

Politechnika Świętokrzyska

Wydział Elektrotechniki, Automatyki i Informatyki

Oprogramowanie dydaktyczne do testowania działania wybranych metod dla modeli szeregów czasowych na różnych zbiorach danych

Dyplomant:

Mateusz Godlewski
Informatyka II stopień

Promotor:

dr inż. Andrzej Kułakowski

Kielce 2021

Cel pracy

Celem pracy było zaprojektowanie oraz stworzenie oprogramowania dydaktycznego, umożliwiającego testowanie działania metod statystycznych oraz eksploracyjnych na wybranych szeregach czasowych.

Plan prezentacji

- Pojęcie szeregu czasowego
- Metody statystyczne oraz eksploracyjne
- Prezentacja oprogramowania dydaktycznego
- Możliwości dalszego rozwoju
- Podsumowanie

Pojęcie szeregu czasowego

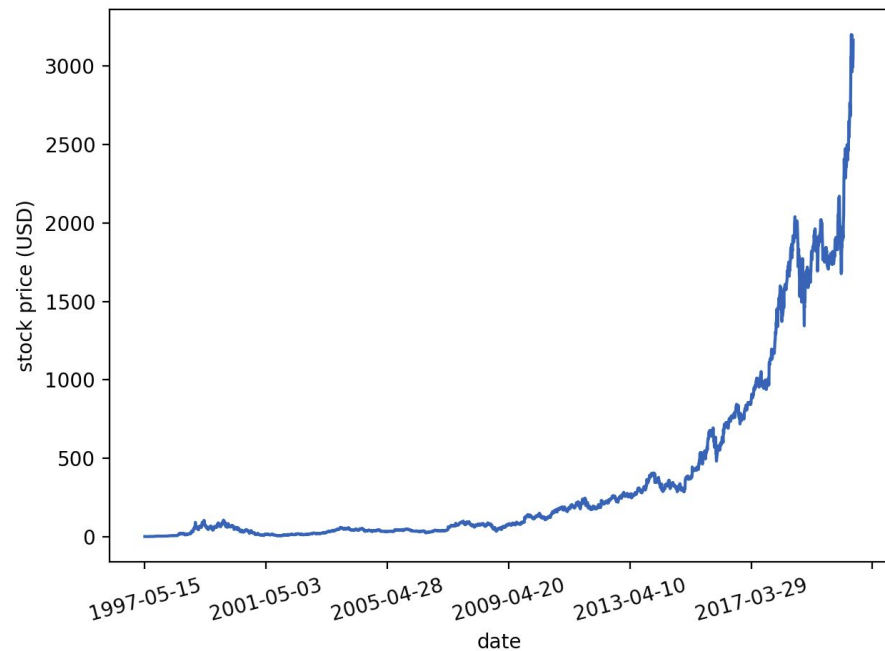
Pojęcie szeregu czasowego

Definicja

Ciąg uporządkowanych obserwacji dokonanych w różnych momentach na zmiennej charakteryzującej pewną jednostkę lub zbiorowość.

Pojęcie szeregu czasowego

Przykład



Pojęcie szeregu czasowego

Właściwości

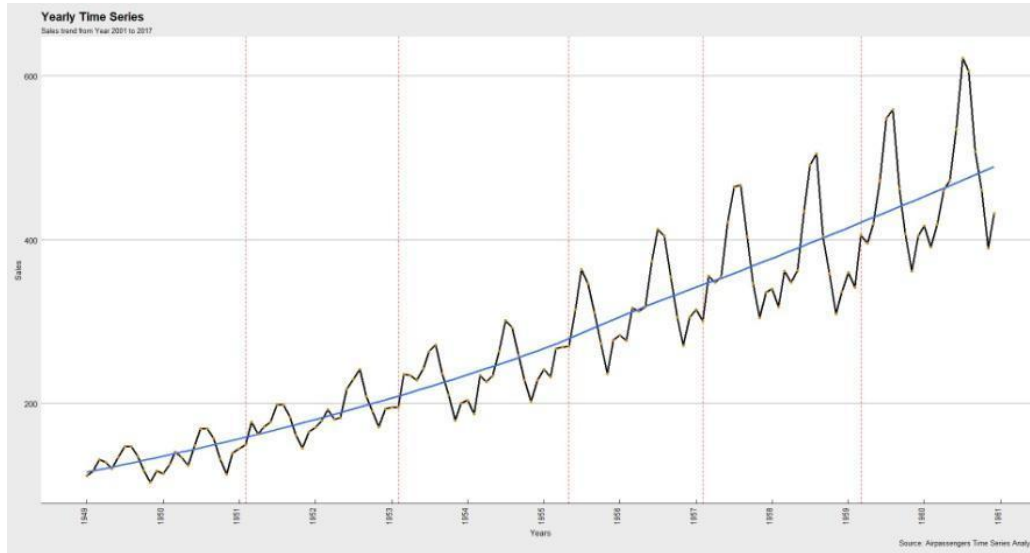
Właściwości szeregów czasowych:

- tendencję rozwojową;
- wahania okresowe;
- wahania cyklu koniunkturalnego;
- wahania przypadkowe.

Pojęcie szeregu czasowego

Właściwości

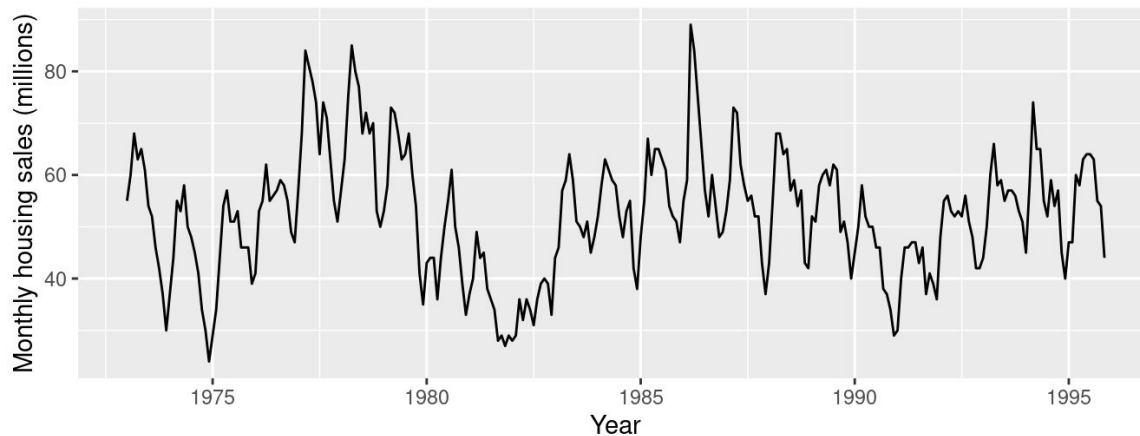
Tendencja rozwojowa



Pojęcie szeregu czasowego

Właściwości

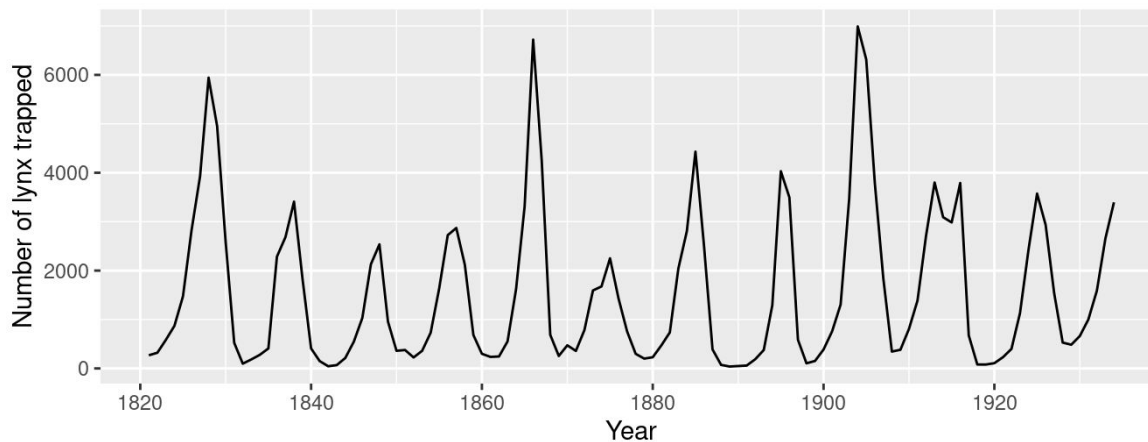
Wahania okresowe



Pojęcie szeregu czasowego

Właściwości

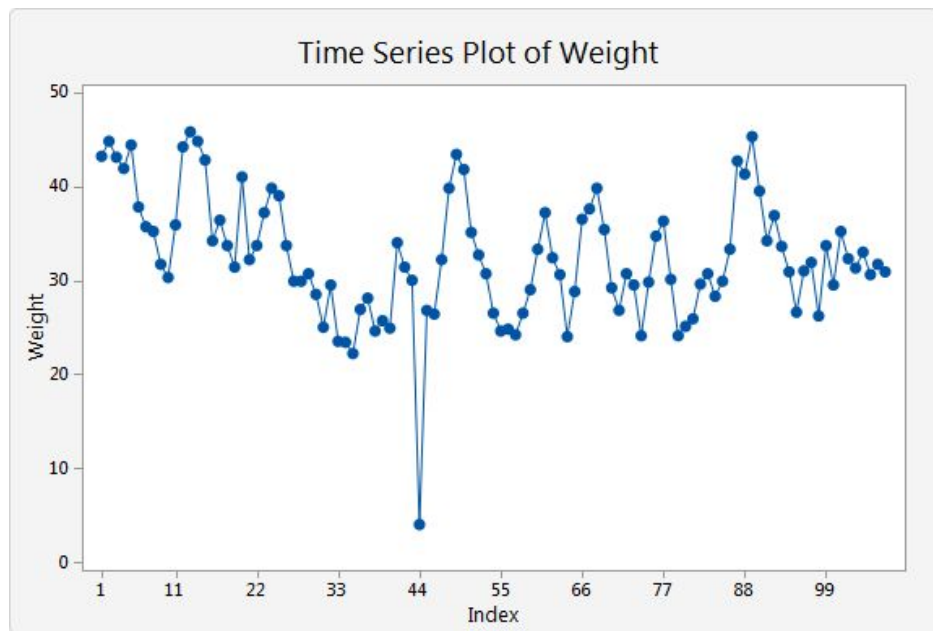
Wahania cyklu koniunkturalnego



Pojęcie szeregu czasowego

Właściwości

Wahania przypadkowe



Metody statystyczne i eksploracyjne

Metody statystyczne i eksploracyjne

Metody statystyczne

- Wyznaczanie wartości minimalnych, maksymalnych oraz średniej arytmetycznej;
- Wyznaczenie odchylenia standardowego;
- Wyznaczenie mediany;
- Wyznaczanie wartości poszczególnego kwantyla;
- Wyznaczanie wartości interkwartyla.

Metody statystyczne i eksploracyjne

Metody eksploracyjne

Wyznaczenie przyszłych wartości szeregów czasowych za pomocą metod predykcyjnych takich jak:

- Model autoregresyjny;
- Średnia ruchoma;
- Zintegrowany model autoregresyjny ze średnią ruchomą.

Metody statystyczne i eksploracyjne

Metody eksploracyjne - model autoregresyjny (AR)

$$X_t = C + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t$$

gdzie:

- o t - liczba obserwacji;
- o X_t - wartość szeregu czasowego dla obserwacji t ;
- o $\varphi_1, \dots, \varphi_p$ - parametry modelu;
- o C - stała (dla uproszczenia formuły, często pomijana);
- o ε_t - szum biały.

Metody statystyczne i eksploracyjne

Metody eksploracyjne - średnia ruchoma (MA)

$$MA = \frac{p_0 + p_1 + \dots + p_{n-1}}{n}$$

gdzie:

- o n – liczba obserwacji
- o p_n - poszczególna obserwacja

Metody statystyczne i eksploracyjne

Metody eksploracyjne - metoda ARIMA

$$Y_t = C + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t$$

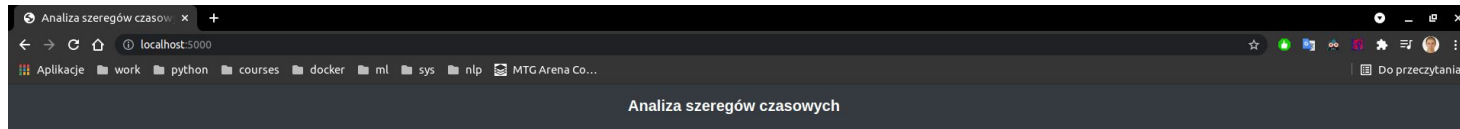
gdzie:

- o Y_t = zmienna objaśniana w dziedzinie czasu;
- o C = stała (dla uproszczenia formuły, często pomijana);
- o ϕ = współczynnik każdego parametru p ;
- o θ = współczynnik każdego parametru q ;
- o e_t = szum biały.

Prezentacja oprogramowania dydaktycznego

Prezentacja oprogramowania dydaktycznego

Strona główna



Prześlij plik z szeregiem czasowym

Wybierz plik

Nie wybrano pliku

Prześlij

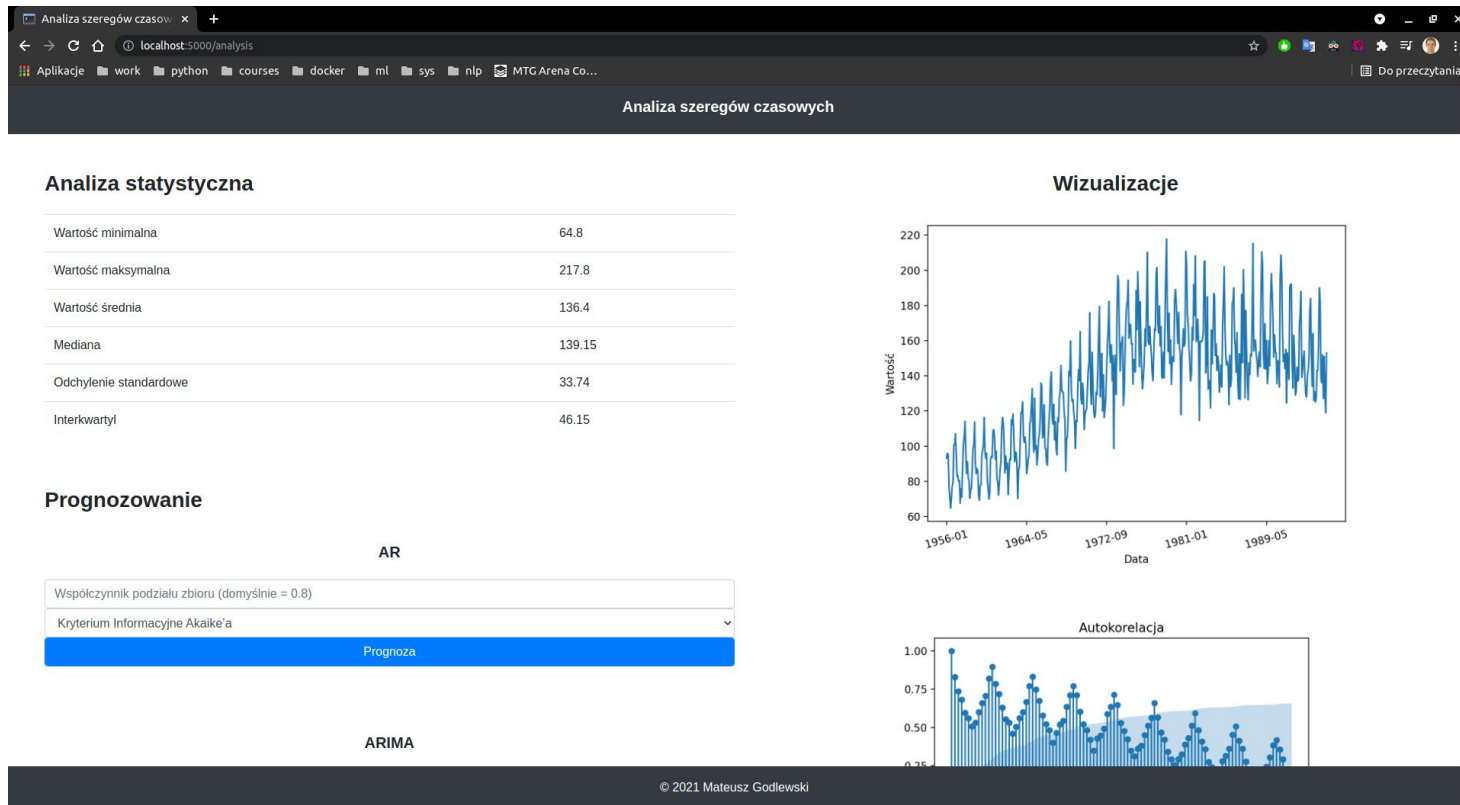
Przesyłany plik powinien posiadać rozszerzenie CSV. Plik powinien posiadać dwie kolumny:

- data - znaczniki czasowe.
- value - wartości.

Kolejność tych kolumn nie ma znaczenia.
Kolumny te powinny być rozdzielone przecinkiem.

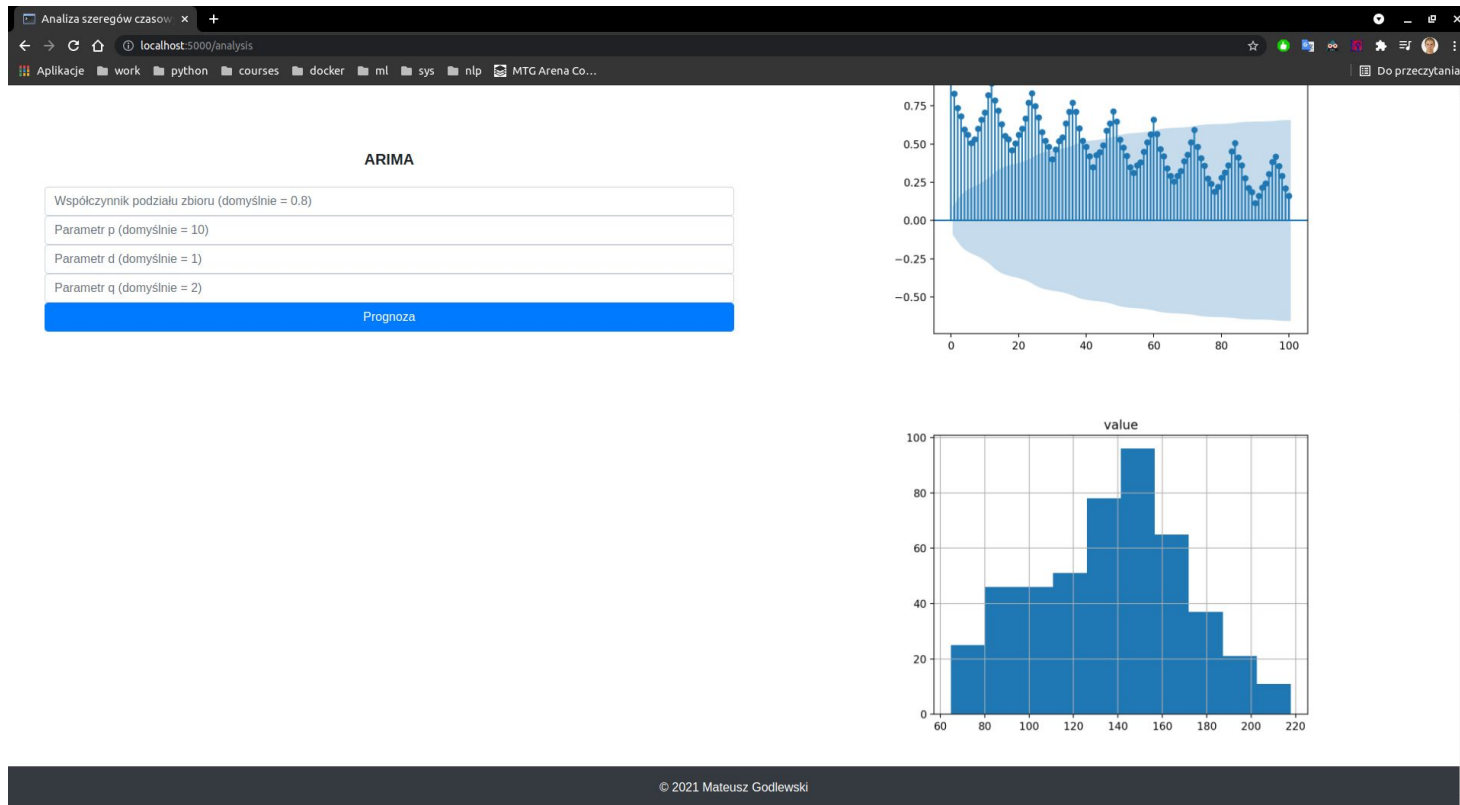
Prezentacja oprogramowania dydaktycznego

Widok analizy statystycznej



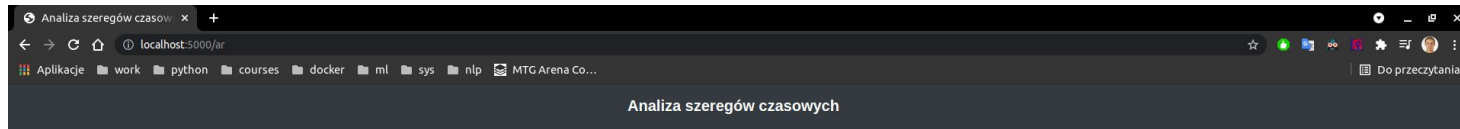
Prezentacja oprogramowania dydaktycznego

Widok analizy statystycznej cd.

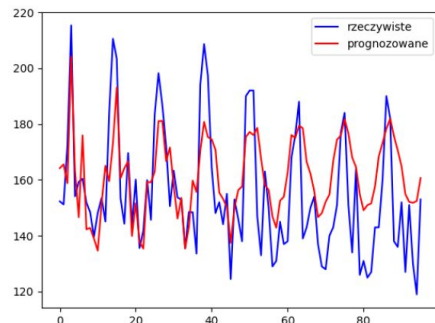


Prezentacja oprogramowania dydaktycznego

Widok analizy eksploracyjnej - prognoza metodą AR



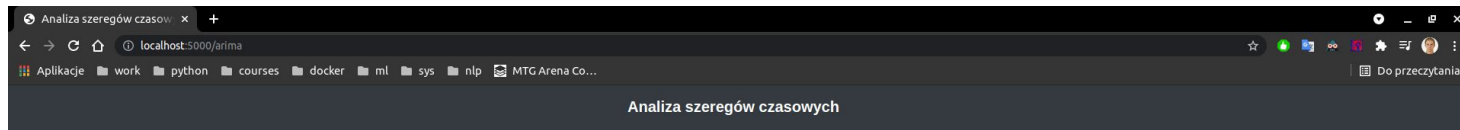
Wyniki prognozy modelu AR



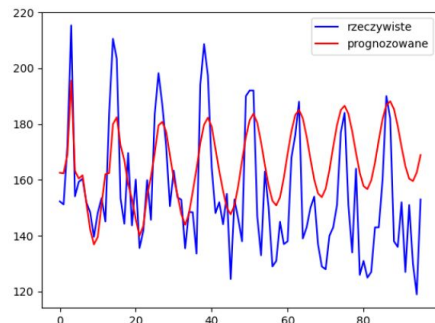
Kryterium doboru rzędu predykcji	Kryterium Informacyjne Akaike'a
Rząd predykcji	16
Ilość trenowanych obserwacji	380
Wartość RMSE - błędu średniokwadratowego	17.568628622993064

Prezentacja oprogramowania dydaktycznego

Widok analizy eksploracyjnej - prognoza metodą ARIMA



Wyniki prognozy metodą ARIMA



Ilość trenowanych obserwacji	379
Parametr p	10
Parametr d	1
Parametr q	2
Błąd prognozy modelu według kryterium AIC	2914.1215425536197
Błąd prognozy modelu według kryterium BIC	2969.2470494247736
Błąd prognozy modelu według kryterium HQIC	2935.997782030747

Możliwości dalszego rozwoju

Możliwości dalszego rozwoju

- Wykrywanie właściwości szeregów czasowych
- Stworzenie interaktywnych wykresów
- Poszerzenie funkcjonalności o większą liczbę metod predykcyjnych

Podsumowanie

Oprogramowanie dydaktyczne do testowania działania wybranych metod dla modeli szeregów czasowych na różnych zbiorach danych

- Prosty oraz intuicyjny interfejs użytkownika
- Możliwość wykorzystania dowolnego szeregu czasowego
- Analiza statystyczna szeregu czasowego
- Prognoza przyszłych wartości szeregu

Dyplomant:

Mateusz Godlewski

Informatyka II stopień

Promotor:

dr inż. Andrzej Kułakowski

Literatura

- https://mfiles.pl/pl/index.php/Szereg_czasowy
- https://en.wikipedia.org/wiki/Autoregressive_model
- https://pl.wikipedia.org/wiki/%C5%9Arednia_ruchoma
- https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average