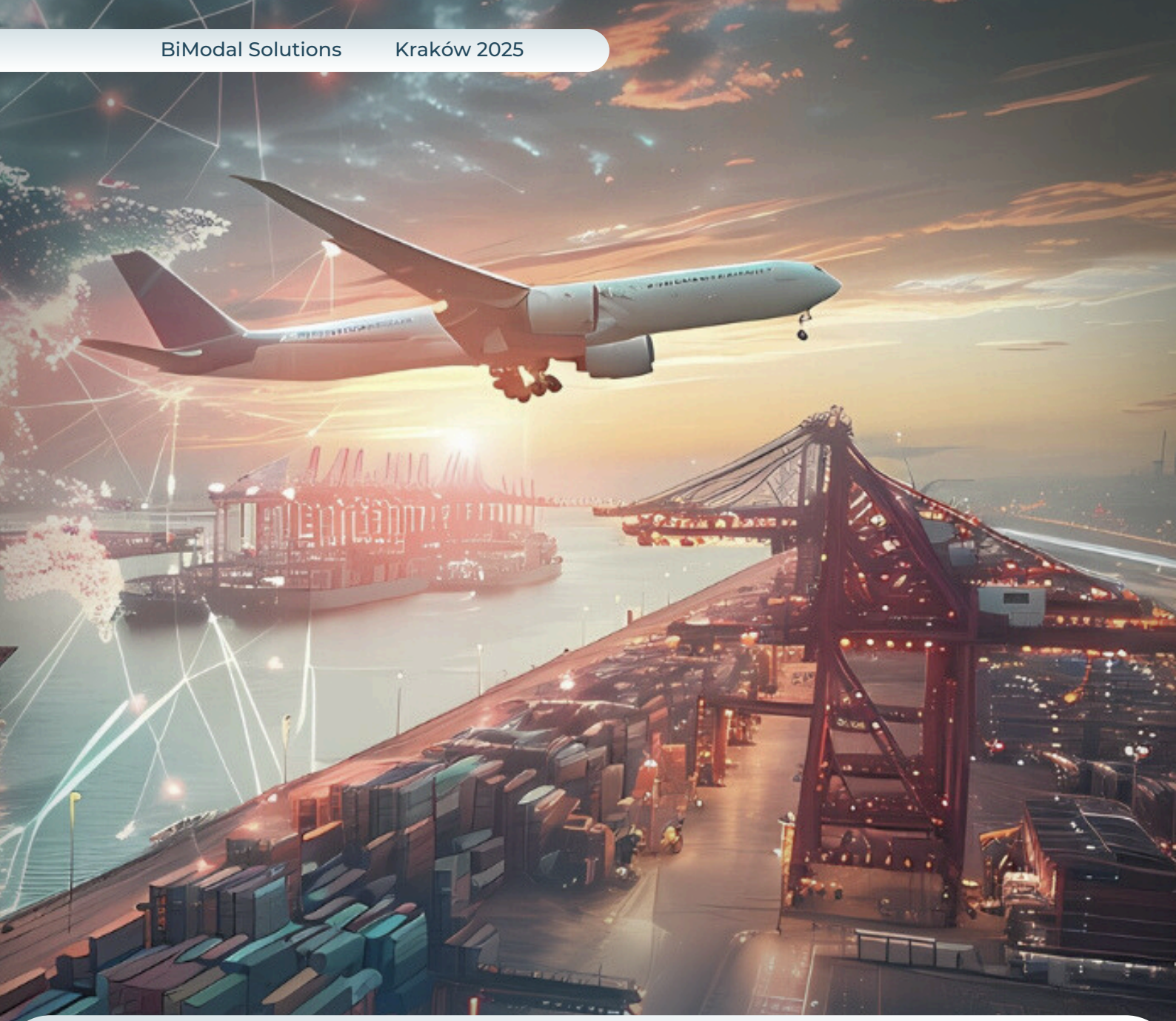


DR MAGDALENA FOLTYN

PROGNOZOWANIE **POPYTU**

WYKORZYSTANIE TECHNOLOGII AI I ML Z KLASYCZNYMI METODAMI STATYSTYCZNYMI W NOWOCZESNYM ŁAŃCUCHU DOSTAW

BiModal SOLUTIONS

1. WPROWADZENIE

DEMAND FORECASTING I SZTUCZNA INTELIGENCJA

W dobie wszechobecnej cyfryzacji każdy słyszał o **sztucznej inteligencji** (ang. ***artificial intelligence***, abb. **AI**). Dla niektórych jest to hasło budzące strach, dla innych— definicja innowacji i rozwoju.

Czym jednak ta sztuczna inteligencja tak naprawdę jest? Odpowiadając na to pytanie jednym zdaniem: to w ogólności zbiór technologii umożliwiający tworzenie rozwiązań naśladujących ludzką inteligencję.

Równie popularnym pojęciem jest **uczenie maszynowe** (ang. ***machine learning***, abb. **ML**). Wbrew powszechnej opinii, nie jest ono tożsame z AI. Prezentuje ono węższy koncept, dotyczący implementacji różnych algorytmów do analizy danych i wnioskowania na tej podstawie.

Przykładowym wnioskiem z analizy danych może być przyporządkowanie psa do danej rasy na podstawie jego wyglądu. Człowiek potrzebuje sporo czasu, by nauczyć się, jak dana rasa wygląda i jak odróżnić jedną od drugiej. Zastosowanie algorytmu uczącego się pozwoli na **odtworzenie tożsamego procesu, który zachodzi u ludzi**.

To, czy algorytm poradzi sobie z takim zadaniem rozpoznawania zależy przede wszystkim od **danych treningowych**. Jeżeli dane treningowe nie są wiarygodne, algorytm również taki nie będzie. W przypadku rozpoznawania ras psów, sprawa jest prosta. Możemy poprosić eksperta, który sprawdzi i sklasyfikuje nasze dane ze zbioru treningowego, a algorytm uczący się na nich będzie wiarygodny.

- Co jednak w przypadku, gdy mówimy o algorytmie uczącym się, którego zadaniem jest predykcja popytu na dany produkt?
- Czy przewidywanie, co się wydarzy w przyszłości może być wiarygodne?
- Jak się wykonuje takie przewidywania dla zmiennych zależnych od czasu (predykowanie szeregu czasowego, ang. *time series forecasting*)?

Odpowiedzi na te i inne pytanie znajdziecie w poniższym artykule.

DEMAND FORECASTING:

DLACZEGO JEST POTRZEBNY W BIZNESIE I PLANOWANIU?

Forecasting, czyli przewidywanie, jest integralną częścią planowania w biznesie. Im precyzyjniej przewidzimy wynik, tym lepiej zaplanujemy działanie naszej firmy.

Skąd jednak się dowiedzieć, co się wydarzy w przyszłości? Możemy **przeanalizować wydarzenia z przeszłości i sprawdzić, czy są one powtarzalne.**

Kwantyfikując takie wydarzenia możemy również uzyskać informację o trendzie danych zjawisk. Oprócz tego, możemy sprawdzić, jaki wpływ na naszą predykowaną zmienną miały inne czynniki.
W niniejszym artykule przedstawiamy **proces od analizy surowych danych** do **predykcji na przykładowym szeregu czasowym**—historii sprzedaży produktu.

WSTĘPNA ANALIZA DANYCH

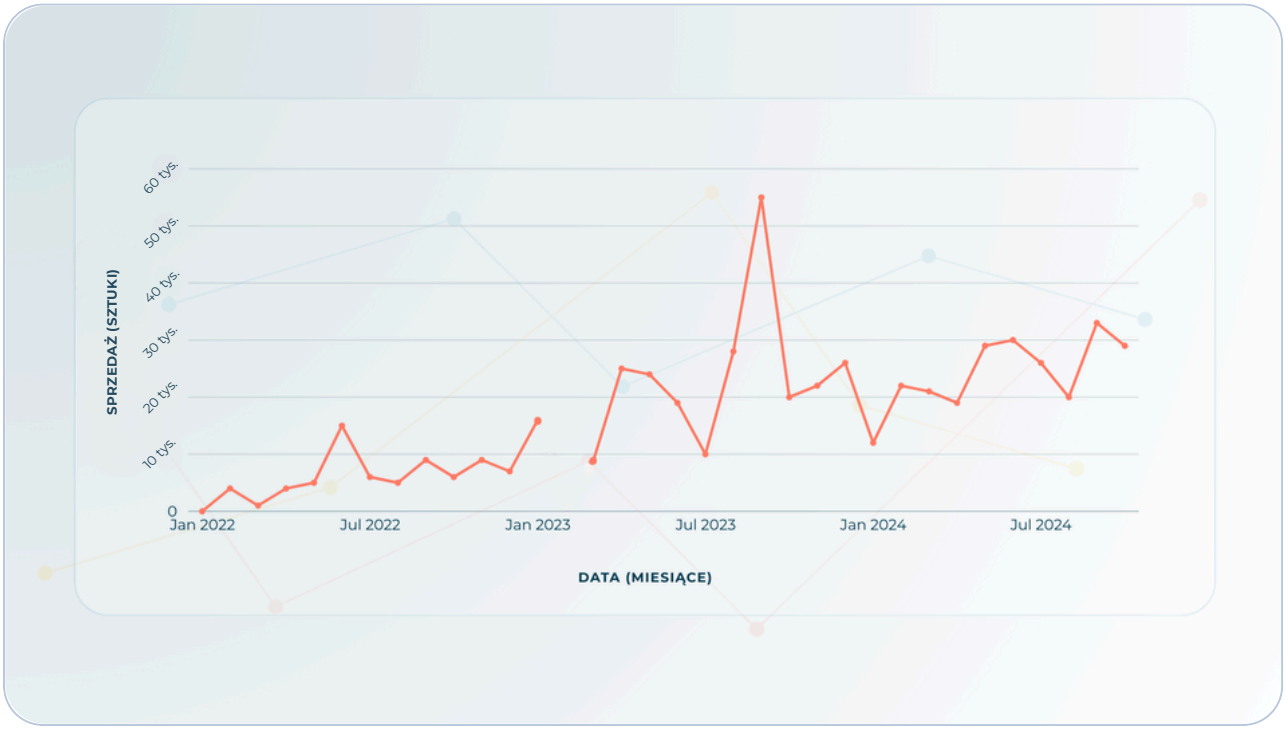
NA JAKICH DANYCH MOŻNA PREDYKOWAĆ? DLACZEGO NALEŻY PRZEANALIZOWAĆ I WYCZYŚCIĆ DANE?

Nie każdy szereg czasowy nadaje się do prognozowania. Dlatego przed wykonaniem predykcji, wpierw należy **ocenić przydatność danych historycznych.**

Czym muszą się one zatem cechować, by można było na nich wykonać predykcję—i co należy z nimi zrobić, aby były one w tym celu przydatne?

Szereg czasowy, na którym omówimy proces forecastingu to przykładowe dane historyczne sprzedaży produktu; załóżmy, że są to rogale z kremem pistacjowym (Rys. 1.).

Dane zostały spreparowane na potrzeby tego artykułu.



Rysunek 1: Historyczna sprzedaż rogali z kremem pistacjowym

ODPOWIEDNI WOLUMEN DANYCH W STOSUNKU DO ICH CZĘSTOTLIWOŚCI

Danych musi być tyle, aby pewne **przewidywane zachowania mogły zostać uchwycone.**

- Przykładowo, w przypadku **danych miesięcznych**, potrzeba danych z przestrzeni **kilku lat**.
- W przypadku zaś danych dziennych— kilku tygodni bądź miesięcy.

Wolumen powinien być również **dostosowany do zmian w historii**. Jeżeli w danych istnieje zauważalna zmiana zachowania, należy wykluczyć te wcześniejsze. Przykładowo, dla danych miesięcznych przyjmuje się minimum rzędu **24 punktów** (2 pełnych okresów rocznych).

Ilość danych wpływa również na to, **jaki okres możemy predykować z dużym prawdopodobieństwem**. Im mniej posiadamy danych, tym mniej okresów możemy przewidzieć w przyszłości.

Dane zostały przedstawione z częstotliwością miesięczna w okresie od stycznia 2022 do grudnia 2024. Mamy więc 36-miesięczny okres, czyli wystarczającą ilość punktów, aby wykonać predykcję.

OCENA PRZEWIDYWALNOŚCI DANYCH (ANG. *FORECASTIBILITY*)

Jeżeli ilość danych jest odpowiednia, następnym krokiem jest ocena przewidywalności danych. Takiej oceny dokonuje się obliczając dwa wskaźniki:

- **Squared Coefficient of Variation** – współczynnik zmienności, mówiący o zmienności danych.

Wzór: $CV^2 = (\sigma/\mu)^{2,}$
gdzie: σ – odchylenie standardowe danych, μ - średnia danych.

- **Average Demand Interval** – średni przedział popytu mówiący o regularności danych.

Wzór: *ADI = x / y,*
gdzie: x – liczba okresów w naszej historii, np. 2 lata, czyli x = 24;
y – liczba obserwacji, czyli zarejestrowanych danych.

Przykładowo, w wybranym okresie dwuletnim mamy 3 braki, więc y = 21.

Mając te dwie zmienne, możemy wstępnie ocenić przewidywalność danych według następującej klasyfikacji:

- **Popyt gładki** (ang. *smooth demand*) – ADI < 1.32, CV^2 < 0.49 – dane bardzo dobrze przewidywalne.
- **Popyt przerywany** (ang. *intermittent demand*) – ADI >= 1.32, CV^2 < 0.49 – dane przewidywalne, ale z brakami.
- **Popyt nierówny** (ang. *erratic demand*) – ADI < 1.32, CV^2 >= 0.49 – dane średnio przewidywalne.
- **Popyt wyboisty** (ang. *lumpy demand*) - ADI >= 1.32, CV^2 >= 0.49 – dane nieprzewidywalne.

Rysunek 2: Wizualna reprezentacja klasyfikacji popytu.
Źródło: Artykuł mojego autorstwa

W danych obserwujemy jeden brak, w lutym 2023. ADI = 1.03, CV^2 = 0.48. Mamy więc do czynienia z popytem gładkim. Oznacza to, że nasze dane są przewidywalne i nadają się do forecastu.

BRAKI DANYCH I ICH UZUPEŁNIENIE

W przypadku popytu przerywanego i nierównego braki danych powinny zostać uzupełnione, jeśli nie jest ich zbyt dużo. Zależy to również od wolumenu danych. Racjonalnym jest akceptować braki **do 10% danych**.

Dużo zależy również od tego, **w jakich miejscach takich danych brakuje**. Jeżeli braki są losowo rozmieszczone na osi czasu (czyli nie za blisko siebie), zawsze można zastosować interpolację do ich uzupełnienia.

Jeżeli występują one obok siebie, tj. w danych jest obecna „dziura", prawdopodobnie trzeba będzie je odrzucić.

W takim wypadku trzeba się też zastanowić, **jakie jest źródło tych braków,** czyli:

- czy sprzedaży nie było w danym czasie (tj. brak jest równy zerowej sprzedaży),
- czy dane nie zostały zebrane.

Inną możliwością jest też predykowanie brakujących punktów.

W lutym 2023 brakuje wartości sprzedaży. Załóżmy, że brak ten wynika z przerwy w dostawie pistacji, sprzedaż zatem wyniosła 0, i tak ten punkt zastępujemy.

ODRZUCENIE PUNKTÓW ODSTAJĄCYCH

Jeżeli w danych nie ma już braków, kolejnym etapem jest odrzucenie **punktów odstających** (ang. ***outliers***). Za punkt odstający uznaje się taki, który znacznie odbiega od ogólnej tendencji danych.

Jednak miejcie na uwadze, że nie każdy punkt odstający od reszty danych jest outlierem. Przykładowo, sprzedaż światełek ledowych jest wyjątkowo wysoka w grudniu; w pozostałych miesiącach jest zdecydowanie niższa.

Dane grudniowe nie są więc odstające, ponieważ wskazują na powtarzalność zjawiska.

KIEDY WARTO TYPOWAĆ PUNKTY ODSTAJĄCE?

Outliery warto typować dla **danych stacjonarnych (zdetrendowanych).** Pozwala to na uchwycenie przypadkowych zmian. Przykładowo, jeśli sprzedaż rogalików z czekoladą z roku na rok rośnie, mamy do czynienia z rosnącym trendem sprzedaży.

Wyobraźmy sobie, że pewien instagramowy celebryta zaczął reklamować rogaliki i produkowane przez obecną w całym kraju sieć piekarni. Zakładając, że ma on duży posłuch wśród fanów, zaczynają oni oblegać piekarnie. W rezultacie sprzedaż wyrobów piekarniczych w tej sieci drastycznie rośnie.

Taka spora zmiana 'nie pasuje' do trendu danych. Jest to **zdarzenie biznesowe**– niepowtarzalne, więc nieprzewidywalne.

PRZEWIDYWALNOŚĆ DANYCH A OUTLIERY

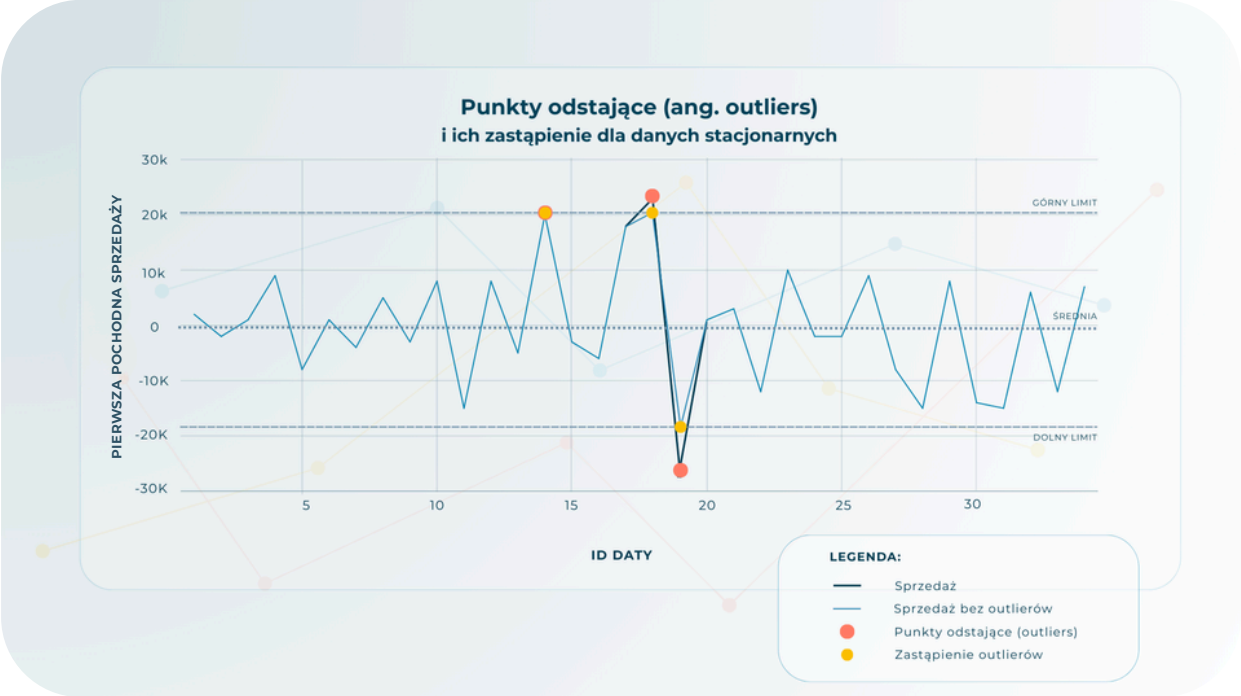
Odrzucenie takiego punktu poprawi **przewidywalność danych** (ang. ***data forecastability***). W przypadku odrzucenia outliera, trzeba taki punkt również zastąpić.
Można się w tym celu posłużyć **interpolacją**, ale można też daną wartość odstającą **zredukować** lub **podnieść do poziomu odpowiadającemu statystykom** granicznych danych.

Jest to dobre podejście w przypadku, gdy outlier leży na samym końcu przedziału danych, gdzie nie da się zastosować interpolacji, a ekstrapolacja jest zbyt losowa.

Można oczywiście taki punkt po prostu **wyrzucić**, ale w przypadku małej ilości danych może się okazać, że nie będą one się nadawać do predykowania.

Krem pistacjowy nie był jeszcze tak popularny w Polsce 3 lata temu. Z wykresu można zaobserwować trend rosnący. Na pierwszy rzut oka nie widać wyraźnej sezonowości. Dane trzeba **zdetrendować**, tj. doprowadzić je do postaci stacjonarnej.

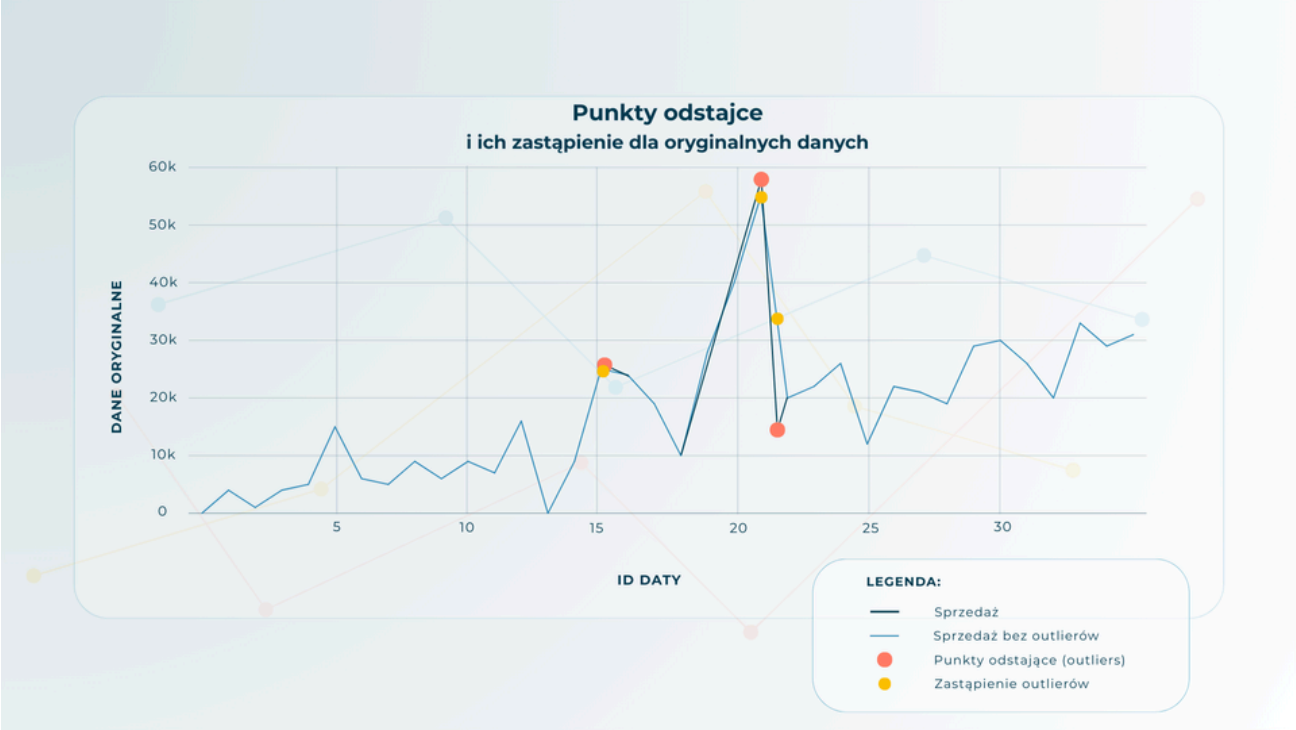
Wyliczamy więc **pierwszą pochodną danych**. Dane traktujemy jak z rozkładu normalnego. Przyjmując, że akceptowalny dla nas rozrzut danych to μ±2σ, wyznaczamy **dolny i górny limit**. Punkty znajdujące się poza limitami uznajemy za outliery.



Rysunek 3. Wyznaczenie outlierów dla danych stacjonarnych

W tym wypadku mamy do czynienia z dwoma takimi punktami (Rys. 3.). Punkty te zastępujemy **wartościami odpowiednich limitów**, a w ostatnim etapie przywracamy ich pierwotną formę (Rys. 4.).

Analizując zaznaczone outliery, możemy stwierdzić, że ta procedura nieznacznie zniwelowała maksima.



Rysunek 4: Wyczyszczone dane

Tak przygotowane dane mogą zostać użyte do predykowania. Proces czyszczenia danych został również opisany w artykule na naszym blogu.

PREDYKCJA POPYTU

PROCES UCZENIA MODELU FORECASTINGOWEGO

Przewidywanie popytu opiera się na **wykorzystaniu konkretnego modelu**. Aby model forecastingowy mógł nam dać informację o tym, co się wydarzy w przyszłości, należy go najpierw **nauczyć, co się działo w przeszłości**— pokazać mu pochodzące z niej dane oraz pozwolić mu je zanalizować, a następnie wyciągnąć wnioski. Modeli możliwych do zastosowania do szeregów czasowych jest dużo. W artykule omawiamy kilka z nich.

W dalszej części artykułu wykorzystam 4 przykładowe modele: **ARIMA, Exponential Smoothing, XGBoost i LGBM.**

ZASTOSOWANIE ALGORYTMÓW AI

DO PREDYKCJI POPYTU

ETAP 1: UCZENIE MODELU

Najważniejszą kwestią jest podział zbioru danych na zbiór treningowy oraz testowy. Zazwyczaj szereg czasowy dzieli się w stosunku 7:3:

- 70% danych, zaczynając od najstarszych, to **zbiór treningowy**,
- 30%, czyli dane najmłodsze, to **zbiór testowy**.

Uczenie modelu to de facto **dopasowanie jego parametrów** (czyli w zasadzie dopasowanie funkcji matematycznej) **do zbioru treningowego**. By sprawdzić, czy model się dobrze nauczył, sprawdzamy go na zbiorze testowym.

Polega to na tym, że mając dopasowaną funkcję modelu, **wyznaczamy wartości tej funkcji dla okresów zbioru testowego**. W tym wypadku model już nie widzi wartości popytu; teraz pokazuje on to, czego się nauczył.

Tak wyznaczoną prognozę porównujemy z danymi testowymi i oceniamy, jak model sobie poradził. Zbiory treningowy i testowy **nie mogą być za małe**.

Podział musi być taki, żeby model miał się na czym nauczyć, a potem mieć odpowiednio dużą próbkę, na której go sprawdzamy.

Do przykładowych danych sprzedaży rogali z nadzieniem pistacjowym po ich wyczyszczeniu zastosowaliśmy modele: ARIMA, Exponential Smoothing, XgBoost, LGBM. Parametry tych modeli zostały dobrane z odpowiednio szerokich przedziałów. Obliczenia wykonaliśmy dla następująco zdefiniowanych wzorów: treningowy - 25 punktów, testowy – 11 punktów.

ETAP 2: WYLICZENIE BŁĘDU PROGNOSTYCZNEGO

Aby ocenić jakość i trafność takiej prognozy, najczęściej wylicza się tzw. **błędy prognozy**. Może to być zwykła średnia z wartości bezwzględnych różnic miedzy wartościami prognozy, a wartościami rzeczywistymi (ang. *mean absolute error, MAE*). Można też wyznaczyć jego wartość procentową w stosunku do danej rzeczywistej (ang. *mean absolute percentage error*). Wiecej na ten temat w artykule dostępnym na naszym blogu.

Do oceny naszej prognozy użyję **SMAPE**, czyli symetrycznego średniego procentowego błędu bezwzględnego (ang. *symmetric mean absolute percentage error*), w konwencji 0-100:

$$\frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - x_i|}{(|x_i| + |y_i|)}$$

gdzie:
n – liczba okresów prognozy,
x_i– wartość prawdziwa,
y_i– wartość prognozy.

Analizujac ten wzór, można zauważyć, że różnica jest odniesiona do średniej z wartości rzeczywistej i predykowanej—stąd w nazwie słowo „symetryczny".

Model wybieramy **na podstawie najmniejszej wartości SMAPE**. Tabela 1. zawiera podsumowanie wartości SMAPE obliczonych dla poszczególnych modeli na zbiorze testowym.

Model	SMAPE [%, 0-100%]
ARIMA	21.7
Exponential Smoothing	15.9
XGBoost	18.2
LSTM	16.0

Tabela 1: *. SMAPE dla modeli zastosowanych na szeregu czasowym sprzedaży rogali z kremem pistacjowym.*

Zgodnie z tabelą, najlepszy forecast otrzymaliśmy dla modelu Exponential Smoothing (Rys. 5.). Wybrany model uwzględnia komponenty trendu, jak też sezonowości (ang. *Holt-Winters method*).

Modele XgBoost i LGBM osiągnęły porównywalne wyniki. ARIMA wykazuje najwyższy wynik błędu.

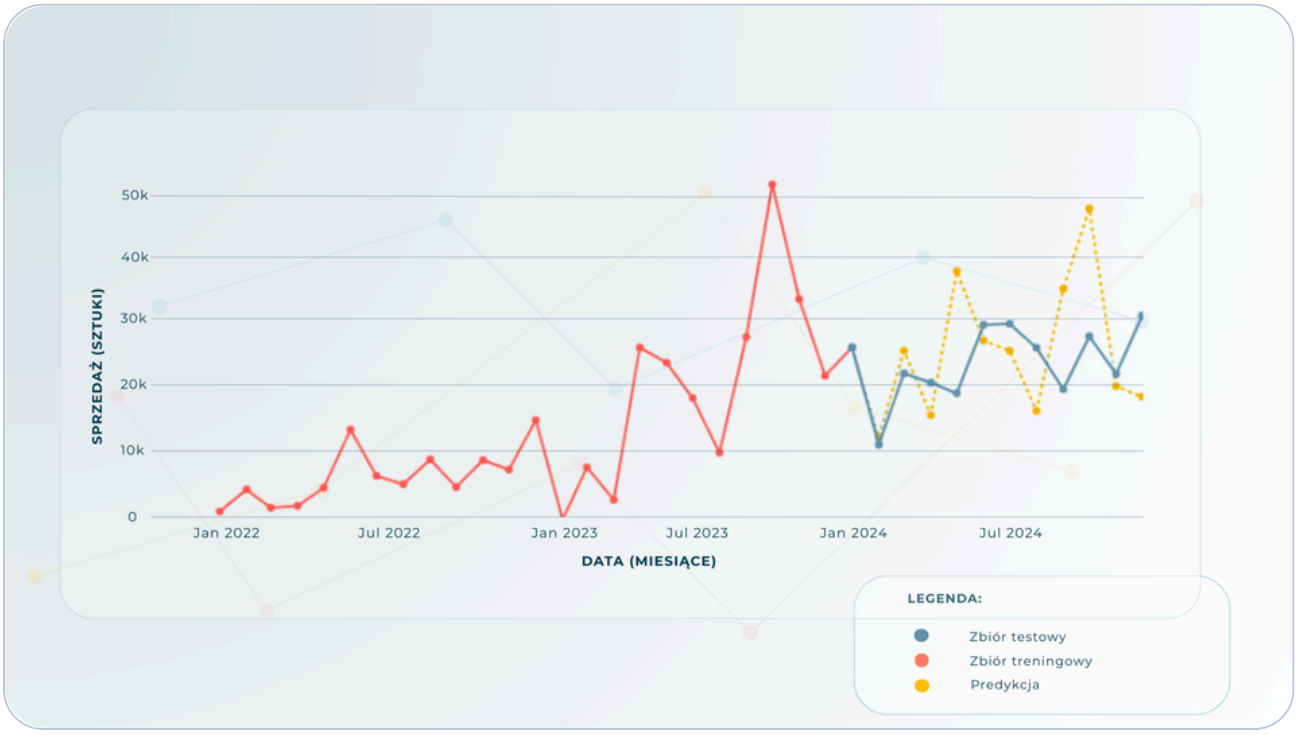
Mając wybrany model, dokonujemy właściwej predykcji z jego użyciem.

ETAP 3: UZYSKANIE WŁAŚCIWEJ PROGNOZY

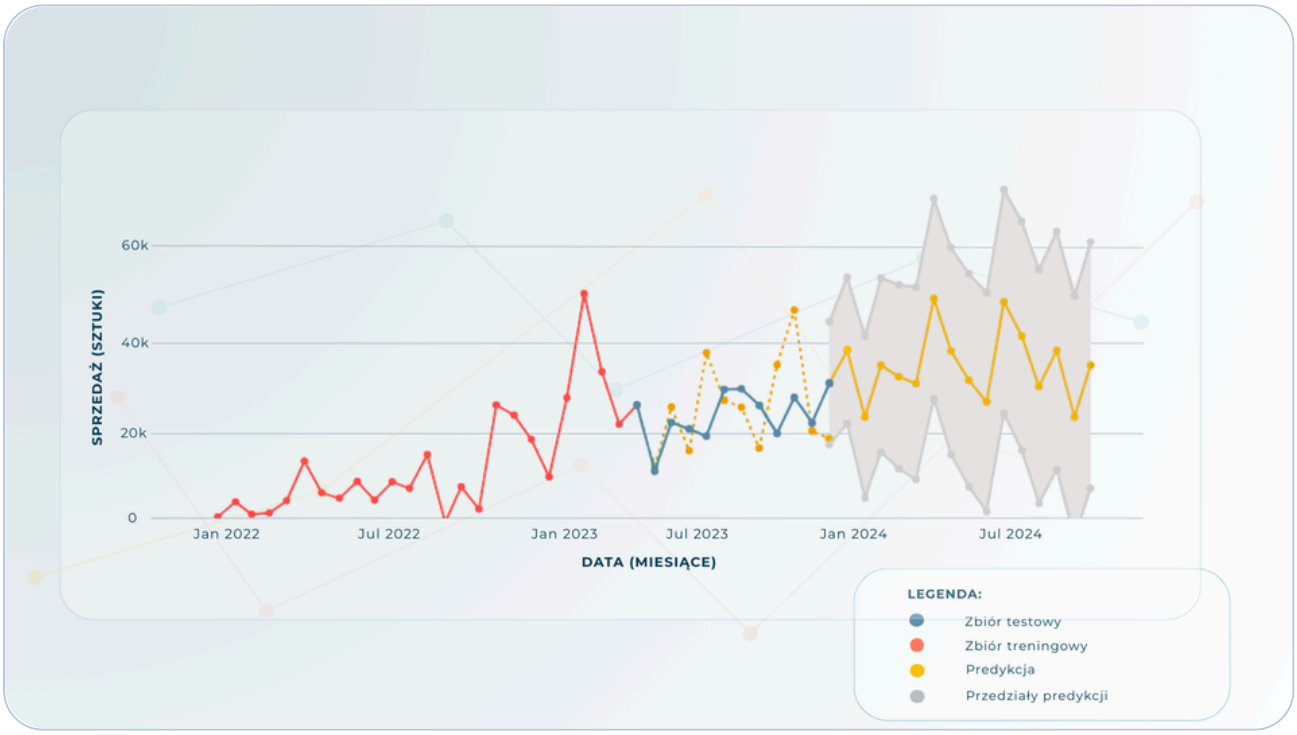
Mamy predykcję testową, czyli opartą na danych z przeszłości. Z tego względu dla celów biznesowych jest ona bezużyteczna.

Aby uzyskać właściwa prognozę, dopasowujemy wyznaczony model do okresów w tzw. **horyzoncie prognozy**—czyli liczby przyszłych okresów, na które prognozujemy.

Predykcji dokonaliśmy na 15 miesięcy (Rys. 6)



Rysunek 5. Prognoza dla zbioru testowego – wygładzanie wykładnicze.



Rysunek 6. Właściwa prognoza wraz z przedziałami predykcji.

ETAP 4: WYLICZENIE BŁĘDU PROGNOSTYCZNEGO

Mamy już wyznaczona prognozę, ale czy możemy być jej pewni? Do takiej oceny służą **przedziały predykcji** (ang. *prediction intervals*), czyli przedziały ufności prognozy. Prognozę, która je zawiera określa się **probabilistyczną** (ang. *probabilistic forecast*), w odróżnieniu od **deterministycznej**, w którym podaje się samą prognozę.
Z punktu widzenia zarówno statystyki, jak i biznesu, forecast deterministyczny jest mało przydatny. Zawsze powinno się podawać niepewność liczonej wartości, aby móc w stanie ocenić jej jakość.

Przedziały predykcji zazwyczaj wyznacza się na poziomach 95% i 80%. Oznacza to, że z określonym prawdopodobieństwem (np. 95%) wartość rzeczywista znajdzie się w tym zakresie.

Jak jednak oblicza się taki margines niepewności? Należy wyznaczyć **rozkład błędów prognozy** (ang. *residuals*), czyli różnic między wartościami rzeczywistymi a prognozowanymi) i policzyć jego charakterystyki. Zwyczajowo przyjmuje się **rozkład normalny**, dla którego liczymy średnią i odchylenie standardowe.

Przedział predykcji zwiększa się wraz z horyzontem prognozy, tj. **im dalej przewidujemy, tym szersze są wyznaczane przedziały**. Wzrost szerokości przedziału ufności wyznaczamy zaś, używając np. prostych metod porównawczych.

Po więcej informacji na ten temat odsyłam do:
https://otexts.com/fpp2/prediction-intervals.html

Przedziały predykcyjne zostały wyliczone na poziomie 80%. Oddaja one kształt prognozy i rozszerzaja sie z czasem. (Rys 6)

PROGNOZA POJEDYNCZEGO PRODUKTU

CZY ZAWSZE JEST SŁUSZNA?

W niniejszym artykule omówiliśmy podstawowy scenariusz prognozowania na bazie historycznych danych sprzedaży pojedynczego produktu. Innymi słowy, dokonaliśmy **predykcji jednowymiarowej** (ang. *univariate forecasting*).
Ale czy jest to zawsze słuszny scenariusz postępowania? Nie zawsze—jeśli historia sprzedaży ma bardzo losowy przebieg, to prognoza na takich danych nie będzie stanowiła wartości dodanej dla biznesu.

Jak można to poprawić? Pierwszą możliwością jest uwzględnienie tzw. **zmiennych objaśniających** (ang. *exogenous variables*) i wykonanie **predykcji wielowymiarowej** (ang. *multivariate forecasting*).

W tej metodzie prognozujemy jednocześnie popyt (zmienna główna) oraz te z dodatkowych zmiennych, które mają na niego wpływ.

WYZNACZENIE ZMIENNYCH OBJAŚNIAJĄCYCH

Ich właśnie dotyczy kolejny problem– bowiem **które z tych zmiennych wpływają na sprzedaż?** Tutaj potrzebna jest przede wszystkim **wiedza biznesowa**, aby przynajmniej zawęzić obszar poszukiwań tych zmiennych objaśniających. Przykładowo, na sprzedaż płytek łazienkowych wpływ powinna mieć sprzedaż nieruchomości lub cena metra kwadratowego mieszkania.

Aby się upewnić, że dana zmienna ma faktycznie wpływ na prognozowany popyt, można policzyć **korelacje między zmienną główną a zmiennymi objaśniającymi**.

Jeżeli występuje korelacja pozytywna lub negatywna, możemy taką zmienna uwzględnić w **forecastingu wielowymiarowym**.

Jego ocenę przeprowadzamy w sposób analogiczny do tej dla pojedynczej zmiennej. Dodatkowo możemy wykonać **analizę wyjaśniającą model** (ang. *model explanation*), w której oblicza się, na ile zmienne wpływają na prognozę.

WYKORZYSTANIE FORECASTU HIERARCHICZNEGO

Może być też jednak tak, że uwzględnienie zmiennych objaśniających nic nie zmienia; w rezultacie, **prognoza nadal nie jest miarodajna**. W takim wypadku warto się zastanowić, czy nie lepiej wykonać prognozę na danych zagregowanych.

Taka agregacja może zostać wykonana na różnych poziomach biznesowych. Dlatego też metodę tę nazywamy **prognozą hierarchiczną** (ang. *hierarchical forecasting*). Suma danych sprzedaży z jednej grupy produktowej może być mniej losowa, niż sprzedaż pojedynczych produktów, więc prognoza na takich danych może okazać się zdecydowanie lepsza.

Ważna jest jednak późniejsza **poprawna dystrybucja takiej prognozy na pojedyncze produkty**. Istnieje wiele metod uzgadniania prognoz hierarchicznych (ang. *forecasts reconciliation*); generalnie, należy wybierać taką, która da najmniejszy błąd prognozy na wybranym poziomie.

PODSUMOWANIE

Przewidywanie popytu jest złożonym procesem, którego poprawne wykonanie pomoże w planowaniu produkcji. Można wykorzystać dedykowane do tego narzędzie, takie jak **BiModal Forecasting—** mogące działać niezależnie, bądź być częścią platformy **BiModal Planning**. BiModal Forecasting jest narzędziem automatyzującym proces prognozowania, które pozwoli oszczędzić czas i dostarczy wiarygodne przewidywania.

Are you interested in our services and planning solutions?

Find out more!

Request a demo

https://www.bimodal-solutions.com/

BiModal Solutions
ul. Wielicka 20, 30-552 Kraków
artur.bula@bimodal-solutions.com