SPRAWOZDANIE

Zajęcia: Uczenie maszynowe

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium Nr 1 Data 12.10.2024

Temat: Praktyczne zastosowanie regresji liniowej w analizie danych.

Implementacja algorytmów klasyfikacji binarnej w Pythonie.

Wariant 8

Imię Nazwisko Hubert Mentel Informatyka II stopień, niestacjonarne, 1 semestr, gr.1a

1. Zadanie:

Opracować przepływ pracy uczenia maszynowego zagadnienia regresji (model regresji liniowej) oraz klasyfikacji binarnej (model SVM) na podstawie zbioru danych według 8 wariantu zadania:

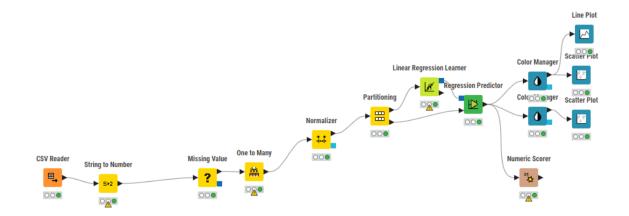
Prostate cancer

https://www.kaggle.com/ashrafalsinglawi/prostate-cancer-survival-data

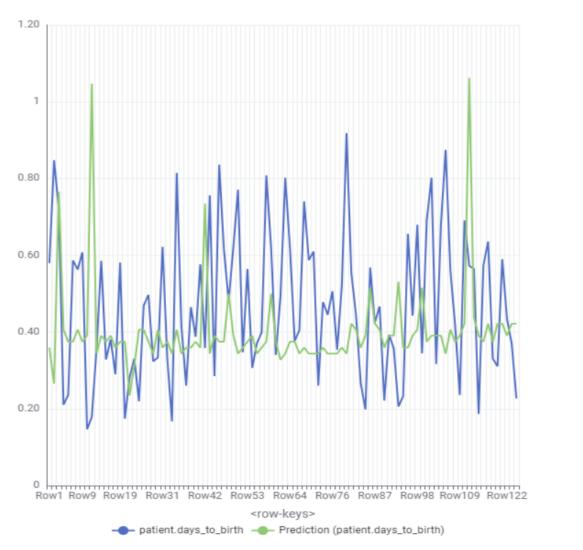
Pliki dostępne są na GitHubie pod linkiem:

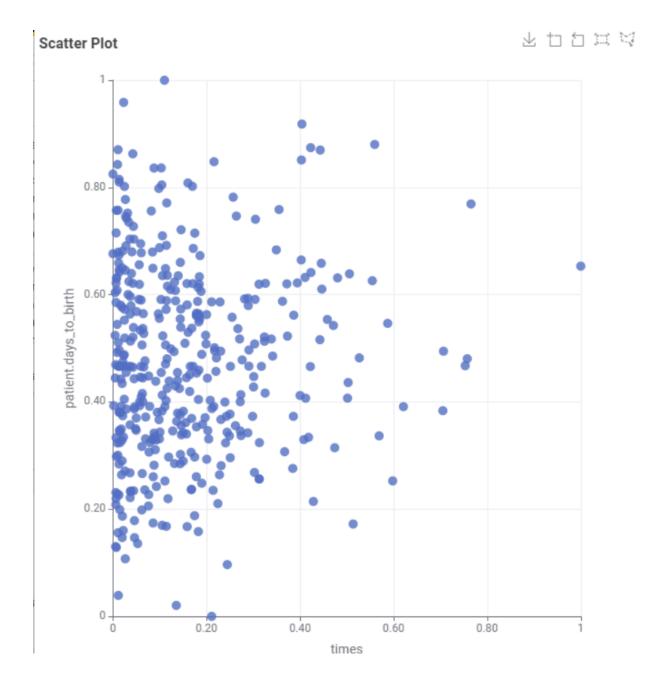
https://github.com/HubiPX/NOD/tree/master/UM/Zadanie%201

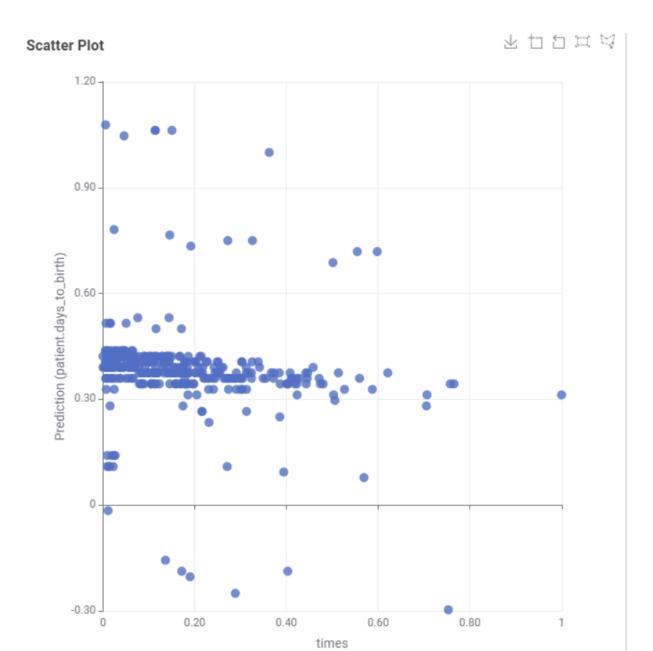
2. Opis programu opracowanego (kody źródłowe, zrzuty ekranu)

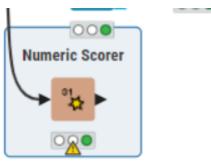






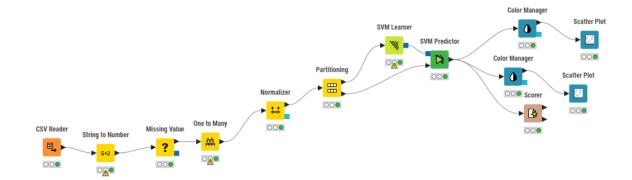






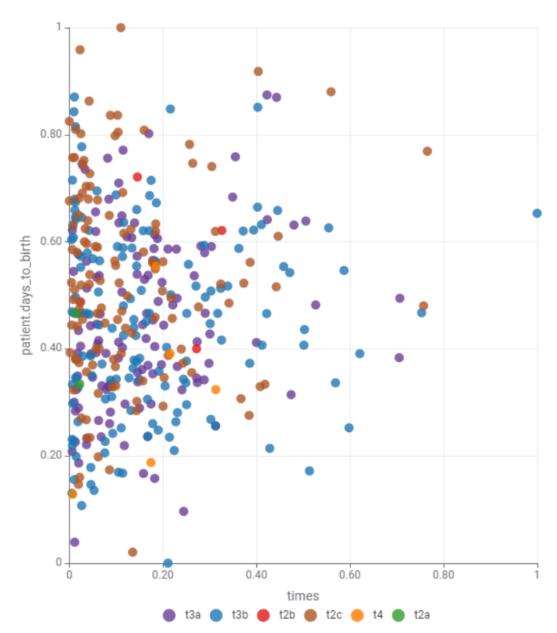
Add comment

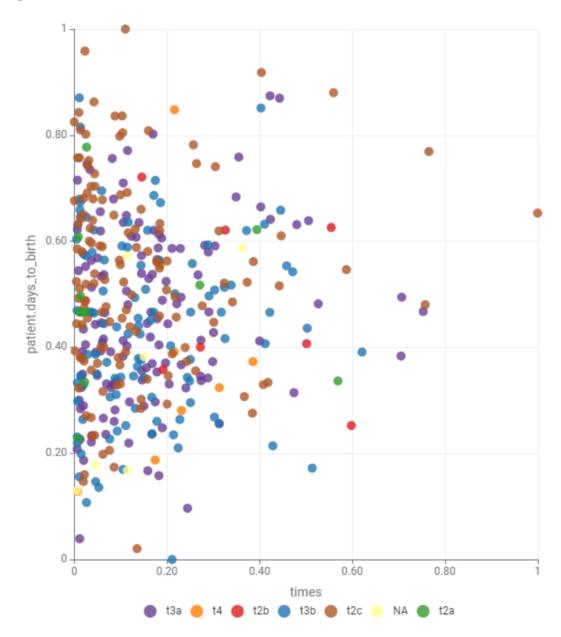
4			
▶ 1	: Sta	tistics 🖂	Flow Variables
Rows	s: 7	Columns:	1
	#	RowID	Prediction (patient.days_to_birth) Number (double)
	1	R^2	-0.921
	2	mea	0.188
	3	mea	0.062
	4	root	0.25
	5	mea	-0.088
	6	mea	NaN
	7	adju	-0.921



Dane Przewidziane









▶ 1:	Confus	sion matri	x ► 2: Accur	racy statisti	cs 💆 Flow Va							
lows	7	Columns:	7			Table	Statistics (9				Q
	#	RowID	t2b Number (integer)	v t3a	nber (integer)	t4 Number (integer)	t3b Number (integer		t2c Number (integer)	t2a Number (integer)	V NA Number (inte	eger)
	1	t2b	4	0		0	4	(0	0	0	
	2	t3a	0	118	3	0	11		0	0	0	
	3	t4	0	0		5	3		0	0	0	
	4	t3b	0	0		0	105		0	0	0	
	5	t2c	0	0		0	9		123	0	0	
	6	t2a	0	0		0	10		0	2	0	
_	_											
	7	NA	0	0	o St Flour Veri	0	6		0	0	0	
	: Confus	NA sion matrix	x ► 2: Accur	0 acy statistic	s 💆 Flow Vari				0	0	0	Q 4
	: Confus	sion matri:	x ► 2: Accur		TrueNega	ables Table [Precision Number (dou	Sensitivity		F-measure Number (dou	
	: Confus	sion matrix	x	acy statistic	TrueNega	Table [Statistics (2)	Precision	Sensitivity	Specificity	F-measure	Accur ~
	: Confus	sion matrix Columns:	x ≥ 2: Accur	FalsePosi.	TrueNega Number (inte_	Table Table Number (inte	Statistics Recall Number (dou	Precision Number (dou	Sensitivity Number (dou_	Specificity Number (dou	F-measure Vumber (dou	Accur \(\frac{1}{Number}\)
	: Confus :: 8 #	sion matri: Columns: RowID t2b	x 2: Accur	FalsePosi. Number (inte.	TrueNega Number (inte_	Table Table Number (inte	Statistics Recall Number (dou > 0.5	Precision Number (dou	Sensitivity Number (dou 0.5	Specificity Number (dou	F-measure Number (dou 0.667	Accur Number
	# 1 2	columns: RowID t2b t3a	x	FalsePosi. Number (inte.	TrueNega Number (inte 392 271	Table Table Number (inte 4	Recall Number (dou. V 0.5 0.915	Precision Number (dou	Sensitivity Number (dou 0.5 0.915	Specificity Number (dou	F-measure Number (dou. 0.667 0.955	Accur Number 3
	# 1 2 3	columns: RowID t2b t3a t4	x	FalsePosi. Number (inte. 0	TrueNega Number (inte	Table FalseNeg Number (inte 4 11 3	Recall Number (dou. 0.5 0.915 0.625	Precision Number (dou 1 1	Sensitivity Number (dou 0.5 0.915 0.625	Specificity Number (dou 1 1 1	F-measure Number (dou.) 0.667 0.955 0.769	Accur Number 3
	# 1 2 3 4	Columns: RowID t2b t3a t4 t3b	x	FalsePosi Number (inte 0 0 43	TrueNega Number (inte	Table Table Table 1 FalseNeg Number (inte 4 11 3 0	Recall Number (dou. 0.5 0.915 0.625	Precision Number (dou 1 1 0.709	Sensitivity Number (dou 0.5 0.915 0.625	Specificity Number (dou 1 1 1 0.854	F-measure Number (dou.) 0.667 0.955 0.769 0.83	Accur Number 7
	# 1 2 3 4 5	columns: RowID t2b t3a t4 t3b t2c	x	FalsePosi Number (inte. 0 0 43	TrueNega Number (inte	Table	Recall Number (dou. 0.5 0.915 0.625 1 0.932	Precision Number (dou 1 1 1 0.709	Sensitivity Number (dou 0.5 0.915 0.625 1 0.932	Specificity Number (dou 1 1 1 0.854 1	F-measure Number (dou. V. 0.667 0.955 0.769 0.83 0.965	Accur Number 7

3. Wnioski

KNIME jest potężnym narzędziem do analizy danych i uczenia maszynowego, które pozwala na szybkie prototypowanie modeli i ich ocenę. Dzięki graficznemu interfejsowi oraz szerokiej bibliotece węzłów, KNIME umożliwia tworzenie zaawansowanych przepływów pracy bez konieczności programowania. Platforma ta wspiera różnorodne techniki uczenia maszynowego, w tym regresję liniową wielowymiarową oraz klasyfikację binarną za pomocą maszyn wektorów nośnych (SVM), co czyni ją wszechstronnym narzędziem dla analityków danych.

Regresja liniowa wielowymiarowa pozwala na modelowanie zależności między wieloma zmiennymi niezależnymi a jedną zmienną zależną, co znajduje zastosowanie w prognozowaniu i analizie trendów. Z kolei klasyfikacja binarna za pomocą SVM umożliwia znalezienie optymalnej hiperpłaszczyzny separującej dwie klasy, co sprawdza się w zadaniach takich jak wykrywanie anomalii czy rozpoznawanie wzorców. Dzięki możliwości integracji różnych metod analizy danych oraz narzędzi do oceny modeli, KNIME jest użytecznym środowiskiem do eksploracji i wdrażania modeli uczenia maszynowego w praktycznych zastosowaniach.