

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:

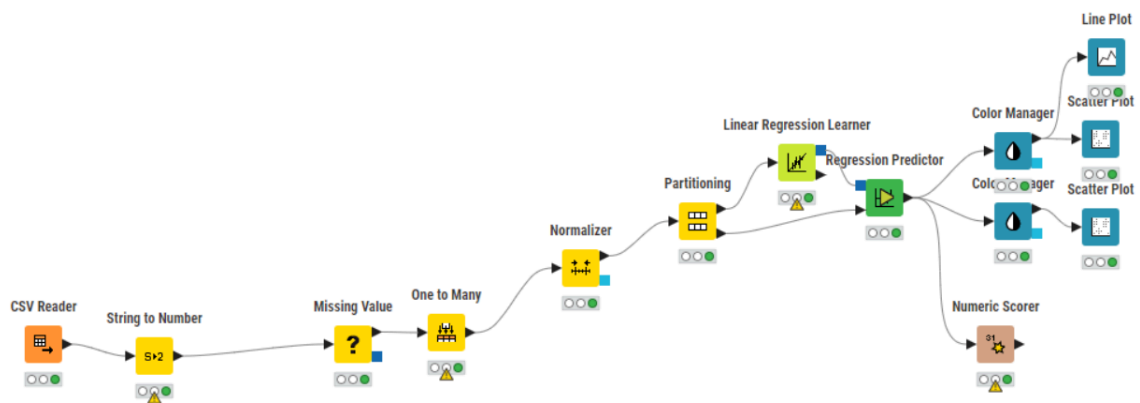
Prostate cancer

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

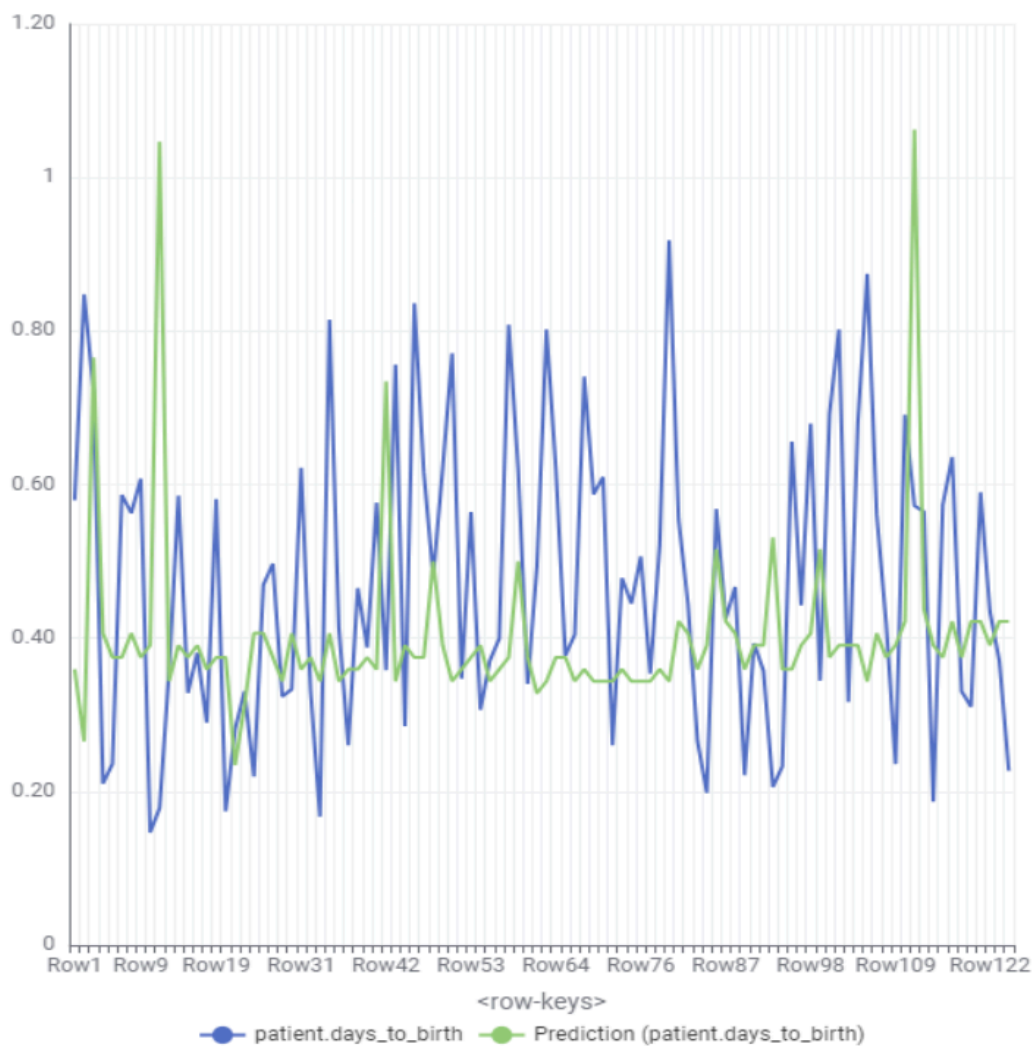
Pliki dostępne są na GitHubie pod linkiem:

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

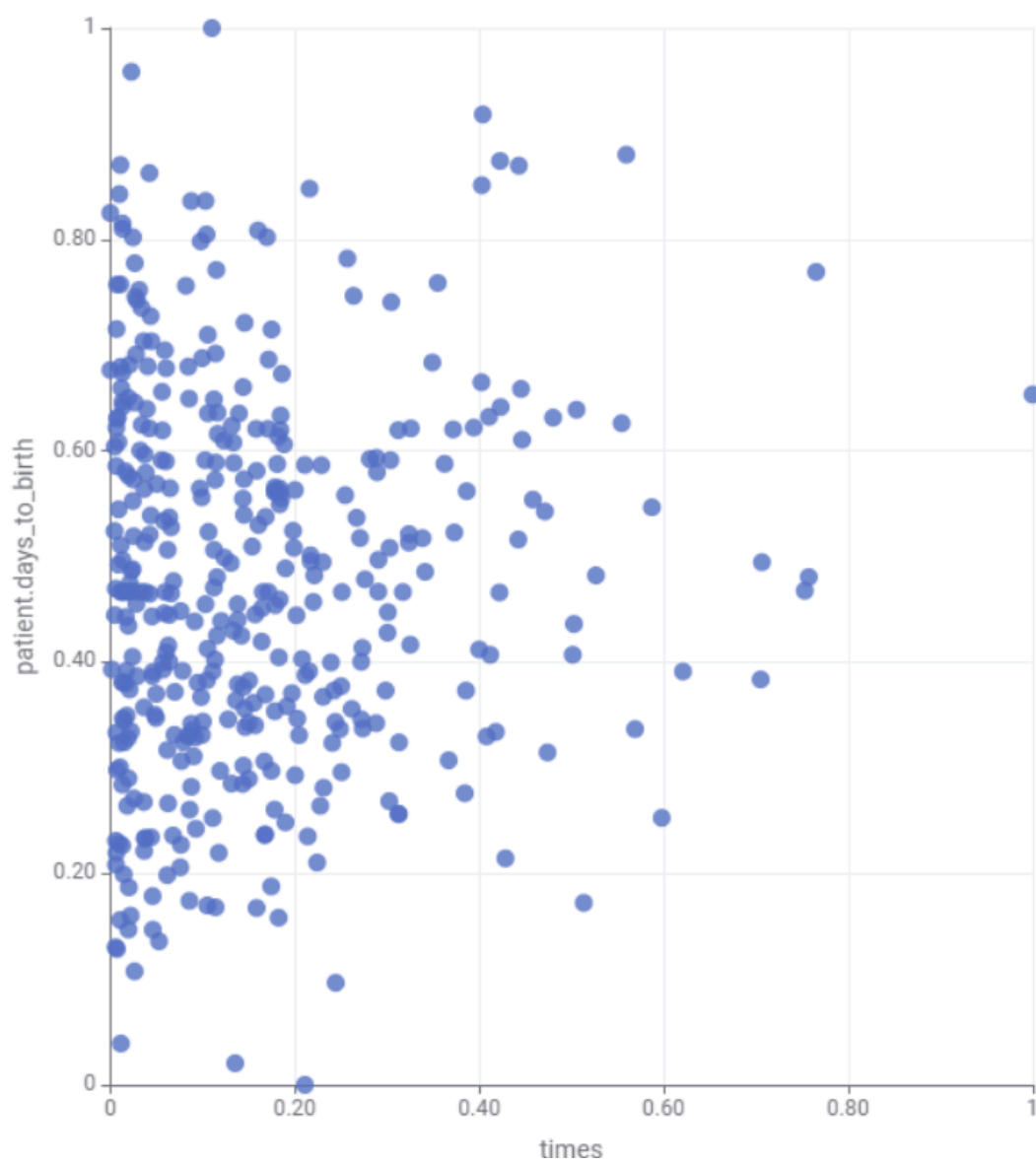
2. Opis programu opracowanego (kody źródłowe, zrzuty ekranu)



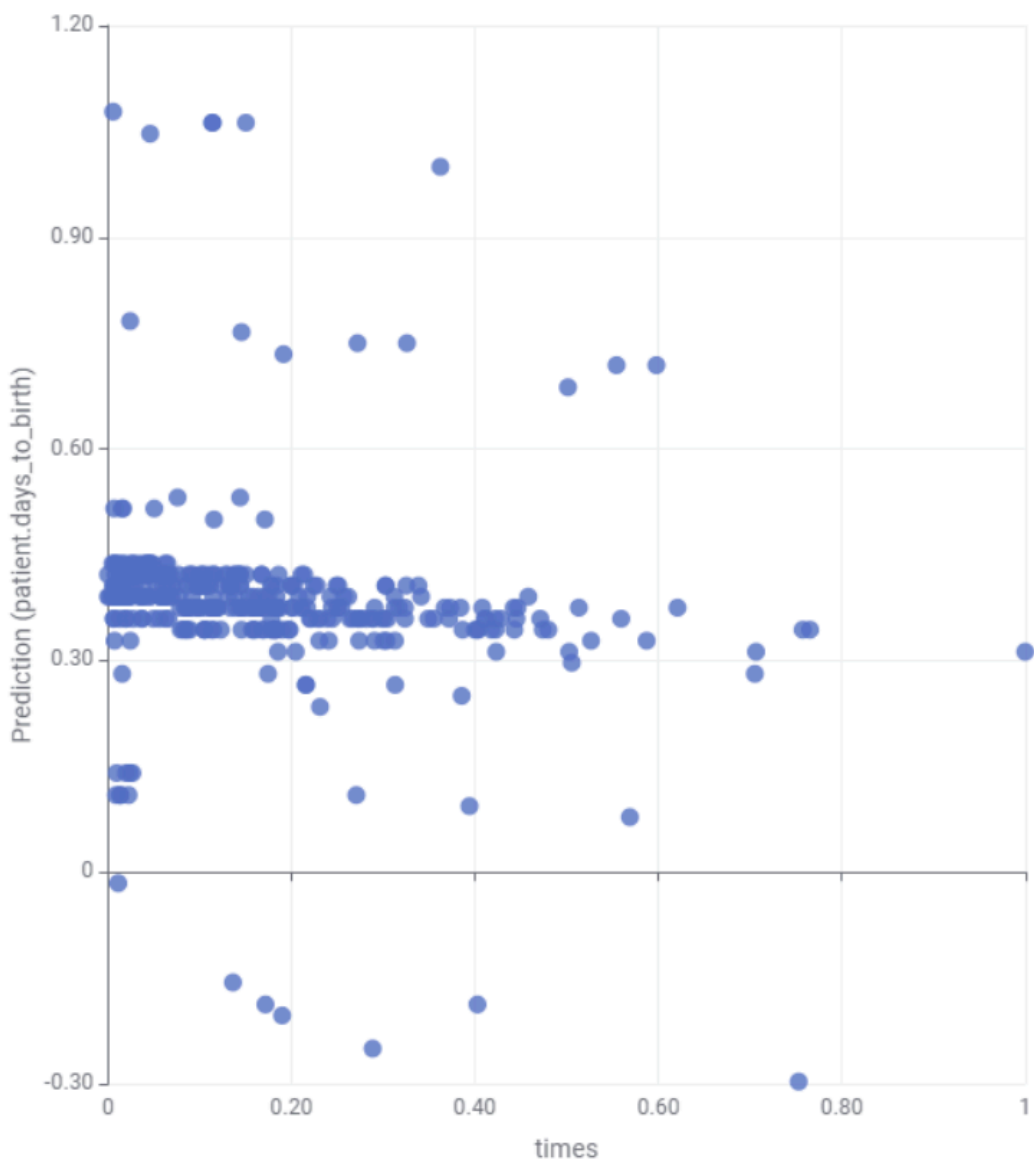
Line Plot

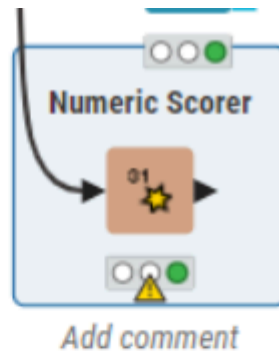


Scatter Plot



Scatter Plot

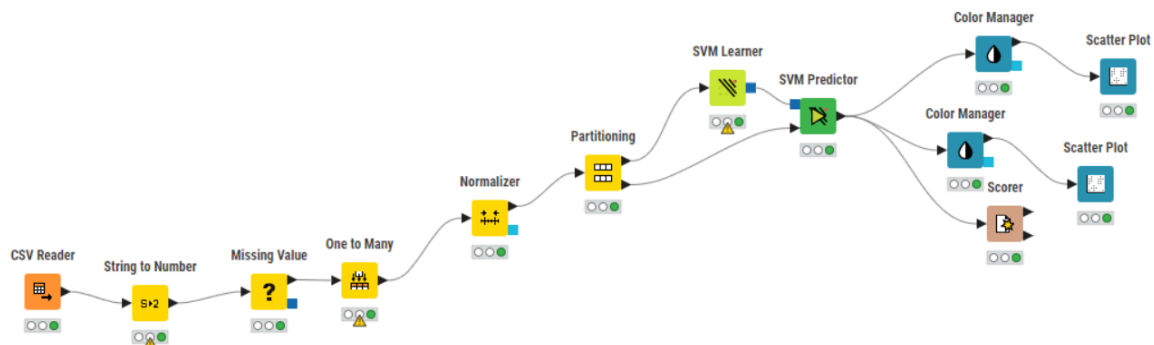




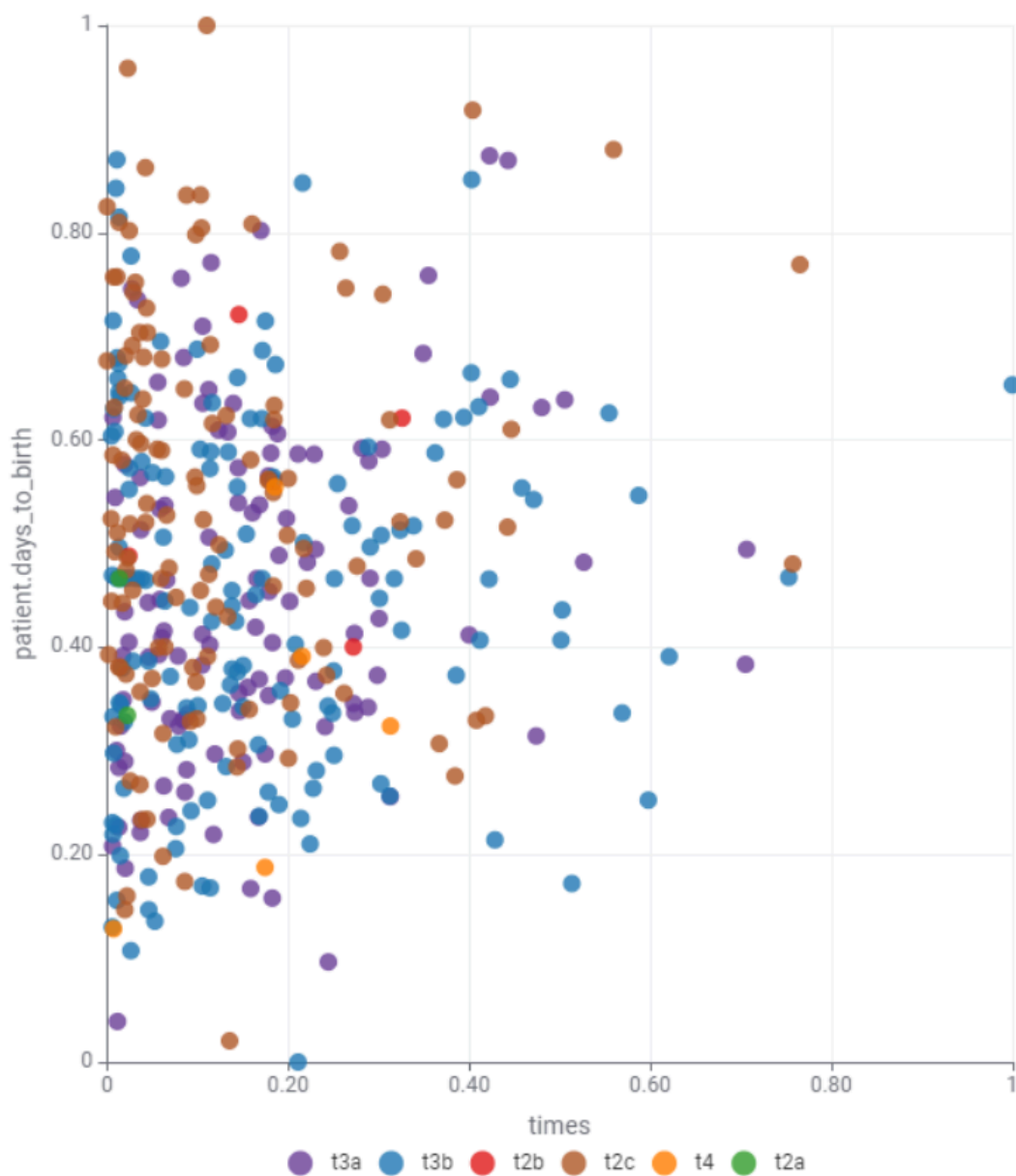
► 1: Statistics  Flow Variables

Rows: 7 | Columns: 1

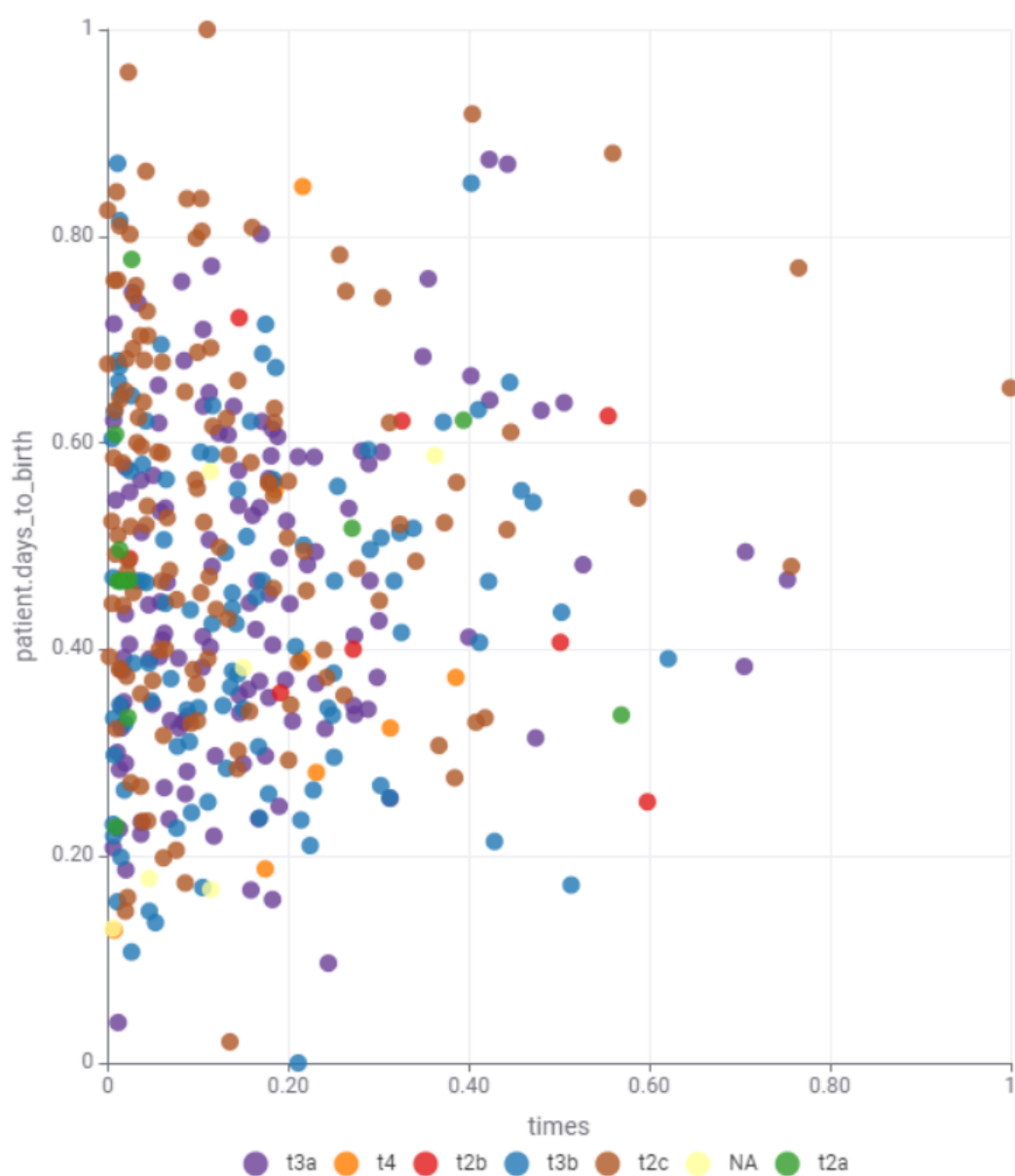
<input type="checkbox"/>	#	RowID	Prediction (patient.days_to_birth) <i>Number (double)</i>
<input type="checkbox"/>	1	R^2	-0.921
<input type="checkbox"/>	2	mea...	0.188
<input type="checkbox"/>	3	mea...	0.062
<input type="checkbox"/>	4	root ...	0.25
<input type="checkbox"/>	5	mea...	-0.088
<input type="checkbox"/>	6	mea...	NaN
<input type="checkbox"/>	7	adju...	-0.921



Dane Przewidziane



Originalne Dane



[illegible]

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.