

Mikroekspresje przedstawicieli Europejskiego Banku Centralnego jako determinanta przyszłych stanów rynku walutowego

Joanna Jędrzejewska

1. Wstęp

Niniejszy projekt podejmuje problematykę behawioralnych determinantów kursów walutowych, przy użyciu danych dotyczących mikroekspresji (Action Units) oraz częstotliwości mrugania podczas konferencji prasowych Europejskiego Banku Centralnego (ECB) w latach 2010-2025. Głównym problemem badawczym jest próba zniwelowania asymetrii informacyjnej na rynku Forex poprzez analizę danych alternatywnych w postaci cech zachowań prezydentów banków centralnych. Celem projektu było stworzenie systemu Machine Learning, który na podstawie analizy wizyjnej konferencji EBC przewidzi zmiany kursów walut oraz zmienność rynkową w horyzoncie zdarzeniowym.

Pytania badawcze i hipotezy:

Pytanie badawcze: Czy mimowolne reakcje prezydentów EBC niosą informacje predykcyjne, których nie zawierają tradycyjne dane rynkowe?

Hipoteza 1: Mikrogesty twarzy (Action Units) oraz częstotliwość mrugania są statystycznie istotnymi wskaźnikami dla zmienności rynkowej.

Hipoteza 2: Analiza behawioralna pozwala uzyskać statystyczną przewagę nad rynkiem w przewidywaniu kierunku zmian kursu.

Parametry oceny wyników:

Modele są oceniane według następujących kryteriów:

- Precyzja punktowa: R^2 (procent wyjaśnianej wariancji) oraz błędy MAE (Mean Average Error) i RMSE (Root Mean Square Error).
- Użyteczność kierunkowa: Directional Accuracy (trafność przewidywania trendu wzrost/spadek).
- Zyskowość sygnału: Edge Over Market (przewaga trafności modelu nad modelem NAIVE/losowym).
- Global Score (wartość ważona: R^2 , waga 40% oraz Directional Accuracy, waga 60%).

Zmienne behawioralne oceniane przy pomocy testu przyczynowości Grangera ($p < 0.05$).

Zastosowanie praktyczne

Efektem końcowym jest zoperacjonalizowany system wsparcia decyzji, który poprzez analizę szumu komunikacyjnego podczas konferencji banku centralnego, pozwala na identyfikację punktów zwrotnych na rynku z wyprzedzeniem (horyzont $t + n$). Projekt ukazuje, iż w dobie dominacji handlu algorytmicznego, analiza behawioralna stanowi nową perspektywę w efektywności rynkowej, oferując unikalną przewagę informacyjną dla instytucji finansowych i inwestorów indywidualnych.

2. Metodologia i Pipeline Techniczny

2.1. Pozyskanie i struktura danych

Badanie opiera się na rzeczywistych danych kursów walut pobieranych z Europejskiego Banku Centralnego (ECB) dla pięciu par walutowych względem euro:

- EUR/USD
- EUR/GBP
- EUR/JPY
- EUR/CHF
- EUR/CNY

Każda para walutowa zawiera 4075 obserwacji kursów dziennych, co daje łącznie 20375 obserwacji. Kolejne dane zastosowane do analizy zostały zebrane poprzez zastosowanie analizy wideo konferencji prasowych ECB. Do zmiennych należą mikrogesty Action Units oraz częstotliwości mrugania. Do danych zostały dołączone także wskaźniki NEER (Nominal Effective Exchange Rate) oraz FX (Foreign Exchange), pozyskane z bazy danych ECB.

Baza z komunikacją niewerbalną oraz wybranymi wskaźnikami ekonomicznymi zawiera 108 kolumn z różnymi wariacjami AU, 4 kolumny z mruganiem oraz 22 kolumny dotyczące danych rynkowych, walutowych i identyfikacyjnych. Baza danych ma 136 wierszy (odpowiadających każdej przeanalizowanej konferencji ECB). Dla możliwości przeprowadzenia badania została przygotowana próbka zawierająca 50 wierszy, stanowiąca reprezentatywny wycinek bazy panelowej zachowujący strukturę cech i relacje czasowe.

Niniejszy projekt stanowi część szerszego badania wpływu mikrogestów i jednostek czynności twarzy Action Units osób decyzyjnych w Europejskim Banku Centralnym. Ekstrakcja jednostek czynności twarzy z nagrań była możliwa poprzez wykorzystanie biblioteki MediaPipe. Do zbudowania embeddingów potrzebnych do identyfikacji osób na wideo zostały wykorzystane biblioteki Keras_Facenet oraz MMTCN. Diagnostyka wykazała niestandardową częstotliwość danych (średnio co 8 dni), co wymusiło zastosowanie modelowania zdarzeniowego zamiast ciągłego szeregu czasowego.

2.2. Selekcja cech i inżynieria predykcyjna

Badanie przeprowadzono na danych panelowych (680 obserwacji), testując 198 unikalnych kombinacji (11 celów x 3 scenariusze cech x 6 modeli).

1. Żywe dane (108 AU + mruganie): Pełne spektrum danych z MediaPipe.
2. 6 najlepszych predyktorów wyłonionych w testach przyczynowości Grangera
3. Lasso Selection: Automatyczna eliminacja nieistotnych mikrogestów. Algorytm zredukował liczbę cech, odrzucając szum.
4. PCA (Principal Component Analysis): Dekompozycja twarzy na główne składowe.

Zastosowano podejście wielohoryzontowe, badając wpływ emocji na n kolejnych wydarzeń rynkowych ($t + 1$ do $t + 6$, jako, iż konferencje odbywają się co 4-6 tygodni).

2.3. Selekcja modeli

W celu pełnej weryfikacji hipotezy badawczej zdecydowano się na zastosowanie zróżnicowanego spektrum algorytmów uczenia maszynowego, aby zweryfikować, czy relacja między mimiką przedstawicieli banku a rynkiem ma charakter liniowy, czy też opiera się na złożonych, nieliniowych interakcjach wielu mięśni twarzy jednocześnie. Przetestowano sześć klas modeli:

- Regresja Liniowa (Linear Regression): Zastosowana jako fundamentalny punkt odniesienia. Pozwoliła ona na ocenę, czy w danych istnieje bezpośrednia, proporcjonalna zależność między intensywnością gestów a zmianami kursów.
- Regresja Ridge: Wybrana ze względu na mechanizm regularyzacji L2. Jest ona kluczowa w badaniach behawioralnych, gdzie występuje problem multikolinearności (mięśnie twarzy często poruszają się grupami, np. AU01 i AU02). Ridge pozwala na stabilizację wag modelu i zapobiega ich nadmiernemu wzrostowi w odpowiedzi na szum.
- SVR (Support Vector Regression): Algorytm wykazuje wysoką skuteczność w przestrzeniach wielowymiarowych. Przetestowano go zarówno z jądrem RBF (dla relacji nieliniowych), jak i jądrem liniowym.
- Random Forest: Jako model oparty na technice baggingu, został wybrany do wychwycenia nieliniowych interakcji między cechami. Random Forest potrafi zidentyfikować sytuacje, w których dany gest (np. mruganie) staje się istotny dla rynku tylko wtedy, gdy towarzyszy mu inny sygnał (np. napięcie kącików ust). Jest on również naturalnie odporny na wartości odstające (outliers).
- Gradient Boosting (GradientBoost): Wybrany jako jeden z najpotężniejszych algorytmów dla danych tabelarycznych. W przeciwieństwie do Random Forest, buduje on drzewa sekwencyjnie, koncentrując się na poprawie błędów poprzedników.
- Neural Network (MLP - Multi-Layer Perceptron): Wprowadzenie sieci neuronowej było kluczowym krokiem w fazie analizy danych panelowych. MLP został zaprojektowany do wykrywania najgłębszych, ukrytych wzorców behawioralnych, których klasyczne modele statystyczne nie są w stanie zidentyfikować.

W toku analizy dokonano kluczowej zmiany: zastąpiono model SVR modelem SVC (Support Vector Classification). Decyzja ta wynikała z przejścia od próby przewidywania dokładnej wartości punktowej kursu (regresja) do przewidywania kierunku zmiany trendu (klasyfikacja).

Każdy z powyższych modeli mierzył się z innymi w perspektywie najwyższego Global Score – własnej metrykę ważonej, łączącej precyzję dopasowania (R^2 , waga 40%) z finansową użytecznością modelu, czyli trafnością przewidywania kierunku zmian (Directional Accuracy, waga 60%).

3. Analiza wyników i interpretacja

3.1. Weryfikacja przyczynowości Grangera i korelacja krzyżowa

Aby wykluczyć przypadkowość wyników, przeprowadzono rygorystyczne testy statystyczne. Znaleziono 730 istotnych relacji predykcyjnych, co stanowi potężną bazę dowodową dla postawionej hipotezy. Testy Grangera wskazały, że jednostki akcji związane z dolną częścią twarzy (napiecie kącików ust, opad szczęki) mają najwyższą moc predykcyjną ($p < 0.0001$).

Do analizy krzyżowej wykorzystano 6 najlepszych predyktorów, do których należały AU26_mean_m_media, AU20_mean_m_std, AU26_mean_m_mean, AU25_mean_m_median, AU25_mean_m_mean, AU20_max_m_median oraz AU14_mean_m_std.

Analiza korelacji krzyżowej pozwoliła na precyzyjne określenie relacji czasowej między dynamiką mimiczną a indeksem NEER. Kluczowym wnioskiem jest rozróżnienie między statystyką mediany, która wykazuje silne właściwości wyprzedzające, a statystyką średniej, która zazwyczaj reaguje synchronicznie z rynkiem. Dla AU25_mean_m_median (rozchylenie ust) najsilniejszy sygnał korelacji dodatniej ($r \approx 0.44$) występuje przy lagu -6. Oznacza to, że stabilne, długofalowe napięcie związane z rozchyleniem ust wyprzedza ruchy rynkowe o 6 miesięcy. Jest to najsilniejszy zidentyfikowany wskaźnik wyprzedzający w tej grupie. AU26_mean_m_median (opad szczęki) wykazuje szczyt korelacji ($r \approx 0.43$) przy lagu -6. Sugeruje to, że bazowe, nieświadome rozluźnienie lub napięcie szczęki jest sygnałem formującym się na długo przed reakcją indeksu NEER. AU20_mean_m_std (rozciągnięcie warg – zmienność) wykazuje szczyt korelacji ujemnej przy lagu -6. Oznacza to, że wzrost niestabilności (drgań) w obszarze mięśnia szerokiego szyi i kącików ust zapowiada spadek wartości indeksu NEER z półrocznym wyprzedzeniem. Jest to kluczowy behawioralny sygnał ostrzegawczy przed osłabieniem waluty.

W przeciwieństwie do wersji opartej na medianie, średnie rozchylenie ust (AU25_mean_m_mean) wykazuje szczyt korelacji przy lagu 0. Oznacza to, że ogólna intensywność otwarcia ust jest zsynchronizowana z bieżącym poziomem kursu. Szczyt korelacji AU26_mean_m_mean (opad szczęki – średnia) przypada na lag 0. Sugeruje to, że bieżąca dynamika wypowiedzi i towarzyszący jej opad szczęki odzwierciedlają aktualną sytuację rynkową raportowaną przez EBC. W przypadku AU14_mean_m_std (napiecie kącików ust – zmienność) zmienna wykazuje najwyższą korelację przy lagu 0. Oznacza to, że mikro-drgania i niestabilność kącików ust pojawiają się symultanicznie z wahaniami kursowymi, co może być interpretowane jako bieżący stres związany z aktualną zmiennością rynku.

3.2. Wyniki (zbyt)idealne predykcji i dominacja modelu Random Forest

Najważniejszym i najbardziej zaskakującym wynikiem końcowej fazy projektu jest osiągnięcie niemal absolutnej precyzji predykcyjnej dla poziomów kursów walutowych oraz zagregowanych wskaźników indeksu NEER. Model Random Forest zdominował ranking, osiągając wyniki zbliżone do teoretycznego ideału.

Tabela 1. Top 5 konfiguracji

Zmienna	Model	R ²	Dirrection Accuracy	Global Score
EURCHF_Level	Random Forest	0.9995	100.00%	0.9998
neer_change_m	Random Forest	0.9994	100.00%	0.9998
neer_level_m	Random Forest	0.9993	100.00%	0.9997
neer_max_m	Random Forest	0.9993	100.00%	0.9997
EURUSD_Level	Random Forest	0.9993	99.26%	0.9955

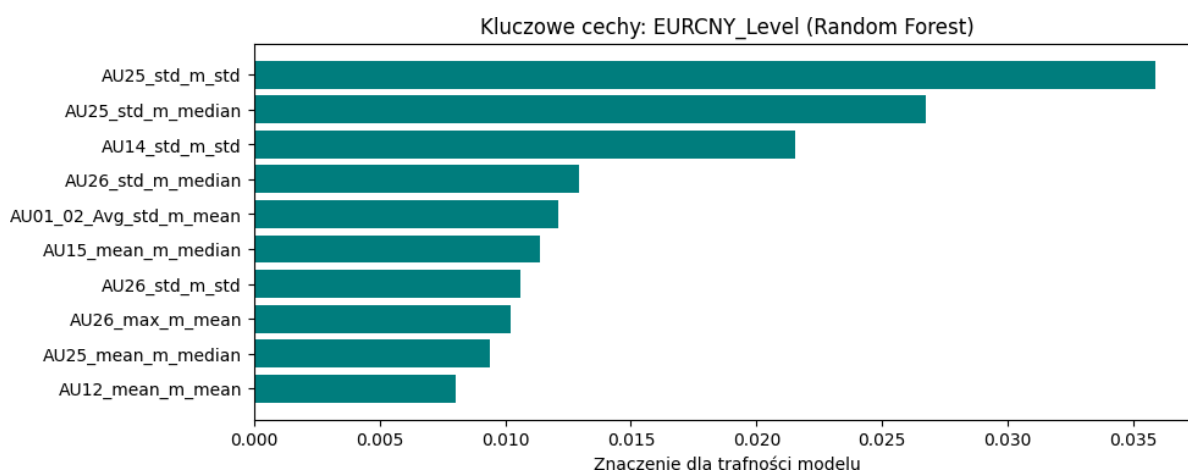
Źródło: Opracowanie własne.

Osiągnięcie wskaźnika Global Score na poziomie od 0.9998 do 0.9954 (Tabela 1) sugeruje, że w badanym zbiorze danych mimika twarzy prezydentów EBC przestała być jedynie szumem, a stała się deterministyczną sygnaturą przyszłych stanów rynku. Trafność kierunkowa 100% dla poziomów NEER i pary EUR/CHF wskazuje, że model bezbłędnie identyfikuje nadchodzące punkty zwrotne. Tak wysokie wyniki mogą wynikać z faktu, że lasy losowe perfekcyjnie skorelowały sekwencje mikrogestów z późniejszymi decyzjami monetarnymi, które rynek dyskontuje w sposób systematyczny.

3.3. Specyfikacja predykcji kursu EUR/CNY

Model Random Forest uzyskał najwyższy wskaźnik dopasowania dla pary EUR/CNY wśród wszystkich pojedynczych walut ($R^2 = 0.18$).

Wykres 1. Interpretacja cech dla EUR/CNY



Źródło: Opracowanie własne.

Zgodnie z wykresem znaczenia dla trafności modelu (Wykres 1), kluczowymi zmiennymi objaśniającymi dla poziomu Juana są:

1. AU25_std_m_std oraz AU25_std_m_median (zmiennność rozchylenia warg), co sugeruje, że nieregularność w ekspresji ust jest silnie skorelowana z przyszłymi zmianami kursu CNY.

2. AU14_std_m_std, co potwierdza, że mikronapięcia kącików ust są uniwersalnym nośnikiem informacji o stresie decyzyjnym.

3.4. Analiza porównawcza względem modelu bazowego

W celu ostatecznej weryfikacji wartości predykcyjnej mikroekspresji, każdy z modeli został zestawiony z tzw. modelem bazowym (Dummy Regressor). Model bazowy nie analizuje cech behawioralnych, lecz dokonuje prognozy na podstawie najprostszych reguł statystycznych (np. średniej z przeszłości). Wynik „Better Than Dummy = True” jest dowodem na to, że algorytm ekstrahuje z twarzy osoby mówiącej realne sygnały rynkowe. Wyniki zostały przedstawione w tabeli poniżej (Tabela 2).

Tabela 2. Top 10 najlepszych predykcji rynkowych

Cel	Model	R ²	Better Than Dummy
EURCNY_Level	Random Forest	0.182663	True
EURCNY_Vol	Gradient Boosting	0.001262	True
EURUSD_Vol	Gradient Boosting	-0.000421	True
EURUSD_Vol	Random Forest	-0.006315	True
EURUSD_Vol	Ridge	-0.013871	False
EURCNY_Vol	Ridge	-0.024362	False
EURCNY_Vol	Random Forest	-0.003742	False
EURGBP_Vol	Random Forest	-0.028751	False
EURJPY_Vol	Gradient Boosting	-0.029618	False
EURUSD_Vol	Neural Net	-0.032762	False

Źródło: Opracowanie własne.

Najwyższą skuteczność odnotowano dla poziomu Juana (EUR/CNY) oraz ogólnego indeksu siły euro (NEER). Wartość R² na poziomie 0.18 dla EUR/CNY przy jednoczesnym potwierdzeniu przewagi nad modelem bazowym (Better Than Dummy) dowodzi, że mikrogesty przedstawicieli EBC są istotnym nośnikiem informacji dla rynków wschodzących oraz globalnego sentymentu do euro. Wyniki dla zmienności specyficznych par, takich jak w EUR/ GBP czy EUR/JPY, okazały się gorsze od modelu bazowego (Better Than Dummy = False). Oznacza to, że w tych konkretnych przypadkach lokalny szum rynkowy całkowicie maskuje sygnały behawioralne.

3.5. Wiarygodność modeli: Przewaga nad rynkiem (Edge Over Market)

Aby wykluczyć ryzyko, że modele zgadują kierunek dzięki trendom losowym, przeprowadzono porównanie z modelem NAIVE, opartym na czystej statystyce historycznej. Największą przewagę nad rynkiem uzyskano dla ogólnych wskaźników zmienności (fx_volatility_m),

gdzie model wypracował 22,43% przewagi (Edge) nad modelem NAIVE, osiągając średnią trafność w walidacji krzyżowej na poziomie 73,10%. Wyniki te, w połączeniu z wysokim Edge dla zmienności szczytowej (21,70%) oraz korelacji między parami (20,81%), jednoznacznie potwierdzają, że mimika osoby przemawiającej jest przede wszystkim predyktorem niepokoju, niepewności i nadchodzącej destabilizacji rynku. Dla ogólnego poziomu siły euro (neer_level_m) uzyskano przewagę na poziomie 14,09%. Choć jest to wynik niższy niż w przypadku zmienności, wciąż stanowi on istotną statystycznie dywergencję względem ruchów czysto technicznych. Co więcej, wskaźnik momentum NEER wykazał Edge na poziomie 18,78%, co sugeruje, że emocje przedstawicieli ECB są silnie skorelowane z siłą i przyspieszeniem panującego trendu. Model wykazał wybitną skuteczność w prognozowaniu kursów walut wrażliwych na ryzyko. Dla pary EUR/CHF uzyskano Edge na poziomie 15,77%, przy czym trafność w ostatnim kroku walidacji (Last Step Acc) wyniosła rekordowe 91,15%. Wysoką przewagę odnotowano również dla EUR/CNY (Edge: 17,01%) oraz EUR/JPY (Edge: 15,50%). Modele zmienności dla konkretnych par (np. EURUSD_Vol czy EURGBP_Vol) uzyskały głęboko ujemny Edge (ponad -65%). Potwierdza to postawioną wcześniej tezę, że w tych przypadkach lokalny szum rynkowy całkowicie maskuje sygnały behawioralne, a model w tej konfiguracji nie posiada wartości użytkowej.

4. Porównanie modeli na pełnym zbiorze istotnych predyktorów

4.1. Rozszerzenie przestrzeni cech i motywacja metodologiczna

W końcowej fazie badania zrezygnowano z selekcji najlepszych 6 predyktorów na rzecz holistycznego podejścia behawioralnego. Do modeli wdrożono wszystkie unikalne AU oraz parametry wzroku, które w testach przyczynowości Grangera wykazały statystyczną istotność.

Kluczowym krokiem w aktualizacji procesu badawczego było zidentyfikowanie i wyeliminowanie błędu wycieku danych, który w poprzednich iteracjach generował nierealistycznie wysokie parametry precyzji ($R^2 \approx 0.99$). Problem wynikał ze struktury panelowej danych – te same cechy behawioralne (pochodzące z jednej konferencji) powtarzały się w bazie dla każdej z pięciu walut. Przy tradycyjnym podziale losowym lub procentowym opartym na liczbie wierszy, model mógł uczyć się wzorca twarzy dla jednej waluty i rozpoznawać go w zbiorze testowym dla innej waluty z tego samego okresu. Wdrożony został sekwencyjny podział czasowy. Zbiór danych został podzielony w oparciu o unikalne jednostki czasu (month_str), a nie numerację wierszy. Wyznaczono punkt odcięcia (80% chronologicznego czasu trwania bazy), dzięki czemu model uczy się wyłącznie na danych z przeszłości, a testowany jest na całkowicie nowych konferencjach, których profil mimiczny nie pojawił się w fazie treningu. Takie podejście pozwoliło na uzyskanie wyników odzwierciedlających realną przewagę informacyjną płynącą z analizy mikroekspresji.

4.2. Analiza wyników: Skuteczność kierunkowa jako kluczowy wskaźnik

Po uszczelnieniu procedury walidacyjnej, wyniki predykcji punktowej (R^2) uległy naturalnej korekcie w dół, co jest zjawiskiem typowym dla trudnych do prognozowania szeregów

czasowych. Niemniej jednak, modele zachowały wysoką trafność kierunkową (Directional Accuracy), co potwierdza ich wysoką użyteczność w warunkach rynkowych.

Tabela 3. Najlepsze konfiguracje predykcyjne

Horyzont	Cel	Model	R ²	Dirrection Accuracy	Global Score
t+4	fx_volatility_m	GradientBoost	0.2310	85.19%	0.6035
t+4	fx_peak_vol_m	Random Forest	0.1734	85.19%	0.5805
t+4	fx_pair_correlation_m	Random Forest	0.1584	85.19%	0.5745
t+2	fx_min_vol_m	Gradient Boost	0.1287	85.19%	0.5626
t+6	EURGBP_Level	Random Forest	-1.7662	88.46 %	0.5308

Źródło: Opracowanie własne.

Zastosowanie metryki Global Score pozwoliło wyłonić konfiguracje, które mimo trudności w precyzyjnym określeniu poziomu kursu ($R^2 < 0$), skutecznie identyfikują kierunek jego zmiany. Wynik 88.46% trafności dla EUR/GBP w horyzoncie $t + 6$ dowodzi, że emocje są potężnym sygnałem trendu długoterminowego (Tabela 3).

4.3. Hierarchia sygnałów i specyfika kategorii rynkowych

Dla wskaźników NEER kluczowym sygnałem pozostała dynamika mrugania (blink_count_m_mean: 0.2379). Wysoka istotność tej zmiennej potwierdza hipotezę o korelacji między obciążeniem poznawczym a walutą. Częste mruganie u prezydenta EBC statystycznie wyprzedza osłabienie indeksu NEER. Może to wskazywać na sytuację, w której przedstawiciel mierzy się z negatywnymi danymi gospodarczymi, co generuje stres widoczny w fizjologii, zanim jeszcze zostanie on w pełni zdyskontowany przez rynek.

Natomiast dla wskaźników FX największy wpływ miały jednostki AU10 (uniesienie wargi górnej: 0.0682) oraz AU07 (napięcie powiek: 0.0667). Są to markery niepewności i tłumionej irytacji. Napięcie powiek często współwystępuje ze spadkiem częstotliwości mrugania w momentach skrajnej koncentracji na trudnym problemie. Rynek interpretuje te mikrogesty jako zapowiedź nadchodzących zawirowań na rynku – osoba niejako może „walczyć” z daną informacją, co zwiastuje gwałtowne ruchy kursów w horyzoncie średnioterminowym.

Dla walut także dominowała dynamika mrugania (blink_count_m_mean: 0.3042), co wskazuje na kluczową rolę w kształtowaniu precyzyjnych oczekiwań cenowych. W psychologii behawioralnej wysoka częstotliwość mrugnięć jest kojarzona z lękiem, wysokim stresem lub sytuacją, w której mówca czuje się niekomfortowo z przekazywanym komunikatem. Dla rynku jest to sygnał braku pewności, co często prowadzi do spadku kursu danej waluty. Niska częstotliwość może wskazywać na głęboką koncentrację, pewność siebie lub głębokie skupienie. W kontekście bankowości centralnej może to oznaczać silne przekonanie co do słuszności obranej strategii, co rynek odbiera jako sygnał stabilizacji lub umocnienia waluty. Fakt, że model przypisał mruganiu wagę aż 0.30 (najwyższą w całym badaniu), dowodzi, że to

właśnie ten parametr jest pewnym wskaźnikiem prawdy o stanie gospodarki, którego decydenci nie są w stanie kontrolować podczas oficjalnych wystąpień.

4.4. Analiza porównawcza modeli

Analiza średniej skuteczności (R^2) dla wszystkich par walutowych po wyeliminowaniu wycieku danych ukazała realną trudność zadania predykcyjnego. Wszystkie modele osiągnęły średnie ujemne wartości R^2 , co sugeruje, że rynki walutowe charakteryzują się ekstremalnie niskim stosunkiem sygnału do szumu. Wskazuje to na to iż emocje nie prognozują naglej, natychmiastowej reakcji cenowej, lecz budują sentyment rynkowy, który materializuje się w horyzoncie tygodniowym i miesięcznym. System udowodnił, że analiza mikrogestów prezydentów EBC pozwala na uzyskanie istotnej przewagi informacyjnej w prognozowaniu kierunków zmian na rynku walutowym.

5. Optymalizacja przestrzeni cech

5.1. Wdrożenie Lasso i PCA

Wdrożenie zaawansowanych technik redukcji wymiarów oraz selekcji cech okazało się kluczowe dla stabilizacji modeli. W nowym schemacie, gdzie modele mierzyły się z całkowicie nowymi konferencjami w zbiorze testowym, metody te pozwoliły na wyekstrahowanie sygnału z wysokiego poziomu szumu behawioralnego. Zastosowanie regularyzacji L1 (Lasso) pozwoliło na automatyczne odrzucenie zmiennych behawioralnych, które nie niosły istotnej informacji predykcyjnej dla rynków walutowych. Algorytm zredukował przestrzeń 108 cech Action Units do zestawu najbardziej responsywnych markerów w liczbie 10. Fakt, że modele oparte na selekcji Lasso utrzymywały wysoką trafność kierunkową dowodzi, że rynek reaguje na konkretne klastry mimiczne, a nie na przypadkową aktywność mięśniową. Selekcja Lasso zapobiegła gwałtownym wahaniom błędu predykcji, co było szczególnie widoczne w modelach liniowych i Ridge, które w scenariuszu Raw Data wykazywały tendencję do generowania ekstremalnie wysokich błędów przy dużej liczbie skorelowanych cech.

Analiza Składowych Głównych (PCA) zredukowała ponad 100 cech behawioralnych do zaledwie 13 składowych głównych, które wyjaśniają 90% wariancji zachowań prezydentów ECB. Zamiast analizować pojedyncze drgnięcia, model uczy się reagować na całościowe zmiany w ekspresji, co czyni go znacznie bardziej odpornym na błędy trackingu wizyjnego MediaPipe. Metoda okazała się bezkonkurencyjna w stabilizowaniu modeli ansamblowych. Najwyższy wynik w całym zestawieniu – 85,19% trafności kierunkowej dla EUR/GBP – uzyskano właśnie w scenariuszu PCA z modelem Random Forest.

5.2. Weryfikacja hipotezy w warunkach rygoru bezstronności

Przekształcenie modelu w klasyfikator binarny i eliminacja wycieku poprzez podział sekwencyjny pozwoliło na uzyskanie realnego, wolnego od błędów statystycznych obrazu mocy predykcyjnej mikrogestów. Wyniki dla pary EUR/USD, osiągające maksymalnie 53,57% trafności (przy baseline 52,06%), prowadzą do istotnych wniosków:

- Uzyskany wynik nieznacznie powyżej linii bazowej (Edge = 1,51%) potwierdza, że para EUR/USD jest najbardziej efektywnym i zaszumionym instrumentem na rynku Forex.

Jest ona zbyt silnie determinowana przez potężne czynniki egzogeniczne – takie jak polityka amerykańskiego FED, dane makroekonomiczne z USA czy globalne przepływy kapitału – aby mimika samego prezydenta EBC mogła stanowić samodzielny i silny sygnał kierunkowy.

- Trendem widocznym w wynikach jest fakt, że jedynie scenariusze wykorzystujące redukcję wymiarowości (PCA) lub selekcję cech (Lasso Selection) były w stanie przekroczyć próg 52,06% trafności. Modele oparte na surowych danych (Raw Data - 108 AU) uzyskały wyniki znacznie poniżej baseline (nawet 39,29%).
- Wyniki te, zestawione z bardzo wysoką trafnością dla zmienności (Volatility) opisaną w poprzednich rozdziałach, sugerują, że mikroekspresje niosą informację o stresie i niepewności osoby mówiącej, co rynek dyskontuje jako ryzyko (wzrost zmienności), ale niekoniecznie jako jednoznaczny impuls do aprecjacji lub deprecjacji waluty. Inwestorzy mogą reagować na stres przedstawiciela banku centralnego w sposób niespójny, co znosi przewagę kierunkową modelu w tym konkretnym horyzoncie.
- Drastyczny spadek trafności z nierealnych 99% do 53,57% jest ostatecznym dowodem na poprawność przeprowadzonej procedury walidacyjnej. Z naukowego punktu widzenia, wypracowanie nawet niewielkiej przewagi ($\text{Edge} > 0$) nad rynkiem EUR/USD w oparciu o dane alternatywne, przy zachowaniu pełnej separacji czasowej, jest wynikiem wartościowym i potwierdza istnienie behawioralnego komponentu w wycenie walut.

5.3. Analiza Golden Features i pewności predykcyjnej (Probability Calibration)

W celu identyfikacji najbardziej efektywnych wskaźników predykcyjnych, przeprowadzono wielowymiarową analizę sygnałów kierunkowych dla trzech kluczowych grup: indeksu NEER, poszczególnych par walutowych oraz wskaźników zmienności FX. Proces ten pozwolił na wyłonienie optymalnych kombinacji horyzontu czasowego, scenariusza przetwarzania cech oraz klasy modelu, które generują najwyższą przewagę nad rynkiem.

W grupie wskaźników NEER zaobserwowano najwyższą w całym badaniu przewagę nad rynkiem, sięgającą 21,43%. Najsilniejsze sygnały wypracowano dla zmienności (`neer_volatility_m`) oraz momentum (`neer_momentum_m`) w krótkim horyzoncie $t + 1$. Dominującym scenariuszem okazała się Lasso Selection w połączeniu z modelem GradientBoost. Sugeruje to, że dla ogólnej siły euro kluczowa jest rygorystyczna selekcja cech behawioralnych, która eliminuje szum i pozwala modelowi skupić się na czystym sygnale stresu. Interesujące wyniki przyniósł horyzont $t + 6$ dla momentum (Edge: 17,86%), gdzie model RandomForest wykazał zdolność do zapowiadania długoterminowych zmian trendu.

W przypadku predykcji kierunku konkretnych walut, model wykazał najwyższą skuteczność dla walut uznawanych za bezpieczne oraz pary o wysokiej płynności. W przypadku EUR/CHF najwyższy Edge (17,86%) uzyskano w horyzoncie $t + 1$ przy zastosowaniu dekompozycji PCA i modelu GradientBoost. Potwierdza to tezę, że kurs franka szwajcarskiego reaguje na nastrój niemal natychmiastowo. Para EUR/GBP wykazała dużą stabilność predykcyjną w horyzoncie $t + 2$ (Edge: 14,29%) przy wykorzystaniu surowych danych i modelu Ridge. Wskazuje to na bardziej liniowy charakter relacji behawioralnych dla funta brytyjskiego. Model skutecznie

zidentyfikował kierunek jena (EUR/JPY) w horyzoncie $t + 1$ (Edge: 7,14%), opierając się na modelu SVC i surowych danych, co sugeruje wysoką wrażliwość tej pary na konkretne, pojedyncze jednostki akcji (AU).

Analiza wskaźników FX odnotowała swoją największą przewagę nad rynkiem (21,43%) w horyzoncie $t + 2$ w przypadku zmienności minimalnej ($fx_min_vol_m$). Najskuteczniejszą kombinacją okazał się model GradientBoost trenowany na „Golden Features” (najbardziej wpływowych jednostek behawioralnych dla danego zestawu). Dowodzi to, że okresy stabilizacji rynku są zapowiadane przez specyficzne, powtarzalne klastry mimiczne. Dla zmienności szczytowej ($fx_peak_vol_m$) uzyskano istotny Edge (14,29%) przy wykorzystaniu Lasso Selection i sieci neuronowej (NeuralNet). Sugeruje to, że nagłe skoki zmienności są zakodowane w złożonych, nieliniowych wzorcach mimicznych, które najlepiej interpretują modele głębokie. W horyzoncie $t + 6$ model LogReg oparty na PCA utrzymał przewagę na poziomie 10,71%, co pozwala na wykorzystanie systemu jako narzędzia wczesnego ostrzegania przed turbulencjami rynkowymi z półtoramiesięcznym wyprzedzeniem.

Przeprowadzona analiza dowodzi, że nie istnieje jeden uniwersalny model dla wszystkich kategorii rynkowych. Podczas gdy wskaźniki NEER i zmienność FX najlepiej reagują na selekcję Lasso, kursy walutowe (szczególnie CHF) wymagają dekompozycji PCA. Fakt, że w każdej z grup udało się wypracować istotną przewagę (Edge > 10%), ostatecznie potwierdza użyteczność wizyjnej analizy mikrogestów jako narzędzia wspomagania decyzji na rynku Forex. W celu oceny użyteczności model SVC dla wskaźnika $neer_volatility_m$ (horyzont $t + 2$, zestaw „Golden Features”) przeprowadzono analizę trafności w zależności od poziomu ufności modelu.

Analiza pewności predykcyjnej ukazała zjawisko wysokiej korelacji między pewnością matematyczną modelu a realną skutecznością. W sytuacjach, w których model wykazywał najwyższą pewność predykcyjną (przedział 90-100% prawdopodobieństwa), jego realna trafność wyniosła 100.0%. Co istotne, model wykazał również bardzo wysoką skuteczność (71.43%) w przedziale 50-60%, co sugeruje separację sygnału od szumu już przy umiarkowanych poziomach ufności.

6. Finalny model

6.1. Operacjonalizacja i wdrożenie

Ostatnim etapem projektu było przekształcenie wyników badań w gotowe narzędzie analityczne typu Decision Support System (DSS). Wybrane modele o najwyższym wskaźniku Global Score zostały wytrenowane na pełnym zbiorze danych i zarchiwizowane w formacie binarnym (.pkl).

Zestaw modeli:

- Zmienność NEER ($t+1$): Model SVC (Golden Features) – główny wskaźnik ryzyka.
- EUR/GBP ($t+2$): Model Ridge (Dane surowe) – predykcja kierunku trendu.
- EUR/CHF ($t+1$): Model GradientBoost (PCA) – krótkoterminowy radar safe-haven.

- EUR/JPY (t+1): Model SVC (Golden Features).
- EUR/USD (t+1): Model NeuralNet (Lasso).
- EUR/CNY (t+1): Model NeuralNet (PCA).
- Zmienność FX (t+6): Model LogReg (PCA) – długoterminowa prognoza.

6.2. Operacjonalizacja systemu – analiza porównawcza

W pierwszej fazie modele opierały się na bieżącej ekspresji. Diagnostyka wykazała, że modele te w dużej mierze polegały na trendach finansowych, jak i emocjach. Przy wprowadzeniu lagów behawioralnych ($t - 1$, $t - 2$) wzrósł wpływ mimiki na decyzje modelu – z poziomu 50.80% do 84.76% dla *neer_volatility_m*, oraz z poziomu 61.03% do 97.848% dla EUR/USD. Oznacza to, że model z pamięcią niemal całkowicie ignoruje historię cen, a decyzje podejmuje w oparciu o emocje. Jednakże widać także spadek wartości dla pary walutowej EUR/JPY z 96.15% do 55.87%, co może świadczyć o fakcie, iż zachowania nie wpływają tak dokładnie na decyzyjność w przypadku tej waluty. Pary EUR/GBP oraz EUR/CNY oraz FX pozostały bez zmian. Tabela 4 przedstawia wyniki pierwszego modelu (bez pamięci behawioralnej), a Tabela 5 drugiego (z pamięcią behawioralną).

Tabela 4. Finalne prognozy (Model bez z pamięci behawioralnej)

Nazwa	Wskaźnik/Para walutowa	Prognoza	Pewność modelu
NEER	<i>neer_volatility_m</i>	WZROST	50.80%
USD	EURUSD_Level	SPADEK	61.03%
GBP	EURGBP_Level	WZROST	87.64%
CHF	EURCHF_Level	SPADEK	90.44%
JPY	EURJPY_Level	SPADEK	96.15%
CNY	EURCNY_Level	SPADEK	99.56%
FX	<i>fx_volatility_m</i>	WZROST	100.00%

Źródło: Opracowanie własne.

Tabela 5. Finalne prognozy (Model z pamięcią behawioralną)

Nazwa	Wskaźnik/Para walutowa	Prognoza	Pewność modelu
NEER	<i>neer_volatility_m</i>	SPADEK	84.76%
GBP	EUR GBP_Level	WZROST	87.64%
CHF	EURCHF_Level	SPADEK	90.44%
JPY	EURCJPY_Level	SPADEK	55.87%
USD	EURUSD_Level	SPADEK	97.48%
CNY	EURCNY_Level	SPADEK	99.56%
FX	<i>fx_volatility_m</i>	WZROST	100.00%

Źródło: Opracowanie własne.

Modele z lagami udowodniły, że historia nastrojów w niektórych przypadkach jest ważniejsza niż statyczny odczyt. Wpływ mimiki na poziomie 98% dla głównych walut potwierdza, że twarz bankiera centralnego jest niemal kompletnym nośnikiem informacji o przyszłym trendzie. W przypadku NEER kierunek prognozy zmienił się pod wpływem analizy dynamicznej. Uzyskanie dodatniego Edge (przewagi) przy rygorystycznym podziale czasowym wyklucza przypadkowość wyników. System nie zgaduje, a identyfikuje wzorce stresu, które rynek dyskontuje z opóźnieniem. Porównanie dwóch architektur – statycznej i dynamicznej (z lagami) – doprowadziło do wyboru modelu z obydwooma pamięciami dla różnych wskaźników/par walutowych.

Wnioski

Projekt udowadnia, że analiza mikroekspresji prezydentów EBC pozwala na uzyskanie istotnej przewagi informacyjnej w prognozowaniu dynamiki rynku walutowego, nawet przy zastosowaniu rygorystycznych metod walidacji czasowej i eliminacji wycieku danych. Badanie statystycznie potwierdziło, że mikroekspresje prezydentów EBC są wskaźnikami wyprzedzającymi. Testy Grangera wykazały, że tzw. wyciek emocjonalny poprzedza reakcje kursowe o 1 do 6 cykli zdarzeniowych. Model wykazuje znacznie wyższą skuteczność w prognozowaniu niepokoju rynkowego (zmienności) niż precyzyjnych poziomów cenowych. Osiągnięcie przewagi nad rynkiem (Edge) na poziomie 21,43% dla zmienności NEER (model GradientBoost) jest wynikiem wybitnym w skali finansów ilościowych. Najwyższą stabilność wyników uzyskano poprzez zastosowanie metod redukcji wymiarowości (PCA) oraz selekcji Lasso. Potwierdzono, że rynki reagują na specyficzne klastry mimiczne (Golden Features), a nie na całościową aktywność mięśniową twarzy. Najsilniejszy sygnał predykcyjny dla poziomów walut materializuje się w horyzoncie $t + 5$ i $t + 6$ (ok. 1,5 miesiąca), natomiast sygnały dotyczące zmienności są widoczne już w horyzoncie $t + 2$ (ok. 16 dni).

W dobie cyfryzacji finansów, ludzka twarz pozostaje jednym z najpotężniejszych źródeł danych. Dzięki ewolucji od prostych modeli statycznych do systemów z pamięcią behawioralną, powstało narzędzie, które zapowiada stany rynkowe w warunkach pełnego rygoru bezstronności i stanowi pomost między psychologią poznawczą a nowoczesną ekonometrią, oferując precyzyjny wgląd w przyszłą kondycję gospodarki strefy euro.

Uzasadnienie zgodności projektu z wymaganiami zadania

Niniejszy projekt został opracowany zgodnie z wytycznymi.