

---

# **Perceptron prosty i Adaline**

## **Sieci Neuronowe 2020**

---

**Jakub Ciszek**  
238035

## Spis treści

<b>1</b>	<b>Opis badań</b>	<b>3</b>
1.1	Plan eksperymentów . . . . .	3
1.2	Charakterystyka zbiorów danych . . . . .	3
<b>2</b>	<b>Ekspertymenty</b>	<b>4</b>
2.1	Wpływ wartości progu theta na szybkość uczenia Perceptronu . . . . .	4
2.2	Wpływ zakresu inicjalizacji wag na szybkość uczenia Perceptronu . . . . .	6
2.3	Wpływ wartości współczynnika uczenia alpha na szybkość uczenia Perceptronu . . . . .	7
2.4	Wpływ funkcji aktywacyjnej (unipolarna, bipolarna) na szybkość uczenia Perceptronu . . . . .	8
2.5	Wpływ zakresu inicjalizacji wag na szybkość uczenia Adaline . . . . .	9
2.6	Wpływ wartości współczynnika uczenia alpha na szybkość uczenia Adaline . . . . .	10
2.7	Wpływ przyjętego dopuszczalnego błędu na wynik uczenia w Adaline . . . . .	11
2.8	Porównanie Perceptronu i Adaline . . . . .	12
<b>3</b>	<b>Wnioski</b>	<b>14</b>

Cały kod wykorzystany w zadaniu znajduje się pod adresem: <https://github.com/Greenpp/sieci-neuronowe-pwr-2020>

## 1 Opis badań

### 1.1 Plan eksperymentów

Wszystkie eksperymenty zostały przeprowadzone 100 razy. Losowość przy inicjalizacji wag oraz generacji danych nie została narzucona żadnym ziarnem. Podczas badań przyjęto górną granicę 1000 epok, po przekroczeniu której, uczenie zostawało przerywane. Jeśli model przekroczył tę granicę przy pierwszych 6 wyuczeniach, zostawał klasyfikowany jako niezdolny do wyuczenia i pomijany. Zgodnie z instrukcją zostały przeprowadzone następujące badania:

- Wpływ wartości progu theta na szybkość uczenia Perceptronu
- Wpływ zakresu inicjalizacji wag na szybkość uczenia Perceptronu
- Wpływ wartości współczynnika uczenia alpha na szybkość uczenia Perceptronu
- Wpływ funkcji aktywacyjnej (unipolarna, bipolarna) na szybkość uczenia Perceptronu
- Wpływ zakresu inicjalizacji wag na szybkość uczenia Adaline
- Wpływ wartości współczynnika uczenia alpha na szybkość uczenia Adaline
- Wpływ przyjętego dopuszczalnego błędu na wynik uczenia w Adaline
- Porównanie Perceptronu i Adaline

### 1.2 Charakterystyka zbiorów danych

Dane użyte w zadaniu są reprezentacją logicznej funkcji AND

Tabela 1: Wartości funkcji AND

$a$	$b$	$a \wedge b$
0	0	0
0	1	0
1	1	1
1	0	0

W trakcie eksperymentów wykorzystano następujące zbiory:

- Zbiór uczący: 4 podstawowe wzorce oraz 4-krotna kopia każdego z nich z przesuniętymi wartościami wejściowymi o  $\pm -0.01 - 0.01$
- Zbiór walidujący: 4 podstawowe wzorce
- Zbiór testowy: 4 podstawowe wzorce

## 2 Eksperymenty

### 2.1 Wpływ wartości progu theta na szybkość uczenia Perceptronu

#### Założenia

Tabela 2: Stałe dla eksperymentu 1

Parametr	Wartość
Bias	Nie
Zakres wag	-0.2 – 0.2
Współczynnik uczenia	0.01

Zmienną w tym eksperymencie była wartość theta dla funkcji aktywacji. Przyjmowała wartości ze zbioru  $\{-1.0, -0.8, -0.5, -0.2, 0.0, 0.2, 0.5, 0.8, 1.0\}$

#### Przebieg

Podczas eksperymentu model został zainicjalizowany 100 razy dla każdej z badanych wartości oraz wyuczony, uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy. Badanie przeprowadzono dla funkcji aktywacyjnej unipolarnej jak i bipolarnej.

#### Wyniki

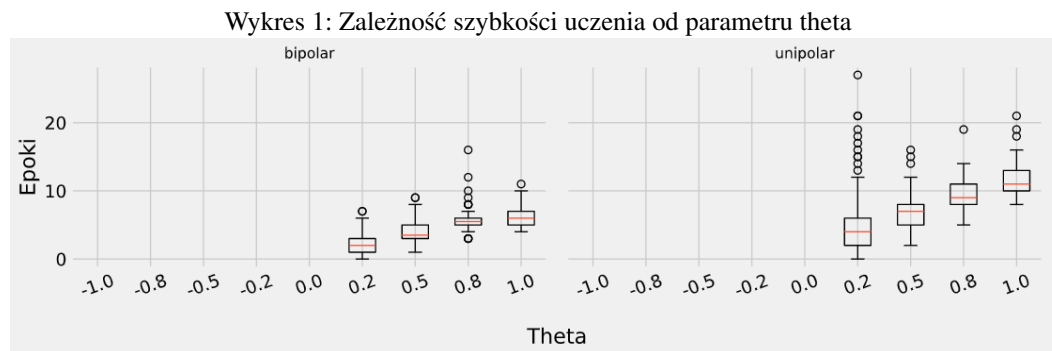


Tabela 3: Średnia ilość epok potrzebna do wyuczenia w zależności od parametru theta

Theta	Epoki	
	bipolarna	unipolarna
-1.0	-	-
-0.8	-	-
-0.5	-	-
-0.2	-	-
0.0	-	-
0.2	<b>2.27</b>	<b>5.47</b>
0.5	3.88	6.77
0.8	5.64	9.46
1.0	6.36	11.57

#### Wnioski

Z otrzymanych wyników, widocznych na wykresie 1 oraz tabeli 3, wynika że wyuczenie modelu nie było możliwe dla progu  $\theta \leq 0$ . Jest to zgodne z jego matematyczną interpretacją jako funkcji

liniowej. Przy progu równym zero nie jest możliwe oderwanie od początku układu współrzędnych co uniemożliwia rozdzielanie zadanych punktów. Podczas gdy ujemne wartości sprawiają, że dla wymaganej orientacji prostej suma ujemnych wag jest większa od wartości progu co powoduje odwrócenie otrzymywanych klas. Kolejną obserwacją jest mniejsza ilość epok potrzebna do wyuczenia przy mniejszym, dodatnim progu. Ta zależność także wynika z charakterystyki równania liniowego, które do zmiany orientacji potrzebuje znacznie mniejszych współczynników  $a$  i  $b$  dla małego  $c$ , co powoduje zmniejszenie liczby kroków, a z tym epok.

## 2.2 Wpływ zakresu inicjalizacji wag na szybkość uczenia Perceptronu

### Założenia

Tabela 4: Stałe dla eksperymentu 2

Parametr	Wartość
Bias	Tak
Theta	0.0
Współczynnik uczenia	0.01

Zmienną w tym eksperymencie była wartość początkowego zakresu wag. Przyjmowała wartości ze zbioru  $\{0.0, -0.1 - 0.1, -0.2 - 0.2, -0.5 - 0.5, -0.8 - 0.8, -1.0 - 1.0\}$

### Przebieg

Podczas eksperymentu model został zainicjalizowany 100 razy dla każdej z badanych wartości oraz wyuczony, uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy. Badanie przeprowadzono dla funkcji aktywacyjnej unipolarnej jak i bipolarnej.

### Wyniki

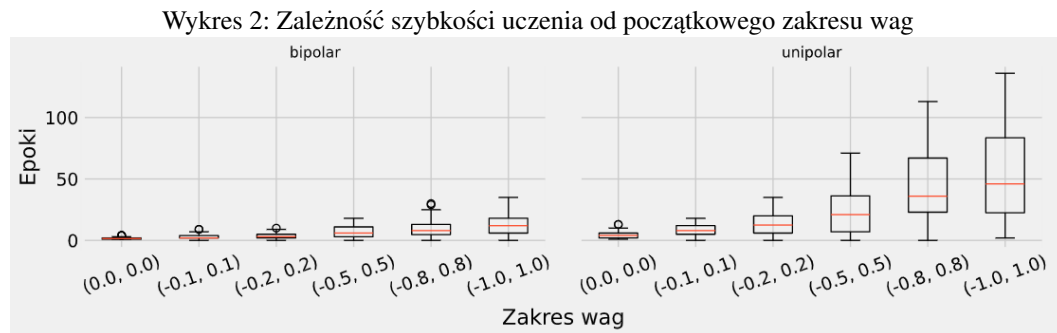


Tabela 5: Średnia ilość epok potrzebna do wyuczenia w zależności od początkowego zakresu wag

Zakres wag	Epoki	
	bipolarna	unipolarna
0.0	<b>1.79</b>	<b>4.62</b>
-0.1 – 0.1	2.83	8.44
-0.2 – 0.2	3.62	13.00
-0.5 – 0.5	7.14	24.01
-0.8 – 0.8	9.29	41.41
-1.0 – 1.0	13.00	52.40

### Wnioski

Z otrzymanych wyników, widocznych na wykresie 2 oraz tabeli 5, wynika że uczenie następowało dużo szybciej dla początkowych wartości wag bliższych 0. Te wyniki zgadzają się z zachowaniem modelu jako równania liniowego - przy mniejszych wartościach potrzebne są mniejsze zmiany aby odpowiednio zorientować prostą. Takie wyniki jednak nie powinny wystąpić podczas badania głębszych modeli gdzie początkowe wartości równe 0 mogą prowadzić do jednakowych zmian wszystkich wag w modelu i upośledzenia jego możliwości uczenia.

## 2.3 Wpływ wartości współczynnika uczenia alpha na szybkość uczenia Perceptronu

### Założenia

Tabela 6: Stałe dla eksperymentu 3

Parametr	Wartość
Bias	Tak
Theta	0.0
Zakres wag	-0.5 – 0.5

Zmienną w tym eksperymencie była wartość współczynnika uczenia. Przyjmowała wartości ze zbioru  $\{0.0001, 0.001, 0.01, 0.1, 1.0\}$

### Przebieg

Podczas eksperymentu model został zainicjalizowany 100 razy dla każdej z badanych wartości oraz wyuczony, uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy. Badanie przeprowadzono dla funkcji aktywacyjnej unipolarnej jak i bipolarnej.

### Wyniki

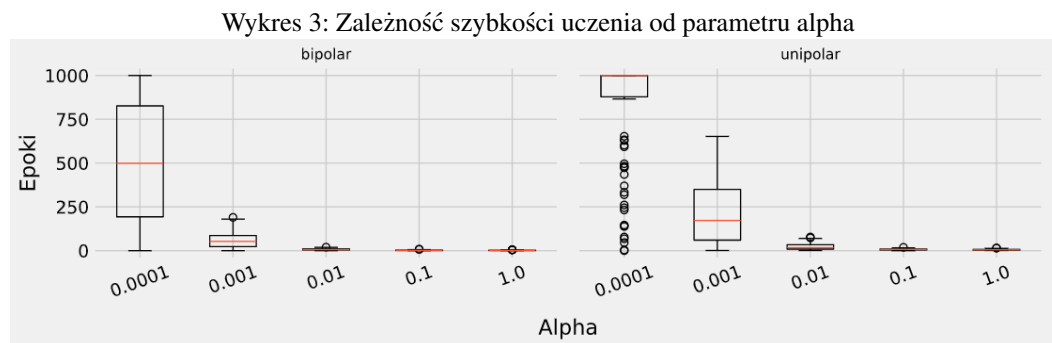


Tabela 7: Średnia ilość epok potrzebna do wyuczenia w zależności od parametru alpha

Alpha	Epoki	
	bipolarna	unipolarna
0.0001	495.44	836.04
0.0010	57.52	223.90
0.0100	6.88	22.60
0.1000	2.33	6.80
1.0000	<b>1.86</b>	<b>5.68</b>

### Wnioski

Z otrzymanych wyników, widocznych na wykresie 3 oraz tabeli 7, wynika że mniejsze wartości współczynnika uczenia spowalniają ten proces. Z jednej strony takie wyniki są poprawne, ponieważ mniejszy współczynnik przyczynia się do mniejszych zmian wag, jednak z drugiej wysoka wartość nie pozwala na tak dobre wyuczenie modelu. W tym przypadku dokładność nie była potrzebna, co powoduje że najlepsze efekty osiągnął wysoki współczynnik alpha.

## 2.4 Wpływ funkcji aktywacyjnej (unipolarna, bipolarna) na szybkość uczenia Perceptronu

### Wyniki

Tabela 8: Średnia ilość epok potrzebna do wyuczenia w zależności od parametru theta

Theta	Epoki	
	bipolarna	unipolarna
-1.0	-	-
-0.8	-	-
-0.5	-	-
-0.2	-	-
0.0	-	-
0.2	<b>2.27</b>	5.47
0.5	<b>3.88</b>	6.77
0.8	<b>5.64</b>	9.46
1.0	<b>6.36</b>	11.57

Tabela 9: Średnia ilość epok potrzebna do wyuczenia w zależności od początkowego zakresu wag

Zakres wag	Epoki	
	bipolarna	unipolarna
0.0	<b>1.79</b>	4.62
-0.1 – 0.1	<b>2.83</b>	8.44
-0.2 – 0.2	<b>3.62</b>	13.00
-0.5 – 0.5	<b>7.14</b>	24.01
-0.8 – 0.8	<b>9.29</b>	41.41
-1.0 – 1.0	<b>13.00</b>	52.40

Tabela 10: Średnia ilość epok potrzebna do wyuczenia w zależności od parametru alpha

Alpha	Epoki	
	bipolarna	unipolarna
0.0001	<b>495.44</b>	836.04
0.0010	<b>57.52</b>	223.90
0.0100	<b>6.88</b>	22.60
0.1000	<b>2.33</b>	6.80
1.0000	<b>1.86</b>	5.68

### Wnioski

Z wyników otrzymanych w poprzednich badaniach, widocznych w tabelach 8, 9 i 10, wynika jednoznacznie, że lepsze efekty otrzymuje się przy użyciu funkcji bipolarnej. Powodem tego jest o wiele mniejsza odległość od początku układu współrzędnych do obszaru w którym powinna znaleźć się separująca dane linia, co zmniejsza liczbę potrzebnych na przesunięcie jej tam epok.



## 2.5 Wpływ zakresu inicjalizacji wag na szybkość uczenia Adaline

### Założenia

Tabela 11: Stałe dla eksperymentu 5

Parametr	Wartość
Bias	Tak
Theta	0.0
Współczynnik uczenia	0.01
Epsilon	0.2

Zmienną w tym eksperymencie była wartość początkowego zakresu wag. Przyjmowała wartości ze zbioru  $\{0.0, -0.1 - 0.1, -0.2 - 0.2, -0.5 - 0.5, -0.8 - 0.8, -1.0 - 1.0\}$

### Przebieg

Podczas eksperymentu model został zainicjalizowany 100 razy dla każdej z badanych wartości oraz wyuczony, uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy.

### Wyniki

Wykres 4: Zależność szybkości uczenia od początkowego zakresu wag

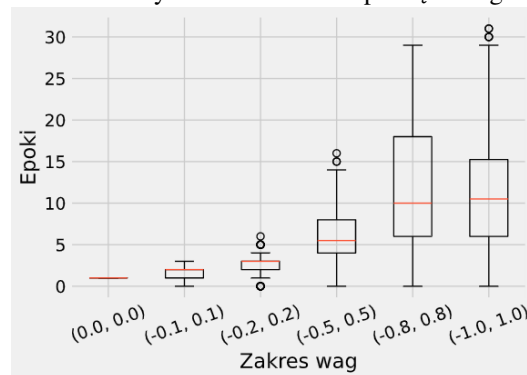


Tabela 12: Średnia ilość epok potrzebna do wyuczenia w zależności od początkowego zakresu wag

Zakres wag	Epoki
0.0	<b>1.00</b>
-0.1 – 0.1	1.61
-0.2 – 0.2	2.51
-0.5 – 0.5	5.93
-0.8 – 0.8	11.54
-1.0 – 1.0	11.95

### Wnioski

Z otrzymanych wyników, widocznych na wykresie 4 oraz tabeli 12, wynika że uczenie następowało dużo szybciej dla początkowych wartości wag bliższych 0. Te wyniki zgadzają się z zachowaniem modelu jako równania liniowego - przy mniejszych wartościach potrzebne są mniejsze zmiany aby odpowiednio zorientować prostą. Takie wyniki jednak nie powinny wystąpić podczas badania głębszych modeli gdzie początkowe wartości równe 0 mogą prowadzić do jednakowych zmian wszystkich wag w modelu i upośledzenia jego możliwości uczenia.

## 2.6 Wpływ wartości współczynnika uczenia alpha na szybkość uczenia Adaline

### Założenia

Tabela 13: Stałe dla eksperymentu 6

Parametr	Wartość
Bias	Tak
Theta	0.0
Zakres wag	-0.5 – 0.5
Epsilon	0.2

Zmienną w tym eksperymencie była wartość współczynnika uczenia. Przyjmowała wartości ze zbioru  $\{0.0001, 0.001, 0.01, 0.1, 1.0\}$

### Przebieg

Podczas eksperymentu model został zainicjalizowany 100 razy dla każdej z badanych wartości oraz wyuczony, uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy.

### Wyniki

Wykres 5: Zależność szybkości uczenia od parametru alpha

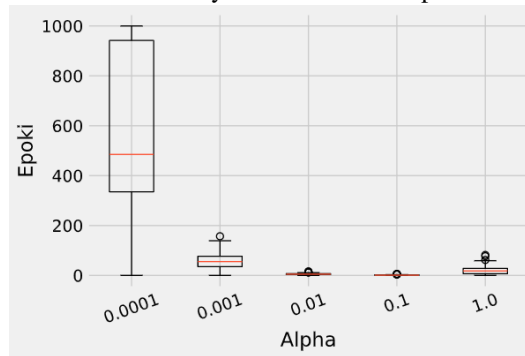


Tabela 14: Średnia ilość epok potrzebna do wyuczenia w zależności od parametru alpha

Alpha	Epoki
0.0001	567.96
0.0010	55.87
0.0100	5.56
0.1000	<b>1.64</b>
1.0000	20.81

### Wnioski

Z otrzymanych wyników, widocznych na wykresie 5 oraz tabeli 14, wynika że optymalnym współczynnikiem w przypadku Adaline jest 0.1. Jest to zapewne spowodowane możliwością otrzymania większych błędów niż w przypadku Perceptronu, gdzie ich zeskalowanie w dół pozwala na szybsze osiągnięcie, bardzo bliskiego środka układu, obszaru umożliwiającego podział danych.

## 2.7 Wpływ przyjętego dopuszczalnego błędu na wynik uczenia w Adaline

### Założenia

Tabela 15: Stałe dla eksperymentu 7

Parametr	Wartość
Bias	Tak
Theta	0.0
Zakres wag	-0.5 – 0.5
Współczynnik uczenia	0.01
Epsilon	0.0

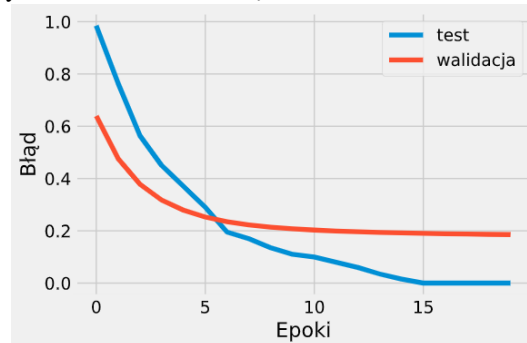
### Przebieg

Podczas eksperymentu model został zainicjalizowany 100 raz, po czym zostało sprawdzone zachowanie błędu. uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy. Zapisane zostały błędy:

- walidacyjny - obliczany w tym samym miejscu co błąd treningowy
- testowy - obliczany już po progowaniu

### Wyniki

Wykres 6: Zachowanie błędu w trakcie uczenia Adaline



Minimalna zaobserwowana wartość błędu walidacji: 0.1303

### Wnioski

Uzyskane wyniki, widoczne na wykresie 6 zostały ograniczone do 20 epok z powodu zerowego błędu testowego podczas dalszego uczenia. Minimalny otrzymany błąd po odwróceniu funkcji błędu daje wartość  $\simeq 1$ . Natomiast na wspomnianym wykresie widać, że błąd testowy przyjmuje zero w miarę zbliżania się błędu walidacyjnego do 0.2, co po takim samym przeliczeniu daje wartość błędu  $\simeq 1.27$  na punkt. Wartości te można by zapewne otrzymać znajdując minimalne oraz optymalne ustawienie linii rozwiązując układ równań, ale z powodu braku czasu ten etap został pominięty.

## 2.8 Porównanie Perceptronu i Adaline

### Wyniki

Wykres 7: Porównanie szybkości uczenia Perceptronu i Adaline w zależności od początkowego zakresu wag

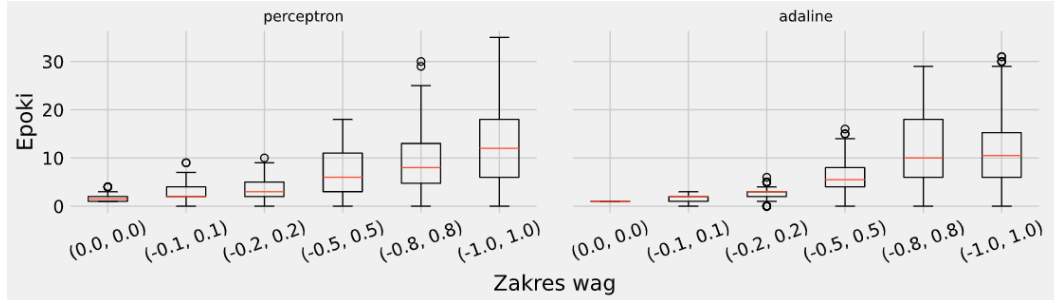


Tabela 16: Porównanie średniej ilości epok potrzebnych do wyuczenia Perceptronu i Adaline w zależności od początkowego zakresu wag

Zakres wag	Epoki	
	Perceptron	Adaline
0.0	1.79	<b>1.00</b>
-0.1 – 0.1	2.83	<b>1.61</b>
-0.2 – 0.2	3.62	<b>2.51</b>
-0.5 – 0.5	7.14	<b>5.93</b>
-0.8 – 0.8	9.29	<b>11.54</b>
-1.0 – 1.0	13.00	<b>11.95</b>

Wykres 8: Porównanie szybkości uczenia Perceptronu i Adaline w zależności od parametru alpha

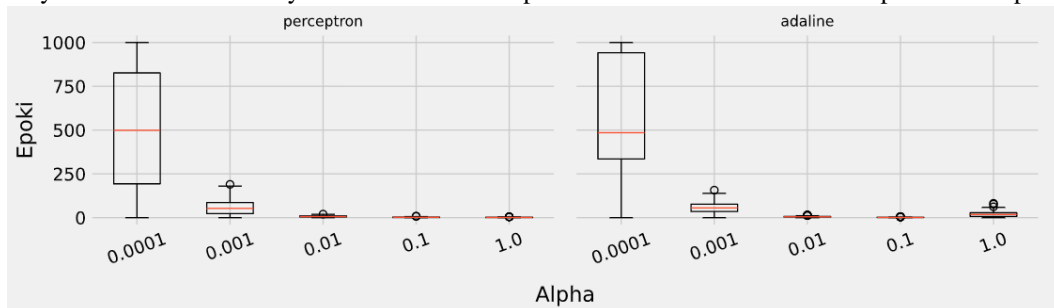


Tabela 17: Porównanie średniej ilości epok potrzebnych do wyuczenia Perceptronu i Adaline w zależności od parametru alpha

Alpha	Epoki	
	Perceptron	Adaline
0.0001	<b>495.44</b>	567.96
0.0010	57.52	<b>55.87</b>
0.0100	6.88	<b>5.56</b>
0.1000	2.33	<b>1.64</b>
1.0000	<b>1.86</b>	20.81

## **Wnioski**

Z wyników otrzymanych w poprzednich badaniach, widocznych na wykresach 16 i 17 oraz w tabelach 7 i 8, wynika, że lepiej radzi sobie Adaline. Powodem tego jest najprawdopodobniej bardziej plastyczny sygnał zwrotny w tym modelu, który nie przyjmuje wartości dyskretnych jak w przypadku Perceptronu.

### 3 Wnioski

- Perceptron tak jak i Adaline umożliwiają klasyfikację w problemach rozdzielnych liniowo.
- Możliwość uczenia względem ciągłego błędu oraz użycie logiki bipolarnej pozwala Adaline na uzyskanie lepszych wyników w tym zadaniu.
- Odpowiednia konfiguracja hiperparametrów może znacząco przyspieszyć proces uczenia, lub w ogóle go umożliwić.
- Odpowiednia reprezentacja danych ma duży wpływ na proces uczenia.