Rozdział 3

Wskaźniki ryzyka i ich predykcja w modelu audytu ciągłego

Jerzy Krawczuk, Izabela Kartowicz-Stolarska, Marek Tabędzki, Piotr Hońko, Marek Krętowski*

Streszczenie. Audyt ciągły jest nowoczesnym sposobem kontrolowania polegającym na stałym monitorowaniu i ocenie procesów oraz operacji w organizacji. Dzięki wykorzystaniu technologii informatycznych audyt ciągły umożliwia bieżące identyfikowanie i analizowanie nieprawidłowości, co pozwala na szybszą reakcję i podejmowanie działań naprawczych. w kontekście projektów bankowych audyt ciągły jest szczególnie ważny ze względu na dynamicznie zmieniające się normy prawne oraz wysokie wymagania dotyczące bezpieczeństwa i zgodności. Kluczowe wskaźniki ryzyka są wykorzystywane do monitorowania zagrożeń, a ich regularna analiza pozwala na wczesne wykrywanie problemów. Referat opisuje narzędzie do oceny ryzyka i prognozowania kluczowych wskaźników ryzyka opracowane w ramach współpracy Wydziału Informatyki Politechniki Białostockiej z jednym z wiodących polskich banków. Przedstawiono w nim elementy architektury zastosowanego rozwiązania, a także metody statystyczne i metody uczenia maszynowego oraz przykładowe wyniki badań.

Słowa kluczowe: audyt ciągły, wskaźniki ryzyka, szeregi czasowe, predykcja

Wprowadzenie

Audyt ciągły (Radziwon, 2013) jest nowoczesnym sposobem kontrolowania polegającym na stałym monitorowaniu i ocenie procesów oraz operacji w organizacji. Dzięki wykorzystaniu technologii informatycznych audyt ciągły umożliwia bieżące identyfikowanie i analizowanie nieprawidłowości, co pozwala na szybszą reakcję i podejmowanie działań naprawczych. W kontekście projektów bankowych audyt ciągły jest szczególnie ważny ze względu na dynamicznie zmieniające się normy prawne oraz wysokie wymagania dotyczące bezpieczeństwa i zgodności operacyjnej.

* Wydział Informatyki, Politechnika Białostocka, Wiejska 45A, 15-351 Białystok, *j.krawczuk@pb.edu.pl, i.stolarska@pb.edu.pl, m.tabedzki@pb.edu.pl, p.honko@pb.edu.pl, m.kretowski@pb.edu.pl*
DOI 10.24427/978-83-68673-08-1_3

Metryki, które pomagają w monitorowaniu potencjalnych zagrożeń i ocenianiu ich wpływu na działalność banku nazywane są kluczowymi wskaźnikami ryzyka (ang. *key risk indicators*, KRI). Przykłady takich wskaźników obejmują: liczbę nieautoryzowanych transakcji, czas przestoju systemów informatycznych, liczbę incydentów bezpieczeństwa oraz wskaźnik zgodności z regulacjami. Regularne monitorowanie KRI pozwala bankom na wcześniejsze wykrywanie problemów i podejmowanie odpowiednich działań prewencyjnych, co jest kluczowe dla utrzymania stabilności i zaufania klientów (Ishtiaq, 2015).

W niniejszym referacie zostaną przedstawione szczegóły realizacji narzędzia wykorzystującego metody statystyczne i uczenia maszynowego do oceny ryzyka i prognozowania przyszłych wartości KRI. Wspomniane narzędzie zostało opracowane przez zespół badawczy Wydziału Informatyki Politechniki Białostockiej na potrzeby projektu *Model audytu oparty na metodyce ciągłej oceny ryzyka i mechanizmów kontrolnych z wykorzystaniem metod statystycznych i sztucznej inteligencji* realizowanego w ramach współpracy z jednym z wiodących polskich banków. W pierwszej części rozdziału opisano wskaźniki ryzyka KRI i metody ich pomiaru oraz kluczowe procesy biznesowe. Następna sekcja opisuje metody statystyczne i metody uczenia maszynowego wykorzystywane przy predykcji wartości KRI oraz metody walidacji modeli. Następnie przedstawiono przykładowe wyniki na rzeczywistych i sztucznych danych. Ostatnia sekcja opisuje zrealizowane narzędzia oraz elementy architektury stworzonego specjalnie dla banku rozwiązania.

3.1. Wskaźniki ryzyka KRI

Kluczowe wskaźniki ryzyka (KRI) stanowią zestawy parametrów procesów biznesowych, które z dużym prawdopodobieństwem odzwierciedlają zmiany poziomu ryzyka. W ramach projektu opracowano ponad 100 wskaźników KRI obejmujących najważniejsze obszary i procesy biznesowe, które prowadzą do osiągnięcia zamierzonego celu i kreują wartość dodaną dla banku lub jego klientów. Do objęcia audytem ciągłym wybrano kilkadziesiąt procesów dotyczących różnych obszarów funkcjonowania banku, które są kluczowe dla jego działalności, w tym bankowości detalicznej, korporacyjnej i inwestycyjnej oraz procesów wsparcia, IT i bezpieczeństwa.

Wypracowane w projekcie wskaźniki KRI mają na celu monitorowanie poziomu ryzyka (Ziora, 2011) w poszczególnych procesach banku i są wykorzystywane do:

- identyfikacji procesów i zagadnień do objęcia badaniami audytowymi,
- wsparcia procesu przygotowania strategicznych i rocznych planów audytów,
- wsparcia procesu wyboru prób audytowych lub testowania umów i operacji na całych populacjach.

Podstawowy typ wskaźników KRI stanowi iloraz dwóch liczb, gdzie licznik to umowy lub operacje, w przypadku których wystąpiły niepożądane zdarzenia bądź wzrosło prawdopodobieństwo ich wystąpienia (np. straty, potencjalne straty, opóźnienia, incydenty, awarie, reklamacje) natomiast mianownik to wszystkie umowy lub operacje.

Wskaźniki KRI mogą mieć charakter ilościowy lub/i wartościowy. w przypadku, gdy istnieje taka możliwość, wskaźniki wyliczane są z zastosowaniem obu tych sposobów. Przykładem może być wskaźnik wypowiedzeń umów kredytowych wyliczany jako liczba umów wypowiedzianych w stosunku do wszystkich umów oraz wartość niespłaconego kapitału w ww. umowach w stosunku do wartości wszystkich kredytów.

Warto wspomnieć, że wskaźniki realizowane w ramach projektu wymagały przygotowania skryptów, które – po integracji z systemem orkiestracji zadań używanym w banku – agregują i aktualizują dane pochodzące z wielu tabel źródłowych do postaci tabel wynikowych. Początkowo przetwarzanie obejmowało pełne zbiory danych z wielomiesięcznych okresów historycznych. W kolejnych iteracjach wdrożono podejście inkrementalne, w którym przetwarzane są wyłącznie nowe lub zmodyfikowane rekordy. Zastosowane podejście umożliwia bieżącą aktualizację agregatów przy znacząco zredukowanym wolumenie przetwarzanych danych, co przekłada się na efektywność i skalowalność rozwiązania.

3.1.1. Metody pomiaru wskaźników KRI

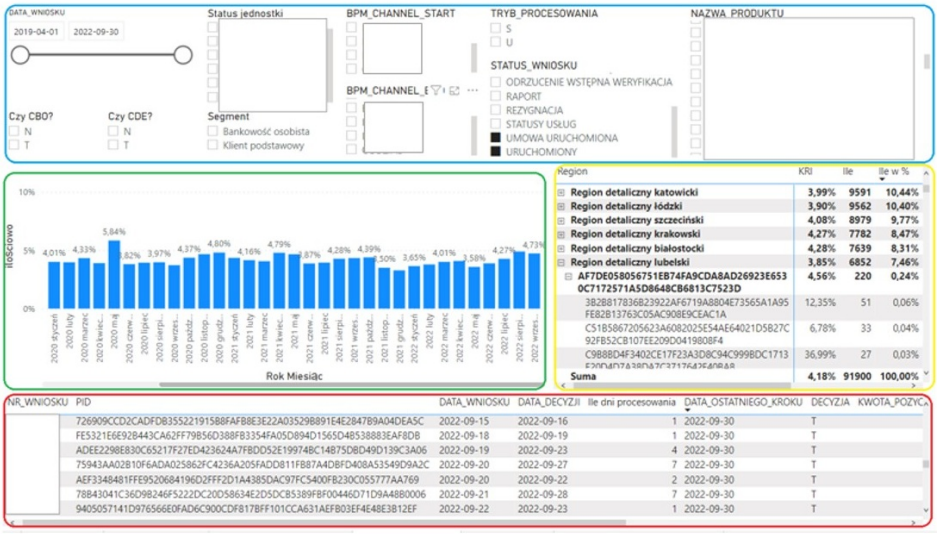
Na podstawie wypracowanych przez zespół badawczy i zespół banku założeń przygotowano opisy dla wskaźników KRI według określonego schematu:

1. Ogólny opis wskaźnika KRI.
2. Opis konstrukcji wskaźnika KRI.
3. Opis filtrów dla wskaźnika KRI (fragmentatorów, funkcji *drill-down* oraz rankingów).

Głównym typem wskaźników KRI jest udział umów lub operacji, w przypadku których wystąpiły niepożądane zdarzenia lub istotne odchylenia (np. straty, opóźnienia). Opis konstrukcji wskaźników KRI zawiera wzór oraz opis danych do ich wyliczania, a opis filtrów – szczegóły związane z zawężeniem danych do określonych przedziałów lub kategorii. Ogólnym założeniem dotyczącym metod pomiaru wszystkich wskaźników była możliwość zawężania zestawu danych na podstawie zdefiniowanych przez pracownika banku filtrów, a także sortowanie i klasyfikowanie danych w celu wizualnego przedstawienia ich w określonej kolejności. z tego powodu zdecydowano o zastosowaniu narzędzia do analizy biznesowej Power BI, które umożliwia przekształcanie danych w interaktywne wizualizacje i raporty (Skender i Manevska, 2022).

3.1.2. Narzędzia wizualizacji i analizy biznesowej

W ramach projektu zrealizowano kilkaset interaktywnych raportów do analizy danych, których źródłem są tabele wynikowe w bazie danych lub kostka analityczna. Raporty prezentują dane w postaci różnorodnych wizualizacji, tj. wykresów, tabel drill-down i rankingów. Przykładowy układ raportu zrealizowanego w projekcie został zaprezentowany na Rysunku 3.1.



Rysunek 3.1: Realizacja przykładowego wskaźnika KRI w narzędziu Power BI

Główną część wizualizacji raportów stanowią wykresy (ramka zielona) prezentujące poziom wskaźników KRI w poszczególnych okresach. Na prawo od wykresów w tabeli przedstawiono szczegółowe wartości KRI dla poszczególnych regionów, jednostek lub pracowników – tzw. rankingi (ramka żółta, w tym przypadku trzypoziomowy ranking kaskadowy). Powyżej wykresów znajdują się tzw. fragmentatory (ramka niebieska) pozwalające na odfiltrowanie analizowanych danych, na przykład według daty zawarcia umów, segmentów klienta, kanałów sprzedaży czy rodzajów produktów. Na dole raportu zaprezentowano szczegółowe dane dotyczące umów lub operacji odbiegających od oczekiwań, tzw. funkcja *drill-down* (ramka czerwona).

Wskaźniki KRI wyliczane są za pomocą reguł DAX (ang. *data analysis expressions*) (Campante, Gonçalves i Gonçalves, 2023), a ich historyczny trend prezentowany jest w postaci wykresów.

Wykresy stanowią podstawowy sposób prezentacji poziomu wskaźnika KRI w poszczególnych obserwowanych okresach, najczęściej miesięcznych.

Na potrzeby prezentacji wybrano wykresy słupkowe jako podstawowe, rzadziej wykorzystywane są wykresy liniowe.

Użytkownicy mogą wchodzić w interakcje z raportami poprzez fragmentatory, która zapewniają im prosty sposób filtrowania danych. Zawężają one zestaw danych, który jest wyświetlany w wizualizacji wskaźników KRI, jak również w rankingach i tabelach *drill-down*.

Fragmentatory umożliwiają analizę KRI według różnych kryteriów wyboru, na przykład z podziałem na segmenty klientów, kanały sprzedaży, grupy i rodzaje produktów, tryby procesowania umów oraz operacji czy skalę odchyleń od oczekiwań.

Rankingi (Skender i Manevska, 2022) prezentują wskaźniki KRI przedstawione na wykresie np. w rozbiciu na jednostki organizacyjne banku. Struktura rankingu ma charakter hierarchiczny, co pozwala na sortowanie wartości wskaźników KRI według regionów, następnie jednostek (oddziałów, centrów, departamentów) aż do poziomu doradców klienta, rodzajów operacji, przyczyn niepożądanych zdarzeń (w zależności od wskaźnika oraz zgodnie z zapotrzebowaniem).

Tabele *drill-down* (Lizotte-Latendresse i Beauregard, 2018) stanowią zestawienie szczegółowych danych o umowach lub operacjach odbiegających od oczekiwań, w tym danych dotyczących skali niepożądanych zdarzeń (np. wartości niespłaconych rat kredytu), numeru umowy lub klienta, wartości, daty i trybu zawarcia umowy, jednostki lub doradcy zawierającego umowę i wszelkich innych parametrów opisujących dany przypadek.

Wskaźniki KRI wyliczane są w okresach sprawozdawczych wynoszących zwykle miesiąc. Powoduje to otrzymanie klasycznego szeregu czasowego o stałym interwale oddzielającym kolejne obserwacje, bez braków w danych. W zależności od obszaru działalności biznesowej wskaźniki mogą być wyliczone dla różnej długości historii, zwykle wynoszącej kilka lat, co daje w efekcie szereg czasowy o kilkudziesięciu obserwacjach, który może posłużyć do budowy modeli predykcyjnych na kolejne okresy sprawozdawcze.

3.2. Predykcja KRI

3.2.1. Modele szeregów czasowych

Zastosowane w projekcie modele predykcyjne obejmowały zarówno klasyczne podejścia statystyczne (np. ARIMA), jak i bardziej zaawansowane metody bazujące na drzewach decyzyjnych. Poniżej przedstawiono krótki opis modeli zaimplementowanych i wdrożonych w trakcie projektu.

Modele statystyczne:

- **ARIMA** – w przypadku modeli ARIMA (ang. *auto regressive integrated moving average*) (Shumway i Stoffer, 2017) istotny jest dobór ich struktury, na przykład ustalenie stopnia autoregresji. w tym celu porównywano różne modele ARIMA za pomocą takich miar, jak AIC (ang. *akaike information criterion*) oraz BIC (ang. *Bayesian information criterion*). Miary te biorą pod uwagę zarówno dopasowanie modelu do danych odzwierciedlone w funkcji największej wiarygodności, jak i stopień skomplikowania modelu opisany liczbą parametrów potrzebnych do jego estymacji.
- **Holt-Winters' Exponential Smoothing** – Holt-Winters (Chatfield, 1978) jedna z najpopularniejszych technik prognozowania szeregów czasowych. Mimo że opracowana została dziesiątki lat temu, wciąż jest obecna w wielu aplikacjach. Holt-Winters to sposób modelowania trzech aspektów szeregów czasowych: typowej wartości (średniej), nachylenia (trendu) w czasie i cyklicznego powtarzania się wzorca (sezonowość).
- **Theta** – metoda ta (Assimakopoulos i Nikolopoulos, 2000) sprawdziła się szczególnie dobrze w wyzwaniu M3-competition i dlatego jest przedmiotem zainteresowania osób zajmujących się prognozowaniem. Jest to skuteczna w wielu zastosowaniach metoda oparta na wygładzaniu wykładniczym z dryfem.

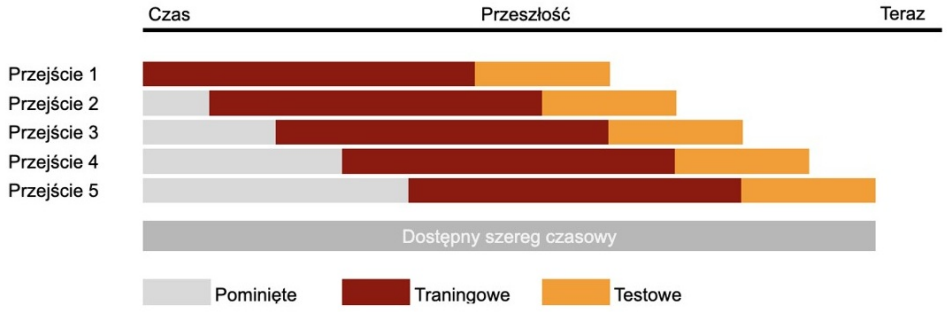
Jako alternatywę dla tradycyjnych metod regresji wykorzystano modele uczenia maszynowego bazujące na drzewach decyzyjnych. w ramach zadania przetestowano następujące modele:

- **LightGBM** (ang. *light gradient boosting machine*) (Ke i in., 2017) – algorytm oparty na drzewach decyzyjnych, używający wzmocnienia gradientowego.
- **Random forest** -– las losowy (Breiman, 2001), komitet drzew decyzyjnych budowanych w sposób quasi-losowy, gdzie na każdym etapie konstrukcji drzewa losowo wybierany jest ograniczony podzbiór cech, które mogą być rozważane jako kryteria podziału aktualnej partycji. Dzięki zastosowaniu tego mechanizmu las losowy charakteryzuje się dużą odpornością na przeuczenie, wysoką precyzją predykcji oraz stosunkowo dobrą wyjaśnialnością.

3.2.2. Walidacja modeli

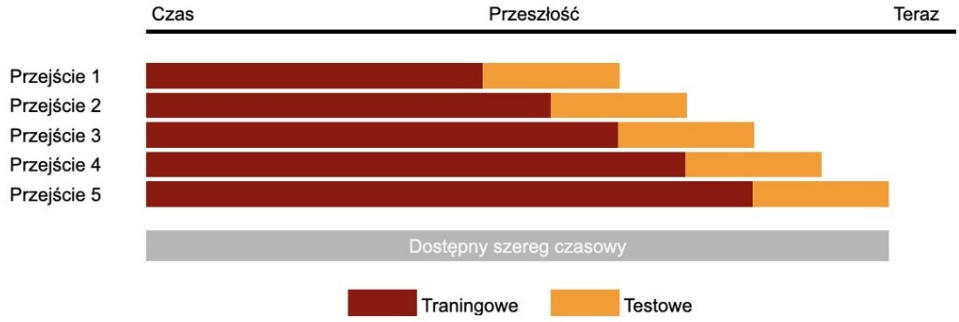
W przypadku szeregów czasowych nie można zastosować standardowej walidacji krzyżowej, gdyż groziłoby to powstaniem tzw. przecieku informacji (predykcja przeszłych wartości wskaźników na podstawie wartości występujących w przyszłości w stosunku do punktu czasowego predykcji). Opracowana ostatecznie koncepcja walidacji modeli opiera się na zastosowaniu okna przesuwnego (ang. *sliding window*) (Taieb, Bontempi, Atiya i Sorjamaa, 2012), które pozwala sprowadzić tę sytuację do problemu uczenia maszynowego nadzorowanego. Predykcja odbywa się zawsze na

okres t+1, z wykorzystaniem danych z okresu [t-T, t], gdzie T oznacza liczbę okresów (miesięcy), na podstawie których budowany jest model predykcyjny (Rysunek 3.2).



Rysunek 3.2: Konstrukcja zbioru treningowego przy użyciu okien czasowych o stałej szerokości

W projekcie zdecydowano się również na budowę modeli zawsze ze wszystkich dostępnych wstecznie danych tj. [0,t] (ang. *expanding window*), gdzie okres danych uczących zawsze stanowi wielokrotność miesiąca (Rysunek 3.3) .

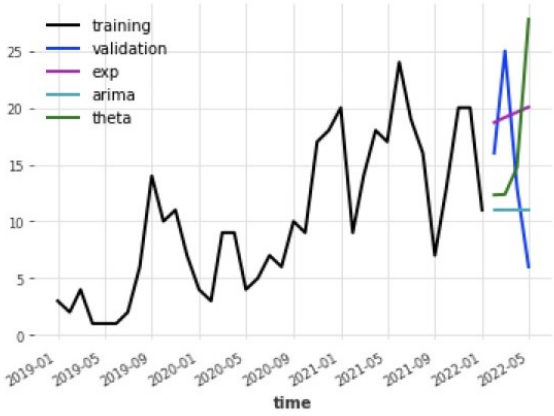


Rysunek 3.3: Konstrukcja zbioru treningowego przy użyciu rosnących okien czasowych

Ze względu na niską precyzję wykrywania miesięcznych fluktuacji w celu predykcji na okres dłuższy niż miesiąc porównywano średnią (lub sumę – tylko dla licznika) wartość wskaźnika KRI ze średnią (lub sumą) wartością predykcji.

3.3. Uśrednianie

Technika uśredniania została zastosowana dwukrotnie. Po pierwsze, w przypadku prognozowania wartości KRI na więcej niż jeden okres (np. cztery kolejne miesiące) istotniejsze od przewidywania wartości wskaźnika w poszczególnych miesiącach okazuje się oszacowanie jego średniego poziomu w całym prognozowanym przedziale czasowym. Wynika to z faktu, że badany wskaźnik odzwierciedla liczbę zdarzeń niepożądanych, a zatem dla celów decyzyjnych wystarczająca może być informacja o łącznej lub przeciętnej liczbie takich zdarzeń w danym okresie, bez konieczności dokładnego przypisywania ich do konkretnych miesięcy.



Rysunek 3.4: Predykcja przykładowego KRI trzema modelami statystycznymi na cztery miesiące do przodu

Na Rysunku 3.4 pokazano predykcje trzema modelami statystycznymi na cztery miesiące do przodu. Wybrany KRI przedstawia liczbę zdarzeń niepożądanych. Można zaobserwować, że żaden z modeli nie wykrył początkowego wzrostu i późniejszego spadku liczby zdarzeń niepożądanych mierzonych za pomocą badanego KRI. Natomiast po zsumowaniu wyników prognoz z czterech miesięcy otrzymano całkowitą wartość oczekiwanej liczby zdarzeń niepożądanych (tabela 3.1). Prawdziwa liczba zdarzeń niepożądanych wynosiła 60. Porównując wyniki, można zaobserwować, że predykcja modelem Theta okazała się najdokładniejsza.

Drugim zastosowanym podejściem uśredniającym było uśrednianie wyników predykcji uzyskanych z wielu modeli. w analizowanym przypadku średnia wartość predykcji wyniosła 62,90, co stanowiło wartość zbliżoną do rzeczywistej. Uśrednianie prognoz pochodzących z różnych modeli często prowadziło do poprawy jakości predykcji, co stanowiło przesłankę do dalszego badania tzw. komitetów modeli. Początkowo zakładano, że analiza dostępnych algorytmów predykcyjnych dla wskaź-

Tabela 3.1: Prognozowana liczba zdarzeń niepożądanych w ciągu czterech miesięcy (bez podziału na poszczególne miesiące)

Model	**Predykcja liczby zdarzeń**
ARIMA	77,49
Exp	44
Theta	67,21

ników KRI pozwoli na identyfikację podzbioru procesów, w których te algorytmy mogłyby zostać skutecznie zastosowane. Jednak przeprowadzone badania wykazały, że wybrany komitet, składający się z pięciu klasyfikatorów, charakteryzuje się wystarczającą uniwersalnością i efektywnością, aby można było go zastosować do predykcji dowolnego wskaźnika KRI.

3.4. Miary jakości predykcji

Dla wybranego okresu oraz każdego z rozważanych modeli lub komitetów modeli przeprowadzono porównanie wyników predykcji z rzeczywistymi wartościami wskaźników KRI. Celem było uzyskanie ilościowej oceny skuteczności modeli, pozwalającej wskazać rozwiązanie generujące najmniejszy błąd prognozy. Do oceny zastosowano tradycyjne metryki jakości modeli regresyjnych, takie jak średni błąd bezwzględny (MAE), błąd średniokwadratowy (MSE) oraz średni bezwzględny błąd procentowy (MAPE) (Cerqueira, Torgo i Mozetič, 2020).

- **MSE**, czyli błąd średnio-kwadratowy (ang. *mean squared error*) – poprzez użycie kwadratu błędu metryka ta zwraca gorszy wynik dla modelu, którego jednostkowe predykcje mogą istotnie odbiegać od wartości referencyjnych. Metryka jest trudna w interpretacji ponieważ jej wartości nie są wyrażone w tych samych jednostkach, w których wyrażona jest przewidywana wartość.
- **MAE** , czyli średni błąd bezwzględny (ang. *mean absolute error*) -– metryka ta kładzie mniejszy nacisk na rażąco odstające predykcje stosując prostą liniową proporcjonalność. Podstawową zaletą metryki MAE jest jej interpretowalność, ponieważ jest wyrażona w tych samych jednostkach, w których wyrażona jest przewidywana wartość.
- **MAPE**, czyli średni bezwzględny błąd procentowy (ang. *mean absolute percentage error*) – średni bezwzględny błąd procentowy, metryka łatwa w interpretacji, ponieważ jej wartość jest niezależna od jednostek, w których wyrażona jest przewidywana wartość.

Sformułowano zalecenia dotyczące doboru metryki ewaluacyjnej w zależności od charakterystyki danego wskaźnika KRI, w szczególności jego zmienności:

- Dla wskaźników charakteryzujących się dużą zmiennością lub wyraźnymi trendami (np. sezonowość, gwałtowne skoki) rekomendowana jest metryka MSE. Jej czułość na duże błędy sprzyja wyborowi modeli, które lepiej radzą sobie z wartościami odstającymi i dynamicznymi zmianami w czasie.
- Dla wskaźników o niskiej zmienności i słabo zarysowanym trendzie rekomendujemy stosowanie metryk opartych na różnicach bezwzględnych. Dodatkowo sugerujemy preferowanie metryki MAE w przypadku wskaźników utrzymujących się na relatywnie wysokim poziomie liczbowym (np. powyżej 1.0), ponieważ zachowuje ona spójność jednostkową i dostarcza intuicyjnych informacji o przeciętnej skali błędu. z kolei dla wskaźników, których wartości są niewielkie (np. poniżej 1.0, w tym ułamki i procenty), bardziej odpowiednia jest metryka MAPE, gdyż jej procentowy charakter pozwala uniknąć zawyżonych względnych ocen błędu typowych dla danych o małej skali liczbowej.

3.5. Przykładowe wyniki

Wymienione wcześniej techniki zostały zaimplementowane tak, aby estymować jakość modeli predykcyjnych. W tabelach 3.2 3.3 i 3.4 zaprezentowano pięć uruchomień algorytmu rosnących okien czasowych.

Tabela 3.2: Przykładowe wyniki predykcji z użyciem rosnących okien na szeregu czasowym z 24 miesięcy dla metod statystycznych. Pogrubieniem zaznaczono wartości rzeczywiste

	Zakres dat	Okno przesuwne	Wart. rzecz.	ARIMA	Exp	Theta
1	10.2020-12.2021	01-05 2022	**3,50**	3,78	3,91	3,67
2	10.2020-01.2022	02-06 2022	**3,53**	3,77	3,84	3,89
3	10.2020-02.2022	03-07 2022	**3,29**	3,78	3,86	3,75
4	10.2020-03.2022	04-08 2022	**3,27**	3,76	3,79	3,59
5	10.2020-04.2022	05-09 2022	**3,52**	3,7	3,01	3,30

Legenda: Zakres dat – zakres dat zbioru treningowego (liczba miesięcy), Okno przesuwne – zakres okna przesuwnego, Wart. rzecz. – wartości rzeczywiste, ARIMA, Exp, Theta – metody prognozowania.

W uruchomieniu pierwszym jako zbiór treningowy użyto danych z 15 miesięcy od października 2020 do grudnia 2021. Zbudowanych zostało sześć modeli:

- trzy modele metod statystycznych, których wyniki przedstawiono w tabeli 3.2;
- trzy modele metod uczenia maszynowego, dla których wyniki przedstawiono w tabeli 3.3.

W tabeli 3.4 przedstawiono średnią wyników uzyskanych za pomocą wszystkich metod.

Każdy z modeli dokonał predykcji na pięć miesięcy (od stycznia do maja 2022). Dodatkowo obliczono średnią wartość predykcji modeli (tabela 3.4), która wyniosła 3.82 (pomijając predykcję skrajną). Prawdziwa wartość średnia badanego KRI w testowanym okresie wyniosła 3.5. w kolejnym wierszu zaprezentowano drugie uruchomienie dla metod statystycznych (tabela 3.2) i uczenia maszynowego (tabela 3.3), w którym zakres czasu predykcji przesunięto o miesiąc do przodu. Predykcja dotyczyła okresu od lutego do czerwca 2022, a okres treningowy zwiększył się o miesiąc (do 16 miesięcy). Tutaj również ponowiono obliczenia, przy czym najlepszy okazał się model ARIMA, który przewidział wartość 3.77, przy wartości prawdziwej wynoszącej 3.53.

Tabela 3.3: Przykładowe wyniki predykcji z użyciem rosnących okien na szeregu czasowym z 24 miesięcy dla metod uczenia maszynowego. Pogrubieniem zaznaczono wartości rzeczywiste

	Zakres dat	Okno przesuwne	Wart. rzecz.	GMB	RF
1	10.2020-12.2021 (15)	01-05 2022	**3,50**	3,85	3,85
2	10.2020-01.2022 (16)	02-06 2022	**3,53**	3,82	3,91
3	10.2020-02.2022 (17)	03-07 2022	**3,29**	3,83	3,95
4	10.2020-03.2022 (18)	04-08 2022	**3,27**	3,80	3,92
5	10.2020-04.2022 (17)	05-09 2022	**3,52**	3,71	3,82

Legenda: Zakres dat – zakres dat zbioru treningowego (liczba miesięcy), Okno przesuwne – zakres okna przesuwnego, Wart. rzecz. – wartości rzeczywiste, GBM - LightGBM , RF - Random Forest – metody prognozowania.

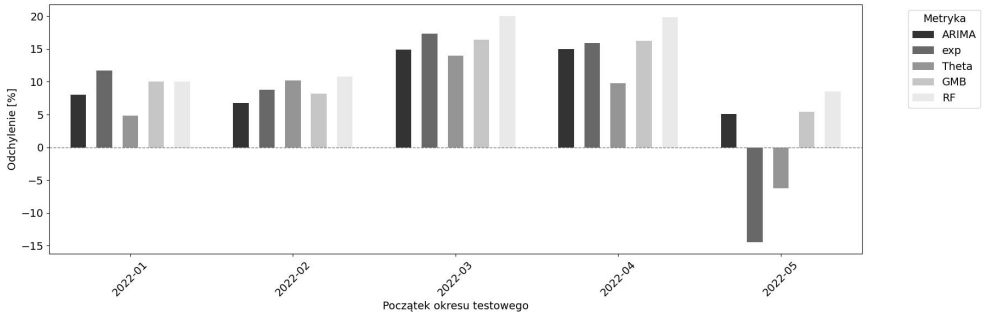
Tabela 3.4: Przykładowe wyniki predykcji z użyciem rosnących okien na szeregu czasowym z 24 miesięcy z uśrednioną wartością wyników modeli statystycznych i modeli uczenia maszynowego. Pogrubieniem zaznaczono wartości rzeczywiste

	Zakres dat	Okno przesuwne	Wart. rzecz.	Średnia
1	10.2020-12.2021 (15)	01-05 2022	**3,50**	3,81
2	10.2020-01.2022 (16)	02-06 2022	**3,53**	3,85
3	10.2020-02.2022 (17)	03-07 2022	**3,29**	3,83
4	10.2020-03.2022 (18)	04-08 2022	**3,27**	3,77
5	10.2020-04.2022 (17)	05-09 2022	**3,52**	3,51

Legenda: Zakres dat – zakres dat zbioru treningowego (liczba miesięcy), Okno przesuwne – zakres okna przesuwnego, Wart. rzecz. – wartości rzeczywiste, Średnia – uśredniony wynik z wszystkich zastosowanych metod.

Przedstawione w tabelach 3.2 i 3.3 wyniki zaprezentowano na wykresie 3.5. Wykres przedstawia procent odchylenia standardowego wartości predykcji uzyskanych dla każdej metody w badanym okresie od wartości rzeczywistej.

Otrzymane predykcje posłużyły do wyliczenia miar ich jakości, takich jak MSE, MAE i MAPE. w tabeli 3.5 zestawiono wyniki dla tych miar. W przedstawionym



Rysunek 3.5: Wykres przedstawiający odchylenie procentowe od wartości rzeczywistej dla każdej z metod statystycznych i uczenia maszynowego. Oś X przedstawia datę końca okresu testowego, natomiast oś Y procentową wartość odchylenia

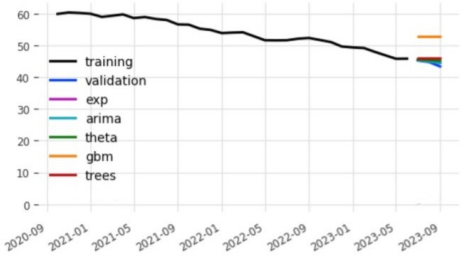
przykładzie model Theta wypadł najlepiej (najmniejszy błąd) pod względem każdej z użytych miar.

Tabela 3.5: Otrzymane miary jakości modeli statystycznych i modeli uczenia maszynowego. Pogrubieniem oznaczono pola z wynikami o najmniejszym błędzie

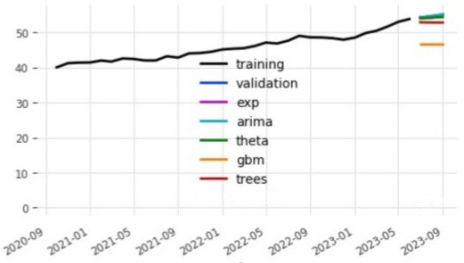
	ARIMA	**Exp**	**Theta**	**GMB**	**RF**	**Średnia**
MSE	0,134	0,226	**0,108**	0,166	0,245	0,160
MAE	0,340	0,466	**0,311**	0,381	0,469	0,359
MAPE	0,101	0,137	**0,092**	0,113	0,139	0,107

3.6. Eksperymenty na danych syntetycznych

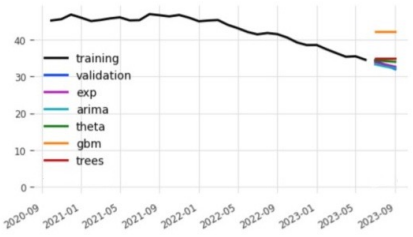
Każdy zaprojektowany eksperyment stanowi symulację określonego zjawiska i przyjmuje postać syntetycznego szeregu czasowego o zdefiniowanej charakterystyce (Viana, Teixeira, Baptista i Pinto, 2024), odpowiadającej długoterminowym zjawiskom, takim jak rosnąca inflacja, gwałtowne załamanie kursu złotego czy spadek poziomu bezrobocia. Na potrzeby realizacji zadania opracowano pięć takich eksperymentów.



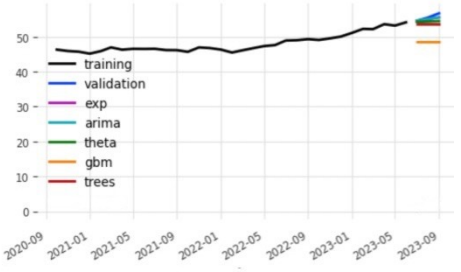
(a) Szereg malejący: szereg czasowy z wyraźną linią monotonicznie malejącego trendu i niewielką wariancją poszczególnych instancji danych wokół trendu



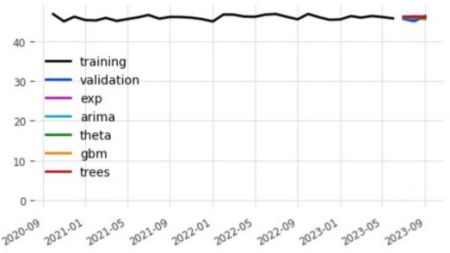
(b) Szereg rosnący: analogicznie do szeregu malejącego, ale linia trendu jest monotonicznie rosnąca



(c) Szereg skokowo malejący: szereg czasowy z wartościami rozrzuconymi wokół wyraźnie zaznaczonej, stałej linii trendu, w którego środku pojawia się wyraźny trend malejący



(d) Szereg skokowo rosnący: analogiczny do szeregu skokowo malejącego, ale linia trendu po zmianie jest monotonicznie rosnąca



(e) Szereg stały: szereg czasowy z wartościami rozrzuconymi losowo wokół wyraźnej, stałej linii trendu

Rysunek 3.6: Zestawienie wykresów prezentujących wyniki zaprojektowanych eksperymentów walidacyjnych dla szeregu malejącego, szeregu rosnącego, szeregu skokowo malejącego, szeregu skokowo rosnącego oraz szeregu stałego

Rysunek 3.6 przedstawia wykresy wizualizujące wyniki opracowanych eksperymentów dla szeregu malejącego, szeregu rosnącego, szeregu skokowo malejącego, szeregu skokowo rosnącego oraz szeregu stałego. Analizując otrzymane wyniki, można zaobserwować, że wszystkie modele, z nieco lepszym lub gorszym skutkiem, są w stanie podążać za syntetyczną linią trendu. Model GBM wydaje się często prognozować średnią z całego okresu, natomiast pozostałe modele reagują poprawnie na zmianę trendu (lub jej stałą wartość).

3.7. Aplikacja WWW predykcji wskaźnika KRI

W trakcie pracy nad projektem zaimplementowano model predykcyjny dla wybranych wskaźników KRI dotyczących różnych procesów bankowych oraz szereg aplikacji wspomagających predykcję, m.in. agregujących dane do predykcji, wizualizujących wyniki oraz umożliwiających ich analizę. Dla każdego z procesów użytkownik korzystający z raportu wizualizacji może:

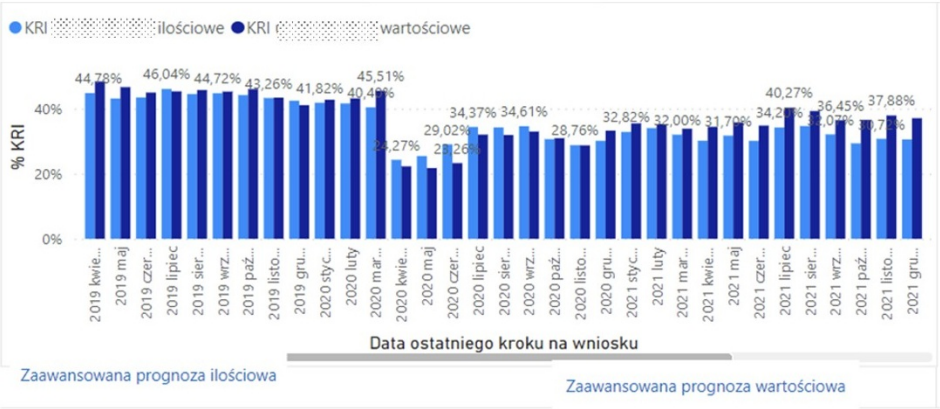
- zdefiniować w raporcie wizualizacji, przy użyciu fragmentatorów, interesujący go zakres danych, dla których wyznaczany jest historyczny trend wartości wskaźnika (np. dla wybranego segmentu klientów);
- następnie przejść do aplikacji WWW prezentującej predykcję wartości wskaźnika w przyszłych okresach poprzez kliknięcie na odnośnik umieszczony w raporcie wizualizacji.

Odnośniki służące do przejścia z raportów wizualizacji dotyczących poszczególnych wskaźników do aplikacji WWW prezentującej ich predykcję znajdują się pod wykresami prezentującymi historyczny trend ich wartości (Rysunek 3.7). Kliknięcie w odnośnik powoduje uruchomienie reguły języka DAX, która generuje miarę zawierającą wartości szeregu czasowego prognozowanego wskaźnika KRI.

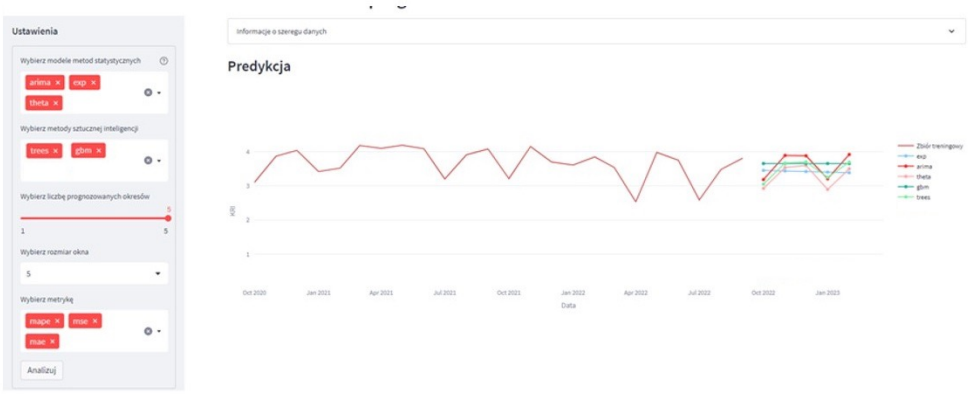
Strona WWW modelu predykcji przetwarza przekazany z raportu szereg czasowy. Szereg czasowy przekazywany jest w formie parametrów GET żądania HTTP. Żądanie HTTP jest wysyłane do serwera WWW po kliknięciu odnośnika w raporcie wizualizacji dotyczącym wybranego wskaźnika. Żądanie HTTP wyświetla stronę z wizualizacją przedstawioną na Rysunku 3.8. Wszystkie obliczenia, prognozy i predykcje są wyznaczane za pomocą biblioteki darts (*Darts: Time Series Made Easy*, 2024). Warto dodać, że stworzona architektura rozwiązania jest uniwersalna i umożliwia pracownikom banku podłączenie w prosty sposób modeli predykcyjnych pod dowolny wskaźnik KRI.

Aplikacja WWW umożliwia:

- analizę (dowolnego) otrzymanego szeregu czasowego,
- prognozę wartości wskaźnika KRI w przyszłości różnymi metodami statystycznymi i uczenia maszynowego,



Rysunek 3.7: Przykładowy wykres słupkowy prezentujący poziom KRI w ujęciu ilościowym i wartościowym oraz odnośniki prowadzące do zaawansowanej prognozy na stronie WWW



Rysunek 3.8: Interfejs użytkownika aplikacji WWW

- testowanie modeli za pomocą okna przesuwnego.

Interfejs użytkownika (Rysunek 3.8) pozwala na wybranie jednego lub kilku modeli predykcji spośród modeli statystycznych i modeli uczenia maszynowego (omówionych w punkcie *predykcja KRI*), liczby prognozowanych okresów, rozmiaru okna przesuwnego oraz metryki do raportowania błędów (omówione w punkcie *Miary jakości predykcji*).

Po wybraniu ustawień aplikacja WWW zwraca predykcję na żądany okres dla każdej wybranej metody, tabelę porównawczą z oceną modeli za pomocą wybranych miar, a także wartości predykcji otrzymane w kolejnych krokach przy zastosowaniu metody okna przesuwnego.

Podsumowanie

W ramach projektu opracowano zestaw aplikacji umożliwiających wizualizację, monitorowanie oraz prognozowanie kluczowych wskaźników ryzyka (KRI). Wszystkie narzędzia wspierają analizę danych z uwzględnieniem podziału na jednostki biznesowe oraz okresy, oferując jednocześnie tabele typu *drill-down*, które identyfikują umowy lub operacje odbiegające od oczekiwań. Funkcjonalność rozwiązań była na bieżąco weryfikowana poprzez testy użyteczności, szybkości działania oraz bezpieczeństwa związanego z przetwarzaniem i prezentacją danych. w toku projektu zoptymalizowano działanie aplikacji zarówno pod kątem wydajności, jak i przejrzystości interfejsu użytkownika.

Wyniki wdrożenia pozostają zgodne z wnioskami płynącymi z wcześniejszych projektów badawczo-wdrożeniowych, takich jak studium przypadku (Alles, Brennan, Kogan i Vasarhelyi, 2018) dotyczące ciągłego monitorowania procesów biznesowych w firmie Siemens. Podobnie jak w tamtym podejściu istotną rolę odegrało zastosowanie wizualizacji oraz bieżącej analizy danych w celu wczesnego wykrywania anomalii i odchyleń od normy. Należy jednak podkreślić istotną różnicę – w odróżnieniu od systemu Siemens opracowane przez nas narzędzia oferują również funkcje predykcyjne, umożliwiające prognozowanie przyszłych poziomów wskaźników ryzyka. Rozszerza to zakres zastosowania systemu – od pasywnego monitorowania do aktywnego wspierania decyzji zarządczych na podstawie przewidywanych trendów. Oba systemy dowodzą, że ciągły audyt oparty na danych operacyjnych może realnie wspierać zarządzanie ryzykiem i poprawę jakości kontroli wewnętrznych.

Podziękowania

Autorzy składają serdecznie podziękowania dr. hab. inż. Mikołajowi Morzemu oraz przedstawicielom banku za nieoceniony wkład w realizację projektu.

Finansowanie

Projekt *Model audytu oparty na metodyce ciągłej oceny ryzyka i mechanizmów kontrolnych z wykorzystaniem metod statystycznych i sztucznej inteligencji* finansowany był ze środków NCBiR na podstawie umowy POIR.01.01.01-00-1284/20.

Bibliografia

Alles, M., Brennan, G., Kogan, A. i Vasarhelyi, M. A. (2018). Continuous monitoring of business process controls: A pilot implementation of a continuous auditing system at siemens. W: *Continuous auditing: Theory and application* 219–246.

Assimakopoulos, V., i Nikolopoulos, K. (2000). The theta model: a decomposition approach to forecasting. *International Journal of Forecasting*, *16*(4), 521–530.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Campante, M. I., Gonçalves, C. T. i Gonçalves, M. J. A. (2023). Business intelligence tools to improve business strategy. W: *International conference on tourism, technology and systems.*

Cerqueira, V., Torgo, L. i Mozetič, I. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, *109*(11), 1997–2028.

Chatfield, C. (1978). The holt-winters forecasting procedure. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *27*(3), 264–279.

Darts: Time series made easy. (2024). `https://pypi.org/project/darts/0.8.1/`. ([Accessed: 2024-09-29])

Ishtiaq, M. (2015). *Risk management in banks: determination of practices and relationship with performance.* (Unpublished or institutional paper)

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. i Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. W: *Advances in neural information processing systems* 30.

Lizotte-Latendresse, S., i Beauregard, Y. (2018). Implementing self-service business analytics supporting lean manufacturing: A state-of-the-art review. *IFAC-PapersOnLine*, *51*(11), 1143–1148.

Radziwon, K. (2013). Ciągły audyt/ciągły monitoring: Kpmg: Audyt i doradztwo w instytucjach finansowych. *Bank*(7-8), 30–31.

Shumway, R. H., i Stoffer, D. S. (2017). Arima models. W: *Time series analysis and its applications: With r examples* 75–163.

Skender, F., i Manevska, V. (2022). Data visualization tools-preview and comparison. *Journal of Emerging Computer Technologies*, *2*(1), 30–35.

Taieb, S. B., Bontempi, G., Atiya, A. i Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition. *Expert Systems with Applications*, *39*(8), 7067–7083.

Viana, D., Teixeira, R., Baptista, J. i Pinto, T. (2024). Synthetic data generation models for time series: A literature review. W: *2024 International Conference on Electrical, Computer and Energy Technologies (ICECET)* 1–6.

Ziora, L. (2011). Systemy business intelligence jako narzędzie wspierające podejmowanie decyzji w przedsiębiorstwach. przegląd studiów przypadków branży finansowej i en. *Informatyka Ekonomiczna*, *20*, 120–130.