

Politechnika Warszawska

W Y D Z I A Ł M E C H A T R O N I K I



Praca dyplomowa inżynierska

na kierunku Automatyka, Robotyka i Informatyka Przemysłowa

Uczenie ciągłe w modelowaniu cen energii elektrycznej

numer pracy według wydziałowej ewidencji prac: 114C-ISP-IP/321116/1251641

Mikołaj Kalisz

numer albumu 321116

promotor

dr inż. Anna Szyber-Betley

konsultacje

—

WARSZAWA 2025

Politechnika Warszawska

W Y D Z I A Ł M E C H A T R O N I K I



Praca dyplomowa inżynierska

na kierunku Automatyka, Robotyka i Informatyka Przemysłowa
w specjalności Informatyka Przemysłowa

Uczenie ciągłe w modelowaniu cen energii elektrycznej

numer pracy według wydziałowej ewidencji prac: 114C-ISP-IP/321255/1278954

Adam Mantiuk

numer albumu 321255

promotor
dr inż. Anna Sztyber-Betley

konsultacje

—

WARSZAWA 2025

PRACA DYPLOMOWA inżynierska

Kierunek studiów: Automatyka, Robotyka i Informatyka Przemysłowa

Specjalność: Informatyka Przemysłowa

Instytut dyplomujący: Instytut Automatyki i Robotyki

Temat pracy: Uczenie ciągłe w modelowaniu cen energii elektrycznej

Zakres pracy:

1. Analiza poszczególnych czynników wpływających na cenę energii elektrycznej
2. Implementacja sieci neuronowych, dobór hiperparametrów
3. Ocena przydatności opracowanych modeli

Literatura:

1. D. Mrowiec, „Analiza merit order sektora wytwórczego w Polsce do 2030 roku,” PSE Innowacje.
2. S. Bhanja i A. Das, „Impact of Data Normalization on Deep Neural Network for Time Series Forecasting”.

Słowa kluczowe: uczenie ciągłe, sieci neuronowe, predykcja, ceny energii

Praca dyplomowa jest realizowana we współpracy z przemysłem

Tak/Nie *

Nazwa firmy:

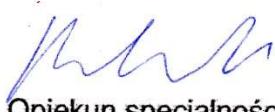
<i>Imię i nazwisko dyplomanta:</i> Mikołaj Kalisz	<i>Imię i nazwisko promotora:</i> dr inż. Anna Sztyber-Betley
<i>Imię i nazwisko konsultanta:</i>	

<i>Temat wydano dnia:</i> 13.05.24r.	<i>Termin ukończenia pracy:</i> 29.01.2025r.
--	--

Miejsce wykonywania praktyki przeddyplomowej:

GOLD LIFT Rafał Chmisi

Zatwierdzenie tematu

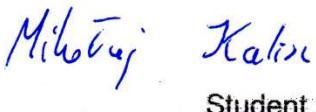


Opiekun specjalności



Z-ca Dyrektora Instytutu

Przyjęcie zakresu pracy



Student



Promotor

PRACA DYPLOMOWA inżynierska

Kierunek studiów: Automatyka, Robotyka i Informatyka Przemysłowa

Specjalność: Informatyka Przemysłowa

Instytut dyplomujący: Instytut Automatyki i Robotyki

Temat pracy: Uczenie ciągłe w modelowaniu cen energii elektrycznej

Zakres pracy:

1. Analiza możliwości wykorzystania uczenia ciągłego w modelowaniu cen energii elektrycznej
2. Stworzenie modelu do predykcji cen opartego o uczenie ciągłe
3. Analiza dokładności modeli z wykorzystaniem uczenia ciągłego względem podejścia klasycznego

Literatura:

1. Z. Chen and B. Liu, Lifelong Machine Learning, Second Edition, Morgan & Claypool, 2018.
2. D. Pyle, Data Preparation for Data Mining, San Francisco: Morgan Kaufmann Publishers, Inc., 1999.

Słowa kluczowe: uczenie ciągłe, sieci neuronowe, predykcja, ceny energii

Praca dyplomowa jest realizowana we współpracy z przemysłem

Tak/Nie *,

Nazwa firmy:

Imię i nazwisko dyplomanta:

Adam Mantiuk

Imię i nazwisko promotora:

dr inż. Anna Szyber-Betley

Temat wydano dnia:

13.05.2024r.

Termin ukończenia pracy:

29.01.2025r.

Miejsce wykonywania praktyki przeddyplomowej:

Transition Technologies – Advanced Solutions Sp. z o.o.

Zatwierdzenie tematu

Opiekun specjalności

Z-ca Dyrektora Instytutu

Przyjęcie zakresu pracy

Student

Promotor

*Niepotrzebne skreślić

Streszczenie

Temat pracy: Uczenie ciągłe w modelowaniu cen energii elektrycznej

Słowa kluczowe: uczenie ciągłe, sieci neuronowe, predykcja, ceny energii

Poniższa praca dotyczy tematyki predykcji cen energii elektrycznej na rynku dnia następnego. Poruszone są kwestie dynamicznej zmienności rynku, jak i globalnych zmian w polityce energetycznej wpływających na produkcję, a co za tym idzie cenę energii na rynku.

W części teoretycznej przedstawiono podstawowe zagadnienia związane z rynkiem energii elektrycznej, w tym istniejące formy transakcji na tym rynku. Opisano sposób wyznaczania ceny energii na rynku dnia następnego i związane z tym konsekwencje. Opisano również strukturę rynku, rozkład mocy na rodzaje jednostek produkcyjnych. Poruszono też podstawowe zagadnienia związane ze sztucznymi sieciami neuronowymi. Zostały omówione kwestie takie jak: architektura sieci, stosowane w pracy techniki oraz istotność odpowiedniego przygotowania danych. Została omówiona idea uczenia ciągłego, oraz najważniejsze wyzwania jakie są stawiane przed tym rodzajem uczenia maszynowego.

W pracy przedstawiono proces przygotowania modeli predykcyjnych bazujących na sieciach neuronowych trenowanych w sposób klasyczny oraz stosując mechanizmy uczenia ciągłego. Pokazany jest proces od wstępnego przygotowania danych, poprzez dokładną analizę czynników wpływających na cenę energii elektrycznej, a kończąc na stworzeniu sieci neuronowych. W pracy porównano wyniki obu podejść.

Otrzymane wyniki wskazują na istniejący potencjał zastosowania uczenia ciągłego do predykcji szeregów czasowych.

Abstract

Thesis topic: Continual Learning in Electricity Price Forecasting

Keywords: continual learning, neural networks, prediction, energy prices

The following study focuses on the prediction of electricity prices on the day-ahead market. It addresses issues related to the dynamic volatility of the market, as well as global changes in energy policy that affect production and, consequently, the price of electricity on the market.

The theoretical section presents fundamental concepts related to the electricity market, including existing forms of transactions. It describes the process of determining the electricity price on the day-ahead market and its associated consequences. The market structure and the distribution of power among different types of production units are also discussed. Additionally, basic concepts related to artificial neural networks are introduced, covering topics such as network architecture, techniques used in the study, and the importance of proper data preparation. The concept of continual learning is explored, along with the key challenges posed by this type of machine learning.

The study presents the process of preparing predictive models based on neural networks trained using both traditional methods and continual learning mechanisms. It details the process from the initial data preparation, through an in-depth analysis of factors influencing electricity prices, to the development of neural networks. A comparison of the results obtained using both approaches is provided.

The obtained results indicate the potential for applying continual learning to time series forecasting.

Spis treści

1	Wstęp.....	17
1.1	Opis problemu	17
1.2	Cel i zakres pracy	17
2	Wprowadzenie.....	19
2.1	Rynek energii elektrycznej.....	19
2.1.1	Cechy rynku energii elektrycznej	19
2.1.2	Formy transakcji na rynku energii elektrycznej	20
2.1.3	Cena energii na rynku dnia następnego	21
2.1.4	Czynniki wpływające na cenę energii.....	22
2.2	Sztuczne sieci neuronowe (SSN)	23
2.2.1	Definicja i podstawy SSN	23
2.2.2	Neuron.....	23
2.2.3	Struktura sieci – perceptron	24
2.2.4	Funkcje aktywacji	25
2.2.5	Normalizacja i standaryzacja danych.....	31
2.2.6	Podział danych	33
2.2.7	Uczenie sieci	34
2.2.8	Regularizacja – dropout	35
2.2.9	Zastosowanie SSN w predykcji szeregów czasowych.....	36
2.3	Uczenie Ciągłe	37
2.3.1	Transfer i adaptacja	39
2.3.2	Katastroficzne zapominanie	40
2.3.3	Wykorzystanie podobieństwa zadań	40
2.3.4	Uczenie niezależnie od zadania	41
2.3.5	Odporność na szum	41
2.3.6	Efektywność zasobowa.....	42
2.3.7	Perspektywa biologiczna	42
2.4	Zastosowanie Uczenia Ciągłego w prognozowaniu i podejmowaniu decyzji na rynku energii elektrycznej	43
2.5	Metryki oceny modeli	45
2.5.1	R^2 – Współczynnik determinacji	46
2.5.2	MSE – średni błąd kwadratowy.....	47
2.5.3	RMSE – pierwiastek średniego błędu kwadratowego	48
2.5.4	MAE – średni błąd bezwzględny	49
2.5.5	MAPE – średni procentowy błąd bezwzględny	49

2.5.6 SMAPE – symetryczny średni procentowy błąd bezwzględny	50
3 Eksploracja danych	52
3.1 Identyfikacja problemu.....	52
3.2 Wstępne przetwarzanie danych.....	53
3.2.1 Znalezienie i identyfikacja danych	53
3.2.2 Agregacja	53
3.2.3 Braki danych	54
3.3 Badania eksploracyjne (EDA)	55
3.3.1 Analiza wstępna danych o cenach energii.....	55
4 Analiza czynników wpływających na ceny energii elektrycznej	65
4.1 Zapotrzebowanie na energię.....	65
4.1.1 Wpływ temperatury na zapotrzebowanie i ceny energii.....	70
4.2 Odnawialne źródła energii	72
4.2.1 Produkcja energii wiatrowej	73
4.2.2 Produkcja energii słonecznej.....	75
4.2.3 Produkcja z innych źródeł.....	78
4.3 Emisyjne źródła energii.....	81
4.3.1 Węgiel.....	81
4.3.2 Gaz.....	85
4.3.3 Ropa	88
4.3.4 Biomasa	89
4.4 Czynniki ekonomiczne i polityczne	90
4.4.1 Opłata za emisję CO₂	90
4.4.2 Wpływ handlu międzynarodowego.....	91
4.4.3 Wpływ ekonomii.....	91
4.5 Inżynieria cech.....	93
4.5.1 Wpływ historycznych cen i wolumenu.....	93
5 Modelowanie cen energii elektrycznej z wykorzystaniem sztucznych sieci neuronowych	94
5.1 Dane.....	94
5.1.1 Wybór cech	94
5.1.2 Podział danych	95
5.1.3 Skalowanie danych	95
5.2 Model	96
5.2.1 Architektura	96
5.2.2 Optymalizator	97

5.2.3	Dobór hiper-parametrów.....	97
6	Modelowanie cen energii elektrycznej z wykorzystaniem uczenia ciągłego	99
6.1	Projektowanie modelu.....	99
6.2	Przegląd strategii zastosowanych w modelowaniu.....	99
6.3	Zastosowane metody przeciwdziałające katastroficzemu zapominaniu	102
6.3.1	Odtwarzanie epizodów.....	103
6.3.2	Elastyczna Konsolidacja Wag	103
7	Ocena modeli oraz porównanie wyników	105
7.1	Wyniki modeli klasycznych	105
7.2	Wyniki modeli z Uczeniem Ciągłym	109
8	Podsumowanie i wnioski	115
8.1	Wnioski z porównania technik uczenia.....	115
8.2	Zalecenia dla przyszłych badań	115
	Bibliografia	116

1 Wstęp

1.1 Opis problemu

Autor: Adam Mantiuk

Modelowanie cen energii elektrycznej nie jest zadaniem prostym. Sytuacja na rynku zawsze jest dynamiczna, a zatem ceny ulegają częstym zmianom. Niektóre wydarzenia takie jak pandemie, wojny, czy katastrofy potrafią zachwiać rynkiem na dłuższy okres czasu. Tradycyjne metody predykacji cen nie są na bieżąco aktualizowane, co w znaczny sposób obniża ich skuteczność. W odpowiedzi na ten problem coraz większą uwagę kładzie się na metody uczenia ciągłego (ang. continual learning). Ten rodzaj uczenia pozwala na ciągła adaptację modeli do zmieniających się warunków środowiskowych, dzięki aktualizacjom opartym o nowe dane. Przed tymi modelami stoją natomiast kolejne wyzwania, a jednym z nich jest przezwyciężenie katastroficznego zapominania, a więc tendencji modeli do zapominania wcześniej zdobytej wiedzy.

1.2 Cel i zakres pracy

Autor: Mikołaj Kalisz

Celem pracy jest ocena skuteczności zastosowania uczenia ciągłego w porównaniu z klasycznym podejściem do trenowania modelu predykcyjnego. W ramach badań zostanie przeprowadzone studium przypadku dotyczące prognozowania cen energii, w którym obie techniki zostaną zastosowane i porównane pod względem ich efektywności.

Praca obejmie proces przygotowania danych, w tym ich oczyszczenie poprzez usunięcie cech zawierających brakujące wartości, które mogłyby negatywnie wpływać na jakość modelu. Następnie zostanie przeprowadzona szczegółowa analiza czynników wpływających na cenę energii, uwzględniająca zmienność rynku dnia następnego oraz analizę korelacji pomiędzy wybranymi zmiennymi.

Na podstawie przetworzonych danych zostaną wytrenowane i ocenione modele predykcyjne o różnych architekturach. Wybrana architektura modelu zostanie następnie wykorzystana w scenariuszu uczenia ciągłego, a jego skuteczność zostanie porównana z tradycyjnym modelem statycznym. Analiza porównawcza obejmie metryki jakości predykcji oraz stabilność działania modeli w zmieniających się warunkach rynkowych.

W ramach projektu zostaną przedstawione wnioski z przeprowadzonej analizy oraz rekomendacje dotyczące przyszłych badań nad zastosowaniem uczenia ciągłego w predykciach cen energii.

Oprogramowanie stworzone w ramach niniejszej pracy inżynierskiej zamieszczono w repozytorium na GitHub¹.

¹ Źródło: [GitHub]
(https://github.com/KaliszM/Praca_inz_Uczenie_Ciagle_W_Modelowaniu_Ceny_Energii)

2 Wprowadzenie

2.1 Rynek energii elektrycznej

2.1.1 Cechy rynku energii elektrycznej

Autor: Adam Mantiuk

Przed transformacją ustroju gospodarczego na początku lat 90. energia elektryczna była dobrem publicznym. Współcześnie energia elektryczna jest towarem na rynku a jej wytwarzanie jest usługą. Rynek ten przez ostatnie trzy dekady był wielokrotnie regulowany. Nie zachodzą tutaj mechanizmy wolnego rynku, pomimo jego liberalizacji po zmianie ustroju państwa. Współcześnie produkcja energii elektrycznej musi mierzyć się z nowymi regulacjami Unii Europejskiej. Również cała Europa Środkowo-Wschodnia, która przez wiele lat opierała swoje bezpieczeństwo energetyczne o paliwa kopalne, będzie musiała przejść transformację [1].

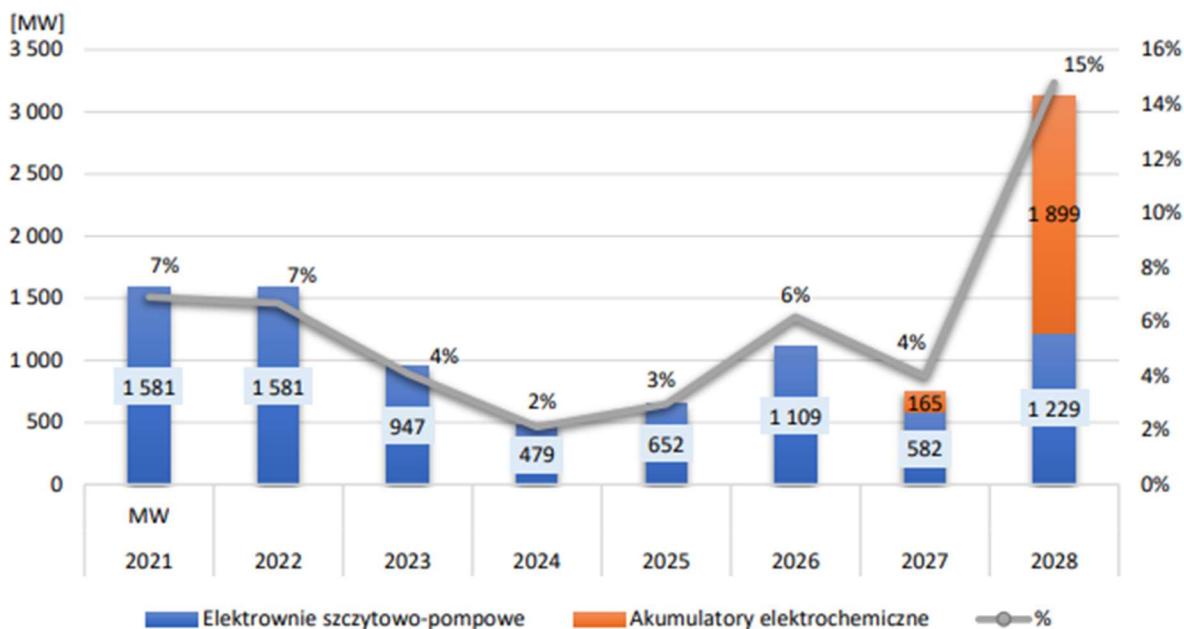
Rynek energii elektrycznej charakteryzuje się pewnymi cechami, które to nadają mu nietypową pozycję zarówno w polskim, jak i w globalnym systemie ekonomicznym. Obecnie nie istnieje odpowiednia infrastruktura pozwalająca na magazynowanie energii elektrycznej. Sprawia to, że rynek energii elektrycznej jest rynkiem czasu rzeczywistego. Nadmiarowa ilość energii elektrycznej (na przykład wyprodukowanej przez odnawialne źródła energii) musi zostać wykorzystana, lub podmioty je produkujące zmuszone są odłączyć swoje urządzenia z sieci. Z drugiej strony elektrownie zasilane paliwami kopalnymi nie mogą zostać odłączone ze względu na nierentowność ich ponownego uruchomienia w przypadku braku wystarczającej podaży z innych źródeł. Przemysł energetyczny wymaga też dalekosieżnych planów inwestycyjnych, a nowi uczestnicy rynku muszą zainwestować znaczne zasoby pieniężne. Spowodowane jest to faktem, iż nie tylko produkcja, ale też transport wymaga pewnej infrastruktury. Stosunkowo niskie koszty dostarczenia energii (oraz rozbudowy istniejącej infrastruktury) prowadzą do kształtowania się naturalnych monopolii [2].

Regulacje na rynku energii elektrycznej sprawiły, że z firm, które postrzegane były za monopole naturalne, wydzielono podmioty odpowiedzialne między innymi za produkcję, czy przesył energii elektrycznej. Regulacje prawne obligują podmioty uczestniczące w tym rynku do przestrzegania pewnych zasad. Zasady te zapewniają między innymi [2]:

- Równoprawny dostęp do sieci przemysłowych
- Prawo wyboru dostawcy energii
- Kontrolę nad nadużyciami siły rynkowej

Europa stawia obecnie na odnawialne źródła energii, które nie są w stanie zapewnić stabilności energetycznej. Spowodowane jest to zmiennymi warunkami pogodowymi, od których zależy zdolność produkcyjna większości jednostek OZE. Na poprawę współpracy sieci dystrybucyjnej z jednostkami OZE mogą w znaczącym stopniu wpływać magazyny energii [1] [3].

Na dzień dzisiejszy magazyny szczytowo-pompowe są jedynymi istniejącymi magazynami energii elektrycznej w Polsce. Jak można zauważyć na rysunku 1, sytuacja ta ma się zmienić w latach 2027-2028, kiedy to na rynek wejdą akumulatory elektrochemiczne [3]. „W wyniku aukcji głównej rynku mocy na rok dostaw 2028, kontrakty mocowe zostały zawarte przez 33 magazyny energii elektrycznej wykorzystujące technologię elektrochemiczną. [...] Sumaryczna wielkość obowiązków mocowych wynikająca z zawartych umów mocowych w wyniku aukcji głównej na rok dostaw 2028 wynosiła 7 070,951 MW. Tym samym zakontraktowana w toku aukcji moc magazynów energii stanowiła ok. 25 proc. sumarycznej wielkości obowiązków mocowych zawartych w toku aukcji głównej.” [3]



Rysunek 1 Moc zakontraktowana przez magazyny energii w aukcjach głównych na lata 2021–2028
Opracowanie pochodzi z [3]

Taka ilość nowych magazynów energii elektrycznej może zachwiać obecnym systemem, który tworzony był przez wiele lat. Magazynowanie energii wyprodukowanej przez odnawialne źródła energii może w znaczącym stopniu spowodować, iż obecne analizy wpływu poszczególnych czynników staną się nieaktualne. Na obecną chwilę nie można jednocześnie stwierdzić, czy cena energii dzięki magazynom będzie bardziej stabilna, czy też właściciele takich magazynów będą w stanie skutecznie manipulować cenami, aby zmaksymalizować swoje zyski.

2.1.2 Formy transakcji na rynku energii elektrycznej

Autor: Adam Mantiuk

Rynek energii elektrycznej umożliwia różnorodne formy transakcji. W Polsce istnieje między innymi rynek giełdowy, w którego skład wchodzi rynek dnia następnego (RDN). Uczestnicy rynku operują na poszczególnych godzinach. Sprzedaż

prowadzona jest dzień przed dobą, w której następuje dostawa energii do odbiorcy. Jako że rynek dnia następnego ma swoje odzwierciedlenie fizyczne (zawarcie transakcji oznacza przepływ energii elektrycznej) to w rynku tym uczestniczą firmy produkujące oraz sprzedające energię elektryczną. Ważną funkcją tej giełdy jest kształtowanie cen kontraktowych oraz wycena firm z branży energetycznej. To tutaj powstają odpowiednie rekomendacje rynkowe odnośnie budowy nowej infrastruktury energetycznej. Oprócz rynku dnia następnego na rynku giełdowym można wyróżnić między innymi rynek terminowy towarowy (RTT) czy rynek dnia bieżącego (RDB). Rynek terminowy towarowy pozwala na zakup energii, która zostanie dostarczona do odbiorcy w przyszłości po cenie określonej w momencie zakupu. Kontrakty na rynku dnia bieżącego pozwalają z kolei na modyfikację zobowiązań handlowych w trakcie dnia dostaw jak i na dzień przed jego rozpoczęciem [2].

Większość krajowego zapotrzebowania na moc zawierana jest w długoterminowych kontraktach. Ten rodzaj transakcji odbywa się na rynku kontraktowym. Udział w nim biorą firmy produkujące energię elektryczną oraz podmioty eksploatujące tą energię [2].

Można wyróżnić również rynek odpowiadający za zachowanie równowagi popytu i podaży energii w systemie elektroenergetycznym. Jest nim rynek bilansujący, na którym odbiorca ma zakupienia brakującej energii powyżej ceny rynkowej, jak i odsprzedania jej nadwyżek poniżej tej ceny. Niekorzystne ceny motywują zatem odbiorców do precyzyjnego przewidywania zapotrzebowania. Warto zaznaczyć, że udział podmiotów kupujących energię (czyli odbiorców) jest tutaj obowiązkowy, a sam rynek nie jest miejscem handlu co czyni go tak zwanym rynkiem technicznym [2].

2.1.3 Cena energii na rynku dnia następnego

Autor: Mikołaj Kalisz

Rynkiem energii elektrycznej dnia następnego, jak każdym rynkiem, rządzi popyt i podaż. Pierwszym etapem ustalenia cen energii jest zaoferowanie ceny przez sprzedawców energii, na podstawie własnych szacunków co do jej produkcji następnego dnia. Następnie uczestnicy rynku mają możliwość składania zleceń kupna lub sprzedaży energii elektrycznej dla poszczególnych godzin. Dane o zleceniach są zbierane i przekształcane w krzywe popytu i podaży. Korzystając z metody „merit order” wyznaczana jest cena kontraktów, jednolita dla każdej godziny [4].

„Merit order” to inaczej mówiąc stos jednostek wytwórczych, uporządkowany od najniższej do najwyższej ceny oferowanej energii. W ten sposób przedstawione informacje o zleceniach kupna mówią nam o istniejącej podaży na rynku. Jednostki produkcyjne o najniższych kosztach wytwórczych oferują najniższą cenę energii, tym samym są pierwsze w kolejce do zapełnienia zapotrzebowania. Jednostki o najwyższych kosztach oferujące najwyższe stawki za MWh energii znajdują się na końcu stosu, to znaczy energia od tych producentów zostanie wykorzystana na

ostatnim miejscu. Po zestawieniu podaży z popytem, w tym wypadku z zapotrzebowaniem na energię, wyznaczana jest cena jednolita za energię elektryczną na daną godzinę. Odczytuje się ją z miejsca przecięcia obu krzywych. Jednostki produkujące energię, których oferta cenowa będzie mniejsza bądź równa wyznaczonej cenie w danej godzinie będą ją zapewniać w danej godzinie. Taka struktura rynku powoduje przyłożenie większej wagi do optymalizacji kosztów, innowacji produkcyjnej oraz przestawiania produkcji na tańsze źródła energii [5].

2.1.4 Czynniki wpływające na cenę energii

Autor: Mikołaj Kalisz

Cena energii jest ustanawiana poprzez istniejące zapotrzebowanie oraz istniejącą podaż. Można wyróżnić kilka głównych czynników wpływających na rynek:

- Inwestycje w odnawialne źródła energii,
- Polityka ekologiczna,
- Koszty paliw,
- Pora dnia, dzień tygodnia, pora roku.

W celu zaoferowania konkurencyjnej energii nowe inwestycje mocniej skupią się na energiach ze źródeł odnawialnych. Można tu wymienić energię słoneczną, wiatrową czy wodną. Koszty inwestycyjne takich przedsięwzięć stanowią większość części kosztów i przeważają nad niskimi kosztami operacyjnymi. Elektrownie oferują najniższe ceny energii na rynku, gdyż źródłem generowanego prądu są darmowe dobra takie jak promienie słoneczne czy wiatr. Mierzącymi czynnikami wpływającym na cenę mogą być: maksymalna dostępna moc produkcyjna źródeł odnawialnych, aktualna moc tych źródeł, warunki atmosferyczne, które bezpośrednio wpływają na wydajność elektrowni. Pogoda jest zmienna na przestrzeni dnia a nawet godziny, co aktualnie uniemożliwia przestawienie się w pełni na źródła odnawialne [6].

Na istniejącą podaż mają także wpływ czynniki polityczne, w tym polityka ekologiczna. Wprowadzone opłaty za emisję CO₂ bezpośrednio przekładają się na wzrost ceny energii ze źródeł nieodnawialnych i są jednym z głównych czynników wpływających na ten rynek [6] [5].

Elektrownie opierające się na źródłach nieodnawialnych ponoszą także koszty paliwa. To dodatkowy zmienny czynnik ceny energii zależny od aktualnego stanu rynku surowców kopalnych takich jak węgiel kamienny, brunatny, gaz ziemny czy uran. [5]

Zapotrzebowanie na energię jest zmienne w zależności od czasu. W ciągu dnia ilość pobieranej mocy energii może zmieniać się diametralnie. Podczas różnorodnych świąt, wydarzeń sportowych, czy weekendów może zmieniać się zapotrzebowanie i odbiegać od przebiegu rynku z tygodnia. Temat został dokładniej przeanalizowany w rozdziale 4.

2.2 Sztuczne sieci neuronowe (SSN)

Autor: Mikołaj Kalisz

Sztuczne sieci neuronowe są odwzorowaniem biologicznych połączeń neuronowych w postaci modeli matematycznych. Pozwalają na odtworzenie rzeczywistości bez znajomości dokładnych wzorów. Podczas nauki same stworzą model „wnioskowania”, czyli zależność między danymi wejściowymi, a wyjściem. Jedynym wymaganiem jest posiadanie danych wejściowych i przykładowych wyjść [7].

2.2.1 Definicja i podstawy SSN

Autor: Mikołaj Kalisz

Sztuczna sieć neuronowa to zbiór połączonych ze sobą węzłów (zwanych neuronami), które przetwarzają dane wejściowe przekształcając je za pomocą funkcji aktywacji i propagując dalej, w celu rozwiązania zadań takich jak, klasyfikacja, regresja czy rozpoznawanie wzorców [7].

Typowa sztuczna sieć neuronowa składa się z [7]:

- Neuronów, inaczej węzłów – wykonywane są w nich operacje matematyczne na wejściach. Dane wejściowe są sumą ważoną z dodanym przesunięciem (biasem). Na otrzymanym wyniku stosuje się funkcję aktywacji.
- Warstw – składających się z wielu neuronów, połączonych między sobą.
- Funkcji aktywacji: Określają, czy neuron ma się aktywować i przekazać sygnał dalej. Różnią się innymi charakterystykami aktywacji. Za dobranie funkcji odpowiada twórca sieci.
- Połączeń i wag: Każde połączenie między neuronami posiada swoją wagę, która określa wpływ danych na neuron.
- Algorytmu uczenia: Proces, podczas którego następuje dobór wag połączeń, np. za pomocą propagacji wstecznej, w celu uzyskania najlepszego wyniku.

2.2.2 Neuron

Autor: Mikołaj Kalisz

Neuron jest podstawowym budulcem sztucznej sieci neuronowej. Wzorowany na uproszczonym funkcjonowaniu biologicznej komórki nerwowej. Połączone neurony w większych ilościach, uporządkowane w warstwach, tworzą między sobą siatkę powiązań.

Neuron można opisać jako funkcję zależności wejść względem wyjścia. Dane na wejściu, przemnożone przez wagę, są sumowane i przemnażane przez funkcję aktywacji danego neuronu. W zależności od charakterystyki danej funkcji na wyjściu uzyskuje się wynik przetwarzania informacji z wejść. W neuronie może być uwzględniony bias, który jest dodawany do sumy ważonej wejść [8].

Poniżej przedstawiony jest wzór na ważoną sumę wejść do neuronu [8]:

$$s = w_0 + \sum_{i=1}^n x_i w_i \quad (1)$$

gdzie:

w_0 – bias

x_i – wartość na wejściu

w_i – waga przypisana do wejścia

n – liczba wejść

Obliczona suma jest wejściem do funkcji aktywacji, która zwraca nam ostateczny wynik [7]:

$$y = f(s) \quad (2)$$

gdzie:

f – funkcja aktywacji

s – suma ważona wejść

W procesie nauczania wartości wag i biasów są zmieniane w celu osiągnięcia jak najlepszego wyniku na wyjściu w sieci.

2.2.3 Struktura sieci – perceptron

Autor: Mikołaj Kalisz

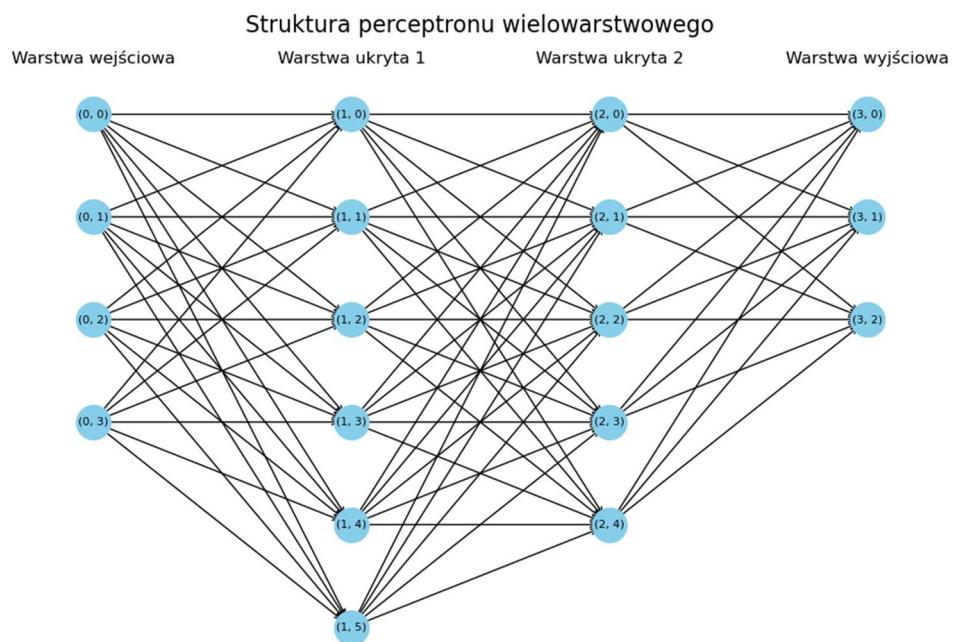
Pojedynczy neuron ma możliwość pozyskania informacji wejściowej i zwrócenia przetworzonej wartości. Jest to prosty sposób na przedstawienie pewnego

wnioskowania, gdzie z n informacji tworzone są nowe wartości y na wyjściu. Taki model można nazwać perceptronem pojedynczym [7].

Istnieje możliwość rozszerzenia tego podejścia do bardziej skomplikowanych struktur łącząc pojedynczy model wnioskujący – neuron, z innymi. W ten sposób powstaje sieć neuronów, każdy z osobnymi wagami. Powstałą strukturę można przedstawić w postaci wielowarstwowej. Taka postać nazywa się perceptronem wielowarstwowym. Składa się on z warstw [7]:

- Wejściowej - odpowiada za przyjęcie danych do sieci,
- Ukrytych – są odpowiedzialne za przekształcanie danych korzystając z połączeń między neuronami,
- Wyjściowej – zwraca wyniki przetwarzania.

Każda warstwa może posiadać inne funkcje aktywacji oraz różną ilość neuronów. Wejściem do następnej warstwy są wyjścia poprzedniej, w ten sposób tworząc sieć powiązań od wejść do wyjść.



Rysunek 2 Przedstawiona jest przykładowa struktura perceptronu składającego się z warstw wejściowej, wyjściowej i dwóch ukrytych. Opracowanie własne

2.2.4 Funkcje aktywacji

Autor: Mikołaj Kalisz

Funkcja aktywacji jest odpowiedzialna za przekształcenie otrzymanej sumy ważonej wejść. O wartości na wyjściu w głównej mierze decyduje charakterystyka tej funkcji. W zależności od kształtu funkcji można otrzymać wartości z różnych

przedziałów wartości, więc dobór jej i dostosowanie do pożąданej wartości na wyjściu sieci jest kluczowe, by osiągnąć jak najlepsze wyniki [7].

Można wyróżnić kilka głównych funkcji aktywacji:

- Liniowa – najprostsza funkcja, zwracająca na wyjściu wartość proporcjonalną do sumy ważonej wejść. Opisana jest poniższym wzorem:

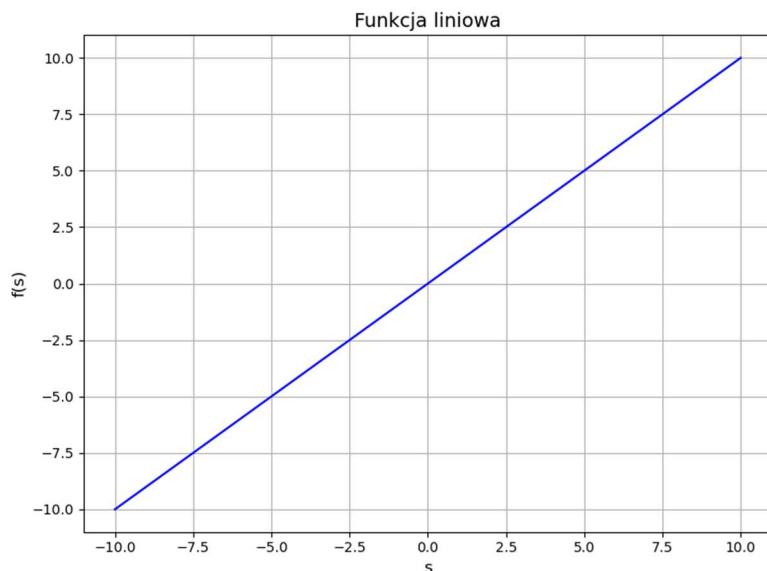
$$f(s) = c \cdot s \quad (3)$$

gdzie:

c – multiplikator

s – suma ważona wejść

Sieci zawierające funkcje liniowe bardzo dobrze się uczą. Niestety w wielu przypadkach zastosowanie funkcji liniowej jest niewystarczające, gdyż model taki jest za prosty i nie uwzględnia nieliniowości, jaka często występuje [7].



Rysunek 3 Wykres funkcji liniowej, pokazuje zależność wejścia s od wyjścia. Opracowanie własne

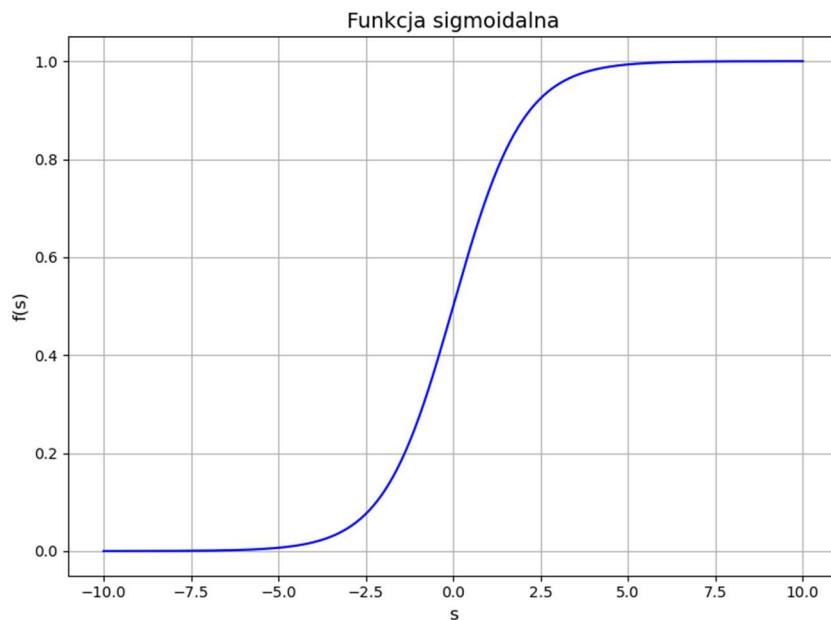
- Sigmoidalna – funkcja nieliniowa, jej wartości są z przedziału od 0 do 1 z łagodnym przejściem między nimi.

$$f(s) = \frac{1}{1 + e^{-s}} \quad (4)$$

gdzie:

s – suma ważona wejścia

Funkcja znalazła swoje zastosowanie w klasyfikacji binarnej ze względu na wyjście w zakresie 0 do 1. Pochodna funkcji przy skrajnych wartościach funkcji jest bliska零 co wskazuje na powolną zmianę wartości. Wynikiem tego zastosowanie funkcji sigmoid może spowolnić uczenie modelu [9].



Rysunek 4 Wykres przedstawiający wartość funkcji sigmoidalnej w zależności od sumy ważonej na wejściu.
Opracowanie własne

- ReLu – Zredukowana jednostka liniowa (ang. Rectified Linear Unit) jest to zmodyfikowana funkcja liniowa, gdzie dla ujemnych wartości przypisywane są wartości zerowe na wyjściu. Jest to funkcja nielinowa.

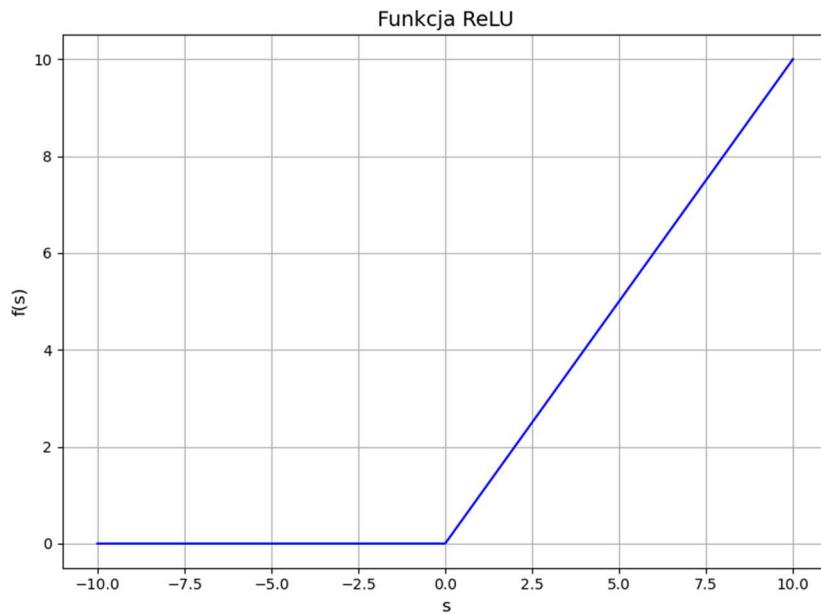
Wygląda następująco:

$$f(s) = \max(0, s) \quad (5)$$

gdzie:

s – suma ważona wejścia

Funkcja ReLu jest jedną z częściej używanych funkcji w sieciach neuronowych. Głównie stosowana w warstwach ukrytych. Z racji na „ucięcie” ujemnych wartości i przypisanie wartości 0, wartość pochodnej dla ujemnych s jest równa zero [10].



Rysunek 5 Wykres zależności wartości funkcji aktywacji ReLu od sumy ważonej na wejściu s. Opracowanie własne.

- LeakyRelu – Nieszczelna zredukowana jednostka liniowa (ang. Leaky Rectified Linear Unit) jest to wariant ReLu, który pozwala “przeciekąć” wartościom ujemnym:

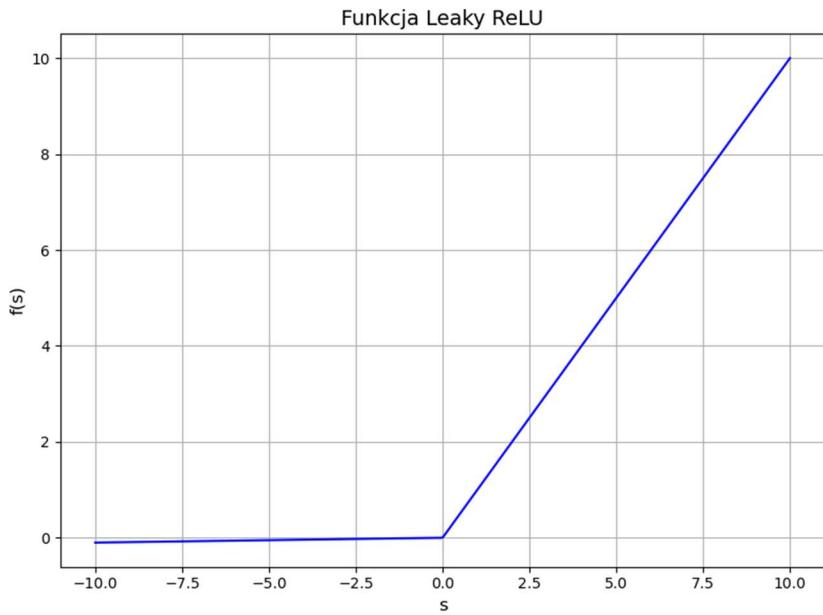
$$f(s) = \max(\alpha \cdot s, s) \quad (6)$$

gdzie:

s – suma ważona wejść

α - stały parametr nachylenia funkcji dla wartości ujemnych

Funkcja LeakyRelu zmienia jedną ważną rzecz względem klasycznego ReLu, jeśli na wejściu są dane ujemne to na wyjściu zostanie przekazana przemnożona wartość wejścia przez alfę. Co za tym idzie pochodna wartości funkcji dla ujemnych wartości s ma wartość alfa, a nie zero. Jest to kluczowa zmiana pod kątem uczenia sieci i liczenia gradientu podczas optymalizacji modelu [10].



Rysunek 6 Funkcja LeakyReLU. Dla wartości s mniejszych od 0 można zauważać delikatny spadek wartości funkcji (w przeciwieństwie do ReLu). Opracowanie własne

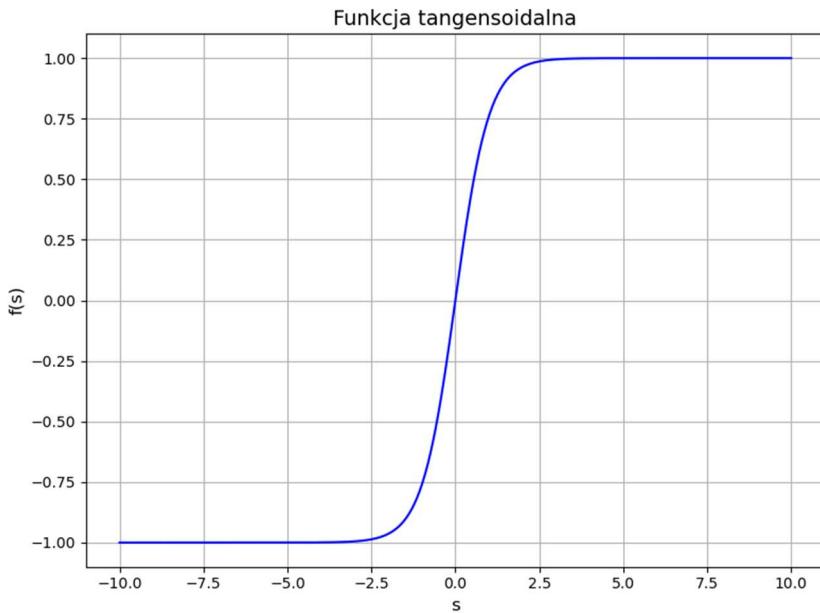
- Tangensoidalna – funkcja podobna do sigmoidalnej z kluczową różnicą, zakres wartości wyjściowych znajduje się w przedziale [-1, 1].

$$f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}} \quad (7)$$

gdzie:

s – suma ważona wejść

Funkcja tangensoidalna swoim wyglądem przypomina funkcję sigmoid. Kluczowymi różnicami jest zakres wartości wyjściowych oraz kąt nachylenia funkcji w środkowym zakresie. Szybsze podejście skutkuje tym, że podczas delikatnej zmiany wartości s, wartość funkcji może zmienić się diametralnie [9].



Rysunek 7 Funkcja tangensoidalna. w zakresie od -2 do 2 można zauważać gwałtowny wzrost wartości funkcji.
Opracowanie własne

- Softmax – funkcja ta na podstawie wartości na wejściu zwraca rozkład prawdopodobieństwa w postaci liczb z zakresu [0,1] o sumie równej 1. Sposób działania softmax jest odmienny od poprzednich funkcji z racji tego, że do obliczeń wymaga informacji o sumach ważonych wejść z innych neuronów z warstwy. Umożliwia to policzenie prawdopodobieństwa sumującego się do 1 ze wszystkimi neuronami w warstwie [9].

Opisuje to funkcja:

$$f_i(s_i) = \frac{e^{s_i}}{\sum_{j=1}^N e^{s_j}} \quad (8)$$

gdzie:

s_i – suma ważona wejść dla i-tego neuronu

s_j - wartość sumy ważonej wejść dla j-tego neuronu

N – liczba neuronów w warstwie

Stosowana jest w klasyfikacji w warstwie wyjścia. Pozyskuje dane z poprzedniej warstwy ukrytej i na podstawie sumy ważonej tworzy rozkład prawdopodobieństwa przynależności do danej klasy. Można zastosować tam, gdzie jest konieczne jednoznaczne wskazanie

klasyfikatora ze względu na sumowanie się prawdopodobieństwa do 1 [9].

2.2.5 Normalizacja i standaryzacja danych

Autor: Mikołaj Kalisz

Dane wejściowe takie jak ceny surowców, kursy walut, czy udział w rynku mogą mieć różne zakresy, mediany czy odchylenie standardowe. W celu ułatwienia nauki modelu oraz zmniejszenia różnic w skali między zmiennymi, które mogły w różnym stopniu wpływać na działanie sieci, stosuje się normalizację oraz standaryzację danych. Oznacza to najczęściej zmianę zakresu wartości danych do określonego przedziału, zmianę wartości średniej, odchylenia standardowego, czy mediany.

Można wyróżnić następujące techniki normalizacji i standaryzacji:

- Min-Max – jest to technika normalizacji opierająca się na wartościach minimalnej i maksymalnej. Prosta zasada działania pozwala na szybkie działanie. Zapewnia zmianę zakresu wartości na [0, 1]. Jest wrażliwa na wartości skrajne, więc w przypadku występowania w danych mocno odbiegających wartości, może to wpływać znaczco na rozkład wartości po standaryzacji [11].

Wartości normalizacji liczone są następująco:

$$x_{norm} = \frac{x - \min(X)}{\max(X) - \min(X)} \quad (9)$$

gdzie:

x – wartość normalizowana

X – zbiór wszystkich wartości

- Standaryzacja danych polega na tym, by osiągnąć średnią równą 0 i odchylenie równe 1. Jest wrażliwa na wartości skrajne. Stosowana w celu przeniesienia dystrybucji danych w okolicę wartości 0. Jednocześnie wpływa to na zniwelowanie dominacji danych o wyższych wartościach. [11].

Wzór wygląda następująco:

$$x_{norm} = \frac{x - \mu(X)}{\sigma(X)} \quad (10)$$

gdzie:

x – wartość normalizowana

X – zbiór wszystkich wartości

$\mu(X)$ – średnia wartość zbioru X

$\sigma(X)$ – odchylenie standardowe zbioru X

- Normalizacja mediany (ang. Median Normalization) polega na podzieleniu wszystkich wartości przez medianę. Nie jest wrażliwa na wartości skrajne [12].

Wartość znormalizowaną liczy się następująco:

$$x_{norm} = \frac{x}{\text{median}(X)} \quad (11)$$

gdzie:

x – wartość normalizowana

X – zbiór wszystkich wartości

- Normalizacja sigmoidalna (ang. Sigmoid Normalization) – do normalizacji wykorzystana jest funkcja sigmoidalna. Przekształca dane na zakres od 0 do 1. Zaletą jest brak potrzeby posiadania rozkładu danych, operacja przeprowadzana jest bezpośrednio na zmiennej normalizowanej [12].

Wzór na normalizację sigmoidalną wygląda następująco:

$$x_{norm} = \frac{1}{1 + e^{-x}} \quad (12)$$

gdzie:

x – wartość normalizowana

- Estymatory tanh (ang. Tanh Estimators) – do normalizacji wykorzystana jest funkcja tangensu hiperbolicznego. Pomaga z eliminacją wpływu wartości skrajnych. Ze względu na swoją budowę jest bardziej wymagający obliczeniowo w porównaniu z MinMax czy standardaryzacją. Przekształca dane na zakres od -1 do 1 [12].

Wzór wygląda następująco [13]:

$$x_{norm} = 0.5 \left(\tanh \left(\frac{0.01(x - \mu(X))}{\sigma(X)} \right) \right) \quad (13)$$

gdzie:

x – wartość normalizowana

X – zbiór wszystkich wartości

$\mu(X)$ – średnia wartość zbioru X

$\sigma(X)$ – odchylenie standardowe zbioru X

2.2.6 Podział danych

Autor: Mikołaj Kalisz

Aby osiągnieć jak najlepsze wyniki model jest zmieniany wielokrotnie, z częstymi zmianami parametrów na podstawie doświadczeń. W celu oceny dokładności modelu przeprowadza się testy. Posiadane dane dzieli się na podzbiory: treningowy oraz testowy. Dane treningowe są przekazywane do modelu w celu nauczania, a dane testowe do sprawdzenia dokładności działania modelu. Używając danych testowych generuje się odpowiedzi modelu, następnie korzystając z metryk omówionych w rozdziale 2.5, sprawdza się dokładność działania powstałego modelu.

Jednakże podział danych na dwa podzbiory, często okazuje się niewystarczający. Do momentu zakończenia procesu nauki nie można w pełni ocenić, jak model poradzi sobie podczas testów. Podczas nauki model zna jedynie swoją dokładność na zbiorze treningowym. Wynik ten jest podatny na przeuczenie (ang. overfitting), czyli zjawisko, w którym model zbyt mocno dopasowuje się do konkretnych wartości występujących w zbiorze treningowym, przez co model wykaże słabe działanie na danych, których wcześniej nie widział.

Rozwiązaniem tego problemu jest wprowadzenie trzeciego podzbioru – zbioru walidacyjnego. Jego zadaniem jest sprawdzanie działania modelu w trakcie procesu nauki. W przeciwieństwie do zbioru treningowego, model nie korzysta z tych danych do nauki, lecz do samej weryfikacji działania między iteracjami [14].

Istotnym jest by zapewnić odpowiedni podział danych, tak by w każdym podzbiorze znajdywały się dane reprezentatywne dla całości zbioru. W ten sposób można zapobiec nierównomiernemu rozłożeniu danych, co mogłoby spowodować nadreprezentację jednej z klas [14].

W celu podziału danych stosuje się następujące techniki:

- Walidacja krzyżowa (ang. cross-validation) – polega na podziale zbioru uczącego na k podzbiorów. Dla każdej iteracji nauczania do nauki używane jest $k - 1$ podzbiorów, a 1 podzbiór do walidacji. Ważne jest podkreślenie faktu, że o ile w danej iteracji zbiory treningowy i walidacyjny są rozdzielne, to z racji na rotacyjną zmianę przypisania danych do zbiorów, występuje tutaj uczenie na danych użytych do walidacji [14].
- Bootstrap – algorytm do próbkowania danych. Polega na losowym wybraniu podzbioru do walidacji, powtarzanym wielokrotnie w każdej iteracji. W ten sposób w zbiorach można osiągnąć taki sam rozkład wartości jak w zbiorze pierwotnym. Może wystąpić ten sam problem z nauczaniem na danych walidacyjnych, czego skutkiem będzie zbyt optymistyczna wartość dokładności modelu na zbiorze walidacyjnym [14].
- Metoda losowa – polega na losowym podziale wartości na dwa podzbiory. Różni się od metody bootstrap, tym, że losowanie występuje raz przed rozpoczęciem nauczania. Zapewnia separację danych między podzbiorami. Z reguły podziału dokonuje się jednorazowo, bez względu na ilość iteracji uczenia. W ten sposób zapewnia się identyczne zbiory danych za każdym razem, lecz nie koniecznie zapewniające najlepsze wyniki.
- Podział względem punktu w czasie – w zastosowaniach predykcji szeregów czasowych możliwe jest zastosowanie podziału względem konkretnego miejsca na osi czasu. W ten sposób model może być testowany na „przyszłych” danych, a uczony na danych z przeszłości. Zapewnia identyczny podział dla każdej iteracji. Umożliwia tworzenie porównań działania różnych modeli, usuwając zależność od podziału danych.

Zastosowanie walidacji umożliwia pozyskanie informacji o najlepszej iteracji nauczania modelu. Tym samym potencjalnie pozwala na cofnięcie się do wersji z najlepszym wynikiem i przekazanie do następnego etapu - testów. W ten sposób można zapobiec korzystaniu z modelu z występującym przeuczeniem, które charakteryzuje się wysokim wynikiem dopasowania do danych treningowych, ale słabym wynikiem na zbiorze walidacyjnym.

2.2.7 Uczenie sieci

Autor: Mikołaj Kalisz

Uczenie, czyli proces dążenia modelu do uzyskania jak najlepszych wyników, opiera się na procesie optymalizacji funkcji kosztu (ang. Loss Function). Funkcja ta

mierzy różnicę między predykcjami modelu, a rzeczywistymi wartościami, stanowiąc miarę błędu [15]. W zależności od rodzaju problemu stosowane są różne funkcje kosztu, takie jak średni błąd kwadratowy MSE opisany w rozdziale 2.5.2.

W trakcie uczenia zmianom podlegają parametry wewnętrzne modelu, takie jak wagi i biasy. Celem optymalizacji sieci neuronowej jest dobranie takich parametrów wewnętrznych modelu, by uzyskać jak najmniejszą wartość funkcji kosztu, co poprawia jakość predykci.

Kluczowym etapem optymalizacji jest obliczanie gradientu funkcji kosztu względem parametrów modelu, który wskazuje na kierunek najszybszego zmniejszania błędu. W sieciach neuronowych gradient można policzyć metodą wstecznej propagacji (ang. Backpropagation), która wykorzystuje regułę łańcuchową do propagowania informacji o błędach z warstw wyjściowych do wejściowych [15].

Wyliczony gradient jest wykorzystywany przez algorytm optymalizacji do aktualizacji parametrów w kierunku minimalizowania wartości funkcji kosztu. Jednym z najczęściej stosowanych i skutecznych optymalizatorów jest algorytm Adam (ang. Adaptive Moment Estimation), który dostosowuje wielkość kroków optymalizacji dla każdego parametru. Dzięki temu zapewnia szybkie zbieżności oraz dobre wyniki w szerokim zakresie zastosowań, w tym sztucznych sieciach neuronowych [16].

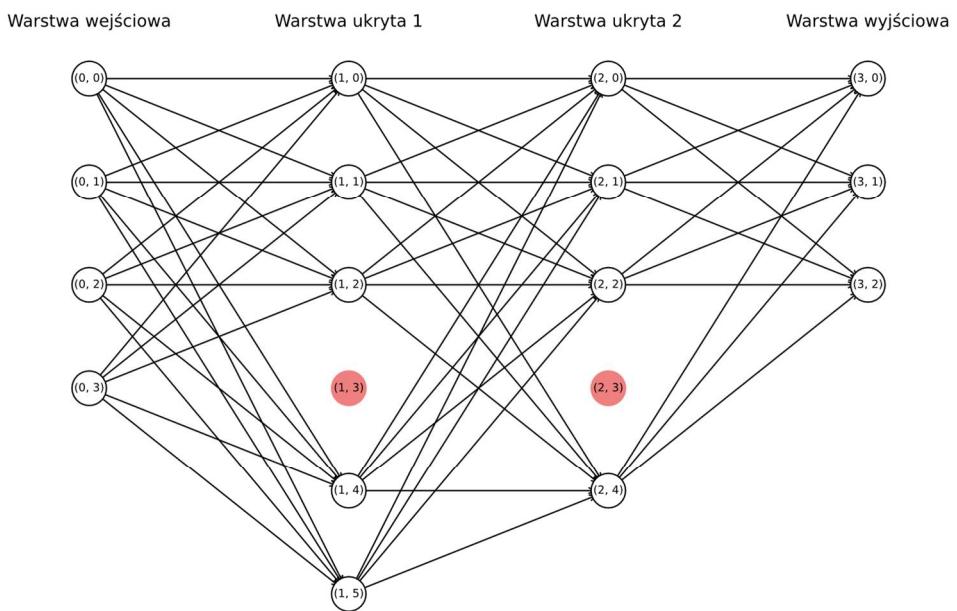
2.2.8 Regularyzacja – dropout

Autor: Mikołaj Kalisz

Jednym z często występujących problemów podczas uczenia sieci neuronowych jest przeuczenie (zdefiniowane w rozdziale 2.2.6). Aby ograniczyć to zjawisko, powstały różne techniki regularyzacji, takie jak L1, L2, wcześnie kończenie (ang. early stopping) oraz dropout. Regularyzacja to technika stosowana w sieciach neuronowych w celu zmniejszenia złożoności modelu oraz poprawy jego generalizacji na nowych danych [17].

Dropout, przedstawiony na rysunku 8, jest jedną z najsukceszniejszych technik regularyzacji. Polega na częściowym wyłączeniu neuronów w wybranych warstwach podczas nauki. Określa się odsetek neuronów, które mają zostać wyłączone, a następnie podczas każdej iteracji losowane są nowe neurony do wyłączenia. Nie uczestniczą one w aktualizacji wag podczas optymalizacji aż do kolejnej iteracji. Zastosowanie tej techniki wydłuża proces treningu modelu. Dropout stosuje się jedynie w fazie nauczania - podczas korzystania z modelu wszystkie neurony są aktywne [17].

Prosta implementacja połączona z obiecującymi wynikami sprawia, że technika ta jest szeroko stosowana i testowana w różnych rodzajach modeli.

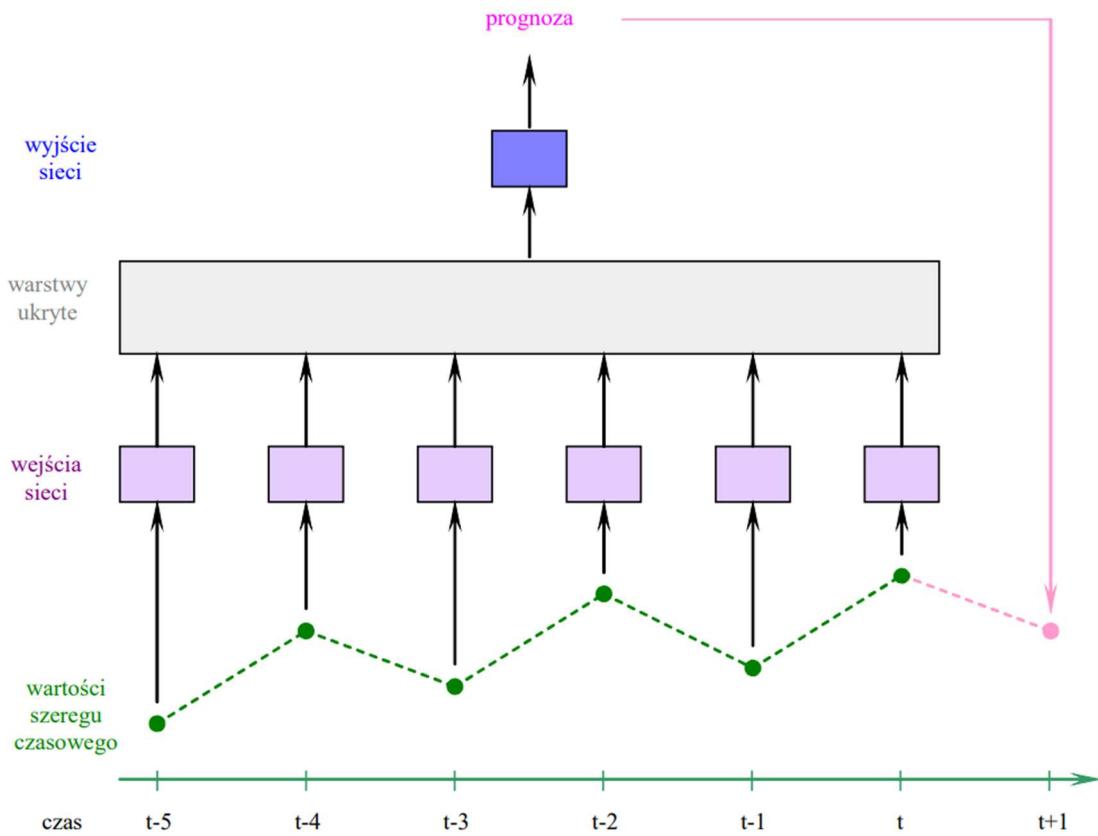


Rysunek 8 Przykładowa architektura sieci pokazująca wyłączone neurony metodą dropout. Opracowanie własne

2.2.9 Zastosowanie SSN w predykcji szeregów czasowych

Autor: Mikołaj Kalisz

W celu użycia SSN w predykcji szeregów czasowych do danych wejściowych dodaje się wartości opóźnione w czasie. W ten sposób sieć ma szansę na odnalezienie powiązań między danymi i dokładniejsze przewidzenie przyszłych wartości. Czas opóźnienia dla danych i ich częstotliwość dobiera się empirycznie [18].



Rysunek 9 „Przykładowe odwzorowanie szeregu czasowego na sieć neuronową wyznaczającą prognozę”
Opracowanie pochodzi z [18]

Na rysunku 9 przedstawiony jest prosty schemat działania sieci neuronowej mającej na celu prognozę następnej wartości w szeregu czasowym. Na podstawie sześciu znanych wartości model ma za zadanie predykcję wartości na następny dzień. Następnego dnia nowa już poznana wartość naszej zmiennej jest podana do modelu, a cały szereg czasowy przesunięty o jedną pozycję w celu predykcji następnej wartości [18]. Podany model można rozszerzyć o inne dane, które można także opóźnić dając dla sieci możliwość pozyskania z nich wiedzy i potencjalnie zwiększenia dokładności predykcji.

2.3 Uczenie Ciągłe

Autor: Adam Mantiuk

Warunki na rynku energii elektrycznej są zmienne, tak jak w każdym innym sektorze gospodarki. Aby stworzone modele uczenia maszynowego mogły działać przez dłuższy okres czasu potrzebne są ich modyfikacje.

Klasyczny paradygmat Uczenia Maszynowego zakłada naukę na określonym zbiorze danych „od zera”, a następnie wdrożeniu stworzonego modelu. Takie podejście

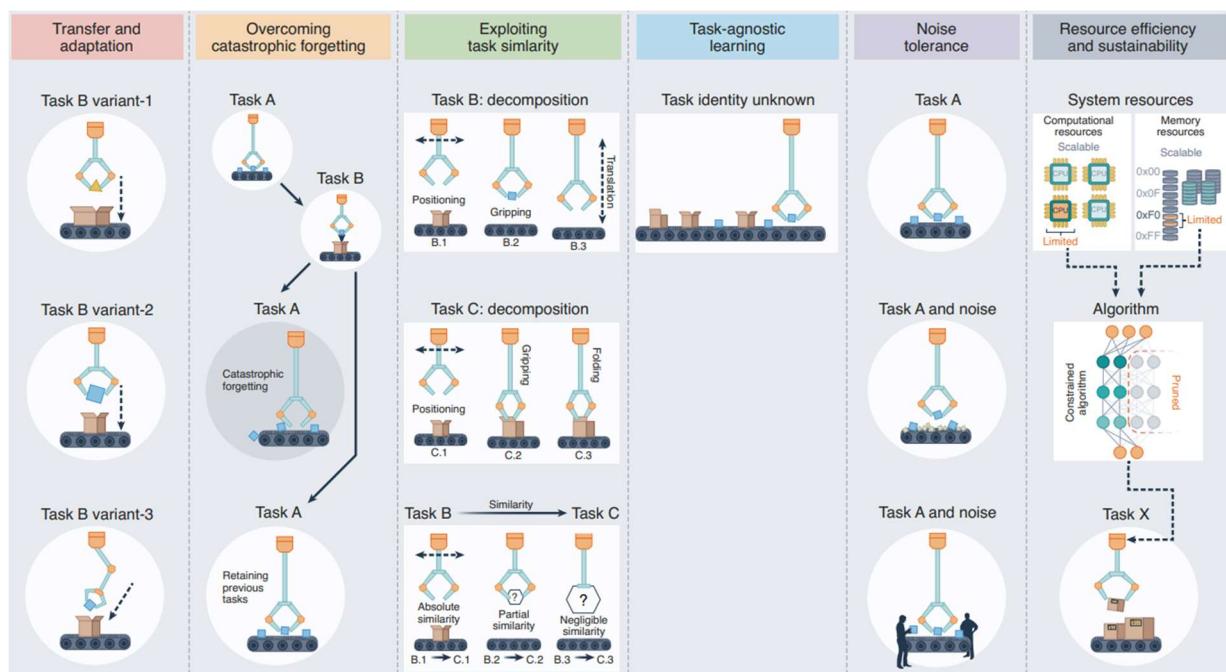
znaczowo różni się od tego jak uczą się ludzie. Człowiek uczy się przez całe życie, dzięki czemu wykorzystuje wcześniej nabycą wiedzę do nauki nowych rzeczy. Praktycznie nigdy nie uczy się „od zera” tak jak modele uczenia maszynowego [19].

Uczenie Ciągłe, nazywane również Uczeniem Przez Całe Życie, zakłada podejście bardziej zbliżone do ludzkiego. Model nauczony na N przykładów nie tylko realizuje określone zadania na przykładzie N+1, ale także kontynuuje na nim naukę. W ciągu takiego procesu model jest aktualizowany o nową zdobytą wiedzę [19].

W zastosowaniu uczenia ciągłego przy predykcji szeregu czasowych można wykorzystywać zatem nowo spływające dane do aktualizacji modelu. Możliwe jest zatem podzielenie roku na kwartały i douczanie go na nowych kwartałach. W takim przypadku model aktualizowany jest cztery razy w roku i może pozwolić to na bardziej dokładne przewidywanie cen.

Przed modelami Uczenia Ciągłego postawionych jest szereg wyzwań. Najważniejsze to [20]:

- Transfer i adaptacja
- Katastroficzne zapominanie
- Wykorzystanie podobieństwa zadań
- Uczenie niezależnie od zadania
- Odporność na szum
- Efektywność zasobowa



Rysunek 10 Key features required to achieve lifelong learning. Opracowanie pochodzi z [20]

2.3.1 Transfer i adaptacja

Autor: Adam Mantiuk

Transfer i adaptacja są bardzo ważnymi aspektami uczenia ciągłego. Modele te z założenia muszą dostosowywać się do zmieniających się warunków. Oznacza to, że takie systemy nie powinny opierać się na kosztownym i czasochłonnym procesie ponownego trenowania [20].

W przypadku szeregow czasowych proces trenowania nowego modelu nie jest, aż tak długi jak na przykład w przypadku klasyfikacji obrazów. Mimo to warto zadbać aby modele uczenia maszynowego były jak najbardziej optymalne i działały sprawnie przez jak najdłuższy czas. Chociażby ze względu na koszty opracowania nowego modelu, który musiałby zostać stworzony na potrzeby nowych warunków środowiskowych.

Transfer bazy wiedzy powinien polegać na wykorzystaniu wcześniejszych nabytych przez model umiejętności do realizacji nowych zadań [20]. Przykładowo, model który jest w stanie monitorować liczbę samochodów na parkingu, może również po minimalnym dostosowaniu wspomagać sterowanie sygnalizacją świetlną przy małym ruchu samochodowym (na przykład w nocy).

Transfer jest bardzo ważnym aspektem zastosowania uczenia ciągłego w predykcji cen energii. Model może być uczony na pewnych ramach czasowych, które to mogą być podzielone na pewne paczki danych. Przykładem takiego podziału może być podział roku na kwartały. Oczekuje się, że model trenowany na pierwszych kwartałach wykorzysta pewne wzorce, których w nich występowały i wykorzysta je do predykcji nowego kwartału.

Adaptacja natomiast powinna móc zapewnić efektywne działanie modelu przy zmienionych warunkach środowiskowych. Nie zawsze możliwe jest przewidzenie wszelkich warunków w jakich będzie pracował model. Dlatego istotnym jest, aby nowe zmienne nie wymuszały ponownego treningu [20]. Przykładowo, jeśli model uczony był do pracy w środowisku pozbawionym opadów śniegu (i tak wyglądał zbiór treningowy) to pojedyncze anomalie pogodowe mogłyby zakłócić działanie modelu, jeśli ten nie byłby w stanie zaadoptować się do nowego środowiska.

Adaptacja jest bardzo ważna w przypadku predykcji szeregow czasowych. Na każdym rynku zdarzają się sytuacje, które w dużym stopniu odmieniają warunki panujące w danym obszarze gospodarki. Przykładem jaki występuje w danych, na których bazuje poniższa praca, jest fakt rozpoczęcia się działań zbrojnych na terenie Ukrainy i nagłe zachwiania na rynku z tym związane. Pozwoli to ocenić czy stworzone modele potrafią odpowiednio zaadaptować się do nowych warunków jakie pojawią się w pierwszym kwartale 2022 roku.

2.3.2 Katastroficzne zapominanie

Autor: Adam Mantiuk

Jednym z najbardziej typowych problemów napotykanych przy tworzeniu modeli Ucznia Ciągłego jest tak zwane „katastroficzne zapominanie”. Objawia się to nadpisywaniem wcześniej wyuczonych wzorców nowymi, pochodzącyimi z nowych danych treningowych [20]. Zjawisko to nazywane jest też dylematem stabilności-plastyczności (ang. stability-plasticity dilemma). Zbyt plastyczny model będzie cierpał na zjawisku katastroficznego zapominania, z kolei model zbyt stabilny będzie w bardzo małym stopniu dostosowywał się do nowych danych treningowych [19].

Zjawisko katastroficznego zapominania można zauważać w typowym przykładzie uczenia transferowego (ang. Transfer Learning). W takim scenariuszu model początkowo trenowany na danych źródłowych jest dostrajany na mniejszym zbiorze danych docelowych, ale tylko w obrębie warstw wyjściowych sieci. Powoduje to dostosowanie modelu głównie do zbioru danych docelowych, przez co model traci swoją wcześniejszą wydajność dla danych źródłowych [19].

Można poddać pod wątpliwość, czy na zmieniającym się z czasem rynku przezwyciężenie katastroficznego zapominania jest czymś istotnym. Z jednej strony wyuczone wzorce mogą się powtarzać, jednak niekoniecznie muszą. Jeśli model jest aktualizowany o nowe dane, to być może zachowywanie w pamięci danych historycznych nie jest nadto istotne. Nadpisywanie starych, wyuczonych schematów może jednak zbyt mocno odmienić model w zbyt krótkim czasie. Istotne jest aby model pamiętał o schematach, które wystąpiły zeszłej jesieni, a nie zostały nadpisane przez wzorce z sezonu wiosenno-letniego.

2.3.3 Wykorzystanie podobieństwa zadań

Autor: Adam Mantiuk

Modele uczenia ciągłego projektowane są do wykonywania wielu zadań jednocześnie, co pozwala na zastosowanie tej samej wiedzy do kilku podobnych sobie przypadków. Takie podejście nie tylko poprawia wydajność systemu, ale również poprawia proces nauki nowych zadań [20].

Transfer wiedzy między zadaniami można podzielić na transfer do przodu (ang. forward transfer) oraz transfer wsteczny (ang. backward transfer) [20].

Transfer do przodu polega na wykorzystaniu wcześniejszej wiedzy do rozwiązania nowych zadań [20]. Przykładowo, model, który potrafi rozpoznawać samochody może wykorzystać tą wiedzę przy procesie uczenia rozpoznawania motocykli.

Transfer wsteczny odnosi się natomiast do przypadków, w których nauka nowego zadania poprawiła wydajność poprzedniego [20]. Zatem nauka

rozpoznawania motocykli mogła pozytywnie wpływać na efektywność rozpoznawania samochodów.

2.3.4 Uczenie niezależnie od zadania

Autor: Adam Mantiuk

Uczenie niezależnie od zadania (ang, task-agnostic learning) jest kluczowe dla modeli ucznia ciągłego. W rzeczywistych warunkach model nie musi być informowany o tym jakie zadanie będzie mu dane wykonać w danym momencie. Środowiska w jakich muszą działać modele uczenia przez całe życie mogą być dynamiczne i nieprzewidywalne. Dlatego muszą skutecznie wykonywać zadaną im pracę bez wiedzy na temat tego w jaki sposób zmieniają się zadania (lub kiedy ta zmiana następuje) [20].

Predykcja ceny energii nie wymaga podziału na zadania, jednak można wyróżnić kilka przypadków, gdzie znalazłyby to zastosowanie. Istnieje możliwość wyróżnienia kilku zadań, na przykład predykcja przy wysokiej cenie gazu. Różne zadania mogłyby być wyróżnione na podstawie sezonów (takim sezonem jest na przykład sezon grzewczy lub sezon jesiennno-zimowy). Model w takim przypadku musiałby rozpoznawać rozpoczynający się sezon grzewczy, na przykład na podstawie danych pogodowych i dostosowywać swoje predykcje do nowych warunków.

2.3.5 Odporność na szum

Autor: Adam Mantiuk

Tradycyjne modele sztucznej inteligencji trenowane są na starannie dobranych i oczyszczonych danych. Pozwala to na uzyskanie wysokiej wydajności w kontrolowanych warunkach. Problem pojawia się w momencie, w którym dane używane do wnioskowania w znaczący sposób odbiegają od danych treningowych. Zmiany te mogą wynikać zarówno z dynamicznych warunków panujących w środowisku, jak i z błędów pojawiających się na czujnikach pomiarowych. Aby modele uczenia ciągłego były skuteczne w rzeczywistych warunkach, muszą być odporne na szумy występujące w spływających danych [20].

Na rynku energii elektrycznej nie bazuje się na danych z czujników pomiarowych, jednak dane mogą być w pewien sposób zaszucone. Zdarzają się chwilowe wahania na rynku, które nie mają ściśle określonego uzasadnienia. Drugim przypadkiem może być chwilowy błąd w zapisie nowych rekordów do bazy danych, zbiór danych zawiera wtedy błędne dane.

2.3.6 Efektywność zasobowa

Autor: Adam Mantiuk

Do Uczenia Przez Całe Życie potrzebne są duże zbiory danych, a to wiąże się z dużym obciążeniem systemów informatycznych. Jednak niektóre dane, takie jak stare dane treningowe, nie są potrzebne dla systemu w czasie rzeczywistym. Natomiast dane, które trzeba przechowywać należy zoptymalizować w taki sposób, aby zajmowały w systemie jak najmniej miejsca [20].

Największe zapotrzebowanie na pamięć mają dane pochodzące z obrazów, bądź filmów video. Dane z szeregow czasowych nie zajmują, aż tak dużej ilości miejsca, a ich proces trenowania nie jest aż tak długi jak proces trenowania modelu klasyfikacji obrazów. Można zatem stwierdzić, że tworzenie modeli uczenia ciągłego w zastosowaniu do predykcji szeregow czasowych nie wiąże się z aż tak dużym zapotrzebowaniem na moc obliczeniową, jak i na samą przestrzeń pamięci dyskowej.

2.3.7 Perspektywa biologiczna

Autor: Adam Mantiuk

Organizmy żywe uczą się przez całe życie. Nowe doświadczenia pozwalają organizmom na adaptację do zmiennych warunków otoczenia. Mechanizmy biologiczne stanowią inspirację dla rozwiązań uczenia maszynowego. Niektóre z nich stosuje się do pokonania zjawiska katastroficznego zapominania [20].

Można zatem stwierdzić, że uczenie ciągłe posiada swoje wzorce biologiczne. Tak samo jak sztuczne sieci neuronowe, które wzorowane są na biologicznych neuronach. Natura może zatem stanowić inspirację przy wielu rozwiązaniach, ponieważ można próbować sprostać wyzwaniom w ten sam sposób w jaki radzą sobie żywe organizmy. Pewne mechanizmy nie mogą zostać dosłownie odwzorowane jednak inspiracja nimi może doprowadzić do znalezienia optymalnego rozwiązania.

Proces tworzenia nowych neuronów, nazywany w biologii neurogenezą, jest mechanizmem chroniącym wcześniej zdobytą wiedzę przy przyswajaniu nowej. Nowy neuron może być odpowiedzialny za reprezentację nowych wspomnień bez ingerencji w poprzednie doświadczenia. W zastosowaniach uczenia maszynowego nowe neurony mogą reprezentować nowe zadania, bądź też nowe obiekty klasyfikacji. Ważna jest też kontrola nad zasobami, co za tym idzie – nowych neuronów nie należy tworzyć w nieskończoność. Rozwiązaniem mogą być mechanizmy selekcyjne. Takie algorytmy miałyby za zadanie odróżniać nowe informacje od tych, które mają już swoją reprezentację w istniejącej strukturze [20].

Zastosowanie neurogenezę może być bardzo efektywne przy próbie klasyfikacji nowych obiektów. Przy predykcji cen energii w szeregach czasowych taki rozwiązanie niekoniecznie musi znaleźć swoje miejsce. Jednak pomysł wartym rozważenia

może być rozpoznawanie zadań. Takim przypadkiem może być na przykład predykcja w okresie grzewczym, bądź predykcja przy niekorzystnym kursie dolara.

Sen jest natomiast mechanizmem, który odgrywa kluczową rolę w procesie uczenia się wszystkich żywych organizmów. Podczas snu zdobyte wcześniej doświadczenia przeplatane są z innymi, wcześniej wyuczonymi wzorcami. W uczeniu ciągłym stosuje się technikę „odtwarzania epizodów” (ang. episodic replay), która jest wzorowana na mechanizmie snu. Ma ona zatem na celu utrwalenie nowej wiedzy z zachowaniem informacji zdobytych podczas wcześniejszego treningu. Takie powtórki mogą być oparte o rzeczywiste paczki danych, które są istotne w procesie uczenia danego modelu [20].

Odtwarzanie epizodów z dużym prawdopodobieństwem może pomóc w przewyciężeniu katastroficznego zapominania. W dalszej części pracy w większym stopniu zostanie poruszona istotność katastroficznego zapominania przy predykcji cen energii elektrycznej. Jednak już w tym momencie warto zwrócić uwagę, iż warunki na rynku są zmienne, więc odtwarzanie starych epizodów w kontekście trenowania modelu na nowych warunkach niekoniecznie musi przynieść korzystny rezultat. Jednak warto rozważyć możliwość odtwarzania epizodów z konkretnych okresów w roku. Powtarzanie okresu jesienno-zimowego przy nauce nowej paczki danych, która zawiera się w tym okresie może okazać się korzystne.

2.4 Zastosowanie Uczenia Ciągłego w prognozowaniu i podejmowaniu decyzji na rynku energii elektrycznej

Autor: Adam Mantiuk

W tym podrozdziale zostaną przedstawione dwie prace poruszające temat zastosowania uczenia ciągłego na rynku energii elektrycznej. Ten, stosunkowo nowy temat, nie został na ten moment szerzej opisany w literaturze naukowej zajmującej się tematyką prognozowania rynku energii elektrycznej. Liczne prace zajmujące się tematyką modelowania cen energii elektrycznej korzystają z metod sztucznej inteligencji, jednak temat uczenia przez całe życie pojawia się dosyć sporadycznie. Omówione tu prace przedstawiają różnorodność możliwości, jakie daje uczenie ciągłe. Pokazują też, że zastosowanie tej strategii może zwiększyć dokładność prognozowania.

Autorzy [21] zastosowali paradygmat uczenia ciągłego do prognozowania zużycia gazu ziemnego. Autorzy zdecydowali się na przetestowanie dwóch wariantów dobierania kolekcji modeli do prognozowania. W jednym przypadku dzielili oni szereg czasowy na pewne odcinki danych. Każdy odcinek charakteryzowała inne cechy statystyczne. Autorzy założyli, że każdy model powinien być trenowany na odcinkach o podobnych właściwościach statystycznych. Punkty które podzieliły dane zostały wyłonione na podstawie znanych sezonów (między innymi sezonu grzewczego) oraz

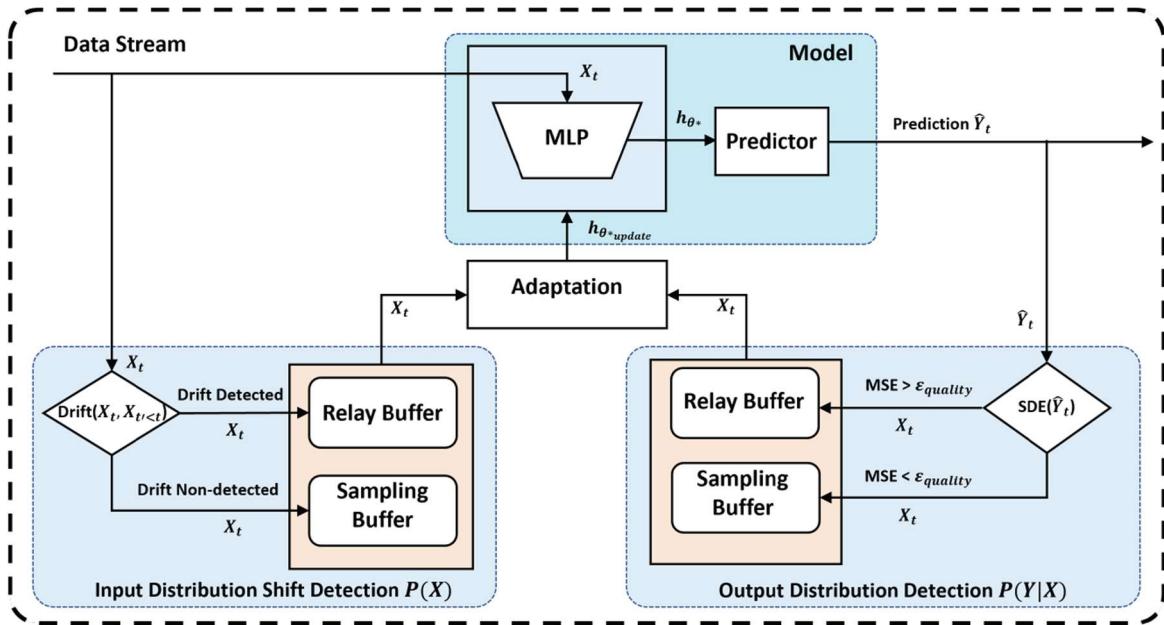
określone za pomocą algorytmów. W drugim przypadku nie dzielili danych na odcinki czasowe, wykorzystali tutaj swoje modele do prognozowania całego zestawu danych.

Podczas przełączania modeli między odcinkami czasowymi, autorzy również zdecydowali się przetestować dwa warianty. W pierwszym po dojściu do punktu zmiany model był natychmiast przełączany. W drugim natomiast wprowadzono obszary wokół punktów zmiany, w których to stosowane były oba modele przejściowe, obliczali błędy prognoz oraz średnią ważoną prognoz.

Jak się okazało wielomodelowe podejście podzielone na kwartały uzyskało lepsze wyniki od podejścia bazowego. Nie była to jednak znacząca różnica. Znacznie lepsze wyniki uzyskały modele oparte o punkty zmiany, tych jednak nie mogło być wykrywanych zbyt wiele. Większa liczba punktów zmiany prowadziła do nadmiernego dopasowania się modelu, co skutkowało zwiększeniem błędów prognozy. Łączenie prognoz wokół punktów zmiany również przyniosło korzyści przy niskiej gęstości punktów zmiany.

Autorzy [22] potraktowali problem przewidywania decyzji transakcyjnych na rynku energii elektrycznej, jako problem klasyfikacji. Model określa, którą ofertę należy przyjąć, a którą odrzucić. Wykorzystana została w tym celu wielowarstwowa sieć neuronowa. Zaproponowana strategia uczenia ciągłego wykorzystuje dwa rodzaje danych. Pierwszym z nich są nowe dane, które charakteryzują odmienne od starych właściwości statystyczne. Wykrywane są zmiany zarówno na danych wejściowych jak i wyjściowych. Drugim rodzajem są stare dane, które mają na celu przezwyciężyć problem katastroficznego zapominania. Model jest aktualizowany jedynie w momencie, gdy środowisko ulegnie zmianie, nowy model zastępuje wtedy swojego poprzednika. Na rysunku 11 przedstawiono schemat zastosowanej strategii. Eksperyment został wykonany na danych pochodzących z Hiszpańskiego rynku.

Zastosowanie powyższej strategii przyczyniło się do poprawy dokładności modelu względem podejścia klasycznego.



Rysunek 11 Schemat adaptacji modelu i wykrywania zmian dystrybucji danych w strumieniach danych
Opracowanie pochodzi z [22]

Mimo iż obie powyższe metody stosują podejście uczenia ciągłego są one odmienne od tej, która zastosowana została w poniższej pracy. Najbardziej zbliżona metoda pojawia się w pracy [21], jednak w przypadku tej pracy jest to predykcja cen energii, a nie zużycia gazu. Szereg czasowy został tutaj odgórnie podzielony na stałe odcinki czasu, a model zmieniany jest natychmiastowo po wykonaniu predykcji na kolejny segment szeregu czasowego.

2.5 Metryki oceny modeli

Autor: Adam Mantiuk

Podczas tworzenia modeli predykcyjnych bardzo ważna jest ocena ich działania. To ona pozwala określić w jakim stopniu stworzone modele są w stanie przewidywać nowe wartości. Co za tym idzie pozwalać na wyłonienie najlepszego modelu. Współcześnie stosuje się wiele metryk, które pozwalają na zmierzenie różnicy między przewidzianymi, a rzeczywistymi wartościami [23].

W niniejszym rozdziale przedstawiono kilka popularnych metryk oceny modeli regresji, które można stosować do prognozowania wartości ciągłych w szeregach czasowych.

Bardzo ważne są też dwie stałe wartości:

- Średnia wartość rzeczywista

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i \quad (14)$$

gdzie:

X_i – wartość rzeczywista

m – liczba obserwacji

- Średnia całkowita suma kwadratów

$$MST = \frac{1}{m} \sum_{i=1}^m (X_i - \bar{X})^2 \quad (15)$$

gdzie:

X_i – wartość rzeczywista

m – liczba obserwacji

\bar{X} – średnia wartość rzeczywista

2.5.1 R^2 – Współczynnik determinacji

Autor: Adam Mantiuk

Współczynnik determinacji (R^2) mierzy, jaka część zmienności zmiennej zależnej jest wyjaśniona przez zmienne niezależne uwzględnione w modelu. Opisuje on, jak dobrze model odwzorowuje dane i wskazuje na stopień dopasowania modelu do obserwacji. Wartość współczynnika mieści się w zakresie od $-\infty$ do 1, gdzie wartości bliższe jedności świadczą o lepszej jakości modelu. Wynik ten można wyrażać w postaci liczbowej lub procentowej.

Wzór na współczynnik determinacji wygląda następująco [23]:

$$R^2 = 1 - \frac{\sum_{i=1}^m (X_i - Y_i)^2}{\sum_{i=1}^m (X_i - \bar{X})^2} \quad (16)$$

gdzie:

X_i – wartość rzeczywista

Y_i – wartość przewidywana przez model

m – liczba obserwacji

\bar{X} – średnia wartość rzeczywista

2.5.2 MSE – średni błąd kwadratowy

Autor: Adam Mantiuk

Średni błąd kwadratowy (MSE, ang. Mean Squared Error) jest miarą najlepiej sprawdzającą się w wykrywaniu wartości odstających. Jest on czuły na duże różnice, ponieważ błędy są podnoszone do kwadratu, co przypisuje większą wagę większym odchyleniom [23].

Metrykę tą wyraża się wzorem:

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (X_i - Y_i)^2 \quad (17)$$

gdzie:

X_i – wartość rzeczywista

Y_i – wartość przewidywana przez model

m – liczba obserwacji

Wyniki średniego błędu kwadratowego zawierają się w przedziale $<0; \infty$). Wartości w tej metryce należy interpretować w następujący sposób: najlepsze wyniki to te dające do 0 (0 jest wynikiem idealnym oznaczającym całkowite dopasowanie), natomiast najgorsze wyniki to te dające do nieskończoności [23].

Metryka MSE zawiera się we współczynniku determinacji R^2 , co oznacza, że R^2 można wyrazić wzorem:

$$R^2 = 1 - \frac{MSE}{MST} \quad (18)$$

gdzie:

MSE – średni błąd kwadratowy

MST – średnia całkowita suma kwadratów

Średnia całkowita suma kwadratów (MST) jest stała dla danej próbki danych. Obydwie metryki są ze sobą powiązane w sposób monotoniczny: wzrost wartości jednej prowadzi do spadku drugiej [23].

2.5.3 RMSE – pierwiastek średniego błędu kwadratowego

Autor: Adam Mantiuk

Pierwiastek średniego błędu kwadratowego (RMSE) jest miarą statystyczną, której szczególną zaletą jest możliwość interpretacji wyników w oryginalnej jednostce [23].

Metrykę tę można wyrazić wzorem:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (X_i - Y_i)^2} \quad (19)$$

gdzie:

X_i – wartość rzeczywista

Y_i – wartość przewidywana przez model

m – liczba obserwacji

Pierwiastek średniego błędu kwadratowego poprzez operację pierwiastkowania jest bezpośrednio powiązany ze średnim błędem kwadratowym. Zatem można go wyrazić wzorem:

$$RMSE = \sqrt{MSE} \quad (20)$$

gdzie:

MSE – średni błąd kwadratowy

Oznacza to, że obie metryki są monotonicznie związane. Zatem posortowanie modeli według MSE da ten sam rezultat co posortowanie modeli według RMSE [23].

2.5.4 MAE – średni błąd bezwzględny

Autor: Adam Mantiuk

Średni błąd bezwzględny (MAE) jest metryką, która oblicza przeciętną różnicę wartości rzeczywistych i przewidywanych. Ignoruje kierunek odchylenia pozwalając na prostą interpretację wielkości błędu [23].

Wzór średniego błędu bezwzględnego jest następujący:

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |X_i - Y_i| \quad (21)$$

gdzie:

X_i – wartość rzeczywista

Y_i – wartość przewidywana przez model

m – liczba obserwacji

Wartości tej metryki mieszczą się w przedziale od 0 do nieskończoności. Im wartość MAE jest bliższa零, tym lepiej model odwzorowuje rzeczywiste dane [23].

Średni błąd bezwzględny nie nadaje dużych wag punktom znacznie oddalonym od rzeczywistych. Sprawia to, że jest szczególnie przydatny, gdy dane zawierają błędnie odstające wartości. Odróżnia to go od średniego błędu kwadratowego [23].

2.5.5 MAPE – średni procentowy błąd bezwzględny

Autor: Mikołaj Kalisz

Średni bezwzględny błąd procentowy (MAPE) to miara oceny dokładności prognozy, która wyraża średni błąd jako procent rzeczywistej wartości. Dzięki swojej definicji, MAPE jest szczególnie zalecana w przypadkach, w których duże znaczenie ma uwzględnienie względnych różnic między prognozowaną a rzeczywistą wartością, a nie samej różnicy bezwzględnej [23].

Wzór na średni bezwzględny błąd procentowy wygląda następująco:

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_i - X_i}{Y_i} \right| \quad (22)$$

gdzie:

X_i – wartość rzeczywista

Y_i – wartość przewidywana przez model

m – liczba obserwacji

Wartości metryki MAPE mieszczą się w przedziale od 0 do nieskończoności. Im wartość MAPE jest bliższa零, tym lepiej model odwzorowuje rzeczywiste dane [23].

Średni bezwzględny błąd procentowy nie nadaje się do danych, których wartości przewidynań są bliskie zeru. Dochodzi wtedy do dzielenia przez bardzo małą wartość, zwiększając wielkość błędu w nieskończoność. Wartość przewidywana nie może być równa zero (dzielenie przez zero) [23].

Średni bezwzględny błąd procentowy ma tendencję do faworyzacji modeli, które generują zaniżone prognozy. Nie należy stosować MAPE dla zastosowań, gdzie mogą wystąpić duże błędy [23].

2.5.6 SMAPE – symetryczny średni procentowy błąd bezwzględny

Autor: Mikołaj Kalisz

Symetryczny średni procentowy błąd bezwzględny (SMAPE) to miara błędu, która eliminuje niektóre ograniczenia klasycznego wskaźnika MAPE, takie jak asymetria w ocenie błędów niedoszacowań i przeszacowań. Podobnie jak w MAPE, SMAPE mierzy względny błąd między przewidywaną a rzeczywistą wartością, ale robi to w sposób symetryczny, uwzględniając tak samo wartości prognozowane, jak i rzeczywiste [23].

Wzór na symetryczny średni procentowy błąd bezwzględny:

$$\text{SMAPE} = \frac{100\%}{m} \sum_{i=1}^m \frac{|X_i - Y_i|}{(|X_i| + |Y_i|) \div 2} \quad (23)$$

gdzie:

X_i – wartość rzeczywista

Y_i – wartość przewidywana przez model

m – liczba obserwacji

Wartości metryki SMAPE mieszczą się w przedziale od 0% do 200%. Im wartość SMAPE jest bliższa零, tym lepiej model odwzorowuje rzeczywiste dane. Wartość 200% oznacza model z najgorszym możliwym prognozowaniem. Wartości bliskie 200% metryka osiąga w przypadku, gdy wartości prognozowane i rzeczywiste są o przeciwnych znakach [23].

Zastosowanie średniej wartości prognozowanej i rzeczywistej w mianowniku eliminuje wpływ przewagi przeszacowań i niedoszacowań, co jest usprawnieniem względem metryki MAPE [23].

3 Eksploracja danych

Autor: Adam Mantiuk

Pierwszym i najważniejszym zadaniem, jakiego należy się podjąć przy projektach uczenia maszynowego, jest przyjrzenie się danym. Wizualizacja danych pozwala na głębsze zrozumienie istoty problemu, oraz pozwala na obranie właściwego kierunku naszego dalszego działania [24].

Aby móc rozpoczęć badanie danych należy je na początku do tego przygotować. I nie chodzi tu jedynie o uzupełnienie brakujących danych, czy wczytanie danych w odpowiednim formacie. Aby zacząć eksplorację danych należy zdefiniować problem, który w następnych krokach ma być rozwiązyany. Określić jaki jest cel ich przetwarzania i jakie rozwiązanie jest poszukiwane [25].

3.1 Identyfikacja problemu

Autor: Adam Mantiuk

Nadrzędnym celem jest wyznaczenie predykcji cen na rynku energii elektrycznej z jak największą dokładnością. Jednak należy odpowiedzieć sobie na pytanie jaka ma być ta dokładność. Na tym etapie należy wstępnie określić jakie dane byłyby nam potrzebne. Co może wpływać na cenę energii na rynku? Jaki ograniczenia mogą nas czekać? Czy posiadane dane posiadają wystarczający zakres [25]?

Z całą pewnością do implementacji rozwiązania potrzebna jest informacja o cenie energii elektrycznej w poszczególnych godzinach. Przydatny może okazać się też wolumen, inaczej mówiąc ilość zakupionej energii po tej cenie. Polska energetyka opiera się na węglu oraz gazie ziemnym, zatem informacja o cenie tych surowców może się okazać kluczowa. Inną ważną zmienną mogą być ceny uprawnień do emisji gazów cieplarnianych (CO₂). Czynnikiem jaki wpływa na cenę energii na rynku jest zapotrzebowanie na nią. Takie informacje można pozyskać z BPKD² (bieżący plan koordynacji dobowej). Można wyciągnąć z niego też inne informacje takie jak wartości rezerw mocy, czy możliwości wytwarzania jednostek, które biorą udział w rynku bilansującym. Ważną kwestią jest też to z czego pozyskiwana była energia elektryczna, zatem informacje o tym z czego produkowana była energia w danym okresie może być kolejną zmiennąwartą przeanalizowania.

Produkcja z odnawialnych źródeł energii powiązana jest z pogodą, jednak analiza tego współczynnika może być bardzo trudna. Tutaj należałoby rozważyć pewne ograniczenie jakimi są dane pogodowe w konkretnej lokalizacji. Czy istnieje możliwość powiązania danych pogodowych z lokalizacją elektrowni wiatrowych czy też fotowoltaicznych? Na pewno byłoby to rozwiązanie czasochłonne. Droga na skróty

² Źródło: [PSE](<https://www.pse.pl/dane-systemowe/plany-pracy-kse/biezacy-plan-koordynacyjny-dobowy-bpkd/wielkosci-podstawowe>)

prowadzi do analizy danych pogodowych z geograficznego środka Polski, jednak silny wiatr na Pomorzu nie implikuje takich samych warunków pogodowych na Podlasiu. Z tego też powodu uśrednione dane pogodowe mogą być słabo skorelowane z danymi o produkcji energii z odnawialnych źródeł energii. Jednak z uwagi na łatwość operacji na tego typu danych można podjąć się takiej analizy i ewentualnie odrzucić te informacje z uwagi na małą istotność.

3.2 Wstępne przetwarzanie danych

Autor: Adam Mantiuk

Wstępne przetwarzanie danych jest etapem poprzedzającym sam proces badania danych, które zostaną użyte do modelowania. To na tym etapie następuje identyfikacja typu pliku w jakich znajdują się dane oraz poddawane są odpowiedniej obróbce. Dzięki temu dane zostaną przygotowane do dalszej eksploracji w projekcie [25].

3.2.1 Znalezienie i identyfikacja danych

Autor: Adam Mantiuk

Dane zostały pozyskane dzięki uprzejmości Mariusza Nowickiego, byłego studenta Politechniki Warszawskiej. Otrzymane pliki z zebranymi danymi należało zidentyfikować i odpowiednio przygotować. Były to pliki Parquet, a więc w formacie który pozwala na przechowywanie danych tabelarycznych w formie kolumnowej. Do ich wczytania niezbędna okazała się biblioteka Apache Arrow. Powiązanie tej biblioteki z językiem Python jest nazywane PyArrow, a sama biblioteka jest w pełni zintegrowana z Numpy czy Pandas [26]. Następnie przy wykorzystaniu tej samej biblioteki dane zostały wczytane do ramek danych (Apache Arrow Table). Tabele zostały przekonwertowane na obiekt pandas DataFrame, a więc do najpopularniejszej struktury danych, która stosowana jest do manipulowania danymi oraz ich analizy. Obiekty Data Frame zostały następnie zapisane w formacie csv.

3.2.2 Agregacja

Autor: Adam Mantiuk

Skuteczne przewidywanie ceny energii na Rynku Dnia Następnego wymaga danych w odpowiedniej rozdzielczości. Aby dokonywać prognoz na każdą godzinę dnia następnego każda zmienna powinna posiadać swoją reprezentację w rekordach odpowiadającym poszczególnym godzinom.

Niektóre informacje, na przykład o cenie gazu czy węgla, aktualizowane były jedynie pierwszego dnia miesiąca i obowiązywały cały miesiąc. Inne z kolei, na przykład kursy walut (euro i dolara) czy ceny uprawnień do emisji gazów

cieplarnianych, aktualizowane były raz dziennie. Inne dane, pochodzące na przykład z bieżących planów koordynacji dobowej, posiadają rozdzielcość godzinową. Chcąc skutecznie złączyć ze sobą te tabele zdecydowano się na implementację funkcji uzupełniania danych do wyższej rozdzielcości. Stworzono dwie funkcje, pierwsza rozszerzała dane o rozdzielcość miesięczną do rozdzielcości godzinowej, druga rozszerzała dane o rozdzielcość dzienną do rozdzielcości godzinowej. Nie zdecydowano się na interpolację czasową, lecz na rozciągnięcie tych samych danych w czasie, a więc te same dane zostały powtórzone dla każdej godziny w ciągu dnia).

Kolumny które pierwotnie posiadały rozdzielcość dzienną:

- kurs_usd_pln
- kurs_eur_pln
- co2_cena
- co2_objetosc

Kolumny które pierwotnie posiadały rozdzielcość miesięczną:

- cena_wegla_pln_gj
- cena_wegla_pln_t
- inflacja_vs_poprzedni_miesiac
- inflacja_vs_poprzedni_rok
- gaz_cena
- gaz_objetosc

Wszystkie tabele z danymi posiadały kolumnę „Timestamp”, która nie wnosiła szczegółowo przydanej informacji w tym projekcie, dlatego zdecydowano się na jej usunięcie. Łącząc dane z wielu tabel to jednej głównej tabeli, zdecydowano się na ustawienie daty (kolumna „datetime”) jako indeks. Do złączenia tabel użyto funkcji „concat” z biblioteki Pandas. Jednak przed tym użyto metody „.sort_index()” obiektu DataFrame dla każdej ramki danych. Zdecydowano się na to, aby uniknąć błędnego łączenia danych i niewłaściwego uzupełnienia wartościami „NaN”.

3.2.3 Braki danych

Autor: Adam Mantiuk

Niektóre kolumny zdarzają się wybrakowane, inne natomiast mogą być w złym formacie. Niekiedy zdarza się też, że cała kolumna została błędnie wczytana do ramki danych i informacja, która była przez nią niesiona została bezpowrotnie utracona. Korzystając z odpowiednich metod obiektów DataFrame można przeanalizować, które kolumny mogą być wadliwe.

Korzystając z metody „dtypes” można sprawdzić w jakim formacie zapisane są wartości w poszczególnych kolumnach. W ten sposób udało się znaleźć kolumnę, która poprzez zły separator części dziesiętnej została uznana za string. W języku Python część dziesiętna liczby zmiennoprzecinkowej oddzielana jest za pomocą kropki, jednak dane w kolumnie „sumaryczna_moc_ladowania_JGMa” miały część

dziesiętną oddzieloną przecinkiem. Korzystając z dostępnych funkcji oraz metod języka Python oraz biblioteki Pandas udało się bezproblemowo przekonwertować dane do odpowiedniego formatu.

Inną metodą która pomaga zlokalizować błędnie zapisane dane jest metoda „describe”. Wywołanie tej metody na obiekcie DataFrame pozwala na obliczenie statystyk dla każdej kolumny. Wśród tych statystyk odnaleźć można między innymi wartość mediany, średniej czy odchylenia standardowego. Dzięki tej metodzie bez większego problemu odnaleziono błędnie zapisane kolumny w tabeli. Niektóre kolumny danych zamiast wartości liczbowej posiadały kreskę. Wśród tych kolumn znalazły się:

- suma_zdolosci_wytwarzczych_JGFWa
- suma_zdolosci_wytwarzczych_JGPVa
- sumaryczna_generacja_JGFWa
- sumaryczna_generacja_JGPVa

Zdecydowano się zatem na usunięcie tych kolumn z obiektu DataFrame, ponieważ nie niosły one żadnej informacji.

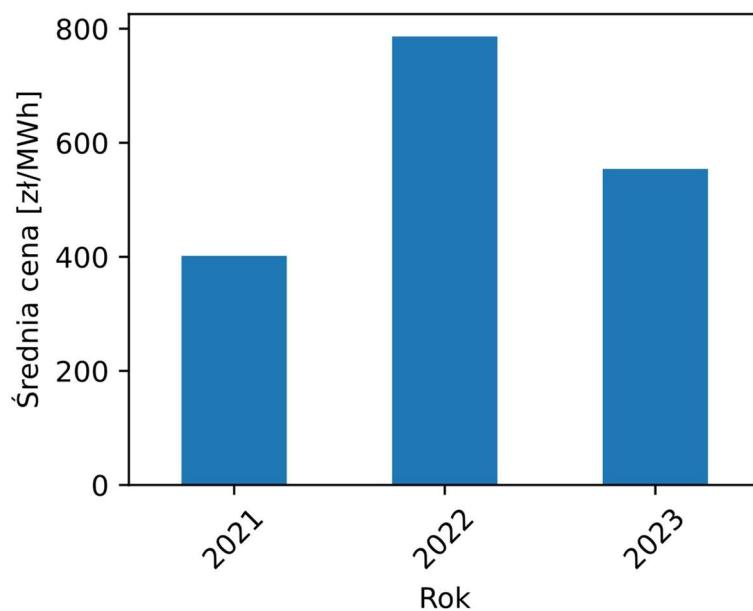
3.3 Badania eksploracyjne (EDA)

3.3.1 Analiza wstępna danych o cenach energii

Autor: Adam Mantiuk

Dane, które zostały przekazane na cel poniższej pracy, obejmują okres od 8 stycznia 2021 roku do 1 września 2023 roku. Jest to ponad 22 tysiące rekordów opisywanych przez ponad 150 kolumn. Mimo tak obszernej ilości danych cały obiekt DataFrame zajmuje w pamięci niecałe 26 MB, co przy mocy obliczeniowej współczesnych komputerów jest wartością niewielką. Dzięki temu analiza zbioru danych jak i trening modelu nie powinny w sposób znaczący obciążać zasobów systemowych.

Zbiór danych zawiera więc dwa pełne lata 2021, 2022 oraz osiem pierwszych miesięcy 2023 roku. Okres ten był dosyć interesujący jeśli chodzi o rynek energii elektrycznej. Przez rok 2021 następowała stopniowa stabilizacja cen po pierwszym roku pandemii COVID-19. Pod koniec roku pojawiły się wahania ze względu na narastające napięcie między Rosją a Ukrainą. Rok 2022 cechową się dużą niepewnością na rynku z powodu trwającego konfliktu. Można to zauważyć na rysunku 12, średnia cena energii w roku 2022 była 2 razy wyższa niż w roku poprzednim



Rysunek 12 Średnia cena energii w poszczególnych latach
Opracowanie własne

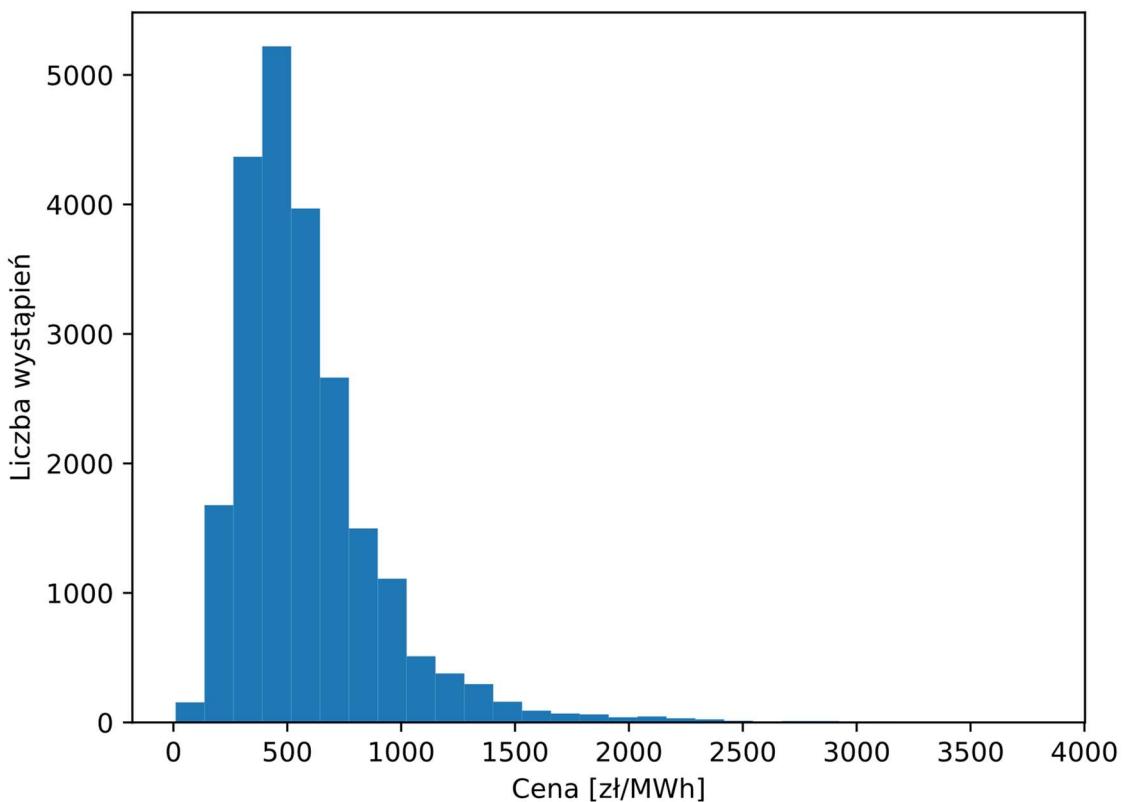
Taki zbiór danych z całą pewnością jest bardzo ciekawy pod kątem analizy, ponieważ prezentuje ekstremalne sytuacje (konflikt zbrojny), które mogą się wydarzyć i spowodować spore zmiany na rynku energii elektrycznej.

W tabeli 1 przedstawione zostały podstawowe statystyki opisowe dla kolumny „fixing_i_price”.

Tabela 1 Podstawowe statystyki opisowe dla kolumny "fixing_i_price"
Opracowanie własne

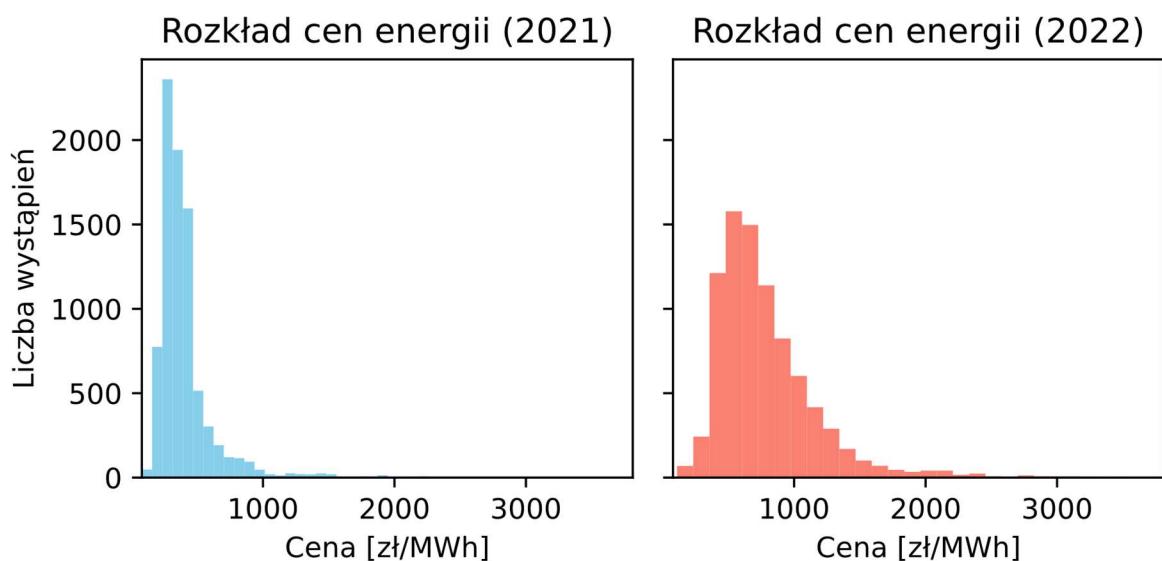
Liczba obserwacji	22435
Średnia cena	585,42 PLN
Mediana	511,89 PLN
Odchylenie standardowe	332,26 PLN
Wartość minimalna	10,00 PLN
Pierwszy kwartyl (25%)	372,00 PLN
Trzeci kwartyl (75%)	704,18 PLN
Wartość maksymalna	3812,45 PLN

Na histogramie przedstawionym na rysunku 13, który przedstawia rozkład dla całego zbioru danych, można na nim zauważać dominujące wartości z przedziału 300-700 PLN. Rozkład wydłużony jest prawostronnie, co wskazuje na występowanie bardzo wysokich wartości. Najprawdopodobniej jest to powiązane z sytuacją panującą na rynku w 2022 roku.



Rysunek 13 Histogram przedstawiający rozkład cen energii na całym zbiorze danych
Opracowanie własne

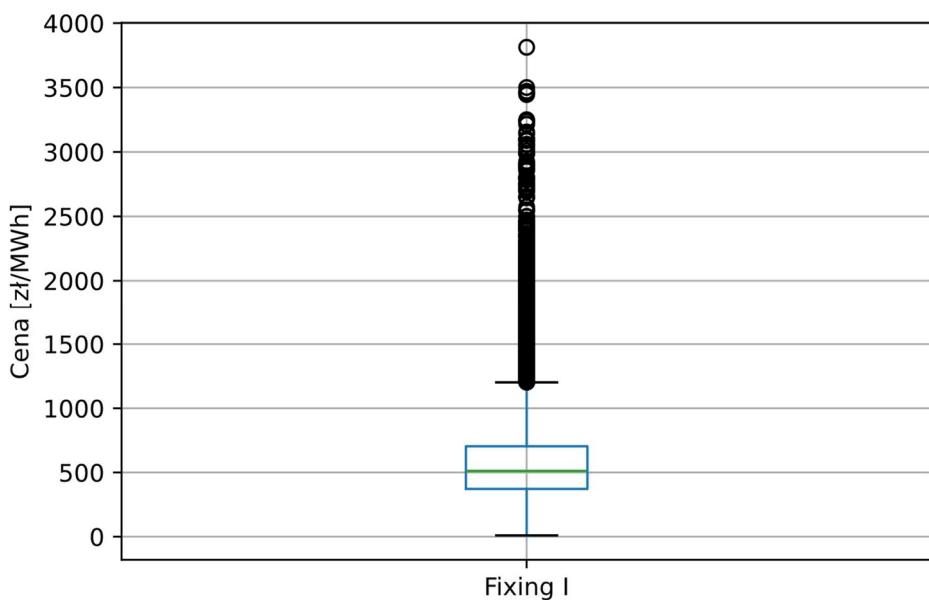
Aby potwierdzić takie założenie należy porównać histogram dla 2021 roku z histogramem dla roku 2022. Na rysunku 14 widać, że w 2022 nie tylko występowały bardziej ekstremalne wartości cen względem roku 2021, ale również przedział dominujących wartości jest znacznie szerszy. Dominowały wtedy wartości od 400 do 1000 PLN. W 2021 roku ceny były znacznie bardziej zwarte i znajdowały się bliżej dolnej granicy. Dominowały wartości z przedziału od 250 do 400 PLN.



Rysunek 14 Histogramy dla lat 2021 oraz 2022
Opracowanie własne

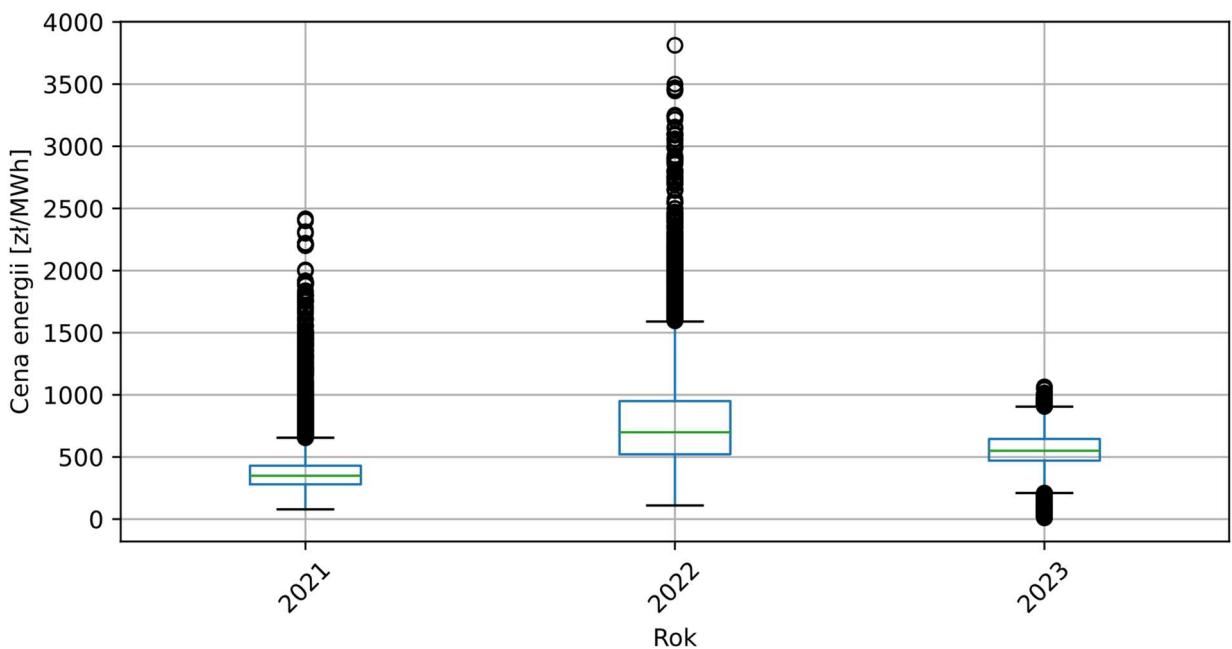
Bardzo pomocne w wizualizacji danych są również wykresy pudełkowe (boxploty). Krawędzie pudełka pokazują jakie wartości znajdują się między 25. a 75. percentylem. Linia wewnętrz pudełka reprezentuje medianę, natomiast „wąsy”, czyli linie wysunięte poza granice pudełek przedstawiają zakres większości danych. Wyjątkiem są dane przekraczające IQR (*interquartile range* - rozstęp międzykwartylowy) 1,5 raza. Rozstęp międzykwartylowy mierzony jest jako różnica między trzecim a pierwszym kwartylem [24].

Z wykresu pudełkowego dla całego okresu danych, przedstawionego na rysunku 15, można odczytać medianę w okolicach 500 PLN. Zgadza się to więc z wcześniejszymi obliczeniami. Można za to zauważyc, iż mimo że pudełko jest dosyć zwarte to wiele wartości wychodzi poza górne wąsy. Zauważalna jest zatem asymetria w kierunku wyższych cen.



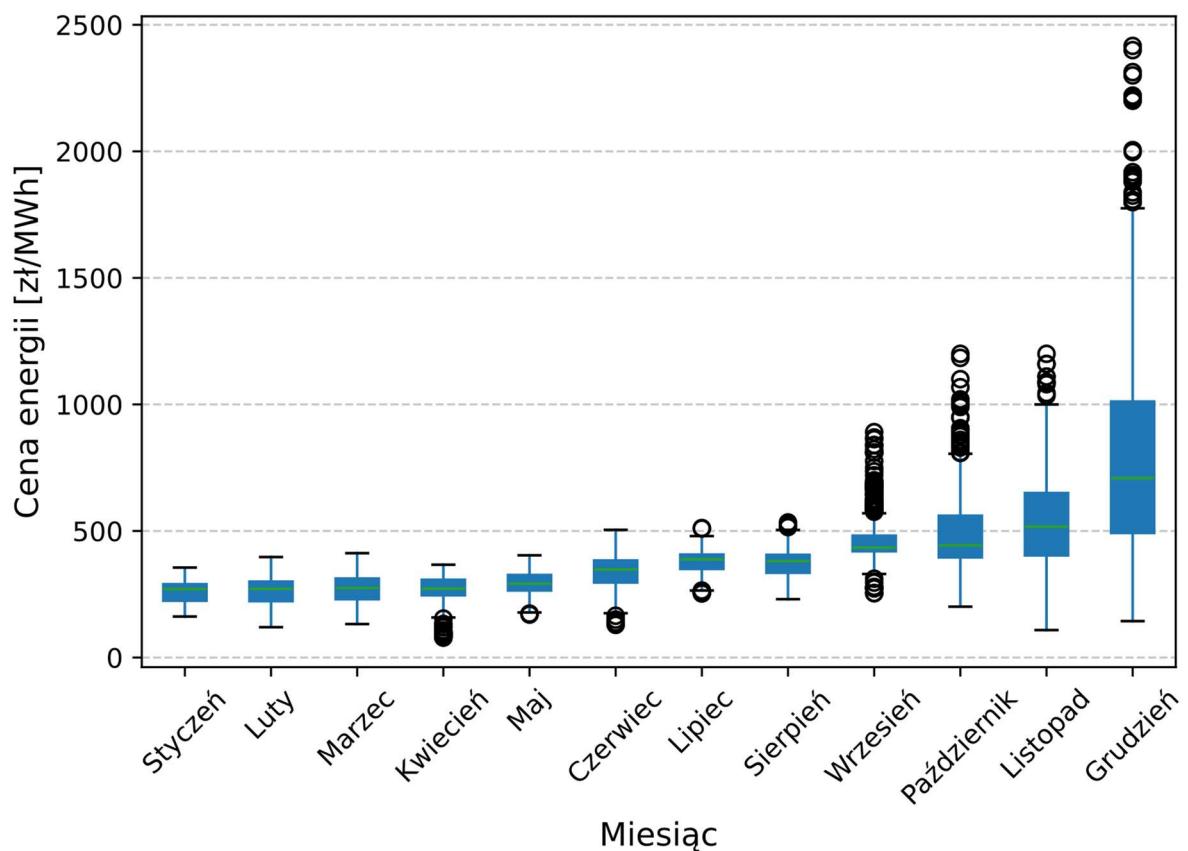
Rysunek 15 Wykres pudełkowy dla całego zbioru danych
Opracowanie własne

Bardziej dokładną analizę można przeprowadzić w momencie, gdy porównamy wykresy pudełkowe poszczególnych lat. Na rysunku 16, przedstawiającym takie właśnie zestawienie można zauważyc, że nie tylko w 2022 roku występowały wyjątkowo wysokie wartości. Już w roku 2021 można zauważyc wiele wartości poza górną granicą wąsów. Jednak wtedy pudełko było znacznie węższe, zatem w roku 2021 rozbieżność cen energii była zdecydowanie mniejsza. W roku 2023 natomiast mediana jest bardzo podobna do tej w roku 2021. Szerokość pudełek w tych latach również jest bardzo podobna, jednak pudełko w roku 2023 przesunięte jest w kierunku wyższych cen. Jednak w 2023 wartości znajdujące się poza wąsami są mniej rozproszone i znajdują się po obu stronach wąsów. Występują zatem odchylenia cen wyższych jak i niższych.



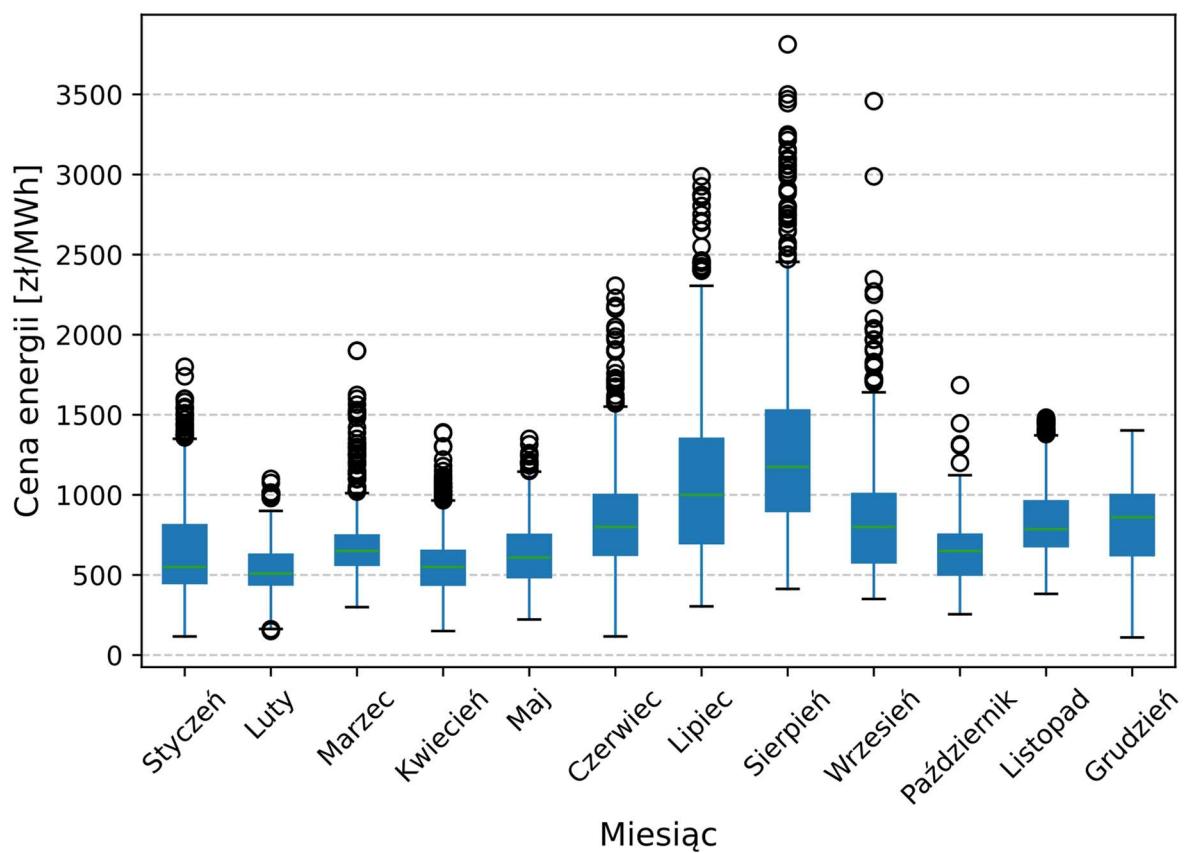
Rysunek 16 Wykresy pudełkowe cen energii w poszczególnych latach
Opracowanie własne

Rozłożenie wykresów pudełkowych na poszczególne miesiące 2021 roku, które przedstawia rysunek 17, pozwala zauważać różnice między początkiem a końcem tego roku. Pierwsze osiem miesięcy charakteryzowało się dużą stabilnością, pudełka były wąskie, a wartości poza wąsami występowały w niewielkich ilościach. Pojawiały się one po obu stronach wąsów z liczną przewagą w kierunku niższych cen, a nie wyższych. Zmianę można zaobserwować już we wrześniu 2021 roku, kiedy pojawia się znacznie więcej wartości wykraczających poza wąsy wykresu pudełkowego. W kolejnych miesiącach szerokość pudełka staje się coraz większa. Również wąsy obejmują znacznie szerszy zakres wartości, a mediana przesuwa się w stronę wyższych cen. Rok 2021 charakteryzował się zatem stabilnością cen w pierwszych trzech kwartałach, jednak w czwartym kwartale pojawiły się znaczne wzrosty cen. Liczne wartości odstające w końcówce tego roku można utożsamiać z dużą niepewnością na rynku spowodowaną narastaniem napięcia za wschodnią granicą Polski.



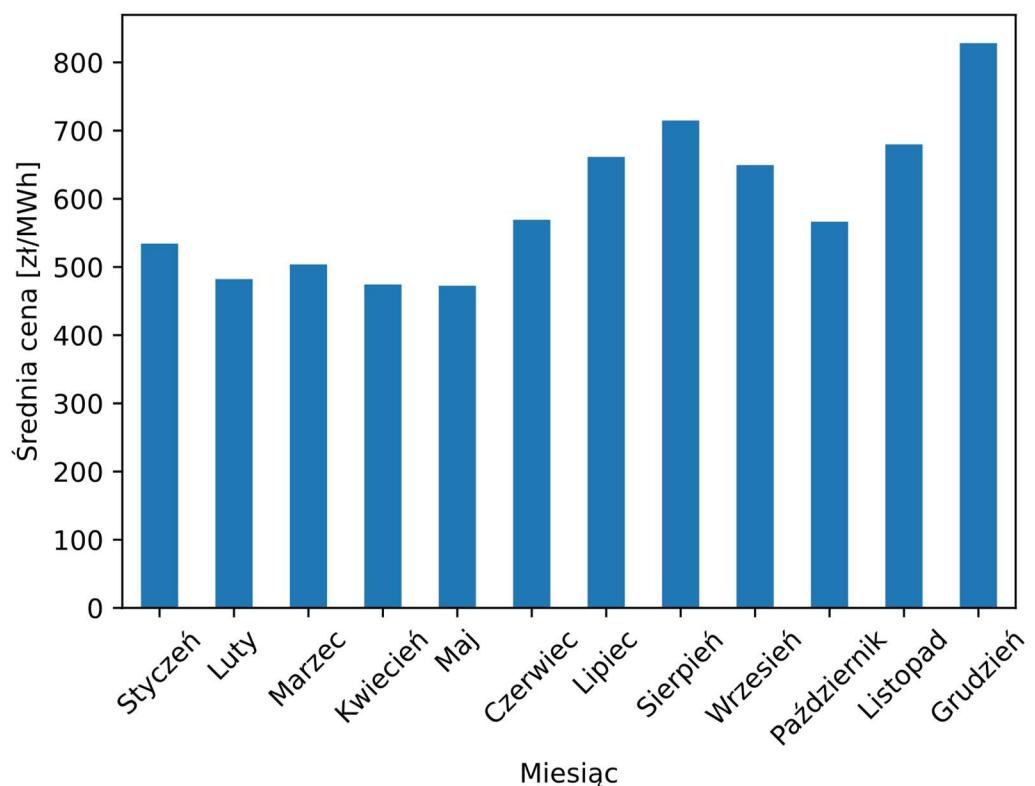
Rysunek 17 Rozkład cen energii w poszczególnych miesiącach 2021 roku
Opracowanie własne

Rysunek 18 przedstawia rozkład cen energii w poszczególnych miesiącach 2022 roku. Można na nim zauważyć, że rozkłady poszczególnych miesięcy są znacznie bardziej zróżnicowane niż w roku poprzednim. Wskazuje to na znacznie większą niestabilność rynku. W pierwszych miesiącach medianą jest wyższa niż w roku poprzednim, jednak utrzymuje się na w miarę stałym poziomie. Pudełka w okresie luty-maj są dosyć wąskie, występują natomiast wartości wykraczające poza zakres wąsów. W miesiącach letnich można zaobserwować nie tylko znaczny wzrost mediany, ale także znacznie większe rozproszenie wartości. Pudełka są szerokie, a zakres wartości, który obejmują wąsy rozciąga się od 200 do prawie 2500 PLN. Liczne wartości odstające sięgają ponad 3500PLN w sierpniu. Taki obraz danych sugeruje dużą niepewność na rynku energii przed zbliżającym się okresem grzewczym. Ten z kolei charakteryzuje się już niższą medianą i mniejszym rozproszeniem danych, co sugeruje stabilizację na rynku w końcowce roku.



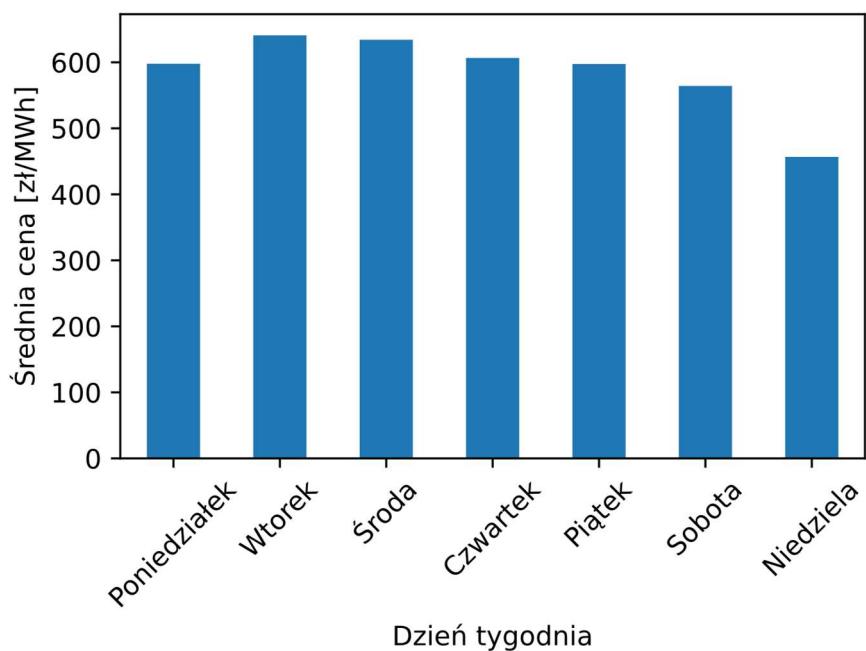
Rysunek 18 Rozkład cen energii w poszczególnych miesiącach 2022 roku
Opracowanie własne

Rysunek 19 przedstawia z kolei średnią cenę w poszczególnych miesiącach, obliczoną na całym zbiorze danych. Można na nim zauważać bardzo wysoką średnią cenę w grudniu oraz okresie letnim. Te wysokie wartości wynikają ze wcześniejszych wspomnianych wahań na rynku w roku 2022 oraz pod koniec roku 2021.



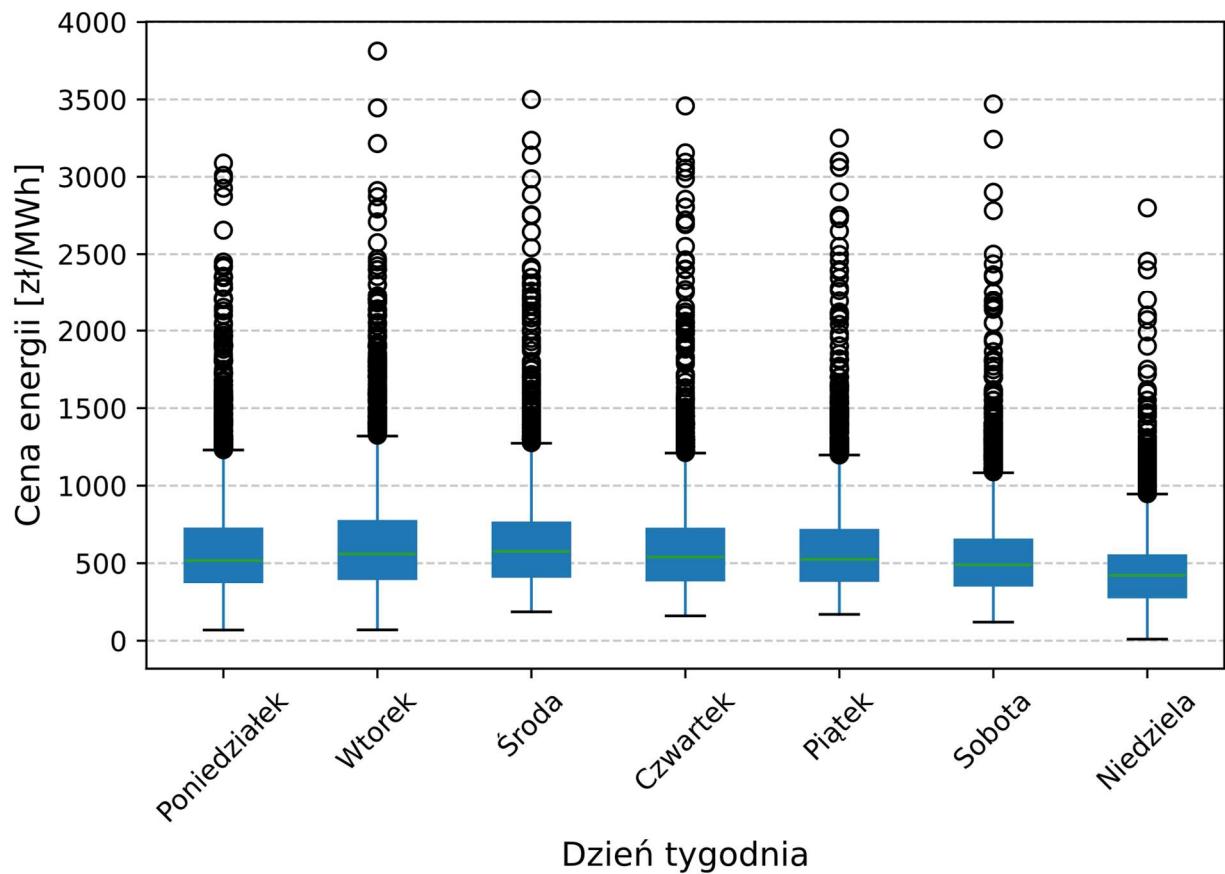
Rysunek 19 Średnie ceny energii w poszczególnych miesiącach
Opracowanie własne

Na rysunku 20 można zauważyć, że średnia cena w ciągu tygodnia utrzymuje się na mniej więcej stałym poziomie. Jest zauważalnie niższa w niedzielę, jednak nie są to wartości znaczące.



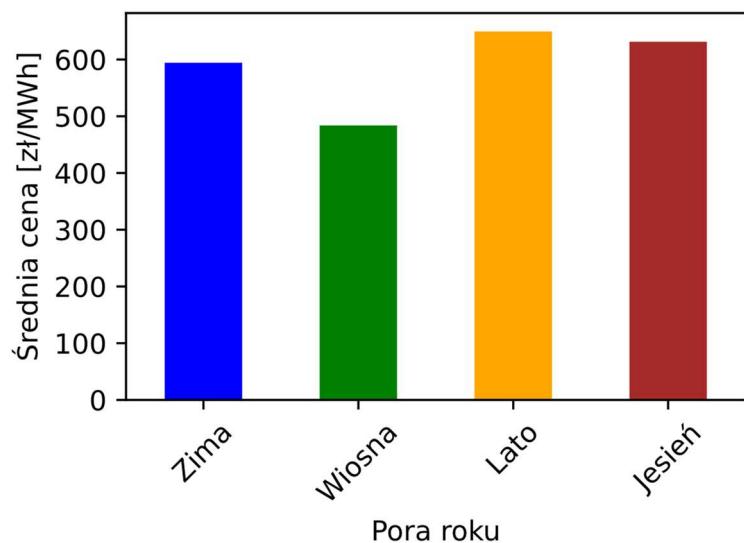
Rysunek 20 Średnia cena energii w poszczególnych dniach tygodnia
Opracowanie własne

Rozkład cen energii w poszczególnych dniach tygodnia, przedstawiony na rysunku 21, potwierdza brak znaczących wahania w ciągu tygodnia. Pudełka są mniej więcej tej samej szerokości, a wartości odstające występują na przestrzeni całego tygodnia. Zakres wąsów również jest w miarę stały.



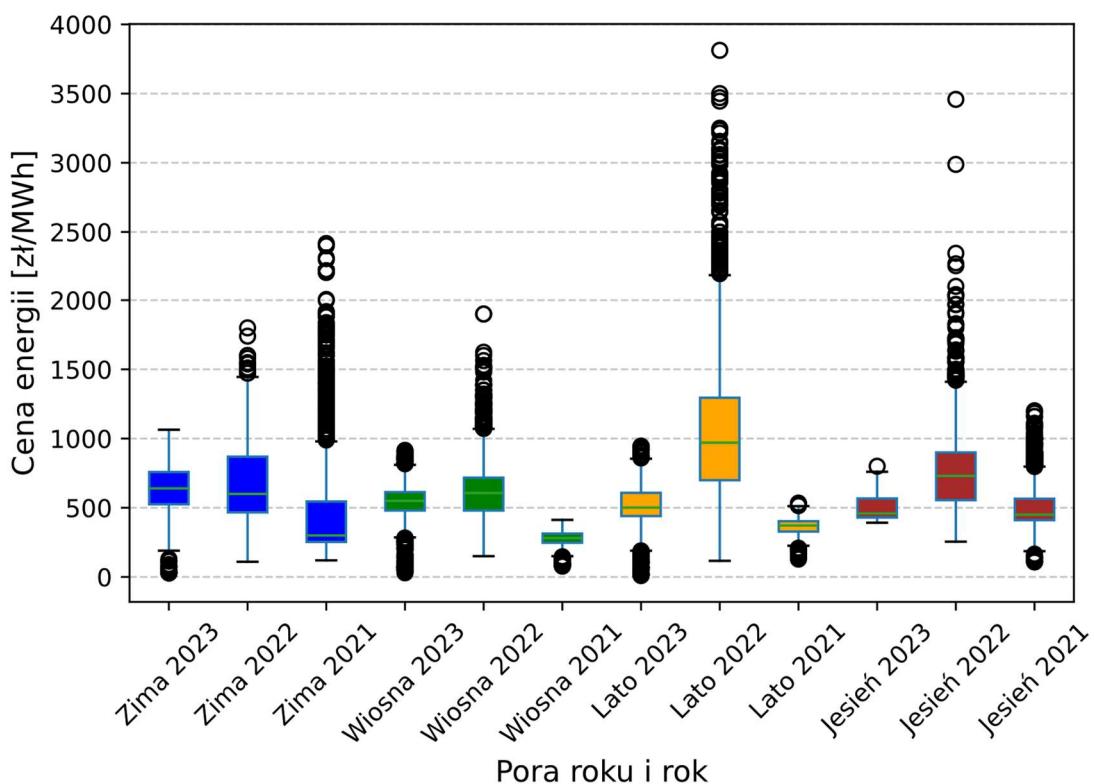
Rysunek 21 Rozkład cen energii w poszczególnych dniach tygodnia
Opracowanie własne

Średnie ceny energii w poszczególnych porach roku są na mniej więcej podobnym poziomie i oscylują w granicach od 500 do nieco ponad 600 PLN. Nieco niższa średnia cena energii wiosną może świadczyć o wahaniach cen pod koniec 2021 roku oraz pod koniec lata 2022 roku.



Rysunek 22 Średnia cena energii w poszczególnych porach roku
Opracowanie własne

Bliższe spojrzenie na pory roku w poszczególnych latach pozwala zauważyc jak duże wahania wystąpiły latem i jesienią 2022 roku oraz zimą 2021 roku. Te trzy sezony cechują się wyjątkowo dużym rozrzutem danych odstających. Rozmiar pudełka dla lata 2022 jest ekstremalnie szeroki, nawet względem jesieni 2022 roku. Porównując do siebie te wykresy pudełkowe należy dostrzec bardzo stabilne okresy w 2021 oraz 2023 roku.



Rysunek 23 Rozkład cen energii w poszczególnych porach roku i latach
Opracowanie własne

4 Analiza czynników wpływających na ceny energii elektrycznej

Autor: Mikołaj Kalisz

Podstawą dobrego modelu jest selekcja danych. Sieci neuronowe wyszukują wzorce i wyodrębniają cechy, które pozwalają uzyskać ostateczny wynik predykcji. Dlatego dane wejściowe muszą być należycie dobrane tak, by wspierały proces uczenia. Wprowadzenie danych niezwiązanych z oczekiwana wartością utrudnia modelowi osiągnięcie dobrych wyników. Z kolei zastosowanie nadmiernej liczby danych, zwłaszcza takich, które silnie ze sobą korelują (a przez to nie wnoszą nowych informacji), może prowadzić do przeuczenia (ang. overfitting). W tym dziale zostanie przeprowadzona analiza dostępnych danych, mająca na celu ułatwienie wyboru odpowiedniego podzbioru do nauki.

Ze względu na zjawiska poza rynkowe, jak wojna, rozkład danych został zaburzony, co pokazano w rozdziale 0. W celu uwzględnienia tychże zaburzeń w analizie zostanie przeprowadzona dokładniejsza analiza korelacji z dodatkowym podziałem na lata, pory roku oraz kombinację tych dwóch wartości. Dodatkowo należy się przyjrzeć wykresom, by uchwycić potencjalne zależności nieliniowe. W ten sposób można zapobiec nieuwzględnieniu kluczowych danych, potencjalnie poprawiając przyszłe wyniki modelu.

4.1 Zapotrzebowanie na energię

Autor: Mikołaj Kalisz

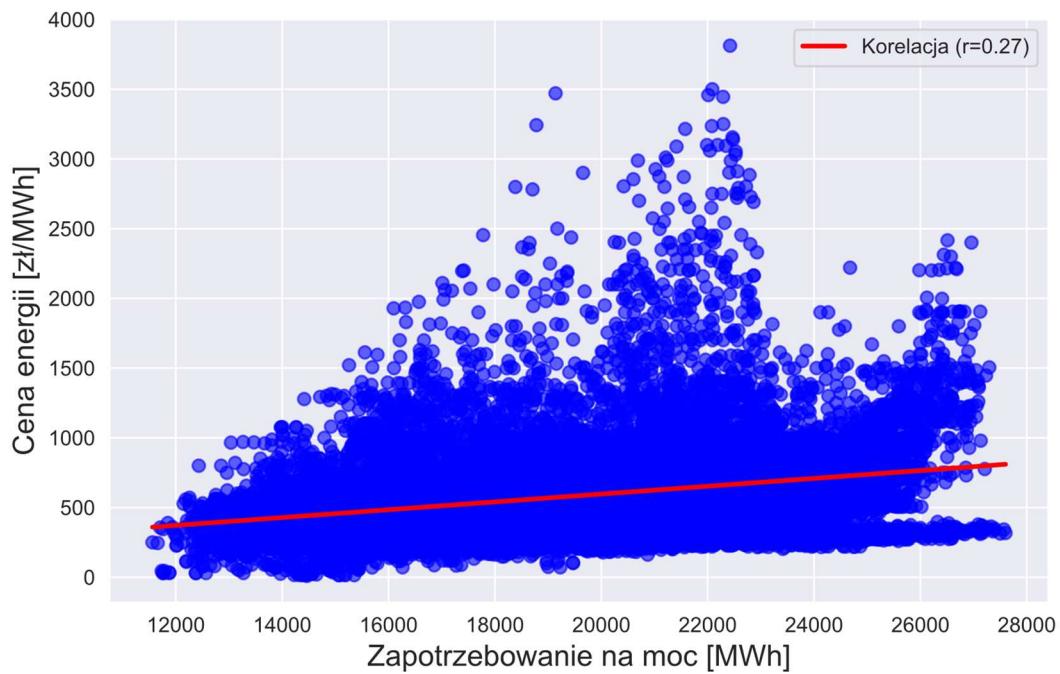
Jako że rynkiem kieruje popyt i podaż, przy analizie czynników wpływających na cenę energii nie może zabraknąć danych o zapotrzebowaniu na energię. W dzisiejszych czasach, w dobie urządzeń elektrycznych, jest to ważny czynnik kierujący ceną.

W zbiorze danych mamy cechy takie jak:

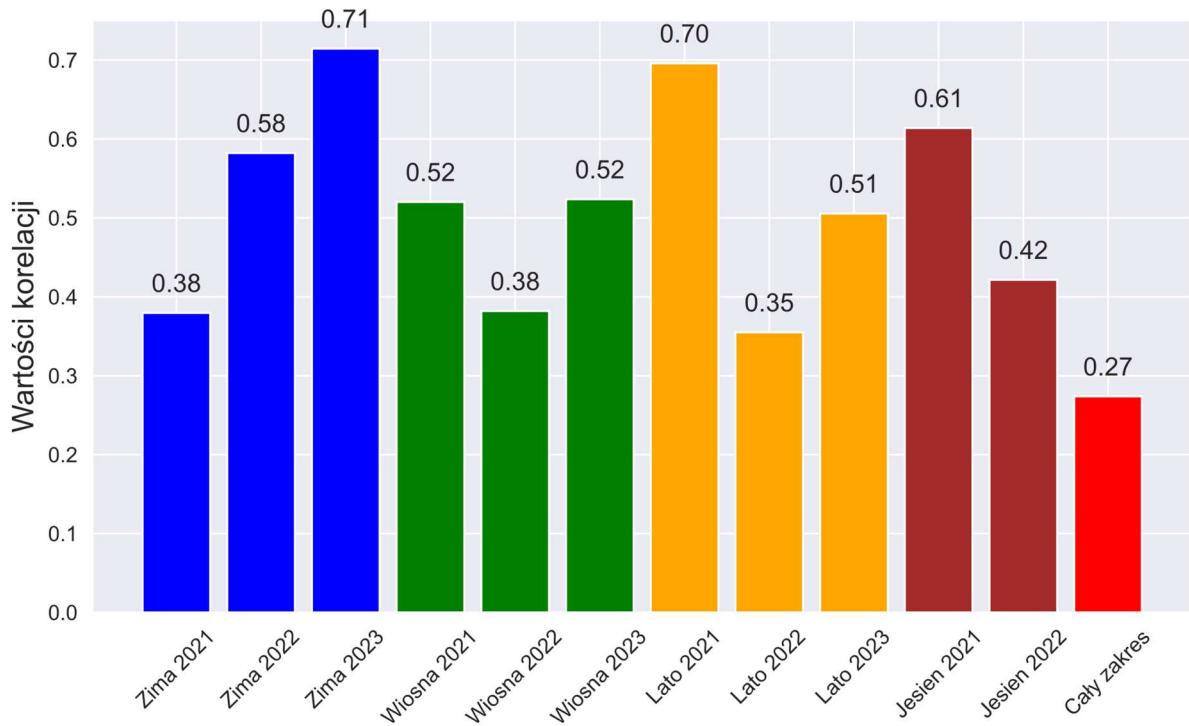
- Zapotrzebowanie na moc.
- Rezerwa mocy ponad zapotrzebowanie.
- Rezerwa mocy poniżej zapotrzebowania.

Należy sprawdzić czy istnieje korelacja tych wartości z ceną energii oraz co wpływa na zapotrzebowanie.

Wykres zależności zapotrzebowania na moc z ceną, wraz z linią najlepszego dopasowania, jest przedstawiony na rysunku 24. Na pierwszy rzut oka nie widać zależności między tymi danymi, korelacja na poziomie $r = 0,27$ jest słabą korelacją, jednakże po pogrupowaniu danych na pory roku w konkretnych latach można zauważać większą zależność tych dwóch wielkości. Wykres słupkowy przedstawiający tą zależność znajduje się na rysunku 25.



Rysunek 24 Wykres zależności ceny energii od zapotrzebowania. Z narysowaną linią najlepszego dopasowania. Opracowanie własne

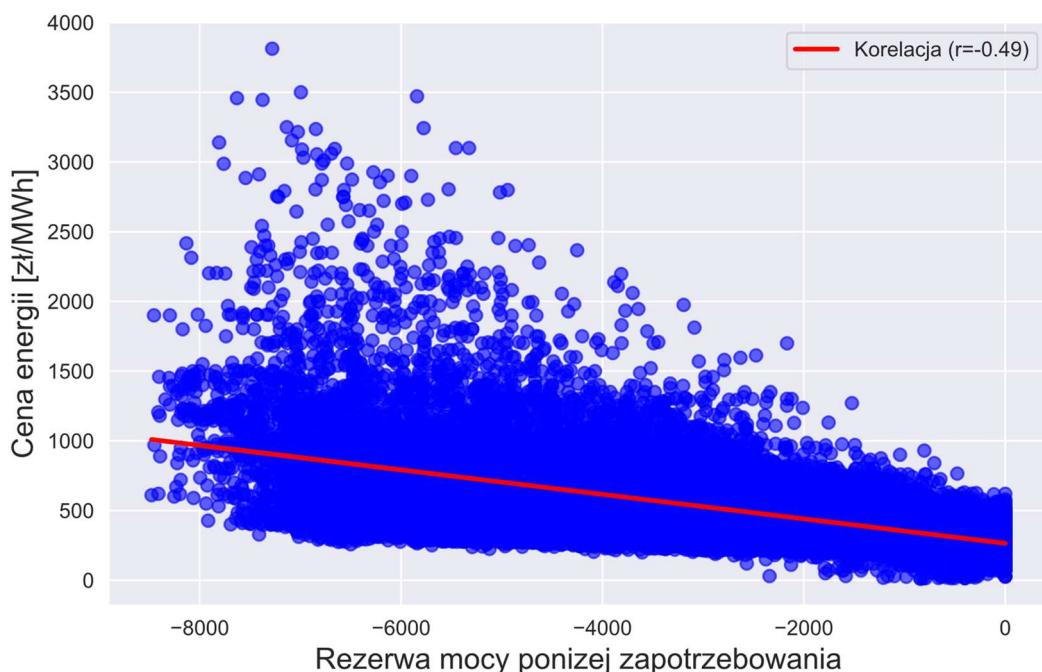


Rysunek 25 Wykres słupkowy przedstawiający wartości korelacji danych o zapotrzebowaniu na moc względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne

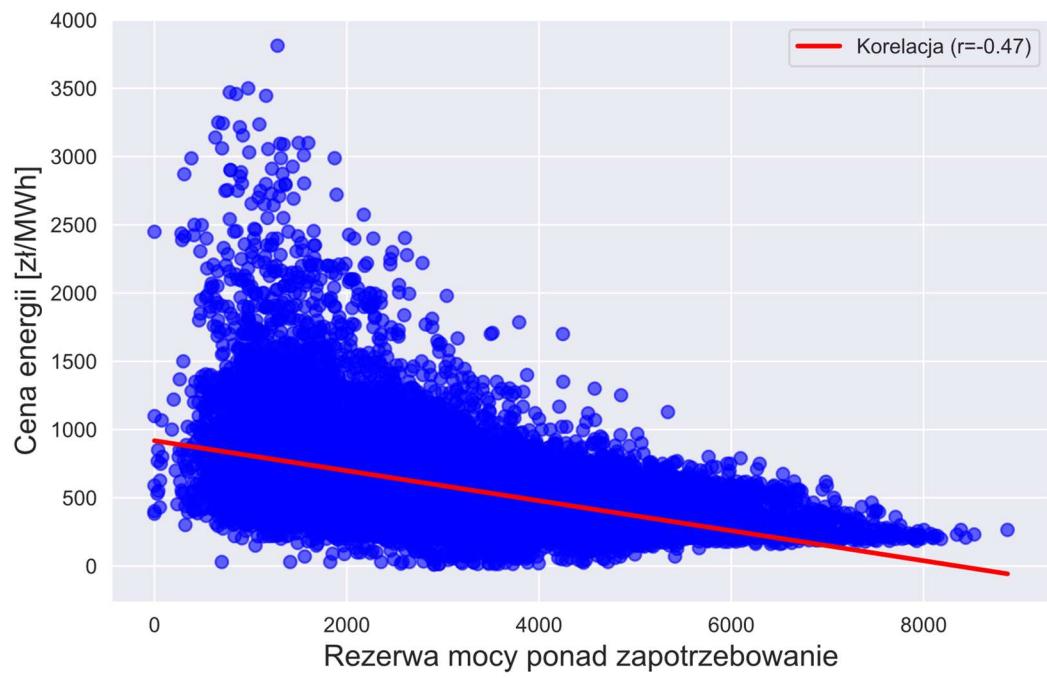
Jak widać korelacja między wartościami może być inna w czasie, co można zaobserwować na rysunku 25. Mimo słabych korelacji w niektórych przedziałach czasowych, w większości występuje korelacja umiarkowana i silna. W tym wypadku

mimo słabej korelacji dla całego zbioru można zauważyc, że informacja o zapotrzebowaniu na moc może być potencjalnym źródłem informacji dla modelu.

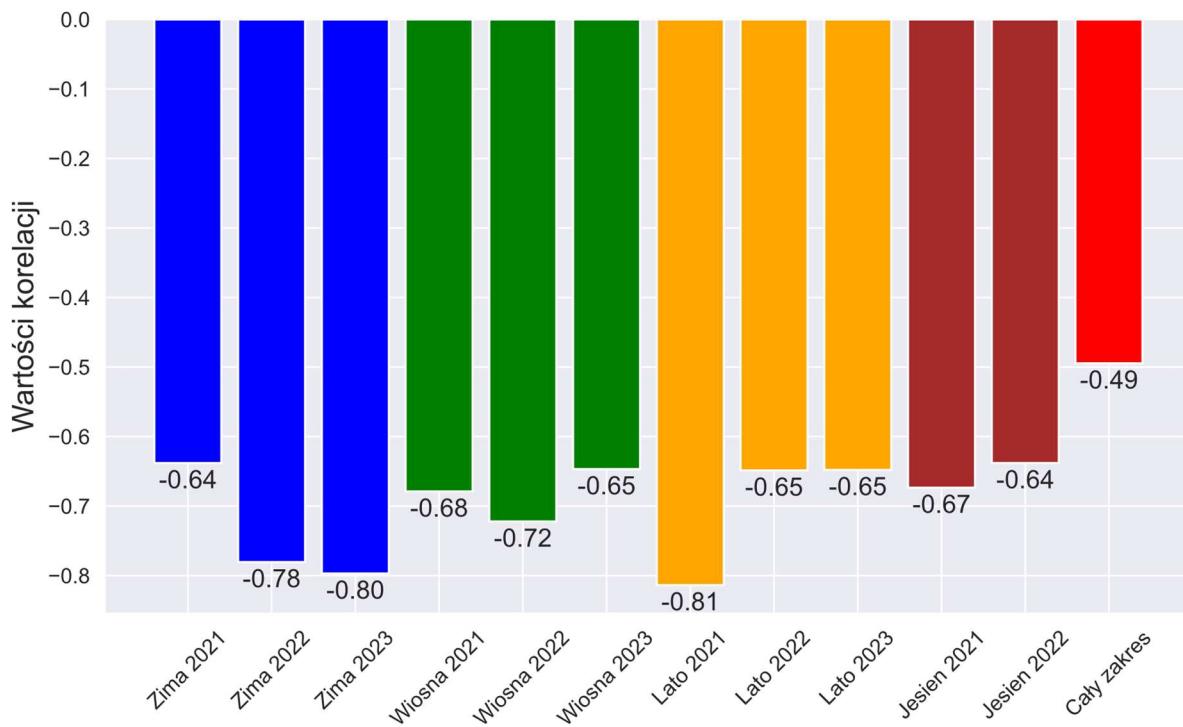
Wartościami powiązanymi z zapotrzebowaniem na moc są cechy rezerwa mocy ponad zapotrzebowanie oraz rezerwa mocy poniżej zapotrzebowania. Charakteryzują się one wyższą bezwzględną wartością korelacji z ceną w pełnym zakresie zbioru danych, odpowiednio $r_{abs} = 0,47$ oraz $r_{abs} = 0,49$, gdzie r_{abs} to wartość bezwzględna (siła) korelacji.. Na rysunkach 26 i 27 możemy zaobserwować relację między tymi cechami. Dodatkowo w celu dokładniejszej analizy porównawczej z cechą „zapotrzebowanie na moc” ponownie zostanie dokonany podział na lata i pory roku.



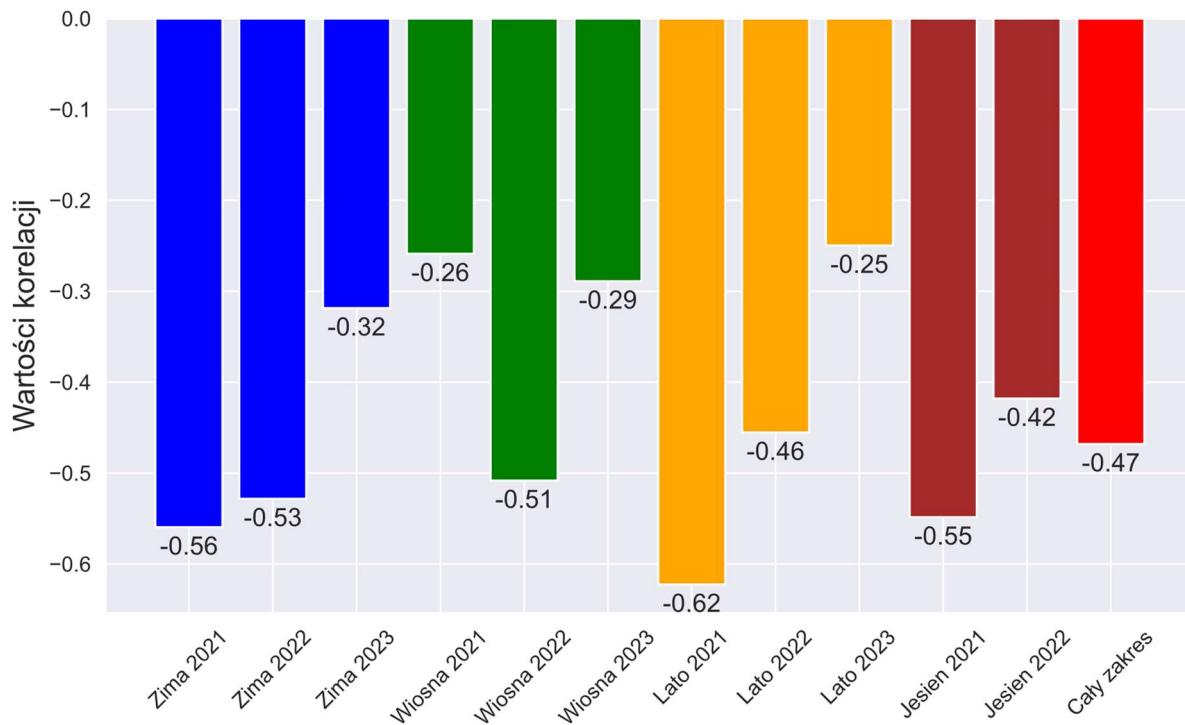
Rysunek 26 Wykres zależności ceny energii od rezerwy mocy poniżej zapotrzebowania, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne



Rysunek 27 Wykres zależności ceny energii od rezerwy mocy ponad zapotrzebowanie, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne



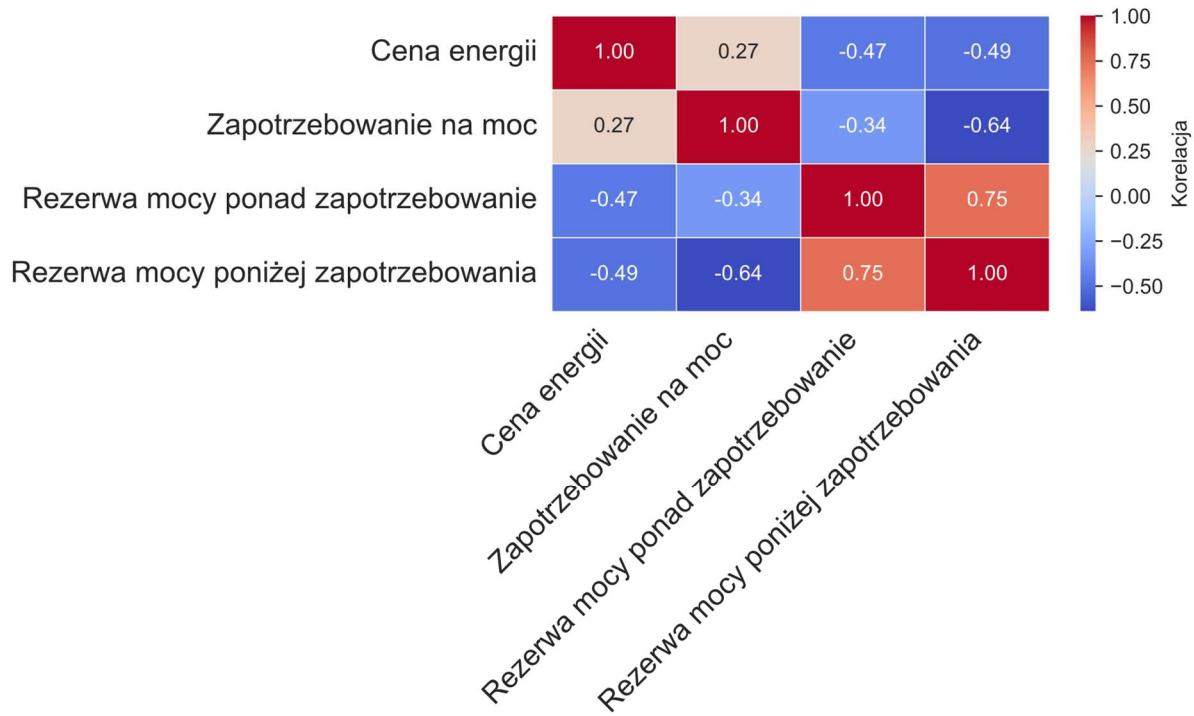
Rysunek 28 Wykres słupkowy przedstawiający wartości korelacji danych rezerwy mocy ponizej zapotrzebowania względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne



Rysunek 29 Wykres słupkowy przedstawiający wartości korelacji danych wartości rezerwy mocy ponad zapotrzebowanie względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne

Na rysunkach 28 i 29 można zobaczyć wartości korelacji dla poszczególnych okresów. Wartości bezwzględne dla danych o rezerwie mocy poniżej zapotrzebowania są zauważalnie wyższe niż dla zapotrzebowania mocy oraz rezerwy mocy ponad zapotrzebowanie. Dla dwóch okresów (Zima 2023, Lato 2021) występuje silna korelacja ($|r| \geq 0.8$, gdzie r to wartość (siła) korelacji).

Na rysunku 30 można zobaczyć macierz korelacji, powstałą w celu sprawdzenia czy nie występuje korelacja między poszczególnymi cechami. Wynika z niej, że podane dane korelują także ze sobą. Szczególnie rezerwa mocy poniżej i ponad zapotrzebowanie oraz rezerwa mocy poniżej zapotrzebowania z zapotrzebowaniem na moc. W takiej sytuacji zalecane jest wybranie tylko jednej cechy do nauki modelu, gdyż potencjalnie są to źródła tej samej informacji.

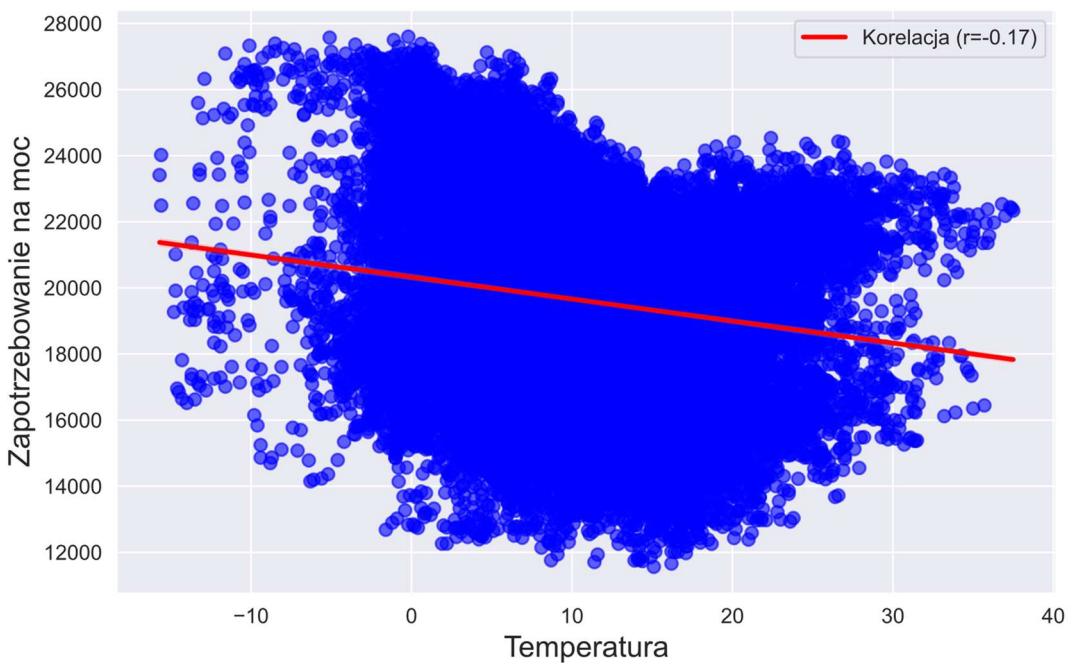


Rysunek 30 Macierz korelacji między wybranymi cechami. Opracowanie własne

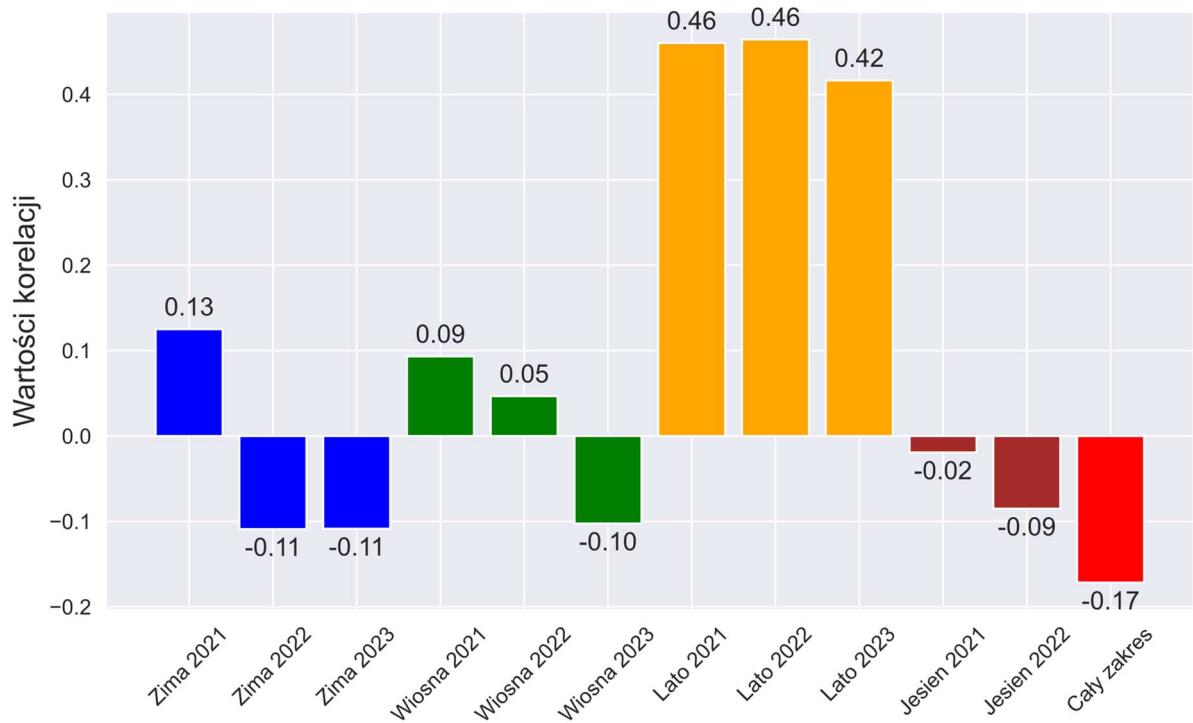
4.1.1 Wpływ temperatury na zapotrzebowanie i ceny energii

Autor: Mikołaj Kalisz

W zależności od temperatury otoczenia, uruchamia się systemy ogrzewania, bądź klimatyzację, potencjalnie wpływając na cenę energii. Jak możemy zauważyć na rysunku 31, dane o temperaturze wydają się nie korelować z zapotrzebowaniem na moc. Jednakże można zauważyć lekkie zmniejszenie wartości zapotrzebowania w okolicach od 12 do 15 stopni. Może to wynikać z faktu, że jest to optymalna temperatura dla człowieka niewymagająca, ani ogrzewania, ani chłodzenia. Poniżej tego zakresu zapotrzebowanie nieznacznie rośnie, a powyżej pozostaje stałe.



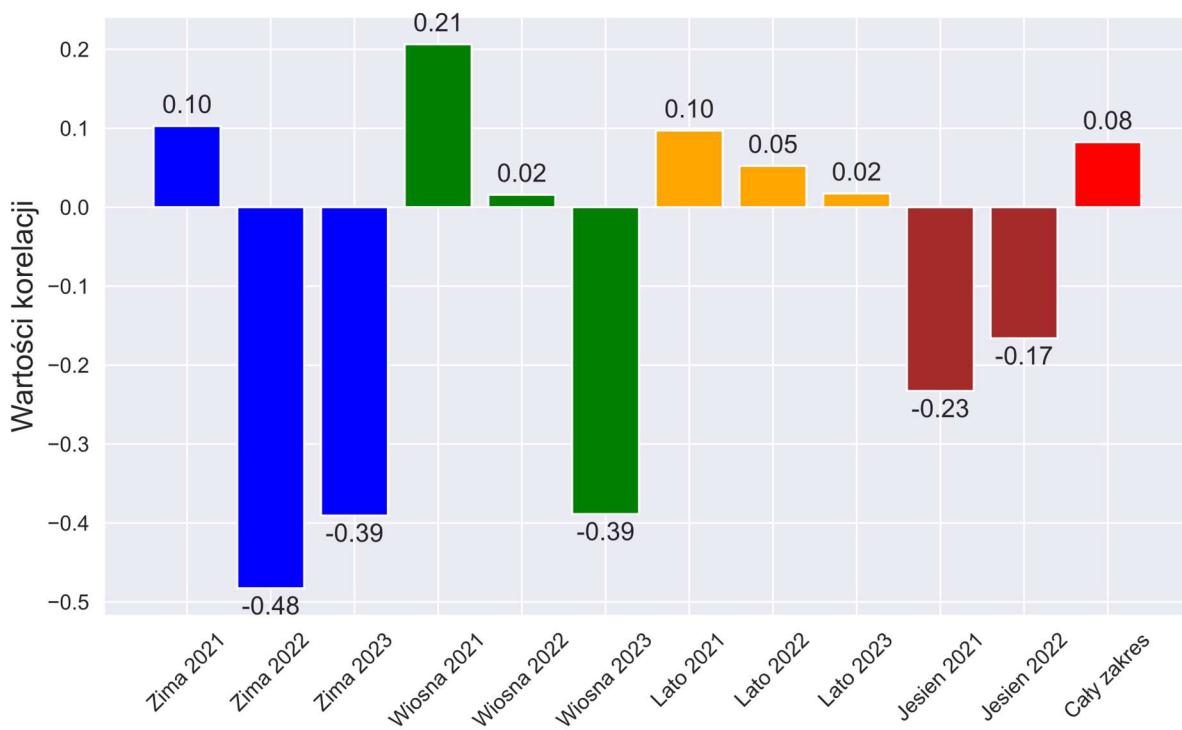
Rysunek 31 Wykres zależności temperatury od zapotrzebowania na moc, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne



Rysunek 32 Wykres słupkowy przedstawiający wartości korelacji danych wartości temperatury względem zapotrzebowania na moc, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne

Na rysunku 32 przedstawiono wartości korelacji temperatury z zapotrzebowaniem na moc dla poszczególnych przedziałów czasowych. Można zauważać średnią dodatnią korelację latem, lecz poza tym sezonem nie ma innej znaczącej wartości.

W przypadku odniesienia do ceny energii także nie widać mocnej korelacji, co pokazano na rysunku 33. Zbliżamy się do wartości średniej siły korelacji zimą 2022, 2023 oraz wiosną 2023, jednakże to zbyt małe wartości, by uwzględnić temperaturę w procesie uczenia modelu.



Rysunek 33 Wykres słupkowy przedstawiający wartości korelacji wartości temperatury względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne

4.2 Odnawialne źródła energii

Autor: Mikołaj Kalisz

Na przestrzeni ostatnich lat odnawialne źródła energii zwiększają swój udział w rynku. W 2024 roku produkcja energii elektrycznej ze źródeł bezemisyjnych wyniosła 26,5%. Dla porównania 4 lata wcześniej, w 2020 roku ten udział wyniósł zaledwie 13,3%³. Rosnący udział tańszej energii potencjalnie zmniejszy jej cenę. Jednocześnie wahania w produkcji ze źródeł bezemisyjnych mogą mocniej wpływać na dzienne zmiany ceny rynkowej. W Polsce głównymi źródłami energii odnawialnej są energia wiatrowa (14%, stan na 2024 rok) oraz energia słoneczna (11,3%, stan na 202 rok)⁴.

³ Źródło: [Energy instrat] (<https://energy.instrat.pl/system-elektroenergetyczny/produkcja-udzial-oze/>)

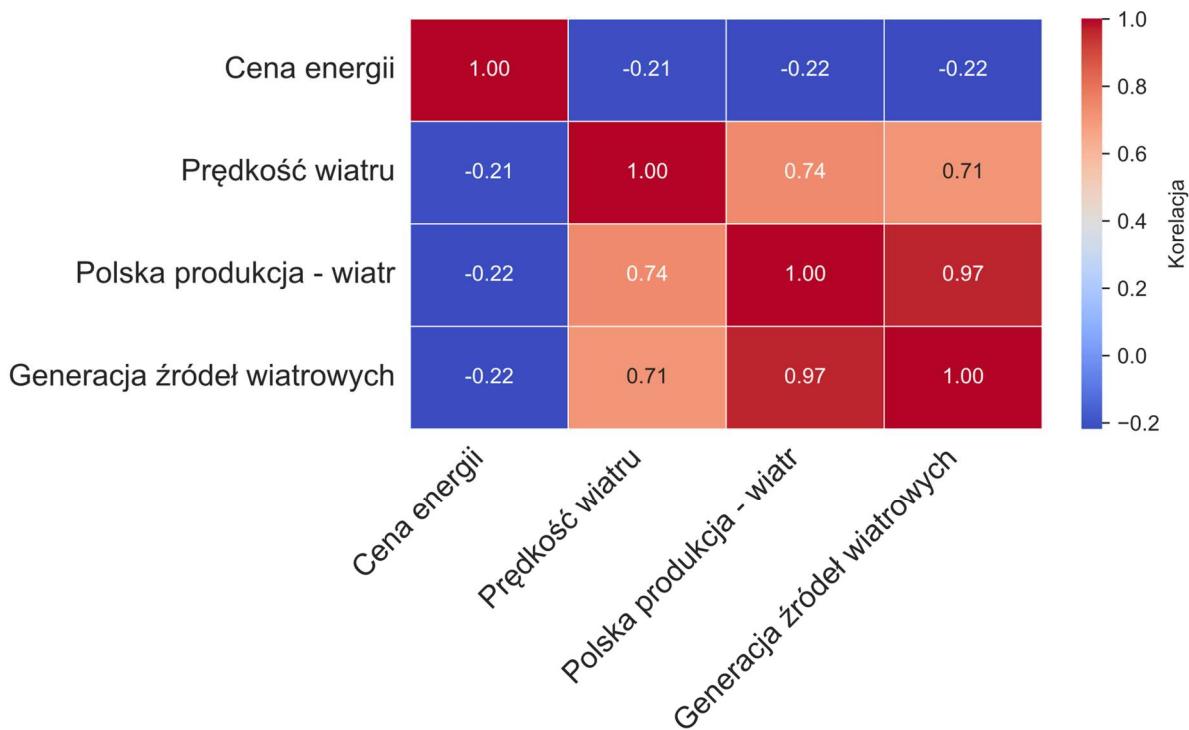
⁴ Źródło: [Energy instrat] (<https://energy.instrat.pl/system-elektroenergetyczny/produkcja-entsoe/>)

4.2.1 Produkcja energii wiatrowej

Autor: Mikołaj Kalisz

W zbiorze danych posiadamy trzy cechy powiązane z wiatrem w Polsce:

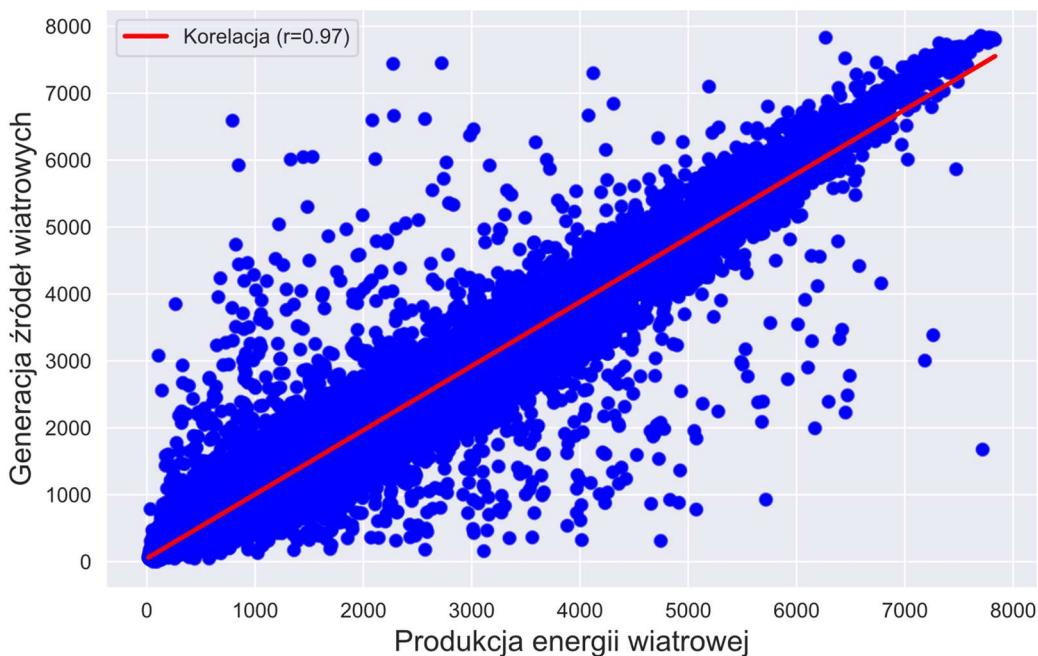
- Prędkość wiatru.
- Polska produkcja - wiatr.
- Generacja źródeł wiatrowych.



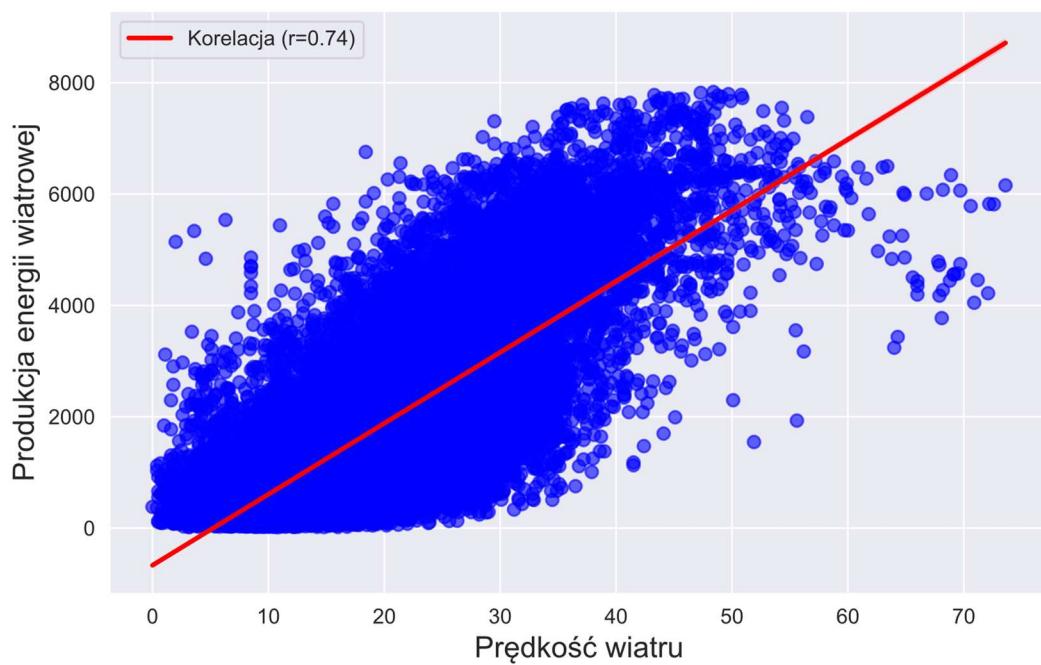
Rysunek 34 Macierz korelacji cech powiązanych z wiatrem oraz ceny energii. Opracowanie własne

Jak możemy zauważyć na macierzy korelacji, przedstawionej na rysunku 34, wszystkie z wymienionych cech mają silną bądź bardzo silną korelację ze sobą. Korelacja między Polska produkcja – wiatr, a Generacja źródeł wiatrowych na poziomie $r = 0,97$ oznacza bardzo silną korelację, a patrząc na rysunku 35 można zauważyć, że większość danych znajduje się przy linii najlepszego dopasowania oznaczającą liniową zależność między zmiennymi. Jednocześnie z macierzy korelacji możemy odczytać korelacje wybranych cech z ceną energii. Mają wartości poniżej $|r| \leq 0,3$ co oznacza słabą siłę korelacji.

Z pokazanej na rysunku 36 zależności produkcji energii wiatrowej od prędkości wiatru można wnioskować, że istnieje zależność między tymi wartościami. Dane są nieznacznie bardziej rozproszone niż w przypadku zależności generacji źródeł wiatrowych i produkcji.



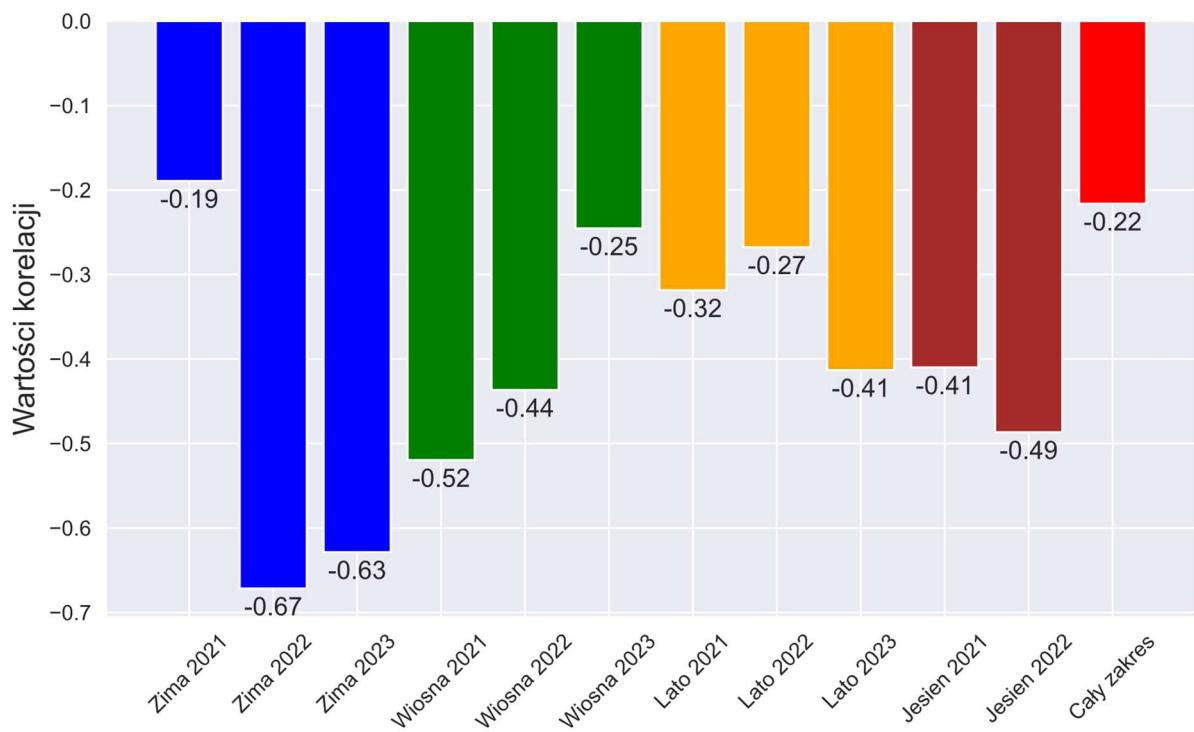
Rysunek 35 Wykres zależności generacji źródeł wiatrowych od produkcji energii wiatrowej, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne



Rysunek 36 Wykres zależności produkcji energii wiatrowej od prędkości wiatru, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne

Z dokładniejszej analizy generacji źródeł wiatrowych przedstawionej na rysunku 37 z podziałem na pory roku, można zauważać umiarkowaną korelację dla paru okresów, a dla większości słabą. Dla cech Polska produkcja - wiatr oraz prędkość

wiatru wartości są podobne. Z racji na silną wzajemną korelację z tych trzech cech preferowany jest wybór jednej, najlepiej skorelowanej z ceną energii – Polska produkcja – wiatr.



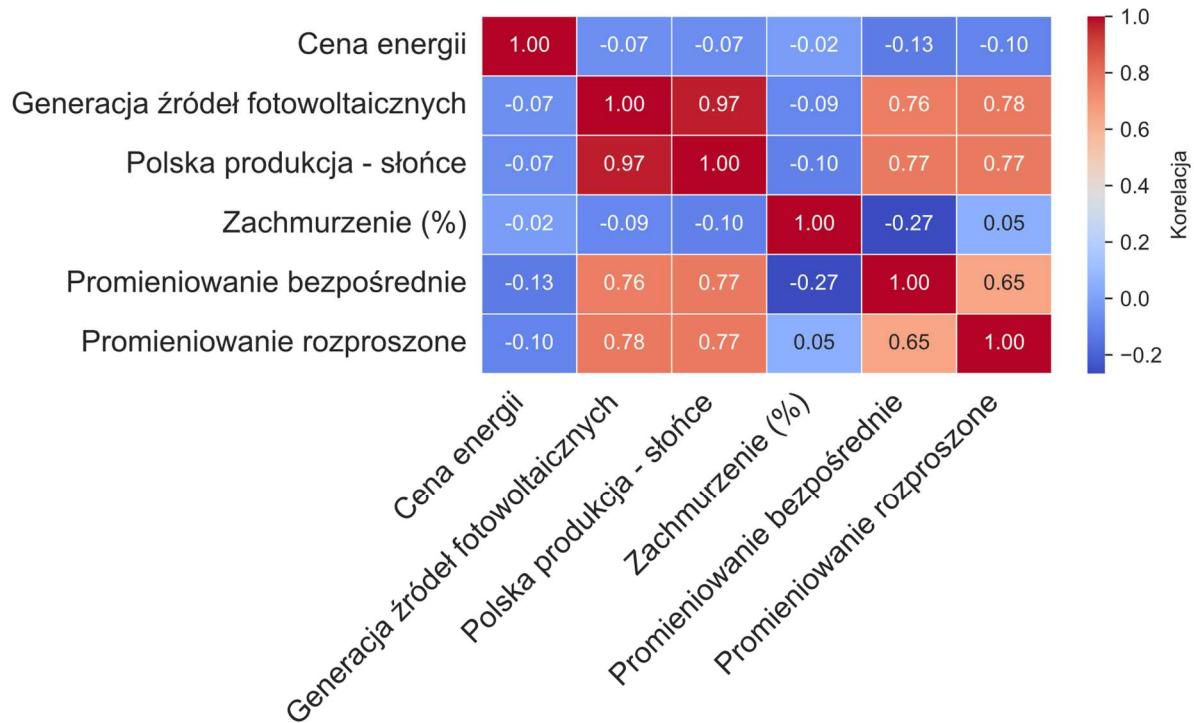
Rysunek 37 Wykres słupkowy przedstawiający wartości korelacji wartości generacji źródeł wiatrowych względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne

4.2.2 Produkcja energii słonecznej

Autor: Mikołaj Kalisz

Energia słoneczna, zaraz po energii wiatrowej, jest drugim największym udziałowcem rynku w kategorii energii odnawialnej. W danych znajdują się następujące cechy powiązane z tym tematem:

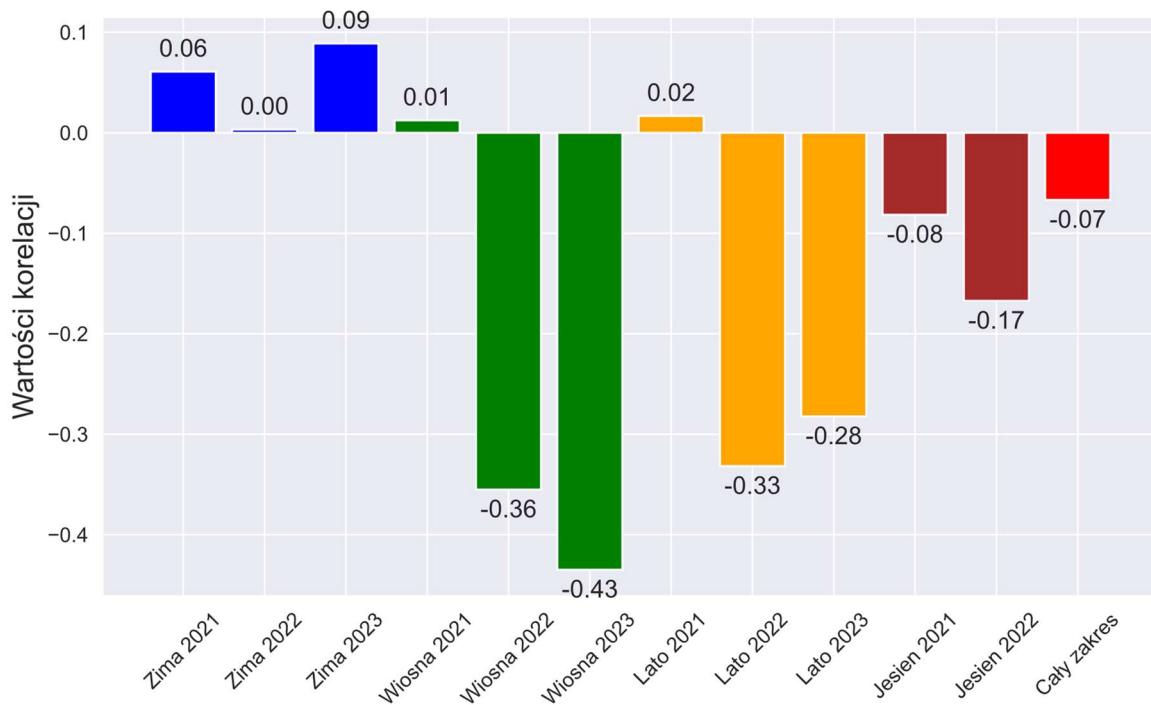
- Generacja źródeł fotowoltaicznych.
- Polska produkcja – słońce.
- Promieniowanie bezpośrednie.
- Promieniowanie rozproszone.
- Zachmurzenie – procent



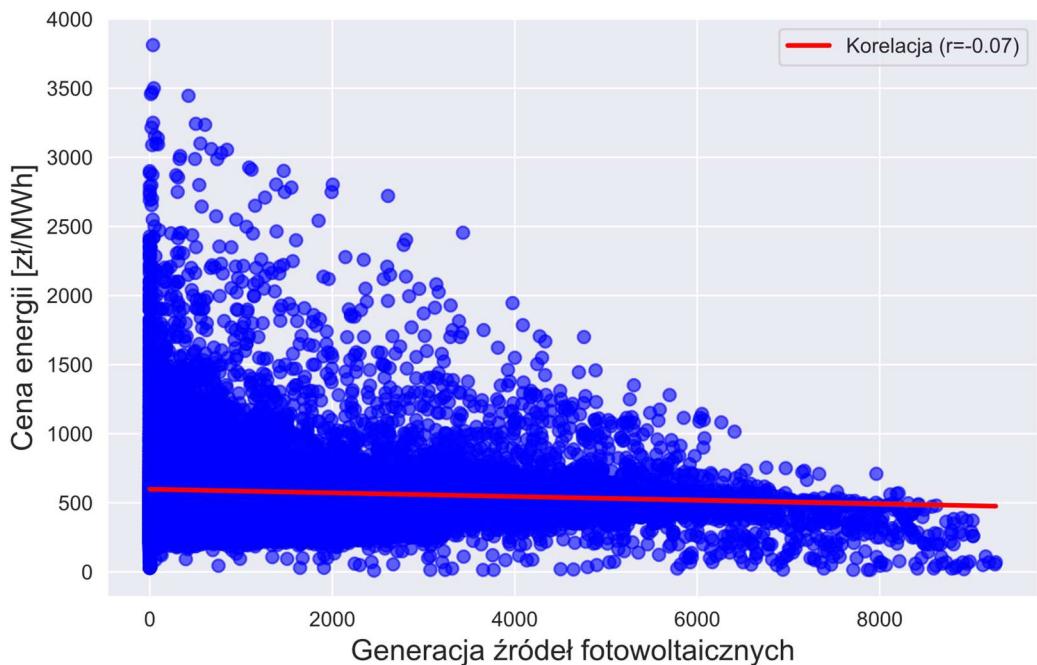
Rysunek 38 Macierz korelacji cech powiązanych ze słońcem oraz ceną energii. Opracowanie własne

Na macierzy korelacji z rysunku 38 można wyczytać, że cecha generacja źródeł fotowoltaicznych ma bardzo silną korelację z Polska produkcja – słońce. Dodatkowo wykazują one silną korelację z cechami promieniowania bezpośredniego i rozproszonego. Między promieniowaniami występuje silna korelacja. Wszystkie z podanych cech wykazują bardzo słabą korelację z ceną energii.

W przypadku podziału danych na pory roku, przedstawionym na rysunku 39 można zauważać wciąż niewielkie wartości siły korelacji, z wyróżniającymi się czterema okresami o umiarkowanej sile korelacji. Jednakże na wykresie zależności wartości ceny i generacji źródeł fotowoltaicznych, na rysunku 40, można zauważać zależność. W przypadku większej generacji energii z tego typu źródła występują niższe wartości ceny energii.



Rysunek 39 Wykres słupkowy przedstawiający wartości korelacji wartości generacji źródeł fotowoltaicznych względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne



Rysunek 40 Wykres zależności ceny energii od generacji źródeł fotowoltaicznych wraz z linią najlepszego dopasowania. Opracowanie własne

Potencjalnie model może wyciągnąć jakąś informację z tych danych, więc przekazanie na wejście przykładowo cechy Generacja źródeł fotowoltaicznych może przynieść lepsze wyniki. Wraz ze wzrostem udziału tego typu źródeł w produkcji może nastąpić zwiększenie znaczenia tych danych.

4.2.3 Produkcja z innych źródeł

Autor: Mikołaj Kalisz

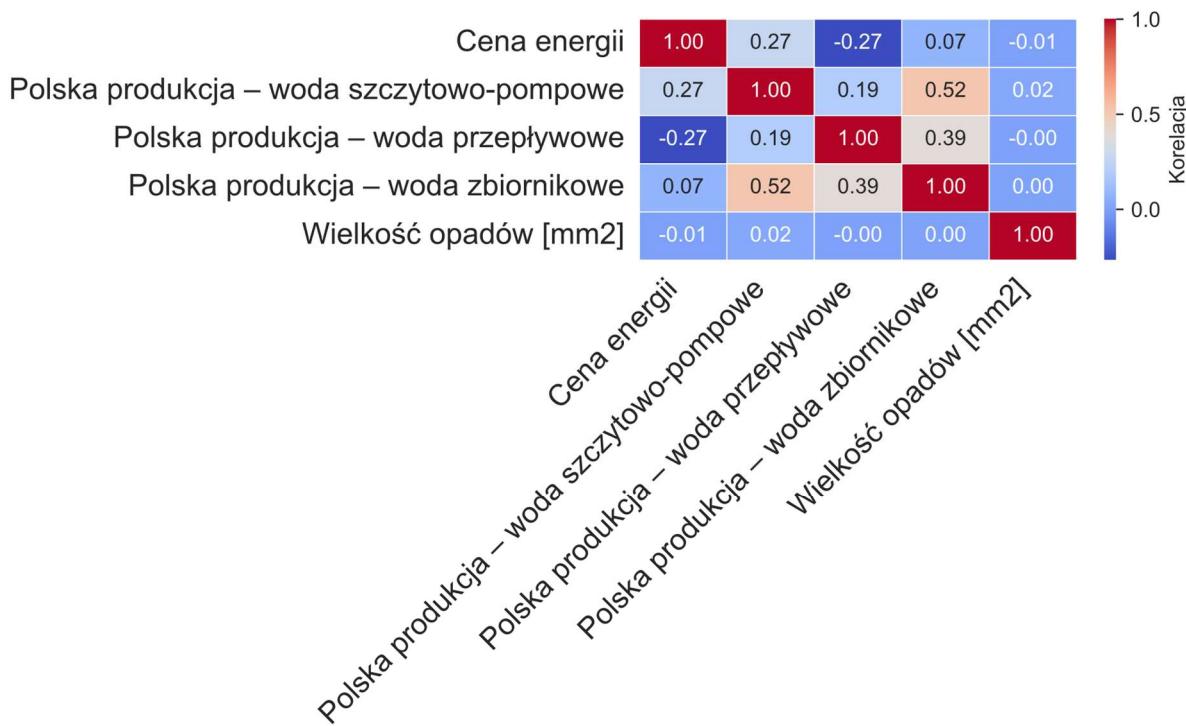
Pozostałymi źródłami energii w Polsce są elektrownie wodne:

- Szczytowo-pompowe.
- Przepływowe.
- Zbiornikowe.

Stanowią one 1,5% produkcji energii w Polsce (stan na rok 2024)⁵. W danych posiadamy trzy cechy z nimi związane:

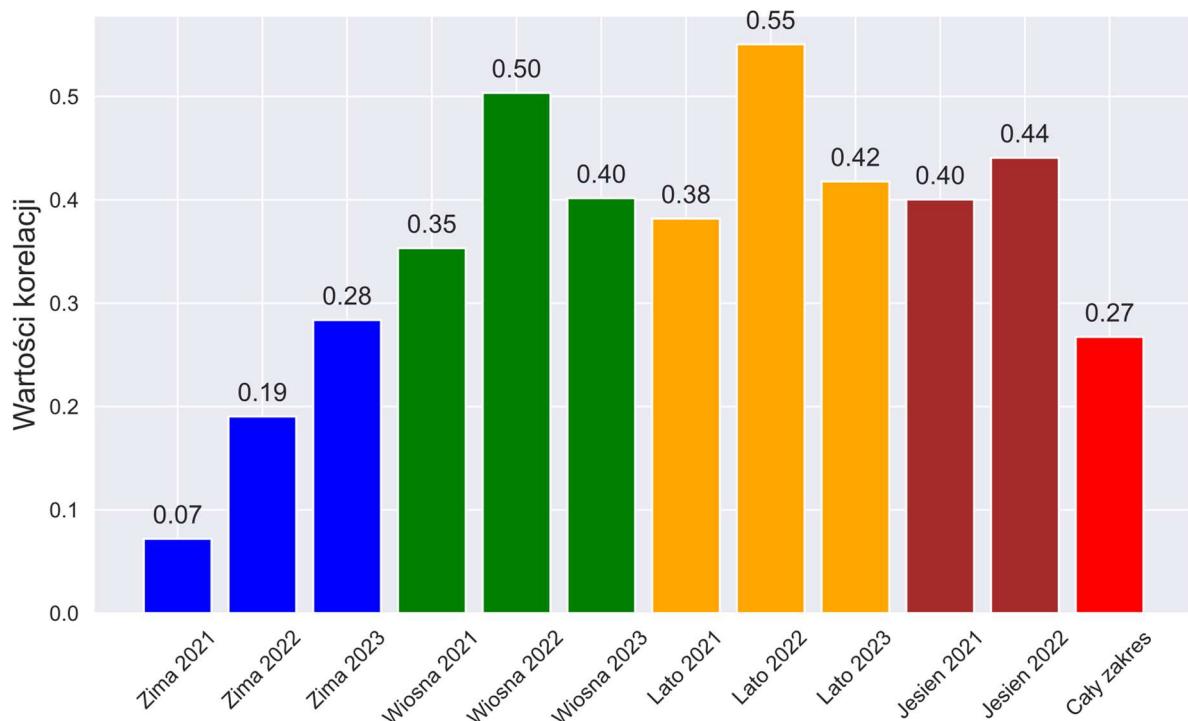
- Polska produkcja – woda szczytowo-pompowe.
- Polska produkcja – woda przepływowe.
- Polska produkcja – woda zbiornikowe.
- Wielkość opadów.

Macierz korelacji przedstawiona na rysunku 41 wskazuje na słabą korelację tych cech z ceną energii. Wielkość opadów wykazuje bardzo słabą korelację, i z ceną, i innymi cechami.

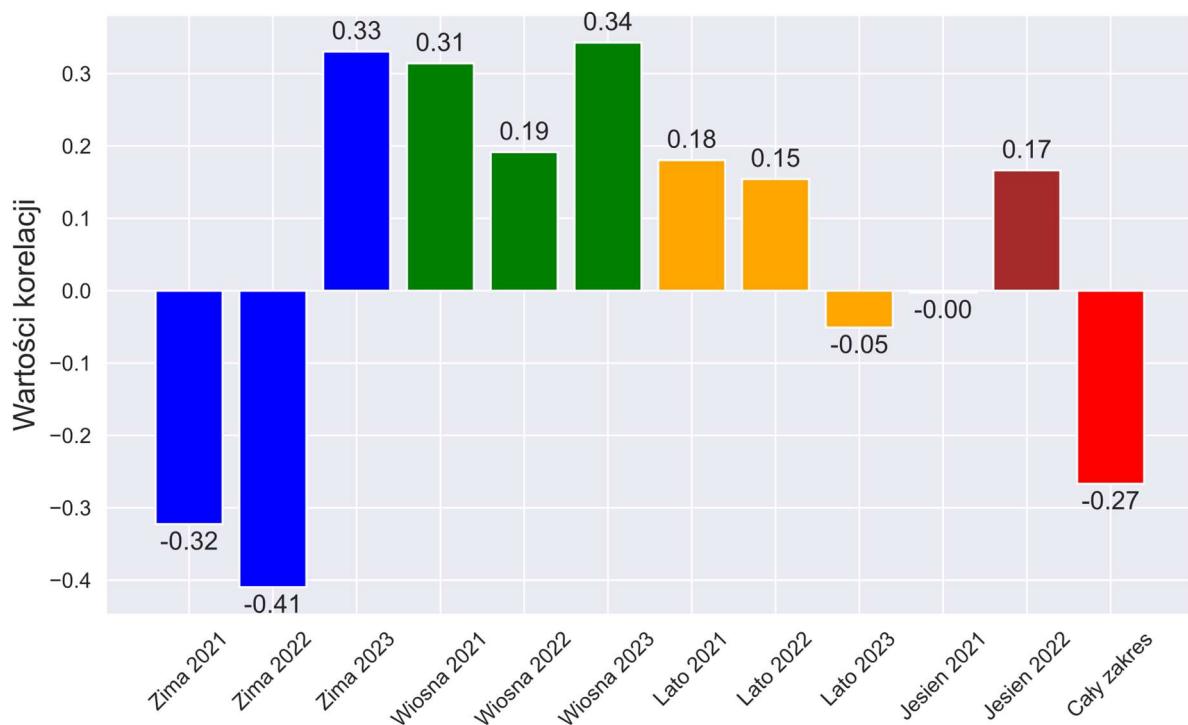


Rysunek 41 Macierz korelacji cech powiązanych z wodą oraz ceny energii. Opracowanie własne

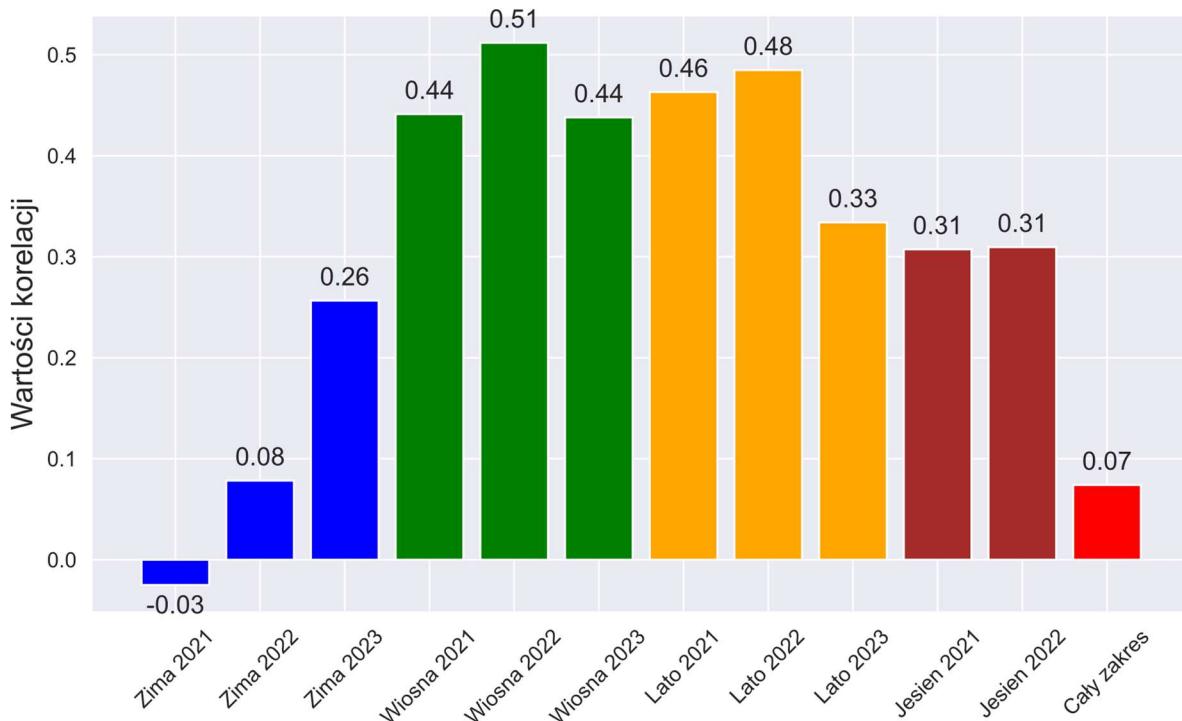
⁵ Źródło: [Energy instrat] (<https://energy.instrat.pl/system-elektroenergetyczny/produkcja-entsoe/>)



Rysunek 42 Wykres słupkowy przedstawiający wartości korelacji produkcji elektrowni szczytowo-pompowych względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne



Rysunek 43 Wykres słupkowy przedstawiający wartości korelacji produkcji elektrowni przepływowej względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne



Rysunek 44 Wykres słupkowy przedstawiający wartości korelacji produkcji elektrowni wodnej - zbiornikowej względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne

Na wykresie 42 można zauważyć, że istnieje umiarkowana zależność między produkcją elektrowni szczytowo-pompowych, a ceną energii. Podobnie na rysunku 44, przedstawiającym korelacje w poszczególnych okresach, ale dla elektrowni wodnej – zbiornikowej. Zima zauważalnie odstaje od reszty pór roku na obu wykresach. Jednocześnie warto zauważyć, że ta korelacja jest dodatnia, w przeciwnieństwie do reszty produkcji. Jest to spowodowane tym, że elektrownie wodne szczytowo-pompowe oraz zbiornikowe są używane jako magazyny energii, więc energia z nich produkowana jest kluczowa w sytuacjach, gdy istnieje wysokie zapotrzebowanie, a co za tym idzie wyższa cena. W momentach gdy cena energii jest niska elektrownie te korzystają z nadmiaru energii do pompowania wody do zbiornika górnego.

Na wykresie 43, przedstawiającym zależności między produkcją elektrowni wodnej przepływowej, a ceną prądu z podziałem na okresy, nie widać silnej jednoznacznej korelacji. Elektrownie przepływowe w Polsce nie mają dużej mocy zainstalowanej w systemie, zaledwie 0,8 GW (stan na 2024 rok) ⁶.

⁶ Źródło: [Energy instrat] (<https://energy.instrat.pl/system-elektroenergetyczny/zainstalowana-moc-entsoe/>)

4.3 Emisyjne źródła energii

Autor: Mikołaj Kalisz

Nieodnawialne źródła energii przez wiele lat posiadały monopol w produkcji. Dopiero w ostatnich latach za sprawą wielu reform i polityki klimatycznej zaczęły tracić na dominacji. Jednakże wciąż w dużej mierze dyktują warunki na rynku posiadając 73,5% udziału w wyprodukowanej energii za rok 2024⁷. Poprawna analiza tych źródeł jest znacząca dla ostatecznie osiągniętych wyników.

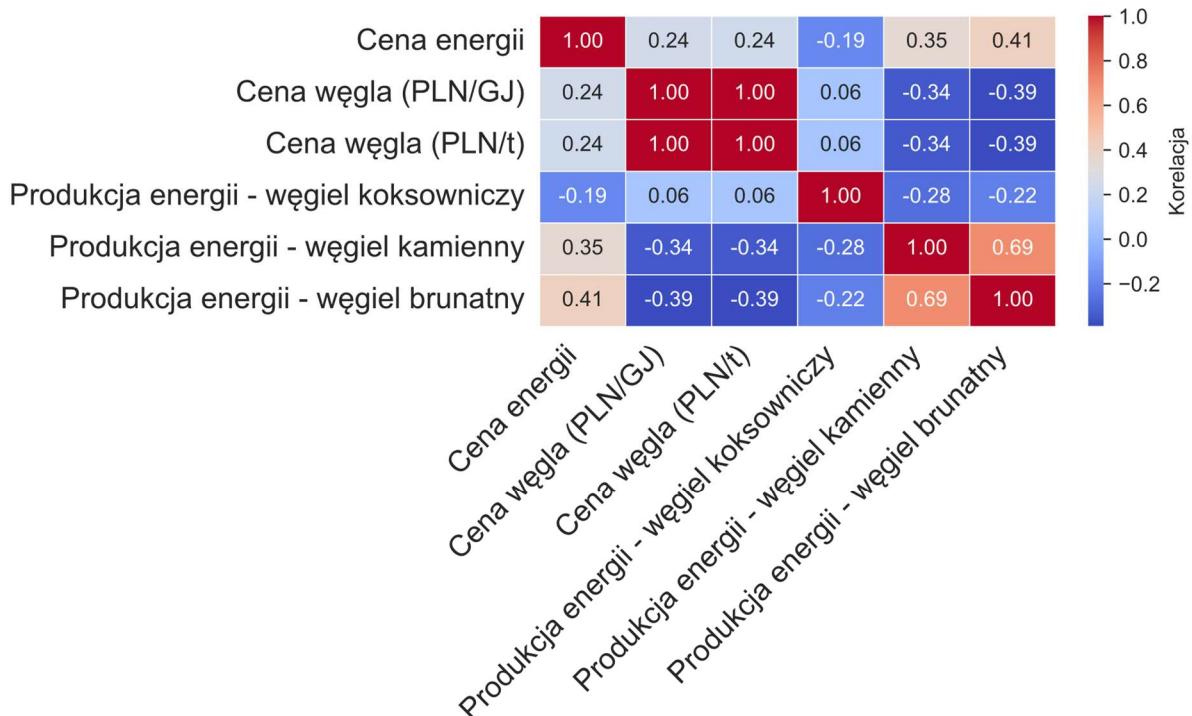
4.3.1 Węgiel

Autor: Mikołaj Kalisz

Węgiel stanowi podstawę polskiej energetyki. W Polsce produkowana jest energia z węgla kamiennego, brunatnego oraz koksowniczego. W zbiorze danych znajdują się następujące cechy:

- Cena węgla [PLN/GJ].
- Cena węgla [PLN/t].
- Produkcja energii z węgla koksowniczego.
- Produkcja energii z węgla kamiennego.
- Produkcja energii z węgla brunatnego.

⁷ Źródło: [Energy instrat] (<https://energy.instrat.pl/system-elektroenergetyczny/produkcja-udzial-oze/>)



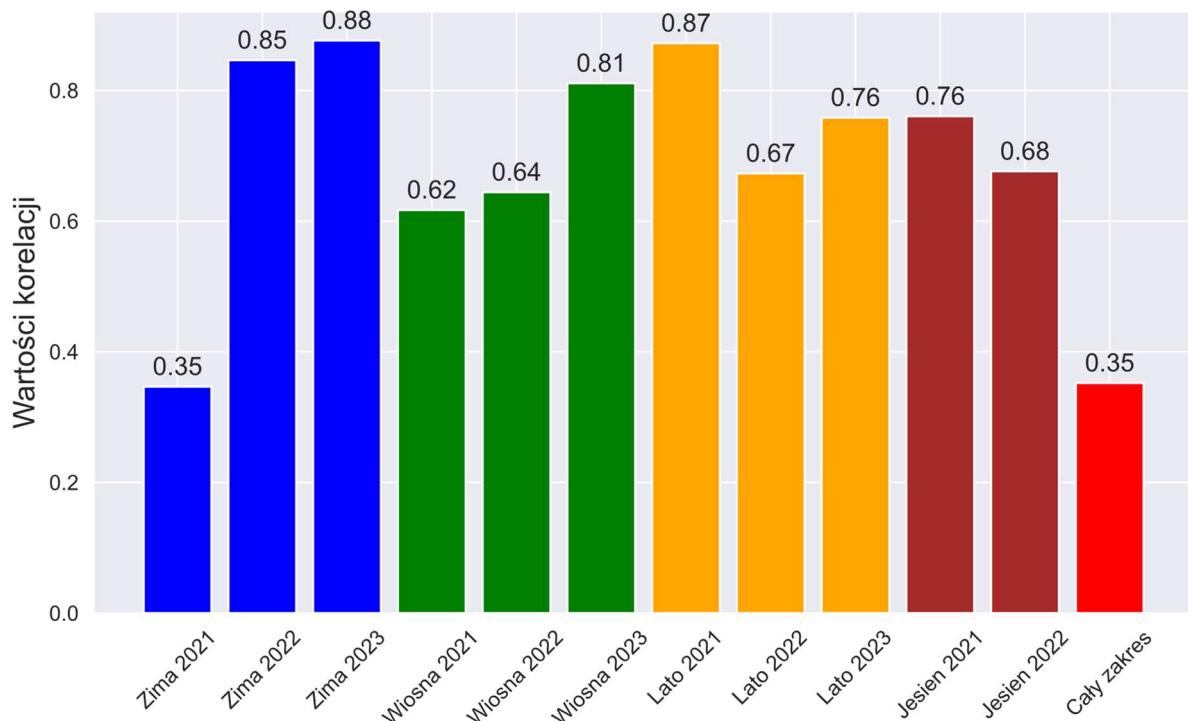
Rysunek 45 Macierz korelacji cech powiązanych z węglem oraz ceny energii. Opracowanie własne

Na rysunku 45 przedstawiono macierz korelacji między wybranymi danymi a ceną energii. Korelacja między cenami węgla jest korelacją pełną, z racji na to, że jest to ta sama cena wyrażona w innych jednostkach. W dalszej analizie należy korzystać tylko z jednej z nich.

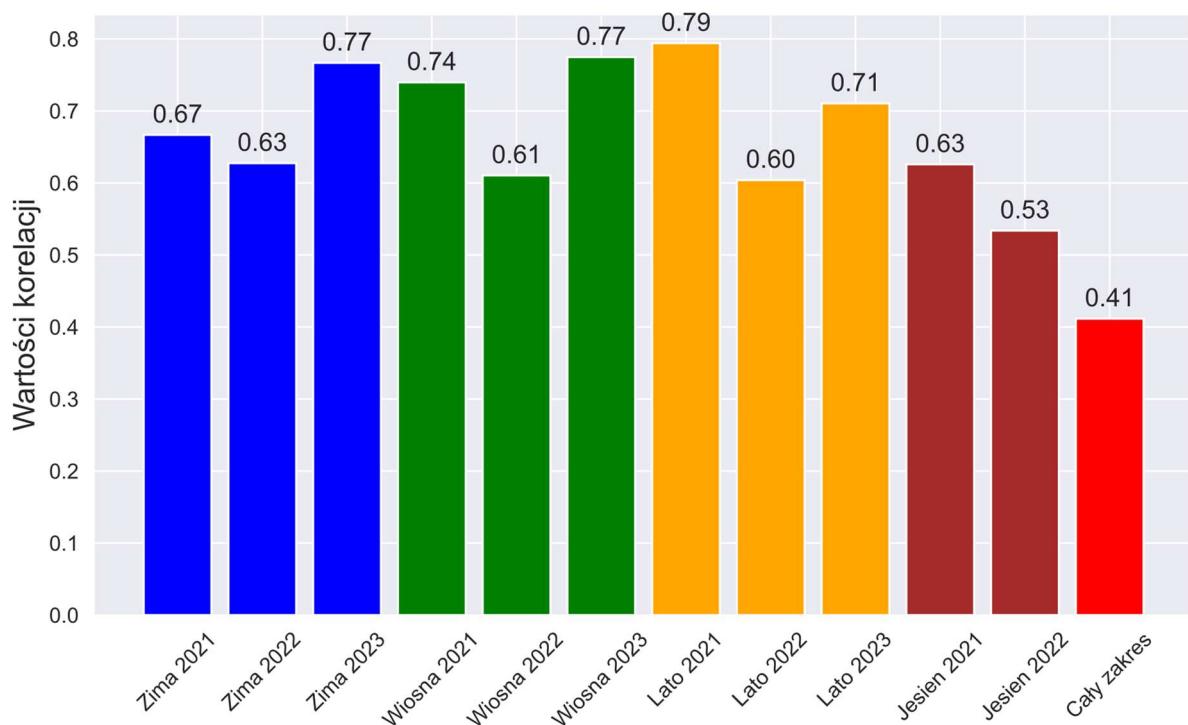
Cena ma umiarkowaną korelację z produkcją energii z węgla brunatnego oraz kamiennego. Korelacja z ceną węgla oraz produkcją energii z węgiel koksowniczego jest słaba.

Na rysunkach 46 oraz 47 przedstawiono wykres siły korelacji dla poszczególnych pór roku i lat odpowiednio dla węgla kamiennego i brunatnego. Możemy zauważyć silną/bardzo silną korelację dla większości miesięcy. Może to oznaczać, że wartości mocy produkcji z węgla są kluczowe w kreowaniu ceny energii na rynku.

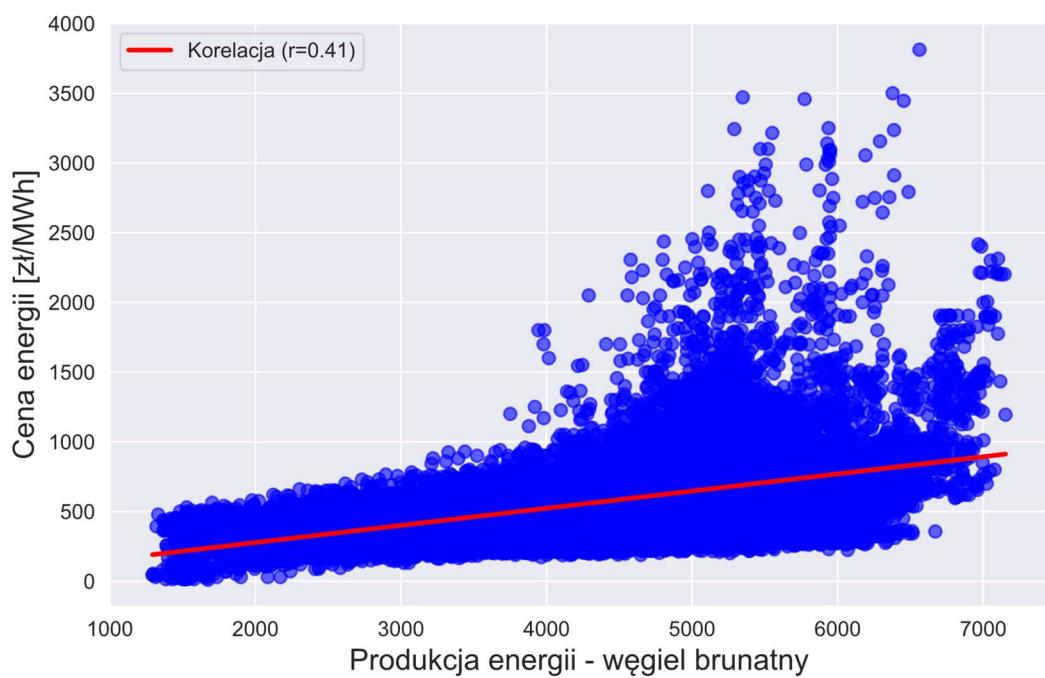
Na rysunkach 48 oraz 49 przedstawiono wykres zależności ceny energii od produkcji energii z węgla odpowiednio brunatnego i kamiennego. Można na nich zauważyć, że im większa jest produkcja energii, tym wyższa cena, co potwierdza uprzednią analizę korelacji z podziałem danych na mniejsze okresy.



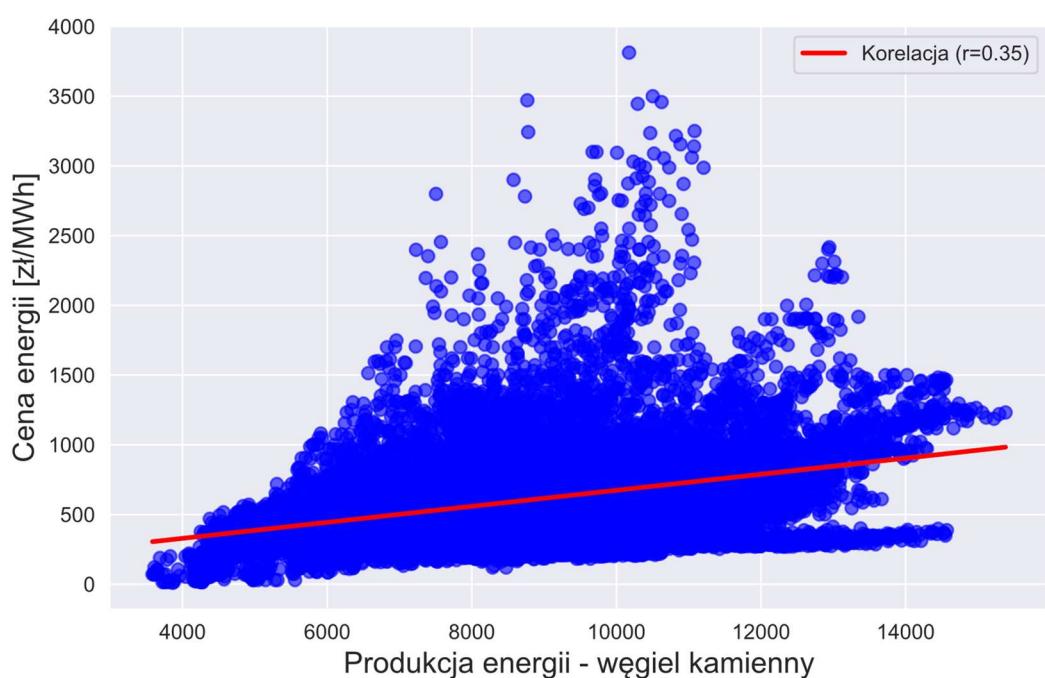
Rysunek 46 Wykres słupkowy przedstawiający siłę korelacji wartości produkcji energii z węgla kamiennego względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne



Rysunek 47 Wykres słupkowy przedstawiający siłę korelacji wartości produkcji energii z węgla brunatnego względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne



Rysunek 48 Wykres zależności ceny energii od produkcji energii z węgla brunatnego wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne



Rysunek 49 Wykres zależności ceny energii od produkcji energii z węgla kamiennego wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne

4.3.2 Gaz

Autor: Mikołaj Kalisz

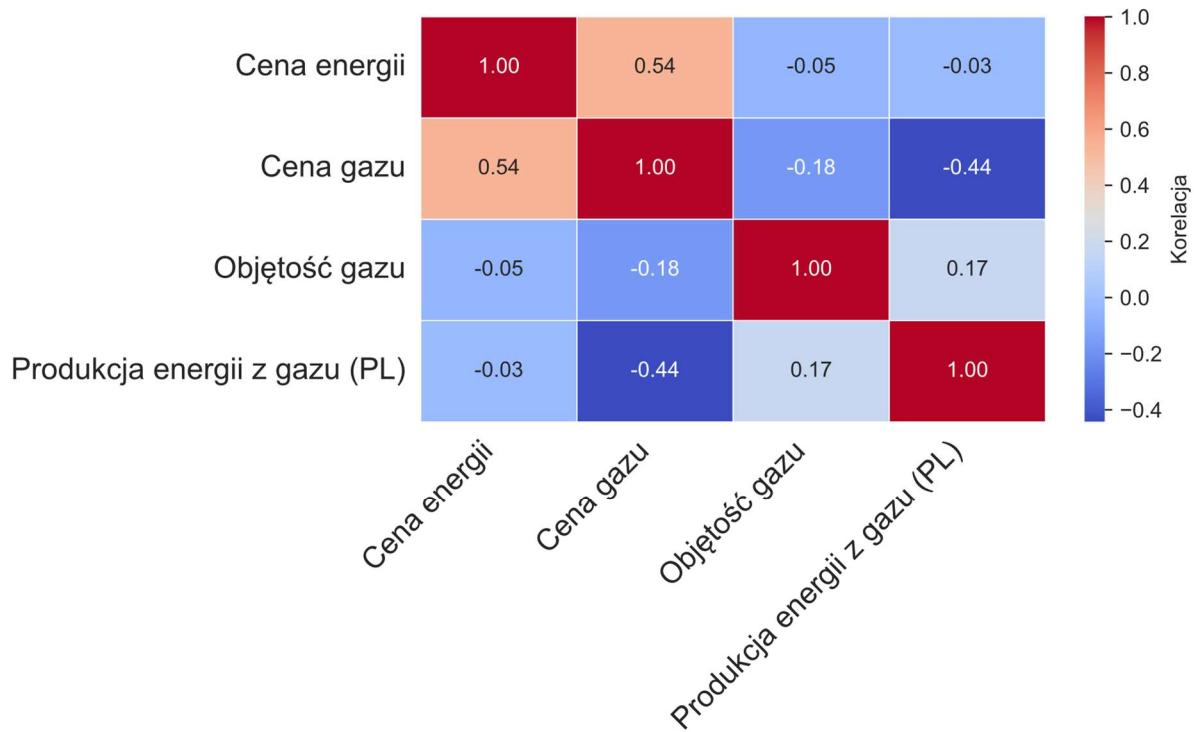
Energia z gazu stanowi 11,4% produkcji w Polsce (stan na 2024 rok). Jest to zdecydowanie mniej niż węgiel, jednakże dalej to znacząca siła na polskim rynku. Na rysunku 50 przedstawiona jest macierz korelacji danych powiązanych z gazem, czyli:

- Cena gazu.
- Objętość wyprodukowanego gazu.
- Produkcja energii z gazu.

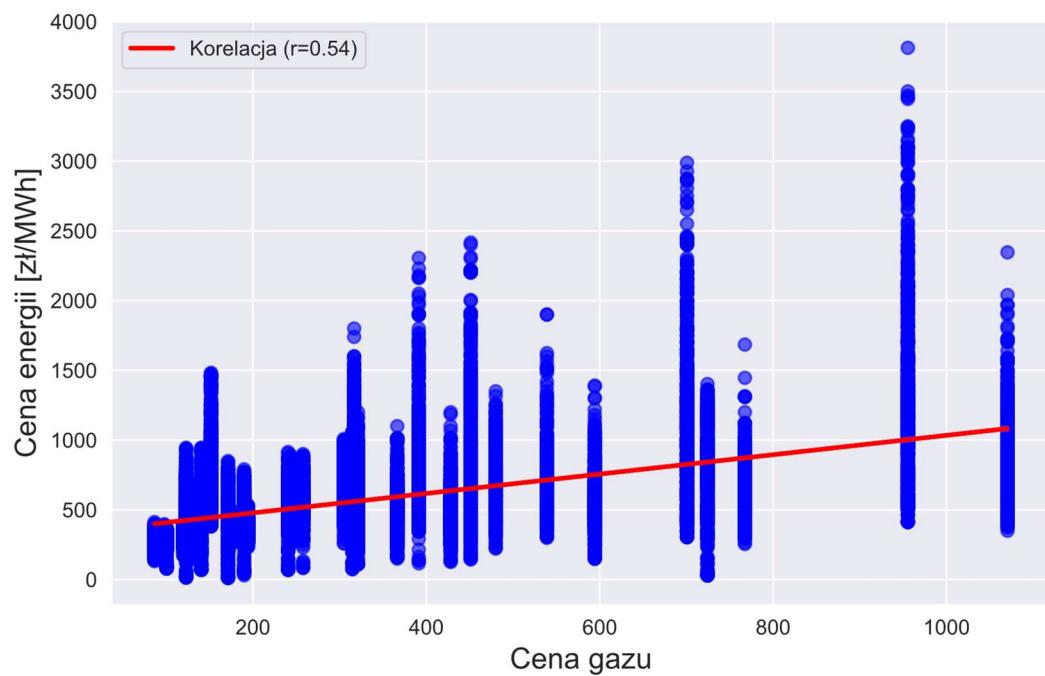
Jak można zauważyć cena gazu ma umiarkowaną silę korelacji z ceną energii oraz z produkcją energii z gazu. Gdy cena gazu rośnie to cena energii także, co można zaobserwować na wykresie 51 i jednocześnie produkcja spada (wykres 52). Objętość produkowanego gazu posiada słabą korelację z innymi cechami.

Z macierzy wynika, że istnieje bardzo słaba korelacja ceny prądu z produkcją energii z gazu, jednakże jak można zauważyć na rysunku 53, dla wielu okresów można zauważać korelacje na umiarkowaną i silną. Ciekawym zjawiskiem jest wystąpienie ujemnej korelacji (bliskiej zeru) dla całości danych, w momencie gdy w poszczególnych okresach występuje dodatnia korelacja. O ile w poszczególnych okresach wzrost produkcji energii z gazu koreluje z ceną energii to w dłuższej perspektywie, prawdopodobnie ze względu na zmianę średniej wartości produkcji, wykres zależności w różnych okresach przesuwa się wzduż osi produkcji, tworząc rozproszone dane bez widocznej zależności, co widać na rysunku 54.

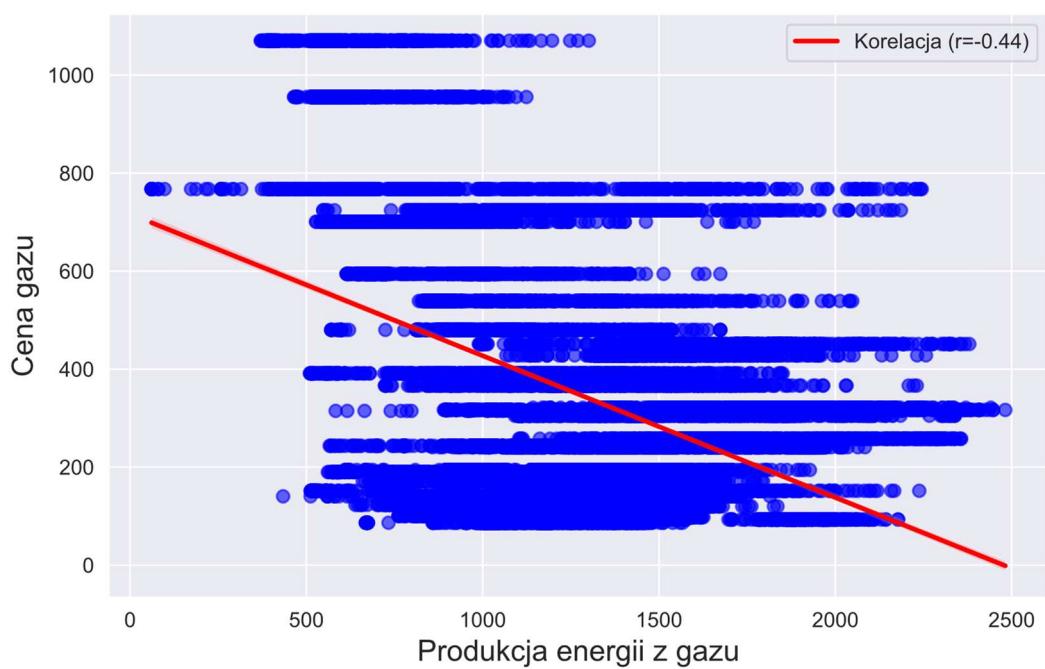
Warto rozważyć uwzględnienie danych o gазie w modelu, zwłaszcza że gaz ma być opcją przejściową na drodze do osiągnięcia neutralności energetycznej. Dodatkowo ze względu na sankcje wobec Rosji, dużego gracza na rynku gazowym, charakter zależności od ceny gazu może być bardzo zmienny.



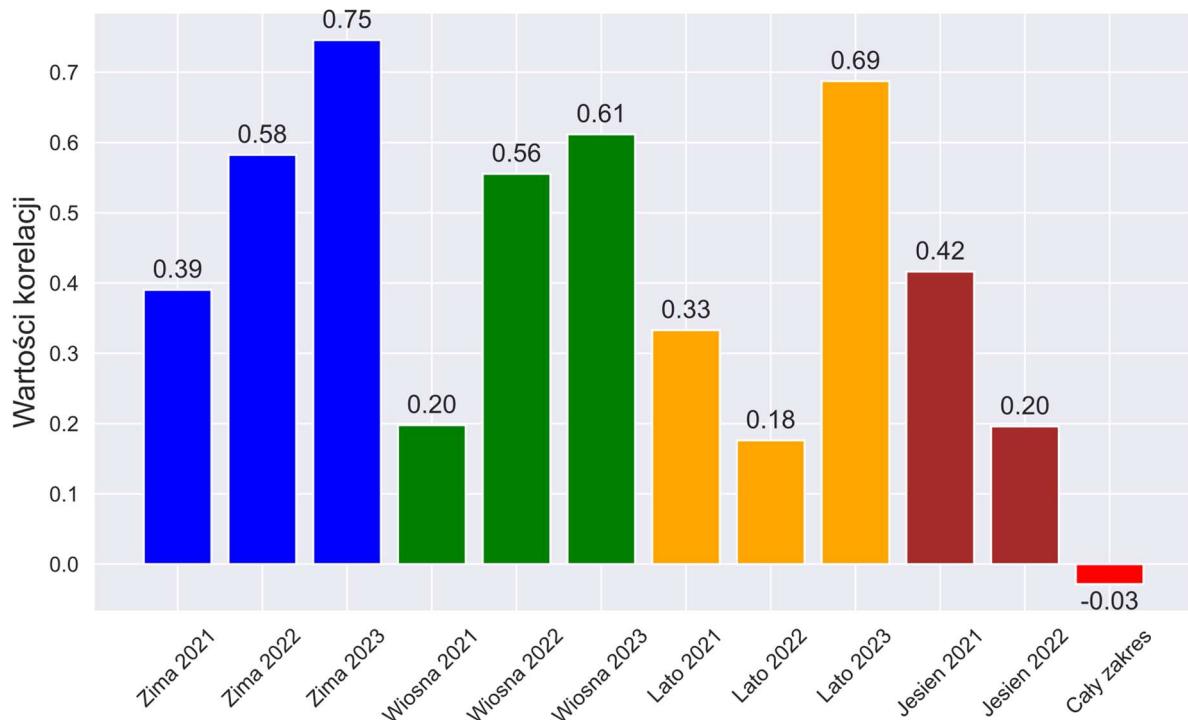
Rysunek 50 Macierz korelacji cech powiązanych z gazem oraz ceny energii. Opracowanie własne



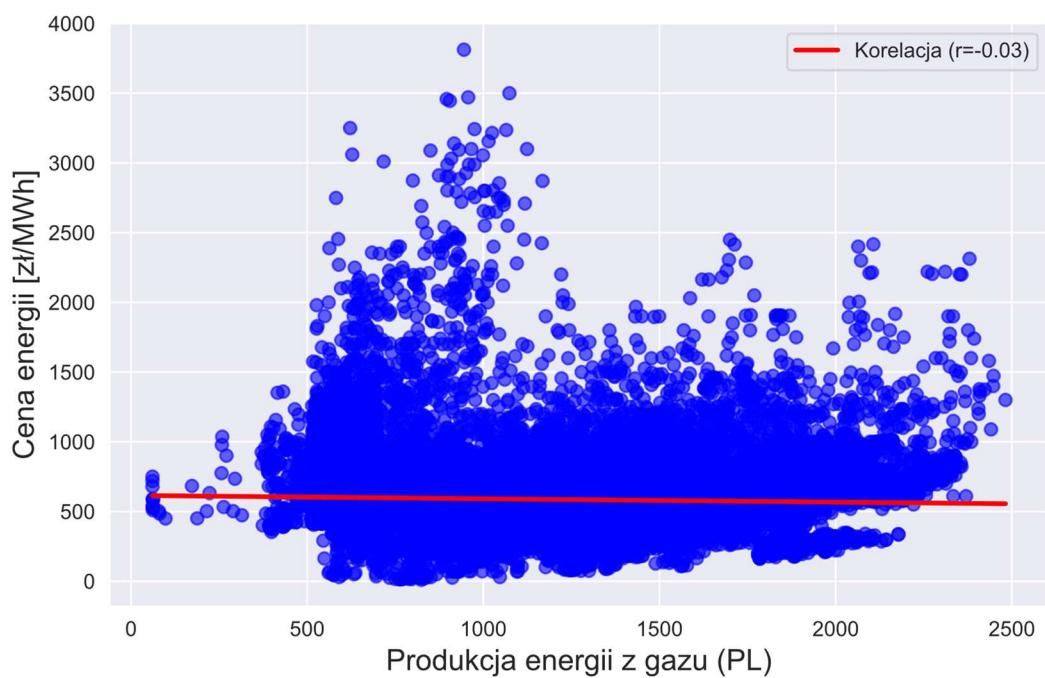
Rysunek 51 Wykres zależności ceny energii od ceny gazu wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne



Rysunek 52 Wykres zależności ceny gazu od produkcji energii z gazu wraz z narysowaną linią najlepszego dopasowania.
Opracowanie własne



Rysunek 53 Wykres słupkowy przedstawiający siłę korelacji wartości produkcji energii z gazu względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne

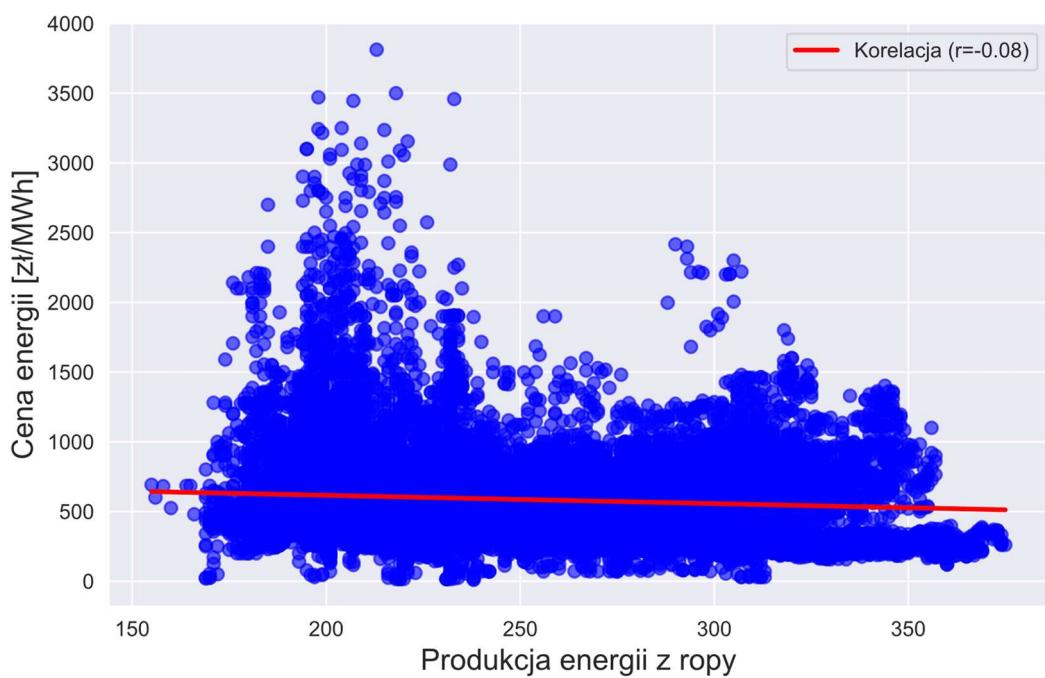


Rysunek 54 Wykres zależności ceny energii od produkcji energii z gazu wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne

4.3.3 Ropa

Autor: Mikołaj Kalisz

Energia z ropy stanowi ułamek produkcji energii w Polsce. W danych posiadamy cechę Produkcja energii z ropy (PL). Jak można zauważyć na wykresie 55, brak zauważalnej korelacji między produkcją energii z ropy, a ceną energii.



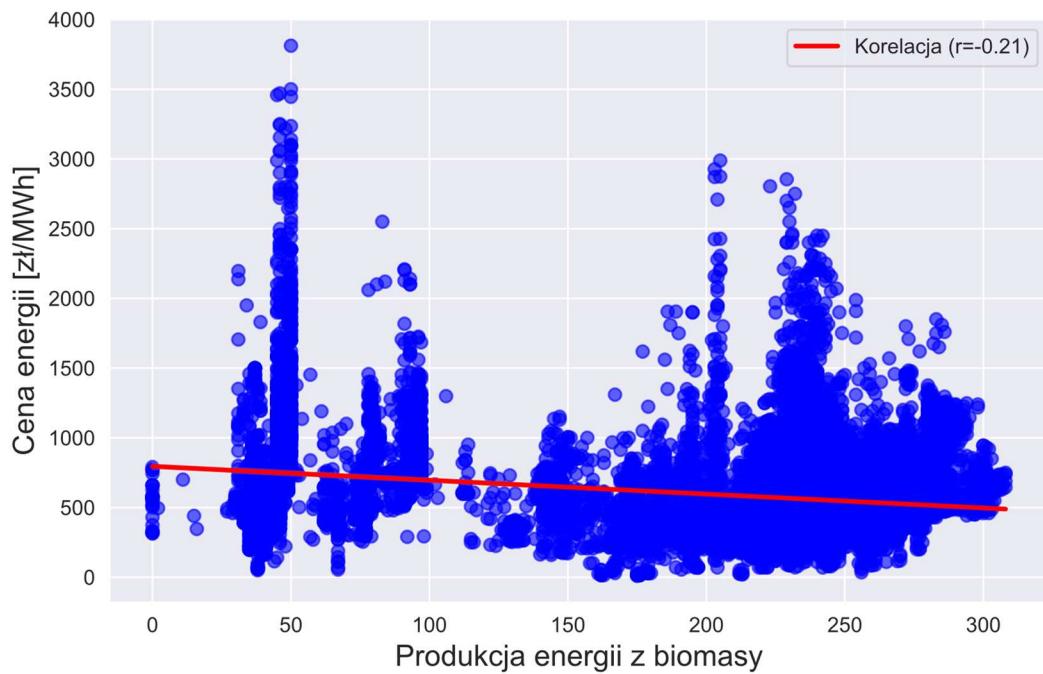
Rysunek 55 Wykres zależności ceny energii od produkcji energii z ropy wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne

4.3.4 Biomasa

Autor: Mikołaj Kalisz

Energia z biomasy stanowi 1,7%⁸ produkcji energii w Polsce (stan na 2024 rok). W danych posiadamy cechę Produkcja energii z biomasy (PL). Jak można zauważyć na wykresie 56, brak zauważalnej korelacji między produkcją energii z biomasy, a ceną energii.

⁸ Źródło: [Energy instrat] (<https://energy.instrat.pl/system-elektroenergetyczny/produkcja-entsoe/>)



Rysunek 56 Wykres zależności ceny energii od produkcji energii z biomasy wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne

4.4 Czynniki ekonomiczne i polityczne

Autor: Mikołaj Kalisz

Ceny surowców, ilość i moc jednostek wytwórczych, czy podatki to jedne z wielu czynników na który ma wpływ polityka. Dodatkowo warto wspomnieć o handlu energią między krajami sąsiadującymi w celu efektywnego zarządzania mocą w sieci oraz bezpieczeństwem energetycznym. Połącznie sieciowe powoduje także połączenie rynków i stworzenie zależności między nimi.

4.4.1 Opłata za emisję CO₂

Autor: Mikołaj Kalisz

Wprowadzona już w 2005 roku opłata za emisję dwutlenku węgla, jest jednym z czynników wpływających na cenę energii emisyjnej. Opłata obowiązuje za każdą wyemitowaną tonę CO₂ i jest zmienna w czasie. W 2024 roku średnia cena uprawnienia do emisji za tonę wynosiła 64,75 euro. Dla porównania dziesięć lat wcześniej, w 2014 roku ta cena średnio wynosiła 5,9 euro⁹.

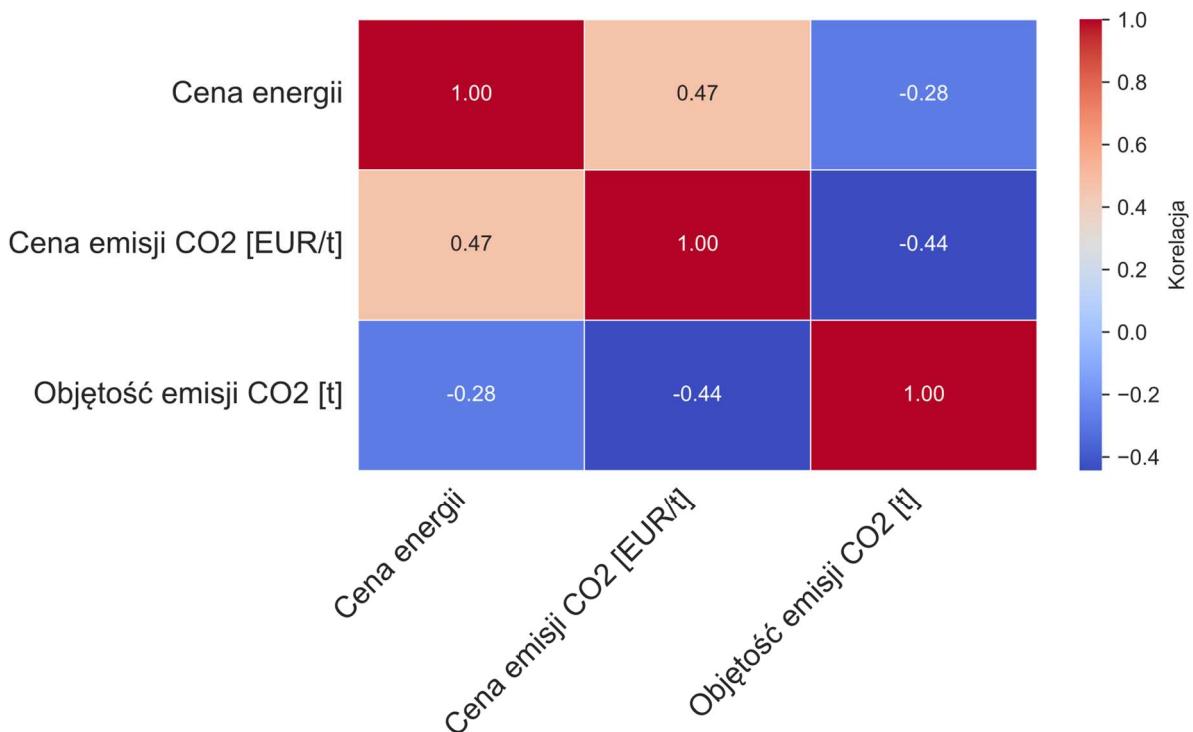
W danych posiadamy dwie kolumny powiązane z tym tematem:

- Cena uprawnień CO₂ [EUR/t]

⁹ Źródło: [Energy instrat] <https://energy.instrat.pl/ceny/eu-ets/>

- Objętość emisji CO₂ [t]

Na rysunku 57 przedstawiono macierz korelacji między tymi cechami oraz ceną energii. Jak można zauważyć istnieje słaba korelacja między objętością a ceną energii, jednakże istnieje umiarkowanie silna korelacja ceny energii z ceną emisji CO₂. Na przestrzeni pojedynczych lat nie ma dużych zmian w cenie emisji, jednakże na przestrzeni kilku lat ta różnica występuje i ma znaczenie w kontekście kreowania się ceny prądu wysokoemisyjnych źródeł energii.



Rysunek 57 Macierz korelacji cech powiązanych z emisją CO₂ oraz ceny energii. Opracowanie własne

4.4.2 Wpływ handlu międzynarodowego

Autor: Mikołaj Kalisz

W zależności od nadmiaru lub niedoboru prądu istnieje możliwość korzystania z sieci energetycznej sąsiednich państw. W danych jest wiele kolumn dotyczących rynków sąsiednich jednakże występuje jedynie bardzo słaba lub słaba korelacja tych danych z ceną energii. Pojedyncze dane wykazują umiarkowaną korelację, jak na przykład Słowacja – udział energii odnawialnych, jednakże nie oznacza to wpływu produkcji energii odnawialnej na Słowację na cenę energii w Polsce. Ze względu na bliskie położenie dwóch krajów, mogą występować podobne warunki atmosferyczne, jak i ekonomiczne wpływając podobnie na cenę.

4.4.3 Wpływ ekonomii

Autor: Mikołaj Kalisz

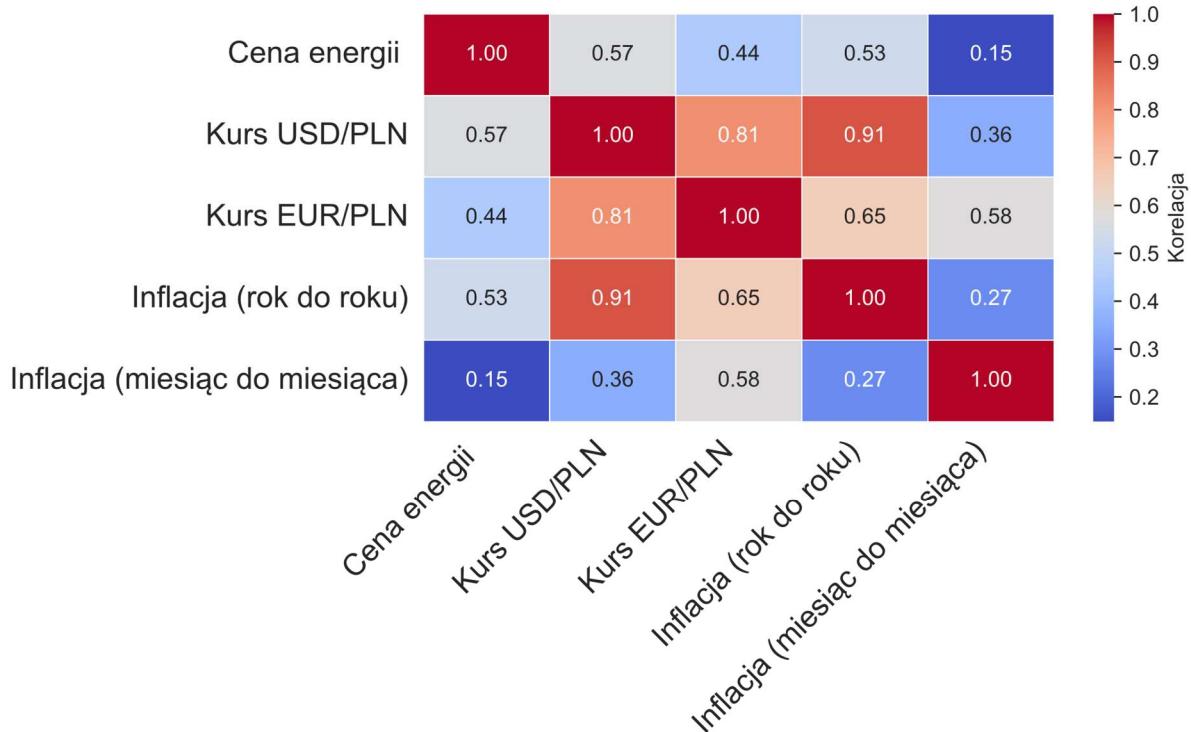
Rynek energii elektrycznej dnia następnego jest także zależny od gospodarki danego kraju oraz siły waluty danego kraju. Duża część surowców potrzebnych do produkcji energii, bądź produktów potrzebnych do budowy nowych jednostek produkcyjnych jest zagraniczna, a co za tym idzie dochodzi dodatkowy czynnik, kurs przewalutowania.

Na poniższym rysunku numer 58 przedstawiona została macierz korelacji następujących cech:

- Kurs USD/PLN
- Kurs EUR/PLN
- Inflacja (rok do roku)
- Inflacja (miesiąc do miesiąca)

Możemy zauważyć, że oba kursy mają umiarkowaną korelację z ceną energii. Silna korelacja występuje między kursami USD/PLN a EUR/PLN. Oba kursy są w odniesieniu do złotówki, więc potencjalnie w danych widoczna jest siła złotówki wobec reszty świata. Inflacja miesiąc do miesiąca wykazuje słabą korelację, w przeciwieństwie do inflacji rok do roku, która ma już zauważalnie umiarkowaną siłę korelacji z ceną energii.

Do modelu można rozważyć dodanie cech kurs USD/PLN oraz inflacja rok do roku. Można przetestować, czy dodanie kursu euro nie spowoduje nadmiernego dopasowania do danych treningowych.



Rysunek 58 Macierz korelacji cech powiązanych z ekonomią oraz ceny energii. Opracowanie własne

4.5 Inżynieria cech

Autor: Mikołaj Kalisz

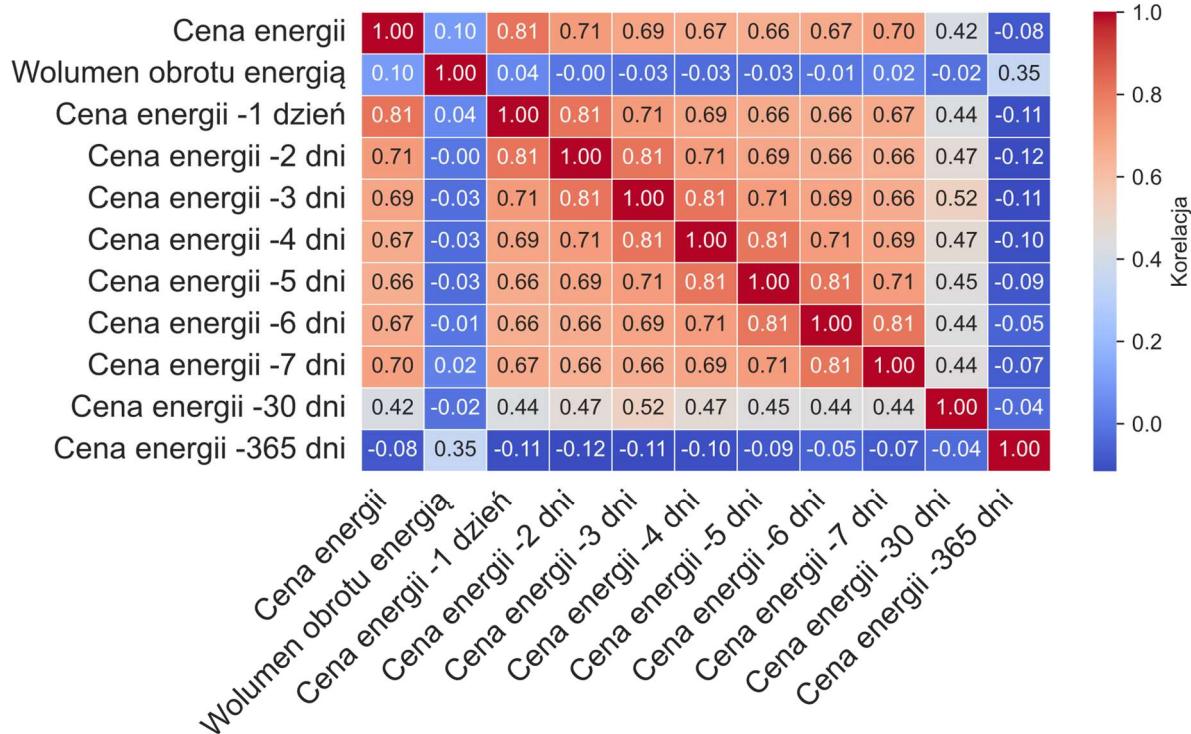
Niektóre dane nie wykazują korelacji w pojedynkę, lecz połączone razem we wspólną cechę mogą okazać się przydatne dla modelu. Drugim sposobem tworzenia nowych cech jest wytworzenie cechy pochodnej, jak dzień tygodnia, czy numer miesiąca z daty. W ten sposób można potencjalnie osiągnąć zależności, z którymi model mógłby mieć trudność, gdyby nie były wcześniej wyodrębnione.

4.5.1 Wpływ historycznych cen i wolumenu

Autor: Mikołaj Kalisz

W celu próby uchwycenia trendów w cenie energii utworzono cechy pomocnicze zawierające wartości historyczne ceny energii. Dodatkowo sprawdzana jest korelacja z wolumenem obrotu energią na giełdzie. Na rysunku 59 przedstawiono macierz korelacji tych wartości. Można zauważyć, że istnieje silna korelacja z danymi przesuniętymi w czasie. Nie widać natomiast zależności ceny od wolumenu.

Do modelu warto zastosować dane historyczne, lecz czy wszystkie cofnięte wartości, czy tylko wybrane, należy sprawdzić empirycznie, mając na uwadze możliwe wystąpienie nadmiernego dopasowania do danych treningowych (ang. overfitting).



Rysunek 59 Macierz korelacji danych historycznych wraz z ceną energii. Opracowanie własne

5 Modelowanie cen energii elektrycznej z wykorzystaniem sztucznych sieci neuronowych

Autor: Mikołaj Kalisz

W tym rozdziale zostanie opisana architektura modelu, dane wykorzystane do nauki, ich podział oraz dobór hiper-parametrów. Celem stworzenia tego modelu jest próba osiągnięcia jak najlepszych wyników predykcji ceny energii.

5.1 Dane

Autor: Mikołaj Kalisz

Wybór odpowiednich danych oraz ich przygotowanie jest kluczowym etapem tworzenia modelu sieci neuronowej. Od nich zależy proces uczenia, a z niewłaściwych danych model nie nauczy się właściwych zależności.

5.1.1 Wybór cech

Autor: Mikołaj Kalisz

Korzystając z analizy czynników wpływających na cenę energii opisanej w rozdziale 0 wybrano listę cech, która zostanie użyta do nauki modelu. Wartości które wybrano wykazywały szczególnie silną korelację z ceną potencjalnie wpływając pozytywnie na wyniki końcowe.

Oto lista cech wykorzystana do nauki:

- Produkcja energii z fotowoltaiki
- Rezerwa mocy poniżej zapotrzebowania
- Cena uprawnień emisyjnych (CO₂)
- Inflacja w porównaniu do poprzedniego roku
- Cena gazu
- Produkcja energii z węgla kamiennego w Polsce
- Produkcja energii z wody (szczytowo-pompowa)
- Produkcja energii z węgla brunatnego w Polsce
- Produkcja energii z wiatru w Polsce
- Kurs USD/PLN
- Sumaryczna generacja JGW-a

- Sumaryczna generacja JGMa
- Sumaryczna generacja JGX-x
- Suma zdolności wytwórczych JGW-a
- Suma zdolności wytwórczych JGMa
- Sumaryczna generacja jednostek nieuczestniczących

5.1.2 Podział danych

Autor: Mikołaj Kalisz

Dane zostały podzielone na trzy podzbiory w następujący sposób:

- Podzbiór treningowy – 08.01.2021 – 09.08.2022
- Walidacyjny – 09.08.2022- 31.12.2022
- Testowy. 01.01.2023 – 01.09.2023

Podział został dokonany w taki sposób, by dane walidacyjne stanowiły około 20% sumarycznego zbioru danych treningowych i walidacyjnych. Stanowi to okres około czterech miesięcy.

W celu zapobiegania przeuczenia zastosowano metodę wcześniego kończenia uczenia. Jeśli w określonej liczbie epok dokładność predykcji na zbiorze walidacyjnym się nie zwiększy to nauka jest przerywana i następuje powrót do najlepszej wartości wskaźnika dokładności predykcji. W analizie z rozdziału 4 wykazano dużą zmienność zależności danych w poszczególnych okresach. W przypadku wystąpienia przeuczenia model zwiększy swoją dokładność dla danych treningowym, lecz straci możliwość wychwycenia ogólnych zależności między danymi, diametralnie pogarszając wyniki na zbiorze walidacyjnym i testowym. Wtedy zadziała metoda wcześniego kończenia i przywróci model do stanu z najlepszym wynikiem na zbiorze walidacyjnym.

5.1.3 Skalowanie danych

Autor: Mikołaj Kalisz

Następnym etapem po podziale danych jest ich skalowanie. W tym celu przetestowano następujące techniki, opisane w rozdziale 2.2.5:

- Min-Max.
- Standaryzacja.

Do implementacji wykorzystano bibliotekę sklearn i zawarte w niej klasy MinMaxScaler oraz StandardScaler. Wartości do skalowania zostały dobrane automatycznie na podstawie danych treningowych. Następnie zostało wykonane skalowanie na wszystkich wartościach, razem z treningowymi. Wyniki modelu są odwrotnie skalowane w celu umożliwienia porównania ich z oryginalnymi wartościami.

5.2 Model

Autor: Mikołaj Kalisz

W celu otrzymania najlepszego wyniku zostało stworzone wiele modeli o różnych hiperparametrach. Na początek wybrano zakres hiperparametrów do testowania, a następnie iteracyjnie testowano różne kombinacje w poszukiwaniu najlepszego wyniku na zbiorze walidacyjnym. Więcej o procesie doboru hiperparametrów zostanie przedstawione w podrozdziale 5.2.3.

5.2.1 Architektura

Autor: Mikołaj Kalisz

Testowano różne architektury systemu, z różną ilością warstw i neuronów w każdej warstwie. Cechami wspólnymi każdej z nich są:

- Neurony na wejściu odpowiadają liczbie cech.
- Sieć jest gęsto połączona.
- Warstwy ukryte korzystają z funkcji aktywacji ReLu.
- Warstwa wyjściowa posiada jeden neuron z liniową funkcją aktywacji.
- Na każdej warstwie ukrytej stosowany jest dropout.

Przykładowe testowane architektury, gdzie każda kolejna wartość to ilość neuronów w każdej kolejnej warstwie, prezentują się następująco:

- (140, 64, 32)
- (128, 64, 32, 16, 8)
- (256, 128, 64, 32)
- (1024, 512, 256, 128, 64)
- (2048, 1024, 512, 256, 128)
- (280, 140, 70, 35)

Do implementacji wykorzystano model sekwencyjny z biblioteki Keras.

5.2.2 Optymalizator

Autor: Mikołaj Kalisz

Jest to jeden z kluczowych elementów tworzenia modelu. Przetestowano zastosowanie trzech z wielu dostępnych optymalizatorów:

- Adam
- SGD
- RMSprop

Dodatkowo przetestowano różne wartości współczynnika uczenia (ang. learning rate). Początkowo zaczęto od przedziału 0,001 do 0,01 i następnie w zależności od wyników zakres był zmniejszany.

5.2.3 Dobór hiper-parametrów

Autor: Mikołaj Kalisz

Z racji na dużą ilość występujących hiperparametrów zdecydowano się na zastosowanie narzędzi do zapisów wyników - Mlflow¹⁰ oraz do automatycznego doboru hiperparametrów - Optuna¹¹.

Metoda działania Optuny jest następująca:

- Losowy wybór hiperparametrów – Ze zbioru wcześniej zasugerowanych hiperparametrów losowane są konkretne do użycia w pierwszej iteracji.
- Trening modelu – Tworzony jest model z architekturą wcześniej wylosowaną i przeprowadzane jest uczenie.
- Optymalizacja – Parametry są dostosowywane do następnej iteracji na podstawie wyników na zbiorze walidacyjnym.
- Następuje kolejna iteracja uczenia i powtórzenie procesu optymalizacji parametrów.
- Na koniec zwraca jest konfiguracja która osiągnęła najlepszy wynik.

Optuna umożliwia wybór sposobu optymalizacji doboru hiperparametrów, jednakże w tej pracy zastosowano domyślny algorytm TPE – drzewiaste estymatory Parzena (ang. Tree-structured Parzen Estimator).

¹⁰ Źródło: [Mlflow] <https://mlflow.org/>

¹¹ Źródło: [Optuna] <https://optuna.org/>

Mlflow został użyty w celu zapisywania wyników, konfiguracji oraz samego modelu. Narzędzie umożliwia dostęp do wejżenia w historię stworzonych iteracji uczenia oraz porównanie ich.

Do oceny postępu nauki użyto metryki MAE opisanej w rozdziale 2.5.4

6 Modelowanie cen energii elektrycznej z wykorzystaniem uczenia ciągłego

Autor: Adam Mantiuk

W tym rozdziale przedstawione zostaną techniki uczenia ciągłego, jakie zostały zastosowane w poniższej pracy. Przedstawione zostaną poszczególne strategie podziału danych na mniejsze paczki, które determinują częstotliwość aktualizacji modeli uczenia ciągłego. Przedstawione zostaną również metody zapobiegania katastroficznemu zapominaniu, które zostały zaimplementowane podczas modelowania.

6.1 Projektowanie modelu

Autor: Adam Mantiuk

Podczas implementacji modelu z wykorzystaniem strategii uczenia ciągłego zdecydowano się na przyjęcie identycznej architektury sieci, która opisana została w rozdziale poprzednim. Liczba warstw sieci neuronowej, jak i liczba neuronów w poszczególnych warstwach pozostała niezmieniona. Tak samo postąpiono z innymi hiperparametrami, które występują w obu rozdziałach.

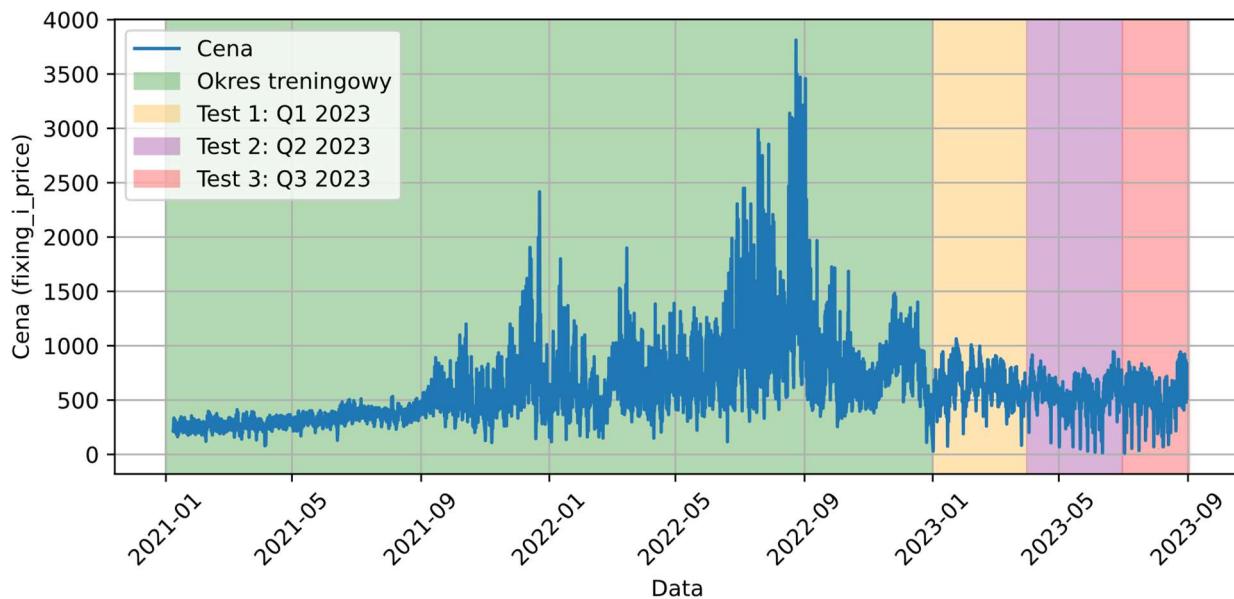
Zastosowano oczywiście architekturę, która osiągnęła najlepsze rezultaty na danych walidacyjnych. Takie podejście pozwoliło na zachowanie spójności i rzetelności wyników. Celem tego rozdziału jest zbadanie możliwości wykorzystania uczenia ciągłego w predykcji szeregów czasowych, zatem uniknięcie różnic wynikających z doboru hiperparametrów jest niezwykle istotne. Zastosowanie innej architektury sieci mogłoby zfałszować rezultaty i doprowadzić do błędnych wniosków.

6.2 Przegląd strategii zastosowanych w modelowaniu

Autor: Adam Mantiuk

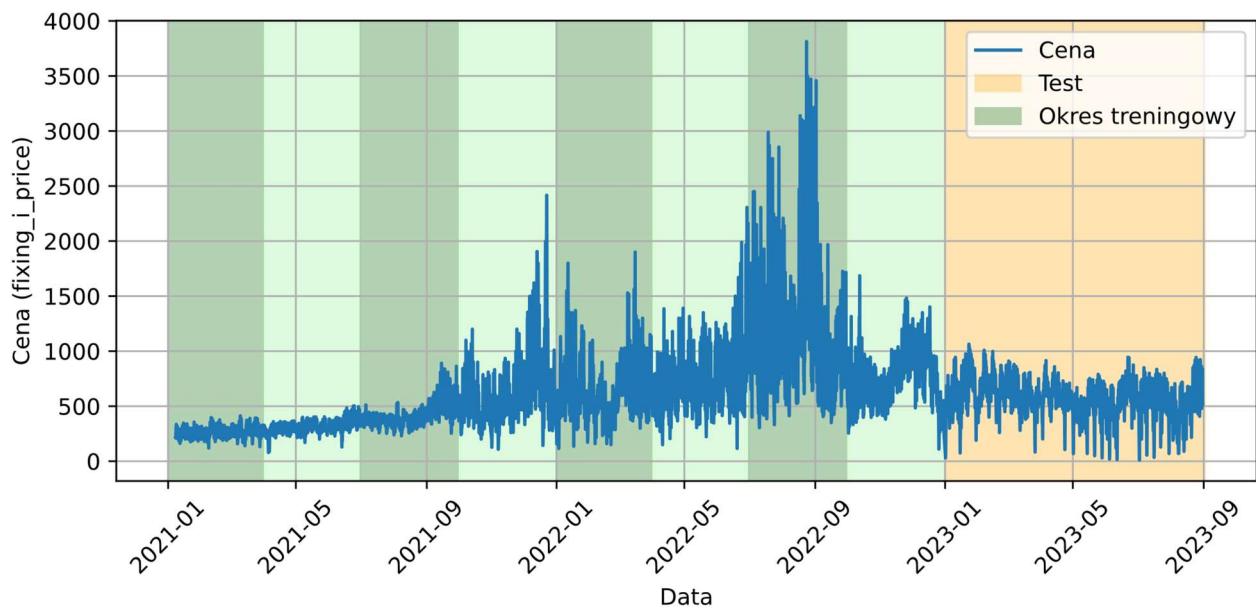
Podczas modelowania kluczową rolę odegrał sposób podzielenia danych. W uczeniu ciągłym model podlega ciągłym aktualizacjom. Należy podjąć zatem decyzję jak często takie aktualizacje powinny występować. W tej pracy przetestowane zostały różne strategie podziału danych.

Pierwsza strategia zakładała naukę na zbiorze treningowym w sposób klasyczny, a więc taki jaki został przedstawiony w poprzednim rozdziale. Zbiór treningowy został jednak podzielony zgodnie z kwartałami roku. W zbiorze testowym znalazły się zatem dwa pełne kwartały oraz dwa pierwsze miesiące trzeciego kwartału 2023 roku. Model aktualizowany był po pierwszym kwartale (danymi z tego kwartału) oraz po drugim kwartale. Przed aktualizacją następowała oczywiście predykcja poprzedniego modelu na ten okres czasu. Strategię numer 1 przedstawia rysunek 60.



Rysunek 60 Strategia numer 1
Opracowanie własne

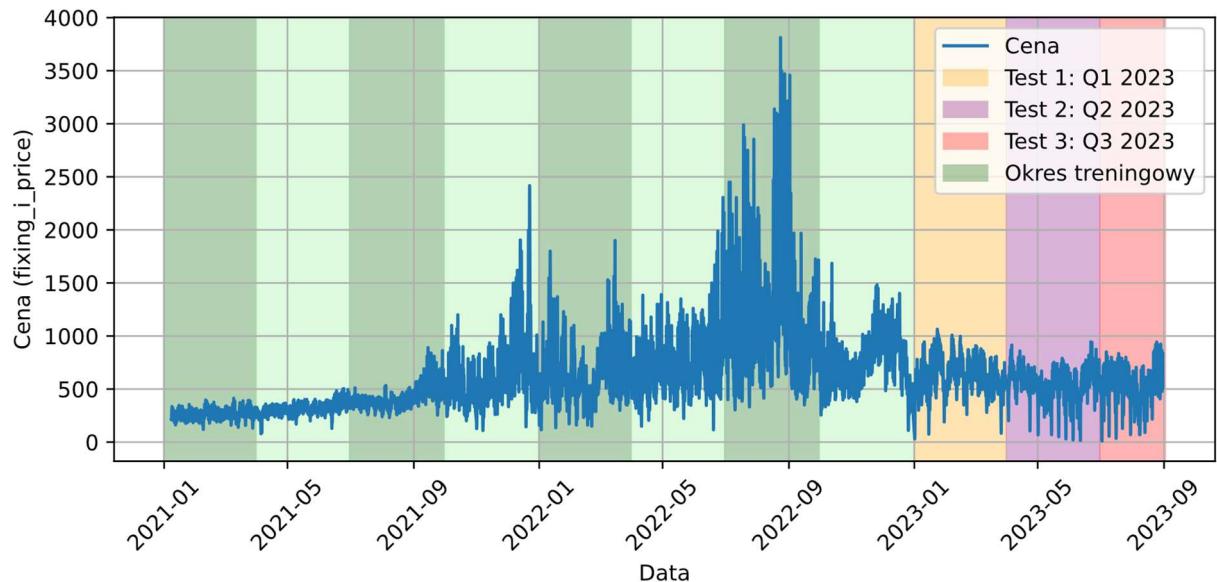
Druga strategia zakładała natomiast odwrotne podejście od strategii pierwszej. Dane treningowe zostały podzielone na trzymiesięczne paczki danych, które odpowiadały kolejnym kwartałom 2021 oraz 2022 roku. Model aktualizowany był zatem po każdym kolejnym kwartale, a następnie dokonywał predykcji na pierwsze osiem miesięcy 2023 roku. Strategię numer 2 przedstawia rysunek 61.



Rysunek 61 Strategia numer 2
Opracowanie własne

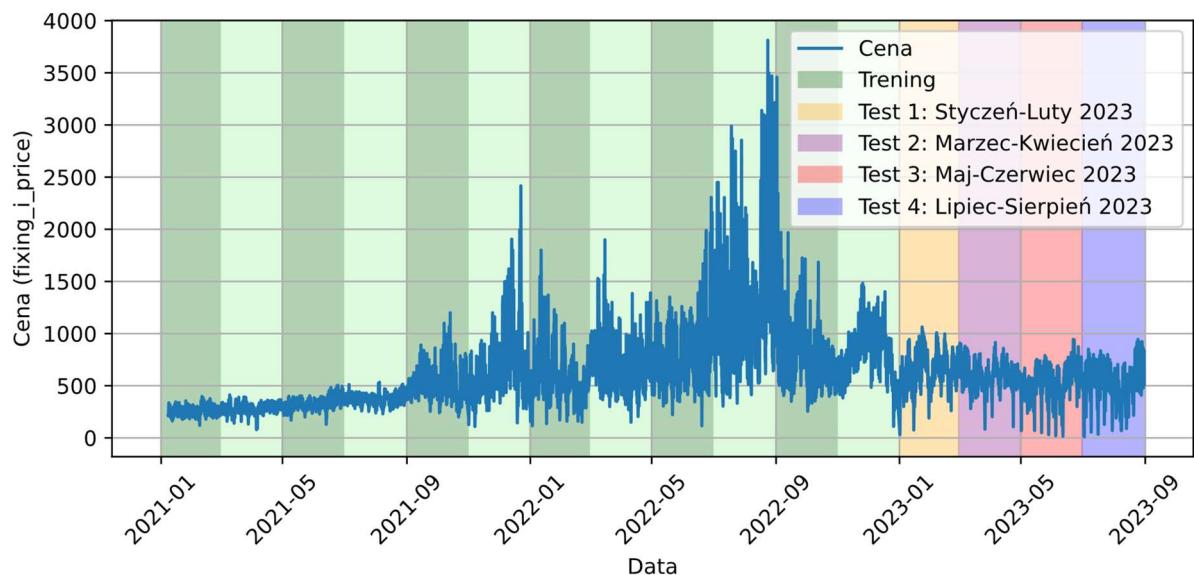
Trzecia strategia zakładała podzielenie danych treningowych na trzymiesięczne paczki danych, które odpowiadały kolejnym kwartałom 2021 oraz 2022 roku. Zatem podejście do danych treningowych zostało zaczerpnięte ze strategii numer 2. Dane

testowe również zostały podzielone według kwartałów, w ten sam sposób, w który został wykorzystany w strategii numer 1. Model aktualizowany był po każdym kwartale, a w przypadku danych testowych przed aktualizacją modelu następuała predykcja za pomocą modelu poprzedniego. Strategię numer 3 przedstawia rysunek 62.

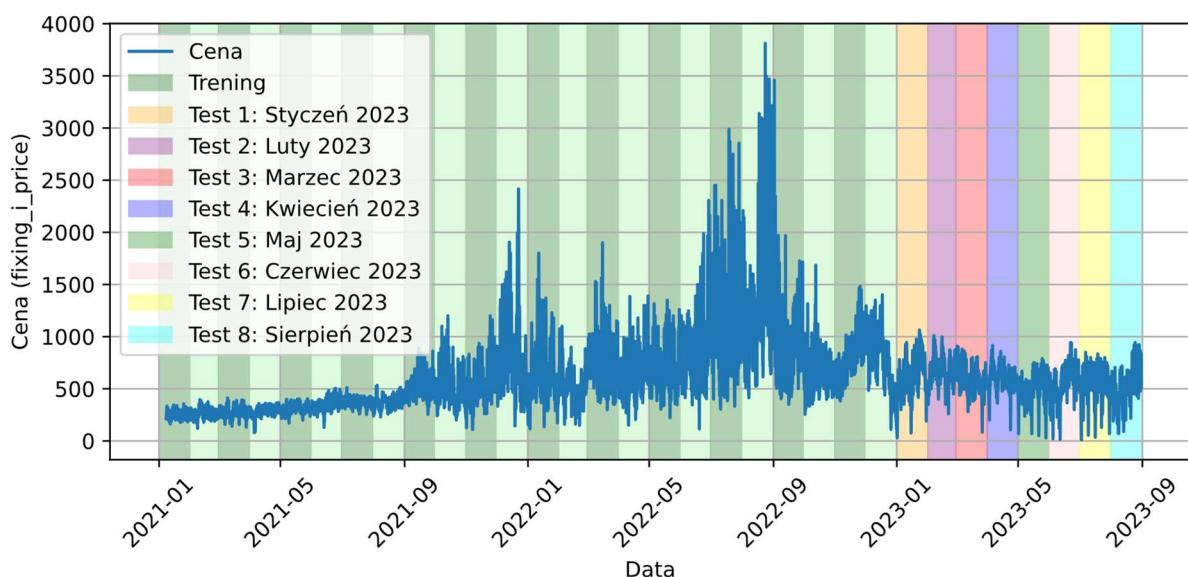


Rysunek 62 Strategia numer 3
Opracowanie własne

W strategii czwartej oraz piątej podejście było bardzo podobne do tego w strategii trzeciej. W tych strategiach paczki danych były mniejsze, i ich rozmiary wynosiły kolejno dwa miesiące oraz miesiąc. Model aktualizowany był przed każdą nową paczką danych, a w przypadku paczek testowych przed aktualizacją modelu dokonywano predykcji. Powyższe strategie zostały przedstawione na rysunku 63 oraz na rysunku 64.



Rysunek 63 Strategia numer 4
Opracowanie własne



Rysunek 64 Strategia numer 5
Opracowanie własne

Poniżej w tabeli 2 zostało przedstawione podsumowanie powyższych strategii zastosowanych w poniższej pracy.

Tabela 2 Podsumowanie strategii uczenia ciągłego
Opracowanie własne

Numer strategii	Dane treningowe	Częstotliwość aktualizacji modelu podczas treningu	Dane testowe	Częstotliwość aktualizacji modelu podczas testowania
1		Brak aktualizacji		Po każdym kwartale
2		Po każdym kwartale		Brak aktualizacji
3	Pełne lata 2021 oraz 2022	Po każdym kwartale	Od stycznia 2023 do końca sierpnia 2023	Po każdym kwartale
4		Co 2 miesiące		Co 2 miesiące
5		Po każdym miesiącu		Po każdym miesiącu

6.3 Zastosowane metody przeciwdziałające katastroficznemu zapominaniu

Autor: Adam Mantiuk

Na najlepszej strategii z poprzedniego podrozdziału zostały dodatkowo zbadane dwa mechanizmy przeciwdziałające katastroficznemu zapominaniu. Były to kolejno: odtwarzanie epizodów oraz elastyczna konsolidacja wag.

6.3.1 Odtwarzanie epizodów

Autor: Adam Mantiuk

Odtwarzanie epizodów nazywane również odtwarzaniem doświadczeń (ang. Experience Replay) jest techniką zapobiegającą katastroficznemu zapominaniu. Polega na mieszaniu starych danych z nowymi podczas aktualizacji modelu [27].

W tej pracy zdecydowano się na zastosowanie tej metody w celu zbadania jej możliwości pod kątem poprawienia wyników predykcji cen energii elektrycznej. W tym celu zaimplementowana została klasa ReplayBuffer. Konstruktor tej klasy inicjalizuje ją z określoną maksymalną liczbą przechowywanych próbek. Przechowywanie danych w buforze realizowane jest za pomocą kolejki dwustronnej, która przechowuje dane wyjściowe oraz wejściowe w postaci krotek. Kiedy kolejka osiągnie maksymalny rozmiar najstarsze dane zostają usunięte, aby stworzyć miejsce w kolejce nowym krotkom. Rozmiar bufora początkowo wynosił 1000, z czasem został stopniowo zwiększony do 8000. Podczas aktualizacji modelu z bufora pobierane były losowe próbki. Ich ilość określał parametr `replay_sample_size`, który początkowo wynosił 64 i stopniowo został zwiększony do 256.

6.3.2 Elastyczna Konsolidacja Wag

Autor: Adam Mantiuk

Metoda Elastycznej Konsolidacji Wag (ang. Elastic Weight Consolidation) jest kolejną metodą zwalczania katastroficznego zapominania. Implementuje ona mechanizm ograniczający zmiany wag, które były istotne dla wcześniejszej wersji modelu. Aby określić, które parametry są najważniejsze obliczane są elementy diagonalne macierzy Fishera, na podstawie drugich pochodnych funkcji strat w minimum (możliwa jest też aproksymacja przez pierwsze pochodne). Aby zmniejszyć odchylenie nowych parametrów od starych, najbardziej istotnych, wprowadza się dodatkowy składnik do funkcji strat [28].

Funkcja strat w takim przypadku obliczana jest według poniższego wzoru [28]:

$$L(\theta) = L_B(\theta) + \sum_i \frac{\lambda}{2} F_i (\theta_i - \theta_{A,i}^*)^2 \quad (24)$$

gdzie:

$L(\theta)$ – funkcja strat

$L_B(\theta)$ – funkcja strat dla nowych danych

λ – współczynnik określający ważność poprzednich zadań w stosunku do nowego zadania

F_i – elementy diagonalne macierzy Fishera, określające, które parametry są najważniejsze

θ_i – parametry modelu B

$\theta_{A,i}^*$ – parametry optymalne dla wcześniejszego modelu A

Aby zaimplementować mechanizm Elastycznej Konsolidacji Wag w tej pracy, zdecydowano się na implementację funkcji `compute_fisher_information`. Oblicza ona aproksymowaną macierz Fishera za pomocą gradientów funkcji strat względem wag modelu. Następnie gradienty podnoszone są do kwadratu. W celu wprowadzenia mechanizmu regularyzacji została zaimplementowana klasa `EWCloss`. Konstruktor tej klasy przyjmuje wszystkie niezbędne parametry, a więc bazową funkcję straty, macierz Fishera, poprzednie wagi modelu oraz współczynnik regularyzacji `lambda` (ten współczynnik kontroluje, jak silnie mają być karane zmiany w wagach). Główną funkcjonalnością tej klasy jest połączenie standardowej funkcji straty ze składnikiem regularizacyjnym. Wszystkie obliczenia dokonywane są w metodzie `call`, której funkcjonalność została zaimplementowana zgodnie z równaniem (24). Wartość współczynnika regularyzacji `lambda` początkowo została ustawiona na 1, a z czasem zwiększona do 500.

7 Ocena modeli oraz porównanie wyników

7.1 Wyniki modeli klasycznych

Autor: Mikołaj Kalisz

Podczas nauki dodatkowo były dokonane zmiany w danych wejściowych. W celu zobaczenia, jak ilość danych o cenie energii z poprzednich dni wpłynie na dokładność modelu. Dla obu wariantów został wykonany wyżej opisany proces automatycznego dobierania hiperparametrów.

Na początku porównane zostaną modele różniące się występowaniem cech

- Cena prądu minus 2 dni
- Cena prądu minus 3 dni
- Cena prądu minus 4 dni
- Cena prądu minus 5 dni
- Cena prądu minus 6 dni

Ze względu na obawę, że liczne wystąpienie skorelowanych ze sobą wartości może spowodować osłabienie wyników poprzez zwiększenia dopasowania do wartości treningowych, porównanie modeli bez i z tymi dodatkowymi cechami może być kluczowe dla dalszych wyników.

Tabela 3 Wyniki modelu bez dodatkowych cech. Opracowanie własne

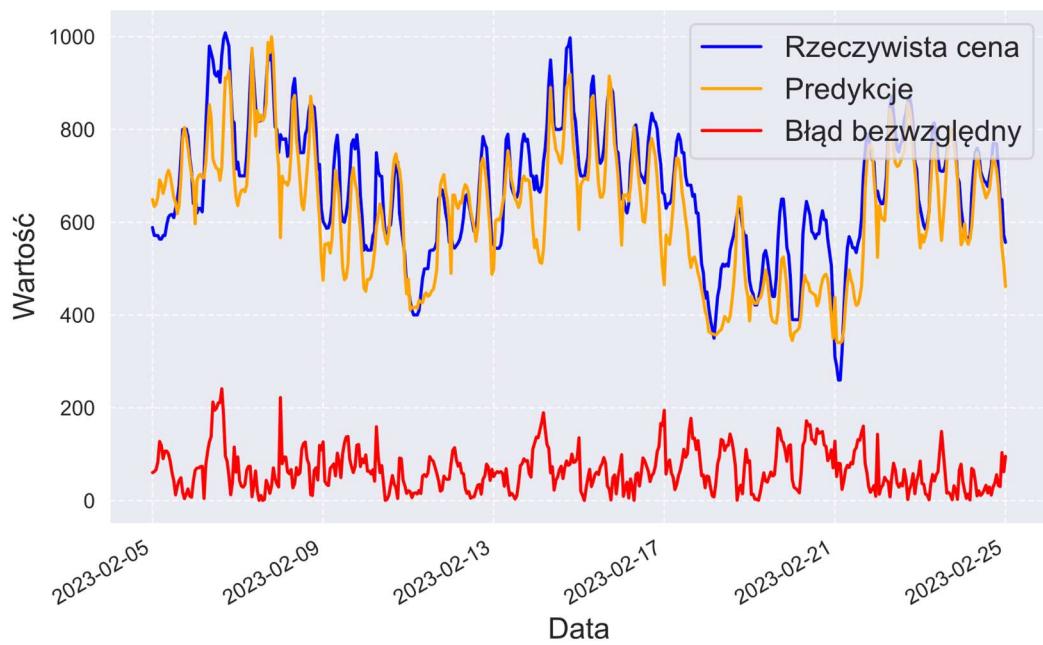
Dane	Treningowe		Walidacyjne		Testowe	
Nazwa metryki	MAE	R2	MAE	R2	MAE	R2
Wartość	88,16	0,843	127,73	0,809	66,5	0,714

Tabela 4 Wyniki modelu z dodatkowymi cechami. Opracowanie własne

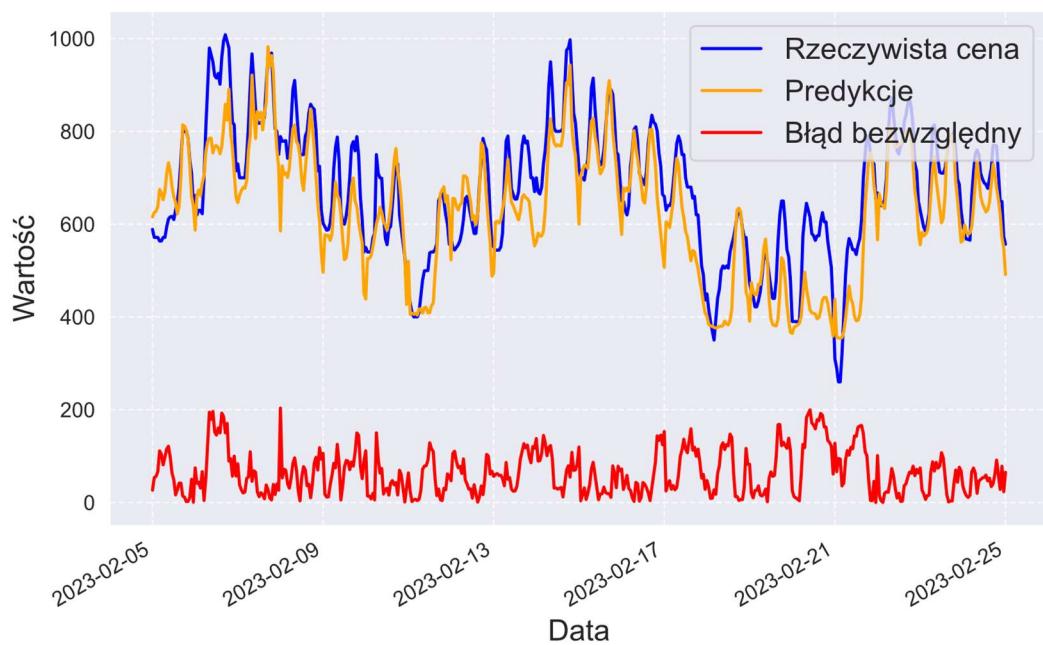
Dane	Treningowe		Walidacyjne		Testowe	
Nazwa metryki	MAE	R2	MAE	R2	MAE	R2
Wartość	79	0,868	133	0,798	66,1	0,716

Tabela 5 Porównanie hiperparametrów modeli. Opracowanie własne

	Model bez dodatkowych cech	Model z dodatkowymi cechami
Warstwy	(256, 128, 64, 32)	(256, 128, 64, 32)
Batch	32	32
Epoki	56	13
Tempo uczenia	0,005	0,005
Dropout	0,2	0,1
Optymalizator	Adam	Adam



Rysunek 65 Przebieg wartości rzeczywistej, predykcji oraz błędu bezwzględnego między tymi wartościami dla modelu bez dodatkowych cech. Opracowanie własne



Rysunek 66 Przebieg wartości rzeczywistej, predykcji oraz błędu bezwzględnego między tymi wartościami dla modelu z dodatkowymi cechami. Opracowanie własne

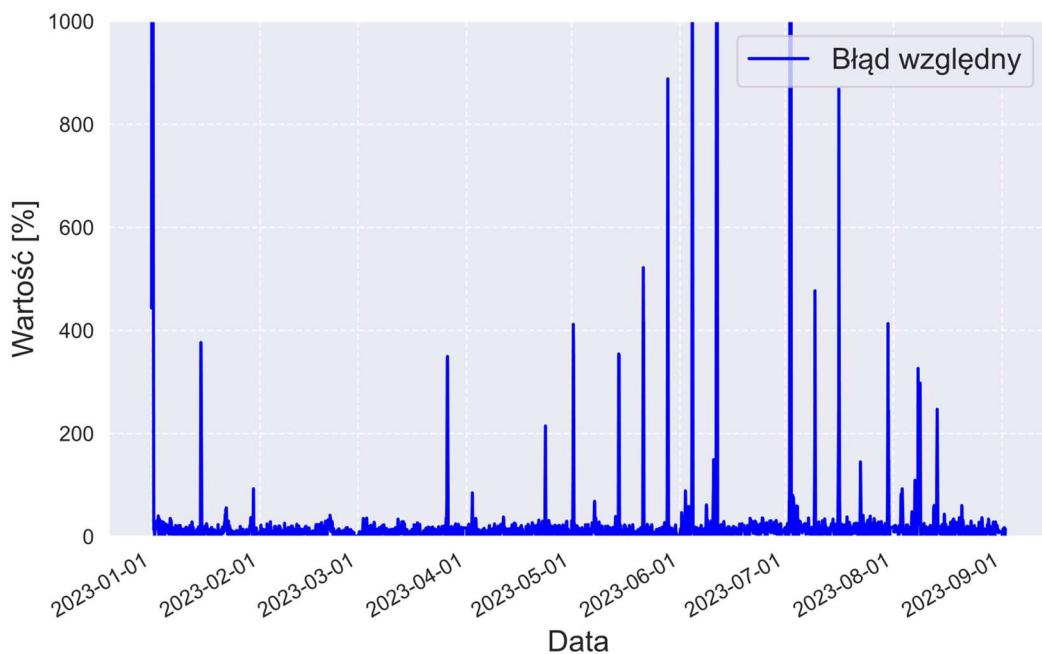
Jak można zauważać na przebiegach:

- na rysunku 65 - dla modelu bez dodatkowych cech,
- na rysunku 66 - dla modelu z dodatkowymi cechami,

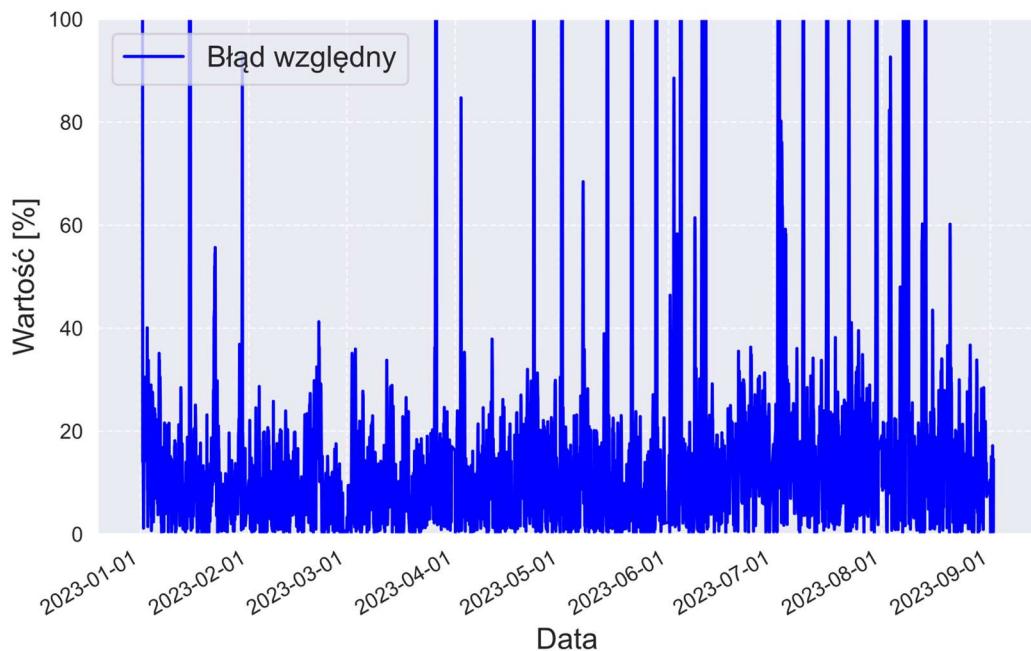
oraz w tabelach z wartościami metryk, odpowiednio w 3 i 4 oba modele mają bardzo zbliżone do siebie wyniki. Największą zmianą jest zauważalna różnica w wartości metryki R2 dla zbioru treningowego. Większa wartość dla modelu z dodatkowymi cechami może sugerować lepsze dopasowanie trendów w danych treningowych, jednakże nie widać zmiany wyników dla pozostałych zbiorów, więc nie jest to przypadek za dużego dopasowania do danych treningowych.

W dalszych wynikach zdecydowano się na korzystanie ze zbioru z dodatkowymi cechami. Warto zwrócić uwagę na fakt zastosowania technik zapobiegających nadmiernemu dopasowaniu do danych treningowych, takich jak wcześnie kończenie, czy dropout.

Korzystając z wybranego modelu zrobiono predykcję na danych testowych z 2023 roku. Dane metryk opisujące dokładność modelu przedstawiono w tabeli 4. Poniżej na rysunku 67 przedstawiono wartości błędu względnego. Jak można zauważać wartości błędu większość czasu są stosunkowo małe, jednakże widoczne są momenty nagłych zmian tego wskaźnika. Po dokładniejszym przyjrzeniu się tym wartościom na rysunku 68, można zauważać że dla większości przedziału błąd względny mieści się w przedziale od 0 do 35%.



Rysunek 67 Zależność wartości błędu względnego w procentach od czasu
Opracowanie własne



Rysunek 68 Zależność wartości błędu względnego w procentach od czasu. Dokładniejsza analiza
Opracowanie własne

Dodatkowym spostrzeżeniem może być nieznaczne zwiększenie się błędu im dalej od pierwszych wartości predykci. Poniżej w tabeli 6 przedstawiono wartości metryk dla poszczególnych miesięcy. W styczniu występuje najlepszy wynik metryki R². W następnych miesiącach wartość ta spada, osiągając minimum w kwietniu. Brak ciągłego aktualizowania parametrów modelu sprawia, że model nie ma możliwości dostosowania się do zmieniającego rynku potencjalnie wpływając negatywnie na dokładność predykacji.

Tabela 6 Zestawienie wartości metryk MAE oraz R2 dla poszczególnych miesięcy
Opracowanie własne

Miesiąc	MAE	R2
Styczeń	76,02	0,76
Luty	57,43	0,71
Marzec	54,00	0,69
Kwiecień	59,66	0,42
Maj	52,47	0,67
Czerwiec	77,59	0,64
Lipiec	86,11	0,60
Sierpień	65,92	0,70

7.2 Wyniki modeli z Uczeniem Ciągły

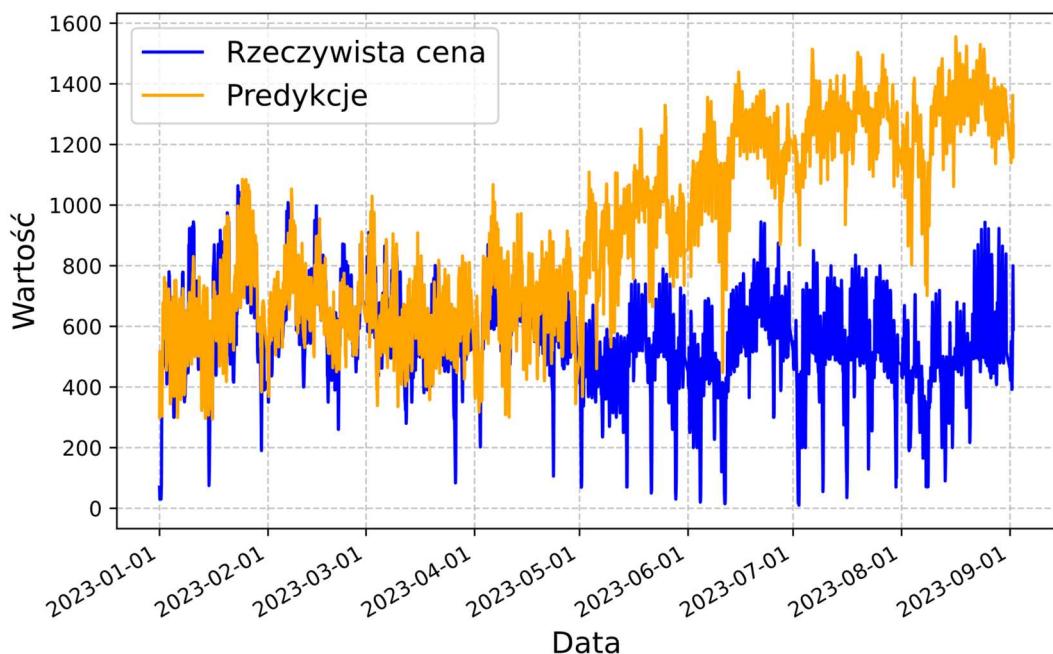
Autor: Adam Mantiuk

W tabeli 7 zostały przedstawione wyniki jakie osiągnęły poszczególne strategie.

Tabela 7 Porównanie wyników poszczególnych strategii
Opracowanie własne

Numer strategii	1		2		3		4		5	
Metryka	MAE	R2	MAE	R2	MAE	R2	MAE	R2	MAE	R2
Wartość	60,12	0,73	277	-3,2	59,32	0,733	69,17	0,61	67,23	0,64

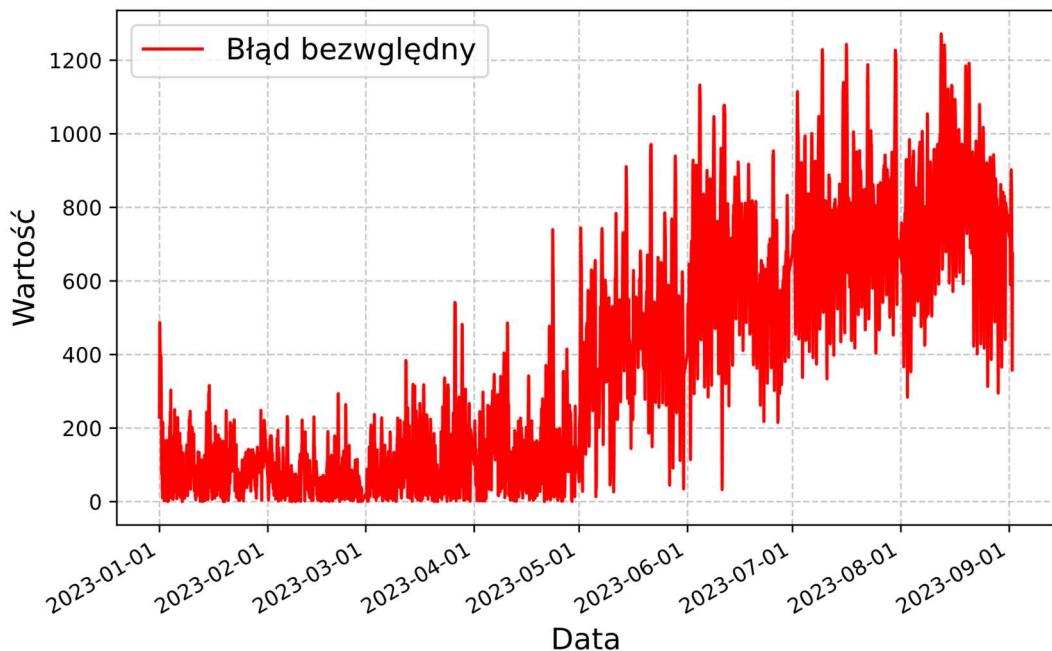
Najgorsze wyniki osiągnął model oparty o strategię numer 2. Na rysunku 69 został przedstawiony przebieg wartości rzeczywistych oraz predykcji.



Rysunek 69 Przebieg wartości rzeczywistych oraz predykcji modelu opartego o strategię numer 2
Opracowanie własne

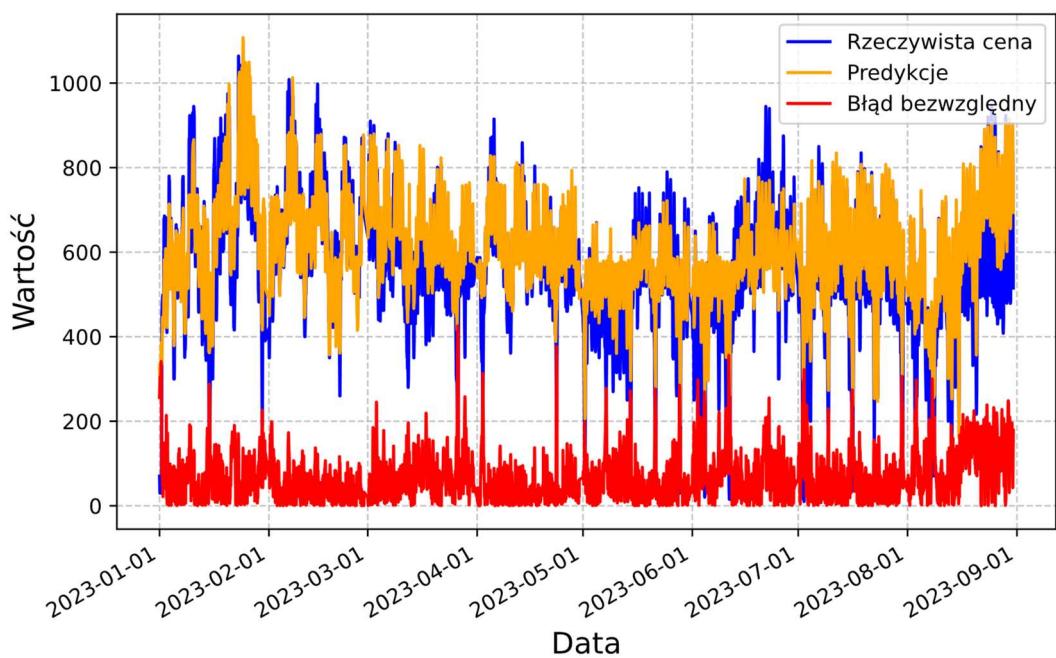
Na rysunku 70 została przedstawiona wartość błędu bezwzględnego. Możemy zauważyc, że jego wartość zwiększa się drastycznie na początku maja. Jest to najprawdopodobniej spowodowane specyfiką strategii numer 2, w której model uczony był na trzymiesięcznych paczkach danych, a następnie nie był poddawany aktualizacjom. Świadczy to o zmienności rynku w czasie, oraz że modele oparte o

modele ciągłe muszą być nieustannie aktualizowane w przypadku predykcji szeregów czasowych.

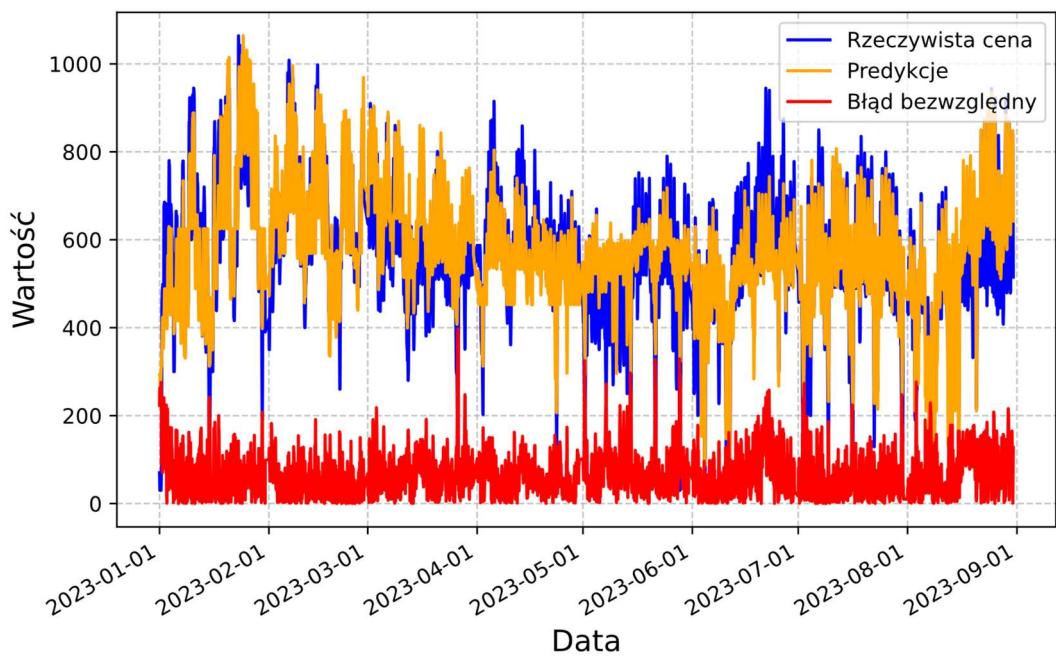


Rysunek 70 Wartość błędu bezwzględnego przy strategii numer 2
Opracowanie własne

Wyniki znacznie lepsze osiągnęły modele oparte o strategie numer 4 oraz 5. Z całą pewnością o lepszych wynikach zadecydowały aktualizacje modelu na kolejnych zbiorach testowych. Jednak w obu przypadkach paczki danych, na których były trenowane i aktualizowane modele, były zbyt małe by osiągnąć najlepsze rezultaty wśród wszystkich strategii. Na rysunku 71 oraz rysunku 72 zostały przedstawione przebiegi wartości rzeczywistych oraz predykcji modeli opartych o powyższe strategie.

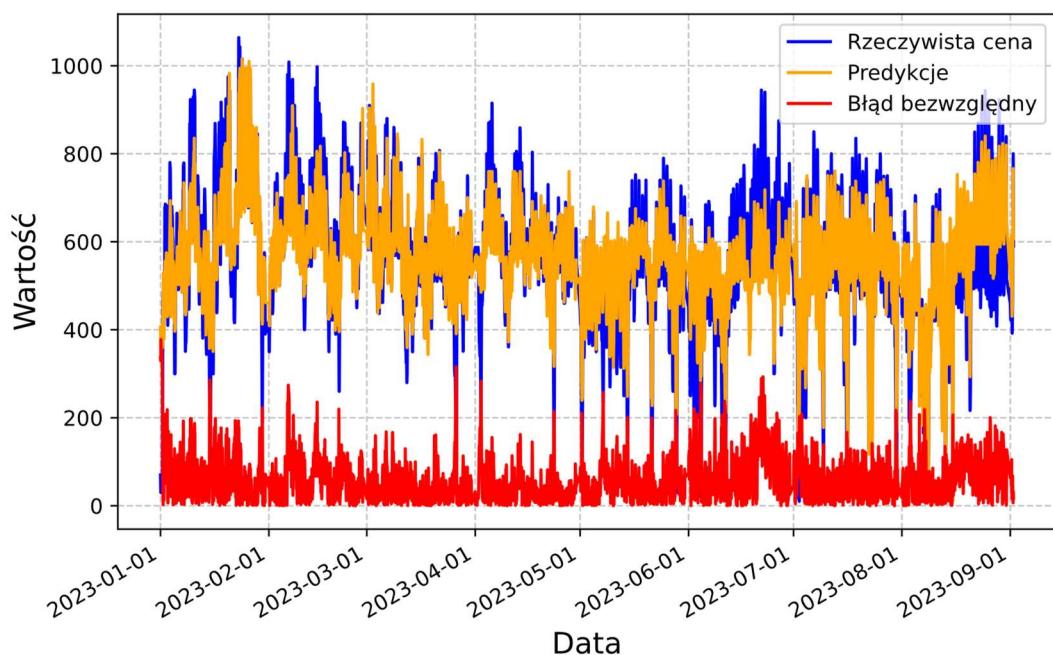


Rysunek 71 Przebiegi modelu opartego o strategię numer 4
Opracowanie własne



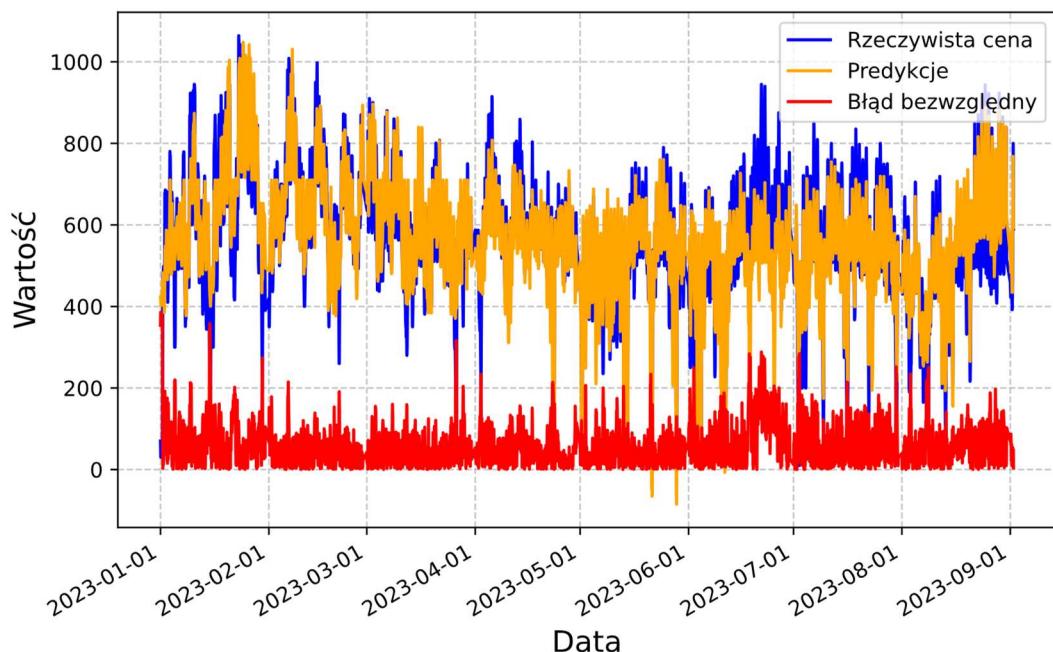
Rysunek 72 Przebiegi modelu opartego o strategię numer 5
Opracowanie własne

Najlepsze wyniki osiągnęły strategie numer 1 oraz numer 3. Przebieg wartości rzeczywistych oraz predykcji modelu opartego o strategię numer 1 przedstawiono na rysunku 73. Na przebiegi zostały też naniesione wartości błędu bezwzględnego.



Rysunek 73 Przebiegi modelu opartego o strategię numer 1
Opracowanie własne

Natomiast przebieg wartości rzeczywistych oraz predykcji modelu opartego o strategię numer 3 przedstawiono na rysunku 74. W tym przypadku również na przebiegi zostały naniesione wartości błędu bezwzględnego.



Rysunek 74 Przebiegi modelu opartego o strategię numer 3
Opracowanie własne

W tabeli 8 zostały przedstawione wyniki (modelu opartego o strategię numer 3) w poszczególnych kwartałach 2023 roku.

*Tabela 8 Wyniki modelu opartego o strategię numer 3 w poszczególnych kwartałach 2023 roku.
Opracowanie własne*

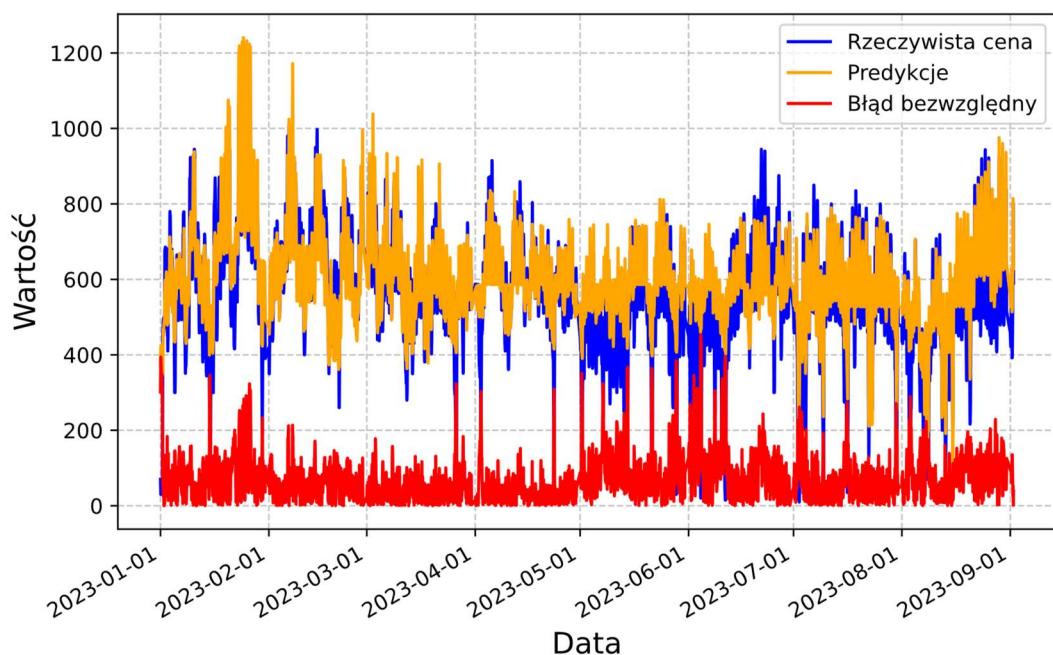
Kwartał	Pierwszy	Drugi	Trzeci
Metryka		R2	
Wartość	0,77	0,68	0,75

Na najlepszej strategii, którą okazała się być strategia numer 3, zostały przetestowane mechanizmy przeciwdziałające katastroficznemu zapominaniu. Jak przedstawia tabela 9, nie poprawiły one wyników, lecz je pogorszyły.

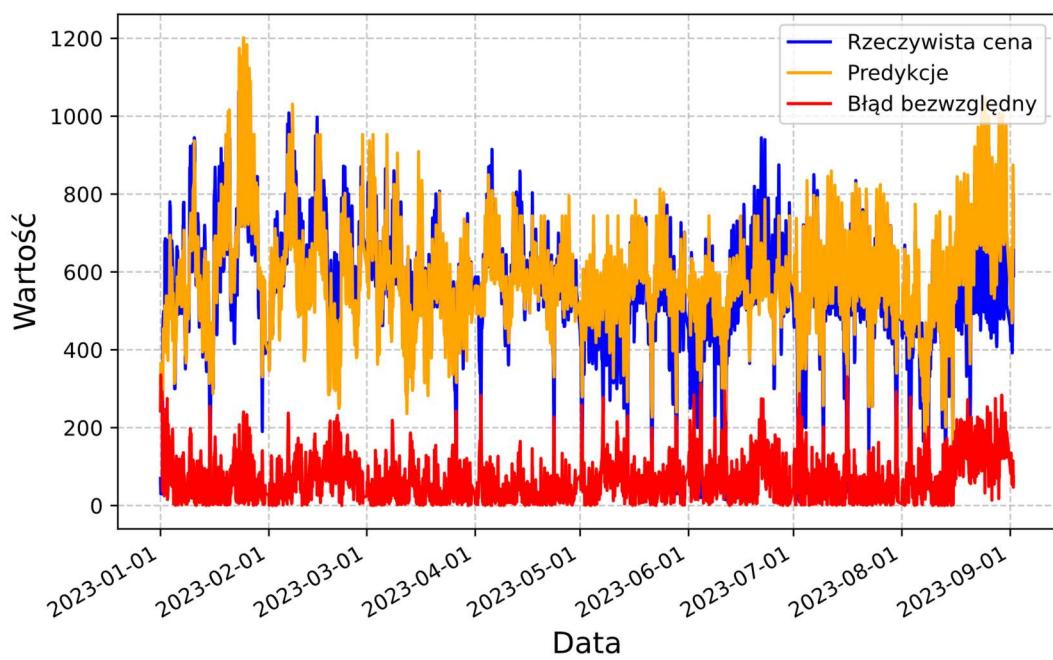
Tabela 9 Wyniki modelu opartego na strategii numer 3 z zastosowanymi mechanizmami przeciwdziałającymi katastroficznemu zapominaniu. Opracowanie własne

Mechanizm	Odtwarzanie epizodów		Elastyczna Konsolidacja Wag	
	MAE	R2	MAE	R2
Metryka				
Wynik	62,49	0,71	71,75	0,59

Przebiegi wartości rzeczywistych oraz predykcji modelu opartego o strategię numer 3 z zastosowanymi mechanizmami katastroficznego zapominania przedstawiono na rysunkach poniżej. Rysunek 75 przedstawia przebiegi z zastosowaniem mechanizmu Elastycznej Konsolidacji Wag, natomiast rysunek 76 przedstawia przebiegi z zastosowaniem mechanizmu odtwarzania epizodów. W tym przypadku również na przebiegi zostały naniesione wartości błędu bezwzględnego.



Rysunek 75 Przebiegi dla modelu opartego o strategię numer 3 z zastosowanym mechanizmem Elastycznej Konsolidacji Wag. Opracowanie własne



Rysunek 76 Przebiegi dla modelu opartego o strategię numer 3 z zastosowanym mechaizmem odtwarzania epizodów. Opracowanie własne

8 Podsumowanie i wnioski

8.1 Wnioski z porównania technik uczenia

Autor: Adam Mantiuk

Modele, które wykorzystały strategie uczenia ciągłego osiągnęły minimalnie lepsze rezultaty od rozwiązania opartego o podejście klasyczne trenowania sztucznych sieci neuronowych. Poziom skomplikowania rozwiązania opartego o techniki uczenia ciągłego nie jest rekompensowany przez znacznie lepsze rezultaty jakie osiąga model. Obserwując wyniki, jakie osiąga model oparty o strategię uczenia ciągłego, w poszczególnych kwartałach 2023 roku można zauważać spadek jakości predykcji w drugim kwartale oraz jej poprawę w trzecim kwartale. Nie można jednak potwierdzić, że poprawa wynika z aktualizacji modelu, ponieważ bardzo podobną zależność można zaobserwować w przypadku podejścia klasycznego.

8.2 Zalecenia dla przyszłych badań

Autor: Mikołaj Kalisz

W pracy porównano wyniki modeli z zastosowaniem uczenia ciągłego oraz bez niego, opierając się na tej samej architekturze, dobranej pod kątem uzyskania jak najlepszych wyników na zbiorze walidacyjnym w klasycznym podejściu. Nie uwzględniono jednak faktu, że otrzymana architektura może nie być optymalna. Zastosowanie takiego samego procesu doboru hiperparametrów potencjalnie pozwoliłoby na osiągnięcie większej różnicy w wynikach końcowych. Forma rywalizacji w dążeniu do optymalnych wyników mogłaby dostarczyć dodatkowych, interesujących obserwacji.

Ze względu na zmianę sposobu raportowania danych przez PSE pojawiły się trudności z uzyskaniem większego przedziału czasowego dla zbioru danych dotyczącego ceny energii. Warto zbadać, jak model stosujący uczenie ciągłe radziłby sobie w dłuższej perspektywie czasowej oraz czy potrafiłby dostosowywać się do zmieniającego się rynku.

Trudnością napotkaną podczas tworzenia pracy była ilość cech do analizy. Z racji na dużą zmienność rynku i wiele czynników wpływających na cenę, odpowiedni dobór i przejrzenie ponad 153 cech jest wyzwaniem. Potencjalnie wprowadza to możliwość przegapienia ważnych zależności, które są widoczne dopiero przy głębszej analizie, przykładowo z podziałem na mniejsze okresy.

Bibliografia

- [1] B. Adamska, M. Bando, K. Bielecki, J. Buzek, T. Dąbrowski, M. Dusiło, R. Gawin, M. Graniżewska, P. Grądzik, A. Grzeszak, A. Jeżewski, L. Juchniewicz, H. Kaliś i M. Krasnodębska-Tomkiel, *W drodze do bezpiecznej i czystej energii, czyli jak napisać rynek na nowo 25 LAT REGULACJI RYNKÓW ENERGII W POLSCE*, Warszawa: Urząd Regulacji Energetyki, 2022.
- [2] R. Kasperowicz, „Struktura rynku oraz formy handlu energią elektryczną w Polsce,” *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Poznaniu*, 2011.
- [3] Urząd Regulacji Energetyki, „Magazynowanie Energii Elektrycznej,” Urząd Regulacji Energetyki, Warszawa, 2024.
- [4] Next Kraftwerke, „Czym jest obrót energią elektryczną na Rynku Dnia Następnego?,” Next Kraftwerke.
- [5] D. Mrowiec, „Analiza merit order sektora wytwórczego w Polsce do 2030 roku,” PSE Innowacje.
- [6] Instytut Energii Odnawialnej, „Analiza trendów cen energii wraz z prognozą do 2030r.,” Instytut Energii Odnawialnej, Wrocław, 2020.
- [7] M. Szaleniec i R. Tadeusiewicz, LEKSYKON SIECI NEURONOWYCH, 2015.
- [8] S. Marsland, Machine Learning - An Algorithmic Perspective.
- [9] B. Kainz, „Deep Learning – Activation Functions,” [Online]. Available: https://www.doc.ic.ac.uk/~bkainz/teaching/DL/L03_DL_activation_and_losses.pdf.
- [10] S. R. Dubey, S. K. Singh and B. B. Chaudhuri, Activation Functions in Deep Learning: A Comprehensive Survey and Benchmark.
- [11] scikit-learn, „scikit-learn,” [Online]. Available: <https://scikit-learn.org/>.
- [12] S. C. Nayak, B. B. Misra and H. S. Behra, Impact of Data Normalization on Stock Index Forecasting.
- [13] S. Bhanja i A. Das, „Impact of Data Normalization on Deep Neural Network for Time Series Forecasting”.
- [14] D. E. Birba, *A Comparative study of data splitting algorithms for machine learning model selection.*
- [15] A. Lopez Davila, *Neural Networks: The Backpropagation Algorithm.*
- [16] D. P. Kingma and J. L. Ba, *ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION.*

- [17] B. H. Pansambal i A. B. Nandgaokar, „Integrating Dropout Regularization Technique at Different Layers to Improve the Performance of Neural Networks,” *International Journal of Advanced Computer Science and Applications*, 2023.
- [18] J. Perzyńska, Zastosowania sztucznych sieci neuronowych do prognozowania szeregow czasowych, Szczecin, 2018.
- [19] Z. Chen and B. Liu, Lifelong Machine Learning, Second Edition, Morgan & Claypool, 2018.
- [20] D. Kuditipudi, M. Aguilar-Simon, J. Babb, M. Bazhenov, D. Blackiston, J. Bongard, A. P. Brna, S. Chakravarthi Raja, N. Cheney, J. Clune, A. Daram, S. Fusi, P. Helfer, L. Kay i N. Ketz, „Biological underpinnings for lifelong learning machines,” *Nature Machine Intelligence*, Marzec 2022.
- [21] R. Svoboda, S. Basterrech, J. Kozal, J. Platoš and M. Woźniak, "A Natural Gas Consumption Forecasting System for Continual Learning Scenarios based on Hoeffding Trees with Change Point Detection Mechanism," *Knowledge-Based Systems*, Wrzesień 2023.
- [22] T. Nguyen, C.-H. Hua, U. Ur Rehman, M. Hussain and S. Lee, "CLEB: A Continual Learning Energy Bidding Framework For An Energy Market Bidding Application," in *In Proceedings of the 2023 6th Artificial Intelligence and Cloud Computing Conference (AICCC '23)*. Association for Computing Machinery, New York, 2024.
- [23] D. Chicco, M. J. Warrens and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, Lipiec 2021.
- [24] P. Bruce, A. Bruce i P. Gedeck, Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python 2nd Edition, Sebastopol: O'Reilly Media, 2020.
- [25] D. Pyle, Data Preparation for Data Mining, San Francisco: Morgan Kaufmann Publishers, Inc., 1999.
- [26] Apache Software Foundation, „arrow.apache,” [Online]. Available: <https://arrow.apache.org/>.
- [27] D. Rolnick, A. Ahuja, J. Schwarz, T. Lillicrap and G. Wayne, "Experience Replay for Continual Learning," in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, Vancouver, 2018.
- [28] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, D. Hassabis, C. Clopath, D. Kumaran i R. Hadsell, „Overcoming Catastrophic Forgetting in Neural Networks,” *Proceedings of the National Academy of Sciences (PNAS)*, 14 Marzec 2017.
- [29] K. Chellapilla, "Neural Networks, Find the Function — A Basic Introduction to Neural Networks," *The Batch*, 25 Maj 2022.

Wykaz symboli i skrótów

CSV – comma-separated values (wartości rozdzielone przecinkiem)

MAE – Mean Absolute Error (średni błąd bezwzględny)

R2 – Współczynnik determinacji R²

Spis Rysunków

Rysunek 1 Moc zakontraktowana przez magazyny energii w aukcjach głównych na lata 2021–2028 Opracowanie pochodzi z [3]	20
Rysunek 2 Przedstawiona jest przykładowa struktura perceptronu składającego się z warstw wejściowej, wyjściowej i dwóch ukrytych. Opracowanie własne	25
Rysunek 3 Wykres funkcji liniowej, pokazuje zależność wejścia s od wyjścia. Opracowanie własne	26
Rysunek 4 Wykres przedstawiający wartość funkcji sigmoidalnej w zależności od sumy ważonej na wejściu. Opracowanie własne.....	27
Rysunek 5 Wykres zależności wartości funkcji aktywacji ReLu od sumy ważonej na wejściu s. Opracowanie własne.....	28
Rysunek 6 Funkcja LeakyReLu. Dla wartości s mniejszych od 0 można zauważać delikatny spadek wartości funkcji (w przeciwnieństwie do ReLu). Opracowanie własne.....	29
Rysunek 7 Funkcja tangensoidalna. w zakresie od -2 do 2 można zauważać gwałtowny wzrost wartości funkcji. Opracowanie własne.....	30
Rysunek 8 Przykładowa architektura sieci pokazująca wyłączone neurony metodą dropout. Opracowanie własne	36
Rysunek 9 „Przykładowe odwzorowanie szeregu czasowego na sieć neuronową wyznaczającą prognozę” Opracowanie pochodzi z [18]	37
Rysunek 10 Key features required to achieve lifelong learning. Opracowanie pochodzi z [20] ...	38
Rysunek 11 Schemat adaptacji modelu i wykrywania zmian dystrybucji danych w strumieniach danych Opracowanie pochodzi z [22].....	45
Rysunek 12 Średnia cena energii w poszczególnych latach Opracowanie własne.....	56
Rysunek 13 Histogram przestawiający rozkład cen energii na całym zbiorze danych Opracowanie własne	57
Rysunek 14 Histogramy dla lat 2021 oraz 2022 Opracowanie własne.....	57
Rysunek 15 Wykres pudełkowy dla całego zbioru danych Opracowanie własne	58
Rysunek 16 Wykresy pudełkowe cen energii w poszczególnych latach Opracowanie własne ...	59
Rysunek 17 Rozkład cen energii w poszczególnych miesiącach 2021 roku Opracowanie własne	60
Rysunek 18 Rozkład cen energii w poszczególnych miesiącach 2022 roku Opracowanie własne	61
Rysunek 19 Średnie ceny energii w poszczególnych miesiącach Opracowanie własne	62
Rysunek 20 Średnia cena energii w poszczególnych dniach tygodnia Opracowanie własne	62
Rysunek 21 Rozkład cen energii w poszczególnych dniach tygodnia Opracowanie własne	63
Rysunek 22 Średnia cena energii w poszczególnych porach roku Opracowanie własne	64
Rysunek 23 Rozkład cen energii w poszczególnych porach roku i latach Opracowanie własne..	64
Rysunek 24 Wykres zależności ceny energii od zapotrzebowania. Z narysowaną linią najlepszego dopasowania. Opracowanie własne	66

Rysunek 25 Wykres słupkowy przedstawiający wartości korelacji danych o zapotrzebowaniu na moc względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych.	64
Opracowanie własne	66
Rysunek 26 Wykres zależności ceny energii od rezerwy mocy poniżej zapotrzebowania, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	67
Rysunek 27 Wykres zależności ceny energii od rezerwy mocy ponad zapotrzebowanie, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	68
Rysunek 28 Wykres słupkowy przedstawiający wartości korelacji danych wartości rezerwy mocy poniżej zapotrzebowania względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	68
Rysunek 29 Wykres słupkowy przedstawiający wartości korelacji danych wartości rezerwy mocy ponad zapotrzebowanie względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	69
Rysunek 30 Macierz korelacji między wybranymi cechami. Opracowanie własne	70
Rysunek 31 Wykres zależności temperatury od zapotrzebowania na moc, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	71
Rysunek 32 Wykres słupkowy przedstawiający wartości korelacji danych wartości temperatury względem zapotrzebowania na moc, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	71
Rysunek 33 Wykres słupkowy przedstawiający wartości korelacji wartości temperatury względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	72
Rysunek 34 Macierz korelacji cech powiązanych z wiatrem oraz ceny energii. Opracowanie własne	73
Rysunek 35 Wykres zależności generacji źródeł wiatrowych od produkcji energii wiatrowej, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	74
Rysunek 36 Wykres zależności produkcji energii wiatrowej od prędkości wiatru, wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	74
Rysunek 37 Wykres słupkowy przedstawiający wartości korelacji wartości generacji źródeł wiatrowych względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	75
Rysunek 38 Macierz korelacji cech powiązanych ze słońcem oraz ceny energii. Opracowanie własne	76
Rysunek 39 Wykres słupkowy przedstawiający wartości korelacji wartości generacji źródeł fotowoltaicznych względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	77
Rysunek 40 Wykres zależności ceny energii od generacji źródeł fotowoltaicznych wraz z linią najlepszego dopasowania. Opracowanie własne	77
Rysunek 41 Macierz korelacji cech powiązanych z wodą oraz ceny energii. Opracowanie własne	78
Rysunek 42 Wykres słupkowy przedstawiający wartości korelacji produkcji elektrowni szczytowo-pompowych względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	79
Rysunek 43 Wykres słupkowy przedstawiający wartości korelacji produkcji elektrowni przepływowej względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	79

Rysunek 44 Wykres słupkowy przedstawiający wartości korelacji produkcji elektrowni wodnej - zbiornikowej względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych.	
Opracowanie własne	80
Rysunek 45 Macierz korelacji cech powiązanych z węglem oraz ceny energii. Opracowanie własne	82
Rysunek 46 Wykres słupkowy przedstawiający siłę korelacji wartości produkcji energii z węgla kamiennego względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych.	
Opracowanie własne	83
Rysunek 47 Wykres słupkowy przedstawiający siłę korelacji wartości produkcji energii z węgla brunatnego względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych.	
Opracowanie własne	83
Rysunek 48 Wykres zależności ceny energii od produkcji energii z węgla brunatnego wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	84
Rysunek 49 Wykres zależności ceny energii od produkcji energii z węgla kamiennego wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	84
Rysunek 50 Macierz korelacji cech powiązanych z gazem oraz ceny energii. Opracowanie własne	86
Rysunek 51 Wykres zależności ceny energii od ceny gazu wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	86
Rysunek 52 Wykres zależności ceny gazu od produkcji energii z gazu wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne.....	87
Rysunek 53 Wykres słupkowy przedstawiający siłę korelacji wartości produkcji energii z gazu względem ceny energii, dla poszczególnych okresów oraz dla wszystkich danych. Opracowanie własne	87
Rysunek 54 Wykres zależności ceny energii od produkcji energii z gazu wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne.....	88
Rysunek 55 Wykres zależności ceny energii od produkcji energii z ropy wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne.....	89
Rysunek 56 Wykres zależności ceny energii od produkcji energii z biomasy wraz z narysowaną linią najlepszego dopasowania. Opracowanie własne	90
Rysunek 57 Macierz korelacji cech powiązanych z emisją CO ₂ oraz ceny energii. Opracowanie własne	91
Rysunek 58 Macierz korelacji cech powiązanych z ekoniemią oraz ceny energii. Opracowanie własne	92
Rysunek 59 Macierz korelacji danych historycznych wraz z ceną energii. Opracowanie własne	93
Rysunek 60 Strategia numer 1 Opracowanie własne	100
Rysunek 61 Strategia numer 2 Opracowanie własne	100
Rysunek 62 Strategia numer 3 Opracowanie własne	101
Rysunek 63 Strategia numer 4 Opracowanie własne	101
Rysunek 64 Strategia numer 5 Opracowanie własne	102
Rysunek 65 Przebieg wartości rzeczywistej, predykcji oraz błędu bezwzględnego między tymi wartościami dla modelu bez dodatkowych cech. Opracowanie własne	106
Rysunek 66 Przebieg wartości rzeczywistej, predykcji oraz błędu bezwzględnego między tymi wartościami dla modelu z dodatkowymi cechami. Opracowanie własne	106
Rysunek 67 Zależność wartości błędu względnego w procentach od czasu Opracowanie własne	107
Rysunek 68 Zależność wartości błędu względnego w procentach od czasu. Dokładniejsza analiza Opracowanie własne	108

Rysunek 69 Przebieg wartości rzeczywistych oraz predykcji modelu opartego o strategię numer 2 Opracowanie własne	109
Rysunek 70 Wartość błędu bezwzględnego przy strategii numer 2 Opracowanie własne	110
Rysunek 71 Przebiegi modelu opartego o strategię numer 4 Opracowanie własne	111
Rysunek 72 Przebiegi modelu opartego o strategię numer 5 Opracowanie własne	111
Rysunek 73 Przebiegi modelu opartego o strategię numer 1 Opracowanie własne	112
Rysunek 74 Przebiegi modelu opartego o strategię numer 3 Opracowanie własne	112
Rysunek 75 Przebiegi dla modelu opartego o strategię numer 3 z zastosowanym mechanizmem Elastycznej Konsolidacji Wag. Opracowanie własne	114
Rysunek 76 Przebiegi dla modelu opartego o strategię numer 3 z zastosowanym mechanizmem odtwarzania epizodów. Opracowanie własne	114

Spis Tabel

Tabela 1 Podstawowe statystyki opisowe dla kolumny "fixing_i_price" Opracowanie własne	56
Tabela 2 Podsumowanie strategii uczenia ciągłego Opracowanie własne	102
Tabela 3 Wyniki modelu bez dodatkowych cech. Opracowanie własne	105
Tabela 4 Wyniki modelu z dodatkowymi cechami. Opracowanie własne	105
Tabela 5 Porównanie hiperparametrów modeli. Opracowanie własne	105
Tabela 6 Zestawienie wartości metryk MAE oraz R2 dla poszczególnych miesięcy Opracowanie własne	108
Tabela 7 Porównanie wyników poszczególnych strategii Opracowanie własne	109
Tabela 8 Wyniki modelu opartego o strategię numer 3 w poszczególnych kwartałach 2023 roku. Opracowanie własne	113
Tabela 9 Wyniki modelu opartego na strategii numer 3 z zastosowanymi mechanizmami przeciwdziałającymi katastroficznemu zapominaniu. Opracowanie własne	113