
Sprawozdanie 1

Konrad Gmyrek
247008

Abstract

Poniżej przeprowadzono kilka badań dla Perceptronu prostego oraz Adaline, w których badany jest wpływ różnych hiperparametrów na szybkość uczenia. Głównym celem tych badań było sprawdzenie jak poszczególne hiperparametry wpływają na modele oraz wyrobienie sobie intuicji odnośnie działania Perceptronu oraz Adaline.

Repozytorium kodu: <https://github.com/KonradGmy/Lab01>

1 Badania Perceptronu

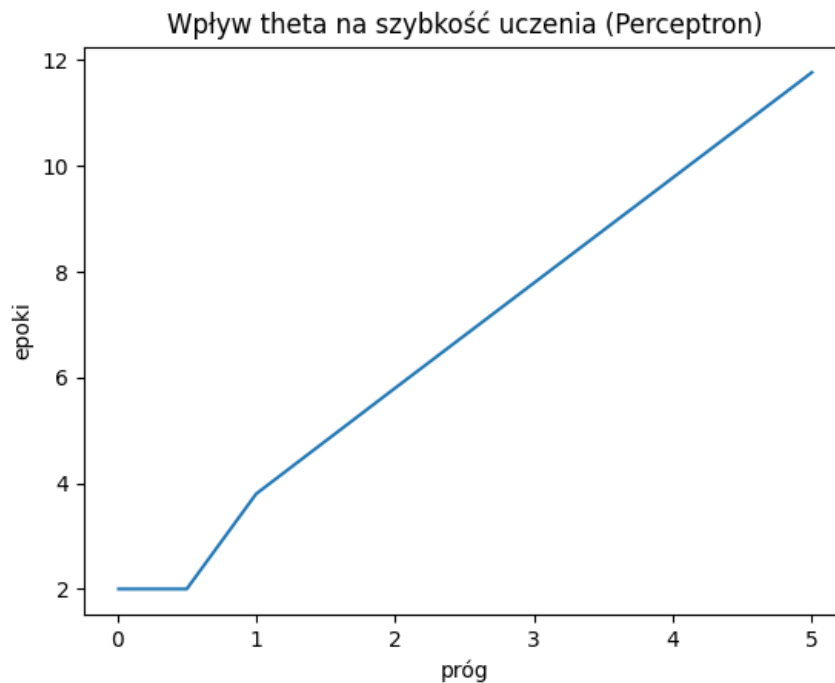
Poniżej zostały przedstawione wyniki badań dla Perceptronu. Perceptron symuluje działanie funkcji AND dla dwóch zmiennych w logice bipolarnej. Perceptron przestaje aktualizować wagi gdy w epoce nie pojawi się błąd. Każde badanie dla klarowności zostało przeprowadzone sto razy (wyniki dla dziesięciu powtórzeń często znacząco się różniły dlatego też zwiększyłem liczbę powtórzeń danego badania)

1.1 Badanie 1

Zbadany został wpływ progu θ na szybkość uczenia Perceptronu prostego, wagi początkowe wektora W zostały wybrane z przedziału $(-0,01, 0,01)$, wartość współczynnika uczenia α wynosi 0,1. Dla każdej wartości progu θ badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 1: Wpływ θ na szybkość uczenia

θ	0.01	0.1	0.5	1	2	5
epoki	2.0	2.0	2.0	3.8	5.8	11.77



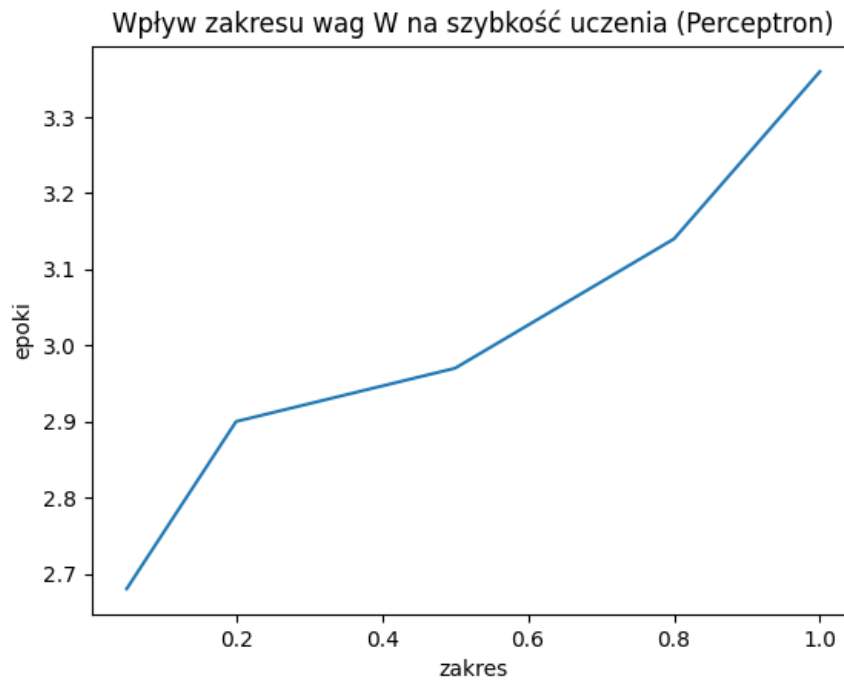
Wnioski: Zależność szybkości uczenia oraz progu θ jest widocznie liniowa, im większy próg tym średnio więcej epok potrzeba do wyuczenia neuronu. Dzieje się tak dlatego, że dla małych wag wektora W oraz dużego progu potrzeba więcej epok na znalezienie odpowiedniej prostej która będzie klasyfikować dane.

1.2 Badanie 2

Zbadany został wpływ zakresu w jakim losowane są początkowe wagi wektora W na szybkość uczenia. Waga początkowa progu θ została wylosowana z przedziału $(-1, 1)$, wartość współczynnika uczenia α wynosi 0,1. Dla każdej wartości zakresu badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 2: Wpływ zakresu wag W na szybkość uczenia

zakres	$(-1, 1)$	$(-0.8, 0.8)$	$(-0.5, 0.5)$	$(-0.2, 0.2)$	$(-0.05, 0.05)$
epoki	3.36	3.14	2.97	2.9	2.68



Wnioski: Zależność szybkości uczenia i zakresu początkowego wektora W jest ponownie liniowa, im bardziej losowe wartości wag tym średnio więcej epok potrzebnych do wyuczenia neuronu. Gdy współczynnik uczenia α jest mały, wartości wektora W zmieniają się powoli, a zatem potrzeba więcej epok zanim znajdziemy odpowiedni wektor W .

1.3 Badanie 3

Zbadany został wpływ współczynnika uczenia α na szybkość uczenia. Waga początkowa progu θ została wylosowana z przedziału $(-1, 1)$, wagi początkowe wektora W zostały wybrane z przedziału $(-0,01, 0,01)$. Dla każdej wartości współczynnika uczenia badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 3: Wpływ współczynnika uczenia na szybkość uczenia

α	0.01	0.1	0.2	0.3	0.4	0.5	1	2	5
epoki	10.01	2.51	2.21	2.34	2.36	2.49	2.56	2.42	2.56



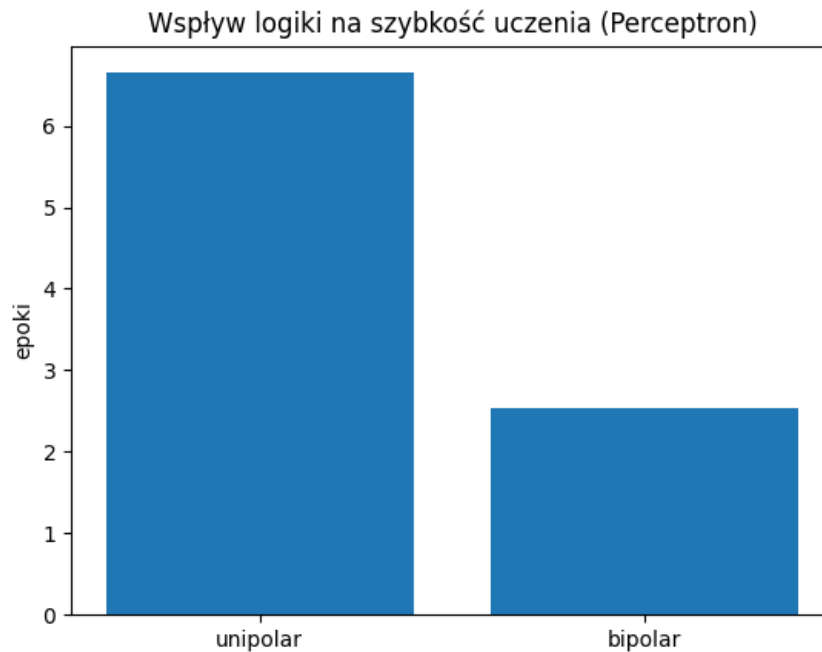
Wnioski: Dla niskich wartości α liczba epok potrzebnych do wyuczenia neuronu jest bardzo duża ponieważ potrzeba więcej czasu na znalezienie odpowiednich wartości wektora W , dalej z kolei dla wartości większych niż 1 liczba epok utrzymuje się mniej więcej poniżej trzech. Optymalna wartość α dla tego zadania wynosi około 0,2.

1.4 Badanie 4

Zbadany został wpływ logiki na szybkość uczenia. Waga początkowa progu θ została wylosowana z przedziału $(-1, 1)$, wagi początkowe wektora W zostały wybrane z przedziału $(-0,01, 0,01)$, wartość współczynnika uczenia α wynosi 0,1. Dla każdej logiki badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 4: Wpływ logiki na szybkość uczenia

	logika unipolarna	logika biloparna
epoki	6.65	2.54



Wnioski: Perceptron prosty dla logiki unipolarnej uczy się zdecydowanie wolniej niż dla logiki bipolarnej. Dzieje się tak dlatego, że w logice bipolarnej przestrzeń w której rozwiązanie jest poprawne jest większa oraz błędy mogą wynosić -2, 0 oraz 2, dlatego też w tym przypadku szybciej jest trafić na odpowiednie rozwiązanie.

2 Badania Adaline

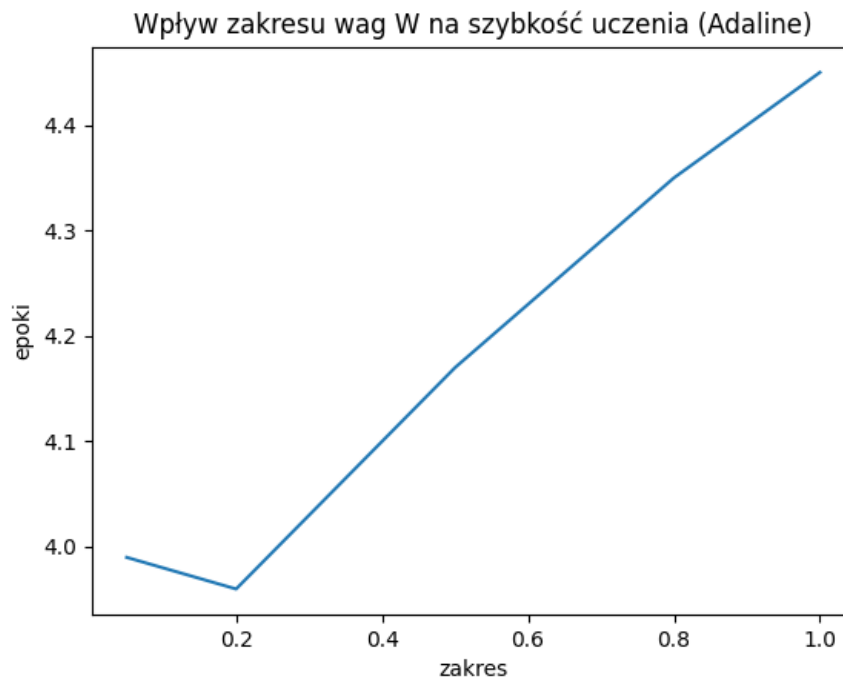
Poniżej zostaną przedstawione wyniki badań dla Adaline. Adaline symuluje działanie funkcji AND and dwóch zmiennych w logice bipolarnej. Każde badanie również zostało przeprowadzone sto razy. Adaline przestaje aktualizować wagi gdy wartość błędu średniokwadratowego spadnie poniżej progu błędu.

2.1 Badanie 5

Zbadany został wpływ zakresu w jakim losowane są początkowe wagi wektora W na szybkość uczenia. Waga początkowa progu θ została wylosowana z przedziału $(-1, 1)$, wartość współczynnika uczenia α wynosi 0,1, wartość progu błędu wynosi 0,4. Dla każdej wartości zakresu badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 5: Wpływ zakresu wag na szybkość uczenia

zakres	$(-1, 1)$	$(-0.8, 0.8)$	$(-0.5, 0.5)$	$(-0.2, 0.2)$	$(-0.05, 0.05)$
epoki	4.45	4.35	4.17	3.96	3.99



Wnioski: Zależność szybkości uczenia i zakresu początkowego wektora W dla Adaline jest tak samo jak w przypadku Perceptronu liniowa, im bardziej losowe wartości wag tym średnio więcej epok potrzebnych do wyuczenia neuronu, powód takiej zależności również jest taki sam jak w przypadku Perceptronu.

2.2 Badanie 6

Zbadany został wpływ współczynnika uczenia α na szybkość uczenia. Waga początkowa progu θ została wylosowana z przedziału $(-1, 1)$, wagi początkowe wektora W zostały wybrane z przedziału $(-0,01, 0,01)$, wartość progu błędu wynosi 0,4. Dla każdej wartości współczynnika uczenia α badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 6: Wpływ współczynnika uczenia na szybkość uczenia

α	0.001	0.01	0.1	1
epoki	100.0	24.94	3.98	67.33



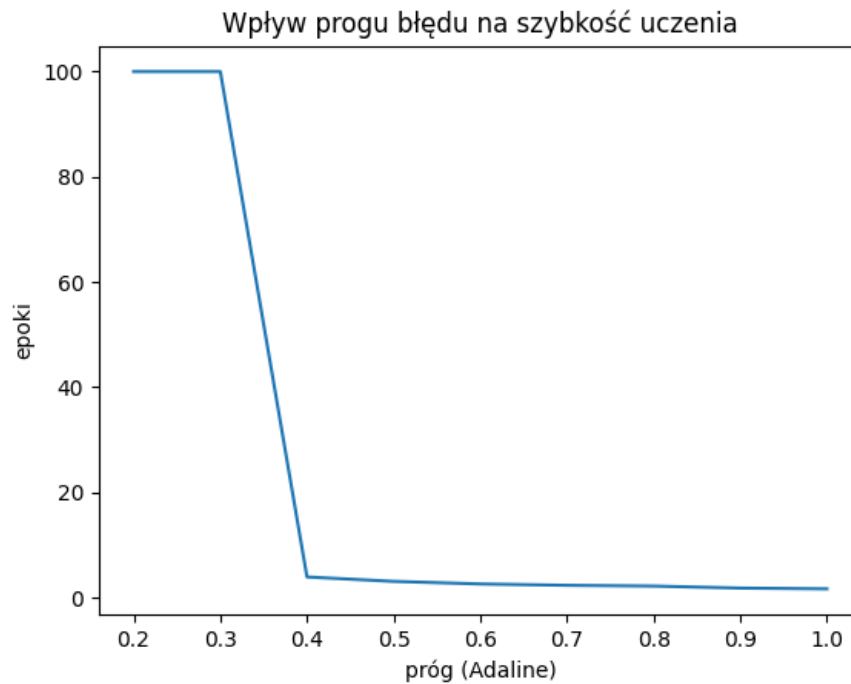
Wnioski: Adaline reaguje o wiele gorzej na zmiany współczynnika uczenia, dla bardzo małych wartości α neuron nie był w stanie wyuczyć się w mniej niż 100 epokach, najlepsza wartość α wynosiła dla tego zadania 0,1. Dla wartości α większych niż 1 nie jesteśmy w stanie znaleźć rozwiązania ponieważ błędy zamiast maleć, rosną w nieskończoność.

2.3 Badanie 7

Zbadany został wpływ wartości progu błędu na szybkość uczenia. Waga początkowa progu θ została wylosowana z przedziału $(-1, 1)$, wagi początkowe wektora W zostały wybrane z przedziału $(-0,01, 0,01)$, wagi początkowe wektora V zostały wybrane z przedziału $(-0,01, 0,01)$. Dla każdej wartości progu błędu badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 7: Wpływ progu błędu na szybkość uczenia

próg	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
epoki	100.0	100.0	4.0	3.04	2.52	2.27	2.06	1.83	1.6

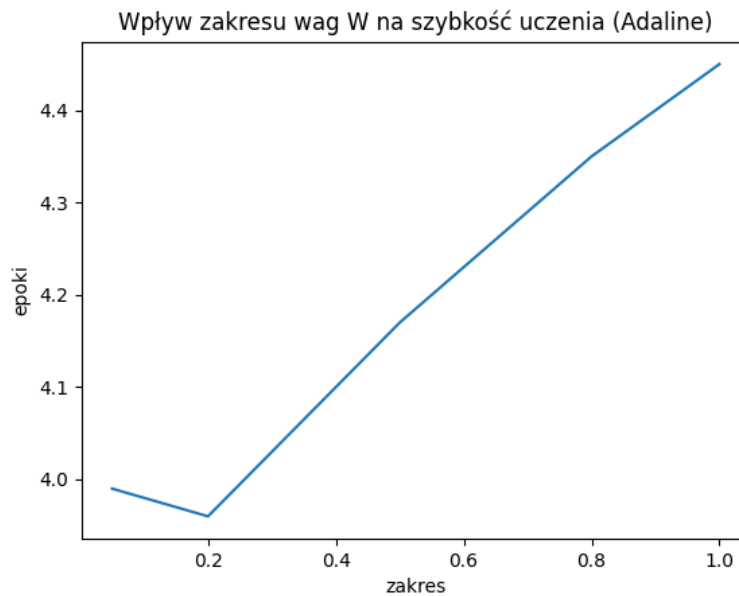
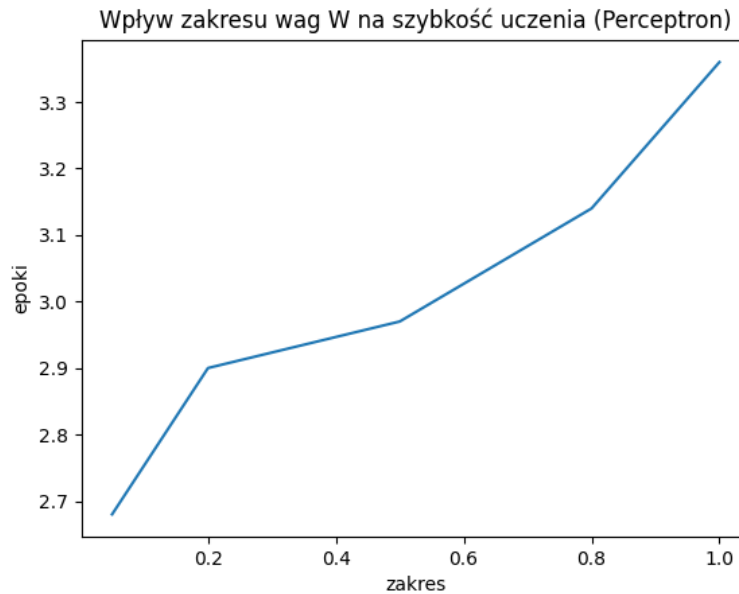


Wnioski: Dla niskich wartości progu błędu (poniżej 0,4) neuron nie jest w stanie osiągnąć błędu mniejszego niż próg w mniej niż 100 epokach. Dla wartości większych niż obserwujemy zależność liniową jednak nie znaczy to że wyższy próg jest lepszy ponieważ dla wartości progu błędu równego neuron kończy proces uczenia pomimo tego że nie zwraca poprawnych wyników.

3 Porównanie wyników badań dla Perceptronu i Adaline

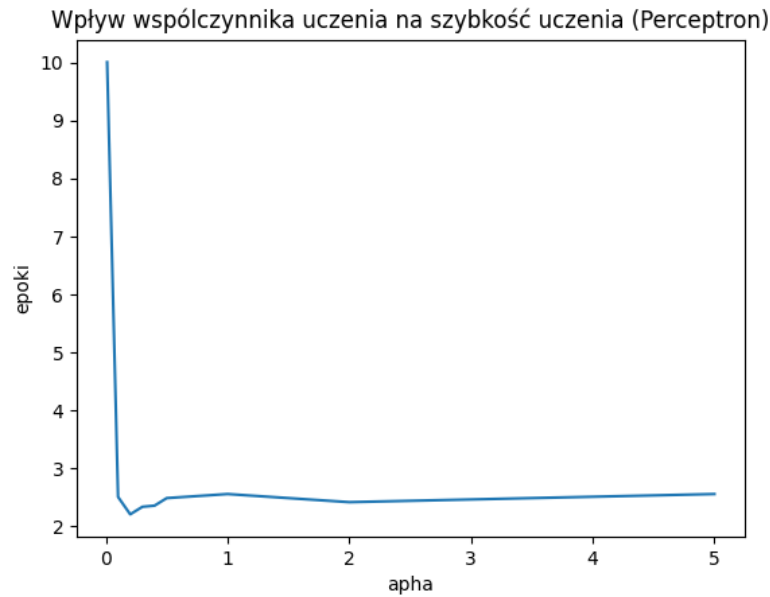
Poniżej przedstawione zostały porównania poszczególnych badań dla Perceptronu prostego i Adaline.

3.1 Porównanie wpływu zakresu wag wektora W na szybkość uczenia



Wnioski: W dla obu modeli obserwujemy liniową zależność pomiędzy zakresem początkowym losowanych wartości wag wektora W , a ilością epok potrzebnych do wyuczenia modelu.

3.2 Porównanie wpływu współczynnika uczenia na szybkość uczenia



Wnioski: Na wykresach widoczna jest wyraźna różnica pomiędzy wpływem współczynnika uczenia α na Perceptron oraz na Adaline, w przypadku tego drugiego, zbyt małe lub zbyt duże α powoduje to, że model nie jest w stanie wyuczyć się w mniej niż 100 epok.

3.3 Porównanie Perceptronu i Adaline dla znalezionych hiperparametrów

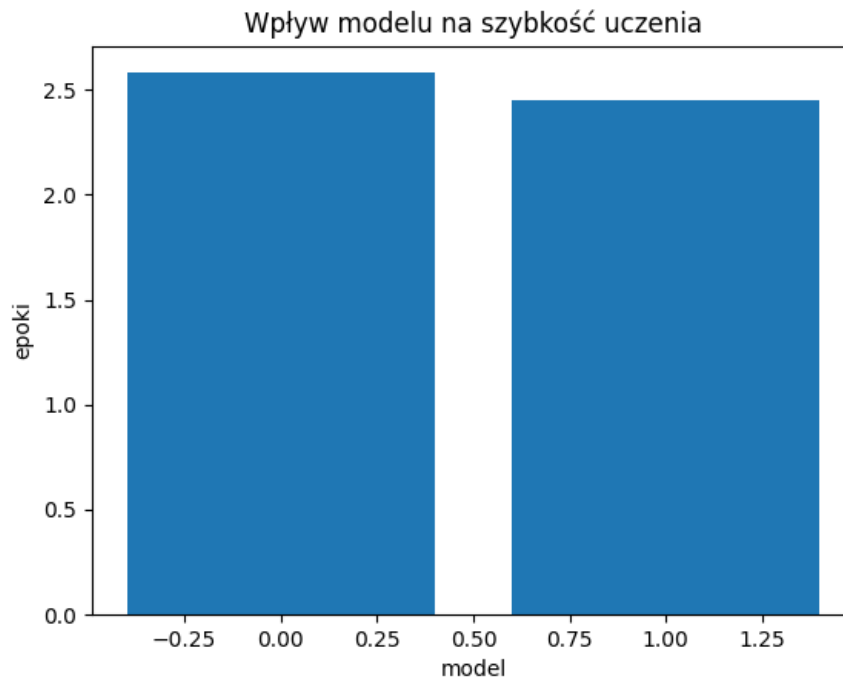
Poniżej przedstawione wyniki badania szybkości uczenia Perceptronu i Adaline dla optymalnych hiperparametrów. Dla każdego modelu badanie zostało powtórzone sto razy, poniżej uśrednione wyniki badania.

Tabela 8: Dobrane hiperparametry

Perceptron	$\theta \in (-0.5, 0.5)$	$\alpha = 0.2$	$W_i \in (-0.1, 0.1)$	model = bipolar
Adaline	$\theta \in (-0.5, 0.5)$	$\alpha = 0.1$	$W_i \in (-0.2, 0.2)$	threshold = 0.6

Tabela 9: Wpływ modelu na szybkość uczenia

Perceptron	Adaline
2.58	2.45



Wnioski: Dla obydwu modeli średnia liczba epok potrzebnych do wyuczenia modelu jest zbliżona, jednak modele różnią się od siebie w znaczący sposób i mają swoje mocne i słabe strony. Perceptron kończy proces uczenia gdy znajdzie pierwsze poprawne rozwiązanie co może prowadzić do tego, że gdy podstawimy mu dane inne niż testowe będzie je klasyfikował błędnie, Adaline z kolei może nigdy nie znaleźć odpowiednio niskiego błędu, a dla złego α może nigdy się nie wyuczyć, za to nie jest mniejsze prawdopodobieństwo problemu zbytniego dopasowania się do danych testowych, ponieważ Adaline znajduje lepszą prostą klasyfikującą dane, nieleżącą w pobliżu żadnej z klas.