



**WYDZIAŁ  
ELEKTROTECHNIKI  
I INFORMATYKI**  
POLITECHNIKI RZESZOWSKIEJ

**Mateusz Fesz**

Implementacja wielowarstwowej sieci neuronowej oraz algorytmu  
wstecznej propagacji błędów z adaptacyjnym współczynnikiem  
uczenia. Badanie wpływu metaparametrów sieci na  
efektywność uczenia

**Praca projektowa z modułu „Sztuczna inteligencja”**

Rzeszów, 2022



# Spis treści

<b>1. Opis projektu</b>	<b>5</b>
1.1. Cel projektu	5
1.2. Opis i przygotowanie wykorzystanych danych	5
<b>2. Zagadnienia teoretyczne związane z wykorzystaną siecią neuronową oraz algorytmem uczenia</b>	<b>7</b>
2.1. Model neuronu	7
2.2. Sieć neuronowa jednowarstwowa	8
2.3. Sieć neuronowa wielowarstwowa	9
2.4. Proces uczenia sieci z wykorzystaniem metody gradientowej	10
2.4.1. Ogólny algorytm gradientowy	11
2.4.2. Zastosowanie algorytmu gradientowego w sieciach wielowarstwowych	13
2.4.3. Metoda stochastycznego spadku gradientu	13
<b>3. Wstęp/wprowadzenie</b>	<b>14</b>
<b>4. Tekst zasadniczy – I</b>	<b>15</b>
4.1. Formatowanie rozdziałów i podrozdziałów	15
<b>5. Tekst zasadniczy – II</b>	<b>16</b>
5.1. Formatowanie tekstu. Należy pamiętać, że na końcu tytułu rozdziału, podrozdziału i zakresu nie umieszcza się kropki	16
5.1.1. Marginesy i akapity	16
5.1.2. Zalecenia co do sposobu pisania jednostek i symboli wielkości fizycznych	17
5.1.3. Rysunki i tabele	19
5.1.4. Wzory matematyczne	20
5.1.5. Listingi programów	22
5.1.6. Numerowanie i punktowanie	22
5.2. Wykaz literatury	23
5.3. Wydruk pracy	23
<b>6. Podsumowanie i wnioski końcowe</b>	<b>25</b>
<b>Załączniki</b>	<b>26</b>
<b>Literatura</b>	<b>27</b>



# 1. Opis projektu

## 1.1. Cel projektu

Celem projektu jest realizacja uniwersalnej sieci neuronowej wielowarstwowej oraz zbadanie wpływu metaparametrów sieci na przebieg i efektywność procesu uczenia. Metaparametry poddane badaniom to:

- S1 - liczba neuronów w pierwszej warstwie sieci
- S2 - liczba neuronów w drugiej warstwie sieci
- lr - współczynnik uczenia sieci
- er - współczynnik maksymalnego dopuszczalnego przyrostu błędu, oznaczany również jako MAX\_PERF\_INC
- lr<sub>dec</sub> - modyfikator współczynnika uczenia w przypadku przekroczeniu maksymalnego dopuszczalnego przyrostu błędu
- lr<sub>inc</sub> - modyfikator współczynnika uczenia w przypadku spadku błędu

## 1.2. Opis i przygotowanie wykorzystanych danych

W celu przeprowadzenia badań, wykorzystany został zbiór danych dostępny pod adresem: <https://archive.ics.uci.edu/ml/datasets/zoo>. Składa się na niego 101 rekordów opisanych 18 parametrami. Opisują one następujące cechy zwierzęcia:

- 1) animal name - wartość tekstowa, stanowiąca nazwę zwierzęcia. Nieuwzględniana w procesie uczenia
- 2) hair - wartość logiczna, określająca występowanie owłosienia na ciele zwierzęcia
- 3) feathers, wartość logiczna, stwierdzająca występowanie piór na ciele zwierzęcia
- 4) eggs - wartość logiczna, niosąca informację o składaniu przez zwierzę jaj
- 5) milk - wartość logiczna, określająca zdolność zwierzęcia do wytwarzania mleka
- 6) airborne - wartość logiczna, stwierdzająca zdolność zwierzęcia do lotu

- 7) aquatic - wartość logiczna, niosąca informację o zdolności zwierzęcia do funkcjonowania w środowisku wodnym
- 8) predator - wartość logiczna, informująca czy zwierzę jest drapieżnikiem
- 9) toothed - wartość logiczna, określająca występowanie uzębienia u zwierzęcia
- 10) backbone - wartość logiczna, stwierdzająca występowanie kręgosłupa
- 11) breathes - wartość logiczna, niosąca informację o sposobie oddychania zwierzęcia
- 12) venomous - wartość logiczna, określająca czy zwierzę jest jadowite
- 13) fins - wartość logiczna, stwierdzająca występowanie płetw
- 14) legs - wartość liczbowa, informująca o liczbie nóg danego zwierzęcia
- 15) tail - wartość logiczna, niosąca informację o występowaniu ogona
- 16) domestic - wartość logiczna, określająca czy zwierzę jest uznawane za domowe
- 17) catsize - wartość logiczna, stwierdzająca czy zwierzę jest porównywalne rozmiarem, bądź większe od kota
- 18) type - wartość liczbowa, określająca przynależność zwierzęcia do jednej z siedmiu klas

W procesie uczenia należy dokonać dopasowania zwierzęcia do odpowiedniej klasy. Jako parametry przyjmuje się kolumny 2-17. Kolumna 18 zawiera oczekiwaną wartość klasyfikacji. Kolumna pierwsza nie jest wykorzystywana w procesie uczenia, gdyż nie opisuje ona fizycznej cechy zwierzęcia, a jedynie jego nazwę.

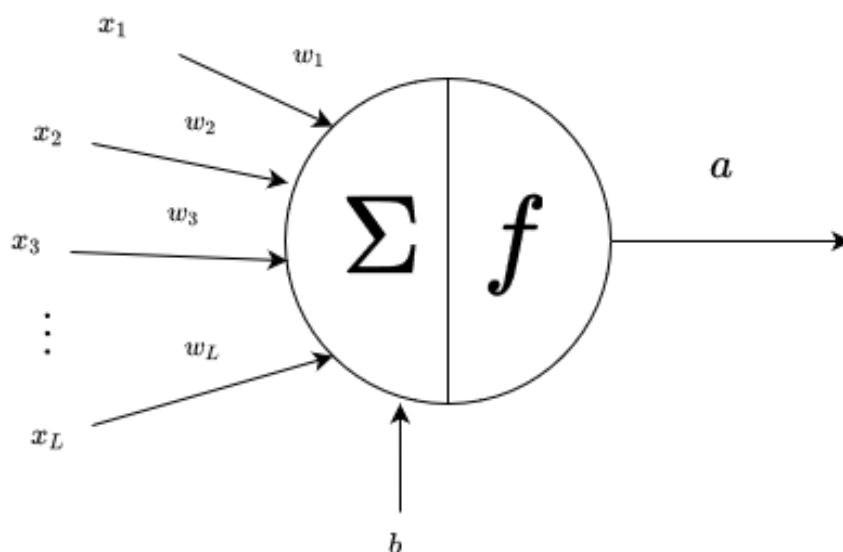
Przed przystąpieniem do procesu uczenia, przeprowadzono normalizację danych wejściowych. Wymagała jej jedynie kolumna „legs” przyjmująca wartości numeryczne z zakresu [0; 8]. Numer klasy stanowiący jedyną daną wyjściową został natomiast przekształcony do postaci siedmio-elementowego wektora, zawierającego wartość 1 na pozycji odpowiadającej danej klasie oraz wartości 0 na pozostałych pozycjach. W ostatnim kroku, dokonano podziału danych na zbiór uczący oraz zbiór walidacyjny. W tym celu przydzielono 25% rekordów należących do danej klasy do zbioru walidacyjnego, zaś pozostałą część do zbioru uczącego.

## 2. Zagadnienia teoretyczne związane z wykorzystaną siecią neuronową oraz algorytmem uczenia

### 2.1. Model neuronu

Neuron jest podstawowym elementem sieci neuronowej. Jego elementami są:

- Zbiór wejść  $x$  długości  $L$
- Zbiór powiązanych z wejściami wag  $w$  długości  $L$
- Wartość przesunięcia  $b$
- Blok sumujący
- Funkcja aktywacji  $f$
- Wyjście  $a$



Rysunek 2.1: Graficzny model neuronu

Wartość sygnału wyjściowego neuronu możemy określić wzorem:

$$a = f \left( \sum_{j=1}^L w_j x_j + b \right) \quad (2.1)$$

W powyższym zapisie,  $x_j$  oraz  $w_j$  oznaczają odpowiednio kolejne wartości wejściowe i powiązane z nimi wagi (współczynniki wagowe). Zapis ten możemy jednakże uprościć, zakładając że  $x = [x_1, x_2, \dots, x_L]^T$  będzie wektorem kolumnowym wejść,  $w = [w_1, w_2, \dots, w_L]$  - macierzą wierszową powiązanych z wejściami wag, natomiast wartości  $a$  oraz  $b$  - skalarami. Wtedy równanie 2.1 przyjmie postać:

$$a = f(wx + b) \quad (2.2)$$

Ponadto istotnym w dalszych rozważaniach elementem modelu neuronu jest wyjście z bloku sumującego. Oznaczając je jako  $z$  otrzymamy:

$$z = \sum_{j=1}^L w_j x_j + b \quad (2.3)$$

Obliczona w ten sposób wartość, stanowi argument funkcji aktywacji neuronu. Funkcja ta przekształca wartość  $z$  na inną wartość uzależnioną od jej charakterystyki. Na potrzeby projektu, przyjęto że rolę tę będzie pełniła funkcja sigmoidalna (logsig). Jej wartość możemy obliczyć zgodnie ze wzorem:

$$f(z) = \frac{1}{1 + \exp(-z)} \quad (2.4)$$

W procesie uczenia istotna będzie także wartość pochodnej tej funkcji w danym punkcie. W przypadku funkcji sigmoidalnej można ją sprowadzić do następującej postaci:

$$f'(z) = \frac{f(z)}{1 - f(z)} \quad (2.5)$$

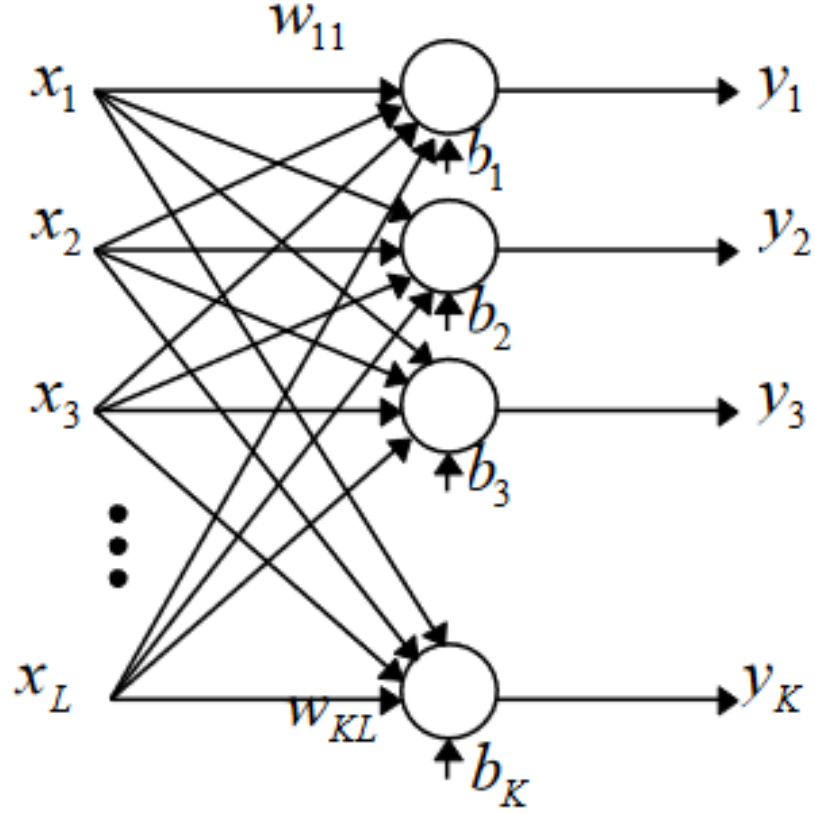
## 2.2. Sieć neuronowa jednowarstwowa

Siecią neuronową jednowarstwową nazywamy taką sieć, w której neurony nie są połączone ze sobą bezpośrednio, a jedynie otrzymują dane na wejściu i podają wynik na wyjście.

Wejściem każdego neuronu jest wektor sygnałów wejściowych  $x$ , zaś wyjściem sieci wektor  $y = [y_1, y_2, y_3, \dots, y_L]^T$ . Ponadto możemy zdefiniować również wektor przesunięć  $b = [b_1, b_2, \dots, b_K]$ . Liczba wyjść z sieci jest tożsama z liczbą neuronów.

Wagi przyporządkowane wejściom można wyrazić w postaci macierzy o rozmia-





Rysunek 2.2: Model prostej sieci jednowarstwowej - Placeholder

rach  $K \times L$  gdzie  $L$  oznacza liczbę wejść sieci, a  $K$  liczbę neuronów.

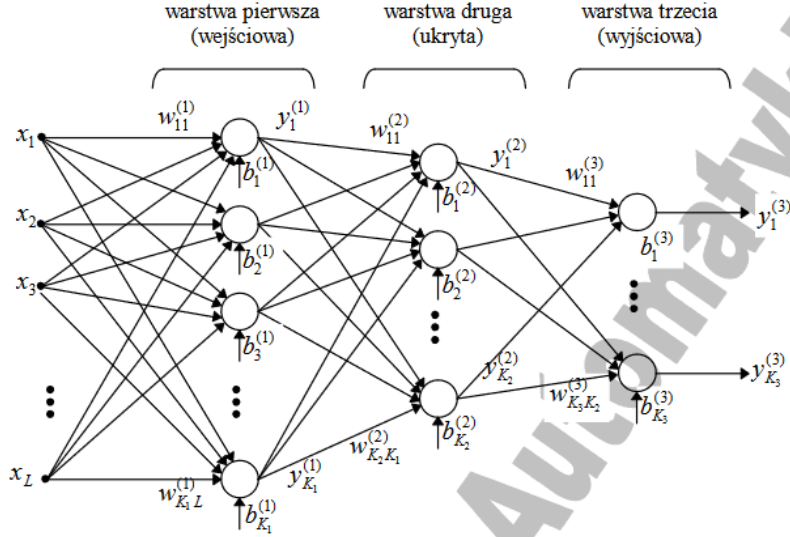
$$\mathbf{w} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1L} \\ w_{21} & w_{22} & \cdots & w_{2L} \\ \vdots & \vdots & & \vdots \\ w_{K1} & w_{K2} & \cdots & w_{KL} \end{bmatrix}$$

Przy założeniu że każdy neuron realizuje tę samą funkcję aktywacji, działanie sieci jednowarstwowej wielowarstwowej możemy wyrazić w postaci macierzowej:

$$y = f(wx + b) \quad (2.6)$$

### 2.3. Sieć neuronowa wielowarstwowa

Siecią neuronową wielowarstwową, nazywamy taką sieć w której neurony ułożone są w dwóch lub więcej połączonych ze sobą warstwach. Możemy zatem powiedzieć że jest to pewna liczba połączonych kaskadowo sieci jednowarstwowych, a wyjście pojedynczej warstwy jest równocześnie wejściem warstwy następnej.



Rysunek 2.3: Model prostej sieci trójwarstwowej - Placeholder

W przypadku tego typu sieci ilość macierzy wag jest równa liczbie warstw neuronów. Oznaczając  $y^{(l)}$  jako wyjście,  $w^{(l)}$  jako macierz wag,  $b^{(l)}$  jako macierz biasów, a  $f^{(l)}$  jako funkcję aktywacji neuronów  $l$ -tej warstwy oraz wykorzystując równanie 2.6, w przypadku sieci trójwarstwowej otrzymamy równania:

$$\begin{aligned} y^{(1)} &= f^{(1)}(w^{(1)}x + b^{(1)}) \\ y^{(2)} &= f^{(2)}(w^{(2)}y^{(1)} + b^{(2)}) \\ y^{(3)} &= f^{(3)}(w^{(3)}y^{(2)} + b^{(3)}) \end{aligned} \quad (2.7)$$

Następnie, wykorzystując równania 2.7, możemy opisać działanie całej sieci trójwarstwowej równaniem:

$$y^{(3)} = f^{(3)}(w^{(3)}f^{(2)}(w^{(2)}f^{(1)}(w^{(1)}x + b^{(1)}) + b^{(2)}) + b^{(3)}) \quad (2.8)$$

## 2.4. Proces uczenia sieci z wykorzystaniem metody gradientowej

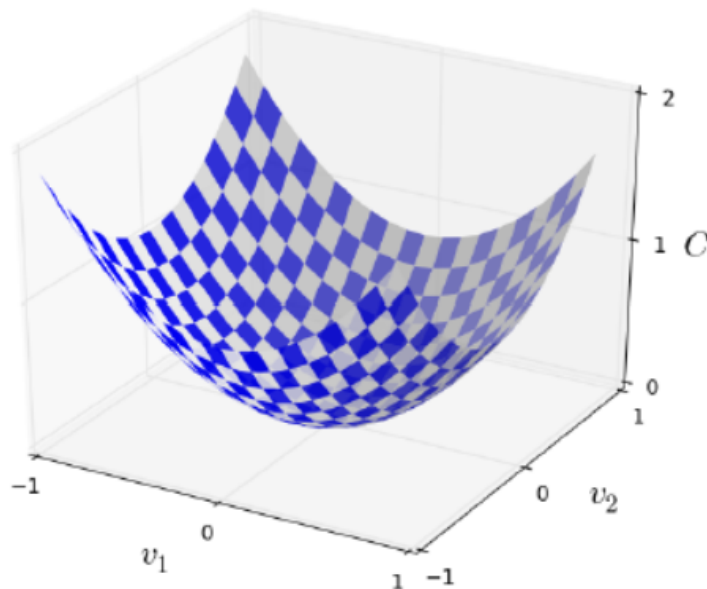
Celem określenia sposobu uczenia sieci, koniecznym jest uprzednie zdefiniowanie tego, co oznacza określenie sieci „nauczona” bądź „nienauczona”. W tym celu definiujemy tzw. funkcję kosztu, określającą poziom rozbieżności pomiędzy wartością otrzymaną na wyjściu sieci, a wartością oczekiwaną. Przykładem takiej funkcji, jest funkcja błędu średniokwadratowego (MSE):

$$C(w, b, x, a) = \frac{1}{2n} \sum_x ||y(x) - a||^2 \quad (2.9)$$

W powyższym równaniu,  $w$  określa zbiór wszystkich wag wewnątrz sieci,  $b$  wszystkich jej biasów,  $x$  zbiór danych wejściowych,  $n$  ilość danych wejściowych,  $y(x)$  zbiór oczekiwanych danych wyjściowych, natomiast  $a$  zbiór oczekiwanych wyjść z sieci. Zakładając, że w procesie uczenia wartości  $x$  oraz  $a$  pozostają stałe, możemy uprościć oznaczenie funkcji do postaci  $C(w, b)$ . Funkcja ta dąży do 0 gdy wartości otrzymywane na wyjściu sieci są bliskie wartościom oczekiwany, i rośnie wraz ze wzrostem różnicy pomiędzy nimi. W związku z powyższym, należy znaleźć taką metodę, która pozwoli na minimalizację wartości funkcji  $C(w, b)$ .

#### 2.4.1. Ogólny algorytm gradientowy

Celem zobrazowania problemu rozważmy prostą funkcję  $C(v_1, v_2)$  zobrazowaną na rysunku 2.4. Naszym celem jest znalezienie jej globalnego minimum. Teoretycznie możliwe jest jego wyznaczenie metodą analityczną, jednakże według literatury [3] jest to rozwiązanie mało wydajne, szczególnie gdy rozważymy funkcje więcej niż kilku zmiennych.



Rysunek 2.4: Przykładowa funkcja 2 zmiennych

Obserwując rysunek 2.4 możemy jednak dojść do znacznie prostszej obliczeniowo metody. Przyjmując bliskie zeru wartości  $\Delta v_1$  oraz  $\Delta v_2$  możemy przyjąć zmianę

wartości funkcji  $C(v_1, v_2)$  na poziomie:

$$\Delta C \approx \frac{\delta C}{\delta v_1} \Delta v_1 + \frac{\delta C}{\delta v_2} \Delta v_2 \quad (2.10)$$

Na podstawie powyższego równania możemy zdefiniować wektor zmian:

$$\Delta v = (\Delta v_1, \Delta v_2)^T \quad (2.11)$$

oraz tzw. wektor gradientu:

$$\nabla C = \left( \frac{\delta C}{\delta v_1}, \frac{\delta C}{\delta v_2} \right)^T \quad (2.12)$$

Wykorzystując powyższe definicje, równanie 2.12 możemy zapisać w postaci:

$$\Delta C \approx \nabla C \cdot \Delta v \quad (2.13)$$

Problemem przy równaniu 2.13 jest wyznaczenie optymalnego wektora  $\Delta v$ , tak aby zagwarantować ujemną wartość  $\Delta C$ . W tym celu, możemy przyjąć wartości opisywane równaniem 2.11 jako równe:

$$\Delta v = -\eta \nabla C \quad (2.14)$$

Wartość  $\eta$  nazywana jest „współczynnikiem uczenia” i przyjmuje bliskie zeru, dodatnie wartości. Wykorzystując powyższą definicję, możemy zapisać równanie 2.13 w postaci:

$$\Delta C \approx -\eta \nabla C \cdot \nabla C$$

$$\Delta C \approx -\eta \|\nabla C\|^2$$

W tej postaci równania widzimy, że dla odpowiednio niskiego współczynnika  $\eta$  zagwarantowany jest spadek wartości funkcji kosztu  $C(v)$ . Wykorzystując właściwość opisaną równaniem 2.14 możemy wyznaczyć nowe wartości zmiennych zawartych w wektorze  $v$ :

$$\begin{aligned} v &\rightarrow v' = v - \eta \nabla C \\ v' &= v + \Delta v \end{aligned} \quad (2.15)$$

Wielokrotnie aplikując powyższą regułę, jesteśmy w stanie podążać za gradientem aż do osiągnięcia minimum. Literatura [3] definiuje ją jako „algorytm spadku gradient”, ale jednocześnie wspomina że w niektórych sytuacjach, może nie być w pełni skuteczna.

### 2.4.2. Zastosowanie algorytmu gradientowego w sieciach wielowarstwowych

Algorytm gradientowy może zostać wykorzystany do minimalizacji funkcji przedstawionej równaniem 2.10, poprzez odnajdowanie wartości wag oraz biasów minimalizujących funkcję kosztu. Wykorzystując oznaczenia  $w_{ij}^{(l)}$  jako  $j$ -tą wagę  $i$ -tego neuronu  $l$ -tej warstwy oraz  $b_i^{(l)}$  jako bias  $i$ -tego neuronu w  $l$ -tej warstwie, możemy wyznaczyć ich wartości w kolejnych iteracjach (epokach) procesu uczenia:

$$w_{ij}^{(l)} \rightarrow w'_{ij}{}^{(l)} = w_{ij}^{(l)} - \eta \frac{\delta C}{\delta w_{ij}^{(l)}} \quad (2.16)$$

$$b_i^{(l)} \rightarrow b'_i{}^{(l)} = b_i^{(l)} - \eta \frac{\delta C}{\delta b_i^{(l)}} \quad (2.17)$$

Przyjmując jako funkcję kosztu błąd średniokwadratowy, przy założeniu wystąpienia wielu danych uczących, możemy ją zapisać w postaci:

$$C = \frac{1}{n} \sum_x C_x \quad (2.18)$$

czyli średniej wartości błędów  $C_x = \frac{\|y(x)-a\|^2}{2}$  dla poszczególnych par uczących.

TODO: Rozpisać wzory na poszczególne pochodne - instrukcja [1]

### 2.4.3. Metoda stochastycznego spadku gradientu

Jednym z problemów algorytmu spadku gradientu, objawiającym się przy dużej liczbie rekordów uczących [3] jest długi czas potrzebny na obliczanie wartości pochodnych w równaniu 2.12. Celem zniwelowania tego problemu można wykorzystać metodę zwaną „stochastycznym spadkiem gradientu”. Polega ona na oszacowaniu rzeczywistego gradientu  $\nabla C$  poprzez obliczenie  $\nabla C_x$  dla losowo wybranej próbki danych uczących. Poprzez odpowiednie dobranie rozmiaru próbek, oraz ich odpowiednie uśrednienie, możemy uzyskać dobre przybliżenie rzeczywistego gradientu  $\nabla C$ . [3] Elementy pojedynczej próbki możemy oznaczyć jako  $X_1, X_2, \dots, X_m$ . Zakładając odpowiednio duży rozmiar  $m$  możemy przyjąć za prawdziwe równanie:

$$\frac{\sum_{j=1}^m \nabla C_{X_j}}{m} \approx \frac{\sum_x \nabla C_x}{n} = \nabla C \quad (2.19)$$

Przenosząc czynnik  $\frac{1}{m}$  przed znak sumy, otrzymamy:

$$\nabla C \approx \frac{1}{m} \sum_{j=1}^m \nabla C_{X_j} \quad (2.20)$$

Wartym zauważenia jest też fakt występowania różnych konwencji uśredniania czy też skalowania gradientu stochastycznego. [3] Niektóre źródła mówią o braku

konieczności ich stosowania gdyż wpływają na działanie sieci jedynie pośrednio, poprzez modyfikację wartości współczynnika uczenia.

### 3. Wstęp/wprowadzenie

1 ÷ 5 stron charakterystyka problematyki w świetle aktualnego stanu wiedzy i techniki, ze wskazaniem na zagadnienia istotne z punktu widzenia realizowanej pracy. Na trzeciej stronie można zamieścić podziękowania dla osób, które przyczyniły się do powstania pracy dyplomowej. Na kolejnej stronie nieparzystej rozpoczyna się spis treści. Po spisie treści zalecane jest umieszczenie wykazu użytych symboli, oznaczeń i akronimów. Od tego miejsca rozpoczyna się numeracja rozdziałów. Na następnej stronie umieszcza się wprowadzenie do pracy (scharakteryzowanie problematyki pracy, uzasadnienie wyboru tematyki) oraz przedstawia: cel i/lub tezę pracy, zakres pracy, przyjęte założenia itp. Ostatni akapit wstępu musi zawierać zwięzłe sformułowanie celu i zakresu pracy.

#### Uwaga:

Jeżeli decydujesz się wykorzystywać  $\text{\LaTeX}$ 'a, ignoruj ogólny dokument dotyczący formatowania pracy dyplomowej na WEiI - jest przeznaczony dla użytkowników innych edytorów tekstu. Korzystaj z załączonego arkusza stylu, stosuj formatowanie znaczeniowe (nie wymuszaj formatowania), a wynikowa praca będzie zgodna z wymaganiami. Zachęcamy do używania  $\text{\LaTeX}$ 'a, czas poświęcony na jego przyswojenie, zwróci się z nawiązką nawet w trakcie tworzenia pracy dyplomowej.

Niniejszy tekst, wykorzystujący styl `weii szablon.sty` zawiera informacje o formatowaniu, wielkości czcionek, wyrównania..., ale **uwaga**, sama treść nie jest istotna (np. opis wielkości czcionek), formatowanie wykona się automatycznie, tu te zapisy są tylko po to, aby dostarczyć dokument zawierający jak najwięcej przykładów użycia  $\text{\LaTeX}$ 'a.

## 4. Tekst zasadniczy – I

Do 20% objętości pracy. W zależności od charakteru pracy ten rozdział powinien zawierać:

- a) opis tematyki zagadnienia – aktualny stan zagadnienia,
- b) metody i rozwiązania,
- c) dyskusja i krytyczna ocena stanu aktualnego,
- d) podsumowanie stanu wiedzy, techniki literaturowe itp.

### 4.1. Formatowanie rozdziałów i podrozdziałów

Rozdziały zaczynają się u góry nowej strony (parzystej lub nieparzystej). Podrozdziały i zakresy mogą zaczynać się w dowolnym miejscu strony. Przy końcu pracy zamieszcza się podsumowanie i wnioski. Ostatni akapit podsumowania musi zawierać wyszczególnienie własnej pracy Autora i zaczynać się od sformułowania: „Autor za własny wkład pracy uważa:”. W tym miejscu kończy się numeracja rozdziałów.

Ewentualne listingi programów, instrukcje obsługi stanowisk lub inne tego rodzaju materiały zaleca się zamieścić w formie dodatków. Kolejno zamieszcza się: wykaz literatury, spis rysunków/tabel oraz streszczenie (zgodne ze „Wzorem streszczenia”). Wykaz literatury rozpoczyna od strony nieparzystej.

Opisując własne dokonania, stosuje się formę bezosobową w czasie przeszłym np. celem pracy było zaprojektowanie..., zakres pracy obejmował wyznaczenie..., w ramach pracy wykonano model... itp.

## 5. Tekst zasadniczy – II

Ponad 50% objętości pracy – część autorska:

- a) założenia – dane,
- b) opis zastosowanej metody rozwiązania lub analizy,
- c) opis proponowanego rozwiązania, wyniki analizy teoretycznej, obliczenia, projekt konstrukcyjny, procesowy, technologiczny,
- d) wyniki badań analitycznych, symulacyjnych lub eksperymentalnych itp.

Przy stosowaniu podziału na rozdziały i podrozdziały zaleca się unikać podziału więcej niż trzystopniowego. Podział tekstu, szczególnie na rozdziały główne, wynikać powinien z zakresu i charakterystyki realizowanej pracy.

### 5.1. Formatowanie tekstu. Należy pamiętać, że na końcu tytułu rozdziału, podrozdziału i zakresu nie umieszcza się kropki

#### 5.1.1. Marginesy i akapity

Marginesy deklaruje się jako „lustrzane” i ustawia na 2 cm, na oprawę 1,5 cm. Nagłówek i stopka 1,25 cm. Tekst podstawowy akapitu: czcionka szeryfowa, styl Times (Times New Roman, Liberation Serif itp.), rozmiar 12 punktów, interlinia 1,5 wiersza. Akapit wyjustowany, wcięcie pierwszego wiersza 1,25 cm.

Na końcu każdego akapitu, którego tekst zaczerpnięto z literatury, musi znajdować się odnośnik do właściwej pozycji w wykazie literatury. W pracy nie stosuje się odnośników w formie przypisów. Liczby w nawiasie kwadratowym oznaczają kolejny numer pozycji w wykazie, np. [1] lub [1, 4, 7] lub [1, 6-8] itp.

Cytaty (dosłowne przytoczenie obcego tekstu w pracy) pisze się czcionką pochylą (kursywą) i ujmuje w cudzysłów. Przykład: „*Współpracując z jednostkami gospodarczymi działającymi w kraju, kształci wysokokwalifikowaną kadrę inżynierów*”.

Fragmenty kodów programów pisze się czcionką o stałej szerokości, styl Courier (Courier New, Liberation Mono itp.) o rozmiarze 10 punktów.



### 5.1.2. Zalecenia co do sposobu pisania jednostek i symboli wielkości fizycznych

Poniższy podrozdział opracowano na podstawie [?]. W trakcie pisania pracy należy zwracać uwagę na sposób oznaczania jednostek i symboli wielkości fizycznych. Przy zapisywaniu jednostek i symboli wielkości fizycznych można wyróżnić zapis w postaci kursywy (pismo pochyłe) oraz antykwy (pismo proste).

1) Kursywę należy stosować w następujących przypadkach:

- symboli wielkości fizycznych niezależnie od tego czy jest to litera alfabetu greckiego (np. przenikalność magnetyczna  $\mu$ ) czy też łacińskiego (np. rezystancja  $R$ ). Należy przestrzegać tej zasady niezależnie od miejsca, w którym pojawia się symbol tj. tekst, wzory matematyczne, rysunki, tabele,
- ogólny symbol zapisu funkcji czyli np.  $f$ , a nie  $f$ . Nie dotyczy to jednak zapisu konkretnych funkcji np.  $\cos\omega t$  a nie  $\cos\omega t$ ,
- macierze, wektory, których elementami są wielkości fizyczne należy zapisywać dodatkowo czcionką półgrubą (bold) np.  $\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ ,  $\mathbf{U} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ ,
- wskaźnik dolny, górny, prawo- i lewostronny, ale tylko gdy odnosi się do konkretnej wielkości fizycznej, czyli np. składowa  $x$ -owa indukcji magnetycznej  $B_x$ , a nie  $B_x$ ,
- wskaźniki górne i dolne oznaczające dowolną liczbę np.  $R_j$ ,  $I^k$ , ale nie  $R_1$ ,  $I^2$ .

2) Antykwę należy stosować w następujących sytuacjach:

- wszystkie cyfry,
- symbole konkretnych funkcji np.  $\operatorname{tg} \omega t$ , a nie  $\operatorname{tg}\omega t$ ,
- operatory operacji matematycznych np. pochodne zwyczajne  $\frac{dx}{dt}$ , a nie  $\frac{dx}{dt}$ ,
- symbole liczb o konkretnej wartości np. przenikalność elektryczna próżni  $\varepsilon_0 = 8,8542 \cdot 10^{-12} \text{ F} \cdot \text{m}^{-1}$ , a nie  $\varepsilon_0 = 8,8542 \cdot 10^{-12} \text{ F} \cdot \text{m}^{-1}$ ,

- indeksy, jeżeli odnoszą się do: obiektów (fizycznych, geometrycznych), czyli, np. natężenie pola elektrycznego w punkcie A to  $E_A$ , a nie  $E_A$ , zjawisk lub stanów fizycznych, np. moment obciążenia to  $T_L$ , a nie  $T_L$ , do nazwisk czy też oznaczeń pierwiastków, np. straty w miedzi to  $P_{Cu}$  a nie  $P_{Cu}$ , do charakteru wielkości symbolizowanej przez literę źródłową, np. wartość maksymalna siły to  $F_{\max}$ , a nie  $F_{max}$ , oznaczeń jednostek miary np.  $M\Omega$ , a nie  $M\Omega$ .
- 3) W przypadku jednostek miar (które zawsze należy pisać antykwą) zapisując konkretną wartość liczbą należy podać jej wartość i jednostkę z zachowaniem następujących zasad:
- zapisując wartość liczbową wielkości fizycznej po spacji należy podać jej jednostkę, ale nie nazwę jednostki np. 10A, ale nie 10 amper czy też 10 amperów,
  - zapisując wartość liczbową słownie należy w tej konwencji podać też jednostkę np. dziesięć omów, ale nie dziesięć  $\Omega$
  - do oznaczeń jednostek nie wolno dopisywać indeksów, np. moc wyjściowa silnika wynosi  $P = 100 \text{ kW}_{\text{out}}$ . W takim przypadku należy zapisać  $P_{\text{out}} = 100 \text{ kW}$ ,
  - jednostek nie należy umieszczać w nawiasach kwadratowych, np.  $I = 1 \text{ [A]}$ . Odstępstwem od tej zasady mogą być tabele, nagłówki kolumn, opisy osi na wykresach oraz w sporadycznych sytuacjach we wzorach matematycznych (ale tylko wówczas, gdy zależność matematyczna nie wskazuje w jakiej jednostce wystąpi wartość liczbowa). Przykłady odstępstw zamieszczono w podrozdziale 5.1.3.
- 4) W trakcie zapisu symboli wielkości matematycznych można stosować również szereg znaków diakrytycznych, jak również należy przestrzegać następujących zaleceń:
- wartości chwilowe podstawowych wielkości fizycznych używanych np. w elektrotechnice należy zapisać małymi literami, np.  $u$ ,  $i$ , lub stosować zapis np.  $u(t)$ , lub stosować indeks „t” przy wielkości, np.  $U_t$ ,
  - wartości skuteczne wielkości okresowych należy zapisać dużą literą np.  $U$ ,  $I$ ,

- wartości szczytowe funkcji zmiennej, amplitudę funkcji sinusoidalnej czasu należy zapisać jako np.  $U_m$ ,
- podkreślenie symboli reprezentujących wielkości fizyczne, których wartość liczbową jest liczbą zespoloną, przy czym podkreślenie dotyczy tylko literki źródłowej np.  $\underline{Z}_1$ , a nie  $\underline{Z}_1$ ,
- kreska nad literą źródłową oznacza wartość średnią, np.  $\bar{I}$  co jest równoważne  $I_{av}$ .

### 5.1.3. Rysunki i tabele

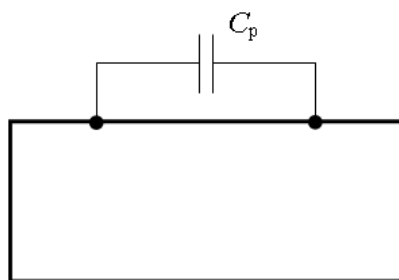
Tekst podstawowy w tabeli pisze się czcionką o rozmiarze 10 punktów, pojedyn-  
cza interlinia. Dane liczbowe – wyśrodkowane, dane tekstowe – wyrównane do lewej.  
Rysunki i tabele zamieszcza się wyśrodkowane na stronie, bez wcięcia pierwszego wiersza.

W akapicie poprzedzającym rysunek lub tabelę musi znajdować się krótki opis, czego dotyczy dany rysunek/tabela (odniesienie do rysunku/tabeli). Tytuły numeruje się zgodnie z kolejnością w danym rozdziale: numer\_rozdziału.numer\_tabeli/rysunku (np. rys. 2.1, tabela 3.5). W tytule rysunku/tabeli, zaczerpniętych z literatury, podaje się odnośnik do właściwej pozycji. Należy zadbać o to, aby opisy na rysunkach były czytelne (czcionka 8 punktów lub większa). Staraj się nie wymuszać numeracji, pozwól aby robił to za ciebie L<sup>A</sup>T<sub>E</sub>X. Stosuj `\label` do znakowania obiektów, do których być może w tekście się będziesz odwoływał (rozdziały, rysunki, tabele, wzory, listingi ...). Odwołuj się do nich w tekście za pomocą funkcji `\ref{NazwaObiektu}`. Pamiętaj, że L<sup>A</sup>T<sub>E</sub>X korzystając z polecenia `latex` nie odczytuje z plików .jpg, .png ich wielkości. Polecenie `latex` generuje plik DVI. Jeżeli chcesz go używać zgłoś stosowny błąd. Aby się go pozbyć zdefiniuj wielkość natywną pliku grafiki. Polecamy jednak używanie zamiast polecenia `latex`, polecenie `pdflatex`, wówczas problem nie wystąpi.

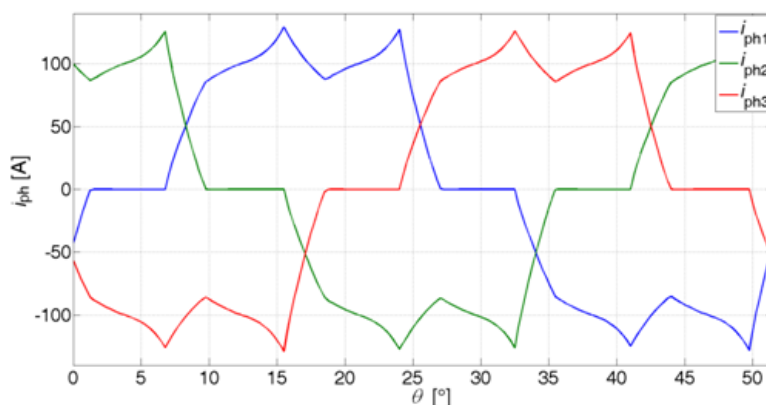
**Przykład:** [...] co umożliwia wyznaczenie wartości napięcia. Na rys. 5.5 przedstawiono schemat obwodu z równolegle dołączoną pojemnością  $C_p$ .

**Przykład:** [...] Na rysunku 5.6 pokazano przykładową zależność prądów pasmowych  $i_{ph}$  bezszczotkowego silnika prądu stałego z magnesami trwałymi w funkcji położenia wirnika  $\theta$ .

**Przykład:** [...] oraz indukcyjności wzajemnej. W tabeli 5.1 przedsta-



Rysunek 5.5: Tytuł rysunku, rozmiar 11 pkt., pojedyncza interlinia, akapit wyśrodkowany, bez wcięcia pierwszego wiersza. Na końcu tytułu rysunku/tabeli nie stawia się kropki [8]



Rysunek 5.6: Tytuł rysunku, rozmiar 11 pkt., pojedyncza interlinia, akapit wyśrodkowany, bez wcięcia pierwszego wiersza. Na końcu tytułu rysunku/tabeli nie stawia się kropki [8]

wiono podstawowe parametry obwodu nieliniowego, zasilanego napięciem trójfazowym.

#### 5.1.4. Wzory matematyczne

Zmienne we wzorach pisze się czcionką pochylą (styl edytora równań „Matematyka”) natomiast symbole, nie będące zmiennymi, czcionką prostą (styl „Tekst”). Rozmiary czcionek: normalny 12 punktów, indeks dolny/górny 9 pkt., indeks podrzędny 7 pkt., symbol 24 pkt., podsymbol 12 pkt. Separatorem dziesiętnym w liczbach jest przecinek, a nie kropka (dotyczy to również liczb pisanych w tekście akapitu). Poddaj się w tym zakresie  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ ’owi - pisz wzór, a poprawnie się utworzy.

Pod wzorem należy zamieścić objaśnienia użytych symboli (chyba, że znajdują

Tablica 5.1: Tytuł tabeli, rozmiar 11 pkt., pojedyncza interlinia, akapit wyrównany do lewej

$U$ [V]	$I$ [mA]	$R$ , [k $\Omega$ ]	$L$ [mH]	$R/R_{20}$
13,6	7,29	3,94	100	1,25

się w wykazie na początku pracy). Wzory umieszcza się wyśrodkowane i numeruje zgodnie z kolejnością w danym rozdziale: (numer\_rozdziału.numer\_wzoru). Numery wzorów wyrównuje się do prawego marginesu. W akapicie poprzedzającym wzór musi znajdować się krótki opis, czego dotyczy dany wzór i – jeżeli potrzeba – odwołanie do literatury.

**Przykład:** [...] wyznacza się, na podstawie wyrażenia (5.21). W nawiasach podano rozmiary czcionek używanych we wzorach

$$A(12) = \sum (24) m_{s(9)} N^{k_p(7)} \quad (5.21)$$

gdzie:  $m_s$  – masa próbki,  $N$  – natężenie oświetlenia,  $k_p$  – wykładnik potęgi ( $k_p = 1, 3 - 2, 1$ ).

### 5.1.5. Listingi programów

W pracy dyplomowej możesz umieszczać fragmenty programów. Pamiętaj, aby umieszczać krótkie, tylko najważniejsze fragmenty kodów źródłowych. Zawsze je komentuj w treści pracy dyplomowej. Typowo w  $\text{\LaTeX}$  kody źródłowe umieszczane są w środowisku `verbatim` (`\begin{verbatim}...\end{verbatim}`). Obecnie istnieje jednak bardziej nowoczesne i bardziej funkcjonalne środowisko `lstlisting` (wymaga zainstalowanego w systemie pakietu `listings`). Zwróć uwagę, że możesz kolorować składnię automatycznie za pomocą parametru `language`. W niniejszym dokumencie przedstawiono dwa przykłady listingów, Listing 1 to przykład kodu źródłowego Matlaba, a poniżej Listing 2 dla Perl'a.

```
1 i = 1
2 p = 3
3 for i = 1:10
4     if i > 3
5         i=i+p
6     else
7         i=i+1
8     end
9 end
```

Listing 1: Listing programu Matlab

```
1 my $url = 'http://pei.prz.edu.pl';
2 use LWP::Simple;
3 my $content = get $url;
4 die "Couldn't get $url" unless defined $content;
5 print $content;
6 print "\n";
7 print "Length " + length($content)
```

Listing 2: Listing programu Perl

Z pewnością przeglądając źródło tego dokumentu zobaczysz, że kody źródłowe powinny mieć zdefiniowane parametry `label`, aby łatwo w tekście do nich się odwoływać. Numeracja linii jest w stylu domyślnie włączona (to przydatne, bo w treści pracy łatwo odwołać się dzięki temu do konkretnego wiersza w kodzie źródłowym), możesz je wyłączyć podając jako parametr `numbers=none`. Więcej szczegółów możesz odnaleźć w sekcji `\lstset` pliku arkusza stylu.

### 5.1.6. Numerowanie i punktowanie

- 1) Pierwszy poziom (stosuje się numerowanie lub punktowanie). Formatowanie: akapit wyjustowany, wcięcie od lewej 0,75 cm, wysunięcie co 0,5 cm.

- 2) Znakiem numerowania jest liczba (z kropką lub nawiasem).
- drugi poziom (stosuje się wyłącznie punktowanie). Formatowanie: akapit wyjustowany, wcięcie od lewej 1,25 cm, wysunięcie co 0,5 cm,
  - znakiem punktowania jest łącznik lub mała litera alfabetu (z nawiasem). Nie zaleca się stosowania kropek, strzałek itp.,
  - punktowane akapity rozpoczyna się minuskulą (małą literą), na końcu akapitu stawia się przecinek, ostatni punktowany akapit kończy się kropką.
- 3) Numerowane akapity rozpoczyna się majuskulą (wielką literą) i kończy kropką.
- 4) Należy zwrócić uwagę, aby nie rozdzielać numerowania/punktowania pomiędzy kolejnymi stronami tekstu.

## 5.2. Wykaz literatury

W wykazie literatury zamieszcza się wyłącznie pozycje, na które powołano się w pracy. Kolejność numerów w wykazie – zgodna z kolejnością pojawiania się danej pozycji w tekście.

Format akapitu: akapit wyjustowany, wysunięcie 0,75 cm. Prawidłowo opracowany wykaz został zaprezentowany w niniejszym dokumencie w odpowiednim rozdziale, oznaczonym jako „Literatura” (pozycja nr [?] to zasoby internetowe, [?] – książka, [?] – artykuł w czasopiśmie, [?] – karta katalogowa).

## 5.3. Wydruk pracy

Przed wydrukiem należy usunąć ewentualne błędy literowe i sprawdzić prawidłową interpunkcję. Przykładowo, łącznik zapisuje się za pomocą krótkiego minusa (np. badawczo-rozwojowy) natomiast myślnik – stosowany w zdaniach wtrąconych – zapisuje się za pomocą długiej pauzy. Dzielenie wyrazów według uznania Autora (można podzielić długie wyrazy, powodujące duże „rozstrzelenie” tekstu w poprzedzającym wierszu. Zaleca się usunięcie pojedynczych znaków na końcu wiersza oraz podwójnych spacji w tekście. Dla przedrostka „mikro” należy unikać stosowania litery „u” zamiast „μ”. Znak „μ” można otrzymać przytrzymując lewy Alt i wpisując na klawiaturze numerycznej 0181 (podobnie „stopień”: Alt-0176). W celu uniknięcia „rozstrzelenia” liczb i ich jednostek zaleca się używanie „twardej” spacji pomiędzy

liczbą i jednostką. Należy sprawdzić, czy tytuły podrozdziałów/zakresów nie zostały jako pojedyncze wiersze na poprzedniej stronie oraz czy rysunki/tabele i ich tytuły nie zostały rozdzielone pomiędzy kolejnymi stronami.

Pracę drukuje się dwustronnie. Zaleca się wydruk w kolorze. Przed wydrukiem należy ponumerować strony (czcionka 10 pkt., dół strony, akapit wyśrodkowany). Strony tytułowej oraz strony z podziękowaniem nie numeruje się. Spis treści rozpoczyna się od strony numer 3 (lub 5, jeżeli zamieszczono podziękowania).



## 6. Podsumowanie i wnioski końcowe

1 ÷ 3 stron merytorycznie podsumowanie najważniejszych elementów pracy oraz wnioski wynikające z osiągniętego celu pracy. Proponowane zalecenia i modyfikacje oraz rozwiązania będące wynikiem realizowanej pracy.

Ostatni akapit podsumowania musi zawierać wykaz własnej pracy dyplomanta i zaczynać się od sformułowania: „Autor za własny wkład pracy uważa: ...”.

## Załączniki

Według potrzeb zawarte i uporządkowane uzupełnienie pracy o dowolny materiał źródłowy (wydruk programu komputerowego, dokumentacja konstrukcyjno-technologiczna, konstrukcja modelu – makiety – urządzenia, instrukcja obsługi urządzenia lub stanowiska laboratoryjnego, zestawienie wyników pomiarów i obliczeń, informacyjne materiały katalogowe itp.).

## Literatura

- [1] <http://materialy.prz-rzeszow.pl/pracownik/pliki/34/sztuczna-inteligencja-cw9-siec-wielowarstw.pdf> (Dostęp 05.06.2022r.)
- [2] <http://materialy.prz-rzeszow.pl/pracownik/pliki/34/sztuczna-inteligencja-cw8-siec-jednowarstw.pdf> (Dostęp 05.06.2022r.)
- [3] Nielsen M.: Neural Networks and Deep Learning.  
<http://neuralnetworksanddeeplearning.com/> (Dostęp. 05.06.2022r.)

**STRESZCZENIE PRACY DYPLOMOWEJ WPISZ-RODZAJ-PRACY**

**IMPLEMENTACJA WIELOWARSTOWEJ SIECI NEURONOWEJ  
ORAZ ALGORYTMU WSTECZNEJ PROPAGACJI BŁĘDU Z  
ADAPTACYJNYM WSPÓŁCZYNNIKIEM UCZENIA. BADANIE  
WPŁYWU METAPARAMETRÓW SIECI NA EFEKTYWNOŚĆ  
UCZENIA**

Autor: Mateusz Fesz, nr albumu: EF-167788

Opiekun: (tytuł naukowy przed) Imię i nazwisko opiekuna (tytuł po)

Słowa kluczowe: (max. 5 słów kluczowych w 2 wierszach, oddzielanych przecinkami)

Treść streszczenia po polsku

**WPISZ-RODZAJ-PRACY THESIS ABSTRACT**

**TEMAT PRACY PO ANGIELSKU**

Author: Mateusz Fesz, nr albumu: EF-167788

Supervisor: (academic degree) Imię i nazwisko opiekuna

Key words: (max. 5 słów kluczowych w 2 wierszach, oddzielanych przecinkami)

Treść streszczenia po angielsku