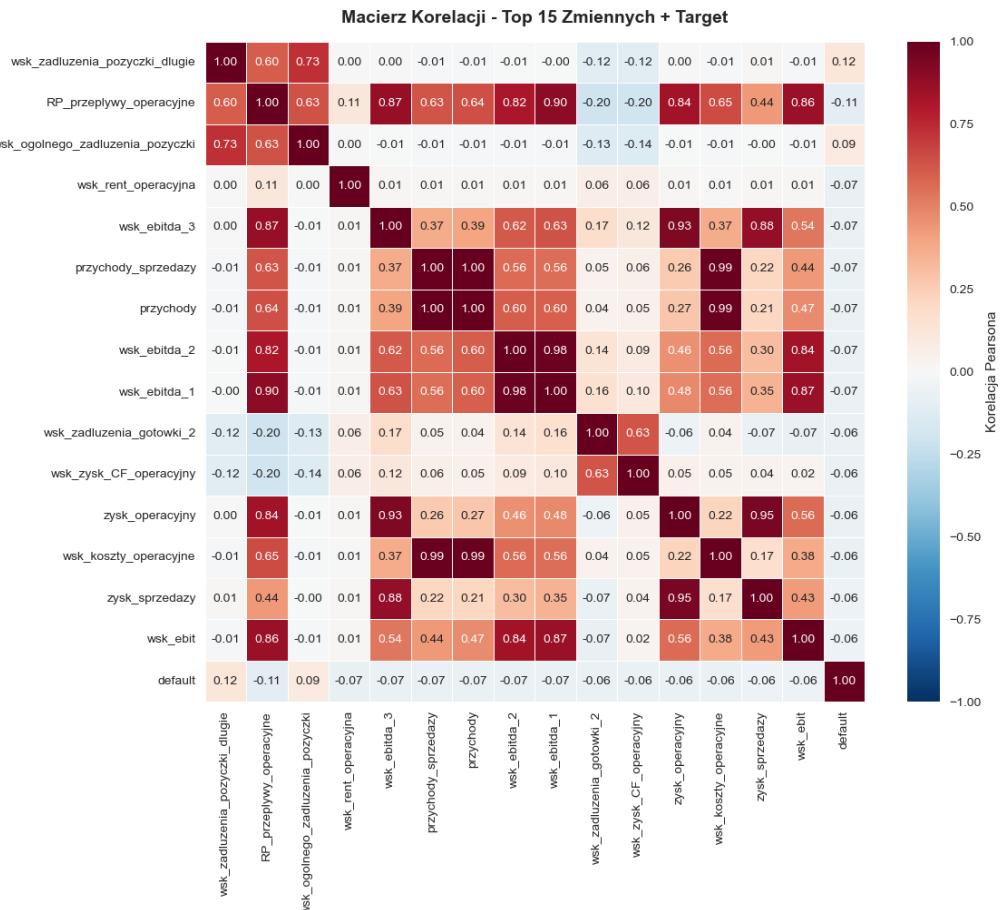
[Rating Mapping AAA-D] → [Decyzja]

Kluczowe komponenty:

WoE Encoding: Transformacja na log-odds (monotoniczność).

Kalibracja: Dopasowanie PD do 4% (średni default rate).

Preprocessing i Feature Engineering



1. Czyszczenie danych:

Missing: Median imputation
Outliers: Winsorization (1-99%)

2. WoE & Selekcja:

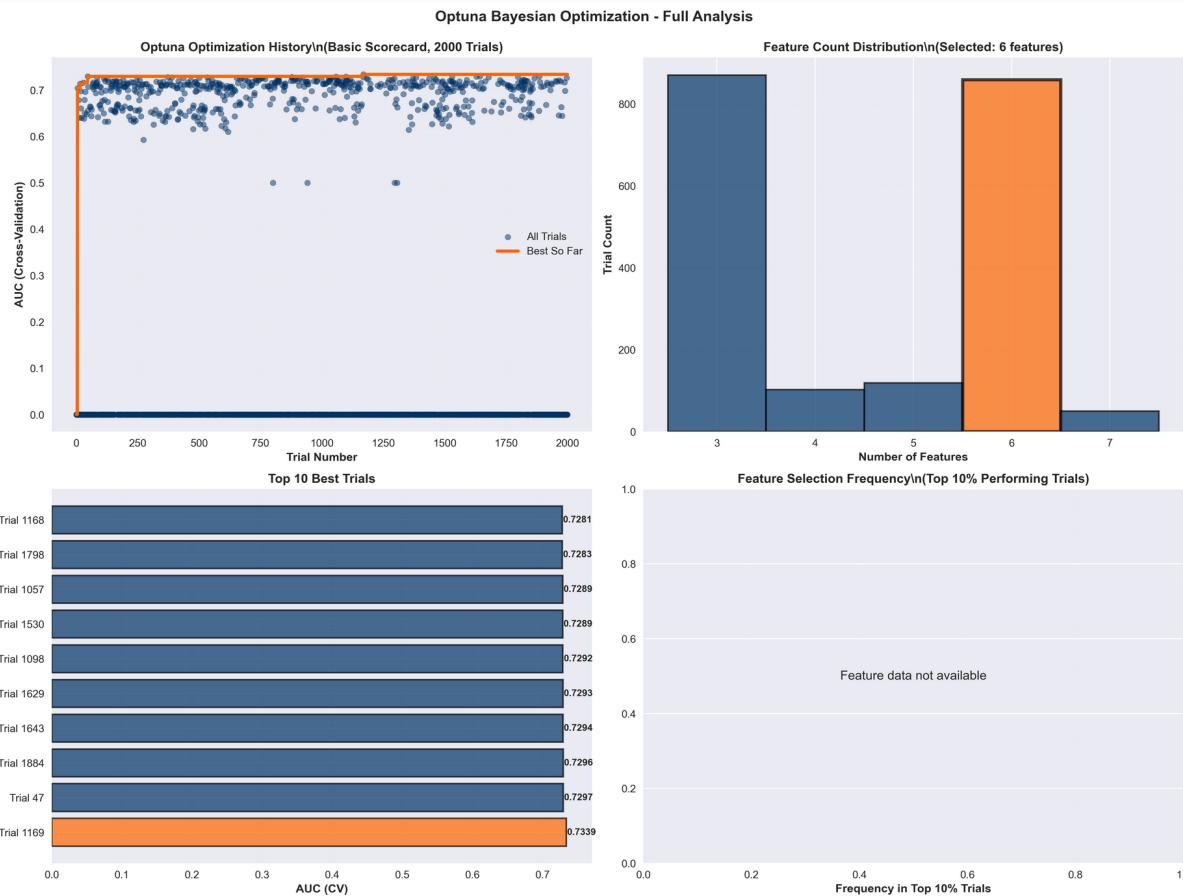
Adaptive Binning (3-10 binów)
Selekcja IV > 0.10 (Silna predykcja)

3. Stabilność (PSI):

Wszystkie zmienne PSI < 0.01

Optuna

- Selekcję cech przeprowadzono algorytmem Optuna w 2000 próbach z 5-krotną walidacją krzyżową
- Za główny cel powzięliśmy maksymalizując AUC przy zachowaniu stabilności znaków współczynników dla zapewnienia interpretowalności.
- Przed optymalizacją z modelu wykluczono zmienne wykazujące istotną niestabilność w czasie ($\text{PSI} > 0.25$): `wsk_rent_kapitalu`, `wsk_marza_brutto`, `wsk_rent_kaptial_podstawowy`, `wsk_udzial_zob_biez_aktywa_2` oraz `wsk_zast_kapitalu_podstawowego`.



Wybrane cechy modelu Basic Scorecard

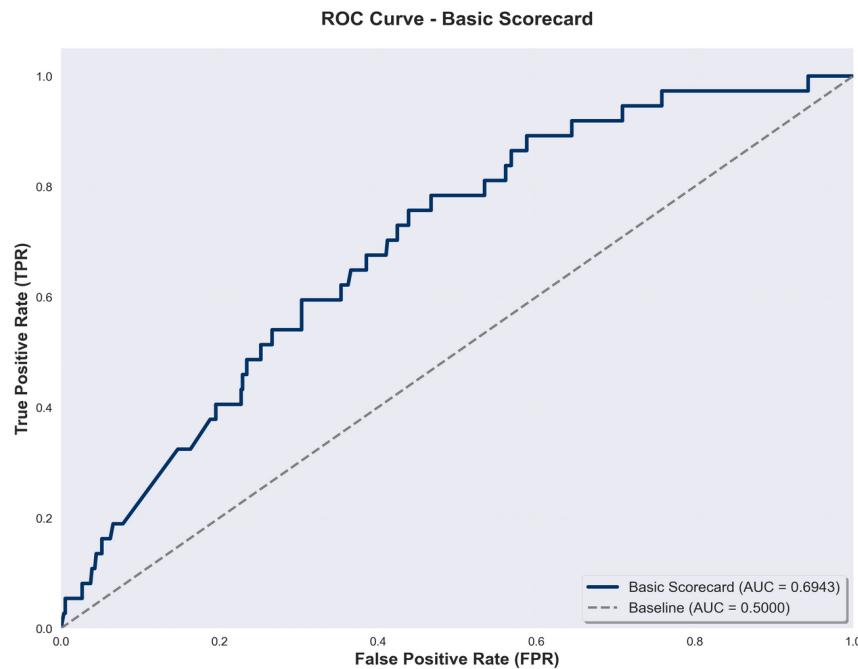
Lp	Zmienna	IV	PSI	Kategoria
1	wsk_ogolnego_zadluzenia_1	0.56	0.001	Zadłużenie
2	wsk_ebitda_3	0.39	0.003	Rentowność
3	wsk_koszty_przychody	0.29	0.002	Efektywność
4	wsk_udzial_kap_wlasnego...	0.26	0.003	Kapitał
5	Naleznosci_dostaw_uslug ...	0.25	0.001	Working Cap
6	Zapasy	0.18	0.002	Płynność

Wszystkie zmienne mają IV > 0.10 (silna predykcja) oraz PSI < 0.01 (wysoka stabilność).

Wyniki na zbiorze testowym

Metryka	Wartość	Interpretacja
AUC (ROC)	0.6943	Dobra zdolność dyskryminacyjna
Gini	0.3886	Standard rynkowy dla modeli behawioralnych
KS Statistic	0.3180*	Maksymalna separacja Good/Bad
Brier Score	0.2227	Niska strata probabilistyczna
PR-AUC	0.1324	Precyza dla klasy mniejszościowej (Default)

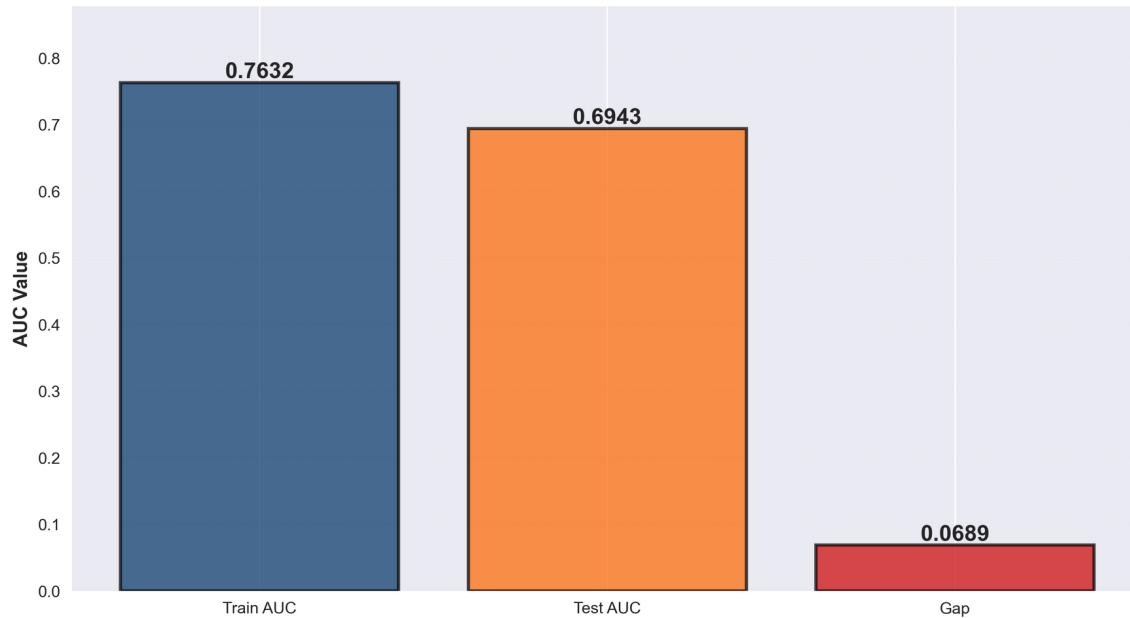
Wyniki modelu - Metryki wydajności



Metryka	Wartość	Interpretacja
AUC (ROC)	0.6943	Dobra dyskryminacja (>0.65)
Gini	0.3886	Standard w credit scoring
KS Statistic	0.3180	Silna separacja (>0.30)
Brier Score	0.2227	Niska strata probabilistyczna
PR-AUC	0.1324	Akceptowalne

Generalizacja modelu - Train vs Test

Train vs Test AUC - Generalization Analysis

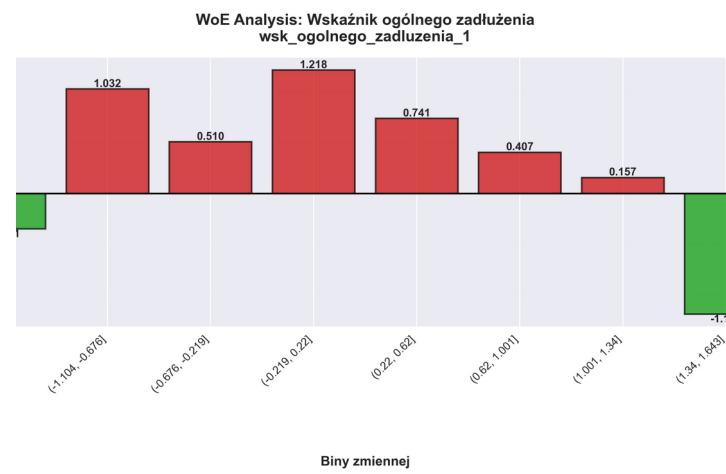
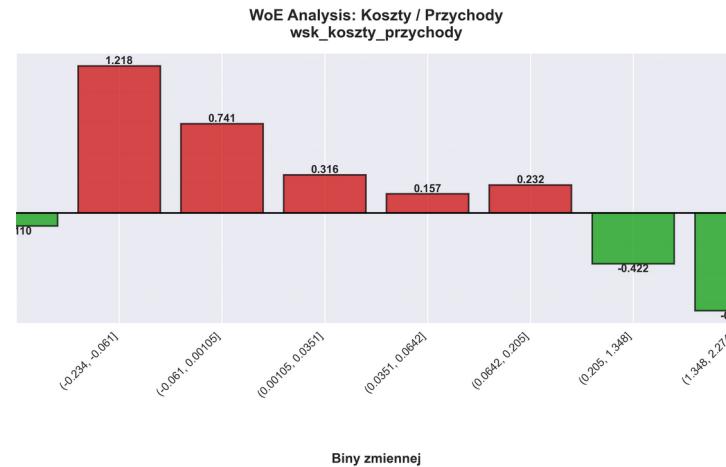
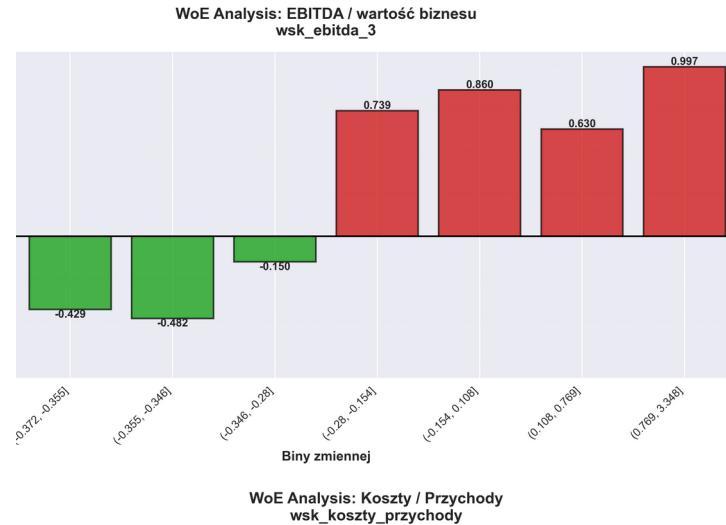


Train AUC	Test AUC	Gap	Status
0.7632	0.6943	0.0689 (9%)	Minimalny Overfitting

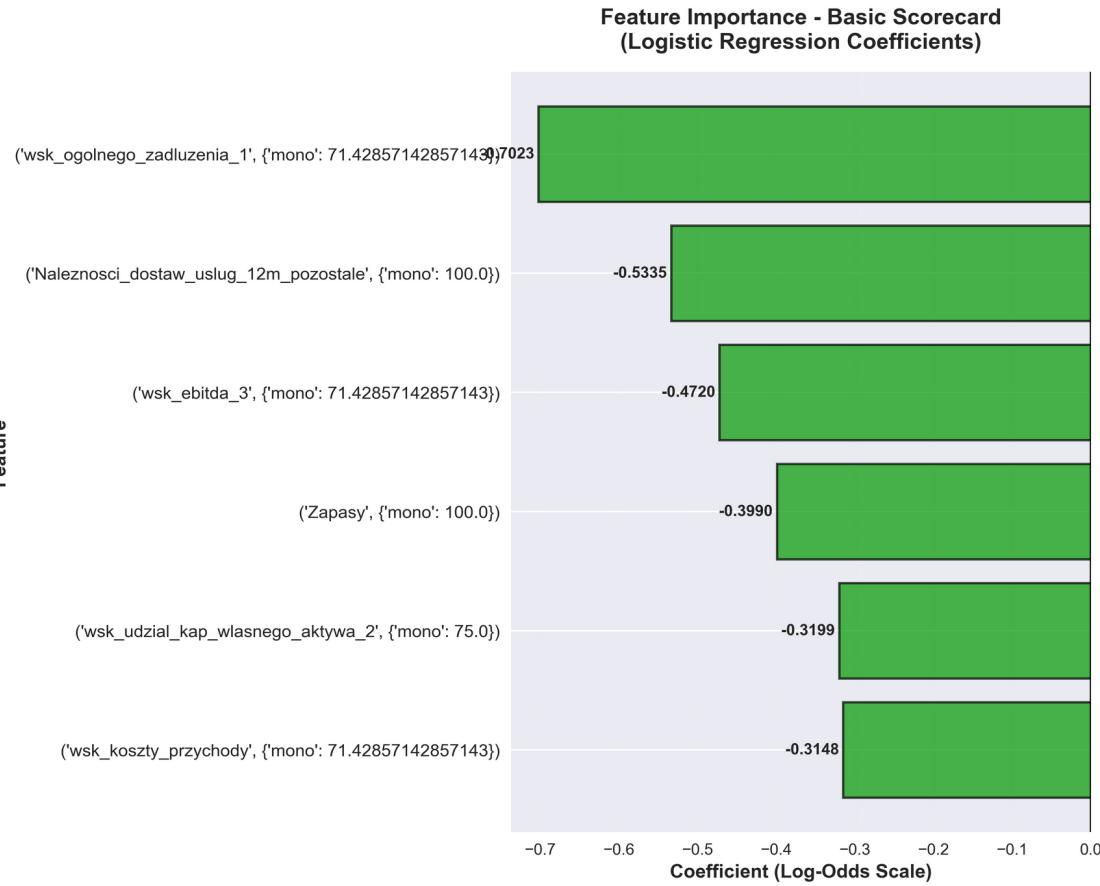
Interpretacja zmiennych

- Analiza WoE

- Zadłużenie (L): Wyższe zadłużenie → Wyższe ryzyko (rosnące WoE).
- EBITDA (Śr): Wyższa rentowność → Niższe ryzyko (malejące WoE).
- Koszty/Przychody (P): Wyższe koszty → Wyższe ryzyko.



Wpływ zmiennych na decyzję



Interpretacja:

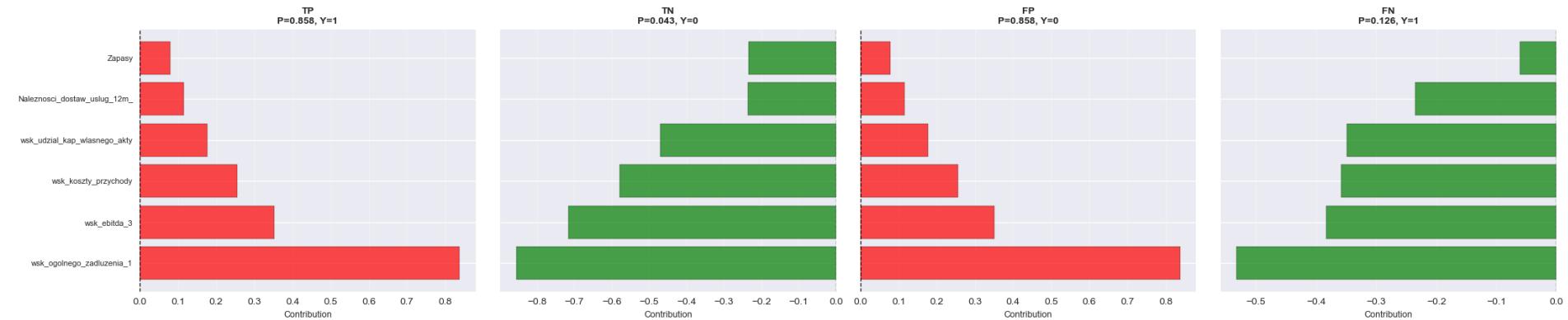
Wysokie zadłużenie (Coef +0.85) drastycznie zwiększa ryzyko.

Wysoka EBITDA (Coef -0.62) silnie obniża ryzyko.

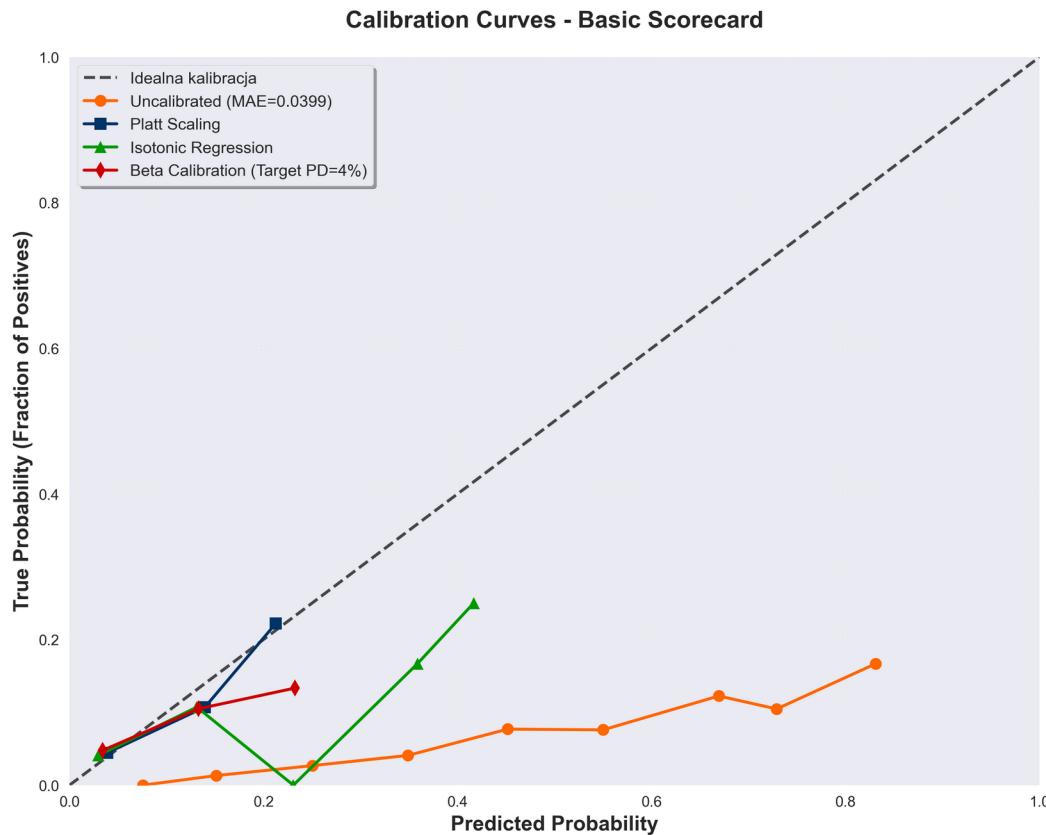
Kierunek wpływu zmiennych jest zgodny z intuicją biznesową.

Case studies

BASIC Scorecard - Porównanie Wkładów Cech



Kalibracja prawdopodobieństw



Metoda: Beta Calibration

Target PD: 4%

Wyniki MAE:

Train MAE: 0.0372

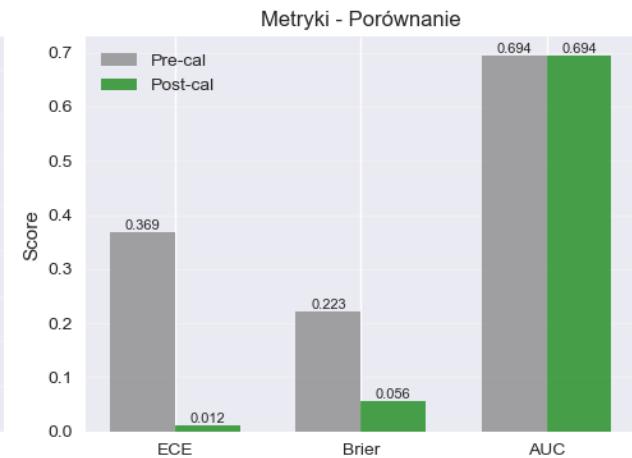
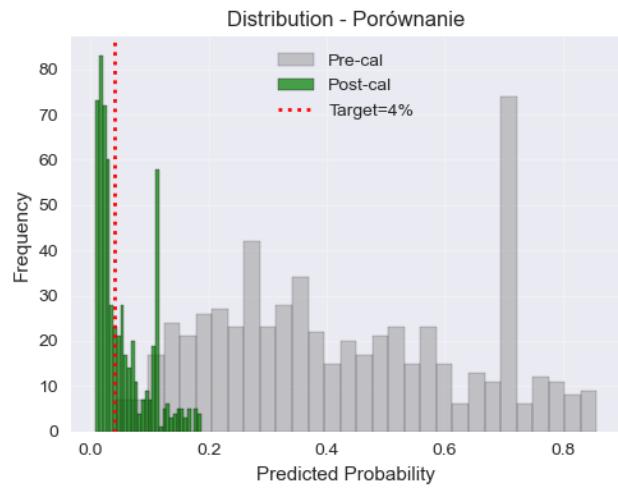
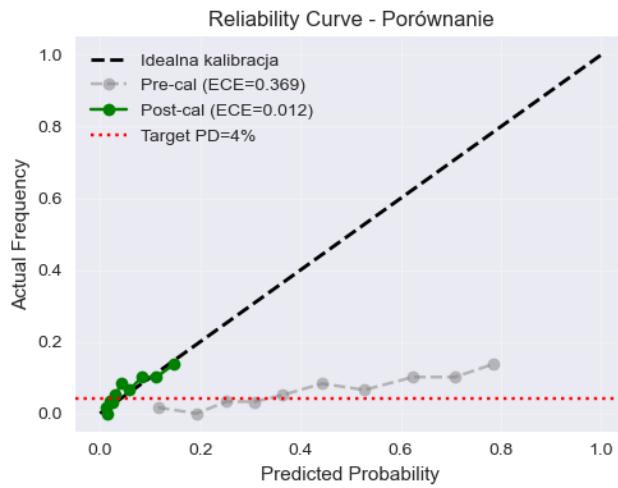
Test MAE: 0.0399

Gap: 0.0027

Wniosek:

Doskonała kalibracja.

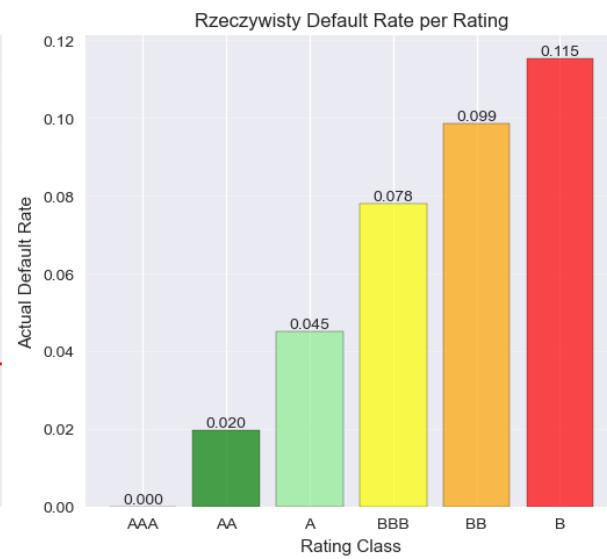
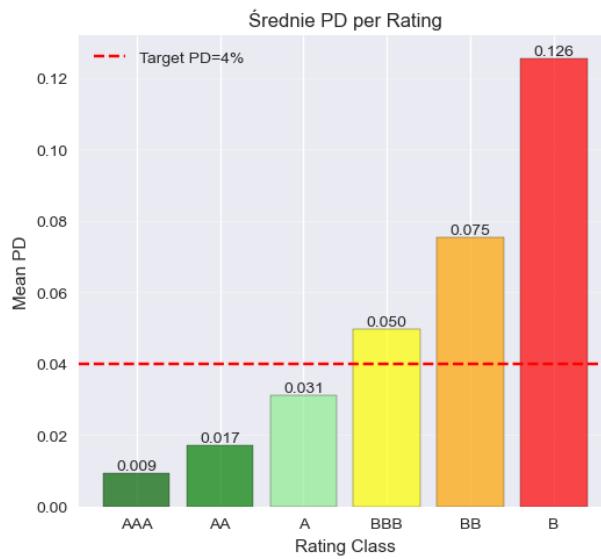
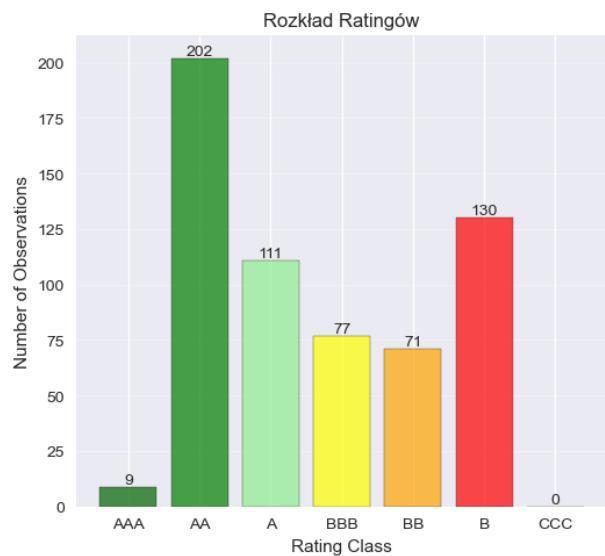
C.D



Rating Scorecard - Mapping AAA-D

Rating	PD Range	Opis	Rekomendacja
AAA	0.0-0.5%	Minimalne	Auto-Akceptacja
AA-A	0.5-2.0%	Niskie	Standard
BBB	2.0-4.0%	Umiarkowane	Analiza
BB	4.0-8.0%	Podwyższone	Zabezpieczenia
B	8.0-15.0%	Wysokie	Silne zabezp.
CCC	15.0-30.0%	Bardzo wysokie	Restrukturyzacja
D	>75.0%	Default	Odrzucenie

C.D



Podsumowanie i wnioski

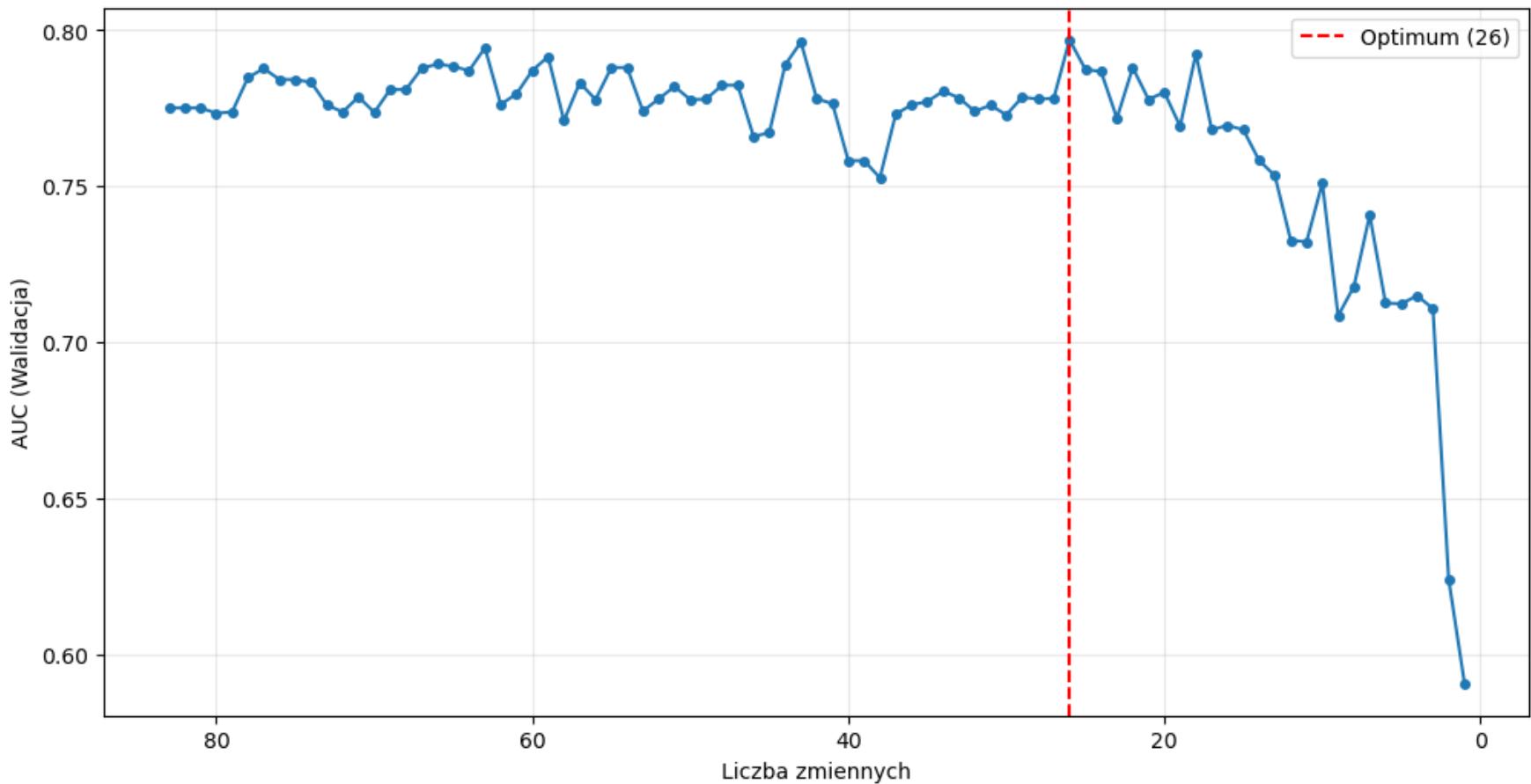
- Jakość modelu: AUC = 0.6943 (Dobry wynik).
- Separacja: KS = 0.3180 (Silna dyskryminacja).
- Kalibracja: Gap MAE = 0.0027 (Bardzo dokładne PD).
- Stabilność: PSI < 0.01 dla wszystkich zmiennych.
- Interpretowalność: Pełna przejrzystość dzięki WoE.
- Gotowość: Model spełnia wymogi do wdrożenia.

Metryka Kalibracji	Przed Kalibracją	Po Kalibracji	Zmiana (Δ)
Mean PD (Średnie ryzyko)	43.05%	5.42%	▼ 37.6 p.p. ✓
ECE (Błąd kalibracji)	0.3688	0.0119	▼ 0.3569 ✓
Brier Score	0.2227	0.0565	▼ 0.1662 ✓

Proces obróbki danych dla modelu blackbox

- Tworzenie nowych zmiennych i wskaźników
- Zastąpienie wartości nieskończonych brakami, a braki medianą
- Winsoryzacja i Yeo-Johnson
- Anomaly score na podstawie Isolation Forrest
- Użycie PCA do obliczenia dźwigni obserwacji
- VIF do wstępnego zawężenia cech
- SVD dla maksymalizacji ekstrakcji informacji z szumu
- RFE w celu redukcji wymiarowości

RFE Path: Validation AUC vs Number of Features



Wyniki RFE

Usuwamy zbędne kolumny
które wprowadzają szum

Model black box

Zostało przetestowanych wiele rozwiązań

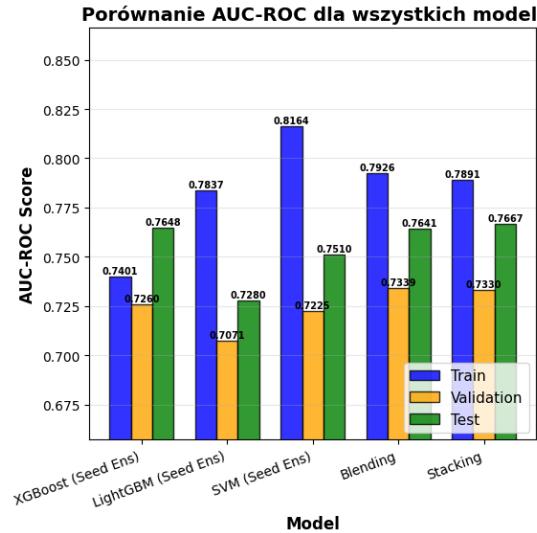
- LightGMB
- XGBoost
- SVM
- Blending
- Stacking

Wyniki Black box

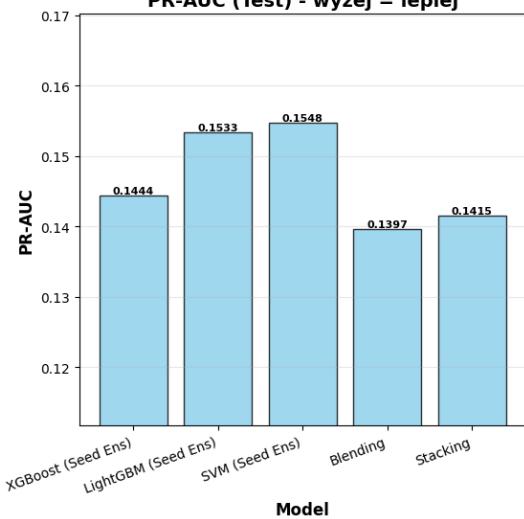
Model	AUC (Tr)	AUC (Val)	AUC (Ts)	PR AUC (Tr)	PR AUC (Val)	PR AUC (Ts)	KS (Tr)	KS (Val)	KS (Ts)
XGBoost (Seed Ens)	0.7401	0.7260	0.7648	0.1505	0.1952	0.1444	0.3983	0.3986	0.4468
LightGBM (Seed Ens)	0.7837	0.7071	0.7280	0.2252	0.1449	0.1533	0.4555	0.4232	0.4043
SVM (Seed Ens)	0.8164	0.7225	0.7510	0.2200	0.1528	0.1548	0.4883	0.3976	0.4297
Blending	0.7926	0.7339	0.7641	0.2076	0.1637	0.1397	0.4814	0.4872	0.4622
Stacking	0.7891	0.7330	0.7667	0.2018	0.1755	0.1415	0.4797	0.4818	0.4639

Porównanie wyników: Seed Ensemble (20 seedów) + Blending + Stacking

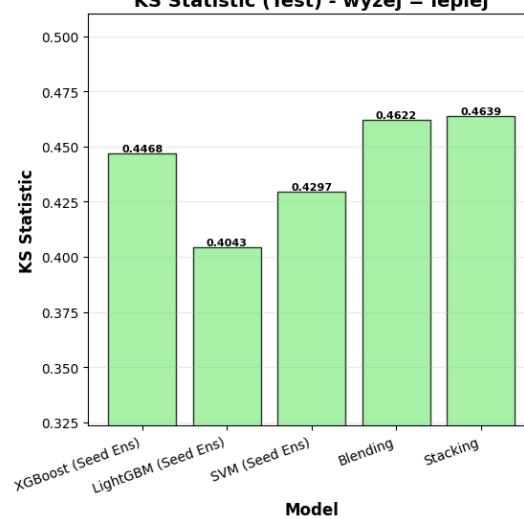
Porównanie AUC-ROC dla wszystkich modeli



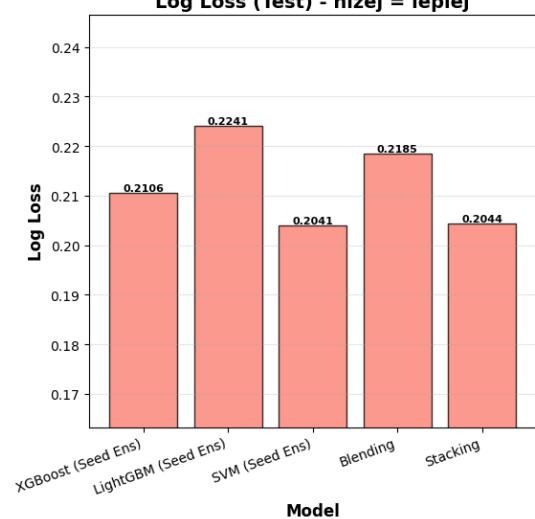
PR-AUC (Test) - wyżej = lepiej



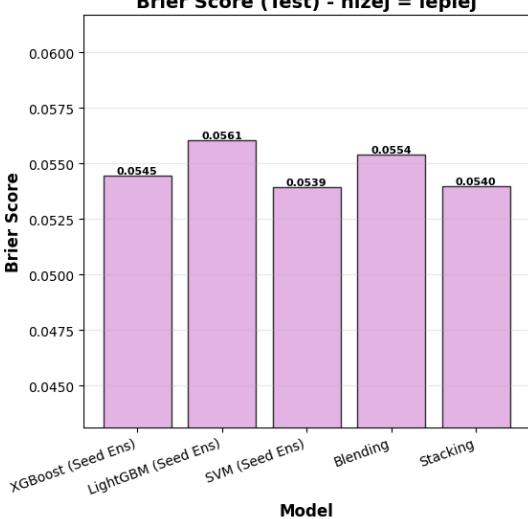
KS Statistic (Test) - wyżej = lepiej



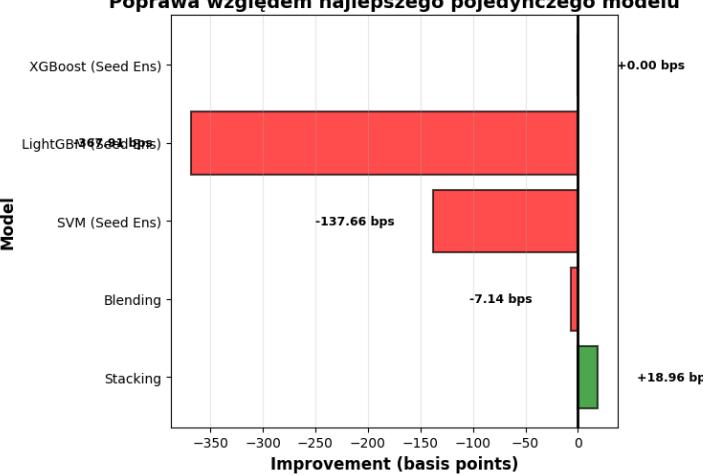
Log Loss (Test) - niżej = lepiej



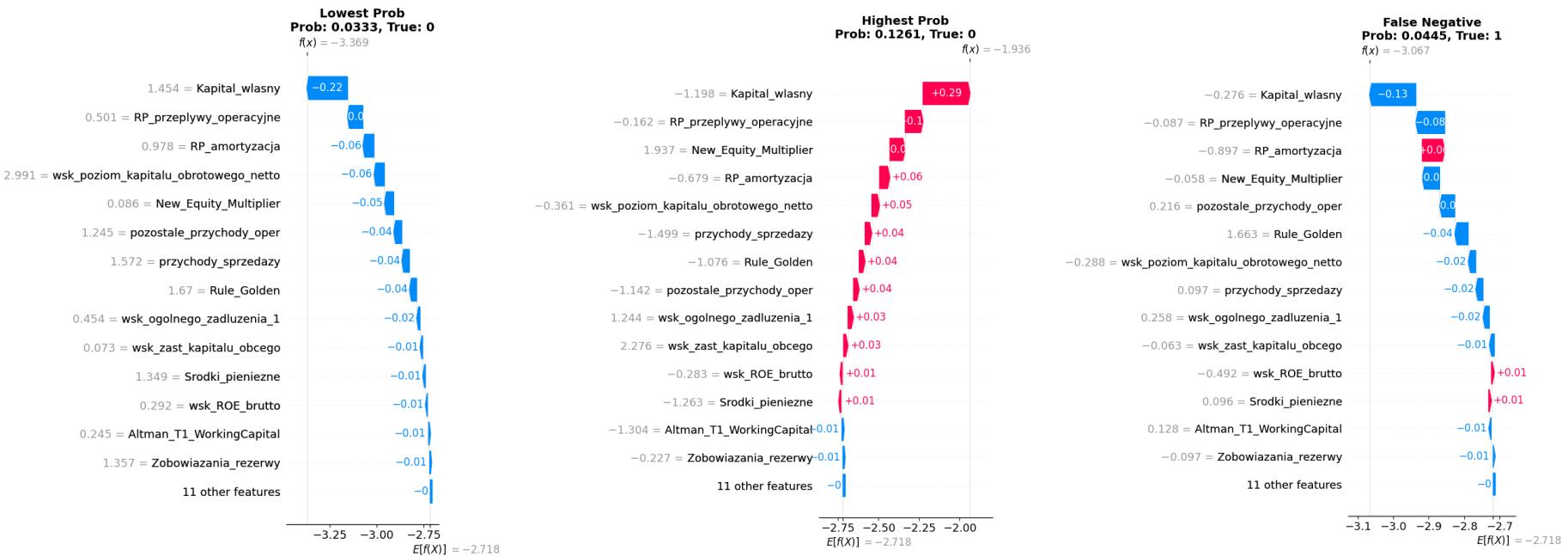
Brier Score (Test) - niżej = lepiej



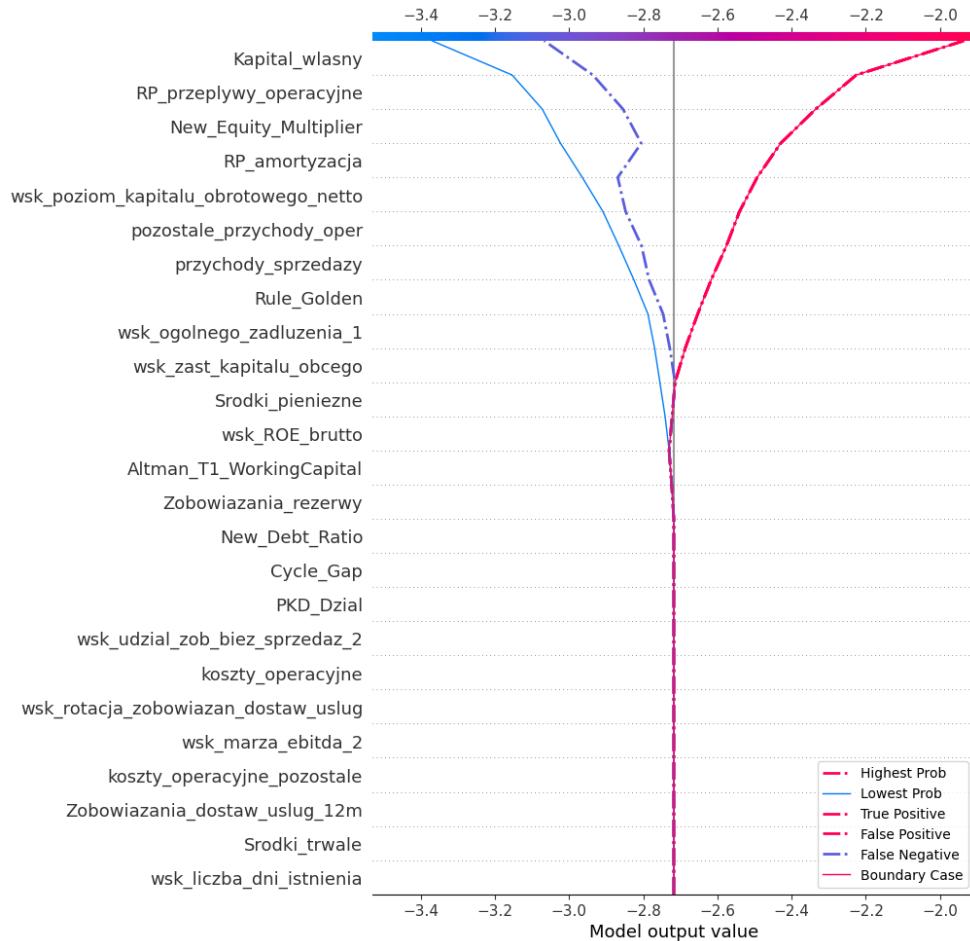
Poprawa względem najlepszego pojedynczego modelu



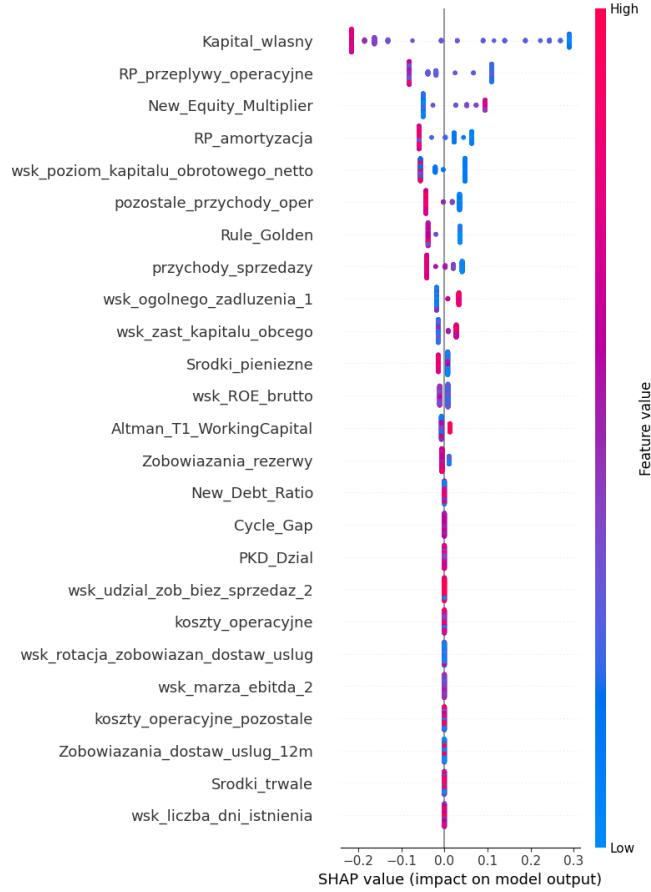
Analiza globalna i lokalna



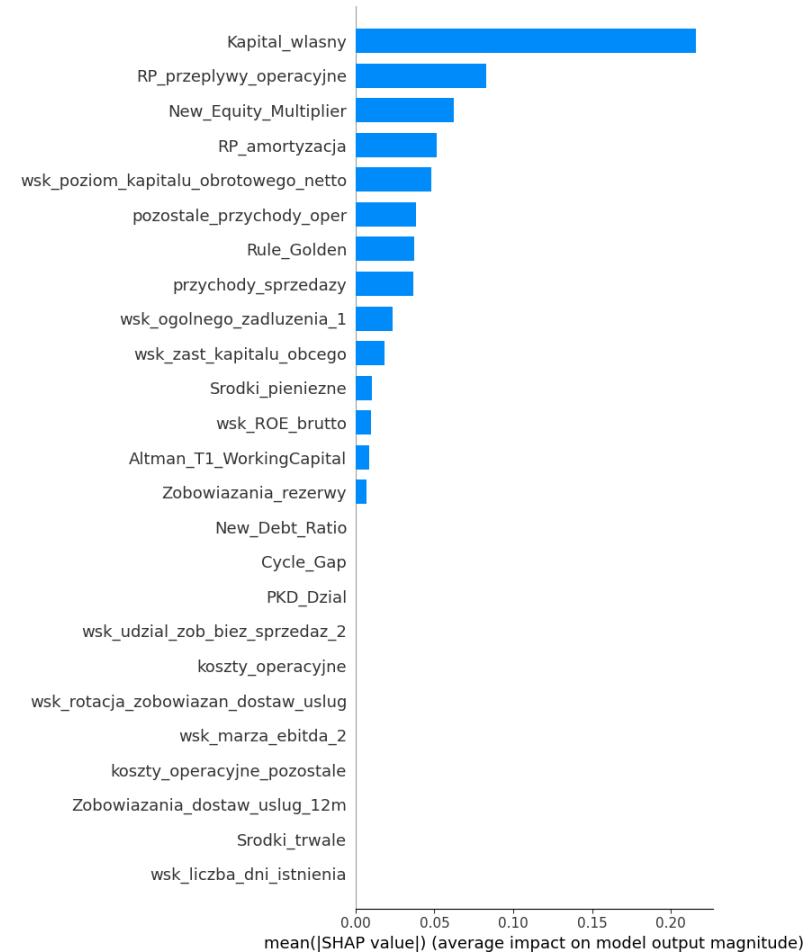
Granice progów decyzyjnych

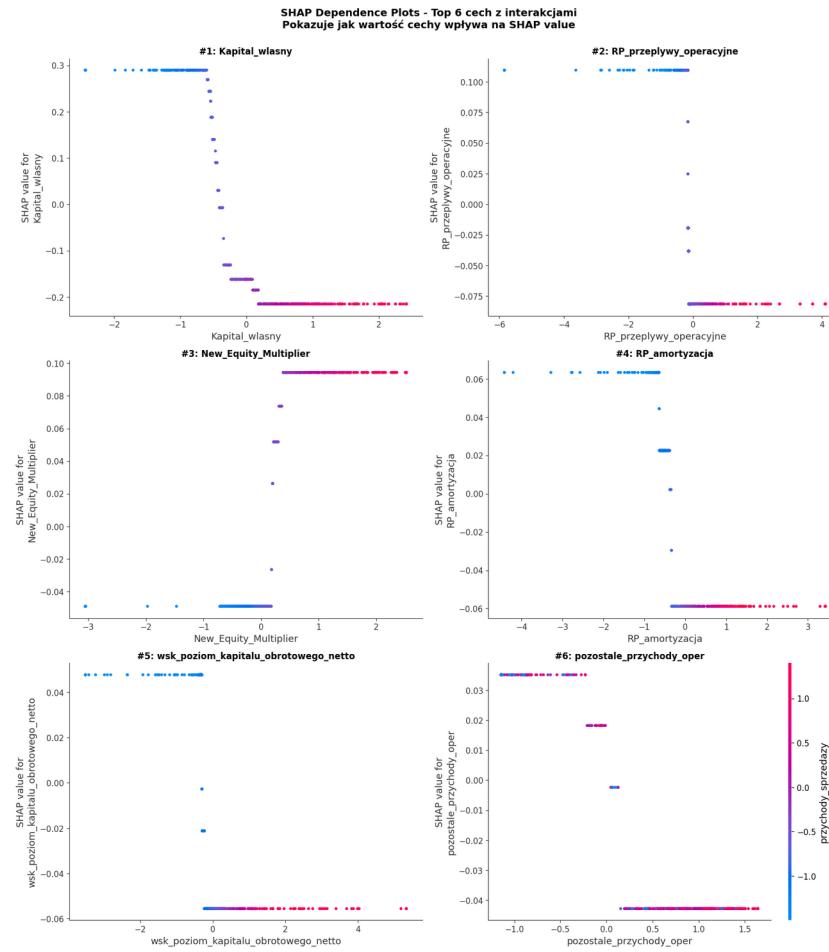
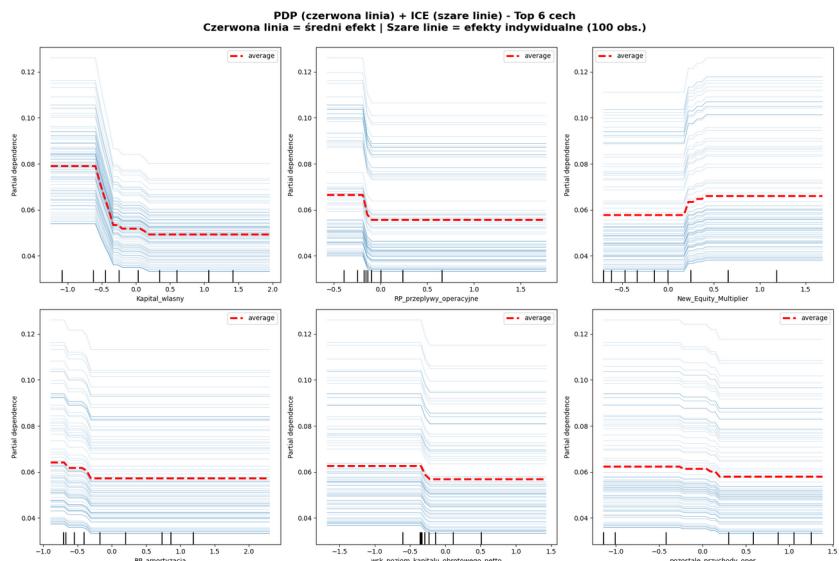


SHAP Summary Plot - Top 30 cech
Każdy punkt to jedna obserwacja, kolor oznacza wartość cechy

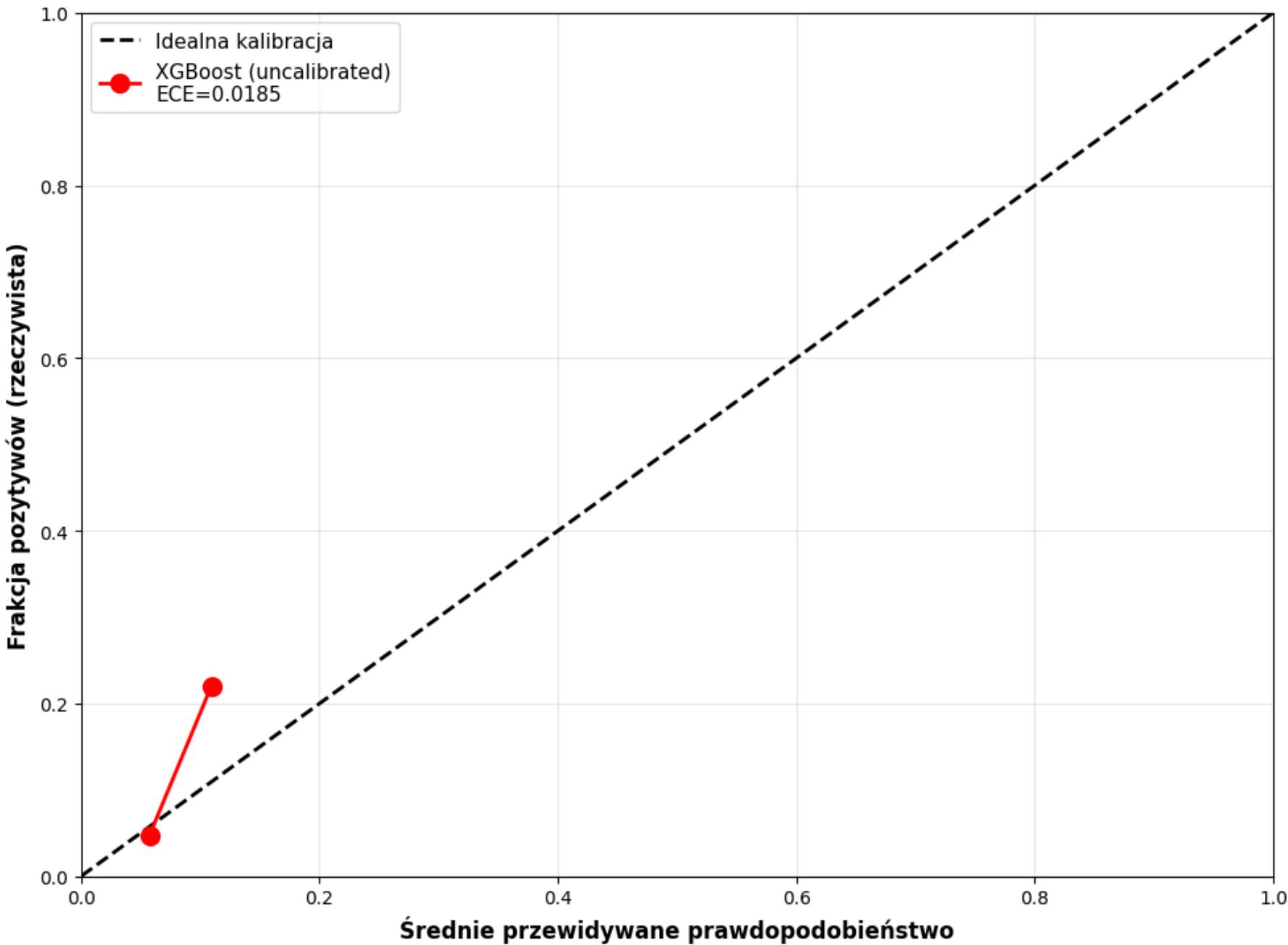


SHAP Bar Plot - Top 30 najważniejszych cech globalnie
Średnia wartość bezwzględna SHAP

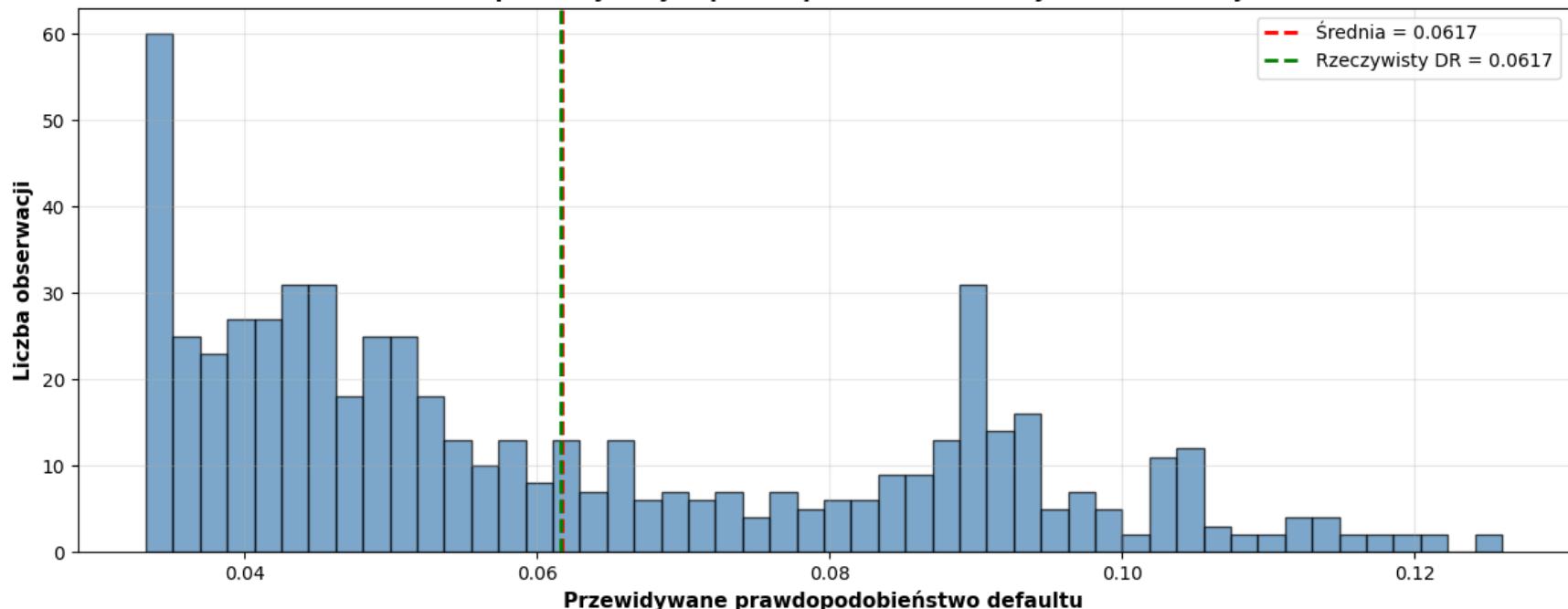




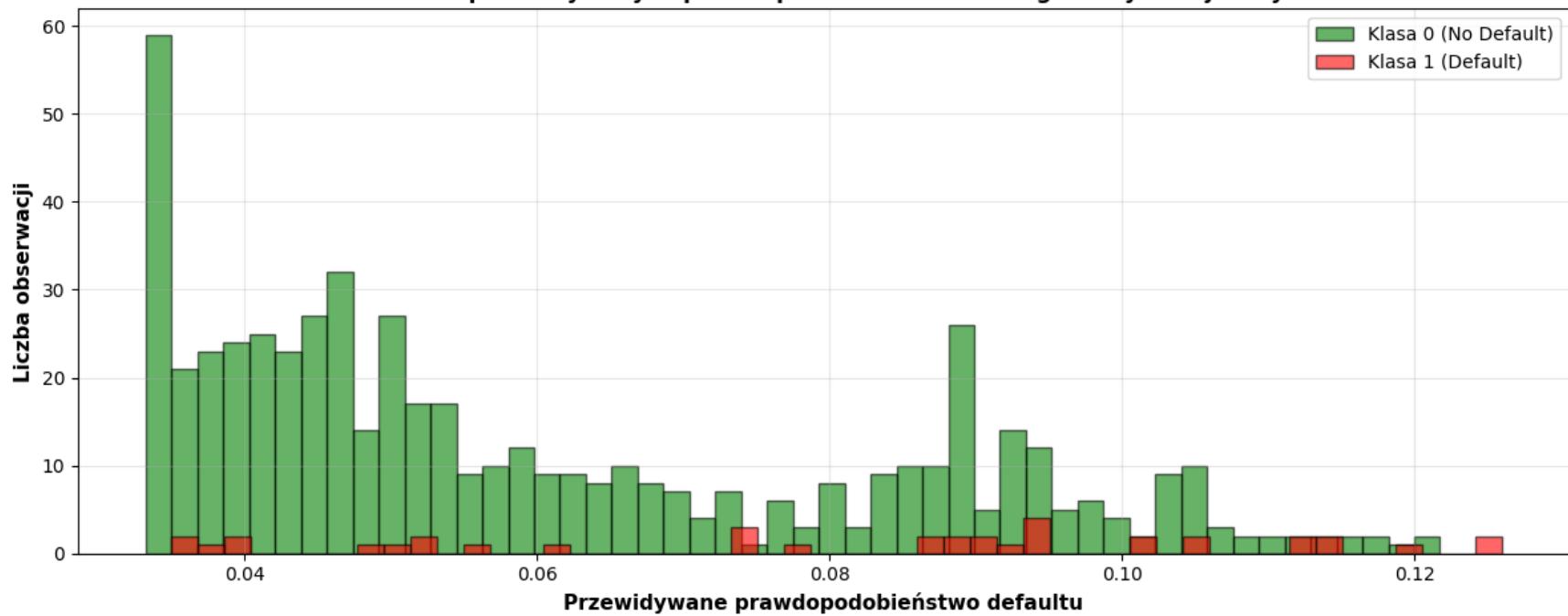
Reliability Curve - Kalibracja prawdopodobieństw przed korektą Im bliżej linii diagonalnej, tym lepsza kalibracja



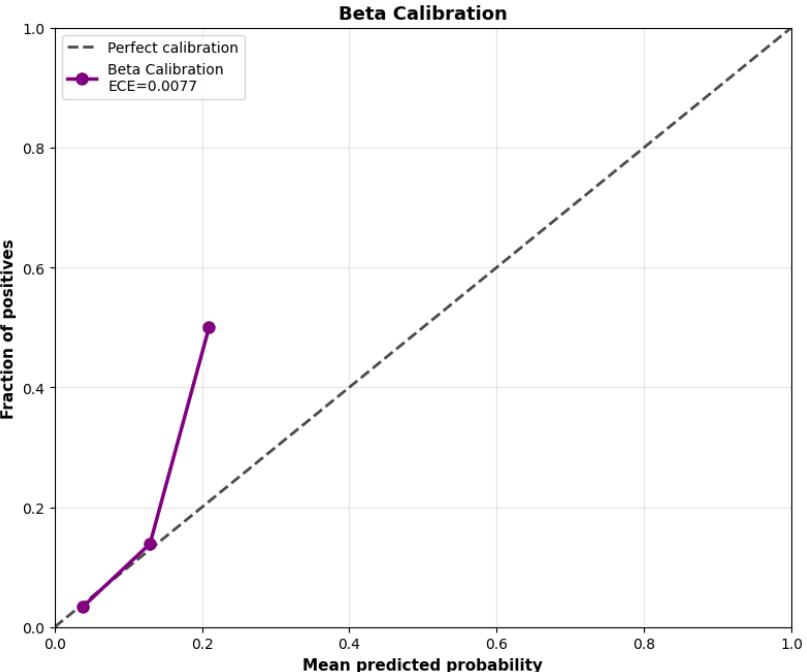
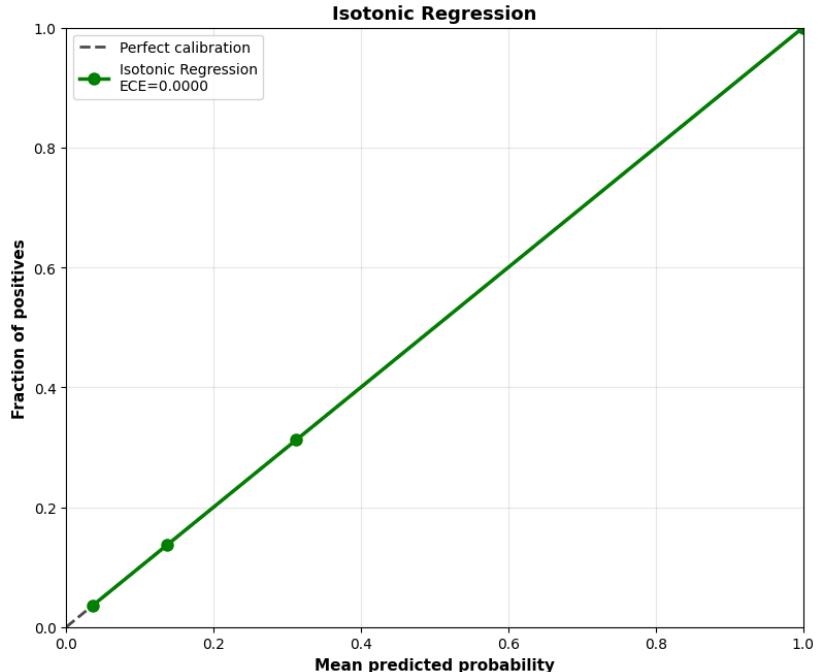
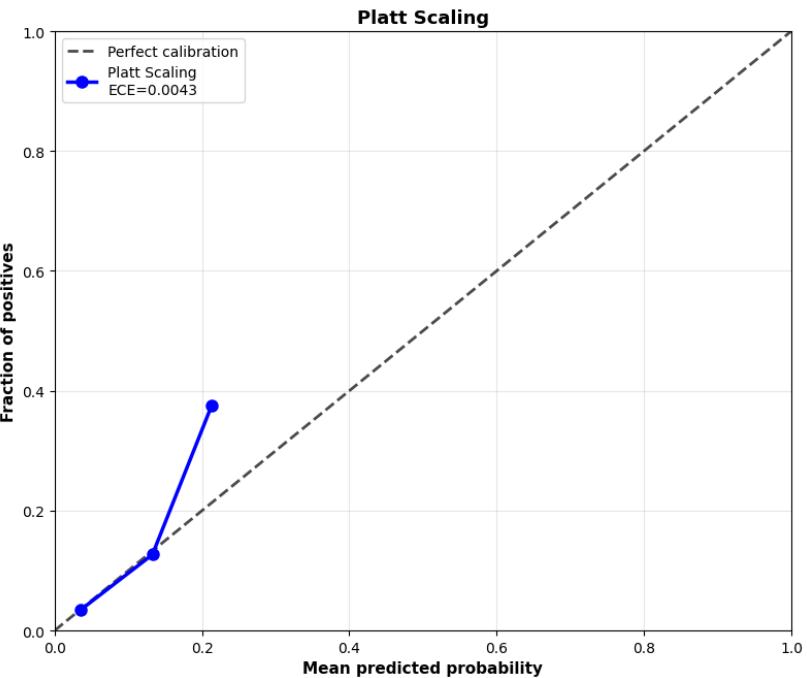
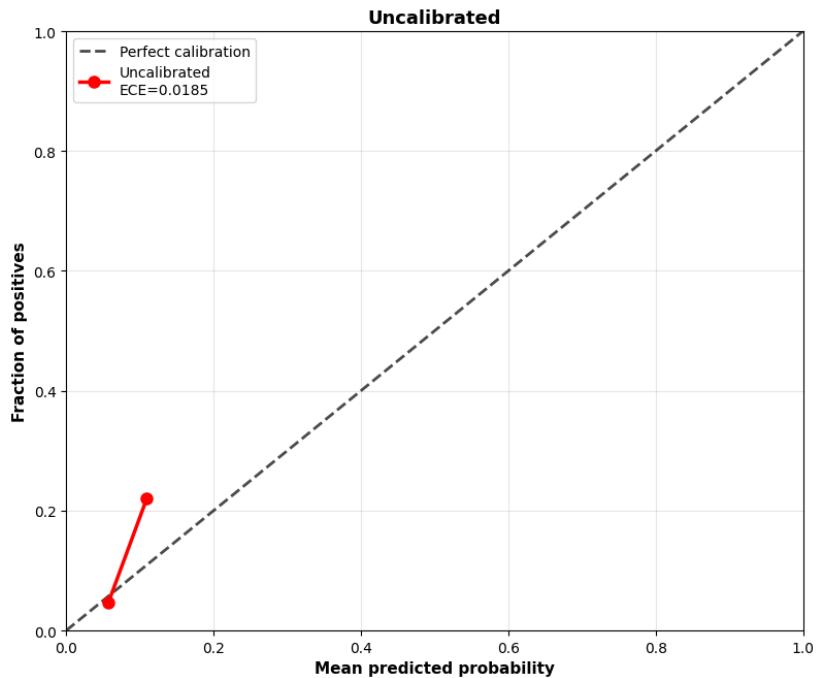
Rozkład przewidywanych prawdopodobieństw - wszystkie obserwacje



Rozkład przewidywanych prawdopodobieństw - według rzeczywistej klasy

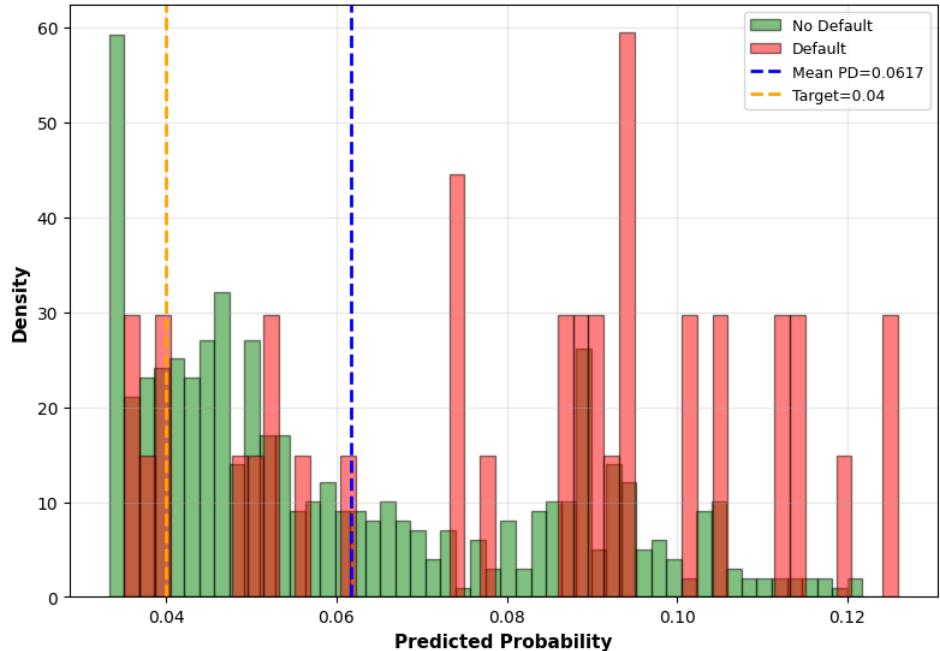


Reliability Curves - Porównanie metod kalibracji Im bliżej diagonali, tym lepsza kalibracja

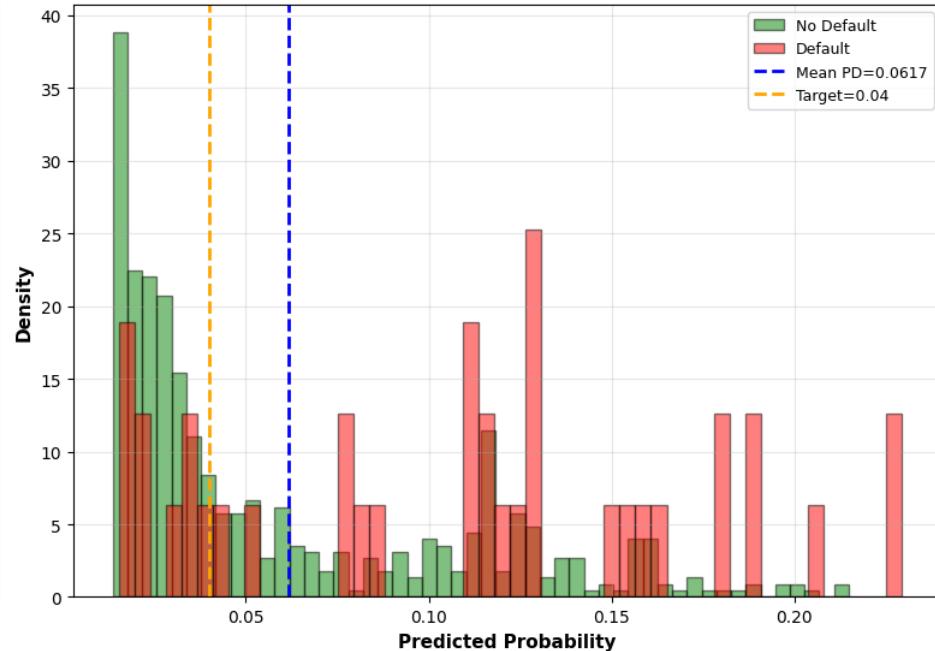


Rozkład przewidywanych prawdopodobieństw - Porównanie metod

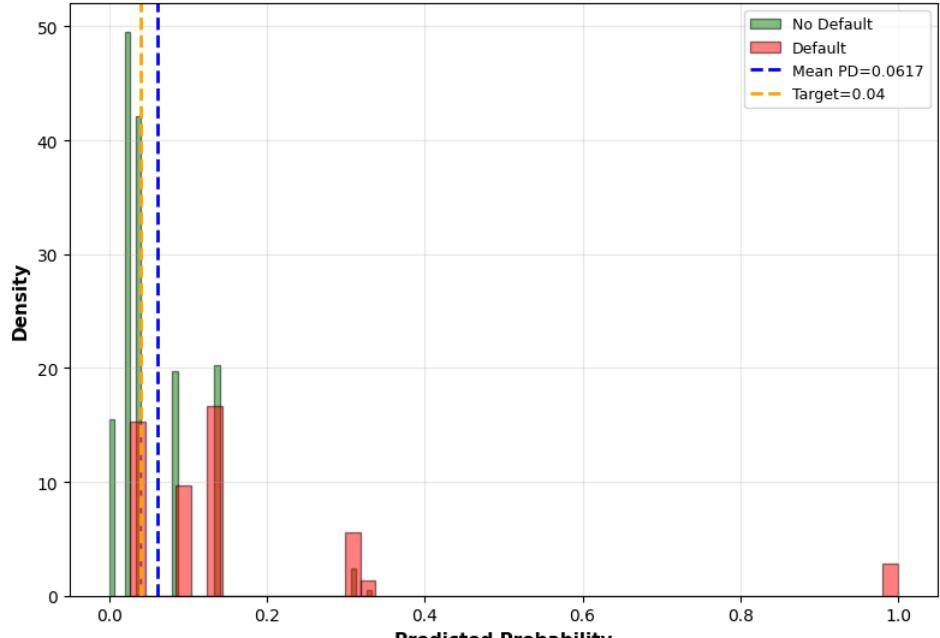
Uncalibrated



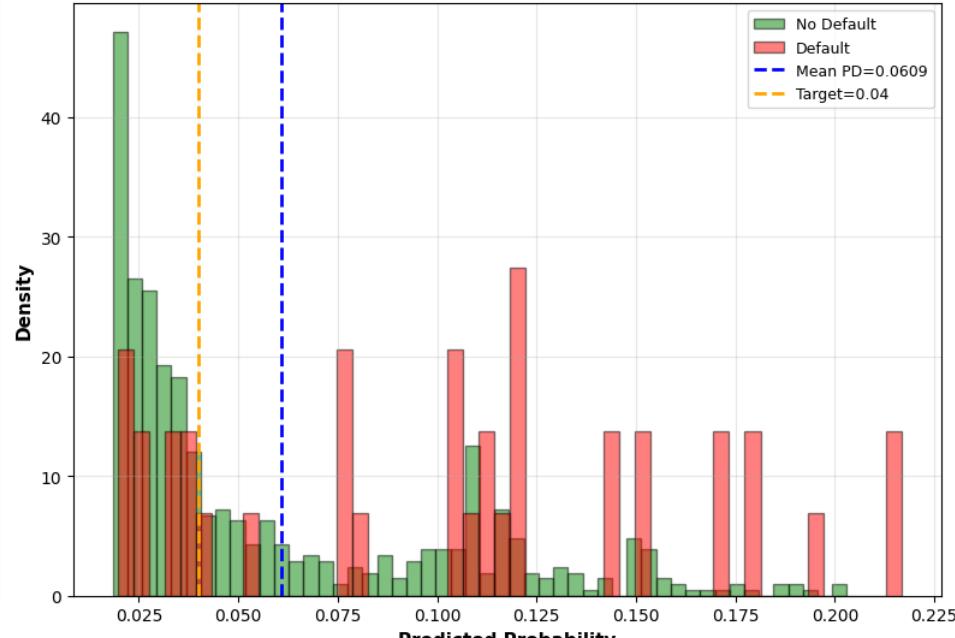
Platt Scaling



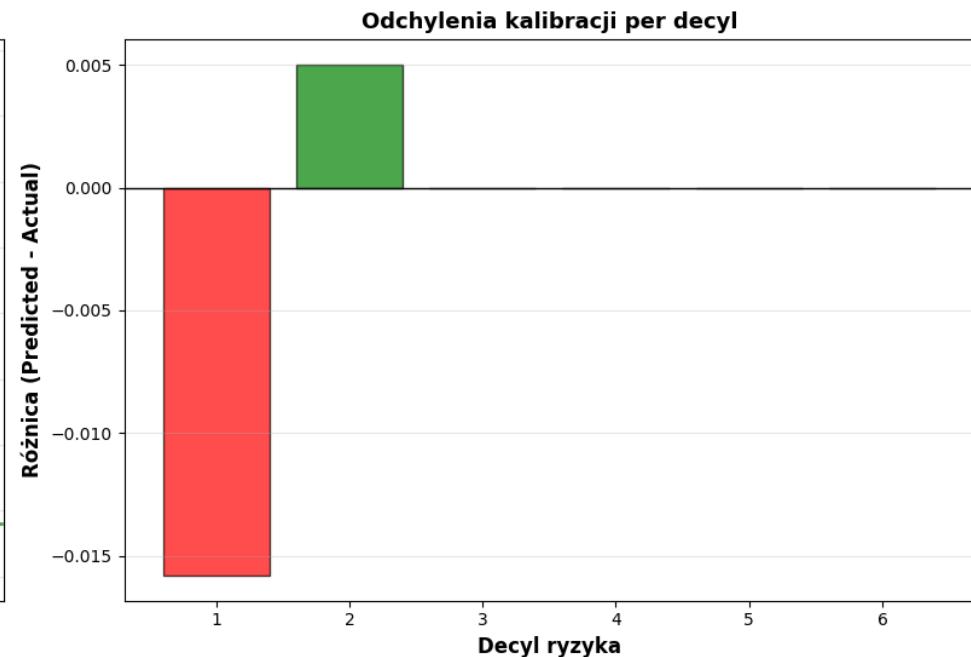
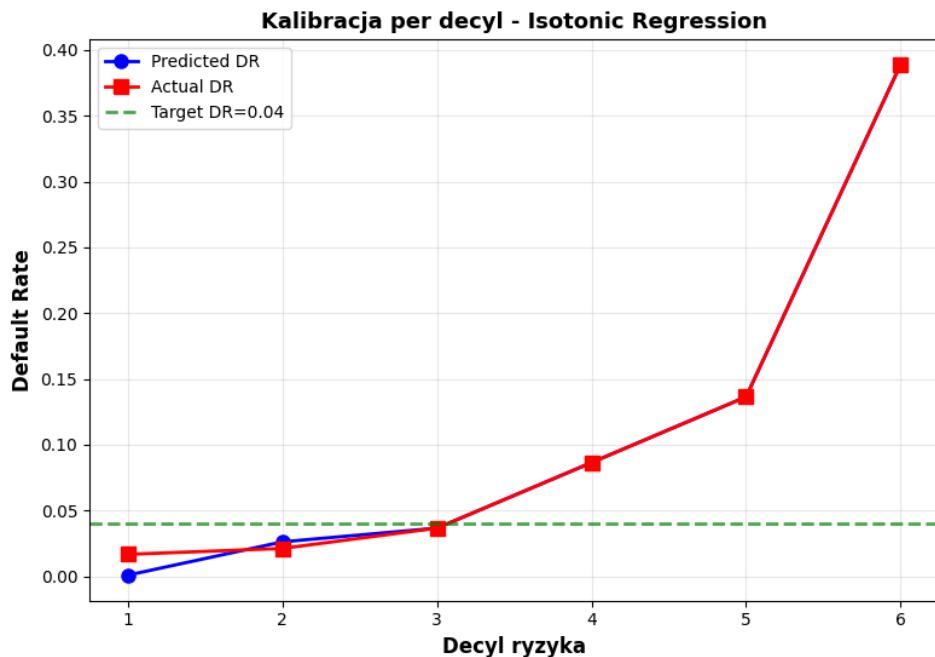
Isotonic Regression



Beta Calibration



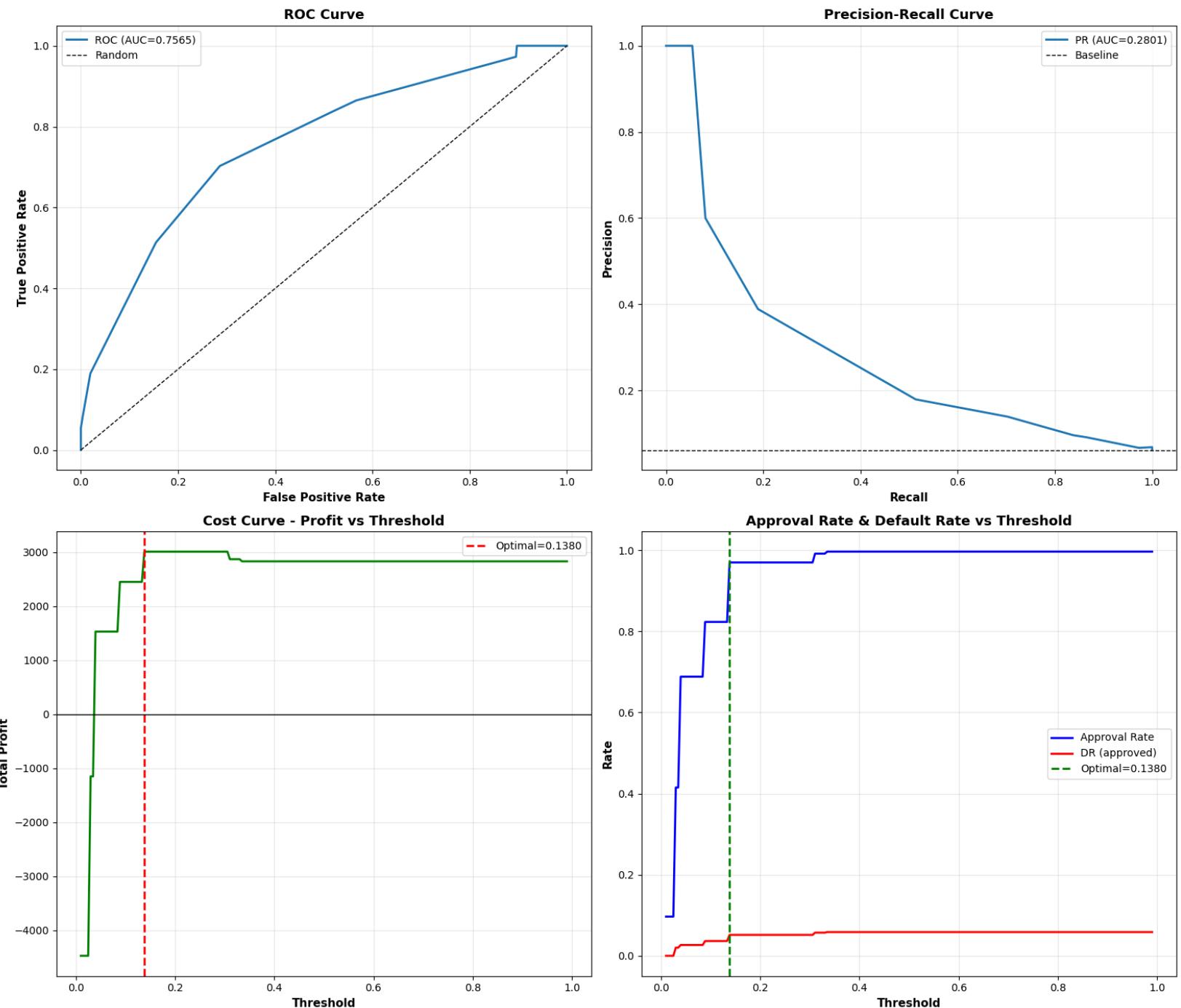
Stabilność kalibracji według podgrup - Isotonic Regression



Funkcja straty

	Spłaci	Nie spłaci
Przewidywanie – Spłaci	10	-80
Przewidywanie – Nie Spłaci	-10	0

Analiza progów decyzyjnych - ROC, PR, Cost

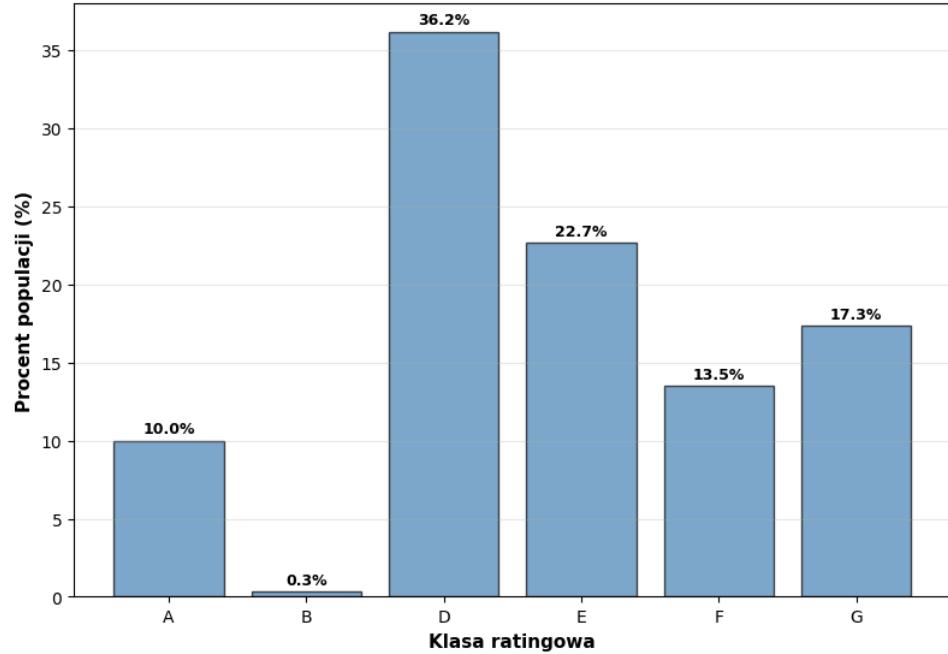


Strategia dla każdej kategorii

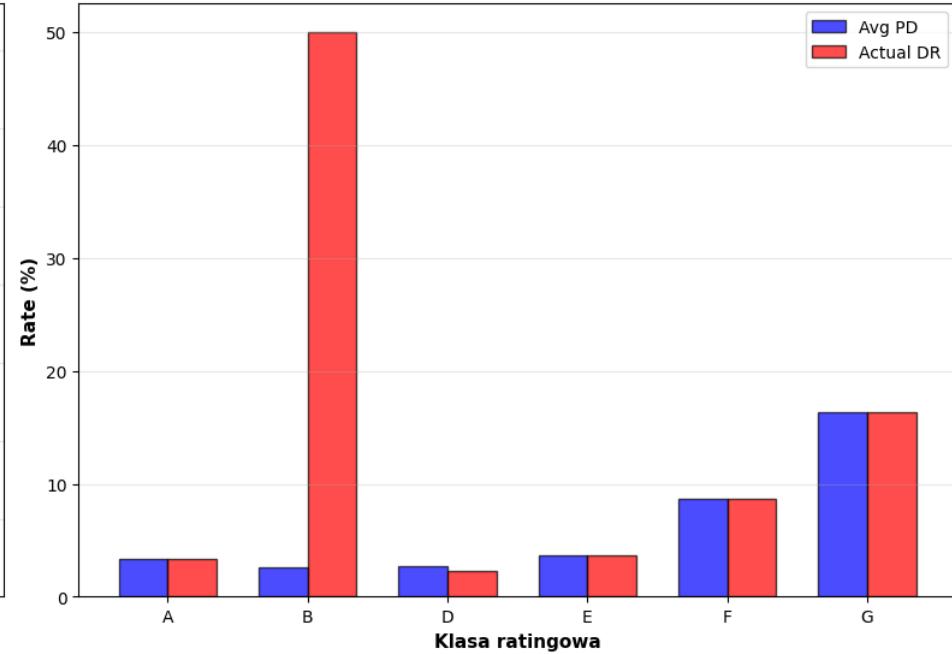
Rating	Średnie PD	Faktyczne PD	Akcja	Zasady Umowy
A	3.33%	3.33%	AUTO-ACCEPT	Standard terms
B	2.62%	50.00%	AUTO-ACCEPT	Standard terms
D	2.74%	2.30%	ACCEPT	Standardowe zasady, możliwe wyższe oprocentowanie
E	3.68%	3.68%	MANUAL REVIEW	Wyższe oprocentowanie
F	8.64%	8.64%	REJECT lub MANUAL	Wysokie oprocentowanie bądź zastaw
G	16.35%	16.35%	AUTO-REJECT	Brak

Analiza klas ratingowych - charakterystyki

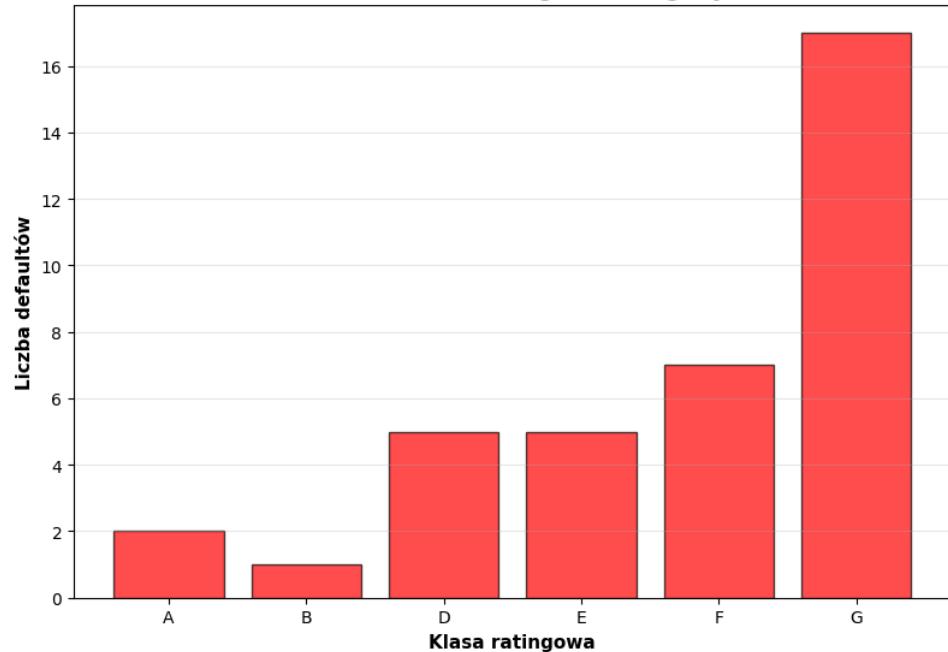
Rozkład populacji według klas ratingowych



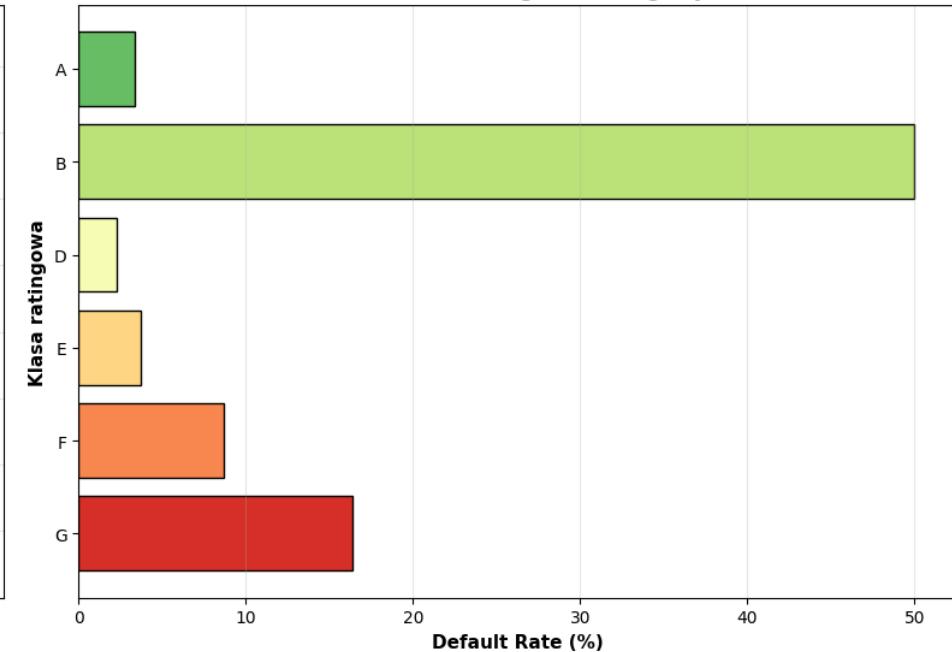
Porównanie PD vs Actual DR



Liczba defaultów według klas ratingowych



Default Rate według klas ratingowych



Wpływ Progu na Zysk i Ryzyko

