

Politechnika Warszawska

WYDZIAŁ ELEKTRONIKI
I TECHNIK INFORMACYJNYCH



Sprawozdanie

Wprowadzenie do sztucznej inteligencji

Ćwiczenie nr. 7

Mikołaj Bańkowski

Numer albumu 310408

prowadzący
Grzegorz Rypeś

Warszawa 2024

1. Titanic

1.1. Atrybuty

Plik test.csv i train.csv składa się z następujących atrybutów:

- PassengerId: Unikalny identyfikator pasażera.
- Pclass: Klasa biletu pasażera (1 = pierwsza klasa, 2 = druga klasa, 3 = trzecia klasa).
- Name: Imię i nazwisko pasażera.
- Sex: Płeć pasażera (male = mężczyzna, female = kobieta).
- Age: Wiek pasażera w latach.
- SibSp: Liczba rodzeństwa lub małżonków pasażera na pokładzie.
- Parch: Liczba rodziców lub dzieci pasażera na pokładzie.
- Ticket: Numer biletu pasażera.
- Fare: Opłata za bilet pasażera.
- Cabin: Numer kabiny pasażera.
- Embarked: Port zaokrętowania pasażera (C = Cherbourg, Q = Queenstown, S = Southampton).

Dla zadań modelowania wybrano następujące pięć atrybutów:

- Pclass: Klasa biletu pasażera, informująca o warunkach podróży i kosztach biletu, co może być istotne dla przeżycia.
- Sex: Płeć pasażera, ponieważ statystycznie kobiety miały większe szanse na przeżycie.
- Age: Wiek pasażera, ponieważ dzieci mogły mieć priorytet w trakcie ewakuacji.
- Fare: Opłata za bilet, co może być związane z zasobami finansowymi pasażera i jego miejscem na statku.
- Embarked: Port zaokrętowania, ponieważ miejsce zaokrętowania może wpłynąć na przeżycie z powodu różnic w dostępnych zasobach ratunkowych.

Atrybut 'Age' został zamieniony na atrybut, którego wartości to {dziecko, dorosły, osoba starsza}.

1.2. Model Naiwny Bayesa

1.2.1. Jaka jest jakość stworzonego klasyfikatora?

Tabela przedstawiająca wynik ewaluacji modelu

Zbiór danych	Dokładność
Zbiór walidacyjny	0.776536
Zbiór treningowy	0.780899

Ocena modelu ze strony
[Kaggle - Titanic - Machine Learning from Disaster](#)

Submissions


All	Successful	Errors	Recent ▾
Submission and Description			Public Score ⓘ
	gender_submission.csv Complete · 4h ago	0.73684	

Tabela przedstawiająca raport klasyfikacji dla zbioru walidacyjnego:

Klasa	Precyzja	Czułość (Recall)	F1-score	Support
0	0.82	0.80	0.81	105
1	0.72	0.74	0.73	74
Średnia dokładność	0.78	0.78	0.78	179
Macro avg	0.77	0.77	0.77	179
Weighted avg	0.78	0.78	0.78	179

Opis poszczególnych wierszy:

Klasa 0 (przeżył)

- Precyzja (Precision): 0.82 - Procent prawidłowo przewidzianych przeżyć spośród wszystkich przewidzianych jako przeżycie.
- Czułość (Recall): 0.80 - Procent rzeczywistych przeżyć, które zostały prawidłowo przewidziane.
- F1-score: 0.81 - Średnia harmoniczna precyzji i czułości.
- Support: 105 - Liczba rzeczywistych przykładów tej klasy w zbiorze walidacyjnym.

Klasa 1 (zmarł)

- Precyzja (Precision): 0.72 - Procent prawidłowo przewidzianych zgonów spośród wszystkich przewidzianych jako zgon.
- Czułość (Recall): 0.74 - Procent rzeczywistych zgonów, które zostały prawidłowo przewidziane.
- F1-score: 0.73 - Średnia harmoniczna precyzji i czułości.
- Support: 74 - Liczba rzeczywistych przykładów tej klasy w zbiorze walidacyjnym.

Średnia dokładność (accuracy): 0.78

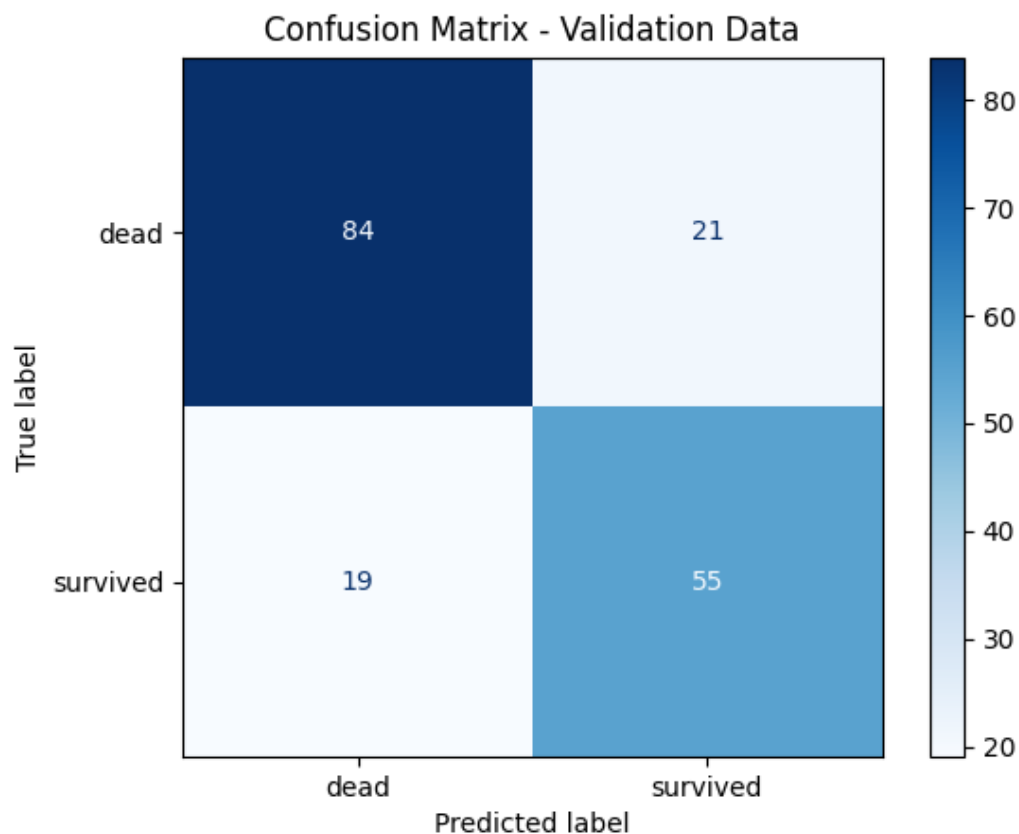
- Średnia dokładność modelu na zbiorze walidacyjnym, wyrażona jako procent poprawnych przewidywań ze wszystkich przewidywań.

Macro avg (średnia arytmetyczna)

- Precyzja: 0.77 - Średnia arytmetyczna precyzji dla obu klas.
- Czułość (Recall): 0.77 - Średnia arytmetyczna czułości dla obu klas.
- F1-score: 0.77 - Średnia arytmetyczna F1-score dla obu klas.
- Support: 179 - Łączna liczba przykładów w zbiorze walidacyjnym.

Weighted avg (średnia ważona)

- Precyzja: 0.78 - Średnia ważona precyzji, z wagami odpowiadającymi liczbie przykładów w każdej klasie.
- Czułość (Recall): 0.78 - Średnia ważona czułości, z wagami odpowiadającymi liczbie przykładów w każdej klasie.
- F1-score: 0.78 - Średnia ważona F1-score, z wagami odpowiadającymi liczbie przykładów w każdej klasie.
- Support: 179 - Łączna liczba przykładów w zbiorze walidacyjnym.



Dead (Nie przeżył): 19 przypadków, gdzie model błędnie przewidział, że osoba nie przeżyła, mimo że faktycznie przeżyła.

Dead (Nie przeżył): 84 przypadki, gdzie model poprawnie przewidział, że osoba nie przeżyła.

Survived (Przeżył): 55 przypadków, gdzie model poprawnie przewidział, że osoba przeżyła.

Survived (Przeżył): 21 przypadków, gdzie model błędnie przewidział, że osoba przeżyła, mimo że faktycznie nie przeżyła.

Tabela przedstawiająca raport klasyfikacji dla zbioru treningowego:

Klasa	Precyzja	Czułość (Recall)	F1-score	Liczebność (Support)
0	0.83	0.81	0.82	444
1	0.70	0.73	0.71	268
Średnia dokładność	0.78	0.78	0.78	712
Macro avg	0.77	0.77	0.77	712
Weighted avg	0.78	0.78	0.78	712

Opis poszczególnych wierszy:

Klasa 0 (przeżył)

- Precyzja (Precision): 0.83 - Procent prawidłowo przewidzianych przeżyć spośród wszystkich przewidzianych jako przeżycie.
- Czułość (Recall): 0.81 - Procent rzeczywistych przeżyć, które zostały prawidłowo przewidziane.
- F1-score: 0.82 - Średnia harmoniczna precyzji i czułości.
- Liczebność (Support): 444 - Liczba rzeczywistych przykładów tej klasy w zbiorze treningowym.

Klasa 1 (zmarł)

- Precyzja (Precision): 0.70 - Procent prawidłowo przewidzianych zgonów spośród wszystkich przewidzianych jako zgon.
- Czułość (Recall): 0.73 - Procent rzeczywistych zgonów, które zostały prawidłowo przewidziane.
- F1-score: 0.71 - Średnia harmoniczna precyzji i czułości.
- Liczebność (Support): 268 - Liczba rzeczywistych przykładów tej klasy w zbiorze treningowym.

Średnia dokładność (accuracy): 0.78

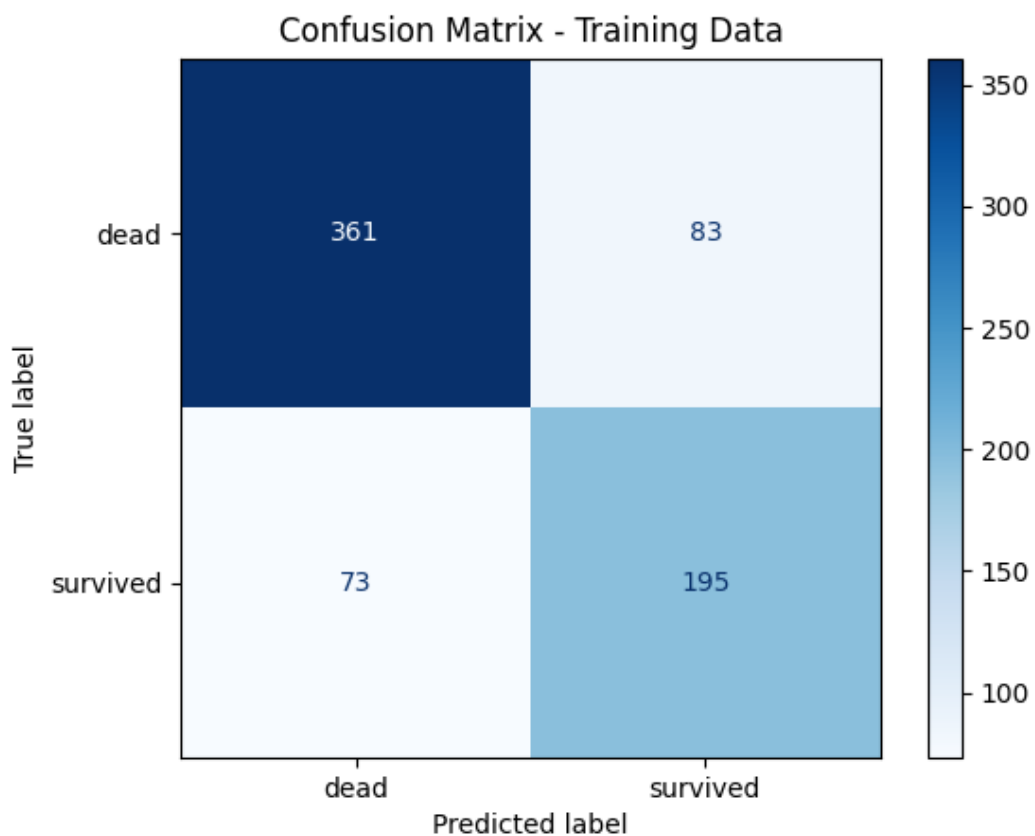
- Średnia dokładność modelu na zbiorze treningowym, wyrażona jako procent poprawnych przewidywań ze wszystkich przewidywań.

Macro avg (średnia arytmetyczna)

- Precyzja: 0.77 - Średnia arytmetyczna precyzji dla obu klas.
- Czułość (Recall): 0.77 - Średnia arytmetyczna czułości dla obu klas.
- F1-score: 0.77 - Średnia arytmetyczna F1-score dla obu klas.
- Liczebność (Support): 712 - łączna liczba przykładów w zbiorze treningowym.

Weighted avg (średnia ważona)

- Precyzja: 0.78 - Średnia ważona precyzji, z wagami odpowiadającymi liczbie przykładów w każdej klasie.
- Czułość (Recall): 0.78 - Średnia ważona czułości, z wagami odpowiadającymi liczbie przykładów w każdej klasie.
- F1-score: 0.78 - Średnia ważona F1-score, z wagami odpowiadającymi liczbie przykładów w każdej klasie.
- Liczebność (Support): 712 - łączna liczba przykładów w zbiorze treningowym.



Dead (Nie przeżył): 361 przypadków, gdzie model błędnie przewidział, że osoba nie przeżyła, mimo że faktycznie przeżyła.

Dead (Nie przeżył): 73 przypadki, gdzie model poprawnie przewidział, że osoba nie przeżyła.

Survived (Przeżył): 195 przypadków, gdzie model poprawnie przewidział, że osoba przeżyła.

Survived (Przeżył): 83 przypadków, gdzie model błędnie przewidział, że osoba przeżyła, mimo że faktycznie nie przeżyła.

Czy jest niedouczenie, albo przeuczenie?

Model nie wydaje się być przeuczony ani niedouczony, biorąc pod uwagę bliskość wyników na zbiorach walidacyjnym i treningowym.

Jeśli różnice te byłyby większe, można by podejrzewać przeuczenie. Wartości są jednak bliskie, co sugeruje, że model generalizuje całkiem dobrze..

Jak to zmierzyć skoro zbiór testowy nie jest otagowany?

Bez otagowanego zbioru testowego nie można bezpośrednio zmierzyć dokładności na danych testowych.

Można używać zbioru walidacyjnego jako przybliżenia do oceny modelu lub przestać zadanie na stornie Kaggle, gdzie nasz model zostanie poddany ocenie.

Alternatywnie, można skorzystać z technik takich jak krosvalidacja (cross-validation) do oceny stabilności modelu. Krosvalidacja to technika oceny modelu, która polega na podzieleniu danych na kilka podzbiorów (zwanym "foldami") w celu lepszego oszacowania jego wydajności i generalizacji.

Tabela przedstawiająca prawdopodobieństwa przeżycia w zależności od płci:

Kategoria	Prawdopodobieństwo przeżycia
Mężczyzna	0.3124150782959129
Kobieta	0.9476264243772848
Chłopiec	0.9827081623380888
Dziewczynka	0.9827081623380888

Funkcja "predict_survival_proba" przekształca cechy pasażera na wektor cech zgodny z danymi treningowymi, a następnie używa modelu Naive Bayes do obliczenia prawdopodobieństwa przeżycia.

2. Snake

2.1. Atrybuty

Obstacles (Przeszkody):

- left: Czy na lewo od głowy węża znajduje się przeszkoda (ciało węża) lub krawędź planszy.
- right: Czy na prawo od głowy węża znajduje się przeszkoda (ciało węża) lub krawędź planszy.
- up: Czy nad głową węża znajduje się przeszkoda (ciało węża) lub krawędź planszy.
- down: Czy pod głową węża znajduje się przeszkoda (ciało węża) lub krawędź planszy.

Food Direction (Kierunek Jedzenia):

- food_left: Czy jedzenie znajduje się na lewo od głowy węża.
- food_right: Czy jedzenie znajduje się na prawo od głowy węża.
- food_up: Czy jedzenie znajduje się powyżej głowy węża.
- food_down: Czy jedzenie znajduje się poniżej głowy węża.

Snake Direction (Kierunek Węża):

- direction_feature: Aktualny kierunek ruchu węża, reprezentowany jako wektor czteroelementowy (dla każdego kierunku jeden element):
 - [1, 0, 0, 0] - Wąż porusza się w górę.
 - [0, 1, 0, 0] - Wąż porusza się w prawo.
 - [0, 0, 1, 0] - Wąż porusza się w dół.
 - [0, 0, 0, 1] - Wąż porusza się w lewo.

2.2. Model Naiwny Bayesa

Tabela przedstawiająca wynik ewaluacji modelu

Zbiór danych	Dokładność
Zbiór treningowy	0.8379739336492891
Zbiór testowy	0.842603550295858

Model nie wydaje się być przeuczony ani niedouczony, biorąc pod uwagę bliskość wyników na zbiorach walidacyjnym i treningowym.

Jeśli różnice te byłyby większe, można by podejrzewać przeuczenie. Wartości są jednak bliskie, co sugeruje, że model generalizuje całkiem dobrze..

Tabela przedstawiająca raport klasyfikacji dla zbioru treningowego:

Class	Precision	Recall	F1-Score	Support
0	0.84	0.83	0.84	849
1	0.83	0.84	0.83	847
2	0.82	0.90	0.86	891
3	0.87	0.78	0.82	789
accuracy			0.84	3376
macro avg	0.84	0.84	0.84	3376
weighted avg	0.84	0.84	0.84	3376

Tabela przedstawiająca raport klasyfikacji dla zbioru testowego:

Class	Precision	Recall	F1-Score	Support
0	0.83	0.83	0.83	188
1	0.85	0.86	0.86	226
2	0.84	0.86	0.85	231
3	0.85	0.81	0.83	200
accuracy			0.84	845
macro avg	0.84	0.84	0.84	845
weighted avg	0.84	0.84	0.84	845

W raporcie klasyfikacyjnym dla modelu Naive Bayes, wartości 0, 1, 2 i 3 odpowiadają różnym kierunkom, w które może poruszać się wąż w grze Snake. Oto co oznaczają te liczby:

- 0: Lewo (Left)
- 1: Prawo (Right)
- 2: Góra (Up)
- 3: Dół (Down)

W tabeli przedstawiono wyniki uzyskane przez model w grze Snake. Wyniki pokazują liczbę rozegranych gier oraz średnią liczbę punktów uzyskiwanych w każdej partii. Analiza wskazuje, że niezależnie od liczby rozegranych gier, średnia liczba punktów utrzymuje się na podobnym poziomie, wynoszącym około 33,5 punktów. Wyjątek stanowi 20 i 30 rozegranych gier, gdzie średnia liczba punktów jest nieco niższa, wynosząc 33,4.

Tabela przedstawia średnią ilość punktów na liczbę rozegranych gier

Liczba rozegranych gier	Średnia liczba punktów
10	33.5
20	33.4
30	33.4
40	33.5
50	33.5
60	33.5
70	33.5
80	33.5
90	33.5
100	33.55

Wyniki sugerują stabilność działania modelu, choć mogą również wskazywać na niewielkie różnice w wydajności w zależności od liczby rozegranych gier.