

Prognozowanie produkcji energii fotowoltaicznej z wykorzystaniem platformy NEXO i VRM Portal

Streszczenie. W pracy przedstawiono wyniki prognozowania produkcji energii fotowoltaicznej z wykorzystaniem dwóch modeli regresyjnych: regresji LASSO oraz lasu losowego. Dane o produkcji energii uzyskano przy użyciu mikroinstalacji Victron. Dzięki połączeniu usługi VRM Portal oraz platformy teleinformatycznej NEXO, możliwe było pozyskanie danych niezbędnych do nauki modeli prognostycznych. W artykule zaprezentowano metodologię modelowania produkcji energii fotowoltaicznej przy użyciu języka Python oraz biblioteki SkForecast. Na podstawie wyników badań stwierdzono, że najlepsze prognozy uzyskuje się stosując model lasów losowych.

Abstract. The paper presents forecasting of solar energy production using two regression models: LASSO regression and random forest. Data on energy production was obtained using Victron micro-installations. Thanks to the combination of the VRM Portal service and the NEXO platform, it was possible to obtain the data necessary to learn forecasting models. The article presents the methodology of solar energy production modeling using the Python language and the SkForecast library. Based on the test results, it was found that the best forecasts are obtained using the random forests model. (**Forecasting of solar energy production using the NEXO platform and VRM Portal**).

Słowa kluczowe: prognozowanie produkcji energii fotowoltaicznej, regresja LASSO, las losowy, Internet Rzeczy.

Keywords: PV energy forecasting, LASSO regression, random forest, NEXO, Internet of Things.

Platforma teleinformatyczna NEXO

Obecnie Internet Rzeczy (ang. *IoT*, *Internet of Things*) jest jedną z najdynamiczniej rozwijających się technologii informatycznych [1]. W dużej mierze rozwój ten jest wynikiem coraz większej dostępności technologii sieciowych i Internetu w wielu gałęziach przemysłu oraz dla odbiorców indywidualnych. Koncepcja Internetu Rzeczy wciąż ewoluuje, a obszar jego zastosowań poszerza się. Rozwojowi *IoT* silnie sprzyja również trend cyfryzacji przemysłu oraz życia codziennego, w którym coraz większą rolę odgrywają inteligentne urządzenia lub usługi cyfrowe [1, 2]. W najprostszej postaci Internet Rzeczy to technologia pozwalająca na komunikację pomiędzy wieloma urządzeniami końcowymi (ang. *edge devices*) oraz usługami sieciowymi (ang. *services*), w tym chmurami obliczeniowymi (ang. *cloud computing*). Na szczególną uwagę zasługują systemy typu *Smart Home*, pozwalające na inteligentne zarządzanie budynkami. Rozwiązania takie pozwalają integrować w ramach jednego systemu automatyki zarządzanie ogrzewaniem, zużyciem energii i innych mediów, bezpieczeństwem itp. [3]. Przykładami systemów automatyki domowej są zarówno dostępne komercyjnie węzły sterujące takie jak: HomeKit (Apple), Google Home (Google) lub Amazon Echo (Amazon), jak i darmowe systemy OpenHub, Domoticz lub HomeAssistant posiadające wsparcie szerokiej społeczności *open source*.

Jednym z dostępnych systemów *IoT* pozwalających na integrację wielu urządzeń automatyki budynkowej, z wykorzystaniem popularnych protokołów *IoT* jest platforma teleinformatyczna NEXO. Architektura platformy dla jednego węzła lokalnego (ang. *local node*) została przedstawiona na rysunku 1. W skład platformy wchodzi:

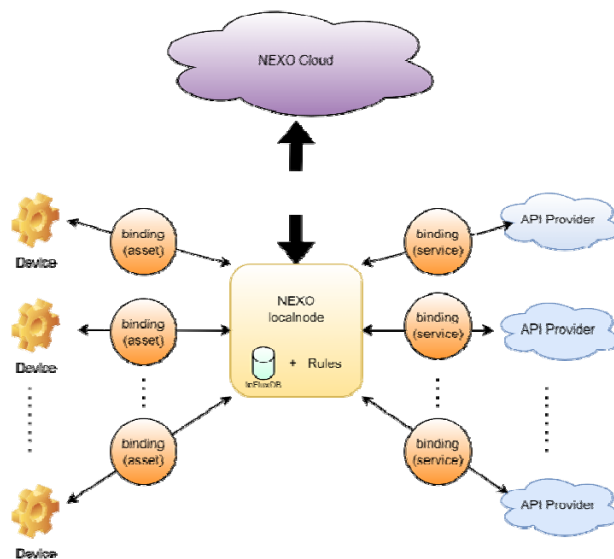
- Węzeł lokalny (*NEXO local node*). Węzeł obliczeniowy oparty na komputerze jednopłytkowym (ang. *single-board computer*), umożliwiający wymianę i przetwarzanie danych pomiędzy chmurą obliczeniową (*NEXO cloud*), internetowymi usługami API (*API providers*) oraz urządzeniami końcowymi. Dodatkowo węzeł lokalny pozwala na utrwalanie danych w lokalnej bazie danych szeregów czasowych InfluxDB oraz implementację reguł pozwalających na zautomatyzowane zarządzanie infrastrukturą budynku.

- Konektor usługi (ang. *binding service*). Specjalizowany moduł oprogramowania umożliwiający komunikację

pomiędzy węzłem lokalnym NEXO, a internetowymi usługami API, takimi jak np. serwisy pogodowe lub obsługa urządzeń. W dalszej części pracy opisano jedną z usług pozwalającą na pozyskiwanie danych o pracy prosumentckiej mikroinstalacji fotowoltaicznej.

- Konektor urządzenia (ang. *binding asset*). Specjalizowany moduł oprogramowania umożliwiający komunikację z zasobami sprzętowymi za pomocą protokołów komunikacyjnych, takich jak MODBUS, MQTT, ZigBee itp.

- Chmura obliczeniowa NEXO Cloud. Oprogramowanie integrujące wiele węzłów lokalnych, pozwalające na zarządzanie wieloma zasobami i usługami np. na poziomie zarządcy budynku wielorodzinnego.



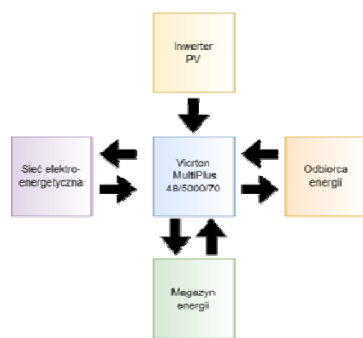
Rys. 1. Architektura platformy NEXO

Akwizycja danych z wykorzystaniem VRM Portal i konektora VICTRON

Mikroinstalacja utworzona z wykorzystaniem przetwornicy VICTRON pozwala na współpracę pomiędzy inwerterem PV, magazynem energii, instalacją odbiorczą oraz siecią elektroenergetyczną. Dzięki zastosowaniu specjalistycznego oprogramowania możliwe jest wdrażanie

optymalnych algorytmów sterowania przepływem energii elektrycznej wewnątrz mikroinstalacji.

W pracy przedstawiono wyniki badań regresyjnych modeli uczenia maszynowego, zastosowanych do prognozowania wielkości produkcji instalacji fotowoltaicznej. Modele uczone na danych uzyskanych z rzeczywistej elektrowni fotowoltaicznej o mocy 33 kWp, na która składały się 132 moduły polikrystaliczne o mocy 250 Wp każdy. Według danych producenta, sprawność elektrowni wynosiła 15,6%. Moduły fotowoltaiczne podłączono do dwóch trójfazowych inwerterów, każdy o nominalnej mocy 17 kW. W obwodzie inwertera nr 1 zainstalowano moduły o mocy 19 kWp, natomiast do inwertera nr 2 przyłączono moduły o łącznej mocy 14 kWp. Obwody wyjściowe z inwertera nr 1 są podłączone do trzech jednofazowych przetwornic Victron Energy tworzących mikroinstalację, w której skład wchodzi również magazyn energii o mocy 25 kW. Dzięki przetwornicom możliwe było sterowanie strumieniem energii elektrycznej pochodzącej z badanej elektrowni fotowoltaicznej w czasie rzeczywistym. Za pomocą ustalonych reguł energia ta może być kierowana bezpośrednio do bieżącego zużycia w zasilanym obiekcie lub przekazywana do magazynu energii. Schemat przepływów energii dla badanej mikroinstalacji zaprezentowano na rysunku 2.



Rys. 2. Schemat blokowy przepływu energii w mikroinstalacji Victron

Elektrownię fotowoltaiczną oraz magazyn energii nadzorują trzy jednofazowe inwertery Victron Energy MultiPlus 48/5000/70. Monitoring i zarządzanie przepływami energii realizowane jest przez oprogramowanie dostarczone wraz z inwerterami. System jest sterowany lokalnie za pomocą panelu Color Control GX. Zdalny dostęp do instalacji jest możliwy z wykorzystaniem oprogramowania Victron Remote Management (VRM) oraz portalu internetowego z dostępnym interfejsem API REST. W ramach procedury badawczej wyniki pomiarów energii wytworzonej przez opisywaną elektrownię fotowoltaiczną były rejestrowane w czasie rzeczywistym przez usługę VRM Portal. W dalszej kolejności pomiary te były pozyskiwane za pośrednictwem konektora Victron VRM binding opisanego w sekcji 1.2 i utrwalane w bazie danych szeregów czasowych platformy NEXO. Ostatecznie, dane odczytane z bazy NEXO posłużyły do uczenia modeli prognostycznych.

Jednym z konektorów usługi dostępnych na platformie NEXO jest konektor Victron API (*VRM binding*), udostępniający dane reprezentujące stan mikroinstalacji prosumenckiej zbudowanej w oparciu o przetwornice firmy Victron i oprogramowanie w chmurze Victron Energy VRM Portal. Struktura konektora VRM binding ściśle odpowiada interfejsowi VICTRON API 2.0. W jej skład wchodzi punkty końcowe (ang. *endpoints*) oraz mechanizmy autentykacji zgodne ze standardem REST API.

Jednym z celów badań przeprowadzonych w ramach niniejszej pracy było opracowanie algorytmu

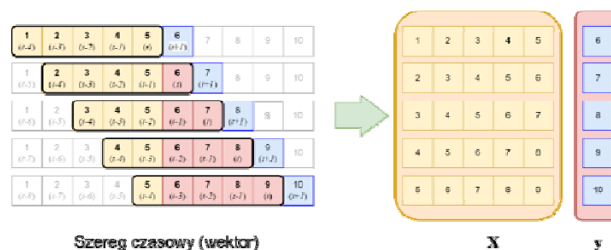
prognozowania produkcji energii fotowoltaicznej. Do realizacji tego zadania zastosowano wybrane modele uczenia maszynowego. Dane do trenowania i testowania modeli pozyskano z wykorzystaniem konektora VICTRON współpracującego z platformą NEXO.

Prognozowanie produkcji energii fotowoltaicznej z wykorzystaniem biblioteki SkForecast

Prognozowanie szeregów czasowych polega na wyznaczeniu przyszłych wartości zmiennej w czasie wielkości na podstawie modelu i wartości tej wielkości zmierzonych w przeszłości [4]. W niniejszej pracy wielkością podlegającą prognozowaniu jest ilość energii elektrycznej produkowanej z wykorzystaniem instalacji fotowoltaicznej będącej jednym z elementów mikroinstalacji prosumenckiej zbudowanej w oparciu o system Victron Energy.

Istnieje wiele metod prognozowania szeregów czasowych używanych do prognozowania produkcji energii fotowoltaicznej, w tym m.in.: *Autoregressive Moving Average (ARMA)*, *Autoregressive Integrated Moving Average (ARIMA)*, *Seasonal Autoregressive Integrated Moving-Average (SARIMA)* [4, 5] lub też model stworzony przez pracowników firmy Facebook, oparty na uogólnionym modelu addytywnym z wykorzystaniem sezonowości i kalendarza zdarzeń [6]. W niniejszej pracy zaproponowano odmienne podejście. Do prognozowania zastosowano znane modele regresyjne, takie jak lasy losowe i regresja z regularyzacją (LASSO) [7]. Dla tak sformułowanego zadania prognozowania niezbędne było przekształcenie szeregu czasowego do formatu typowego dla algorytmów regresyjnych z wykorzystaniem biblioteki SkForecast.

Otwartoźródłowe oprogramowanie Skforecast, stworzone w języku Python pozwala na prognozowanie szeregów czasowych przy użyciu algorytmów regresji wchodzących w skład biblioteki scikit-learn, stanowiącej światowy standard w dziedzinie uczenia maszynowego w języku Python [8, 9]. Podstawowym zadaniem procedur zaimplementowanych w ramach biblioteki SkForecast jest transformacja szeregu czasowego zapisanego w postaci wektora, do formy macierzowej, wymaganej przez algorytm regresji. Schemat przekształcenia przedstawia rysunek 3.



Rys. 3. Schemat przekształcenia szeregu czasowego w postaci wektora do formy macierzowej

Zgodnie ze schematem przedstawionym na rysunku 3, wejściowy szereg czasowy reprezentuje 10-cio elementowy wektor próbek pomiarowych. W trakcie działania algorytmu transformacji wartości elementów wektora przepisywane są do kolejnych wierszy macierzy wejściowej danych uczących X oraz wektora wyjściowych danych uczących y . W wierszach macierzy X znajdują się więc kolejne wielowymiarowe obserwacje zapisane w kolumnach cech. Cechy te reprezentują opóźnione wartości szeregu czasowego. Działanie algorytmu można zobrazować poprzez przesuwanie okna czasowego o stałej długości po wektorze wejściowym i przepisywanie wartości próbek szeregu od $(t - n - 1)$ do t do macierzy X , przy czym przez n

oznaczono liczbę próbek okna (np. dla wektora z rysunku 3, $n = 5$). Dodatkowo, w każdym kroku algorytmu próbka ($t + 1$), traktowana jako bieżąca prognoza, jest zapisywana do wektora kolumnowego reprezentującego docelowe wartości wyjściowe. Dzięki takiemu przekształceniu danych, możliwe jest prognozowanie szeregów czasowych za pomocą regresyjnych algorytmów uczenia maszynowego.

Regresja jest jednym z rodzajów uczenia nadzorowanego pozwalającym na modelowanie zależności wejście-wyjście w przypadku ciągłej zmiennej objaśnianej. Model regresyjny może być więc traktowany jak funkcja aproksymująca wielowymiarową zależność pomiędzy zmiennymi objaśniającymi (wejścia modelu) a zmienną objaśnianą (wyjście modelu). W niniejszej pracy problem prognozowania produkcji energii przedstawiono za pomocą schematu regresyjnego zbudowanego w oparciu o bibliotekę SkForecast. Do rozwiązania problemu zastosowano model regresji regularyzowanej LASSO oraz model lasu losowego.

Model regresji LASSO to model liniowy, pozwalający na oszacowanie współczynników regresji przy jednoczesnym karaniu (ograniczaniu) ich bezwzględnych wartości. Z tego powodu model ten preferuje rozwiązania o mniejszej liczbie niezerowych współczynników, skutecznie zmniejszając liczbę cech, od których dane rozwiązanie jest zależne. Model LASSO minimalizuje następującą funkcję celu [7]:

$$(1) \quad \min_{\omega} \frac{1}{2n_s} \|X\omega - y\|_2^2 + \alpha \|\omega\|_1,$$

gdzie: n_s – liczba próbek uczących, X – macierz danych uczących o rozmiarze liczba obserwacji \times liczba cech, y – wektor wyjściowych danych uczących, $\|\omega\|_1$ – norma l_1 wektora parametrów ω .

W przypadku regresyjnym model losowego lasu decyzyjnego pozwala na budowę zbioru drzew decyzyjnych, dla którego następnie wyznacza się przewidywaną średnią tego zbioru [10]. Podstawowym składnikiem lasu losowego jest drzewo decyzyjne. Pozwala ono na zapis decyzji oraz związanych z nimi prawdopodobieństw odnoszących się do zadania klasyfikacji lub regresji. Jednym z algorytmów budowy drzewa decyzyjnego, stosowanym w uczeniu maszynowym jest algorytm CART (ang. *classification and regression tree*). W pierwszym kroku rozdziela on dane uczące na podzbiory: lewy i prawy minimalizując następującą funkcję kosztu MSE:

$$(2) \quad J(k, t_k) = \frac{m_l}{m} MSE_l + \frac{m_r}{m} MSE_r,$$

gdzie: m_l – licznosc zbioru lewego, m_r – licznosc zbioru prawego, m – całkowita liczba próbek uczących, natomiast błąd MSE_{node} w wybranym węzle jest zdefiniowany następująco:

$$(3) \quad MSE_{node} = \sum_{i \in node} (\hat{y}_{node} - y^{(i)})^2,$$

$$(4) \quad \hat{y}_{node} = \frac{I}{m_{node}} \sum_{i \in node} y^{(i)}$$

Kolejne kroki algorytmu prowadzą do dalszych podziałów zbioru uczącego. Liczba kolejnych podziałów (głębokość drzewa) może być kontrolowana za pomocą jednego z hiperparametrów modelu. Algorytmy uczenia maszynowego działające w oparciu o równania (1) – (4) zostały użyte do budowy i trenowania modeli prognozujących produkcję energii fotowoltaicznej. Wyniki badań modeli opisano w dalszej części pracy.

Wyniki badań

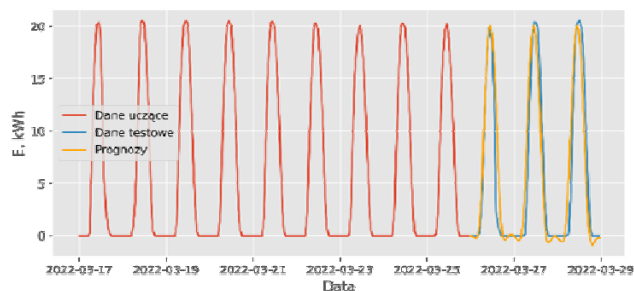
W badaniach zastosowano dwa rodzaje modeli uczenia maszynowego tj. regularyzowaną regresję liniową LASSO oraz lasy losowe. Do badań użyto szeregów czasowych reprezentujących produkcję energii elektrowni fotowoltaicznej w marcu i maju 2022 roku, pozyskanych z użyciem platformy NEXO i konektora VICTRON. W pierwszym etapie badań przeprowadzono czyszczenie danych, w tym uzupełnianie danych brakujących, a także usunięcie wartości odstających. Przetwarzanie wstępne przeprowadzono z użyciem bibliotek *Numpy* i *Pandas* języka *Python*. Kolejnym etapem był podział szeregów czasowych na zbiory uczące i testowe. W przypadku modelu LASSO dodatkowo zastosowano standaryzację danych uczących, zgodnie z następującą formułą skalującą:

$$(5) \quad z = \frac{x - \mu}{\sigma},$$

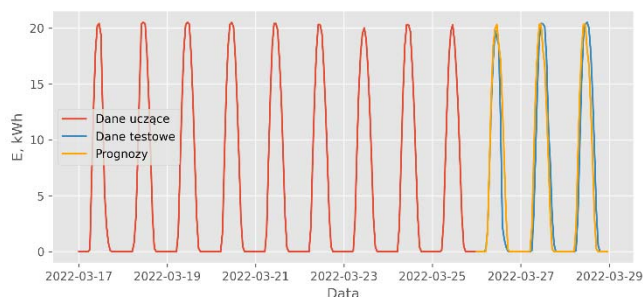
gdzie: x – pojedyncza wartość próbki uczącej, μ – wartość średnia zbioru uczącego, σ – odchylenie standardowe zbioru uczącego.

W dalszej kolejności zrealizowano trening modeli uczenia maszynowego i strojenie hiperparametrów. W trakcie strojenia, dla obu z testowanych typów modeli utworzono siatkę parametrów oraz siatkę opóźnień. Dla każdej kombinacji wartości parametrów i liczby opóźnień uczono i ewaluowano model za pomocą procedury testowania wstecznego (ang. *backtesting*) przy użyciu metryk *MSE* (ang. *mean squared error*) i *MAE* (ang. *mean absolute error*). W przypadku modelu LASSO przyjęto dwie wartości liczby opóźnień: $l = 24$, $l = 36$ oraz 100 wartości hiperparametru α w przedziale: $(10^{-3}, 10^3)$, dzięki czemu utworzono 200 modeli. Dla modelu lasu losowego przyjęto $l = 24$ i $l = 36$ oraz cztery wartości liczby drzew ($n_{estimators} = [2, 8, 20, 50]$) i pięć wartości maksymalnej głębokości drzewa ($max_{depth} = [2, 4, 6, 8, 10]$), co pozwoliło na utworzenie 60 modeli.

Wyniki prognozowania produkcji energii w wybranym tygodniu marca 2022 r. zaprezentowano na rysunkach 4 i 5.

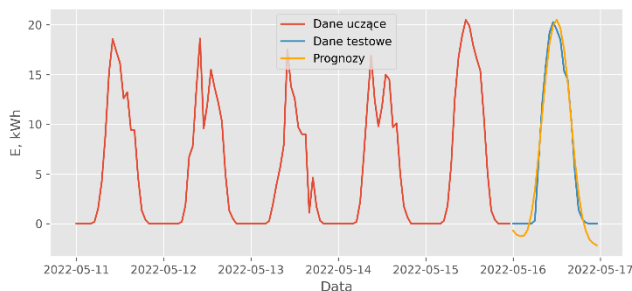


Rys. 4. Wyniki prognozowania produkcji energii (marzec, model LASSO, $\alpha = 0.001$, $l = 36$)

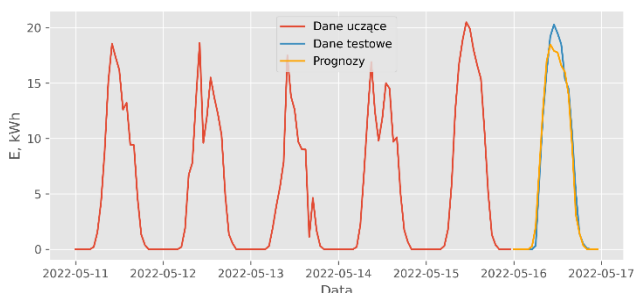


Rys. 5. Wyniki prognozowania produkcji energii (marzec, las losowy, $max_{depth} = 8$, $n_{estimators} = 2$, $l = 36$).

Prognozy uzyskano przy użyciu modelu LASSO (rys. 4) oraz modelu lasu losowego (rys. 5). Na rysunkach 6 i 7 zaprezentowano wyniki prognozowania produkcji energii w wybranym tygodniu maja 2022 r. Prognozy uzyskane przy użyciu modelu LASSO przedstawiono na rysunku 6 natomiast na rysunku 7 pokazano wyniki dla modelu lasu losowego.



Rys. 6. Wyniki prognozowania produkcji energii (maj, model LASSO, $\alpha = 0.014$, $l = 36$)



Rys. 7. Wyniki prognozowania produkcji energii (maj, las losowy, $max_{depth} = 6$, $n_{estimators} = 8$, $l = 36$)

Wartości miar MSE i MAE uzyskane na danych testowych na etapie ewaluacji modeli dla trzech wybranych zestawów hiperparametrów zestawiono w tabelach 1, 2. Dodatkowo w tabelach przedstawiono wartości miary MSE_{bck} obliczonej w procesie testowania wstecznego. W tabeli 1 zaprezentowano wyniki ewaluacji modeli typu LASSO, natomiast wyniki dla modeli typu las losowy przedstawiono w tabeli 2.

Tabela 1. Wybrane wyniki ewaluacji modeli typu Lasso

Miesiąc	Hiperparametry	MSE	MAE	MSE_{bck}
Marzec	$\alpha = 0,001$, $l = 36$	7,7	1,7	0,8
	$\alpha = 0,335$, $l = 24$	23,9	4,2	19,6
	$\alpha = 100,0$, $l = 36$	60,3	7,0	62,5
Maj	$\alpha = 0,012$, $l = 36$	1,8	1,2	13,1
	$\alpha = 12,04$, $l = 24$	62,6	6,9	51,2
	$\alpha = 1000$, $l = 36$	62,6	6,9	51,3

Wnioski

W pracy przeprowadzono badania mające na celu uzyskanie prognoz produkcji energii elektrycznej dla elektrowni fotowoltaicznej. Do prognozowania użyto modeli LASSO i lasu losowego, przy czym modele trenowano na danych pozyskanych z wykorzystaniem platformy NEXO. W trakcie analizy wyników badań stwierdzono m.in., że dokładność obu badanych modeli ściśle zależy od liczby opóźnień l . Dodatkowo stwierdzono, że w przypadku modelu LASSO duży wpływ na dokładność ma współczynnik kary α , natomiast w przypadku lasu losowego istotną rolę grają odpowiednio liczba drzew i maksymalna głębokość drzewa. Podsumowując, należy stwierdzić, że zastosowanie regresji LASSO pozwala na uzyskanie podobnej dokładności prognozowania jak w przypadku

klasycznych metod, takich jak SARIMA czy też uogólniony model addytywny.

Tabela 2. Wybrane wyniki ewaluacji modeli typu las losowy

Miesiąc	Hiperparametry	MSE	MAE	MSE_{bck}
Marzec	$max_{depth} = 8$, $n_{estimators} = 2$, $l = 36$	7,6	1,6	0,096
	$max_{depth} = 4$, $n_{estimators} = 6$, $l = 24$	6,9	1,5	0,7
	$max_{depth} = 2$, $n_{estimators} = 6$, $l = 36$	8,3	1,6	3,7
Maj	$max_{depth} = 6$, $n_{estimators} = 8$, $l = 36$	1,0	0,7	7,2
	$max_{depth} = 6$, $n_{estimators} = 20$, $l = 24$	4,9	1,4	9,6
	$max_{depth} = 8$, $n_{estimators} = 6$, $l = 24$	3,04	1,0	10,9

Zdaniem autorów, bardzo wysoką dokładność prognoz w porównaniu do modeli klasycznych uzyskuje się z wykorzystaniem lasów losowych, przy czym niezbędne są dalsze badania dotyczące w szczególności właściwego doboru cech i metod wstępnego przetwarzania danych uczących.

Artykuł powstał w ramach działań mających na celu rozpowszechnienie rezultatów osiągniętych w ramach projektu POIR.01.01.01-00-0767/17 „SmartX - platforma do integracji rozwiązań inteligentnego budynku” współfinansowanego ze środków Europejskiego Funduszu Rozwoju Regionalnego.



Fundusze Europejskie
Inteligentny Rozwój



Rzeczpospolita
Polska

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego



Autorzy: dr hab. inż. Sebastian Dudzik, Politechnika Częstochowska, Wydział Elektryczny, ul. Armii Krajowej 17, 42-200 Częstochowa, E-mail: sebastian.dudzik@pcz.pl; dr inż. Bartosz Kowalczyk, ConnectPoint sp. z o.o. ul. Zagadki 21, 02-227 Warszawa, E-mail: bartosz.kowalczyk@connectpoint.pl.

LITERATURA

- [1] Majid, M.; Habib, S.; Javed, A.R.; Rizwan, M.; Srivastava, G.; Gadekallu, T.R.; Lin, J.C.-W. *Applications of Wireless Sensor Networks and Internet of Things Frameworks in the Industry Revolution 4.0: A Systematic Literature Review*. Sensors 2022, 22, 2087. <https://doi.org/10.3390/s22062087>
- [2] Miller M., *Internet rzeczy. Jak inteligentne telewizory, samochody, domy i miasta zmieniają świat*, Wydawnictwa Naukowe PWN, Warszawa 2016
- [3] Saranya T., Deisy C., Sridevi S., Kalaiarasi Sonai Muthu Anbananthen, *A comparative study of deep learning and Internet of Things for precision agriculture*, Engineering Applications of Artificial Intelligence, Volume 122, 2023, 106034, ISSN 0952-1976
- [4] Box G. E. P., Jenkins G. M., Reinsel G. C., Ljung, G. M., *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Hoboken, New Jersey, 2015
- [5] Nespoli A., Ogliari E., Leva S., Massi Pavan A., Mellit A., Lughì V., Dolara A., *Day-Ahead Photovoltaic Forecasting: A Comparison of the Most Effective Techniques*, *Energies*, 12, 1621, 2019
- [6] Harvey, A., Peters, S., *Estimation procedures for structural time series models*, *Journal of Forecasting* 1990, 9, pp. 89–108
- [7] Kim S. J., Koh K., Lustig M., Boyd S., Gorinevsky D., *An Interior-Point Method for Large-Scale L1-Regularized Least Squares*, *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007
- [8] Strona internetowa biblioteki scikit-learn. <https://scikit-learn.org/>
- [9] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., *Scikit-learn: Machine Learning in Python*, *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011
- [10] Breiman, L., *Random Forests*. *Machine Learning*, vol. 45, pp. 5–32, 2001. <http://dx.doi.org/10.1023/A:1010933404324>