

Modele predykcyjne w danych finansowych



Czym jest modelowanie predykcyjne?

Modelowanie predykcyjne to technika statystyczna wykorzystująca uczenie maszynowe (ML) i eksplorację danych do przewidywania i prognozowania prawdopodobnych przyszłych wyników przy użyciu zarówno danych historycznych, jak i istniejących.¹

Modele stosowane do predykcji danych finansowych.

- Regresja liniowa
- Drzewa decyzyjne i las losowy
- ARMA, ARIMA
- Sieci LSTM

Regresja liniowa

Regresja liniowa to jedna z najprostszych technik uczenia nadzorowanego. Modeluje ona zależność między zmienną zależną (ceną akcji) a jedną lub kilkoma zmiennymi niezależnymi (danymi rynkowymi). Celem jest dopasowanie równania liniowego do danych i prognozowanie przyszłej ceny na podstawie tej zależności.²

Drzewa decyzyjne i las losowy

- Drzewa decyzyjne pokazują decyzje w formie hierarchicznej struktury, co upraszcza interpretację wyników. Model dzieli dane na mniejsze grupy według wartości cech, wskazując optymalne ścieżki decyzyjne.³
- Las losowy to zaawansowany model łączący wiele drzew decyzyjnych, zwiększający dokładność prognoz i redukujący ryzyko przeuczenia. Agreguje wyniki z różnych drzew, zapewniając stabilniejsze prognozy. Znajduje zastosowanie w diagnostyce medycznej i analizie finansowej.³

ARMA, ARIMA

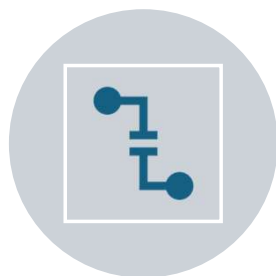
- **ARMA**

Model autoregresyjny (AR) i model średniej ruchomej (MA) mają swoje ograniczenia i mogą podlegać wpływom trendów. Aby przezwyciężyć wady każdego z nich i złagodzić ich ograniczenia, połączono je, tworząc model autoregresyjny średniej ruchomej (ARMA). Model ARMA ma jednak problemy z uwzględnianiem losowych wahań, dlatego też w tym modelu wymagana jest stacjonarność.⁴

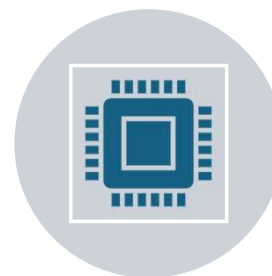
- **ARIMA**

Jeżeli dane nie są stacjonarne, możemy posłużyć się modelem ARIMA (Autoregressive Integrated Moving Average). Litera "I" w nazwie oznacza "zintegrowany". Odnosi się to do liczby razy jaką trzeba różnicować dane w celu osiągnięcia stacjonarności.⁴

LSTM (Long Short-Term Memory)



Pamięć długoterminowa (LSTM) to udoskonalona wersja rekurencyjnej sieci neuronowej (RNN). Sieci LSTM mogą rejestrować długoterminowe zależności w danych sekwencyjnych, co czyni je idealnymi do zadań takich jak na przykład rozpoznawanie mowy i prognozowanie szeregów czasowych.⁵

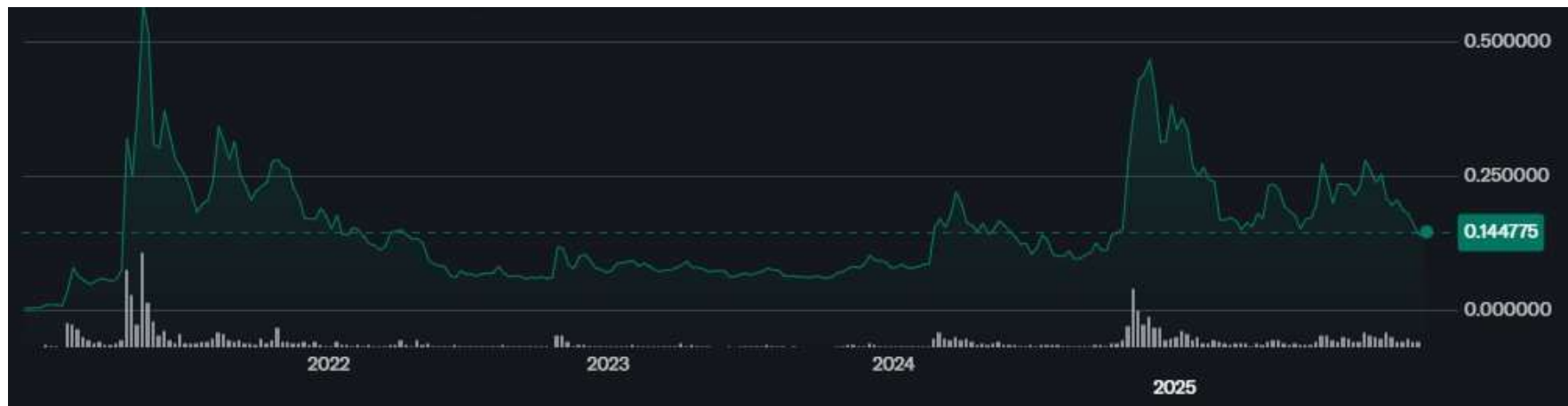


LSTM to typ sieci RNN, który ma wbudowaną komórkę pamięci i „bramki” (wejścia, wyjścia, zapominania), dzięki czemu sieć może świadomie decydować, jakie informacje zapamiętać, a jakie zapomnieć. Dzięki temu dobrze radzi sobie z długimi sekwencjami danych

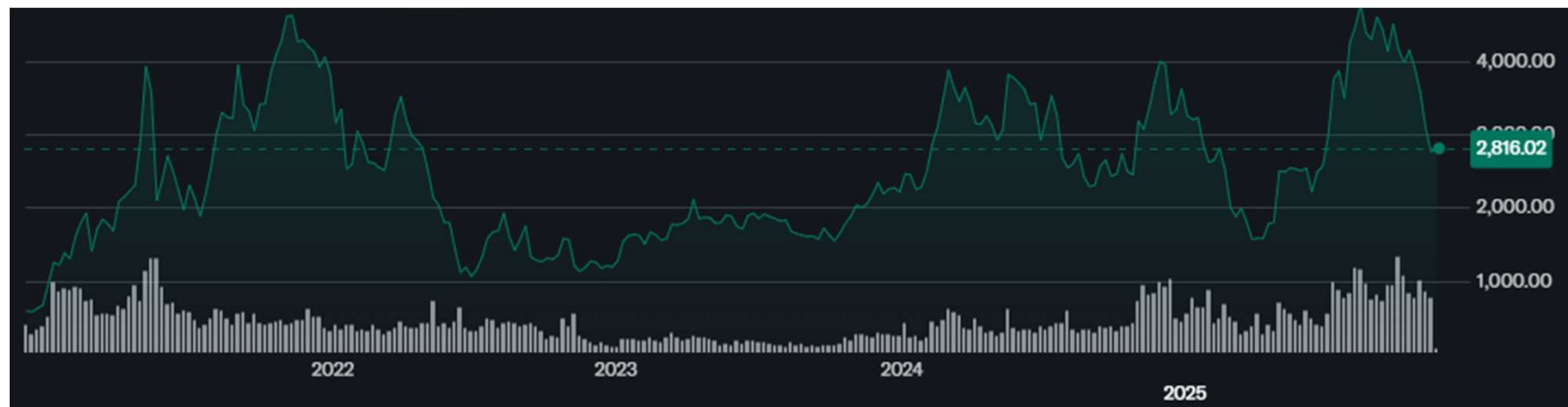
Nasze dane


Ticker	DOGE-USD	BTC-USD	ETH-USD	XRP-USD
23.11.2020	0.003648	18364.12109	608.4540405	0.613979995
24.11.2020	0.004204	19107.46484	603.8977661	0.692103028
25.11.2020	0.003727	18732.12109	570.6866455	0.635483027
26.11.2020	0.00325	17150.62305	518.8011475	0.533019006
27.11.2020	0.003275	17108.40234	517.4937134	0.559267998
28.11.2020	0.003405	17717.41406	538.2297974	0.625621974
29.11.2020	0.003448	18177.48438	575.7580566	0.606245995
30.11.2020	0.003551	19625.83594	614.8425293	0.664336979
01.12.2020	0.003335	18802.99805	587.3241577	0.612923026
02.12.2020	0.003374	19201.0918	598.352356	0.629893005
03.12.2020	0.003428	19445.39844	616.7087402	0.631483018
04.12.2020	0.003298	18699.76563	569.354187	0.557262003
05.12.2020	0.003397	19154.23047	596.595459	0.584420979
06.12.2020	0.003388	19345.12109	601.9089966	0.619850993

Dogecoin



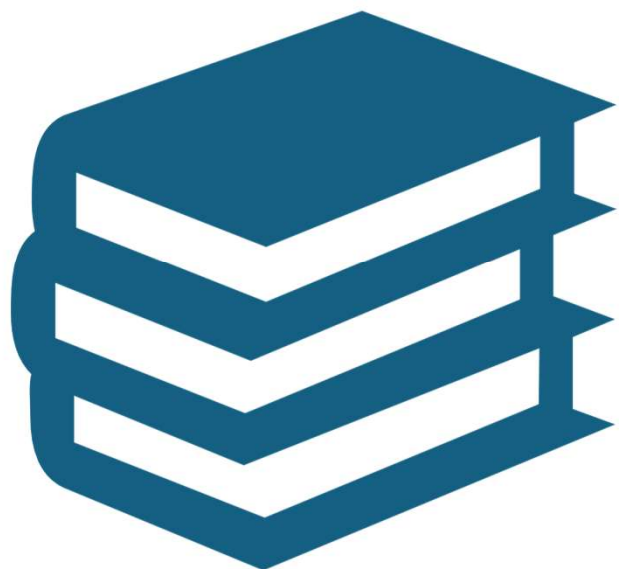
Ethereum





Wybrany
przez Nas
model

ARIMA



Źródła

1. <https://www.netsuite.com/portal/resource/articles/financial-management/predictive-modeling.shtml>
2. <https://leomercanti.medium.com/machine-learning-models-for-stock-price-prediction-70eeae94cc7b>
3. <https://aitrends.pl/model-predykcyjny-biznes>
4. <https://iamkartikeya.medium.com/time-series-modeling-13f076d6d6ca>
5. <https://www.geeksforgeeks.org/deep-learning/deep-learning-introduction-to-long-short-term-memory/>
6. <https://finance.yahoo.com>
7. <https://www.mathworks.com/discovery/lstm.html>