

Analiza dokładności prognoz pogodowych na podstawie danych pomiarowych instalacji

Julia Marzec^a, Jarosław Małecki^b and Paweł Kwinta^{c,*}

^aWydział Elektrotechniki, Automatyki i Informatyki, Politechnika Opolska, Opole, Polska

ARTICLE INFO

Keywords:
fotowoltaika
prognozy pogodowe
agregacja danych
produkcja energii
zarządzanie energią

ABSTRACT

Program przetwarza dane pomiarowe i prognozy pogodowe z instalacji fotowoltaicznych w celu przewidywania produkcji energii za pomocą modelu N-BEATS. Łączy i agreguje dane w interwałach czasowych, przeprowadza ich analizę i czyszczenie, a następnie trenuje model na przekształconych sekwencjach czasowych. Wyniki prognoz są oceniane i wizualizowane, wspierając optymalizację zarządzania energią w systemach PV.

1. Wstęp

Dynamiczny rozwój odnawialnych źródeł energii, takich jak instalacje fotowoltaiczne (PV), stanowi kluczowy element globalnej transformacji energetycznej. Wzrost znaczenia energii słonecznej jako źródła zrównoważonej energii wymaga precyzyjnych narzędzi do analizy i prognozowania wydajności instalacji PV w zmieniających się warunkach atmosferycznych.

Celem projektu jest stworzenie programu, który integruje dane pomiarowe z instalacji PV z prognozami pogodowymi oraz wykorzystuje model N-BEATS do przewidywania produkcji energii. Model N-BEATS, opracowany z myślą o analizie szeregów czasowych, umożliwia dokładne prognozowanie, co jest kluczowe dla optymalizacji zarządzania energią i zwiększenia efektywności systemów fotowoltaicznych.

Dane wejściowe obejmują:

- Pomiarowe dane z instalacji PV, które zawierają informacje o produkcji energii oraz obciążeniach systemów.
- Prognozy pogodowe, dostarczające szczegółowych danych o promieniowaniu słonecznym, temperaturze powietrza i innych kluczowych zmiennych atmosferycznych.

Program realizuje następujące kroki:

1. Wczytanie danych pomiarowych i prognoz pogodowych.
2. Przetwarzanie danych, w tym ich synchronizację, interpolację i agregację do ustalonych interwałów czasowych.
3. Eksploracyjna analiza danych (EDA), mająca na celu identyfikację kluczowych zmiennych oraz oceny jakości danych.
4. Budowa i trening modelu N-BEATS na danych historycznych.
5. Ocena wyników oraz wizualizacja prognoz w odniesieniu do rzeczywistych pomiarów.

W projekcie zastosowano zaawansowane metody przetwarzania danych, które pozwalają na efektywne łączenie pomiarów z prognozami. Dzięki temu możliwe jest nie tylko

przewidywanie wydajności instalacji PV, ale także lepsze planowanie ich pracy w oparciu o zmieniające się warunki atmosferyczne. Praca ta wpisuje się w nurt badań nad optymalizacją systemów energetycznych, co znajduje potwierdzenie w literaturze przedmiotu (1; 2; 3).

2. Cel i zakres

Głównym celem niniejszej pracy było opracowanie i implementacja programu do analizy danych pomiarowych i prognozowych związanych z instalacjami fotowoltaicznymi. Zastosowano model **N-BEATS**, aby przewidywać produkcję energii w oparciu o dane pogodowe i pomiarowe, co umożliwia efektywniejsze zarządzanie energią.

2.1. Integracja danych

Pierwszym krokiem była integracja danych pomiarowych i prognoz pogodowych. W tym celu dokonano harmonizacji czasowej, interpolacji brakujących wartości oraz agregacji danych do ustalonych interwałów czasowych. Proces ten pozwolił na precyzyjne połączenie danych o różnej częstotliwości pomiarów, co zwiększyło ich przydatność analityczną.

Ważnym aspektem było zachowanie spójności czasowej między danymi. Dzięki zastosowaniu zaawansowanych technik interpolacji możliwe było uzupełnienie brakujących wartości w sposób minimalizujący wpływ na jakość danych. Agregacja danych w ustalonych interwałach czasowych, takich jak 30 minut, pozwoliła na dostosowanie ich do wymagań modelu N-BEATS.

2.2. Modelowanie predykcyjne

Następnym etapem było wykorzystanie modelu **N-BEATS**, który został wytrenowany na przygotowanych danych w celu przewidywania przyszłej produkcji energii. Model ten wykorzystuje struktury sieci neuronowych zoptymalizowane do analizy szeregów czasowych, co pozwala na uchwycenie zarówno krótkoterminowych, jak i długoterminowych trendów w danych.

Aby ocenić skuteczność modelu, przeprowadzono ewaluację z użyciem wskaźników takich jak *Mean Squared Error (MSE)* i *Mean Absolute Error (MAE)*. Wyniki modelu

ORCID(s):

zostały porównane z rzeczywistymi pomiarami, co pozwoliło na weryfikację dokładności prognoz.

2.3. Zastosowania praktyczne

Rezultaty pracy znajdują zastosowanie w optymalizacji zarządzania energią w systemach fotowoltaicznych. Prognozy produkcji energii mogą wspierać operatorów w lepszym planowaniu obciążeń sieci, minimalizowaniu strat energii oraz zwiększaniu wydajności instalacji PV.

Dzięki dokładnym prognozom możliwe jest również lepsze dostosowanie pracy instalacji PV do zmieniających się warunków atmosferycznych, co ma kluczowe znaczenie dla poprawy efektywności operacyjnej i ekonomicznej systemów energetycznych. Takie podejście wpisuje się w globalne dążenia do zwiększenia udziału odnawialnych źródeł energii w miksie energetycznym.

3. Zbiór danych (Dataset)

Dane wykorzystane w projekcie pochodzą z dwóch głównych źródeł:

1. Dane pomiarowe z instalacji PV:

- Zawierają informacje o produkcji energii przez różne komponenty systemu, takie jak *PV1 Pd*, *PV2 Wsch*, *PV3 Zach*.
- Uwzględniają obciążenia w liniach systemu (*Obciążenie L1*, *Obciążenie L2*, *Obciążenie L3*).
- Dane były zbierane w interwałach pięciominutowych, co umożliwia szczegółową analizę wydajności instalacji.

2. Dane prognozowe:

- Prognozy pogodowe, takie jak *ghi* (Global Horizontal Irradiance), *dni* (Direct Normal Irradiance), *cloud_opacity*, *air_temp*, *zenith*, *azimuth*.
- Dane generowane co dwie godziny, zawierające prognozy na kolejne godziny.
- Wartości zostały zharmonizowane z danymi pomiarowymi przez interpolację i agregację.

Proces przetwarzania danych obejmował usuwanie błędnych i niekompletnych rekordów oraz harmonizację czasową w celu uzyskania zestawów danych gotowych do modelowania predykcyjnego.

Tak zorganizowany zbiór danych okazał się bogatym źródłem informacji, pozwalającym na analizę efektywności treningowej, ocenę wydolności w zależności od dyscypliny oraz planowanie optymalnych obciążeń treningowych.

4. Algorytm N-BEATS

Model **N-BEATS** (Neural Basis Expansion Analysis for Time Series) jest zaawansowaną architekturą opartą na sieciach neuronowych, zaprojektowaną specjalnie do analizy i prognozowania szeregów czasowych. Kluczowym elementem tego modelu jest jego elastyczność i zdolność do

przewidywania zarówno krótkoterminowych, jak i długoterminowych trendów w danych. W odróżnieniu od innych modeli, N-BEATS nie wymaga skomplikowanej inżynierii cech, co czyni go szczególnie efektywnym przy pracy z surowymi danymi.

4.1. Architektura modelu

Model N-BEATS opiera się na blokach prognostycznych, które składają się z dwóch głównych komponentów:

- **Backward Model:** Odpowiada za analizę wzorców w danych historycznych, które są następnie używane do prognozowania.
- **Forward Model:** Umożliwia generowanie prognoz w oparciu o wnioski wyciągnięte przez model backward.

Każdy blok składa się z warstw w pełni połączonych, które transformują dane wejściowe za pomocą funkcji aktywacji ReLU. Model wykorzystuje również podejście typu ensemble, gdzie wiele bloków działa równolegle, aby zwiększyć dokładność prognoz i ograniczyć przeuczenie.

4.2. Zastosowanie do zbioru danych

W analizowanym projekcie model N-BEATS został zastosowany do przewidywania produkcji energii w systemach PV w oparciu o:

- Dane pomiarowe, takie jak obciążenia systemu i produkcja energii w różnych komponentach instalacji.
- Prognozy pogodowe, w tym promieniowanie słoneczne, temperatura powietrza i opakowanie chmur.

Dane wejściowe zostały przygotowane w formie sekwencji czasowych, gdzie każda sekwencja zawierała 24 godziny danych historycznych, a model prognozował wartości dla następnych godzin. Dzięki zastosowaniu N-BEATS możliwe było uchwycenie sezonowych wzorców oraz dynamicznych zmian w produkcji energii, co pozwoliło na zwiększenie dokładności prognoz.

4.3. Wyniki i efektywność

Model N-BEATS okazał się skuteczny w analizie danych fotowoltaicznych, osiągając niskie wartości błędów prognostycznych (MSE i MAE). Prognozy generowane przez model były zbliżone z rzeczywistymi pomiarami, co potwierdziło jego przydatność w praktycznych zastosowaniach. Elastyczność modelu pozwoliła również na szybkie dostosowanie do zmian w danych wejściowych, takich jak różna częstotliwość pomiarów.

5. Zastosowane metryki oceny

Aby ocenić skuteczność modelu N-BEATS, zastosowano następujące metryki oceny jakości prognoz:

- **Mean Squared Error (MSE):**

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

Metryka ta mierzy średnią kwadratową różnic między prognozowanymi a rzeczywistymi wartościami. Niższe wartości MSE wskazują na wyższą dokładność prognoz.

- **Mean Absolute Error (MAE):**

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2)$$

MAE mierzy średni błąd absolutny między prognozowanymi a rzeczywistymi wartościami, co pozwala na ocenę bezwzględnych odchyleń prognoz od obserwacji.

- **Symmetric Mean Absolute Percentage Error (sMAPE):**

$$sMAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(y_i + \hat{y}_i)/2} \quad (3)$$

Metryka ta wyraża błąd w procentach, co ułatwia interpretację wyników w kontekście względnym.

- **Coefficient of Determination (R²):**

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

Wskaźnik R² ocenia, jaka część wariancji danych jest wyjaśniana przez model. Wartość bliska 1 oznacza wysoką jakość dopasowania modelu.

Metryki te zostały wykorzystane do kompleksowej oceny jakości prognoz generowanych przez model N-BEATS, co pozwoliło na dokładne określenie jego skuteczności w analizie danych fotowoltaicznych.

6. Zastosowanie i implementacja algorytmu

W niniejszym projekcie algorytm N-BEATS został zaimplementowany przy użyciu biblioteki *PyTorch*, która umożliwia tworzenie i trenowanie zaawansowanych modeli sieci neuronowych. Implementacja obejmowała następujące etapy:

6.1. Przygotowanie danych

Dane wejściowe zostały przygotowane w formie sekwencji czasowych. Każda sekwencja zawierała dane historyczne z ostatnich 24 godzin, które były używane do prognozowania wartości na kolejne godziny. Dane zostały:

- Znormalizowane przy użyciu *MinMaxScaler*, co pozwoliło na ujednolicenie skali wartości.
- Podzielone na zbiór treningowy (80%) i testowy (20%).
- Sformatowane w taki sposób, aby pasowały do wymagań modelu N-BEATS.

6.2. Budowa modelu

Model N-BEATS został skonfigurowany z następującymi parametrami:

- Liczba bloków prognostycznych: 4.
- Rozmiar wewnętrznej reprezentacji: 256 neuronów na warstwę.
- Funkcja aktywacji: ReLU.
- Optymalizator: Adam, z początkową szybkością uczenia *learning rate* równą 0.001.

Każdy blok modelu generował prognozy, które były następnie sumowane, aby uzyskać końcowy wynik.

6.3. Trenowanie modelu

Model został wytrenowany przez 100 epok z użyciem wskaźnika błędu *Mean Squared Error* (MSE) jako funkcji kosztu. W trakcie trenowania monitorowano błędy na zbiorze testowym, co pozwoliło na optymalizację parametrów modelu.

6.4. Prognozowanie

Po zakończeniu trenowania model był używany do generowania prognoz produkcji energii na podstawie danych wejściowych. Wyniki prognoz były porównywane z rzeczywistymi wartościami w celu oceny skuteczności modelu.

6.5. Zastosowanie praktyczne

Implementacja algorytmu N-BEATS pozwala na:

- Dokładne prognozowanie produkcji energii, co wspiera operatorów systemów PV w podejmowaniu decyzji.
- Identyfikację wzorców sezonowych i dynamicznych zmian w danych.
- Zastosowanie w praktyce, takich jak planowanie obciążeń sieci i optymalizacja pracy instalacji PV.

7. Wyniki

7.1. Proces uczenia modelu

W procesie uczenia modelu N-BEATS obserwowano stopniowe zmniejszanie wartości funkcji kosztu na zbiorze treningowym, co wskazuje na skuteczne dopasowanie modelu do danych. Na zbiorze walidacyjnym wartości te były stabilne, co świadczy o braku przeuczenia modelu. Wartość końcowa funkcji kosztu dla zbioru walidacyjnego wyniosła około 0.09, a dla zbioru treningowego spadła do 0.05.

7.2. Porównanie rzeczywistych i prognozowanych wartości

Analiza wyników prognozowania modelu N-BEATS wskazuje na dużą zgodność między wartościami rzeczywistymi a prognozowanymi. Na podstawie wartości wskaźników błędu, takich jak Mean Absolute Error (MAE) na poziomie 0.19 oraz Mean Squared Error (MSE) na poziomie 0.09, można stwierdzić, że model charakteryzuje się wysoką precyzją w prognozowaniu produkcji energii.

7.3. Wnioski z metryk oceny

Wyniki uzyskane na podstawie metryk oceny wskazują na następujące wnioski:

- **Mean Squared Error (MSE):** Wartość MSE wynosząca 0.09 potwierdza, że model N-BEATS skutecznie minimalizuje większe odchylenia między wartościami prognozowanymi a rzeczywistymi, co wskazuje na dokładne odwzorowanie trendów danych.
- **Mean Absolute Error (MAE):** Niska wartość MAE (0.19) oznacza, że średnie różnice między prognozami a rzeczywistymi danymi są niewielkie, co czyni model odpowiednim do praktycznych zastosowań, takich jak zarządzanie energią.
- **Brak przeuczenia:** Stabilne wyniki na zbiorze walidacyjnym wskazują na dobrą ogólną zdolność uogólniania modelu bez ryzyka nadmiernego dopasowania do danych treningowych.
- **Potencjał praktyczny:** Wysoka precyzja modelu oraz jego zdolność do przewidywania krótkoterminowych i długoterminowych trendów pozwalają na zastosowanie w optymalizacji pracy instalacji fotowoltaicznych.

Model N-BEATS udowodnił swoją skuteczność w analizie danych fotowoltaicznych, co czyni go wartościowym narzędziem w kontekście prognozowania produkcji energii i zarządzania zasobami energetycznymi.

8. 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.

Wnioski płynące z projektu obejmują:

- Model N-BEATS charakteryzuje się wysoką precyzją prognozowania, co potwierdzają niskie wartości błędów (MSE: 0.09, MAE: 0.19).
- Brak przeuczenia modelu, co świadczy o jego zdolności do uogólniania i przewidywania wartości dla nowych danych.
- Możliwość praktycznego zastosowania wyników w zarządzaniu energią, w tym optymalizacji pracy instalacji PV oraz planowaniu produkcji energii w zmieniających się warunkach pogodowych.

W przyszłości warto rozważyć:

- Rozszerzenie zbioru o dodatkowe dane wejściowe, takie jak sezonowe zmienności w produkcji energii czy wpływ ekstremalnych warunków pogodowych.
- Porównanie skuteczności modelu N-BEATS z innymi metodami prognozowania, np. LSTM czy Prophet, co może pomóc w jeszcze lepszym dostosowaniu narzędzi analitycznych.
- Zastosowanie modelu w szerszym kontekście, np. w integracji z systemami zarządzania energią w budynkach inteligentnych.

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.

9. Bibliografia

References

- [1] Oreshkin B. et al., *N-BEATS: Neural Basis Expansion Analysis for Time Series Forecasting*, 2020.
- [2] Liu K. et al., *A review of photovoltaic systems performance prediction and its influencing factors*, Renewable and Sustainable Energy Reviews, 2021.
- [3] Little R.J.A., Rubin D.B., *Statistical Analysis with Missing Data*, Wiley, 2002.