

Analiza dyskryminacyjna

Damian Guzek i Emilia Koniszewska

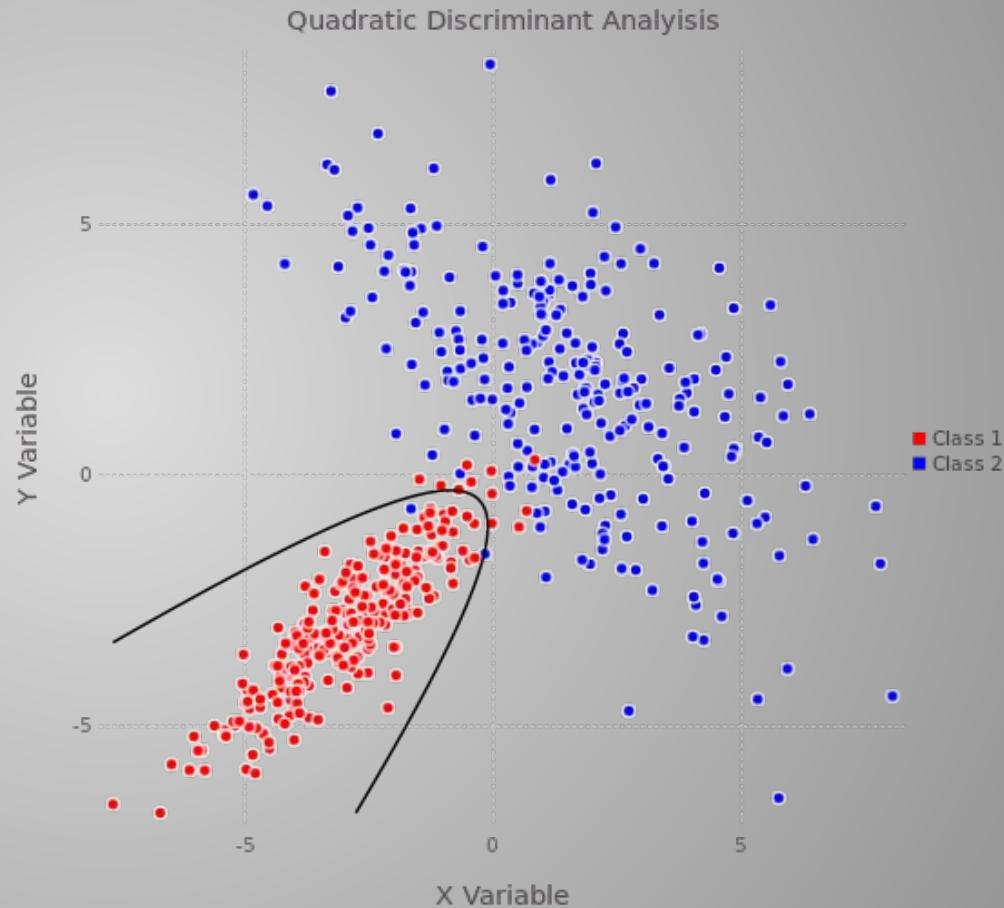
Dyskryminacja (łac. *discrimino* – rozdzielim) w ogólności oznacza odmienne traktowanie różnych obiektów, które znajdują się w podobnej sytuacji.

Analiza dyskryminacyjna to metoda pozwalająca badać różnice między grupami obiektów w oparciu o zbiór wybranych zmiennych, zwanyimi również predyktorami.

W szczególności, jest ona stosowana do rozstrzygania, które zmienne pozwalaj w najlepszy sposób dzielić dany zbiór przypadków na występujące w naturalny sposób grupy.

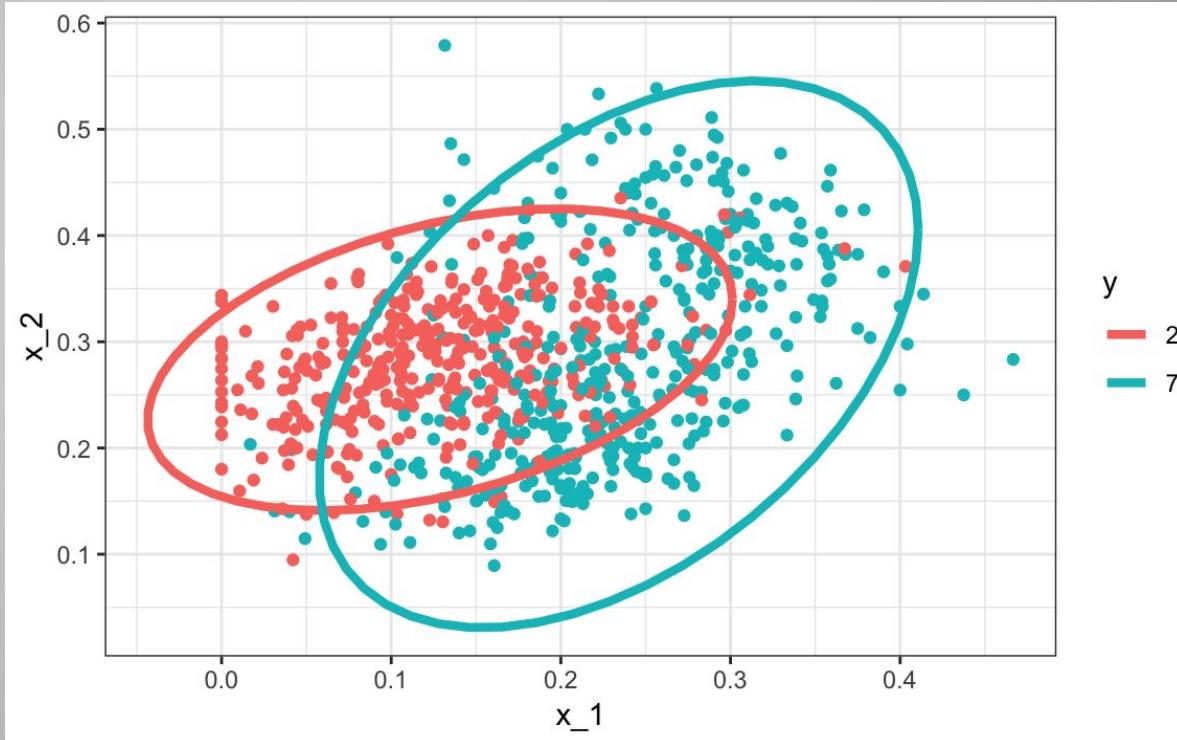
Graficznie przedstawiony cel analizy dyskryminacji

przypadek prosty:



Graficznie przedstawiony cel analizy dyskryminacji

przypadek trudny:



Analiza dyskryminacyjna jest efektywnie stosowana w dziedzinach nauki oraz praktyki społecznej takich, jak:

- nauki biologiczno-medyczne,
- psychologia,
- socjologia,
- politologia,
- ekonomia,
- prawo

Przykłady:

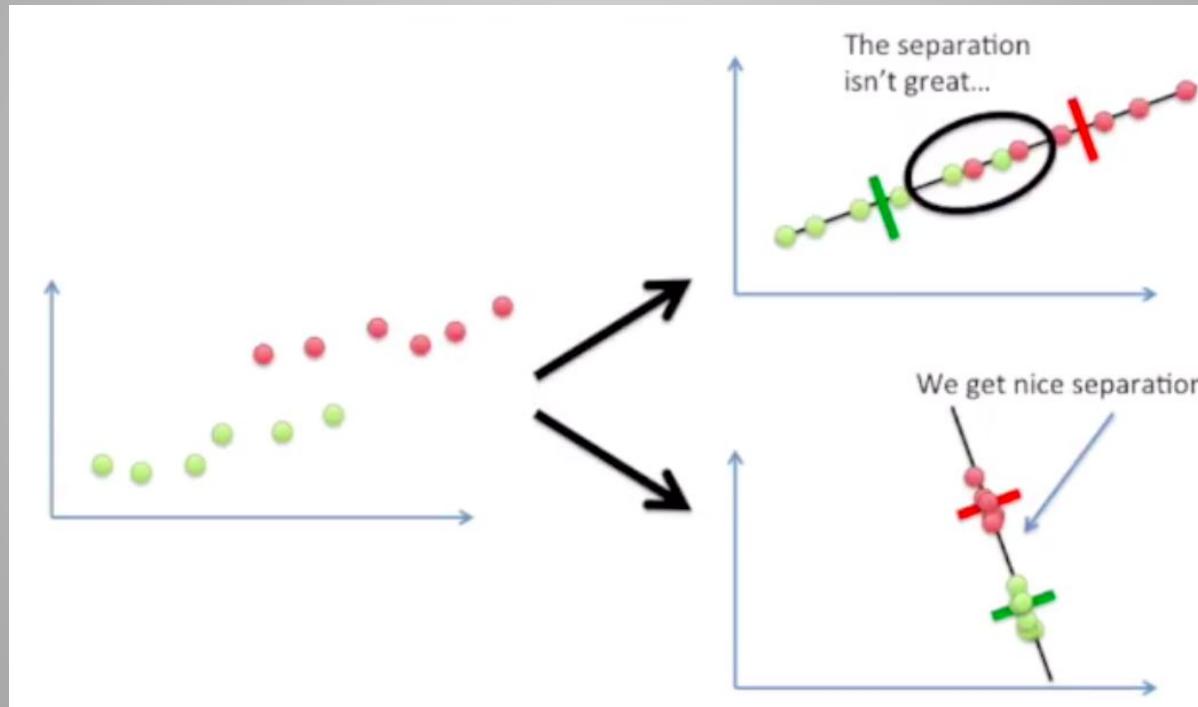
- wykrywanie grupy

| | | |
|---|---|-----------------------|
|  | Flu Forecasting Predict when, where and how strong the flu will be Masters · 5 years ago · Limited | \$125,000 50 teams |
|---|---|-----------------------|

- znajdowanie złóż soli

| | | |
|---|--|--------------------------|
|  | TGS Salt Identification Challenge Segment salt deposits beneath the Earth's surface Featured · 5 months ago · geology, image data | \$100,000 3,234 teams |
|---|--|--------------------------|

Przykład separacji danych



Założenia analizy dyskryminacyjnej

Rozkład normalny

Zakłada się, że dane (ujęte w postaci zmiennych) reprezentują próbę z wielowymiarowego rozkładu normalnego. Łatwo można zbadać, czy zmienne podlegają bądź nie podlegają rozkładowi normalnemu za pomocą histogramów rozkładów liczności. Zauważmy jednak, że naruszanie założenia o normalności zazwyczaj nie jest "zgubne" w tym sensie, że otrzymywane testy istotności itd. pozostają "godne zaufania". Dodatkowo (obok wykresów), możemy także użyć specjalnych testów służących do oceny normalności rozkładu.

Jednorodność wariancji/kowariancji

Zakłada się, że macierze wariancji/kowariancji zmiennych są jednorodne (homogeniczne) w grupach. I znów, nieznaczne odchylenia nie są aż tak ważne. Zanim jednak przyjmiemy ostateczne wnioski z ważnego badania, warto przejrzeć macierze wariancji wewnętrzgrupowych i macierze korelacji. W szczególności, warto w tym celu wykorzystać macierz wykresów rozrzutu. Jeśli mamy wątpliwości, możemy powtórzyć analizy eliminując jedną lub dwie mniej interesujące grupy. Jeśli ogólne wyniki (interpretacje) pozostaną takie same, to prawdopodobnie wszystko jest w porządku. Wielowymiarowy test *M Boxa* służący do oceny jednorodności wariancji/kowariancji jest szczególnie wrażliwy na odchylenia od wielowymiarowej normalności i nie powinien być traktowany zbyt "poważnie".

Korelacje między średnimi i wariancjami

Podstawowe "rzeczywiste" zagrożenie dla trafności testów istotności pojawia się wówczas, gdy średnie zmiennych w grupach są skorelowane z wariancjami (lub odchyleniami standardowymi). Intuicyjnie, jeśli w grupie występuje duża zmienność przy szczególnie wysokich średnich niektórych zmiennych, to te wysokie średnie nie są rzetelne. Ogólne testy istotności są jednak oparte na łącznych wariancjach, to znaczy na przeciętnej wariancji z wszystkich grup. Zatem testy istotności dla relatywnie większych średnich (przy dużych wariancjach) byłyby oparte na relatywnie mniejszych wariancjach łącznych, dając w efekcie mylny obraz istotności statystycznej. W praktyce taki układ może się pojawić wtedy, gdy jedna z badanych grup zawiera kilka odstających obserwacji, które mają duży wpływ na średnie, a także zwiększą zmienność. Aby ustrzec się przed tym problemem należy na okoliczność takich powiązań zbadać statystyki opisowe, to znaczy średnie i odchylenia standardowe lub wariancje.

Problem złego uwarunkowania macierzy

Inne założenie analizy funkcji dyskryminacyjnej wymaga, by zmienne wykorzystane do dyskryminacji grup nie były całkowicie redundantne. W ramach obliczeń związanych z analizą dyskryminacyjną występuje operacja odwracania macierzy wariancji/kowariancji zmiennych w modelu. Jeśli któraś ze zmiennych jest całkowicie redundantna wobec innych zmiennych, to o macierzy mówi się, że jest *źle uwarunkowana* i nie może być odwrócona. Na przykład, jeśli zmienna jest sumą trzech innych zmiennych, które także znajdują się w modelu, to macierz jest źle uwarunkowana.

Wartości tolerancji

Aby uniknąć złego uwarunkowania macierzy, stale będzie sprawdzane dla każdej zmiennej tak zwane wartości tolerancji. Wartość tolerancji jest obliczana jako $1 - R\text{-kwadrat}$ danej zmiennej przy włączeniu do bieżącego modelu wszystkich innych zmiennych. Jest to zatem część wariancji wyjaśniana przez daną zmienną.

Funkcja dyskryminacyjna

Funkcja predyktorów, która tworzy nową zmienną i służy do określania przynależności danego obiektu do klasy nazywana jest funkcją dyskryminacyjną (rozpoznawczą).

W przeprowadzanej metodzie funkcji dyskryminacyjnych może być więcej niż jedna. Liczba funkcji musi być mniejsza od liczby klas oraz nie większa niż liczba predyktorów. Najczęściej używane są funkcje liniowe.

Cechy funkcji dyskryminacyjnej:

- daje wynik w postaci jednej liczby,
- osiąga największą wartość dla klasy, do której ma przynależeć badany obiekt.

Funkcją dyskryminacyjną może być liniowa kombinacja zmiennych niezależnych, służąca jako kryterium przyporządkowywania obserwacji do grup. Informacje, których nośnikami są predyktory, zapisywane są jako wyniki funkcji dyskryminacyjnej.

Równanie tej funkcji przypomina analizę regresji wielokrotnej i wygląda następująco:

$$D_{tm} = b_0 + b_1 \cdot x_{1\ tm} + b_2 \cdot x_{2\ tm} + \dots + b_p \cdot x_{p\ tm}$$

gdzie:

D_{tm} – wartość funkcji dyskryminacyjnej dla obserwacji m w grupie t (wynik dyskryminacyjny)

$x_{k\ tm}$ – wartość k -tej zmiennej dyskryminującej dla obserwacji m w grupie t

b_k – współczynnik dyskryminacyjny dla zmiennej x_k reprezentujący jej ważony efekt, $k = 1, 2, \dots, p$

b_0 – stała.

Funkcja klasyfikacyjna

Funkcja, która może być wykorzystana do rozstrzygnięcia, do której grupy najprawdopodobniej należą poszczególne przypadki, nazywana jest funkcją klasyfikacyjną.

Jest tyle funkcji klasyfikacyjnych, ile jest grup.

Funkcje klasyfikacyjne separują poszczególne grupy obiektów w możliwie najlepszy sposób.

Funkcje klasyfikacyjne pozwalają obliczyć wartości klasyfikacyjne (K) dla każdego przypadku w każdej grupie. Wzór funkcji klasyfikacyjnej może wyglądać następująco

$$K_t = c_t + d_{t1} \cdot x_1 + d_{t2} \cdot x_2 + \dots + d_{tp} \cdot x_p,$$

gdzie:

t - określa daną grupę

p - liczba zmiennych

c_t - stała dla grupy t

d_{tk} - waga dla k -tej zmiennej przy obliczaniu wartości klasyfikacyjnej

dla grupy t

x_k - wartość obserwowana dla danego przypadku dla k -tej zmiennej, $k = 1, 2, \dots, p$

Autorem koncepcyjnych i matematycznych podwalin analizy dyskryminacyjnej był Ronald Fisher (1890-1962), wybitny genetyk oraz matematyk, pionier statystyki, znany także z innej metody analizy statystycznej, znacznie bardziej rozpowszechnionego modelu analizy wariancji.

Jądrowa estymacja gęstości

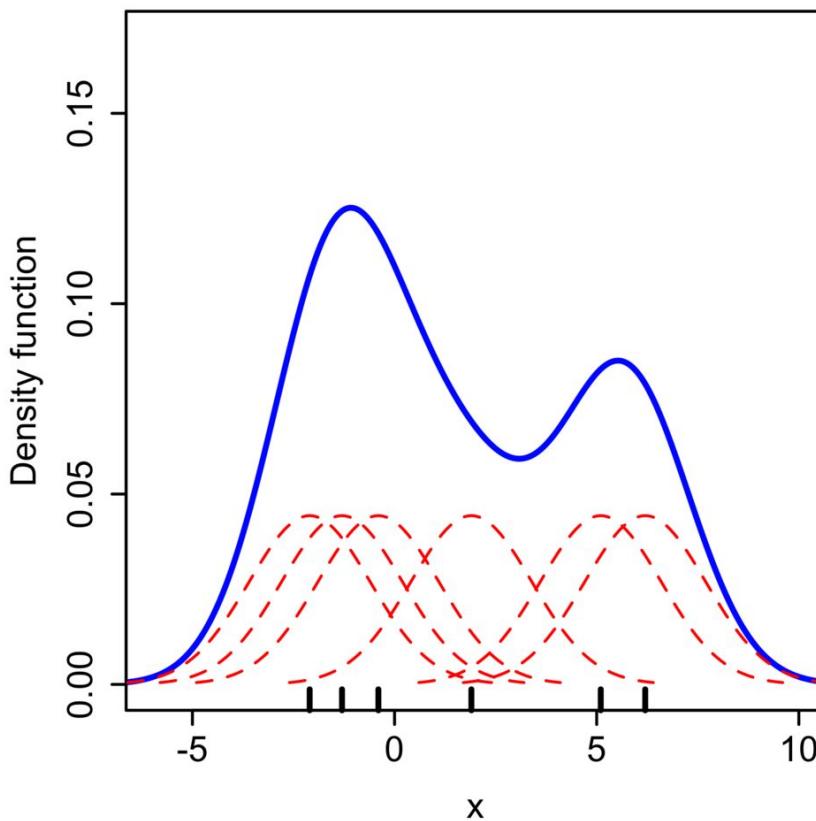
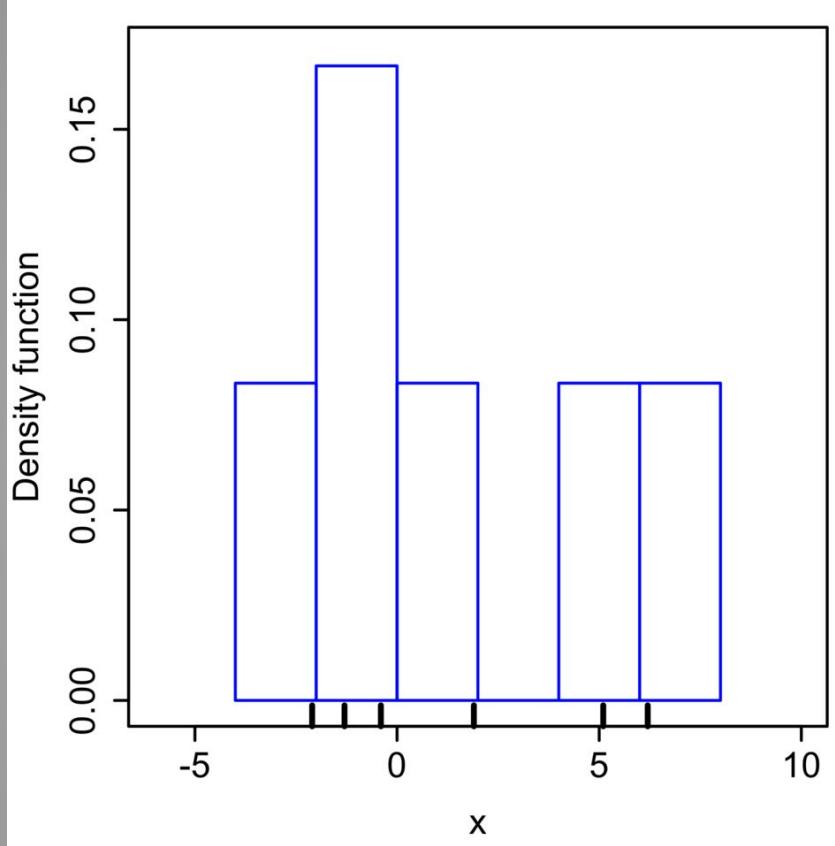
Niech dana będzie n -wymiarowa zmienna losowa X , której rozkład posiada gęstość f . Jej estymator jądrowy $\hat{f}: R^n[0, \infty)$ wyznacza się w oparciu o wartości m -elementowej próby losowej $x_1, x_2, x_3, \dots, x_m$ uzyskanej ze zmiennej X_i w swej podstawowej formie jest on definiowany wzorem:

$$\hat{f}(x) = \frac{1}{mh^n} \sum_{i=1}^m K\left(\frac{x - x_i}{h}\right)$$

gdzie:

h - parametr wygładzania

K - jądro (funkcja)



Interpretacja

1. Ze statystycznego punktu widzenia, postać jądra nie ma istotnego znaczenia i wybór funkcji K może być arbitralny, uwzględniając jednak pożądane własności takie jak ciągłość, różniczkowalność itp.
2. W przypadku jednowymiarowym jako funkcję K przyjmuje się klasyczne postacie gęstości rozkładów probabilistycznych np. rozkład normalny, Cauchy'ego.

3. W przeciwnieństwie do postaci jądra przyjęta wartość parametru wygładzania ma istotny wpływ na jakość otrzymanego estymatora.

- zbyt mała wartość parametru h powoduje pojawienie się znacznej ilości ekstremów lokalnych estymatora f , co jest sprzeczne z faktycznymi własnościami realnych populacji
- za duże wartości skutkują nadmiernym wygładzeniem tego estymatora, maskując przy tym specyficzne cechy badanego rozkładu

Opracowane zostały dogodne algorytmy umożliwiające obliczenie zbliżonej do optymalnej wartości parametru wygładzania, na podstawie próby losowej $x_1, x_2, x_3, \dots, x_m$. Uniwersalną metodą jest **walidacja krzyżowa**, czyli metoda statystyczna polegająca na podziale próby statystycznej na podzbiory, a następnie przeprowadzeniu wszelkich analiz na niektórych z nich, podczas gdy pozostałe służą do potwierdzenia wiarygodności jej wyników.

Bez zastosowania walidacji krzyżowej nie mamy pewności czy unikniemy zjawiska nadmiernego dopasowania modelu.

Wyróżniamy parę podstawowych rodzajów walidacji krzyżowych:

1. Prosta walidacja

Jest to najbardziej typowy rodzaj walidacji, w którym próbę dzieli się losowo na rozłączne zbiory: uczący i testowy. Zwykle zbiór testowy stanowi mniej niż 1/3 próby. Niektórzy nie zaliczają tego typu walidacji do metody sprawdzianu krzyżowego

2.

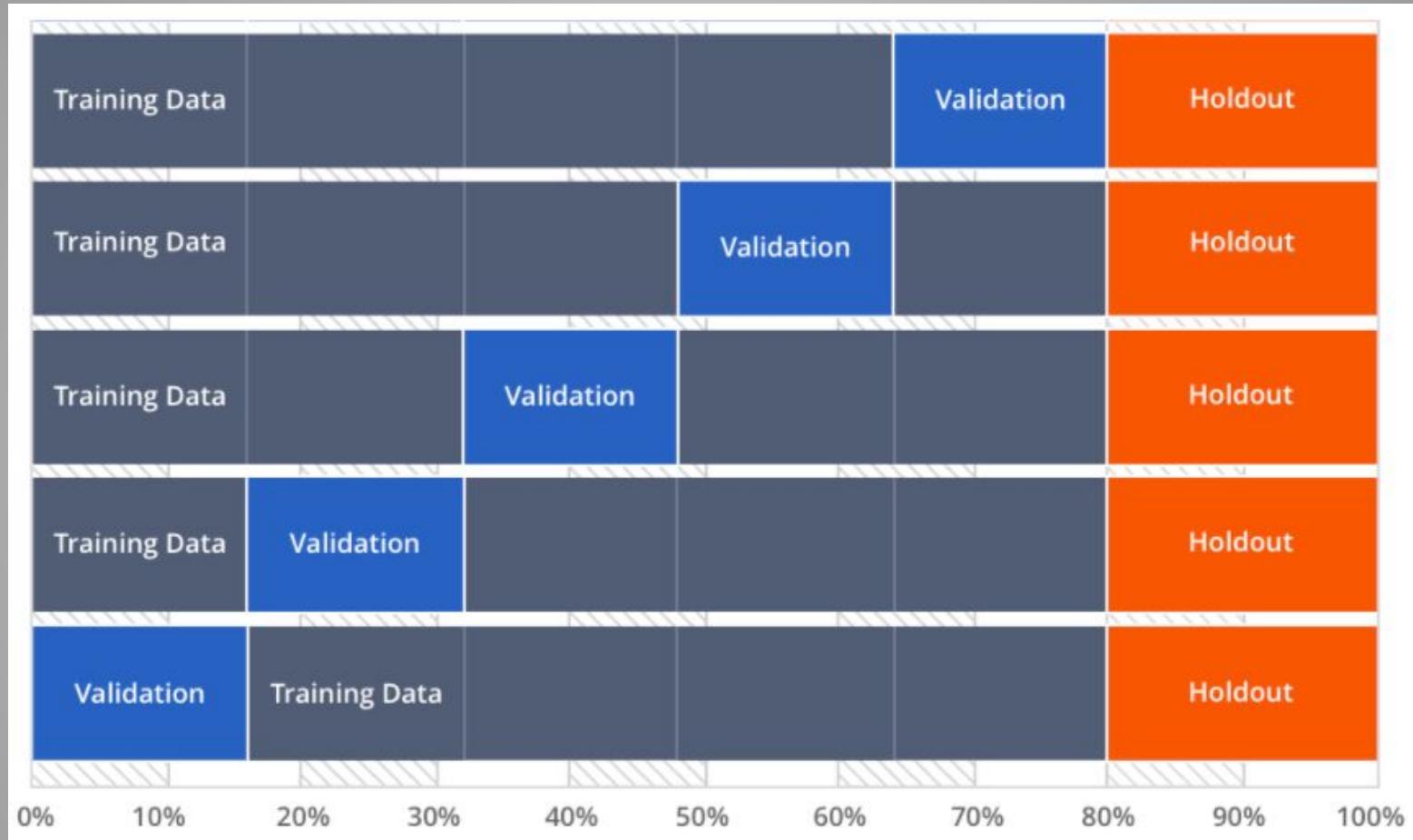
K-krotna

walidacja

W tej metodzie, oryginalna próba jest dzielona na K podzbiorów. Następnie kolejno każdy z nich bierze się jako zbiór testowy, a pozostałe razem jako zbiór uczący i wykonuje analizę. Analiza jest więc wykonywana K razy. K rezultatów jest następnie uśrednianych (lub łączonych w inny sposób) w celu uzyskania jednego wyniku.

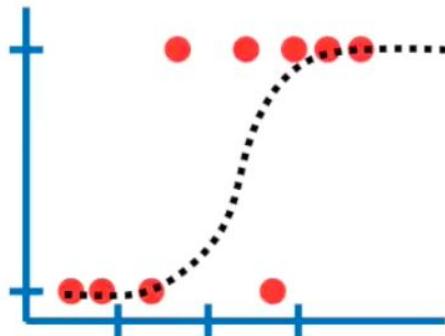
3. Leave-one-out

Jest to odmiana walidacji K -krotnej, gdy N -elementowa próba jest dzielona na N podzbiorów, zawierających po jednym elemencie. Stosowana często dla małych zbiorów danych.



In the end, every block of data is used for testing and we can compare methods by seeing how well they performed.

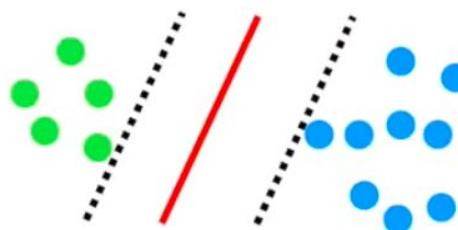
Logistic Regression



Correct
16

Incorrect
8

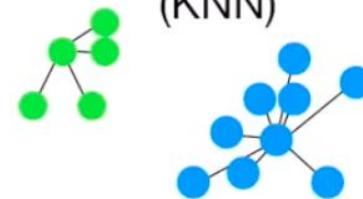
Support Vector machines (SVM)



Correct
18

Incorrect
6

K-nearest neighbors (KNN)



Correct
10

Incorrect
12

Centroid

Centroid to swojego rodzaju reprezentantem danego skupienia, przydatnym w szczególności przy interpretacji wyników analizy w formie zrozumiałej dla klienta. Można wówczas wybrać typowego reprezentanta każdego ze skupień i po sprawdzeniu jego cech nadać nazwę całemu skupieniu. Należy jednak pamiętać, aby nie stosować tego pojęcia w przypadku zbyt niejednorodnych grup.

Na centroidach opiera się większość hierarchicznej analizy skupień, gdzie coraz większe skupienia są w kolejnych krokach analizy zastępowane centroidami.

Odległość pomiędzy grupami

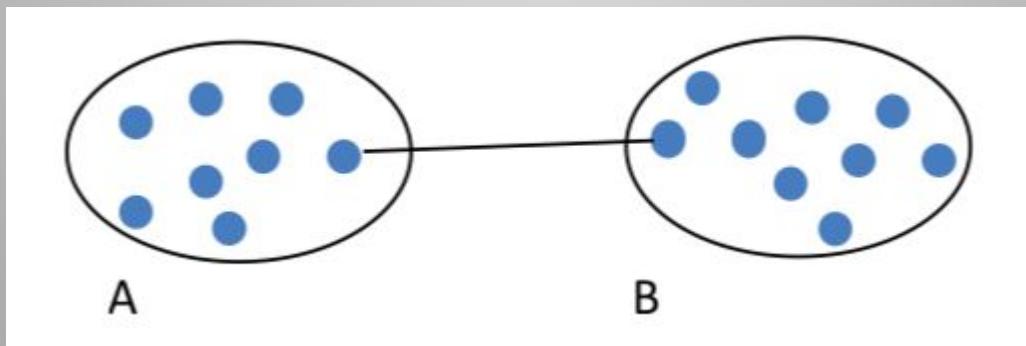
Odległość pomiędzy grupami możemy mierzyć na kilka sposobów. Opiszemy najpopularniejsze z metod.

Metoda

pojedynczego

wiązania

W tej metodzie odległość między dwoma skupieniami określona jest przez odległość między najbliższymi obiektami należącymi do różnych grup.

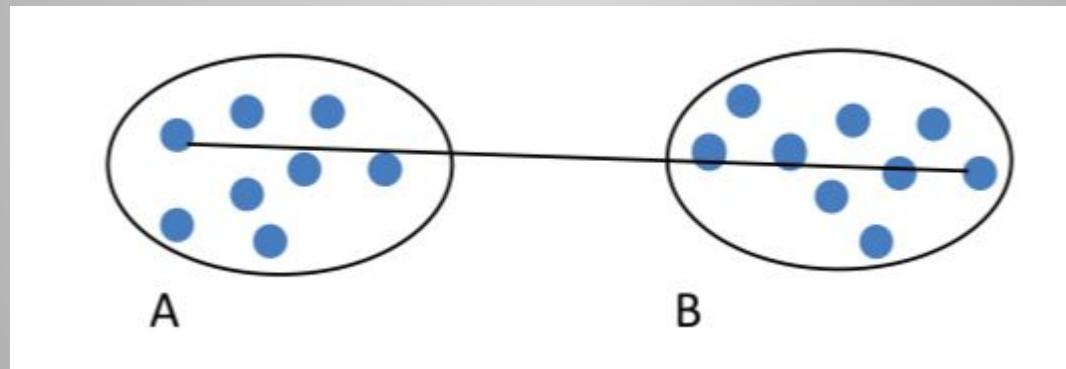


Metoda

pełnego

wiązania

W tej metodzie za odległość między grupami uznajemy odległość pomiędzy najdalej położonymi obiektami należącymi do różnych grup

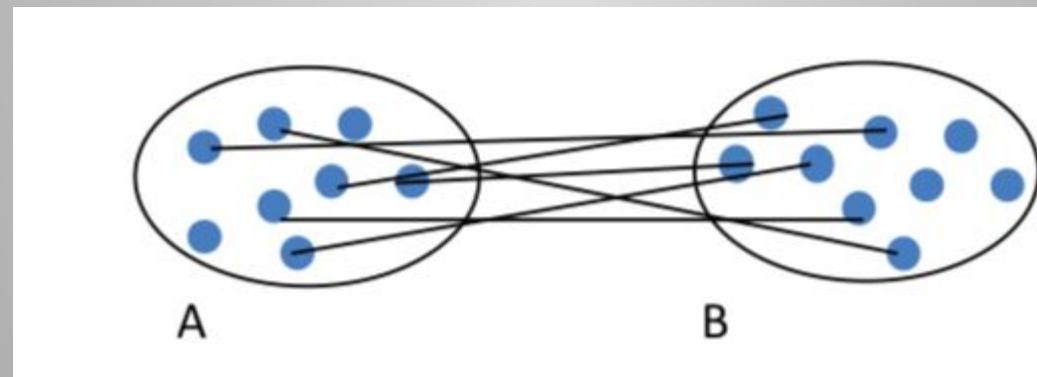


Metoda

średnich

połączeń

Odległość między dwoma grupami oblicza się za pomocą średniej arytmetycznej wyznaczonej ze wszystkich odległości obiektów należących do dwóch grup

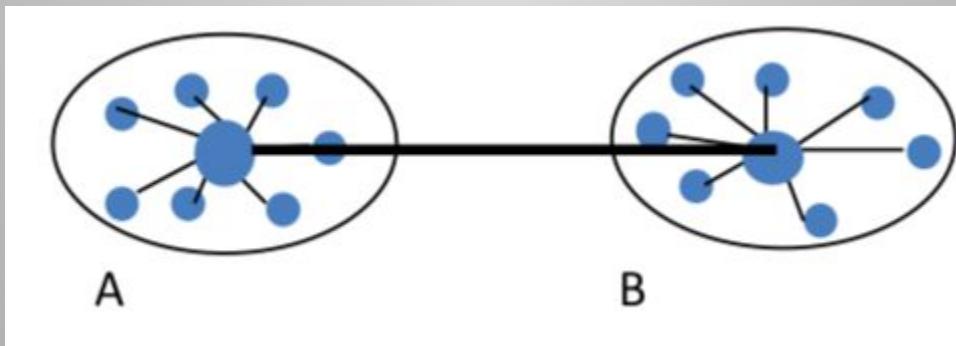


Metoda

środków

ciężkości

Odległość między dwoma grupami jest równa odległości pomiędzy ich środkami ciężkości

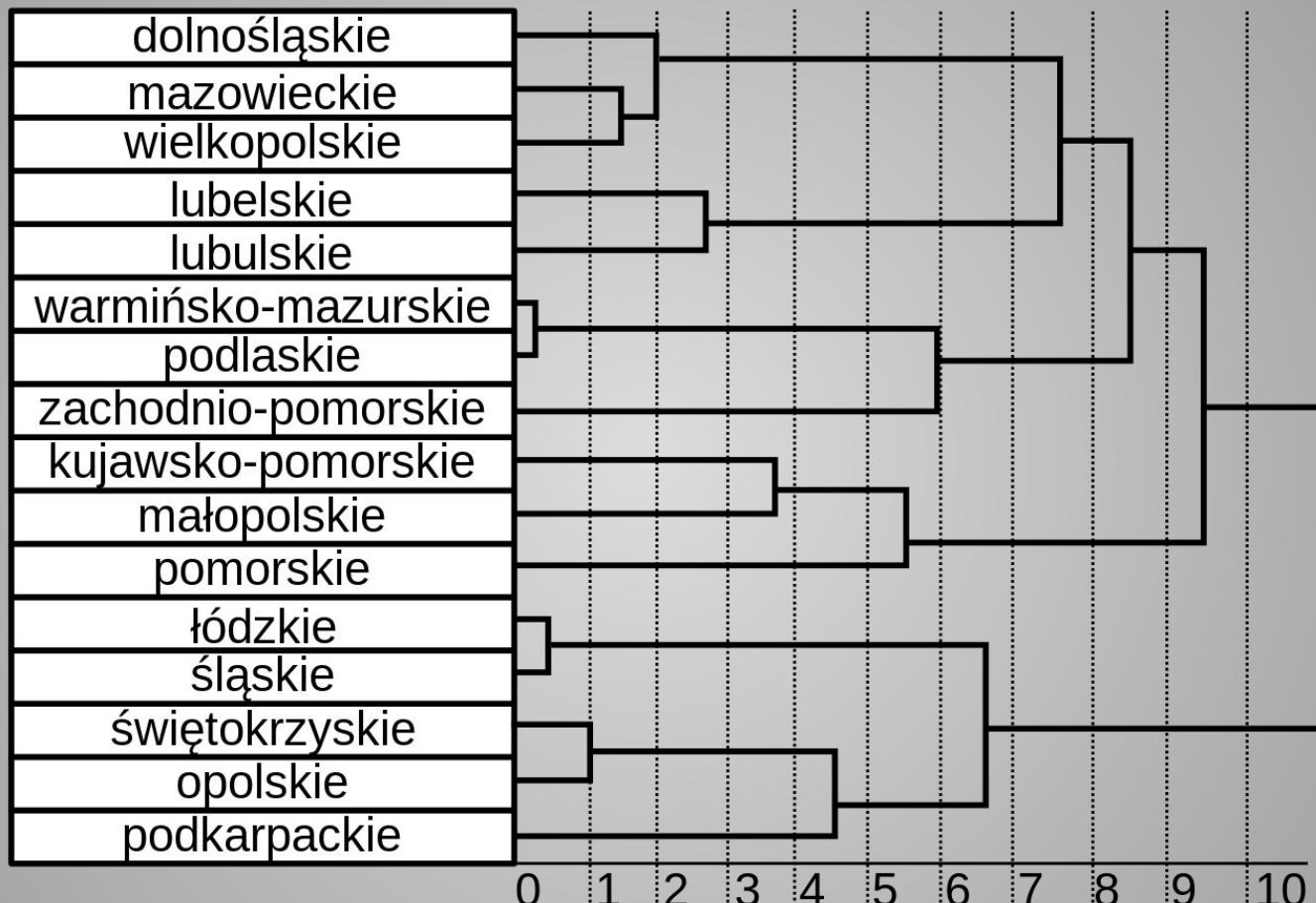


Metoda Warda

Do oszacowania odległości między grupami wykorzystuje podejście analizy wariancji. Metoda ta zmierza do minimalizacji sumy kwadratów odchyleń wewnętrz grup. Miarą zróżnicowania skupienia względem wartości średnich jest ESS (Error Sum of Squares).

$$ESS = \sum_{i=1}^k (x_i - \bar{x})^2$$

Metoda Warda jest traktowana jako najbardziej efektywna - mimo tendencji do tworzenia skupień o małej wielkości. W wyniku zastosowania tych miar odległości otrzymujemy dendrogram obrazujący odległości pomiędzy grupami.



Analiza dyskryminacyjna w SAS

Procedura STAT w programie SAS dla analizy dyskryminacyjnej dopasowuje dane z jedną zmienną klasyfikującą oraz wieloma zmiennymi analizowanymi. Celem analizy dyskryminacyjnej może być znalezienie jednej lub wielu z następujących rzeczy:

- matematycznej reguły lub funkcji dyskryminacyjnej, która przydziela daną obserwację do jednej z klas, bazując jedynie na danych pochodzących od zmiennych analizowanych,
- zbioru liniowych kombinacji zmiennych analizowanych, który najlepiej odzwierciedla różnice wśród klas,
- podzbioru zmiennych analizowanych, który najlepiej ukazuje różnice wśród klas.

Procedury analizy dyskryminacyjnej w SAS

Do procedur analizy dyskryminacyjnej SAS/STAT należą:

- CANDISC — przeprowadza kanoniczną analizę dyskryminacyjną, oblicza kwadrat odległości Mahalanobisa pomiędzy klasami średnich, oraz dokonuje jednoczynnikowej i wieloczynnikowej jednostronnej analizy wariancji,
- DISCRIM - rozwija kryterium dyskryminacyjne przez klasyfikowanie każdej obserwacji do jednej z grup,
- STEPDISC — mając zmienną klasyfikacji i zmienne analizowane, procedura wykonuje krokową analizę dyskryminacyjną, aby wybrać najlepszy podzbiór zmiennych analizowanych dyskryminujący klasy.

Ogólna forma PROC DISCRIM

Schemat

procedury:

```
PROC DISCRIM <opcje>;
  VAR zmienne analizowane;
  CLASS zmienna klasyfikujaca;
  PRIORS prawdopodobienstwa;
RUN;
```

Jedynie PROC DISCRIM oraz zmienna CLASS są wymagane.

Opcje procedury DISCRIM (dostępne w SAS Enterprise Guide):

DATA - zbiór analizowany

TESTLIST - wyniki klasyfikacji dla zbioru testującego

CROSSVALIDATE - rezultaty zbiorcze dla walidacji krzyżowej

CROSSLIST - wyniki walidacji krzyżowej dla każdej obserwacji

LIST - wyniki resubstytucji dla każdej obserwacji

POSTERR - oceny błędu prawd. a posteriori

ANOVA - jednowymiarowy test równości średnich

MANOVA - wielowymiarowy test równości średnich

LISTERR - błędne sklasyfikowane wyniki resubstytucji

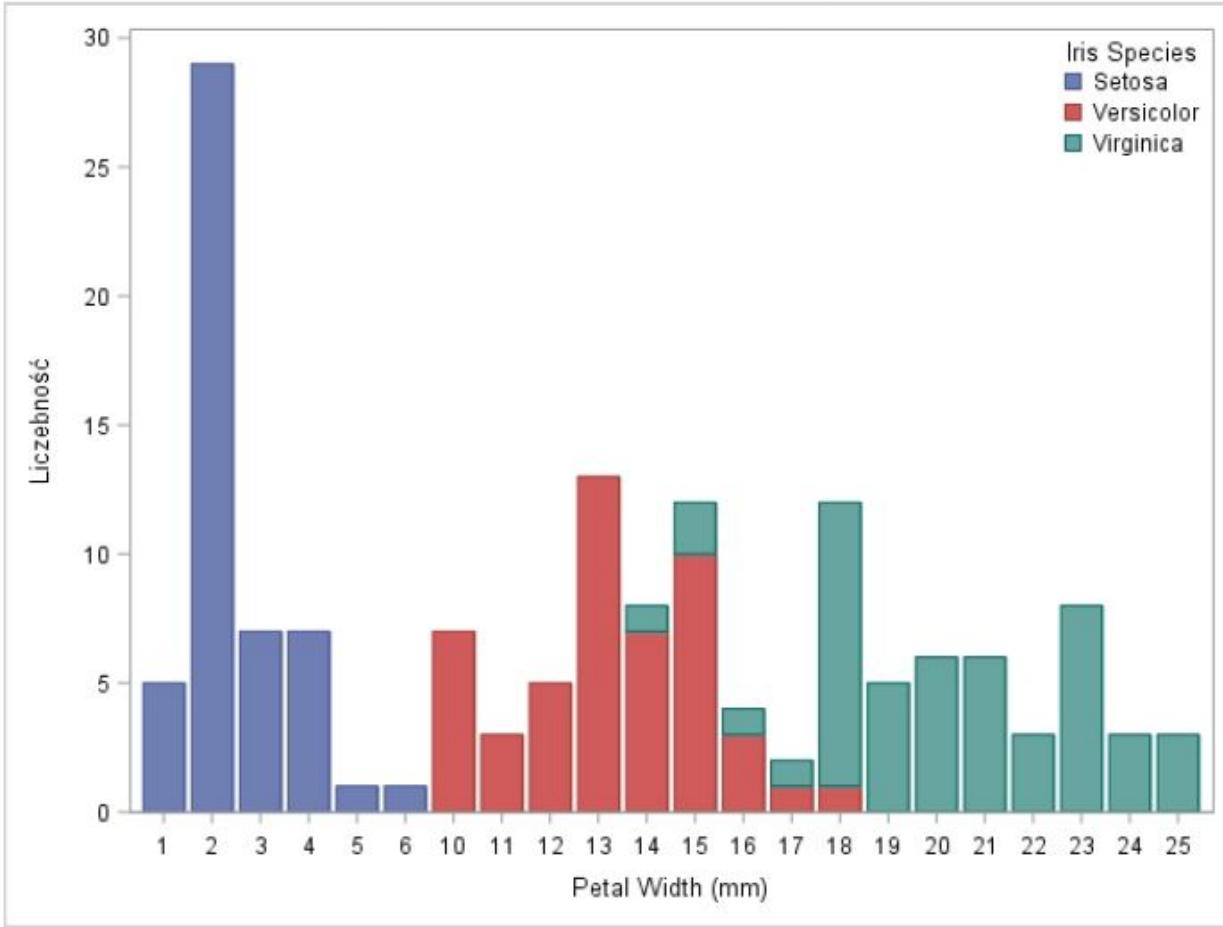
CROSSLISTERR - błędne sklasyfikowane wyniki walidacji krzyżowej

CAN - kanoniczna analiza dyskryminacyjna

TESTDATA=zbior testujący

Kod 1

Analiza dyskryminacji Fishera (1936) Iris



Kod 3

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego ze stałą wariancją

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 1 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 3

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego ze stałą wariancją

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Wyniki walidacji krzyżowej za pomocą Liniowa funkcja rozpoznawcza

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | | | | | |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.0952 | 0.9048 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.3828 | 0.6172 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.9610 | 0.0390 |
| 124 | Virginica | Versicolor | * | 0.0000 | 0.9940 | 0.0060 |
| 130 | Virginica | Versicolor | * | 0.0000 | 0.8009 | 0.1991 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.9610 | 0.0390 |

* Obserwacja błędnie zaklasyfikowana

Kod 3

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego ze stałą wariancją

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Podsumowanie walidacji krzyżowej za pomocą Liniowa funkcja rozpoznawcza

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 4 | 46 | 50 |
| | 0.00 | 8.00 | 92.00 | 100.00 |
| Razem | 50 | 52 | 48 | 150 |
| | 33.33 | 34.67 | 32.00 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0000 | 0.0400 | 0.0800 | 0.0400 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 3

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego ze stałą wariancją

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTTDATA

Podsumowanie klasyfikacji za pomocą Liniowa funkcja rozpoznawcza

| Profile obserwacji dla danych testowych | |
|---|----|
| Liczba obserwacji wczytanych | 71 |
| Liczba obserwacji użytych | 71 |

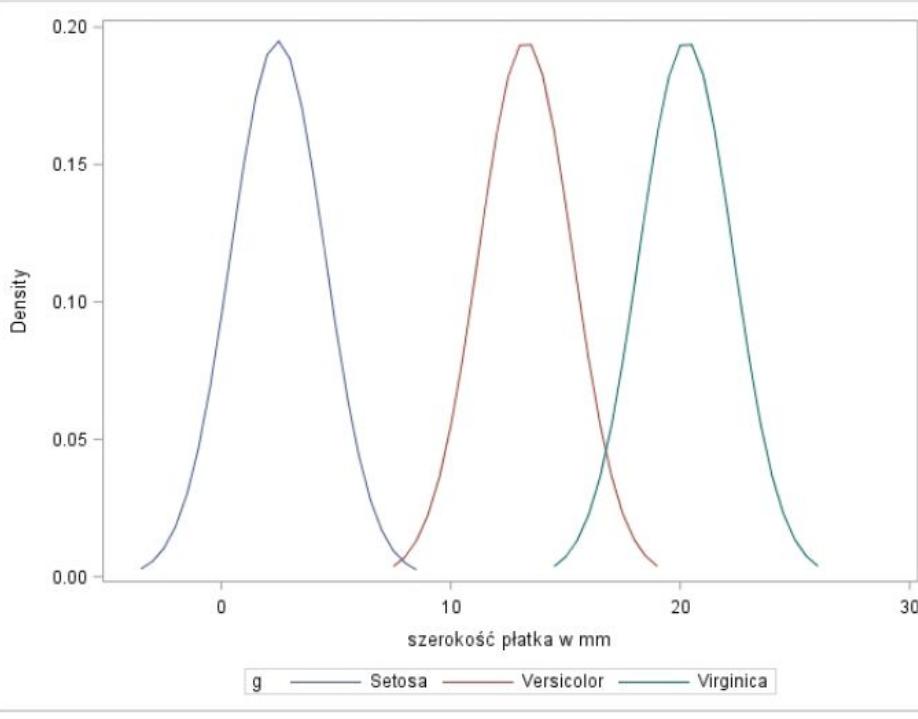
| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Razem |
| Razem | 26 | 18 | 27 | 71 |
| | 36.62 | 25.35 | 38.03 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

Kod 3

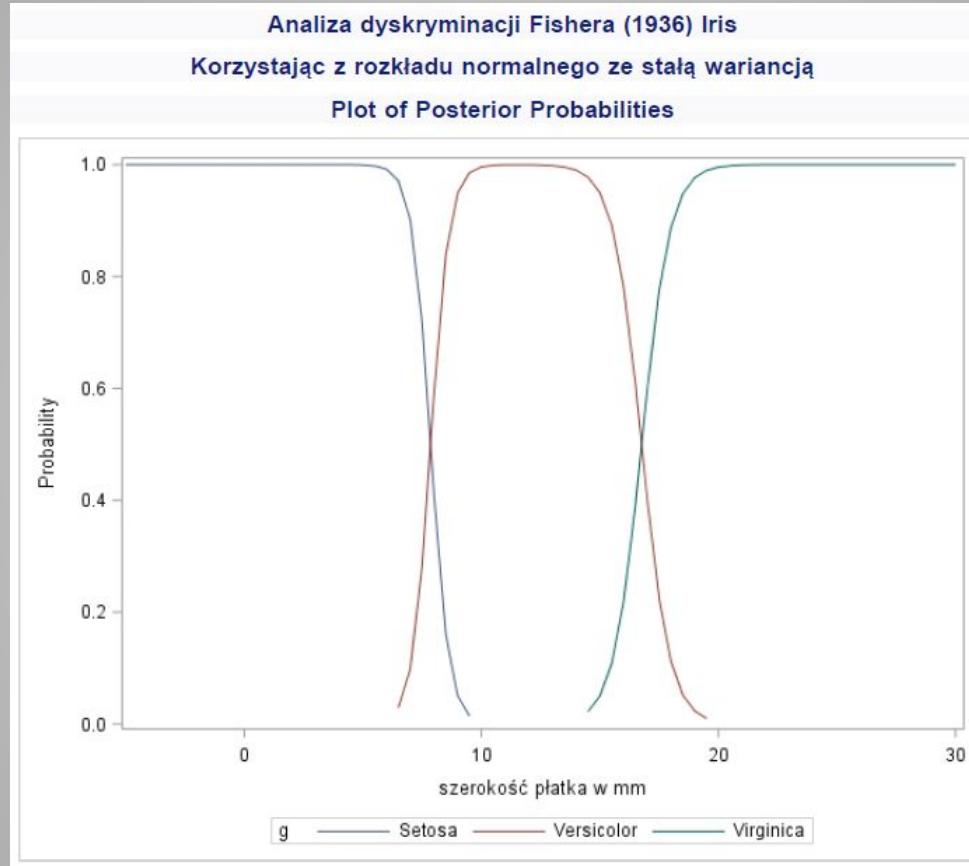
Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego ze stałą wariancją

Plot of Estimated Densities



Kod 3



Kod 4

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego bez zał. o stałości wariancji

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 1 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 4

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego bez zał. o stałości wariancji

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Wyniki walidacji krzyżowej za pomocą kwadratowej funkcji rozpoznawczej

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|------------|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | Setosa | Versicolor | | | |
| 10 | Setosa | Versicolor | * | 0.4923 | 0.5073 | 0.0004 |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.0686 | 0.9314 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.2871 | 0.7129 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.8740 | 0.1260 |
| 124 | Virginica | Versicolor | * | 0.0000 | 0.9602 | 0.0398 |
| 130 | Virginica | Versicolor | * | 0.0000 | 0.6558 | 0.3442 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.8740 | 0.1260 |

* Obserwacja błędnie zaklasyfikowana

Kod 4

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego bez zał. o stałości wariancji

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Podsumowanie walidacji krzyżowej za pomocą kwadratowej funkcji rozpoznawczej

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | | |
|--|---------|------------|-----------|--------|--|
| Od Species | Setosa | Versicolor | Virginica | Razem | |
| Setosa | 49 | 1 | 0 | 50 | |
| | 98.00 | 2.00 | 0.00 | 100.00 | |
| Versicolor | 0 | 48 | 2 | 50 | |
| | 0.00 | 96.00 | 4.00 | 100.00 | |
| Virginica | 0 | 4 | 46 | 50 | |
| | 0.00 | 8.00 | 92.00 | 100.00 | |
| Razem | 49 | 53 | 48 | 150 | |
| | 32.67 | 35.33 | 32.00 | 100.00 | |
| A priori | 0.33333 | 0.33333 | 0.33333 | | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0200 | 0.0400 | 0.0800 | 0.0467 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 4

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z rozkładu normalnego bez zał. o stałości wariancji

Procedura DISCRIM

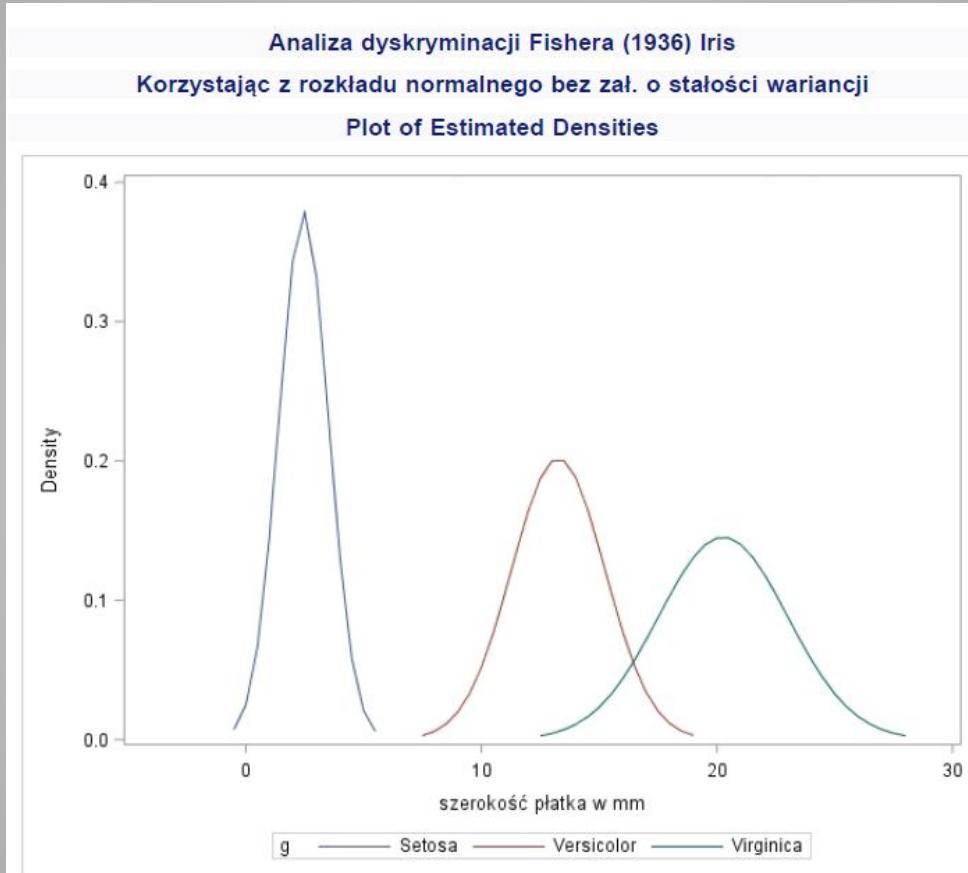
Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTTDATA

Podsumowanie klasyfikacji za pomocą kwadratowej funkcji rozpoznawczej

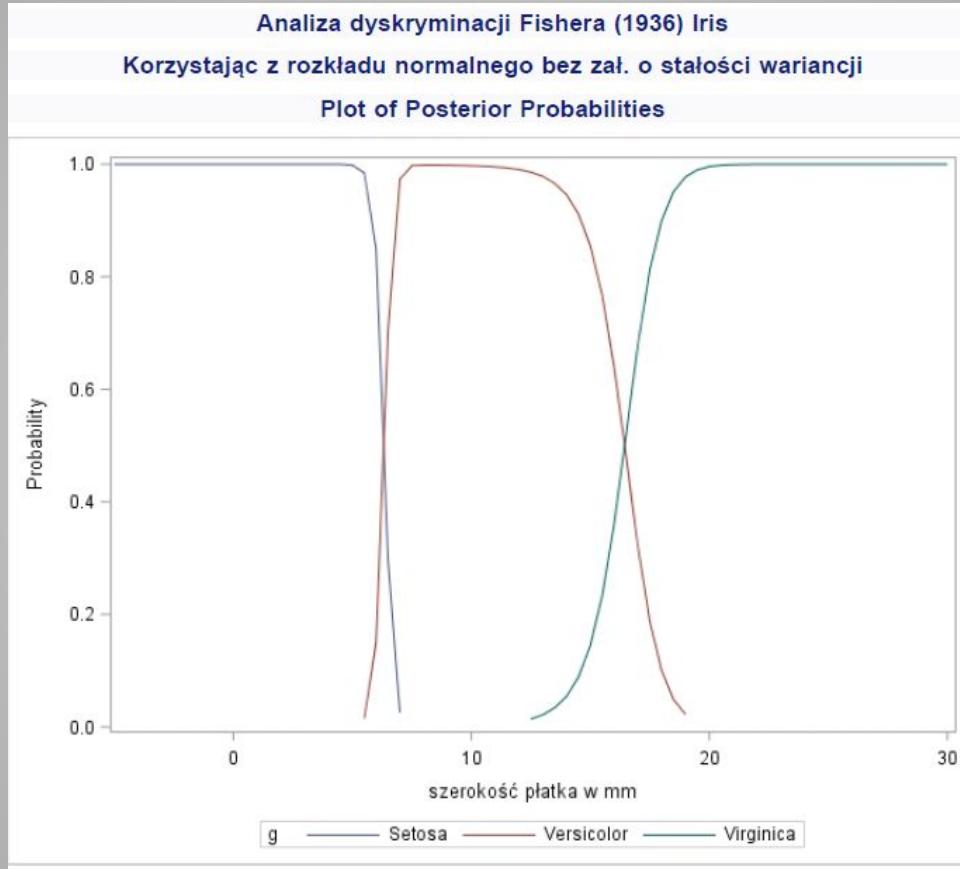
| Profile obserwacji dla danych testowych | |
|---|----|
| Liczba obserwacji wczytanych | 71 |
| Liczba obserwacji użytych | 71 |

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Razem |
| Razem | 23 | 20 | 28 | 71 |
| | 32.39 | 28.17 | 39.44 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

Kod 4



Kod 4



Kod 5

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 1 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 5

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Wyniki walidacji krzyżowej za pomocą gęstości jądra normalnego

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | | | | | |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.0438 | 0.9562 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.2586 | 0.7414 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.8827 | 0.1173 |
| 124 | Virginica | Versicolor | * | 0.0000 | 0.9472 | 0.0528 |
| 130 | Virginica | Versicolor | * | 0.0000 | 0.8061 | 0.1939 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.8827 | 0.1173 |

* Obserwacja błędnie zaklasyfikowana

Kod 5

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Podsumowanie walidacji krzyżowej za pomocą gęstości jądra normalnego

Liczba obserwacji i procent zaklasyfikowanych
do Species

| Od Species | Setosa | Versicolor | Virginica | Razem |
|------------|---------|------------|-----------|--------|
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 4 | 46 | 50 |
| | 0.00 | 8.00 | 92.00 | 100.00 |
| Razem | 50 | 52 | 48 | 150 |
| | 33.33 | 34.67 | 32.00 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

Oceny liczby błędów dla Species

| | Setosa | Versicolor | Virginica | Suma |
|----------|--------|------------|-----------|--------|
| Stopień | 0.0000 | 0.0400 | 0.0800 | 0.0400 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 5

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

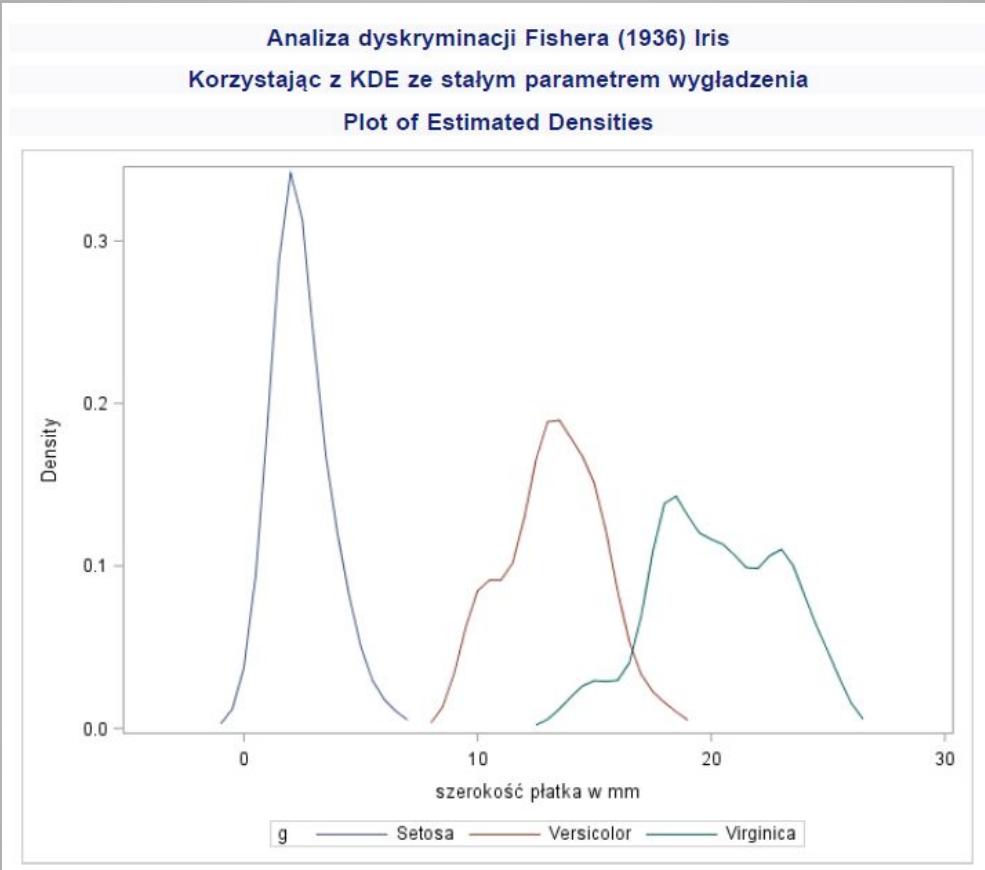
Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTDATA

Podsumowanie klasyfikacji za pomocą gęstości jądra normalnego

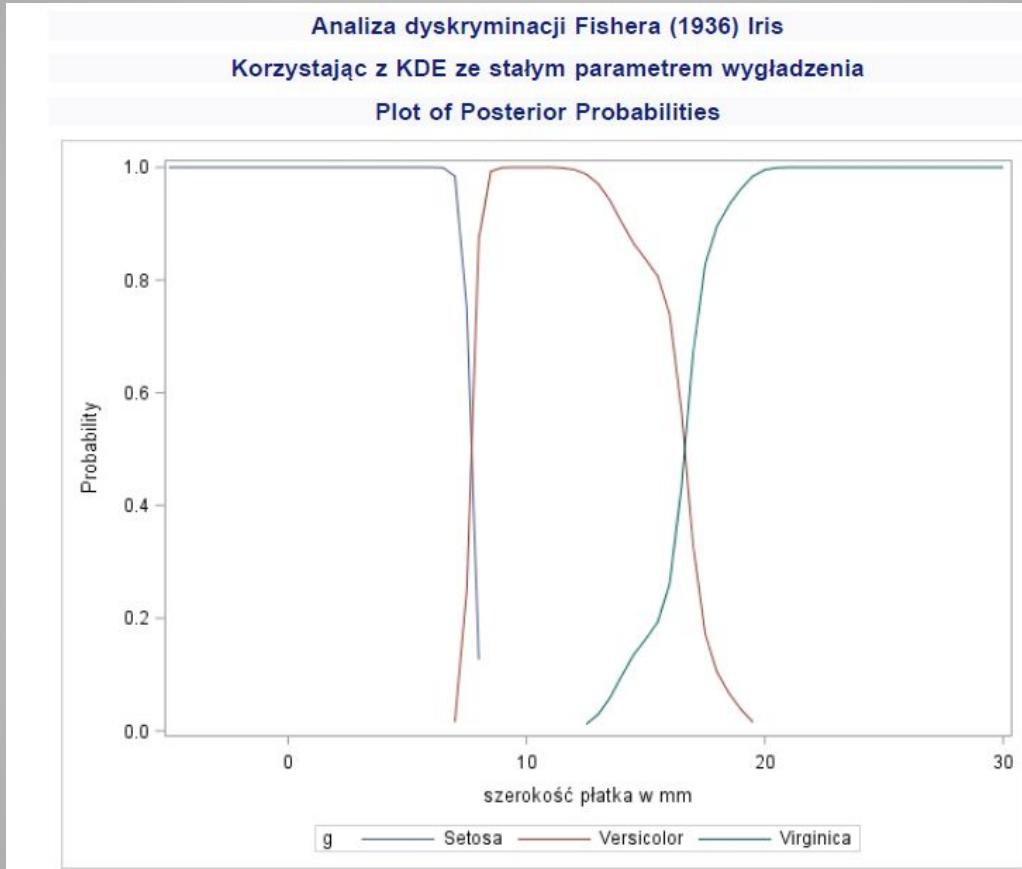
| Profile obserwacji dla danych testowych | |
|---|----|
| Liczba obserwacji wczytanych | 71 |
| Liczba obserwacji użytych | 71 |

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Razem |
| Razem | 26 | 18 | 27 | 71 |
| | 36.62 | 25.35 | 38.03 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

Kod 5



Kod 5



Kod 6

Analiza dyskryminacji Fishera (1936) Iris
Korzystając z KDE bez stałego parametru wygładzenia
Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 1 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

| Informacje o poziomach klasyfikacji | | | | | |
|-------------------------------------|----------------|------------|---------|----------|-----------------------------|
| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 6

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Wyniki walidacji krzyżowej za pomocą gęstości jądra normalnego

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | Virginica | * | | | |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.0475 | 0.9525 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.2310 | 0.7690 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.8805 | 0.1195 |
| 124 | Virginica | Versicolor | * | 0.0000 | 0.9394 | 0.0606 |
| 130 | Virginica | Versicolor | * | 0.0000 | 0.7193 | 0.2807 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.8805 | 0.1195 |

* Obserwacja błędnie zaklasyfikowana

Kod 6

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Podsumowanie walidacji krzyżowej za pomocą gęstości jądra normalnego

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|--|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 4 | 46 | 50 |
| | 0.00 | 8.00 | 92.00 | 100.00 |
| Razem | 50 | 52 | 48 | 150 |
| | 33.33 | 34.67 | 32.00 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0000 | 0.0400 | 0.0800 | 0.0400 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 6

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTDATA

Podsumowanie klasyfikacji za pomocą gęstości jądra normalnego

Profile obserwacji dla danych testowych

| | |
|------------------------------|----|
| Liczba obserwacji wczytanych | 71 |
| Liczba obserwacji użytych | 71 |

Liczba obserwacji i procent zaklasyfikowanych do Species

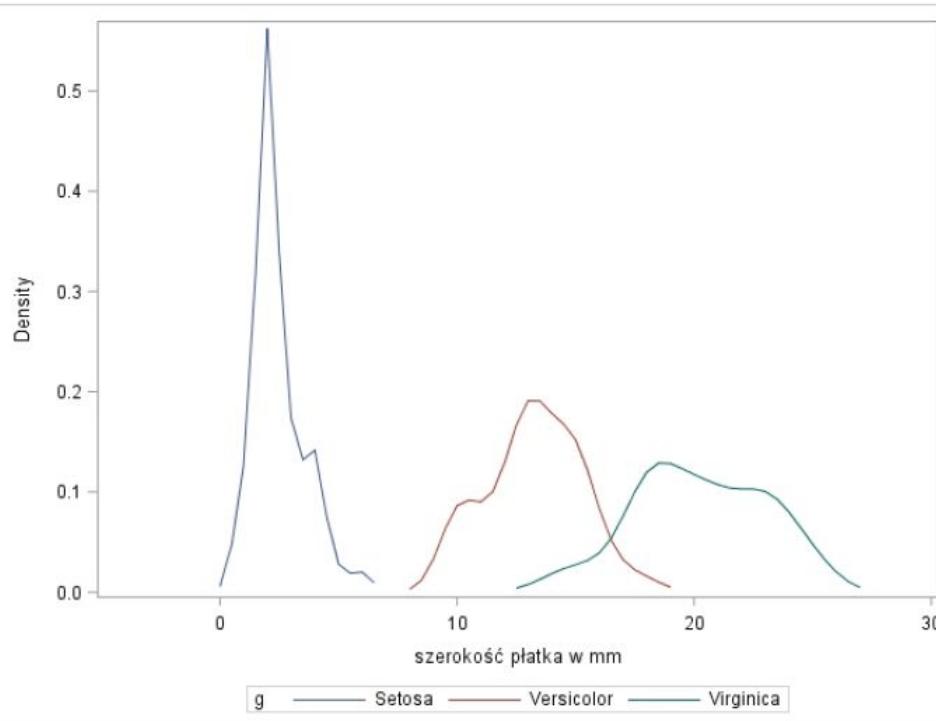
| | Setosa | Versicolor | Virginica | Razem |
|----------|---------|------------|-----------|--------|
| Razem | 25 | 18 | 28 | 71 |
| | 35.21 | 25.35 | 39.44 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

Kod 6

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Plot of Estimated Densities

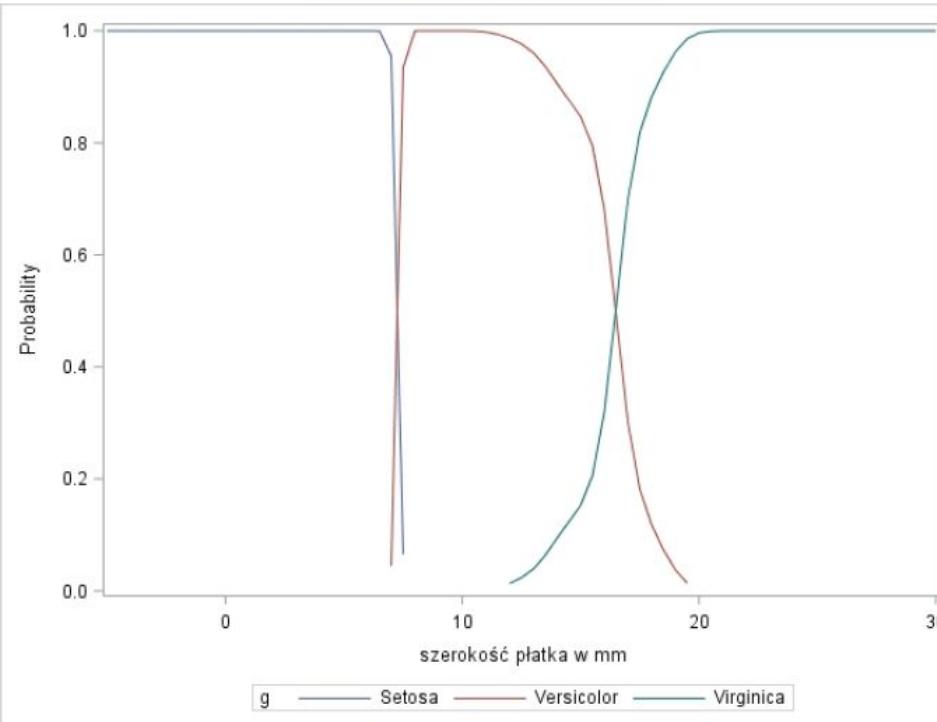


Kod 6

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Plot of Posterior Probabilities

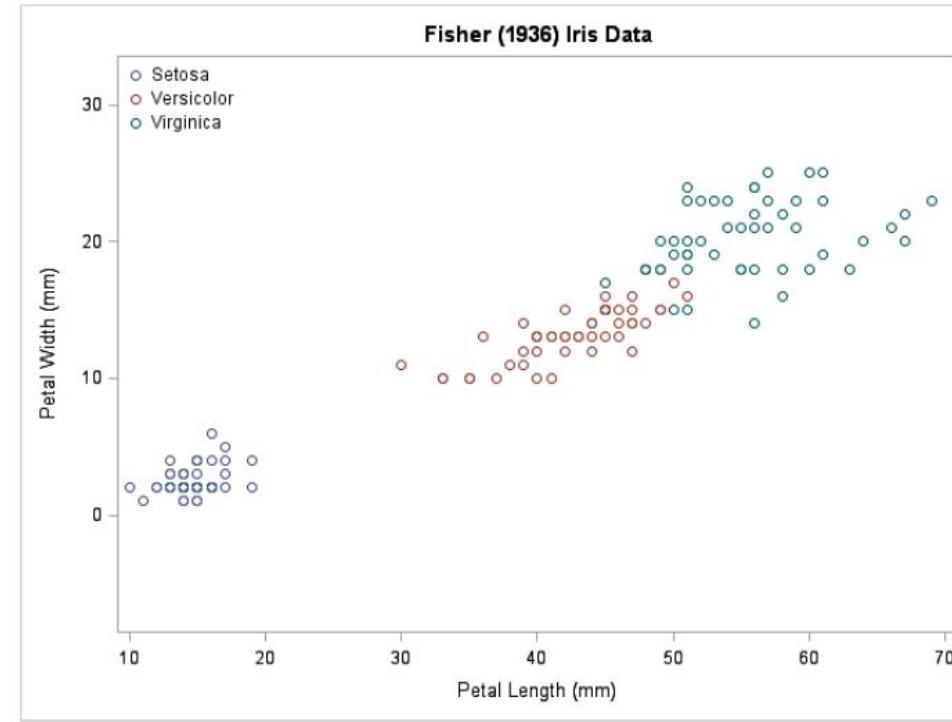


Kod 7

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Plot of Posterior Probabilities



Kod 9

Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego i stałości wariancji

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Calk. liczebność | 150 | DF | 149 |
| Zmienne | 2 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 9

Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego i stałości wariancji

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Wyniki walidacji krzyżowej za pomocą Liniowa funkcja rozpoznawcza

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | Virginica | * | | | |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.2130 | 0.7870 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.3118 | 0.6882 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.8453 | 0.1547 |
| 113 | Virginica | Versicolor | * | 0.0000 | 0.8322 | 0.1678 |
| 124 | Virginica | Versicolor | * | 0.0000 | 0.8057 | 0.1943 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.8903 | 0.1097 |

* Obserwacja błędnie zaklasyfikowana

Kod 9

Analiza dyskryminacji Fishera (1936) Iris korzystając z estymacji rozkładu normalnego i stałości wariancji

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Podsumowanie walidacji krzyżowej za pomocą Liniowa funkcja rozpoznawcza

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 4 | 46 | 50 |
| | 0.00 | 8.00 | 92.00 | 100.00 |
| Razem | 50 | 52 | 48 | 150 |
| | 33.33 | 34.67 | 32.00 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0000 | 0.0400 | 0.0800 | 0.0400 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 9

Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego i stałości wariancji

Procedura DISCRIM

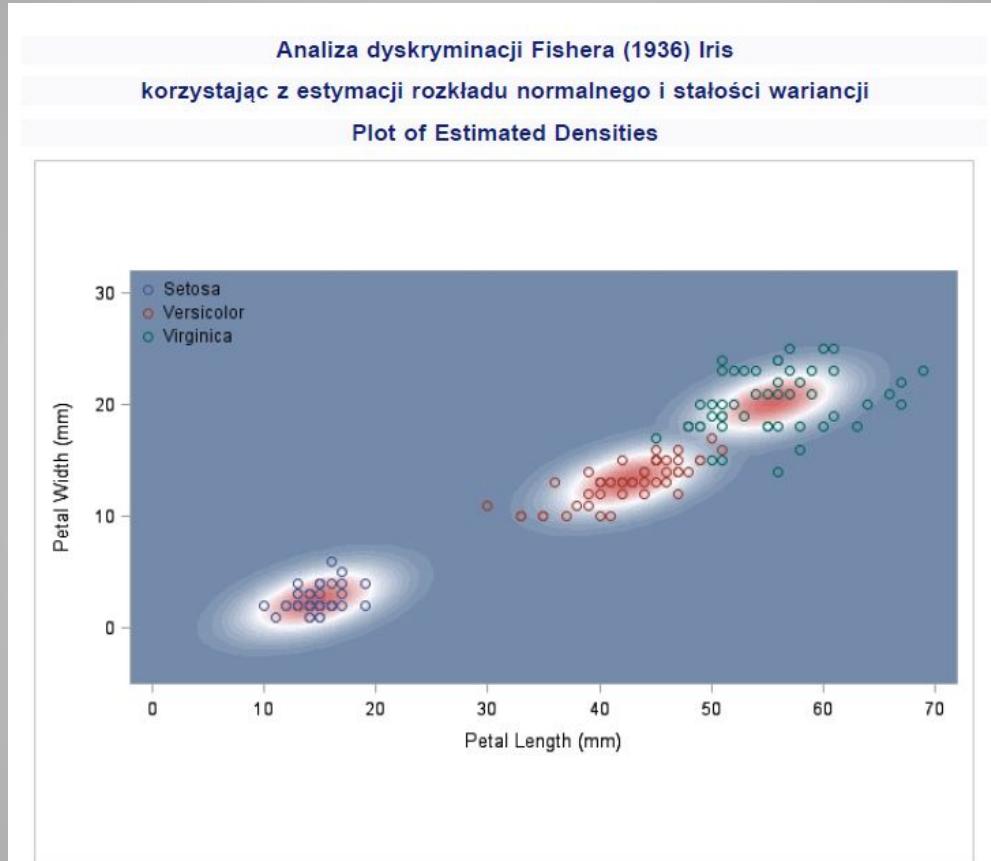
Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTTDATA

Podsumowanie klasyfikacji za pomocą Liniowa funkcja rozpoznawcza

| Profile obserwacji dla danych testowych | |
|---|-------|
| Liczba obserwacji wczytanych | 11175 |
| Liczba obserwacji użytych | 11175 |

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Razem |
| Razem | 3670 | 4243 | 3262 | 11175 |
| | 32.84 | 37.97 | 29.19 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

Kod 9

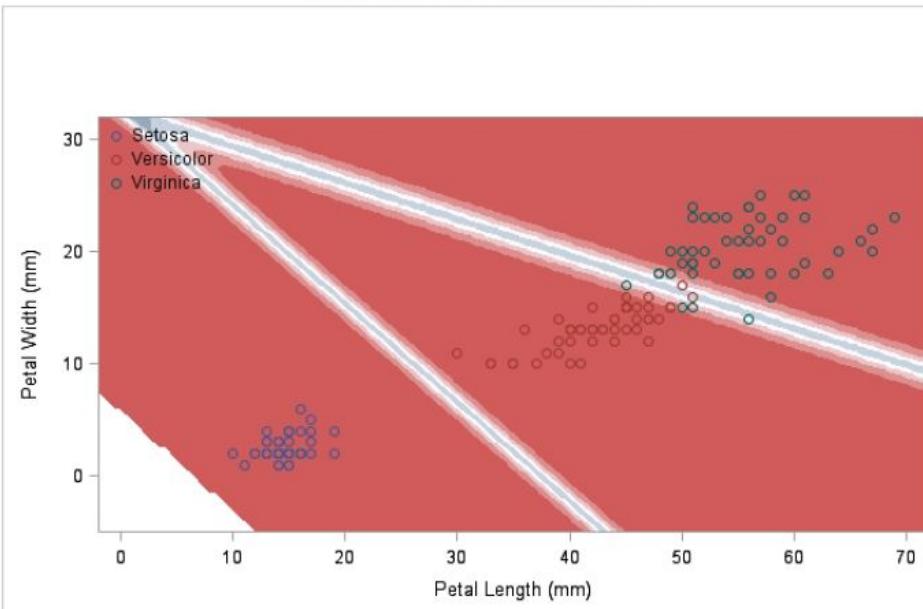


Kod 9

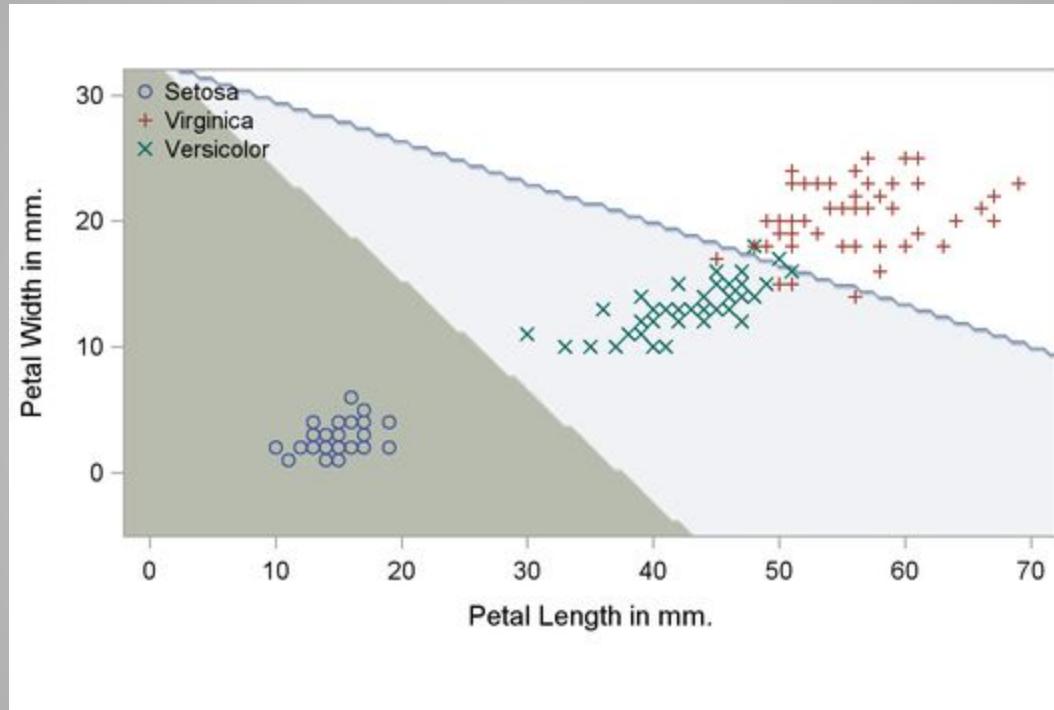
Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego i stałości wariancji

Plot of Posterior Probabilities



Kod 9



Kod 10

Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego bez stałości wariancji

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 2 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 10

Analiza dyskryminacji Fishera (1936) Iris korzystając z estymacji rozkładu normalnego bez stałości wariancji

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Wyniki walidacji krzyżowej za pomocą kwadratowej funkcji rozpoznawczej

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | Virginica | * | | | |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.0903 | 0.9097 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.4675 | 0.5325 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.7288 | 0.2712 |
| 113 | Virginica | Versicolor | * | 0.0000 | 0.5196 | 0.4804 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.8335 | 0.1665 |

* Obserwacja błędnie zaklasyfikowana

Kod 10

Analiza dyskryminacji Fishera (1936) Iris korzystając z estymacji rozkładu normalnego bez stałości wariancji

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Podsumowanie walidacji krzyżowej za pomocą kwadratowej funkcji rozpoznawczej

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 3 | 47 | 50 |
| | 0.00 | 6.00 | 94.00 | 100.00 |
| Razem | 50 | 51 | 49 | 150 |
| | 33.33 | 34.00 | 32.67 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

| Oceny liczby błędów dla Species | | | |
|---------------------------------|--------|------------|-----------|
| | Setosa | Versicolor | Virginica |
| Stopień | 0.0000 | 0.0400 | 0.0600 |
| A priori | 0.3333 | 0.3333 | 0.3333 |

Kod 10

Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego bez stałości wariancji

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTDATA

Podsumowanie klasyfikacji za pomocą kwadratowej funkcji rozpoznawczej

| Profile obserwacji dla danych testowych | |
|---|-------|
| Liczba obserwacji wczytanych | 11175 |
| Liczba obserwacji użytych | 11175 |

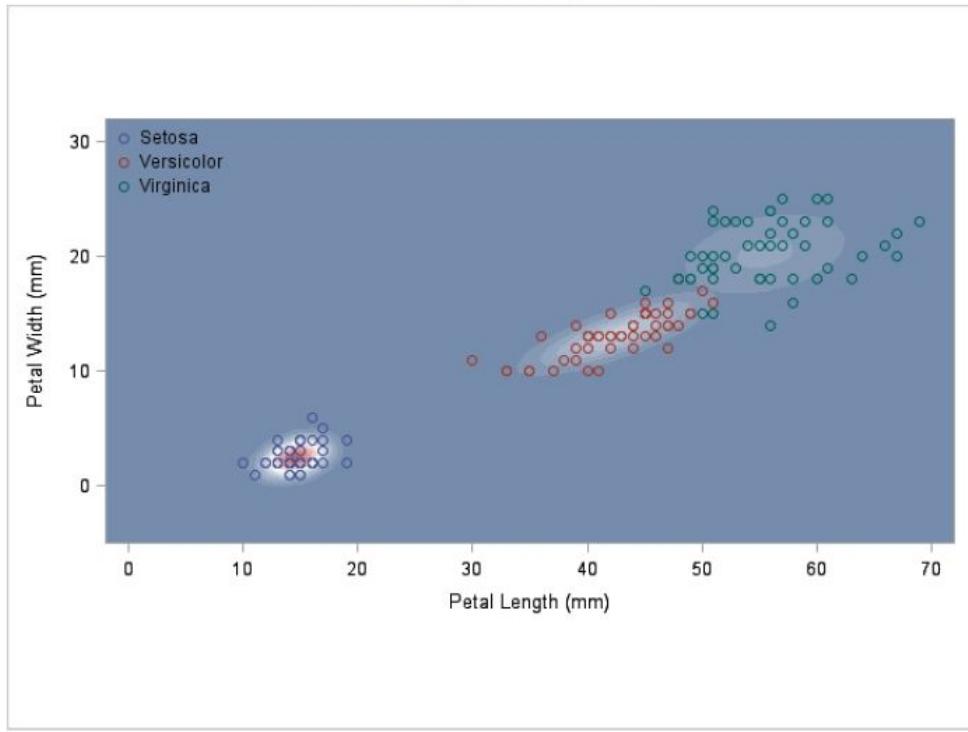
| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|---|---------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Razem |
| Razem | 1382 | 1345 | 8448 | 11175 |
| | 12.37 | 12.04 | 75.60 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

Kod 10

Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego bez stałości wariancji

Plot of Estimated Densities

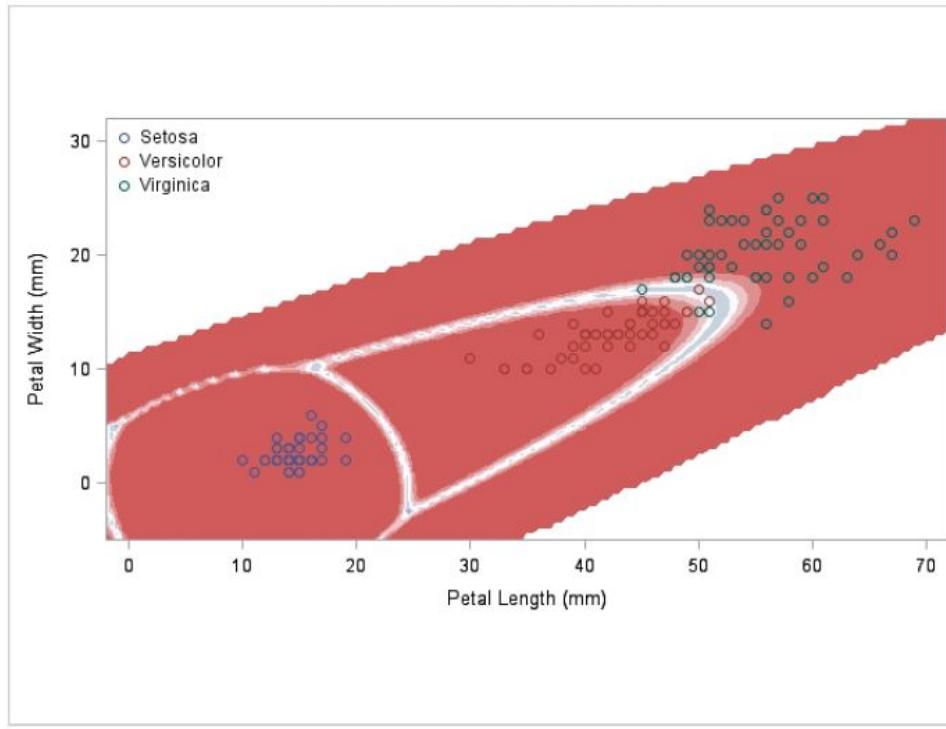


Kod 10

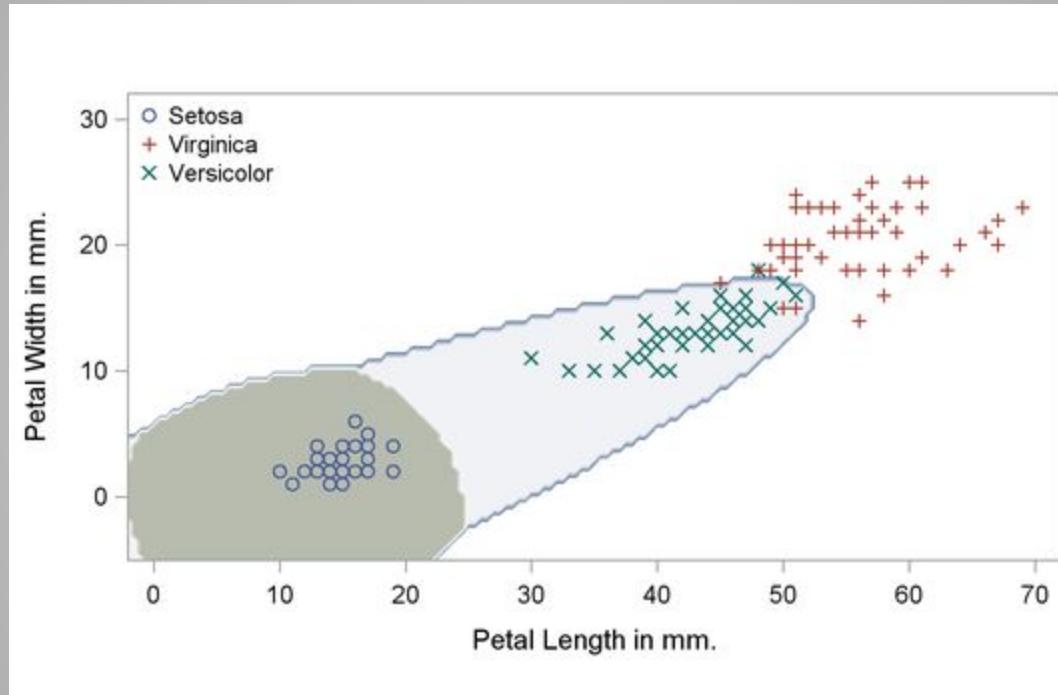
Analiza dyskryminacji Fishera (1936) Iris

korzystając z estymacji rozkładu normalnego bez stałości wariancji

Plot of Posterior Probabilities



Kod 10



Kod 11

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 2 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 11

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Wyniki walidacji krzyżowej za pomocą gęstości jądra normalnego

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | Virginica | * | | | |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.0800 | 0.9200 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.4123 | 0.5877 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.7474 | 0.2526 |
| 113 | Virginica | Versicolor | * | 0.0000 | 0.5863 | 0.4137 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.8358 | 0.1642 |

* Obserwacja błędnie zaklasyfikowana

Kod 11

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Podsumowanie walidacji krzyżowej za pomocą gęstości jądra normalnego

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|--|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 3 | 47 | 50 |
| | 0.00 | 6.00 | 94.00 | 100.00 |
| Razem | 50 | 51 | 49 | 150 |
| | 33.33 | 34.00 | 32.67 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0000 | 0.0400 | 0.0600 | 0.0333 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 11

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTTDATA

Podsumowanie klasyfikacji za pomocą gęstości jądra normalnego

Profile obserwacji dla danych testowych

| | |
|------------------------------|-------|
| Liczba obserwacji wczytanych | 11175 |
| Liczba obserwacji użytych | 11175 |

Liczba obserwacji i procent zaklasyfikowanych do Species

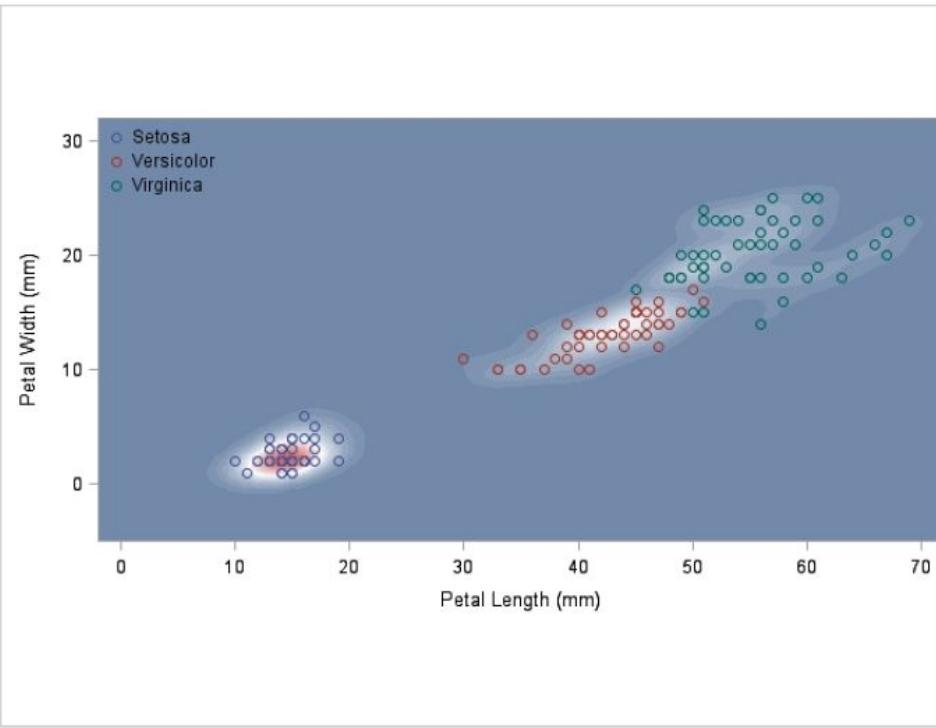
| | Setosa | Versicolor | Virginica | Razem |
|----------|---------|------------|-----------|--------|
| Razem | 3195 | 2492 | 5488 | 11175 |
| | 28.59 | 22.30 | 49.11 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

Kod 11

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Plot of Estimated Densities

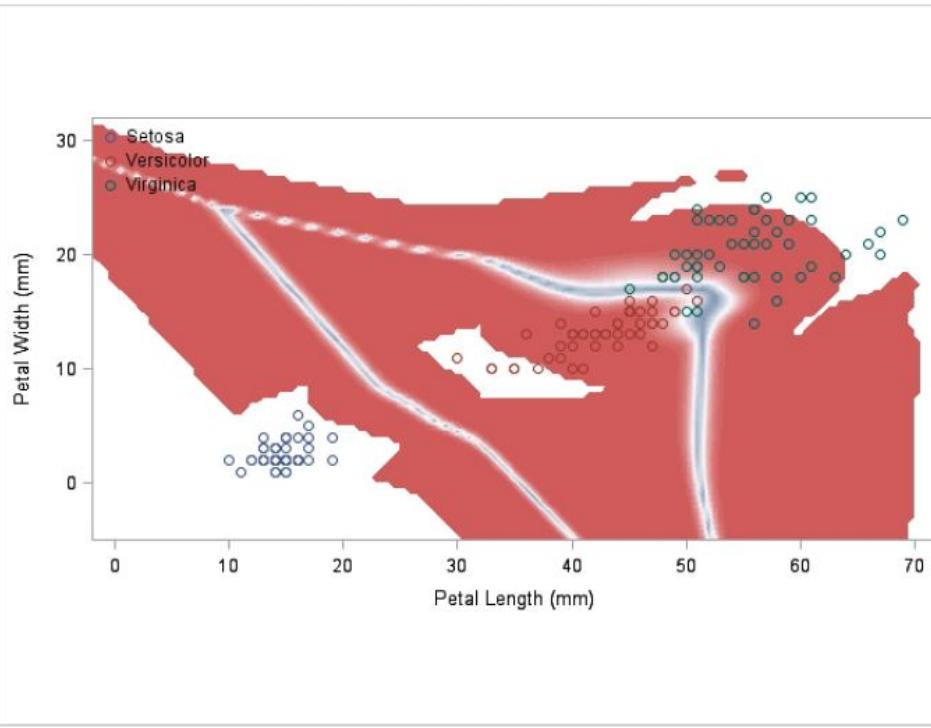


Kod 11

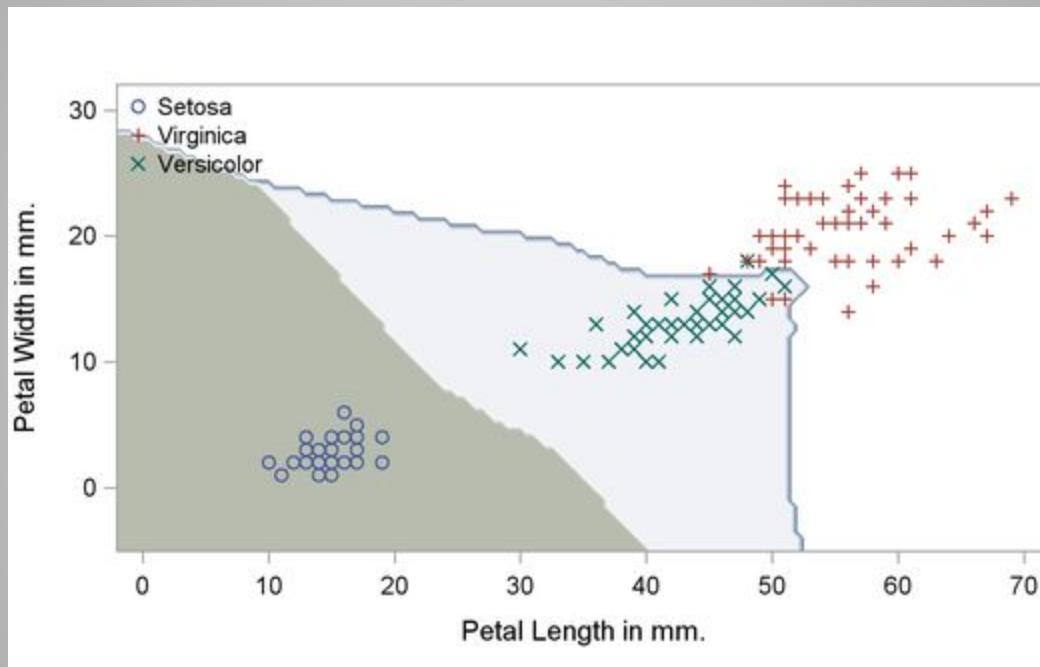
Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE ze stałym parametrem wygładzenia

Plot of Posterior Probabilities



Kod 11



Kod 12

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 2 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 12

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Wyniki walidacji krzyżowej za pomocą gęstości jądra normalnego

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | | | | | |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.0516 | 0.9484 |
| 100 | Versicolor | Virginica | * | 0.0000 | 0.3773 | 0.6227 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.7826 | 0.2174 |
| 136 | Virginica | Versicolor | * | 0.0000 | 0.8802 | 0.1198 |

* Obserwacja błędnie zaklasyfikowana

Kod 12

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Podsumowanie walidacji krzyżowej za pomocą gęstości jądra normalnego

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|--|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 2 | 48 | 50 |
| | 0.00 | 4.00 | 96.00 | 100.00 |
| Razem | 50 | 50 | 50 | 150 |
| | 33.33 | 33.33 | 33.33 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0000 | 0.0400 | 0.0400 | 0.0267 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 12

Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych testowych: WORK.PLOTTDATA

Podsumowanie klasyfikacji za pomocą gęstości jądra normalnego

Profile obserwacji dla danych testowych

| | |
|------------------------------|-------|
| Liczba obserwacji wczytanych | 11175 |
| Liczba obserwacji użytych | 11175 |

Liczba obserwacji i procent zaklasyfikowanych do Species

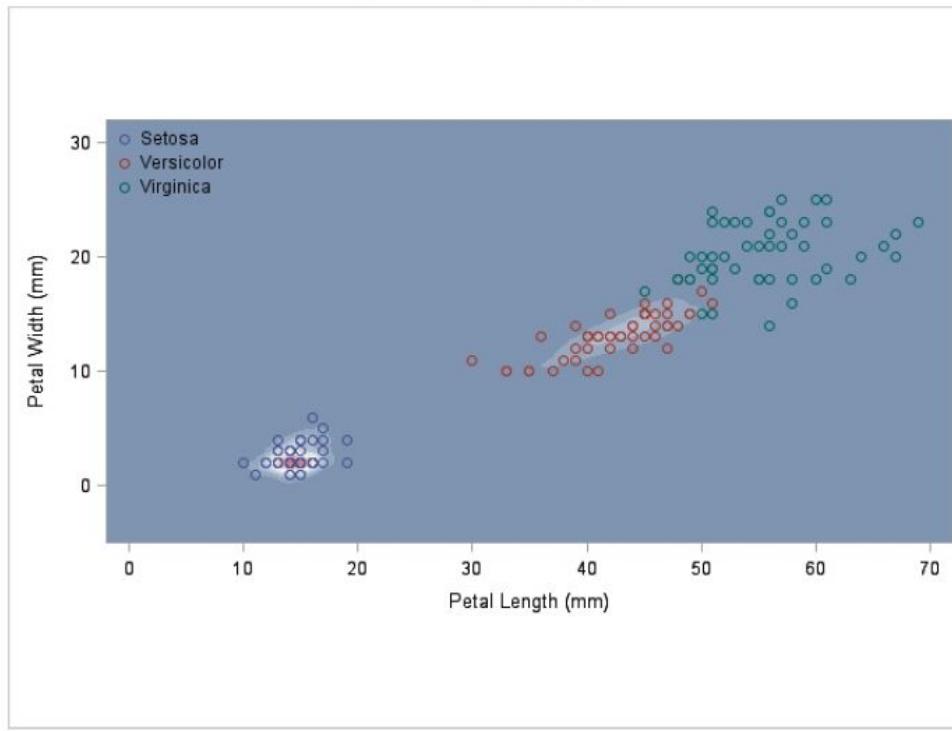
| | Setosa | Versicolor | Virginica | Razem |
|----------|---------|------------|-----------|--------|
| Razem | 1370 | 1505 | 8300 | 11175 |
| | 12.26 | 13.47 | 74.27 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

Kod 12

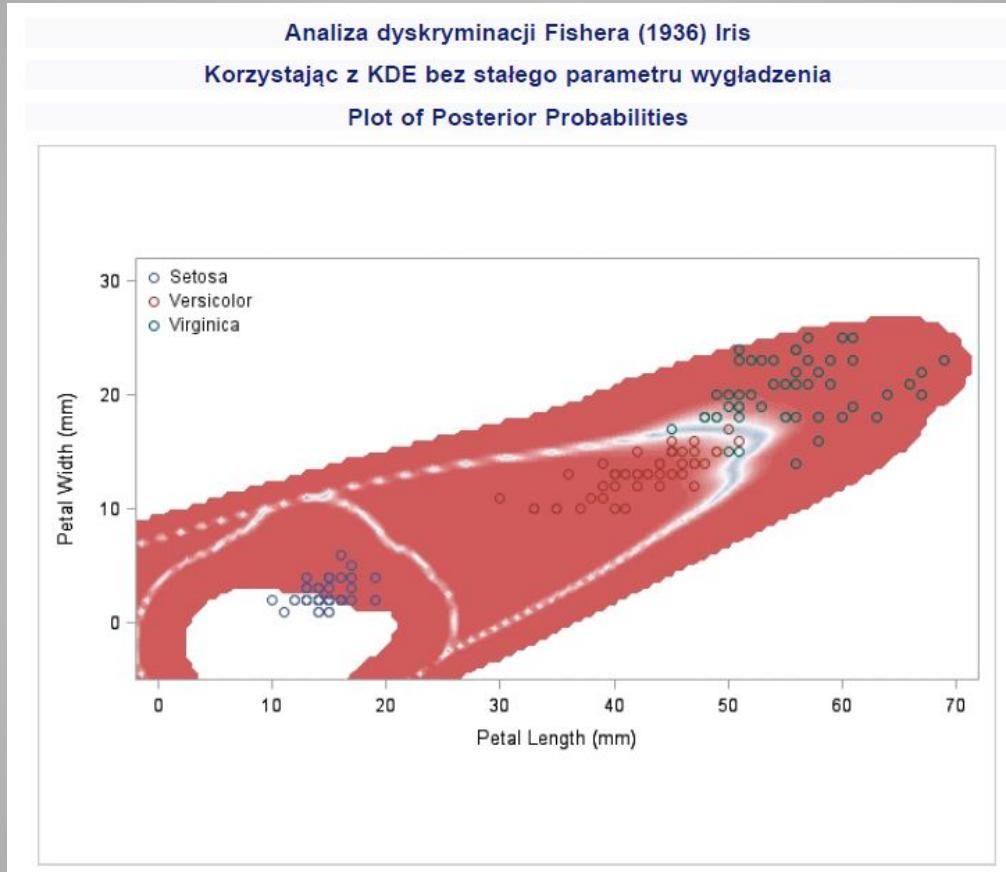
Analiza dyskryminacji Fishera (1936) Iris

Korzystając z KDE bez stałego parametru wygładzenia

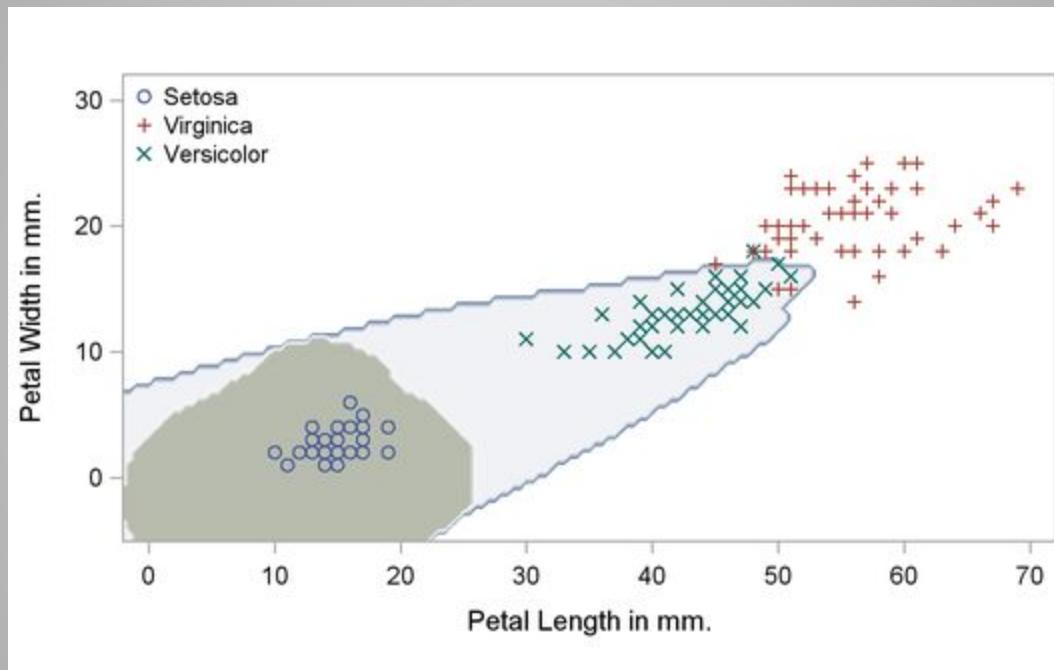
Plot of Estimated Densities



Kod 12



Kod 12



Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

| | | | |
|------------------|-----|--------------|-----|
| Całk. liczebność | 150 | DF | 149 |
| Zmienne | 4 | DF w klasach | 147 |
| Klasy | 3 | DF klas | 2 |

| | |
|------------------------------|-----|
| Liczba obserwacji wczytanych | 150 |
| Liczba obserwacji użytych | 150 |

Informacje o poziomach klasyfikacji

| Species | Nazwa zmiennej | Liczebność | Waga | Udział | Prawdopodobieństwo a priori |
|------------|----------------|------------|---------|----------|-----------------------------|
| Setosa | Setosa | 50 | 50.0000 | 0.333333 | 0.333333 |
| Versicolor | Versicolor | 50 | 50.0000 | 0.333333 | 0.333333 |
| Virginica | Virginica | 50 | 50.0000 | 0.333333 | 0.333333 |

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

Macierze kowariancji wewnętrz klas

| Species = Setosa, DF = 49 | | | | | | |
|---------------------------|-------------------|-------------|-------------|-------------|------------|--|
| Zmienna | Etykieta | SepalLength | SepalWidth | PetalLength | PetalWidth | |
| SepalLength | Sepal Length (mm) | 12.42489796 | 9.92163265 | 1.63551020 | 1.03306122 | |
| SepalWidth | Sepal Width (mm) | 9.92163265 | 14.36897959 | 1.16979592 | 0.92979592 | |
| PetalLength | Petal Length (mm) | 1.63551020 | 1.16979592 | 3.01591837 | 0.60693878 | |
| PetalWidth | Petal Width (mm) | 1.03306122 | 0.92979592 | 0.60693878 | 1.11061224 | |

| Species = Versicolor, DF = 49 | | | | | | |
|-------------------------------|-------------------|-------------|------------|-------------|------------|--|
| Zmienna | Etykieta | SepalLength | SepalWidth | PetalLength | PetalWidth | |
| SepalLength | Sepal Length (mm) | 26.64326531 | 8.51836735 | 18.28979592 | 5.57795918 | |
| SepalWidth | Sepal Width (mm) | 8.51836735 | 9.84693878 | 8.26530612 | 4.12040816 | |
| PetalLength | Petal Length (mm) | 18.28979592 | 8.26530612 | 22.08163265 | 7.31020408 | |
| PetalWidth | Petal Width (mm) | 5.57795918 | 4.12040816 | 7.31020408 | 3.91061224 | |

| Species = Virginica, DF = 49 | | | | | | |
|------------------------------|-------------------|-------------|-------------|-------------|------------|--|
| Zmienna | Etykieta | SepalLength | SepalWidth | PetalLength | PetalWidth | |
| SepalLength | Sepal Length (mm) | 40.43428571 | 9.37632653 | 30.32897959 | 4.90938776 | |
| SepalWidth | Sepal Width (mm) | 9.37632653 | 10.40040816 | 7.13795918 | 4.76285714 | |
| PetalLength | Petal Length (mm) | 30.32897959 | 7.13795918 | 30.45877551 | 4.88244898 | |
| PetalWidth | Petal Width (mm) | 4.90938776 | 4.76285714 | 4.88244898 | 7.54326531 | |

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

| Połączona międzyklasowa macierz kowariancji, DF = 147 | | | | | | |
|---|-------------------|-------------|------------|-------------|------------|------------|
| Zmienna | Etykieta | SepalLength | SepalWidth | PetalLength | PetalWidth | |
| SepalLength | Sepal Length (mm) | 26.50081633 | 9.27210884 | 16.75142857 | 3.84013605 | |
| SepalWidth | Sepal Width (mm) | | 9.27210884 | 11.53877551 | 5.52435374 | 3.27102041 |
| PetalLength | Petal Length (mm) | 16.75142857 | 5.52435374 | 18.51877551 | 4.26653061 | |
| PetalWidth | Petal Width (mm) | 3.84013605 | 3.27102041 | 4.26653061 | 4.18816327 | |

| Informacje o macierzach kowariancji wewnętrzgrupowych | | |
|---|---------------------------|---|
| Species | Rząd macierzy kowariancji | Logarytm naturalny wyznacznika macierzy kowariancji |
| Setosa | 4 | 5.35332 |
| Versicolor | 4 | 7.54636 |
| Virginica | 4 | 9.49362 |
| Polączone | 4 | 8.46214 |

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

Test homogeniczności macierzy kowariancji wewnętrzgrupowych

| Chi-kwadrat | DF | Pr. > chi-kw. |
|-------------|----|---------------|
| 140.943050 | 20 | <.0001 |

Ponieważ wartość Chi-kwadrat jest istotna na poziomie 0.1, w funkcji dyskryminacyjnej zostaną użyte macierze kowariancji wewnętrzgrupowych.

Reference: Morrison, D.F. (1976) Multivariate Statistical Methods p252.

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

| Odległość kwadratowa od Species | | | |
|---------------------------------|-----------|------------|-----------|
| Od Species | Setosa | Versicolor | Virginica |
| Setosa | 0 | 103.19382 | 168.76759 |
| Versicolor | 323.06203 | 0 | 13.83875 |
| Virginica | 706.08494 | 17.86670 | 0 |

| Uogólniona odległość kwadratowa od Species | | | |
|--|-----------|------------|-----------|
| Od Species | Setosa | Versicolor | Virginica |
| Setosa | 5.35332 | 110.74017 | 178.26121 |
| Versicolor | 328.41535 | 7.54636 | 23.33238 |
| Virginica | 711.43826 | 25.41306 | 9.49362 |

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

| Statystyki testu jednowymiarowego | | | | | | | | |
|------------------------------------|-------------------|-------------------------------------|-------------------------------------|-------------------------------------|-----------|-----------------------|-----------|-------|
| Statystyki F, Num DF=2, Den DF=147 | | | | | | | | |
| Zmienna | Etykieta | Calkowite odchylenie standaryzowane | Połączone odchylenie standaryzowane | Pośrednie odchylenie standaryzowane | R-kwadrat | R-kwadrat / (1-R-kw.) | Wartość F | Pr. > |
| SepalLength | Sepal Length (mm) | 8.2807 | 5.1479 | 7.9506 | 0.6187 | 1.6226 | 119.26 | <.000 |
| SepalWidth | Sepal Width (mm) | 4.3587 | 3.3969 | 3.3682 | 0.4008 | 0.6688 | 49.16 | <.000 |
| PetalLength | Petal Length (mm) | 17.6530 | 4.3033 | 20.9070 | 0.9414 | 16.0566 | 1180.16 | <.000 |
| PetalWidth | Petal Width (mm) | 7.6224 | 2.0465 | 8.9673 | 0.9289 | 13.0613 | 960.01 | <.000 |

| Przeciętne R-kwadrat | |
|----------------------|----------|
| Nieważone | 0.722435 |
| Ważone wg wariancji | 0.868944 |

| Statystyki wielowymiarowe i przybliżenia F | | | | | | |
|--|-------------|-----------|-----------|----------|--------|--|
| | S=2 | M=0.5 | N=71 | | | |
| Statystyka | Wartość | Wartość F | DF liczn. | DF mian. | Pr. > | |
| Lambda Wilksa | 0.02343863 | 199.15 | 8 | 288 | <0.001 | |
| Ślad Pillai'a | 1.19189883 | 53.47 | 8 | 290 | <0.001 | |
| Ślad Hotellinga-Lawleya | 32.47732024 | 582.20 | 8 | 203.4 | <0.001 | |
| Najw. pierwiastek Roya | 32.1919290 | 1166.96 | 4 | 145 | <0.001 | |

UWAGA: Statystyka F dla największego pierwiastka Roya jest górnym kresem

UWAGA: Statystyka F dla lambdy Wilksa jest dokładna.

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacyjnej

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Wyniki resubstytucji za pomocą kwadratowej funkcji rozpoznawczej

Prawdopodobieństwo a posteriori znalezienia się w Species

| Przypisanie do gatunku na podstawie klasyfikacji | | | | | | |
|--|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.3359 | 0.6641 |
| 55 | Versicolor | Virginica | * | 0.0000 | 0.1543 | 0.8457 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.6050 | 0.3950 |

* Obserwacja błędnie zaklasyfikowana

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Podsumowanie resubstytucji za pomocą kwadratowej funkcji rozpoznawczej

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|--|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 48 | 2 | 50 |
| | 0.00 | 96.00 | 4.00 | 100.00 |
| Virginica | 0 | 1 | 49 | 50 |
| | 0.00 | 2.00 | 98.00 | 100.00 |
| Razem | 50 | 49 | 51 | 150 |
| | 33.33 | 32.67 | 34.00 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0000 | 0.0400 | 0.0200 | 0.0200 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

Wyniki klasyfikacji dla danych kalibracyjnych: WORK.IRIS

Wyniki walidacji krzyżowej za pomocą kwadratowej funkcji rozpoznawczej

| Prawdopodobieństwo a posteriori znalezienia się w Species | | | | | | |
|---|------------|----------------------------|---|--------|------------|-----------|
| Obs. | Od Species | Zaklasyfikowane do Species | | Setosa | Versicolor | Virginica |
| | | | * | | | |
| 52 | Versicolor | Virginica | * | 0.0000 | 0.3134 | 0.6866 |
| 53 | Versicolor | Virginica | * | 0.0000 | 0.1616 | 0.8384 |
| 55 | Versicolor | Virginica | * | 0.0000 | 0.0713 | 0.9287 |
| 103 | Virginica | Versicolor | * | 0.0000 | 0.6632 | 0.3368 |

* Obserwacja błędnie zaklasyfikowana

Kod 13

Analiza dyskryminacyjna irysów

Korzystając z kwadratowej funkcji dyskryminacji

Procedura DISCRIM

Podsumowanie klasyfikacji dla danych kalibracyjnych: WORK.IRIS
Podsumowanie walidacji krzyżowej za pomocą kwadratowej funkcji rozpoznawczej

| Liczba obserwacji i procent zaklasyfikowanych do Species | | | | |
|--|---------|------------|-----------|--------|
| Od Species | Setosa | Versicolor | Virginica | Razem |
| Setosa | 50 | 0 | 0 | 50 |
| | 100.00 | 0.00 | 0.00 | 100.00 |
| Versicolor | 0 | 47 | 3 | 50 |
| | 0.00 | 94.00 | 6.00 | 100.00 |
| Virginica | 0 | 1 | 49 | 50 |
| | 0.00 | 2.00 | 98.00 | 100.00 |
| Razem | 50 | 48 | 52 | 150 |
| | 33.33 | 32.00 | 34.67 | 100.00 |
| A priori | 0.33333 | 0.33333 | 0.33333 | |
| | | | | |

| Oceny liczby błędów dla Species | | | | |
|---------------------------------|--------|------------|-----------|--------|
| | Setosa | Versicolor | Virginica | Suma |
| Stopień | 0.0000 | 0.0600 | 0.0200 | 0.0267 |
| A priori | 0.3333 | 0.3333 | 0.3333 | |

Kod 13

Analiza dyskryminacyjna irysów

Wyniki analizy dyskryminacji

| Obs. | Species | TYPE | NAME | SepalLength | SepalWidth | PetalLength | PetalWidth |
|------|------------|-------------|-------------|-------------|------------|-------------|------------|
| 1 | N | | | 150.00 | 150.00 | 150.00 | 150.00 |
| 2 | Setosa | N | | 50.00 | 50.00 | 50.00 | 50.00 |
| 3 | Versicolor | N | | 50.00 | 50.00 | 50.00 | 50.00 |
| 4 | Virginica | N | | 50.00 | 50.00 | 50.00 | 50.00 |
| 5 | MEAN | | | 58.43 | 30.57 | 37.58 | 11.99 |
| 6 | Setosa | MEAN | | 50.06 | 34.28 | 14.82 | 2.46 |
| 7 | Versicolor | MEAN | | 59.38 | 27.70 | 42.80 | 13.26 |
| 8 | Virginica | MEAN | | 65.88 | 29.74 | 55.52 | 20.26 |
| 9 | Setosa | PRIOR | | 0.33 | 0.33 | 0.33 | 0.33 |
| 10 | Versicolor | PRIOR | | 0.33 | 0.33 | 0.33 | 0.33 |
| 11 | Virginia | PRIOR | | 0.33 | 0.33 | 0.33 | 0.33 |
| 12 | Setosa | CSSCP | SepalLength | 608.82 | 486.16 | 80.14 | 50.62 |
| 13 | Setosa | CSSCP | SepalWidth | 498.16 | 704.08 | 57.32 | 45.56 |
| 14 | Setosa | CSSCP | PetalLength | 80.14 | 57.32 | 147.78 | 29.74 |
| 15 | Setosa | CSSCP | PetalWidth | 50.62 | 45.56 | 29.74 | 54.42 |
| 16 | Versicolor | CSSCP | SepalLength | 1305.52 | 417.40 | 896.20 | 273.32 |
| 17 | Versicolor | CSSCP | SepalWidth | 417.40 | 482.50 | 405.00 | 201.90 |
| 18 | Versicolor | CSSCP | PetalLength | 896.20 | 405.00 | 1082.00 | 358.20 |
| 19 | Versicolor | CSSCP | PetalWidth | 273.32 | 201.90 | 368.20 | 191.62 |
| 20 | Virginia | CSSCP | SepalLength | 1981.28 | 459.44 | 1486.12 | 240.56 |
| 21 | Virginia | CSSCP | SepalWidth | 459.44 | 509.62 | 349.78 | 233.38 |
| 22 | Virginia | CSSCP | PetalLength | 1486.12 | 349.78 | 1492.48 | 239.24 |
| 23 | Virginia | CSSCP | PetalWidth | 240.56 | 233.38 | 239.24 | 389.62 |
| 24 | PSSCP | SepalLength | 3895.62 | 1363.00 | 2462.46 | 564.50 | |
| 25 | PSSCP | SepalWidth | 1363.00 | 1664.20 | 812.08 | 480.84 | |
| 26 | PSSCP | PetalLength | 2462.46 | 812.08 | 2722.26 | 627.18 | |
| 27 | PSSCP | PetalWidth | 564.50 | 480.84 | 627.18 | 615.66 | |
| 28 | BSSCP | SepalLength | 6321.21 | -1995.27 | 16524.84 | 7127.93 | |
| 29 | BSSCP | SepalWidth | -1995.27 | 1134.49 | -5723.96 | -2293.27 | |
| 30 | BSSCP | PetalLength | 16524.84 | -5723.96 | 43710.28 | 18677.40 | |
| 31 | BSSCP | PetalWidth | 7127.93 | -2293.27 | 18677.40 | 8041.33 | |
| 32 | CSSCP | SepalLength | 10216.83 | -632.27 | 18987.30 | 7692.43 | |
| 33 | CSSCP | SepalWidth | -632.27 | 2830.69 | -4911.88 | -1812.43 | |
| 34 | CSSCP | PetalLength | 18987.30 | -4911.88 | 46432.54 | 19304.58 | |
| 35 | CSSCP | PetalWidth | 7692.43 | -1812.43 | 19304.58 | 8856.99 | |
| 36 | RSQUARED | | | 0.62 | 0.40 | 0.04 | 0.93 |
| 37 | Setosa | COV | SepalLength | 12.42 | 9.92 | 1.84 | 1.03 |
| 38 | Setosa | COV | SepalWidth | 9.92 | 14.37 | 1.17 | 0.93 |
| 39 | Setosa | COV | PetalLength | 1.64 | 1.17 | 3.02 | 0.61 |
| 40 | Setosa | COV | PetalWidth | 1.03 | 0.93 | 0.61 | 1.11 |
| 41 | Versicolor | COV | SepalLength | 26.64 | 8.82 | 18.29 | 5.58 |
| 42 | Versicolor | COV | SepalWidth | 8.52 | 9.85 | 8.27 | 4.12 |
| 43 | Versicolor | COV | PetalLength | 18.29 | 8.27 | 22.08 | 7.31 |
| 44 | Versicolor | COV | PetalWidth | 5.58 | 4.12 | 7.31 | 3.91 |
| 45 | Virginica | COV | SepalLength | 40.43 | 9.38 | 30.33 | 4.91 |
| 46 | Virginica | COV | SepalWidth | 9.38 | 10.40 | 7.14 | 4.76 |
| 47 | Virginica | COV | PetalLength | 30.33 | 7.14 | 30.46 | 4.88 |
| 48 | Virginica | COV | PetalWidth | 4.91 | 4.76 | 4.88 | 7.54 |
| 49 | PCOV | SepalLength | 26.50 | 9.27 | 16.75 | 3.84 | |

Analiza dyskryminacyjna irysów

Wyniki analizy dyskryminacji

| Obs. | Species | TYPE | NAME | SepalLength | SepalWidth | PetalLength | PetalWidth |
|------|------------|----------|-------------|-------------|------------|-------------|------------|
| 50 | | PCOV | SepalWidth | 9.27 | 11.54 | 5.52 | 3.27 |
| 51 | | PCOV | PetalLength | 16.75 | 5.52 | 18.52 | 4.27 |
| 52 | | PCOV | PetalWidth | 3.84 | 3.27 | 4.27 | 4.19 |
| 53 | | BCOV | SepalLength | 63.21 | -19.95 | 165.25 | 71.28 |
| 54 | | BCOV | SepalWidth | -19.95 | 11.34 | -57.24 | -22.93 |
| 55 | | BCOV | PetalLength | 165.25 | -57.24 | 437.10 | 188.77 |
| 56 | | BCOV | PetalWidth | 71.28 | -22.93 | 188.77 | 80.41 |
| 57 | | COV | SepalLength | 68.57 | -4.24 | 127.43 | 51.63 |
| 58 | | COV | SepalWidth | -4.24 | 19.00 | -32.97 | -12.16 |
| 59 | | COV | PetalLength | 127.43 | -32.97 | 311.63 | 129.56 |
| 60 | | COV | PetalWidth | 51.63 | -12.16 | 129.56 | 58.10 |
| 61 | Setosa | STD | | 3.52 | 3.79 | 1.74 | 1.05 |
| 62 | Versicolor | STD | | 5.16 | 3.14 | 4.70 | 1.98 |
| 63 | Virginia | STD | | 6.38 | 3.22 | 5.52 | 2.75 |
| 64 | | PSTD | | 5.15 | 3.40 | 4.30 | 2.05 |
| 65 | | BSTD | | 7.95 | 3.37 | 20.91 | 8.97 |
| 66 | | STD | | 8.28 | 4.36 | 17.65 | 7.62 |
| 67 | Setosa | CORR | SepalLength | 1.00 | 0.74 | 0.27 | 0.28 |
| 68 | Setosa | CORR | SepalWidth | 0.74 | 1.00 | 0.18 | 0.23 |
| 69 | Setosa | CORR | PetalLength | 0.27 | 0.18 | 1.00 | 0.33 |
| 70 | Setosa | CORR | PetalWidth | 0.28 | 0.23 | 0.33 | 1.00 |
| 71 | Versicolor | CORR | SepalLength | 1.00 | 0.53 | 0.75 | 0.55 |
| 72 | Versicolor | CORR | SepalWidth | 0.53 | 1.00 | 0.56 | 0.66 |
| 73 | Versicolor | CORR | PetalLength | 0.75 | 0.56 | 1.00 | 0.79 |
| 74 | Versicolor | CORR | PetalWidth | 0.55 | 0.66 | 0.79 | 1.00 |
| 75 | Virginia | CORR | SepalLength | 1.00 | 0.46 | 0.86 | 0.28 |
| 76 | Virginia | CORR | SepalWidth | 0.46 | 1.00 | 0.40 | 0.54 |
| 77 | Virginia | CORR | PetalLength | 0.98 | 0.40 | 1.00 | 0.32 |
| 78 | Virginia | CORR | PetalWidth | 0.28 | 0.54 | 0.32 | 1.00 |
| 79 | | PCORR | SepalLength | 1.00 | 0.53 | 0.76 | 0.36 |
| 80 | | PCORR | SepalWidth | 0.53 | 1.00 | 0.38 | 0.47 |
| 81 | | PCORR | PetalLength | 0.76 | 0.38 | 1.00 | 0.48 |
| 82 | | PCORR | PetalWidth | 0.36 | 0.47 | 0.48 | 1.00 |
| 83 | | BCORR | SepalLength | 1.00 | -0.75 | 0.99 | 1.00 |
| 84 | | BCORR | SepalWidth | -0.75 | 1.00 | -0.81 | -0.76 |
| 85 | | BCORR | PetalLength | 0.99 | -0.81 | 1.00 | 1.00 |
| 86 | | BCORR | PetalWidth | 1.00 | -0.76 | 1.00 | 1.00 |
| 87 | | CORR | SepalLength | 1.00 | -0.12 | 0.87 | 0.82 |
| 88 | | CORR | SepalWidth | -0.12 | 1.00 | -0.43 | -0.37 |
| 89 | | CORR | PetalLength | 0.87 | -0.43 | 1.00 | 0.96 |
| 90 | | CORR | PetalWidth | 0.82 | -0.37 | 0.98 | 1.00 |
| 91 | Setosa | STDMEAN | | -1.01 | 0.85 | -1.30 | -1.25 |
| 92 | Versicolor | STDMEAN | | 0.11 | -0.66 | 0.28 | 0.17 |
| 93 | Virginia | STDMEAN | | 0.90 | -0.19 | 1.02 | 1.08 |
| 94 | Setosa | PSTDMEAN | | -1.63 | 1.09 | -5.34 | -4.66 |
| 95 | Versicolor | PSTDMEAN | | 0.18 | -0.85 | 1.17 | 0.62 |
| 96 | Virginia | PSTDMEAN | | 1.45 | -0.25 | 4.17 | 4.04 |
| 97 | | LNDETERM | | 8.46 | 8.46 | 8.46 | 8.46 |
| 98 | Setosa | LNDETERM | | 5.35 | 5.35 | 5.35 | 5.35 |

Analiza dyskryminacyjna irysów

Wyniki analizy dyskryminacji

| Obs. | Species | TYPE | NAME | SepalLength | SepalWidth | PetalLength | PetalWidth |
|------|------------|---------|-------------|-------------|------------|-------------|------------|
| 99 | Versicolor | LNDTERM | | 7.55 | 7.55 | 7.55 | 7.55 |
| 100 | Virginica | LNDTERM | | 9.49 | 9.49 | 9.49 | 9.49 |
| 101 | Setosa | QUAD | SepalLength | -0.09 | 0.06 | 0.02 | 0.02 |
| 102 | Setosa | QUAD | SepalWidth | 0.06 | -0.08 | -0.01 | 0.01 |
| 103 | Setosa | QUAD | PetalLength | 0.02 | -0.01 | -0.19 | 0.09 |
| 104 | Setosa | QUAD | PetalWidth | 0.02 | 0.01 | 0.06 | -0.53 |
| 105 | Setosa | LINEAR | | 4.46 | -0.76 | 3.36 | -3.13 |
| 106 | Setosa | QUAD | CONST_ | -121.83 | -121.83 | -121.83 | -121.83 |
| 107 | Versicolor | QUAD | SepalLength | -0.05 | 0.02 | 0.04 | -0.03 |
| 108 | Versicolor | QUAD | SepalWidth | 0.02 | -0.10 | -0.01 | 0.10 |
| 109 | Versicolor | QUAD | PetalLength | 0.04 | -0.01 | -0.10 | 0.13 |
| 110 | Versicolor | QUAD | PetalWidth | -0.03 | 0.10 | 0.13 | -0.44 |
| 111 | Versicolor | LINEAR | | 1.80 | 1.60 | 0.33 | -1.47 |
| 112 | Versicolor | QUAD | CONST_ | -76.55 | -76.55 | -76.55 | -76.55 |
| 113 | Virginia | QUAD | SepalLength | -0.05 | 0.02 | 0.05 | -0.01 |
| 114 | Virginia | QUAD | SepalWidth | 0.02 | -0.08 | -0.01 | 0.04 |
| 115 | Virginia | QUAD | PetalLength | 0.05 | -0.01 | -0.07 | 0.01 |
| 116 | Virginia | QUAD | PetalWidth | -0.01 | 0.04 | 0.01 | -0.10 |
| 117 | Virginia | QUAD | LINEAR_ | 0.74 | 1.32 | 0.62 | 0.97 |
| 118 | Virginia | QUAD | CONST_ | -75.82 | -75.82 | -75.82 | -75.82 |

Kod 13

| Obs. | Species | _TYPE_ | _NAME_ | SepalLength | SepalWidth | PetalLength | PetalWidth |
|------|------------|--------|-------------|-------------|------------|-------------|------------|
| 1 | | N | | 150.00 | 150.00 | 150.00 | 150.00 |
| 2 | Setosa | N | | 50.00 | 50.00 | 50.00 | 50.00 |
| 3 | Versicolor | N | | 50.00 | 50.00 | 50.00 | 50.00 |
| 4 | Virginica | N | | 50.00 | 50.00 | 50.00 | 50.00 |
| 5 | | MEAN | | 58.43 | 30.57 | 37.58 | 11.99 |
| 6 | Setosa | MEAN | | 50.06 | 34.28 | 14.62 | 2.46 |
| 7 | Versicolor | MEAN | | 59.36 | 27.70 | 42.60 | 13.26 |
| 8 | Virginica | MEAN | | 65.88 | 29.74 | 55.52 | 20.26 |
| 9 | Setosa | PRIOR | | 0.33 | 0.33 | 0.33 | 0.33 |
| 10 | Versicolor | PRIOR | | 0.33 | 0.33 | 0.33 | 0.33 |
| 11 | Virginica | PRIOR | | 0.33 | 0.33 | 0.33 | 0.33 |
| 12 | Setosa | CSSCP | SepalLength | 608.82 | 486.16 | 80.14 | 50.62 |
| 13 | Setosa | CSSCP | SepalWidth | 486.16 | 704.08 | 57.32 | 45.56 |
| 14 | Setosa | CSSCP | PetalLength | 80.14 | 57.32 | 147.78 | 29.74 |
| 15 | Setosa | CSSCP | PetalWidth | 50.62 | 45.56 | 29.74 | 54.42 |
| 16 | Versicolor | CSSCP | SepalLength | 1305.52 | 417.40 | 896.20 | 273.32 |

Koniec.