

SPRAWOZDANIE

Zajęcia: Uczenie maszynowe

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium Nr 1 Data 19.10.2024 Temat: "Praktyczne zastosowanie regresji liniowej w analizie danych. Implementacja algorytmów klasyfikacji binarnej w Pythonie" Wariant 8	Tomasz Pietrzyk Informatyka II stopień, niestacjonarne, 1semestr, gr.1a
--	--

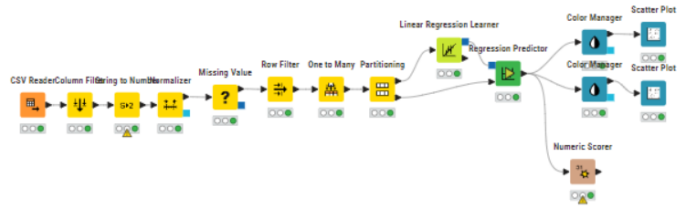
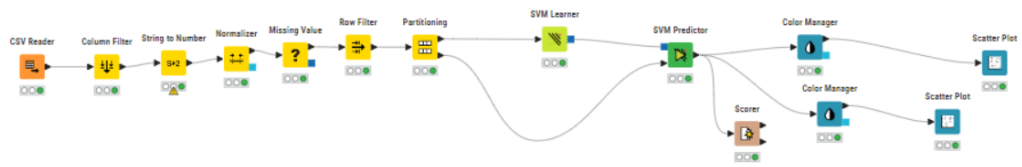
1. Polecenie: wariant 8 zadania

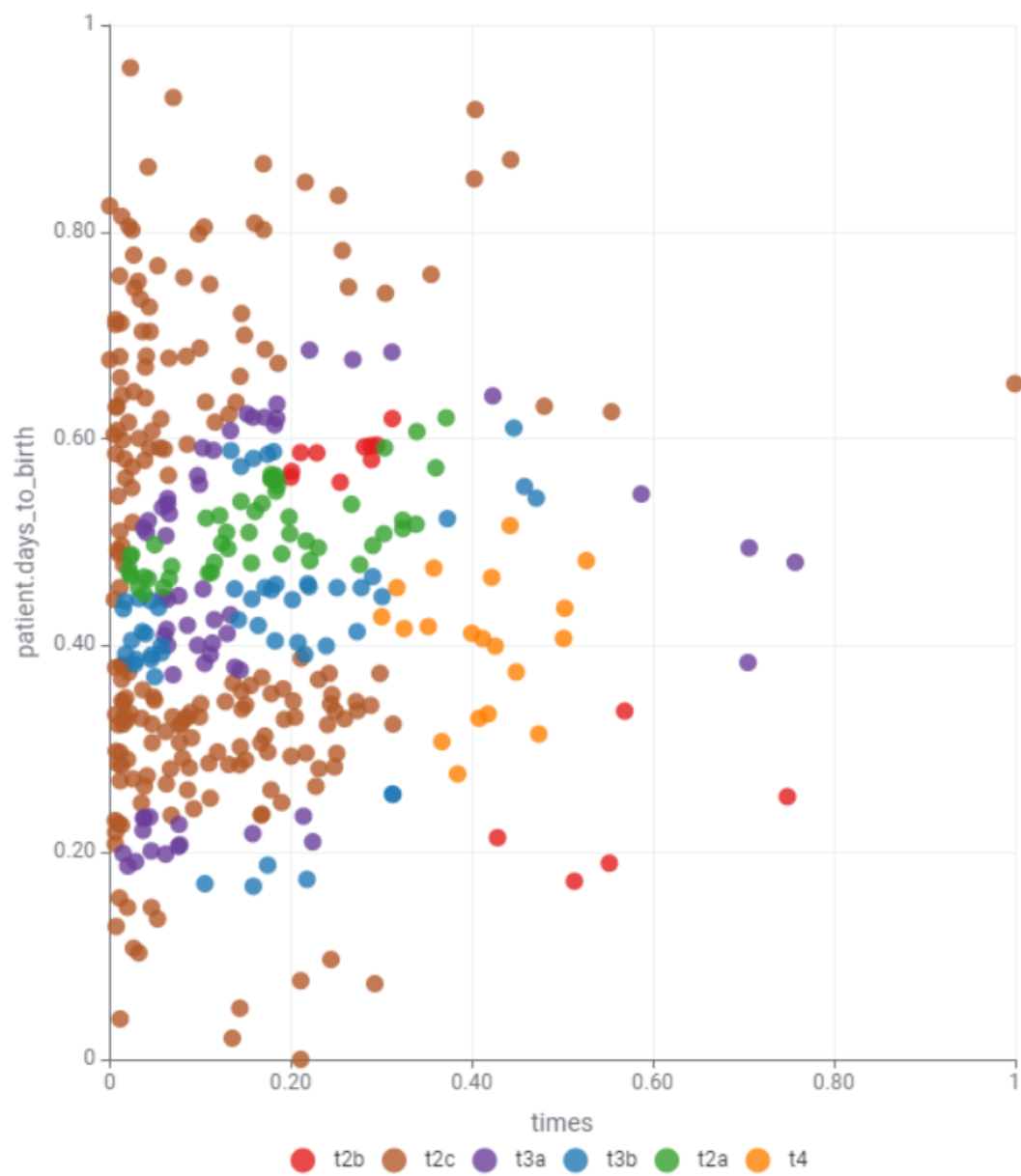
Opracować przepływ pracy uczenia maszynowego zagadnienia regresji (model regresji liniowej) oraz klasyfikacji binarnej (model SVM) na podstawie zbioru danych według wariantu zadania:

Prostate Cancer <https://www.kaggle.com/datasets/ashrafalsinglawi/prostate-cancer-survival-data>

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

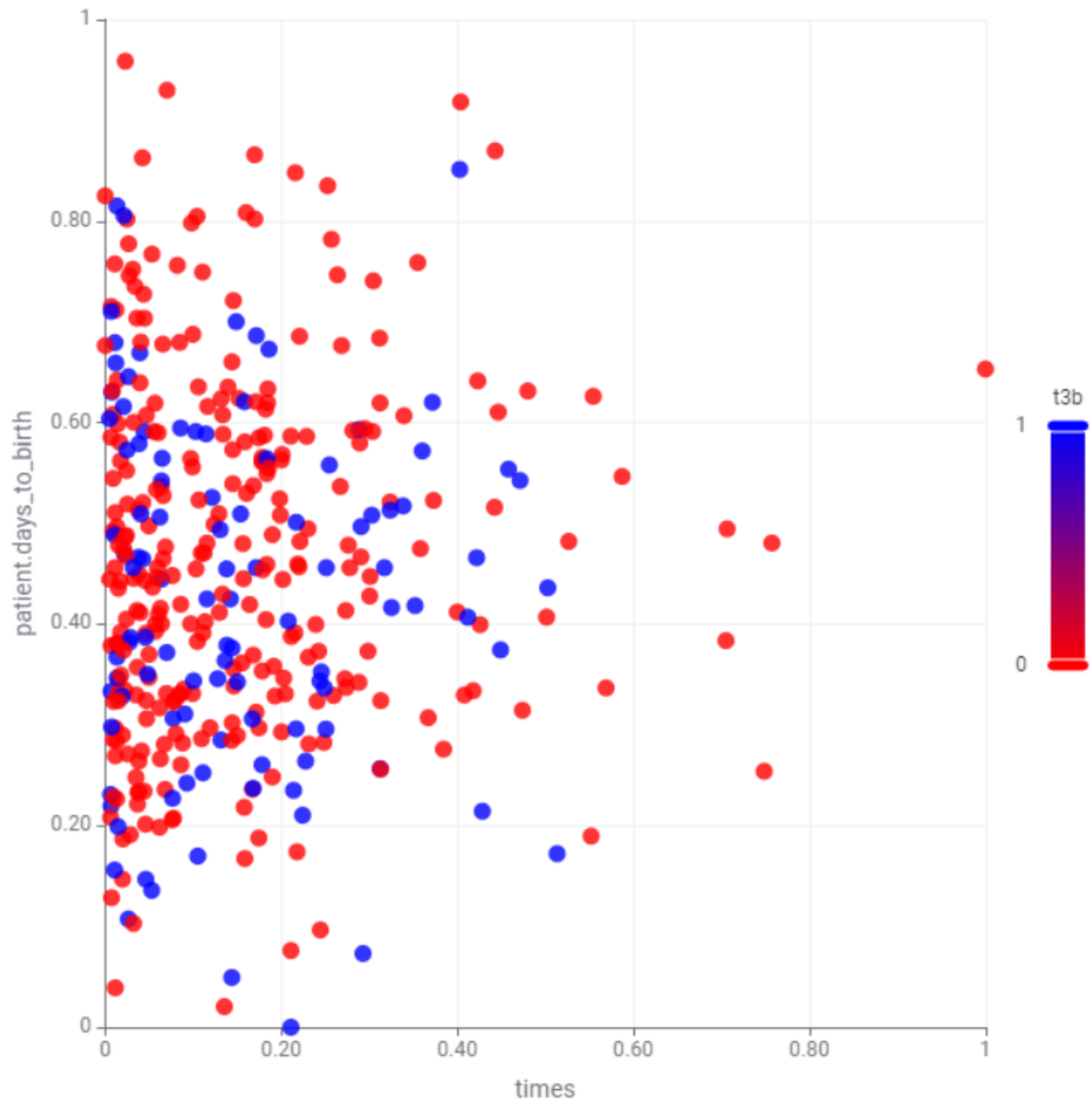
GitHub: https://github.com/TomekPietrzyk/UM_I_2024_NS.git



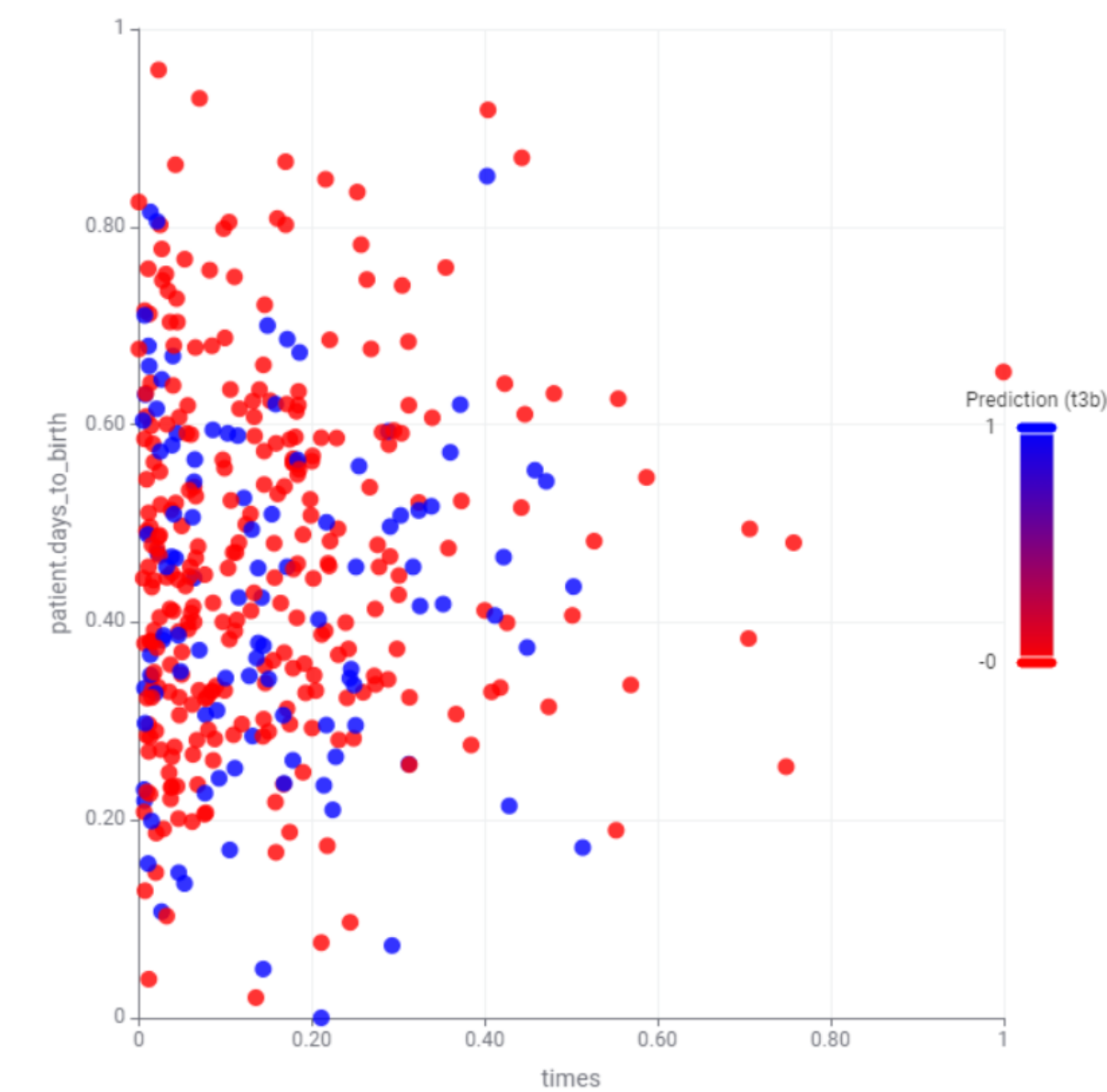


[illegible]

<input type="checkbox"/>	#	RowID	TruePosit... Number (inte...	FalsePosi... Number (inte...	TrueNeg... Number (inte...	FalseNeg... Number (inte...	Recall Number (dou...	Precision Number (dou...	Sensitivity Number (dou...	Specificity Number (dou...	F-measure Number (dou...	Accur Number
<input type="checkbox"/>	1	t2b	0	15	373	6	0	0	0	0.961	②	②
<input type="checkbox"/>	2	t3a	22	38	224	110	0.167	0.367	0.167	0.855	0.229	②
<input type="checkbox"/>	3	t4	0	19	366	9	0	0	0	0.951	②	②
<input type="checkbox"/>	4	t3b	12	38	250	94	0.113	0.24	0.113	0.868	0.154	②
<input type="checkbox"/>	5	t2c	69	131	131	63	0.523	0.345	0.523	0.5	0.416	②
<input type="checkbox"/>	6	t2a	1	49	336	8	0.111	0.02	0.111	0.873	0.034	②
<input type="checkbox"/>	7	NA	0	0	394	0	②	②	②	1	②	②
<input type="checkbox"/>	8	Overall	②	②	②	②	②	②	②	②	②	0.264



Po klasyfikacji



Rows: 7 | Columns: 1

Table | Statistics

<input type="checkbox"/>	#	RowID	Prediction (t3b) <small>Number (double)</small>
<input type="checkbox"/>	1	R^2	1
<input type="checkbox"/>	2	mea...	0
<input type="checkbox"/>	3	mea...	0
<input type="checkbox"/>	4	root ...	0
<input type="checkbox"/>	5	mea...	-0
<input type="checkbox"/>	6	mea...	NaN
<input type="checkbox"/>	7	adju...	1

3. Wnioski

Klasyfikacja za pomocą SVN przebiegła miernie i jej efekt był fatalny, bardzo niska predykcja, duży wkład w tak słaby wynik niewątpliwie ma mała ilość danych cząstkowych opisujących pacjenta. Jeśli chodzi o regresję liniową rozłożenie metody leczenia za pomocą bloku one to many pomogło osiągnąć R^2 na poziomie 1