**Mateusz Godlewski  
Numer albumu: 85163  
  
  
  
  
  
Oprogramowanie dydaktyczne do testowania działania wybranych metod dla modeli szeregów czasowych na różnych zbiorach danych.  
  
Praca magisterska  
na studiach II-go stopnia  
na kierunku Informatyka**

Opiekun pracy dyplomowej:  
dr inż. Andrzej Kułakowski  
Katedra Systemów Informatycznych

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

**Streszczenie**

Celem mojej pracy dyplomowej było zaprojektowanie oraz stworzenie oprogramowania dydaktycznego, umożliwiającego testowanie działania wybranych metod predykcji szeregów czasowych. Oprogramowanie to umożliwia wykonanie analizy statystycznej oraz eksploracyjnej na dostarczonym przez użytkownika szeregu czasowym.

W ninejszej pracy zawarty jest opis wybranego oprogramowania służącego do testowania działania metod matematycznych, charakterystyka zagadnień których dotyczy oprogramowanie oraz charakterystyka tworzonego oprogramowania.

Oprogramowanie zostało dokładnie udokumentowane w postaci instrukcji obsługi użytkownika. Natomiast projekt został zilustrowany w postaci wykresów i diagramów, a jego implementacja została przedstawiona poprzez fragmenty kodów źródłowych i zrzuty ekranu.

Słowa kluczowe: Szereg czasowy, statystyka, autoregresja, średnia ruchoma, Python, Flask.

**Didactic software for testing selected methods for time series models on various data sets**

**Summary**

Purpose of my thesis was to design and create didactic software that would enable user to test selected time series prediction methods. This software enables performing statistical and exploratory analysis on the time series given by user.

Present work contains description of selected software that is using for testing mathematical operations, the characteristic of issues related to the software and characteristic of developed program itself.

The software was documented in the form of user manual. Project of software has been detailed illustrated in the form of charts and diagrams, and its implementation has been presented through source code fragments and screenshots.

Keywords: Time series, statistics, autoregression, moving average, Python, Flask.

**SPIS TREŚCI**

1. WSTĘP11

2. CHARAKTERYSTYKA WYBRANEGO OPROGRAMOWANIA DYDAKTYCZNEGO DO TESTOWANIA DZIAŁANIA METOD MATEMATYCZNYCH13

2.1 Oprogramowania matematyczne13

2.2 Opis programu MATLAB00

3. OPRACOWANIE WYBRANYCH ZAGADNIEŃ DO KONSTRUOWANIA OPROGRAMOWANIA TESTUJĄCEGO00

3.1 Test00

3.2 Test00

3.2.1 Test00

4. CHARAKTERYSTYKA WYBRANEGO DO REALIZACJI ZADANIA OPROGRAMOWANIA00

4.1 Test00

4.2 Test00

4.2.1 Test00

5. PROJEKT OPROGRAMOWANIA DYDAKTYCZNEGO DO TESTOWANIA DZIAŁANIA WYBRANYCH METOD DLA MODELI SZEREGÓW CZASOWYCH00

6. IMPLEMENTACJA OPROGRAMOWANIA DYDAKTYCZNEGO00

6.1 Test00

7. TESTOWANIE DZIAŁANIA WYKONANEGO OPROGRAMOWANIA00

7.1 Test00

8. PODSUMOWANIE00

LITERATURA00

Spis rysunków00

Spis listingów00

**Wstęp**

Dynamiczny rozwój szeroko pojętej sztucznej inteligencji w ostatnim dziesięcioleciu sprawił, iż znalazła swoje zastosowanie w prawie każdej sferze życia codziennego. Uczenie maszynowe, sztuczne sieci neuronowe, czy analiza danych, wykorzystywane są w najbardziej popularnych serwisach internetowych czy aplikacjach mobilnych. Rozpoznawanie obrazów, rozpoznawanie mowy, eksploracja danych czy prognozowanie przyszłości, to jedynie kilka z wielu funkcjonalności dostarczanych przez algorytmy sztucznej inteligencji.

Celem mojej pracy dyplomowej było stworzenie oprogramowania w formie aplikacji internetowej, umożliwiającego testowanie wybranych metod prognozy szeregów czasowych, na różnych zbiorach danych. Użytkownik ma możliwość wykonania analizy statystycznej, na dostarczonym przez niego zbiorze danych oraz wizualizacji zawartych w nim danych w postaci wykresów. Główną funkcjonalnością stworzonego oprogramowania, jest możliwość testowania metod predykcji przyszłych wartości szeregów czasowych, za pomocą Autoregresji lub Autoregresji ze zintegrowaną średnią ruchomą.

Dokument ten składa się z ośmiu rozdziałów przedstawiających najważniejsze aspekty pracy. Na wstępie przedstawiony został cel pracy oraz krótki opis poszczególnych rozdziałów. W rozdziale drugim pt. "Charakterystyka wybranego oprogramowania dydaktycznego do testowania działania metod matematycznych." przedstawiona została charakterystyka najpopularniejszego oprogramowania które jest wykorzystywane przy testowaniu metod matematycznych. Następny rozdział zatytułowany "Opracowanie wybranych zagadnień do konstruowania oprogramowania testującego." zawiera zagadnienia oraz definicje pojęć, które najczęściej występują w stworzonym oprogramowaniu. Rozdział numer cztery "Charakterystyka wybranego do realizacji zadania oprogramowania." Zostały zaprezentowane narzędzia oraz technologie wykorzystane do implementacji systemu. W kolejnym rozdziale pt. „Projekt oprogramowania dydaktycznego do testowania działania wybranych metod dla modeli szeregów czasowych.” opisany jest projekt architektury stworzonego oprogramowania w postaci wykresów i diagramów. Następnie w rozdziale szóstym „Implementacja oprogramowania dydaktycznego.” znajduje się przedstawienie zaimplementowanych funkcjonalności systemu, wraz z kodami źródłowymi oraz zrzutami ekranu stworzonego rozwiązania. Rozdział siódmy "Testowanie działania wykonanego oprogramowania" poświęcony został opisowi testów stworzonego systemu. Ostatni, ósmy rozdział zawiera podsumowanie pracy.

**1. CHARAKTERYSTYKA WYBRANEGO OPROGRAMOWANIA DYDAKTYCZNEGO DO TESTOWANIA DZIAŁANIA METOD MATEMATYCZNYCH**

**1.1 Oprogramowania matematyczne**

Wsród najpopularniejszych rozwiązań, dostępnych na rynku są:

* MATLAB
* Magnolia
* Maple
* Scilab

**1.2 Opis programu MATLAB**

…

**2. OPRACOWANIE WYBRANYCH ZAGADNIEŃ DO KONSTRUOWANIA OPROGRAMOWANIA TESTUJĄCEGO**

**2.1 Zagadnienia matematyczne**

W stworzonym oprogramowaniu pojawiają się pojęcia z dziedzin matematyki oraz statystyki. Aby w pełni zrozumieć logikę aplikacji, oraz operacje, które są wykonywane przez stworzone oprogramowanie należy zapoznać się z definicjami taki zagadnień jak:

* Szereg czasowy
* Analiza statystyczna
* Autoregresja
* Średnia ruchoma

**Szereg czasowy** to seria punktów danych zindeksowanych w porządku czasowym. Najczęściej jest on sekwencją wykonywaną w kolejnych, równych odstępach czasu. Zatem jest to sekwencja danych w czasie dyskretnym. Przykładami szeregów czasowych są wartości akcji na dziełdzie, temperatura powietrza, monitorowanie ciśnienia krwi itd. Szeregi czasowe są bardzo często przedstawiane za pomocą wykresów przebiegu (czasowego wykresu liniowego). Szeregi czasowe są najczęściej wykorzystywane w statystyce, przetwarzaniu sygnałów, rozpoznawaniu wzorców, ekonometrii, finansach matematycznych, prognozowaniu pogody, przewidywaniu trzęsień ziemi, inżynierii sterowania, astronomii i w każdej dziedzinie nauk stosowanych i inżynierii, obejmującej pomiary czasowe.

****  
*Rysunek 2.X Wartości akcji firmy Amazon w latach 1997 – 2020*

Szeregi czasowe mogą (lecz nie muszą) posiadać własności takie jak:

* **Tendencja rozwojowa (trend)** - czyli długookresowe, systematyczne i jednokierunkowe zmiany wielkości zjawiska. Są wywołane przez stałe przyczyny działające w całym okresie.
* **Wahania okresowe (wahania sezonowe)** - są to rytmiczne, powtarzające się zmiany poziomu zjawiska, które są wywoływane przez czynniki przyrodnicze (zmiana pór roku). Długość cyklu wahań sezonowych, to liczba okresów składających się na pełen cykl wahań okresowych
* **Wahania cyklu koniunkturalnego (wahania cykliczne)** - to oscylacje, które wymagają wieloletnich obserwacji, posiadające nieregularny charakter i różną długość cyklu. Wyróżnia się okresy recesji, depresji, ożywienia i dobrobyt. Brakuje ugruntowanej statystycznej metody ich analizy oraz ich przebieg jest trudny do przewidzenia. Są to to systemowe, falowe wahania rozwoju gospodarki obserwowane w okresach dłuższych od roku i ich analiza wymaga wieloletnich obserwacji.
* **Wahania przypadkowe** - czyli nieregularne odchylenia wielkości zjawiska od poziomu, którego spodziewalibyśmy się na podstawie oddziaływania innych czynników (trendu, wahań sezonowych i cyklicznych). Ich natura jest trudna do zdefiniowania. Mogą być wywoływane przez przypadkowe różnokierunkowe działania czynników, które nie są uwzględnione w analizie.