

P O L I T E C H N I K A R Z E S Z O W S K A

im. Ignacego Łukasiewicza

WYDZIAŁ MATEMATYKI I FIZYKI STOSOWANEJ

**Aldona Świrad**

Predykcja szeregów czasowych: metody klasyczne i elementy uczenia maszynowego

**Praca dyplomowa inżynierska**

Opiekun pracy:  
dr Ewa Rejwer-Kosińska

Rzeszów, 2024

Spis treści

[Wstęp 4](#_Toc182556548)

[Cel pracy 4](#_Toc182556549)

[Zakres pracy 4](#_Toc182556550)

[2. Wprowadzenie teoretyczne 5](#_Toc182556551)

[2.1 Definicja szeregów czasowych 5](#_Toc182556552)

[2.2 Model ARIMA 5](#_Toc182556553)

Wstęp

1. Wprowadzenie teoretyczne
   1. Definicja szeregów czasowych
   2. Model ARIMA
2. Analiza wybranych szeregów czasowych metodami klasycznymi
   1. Analiza – opis, wykresy, zakres zmienności
   2. Dobór modelu
3. Metody uczenia maszynowego dla szeregów czasowych

4.1 Model RNN

4.2 Model LSTM

1. Prognozowanie szeregów czasowych

5.1 Prognozowanie metodą ARIMA

5.2 Prognozowanie metodą LSTM

5.3 Porównanie

Podsumowanie

Bibliografia

Dodatek

# Wstęp

Tekst akapitu.

Tekst akapitu.

## Cel pracy

Celem niniejszej pracy inżynierskiej jest zbadanie i porównanie skuteczności metod analizy szeregów czasowych: metod klasycznych oraz metod opartych na uczeniu maszynowym, ze szczególnym uwzględnieniem modeli LSTM. Głównym celem pracy jest zastosowanie tych modeli do prognozowania szeregów czasowych oraz ocena ich przydatności w różnych scenariuszach analizy danych czasowych.

## Zakres pracy

W pierwszym rozdziale pracy podane zostaną najważniejsze pojęcia z zakresu modelowania i predykcji szeregów czasowych. Rozdział drugi poświęcony zostanie analizie własności i modelowaniu wybranych szeregów czasowych, z zastosowaniem metod klasycznych. W rozdziale trzecim podane zostaną podstawowe pojęcia i metody uczenia maszynowego (w tym modele RNN i LSTM), stosowane w analizie szeregów czasowych. Porównanie wybranych metod prognozowania szeregów czasowych pod kątem ich skuteczności, wraz z wnioskami dotyczącymi przydatności poszczególnych modeli w różnych scenariuszach analizy szeregów czasowych, będą zawarte w ostatnim rozdziale pracy. W podsumowaniu zawarte będą sugestie dotyczące dalszych badań oraz potencjalnych zastosowań praktycznych uzyskanych wyników.

Praca będzie opierać się głównie na analizie literatury naukowej oraz eksperymentach przeprowadzonych na rzeczywistych danych, co umożliwi kompleksowe zbadanie tematu oraz wyciągnięcie trafnych wniosków dotyczących analizy szeregów czasowych za pomocą metod klasycznych i uczenia maszynowego.

# 1. Wprowadzenie teoretyczne

Dodać: (trend, sezonowość, proces stochastyczny, biały szum, def. Parametry, szereg czasowy – parametry, wykresy i rysunki obrazujące szereg czasowy, poprawić wzory)

Wprowadzenie teoretyczne ma na celu przybliżenie podstawowych pojęć związanych z szeregami czasowymi oraz zaprezentowanie jednego z najczęściej wykorzystywanych modeli do ich analizy – modelu ARIMA. Zrozumienie tych pojęć oraz metodologii jest kluczowe do poprawnej analizy i prognozowania szeregów czasowych.

2.1 Definicja i własności szeregów czasowych

Szereg czasowy (ang. *time series*) to sekwencja danych gromadzonych w równych odstępach czasu, np. dziennie, miesięcznie, rocznie. Każda obserwacja w szeregu czasowym składa się z dwóch elementów: momentu w czasie, do którego się odnosi, oraz wartości zmiennej mierzonej w tym momencie. Przy czym, kolejność obserwacji ma istotne znaczenie. Przykładami szeregów czasowych mogą być kursy walut, wartości indeksów giełdowych, liczba sprzedanych produktów w danym okresie, czy zmiany temperatury w ciągu roku.

W praktyce, analiza szeregów czasowych ma na celu identyfikację struktury danych, wyodrębnienie trendów, sezonowości oraz innych wzorców, a także prognozowanie przyszłych wartości na podstawie dostępnych danych historycznych, czym będziemy się zajmować w tej pracy.

***Stacjonarność szeregu czasowego***

Szeregi czasowe mogą być stacjonarne lub niestacjonarne. Szereg stacjonarny charakteryzuje się stałą średnią, wariancją i autokorelacją w czasie, natomiast szereg niestacjonarny może zawierać zmieniające się średnie, wariancje lub inne nieliniowe zależności. Aby skutecznie analizować szeregi czasowe i budować modele prognostyczne, często stosuje się metody przekształcania niestacjonarnych szeregów na stacjonarne. **Proszę wymienić i bardziej szczegółowo omówić kilka z tych metod**

***Trend i sezonowość***

**Zamiast podrozdziałów w podrozdziale w taki sposób proszę odnosić się do innych charakterystyk szeregów czasowych, które będzie Pani opisywać (czyli biały szum, itp.)**

***Różnicowanie niestacjonarnego szeregu czasowego***

Różnicowanie służy do przekształcania niestacjonarnych szeregów czasowych   
w szeregi stacjonarne. Proces ten eliminuje trend i sezonowość, co ułatwia modelowanie fluktuacji w krótszych okresach.

Proces różnicowania polega na obliczeniu różnicy między kolejnymi obserwacjami:

**(Różnicowanie będzie Pani wykonywać podczas analizy szeregów czasowych, także być może tutaj będzie trzeba coś więcej dopisać)**

2.2 Model ARIMA

Model ARIMA (ang. *AutoRegressive Integrated Moving Average*) to jeden   
z najpopularniejszych modeli stosowanych w analizie szeregów czasowych. Jest to model liniowy, który łączy trzy kluczowe elementy: autoregresję (AR), różnicowanie (I) i średnią ruchomą (MA).

***Model autoregresyjny AR(p)***

Autoregresja polega na przewidywaniu wartości zmiennej w danym momencie na podstawie wcześniejszych obserwacji tej samej zmiennej. Model *AR(p)*, gdzie parametr *p* oznacza liczbę opóźnionych obserwacji (lagów), zakłada, że wartość zmiennej zależy od jej wcześniejszych wartości:

**(Proszę pisać równania w kreatorze równań, nie jako wstawione zdjęcia ! Ewentualne poprawki/zmiany będzie wtedy można wstawić bez problemu)**

gdzie:

- wartość zmiennej w czasie t,

– współczynniki modelu,

– błąd losowy w czasie t.

***Model średniej ruchomej MA(q)***

Średnia ruchoma zakłada, że bieżąca wartość zmiennej jest kombinacją poprzednich błędów losowych. Model *MA(q)*, gdzie *q* oznacza liczbę opóźnień błędów losowych, zakłada, że wartość zmiennej w danym momencie zależy od wcześniejszych błędów:

gdzie:

– błędy losowe,

– współczynniki modelu.

***Model ARIMA(p,d,q)***

Model *ARIMA* łączy w sobie: autoregresję (*p*), różnicowanie (*d*) oraz średnią ruchomą (*q*). Model ten opisuje szereg czasowy za pomocą kombinacji wcześniejszych wartości zmiennej oraz błędów losowych, przy jednoczesnym uwzględnieniu odpowiedniej liczby różnicowań, w celu zapewnienia stacjonarności. Ogólna postać modelu *ARIMA(p, d, q)* wygląda następująco:

gdzie:

– zróżnicowanie wartości zmiennej,

– współczynniki autoregresji,

– współczynniki średniej ruchomej,

– błąd losowy w czasie t.

Model ARIMA jest szeroko stosowany w prognozowaniu szeregów czasowych ze względu na swoją elastyczność oraz zdolność do modelowania różnych wzorców, takich jak trendy i sezonowość. Jednak, dobór odpowiednich wartości parametrów modelu wymaga analizy danych oraz testów diagnostycznych, takich jak autokorelacja reszt czy testy stacjonarności.

# 4. Metody uczenia maszynowego dla szeregów czasowych

## 4.1 Model RNN

Rekurencyjne sieci neuronowe (RNN, ang. Recurrent Neural Networks) to sieci neuronowe zaprojektowane w ten sposób, aby przetwarzać dane sekwencyjne. Ich charakterystycznym elementem są komórki pamięci. Odróżnia je to od tradycyjnych sieci neuronowych. RNN potrafią „pamiętać” wcześniejsze stany, co jest kluczowe w dalszym podejmowaniu decyzji, np.: jak w przetwarzaniu języka naturalnego, rozpoznawaniu mowy, analizie wideo, czy analizie szeregów czasowych.

Struktura RNN umożliwia rekurencję, co oznacza, że wynik jednej operacji jest wykorzystywany jak wejście dla kolejnej.

Każdy krok czasowy w RNN można opisać jako:

W tym wzorze:

– macierz zawierająca dane wyjściowe w takcie t dla poszczególnego elementu, gdzie m to ilość próbek, a n liczba neuronów

- macierz zawierająca dane wejściowe dla wszystkich przykładów, gdzie n to liczba wejść

– macierz wag dla wejść w danym t

- wag dla danych wyjść z poprzedniego taktu

– wektor o wymiarze będący obciążeniem (*ang. bias*) dla każdego neuronu

W takcie t=0 nie istnieją żadne dane wyjściowe, więc zakłada się, że ich wartość równa jest 0.

Neuron RNN zachowuje pamięć o wartości stanu w poszczególnych taktach w **komórkach pamięci**. Jeden neuron lub warstwa RNN zachowuje w pamięci tylko krótkie sekwencje. Funkcja (*ang. hidden*) określa dane wejściowe w aktualnym kroku czasowym oraz stan z poprzedniego, t.zn.:

Obraz zawierający tekst, diagram, Czcionka, szkic

Opis wygenerowany automatycznie

Neuron można wyobrazić sobie jako zapętlony sam w sobie, czyli działający rekurencyjnie. Można to też przedstawić w postaci rozwijającego się neuronu na kolejne neurony. Ta wizualizacja jest bardziej przejrzysta dla dalszych rozumowań.

Obraz zawierający diagram, tekst, szkic, linia

Opis wygenerowany automatycznie

***Rodzaje sieci RNN***

Modele RNN możemy podzielić na różne rodzaje sieci:

1. Sieć sekwencyjna (*ang. sequence-to-sequence network*)

Dostaje informację i dokonuje predykcji na ciąg kolejnych taktów, np.: podając cenę produktu z N ostatnich dni przewiduje jakie ceny będzie miał w kolejnych dniach (od N-1 do jutra).

1. Sieć sekwencyjno-wektorowa (*ang. sequence-to-vector network*)

Otrzymuje dane i zbierając ich informacje może dac końcowy wynik, np.: zbierając sekwencję zdań można określić , czy są one pozytywne czy negatywne.

1. Sieć wektorowo sekwencyjna (*ang. vector-to-sequence network*)

Stale jest jej przekazywana ta sama informacja na postawie której otrzymujemy wynik, np.: przedstawiamy sygnał dźwiękowy i dostajemy jego opis.

1. Koder-dekoder

Bardziej rozbudowana sieć przetwarzająca część danych na podstawie poprzednich. Stosowana między innymi do przetwarzania języka naturalnego.

Tutaj można wstawić wizualizację opisanych sieci.

Dla modelu RNN, charakterystyczne jest stosowanie BPTT (*ang. backpropagation through time*) – propagacji wstecznej w czasie, zamiast jej klasycznego odpowiednika.

Najpierw dokonywany jest przebieg do przodu poprzez sieć. Następnie sekwencja jest oceniana poprzez funkcję straty. Czy mogę tak zapisać ten wzór?

gdzie:

– wartości znane z zestawu danych,

– wartości przewidywane,

– maksymalny takt.

Po zakończeniu BPTT przeprowadza etap gradientu prostego w celu zaktualizowania parametrów.

**Dodanie informacji o exploding i vanishing gradient i wyjaśnienie dlaczego RNN nie działa na dłuższą metę.**

**Vanishing and exploding gradient (nie wiem czy tłumaczenie jest poprawne)**

***Zanikający i wybuchający gradient***

Jeśli pojawi się coś nowego z RNN dopisać tutaj!

## 4.2 Model LSTM

Jednym z bardziej rozbudowanych rekurencyjnych sieci neuronowych jest model LSTM (*ang. long-short term memory*). Jest on odpowiedzią na problem zanikającego i wybuchającego gradientu, zaprojektowany tak, aby efektywnie gospodarować „pamięcią”. Struktura LSTM jest o wiele bardziej skomplikowana, ale dzięki temu model ten może swobodnie działać na długich sekwencjach danych.

***Jednostka LSTM – struktura***

Podstawowym założeniem jednostki LSTM jest utrzymanie informacji o stanach wcześniejszych. Do tego celu struktura modelu posiada dwa wektory przekazywania informacji. Na rysunku X.X są one kolejno zaznaczone kolorami: zielony: stan ukryty (*ang.hidden*) - długotrwały i czerwony – stan komórkowy (*ang.cell*)– krótkotrwały.

Stan został pozbawiony niewygodnego mnożenia wag, dzięki czemu problem zanikającego i wybuchającego gradientu został usunięty. Stan posiada wagi dla każdej kolejnej bramki czemu w wygodny sposób będzie je modyfikował, nie tracąc informacji.

Śledząc drogę jaką pokonuje informacja można zauważyć, że stan napotyka **bramkę zapominającą.** *(ang. forget gate).* Oczywiście jej rola zapominania jest wykonywana, lecz co ciekawe ma ona też wpływ na zapamiętywanie poprzez dodanie nowych „wspomnień” za pomocą operacji sumowania, a następnie mnożenia. Funkcja aktywacji, która jest w niej zazwyczaj używana, sigmoid, pozwala na utrwalenie ile pamięci zostanie zachowane, np. jeśli wartości niewiele się wahają to zmiany będą widoczne, gdyż będą pomnożone przez ułamek, jeśli wynik będzie skrajny to albo usunie wartości w pamięci będący bliski zeru, lub zostawi wynik prawie bez zmian gdyż dokona się mnożenie przez 1.

**Bramka wyjściowa** *(ang. input gate)* decyduje, na podstawie bieżącego wejścia i poprzedniego stanu ukrytego, jak dużo informacji z komórki pamięci ma zostać ujawnione na wyjściu*.* Cześć odpowiadająca za zapamiętanie w jakim procencie nowy potencjalny stan długotrwały ma zostać zachowany, korzysta z funkcji aktywacji - sigmoid, lecz cześć odpowiadająca za zapamiętanie w jaka wartość nowego potencjalnego stanu długotrwałego ma zostać zachowany, korzysta z funkcji aktywacji tangensa hiperbolicznego. Funkcja ta może przekazać wartości z zakresu [-1, 1]. W ten sposób otrzymujemy nowy stan . Korzystanie z innej funkcji, niż sigmoid może również wskazywać na inny charakter tego elementu. Można to interpretować, że jest to główna warstwa generująca pamięć długotrwałą, a więc tą która w większym stopniu wpływa na wynik końcowy i jest kluczowa w strukturze całego modelu. Pozostałe warstwy nazywamy **kontrolerami bramek** *(ang. gate controllers)* i wpływają one na to ile danej wartości w pamięci zostanie zachowane.

Uzyskany nowy stan długotrwały wykorzystamy ponownie w **bramce wyjściowej** *(ang. output gate)*.

Działanie komórek wpływa na powstawanie nowych wspomnień – informacji przez nie przekształconych i utrwalonych.

|  |  |
| --- | --- |
| POLITECHNIKA RZESZOWSKA im. I. Łukasiewicza | Rzeszów, Rok |
| Wydział Elektrotechniki i Informatyki |  |
|  |  |

**STRESZCZENIE (wybierz rodzaj pracy)**

**TYTUŁ PRACY**

Autor: Imię Nazwisko, nr albumu: (wybierz symbol studiów)-123456 Opiekun: (tytuł naukowy przed) Imię Nazwisko(tytuł naukowy po)

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

(tekst streszczenia - max. 10 wierszy)

RZESZOW UNIVERSITY OF TECHNOLOGY Rzeszow, Rok

Faculty of Electrical and Computer Engineering

**DIPLOMA THESIS (wybierz rodzaj pracy) ABSTRACT**

**TYTUŁ PRACY W WERSJI ANGIELSKIEJ**

Author: Imię Nazwisko, code: (wybierz symbol studiów) -123456 Supervisor: (tytuł naukowy przed) Imię Nazwisko (tytuł naukowy po)

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

(tekst streszczenia w jęz. angielskim - max. 10 wierszy)