

Wstęp

Cel i zakres pracy

Przegląd literatury i analiza istniejących rozwiązań

Rozdział ten stanowi techniczne wprowadzenie do zagadnienia, oparte na analizie literatury i istniejących rozwiązań inżynierskich. Analiza literatury i istniejących rozwiązań stanowi istotny etap procesu projektowego, umożliwiając lepsze zrozumienie kontekstu danego problemu oraz identyfikację potencjalnych obszarów doskonalenia. Niniejszy rozdział skupia się na przeglądzie literatury związanej z tematyką pracy inżynierskiej oraz analizie istniejących rozwiązań, mającej na celu dostarczenie solidnej podstawy teoretycznej i technologicznej dla dalszych etapów badawczych. W tym rozdziale zostaną dogłębnie poruszone teoretyczne kwestie związane z tematem pracy.

Analiza szeregów czasowych i model ARIMA

Istota analizy szeregu czasowego

Szereg czasowym nazywamy uporządkowany zestaw danych, gdzie każda wartość jest przypisana do określonego momentu czasowego. Składa się z punktów danych, zwykle zbieranych w regularnych odstępach czasowych, co pozwala na analizę zmian w czasie. W ramach szeregów czasowych można identyfikować różne wzorce, trendy, sezonowe wahania oraz nieregularne zdarzenia. W kontekście informatyki szeregi czasowe są szeroko stosowane w analizie danych i prognozowaniu. Mogą obejmować dane z różnych dziedzin, takich jak gospodarka, nauki przyrodnicze, zdrowie, finanse czy technologia. Szeregi czasowe są używane do monitorowania i prognozowania zmian, co pozwala wspierać procesy decyzyjne. Przykłady zastosowań szeregów czasowych w informatyce to prognozowanie ruchu w sieciach komputerowych, monitorowanie wydajności systemów, analiza logów serwerów, predykcja awarii sprzętu czy prognozowanie trendów w danych ekonomicznych. Szeregi czasowe stanowią istotny element analizy danych w informatyce, pomagając w zrozumieniu dynamiki zjawisk w czasie.

Model ARIMA

W czasach przed opracowaniem modelu ARIMA prognozowanie wymagały posiadania wiedzy na temat matematycznego modelu procesu. W praktyce badawczej struktura szeregu czasowego bywa często niejednoznaczna, a wariancja składnika losowego jest znaczna. Pomimo tych trudności istnieje potrzeba nie tylko odkrywania ukrytych wzorców, ale również generowania prognoz. W tym celu została opracowana metodyka ARIMA, rozwinięta przez

Boxa i Jenkinsa (1976), która zdobyła znaczną popularność w różnych dziedzinach.

Model ARIMA *AutoRegressive Integrated Moving Average* to model używany do analizy i prognozowania przyszłych wartości w oparciu o historyczne dane.

Nazwa ARIMA opisuje trzy główne składowe tego modelu: AutoRegressive (**AR**), Integrated (**I**) i Moving Average (**MA**).

- AutoRegressive (AR): Ta część modelu odnosi się do autoregresji, czyli zależności między bieżącą wartością szeregu a jego wcześniejszymi wartościami. Model AR opiera się na przekonaniu, że bieżąca wartość szeregu czasowego jest funkcją jej poprzednich wartości, co uwzględnia wpływ autokorelacji.

$$x(t) = \phi_1 \cdot x(t-1) + \phi_2 \cdot x(t-2) + \dots + \phi_p \cdot x(t-p) + a(t) \quad (1)$$

- Integrated (I): Integracja dotyczy transformacji szeregu czasowego w celu uzyskania stacjonarności. Stacjonarność oznacza, że statystyki szeregu nie zmieniają się w czasie, co ułatwia analizę. Proces integracji polega na różniczkowaniu danych, czyli odejmowaniu od każdej wartości szeregu jej poprzedniej wartości.
- Moving Average (MA): Ta część modelu odnosi się do średniej ruchomej, czyli uwzględnienia pewnej liczby poprzednich składników losowych w modelu. Model MA zakłada, że bieżąca wartość szeregu czasowego jest sumą wcześniejszych błędów losowych.

$$x(t) = b_1 \cdot a(t-1) + b_2 \cdot a(t-2) + \dots + b_q \cdot a(t-q) + \varepsilon(t) \quad (2)$$

Wykrywanie wyjątków

Podział wyjątków

Efekt Maskowania

Konteneryzacja aplikacji

Bibliografia