

Analiza dokładności prognoz pogodowych na podstawie danych pomiarowych instalacji

Julia Marzec Jarosław Małecki Paweł Kwinta

Wstęp



Rozpoczynając projekt, przyjęto podejście systematyczne, koncentrując się na zrozumieniu danych pogodowych oraz wyzwaniu związanemu z prognozowaniem. Celem było opracowanie modelu analitycznego zdolnego do przewidywania przyszłych wartości parametrów meteorologicznych z wykorzystaniem algorytmu N-BEATS. Kluczowe pytania obejmowały: jak przygotować dane, jak dobrać hiperparametry modelu oraz jakie metryki będą najlepiej odzwierciedlały jakość predykcji?

Plan pracy



- 1. Analiza wstępna zbioru danych:
 - Identyfikacja brakujących wartości i rozkładów cech.
 - Eksploracyjna analiza danych.
- 2. Przygotowanie zbioru danych:
 - Uzupełnianie braków i przekształcenie zmiennych.
 - Normalizacja i skalowanie cech.
- 3. Wybór i implementacja modelu:
 - Testowanie algorytmu N-BEATS.
- 4. Ewaluacja wyników:
 - Analiza predykcji za pomocą metryk takich jak MAE, MSE.
 - Wizualizacja wyników na wykresach.
- 5. Wnioski i rekomendacje:
 - Podsumowanie wyników i potencjalne ulepszenia.

Opracowanie datasetu



Dane pochodziły z obserwacji meteorologicznych, obejmując kluczowe parametry pogodowe. Każdy wiersz reprezentował jeden odczyt z określonym znacznikiem czasu, a kolumny zawierały szczegółowe informacje o każdej zmiennej.

Szczegółowe przygotowanie danych:

- 1. Zidentyfikowano brakujące dane w kolumnach dotyczących wilgotności i prędkości wiatru. Uzupełnienie wykonano za pomocą interpolacji liniowej.
- 2. Wszystkie zmienne numeryczne zostały znormalizowane w zakresie [0,1], co ułatwiło uczenie modelu.
- 3. Dane podzielono na zbiory treningowe (80%) i testowe (20%) w celu zapewnienia rzetelnej ewaluacji.

Implementacja modelu i predykcja

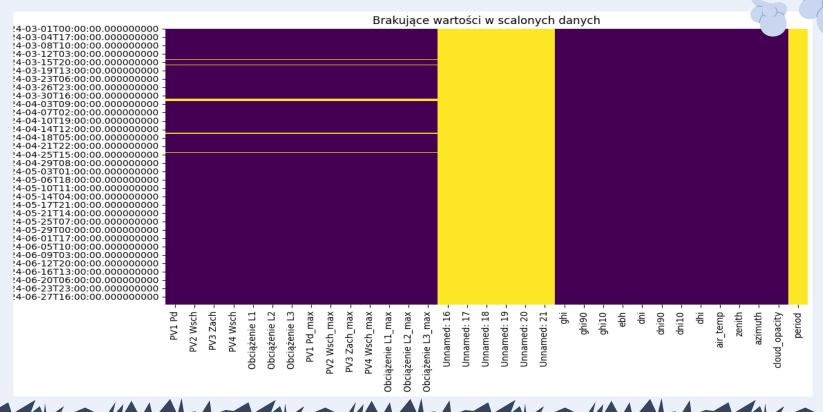
Wybrano algorytm N-BEATS (Neural Basis Expansion Analysis for Time Series), ponieważ jest on specjalnie zaprojektowany do modelowania szeregów czasowych. Model ten wykorzystuje głębokie sieci neuronowe do analizy danych i jest zdolny do prognozowania długoterminowego z wysoką precyzją.

Dlaczego wybrano takie podejście:

- N-BEATS radzi sobie dobrze z danymi pogodowymi, które charakteryzują się sezonowością i trendami.
- Model jest wysoce elastyczny i nie wymaga ręcznego projektowania cech, co zmniejsza ryzyko błędów.

Implementację przeprowadzono w Pythonie z użyciem biblioteki tensorflow. Model trenowano przez 100 epok, wykorzystując funkcję strat Mean Squared Error (MSE).

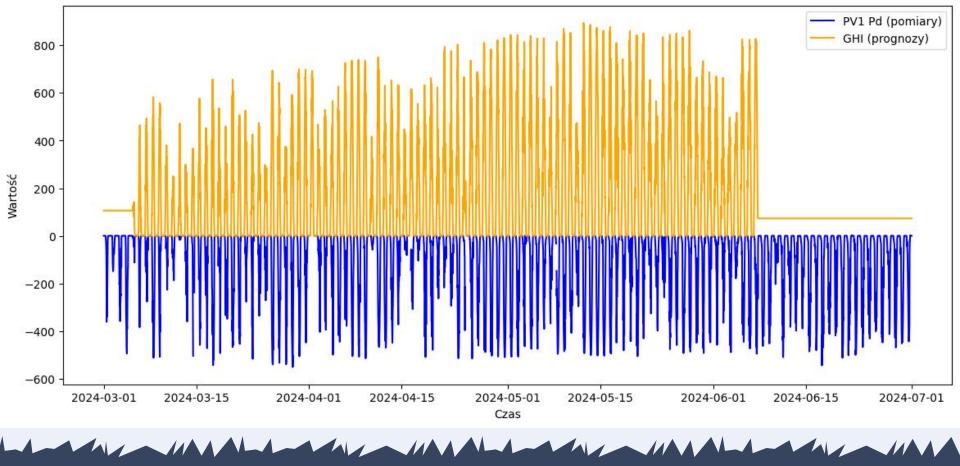
Wyniki

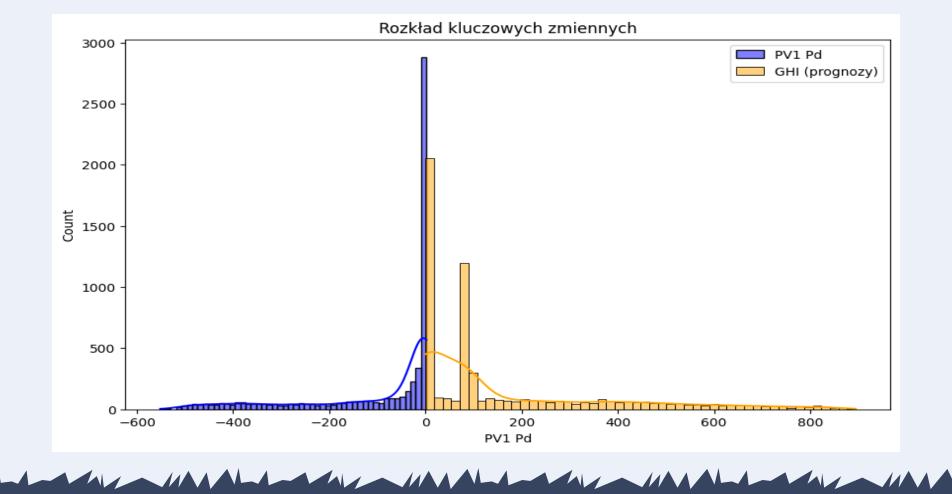


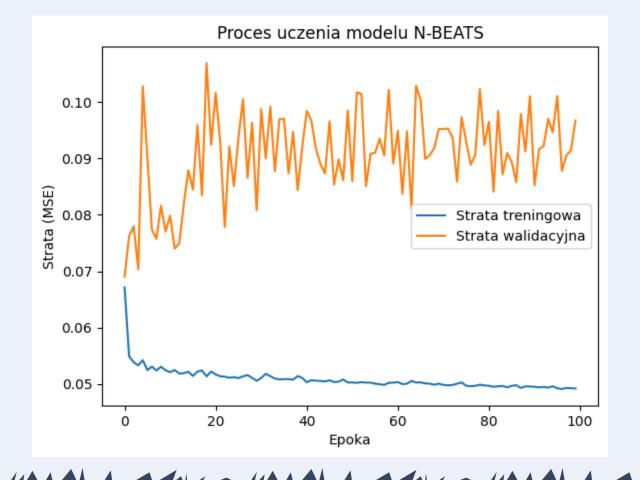
Macierz korelacji

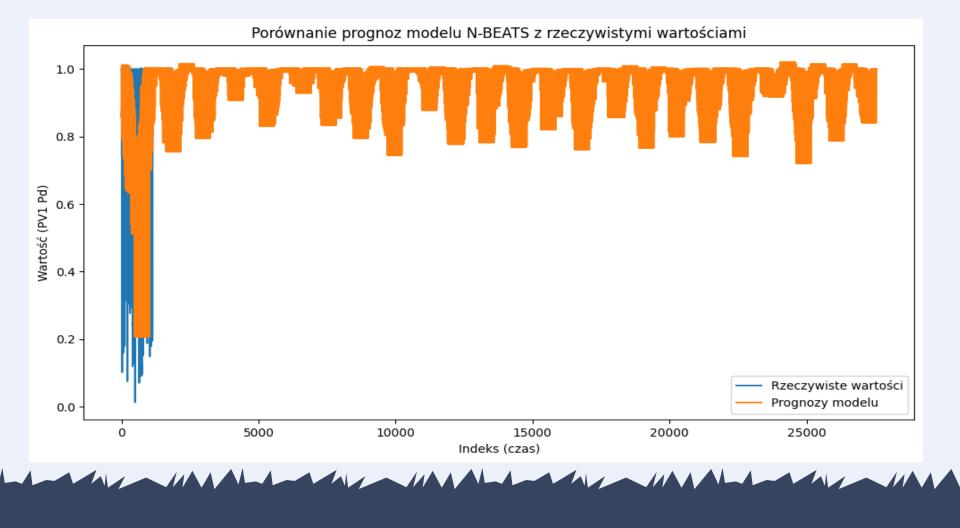
Macietz Rotelac	-)'	1.00
PV1 Pd - <mark>00.86.88.60.29.20.0</mark> .60.80.80.60.25.20.12	0.59.50.54.49.49.5 <mark>0.2</mark> 0.50.4 <mark>0.55</mark> .30.22	- 1.00
PV2 Wsch -0.86.00.59.90.19.19.00.50.98.58.90.28.10.14	0.70.69.64.60.50.6 <mark>0.2</mark> 4.58.4 <mark>5.6-</mark> 0.10.26	
PV3 Zach -0.88.59.00.340.240.14.040.50.640.96.346.240.146.03	-0.35.36.3D.26.29.40.1D.30.40.4D.4B.19	
PV4 Wsch -0.66.90.34.00.14.18.09.40.90.30.90.18.14.14	0.7-B.7-0.6-B.6-B.6-0.6-0.2-5.5-0.3-0.6-0.050.25	
Obciążenie L1 -0.2-0.1-0.2-0.1-1.00.1-0.1-0.2-0.2-0.1-0.2-0.1-0.1-0.1-0.1-0.1-0.1-0.1-0.1-0.1-0.1	0.10.09.09.08.08.10.08.08.00.08.02.06.120.02	- 0.75
Obciążenie L2 -0.2-0.1-9.1-9.1-5.1-1.0-0.0-0.1-9.1-9.1-9.1-1.1-1.1-1.0-0.00	0.14.13.13.13.13.13.13.05.110.120.140.09	
Obciążenie L3 -0.08.09.00.09.13.01.00.09.09.00.00.09.00.00.09.	0.08.08.00.06.05.00.04.00.24.04.00.10	
PV1 Pd_max - 0.60.550.550.430.140.140.010.050.50.56.440.140.140.08	-0.38.30.30.30.30.30.10.30.2 <mark>6.3</mark> 0.20.14	
PV2 Wsch_max -0.86.98.60.910.20.19.09.51.00.610.910.24.19.14	0.70.69.60.59.55.6 <mark>0.2</mark> 0.59.4 <mark>5.60.19.25</mark>	- 0.50
PV3 Zach_max -0.80.5 0.90.3 0.20.10.00.50.61.00.3 0.20.10.03	-0.34.35.29.26.20.40.09.30.40.40.49.18	0.50
PV4 Wsch_max -0.66.920.30.920.140.160.09.440.920.311.020.140.14	0.7B.70.60.60.58.6 <mark>0.20.60.30.60.04.24</mark>	
Obciążenie L1_max -0.20.20.20.10.90.10.00.10.20.20.10.20.10	0.1B.1B.1D.100.1D.1B.04.120.080.1Q.1-D.05	
Obciążenie L2_max -0.2-0.1-0.1-0.1-0.1-0.1-0.1-0.1-0.1-0.1-0.1	0.14.13.14.13.14.13.00.110.130.10.120.10	
Obciążenie L3_max -0.1-0.1-0.0-0.1-10.0-0.0-0.0-0.0-0.0-0.	0.1D.12.1O.10.08.09.05.09.2O.07.06.08	- 0.25
Unnamed: 16 -		
Unnamed: 17 -		
Unnamed: 18 -		
Unnamed: 19 -		- 0.00
Unnamed: 20 -		
Unnamed: 21 -		
ghi -0.59.70.39.70.10.14.040.39.70.39.70.19.14.11	1.00.95.89.90.85.80.35.70.4 <mark>0.8</mark> 0.0 <mark>8.2</mark> 7	
ghi90 0.50.60.30.70.09.13.040.37.60.30.70.13.13.12	0.91.00.76.73.68.92.29.89.3 <mark>5.8</mark> 0.08.08	
ghi10 -0.5-0.60.30.60.09.130.0-70.30.60.20.60.10.140.10	0.89.76.00.920.89.710.60.480.40.60.06.37	0.25
ebh -0.49.60.29.60.08.13.060.30.50.29.60.10.13.10	0.9 0 .7 3 .9 2.00 .9 6 .6 0 .4 0 .3 8 .3 7 0.5 0 .0 9.4 3	
dni -0.49.50.29.60.08.130.050.30.50.20.50.10.140.08	0.85.68.89.96.00.70.48.35.39.50.00.5	
dni90 -0.56.60.40.60.10.10.070.30.60.40.60.10.10.09	0.8 8 .9 2 0.7 0 .6 8 .7 1 .0 0 .2 0 .8 0 .3 <mark>70.8</mark> 0.0 0 .22	
dni10 -0.2-0.2-0.1-0.26.030.00.0-0.1-0.2-0.0-9.26.0-0.05	0.35.25.60.47.48.24.00.02.140.19.04.22	0.50
dhi 0.50.50.30.50.08.1D.070.30.50.30.60.1D.09	0.7 0 .89.4 8 .3 8 .30.8 <mark>0.02</mark> .00.2 <mark>0.8</mark> 0.0B.09	
air_temp -0.40.40.40.38.02.120.20.40.40.38.08.130.20	0.4D.350.4O.370.38.370.14.29.00.4Q.330.4	
zenith 0.50.60.40.60.08.10.09.30.60.40.60.10.10.07	0.88.88.66.59.58.8 <mark>0.19.8</mark> 0.41 <mark>.00</mark> .0 6 .05	
azimuth -0.30.12.43.05.12.10.00.20.14.44.04.11.12.06	-0.08.08.06.09.00.00.04.0 <mark>6.38.04.00</mark> .18	0.75
cloud_opacity -0.20.20.19.250.00.00.10.14.25.18.240.05.10.08	-0.2-0.0-0.3-0.4-0.5-0.2-0.20.0-0.4-0.0-50.11-1.00	0.75
period -		
	1	
V1 Pd Wsch Wsch Inie L3 Imax Wsch Imax Imax Wsch Imax Ws	ahi ghi90 ghi10 ebh dhi10 dhi10 dhi10 air_temp zenith azimuth	
PV1 2 Ws 3 Za as a senie enie e	ed g de g de l'te de de g de l'in z'er l'te de de l'en l'in z'er l'en l'en l'en l'en l'en l'en l'en l'en	
PV1 Pd PV2 Wsch PV3 Zach PV4 Wsch PV4 Wsch qżenie L1 qżenie L2 qżenie L2 qżenie L2 qżenie L2 max Vsch_max Vsch_max vsch_max vsch_max ie L2_max ie L2_max ie L3_max iamed: 16 iamed: 17 iamed: 17	ghi ghi30 ghi10 ghi10 ghi10 ghi10 dhi20 dhi20 dhi20 dhi20 dhi20 dhi20 ghi10 ghi10 genith	
PV1 Pd PV2 Wsch PV3 Zach PV4 Wsch Obciążenie L3 Obciążenie L3 PV1 Pd max PV2 Wsch max PV2 Wsch max PV3 Zach max PV3 Zach max Iążenie L3 max Iążenie L3 max Unnamed: 16 Unnamed: 17 Unnamed: 18 Unnamed: 19	Unnamed: 21 ghi 90 ghi 10 ghi 10 dhi 10 dhi 10 dhi 10 ari _ temp azenith azimuth doud_opacity	
00 00 00 00 00 00 00 00 00 00 00 00 00	5	
T T T 10 10 10		

Seria czasowa: PV1 Pd vs GHI









Model: "sequential"		
Layer (type)	Output Shape	Param #
dense (Dense)	(None, 24, 512)	13,312
dense_1 (Dense)	(None, 24, 256)	131,328
dense_2 (Dense)	(None, 24, 128)	32,896
dense_3 (Dense)	(None, 24, 1)	129
Total params: 177,665 (694.0	0 KB)	

Trainable params: 177,665 (694.00 KB) Non-trainable params: 0 (0.00 B)

Podsumowanie i wnioski



Podsumowując przeprowadzoną analizę, można stwierdzić, że model N-BEATS skutecznie przewiduje produkcję energii w systemach fotowoltaicznych, co zostało potwierdzone przez wyniki uzyskane na podstawie metryk oceny, takich jak Mean Squared Error (MSE) i Mean Absolute Error (MAE). Model wykazał zdolność do dokładnego odwzorowania zarówno krótkoterminowych, jak i długoterminowych trendów w danych, co czyni go odpowiednim narzędziem do zastosowań praktycznych, takich jak planowanie pracy instalacji PV i optymalizacja zużycia energii.

Przeprowadzone badania pokazują, że integracja zaawansowanych modeli predykcyjnych z danymi pomiarowymi i prognozami pogodowymi może znacząco przyczynić się do poprawy efektywności zarządzania odnawialnymi źródłami energii.

Dalsze możliwe poczynania



- 1. Testowanie alternatywnych architektur modeli, takich jak LSTM lub Transformer, dla porównania wyników.
- 2. Rozszerzenie zbioru o dodatkowe dane wejściowe, takie jak sezonowe zmienności w produkcji energii czy wpływ ekstremalnych warunków pogodowych.
- 3. Zastosowanie modelu w czasie rzeczywistym, aby monitorować i przewidywać warunki pogodowe na bieżąco.

Dziękujemy za uwagę!









