

## PODSUMOWANIE RAPORTU

<b>Najlepszy model</b>	<b>Regresja Logistyczna</b>
<b>Wynik walidacji</b>	<b>0.7948</b>
<b>AUROC</b>	<b>0.5258</b>
<b>Gini</b>	<b>0.0516</b>

## Raport klasyfikacji

	Precision	Recall	F1-Score	Support
0	0.79	1.00	0.89	2282
1	0.00	0.00	0.00	589
macro avg	0.40	0.50	0.44	2871
weighted avg	0.63	0.79	0.70	2871

## Statystyki opisowe danych

Feature	mean	std	min	25%	50%	75%	max
CreditScore	650.53	96.65	350.0	584.0	652.0	718.0	850.0
Geography	nan	nan	nan	nan	nan	nan	nan
Gender	nan	nan	nan	nan	nan	nan	nan
Age	38.92	10.49	18.0	32.0	37.0	44.0	92.0
Tenure	5.01	2.89	0.0	3.0	5.0	7.0	10.0
Balance	76485.89	62397.41	0.0	0.0	97198.54	127644.24	250898.09
NumOfProducts	1.53	0.58	1.0	1.0	1.0	2.0	4.0
HasCrCard	0.71	0.46	0.0	0.0	1.0	1.0	1.0
IsActiveMember	0.52	0.5	0.0	0.0	1.0	1.0	1.0
EstimatedSalary	100090.24	57510.49	11.58	51002.11	100193.92	149388.25	199992.48
Exited	0.2	0.4	0.0	0.0	0.0	0.0	1.0
Complain	0.2	0.4	0.0	0.0	0.0	0.0	1.0
Satisfaction Score	3.01	1.41	1.0	2.0	3.0	4.0	5.0
Card Type	nan	nan	nan	nan	nan	nan	nan
Point Earned	606.52	225.92	119.0	410.0	605.0	801.0	1000.0

Profil Klienta na Podstawie Predykcji Modelu

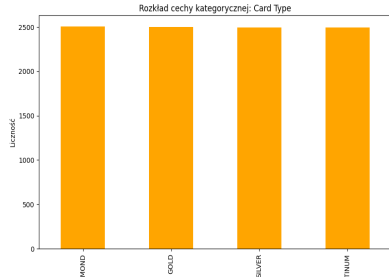
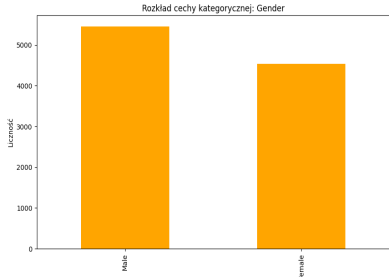
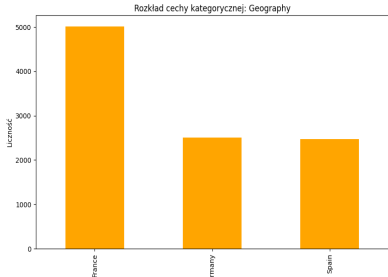
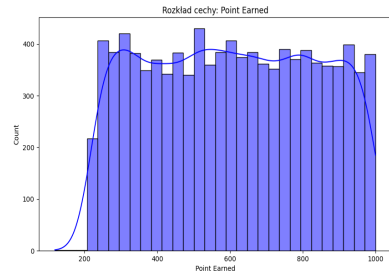
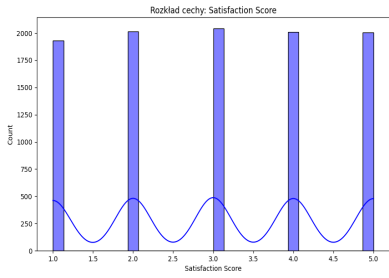
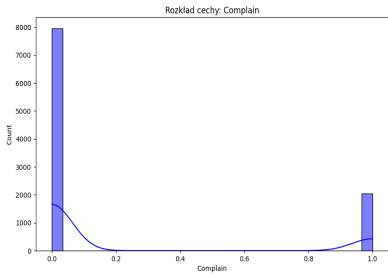
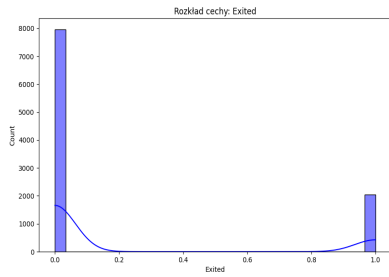
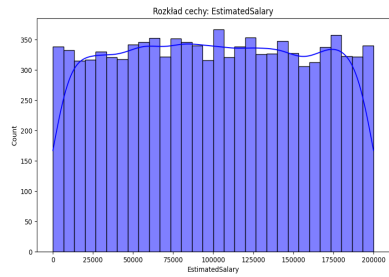
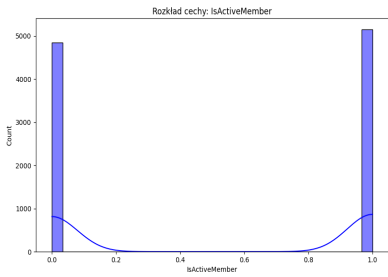
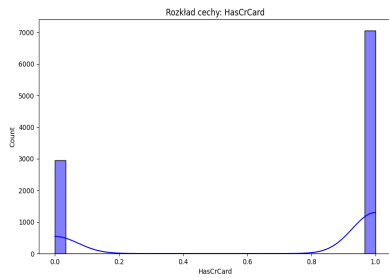
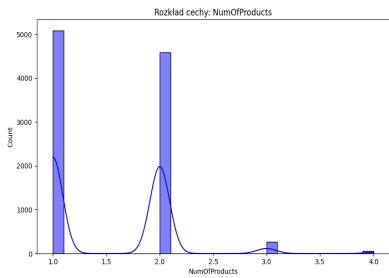
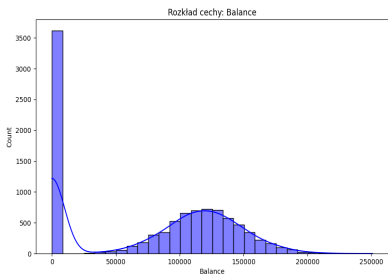
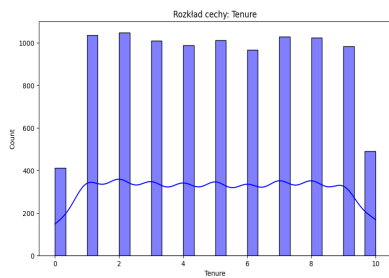
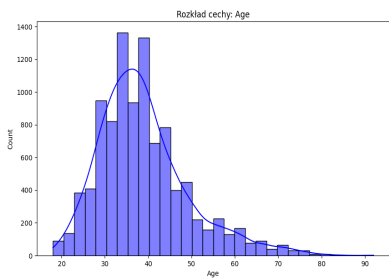
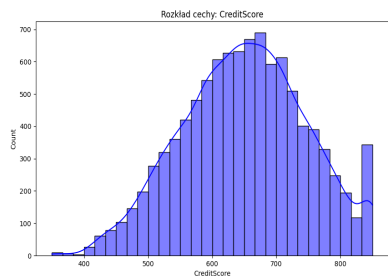
Feature	Overall Average	Predicted High Probability (Target=1)
CreditScore	649.88	669.79
Age	37.75	39.11
Tenure	5.05	4.76
Balance	76266.02	71214.74
NumOfProducts	1.51	1.42
EstimatedSalary	99550.83	108760.69
Satisfaction Score	2.99	2.96
Point Earned	604.95	595.18
Geography_France	0.49	0.55
Geography_Germany	0.25	0.15
Geography_Spain	0.26	0.30
Gender_Male	0.56	0.55
Card Type	1.52	0.00
HasCrCard	0.71	0.71
IsActiveMember	0.50	0.55
Complain	0.19	0.13

Tabela przedstawia znaczenie cech modelu wykorzystywanego do predykcji profilu klienta.

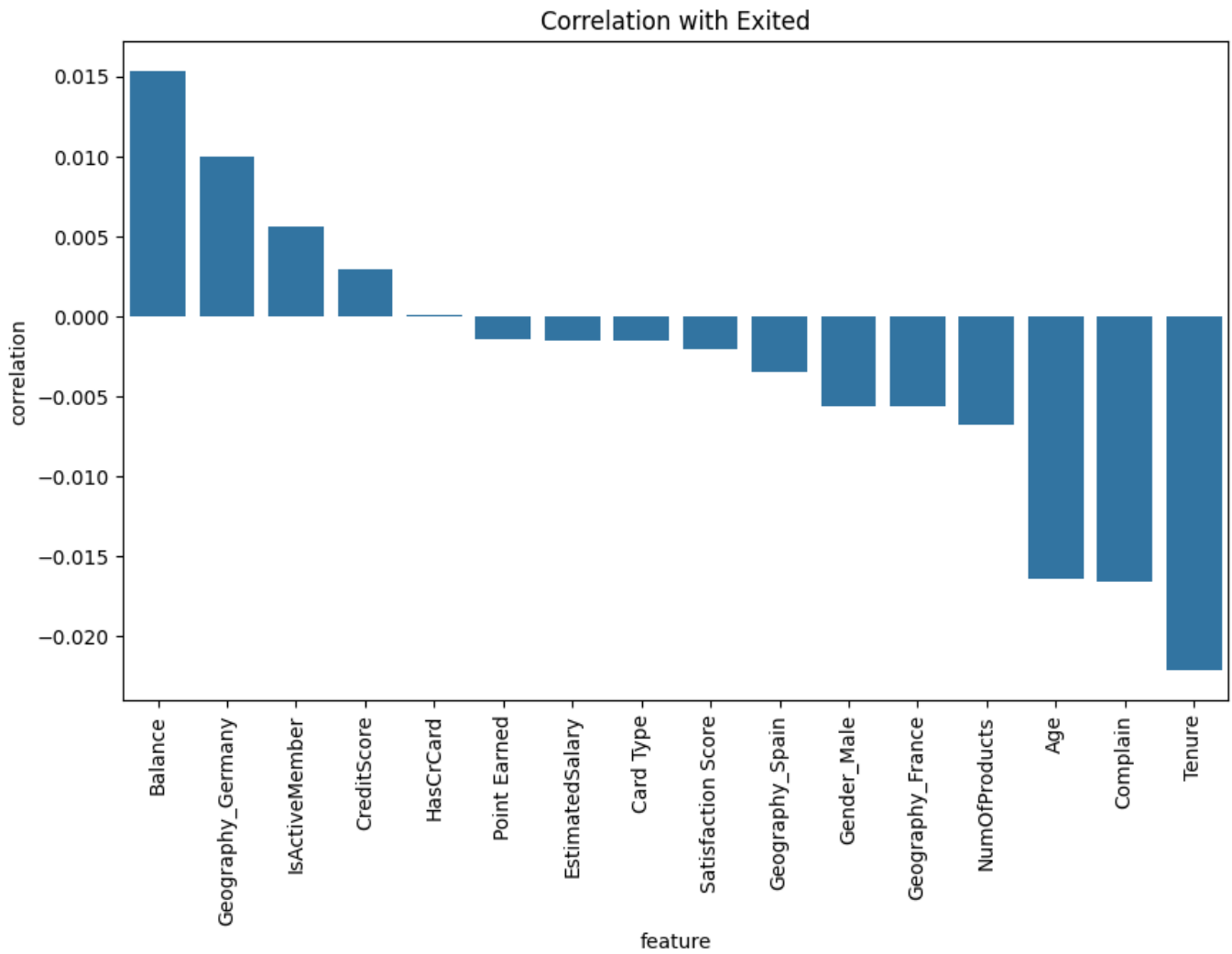
- 1. Feature (Cecha): Nazwa cechy.
- 2. Overall Average (Srednia Ogolna): Srednia wartosc cechy dla wszystkich klientow.
- 3. Predicted High Probability (Target=1) (Przewidywana Wysoka Prawdopodobienstwo dla Target=1):Przewidywane prawdopodobienstwo, ze zmienna docelowa wynosi 1 dla danej cechy.

Wiersze z 3 najwazniejszymi cechami sa podswietlone na zielono.

# Wizualizacja rozkładu danych

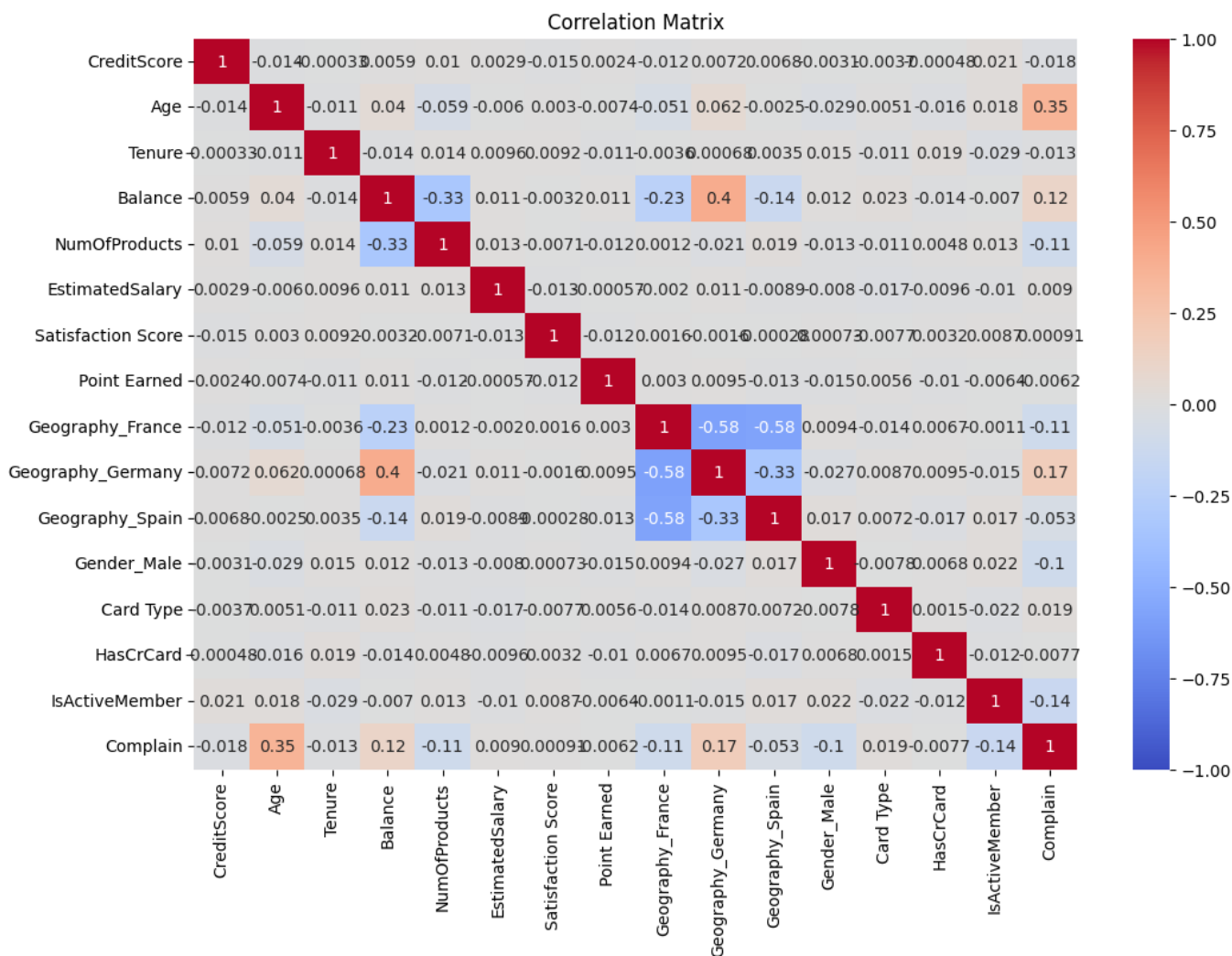


## Korelacja cech z celem



Wnioski: Wykres pokazuje, które cechy są skorelowane z celem. Wartości bliskie 1 oznaczają silną korelację dodatnią, a bliskie -1 silną korelację ujemną. Cechy o wysokiej korelacji z celem mogą być kluczowe w modelowaniu. Wartości poniżej 0.1 są zazwyczaj uznawane za niskie i nieistotne. Najwyższą korelację z celem miały cechy: [].

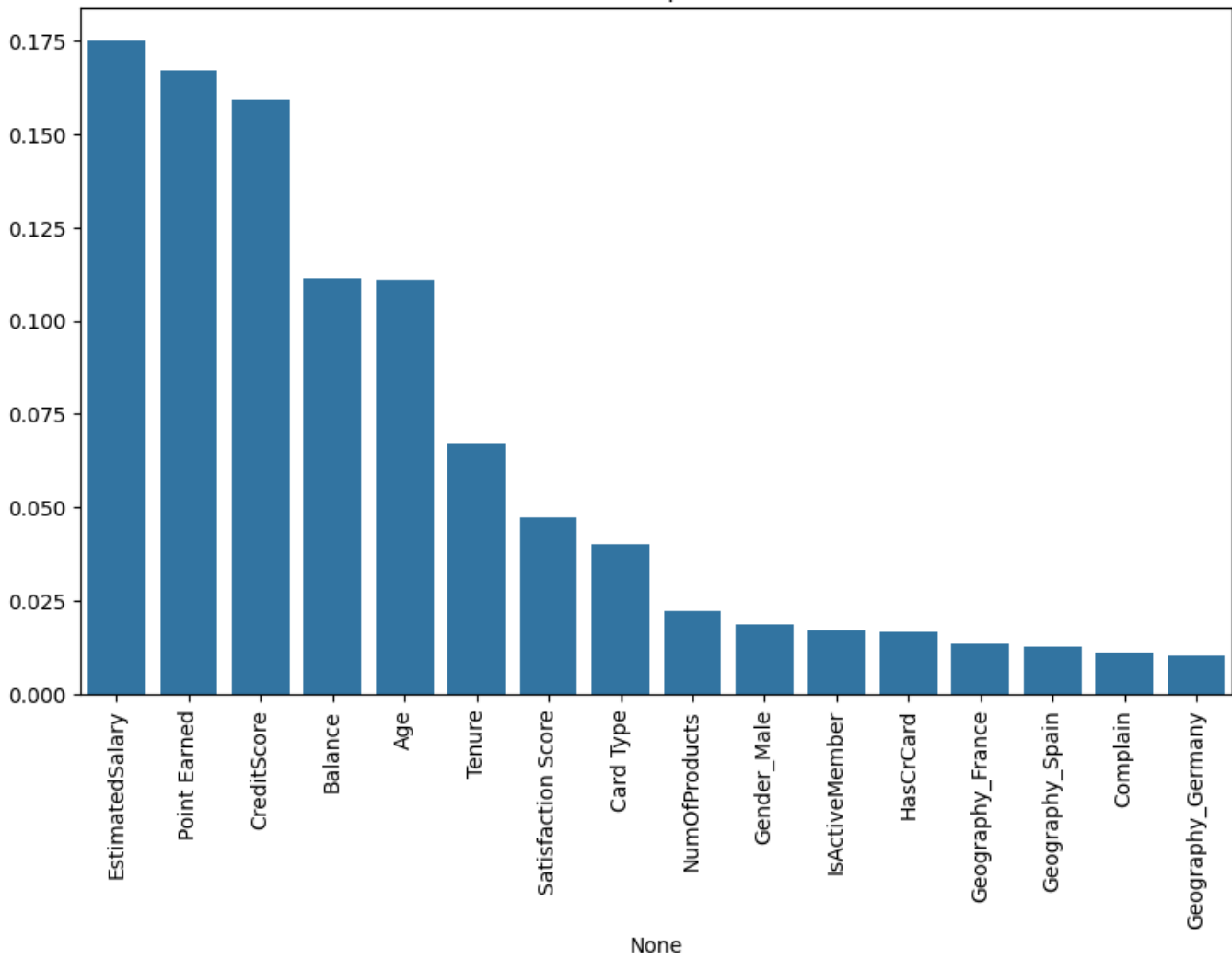
## Macierz korelacji cech



Wnioski: Wykres pokazuje, które cechy są ze sobą skorelowane. W wyniku analizy usunięto cechy o wysokiej korelacji między sobą: ['Geography\_Spain', 'Geography\_France', 'Geography\_Germany'].

## Waznosc cech

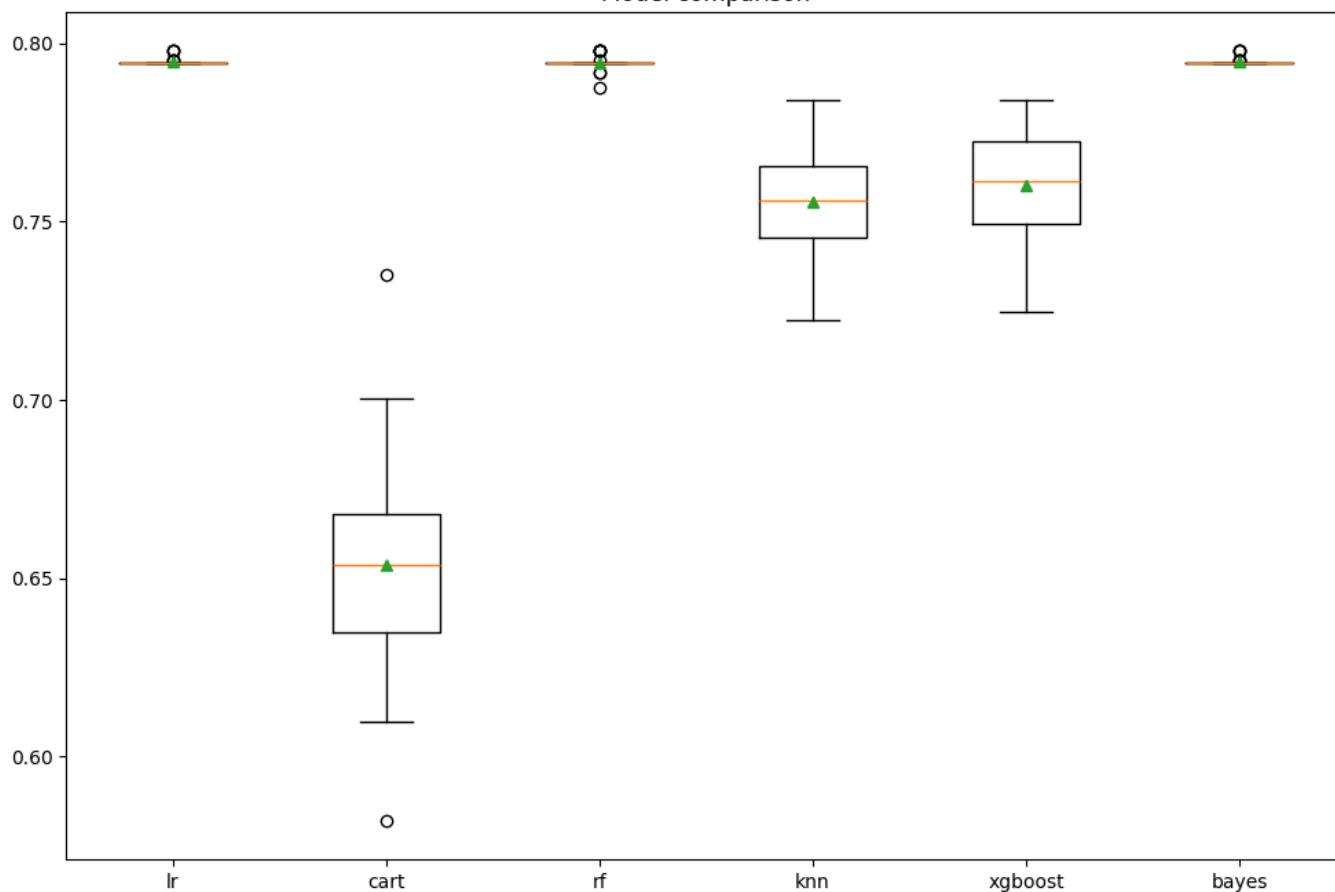
Feature Importance



Wnioski: Wykres pokazuje, które cechy miały największy wpływ na predykcje modelu. Można je rozważyć jako kluczowe w dalszych analizach.

## Porownanie modeli

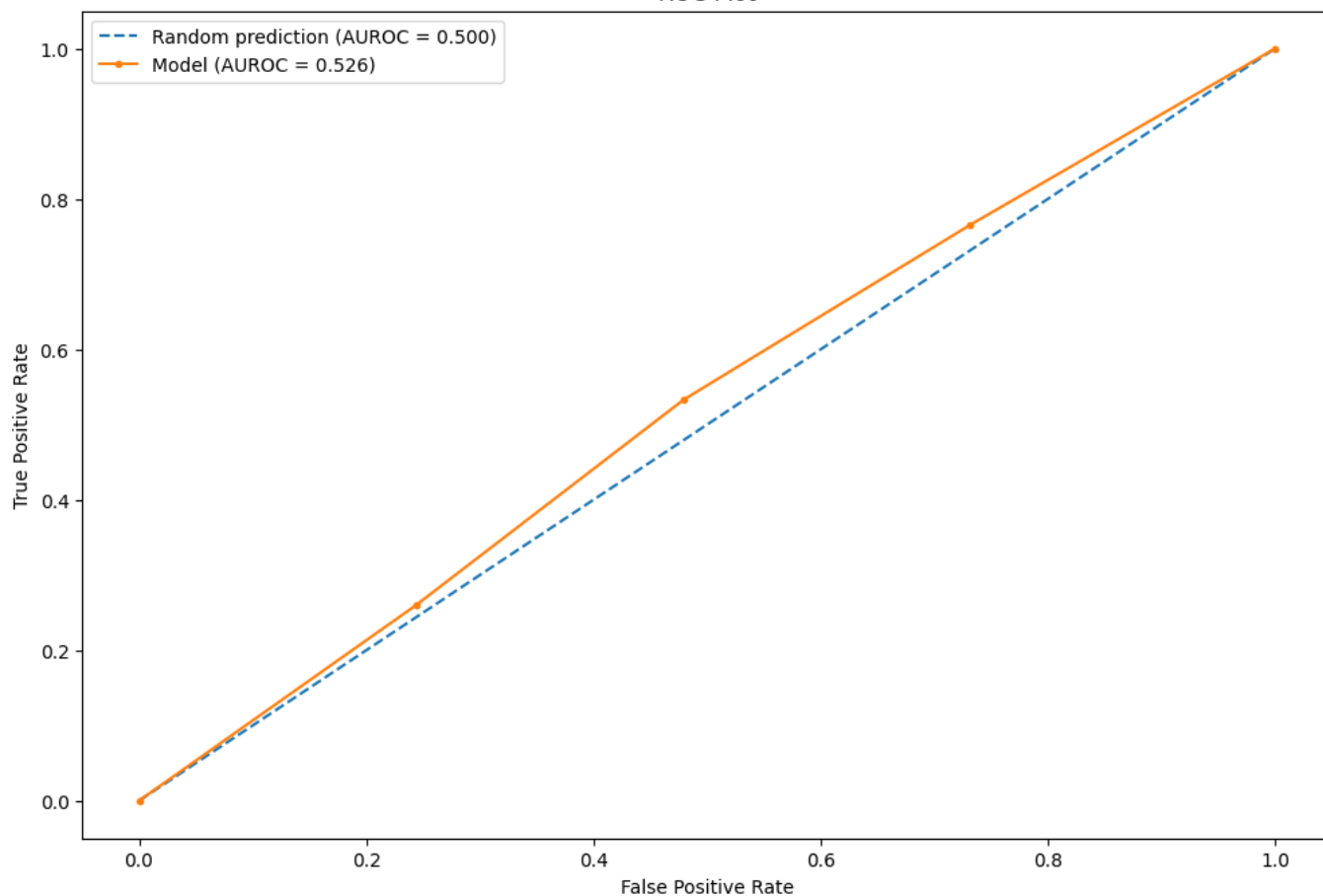
Model comparison



Wnioski: Wykres przedstawia porownanie roznych modeli pod wzgledem ich skuteczności. Widac, ze model 'lr' osiagnal najwyzsze wyniki w walidacji. Z wynikiem 0.7948 jest to najlepszy model do dalszych analiz.

# Krzywa ROC

ROC Plot



Wnioski: Wykres przedstawia krzywa ROC dla najlepszego modelu 'lr'. Im wyższa powierzchnia pod krzywą (AUC), tym lepszy model. W tym przypadku AUC wynosi 0.5258.