

WYDZIAŁ FIZYKI I INFORMATYKI STOSOWANEJ



SZTUCZNE SIECI NEURONOWE
Sprawozdanie z projektu

Aleksandra Poręba
Grzegorz Podsiadło

16 czerwca 2020

Spis treści

1	Wstęp	3
2	Poszukiwanie konfiguracji sieci neuronowej	4
3	Badanie wpływu ilości kolumn na wyniki sieci	9
3.1	Korelacja pomiędzy danymi	9
3.2	Testowanie sieci neuronowej z pomniejszonym zbiorem uczącym	11
3.3	Wnioski	15
4	Przewidywanie wyniku bazując wynikach pozostałych egzaminów	16
5	Podsumowanie	17

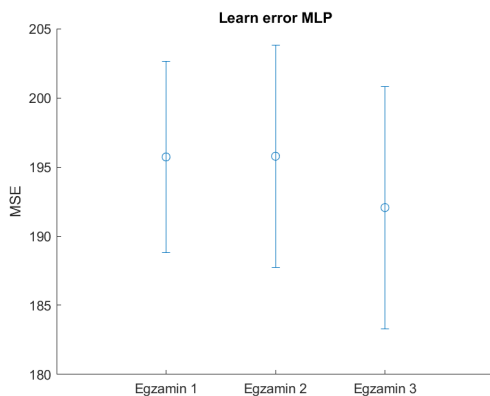
1 Wstęp

2 Poszukiwanie konfiguracji sieci neuronowej

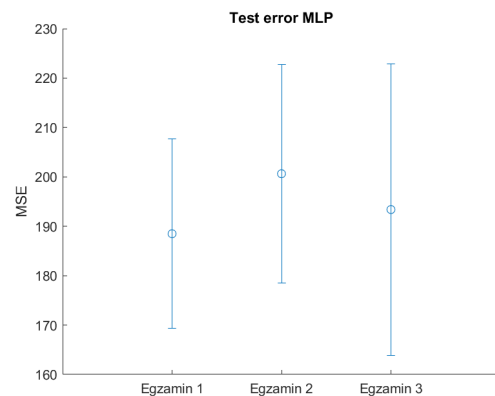
Przetestowano konfigurację sieci o jednej oraz dwóch warstwach, z wykorzystaniem różnych kombinacji funkcji:

1. `logsig`
2. `tansig`
3. `purelin`
4. `radbas`

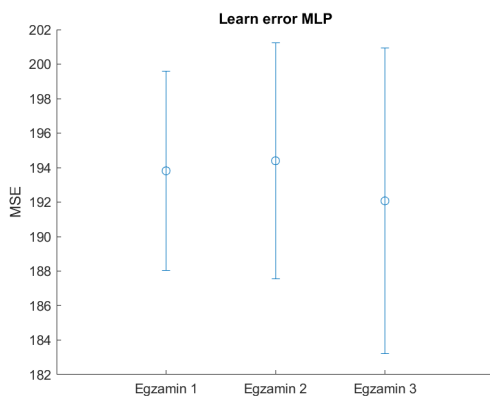
Dla każdej konfiguracji zbadano również wpływ ilości neuronów w poszczególnych warstwach. Ilość neuronów zmieniano między 10 a 100 ze skokiem 10. Wykresy błędów dla wybranych konfiguracji znajdują się na następnych stronach.



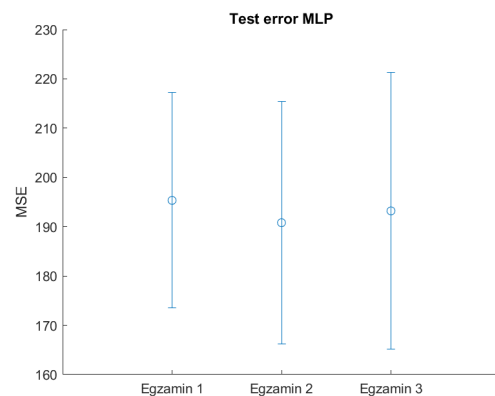
(a) MSE uczenia dla 20 neuronów.



(b) MSE testu dla 20 neuronów.



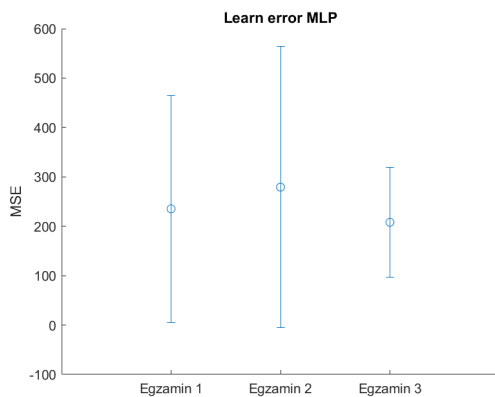
(c) MSE uczenia dla 50 neuronów.



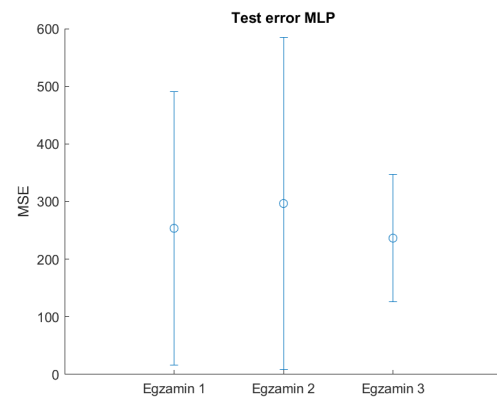
(d) MSE testu dla 50 neuronów.

Rysunek 1: Błąd uczenia i testowania sieci dla X prób dla kolejnych egzaminów, **funkcje: purelin, purelin.**

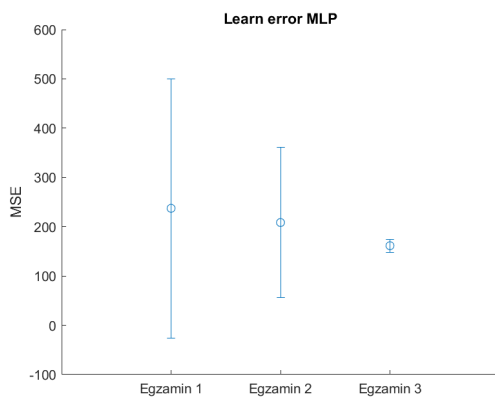
Pierwszym omawianym zestawieniem jest prosty przykład dla kombinacji funkcji purelin oraz purelin. Otrzymane wartości błędów były jednymi z najmniejszych wśród wszystkich konfiguracji, z akceptowalnym odchyleniem. W przypadku tak prostej sieci ilość neuronów nie wpłynęła praktycznie wcale na otrzymane wyniki.



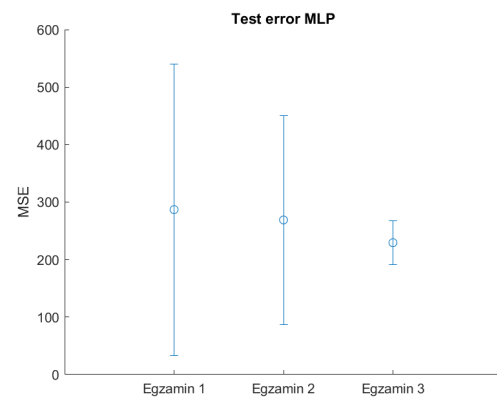
(a) MSE uczenia dla 20 neuronów.



(b) MSE testu dla 20 neuronów.



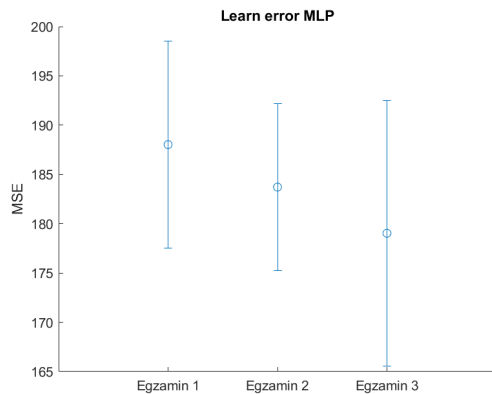
(c) MSE uczenia dla 50 neuronów.



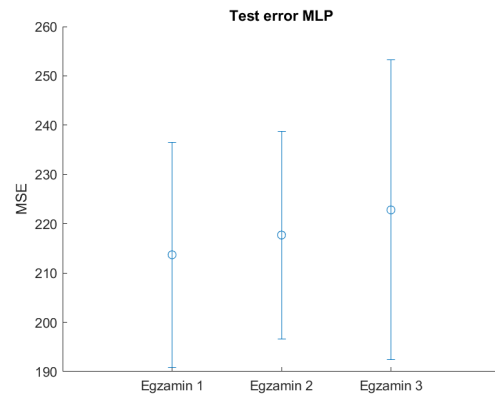
(d) MSE testu dla 50 neuronów.

Rysunek 2: Błąd uczenia i testowania sieci dla X prób dla kolejnych egzaminów, **funkcje: tansig, tansig**.

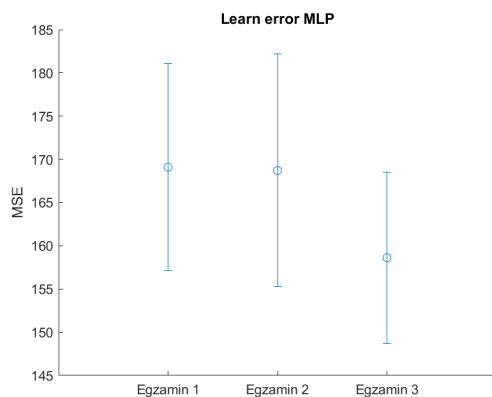
Przykładem jednej z konfiguracji, która daje mierne rezultaty jest konfiguracja z wykorzystaniem funkcji tansig oraz tansig na wyjściu. Wartości błędów były jednymi z większych, jednak istotny jest fakt, że wyniki dla trzeciego egzaminu były zbieżne, w przeciwieństwie do większości pozostałych konfiguracji. Ilość neuronów nie miała dużego wpływu na otrzymywane błędy.



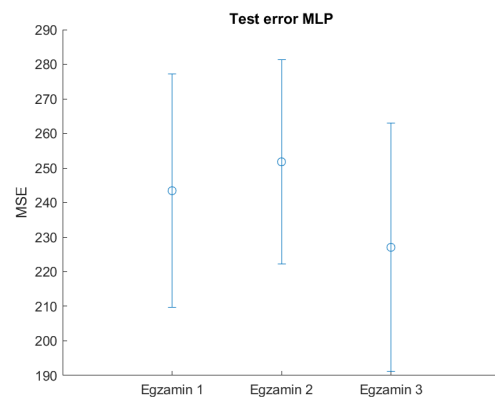
(a) MSE uczenia dla 20 neuronów.



(b) MSE testu dla 20 neuronów.



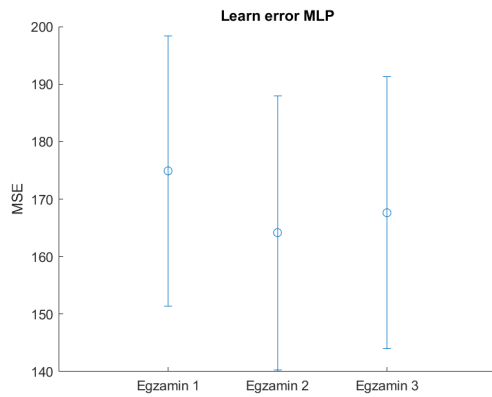
(c) MSE uczenia dla 50 neuronów.



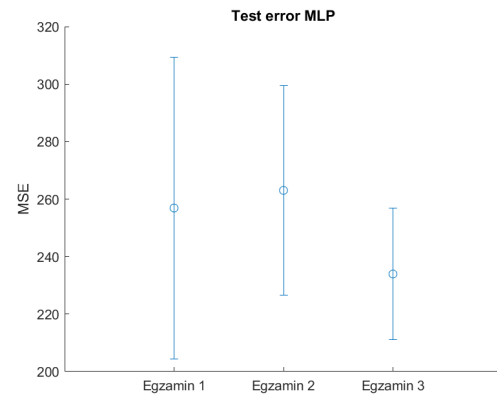
(d) MSE testu dla 50 neuronów.

Rysunek 3: Błąd uczenia i testowania sieci dla X prób dla kolejnych egzaminów, **funkcje: tansig, purelin**.

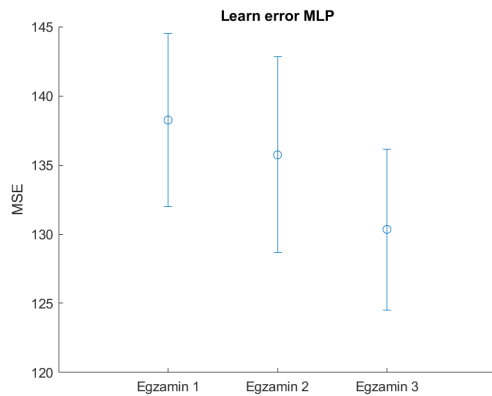
Kombinacja tansig oraz purelin daje bardzo niskie wartości błędów w porównaniu z pozostałymi metodami, jednak z dość dużym rozstrzałem między otrzymywanymi wartościami. Błędy uczenia nieznacznie spadały wraz ze wzrostem ilości neuronów w warstwie, jednak tendencja dla testu była odwrotna. Najlepsze wyniki udało się uzyskać dla 20 oraz 30 neuronów.



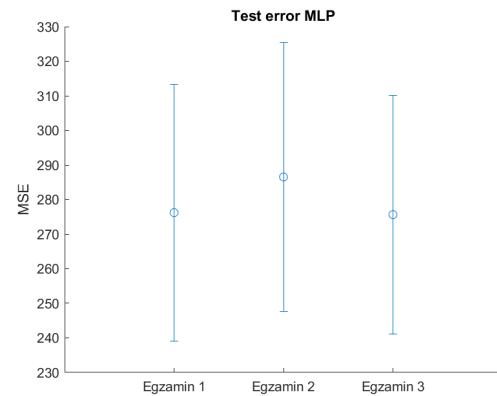
(a) MSE uczenia dla 40 neuronów w pierwszej warstwie, 20 w drugiej.



(b) MSE testu dla 40 neuronów w pierwszej warstwie, 20 w drugiej.



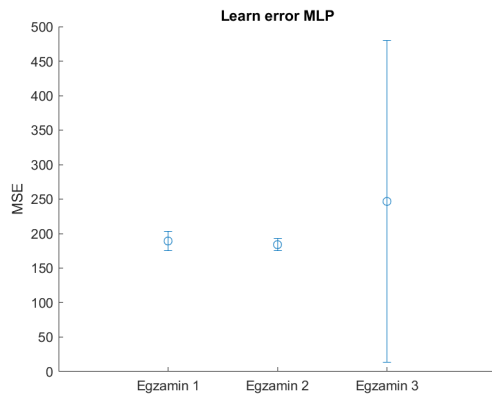
(c) MSE uczenia dla 100 neuronów w pierwszej warstwie, 50 w drugiej.



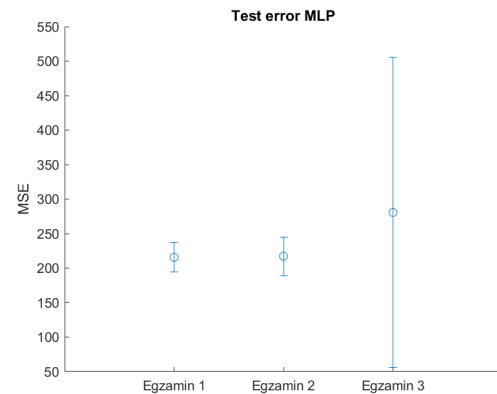
(d) MSE testu dla 100 neuronów w pierwszej warstwie, 50 w drugiej.

Rysunek 4: Błąd uczenia i testowania sieci dla X prób dla kolejnych egzaminów, **funkcje: radbas, tansig, purelin**.

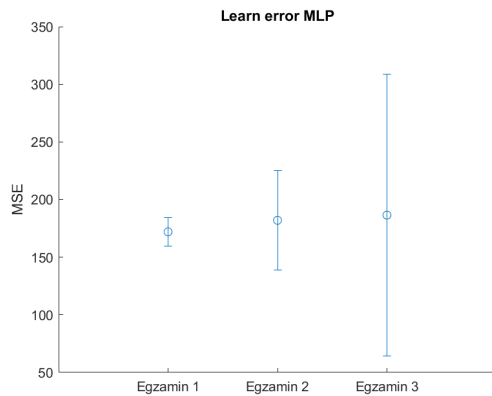
Postanowiono również przetestować sieci o dwóch warstwach w bardziej egzotycznych konfiguracjach, dla przykładu konfiguracja radbas, tansig, purelin dała zadziwiająco niskie wartości błędów dla próby uczenia, który gwałtownie zmieniał się wraz z dokładaniem ilości neuronów. Pod względem błędu uczenia była to najlepsza konfiguracja.



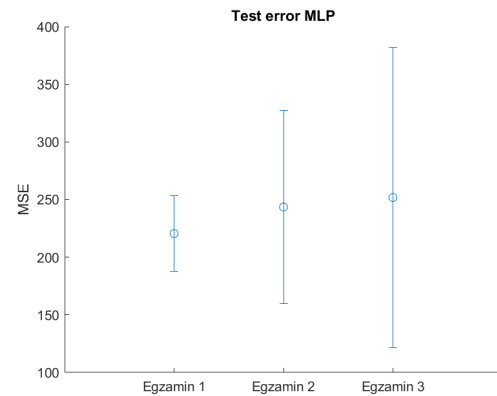
(a) MSE uczenia dla 50 neuronów w pierwszej warstwie, 50 w drugiej.



(b) MSE testu dla 50 neuronów w pierwszej warstwie, 50 w drugiej.



(c) MSE uczenia dla 50 neuronów w pierwszej warstwie, 50 w drugiej.



(d) MSE testu dla 50 neuronów w pierwszej warstwie, 50 w drugiej.

Rysunek 5: Błąd uczenia i testowania sieci dla X prób dla kolejnych egzaminów, **funkcje: logsig, tansig, tansig**.

Kombinacja logsig, tansig oraz tansig dała ciekawe rezultaty, wartości błędu dla wszystkich prób nie licząc trzeciego egzaminu były do siebie zbliżone. Wyniki dla trzeciego egzaminu jednak zbyt odstają od reszty, by móc brać tą konfigurację w kolejnych badaniach.

W zestawieniu nie objęto wykresów dla pozostałych kombinacji (na przykład tansig, tansig), gdyż niezależnie od ilości neuronów wyniki były bardzo złe.

3 Badanie wpływu ilości kolumn na wyniki sieci

3.1 Korelacja pomiędzy danymi

W ramach projektu został obliczony współczynnik korelacji *Pearsona* pomiędzy danymi z repozytorium. Dzięki niemu możemy sprawdzić jak bardzo zależne są od siebie dane, a dzięki temu dostosować informacje użyte do uczenia sieci.

Do obliczenia współczynnika korelacji została użyta funkcja pakietu MATLAB `corr`.

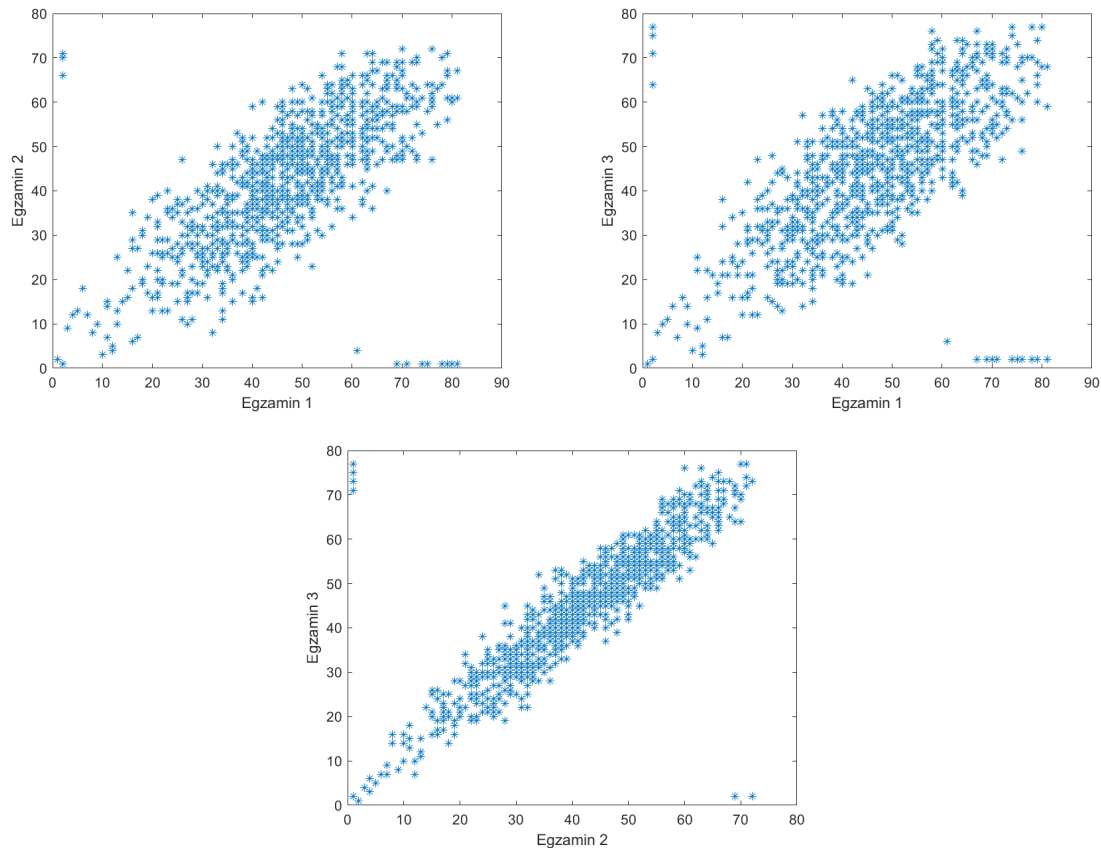
3.1.1 Korelacja pomiędzy wynikami egzaminów

Na początku został wyznaczony współczynnik korelacji pomiędzy wynikami egzaminów. Otrzymane wyniki przedstawiono w tabeli poniżej.

Egzamin 1 i 2	Egzamin 1 i 3	Egzamin 2 i 3
0.6453	0.6435	0.8578

Dla wszystkich kombinacji otrzymaliśmy wartości większe od 0.5, możemy więc uznać, że dane są od siebie zależne. Największa korelacja występuje pomiędzy egzaminem 2 oraz 3 - współczynnik jest równy 0.86.

Na rysunkach poniżej zależności pomiędzy zbiorami zostały przedstawione w sposób graficzny.



Rysunek 6: Zależności pomiędzy wynikami egzaminów przedstawione w sposób graficzny.

3.1.2 Korelacja pomiędzy czynnikami środowiskowymi a wynikami egzaminów

W dalszej części analizy zbioru danych została zbadana zależność pomiędzy czynnikami, będącymi wejściem sieci neuronowej, a wynikami kolejnych egzaminów. Otrzymane współczynniki korelacji przedstawiono w tabeli poniżej.

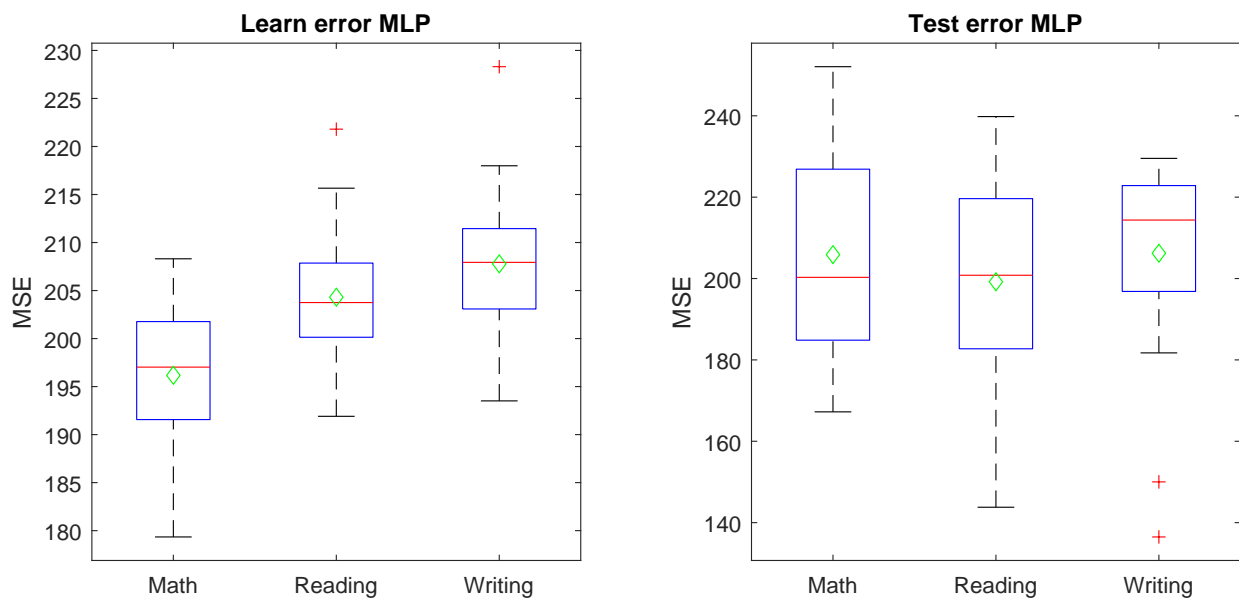
Czynnik środowiskowy	Egzamin 1	Egzamin 2	Egzamin 3
Płeć	0.1558	-0.1886	-0.2396
Rasa	0.1771	0.0770	0.0907
Wykształcenie rodzica	-0.0584	-0.0306	-0.0571
Przystąpienie do kursu	0.3269	0.1906	0.2182
Dieta	-0.1564	-0.1838	-0.2647

Otrzymane wartości są dość niskie, nie istnieje wyraźna korelacja pomiędzy którąś z tych cech, a wynikami. Najmniejszą zależność obserwujemy pomiędzy wynikami, wykształceniem rodziców - są one najbliższe zeru. Największe znaczenie ma przystąpienie do kursu przygotowawczego, w dalszej kolejności dieta oraz płeć.

3.2 Testowanie sieci neuronowej z pomniejszonym zbiorem uczącym

Dla wybranych najlepszych parametrów sieci zostało przeprowadzone uczenie ze zmniejszoną ilością kolumn. Celem tego zabiegu było zbadanie, jaki wpływ mają te czynniki na poprawność działania sieci.

3.2.1 Uzyskane błędy w przypadku usunięcia pierwszej kolumny

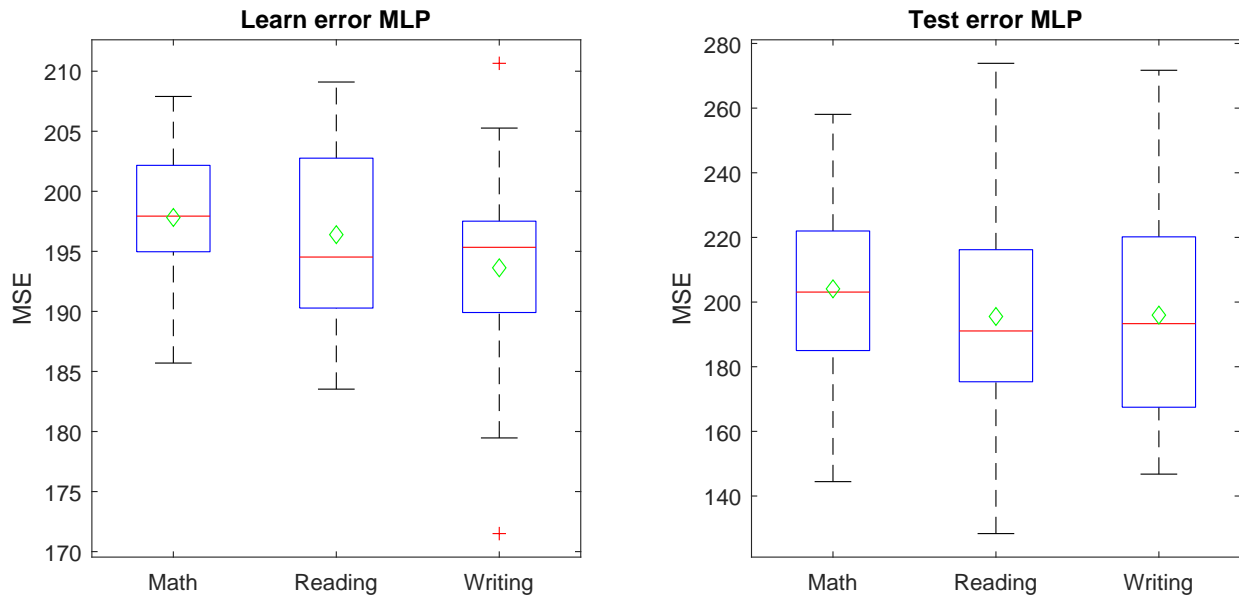


(a) MSE uczenia dla 20 neuronów w pierwszej warstwie.

(b) MSE testu dla 20 neuronów w pierwszej warstwie.

Rysunek 7: Błąd uczenia i testowania sieci w przypadku usunięcia pierwszej kolumny z danymi, funkcje: **purelin**, **purelin**.

3.2.2 Uzyskane błędy w przypadku usunięcia drugiej kolumny

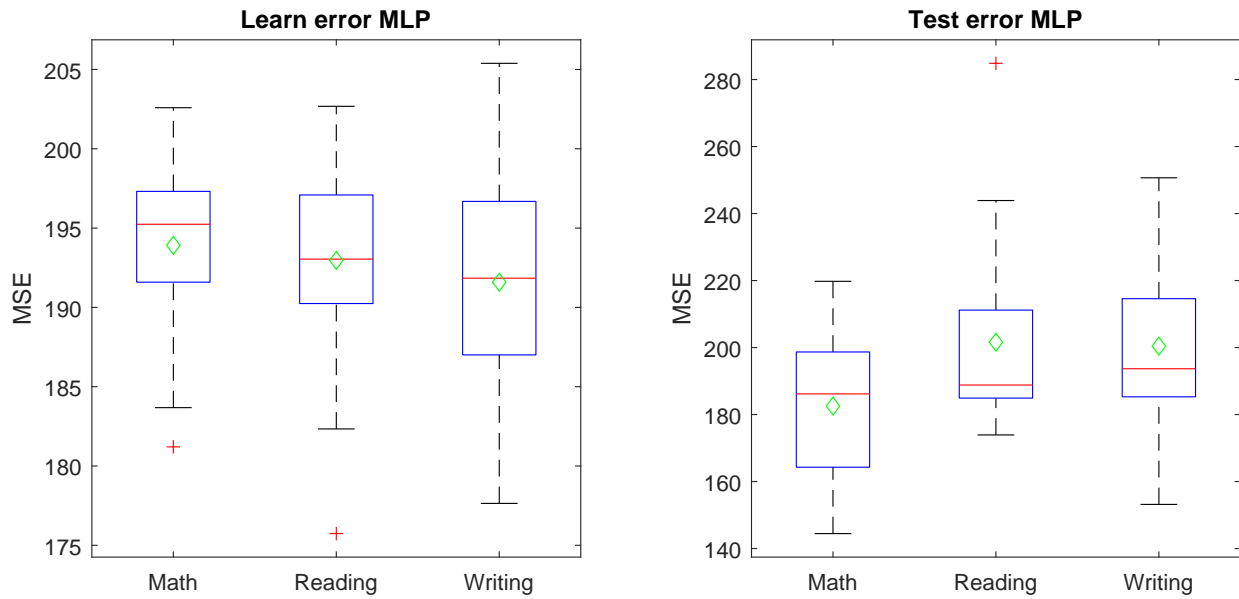


(a) MSE uczenia dla 20 neuronów w pierwszej warstwie.

(b) MSE testu dla 20 neuronów w pierwszej warstwie.

Rysunek 8: Błąd uczenia i testowania sieci w przypadku usunięcia drugiej kolumny z danymi, funkcje: **purelin**, **purelin**.

3.2.3 Uzyskane błędy w przypadku usunięcia trzeciej kolumny

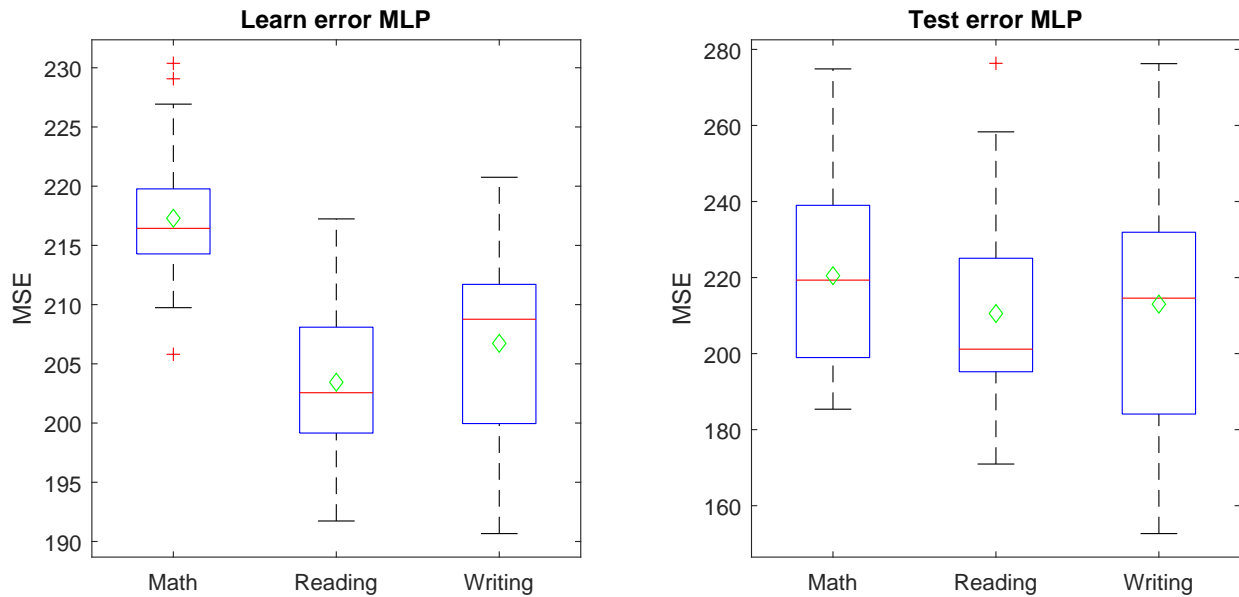


(a) MSE uczenia dla 20 neuronów w pierwszej warstwie.

(b) MSE testu dla 20 neuronów w pierwszej warstwie.

Rysunek 9: Błąd uczenia i testowania sieci w przypadku usunięcia trzeciej kolumny z danymi, funkcje: **purelin**, **purelin**.

3.2.4 Uzyskane błędy w przypadku usunięcia czwartej kolumny

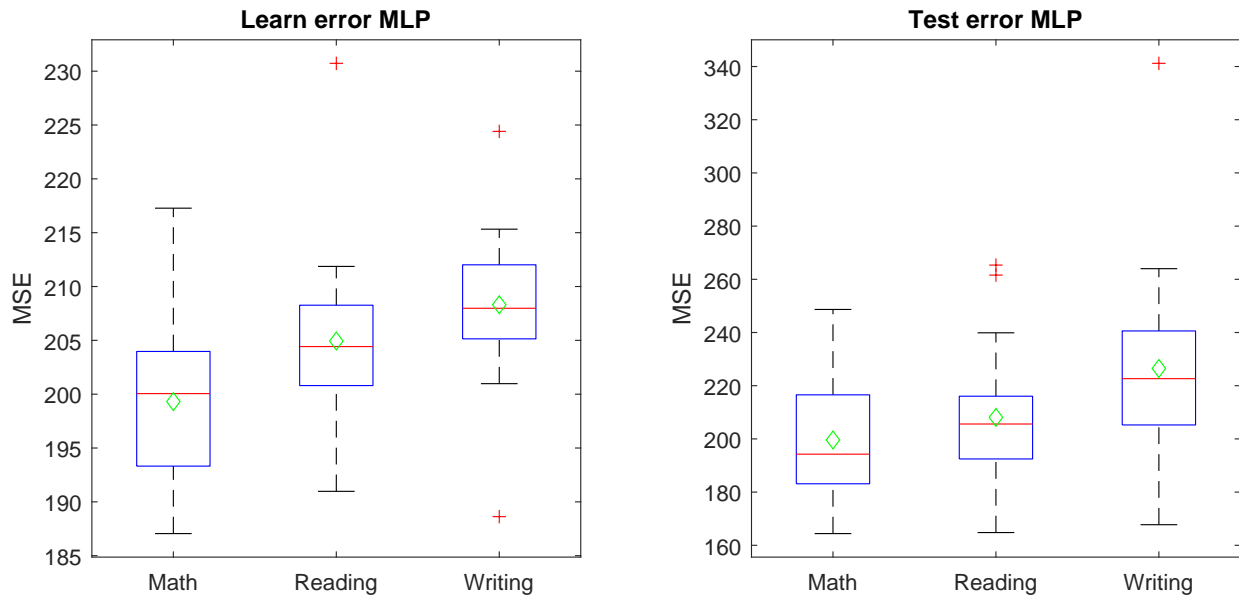


(a) MSE uczenia dla 20 neuronów w pierwszej warstwie.

(b) MSE testu dla 20 neuronów w pierwszej warstwie.

Rysunek 10: Błąd uczenia i testowania sieci w przypadku usunięcia czwartej kolumny z danymi, funkcje: **purelin**, **purelin**.

3.2.5 Uzyskane błędy w przypadku usunięcia piątej kolumny



(a) MSE uczenia dla 20 neuronów w pierwszej warstwie.

(b) MSE testu dla 20 neuronów w pierwszej warstwie.

Rysunek 11: Błąd uczenia i testowania sieci w przypadku usunięcia piątej kolumny z danymi, funkcje: purelin, purelin.

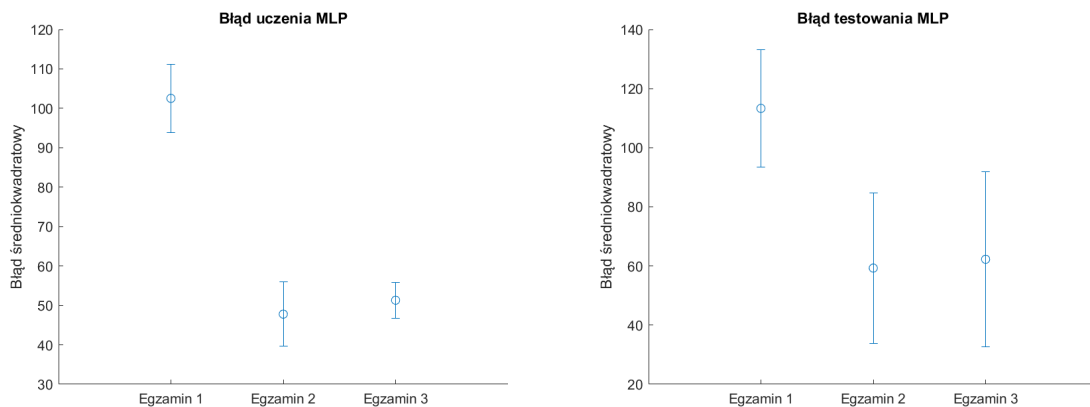
3.3 Wnioski

Badano wersję z funkcjami purelin, purelin oraz 20 neuronami w pierwszej warstwie.

4 Przewidywanie wyniku bazując wynikach pozostałych egzaminów

Jak zostało zauważone podczas badania korelacji, wyniki egzaminów są od siebie zależne (wysoki współczynnik korelacji). Korzystając z wyników z dwóch egzaminów podjęto próbę stworzenia sieci obliczającą wynik z trzeciego egzaminu.

Otrzymane błędy testowania zostały przedstawione poniżej.



Rysunek 12: Błąd uczenia i testowania sieci dla X prób dla kolejnych egzaminów

Na podstawie obserwacji wielkości błędów, można zauważyć, że dla egzaminów 2 oraz 3 sieć osiąga dobre wyniki - na poziomie wartości 50. Przeprowadzono badanie błędu średniokwadratowego dla uczenia oraz testu różnych konfiguracji funkcji aktywacji, ilości warstw oraz ilości neuronów w poszczególnych warstwach.

Otrzymane wyniki można powiązać z analizowaną wcześniej korelacją danych - wyniki egzaminów 2 i 3 są ze sobą w większym stopniu powiązane, niż egzamin 1 z egzaminem 2 lub 3.

5 Podsumowanie

Literatura