POLITECHNIKA WARSZAWSKA

WYDZIAŁ MATEMATYKI  
 I NAUK INFORMACYJNYCH

Przewidywanie szeregów czasowych i indeksów giełdowych

Wyniki i analiza testów

Projekt nr 2 na potrzeby przedmiotu „Sieci Neuronowe”

Maciej Kachniarz

Tomasz Luśtyk

Spis treści

[1 Finalna aplikacja 4](#_Toc410632951)

[2 Pliki wejściowe 4](#_Toc410632952)

[2.1 Szeregi czasowe 4](#_Toc410632953)

[2.2 Indeks giełdowy WIG20 4](#_Toc410632954)

[3 Założenia testów 5](#_Toc410632955)

[4 Testy szeregów czasowych 5](#_Toc410632956)

[4.1 Test wstępny 5](#_Toc410632957)

[4.1.1 Dane wejściowe 5](#_Toc410632958)

[4.1.2 Wyniki i wnioski 5](#_Toc410632959)

[4.2 Test zwiększenia iteracji 6](#_Toc410632960)

[4.2.1 Dane wejściowe 6](#_Toc410632961)

[4.2.2 Wyniki i wnioski 7](#_Toc410632962)

[4.2.3 Weryfikacja empiryczna 7](#_Toc410632963)

[4.3 Test zmiany współczynników 7](#_Toc410632964)

[4.3.1 Dane wejściowe 7](#_Toc410632965)

[4.3.2 Wyniki i wnioski 7](#_Toc410632966)

[4.4 Wnioski końcowe 8](#_Toc410632967)

[5 Testy Indeksu giełdowego WIG20 8](#_Toc410632968)

[5.1 Oznaczenia 8](#_Toc410632969)

[5.2 Obserwacje 9](#_Toc410632970)

[5.3 Test wstępny dla danych 1-wymiarowych 9](#_Toc410632971)

[5.3.1 Dane wejściowe 9](#_Toc410632972)

[5.3.2 Wyniki i wnioski 9](#_Toc410632973)

[5.3.3 Wykresy 11](#_Toc410632974)

[5.4 Test wstępny dla danych 4-wymiarowych 11](#_Toc410632975)

[5.4.1 Dane wejściowe 11](#_Toc410632976)

[5.4.2 Wyniki i wnioski 12](#_Toc410632977)

[5.4.3 Wykresy 13](#_Toc410632978)

[5.5 Test dla danych wielowymiarowych 14](#_Toc410632979)

[5.5.1 Dane wejściowe 14](#_Toc410632980)

[5.5.2 Wyniki i wnioski 14](#_Toc410632981)

[5.6 Test dla danych wielowymiarowych – sprawdzenie PCA 18](#_Toc410632982)

[5.6.1 Dane wejściowe 18](#_Toc410632983)

[5.6.2 Wyniki i wnioski 19](#_Toc410632984)

# Finalna aplikacja

Utworzono aplikację do predykcji wartości indeksów giełdowych oraz szeregów czasowych. Podana aplikacja posiada możliwość ustalania następujących parametrów.

Parametry budowy sieci:

* Rodzaj sieci (do wyboru: MLP, Elman, Jordan)
* Ilość neuronów w poszczególnych warstwach (tzw. architektura sieci)
* Rodzaj funkcji pobudzenia (do wyboru: unipolarna, bipolarna)

Parametry algorytmu uczącego:

* Współczynnik uczenia
* Współczynnik bezwładności
* Współczynnik podziału na zbiór testowy i treningowy

Parametry przeprowadzanego testu:

* Ilość iteracji

Parametry danych testowych:

* Rodzaj przewidywanego indeksu (do wyboru: szereg czasowy, WIG20, S&P 500)
* Optymalizacja danych wejściowych za pomocą algorytmu PCA (do wyboru: brak, liczba kolumn)
* Okno czasowe danych wejściowych (do wyboru: z ostatniego dnia, tygodnia, 2 tygodni, miesiąca, 3 miesięcy, roku)
* Częstość próbkowania danych wejściowych
* Krok

# Pliki wejściowe

## Szeregi czasowe

Pojedyncza wartość szeregu określonego wzorem dla dla kolejnych 1000 wyrazów

## Indeks giełdowy WIG20

Pojedynczy dzień opisany jest następującymi wartościami:

* Otwarcie (wartość indeksu giełdowego podczas otwarcia giełdy, poniżej analogicznie)
* Maksimum
* Minimum
* Zamknięcie
* Zmiana procentowa (od poprzedniego dnia)
* Wolumen
* SMA15 (Oscylator)
* SMA30 (Oscylator)
* SMA45 (Oscylator)
* WMA15 (Oscylator)
* WMA30 (Oscylator)
* WMA45 (Oscylator)
* EMA (Oscylator)
* RSI7 (Oscylator)
* RSI9 (Oscylator)
* RSI14 (Oscylator)
* 14-dniowy oscylator stochastyczny (Oscylator)
* Kurs euro / złoty
* Kurs dolar / złoty
* Kurs frank szwajcarski / złoty
* Kurs funta brytyjskiego / złoty
* Kurs złota

Dane obejmują okres od 2010-03-05 – 2014-10-30. Dla niektórych wartości (kursów walut, złota, wolumen) od dat późniejszych. Wartości oznaczone atrybutem „Oscylator” przekazywane są do sieci tylko z pierwszego dnia okna czasowego, pozostałe z całego okna czasowego.

# Założenia testów

Do wszystkich przeprowadzonych testów użyto stałych parametrów:

* Rodzaj funkcji pobudzenia: unipolarna
* Współczynnik podziału na zbiór testowy i treningowy (treningowy zawiera walidacyjny): 0.7
* Częstość próbkowania danych wejściowych: dzień (dla szeregów czasowych kolejna wartość)
* Krok: dzień (dla szeregów czasowych kolejna wartość)

Funkcją błędu jest błąd średniokwadratowy.

# Testy szeregów czasowych

Opis przeprowadzonych testów szeregów czasowych.

## Test wstępny

Początkowo przeprowadzono testy ogólne aby wstępnie wyłonić najlepsze parametry startowe.

### Dane wejściowe

Testowano dla następujących kombinacji parametrów:

* Rodzaj sieci: MLP, Elman, Jordan
* Architektura sieci:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3-1 | 5-1 | 10-1 | 20-1 | 30-1 |
| 3-3-1 | 5-5-1 | 10-10-1 | 20-20-1 | 30-30-1 |

* Współczynnik uczenia: 0,2; 0.4; 0,6; 0,8; 1.0
* Współczynnik bezwładności: 0; 0,2; 0.4; 0,6; 0,8
* Ilość iteracji: 100
* Liczba powtórzeń: 5

### Wyniki i wnioski

* Najlepszy wynik otrzymywano dla danych wejściowych:
  + Rodzaj sieci: MLP
  + Architektura: 10-10-1
  + Współczynnik uczenia: 1.0
  + Współczynnik bezwładności: 0.0

I wynosił on 1.35E-05

* Najlepsze wyniki otrzymano kolejno dla sieci typu

|  |  |
| --- | --- |
| **Typ sieci** | **Średni błąd** |
| Elman | 0,000488 |
| Jordan | 0,000526 |
| MLP | 0,000427 |

Pomimo tej różnicy najlepsze wyniki (rzędu 1E-05) osiągane są przez wszystkie sieci. Parametr ten zdaje się nie mieć bardzo dużego znaczenia.

* Najlepsze wyniki otrzymano kolejno dla architektury:

|  |  |
| --- | --- |
| **Architektura** | **Średni błąd** |
| 10~10~1 | 0,000032 |
| 20~20~1 | 0,000044 |
| 30~30~1 | 0,000071 |
| 20~1 | 0,000275 |
| 30~1 | 0,000286 |
| 5~5~1 | 0,000363 |
| 10~1 | 0,000523 |
| 5~1 | 0,000937 |
| 3~1 | 0,001256 |
| 3~3~1 | 0,001467 |

|  |  |
| --- | --- |
| **Architektura** | **Minimalny błąd** |
| 10~10~1 | 0,000014 |
| 20~20~1 | 0,000014 |
| 5~5~1 | 0,000025 |
| 30~30~1 | 0,000025 |
| 20~1 | 0,000122 |
| 30~1 | 0,000123 |
| 3~1 | 0,000294 |
| 10~1 | 0,000332 |
| 5~1 | 0,000851 |
| 3~3~1 | 0,001112 |

Zdecydowanie najlepsze wyniki otrzymano dla architektury typu X-X-1 dla X równych 10, 20 itp. W porównaniu do pozostałych różnica wynosi nawet do 2 rzędów.

* Dla współczynnika uczenia zaobserwowano zależność im wyższy współczynnik tym lepszy wynik. Dla współczynnika bezwładności zaobserwowano zależność odwrotną. Najlepsze wyniki otrzymano dla współczynnika uczenia równego 1 i współczynnika bezwładności równego 0.

## Test zwiększenia iteracji

Celem testu jest sprawdzenie czy zwiększenie ilości iteracji polepszy wynik.

### Dane wejściowe

Testowano dla następujących kombinacji parametrów (najlepsze z testu wstępnego):

* Rodzaj sieci: Elman
* Architektura sieci: 10-10-1
* Współczynnik uczenia: 1
* Współczynnik bezwładności: 0
* Ilość iteracji: 100, 500, 1000
* Liczba powtórzeń: 5

### Wyniki i wnioski

Najlepszy wynik otrzymano dla 1000 iteracji i wynosił on 8,14604E-06

|  |  |
| --- | --- |
| **Ilość iteracji** | **Błąd** |
| 1000 | 0,000008 |
| 500 | 0,000018 |
| 100 | 0,000051 |

Zwiększenie ilości iteracji zmniejsza wartość błędu.

### Weryfikacja empiryczna

W celu sprawdzenia czy dalsze zwiększanie ilości iteracji zmniejsza wartość błędu postanowiono przeprowadzić jednorazowy (liczba powtórzeń równa 1) test dla ilości iteracji równej 10k.

Otrzymany błąd wynosił 8,05E-07

|  |  |
| --- | --- |
| **Numer iteracji** | **Błąd** |
| 1k | 2,30E-06 |
| 2k | 1,67E-06 |
| 3k | 1,83E-06 |
| 4k | 1,69E-06 |
| 5k | 1,47E-06 |
| 6k | 1,27E-06 |
| 7k | 1,10E-06 |
| 8k | 9,97E-07 |
| 9k | 8,83E-07 |
| 10k | 8,05E-07 |

Poprawa wyniku była jednak bardzo powolna i istniały iteracje (np. 2000-3000) podczas których dochodziło do pogorszenia wyniku.

## Test zmiany współczynników

Podczas testu wstępnego ustalono, zależność im wyższy współczynnik tym lepszy wynik. Celem testu jest sprawdzenie czy dalsze zwiększenie współczynnika uczenia polepszy wynik.

### Dane wejściowe

Testowano dla następujących kombinacji parametrów (najlepsze z testu wstępnego):

* Rodzaj sieci: Elman
* Architektura sieci: 10-10-1
* Współczynnik uczenia: od 1 do 2 co 0.1
* Współczynnik bezwładności: 0
* Ilość iteracji: 100
* Liczba powtórzeń: 10

### Wyniki i wnioski

Najlepszy wynik otrzymano dla współczynnika uczenia równego … i wynosił on 0,000050

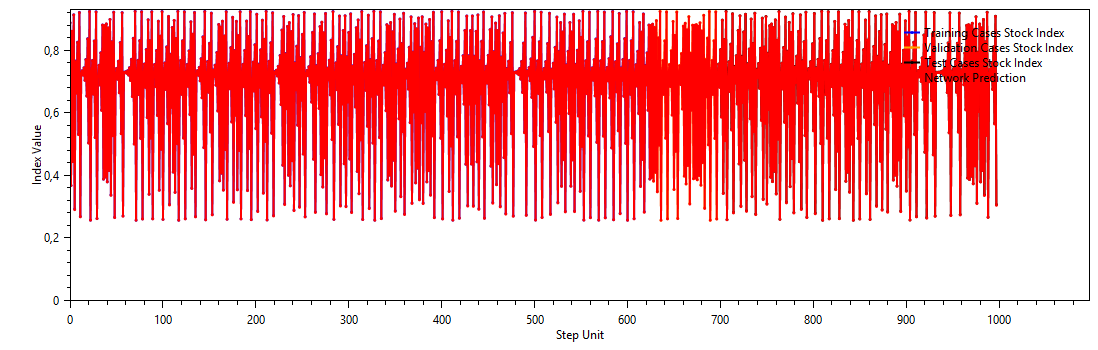
|  |  |
| --- | --- |
| **Współczynnik uczenia** | **Średni błąd** |
| 1,4 | 0,000093 |
| 1,2 | 0,000105 |
| 1 | 0,000112 |
| 1,5 | 0,000118 |
| 1,6 | 0,000169 |
| 1,1 | 0,000180 |
| 1,3 | 0,000268 |
| 1,7 | 0,000311 |
| 1,9 | 0,000335 |
| 2 | 0,001044 |

## Wnioski końcowe

Najlepsze parametry:

* Rodzaj sieci: Elman
* Architektura sieci: 10-10-1
* Współczynnik uczenia: 1,2
* Współczynnik bezwładności: 0
* Ilość iteracji: im więcej tym lepiej

Najlepszy osiągnięty wynik: 8,05E-07



Rysunek 1. Pokrycie wykresu szeregu czasowego przez test z punktu 4.2.3.

# Testy Indeksu giełdowego WIG20

Początkowe testy dla wielowymiarowych danych wejściowych zwracały predykcję sieci nie pokrywającą się w ogóle z rzeczywistością w związku z tym przeprowadzono testy dla danych wejściowych o jak najmniejszym rozmiarze i rozmiar ten metodycznie zwiększano.

## Oznaczenia

* WPPZK – współczynnik poprawnych przewidzeń zmiany kierunku
* WPPZK\_UP – współczynnik poprawnych przewidzeń zmiany kierunku dla rzeczywistej zmiany do góry
* WPPZK\_DOWN – współczynnik poprawnych przewidzeń zmiany kierunku dla rzeczywistej zmiany do dołu
* WPPZK\_TRAIN – współczynnik poprawnych przewidzeń zmiany kierunku na zbiorze treningowym

## Obserwacje

* Jeśli WPPZK\_TRAIN jest znacznie niższy niż WPPZK może to oznaczać, że sieć nie tyle osiągnęła wysoki wynik ze względu na poprawne działanie, ale mogło to wynikać z innych okoliczności. Przypadek ten oznaczany będzie w poniższym dokumencie jako CASE1 .
* Jeśli ze współczynników WPPZK\_UP i WPPZK\_DOWN jest bardzo niski a drugi bardzo wysoki a WPPZK będący ich średnią ważoną wynosi więcej niż 50% i jest wielkości trendu oznacza to, że sieć zwraca wartość maksymalną lub minimalną i nie przewiduje poprawnie. Przypadek ten oznaczany będzie w poniższym dokumencie jako CASE2 .

## Test wstępny dla danych 1-wymiarowych

Początkowo przeprowadzono testy ogólne dla pliku wejściowego jedynie z wartością zamknięcia.

### Dane wejściowe

Testowano dla następujących kombinacji parametrów:

* Rodzaj sieci: MLP, Elman, Jordan
* Architektura sieci:

|  |  |  |
| --- | --- | --- |
| 3-1 | 10-1 | 30-1 |
| 3-3-1 | 10-10-1 | 30-30-1 |

* Współczynnik uczenia: 0,25; 0,5; 0,75; 1
* Współczynnik bezwładności: 0; 0,25; 0,5; 0,75; 1
* Ilość iteracji: 1000
* Liczba powtórzeń: 2
* Okno czasowe danych wejściowych: dzień, tydzień, 2 tygodnie, miesiąc

### Wyniki i wnioski

* Najmniejszy błąd otrzymywano dla danych wejściowych:
  + Rodzaj sieci: Elman
  + Architektura: 10-10-1
  + Współczynnik uczenia: 0,75
  + Współczynnik bezwładności: 0,25
  + Okno czasowe danych wejściowych: miesiąc

I wynosił on 885 (WPPZK 53%, WPPZK\_DOWN 8%, WPPZK\_UP 97%, WPPZK\_TRAIN 56%). Przypadek typu CASE2.

* Największy WPPZK:
  + Rodzaj sieci: Elman
  + Architektura: 3-1
  + Współczynnik uczenia: 1
  + Współczynnik bezwładności: 0,75
  + Okno czasowe danych wejściowych: 2 tygodnie

I wynosił on 59,94% (błąd 1021, WPPZK\_DOWN 54%, WPPZK\_UP 66%, WPPZK\_TRAIN 53%)

* Największy WPPZK\_TRAIN:
  + Rodzaj sieci: Elman
  + Architektura: 3-3-1
  + Współczynnik uczenia: 1
  + Współczynnik bezwładności: 0
  + Okno czasowe danych wejściowych: miesiąc

I wynosił on 57,12% (błąd 1371, WPPZK 49%, WPPZK\_DOWN 44%, WPPZK\_UP 54%)

* Najlepsze wyniki otrzymano dla sieci typu:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Rodzaj sieci** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| Elman | **1326** | **885** | **53,55%** | **59,94%** |
| Jordan | 2005 | 896 | 53,19% | 59,88% |
| MLP | 1974 | 900 | 53,25% | 59,06% |

Sieci typu Elman daje nieznacznie, ale za to prawie zawsze, lepsze rezultaty niż inne typy sieci.

* Najlepsze wyniki otrzymano dla architektury:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Architektura sieci** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| 3~1 | **1150** | 946 | 53,03% | **59,94%** |
| 10~1 | 1249 | 943 | 53,10% | 58,77% |
| 10~10~1 | 1745 | **885** | 52,93% | 59,06% |
| 30~30~1 | 1900 | 907 | 53,44% | 59,88% |
| 3~3~1 | 2254 | 953 | **53,92%** | 59,88% |
| 30~1 | 2311 | 942 | 53,57% | 59,36% |

Architektury zwracają zbliżone wyniki.

* Najlepsze wyniki otrzymano dla współczynnika uczenia:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Współczynnik uczenia** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| 1 | **1452** | 895 | **53,85%** | **59,94%** |
| 0,75 | 1576 | **885** | 53,50% | 59,65% |
| 0,5 | 1760 | 897 | 53,30% | 59,88% |
| 0,25 | 2285 | 936 | 52,68% | 59,88% |
|  |  |  |  |  |

Wartości dla różnych współczynników uczenia są zbliżone, ale widać wyraźną tendencję, że współczynniki wyższe osiągają wyniki lepsze.

* Najlepsze wyniki otrzymano dla współczynnika bezwładności:

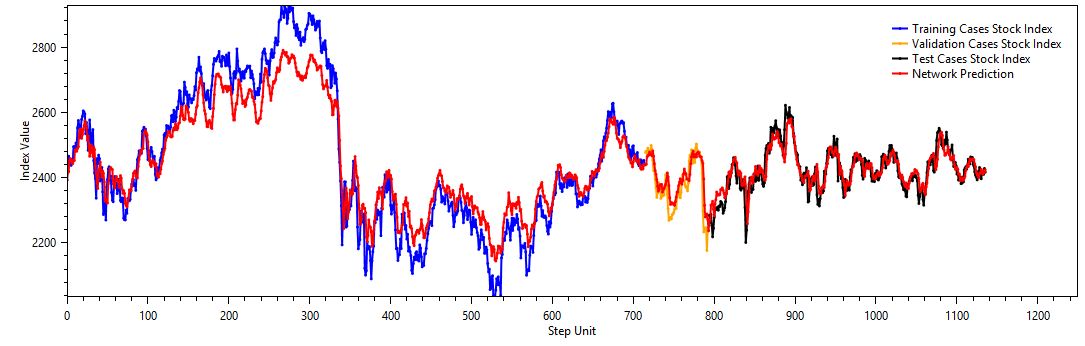
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Współczynnik bezwładności** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| 0 | **1048** | 916 | **54,94%** | 59,74% |
| 0,25 | 1061 | **885** | 53,26% | 59,88% |
| 0,5 | 1297 | 955 | 52,51% | 59,06% |
| 0,75 | 2165 | 965 | 52,83% | **59,94%** |
| 1 | 3271 | 1026 | 53,12% | 59,88% |

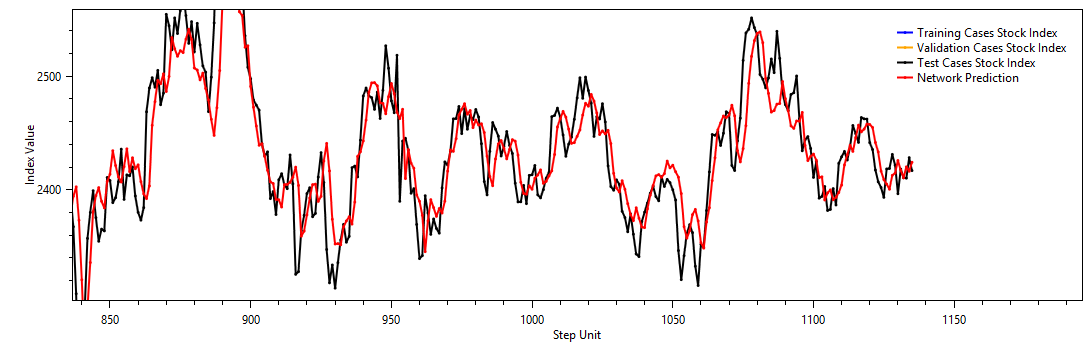
Odwrotnie niż dla współczynnika uczenia, najlepsze wyniki uzyskano dla współczynnika mniejszego bezwładności, zbliżonego do zera.

* Najlepsze wyniki otrzymano dla okna czasowego:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Okno czasowe** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| Day | **1563** | 988 | 52,78% | 59,88% |
| Week | 1713 | 960 | 54,69% | 59,88% |
| TwoWeeks | 1864 | 957 | **54,73%** | **59,94%** |
| Month | 1933 | **885** | 51,13% | 58,55% |
|  |  |  |  |  |

### Wykresy

Rysunek 2. Pokrycie wykresu WIG20 przez sieć dla danych dla największego WPPZK

Rysunek 3. Rysunek 2 na zbliżeniu

## Test wstępny dla danych 4-wymiarowych

Plik wejściowy zawiera wartość zamknięcia, wolumen, RSI9 i oscylator 14-dniowy.

### Dane wejściowe

Testowano dla następujących kombinacji parametrów:

* Rodzaj sieci: MLP, Elman, Jordan
* Architektura sieci:

|  |  |  |
| --- | --- | --- |
| 3-1 | 10-1 | 30-1 |
| 3-3-1 | 10-10-1 | 30-30-1 |

* Współczynnik uczenia: 0,25; 0,5; 0,75; 1
* Współczynnik bezwładności: 0; 0,25; 0,5; 0,75; 1
* Ilość iteracji: 1000
* Liczba powtórzeń: 2
* Okno czasowe danych wejściowych: dzień, tydzień, 2 tygodnie, miesiąc

### Wyniki i wnioski

* **Wnioski są identyczne jak dla testów z danymi 1-wymiarowymi**
* Najmniejszy błąd otrzymywano dla danych wejściowych:
  + Rodzaj sieci: Elman
  + Architektura: 10-10-1
  + Współczynnik uczenia: 0,75
  + Współczynnik bezwładności: 0,25
  + Okno czasowe danych wejściowych: miesiąc

I wynosił on 912 (WPPZK 56%, WPPZK\_DOWN 78%, WPPZK\_UP 35%, WPPZK\_TRAIN 55%).

* Największy WPPZK:
  + Rodzaj sieci: MLP
  + Architektura: 30-30-1
  + Współczynnik uczenia: 0,5
  + Współczynnik bezwładności: 0,25
  + Okno czasowe danych wejściowych: tydzień

I wynosił on 60,02% (błąd 1028, WPPZK\_DOWN 67%, WPPZK\_UP 53%, WPPZK\_TRAIN 52%)

* Największy WPPZK\_TRAIN:
  + Rodzaj sieci: Elman
  + Architektura: 10-10-1
  + Współczynnik uczenia: 0,5
  + Współczynnik bezwładności: 0,25
  + Okno czasowe danych wejściowych: miesiąc

I wynosił on 55,64% (błąd 919, WPPZK 53%, WPPZK\_DOWN 16%, WPPZK\_UP 90%)

* Najlepsze wyniki otrzymano dla sieci typu:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Etykiety wierszy** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| Elman | **1441** | **913** | **51,86%** | 59,30% |
| Jordan | 2202 | 960 | 50,65% | 58,33% |
| MLP | 2161 | 965 | 50,92% | **60,03%** |

* Najlepsze wyniki otrzymano dla architektury:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Etykiety wierszy** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| 10~1 | **1414** | 957 | 50,50% | 58,43% |
| 10~10~1 | 2090 | **913** | 51,61% | 58,33% |
| 3~1 | 1458 | 960 | 50,32% | 58,28% |
| 3~3~1 | 2692 | 958 | **52,02%** | 59,30% |
| 30~1 | 1679 | 963 | 50,55% | 58,28% |
| 30~30~1 | 2276 | 931 | 51,86% | **60,03%** |

* Najlepsze wyniki otrzymano dla współczynnika uczenia:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Etykiety wierszy** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| 0,25 | 2450 | 957 | 50,45% | 57,41% |
| 0,5 | 1917 | 919 | 51,02% | **60,03%** |
| 0,75 | 1719 | **913** | 51,45% | 59,88% |
| 1 | **1653** | 940 | **51,64%** | 59,30% |

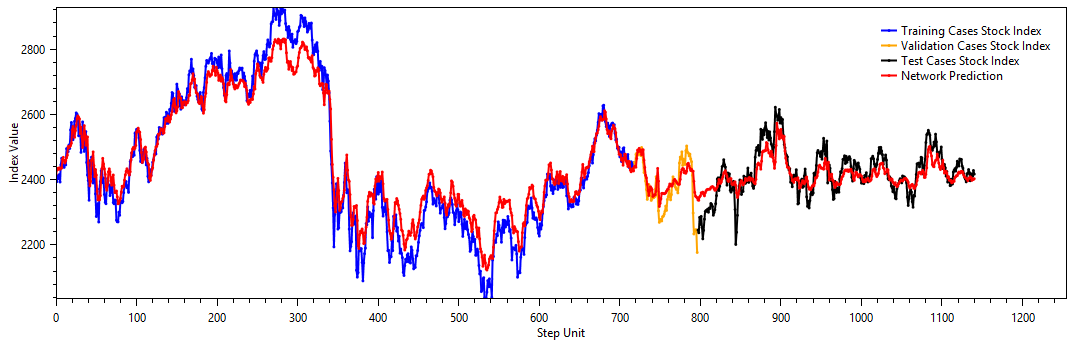
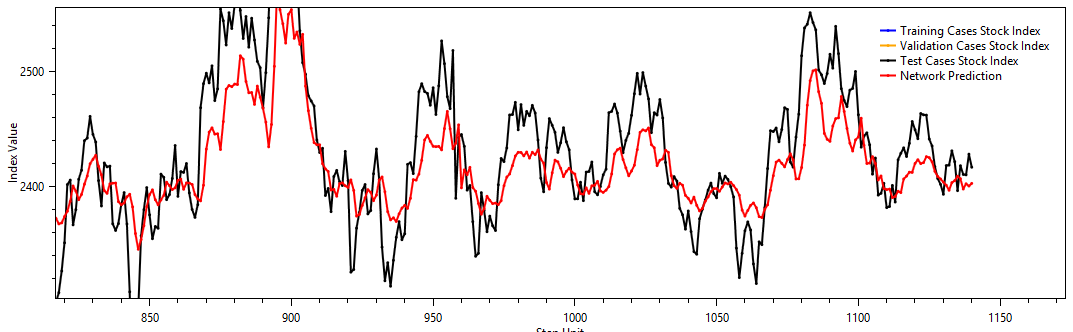
* Najlepsze wyniki otrzymano dla współczynnika bezwładności:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Etykiety wierszy** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| 0 | **1079** | 931 | **51,58%** | 58,33% |
| 0,25 | 1173 | **913** | 50,93% | **60,03%** |
| 0,5 | 1588 | 958 | 50,72% | 58,43% |
| 0,75 | 2347 | 1015 | 51,11% | 59,30% |
| 1 | 3487 | 1161 | 51,39% | 58,43% |

* Najlepsze wyniki otrzymano dla okna czasowego:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Etykiety wierszy** | **Średni błąd** | **Minimalny błąd** | **Średni WPPZK** | **Maksymalny WPPZK** |
| Day | 1744 | 971 | **51,35%** | 58,28% |
| Month | 2390 | **913** | 50,85% | 58,11% |
| TwoWeeks | 1882 | 926 | 51,06% | 58,33% |
| Week | **1723** | 965 | 51,31% | **60,03%** |

### Wykresy

Rysunek 4. Pokrycie wykresu WIG20 przez sieć dla danych dla największego WPPZK

Rysunek 5. Rysunek 4 na zbliżeniu

## Test dla danych wielowymiarowych

### Dane wejściowe

Testowano dla następujących kombinacji parametrów:

* Rodzaj sieci: MLP, Elman, Jordan
* Architektura sieci:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3-1 | 10-1 | 30-1 |  |  |
| 3-3-1 | 10-10-1 | 30-30-1 |  |  |

* Współczynnik uczenia: 0,25; 0,5; 0,75; 1.0
* Współczynnik bezwładności: 0; 0,25; 0.5; 0,75
* Ilość iteracji: 1000
* Liczba powtórzeń: 3
* Brak PCA
* Długość okna: Dzień; Tydzień

### Wyniki i wnioski

* Najmniejszy błąd otrzymano dla następujących danych wejściowych:
  + **Sieć** Elman
  + **Współczynnik uczenia**: 1.0
  + **Bezwładność**: 0.0
  + **Struktura sieci**: 30-30-1
  + **Długość okna**: Tydzień

|  |  |
| --- | --- |
| **BŁĄD** | 764,74 |
| **WPPZK** | 0,5208 |
| **WPPZK\_DOWN** | 0,88 |
| **WPPZK\_UP** | 0,17 |
| **WPPZK\_TRAIN** | 0,62 |

* Minimalne błędy dla poszczególnych rodzajów sieci:

|  |  |  |  |
| --- | --- | --- | --- |
| **Struktura sieci** | **Minimalny błąd** | **Max z WPPZK** | **Max z WPPZK\_TRAIN** |
| 30~30~1 | 764,73779 | 0,56548 | 0,62134 |
| 10~10~1 | 823,77008 | 0,56548 | 0,62229 |
| 10~1 | 897,89976 | 0,53968 | 0,59537 |
| 3~1 | 902,29561 | 0,55258 | 0,58617 |
| 30~1 | 922,78199 | 0,57143 | 0,59375 |
| 3~3~1 | 932,67304 | 0,57594 | 0,57670 |
| **Suma końcowa** | **764,73779** | **0,45437** | **0,49764** |
| **Wniosek: sieci o strukturze 30-30-1 i 10-10-1 dają obiecujące rezultaty – minimalne błędy.** | | | |

* Średnie błędy dla poszczególnych rodzajów sieci:

|  |  |  |  |
| --- | --- | --- | --- |
| **Struktura sieci** | **Średni błąd sieci** | **Średnia z WPPZK** | **Średnia z WPPZK\_TRAIN** |
| 10~1 | 1143,3098 | 0,50734692 | 0,563219058 |
| 10~10~1 | 1729,8185 | 0,51772482 | 0,566402533 |
| 3~1 | 1286,713408 | 0,51884915 | 0,556164919 |
| 3~3~1 | 2721,4134 | 0,52755012 | 0,536529018 |
| 30~1 | 1300,964698 | 0,50736331 | 0,557597180 |
| 30~30~1 | 1576,942359 | 0,51648113 | 0,567694557 |

* Maksymalne błędy dla poszczególnych rodzajów sieci

|  |  |  |  |
| --- | --- | --- | --- |
| **Struktura sieci** | **Maksymalny błąd** | **Maksimum z WPPZK** | **Maksimum z WPPZK\_TRAIN** |
| 10~1 | 2953,57113 | 0,53968 | 0,59537 |
| 3~1 | 3861,62135 | 0,55258 | 0,58617 |
| 30~1 | 3898,94912 | 0,57143 | 0,59375 |
| 30~30~1 | 4805,90489 | 0,56548 | 0,62134 |
| 10~10~1 | 5533,24181 | 0,56548 | 0,62229 |
| 3~3~1 | 9985,73479 | 0,57594 | 0,57670 |
| **Suma końcowa** | **9985,73479** | **0,57594** | **0,62229** |
| **Wnioski: Sieci z dwoma warstwami ukrytymi mają bardzo duży maksymalny błąd, który mocno zawyża błąd średni** | | | |

* Wpływ współczynnika bezwładności na proces uczenia sieci (wpływ na błąd)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Średni błąd** | **InertiaCoefficient** |  |  |  |
| **Struktura sieci** | **0** | **0,25** | **0,5** | **0,75** |
| 10~1 | 946,3 | 947,6 | 1063,3 | 1616,0 |
| 10~10~1 | 902,5 | 1045,3 | 1920,0 | 3051,4 |
| 3~1 | 984,6 | 1006,2 | 1167,0 | 1989,0 |
| 3~3~1 | 1519,0 | 2269,2 | 3023,8 | 4073,6 |
| 30~1 | 950,4 | 968,1 | 1129,5 | 2155,9 |
| 30~30~1 | 881,7 | 981,5 | 1614,1 | 2830,5 |
| **Suma końcowa** | **1030,8** | **1203,0** | **1653,0** | **2619,4** |
| **Wniosek: najlepsze wyniki dla współczynnika bezwładności bliskiego 0** | | | | |

* Wpływ współczynnika uczenia na proces uczenia sieci (wpływ na błąd)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Średni błąd** | **LearningCoefficient** |  |  |  |
| **Struktura sieci** | **0,25** | **0,5** | **0,75** | **1** |
| 10~1 | 1242,4 | 1104,0 | 1115,6 | 1111,2 |
| 10~10~1 | 1978,8 | 1787,8 | 1584,9 | 1567,6 |
| 3~1 | 1442,3 | 1221,6 | 1226,4 | 1256,6 |
| 3~3~1 | 3578,8 | 2719,6 | 2355,3 | 2231,8 |
| 30~1 | 1415,4 | 1211,2 | 1250,3 | 1326,9 |
| 30~30~1 | 1885,8 | 1475,9 | 1435,5 | 1510,6 |
| **Suma końcowa** | **1923,9** | **1586,7** | **1494,7** | **1500,8** |
| **Wniosek: Warte zauważenia jest, że dla sieci z jedną warstwą ukrytą wartość współczynnika najlepsza jest między 0,25-0,5, a dla sieci z dwiema sieciami ukrytymi – między 0,75 a 1** | | | | |

* Porównanie błędu dla różnych długości okna

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Długość okna** | **Średni błąd** | **Minimalny błąd** | **Maksymalny błąd** | |
| Day | 1506,7 | 823,8 | | 8662,2 |
| Week | 1746,3 | 764,7 | | 9985,7 |
| **Wniosek: brak wpływu okna czasowego na błąd (wartości porównywalne)** | | | | |

Tabela 20 najlepszych wartości błędu

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Network**  **Type** | **Learning**  **Coefficient** | | **Inertia**  **Coefficient** | | **Neuron**  **Structure** | **Window**  **Length** | **Result**  **Error** | **Test**  **CorrectDirection**  **PredictionsRate** | **LastTraining**  **Correct**  **Direction**  **PredictionsRate** |
| Elman | | 1 | | 0 | 30~30~1 | Week | 764,7377938 | 0,5208333 | 0,616003788 |
| Elman | | 0,75 | | 0 | 30~30~1 | Week | 804,0105832 | 0,5039683 | 0,618371212 |
| Elman | | 0,75 | | 0,25 | 10~10~1 | Day | 823,7700782 | 0,525641 | 0,617563739 |
| Elman | | 1 | | 0,25 | 10~10~1 | Day | 828,1733505 | 0,535503 | 0,622285175 |
| Elman | | 0,5 | | 0,25 | 10~10~1 | Day | 833,8008934 | 0,5295858 | 0,609065156 |
| Elman | | 0,75 | | 0,25 | 10~10~1 | Week | 834,9287238 | 0,5277778 | 0,603693182 |
| Elman | | 0,25 | | 0,25 | 10~10~1 | Day | 838,7697123 | 0,5128205 | 0,601038716 |
| Jordan | | 1 | | 0 | 30~30~1 | Week | 838,9440853 | 0,5277778 | 0,607481061 |
| MLP | | 1 | | 0 | 30~30~1 | Week | 841,8994393 | 0,5357143 | 0,60842803 |
| Elman | | 0,5 | | 0,25 | 10~10~1 | Week | 849,594158 | 0,5168651 | 0,610795455 |
| Jordan | | 1 | | 0,25 | 30~30~1 | Day | 851,5002856 | 0,5197239 | 0,607176582 |
| Elman | | 1 | | 0,25 | 10~10~1 | Week | 852,3522926 | 0,5128968 | 0,603219697 |
| Elman | | 0,25 | | 0,25 | 30~30~1 | Week | 853,5560657 | 0,5277778 | 0,589488636 |
| Elman | | 0,5 | | 0 | 30~30~1 | Week | 854,6482007 | 0,5128968 | 0,599905303 |
| Elman | | 0,5 | | 0,25 | 30~30~1 | Week | 863,1907744 | 0,5178571 | 0,599431818 |
| Elman | | 1 | | 0,25 | 30~30~1 | Day | 864,6316491 | 0,5394477 | 0,603399433 |
| MLP | | 0,75 | | 0 | 30~30~1 | Week | 865,6286086 | 0,5625 | 0,60842803 |
| Elman | | 1 | | 0 | 10~10~1 | Week | 870,7196273 | 0,5367063 | 0,607954545 |
| Jordan | | 0,5 | | 0 | 30~30~1 | Week | 872,0351261 | 0,5436508 | 0,598011364 |
| MLP | | 1 | | 0 | 30~30~1 | Day | 872,5298053 | 0,5493097 | 0,60434372 |
| Wnioski:   * Najlepsze wyniki dają struktury o dwóch warstwach ukrytych * Wpływ długości okna jest niewielki (mało znaczący) * Inertia na poziomie 0 – 0,25 daje najlepsze wyniki * Współczynnik uczenia – bliski 1 | | | | | | | | | |

* Tabela najlepszych 20 wyników uczenia na zbiorze treningowym

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Network**  **Type** | **Learning**  **Coefficient** | **Inertia**  **Coefficient** | **Neuron**  **Structure** | **Window**  **Length** | **Result**  **Error** | **TestCorrect**  **Direction**  **PredictionsRate** | **LastTraining**  **CorrectDirection**  **PredictionsRate** |
| Elman | 1 | 0,25 | 10~10~1 | Day | 828,17335 | 0,535503 | 0,622285175 |
| MLP | 1 | 0,25 | 30~30~1 | Day | 897,60798 | 0,4901381 | 0,621340888 |
| Elman | 0,75 | 0 | 30~30~1 | Week | 804,01058 | 0,5039683 | 0,618371212 |
| Elman | 0,75 | 0,25 | 10~10~1 | Day | 823,77007 | 0,525641 | 0,617563739 |
| Elman | 1 | 0 | 30~30~1 | Week | 764,73779 | 0,5208333 | 0,616003788 |
| Elman | 0,5 | 0,25 | 10~10~1 | Week | 849,59415 | 0,5168651 | 0,610795455 |
| Jordan | 0,75 | 0 | 30~30~1 | Week | 873,68290 | 0,5357143 | 0,61032197 |
| Jordan | 1 | 0 | 30~30~1 | Day | 888,87859 | 0,5453649 | 0,609065156 |
| Elman | 0,5 | 0,25 | 10~10~1 | Day | 833,80089 | 0,5295858 | 0,609065156 |
| MLP | 0,75 | 0 | 30~30~1 | Day | 890,73265 | 0,5177515 | 0,609065156 |
| MLP | 0,75 | 0 | 30~30~1 | Week | 865,62860 | 0,5625 | 0,60842803 |
| MLP | 1 | 0 | 30~30~1 | Week | 841,89943 | 0,5357143 | 0,60842803 |
| Elman | 1 | 0 | 10~10~1 | Week | 870,71962 | 0,5367063 | 0,607954545 |
| MLP | 1 | 0 | 10~10~1 | Day | 905,04910 | 0,5522682 | 0,607648725 |
| Jordan | 1 | 0 | 30~30~1 | Week | 838,94408 | 0,5277778 | 0,607481061 |
| MLP | 0,5 | 0 | 10~10~1 | Week | 902,12748 | 0,5238095 | 0,607481061 |
| Jordan | 1 | 0,25 | 30~30~1 | Day | 851,50028 | 0,5197239 | 0,607176582 |
| Jordan | 0,75 | 0 | 10~10~1 | Week | 894,06215 | 0,5367063 | 0,607007576 |
| Elman | 0,75 | 0 | 10~10~1 | Week | 875,02288 | 0,5228175 | 0,606534091 |
| Elman | 1 | 0,25 | 30~30~1 | Week | 888,54774 | 0,5238095 | 0,604640152 |
| Wniosek: Rodzaj sieci nie ma dużego wpływu na naukę sieci, kolejne potwierdzenie wyższości sieci z dwiema warstwami ukrytymi.  **Przykład wykresu dla sieci Elmana, struktura 30-30-1****, LC = 0,8 IC=0,25** | | | | | | | |

## Test dla danych wielowymiarowych – sprawdzenie PCA

### Dane wejściowe

Testowano dla następujących kombinacji parametrów:

* Rodzaj sieci: MLP, Elman, Jordan
* Architektura sieci:

|  |  |  |  |
| --- | --- | --- | --- |
| 10-1 | 30-1 |  |  |
| 10-10-1 | 30-30-1 |  |  |

* Współczynnik uczenia: 0,5; 0,75; 1.0
* Współczynnik bezwładności: 0; 0,25
* Ilość iteracji: 1000
* Liczba powtórzeń: 3
* Brak PCA, PCA=10,20
* Długość okna: Dzień

### Wyniki i wnioski

* Najmniejszy błąd otrzymano dla następujących danych wejściowych:
  + **Sieć** MLP
  + **Współczynnik uczenia**: 1.0
  + **Bezwładność**: 0.0
  + **Struktura sieci**: 30-30-1
  + **PCA:** 20

|  |  |
| --- | --- |
| **BŁĄD** | 823,25 |
| **WPPZK** | 0,54 |
| **WPPZK\_DOWN** | 0,80 |
| **WPPZK\_UP** | 0,29 |
| **WPPZK\_TRAIN** | 0,60 |

* Porównanie błędu w zależności od PCA

|  |  |  |
| --- | --- | --- |
| **PCA** | **Średni błąd** | **Minimalny błąd** |
| 20 | 1007,811798 | **823,254576** |
| no PCA (31 kolumn) | **927,820205** | 845,2692078 |
| 10 | 1109,478528 | 904,8500989 |
| **Suma końcowa** | **1015,036844** | **823,254576** |
| **Wnioski: brak PCA średnio daje lepsze rezultaty, ale minimalny błąd jest dla PCA = 20** | | |

* Wpływ PCA i Struktury sieci na błąd sieci

|  |  |  |
| --- | --- | --- |
| **PCA** | **Średni błąd** | **Minimalny błąd** |
| **20** | **1007,811798** | **823,254576** |
| 10~1 | 1018,659983 | 929,6886791 |
| 10~10~1 | 1000,089286 | 826,6177631 |
| 30~1 | 1029,141047 | 951,3805214 |
| 30~30~1 | 983,3568765 | 823,254576 |
| **no PCA** | **927,820205** | **845,2692078** |
| 10~1 | 935,0889366 | 907,6026364 |
| 10~10~1 | 928,7377791 | 845,2692078 |
| 30~1 | 947,6100609 | 922,7315533 |
| 30~30~1 | 899,8440434 | 853,9089185 |
| **10** | **1109,478528** | **904,8500989** |
| 10~1 | 1147,548822 | 995,7456998 |
| 10~10~1 | 1032,690295 | 904,8500989 |
| 30~1 | 1183,23612 | 1020,108452 |
| 30~30~1 | 1074,438876 | 998,2048915 |
| **Suma końcowa** | **1015,036844** | **823,254576** |
| **Wniosek:**   * **Lepsze wartości błędu zarówno średniego jak i minimalnego dla sieci z dwoma warstwami ukrytymi,** * **PCA na poziomie 10 daje najgorsze rezultaty** * **Brak PCA daje lepszy średni błąd, ale minimalny błąd minimalnie lepszy przy zmniejszeniu liczby kolumn** | | |

* PCA a maksymalna predykcja kierunku na zbiorze testowym i treningowym

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PCA** | **Maksymalne WPPZK** | **Maksymalne WPPZK\_TRAIN** | **Średni WPPZK** | **Średni WPPZK\_TRAIN** |
| 10 | 0,55 | 0,57 | 0,51 | 0,55 |
| 20 | 0,57 | 0,63 | 0,52 | 0,58 |
| no PCA | 0,57 | 0,61 | 0,51 | 0,59 |
| **Suma końcowa** | **0,57** | **0,63** | **0,51** | **0,57** |
| **Wniosek:**   * **Brak PCA i PCA = 20 daje najlepsze maksymalne i średnie wartości predykcji kierunku** * **PCA=20 daje minimalne lepsze wyniki (maksymalne)** | | | | |