

## Task1

iris.tab

## Task2

Zestaw danych Iris zawiera informacje dotyczące trzech różnych gatunków irysów: setosa, versicolor i virginica. Każdy gatunek reprezentowany jest przez 50 próbek. Dla każdej próbki zbierane są cztery cechy: długość i szerokość płatków oraz długość i szerokość działek kielicha. Zestaw danych Iris jest często wykorzystywany do zadania klasyfikacji, gdzie celem jest przewidywanie gatunku irysa na podstawie podanych cech. Jest to popularny zestaw danych do nauki maszynowej ze względu na swoją prostotę i dobrze zdefiniowane klasy.

## Task3

Classification Tree

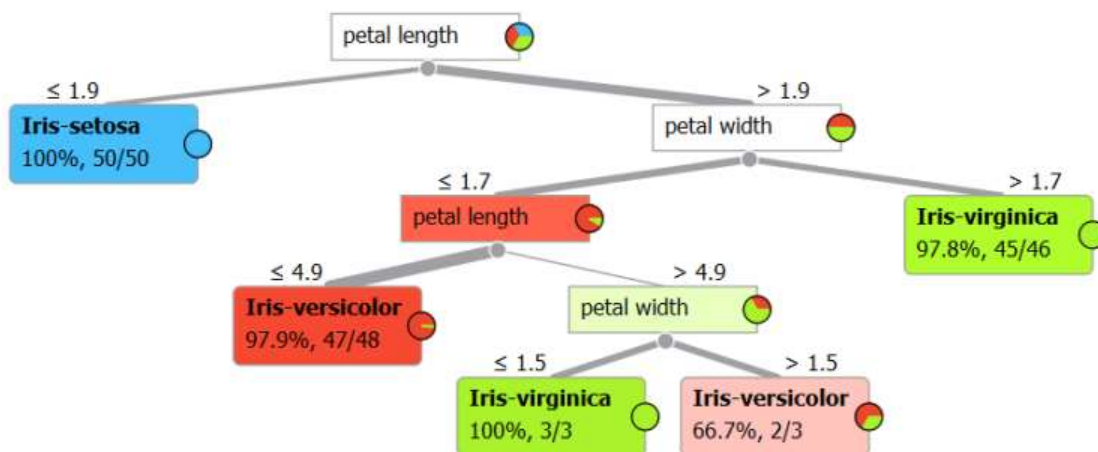
## Task4

Drzewo klasyfikacyjne w widoku przedstawia kolejne podziały danych na podstawie różnych cech (petal length, petal width) oraz wartości progowych. Te podziały umożliwiają klasyfikację obserwacji do odpowiednich klas (Iris-setosa, Iris-versicolor, Iris-virginica) na podstawie wartości cech. Np. jeśli petal length mniej lub równa się 1.9, obserwacje są przypisane do klasy "Iris-setosa", jeżeli większa to następuje kolejne rozgałęzienie na podstawie atrybutu petal width.

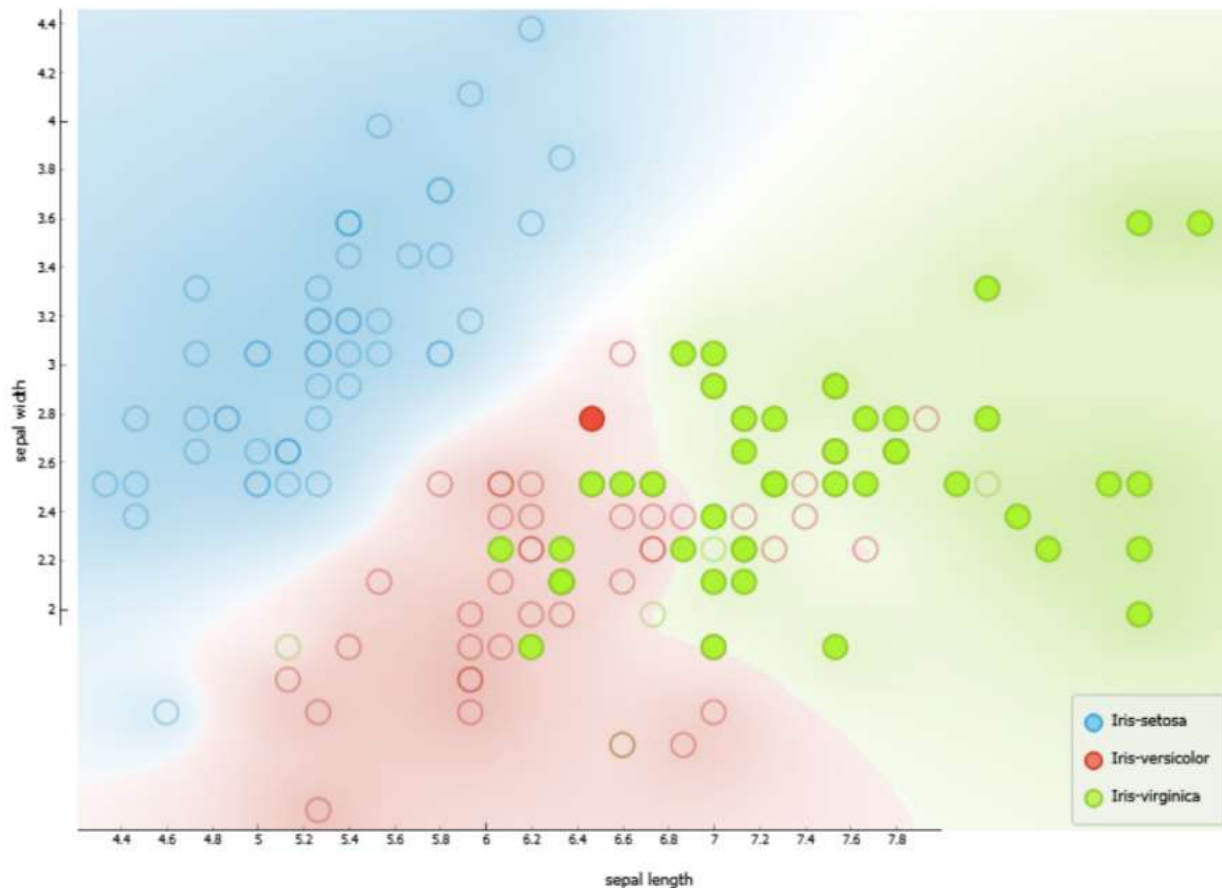
## Task5

5 levels

## Task6



## Task7



## Task8

150

## Task9

Na podstawie wyników można stwierdzić, że model regresji logistycznej dobrze przewiduje gatunek. Wysoka wartość AUC (0.998) wskazuje na bardzo dobrą zdolność modelu do rozróżniania między różnymi gatunkami. Wysoka dokładność (CA) o wartości 0.967 wskazuje na wysoki odsetek poprawnych predykcji gatunku przez model. Wysokie wartości F1, Prec, Recall wskazują na dobrą równowagę między precyzją i czułością modelu w przewidywaniu gatunku. Wartość MCC (0.950) sugeruje, że model ma silną korelację między przewidywaniami a rzeczywistymi wartościami gatunku.

## Task10

Radzi sobie najlepiej Logistic Regression z powodu że ma 5 ze 150 złe zaklasyfikowanych rekordów, podczas gdy Tree oraz SVM mają 6 ze 150. Z tego też wynika że różnica między nimi nie jest duża.

## Task11

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

## Task12

"Data sampler" służy do tworzenia próbek lub podzbiorów danych ze zbioru danych wejściowych. Jest to narzędzie, które umożliwia użytkownikom wybór losowych lub systematycznych próbek z danego zbioru danych w celu analizy, wizualizacji lub przetwarzania tych próbek w dalszej części analizy danych. "Data sampler" pozwala dostosować rozmiar próbki oraz sposób wyboru próbek, takie jak losowe próbkowanie, próbkowanie systematyczne, próbkowanie proporcjonalne itp.

## Task13

Algorytm "Test and Score" umożliwia ocenę skuteczności modelu predykcyjnego na podstawie danych testowych i dostarcza różne metryki oceny, które pozwalają porównać różne modele i zrozumieć, jak dobrze radzą sobie w przewidywaniu odpowiednich wartości na podstawie dostępnych danych.

## Task14

Algorytm "Test and Score" pozwala na ocenę skuteczności modelu predykcyjnego na większym i bardziej zróżnicowanym zbiorze danych poprzez połączenie danych zawartych w dwóch plikach. Dzięki temu algorytmowi możliwe jest scalenie tych danych w jeden większy zbiór danych, który może być wykorzystany do testowania modelu. W ten sposób użytkownicy mogą ocenić, jak dobrze model radzi sobie z przewidywaniem wyników na rozszerzonym i bardziej zróżnicowanym zbiorze danych.

## Task15

5

## Task16

Naive Bayes jest najgorzy i jego CA wynosi 0,873

## Task17

Dla gatunku Iris-setosa wszystkie modele okazały się idealnie

## Task18

19

## Task19

Zostały sklasyfikowane gatunki Iris-setosa oraz Iris-virginica

Predictions - Orange

Predictions - Orange

— — — — —

[illegible]

dołączyłem plik csv oraz xlsx

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	Logistic R <sub>0</sub>	Logistic R <sub>1</sub>	Logistic R <sub>2</sub>	Logistic R <sub>3</sub>	Logistic R <sub>4</sub>	Tree	Iris-v	Tree	Iris-v	Neural Net	Neural Net	Neural Net	ANN	ANN	Iris-v	ANN	Iris-v	Random F	Random F	Random F	Gradient B	Gradient B	Gradient B	Naive Bay	Naive Bay	Naive Bay
2	Iris-setosa	continuous	continuous	continuous	Iris-setosa	continuous	continuous	Iris-setosa	continuous	Iris-setosa	continuous	continuous	Iris-setosa	continuous	Iris-v	ANN	Iris-v	Random F	Random F	Random F	Gradient B	Gradient B	Gradient B	Naive Bay	Naive Bay	Naive Bay
3	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta	meta
4	Iris-setosa	0.982114	0.017886	1.72E-08	Iris-setosa	1	0	0	Iris-setosa	0.990757	0.007712	0.001531	Iris-setosa	1	0	0	Iris-setosa	1	0	0	Iris-setosa	0.999953	4.44E-05	2.7E-06	Iris-setosa	0.995843
5	Iris-setosa	0.966838	0.033162	1.93E-08	Iris-setosa	1	0	0	Iris-setosa	0.038874	0.909468	0.051658	Iris-setosa	1	0	0	Iris-setosa	0.5	0.4	0.1	Iris-setosa	0.566312	0.43087	0.002818	Iris-virginica	0.059511
6	Iris-virginica	6.98E-06	0.032683	0.96731	Iris-virginica	0	0.021739	0.978261	Iris-virginica	2.66E-06	0.02205	0.977947	Iris-virginica	0	0.2	0.8	Iris-virginica	0	0	1	Iris-virginica	1.04E-06	4.48E-06	0.999994	Iris-virginica	1.37E-05