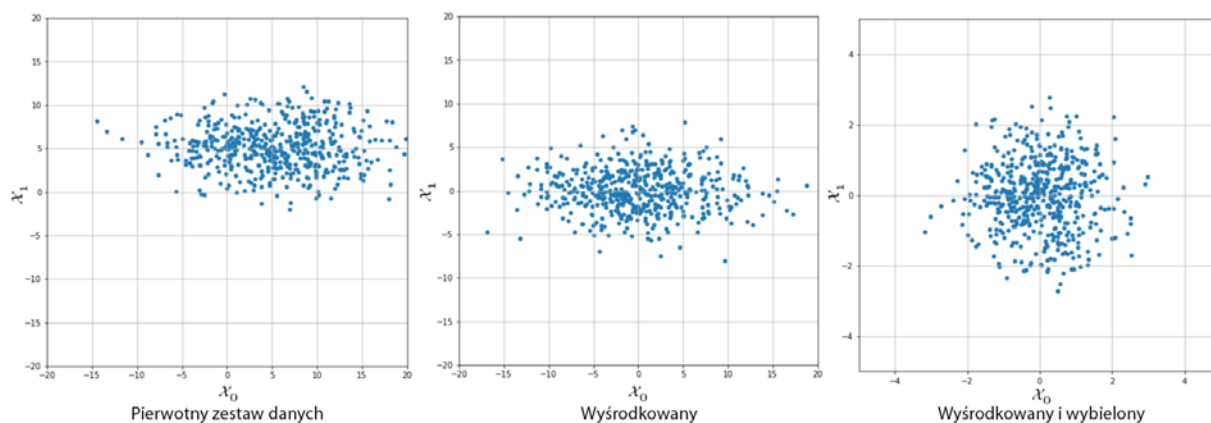
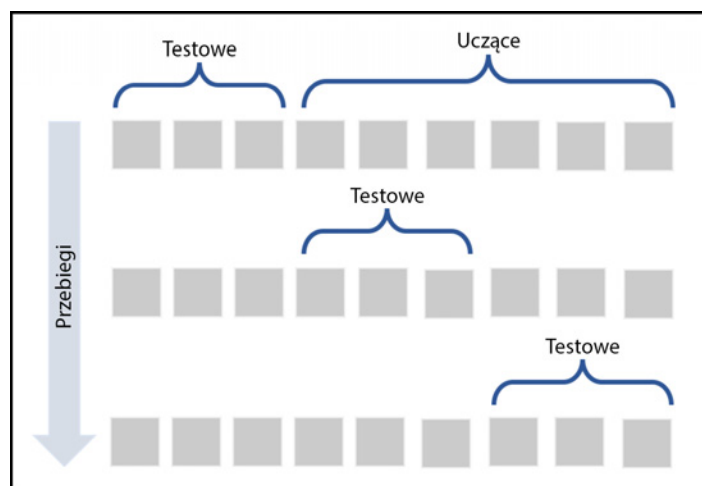


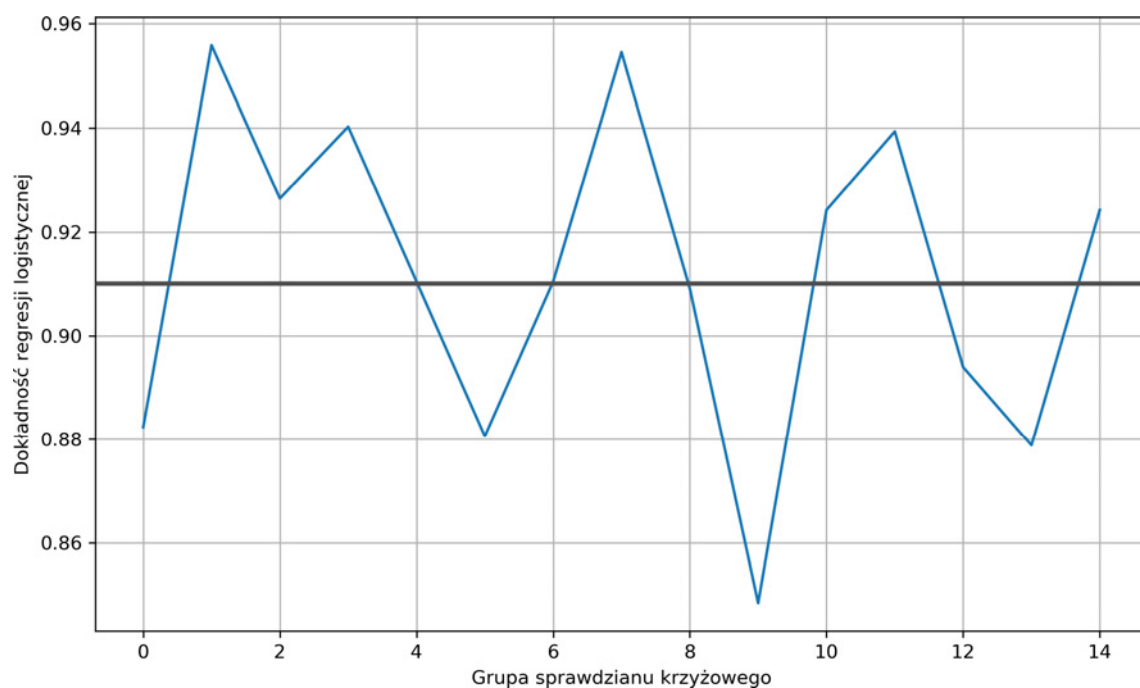
Rysunek 1.1. Schemat typowego modelu parametryzowanego za pomocą wektora  $\theta$



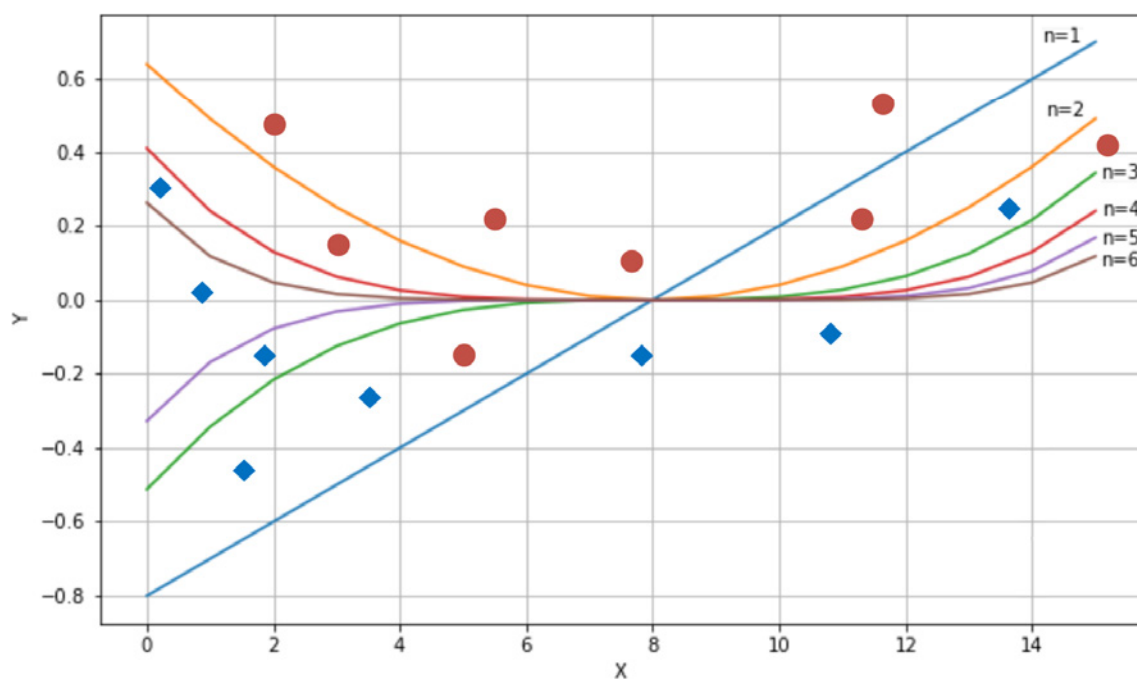
Rysunek 1.2. Pierwotny zestaw danych (po lewej), jego wersja wyśrodkowana (w środku) i wybielona (po prawej)



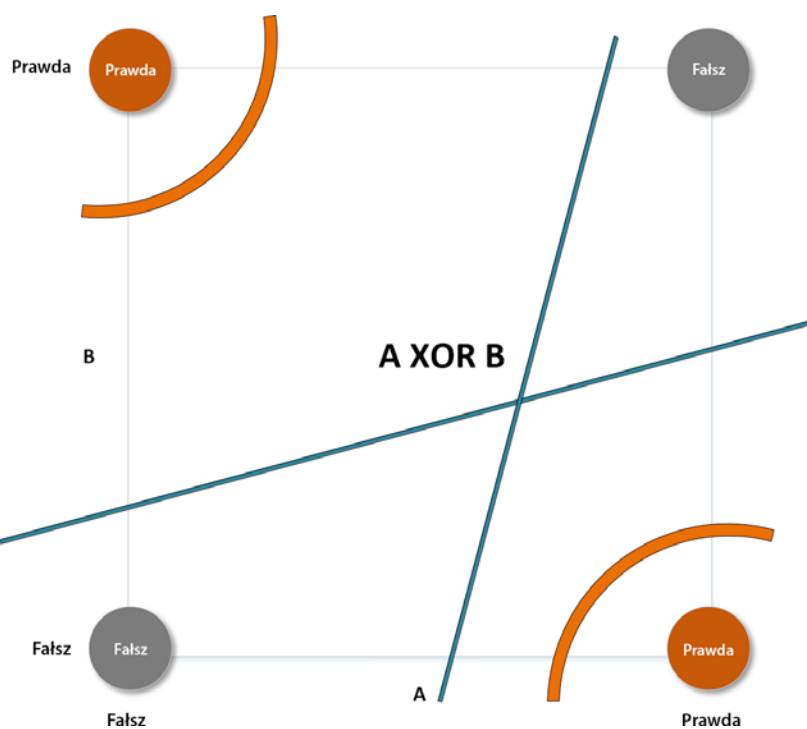
Rysunek 1.3. Schemat k-krotnego sprawdzianu krzyżowego



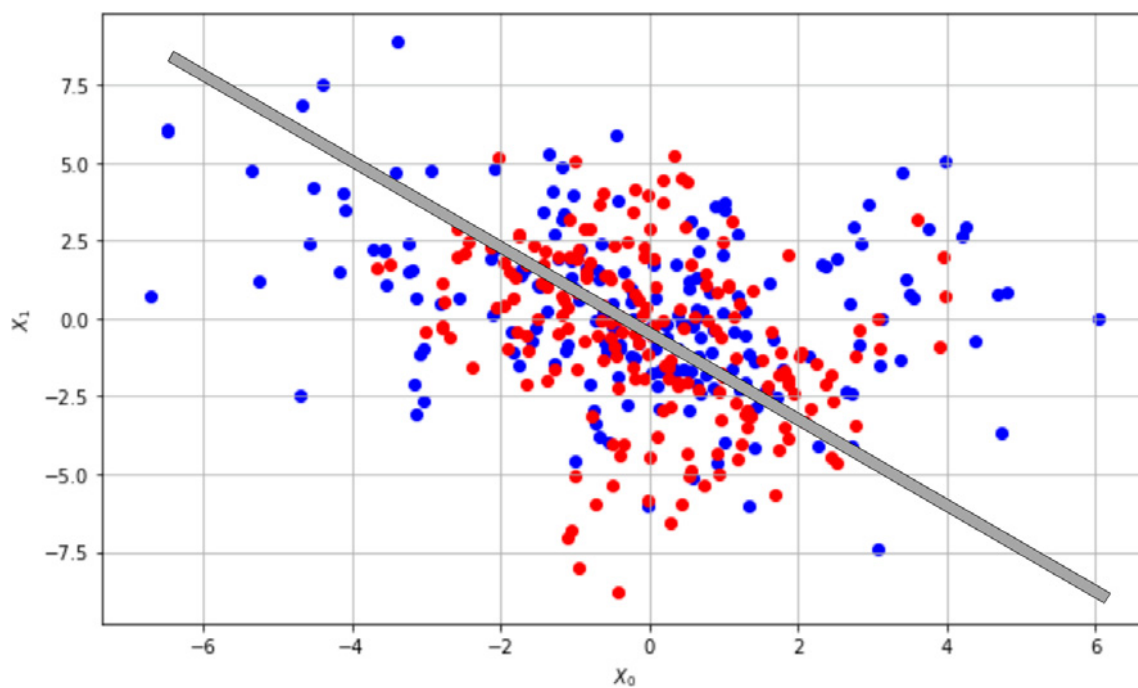
Rysunek 1.4. Dokładności uzyskane za pomocą sprawdzianu krzyżowego



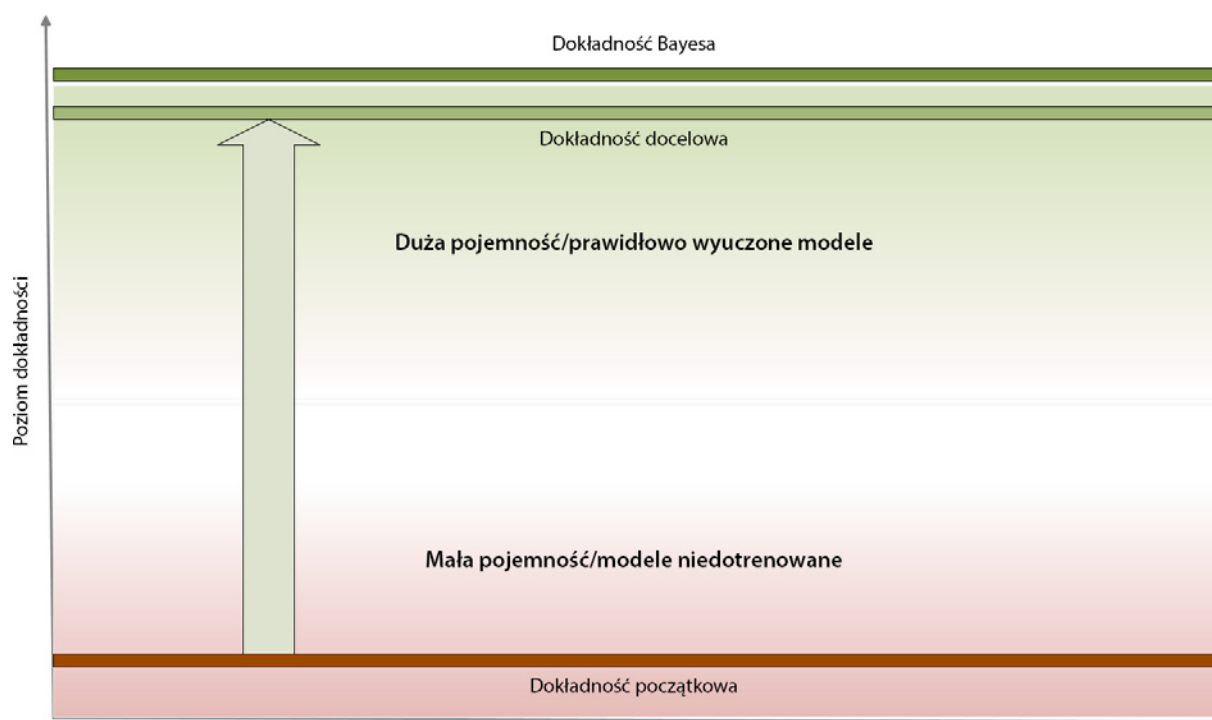
Rysunek 1.5. Przebiegi sześciu różnych, wielomianowych krzywych rozdzielających



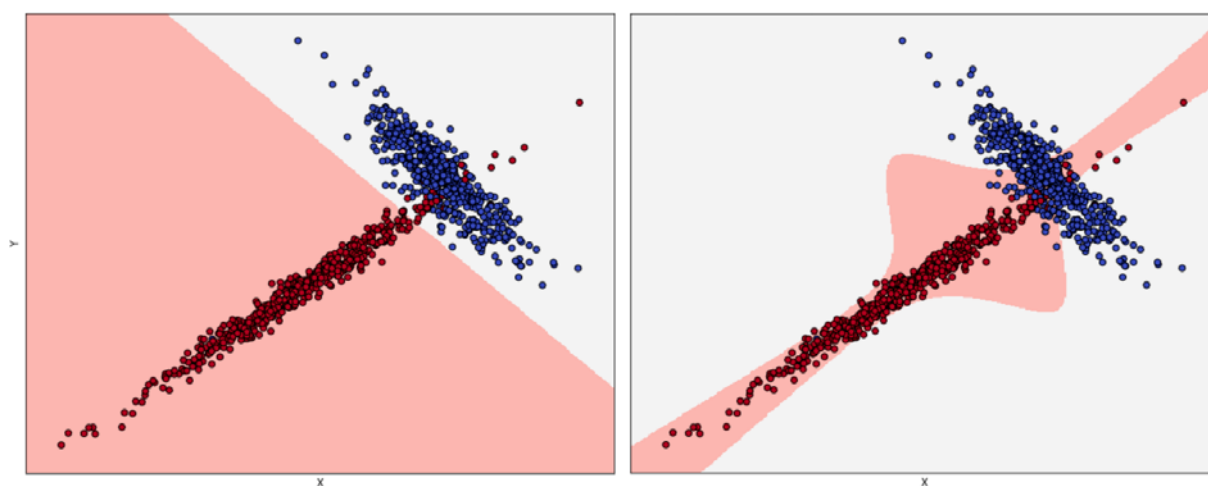
Rysunek 1.6. Problem XOR z ukazanymi różnymi krzywymi rozdzielającymi



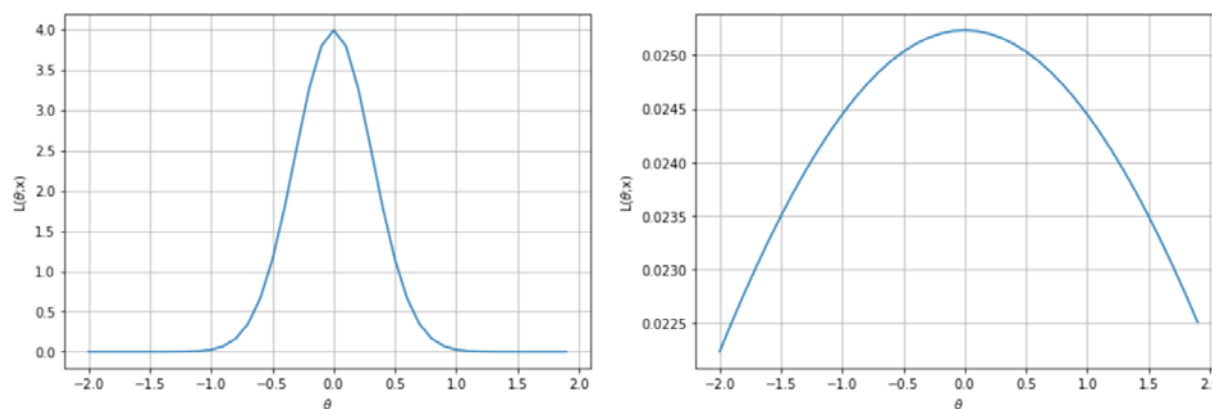
Rysunek 1.7. Klasyfikator niedotrenowany: nie jest w stanie prawidłowo rozdzielić obydwu klas



Rysunek 1.8. Schemat poziomów dokładności

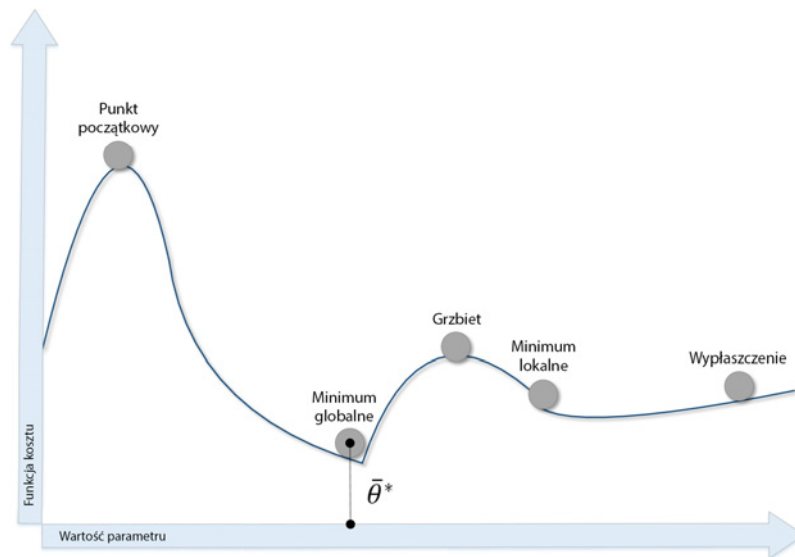


Rysunek 1.9. Prawidłowe dopasowanie (po lewej), klasyfikator przetrenowany (po prawej)

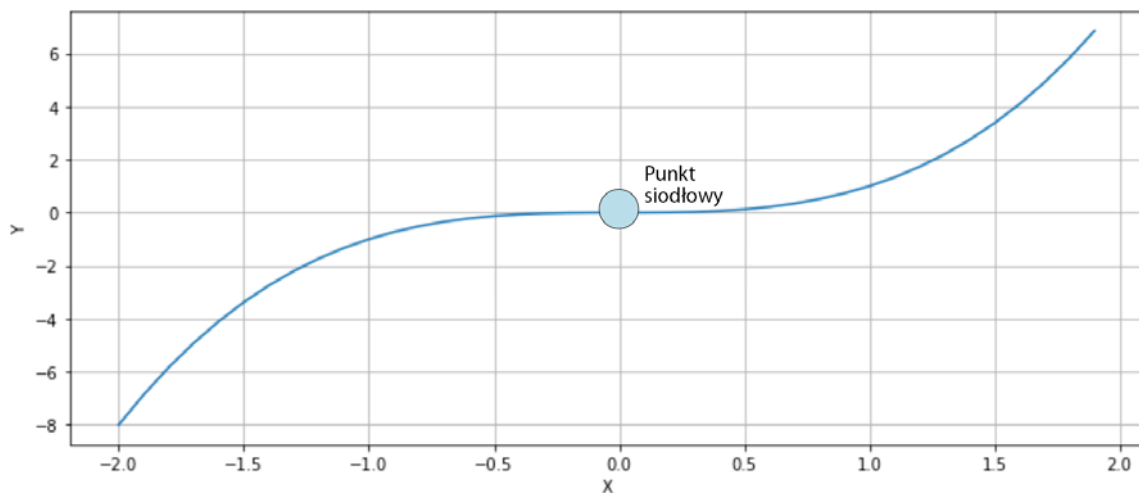


Rysunek 1.10. Spiczasty (po lewej) i spłaszczony (po prawej) przebieg prawdopodobieństwa

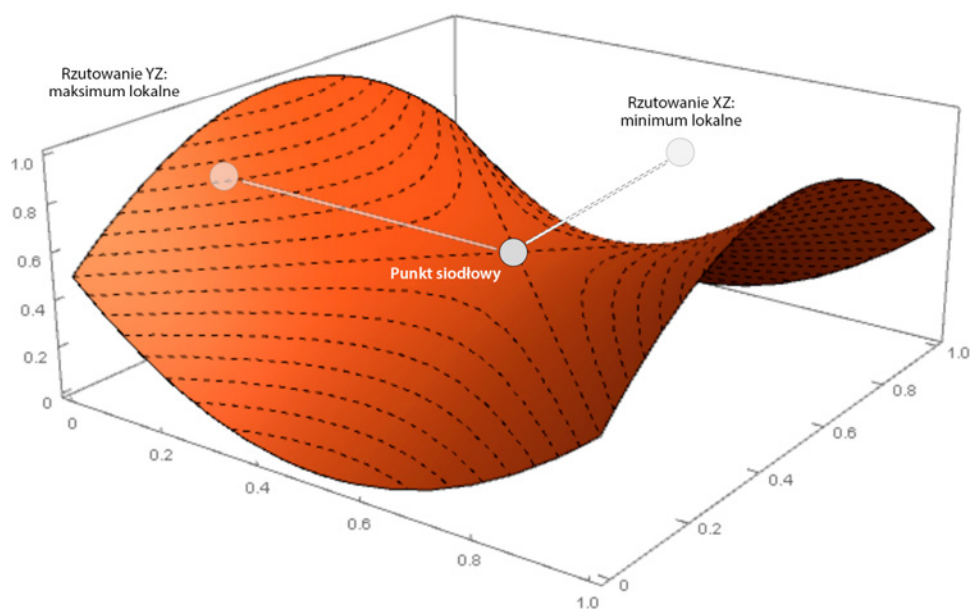




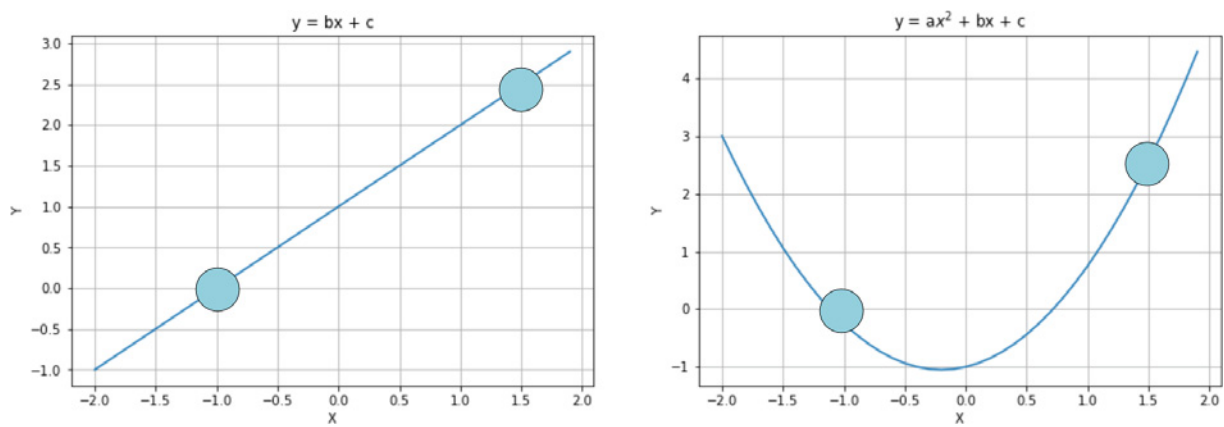
Rysunek 1.11. Różne rodzaje punktów na wykresie dwuwymiarowym



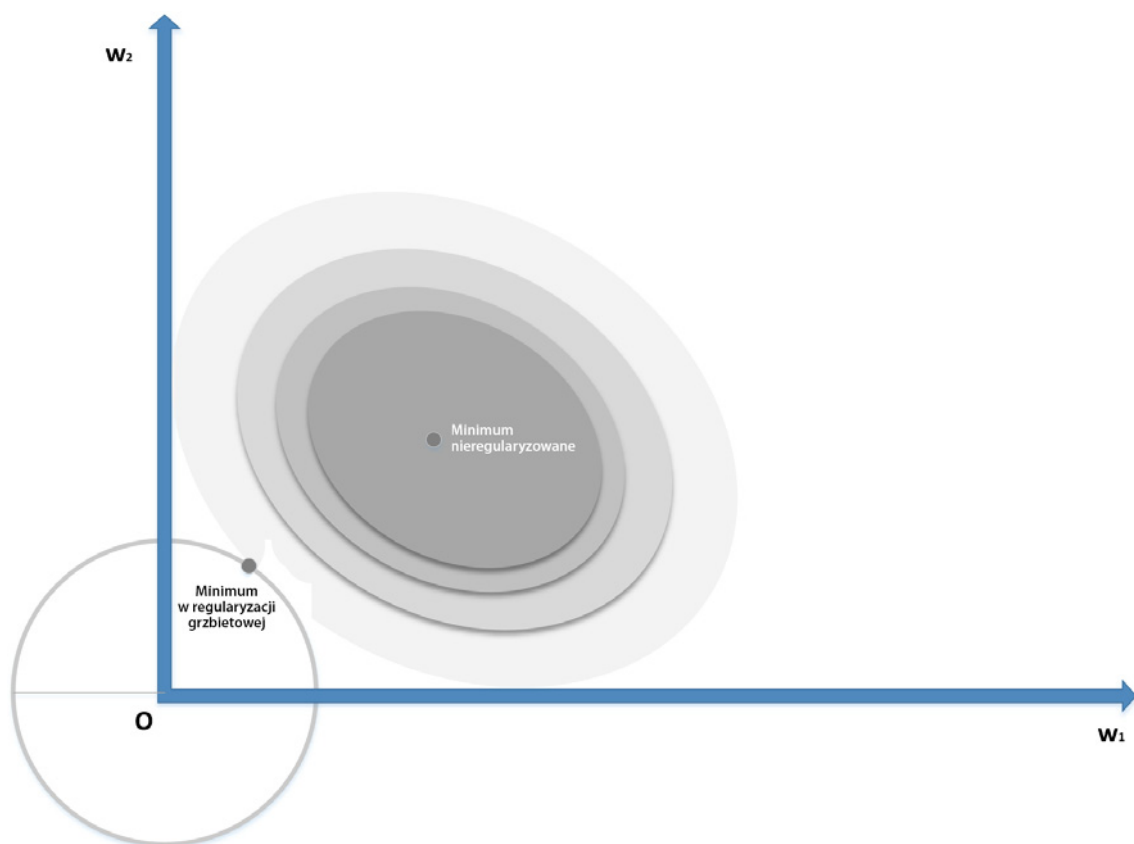
Rysunek 1.12. Punkt siodłowy na wykresie dwuwymiarowym



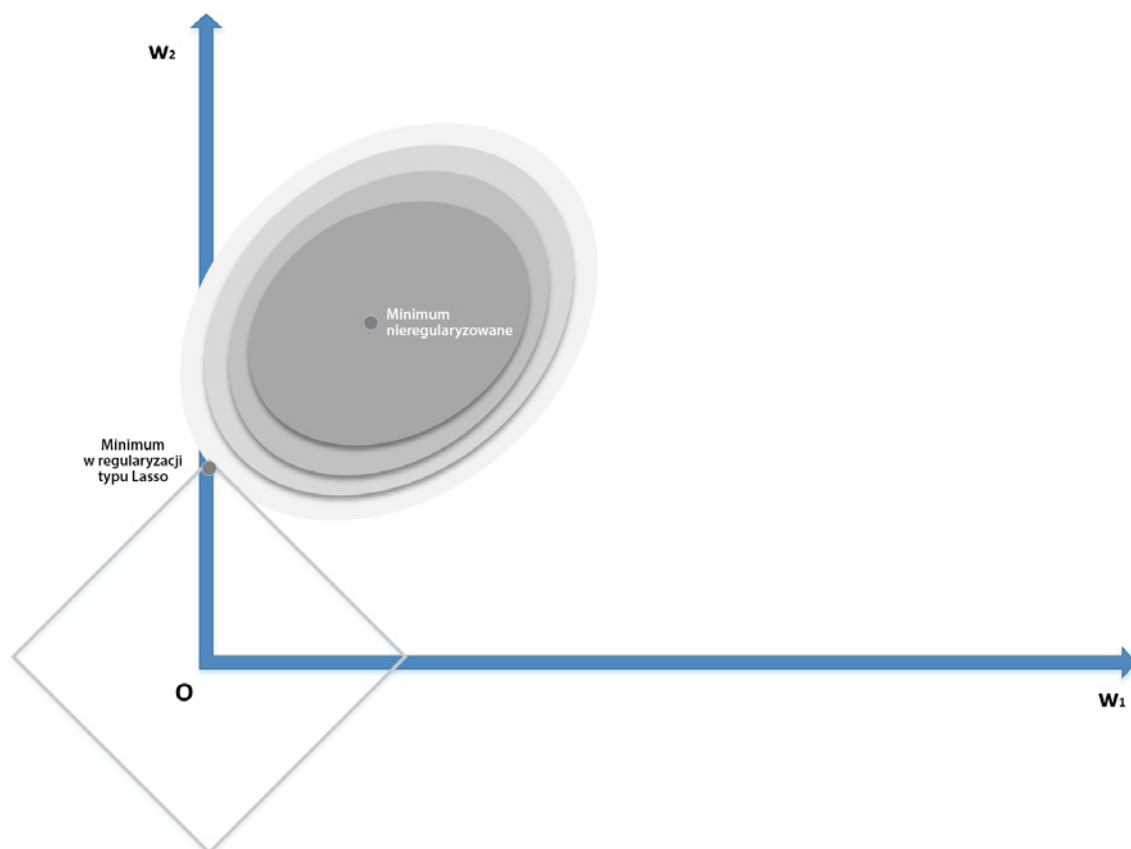
Rysunek 1.13. Punkt siodłowy na wykresie trójwymiarowym



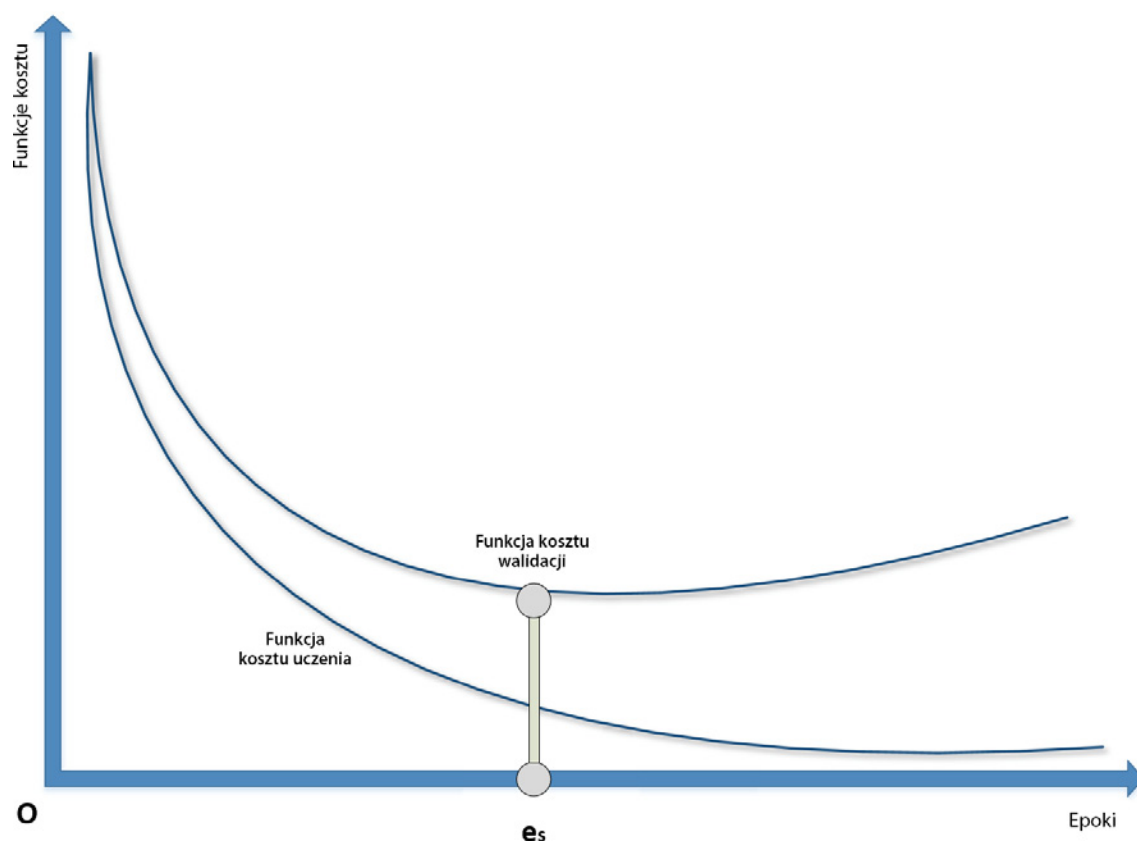
Rysunek 1.14. Interpolacja za pomocą prostej (po lewej) i paraboli (po prawej)



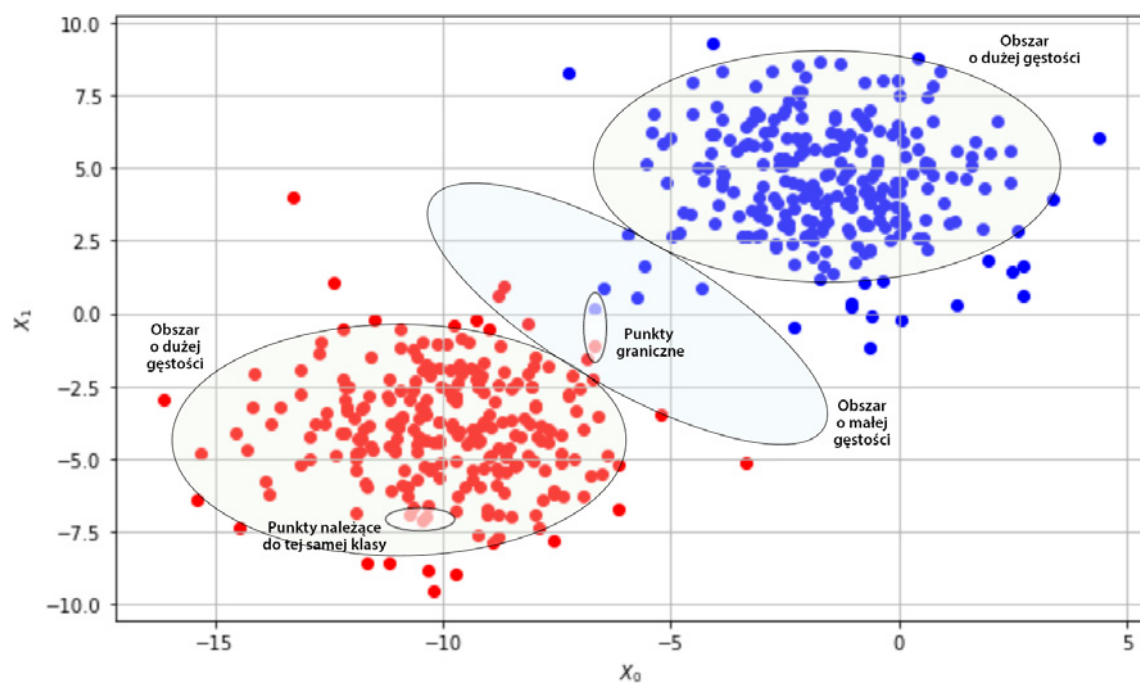
Rysunek 1.15. Regresja grzbietowa (L2)



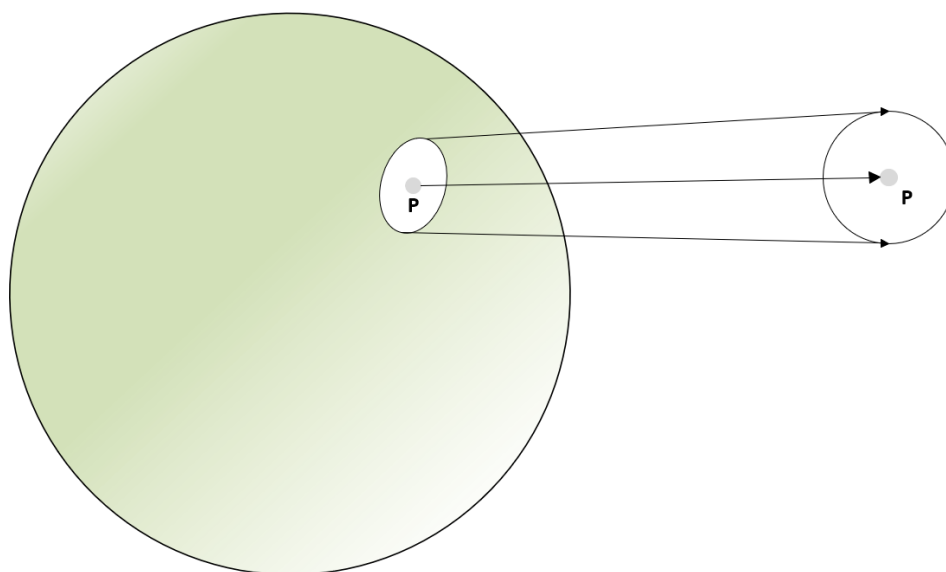
Rysunek 1.16. Regularyzacja typu Lasso (L1)



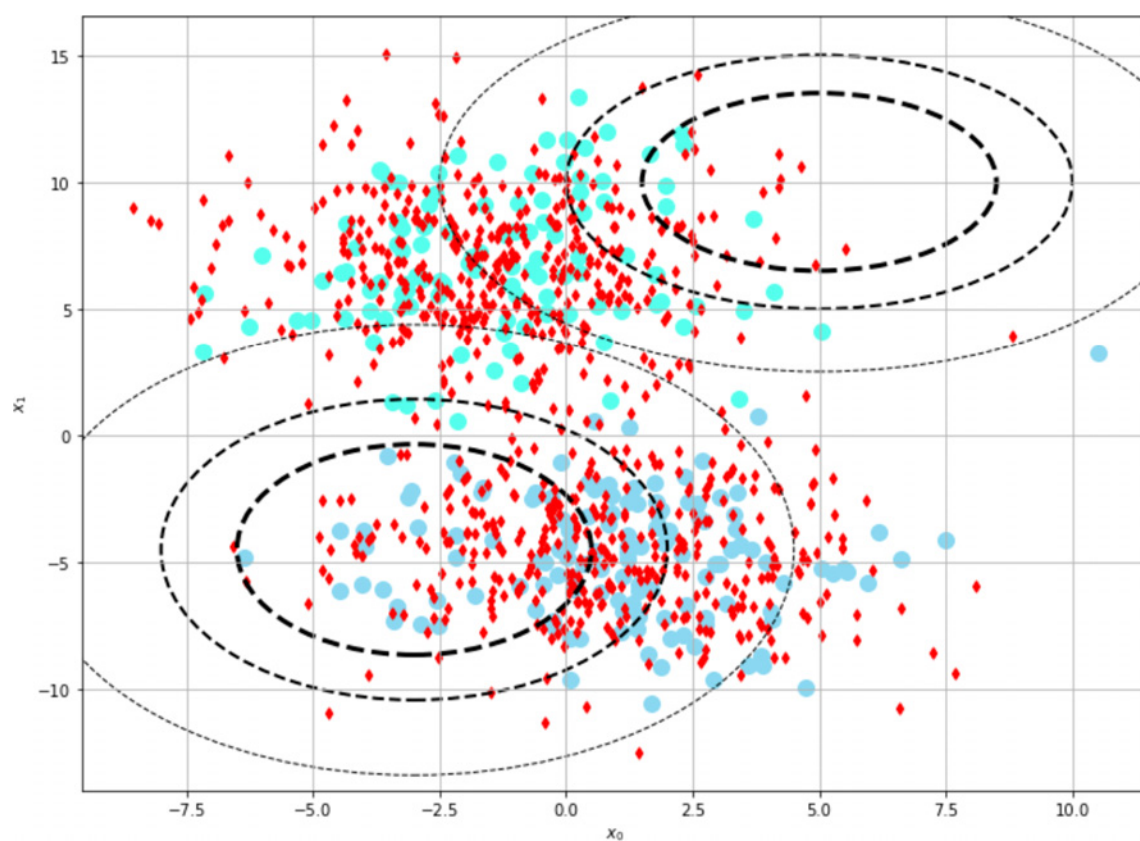
Rysunek 1.17. Przykład zastosowania wczesnego zatrzymywania przed początkiem wznoszenia się krzywej u-kształtnej



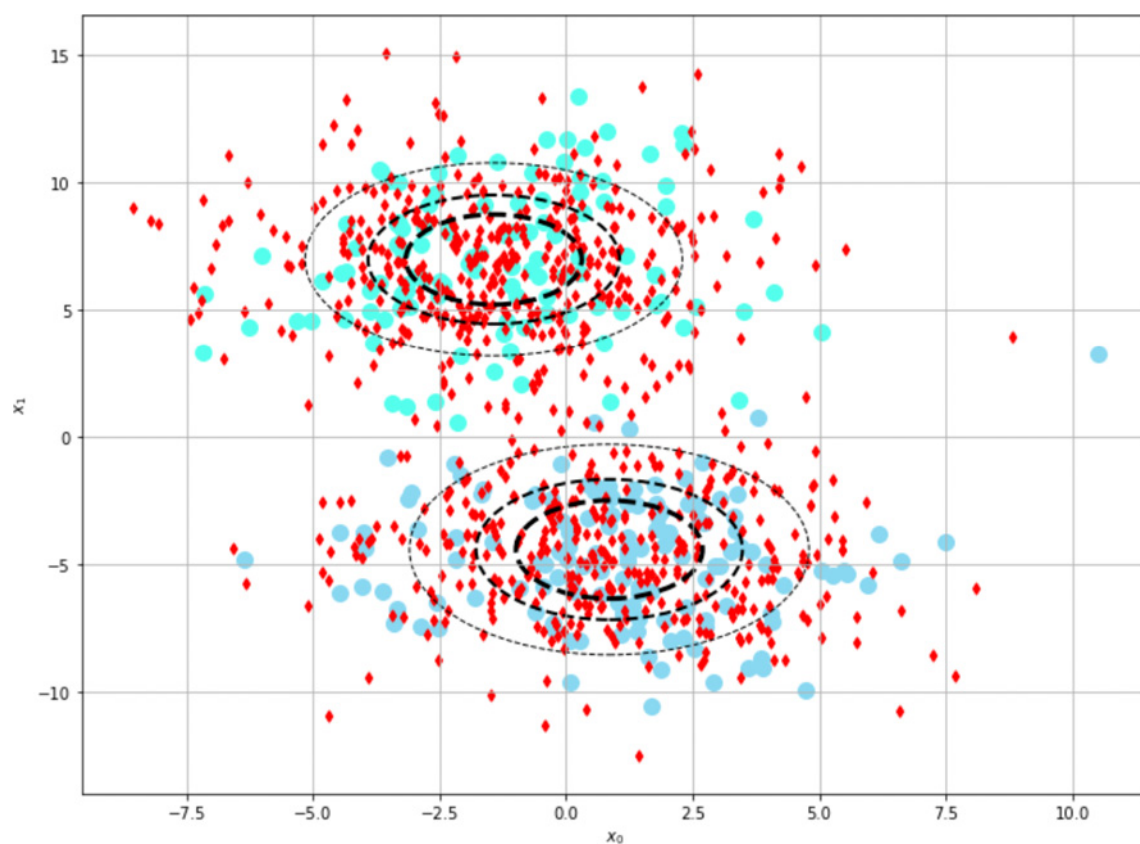
Rysunek 2.1. Maszyna wektorów nośnych w przykładzie ukazującym założenie skupień



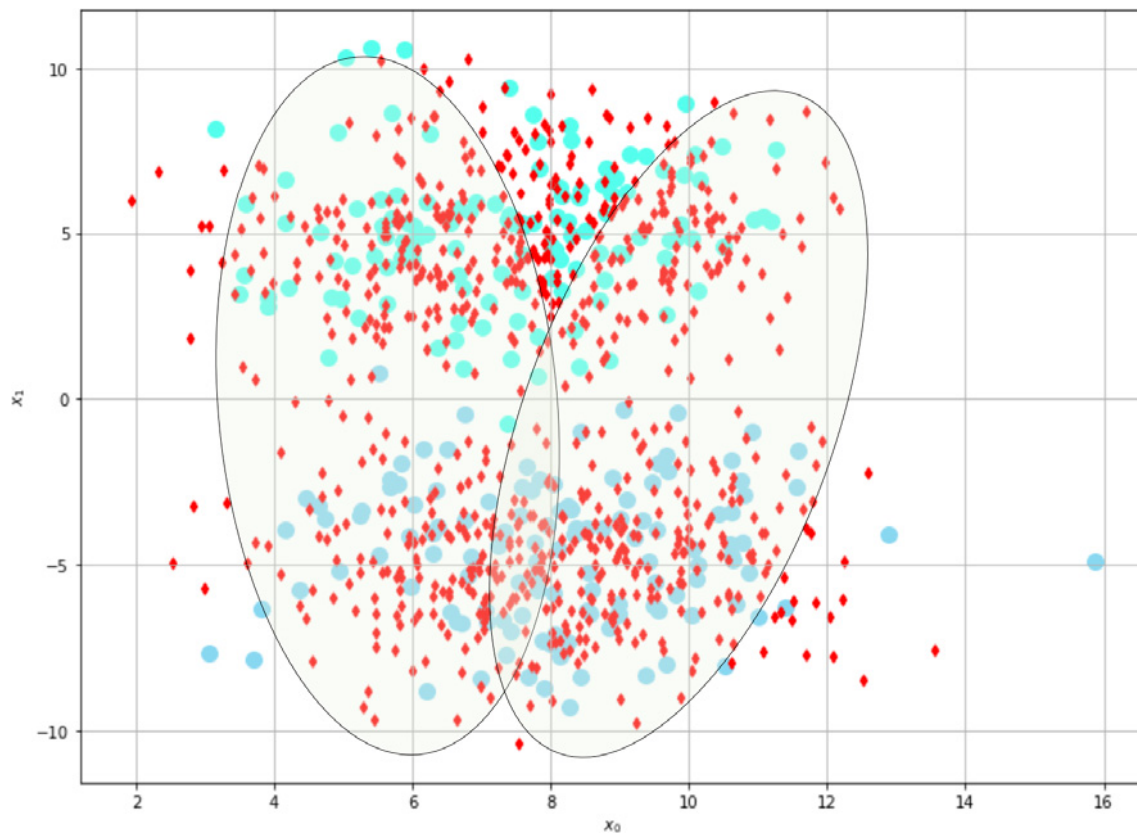
Rysunek 2.2. Rozmaitość dwuwymiarowa uzyskana z powierzchni sferycznej



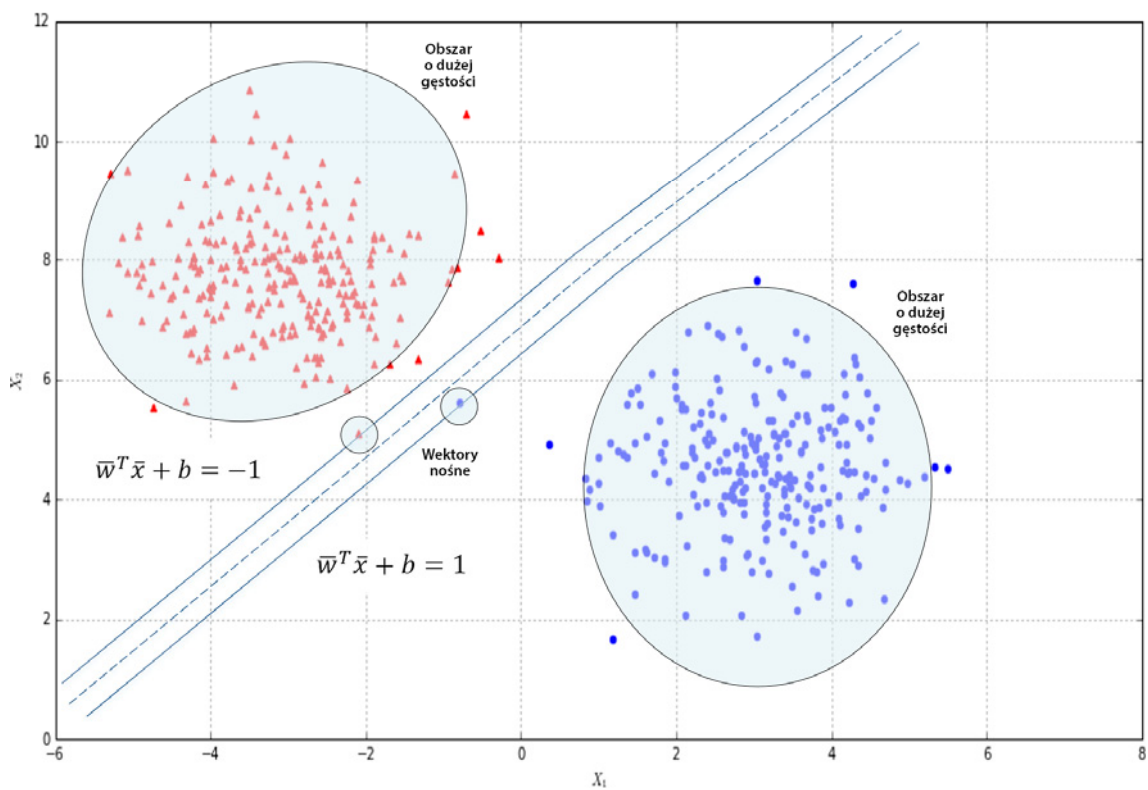
Rysunek 2.3. Konfiguracja początkowa mieszaniny gaussowskiej



Rysunek 2.4. Stan stabilny naszego modelu

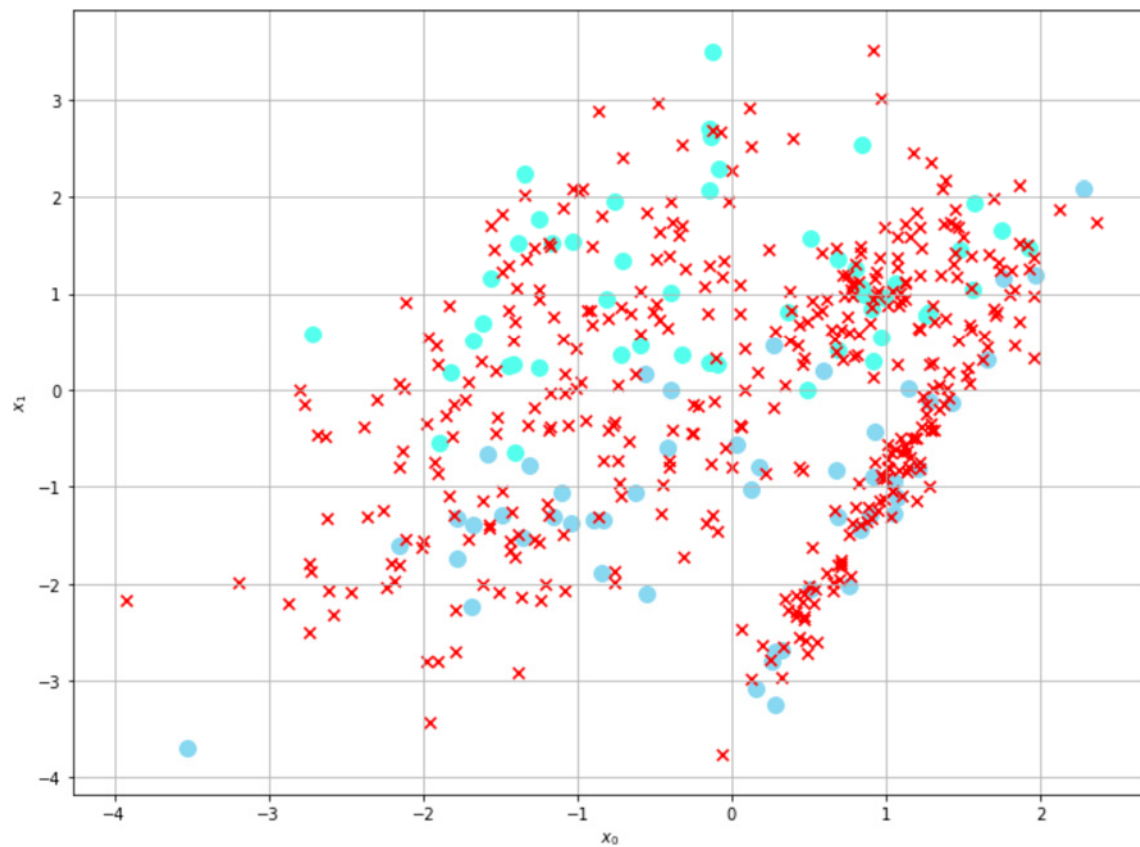


Rysunek 2.5. Obciążona ostateczna konfiguracja mieszaniny gaussowskiej

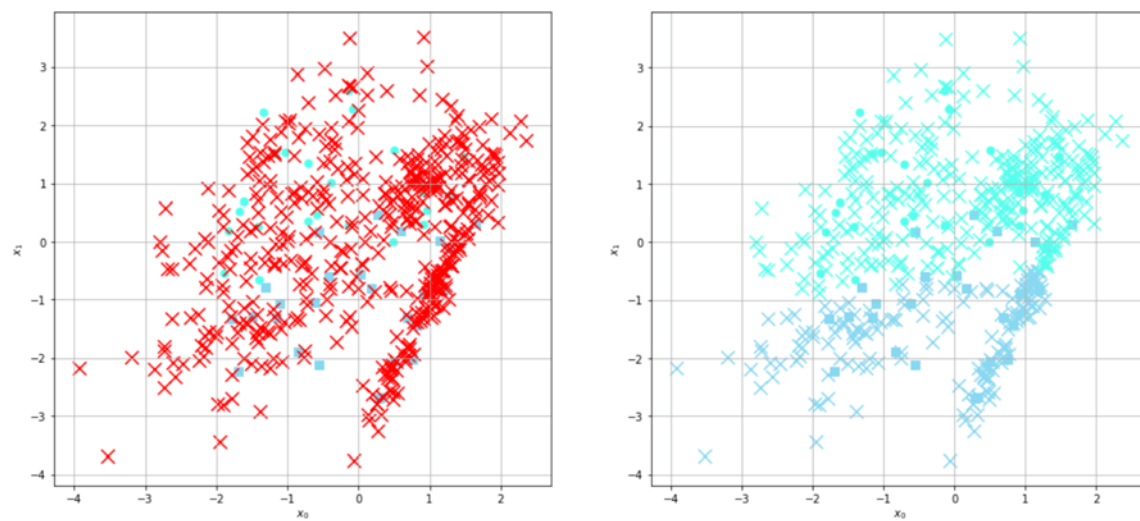


Rysunek 2.6. Ogólny schemat maszyny SVM

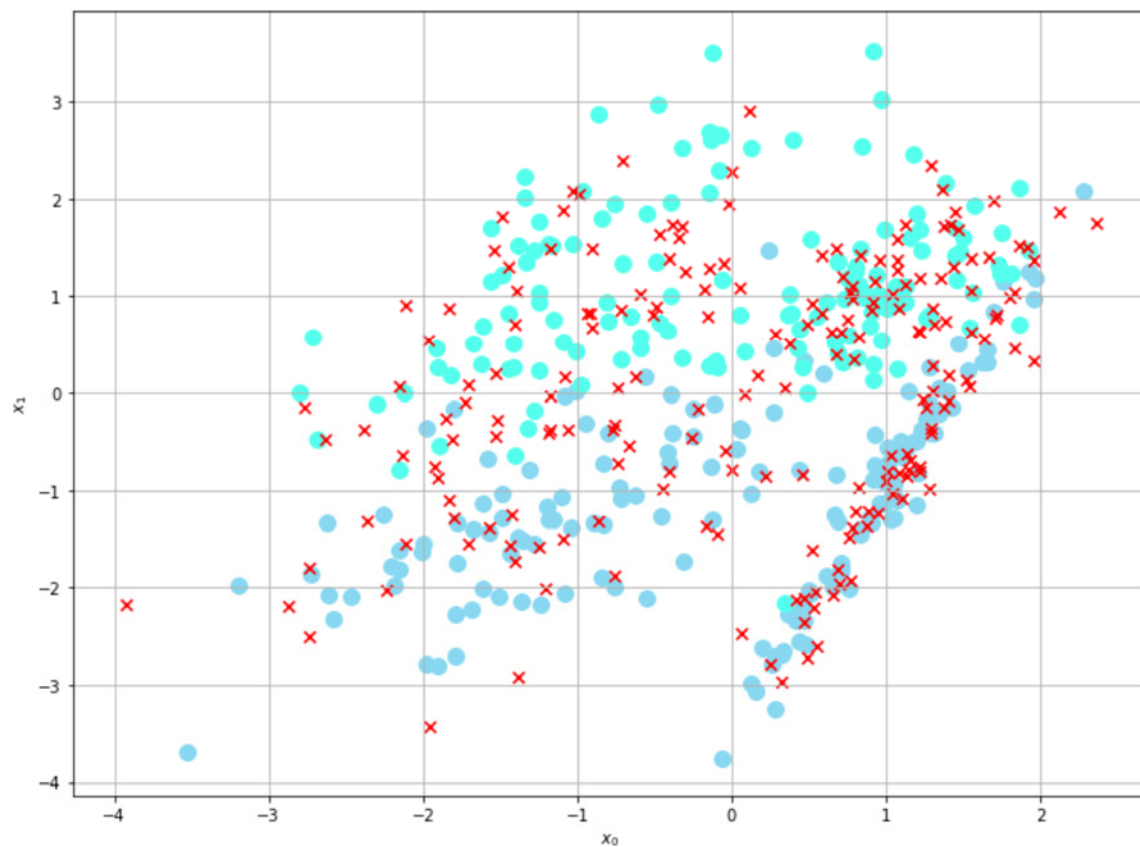




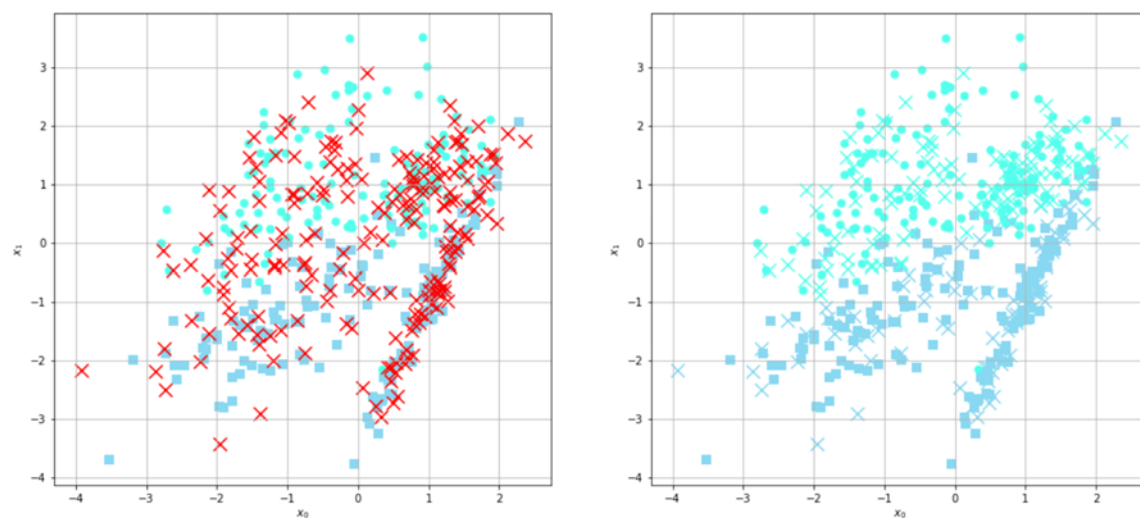
Rysunek 2.7. Pierwotny zestaw danych oznakowanych i nieoznakowanych



Rysunek 2.8. Porównanie pierwotnego zestawu danych (po lewej) z rezultatami klasyfikacji (po prawej)

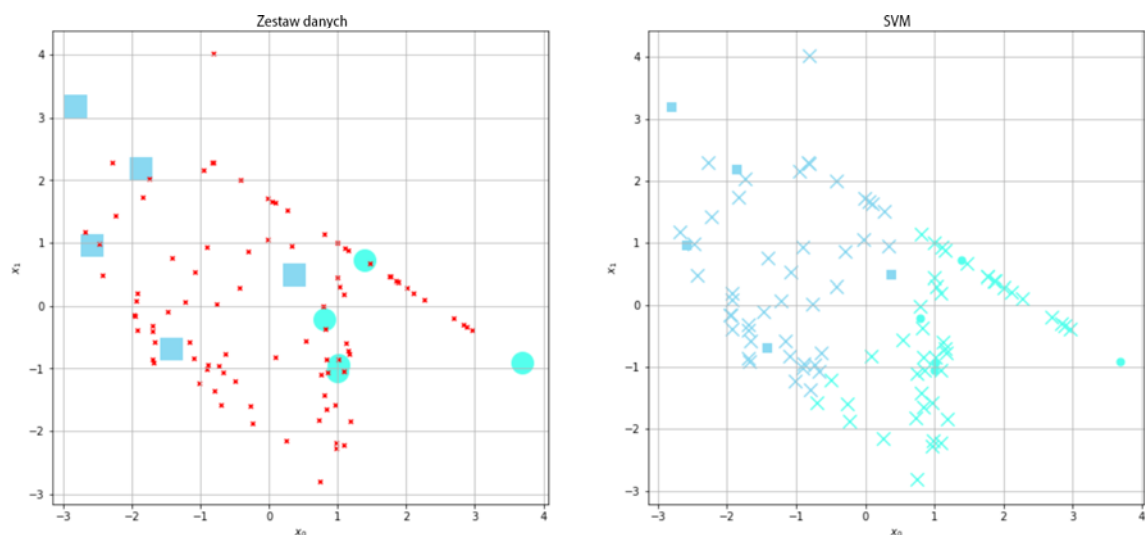


Rysunek 2.9. Pierwotny zestaw danych oznakowanych i nieoznakowanych

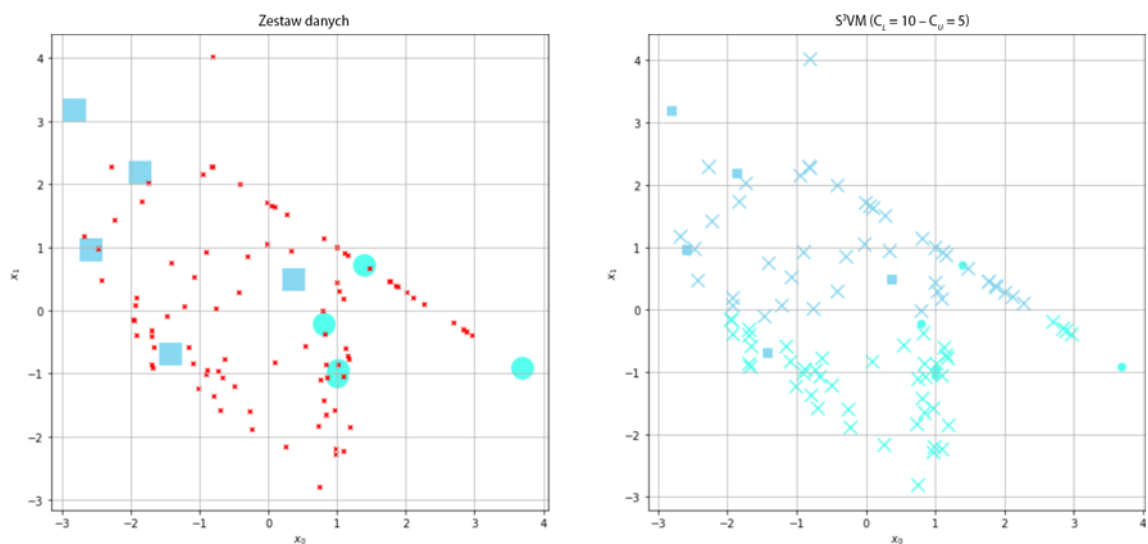


Rysunek 2.10. Pierwotny zestaw danych (po lewej) i ostateczny zestaw danych oznakowanych (po prawej)

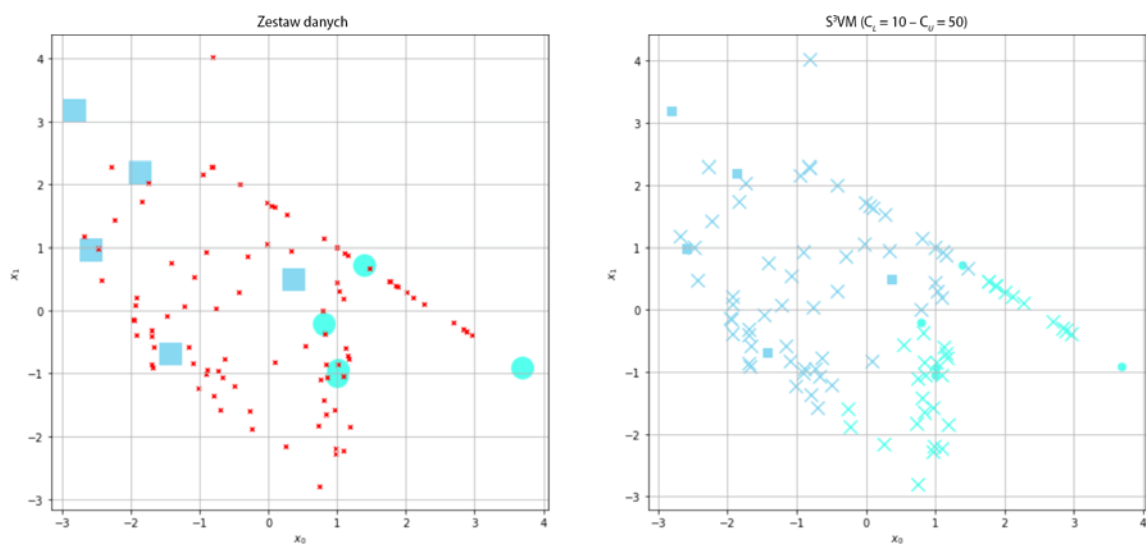




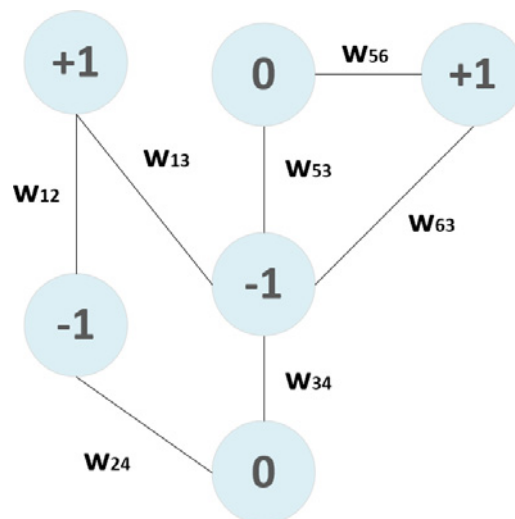
Rysunek 2.11. Pierwotny zestaw danych (po lewej) i ostateczny zestaw danych oznakowanych (po prawej) z parametrem  $C = 1,0$



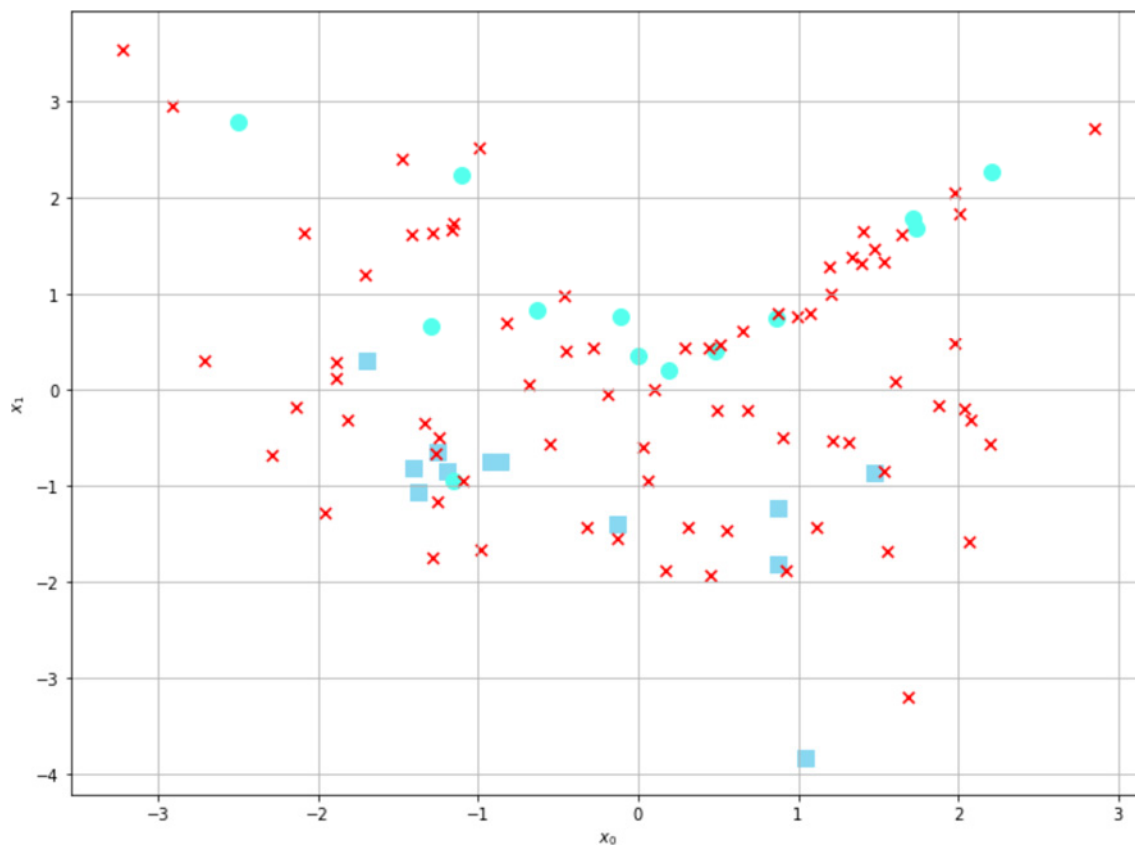
Rysunek 2.12. Pierwotny zestaw danych (po lewej) i ostateczny zestaw danych oznakowanych (po prawej) z parametrami  $C_l = 10$  i  $C_u = 5$



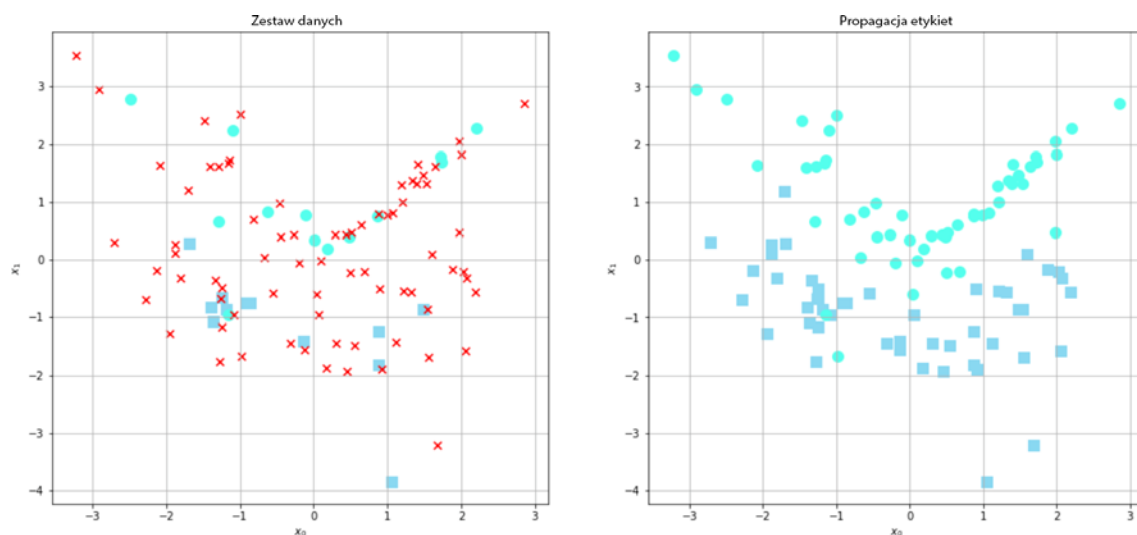
Rysunek 2.13. Pierwotny zestaw danych (po lewej) i ostateczny zestaw danych oznakowanych (po prawej) z parametrami  $C_l = 10$  i  $C_u = 50$



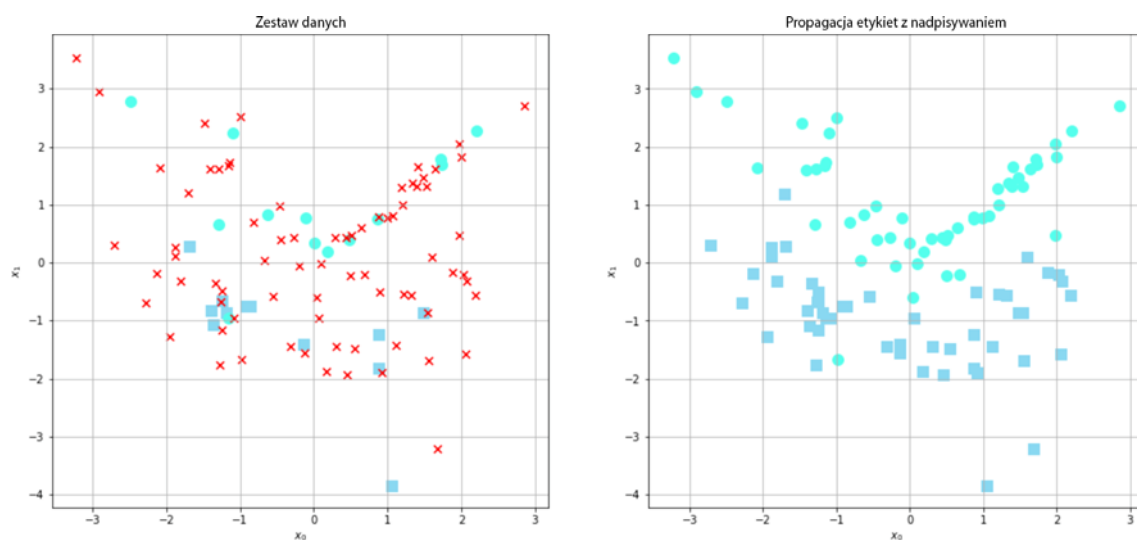
Rysunek 3.1. Przykładowy graf binarny



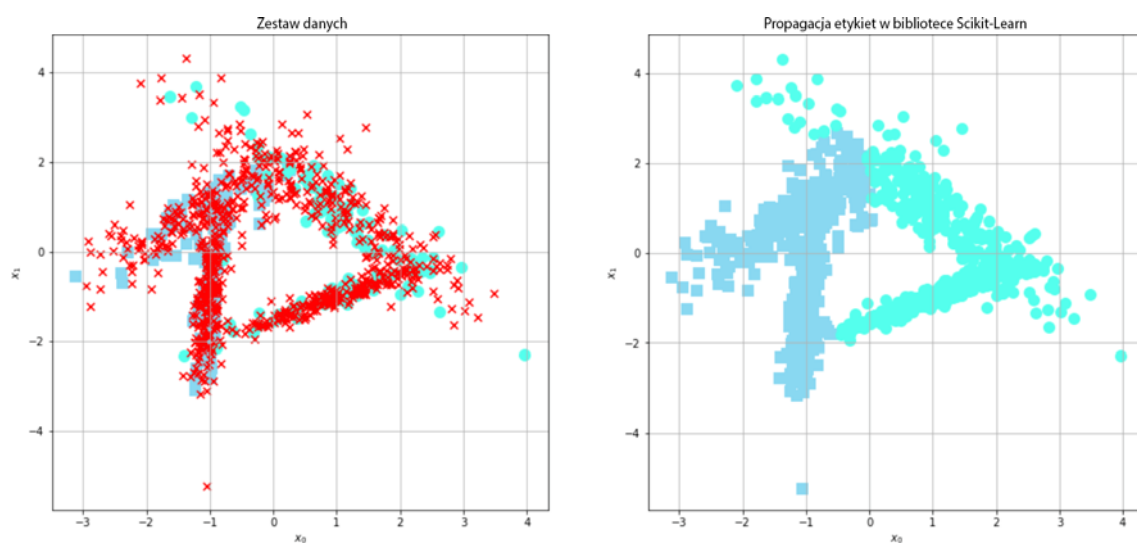
Rysunek 3.2. Zestaw częściowo oznakowanych danych



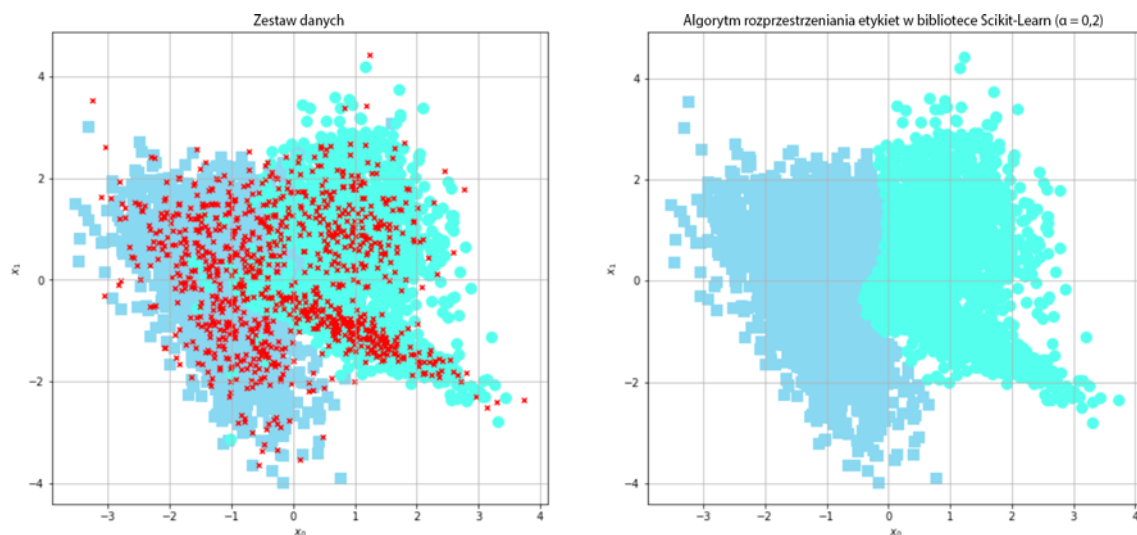
**Rysunek 3.3.** Pierwotny zestaw danych (po lewej); zestaw danych po zrealizowaniu algorytmu propagacji etykiet (po prawej)



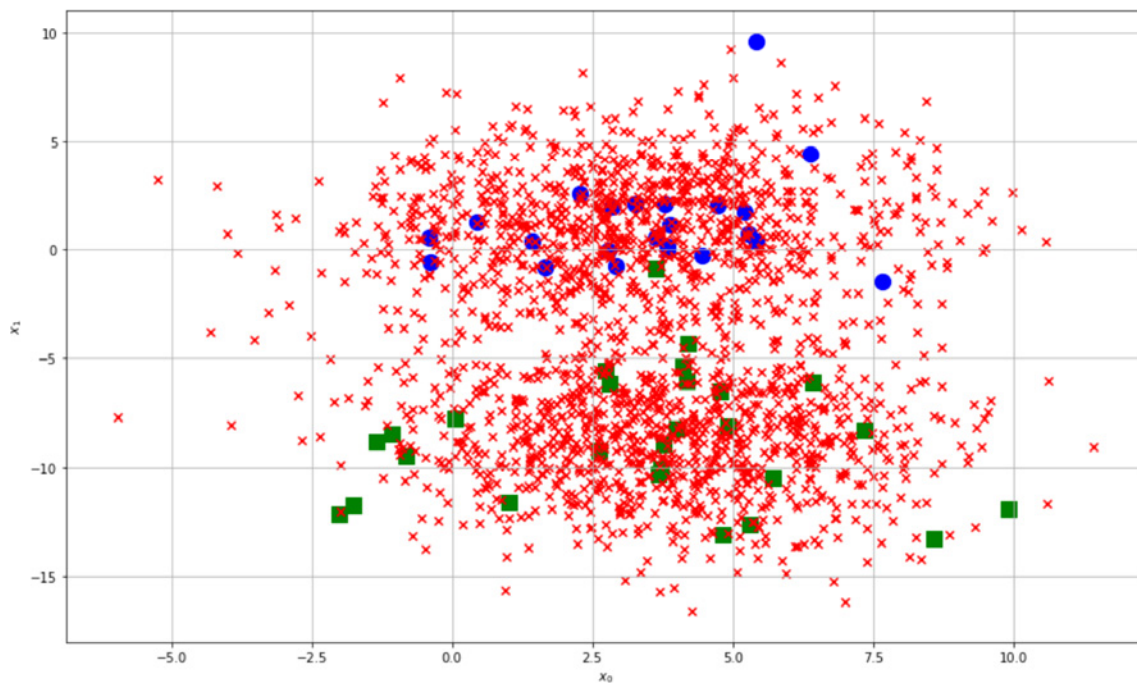
**Rysunek 3.4.** Pierwotny zestaw danych (po lewej); zestaw danych po zrealizowaniu algorytmu propagacji etykiet z nadpisywaniem (po prawej)



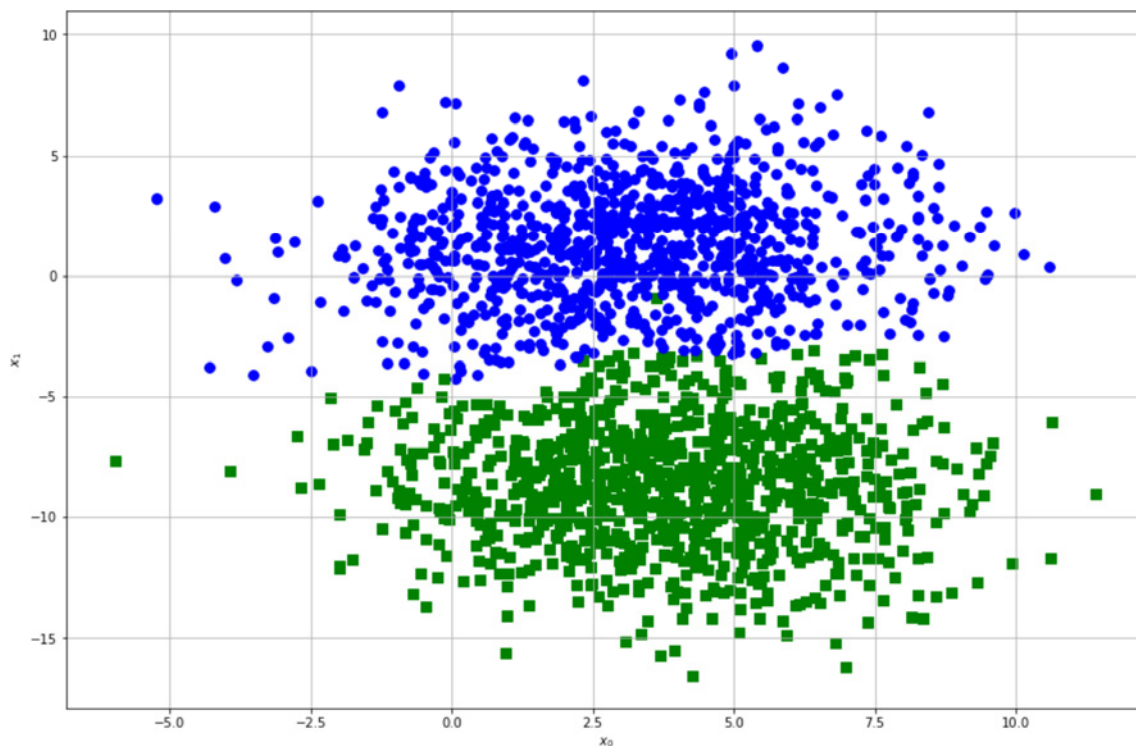
**Rysunek 3.5.** Pierwotny zestaw danych (po lewej); zestaw danych po zrealizowaniu algorytmu propagacji etykiet w bibliotece Scikit-Learn (po prawej)



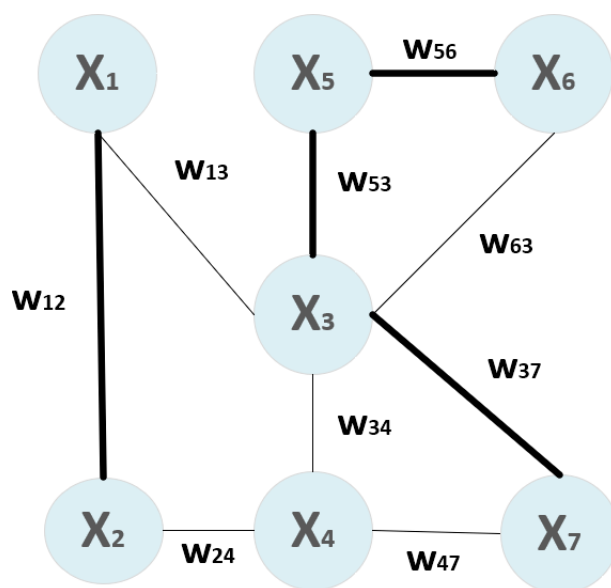
Rysunek 3.6. Pierwotny zestaw danych (po lewej); zestaw danych po pełnej realizacji algorytmu rozprzestrzeniania etykiet (po prawej)



Rysunek 3.7. Zestaw częściowo oznakowanych danych



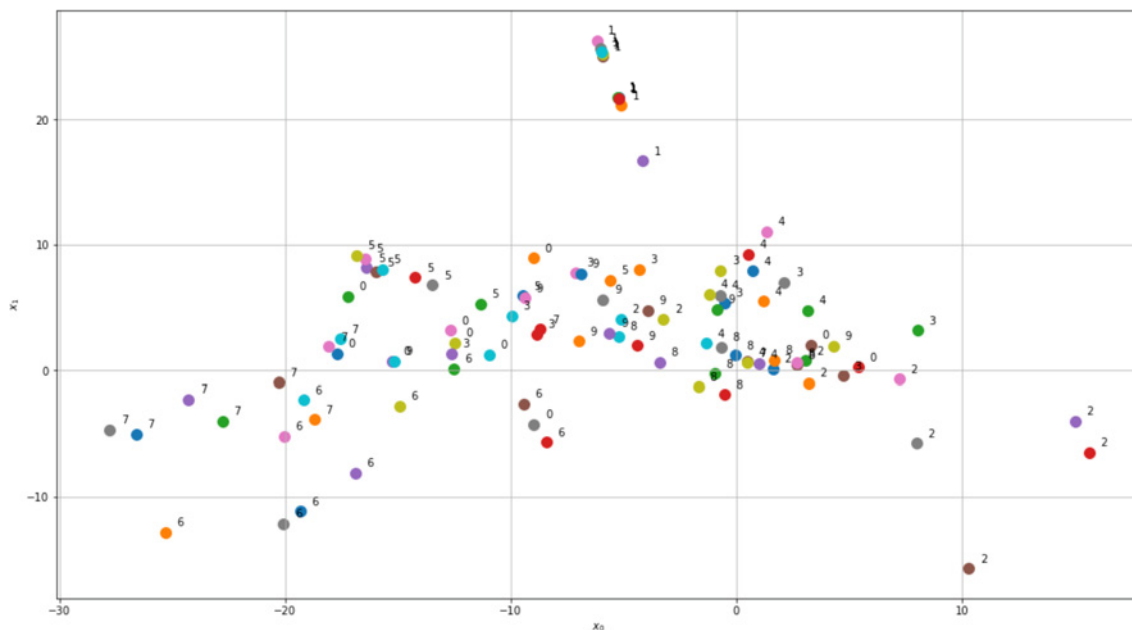
Rysunek 3.8. Zestaw danych po pełnej realizacji algorytmu propagacji etykiet przy użyciu błędzenia losowego Markowa



Rysunek 3.9. Przykład grafu zawierającego zaznaczone najkrótsze odległości pomiędzy parami przykładów



Rysunek 3.10. Podzbiór zestawu danych Olivetti



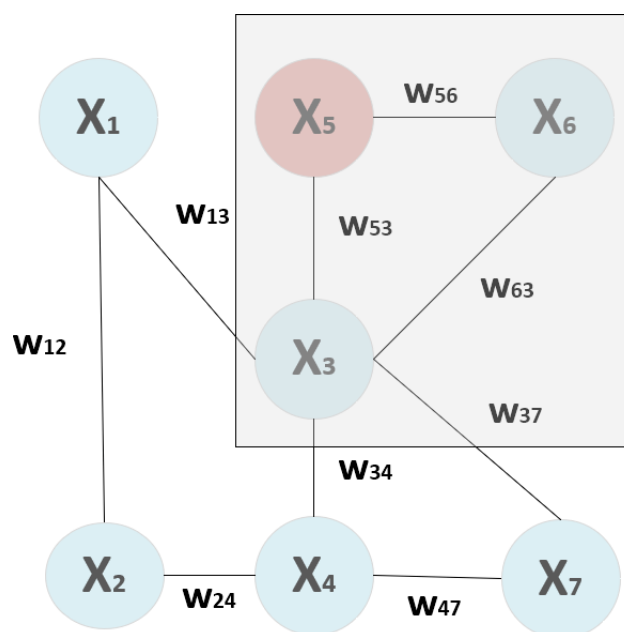
Rysunek 3.11. Wynik uczenia algorytmu Isomap na stu przykładach z zestawu danych Olivetti



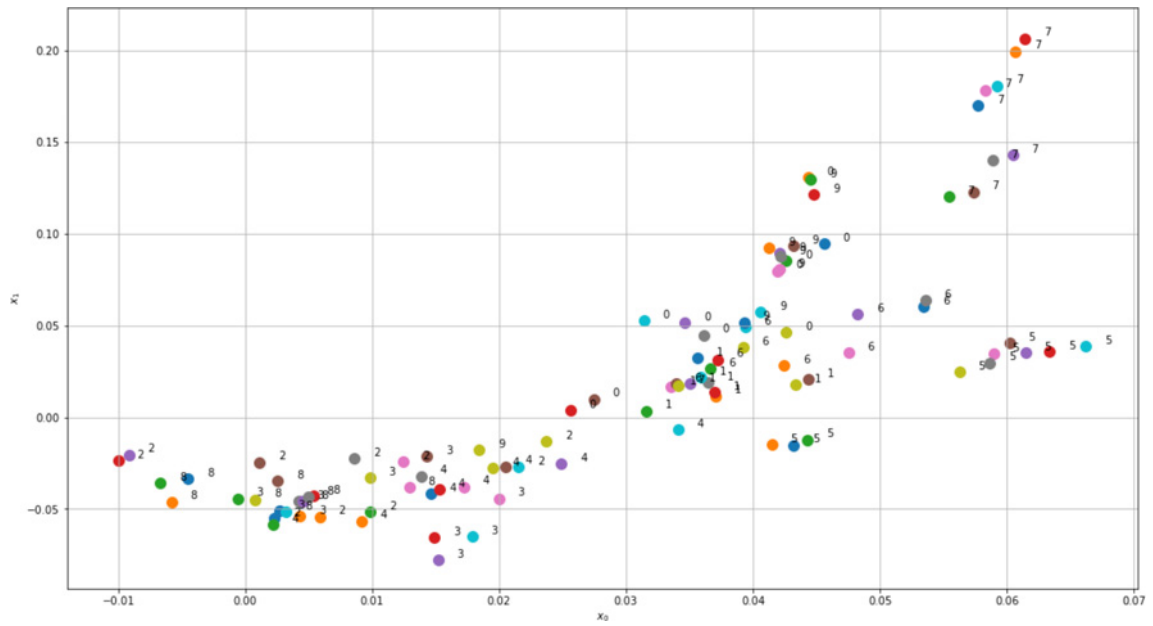
Rysunek 3.12. Przykłady należące do klasy 7.



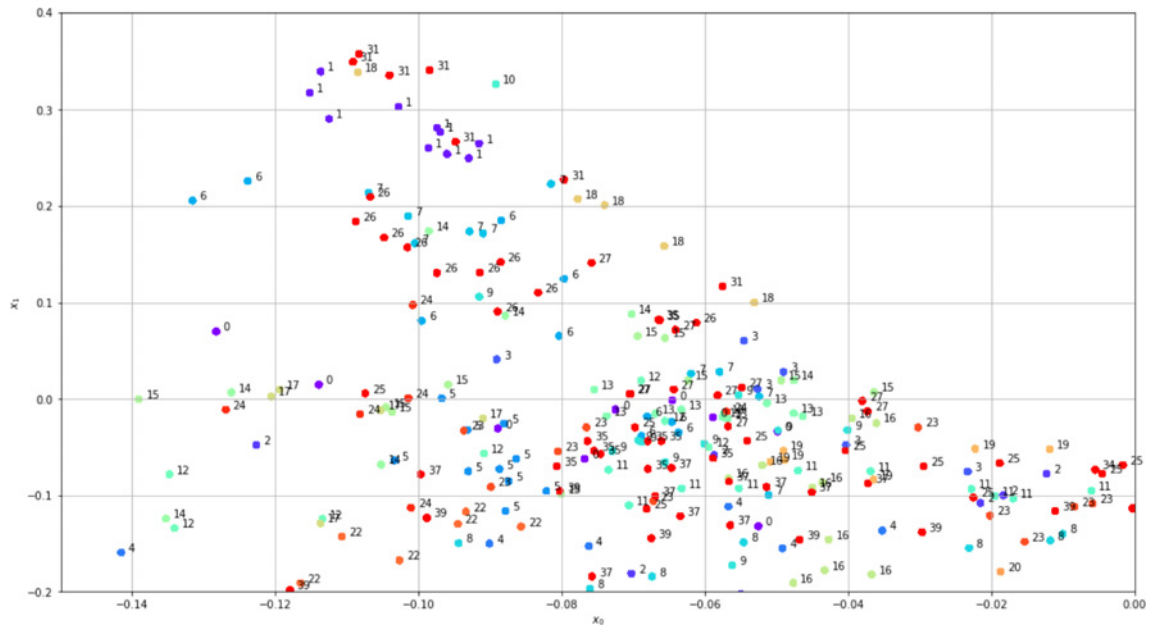
Rysunek 3.13. Przykłady należące do klasy 1.



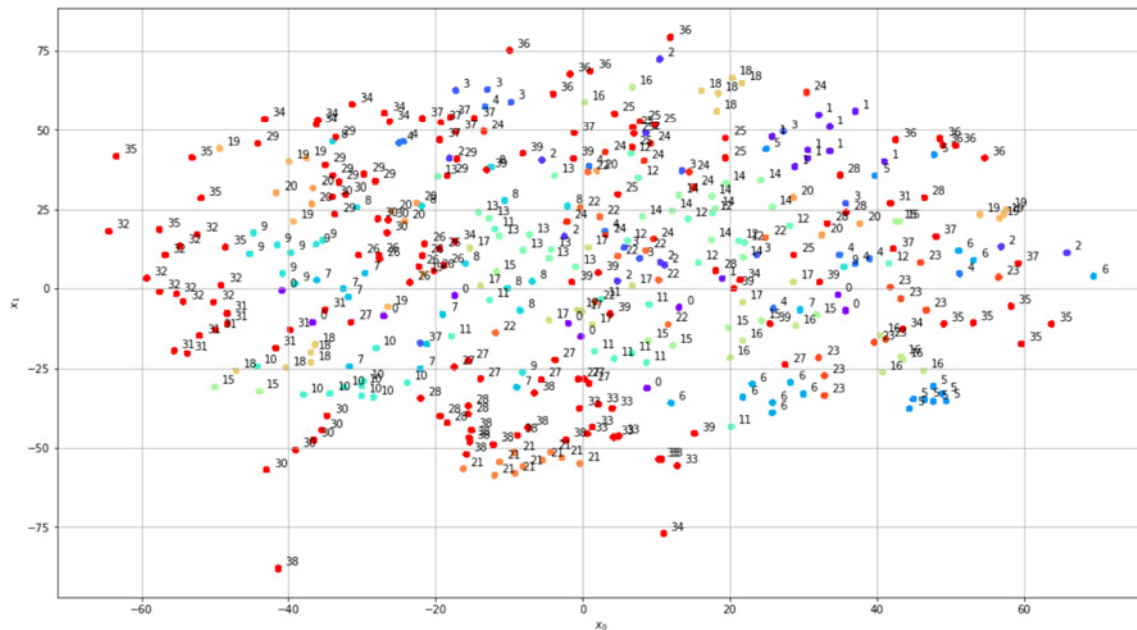
Rysunek 3.14. Graf, w którym sąsiedztwo zostało oznaczone prostokątem



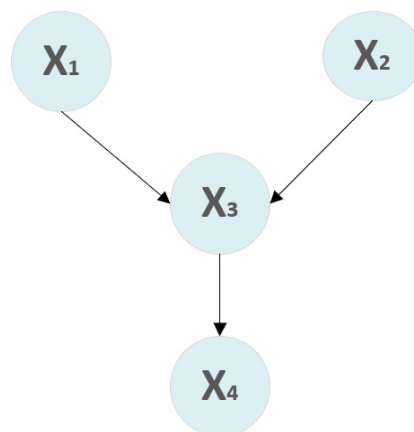
Rysunek 3.15. Osadzanie lokalnie liniowe zastosowane wobec 100 przykładów stanowiących część zestawu danych Olivetti



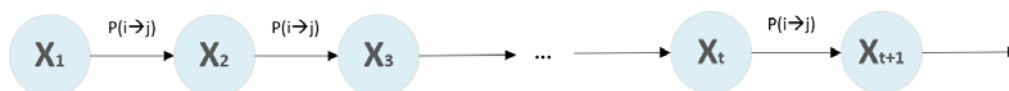
Rysunek 3.16. Osadzanie widmowe Laplace'a użyte na zestawie danych Olivetti



Rysunek 3.17. Algorytm t-SNE użyty wobec zestawu danych Olivetti

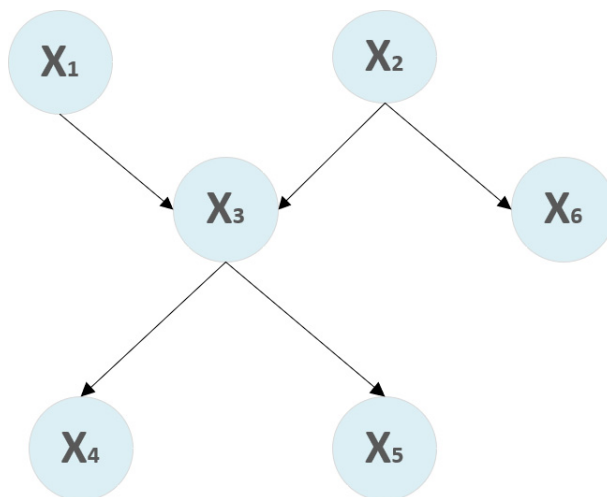


Rysunek 4.1. Przykładowa sieć bayesowska

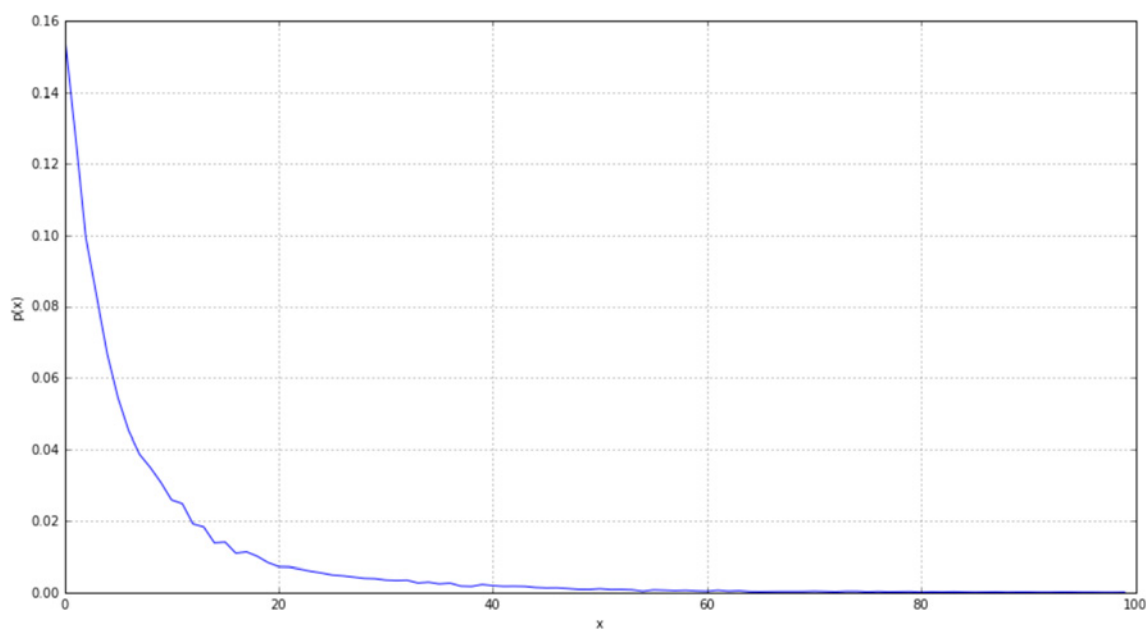


Rysunek 4.2. Struktura ogólnego łańcucha Markowa

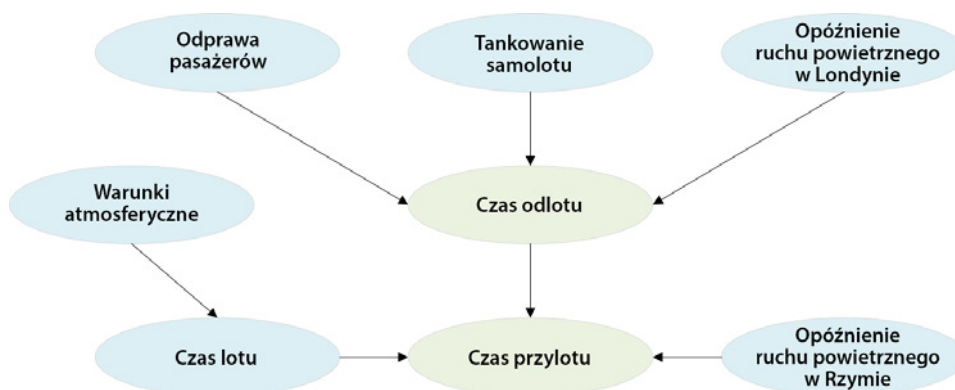




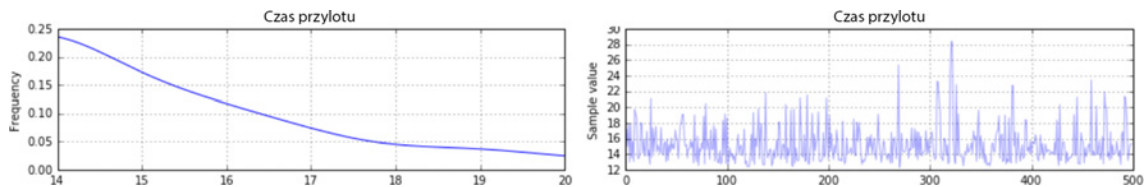
Rysunek 4.3. Sieć bayesowska służąca do wyjaśnienia próbkowania Gibbsa



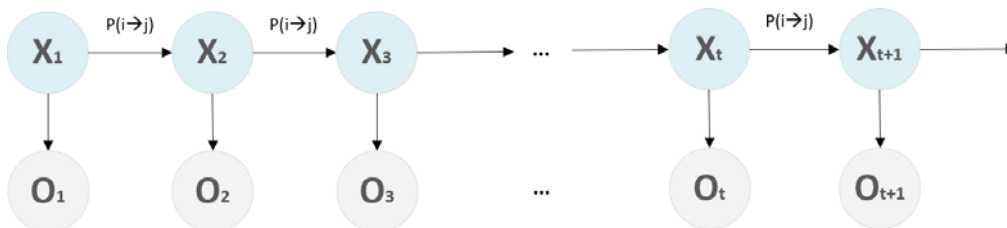
Rysunek 4.4. Próbkowana funkcja gęstości prawdopodobieństwa



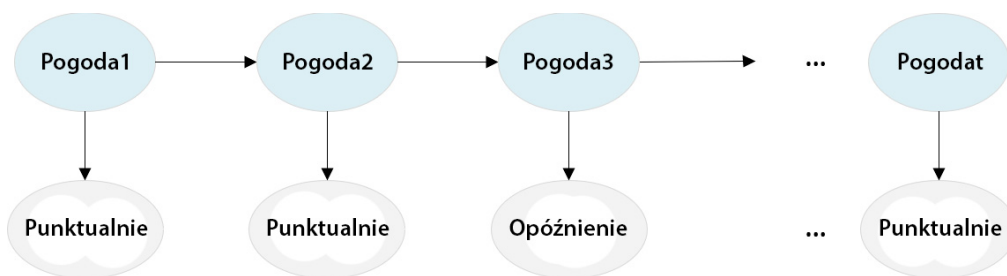
Rysunek 4.5. Sieć bayesowska reprezentująca problem kontroli ruchu powietrznego



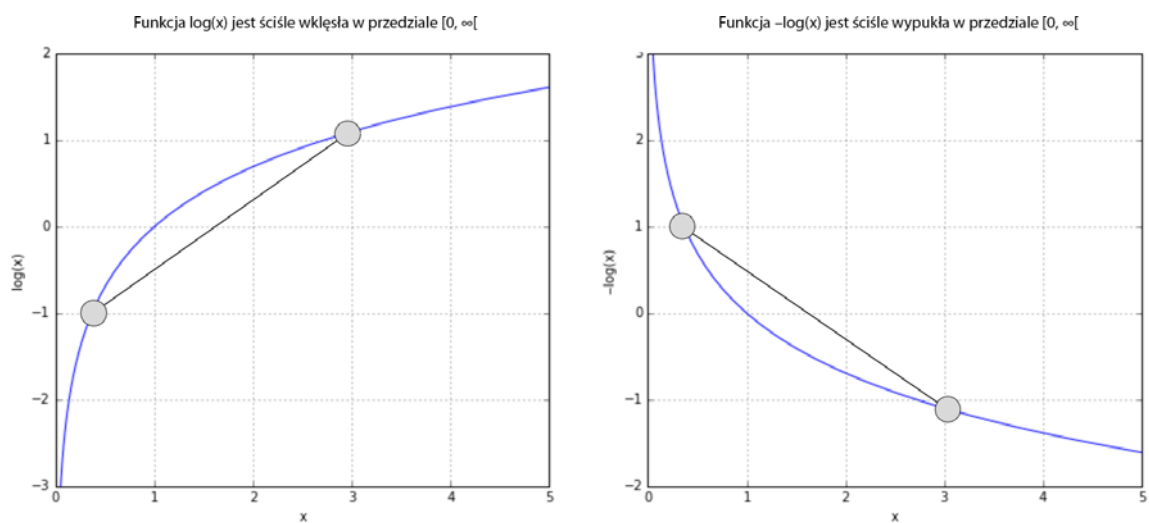
Rysunek 4.6. Rozkład i przykłady uzyskane dla zmiennej losowej Czas przylotu



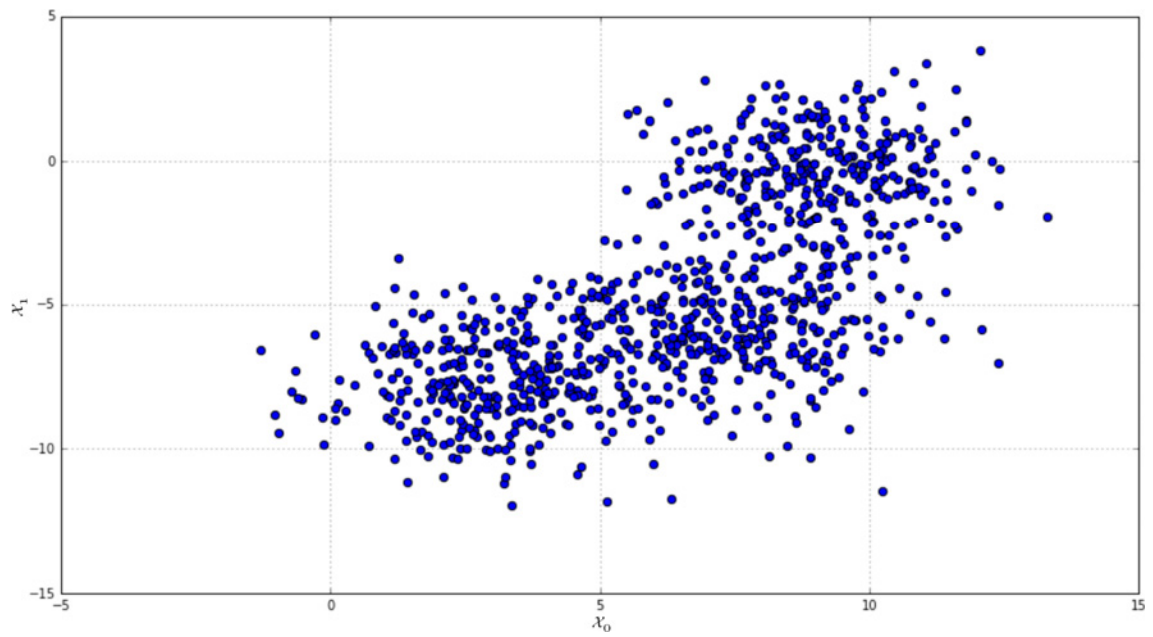
Rysunek 4.7. Schemat ogólny ukrytego modelu Markowa



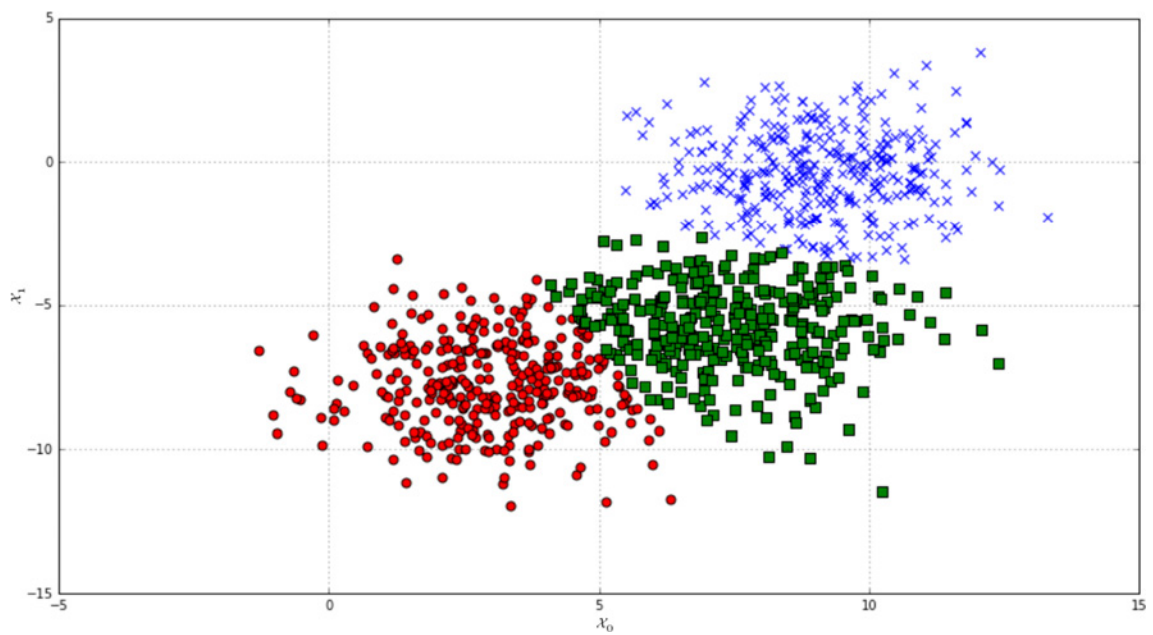
Rysunek 4.8. Model HMM zastosowany dla problemu opóźnień lotu, wynikających z warunków pogodowych



Rysunek 5.1. Przykład funkcji wypukłej i wklęsłej



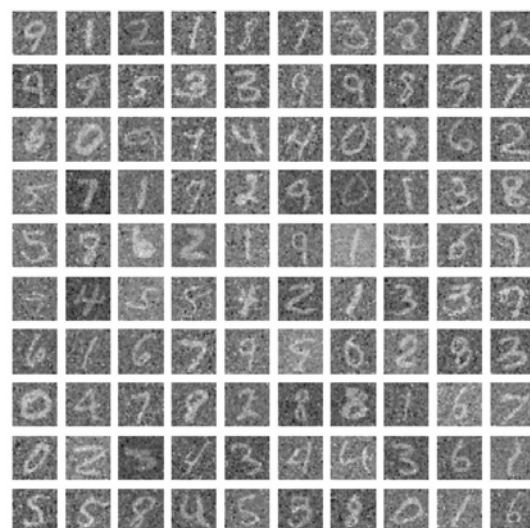
Rysunek 5.2. Przykładowy zestaw danych użyty w implementacji algorytmu mieszanin gaussowskich



Rysunek 5.3. Oznakowany zestaw danych otrzymany poprzez zastosowanie mieszaniny gaussowskiej zawierającej trzy składowe

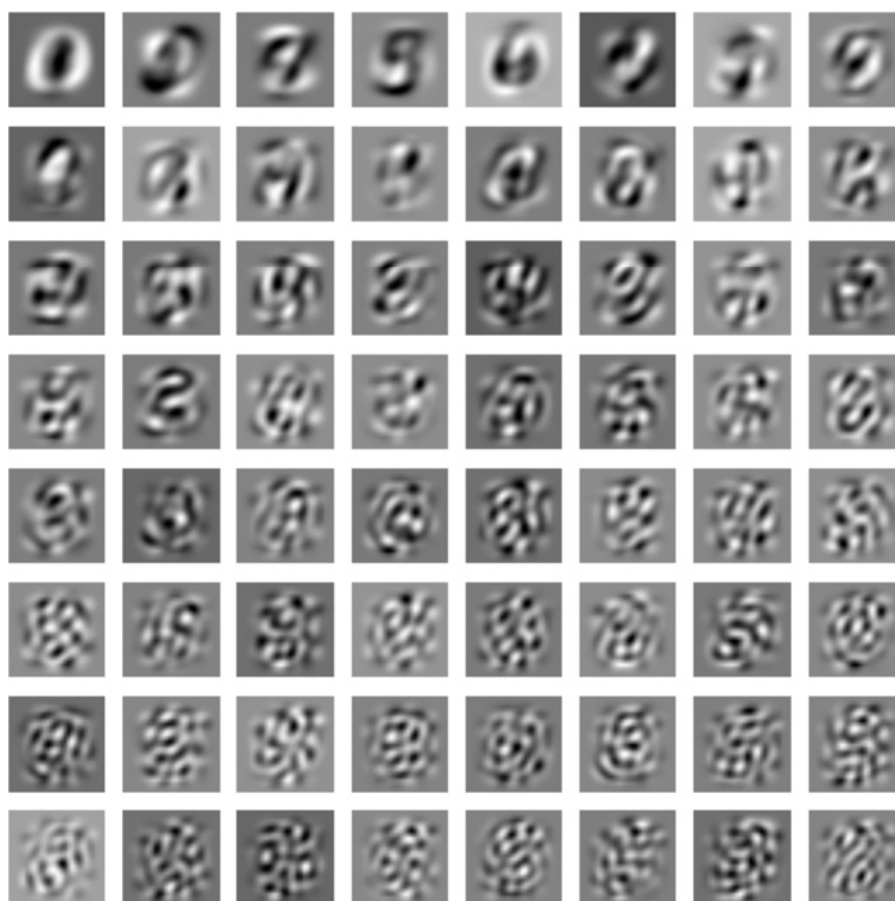


Wersja pierwotna

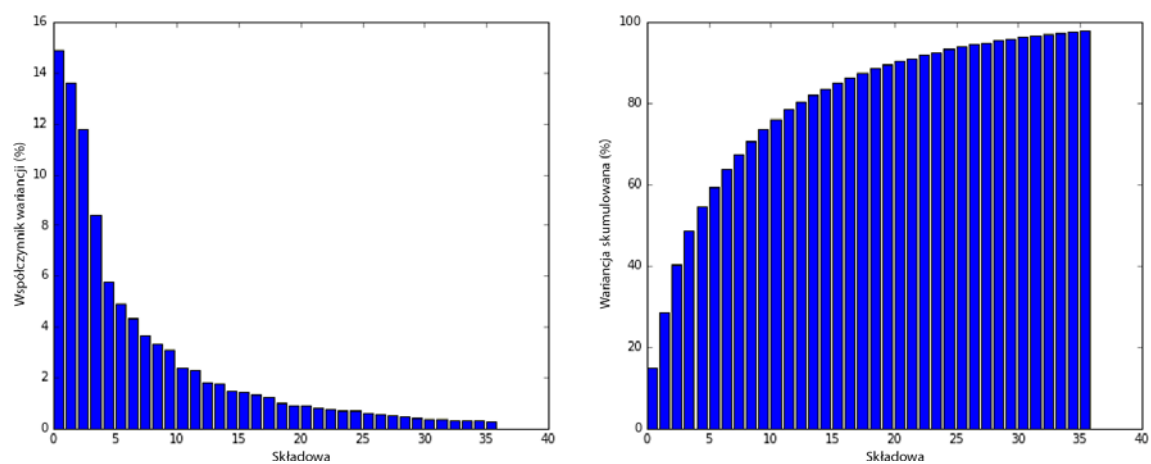


Szum heteroskedastyczny ( $w_i \in [0, 0,75]$ )

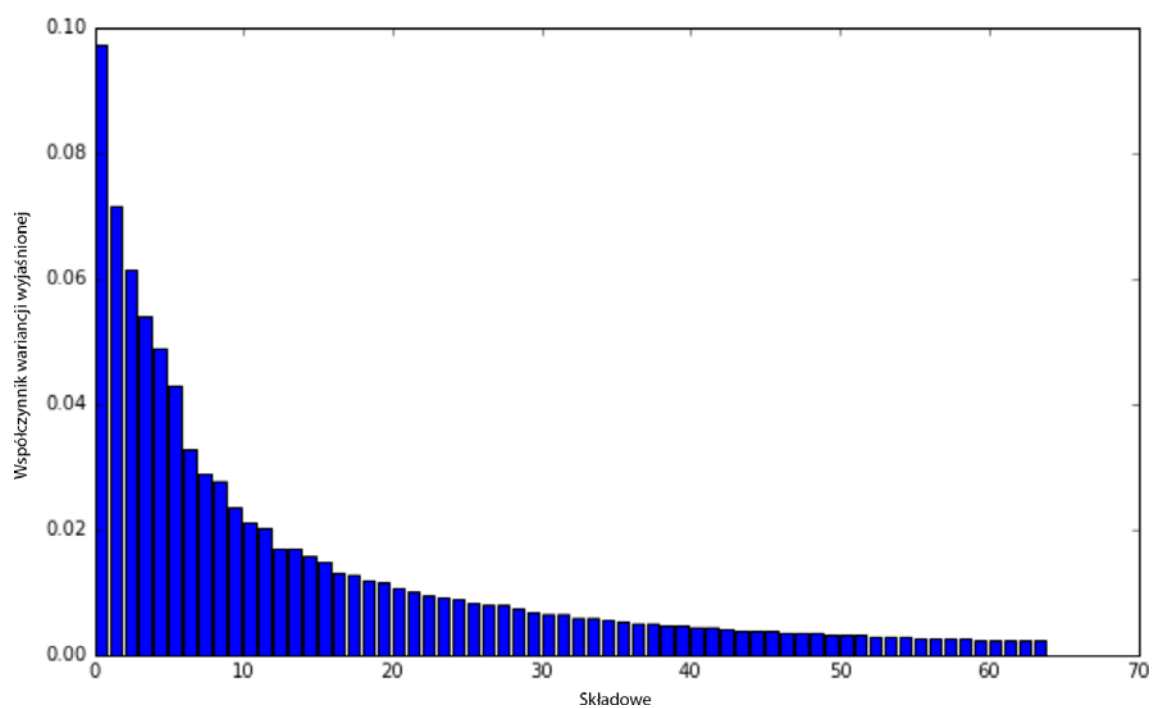
Rysunek 5.4. Pierwotny zestaw danych (po lewej) i jego zaszumiona wersja (po prawej)



Rysunek 5.5. Wykres 64 składowych wydobytych z pierwotnego zestawu danych za pomocą analizy czynnikowej



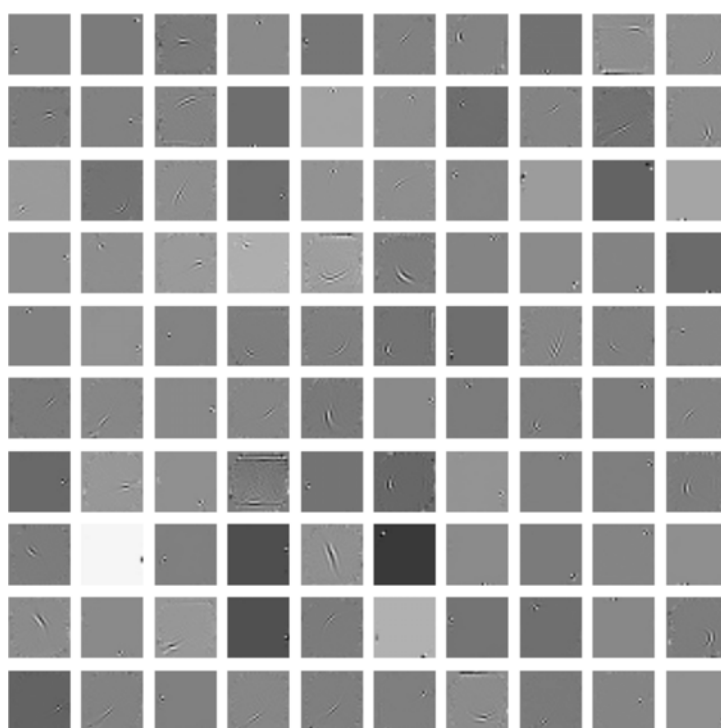
Rysunek 5.6. Wariancja wyjaśniona na każdą składową (po lewej) i wariancja skumulowana na każdą składową (po prawej)



Rysunek 5.7. Całkowita wariancja wyjaśniona uzyskana za pomocą analizy PCA



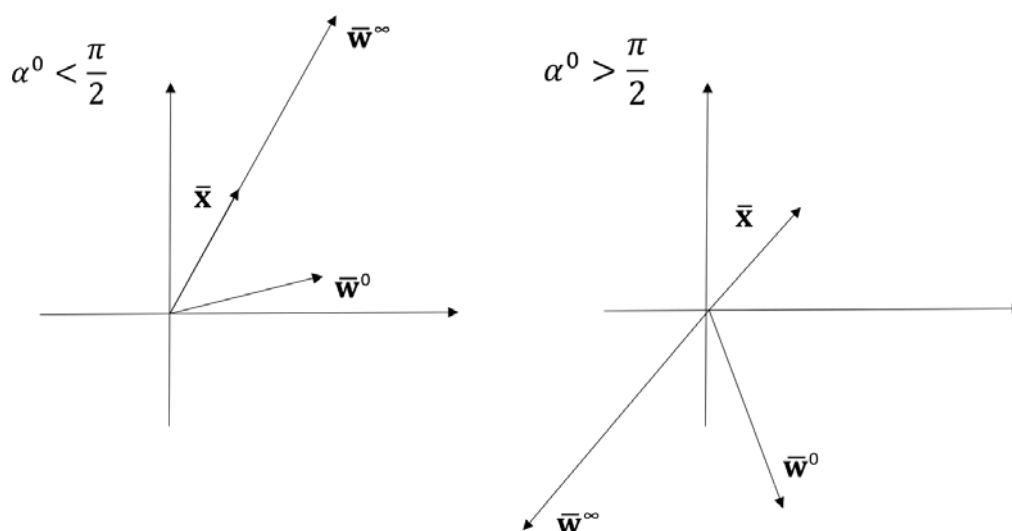
Rysunek 5.8. Składowe niezależne zestawu danych MNIST wydobyte za pomocą algorytmu FastICA (64 składowe)



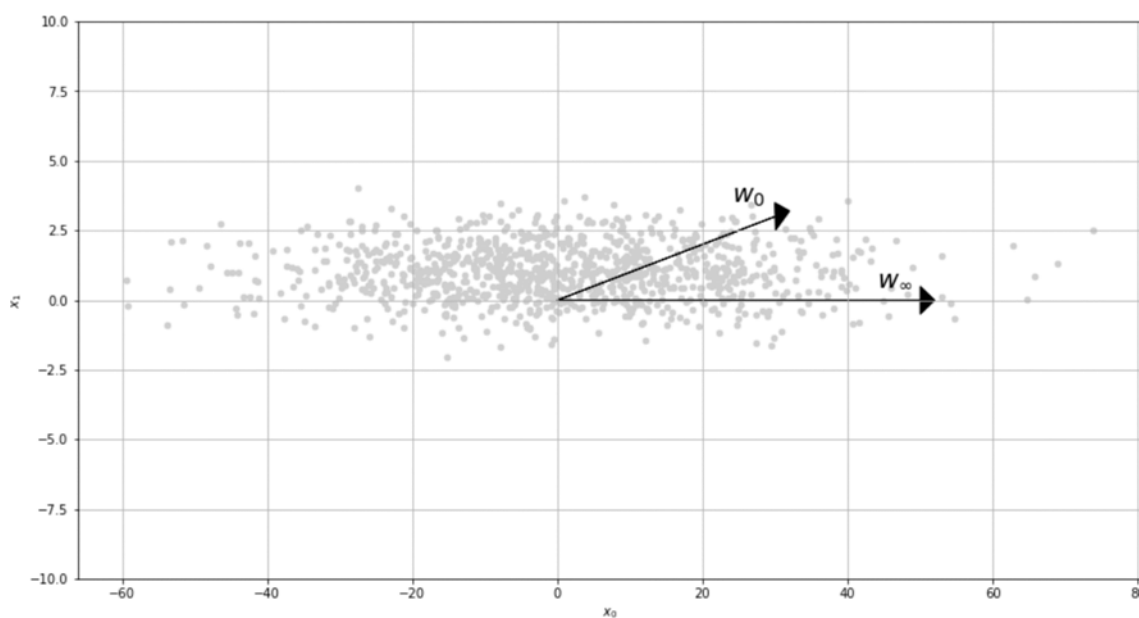
Rysunek 5.9. Składowe niezależne zestawu danych MNIST wydobyte za pomocą algorytmu FastICA (640 składowych)



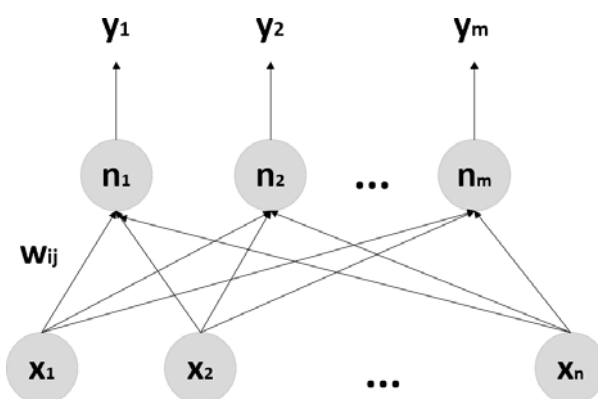
Rysunek 6.1. Ogólny schemat neuronu hebbowskiego z wejściem w postaci wektorowej



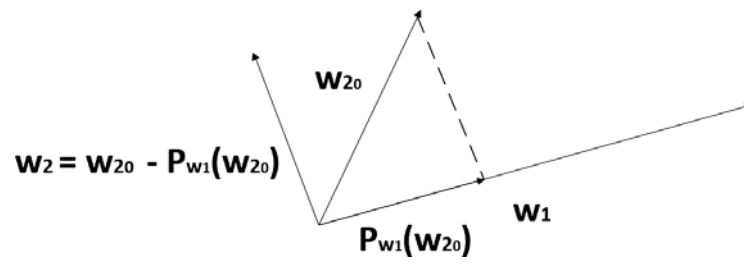
Rysunek 6.2. Analiza wektorowa reguły Hebba



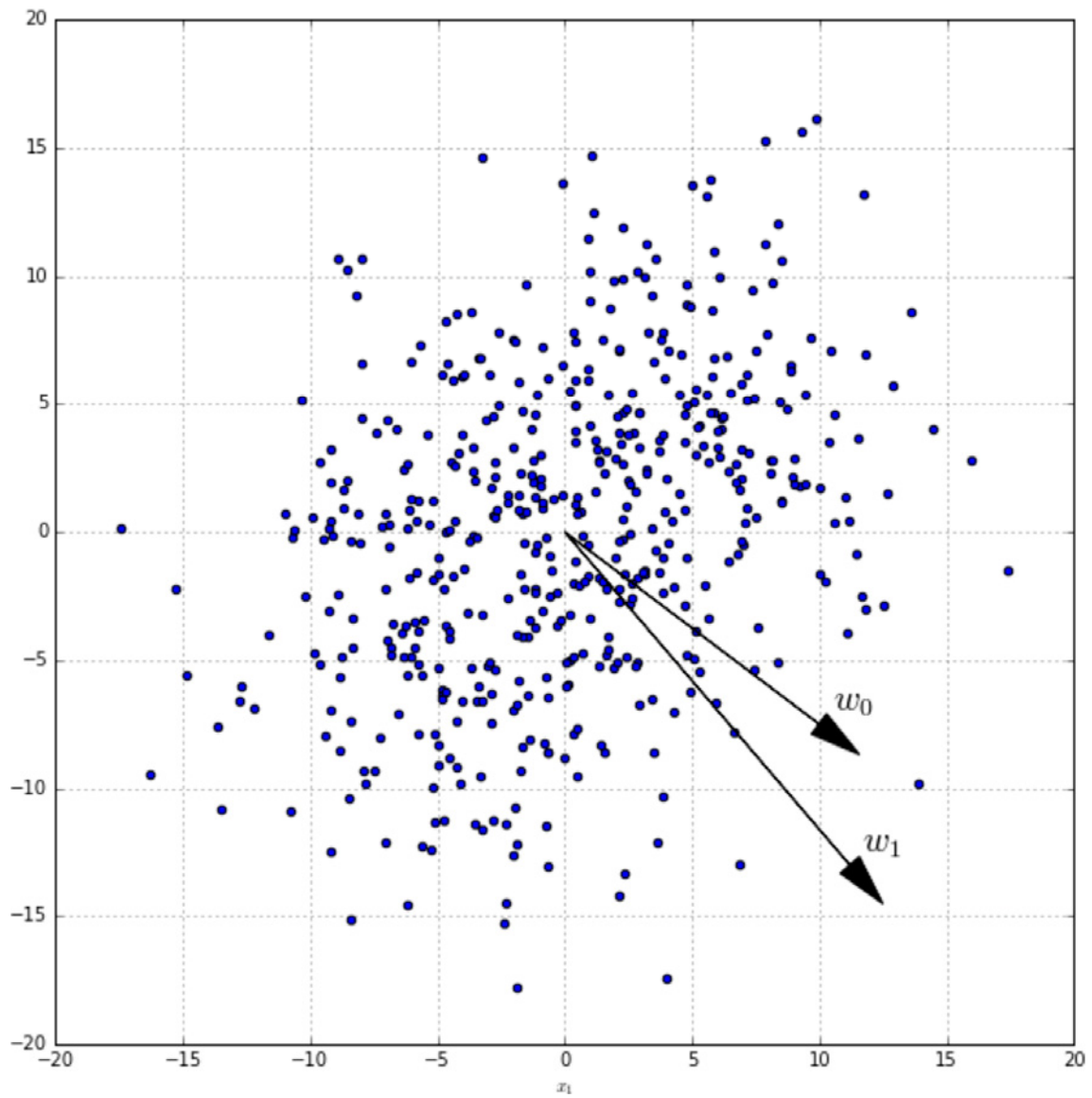
Rysunek 6.3. Zastosowanie reguły kowariancji. Wektor  $w_\infty$  jest wprost proporcjonalny do pierwszej głównej składowej



Rysunek 6.4. Schemat sieci Sangera

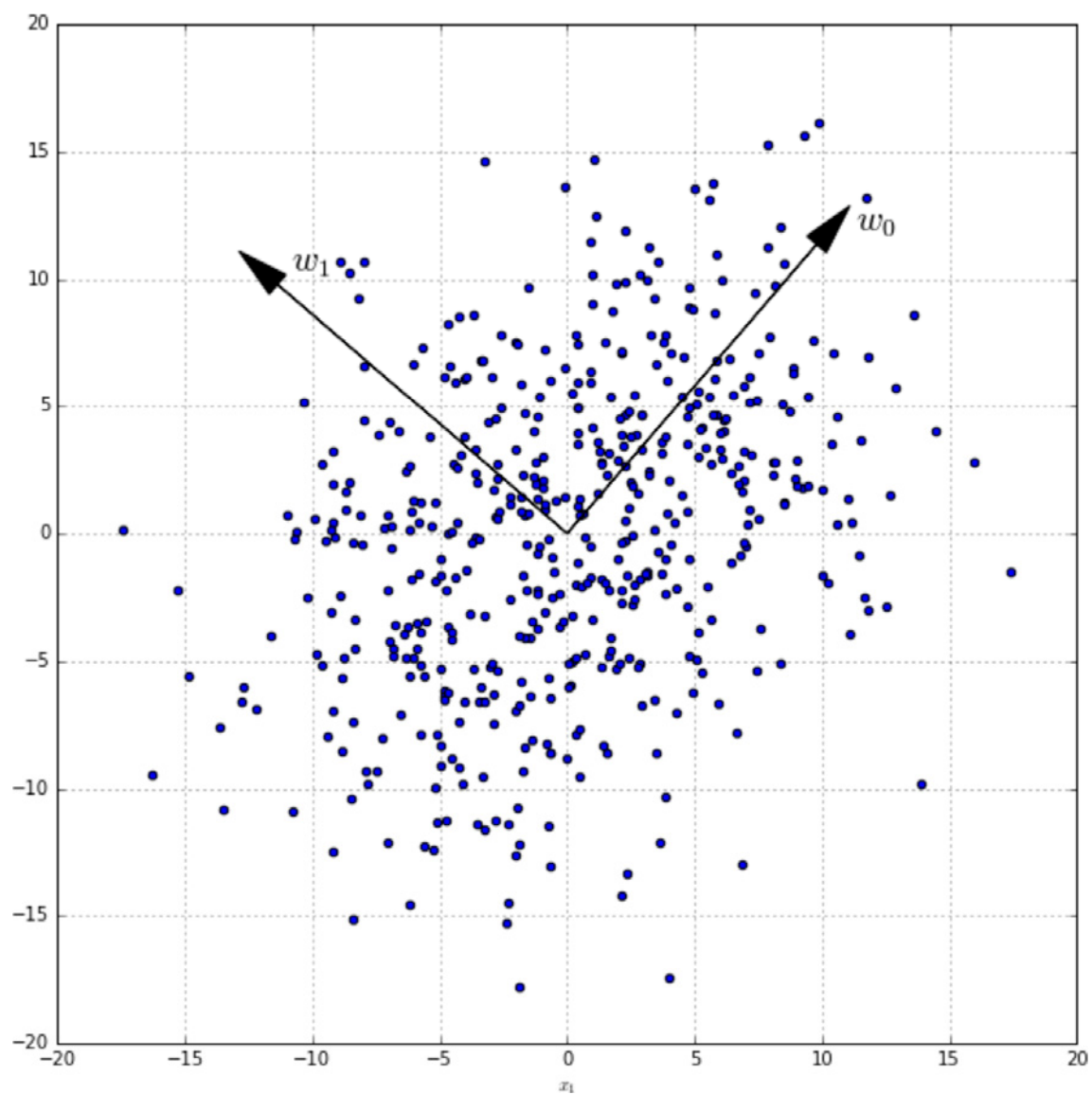


Rysunek 6.5. Ortogonalizacja dwóch wektorów wag

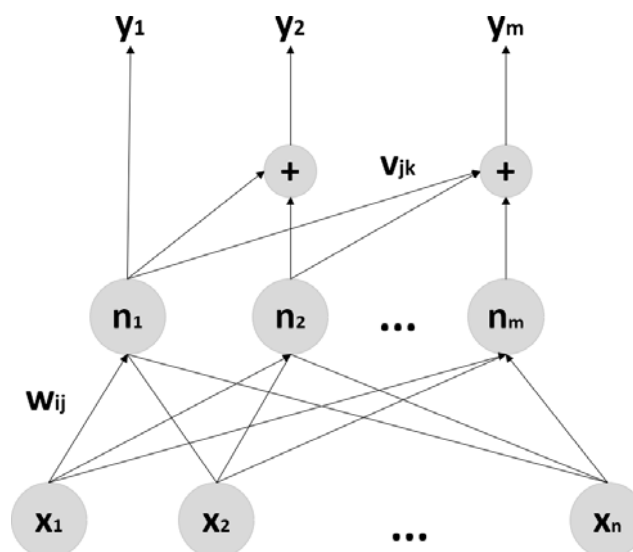


Rysunek 6.6. Zestaw danych wraz z warunkiem początkowym W

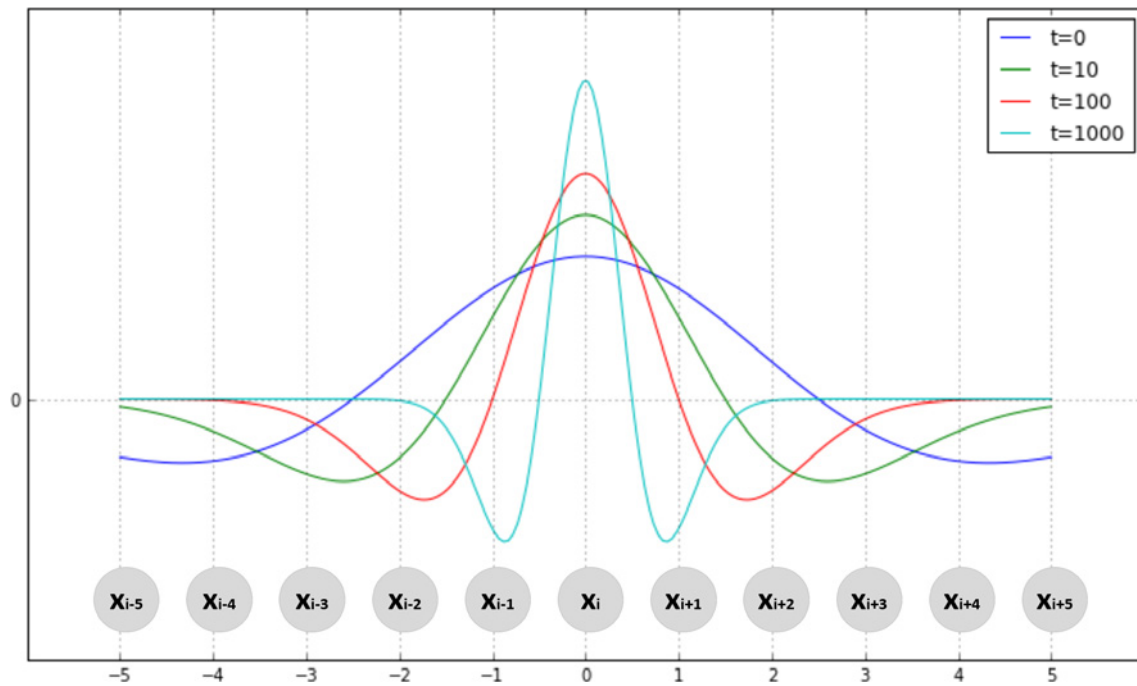




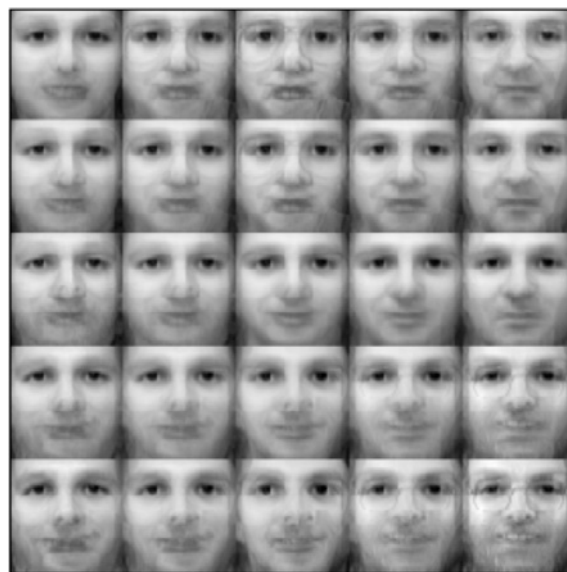
Rysunek 6.7. Stan końcowy modelu; wektor uzyskał zbieżność z dwiema głównymi składowymi



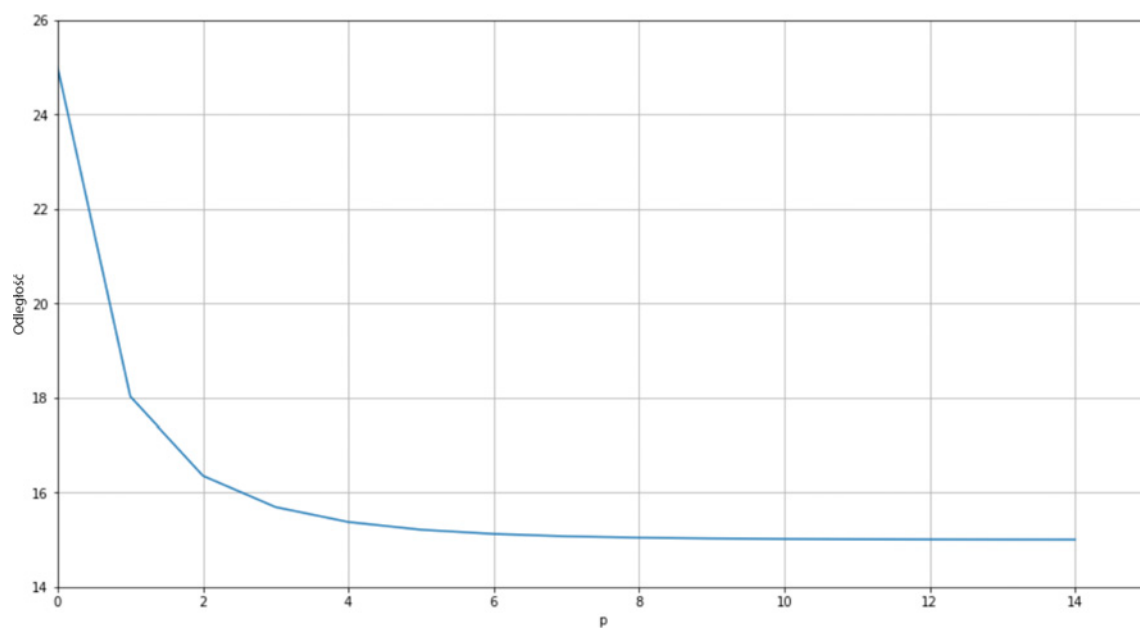
Rysunek 6.8. Sieć Rubnera-Tavana. Połączenia  $v$ , bazują na przeciwieństwie reguły Hebba



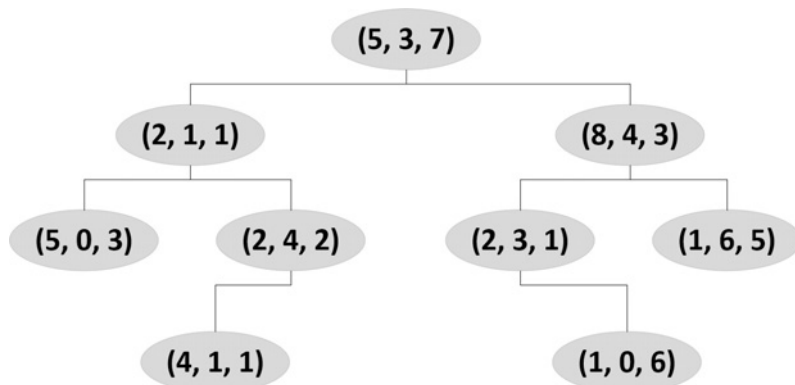
Rysunek 6.9. Przebieg funkcji w postaci „kapelusza meksykańskiego” zaimplementowany w sieci Kohonena



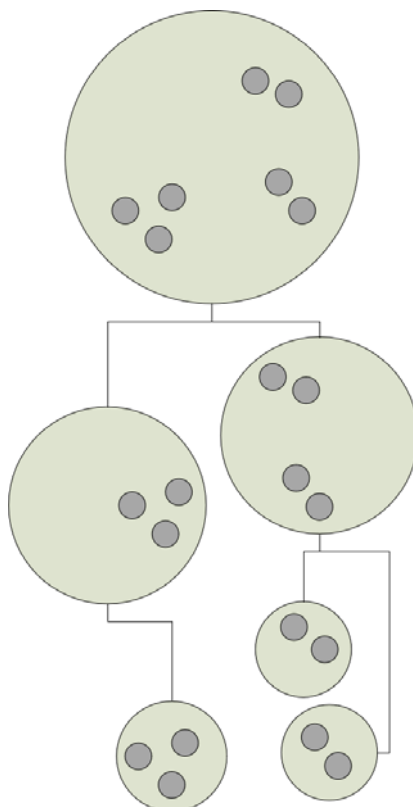
Rysunek 6.10. Rezultaty uzyskane za pomocą algorytmu mapy samoorganizującej



Rysunek 7.1. Odległość Minkowskiego pomiędzy współrzędnymi (0, 0) i (15, 10) jako funkcja parametru p



Rysunek 7.2. Przykład trójwymiarowego drzewa KD



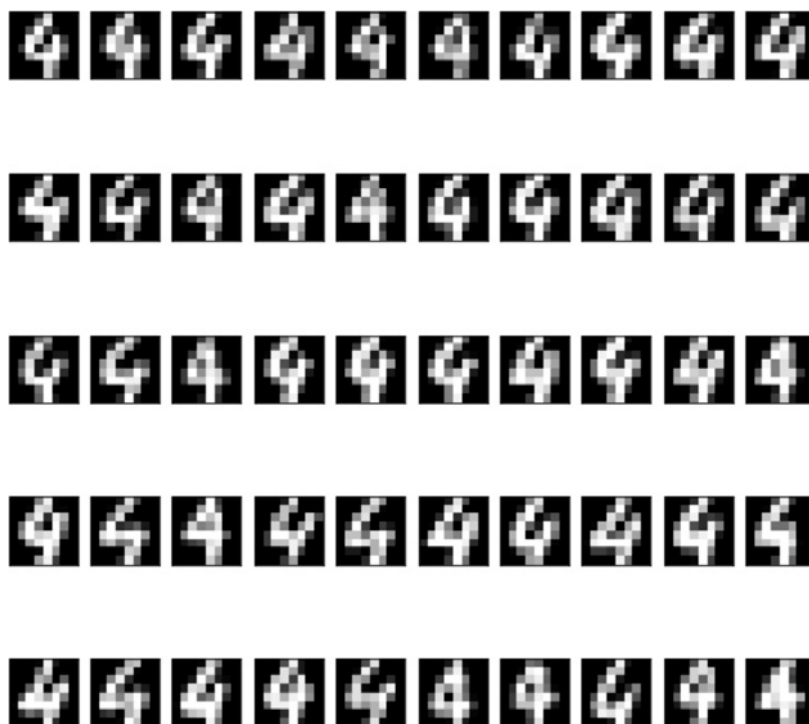
---

**Rysunek 7.3.** Przykład drzewa kulistego zawierającego siedem dwuwymiarowych punktów i dwa poziomy

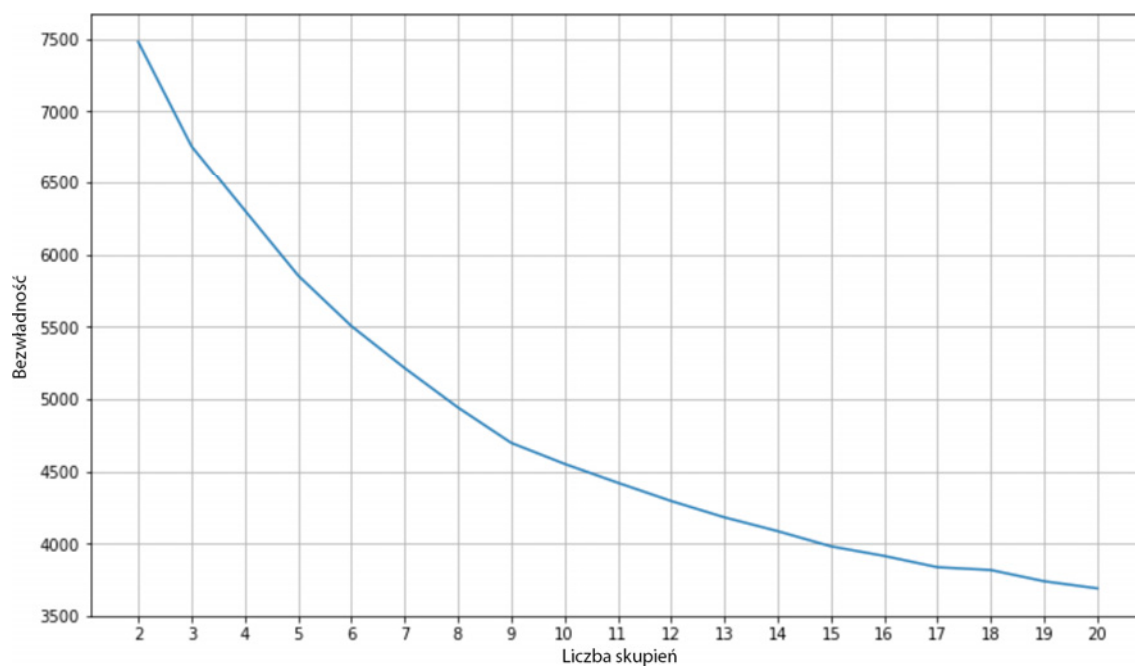


---

**Rysunek 7.4.** Przykładowa cyfra użyta do przepytania modelu KNN



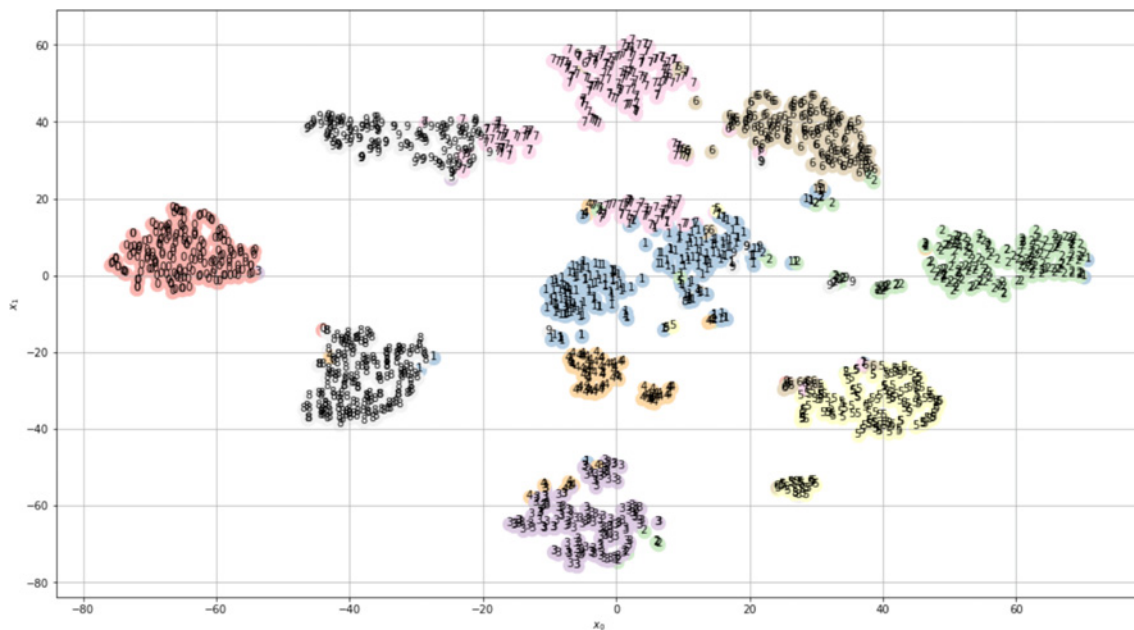
Rysunek 7.5. 50 sąsiadów wybranych przez model KNN



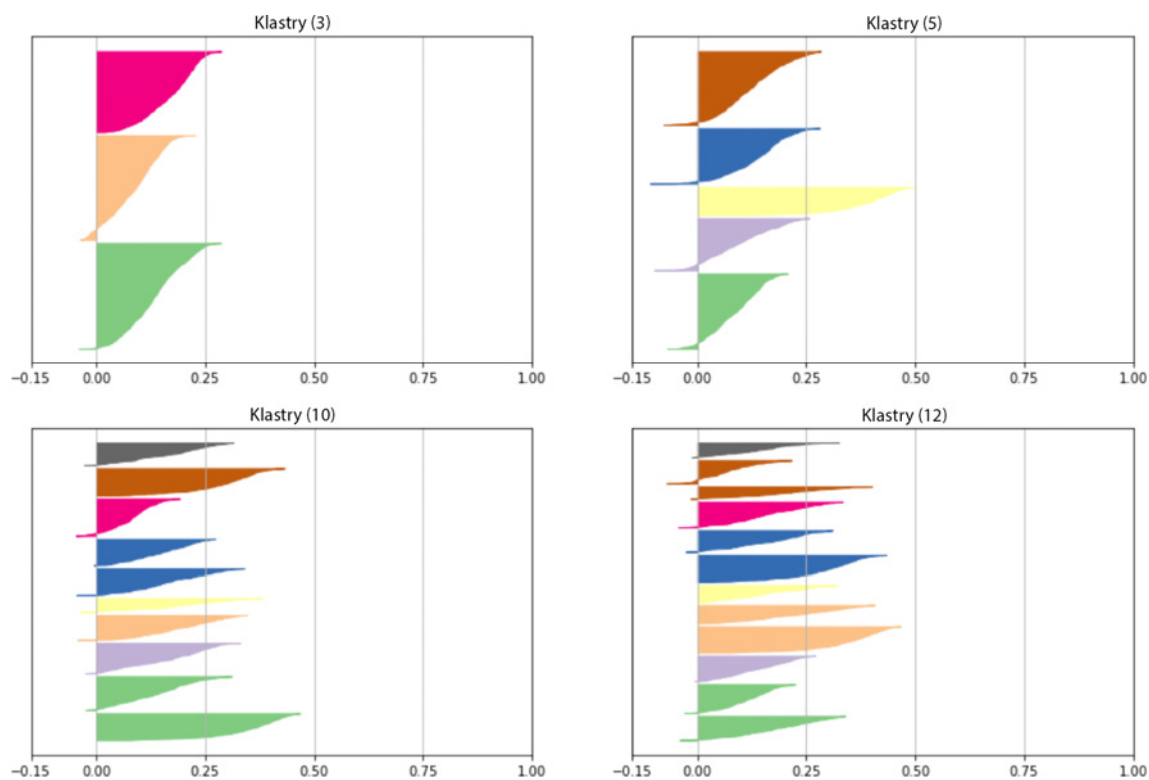
Rysunek 7.6. Bezwładność jako funkcja liczby skupień



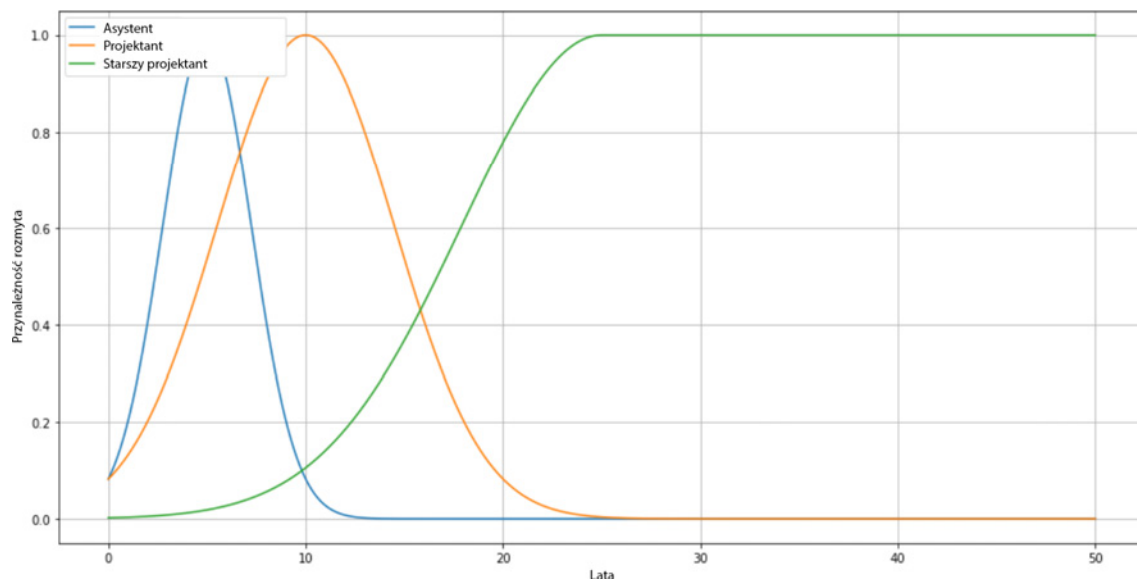
Rysunek 7.7. Centroidy w algorytmie k-średnich na końcu procesu uczenia



Rysunek 7.8. Zestaw danych MNIST przetworzony za pomocą algorytmu t-SNE; etykiety symbolizują poszczególne skupienia



Rysunek 7.9. Wykresy profilu dla różnych liczb skupień



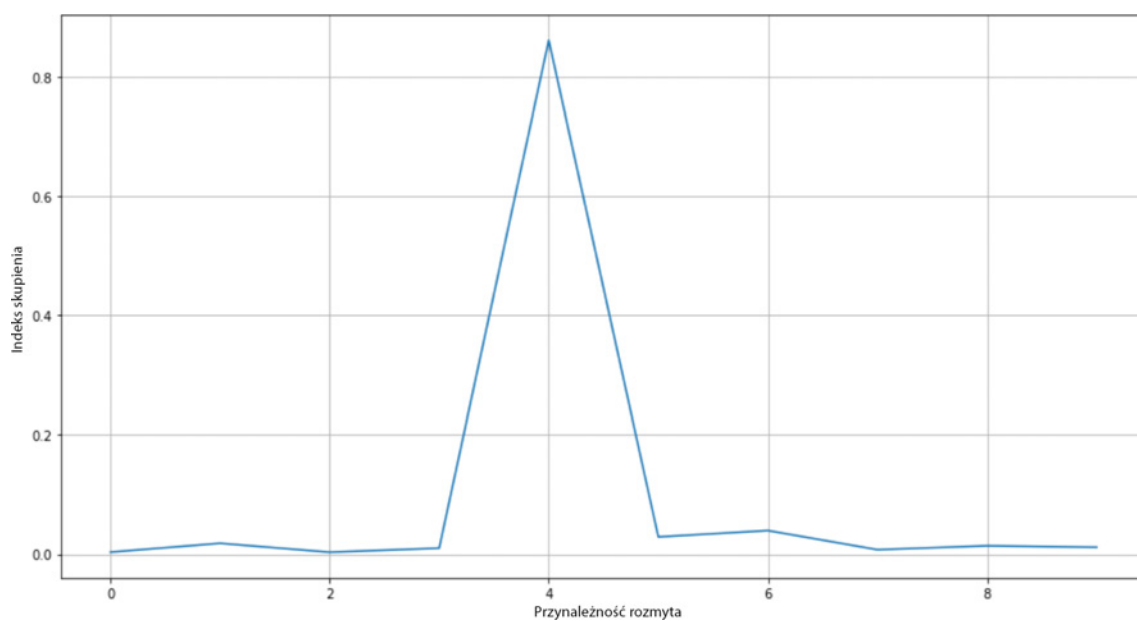
Rysunek 7.10. Przykład zbiorów rozmytych ukazujący staż pracy w stosunku do czasu zdobywania doświadczenia



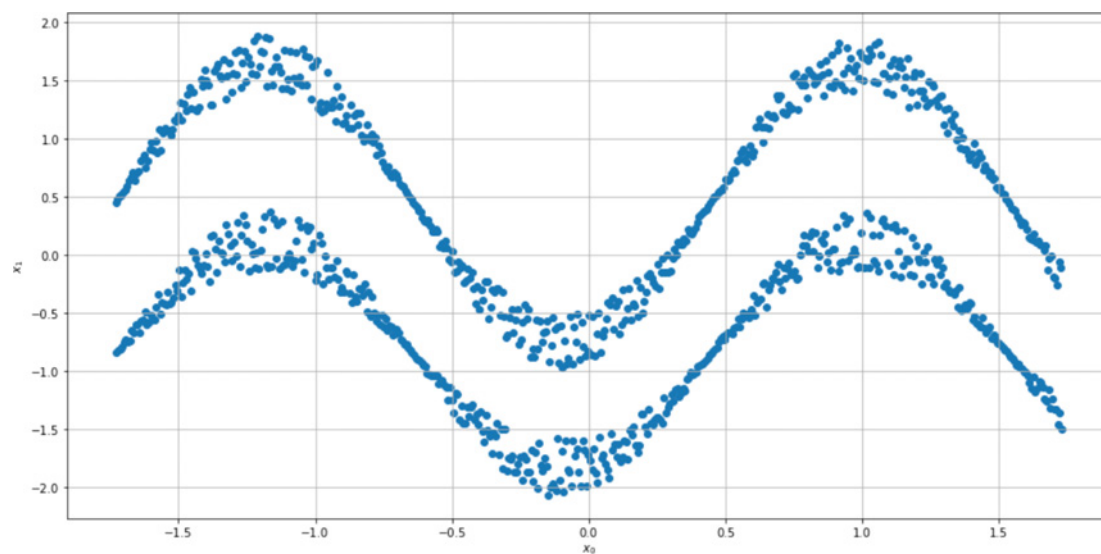
Rysunek 7.11. Centroidy uzyskane za pomocą algorytmu rozmytych c-średnich



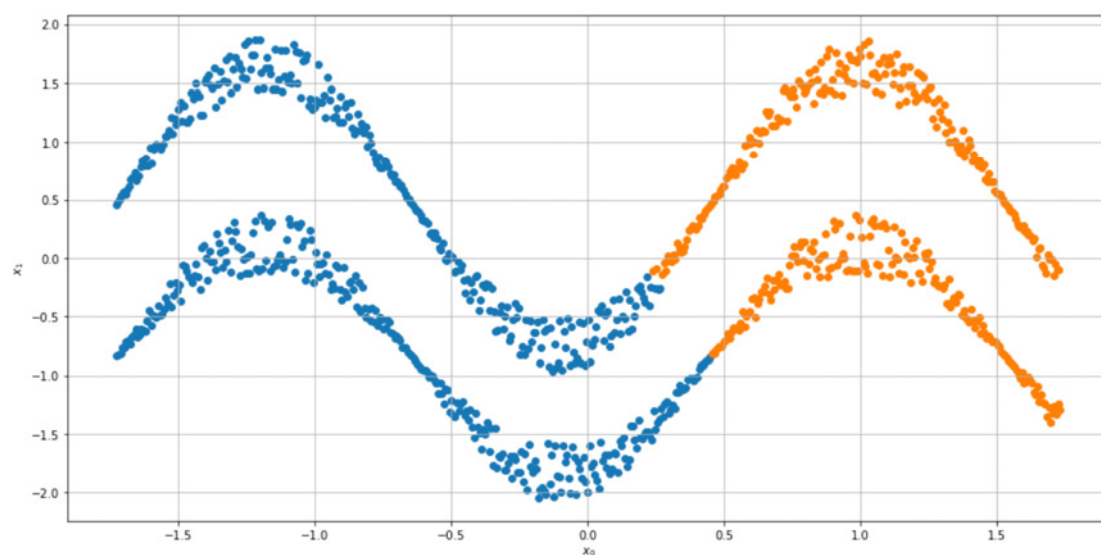
Rysunek 7.12. Przykładowa cyfra (7) wybrana do przetestowania rozmycia



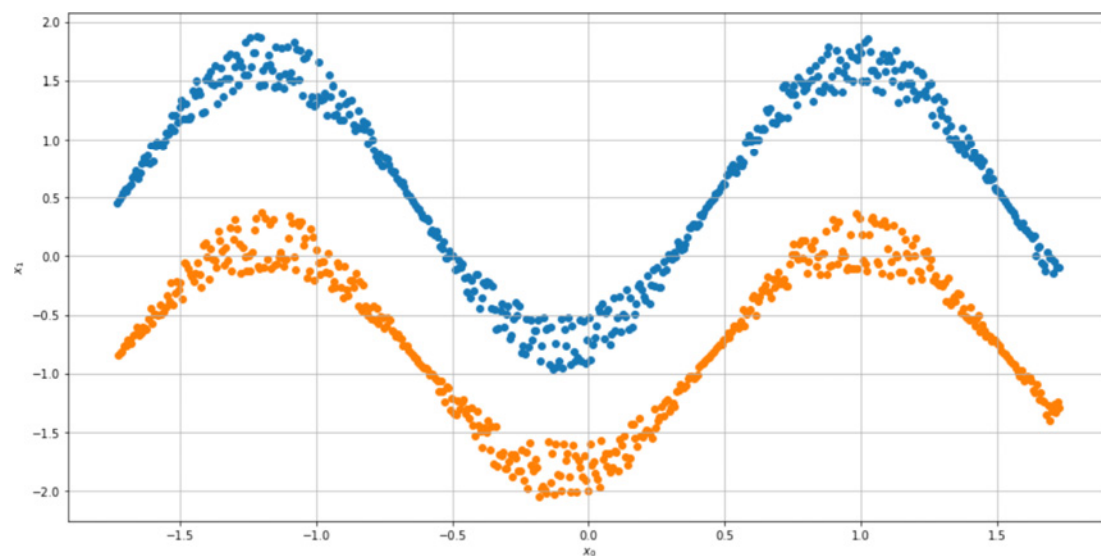
Rysunek 7.13. Wykres przynależności rozmytej odpowiadającej przykładowi reprezentującemu cyfrę 7



Rysunek 7.14. Sinusoidalny zestaw danych

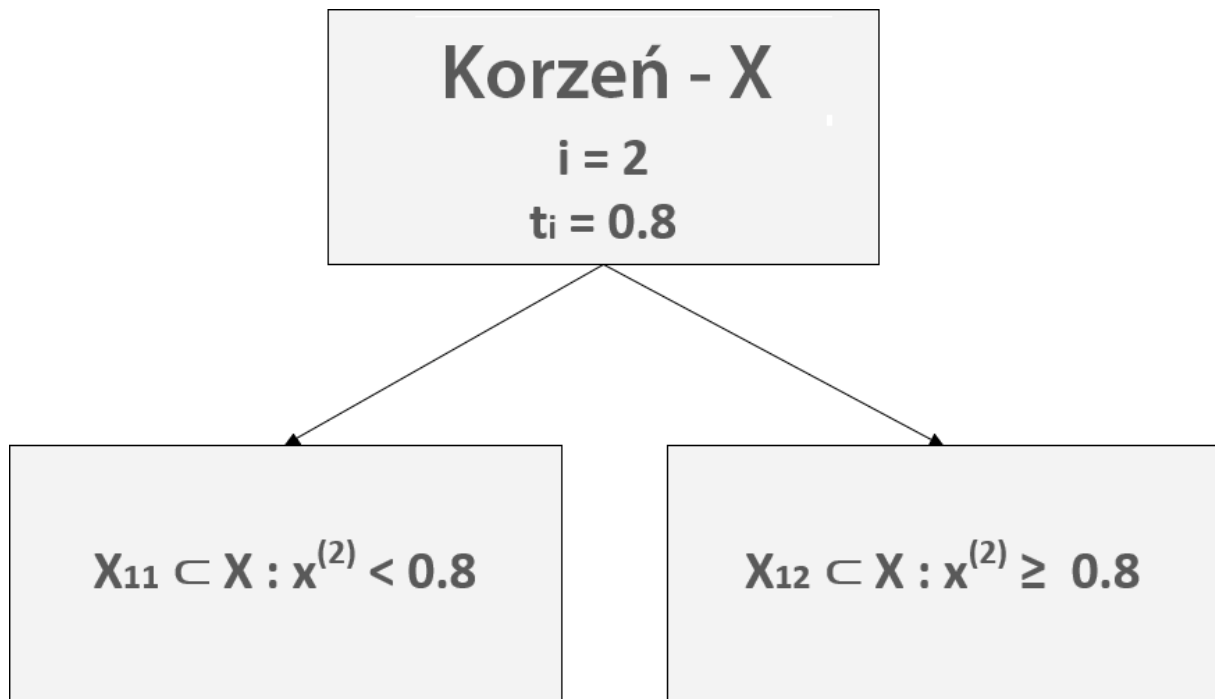


Rysunek 7.15. Wynik analizy skupień przeprowadzonej na sinusoidalnym zestawie danych za pomocą algorytmu centroidów

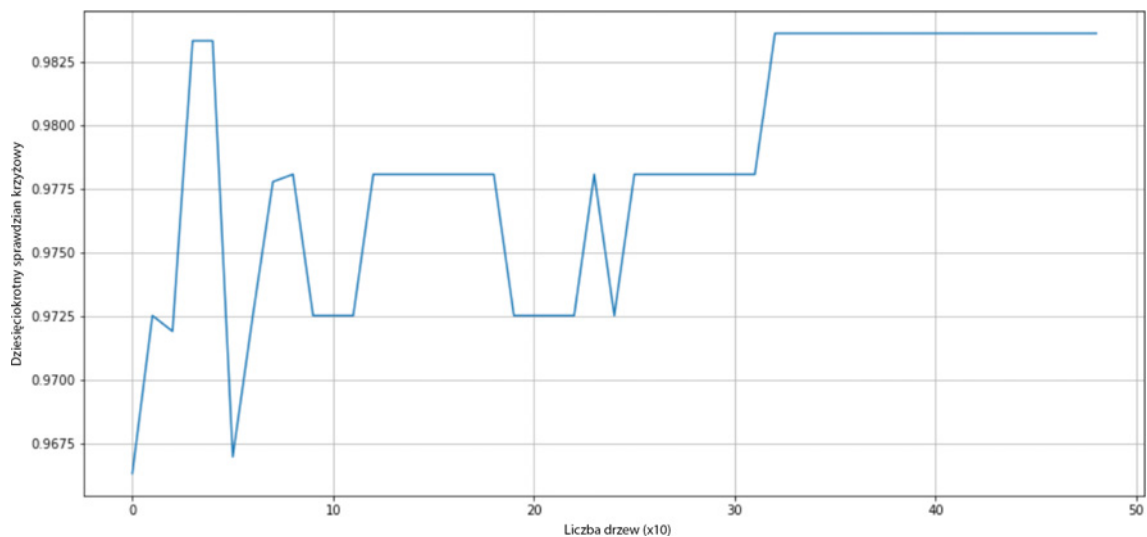


Rysunek 7.16. Wynik klasteryzacji widmowej dla sinusoidalnego zestawu danych

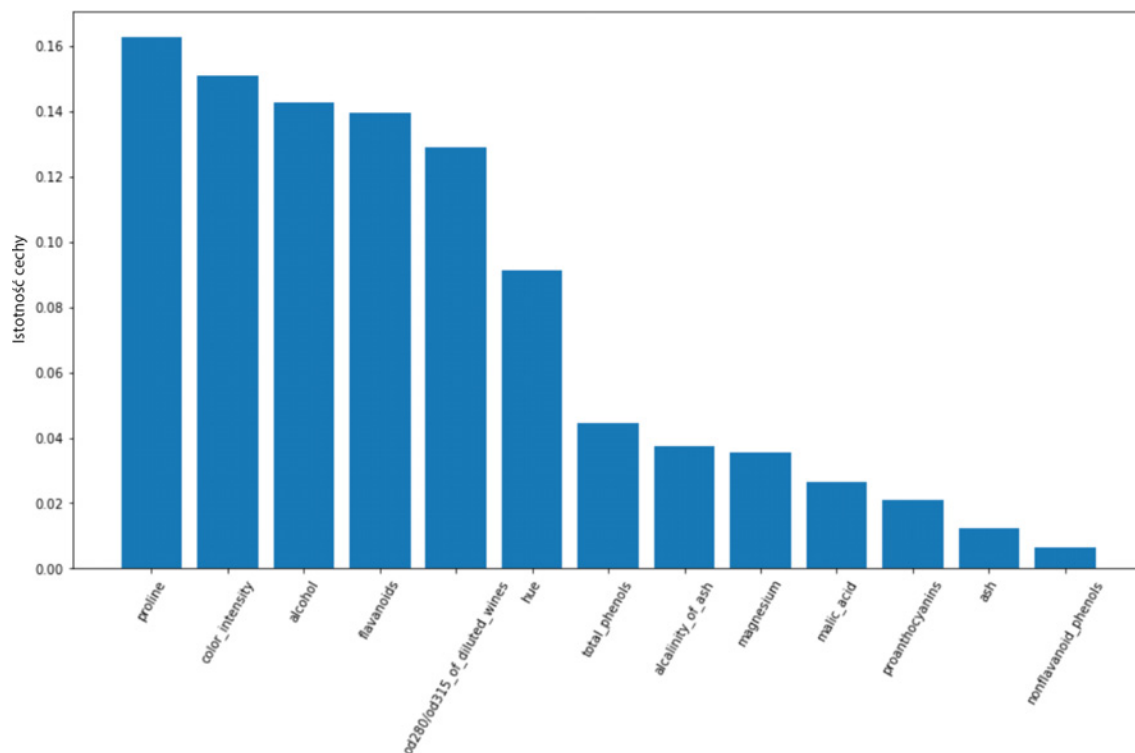




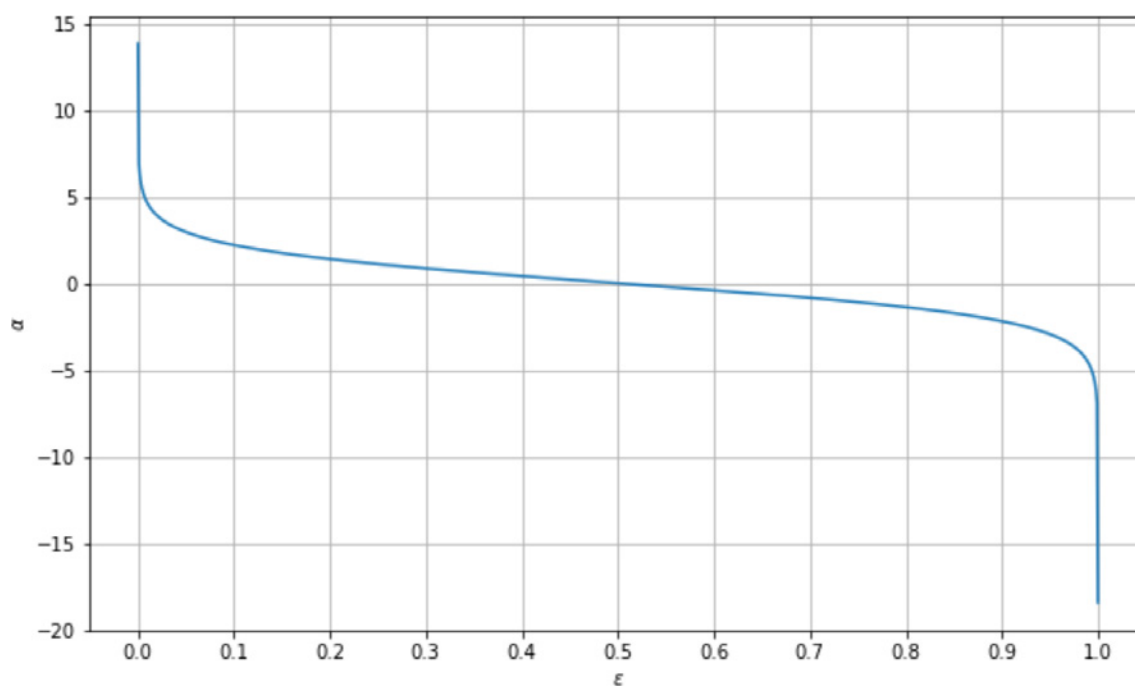
Rysunek 8.1. Przykładowy pierwszy podział w drzewie decyzyjnym



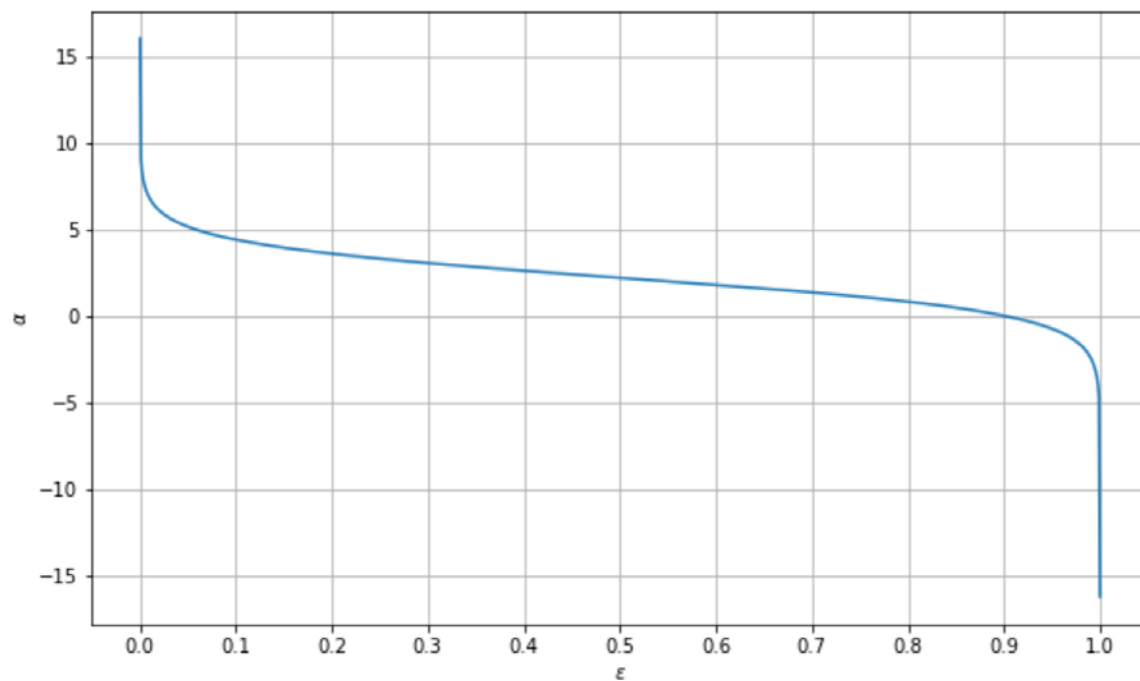
Rysunek 8.2. Dokładność sprawdzianu krzyżowego algorytmu lasu losowego jako funkcja liczby drzew



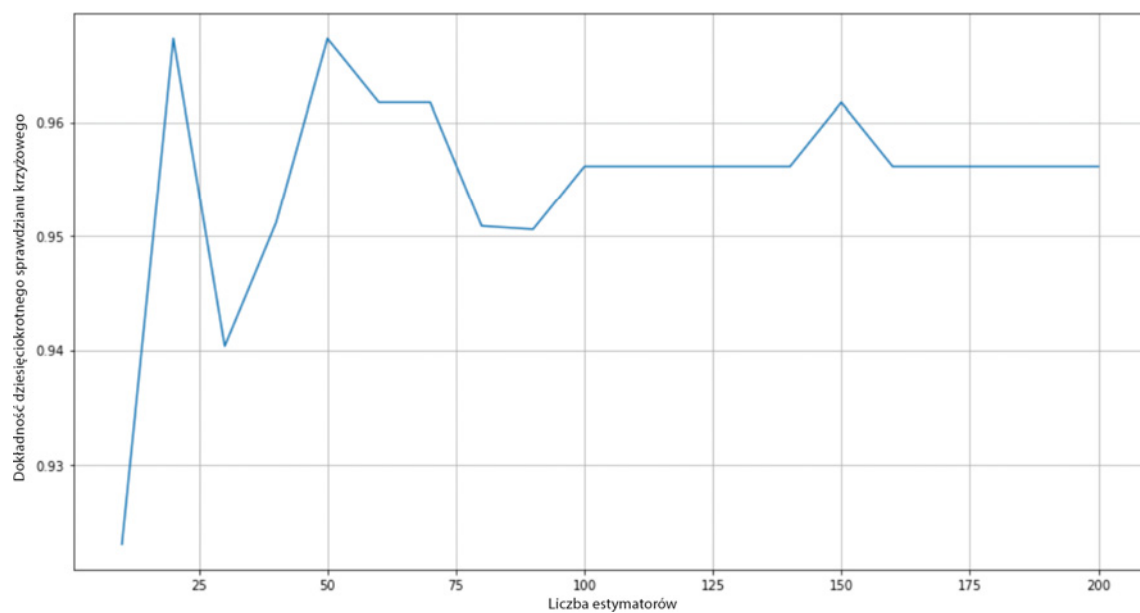
Rysunek 8.3. Istotność cech w zestawie danych Wine



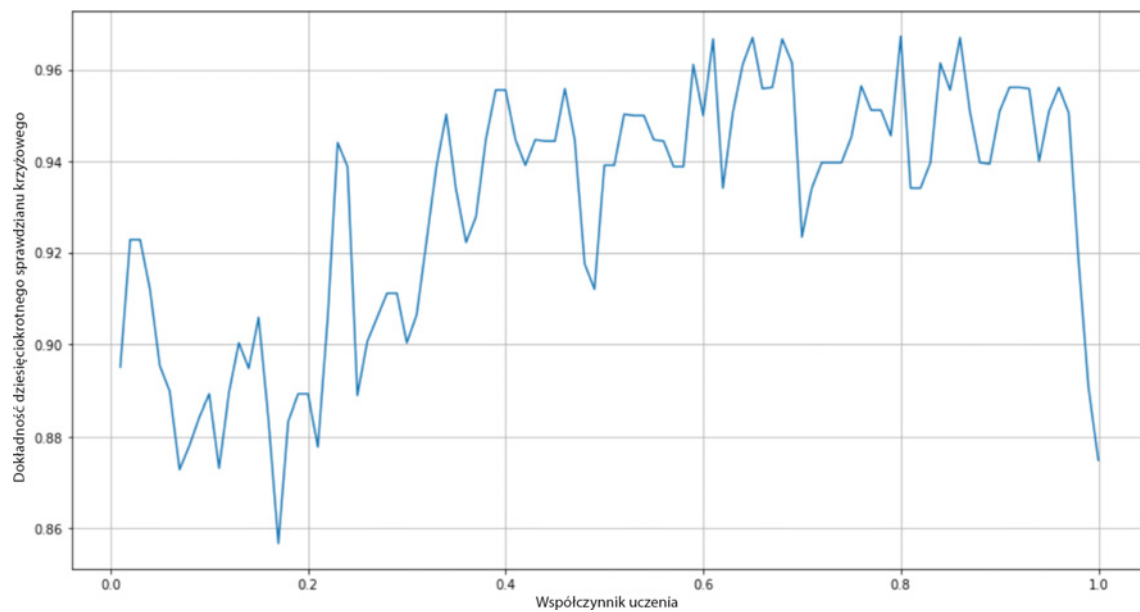
Rysunek 8.4. Wykres wag estymatora jako funkcja znormalizowanej sumy ważonej błędów



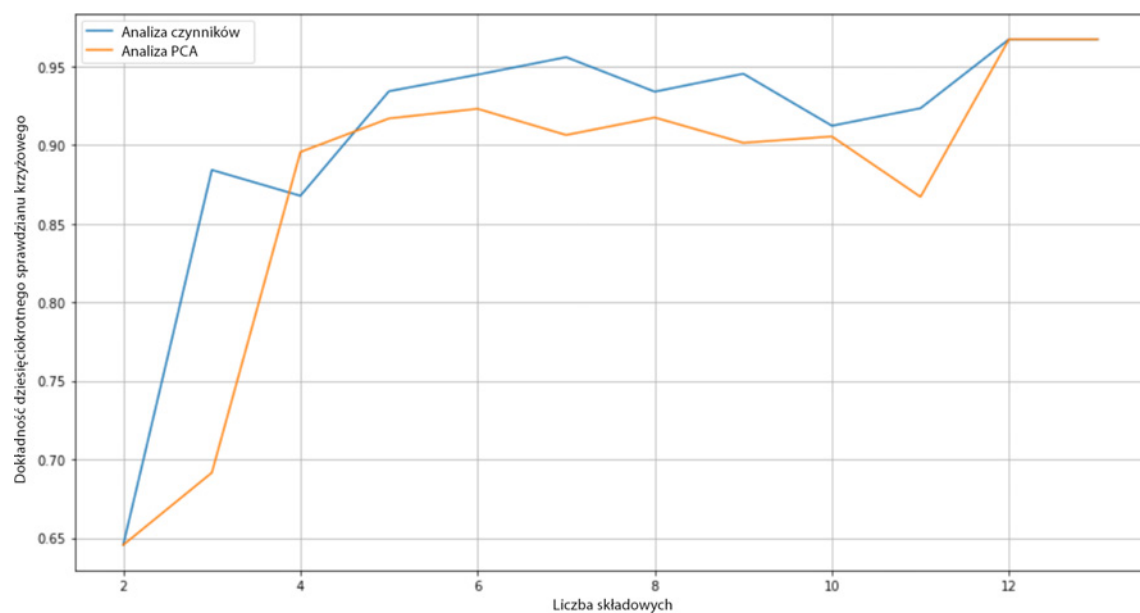
Rysunek 8.5. Wykres wagi estymatora jako funkcji znormalizowanej sumy ważonej błędów dla  $p = 10$



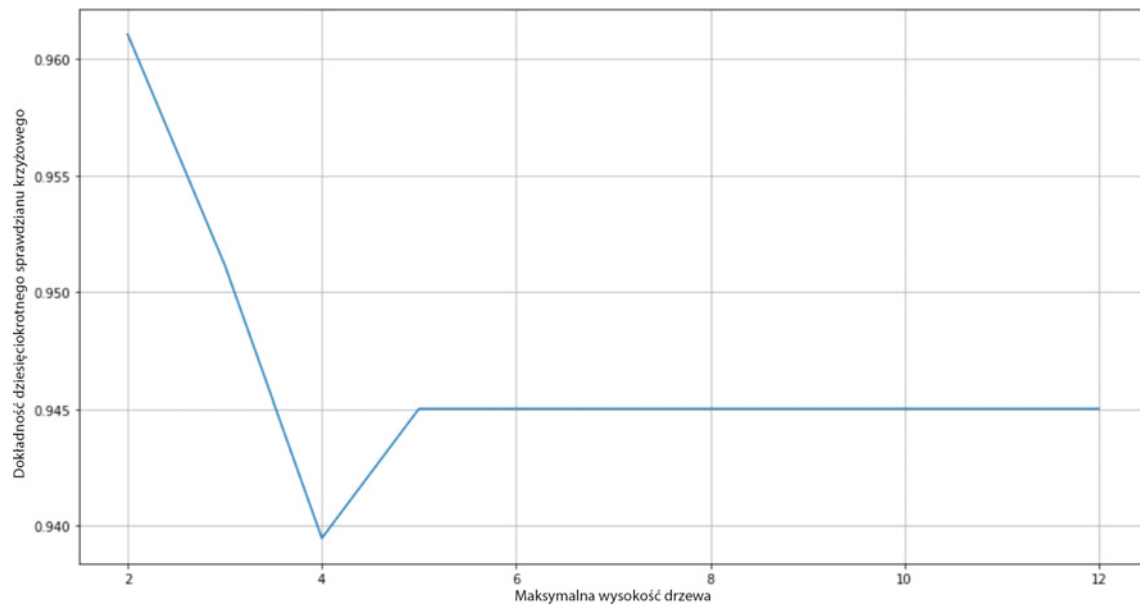
Rysunek 8.6. Dokładność dziesięciokrotnego sprawdzianu krzyżowego jako funkcja liczby estymatorów



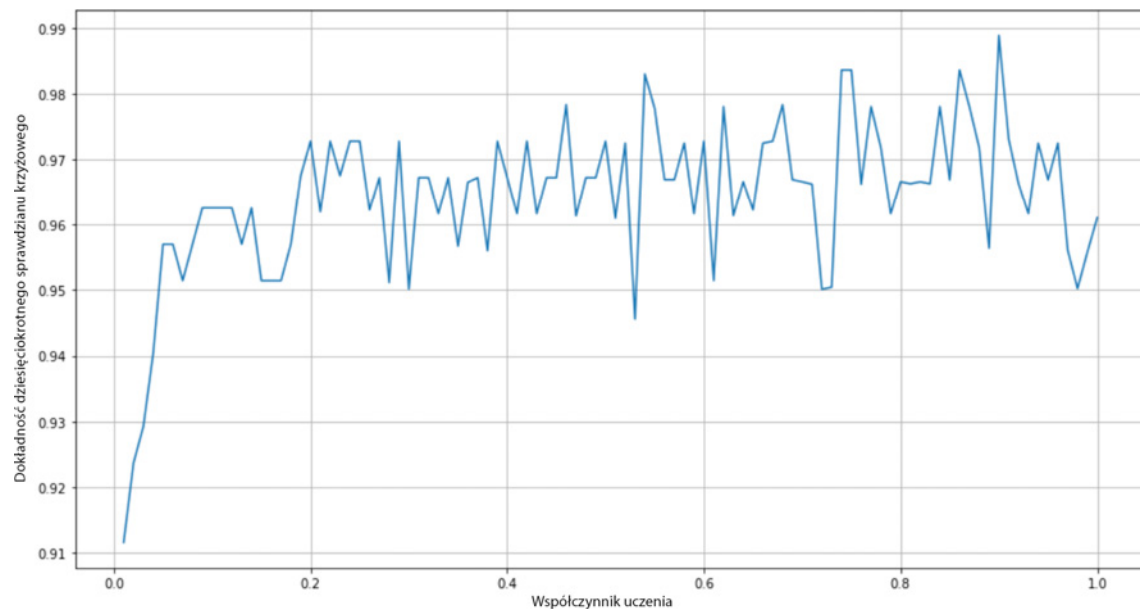
**Rysunek 8.7.** Dokładność dziesięciokrotnego sprawdzianu krzyżowego jako funkcja współczynnika uczenia (liczba estymatorów = 50)



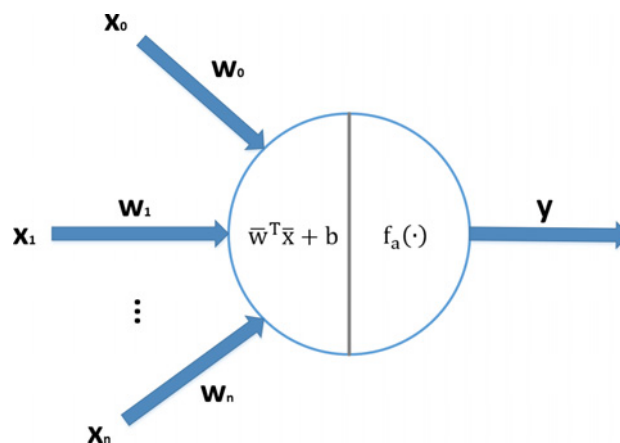
**Rysunek 8.8.** Dokładność dziesięciokrotnego sprawdzianu krzyżowego jako funkcja liczby składowych (uzyskanych metodami PCA i FC)



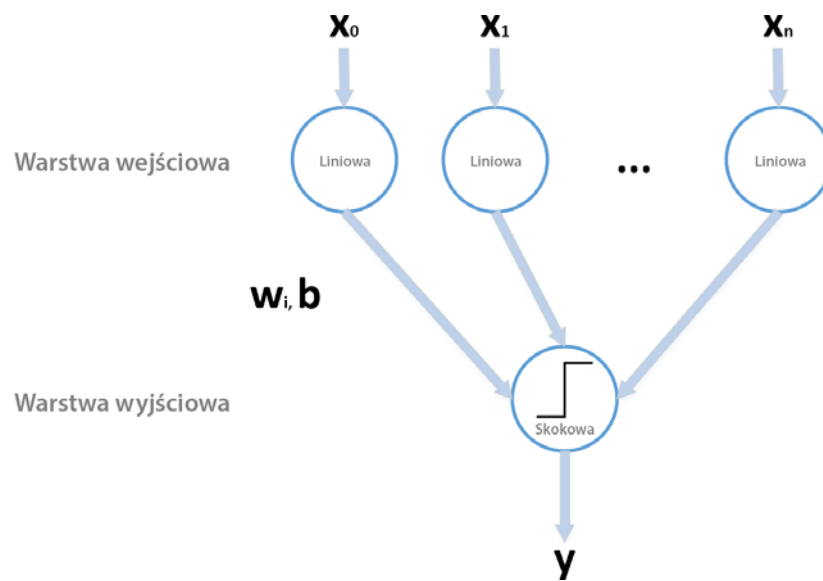
Rysunek 8.9. Dokładność dziesięciokrotnego sprawdzianu krzyżowego jako funkcja maksymalnej wysokości drzewa



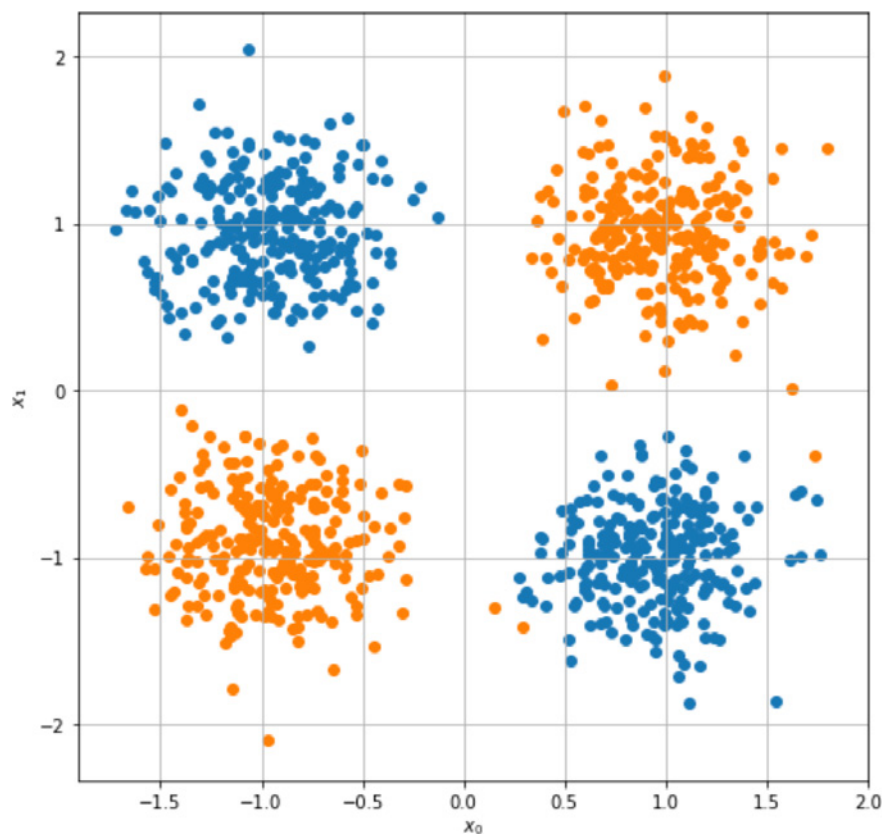
Rysunek 8.10. Dokładność dziesięciokrotnego sprawdzianu krzyżowego jako funkcja współczynnika uczenia (maksymalna wysokość drzew jest równa 2)



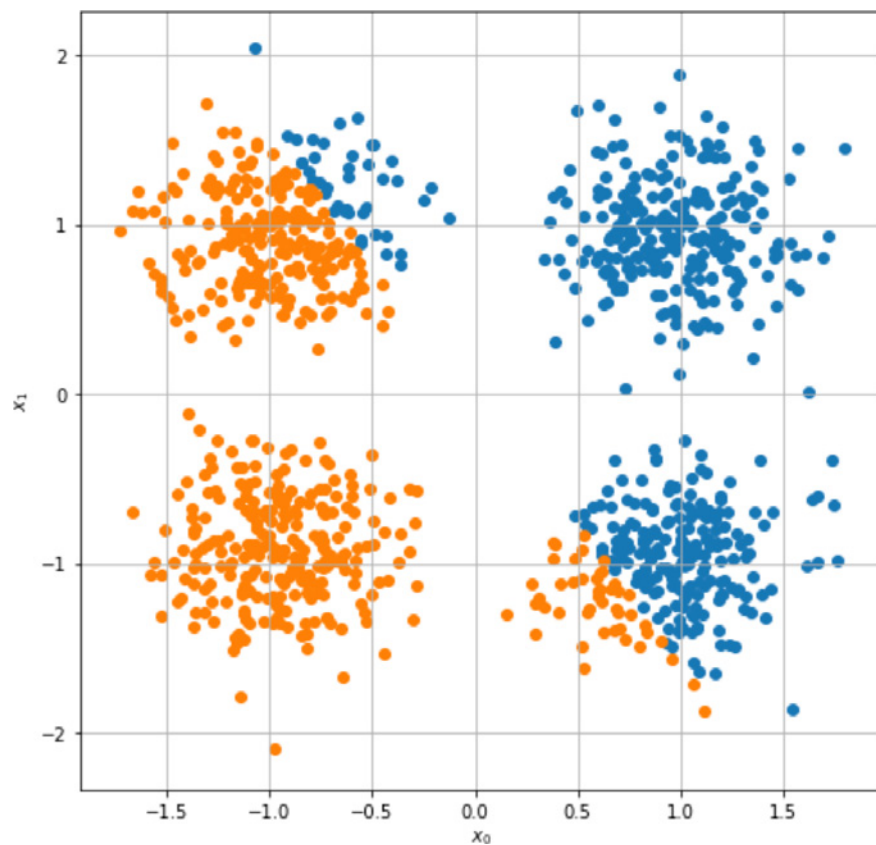
Rysunek 9.1. Schemat budowy podstawowego sztucznego neuronu



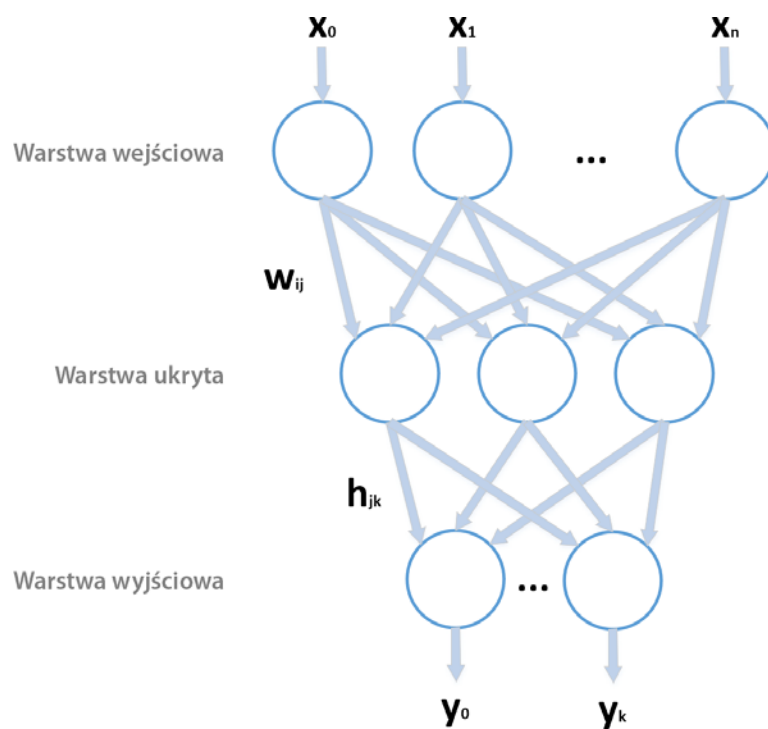
Rysunek 9.2. Schemat budowy perceptronu



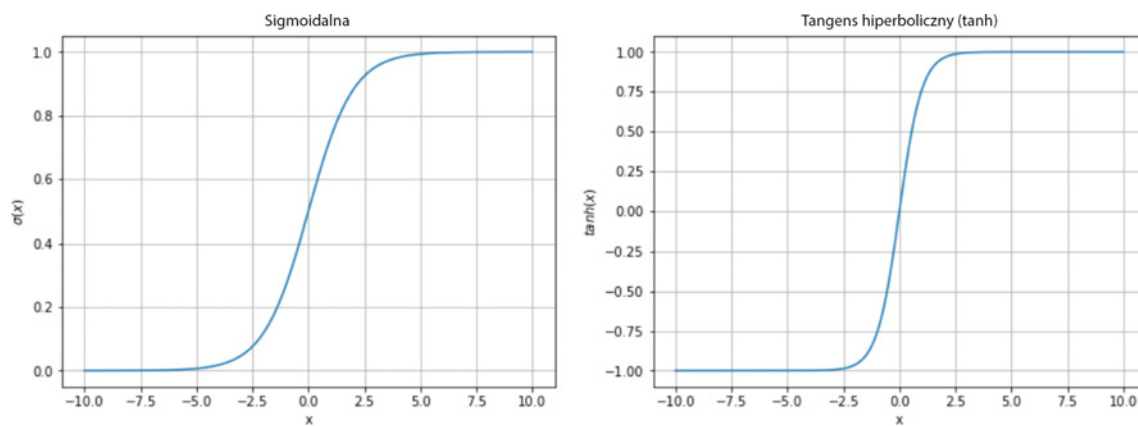
Rysunek 9.3. Przykładowy zestaw danych XOR



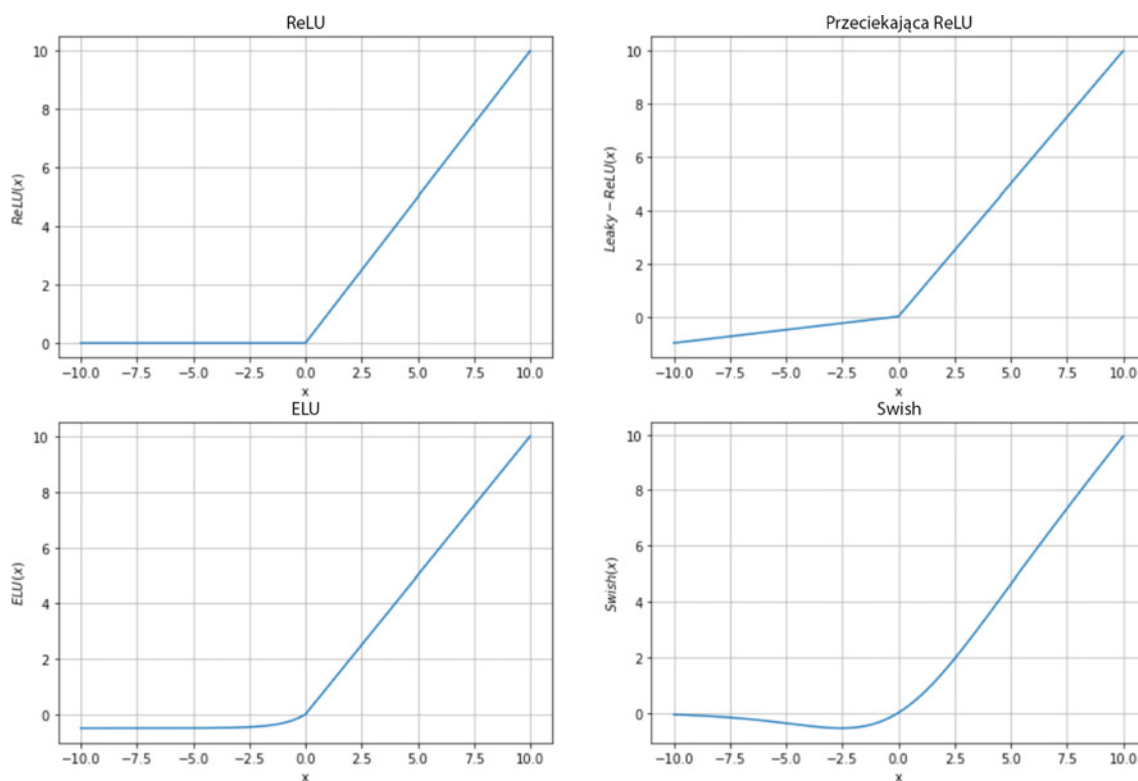
Rysunek 9.4. Zestaw danych XOR oznakowany przez perceptron



Rysunek 9.5. Schemat perceptronu wielowarstwowego zawierającego jedną warstwę ukrytą

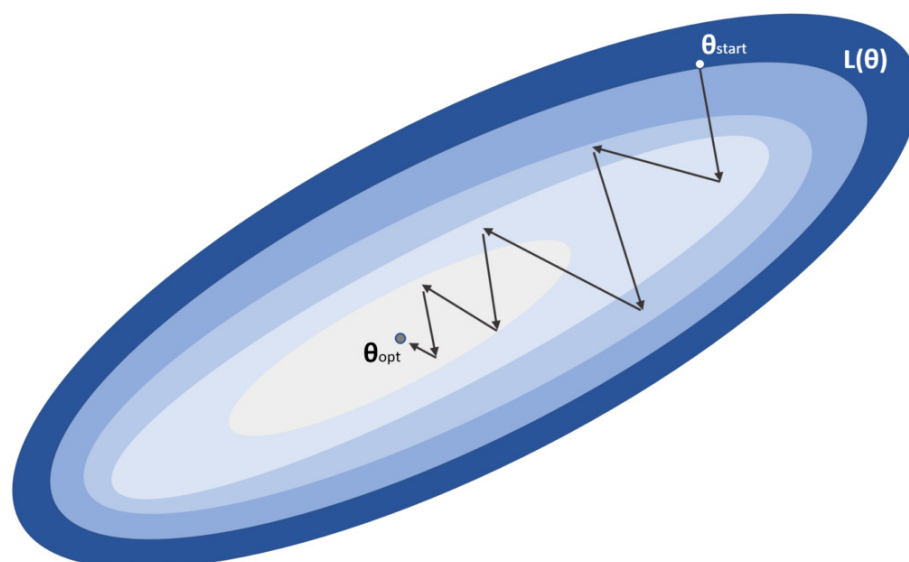


Rysunek 9.6. Wykresy funkcji sigmoidalnej i tangens hiperboliczny

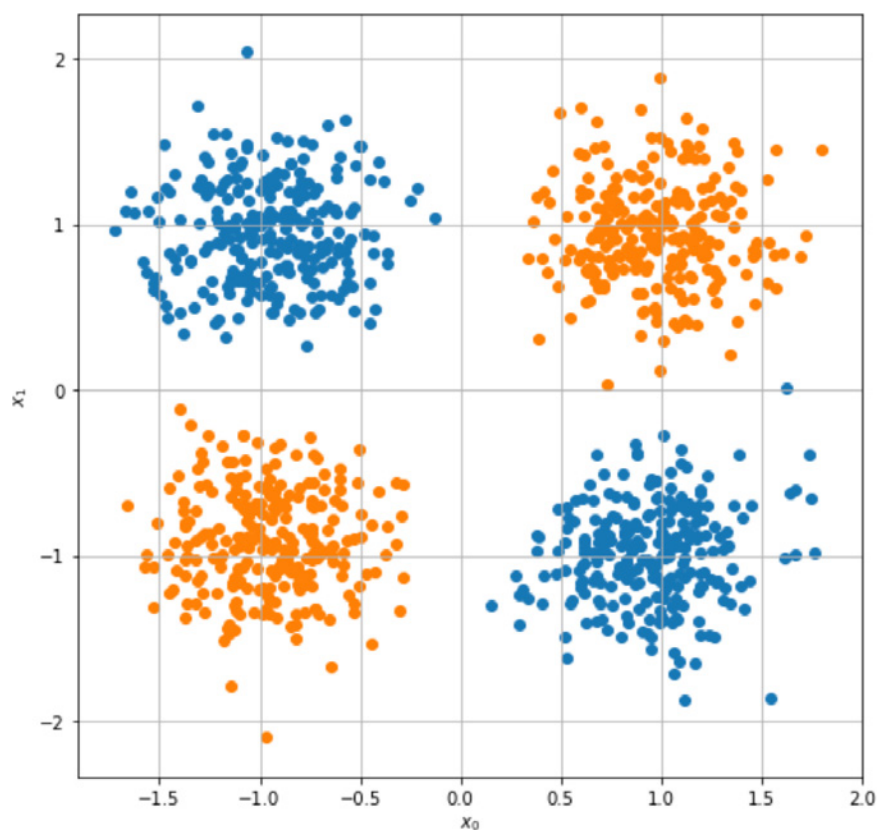


Rysunek 9.7. Wykresy najpopularniejszych funkcji prostujących

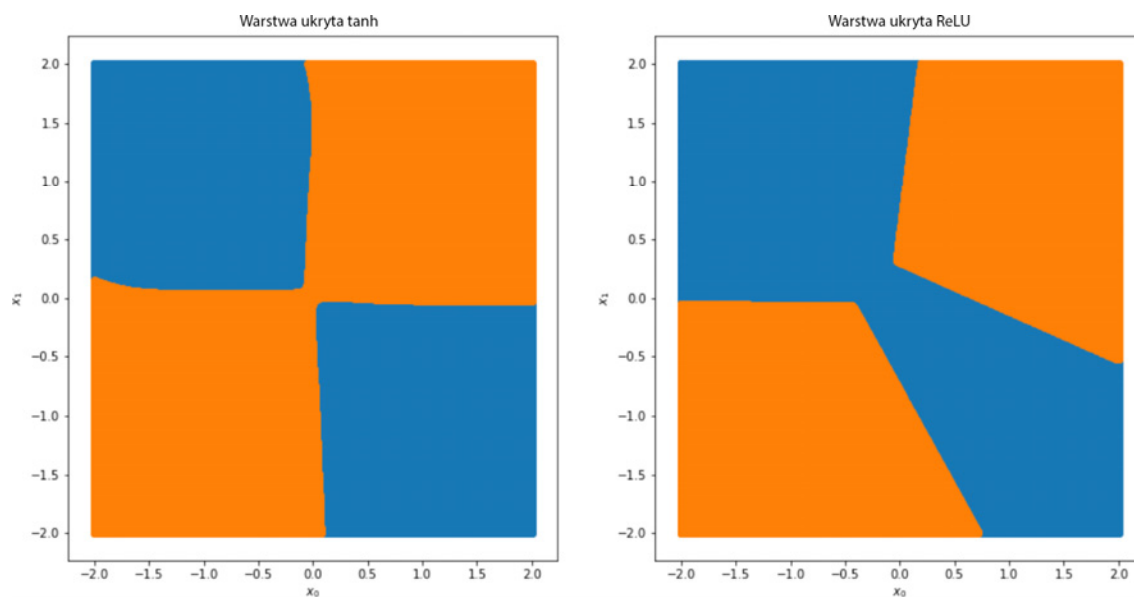




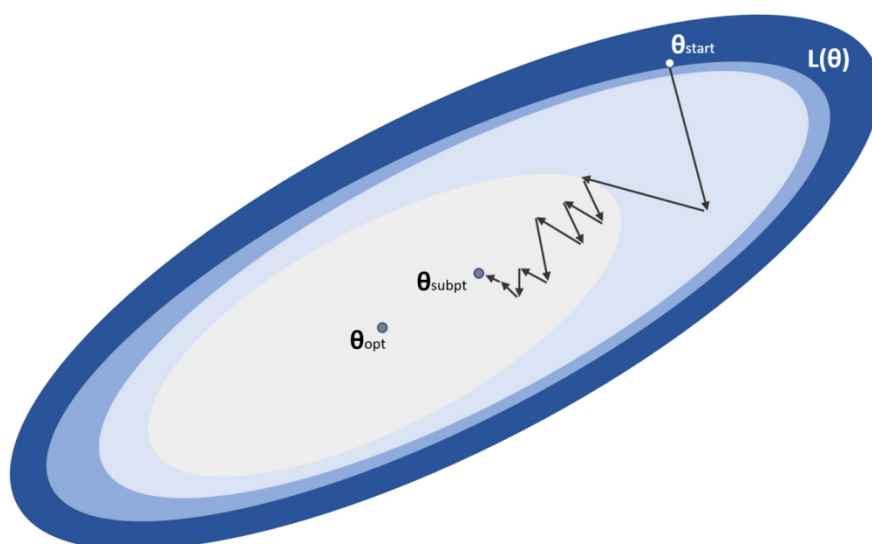
Rysunek 9.8. Graficzne przedstawienie stochastycznego spadku wzdłuż gradientu



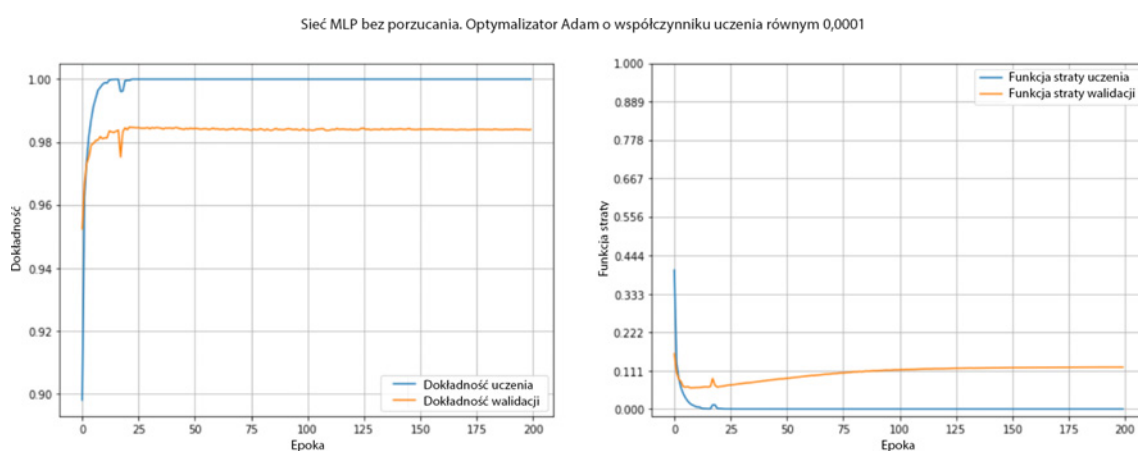
Rysunek 9.9. Klasyfikacja zestawu danych XOR przeprowadzona przez sieć MLP



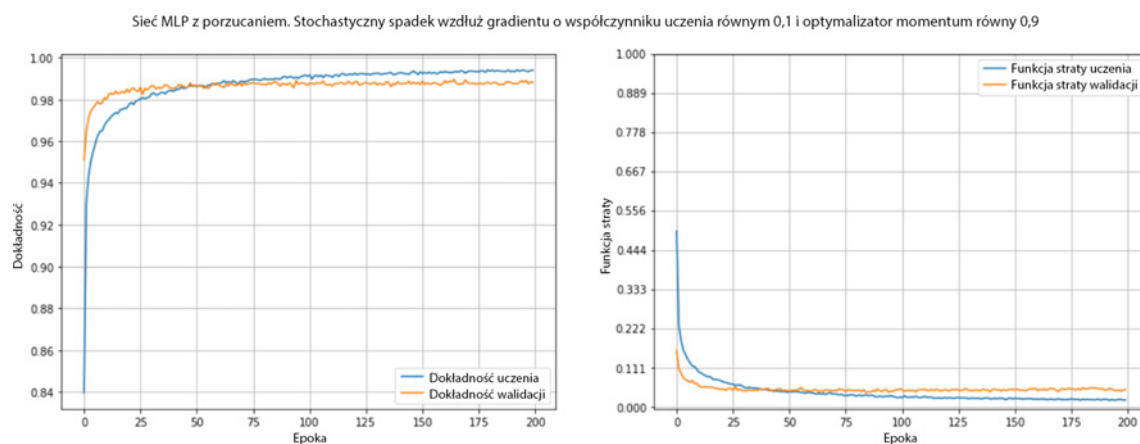
Rysunek 9.10. Powierzchnie decyzyjne sieci MLP w warstwach ukrytych tanh (po lewej) i ReLU (po prawej)



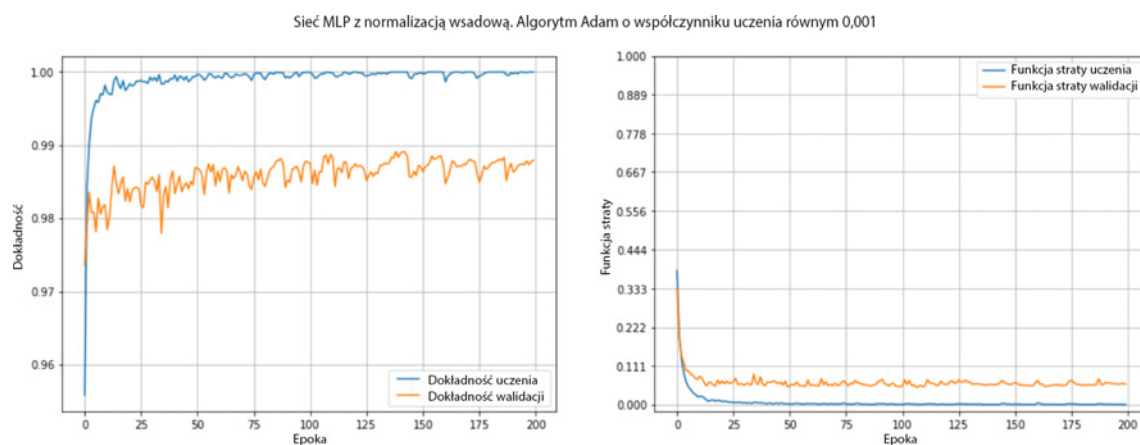
Rysunek 9.11. Często spotykany problem podczas korzystania ze stochastycznego spadku wzdłuż gradientu



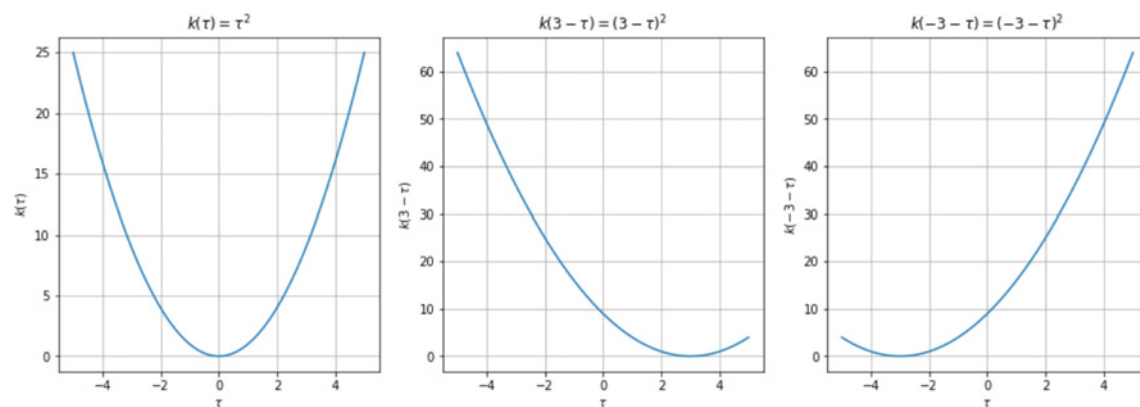
Rysunek 9.12. Wykres dokładności (po lewej) i funkcji straty (po prawej) jako funkcji liczby epok; model bez porzucania



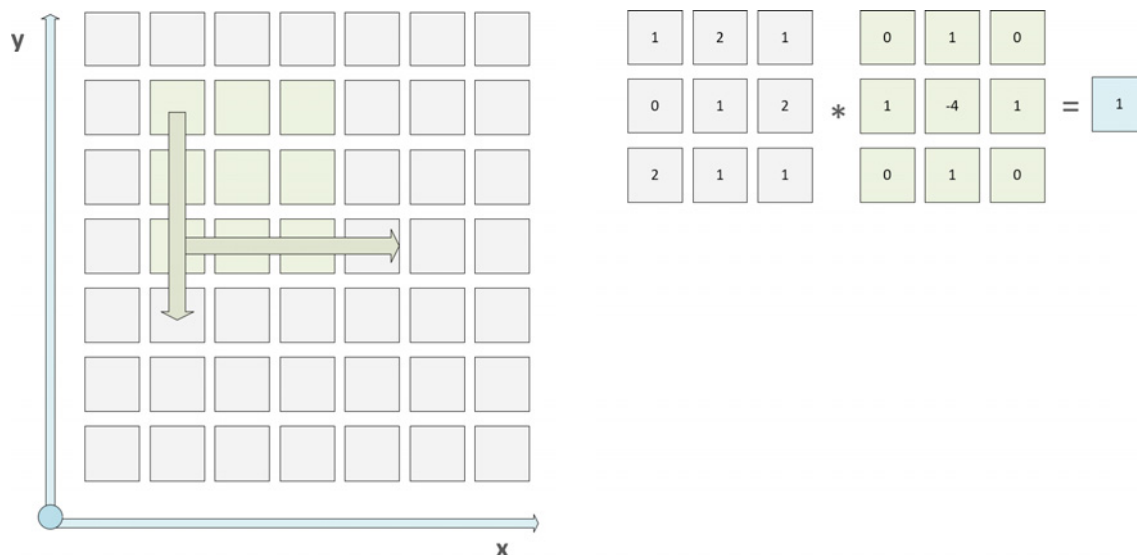
Rysunek 9.13. Wykres dokładności (po lewej) i funkcji straty (po prawej) jako funkcji liczby epok; model z porzucaniem



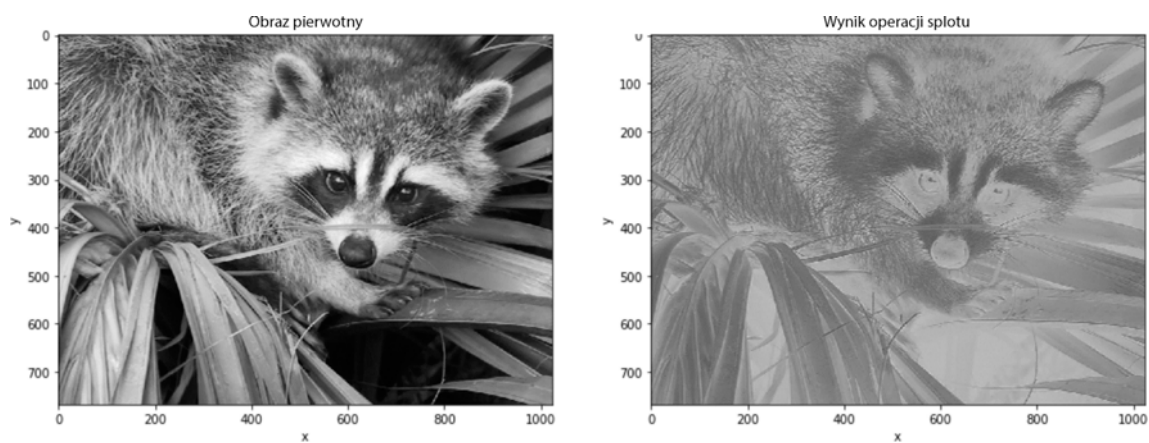
Rysunek 9.14. Wykres dokładności (po lewej) i funkcji straty (po prawej) jako funkcji liczby epok; model MLP z normalizacją wsadową



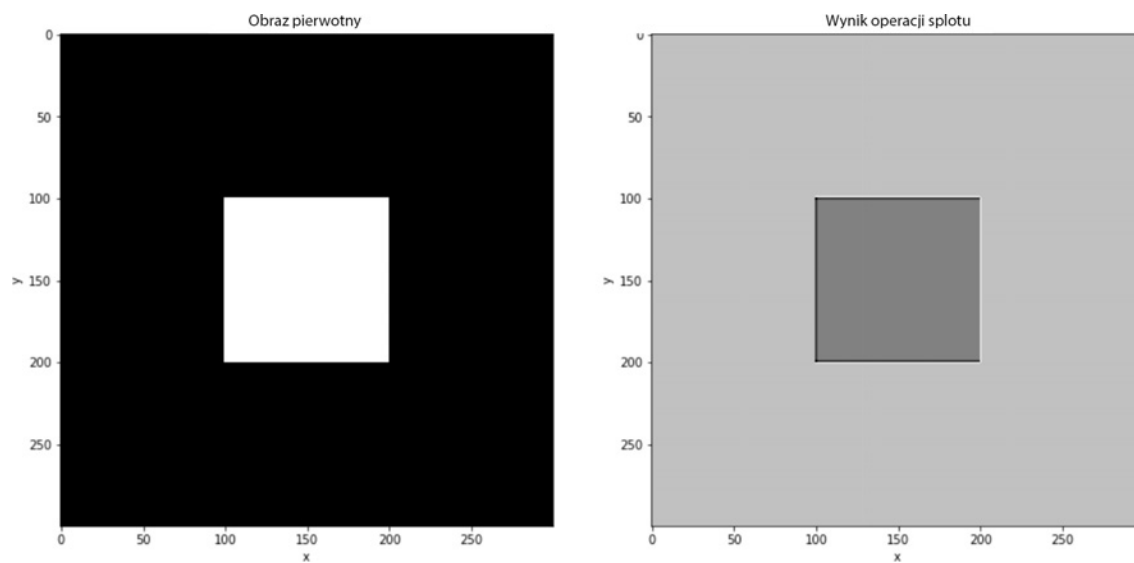
Rysunek 10.1. Przykład przesunięcia jądra



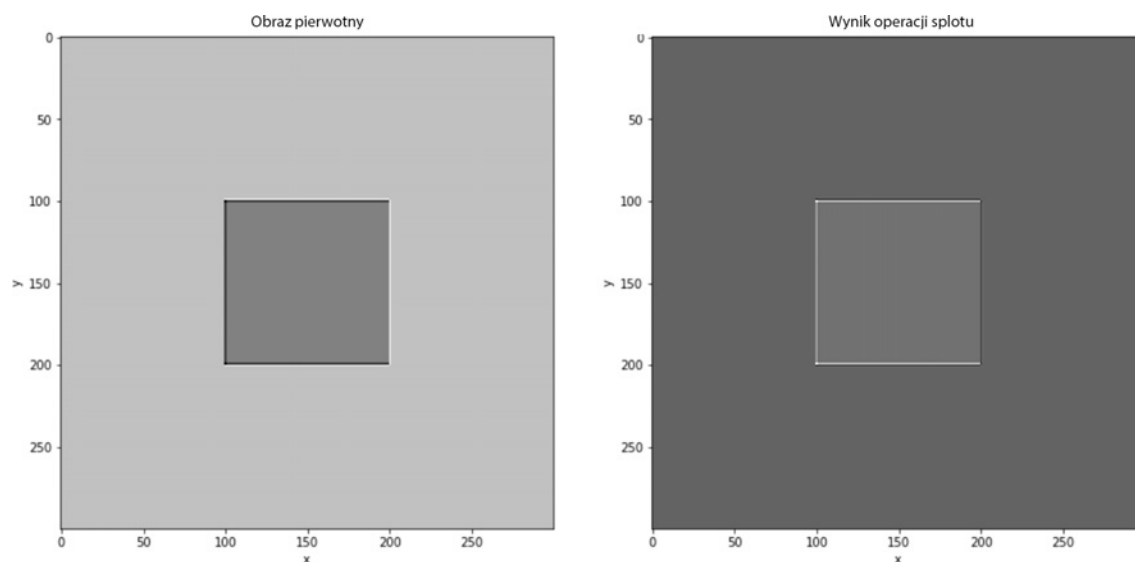
Rysunek 10.2. Przykład splotu dwuwymiarowego zawierającego jądro  $3 \times 3$



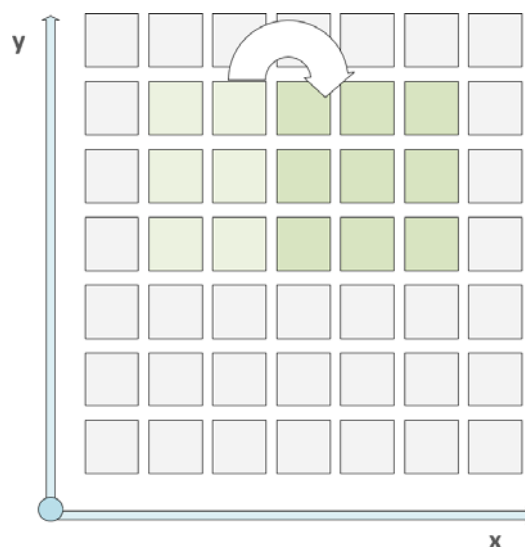
Rysunek 10.3. Przykład operacji splotu za pomocą jądra laplasjanu dyskretnego



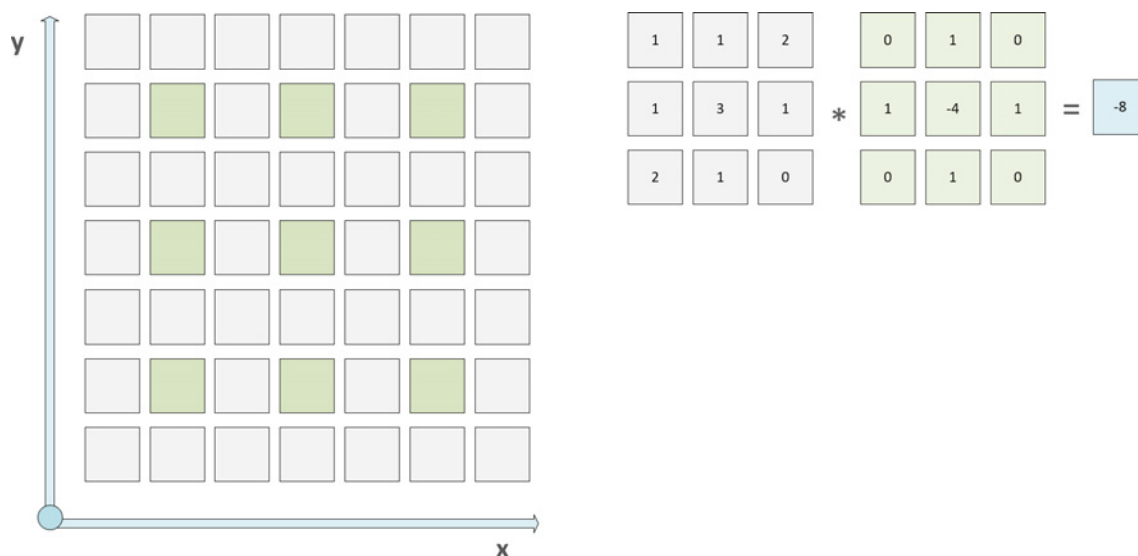
Rysunek 10.4. Obraz pierwotny (po lewej); wynik operacji splotu przy użyciu jądra laplasjanu (po prawej)



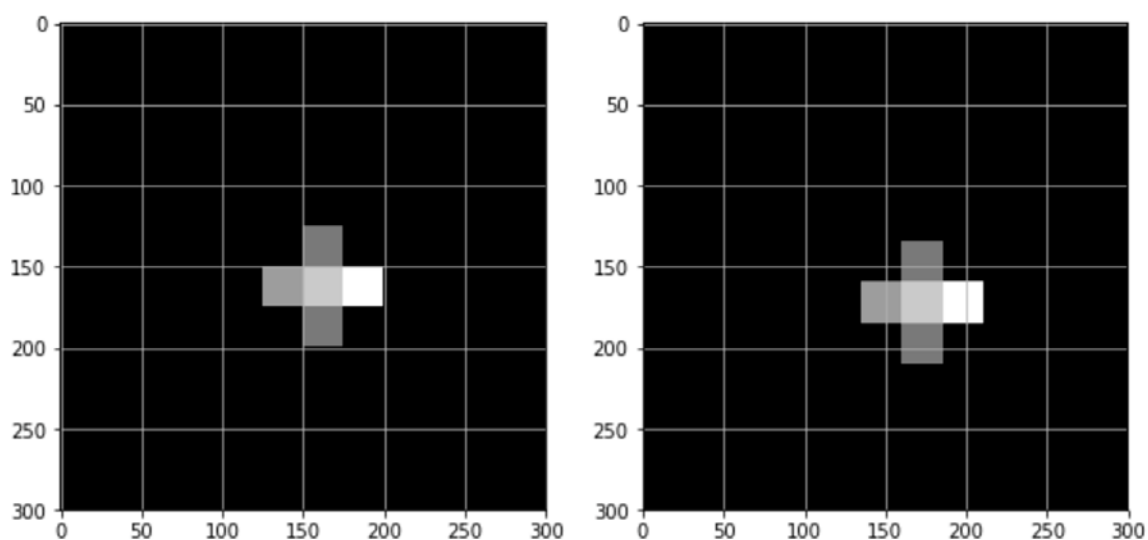
Rysunek 10.5. Drugie zastosowanie jądra laplasjanu



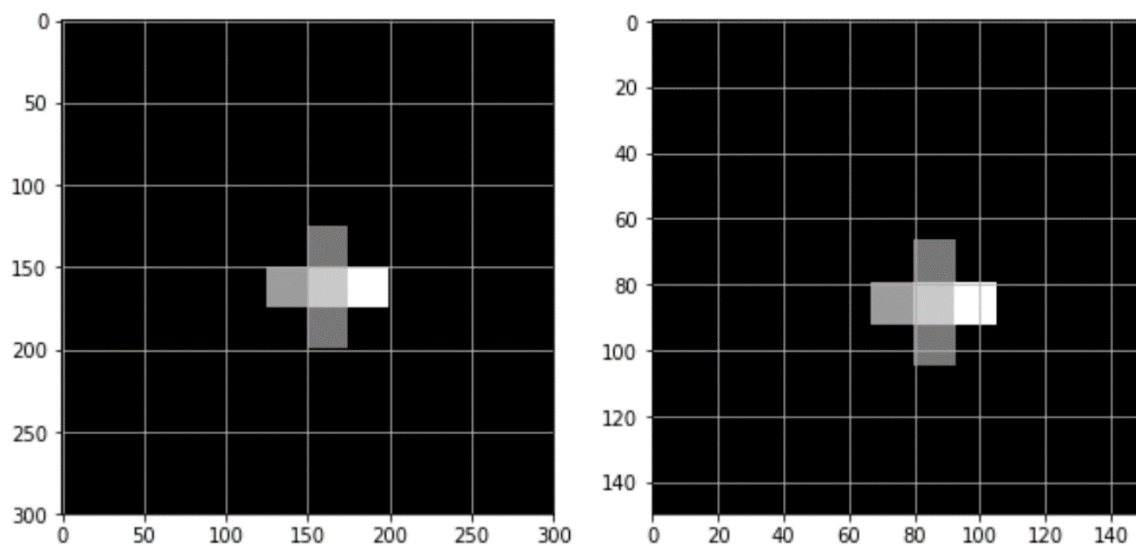
Rysunek 10.6. Przykład splotu dwuwymiarowego o kroku równym 2 na osi x



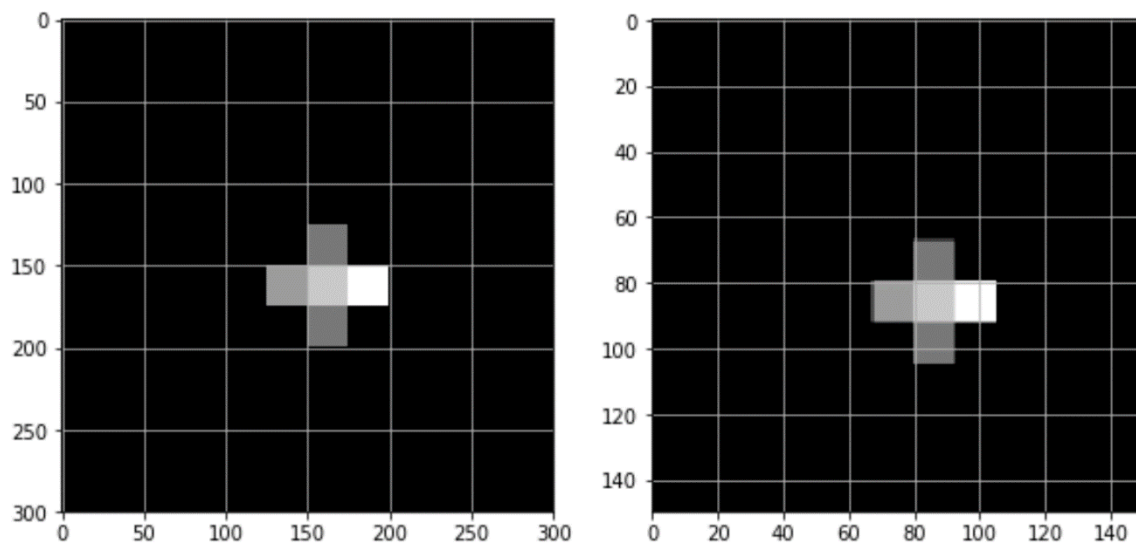
Rysunek 10.7. Przykład operacji splotu rozrzedzonego z jądrem laplasjanu



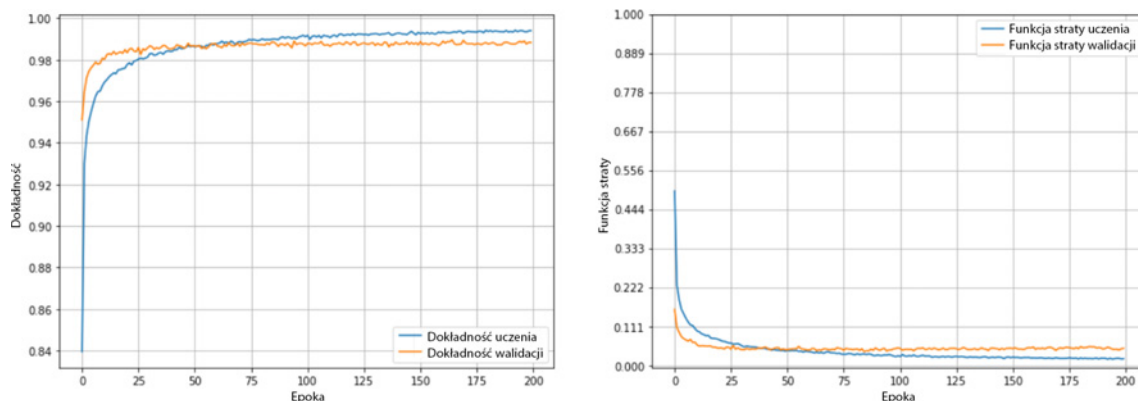
Rysunek 10.8. Obraz pierwotny (po lewej); obraz po zastosowaniu przesunięcia diagonalnego (po prawej)



Rysunek 10.9. Obraz pierwotny (po lewej); wynik zastosowania maksymalizującej warstwy łączącej na przesuniętym obrazie (po prawej)



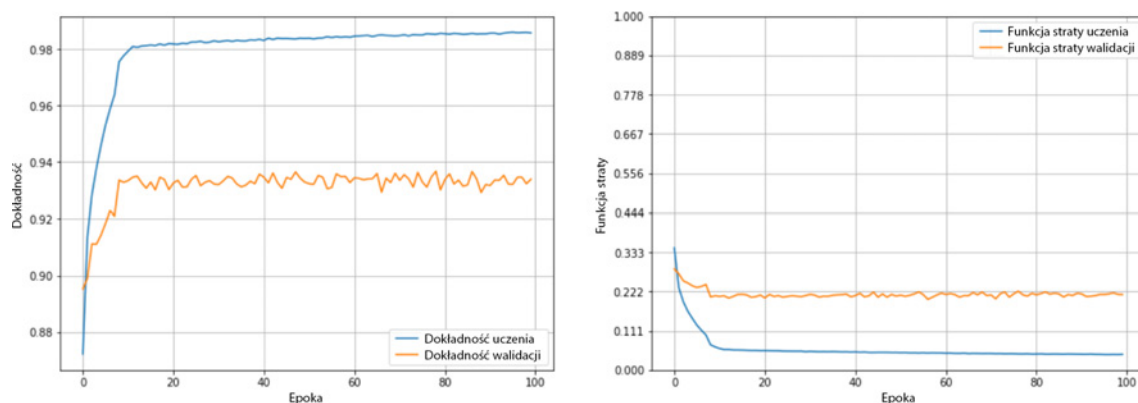
**Rysunek 10.10.** Obraz pierwotny (po lewej); wynik zastosowania uśredniającej warstwy łączącej na przesuniętym obrazie (po prawej)



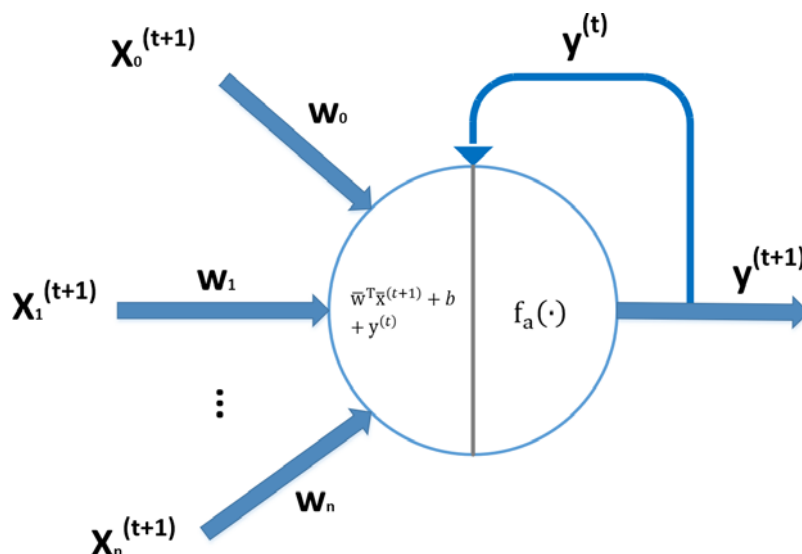
**Rysunek 10.11.** Wykres dokładności (po lewej) i funkcji straty (po prawej) jako funkcji liczby epok w zaimplementowanej głębokiej sieci splotowej



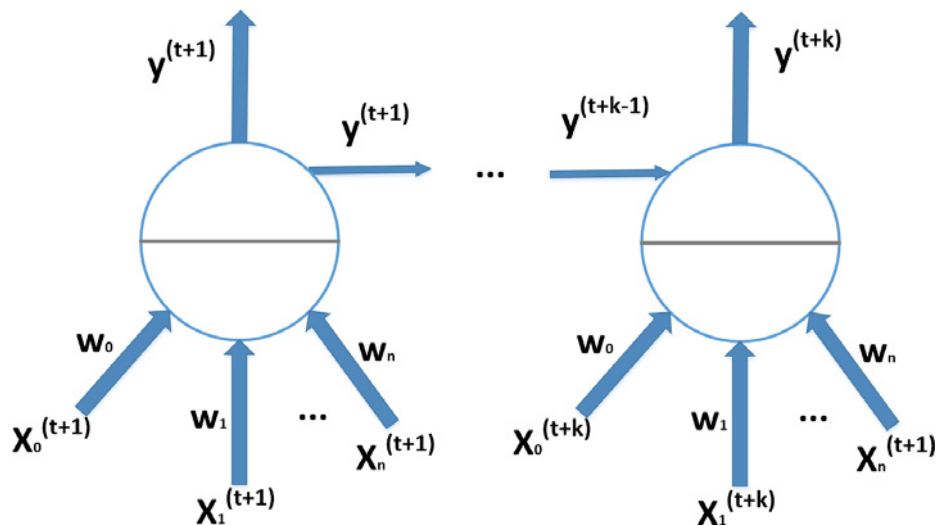
**Rysunek 10.12.** Przykładowe obrazy z zestawu danych Fashion MNIST



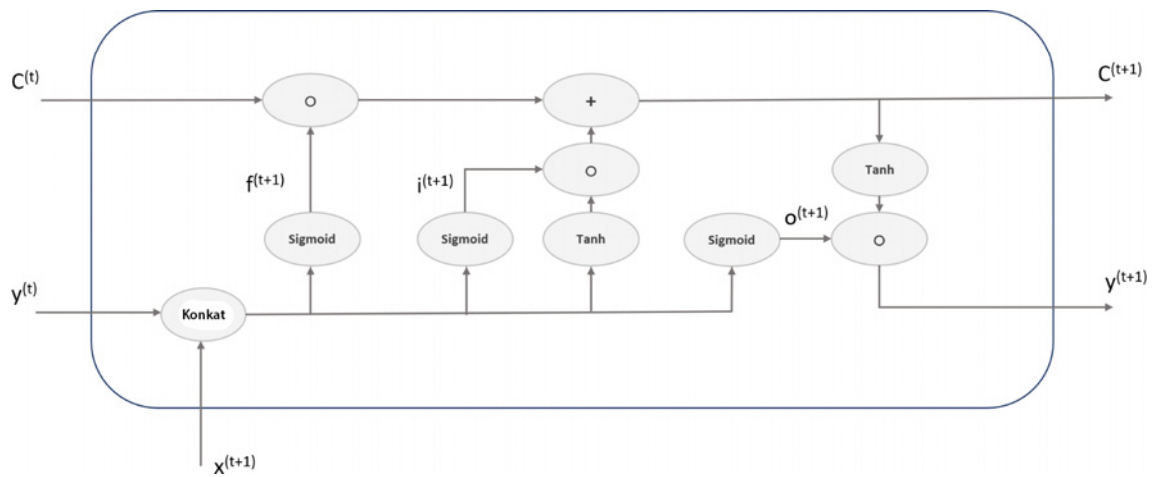
**Rysunek 10.13.** Wykres dokładności (po lewej) i funkcji straty (po prawej) jako funkcji liczby epok w zaimplementowanej głębokiej sieci splotowej z dogenerowaniem danych



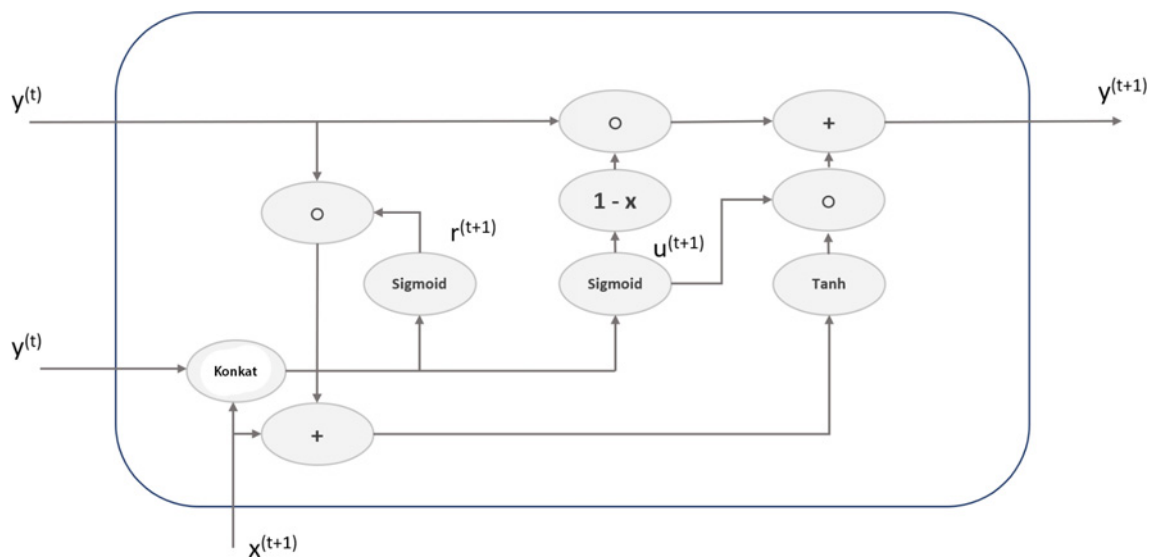
**Rysunek 10.14.** Schemat przykładowej sieci rekurencyjnej



Rysunek 10.15. Schemat rozwiniętej sieci rekurencyjnej

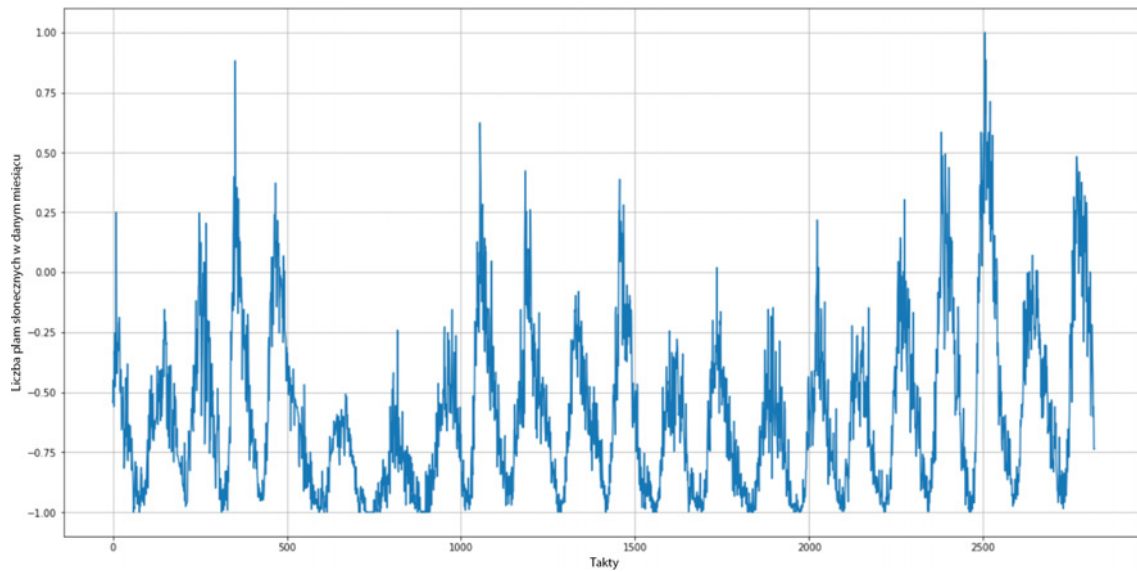


Rysunek 10.16. Schemat budowy jednostki LSTM

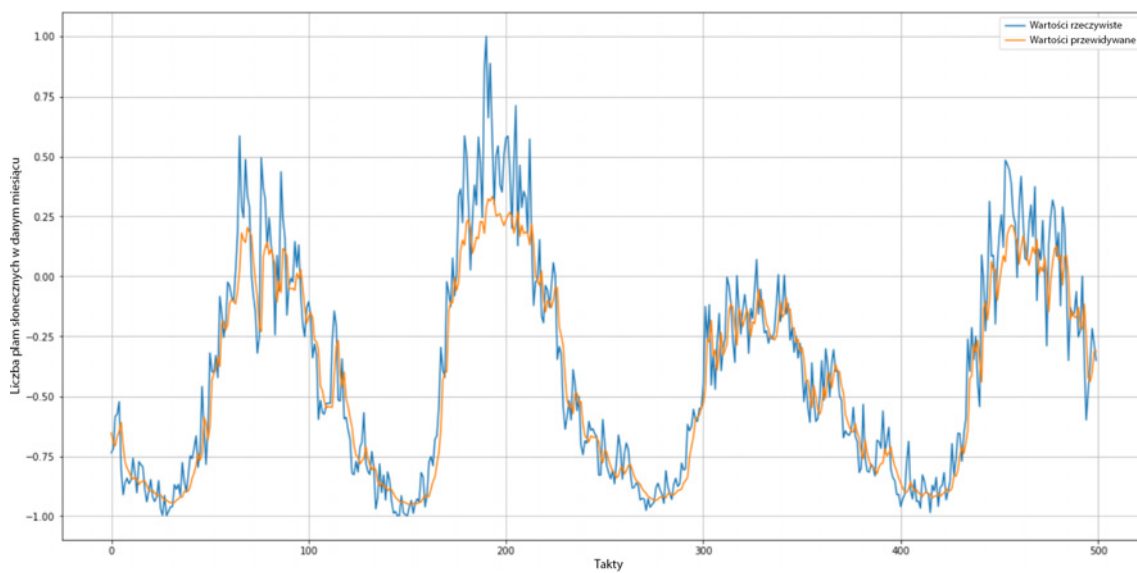


Rysunek 10.17. Schemat jednostki GRU

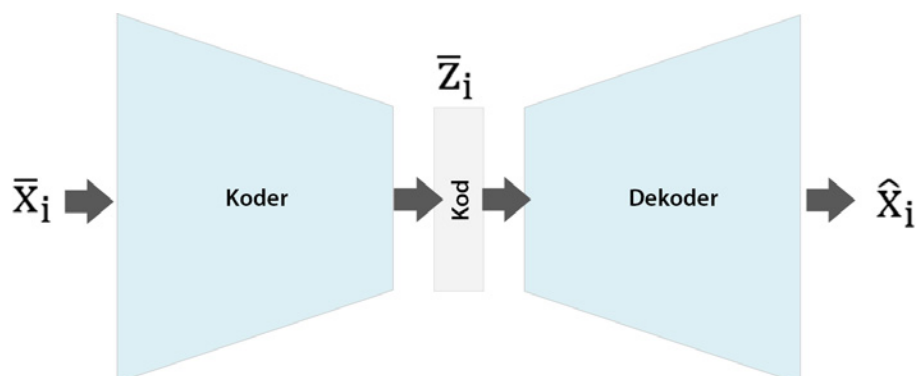




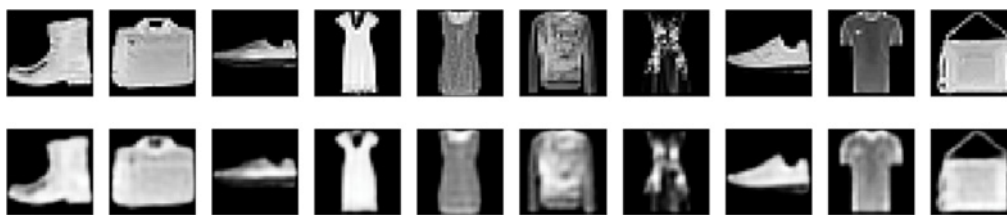
Rysunek 10.18. Zestaw danych Zuerich Monthly Sunspots



Rysunek 10.19. Prognozy uzyskane przez model LSTM za pomocą zestawu danych Zuerich



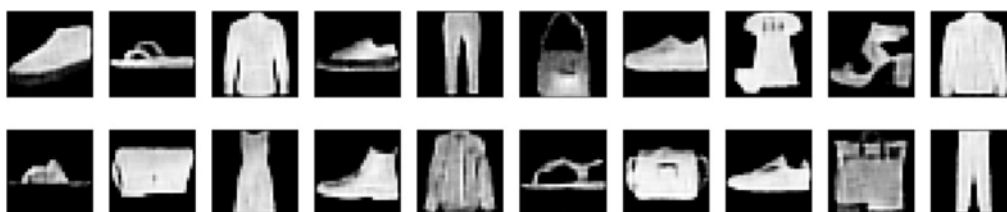
Rysunek 11.1. Ogólny schemat autokodera



Rysunek 11.2. Obrazy pierwotne (na górze); obrazy zdekodowane (na dole)



Rysunek 11.3. Zaszumione przykłady (na górze); przykłady odzsumione (na dole)



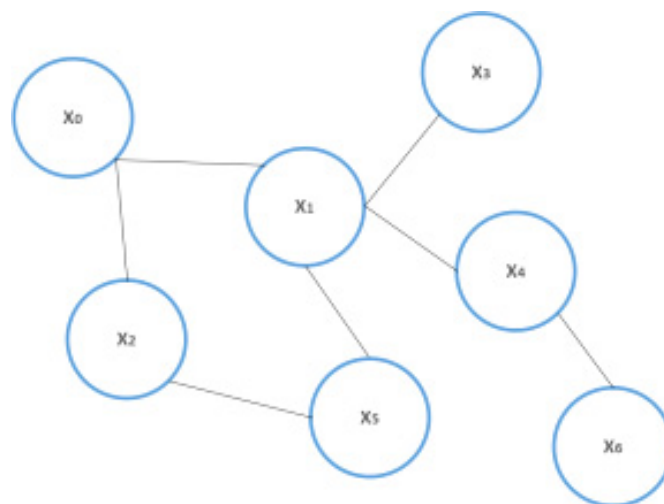
Rysunek 11.4. Wynik autokodera wariacyjnego



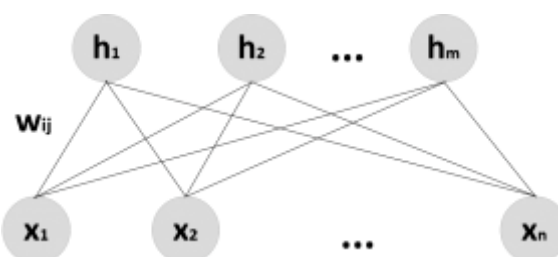
Rysunek 12.1. Przykłady wygenerowane przez sieć DCGAN wytrenowaną za pomocą zestawu danych Fashion MNIST



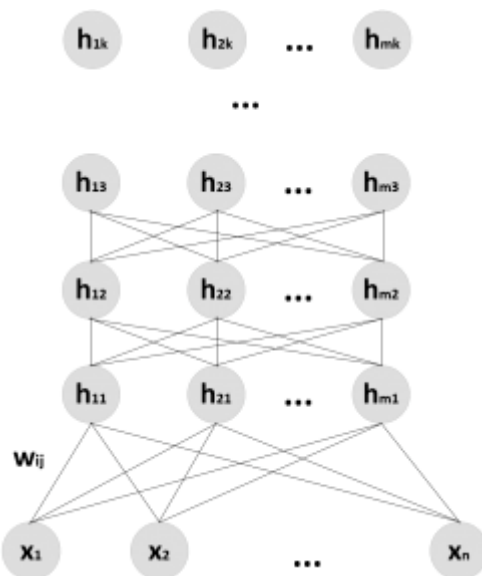
Rysunek 12.2. Przykłady wygenerowane przez sieć WGAN za pomocą zestawu danych Fashion MNIST



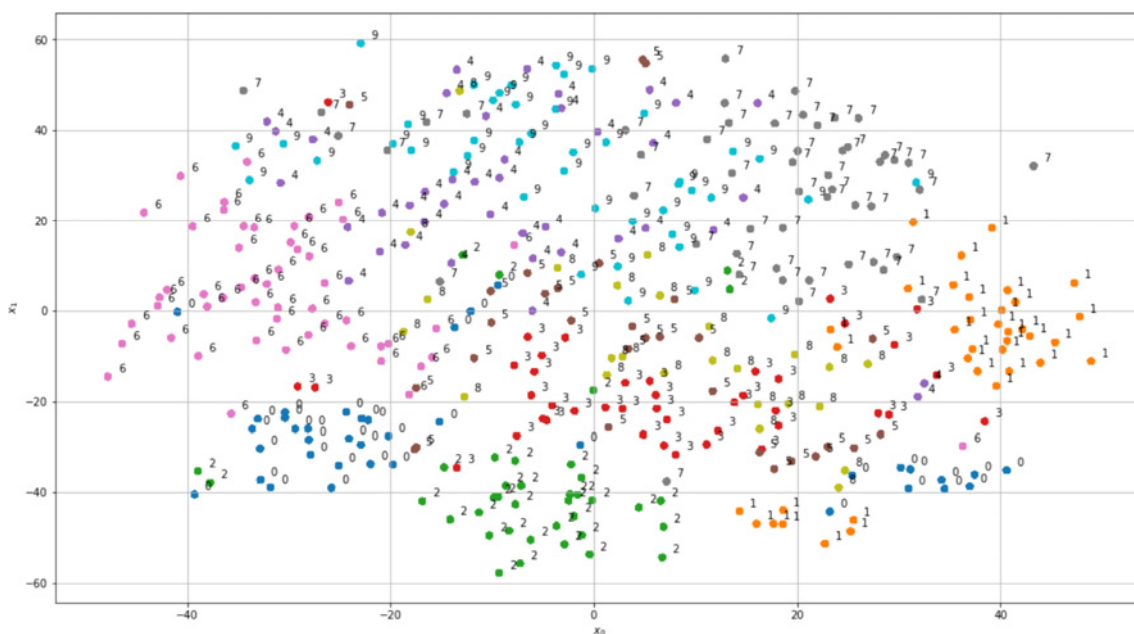
Rysunek 13.1. Przykład nieukierunkowanego grafu probabilistycznego



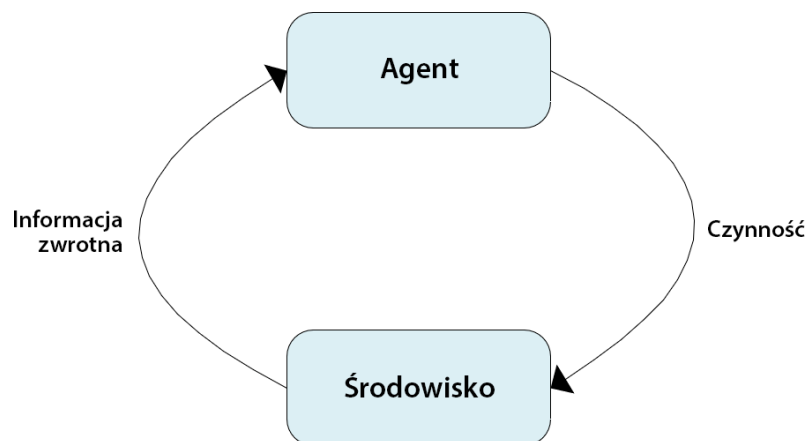
Rysunek 13.2. Struktura ograniczonej maszyny Boltzmann



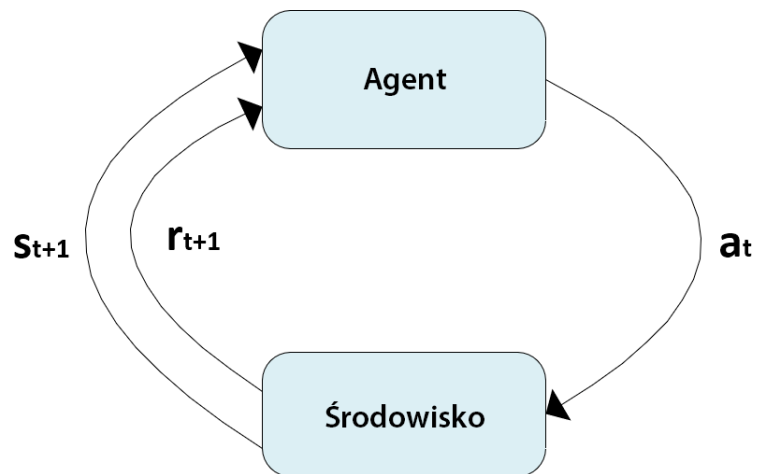
Rysunek 13.3. Ogólna struktura głębokiej sieci przekonań



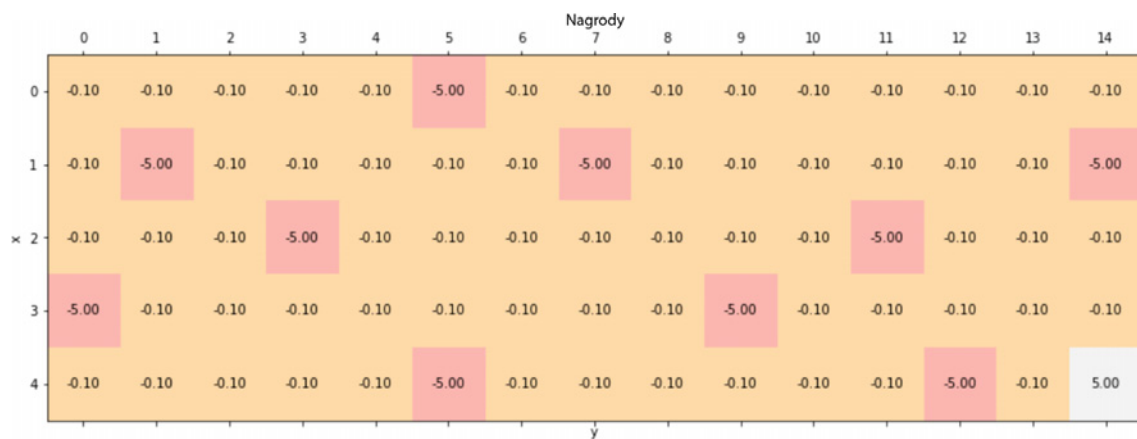
Rysunek 13.4. Wykres t-SNE rozkładu w ostatniej warstwie ukrytej (64-wymiarowej) sieci DBN



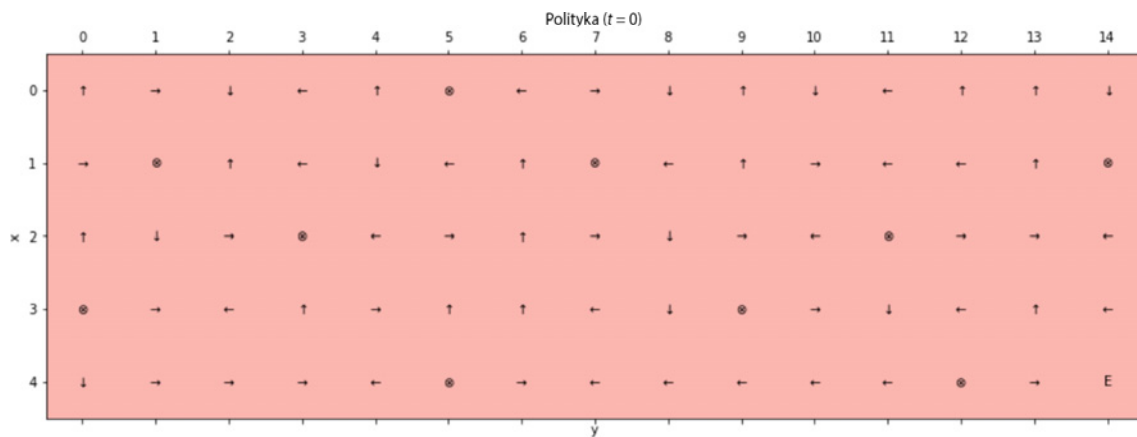
Rysunek 14.1. Ogólny schemat uczenia przez wzmacnianie



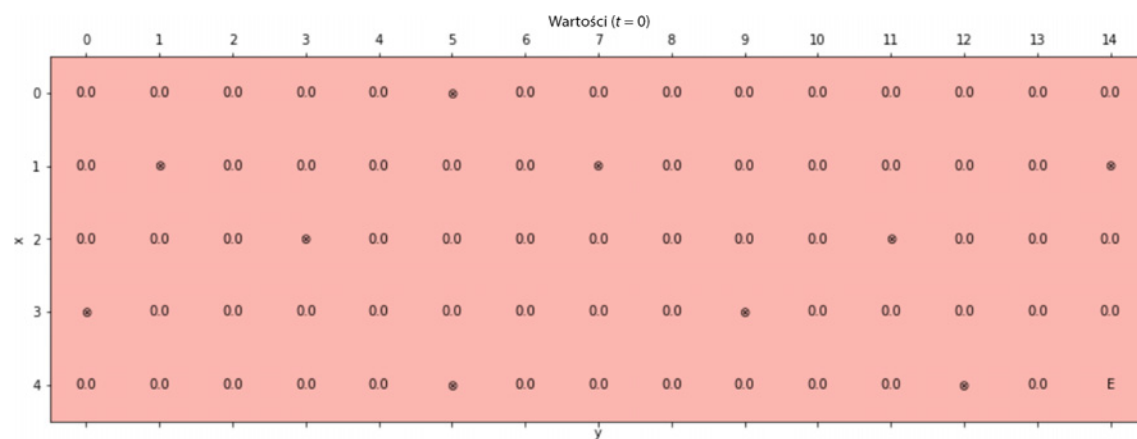
Rysunek 14.2. Schemat oddziaływania pomiędzy agentem a środowiskiem



Rysunek 14.3. Nagrody w środowisku tunelowym (szachownicy)



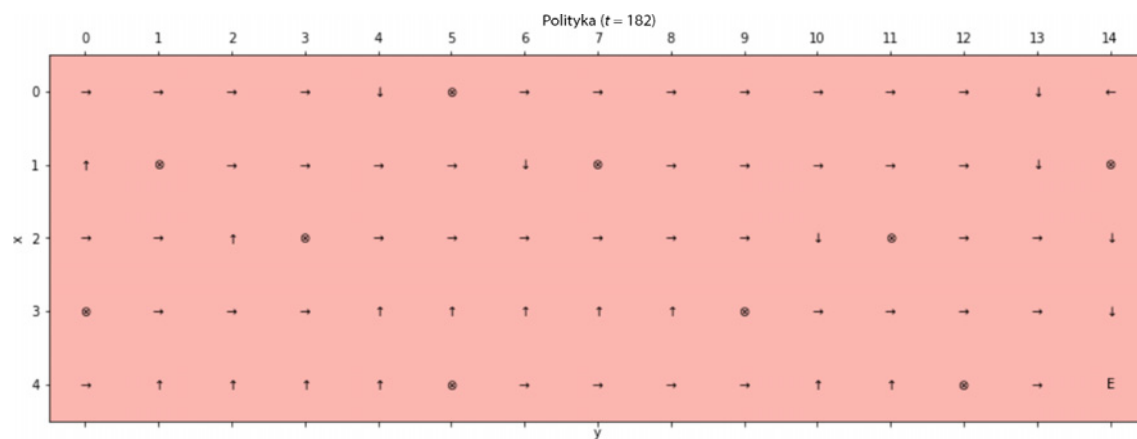
Rysunek 14.4. Początkowa ( $t = 0$ ), losowa polityka



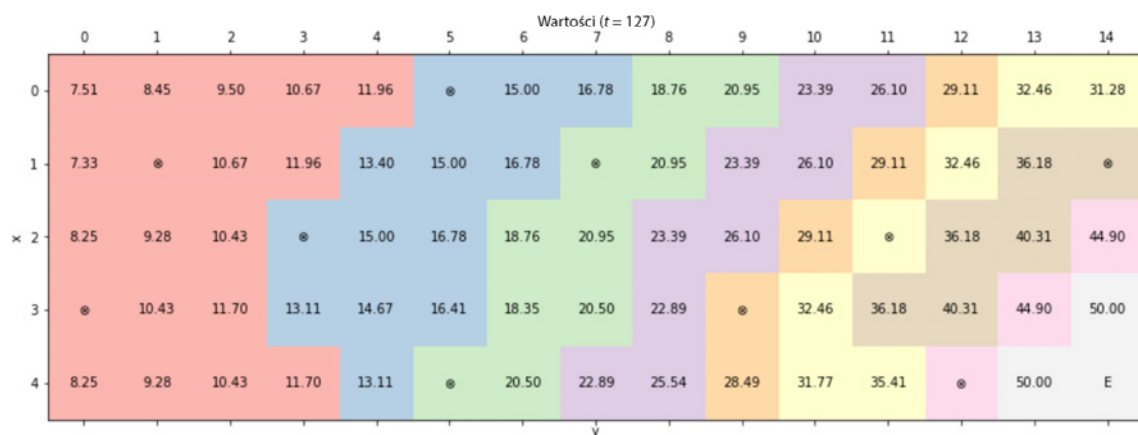
Rysunek 14.5. Początkowa ( $t = 0$ ) macierz wartości



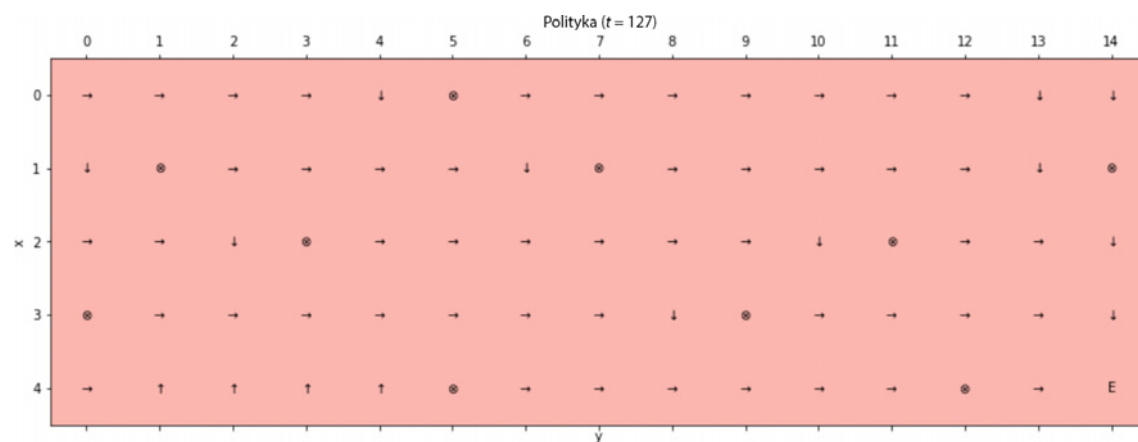
Rysunek 14.6. Końcowa macierz wartości



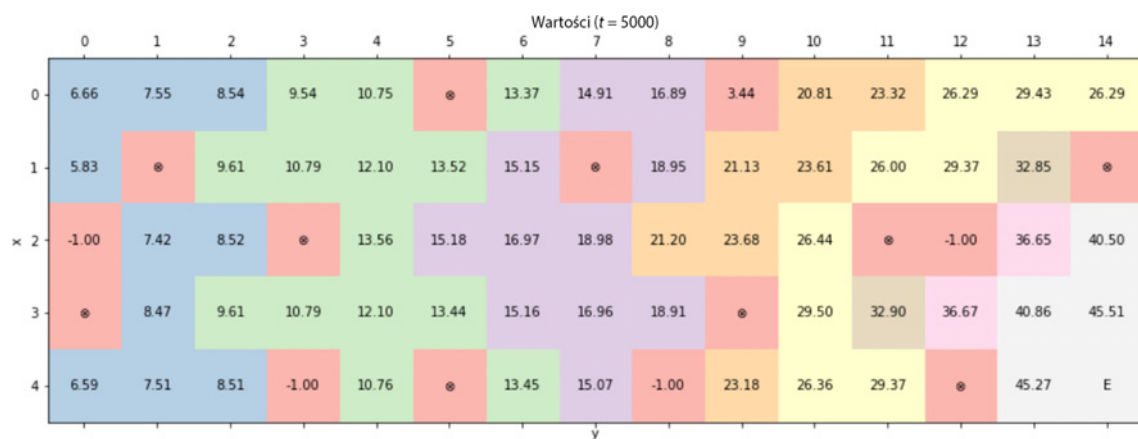
Rysunek 14.7. Polityka końcowa



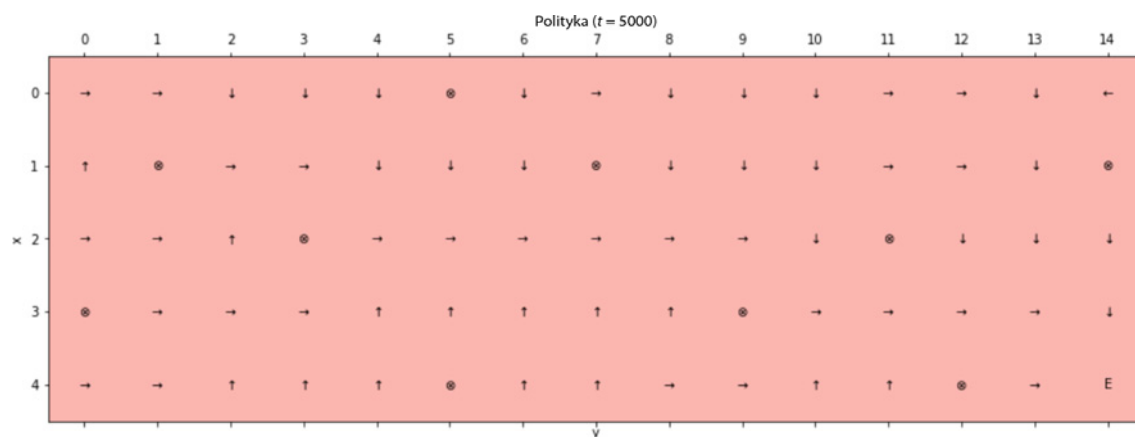
Rysunek 14.8. Końcowa macierz wartości



Rysunek 14.9. Wykres polityki końcowej



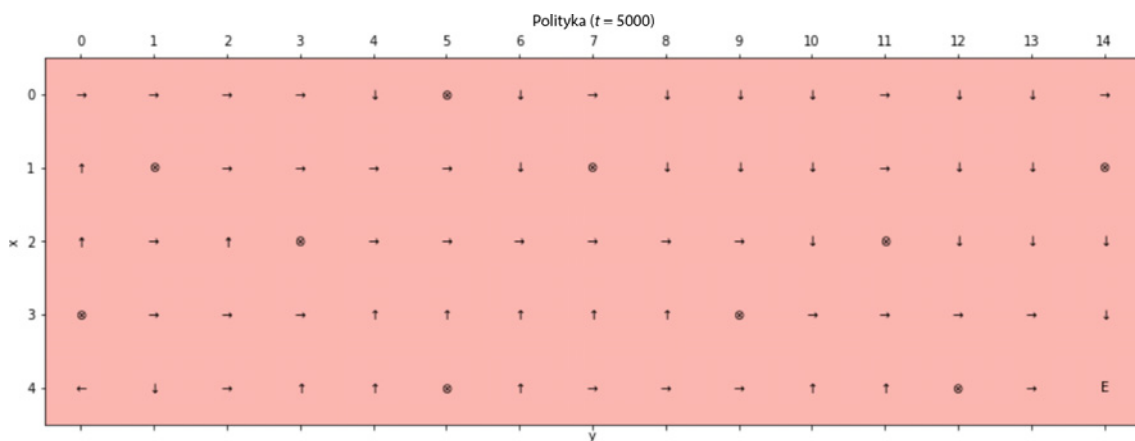
Rysunek 14.10. Macierz wartości końcowych dla losowych komórek początkowych



Rysunek 14.11. Polityka końcowa dla losowych komórek początkowych

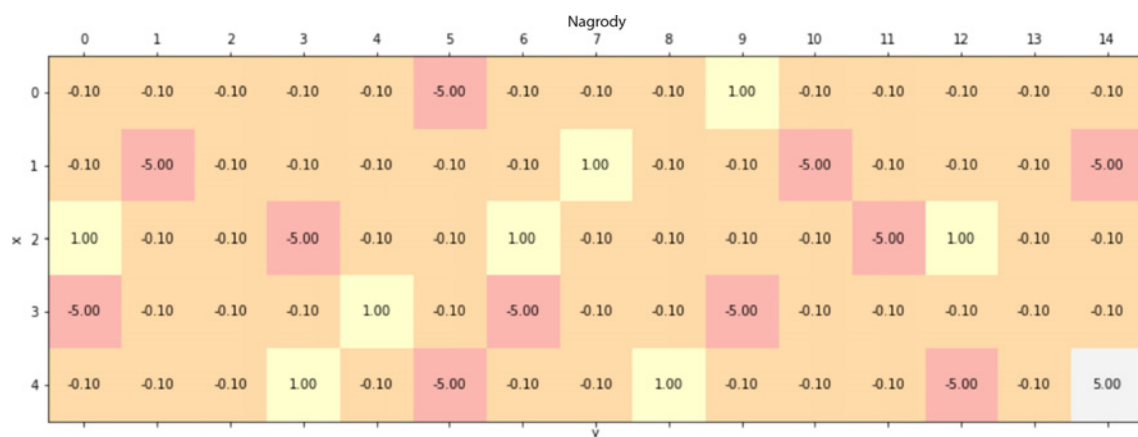


Rysunek 14.12. Macierz wartości końcowych dla wyznaczonego stanu początkowego (0, 0)

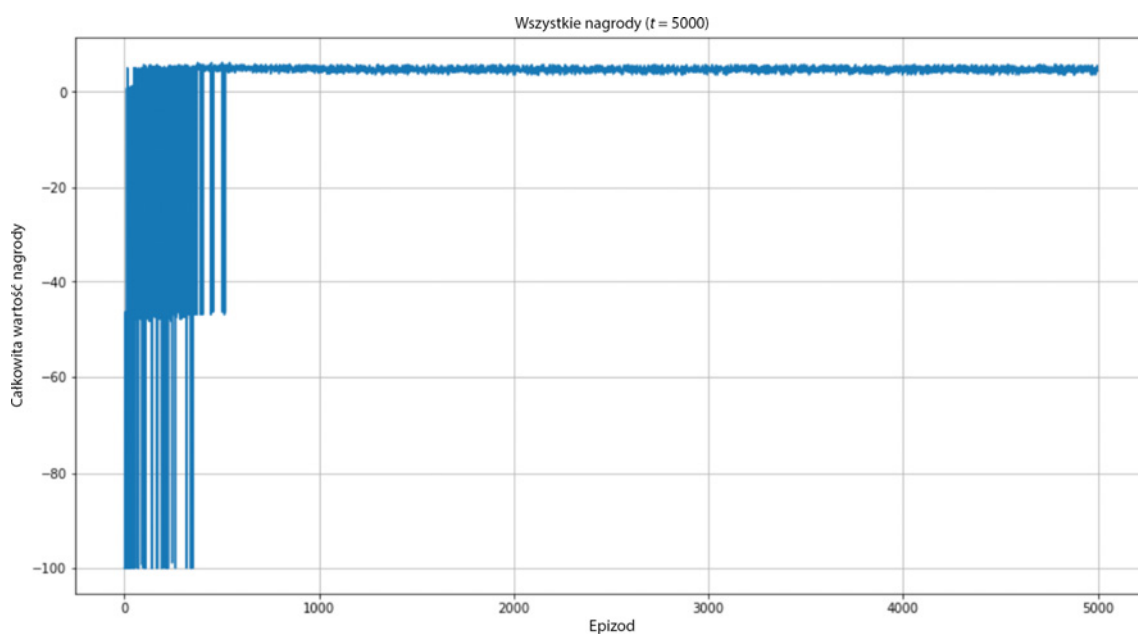


Rysunek 14.13. Polityka końcowa dla wyznaczonego stanu początkowego (0, 0)

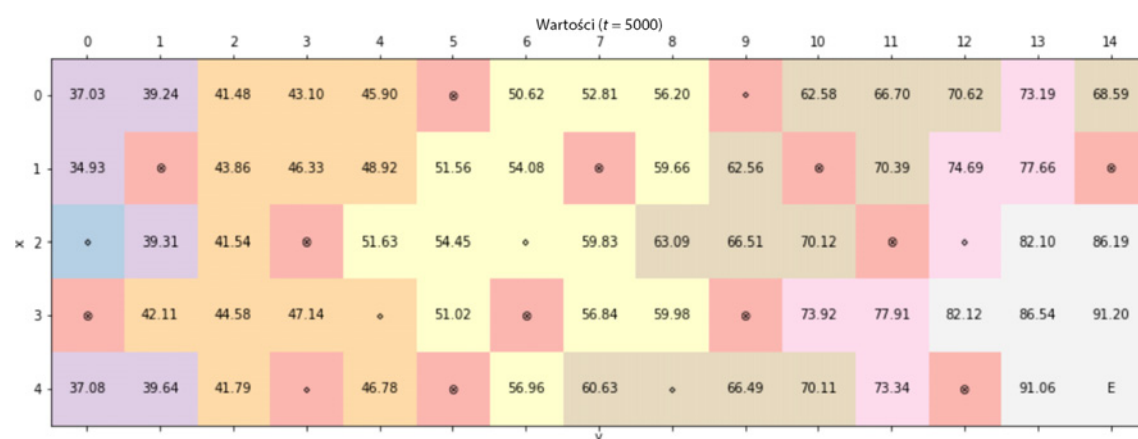




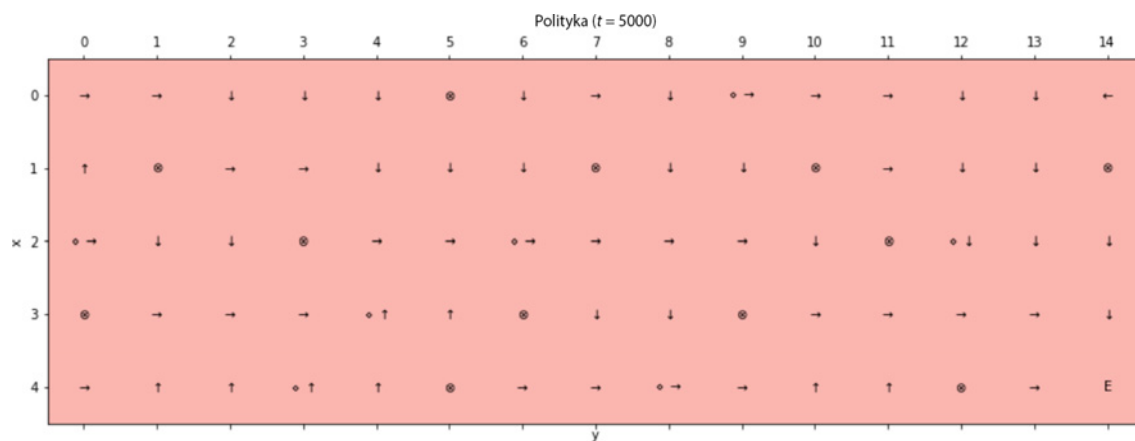
Rysunek 15.1. Schemat nagród w nowym środowisku tunelowym



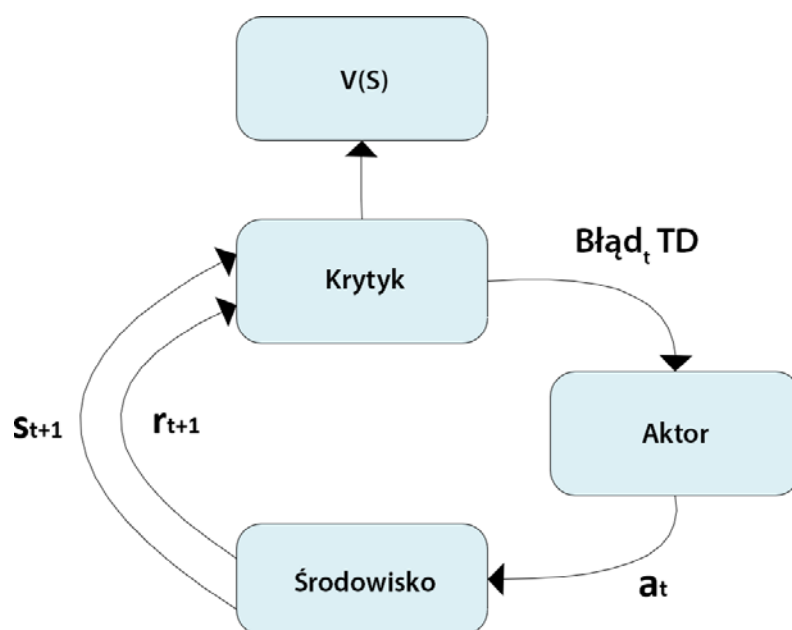
Rysunek 15.2. Wszystkie nagrody uzyskane przez agenta



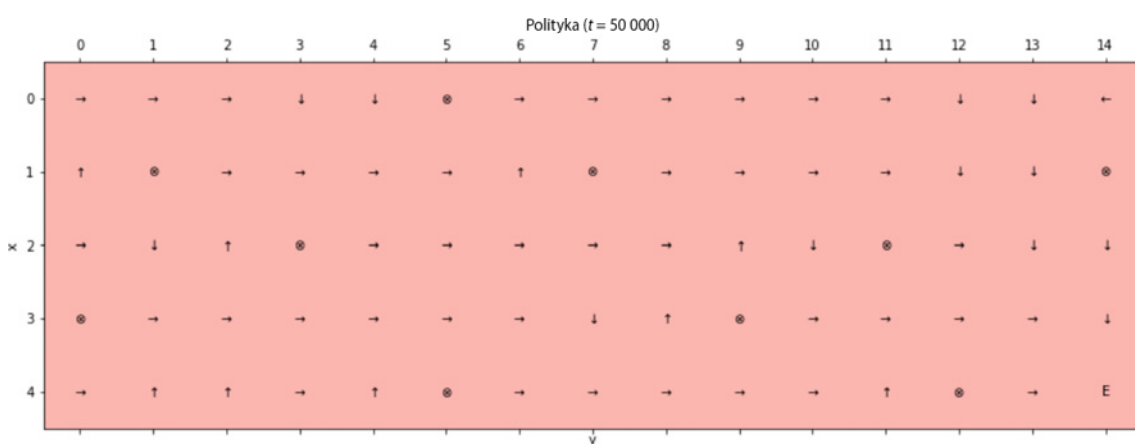
Rysunek 15.3. Końcowa macierz wartości



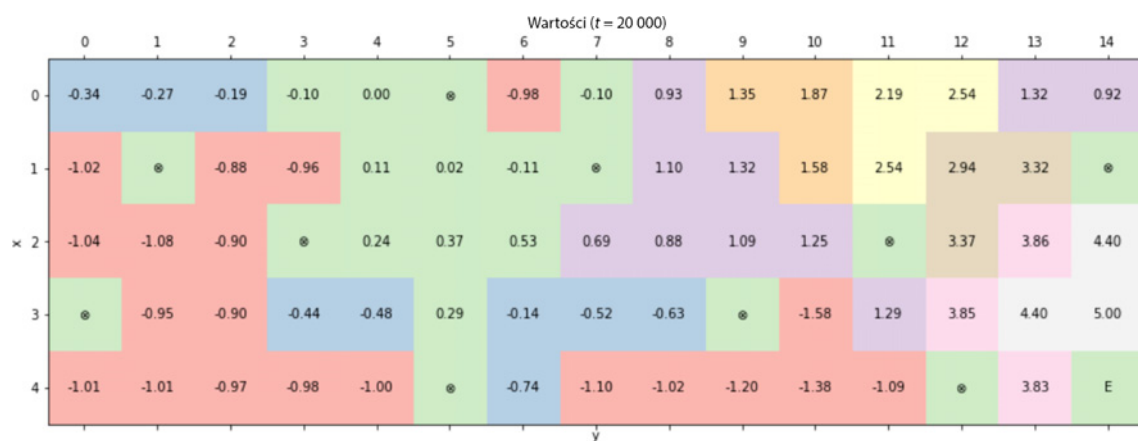
Rysunek 15.4. Polityka końcowa



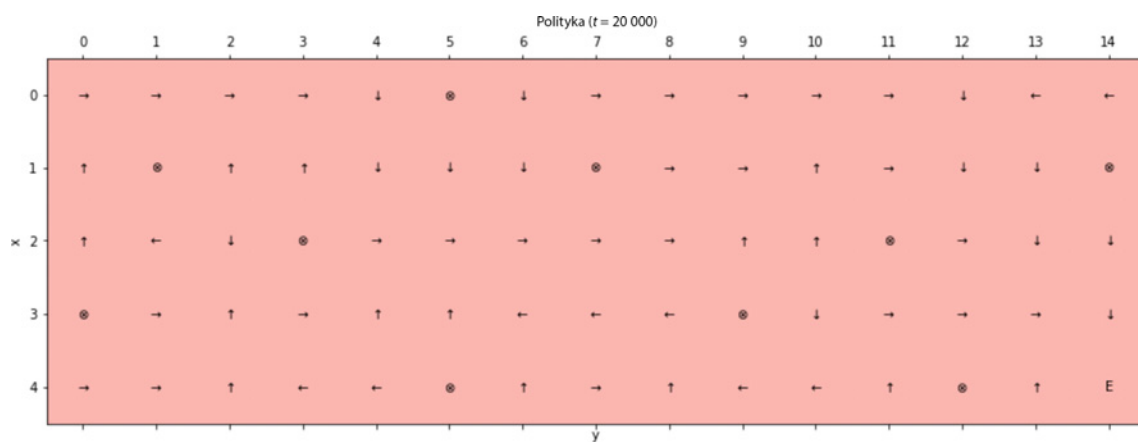
Rysunek 15.5. Schemat algorytmu aktor-krytyk



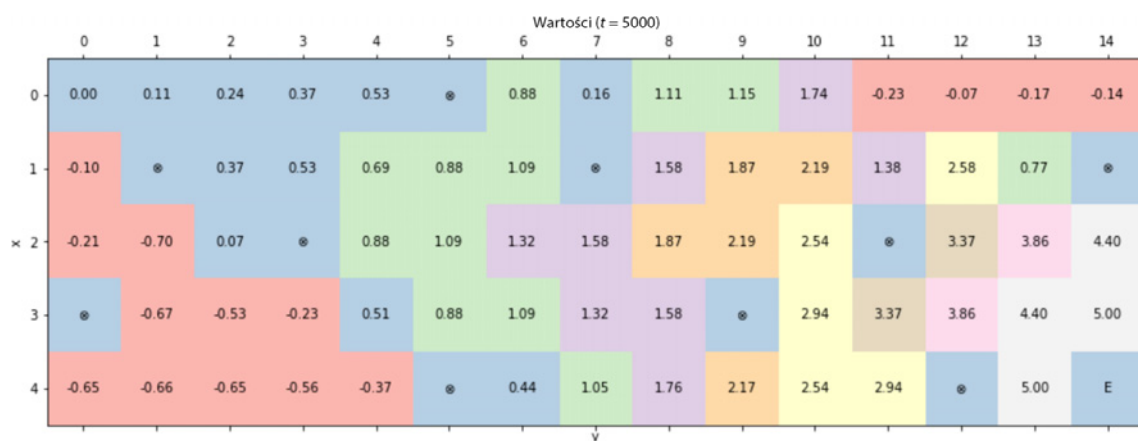
Rysunek 15.6. Końcowa polityka zachłanna



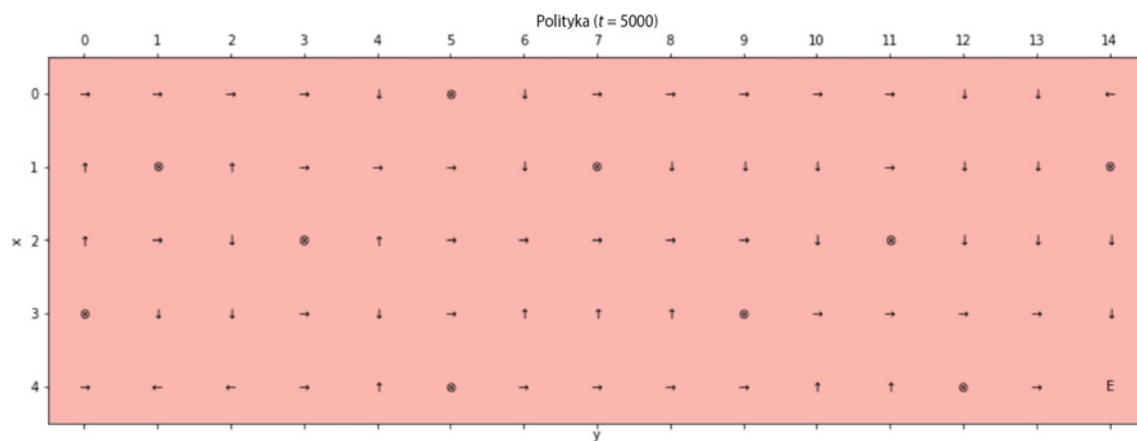
Rysunek 15.7. Końcowa macierz wartości (wykres funkcji  $V(s) = \max_a Q(s, a)$ )



Rysunek 15.8. Polityka końcowa



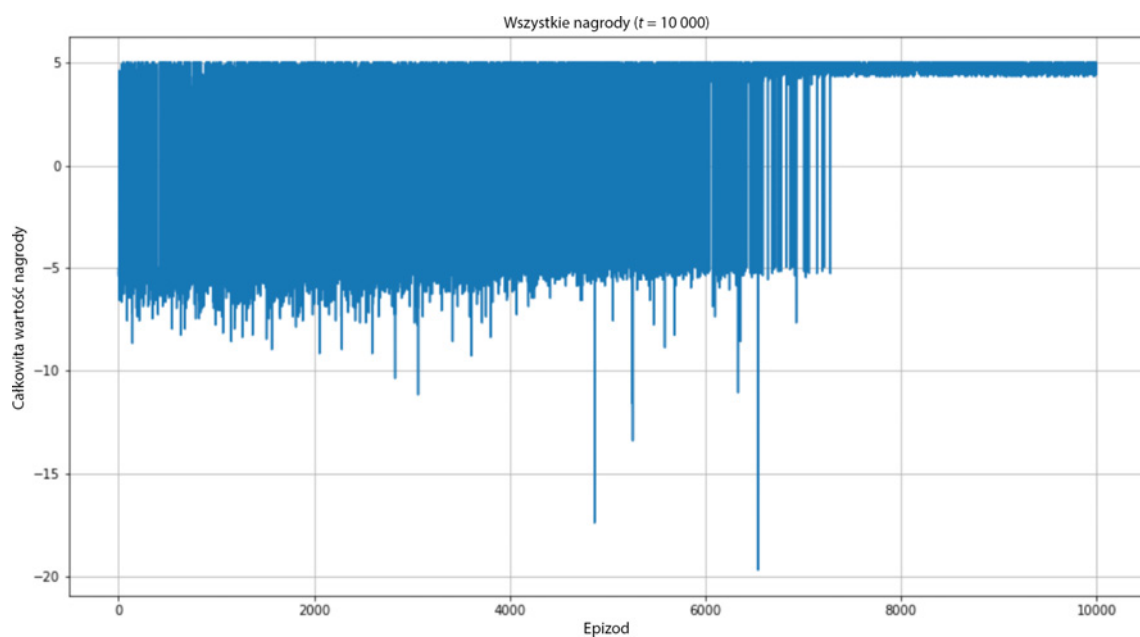
Rysunek 15.9. Macierz wartości końcowych



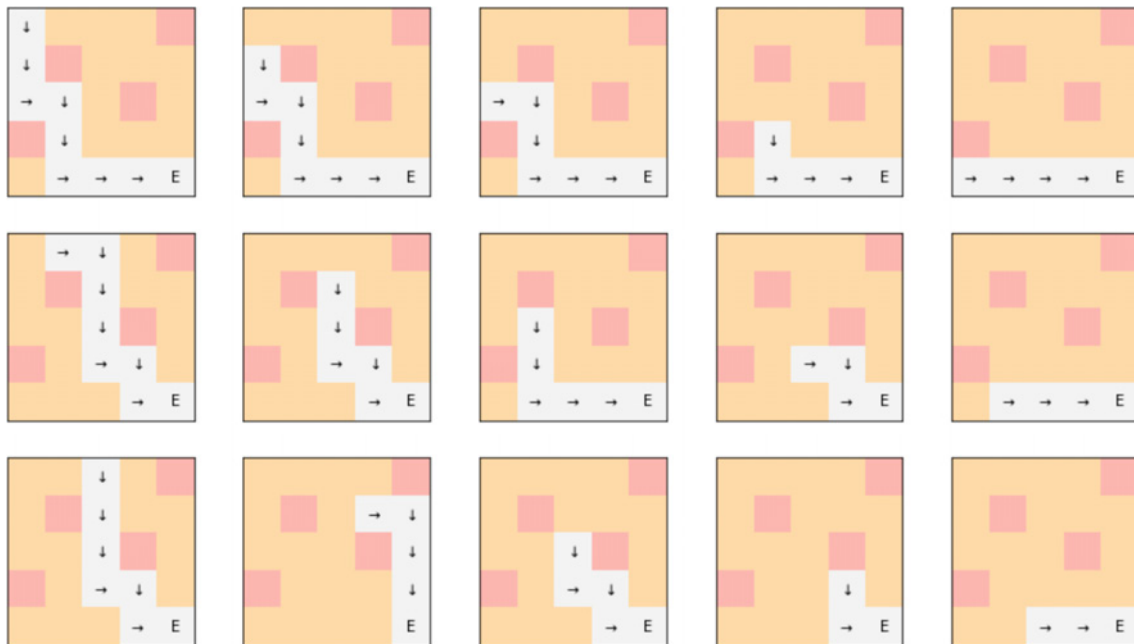
Rysunek 15.10. Polityka końcowa



Rysunek 15.11. Nagrody w mniejszym środowisku szachownicy



Rysunek 15.12. Suma nagród uzyskana za pomocą algorytmu Q-uczenia w postaci sieci neuronowej



Rysunek 15.13. Dwanaście trajektorii wygenerowanych przez politykę zachłanną