

TASK 1

iris.tab

TASK 2

Zawiera dataset z 150 instancjami setozy, Iris virginica i Iris versicolor.

TASK 3

Zastosowano drzewo klasyfikacyjne.

TASK 4

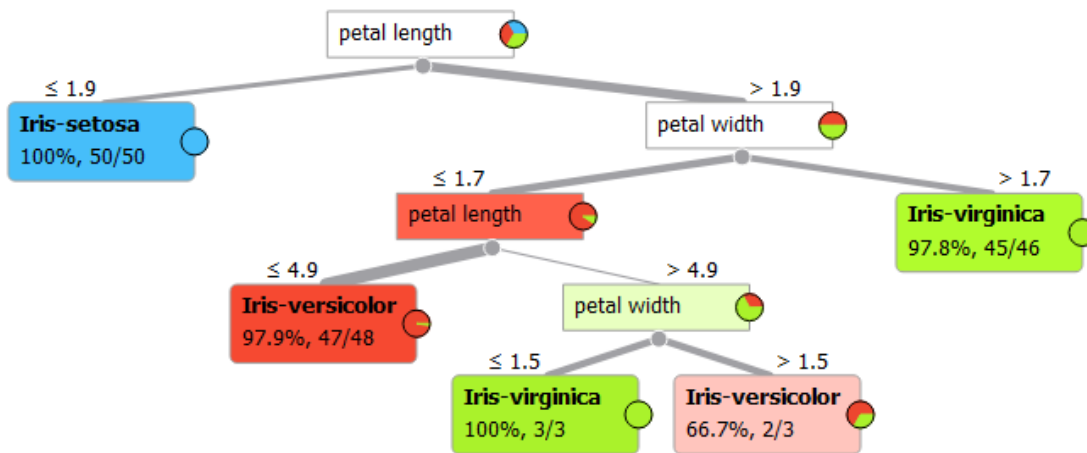
Na widoku są dane rozłożone wg drzewa logicznego, według kryteriów podanych nad węzłami (>1.9 ; >1.7 ; ≤ 1.7 ; ...).

Na początku analizowana jest długość płatków, potem szereokość, a na samym końcu klasyfikowane są do Iris-versicolor lub Iris-virginica.

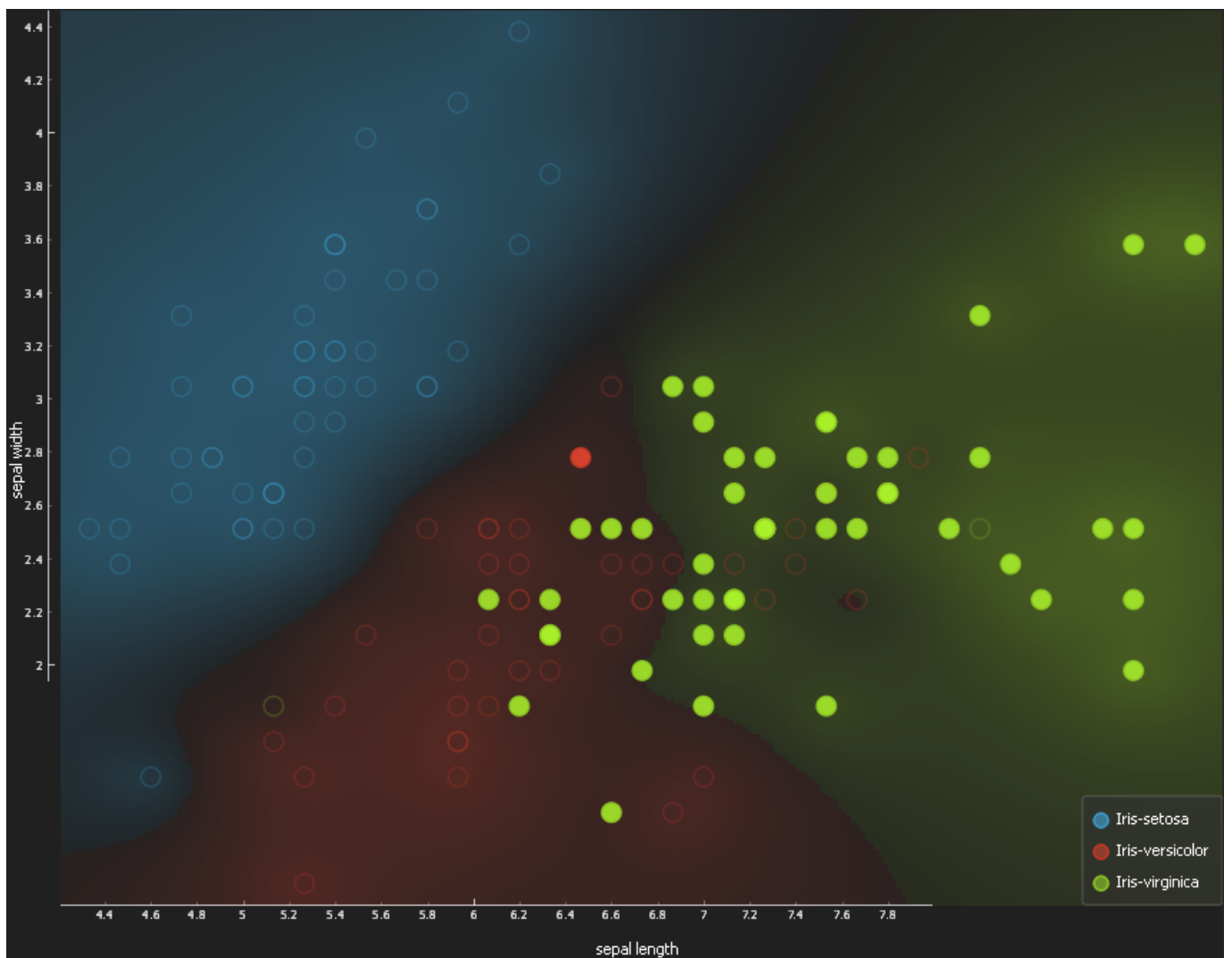
TASK 5

Unlimited - nieograniczona

TASK 6



TASK 7



TASK 8

150

TASK 9

Model dobrze przewiduje gatunek, ponieważ F-miara nie schodzi poniżej 0,949.

TASK 10

Najlepsza wydaje się regresja logistyczna. Różnice między modelami nie są duże.

TASK 11

Iris-virginica	Iris-versicolor	4.9	2.5	4.5	1.7
Iris-versicolor	Iris-virginica	5.9	3.2	4.8	1.8
Iris-versicolor	Iris-virginica	6.0	2.7	5.1	1.6
Iris-versicolor	Iris-virginica	6.7	3.0	5.0	1.7
Iris-virginica	Iris-versicolor	6.0	2.2	5.0	1.5

TASK 12

Jest to widżet umożliwiający implementację kilku metod próbkowania danych.

Zbiór wyjściowy to próbkowany i uzupełniający zbiór danych (z instancjami ze zbioru wejściowego, które nie są zawarte w próbkowanym zbiorze).

Służy m.in. do podziału zbioru na podzbiór treningowy i testowy (walidacyjny), co przy klasyfikacji i regresji daje nam wiarygodną reprezentację całego zbioru i pozwala uzyskać najwierniejszy wynik.

TASK 13

Wyniki nie dają widocznie zauważalnej poprawy względem poprzedniego podejścia.

Metoda Test data działa poprzez uwzględnienie datasetu treningowego i testowego. Testowanie jest na zbiorze testowym, oraz na tej podstawie wyliczane są miary.

TASK 14

- Pobiera podzbiór testowy z nowego zbioru testowego `irist_test.tab`
 - Wyniki są trochę gorsze od testów na danych próbkowanych.
- Poprzednio robiliśmy to na zbiorze `iris.tab`, który mógł dawać różne wyniki z powodu losowego próbkowania

TASK 15

$2+3 = 5$

Pięć rekordów.

TASK 16

Najslabsze wyniki otrzymałem na bazie modelu Naive Bayes z CA = 0.873

TASK 17

Iris-setosa.

TASK 18

$11 + 8 = 19$

19 błędów.

TASK 19

- Iris-setosa
- Iris-setosa
- Iris-virginica

TASK 20

Prawdopodobieństwa to wiersze, w kolumnach podane są gatunki.

Logistic regresion

	Iris-setosa	Iris-versicolor	Iris-virginica
1.	0,00	0,03	0,97
2.	0,97	0,03	0,00
3.	0,98	0,02	0,00

Tree

	Iris-setosa	Iris-versicolor	Iris-virginica
1.	0,00	0,02	0,98
2.	1	0,00	0,00
3.	1	0,00	0,00

TASK 21

Tak, model Neural Networ sklasyfikował do grupy Iris-versicolor drugą pozycję zamiast Iris-setosa

TASK 22

Tree	Tree (Iris-setosa)	Tree (Iris-versicolor)	Tree (Iris-virginica)	Logistic Regression	Logistic Regression (Iris-setosa)	Logistic Regression (Iris-versicolor)
Iris-setosa Iris-versicolor Iris-virginica	continuous	continuous	continuous	Iris-setosa Iris-versicolor Iris-virginica	continuous	continuous
meta	meta	meta	meta	meta	meta	meta
Iris-setosa	1	0	0	Iris-setosa	0,982114059	0,01
Iris-setosa	1	0	0	Iris-setosa	0,966837812	0,03
Iris-virginica	0	0,02173913	0,97826087	Iris-virginica	6,97904E-06	0,03

Wrzucam też zrzut ekranu, ponieważ przycina wklejoną z `.csv` tabelę.

Tree	Tree (Iris-setosa)	Tree (Iris-versicolor)	Tree (Iris-virginica)	Logistic Regression	Logistic Regression (Iris-setosa)	Logistic Regression (Iris-versicolor)	Logistic Regression (Iris-virginica)	sepal length	sepal width	petal length	petal width	iris
Iris-setosa Iris-versicolor Iris-virginica	continuous	continuous	continuous	Iris-setosa Iris-versicolor Iris-virginica	continuous	continuous	continuous	continuous	continuous	continuous	continuous	continuous
meta	meta	meta	meta	meta	meta	meta	meta					
Iris-setosa	1	0	0	Iris-setosa	0,982114059	0,017885924	1,72142E-08	5,2	3,6	1,4	0,3	
Iris-setosa	1	0	0	Iris-setosa	0,966837812	0,033162168	1,93273E-08	7,2	3,1	0,5	1,4	
Iris-virginica	0	0,02173913	0,97826087	Iris-virginica	6,97904E-06	0,032682651	0,96731037	6,1	1,3	5	2,5	