

Detekcja oszustw z wykorzystaniem metod wrażliwych na koszt

Patryk Wielopolski

5 listopada 2019

Rozdział 1

Wstęp

Tutaj będzie wstęp.

Rozdział 2

Wprowadzenie teoretyczne

W tej części zostaną wprowadzone wszelkie potrzebne miary skuteczności modeli oraz modele predykcyjne, które zostaną wykorzystane do przeprowadzenia eksperymentu.

2.1 Miary skuteczności modeli

2.1.1 Macierz pomyłek

Na podstawie podanej macierzy pomyłek w tabeli 2.1 definiujemy następujące miary skuteczności modeli:

$$\text{Skuteczność} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Precyzja} = \frac{TP}{TP + FP}$$

$$\text{Czułość} = \frac{TP}{TP + FN}$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precyzja} \cdot \text{Czułość}}{\text{Precyzja} + \text{Czułość}}$$

		Predykcja	
		Oszustwo	Normalna
Prawda	Oszustwo	TP	FN
	Normalna	FP	TN

Tabela 2.1: Macierz pomyłek

2.1.2 Metryki wrażliwe na koszt

2.2 Standardowe modele

2.2.1 Regresja logistyczna

2.2.2 Drzewo decyzyjne

2.2.3 Las losowy

2.2.4 XGBoost

2.3 Cost Sensitive Training

2.3.1 Regresja logistyczna wrażliwa na koszt

2.3.2 Drzewo decyzyjne wrażliwe na koszt

2.4 Cost Dependent Classification

2.4.1 Optymalizacja progu

2.4.2 Bayesian Minimum Risk

Rozdział 3

Eksperyment

Rozdział 4

Rezultaty

Rozdział 5

Podsumowanie

Rozdział 6

Dalsze prace