Polski Instytut Ekonomiczny

GRUDZIEŃ 2020

WARSZAWA

Zwroty z akcji a sentyment w mediach społecznościowych

ISBN 978-83-66698-22-2

Cytowanie:
Batorski, D., Baranowski, M., Bławat, B., Krawiec, J., Młodożeniec, M., Talarek, P. (2020),
Zwroty z akcji a sentyment w mediach społecznościowych, Polski Instytut Ekonomiczny, Warszawa.

Poglądy i analizy prezentowane w publikacji są odzwierciedleniem stanowiska autorów.

Publikacja dotowana w ramach projektu „Algorytmiczny system oceny cech kwantytatywnych i kwalitatywnych emitenta instrumentów rynku kapitałowego" finansowanego w ramach programu Ministra Nauki i Szkolnictwa Wyższego pod nazwą „DIALOG" nr rej 0180/20117

Warszawa, grudzień 2020 r.

Autorzy: Dominik Batorski, Maciej Baranowski, Bogusław Bławat, Jakub Krawiec,
Marek Młodożeniec, Paweł Talarek
Redakcja: Jakub Nowak, Małgorzata Wieteska
Projekt graficzny: Anna Olczak
Współpraca graficzna: Liliana Gałązka, Tomasz Gałązka, Joanna Cisek
Skład i łamanie: Sławomir Jarząbek
Polski Instytut Ekonomiczny
Al. Jerozolimskie 87
02-001 Warszawa
© Copyright by Polski Instytut Ekonomiczny

ISBN 978-83-66698-22-2

Spis treści

Kluczowe liczby 4
Kluczowe wnioski 5
Wprowadzenie 6
Zakres analizowanych danych 8
Kompletność danych i selekcja profili 9
Wstępna eksploracja danych 12
Selekcja i ocena wydźwięku treści 17
Agregacja danych i inżynieria cech 19
Modelowanie 20
Model klasyfikacyjny 23
Konstrukcja ratingu 27
Walidacja metody na danych z giełdy wiedeńskiej 34
Bibliografia 41
Spis infografik, wykresów i tabel 42

Kluczowe liczby

462	spółek notowanych na GPW poddanych analizie.
8 makrosektorów i 83 branże	były reprezentowane przez badane spółki.
956	profili na Facebooku zidentyfikowano dla spółek.
499 050	unikalnych postów pięciu typów (wydarzenie, link, zdjęcie, status, wideo) spółki opublikowały na Facebooku.
83,3 proc.	trafność predykcji zwrotu z inwestycji w akcje przy pomocy danych z mediów społecznościowych, mierzona na grupie testowej.
84,7 proc.	trafność przypisania spółek do właściwej klasy ratingu (dla wersji ratingu składającej się z trzech klas).

Kluczowe wnioski

- Tworzenie unii rynków kapitałowych w Europie jest dla Polski sporym wyzwaniem. Nie tylko ze względu na niski poziom zaufania w ramach kapitału społecznego, ale także z powodu luki w ofercie ratingowej. Sytuację na rynku kapitałowym małych i średnich przedsiębiorstw może poprawić wykorzystanie algorytmicznego systemu oceny atrakcyjności inwestycyjnej emitenta instrumentów finansowych, który powstał na podstawie danych dostępnych w mediach społecznościowych.
- Algorytmiczny system oceny cech ilościowych i jakościowych emitenta instrumentów rynku kapitałowego – jako alternatywna metoda ratingowa – umożliwia skuteczne przewidywanie przyszłego zachowania kursów akcji i stopy zwrotu. Zaś korzystanie z łatwo dostępnych danych – np. z Facebooka – pozwala również oceniać podmioty nienotowane na giełdzie.
- Upowszechnienie tej metody może ogólnie poprawić wiarygodność firm na rynku papierów wartościowych a także ułatwiać do niego dostęp małym i średnim przedsiębiorstwom.
- Niepowodzeniem zakończyły się próby zbudowania ratingu na podstawie przewidywanej miesięcznej stopy zwrotu z inwestycji w akcje danej spółki na bazie danych z mediów społecznościowych. Żaden z testowanych modeli predykcyjnych nie dawał zadowalających wyników.
- Pozytywne rezultaty przyniosło podejście oparte na modelach klasyfikacyjnych. Użycie metody *lasów losowych* pozwoliło osiągnąć 83,3 proc. trafności przewidywania. Na podstawie tego modelu predykcyjnego stworzono 5- i 3-wartościowy rating, osiągając 84,7 proc. poprawności przy klasyfikacji dla 3-wartościowego ratingu.
- Dzięki wykorzystaniu danych z giełdy wiedeńskiej, algorytmiczny system oceny atrakcyjności inwestycyjnej emitenta instrumentów finansowych pomyślnie przeszedł test ogólności zastosowania.
- Wnioski sformułowane w trakcie przeprowadzonych badań stanowią istotne źródło informacji w zakresie wykorzystywania narzędzi mediów społecznościowych przez podmioty publiczne. Może to stanowić punkt wyjścia do upowszechnienia wiedzy w zakresie istotności roli mediów społecznościowych w postrzeganiu spółek publicznych przez inwestorów giełdowych. Powinno to prowadzić do wzrostu aktywności podmiotów publicznych objętych ratingiem, a następnie zainteresowania tymi zagadnieniami innych spółek publicznych, które obecnie wykazują się brakiem takiej aktywności, bądź charakteryzują się jedynie sporadyczną aktywnością w tym zakresie.

Wprowadzenie

Rozwój pozakredytowych form finansowania firm zwiększa innowacyjność i konkurencyjność gospodarek. Jednak w Polsce, w porównaniu z innymi krajami UE, istotnymi barierami są ogólnie niskie zaufanie społeczne i duża luka w ofercie ratingowej. W ramach programu „Dialog" MNiSW zrealizowaliśmy projekt badawczy, którego efektem było **opracowanie algorytmicznego systemu oceny cech kwantytatywnych i kwalitatywnych emitenta instrumentów rynku kapitałowego bazującego na danych z mediów społecznościowych.**

Opracowanie algorytmicznego systemu oceny atrakcyjności inwestycyjnej emitenta jest jednym z kryteriów wykreowania wiarygodnego rynku dłużnych papierów wartościowych i ułatwienia dostępu do niego małym i średnim przedsiębiorstwom. Na podstawie obiecujących wyników prac naukowych dotyczących wykorzystania dużych danych dotyczących wydźwięku informacji medialnych i komentarzy użytkowników do przewidywania zachowań i wyników na rynkach finansowych **podjęliśmy próbę zdefiniowania nowej metodyki ratingu przy wykorzystaniu danych z mediów społecznościowych.** Założeniem było zbudowanie modelu, który pozwoli zaklasyfikować również firmy, które jeszcze nie są notowane na Giełdzie Papierów Wartościowych (GPW) w Warszawie, ale są obecne na Facebooku.

Pośrednim celem badań była **weryfikacja hipotezy o możliwości predykcji stopy zwrotu z inwestycji w akcje spółek publicznych za pomocą danych uzyskanych z mediów społecznościowych.** Wypracowane rozwiązanie było walidowane danymi z kolejnych okresów oraz danymi z innego rynku, tj. giełdy papierów wartościowych w Wiedniu (Wiener Börse AG).

Jednym ze sposobów zwiększających innowacyjność i konkurencyjność gospodarek jest tworzenie zachęt do inwestowania i mechanizmów efektywnej alokacji kapitału – m.in. rozwój pozakredytowych form finansowania firm (*Strategia Rozwoju Rynku Kapitałowego*, 2019). Komisja Europejska wskazuje, że relatywnie duże uzależnienie podmiotów gospodarczych od źródeł bankowych jest jednym z czynników ograniczających ich rozwój. W ramach tworzenia unii rynków kapitałowych, KE będzie dążyć do różnicowania kanałów finansowania firm *(Plan działania na rzecz tworzenia unii rynków kapitałowych*, 2015). Polska boryka się z niskim poziomem zaufania i dużą luką w ofercie ratingowej, szczególnie w odniesieniu do przedsiębiorstw. Jako kraj będzie zatem miała znacznie poważniejsze wyzwania niż większość krajów UE, w których rola rynków kapitałowych jest dziś zdecydowanie większa. Dostępne usługi ratingowe są oferowane zwykle przez międzynarodowe korporacje nieposiadające konkurencji na polskim rynku. Nie uwzględniają zatem w optymalny sposób specyfiki związanej z ich przynależnością do określonych sektorów czy branż.

W raporcie przedstawiliśmy etapy prac prowadzących do stworzenia modeli ratingowych, zaczynając od pokazania źródeł danych oraz eksploracji danych i inżynierii cech. Pokazaliśmy różne podejścia do konstrukcji ratingu oraz wyniki analiz w przypadku wykorzystania modeli klasyfikacyjnych. Opisaliśmy testowanie modeli, w szczególności dotyczące analiz stabilności ratingu w czasie oraz jego zdolności predykcji wyników giełdowych. Dodatkowe testy przeprowadziliśmy przy użyciu danych z giełdy wiedeńskiej.

Użycie metody lasów losowych pozwoliło osiągnąć 84,7 proc. trafności w klasyfikacji

dla 3-wartościowego ratingu. Potwierdziliśmy też możliwości jego wykorzystania jako instrumentu oceny atrakcyjności inwestycyjnej, umożliwiającego skuteczne przewidywanie przyszłego zachowania kursów akcji i stopy zwrotu. Zastosowanie opracowanego systemu alternatywnej oceny stopnia atrakcyjności inwestycyjnej emitenta może przyczynić się do wykreowania wiarygodnego rynku dłużnych papierów wartościowych, ułatwiając dostęp do finansowania małym i średnim przedsiębiorstwom.

Zakres analizowanych danych

Realizacja projektu wymagała łącznego wykorzystania dwóch typów danych. Z jednej strony sięgnęliśmy po dane dotyczące wyników notowań spółek na Giełdzie Papierów Wartościowych, a z drugiej – po dane z mediów społecznościowych, które dotyczą tych spółek. Wybranie jako przedmiotu analizy spółek publicznych wynika z większej dostępności informacji na ich temat. Zgromadzone dane posłużyły jako zbiór testowy do wytworzenia odpowiedniej metody oceny atrakcyjności inwestycyjnej spółek, którą w przyszłości będzie można zastosować również do ratingu firm nienotowanych na giełdzie.

Dane giełdowe obejmowały notowania akcji wszystkich emitentów obecnych na rynku głównym oraz NewConnect Giełdy Papierów Wartościowych w Warszawie. Tworzyły je statystyki dzienne zawierające informacje o kursach, wolumenie transakcji, liczbie akcji znajdujących się w wolnym obrocie (*free float*), wypłacanej dywidendzie czy emisji prawa poboru, a także strukturze akcjonariatu. Analiza obejmowała dane za lata 2018-2019 oraz za wcześniejsze okresy. Dane te stały się źródłem dla utworzenia szeregu zmiennych pomocniczych, zwłaszcza tych o charakterze analizy technicznej (opartej na wynikach badań historycznych przebiegów szeregów czasowych i prognozowaniu ich przyszłych zachowań). Z przeprowadzonych obserwacji wynika, że duża część inwestorów obserwuje wskaźniki analizy technicznej, dlatego notowania spółek nie zawsze odzwierciedlają ich wskaźniki fundamentalne, jak np. zysk czy poziom zadłużenia. Mając to na uwadze dołączyliśmy do zbioru zmiennych wejściowych modelu szereg zmiennych najczęściej używanych przez inwestorów, wyliczonych na podstawie analizy technicznej.

Przyjęliśmy, że sygnałem wejściowym w procesie budowy ratingu będą **dane dotyczące aktywności na portalu Facebook (FB)** na profilach należących do spółek notowanych na GPW oraz NC. Decyzja ta wynikała z analizy dotyczącej korzystania z mediów społecznościowych w Polsce. Według wyników badań Gemius/PBI Facebook jest najpopularniejszą platformą mediów społecznościowych w Polsce. Co miesiąc odwiedza ją ponad 21 mln użytkowników, z czego prawie 18 mln ma konto w serwisie. Dodatkowo za wyborem tego portalu przemawia to, że swoje profile w tym medium na całym świecie ma 65 mln firm (Interaktywnie.com, 2019). Inne platformy, jak choćby Twitter czy LinkedIn, są w Polsce zdecydowanie mniej popularne i używane przez dość wąskie grupy odbiorców. Również bardziej specjalistyczne fora przeznaczone dla inwestorów są używane przez stosunkowo niewielką liczbę osób.

Analiza danych z mediów społecznościowych może obejmować różne rodzaje informacji. Wyodrębniliśmy dwa rodzaje treści związanych z danym podmiotem. Po pierwsze są to treści publikowane w kanałach należących do tego podmiotu, a także reakcje użytkowników względem nich. Po drugie mogą to być wszystkie wzmianki publikowane w serwisach społecznościowych dotyczące danego podmiotu. W przypadku serwisu Facebook dane pierwszego typu są łatwo dostępne, zaś drugiego znacznie mniej, ze względu na brak możliwości monitorowania treści na profilach prywatnych. Choć narzędzia *social monitoring* umożliwiają monitorowanie mediów społecznościowych, to jednak w praktyce pozwalają zbierać tylko część treści publikowanych przez użytkowników. Z tego też względu zdecydowano się skupić na pierwszym rodzaju informacji, uwzględniając to co piszą firmy oraz komentarze użytkowników.

Kompletność danych i selekcja profili

Podczas badań potwierdziliśmy obserwacje: nie wszystkie spółki giełdowe posiadają swój profil w serwisie Facebook, a niektóre posiadają takich profili wiele. Mogą być one związane z poszczególnymi markami należącymi do spółki, bądź też być profilami należącymi do spółek zależnych. Dlatego też w niniejszym projekcie jako profile należące do danej spółki zdefiniowaliśmy wszystkie te, które dotyczą danej spółki, jednej z jej spółek zależnych lub jednej z marek należących do spółki lub jej spółek składowych.

Łącznie przeanalizowaliśmy dane dla **462 spółek notowanych na GPW przypisanych do 83 branż w 8 sektorach.** Rozkład spółek w poszczególnych sektorach oraz liczbę spółek, które nie są obecne na Facebooku przedstawiliśmy na infografice 1. Udział spółek z danego sektora prowadzących przynajmniej jeden profil jest różny w poszczególnych branżach – najniższy w tych, w których notowano najwięcej spółek.

Największa grupa spółek prowadzi 1-2 profile facebookowe. Drugą grupą pod względem liczebności są spółki nieposiadające żadnego profilu na analizowanym portalu społecznościowym.

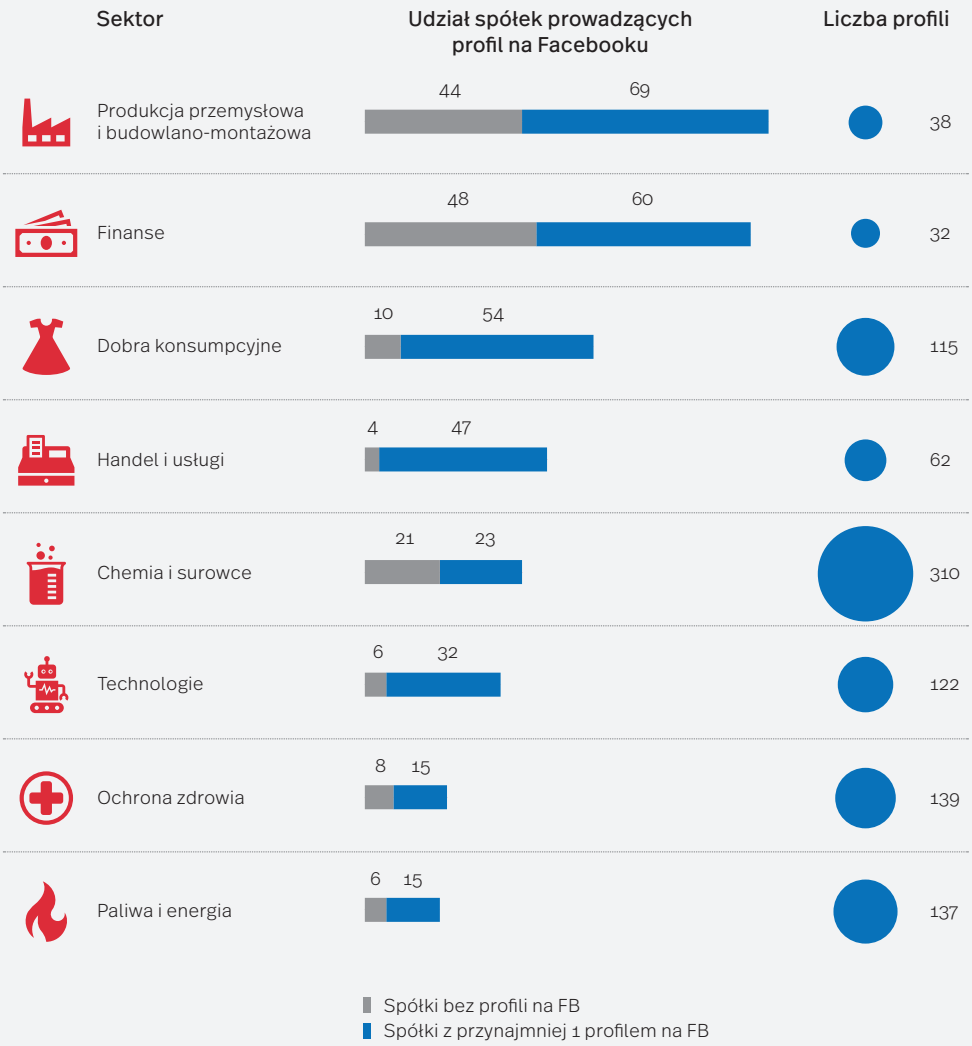
Przy wykorzystaniu specjalnych narzędzi pobraliśmy z API serwisu Facebook dane trzech typów: statystyki dzienne profili, dane dotyczące publikowanych postów, dane dotyczące publikowanych komentarzy.

Zgromadzone dane obejmowały okres 01.2018-12.2019 r. Proces zbierania danych przebiegał w dwóch etapach. Na początku 2019 r. pobraliśmy dane historyczne za cały 2018 r., a dalsze dane gromadziliśmy na bieżąco w ramach procesu comiesięcznej aktualizacji danych. Ze względu na rozpoczęcie prac nad ratingiem, przyjęliśmy że wykorzystane zostaną dane za okres 01.2018-04.2019 r. włącznie. Dane zbierane w kolejnych okresach wykorzystaliśmy do badania stabilności ratingu w czasie.

Rodzaje zgromadzonych danych:

1. **Statystyki dzienne profili,** które agregują wszystkie aktywności na profilach do poziomu poszczególnych dni. Zawierają następujące informacje: liczba postów dodanych na profilu, zaangażowanie fanów w formie wskaźników Interactivity Index i People Talking About That, liczba reakcji zebranych przez posty opublikowane na profilu (komentarze, udostępnienia oraz reakcje typu: „Lubię to", „Super", „Ha-ha", „Smutek", „Złość" i „Zaskoczenie"), dane o liczbie reakcji i komentarzy dodanych przez stronę wraz z informacją o responsywności danego profilu względem aktywności użytkowników. Dodatkowo dane dzienne zawierają również informacje o liczbie fanów profilu wraz z ich dzienną zmianą.
2. **Dane dotyczące publikowanych postów,** które zawierają szczegółowe informacje o postach opublikowanych na rozpatrywanych profilach. Podstawowego podziału jakościowego można dokonać ze względu na autora posta: na posty publikowane przez profile, na których się znalazły oraz posty publikowane przez użytkowników. Dane w tym przypadku zawierały następujące informacje: data i godzina publikacji treści, typ posta (materiał wideo, zdjęcie lub

↘ Infografika 1. **Analizowane spółki i ich aktywność na Facebooku**



Źródło: opracowanie własne PIE.

tekst), treść tekstowa, liczba zebranych reakcji (komentarze, udostępnienia oraz reakcje typu: „Lubię to”, „Super”, „Ha ha”, „Smutek”, „Złość” i „Zaskoczenie”), zaangażowanie w formie wskaźnika Interactivity Index oraz link URL do posta. Dodatkowo, jeśli był to wpis użytkownika, została również zawarta informacja czy post ten uzyskał odpowiedź administratora strony.

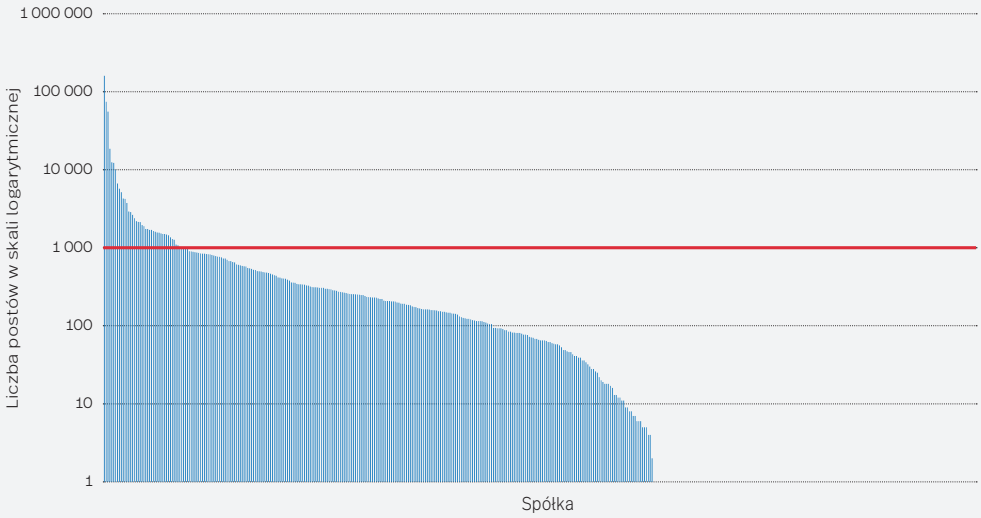
3. **Dane dotyczące publikowanych komentarzy** zawierają szczegółowe informacje o komentarzach do postów, które opublikowano na rozpatrywanych profilach. Podobnie jak w przypadku postów, podstawowego podziału jakościowego komentarzy można dokonać ze względu na autora: na komentarze opublikowane przez właścicieli profili, na których się znalazły oraz komentarze publikowane przez użytkowników. Dane dotyczące komentarzy zawierają następujące informacje: data i godzina publikacji komentarza, treść komentarza, link URL oraz nazwa profilu, na którym opublikowano komentarz.

Wstępna eksploracja danych

Przed przystąpieniem do dalszych prac wykonaliśmy wstępną analizę eksploracyjną. Jej znaczenie jest związane z ustaleniem specyfiki danych, w tym w szczególności takich, które są istotne na etapie modelowania. Przedstawiamy jedynie kluczowe ustalenia mające konsekwencje dla dalszych prac i podejmowanych decyzji.

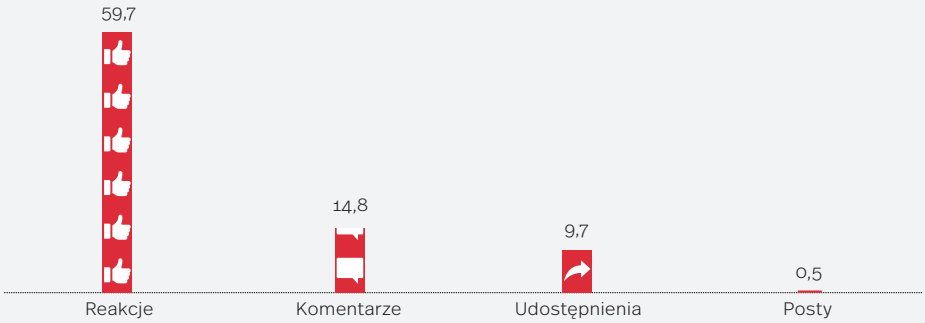
Pierwszym wnioskiem, na który należy zwrócić uwagę, jest bardzo nierównomierny rozkład aktywności między profilami należącymi do poszczególnych spółek. Na wykresie 1 przedstawiliśmy rozkład liczby postów opublikowanych w okresie nieco ponad roku przez wszystkie profile FB przypisane do spółek z głównego rynku GPW, uszeregowane od spółek, które publikowały najwięcej do spółek, które nie publikowały wcale. Każdy słupek obrazuje łączną liczbę postów opublikowaną przez daną spółkę. Poziomą linią zaznaczyliśmy średni wynoszącą 1053 postów. Wyraźnie widać, że kilka spółek publikuje dużo więcej postów niż reszta. Najwięcej postów w rozpatrywanym okresie – 159 823 – opublikowały profile należące do spółki Wirtualna Polska Holding Spółka Akcyjna (WPL). Tylko 40 spółek (8,66 proc. wszystkich) opublikowało więcej postów niż wynosi średnia, zaś aż 170 spółek (36,8 proc.) nie opublikowało żadnego posta (uwzględniliśmy tu spółki, które nie prowadzą żadnego profilu na Facebooku).

↘ Wykres 1. **Rozkład liczby postów opublikowanych przez wszystkie profile na Facebooku przypisane do spółek z głównego parkietu GPW, uszeregowane od spółek, które publikowały najwięcej do spółek, które nie publikowały wcale**

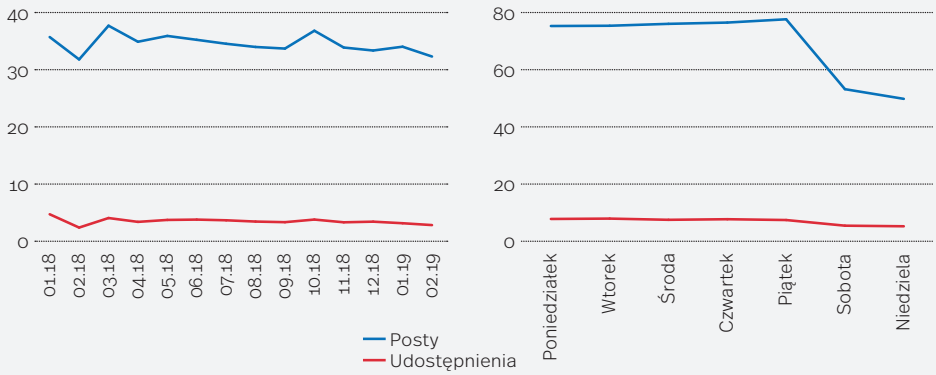


Źródło: opracowanie własne PIE.

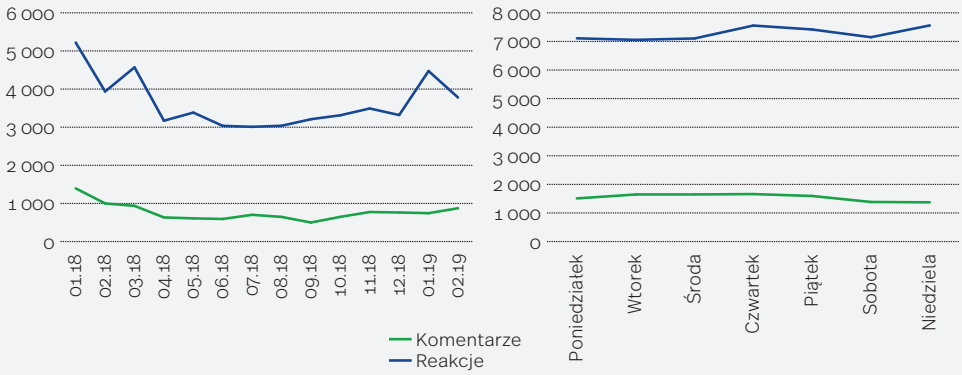
↘ Infografika 2. Rozkład aktywności internautów na posty firm ma Facebooku według miesięcy i dni tygodnia (w mln)



Rozkład liczby postów i udostępnień w kolejnych miesiącach i według dni tygodnia (w tys.)



Rozkład liczby reakcji i komentarzy w kolejnych miesiącach i według dni tygodnia (w tys.)



Źródło: opracowanie własne PIE.

Analogiczną analizę wykonaliśmy z podziałem na poszczególne profile. Rozkład nadal jest bardzo nierównomierny, ale w wyraźnie mniejszym stopniu niż to miało miejsce w przypadku całych spółek. Najwięcej postów w rozpatrywanym okresie (21 800) było publikowanych na profilu Wirtualnej Polski. Średnia dla całej populacji wynosi 514. Powyżej średniej publikowało 130 profili, co daje 13,84 proc. wszystkich rozpatrywanych profili, a żadnych postów nie opublikowało 104 profili (11,08 proc. wszystkich profili).

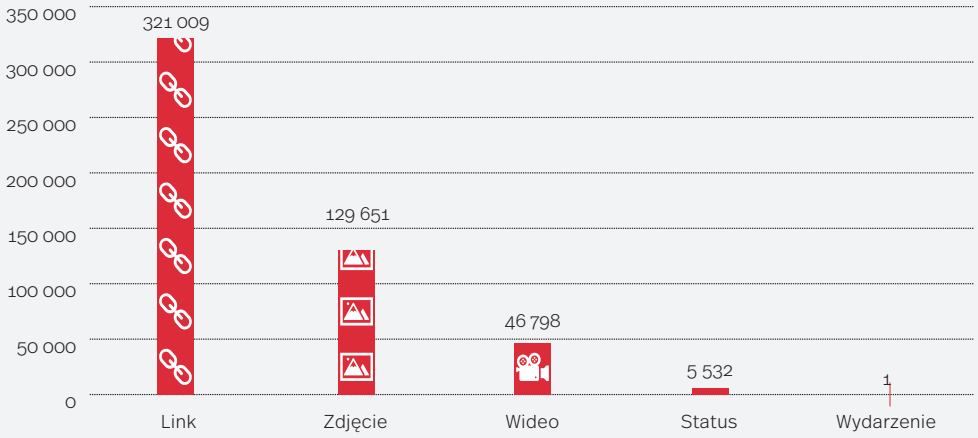
Rozkład sumy liczby postów oraz udostępnień w kolejnych miesiącach zmienia się w niewielkim stopniu, co jest jednym z czynników wskazujących na stabilność danych w czasie. Ma to znaczenie w przypadku budowania modeli predykcyjnych na podstawie danych historycznych. Najniższe wartości obie cechy przyjmują w lutym 2018 r. i 2019 r. (infografika 2). Należy przypuszczać, że wynika to z długości tego miesiąca.

Po rozbiciu danych na poszczególne dni tygodnia wyraźnie widać, że zarówno łączna liczba postów publikowanych przez wszystkie profile jak i udostępnień użytkowników Facebooka maleje w dni wolne od pracy (weekendowe). Dla liczby postów można także zaobserwować lekką tendencję wzrostową w dni robocze.

Rozbicie danych na poszczególne dni tygodnia pozwala zauważyć dość nieregularny podział reakcji z dwoma szczytami: w czwartek i niedzielę.

Nieco inną charakterystykę przyjmuje rozkład sumy reakcji oraz komentarzy w poszczególnych miesiącach. Załamanie pośrodku wykresu prawdopodobnie można tłumaczyć sezonowością aktywności użytkowników Facebooka.

↘ Wykres 2. **Rozkład liczby postów według poszczególnych rodzajów postów możliwych do opublikowania na Facebooku**



Źródło: opracowanie własne PIE.

Dla 462 badanych spółek zidentyfikowano 956 profili na Facebooku. W analizowanym okresie 839 z nich (88 proc.) opublikowało przynajmniej 1 post na Facebooku. Łącznie profile te

opublikowały 499 050 unikalnych postów pięciu typów: wydarzenie, link, zdjęcie, status, wideo. Na wykresie 2 przedstawiliśmy rozkład liczby postów poszczególnych rodzajów. Wśród wszystkich opublikowanych przez profile spółek notowanych na głównym parkiecie GPW, najwięcej postów zawierało linki (321 tys.), druga w kolejności była liczba zdjęć (130 tys.).

Oprócz samych postów elementami treści dostarczającymi istotnych danych są reakcje względem tych postów. Użytkownicy mogą je komentować, zostawiać tzw. reakcje, tj. polubienia i inne oceny jak „Super", „Ha ha", „Przykro mi". Mogą również udostępniać dalej posty publikowane przez innych. Średnią liczbę aktywności użytkowników, które zostały pozostawione pod określonymi typami postów opublikowanych na profilach spółek, prezentujemy na infografice 3. W przypadku wszystkich rodzajów postów najczęstszym typem aktywności były reakcje, a drugim w kolejności – komentarze. Warto zwrócić uwagę na niewielką różnicę między średnią liczbą komentarzy i udostępnień w przypadku treści wideo, które były najchętniej udostępnianymi treściami przez użytkowników.

Analiza liczby komentarzy w zależności od typu posta, wykazała że najwyższą średnią liczbę komentarzy uzyskały posty zawierające treści wideo (średnio ponad 48 komentarzy). Drugimi najchętniej komentowanymi treściami były statusy (średnio 37), na trzecim miejscu znalazły się zdjęcia (29), a na czwartym posty zawierające linki (ponad 26).

Jedną z istotnych statystyk postów publikowanych w mediach społecznościowych jest liczba ich udostępnień. Ma ona duży wpływ na to, jak wiele osób zetknie się z daną treścią w wyniku polecenia przez znajomych (tzw. zasięg wirusowy). Najwięcej wszystkich udostępnień wygenerowały posty zawierające linki – niemal 6 mln pojedynczych udostępnień. Na kolejnych miejscach w zestawieniu znalazły się posty typu wideo (ponad 2 mln), posty ze zdjęciami (1,77 mln) oraz statusy (prawie 15 tys.).

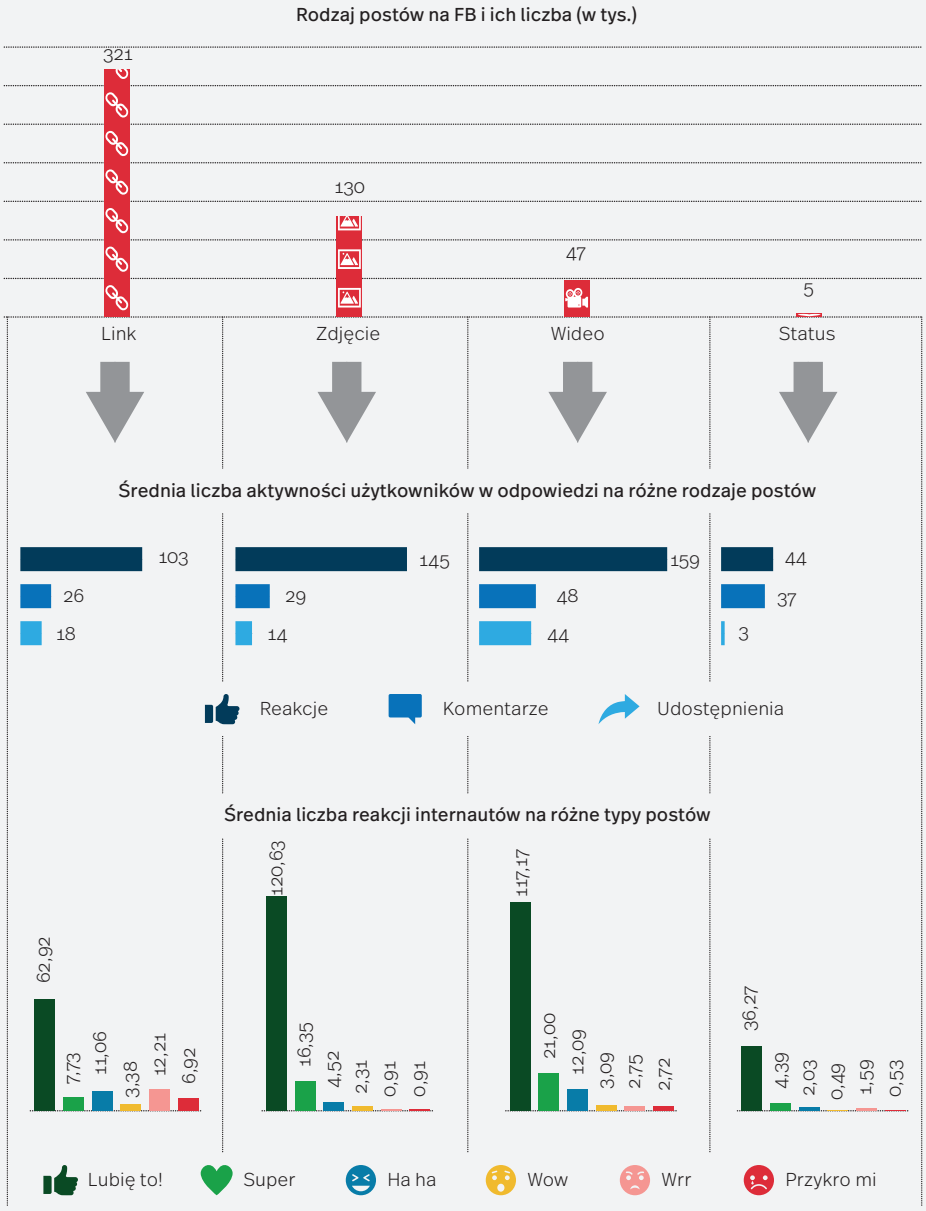
Najwyższą średnią liczbę udostępnień uzyskały posty typu wideo (44 udostępnienia na jeden post), drugie w kolejności były posty z linkami (18), trzecie zdjęcia (14), a czwarte statusy (2,7).

Oprócz komentarzy i dalszego udostępniania postów, użytkownicy Facebooka mogą reagować na oglądaną treść również w inny sposób. Oprócz od dawna dostępnej możliwości polubienia danego posta lub komentarza, dostępnych jest również pięć innych reakcji: „Ha ha", „Przykro mi", „Super", „Wow", „Wrr".

Najwięcej wszystkich reakcji wygenerowały posty zawierające linki – ponad 33 mln. Dalej w kolejności były zdjęcia (prawie 19 mln reakcji), wideo (7,4 mln) i statusy (246 tys.). Średnia liczba wszystkich reakcji, które zebrały posty danego typu, jest dość zróżnicowana. Najwięcej reakcji na jeden post zebrały treści typu wideo – średnio 158,5 reakcji. Drugie w kolejności były zdjęcia (145,4), trzecie – posty zawierające linki do zewnętrznych stron (103,4), a czwarte – kolejne posty typu status (44,5).

W przypadku wszystkich rodzajów postów najczęstszym typem reakcji były polubienia – najwięcej z nich otrzymywały zdjęcia oraz treści wideo. W przypadku zdjęć, statusów oraz treści wideo drugą najczęściej pojawiającą się reakcją była „Super". Z kolei w przypadku postów zawierających link do zewnętrznego źródła, kolejnymi najczęstszymi reakcjami były „Wrr" (symbolizująca złość) oraz „Ha Ha". Posty z linkami zawierały również średnio więcej reakcji – „Przykro mi" oraz „Wow" – niż pozostałe rodzajów postów.

↘ Infografika 3. Rodzaje postów zamieszczanych przez firmy na Facebooku i reakcje internautów



Uwaga: jedna z firm zamieściła również post typu wydarzenie, nieujęty w infografice.
Źródło: opracowanie własne PIE.

Selekcja i ocena wydźwięku treści

Zarówno w przypadku rynku głównego GPW, jak i NewConnect, ocena wydźwięku rozpoczynała się od selekcji treści istotnych. W następnych krokach prowadziliśmy agregację i standaryzację zmiennych uzyskanych w wyniku oceny wydźwięku, jako element inżynierii cech, które w dalszej kolejności wykorzystywaliśmy do budowania modelu.

Jak wspomnieliśmy wcześniej, nie wszystkie treści publikowane w mediach społecznościowych na profilach spółek oraz w komentarzach do nich dotyczą bezpośrednio tych podmiotów, ich produktów bądź usług. Dlatego też, aby wnioskować o kondycji danego podmiotu, istotne jest wyselekcjonowanie tych treści z mediów społecznościowych, które faktycznie się do danego podmiotu odnoszą.

Za istotne uznaliśmy wszystkie posty, które zawierały nazwę spółki, jej profilu Facebook lub przynajmniej jednej z marek do niej należących, a także komentarze do nich. Taka procedura wyboru została zastosowana do spółek notowanych na rynku głównym GPW i NewConnect oraz ich spółek zależnych. Udział postów istotnych we wszystkich postach opublikowanych przez profile na poszczególnych rynkach wynosił prawie 20 proc.

Przed przystąpieniem do analizy wydźwięku sprawdzaliśmy rozkład języków w jakich napisano posty. Zdecydowaną większość (88 proc.) stanowiły treści w języku polskim, nieco ponad 6 proc. w języku angielskim, a resztę – treści w innych językach.

Podobnie przedstawia się rozkład dla istotnych treści związanych ze spółkami notowanymi na NewConnect, choć w ich przypadku praktycznie nie ma innych języków poza polskim, stanowiącym ponad 91 proc. oraz angielskim (ponad 8 proc.). Na podstawie przedstawionych wyników podjęliśmy decyzję o analizowaniu wydźwięku wypowiedzi nie tylko dla języka polskiego, ale również dla angielskiego.

Każdy z istotnych postów oraz komentarzy do tych postów poddaliśmy analizie za pomocą modelu do wykrywania języka danego tekstu. Następnie za pomocą jednego z dwóch modeli klasyfikacyjnych (osobno dla polskiego i angielskiego) przydzielaliśmy tekst do jednej z trzech klas wydźwięku (sentymentu):

- neutralny – post niezawierający jednoznacznie pozytywnego lub negatywnego wydźwięku lub zawierający zarówno pozytywny, jak i negatywny wydźwięk o podobnym natężeniu;
- pozytywny – post zawierający jednoznacznie pozytywny wydźwięk;
- negatywny – post zawierający jednoznacznie negatywny wydźwięk.

Proces działania modelu do oceny wydźwięku rozpoczęliśmy od wstępnego przetworzenia (*preprocessingu*) analizowanego tekstu. W szczególności przeprowadzaliśmy tokenizację, która polegała na podziale tekstu na słowa i inne znaczące symbole, np. znaki przestankowe i emoji. Następnie przeprocesowany tekst wykorzystywaliśmy do stworzenia reprezentacji wektorowej. Teksty o podobnym znaczeniu powinny mieć podobną reprezentację wektorową. Algorytmy tworzące te wektory, czyli tzw. *embeddingi*, są tworzone na podstawie baz tekstowych, by mogły nauczyć się, które słowa, wyrażenia czy zdania mają podobne znaczenie. Do analizowania tekstów z mediów społecznościowych utworzyliśmy nową bazę wykorzystującą algorytmy *embedingowe* typu *fastText* stworzone przez Facebook AI Research (FAIR) (https://fasttext.cc/, 2020) i dostosowane do specyfiki tekstów z mediów

społecznościowych przy wykorzystaniu ogromnych zbiorów danych z Facebooka i Instagrama. Jest to o tyle istotne, że modele do tworzenia *embedingów* są zazwyczaj uczone na danych z tekstów formalnych, albo z Wikipedii – czyli takich, które nie przypominają tekstów pisanych w *social* mediach.

W dalszej kolejności otrzymane wektory wykorzystaliśmy jako warstwę wejściową do sieci neuronowej. W zależności od języka użyte sieci mają różne architektury:

- dla języka polskiego użyliśmy BiGRU (Gated Recurrent Unit) (Cho i in., 2014);
- dla języka angielskiego użyliśmy modelu opartego na sieci typu MLP (Mulitlayer Perceptron), którego cechą wejściową jest suma wektorów dla słów w zdaniu).

Na wyjściu modelu wygenerowaliśmy prawdopodobieństwo przynależności do każdej z grup sentymentu. Wybieraliśmy najbardziej prawdopodobną grupę i na tej podstawie określaliśmy sentyment.

Agregacja danych i inżynieria cech

Kolejnym krokiem prac było stworzenie cech opisujących spółki na podstawie danych z mediów społecznościowych. Prezentowane wyżej dane nieprzetworzone zawierają statystyki na poziomie poszczególnych jednostek treści – postów i komentarzy. Dla jednych i drugich mamy również statystyki reakcji. Treści te są także powiązane z profilami, na których były publikowane. Tym samym dla każdej spółki, prowadzącej swoje profile na Facebooku, mamy przynajmniej jeden profil oraz różnego rodzaju treści na nim publikowane i ich statystyki. Docelowo podejmowane działania służyły zdefiniowaniu zmiennych, które złożą się na opis spółki.

Ze względu na znaczące zróżnicowanie poszczególnych spółek pod względem aktywności w mediach społecznościowych, konieczne jest skonstruowanie takich miar, które będą miały charakter relatywny. Ponadto powinny w spójny sposób dobrze opisywać zarówno podmioty, które publikują bardzo wiele (np. spółki medialne), jak i takie, których aktywność w mediach społecznościowych jest niewielka. Stąd też zmienne otrzymane w wyniku procesu agregacji opisanego powyżej poddaliśmy następującym przekształceniom:

- wyliczenie udziału poszczególnych reakcji użytkowników pod postami we wszystkich reakcjach;
- standaryzacja zmiennych w dwóch wariantach: względem średniej kwartalnej oraz względem średniej rocznej.

Wykorzystanie zmiennych, które pokazują udziały konkretnych wartości, np. procent reakcji „super" na posty publikowane przez spółkę, pozwala na porównywanie spółek o bardzo różnej liczbie aktywności. Gdybyśmy posługiwali się bezwzględną liczbą reakcji tego typu, to spółki, które publikują więcej lub mają relatywnie więcej fanów byłyby znacząco uprzywilejowane. Podobnie jest ze standaryzacją. Zamiast koncentracji na bezwzględnych wartościach, używamy zmiennych, które pokazują odchylenia od wartości przeciętnych dla danej spółki.

Ostatnim elementem przygotowania cech (zmiennych) opartych na danych z mediów społecznościowych było ustalenie właściwych skal czasowych. Dane dzienne nie nadają się bowiem do przewidywania zjawisk, które zachodzą w dłuższych okresach. Modelowanie długookresowego zwrotu z inwestycji za pomocą danych z mediów społecznościowych wymaga więc uwzględnienia również danych, które są zagregowane w czasie i przyjmują wartości opisujące co się działo w mediach społecznościowych w dłuższym okresie.

W projekcie uwzględniliśmy i przetestowaliśmy różne skale czasowe, agregując dzienne dane z mediów społecznościowych do okresów liczących od kilku dni do trzech miesięcy. W ostatecznych wersjach modeli wykorzystaliśmy zmienne z okresów 21- i 28-dniowych oraz 3-miesięcznych, ze względu na ich wartość predykcyjną i większą stabilność. Dodatkową zaletą przyjęcia takich okresów było rzadsze występowanie zmiennych przyjmujących wartości zerowe, co stanowiło szczególny problem dla firm mniej aktywnych na Facebooku.

Agregację wartości w czasie przeprowadziliśmy na kilka różnych sposobów. Dla każdego z przyjętych okresów wyliczaliśmy wartości minimalne, maksymalne oraz różne rodzaje średnich (między innymi w celu uwzględnienia zależności krzywoliniowych).

Modelowanie

Celem podjęcia prac badawczych było opracowanie algorytmu, który dla określonego z góry przedziału czasu (np. miesiąca) przypisze każdemu z emitentów akcji na polskiej Giełdzie Papierów Wartościowych (dalej używamy określenia „spółka") oznaczenie, korespondujące ze skalą ocen ryzyka kredytowego, np. A, B i C. Algorytmiczny sposób przypisania spółek do tych rozłącznych i wyczerpujących podzbiorów (dalej: klasyfikacja, rating) powinien spełniać następujące warunki:

- oznaczenia są silnie powiązane z wynikami giełdowymi badanych emitentów, spółki z oznaczeniem A osiągają wyraźnie lepsze wyniki giełdowe niż spółki z grupy B, a te z kolei – istotnie wyższe stopy zwrotu niż spółki z grupy C;
- przy opracowywaniu metody klasyfikacji mogą być wykorzystywane informacje finansowe, np. o kursach akcji, ale ostateczna wersja algorytmu ma korzystać tylko z informacji niezwiązanych z wynikami giełdowymi; w opracowanej metodzie korzystaliśmy wyłącznie ze wskaźników aktywności na stronach sieci społecznościowych (konkretnie: na stronach portalu Facebook), które reprezentują spółkę lub marki do niej należące;
- rating spółek ma charakter prospektywny: korzysta z danych na temat przeszłości, ale wykazuje się trafnością w przewidywaniu wyników giełdowych notowanych obecnie.

Założenia przyjęte podczas tworzenia ratingu spółek giełdowych na podstawie ich działalności w mediach społecznościowych są odpowiedzią na **potrzebę stworzenia metody oceny atrakcyjności inwestycyjnej nadającej się do szerszego zastosowania. Metoda ta powinna być oparta na powszechnie dostępnych danych, tak aby możliwe było jej zastosowanie również do podmiotów, które nie są notowane na giełdzie.** Powinna być także powiązana z informacjami, które są użyteczne z punktu widzenia potencjalnego inwestora. Stąd w jej stworzeniu zdecydowaliśmy się na powiązanie ratingu z notowaniami spółek na giełdzie. Zrealizowanie tych założeń wymaga **zbudowania modelu predykcyjnego, który będzie przewidywać stopę zwrotu dla danej spółki na podstawie danych z mediów społecznościowych.**

W modelowaniu wykorzystaliśmy różne podejścia, m. in.:

- modele predykcyjne,
- podejście oparte na wartościach rezydualnych,
- eksperckie modele ratingu,
- modele klasyfikacyjne.

Ostatecznie okazało się, że metodą, która wyróżniała się największą trafnością przewidywania wyników giełdowych jest podejście oparte na modelach klasyfikacyjnych, dlatego też w tym opracowaniu opisaliśmy jedynie tę metodę oraz uzyskane na jej podstawie wyniki.

Budowa dowolnego modelu predykcyjnego wymaga określenia i wyliczenia wartości jednej zmiennej zależnej (inaczej: zmiennej wyjaśnianej) oraz zestawu zmiennych niezależnych (tzw. predyktorów). Poniżej opisaliśmy zmienne, które zostały włączone do modelu w charakterze jednych i drugich.

Zmienną wyjaśnianą w modelu jest miesięczna stopa zwrotu z inwestycji w akcje. Wyrażona jest wartością procentową, opisaną wzorem:

$$stopa\ zwrotu\ =$$

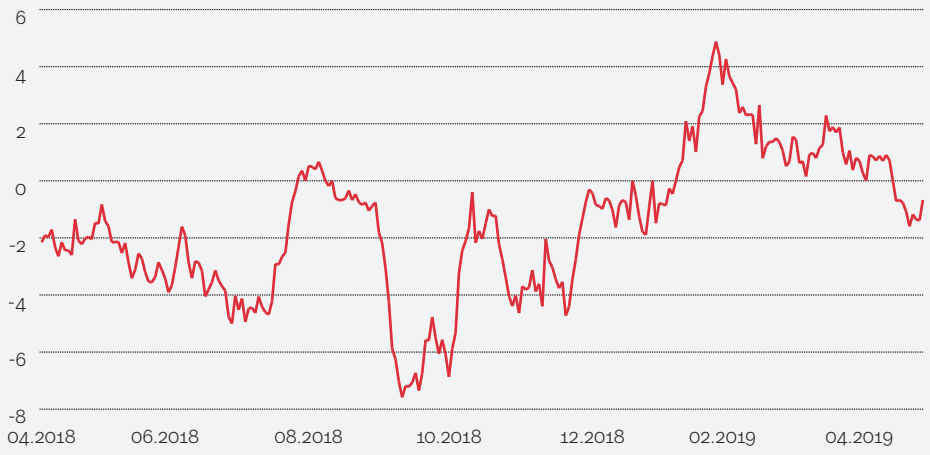
$$= 100\% \cdot \left(\frac{kurs\ zamknięcia}{kurs\ otwarcia} - 1\right)$$

Wartość stopy zwrotu można wyliczyć dla dowolnego okresu. Przy wyliczaniu dziennej stopy zwrotu uwzględnia się kursy otwarcia i zamknięcia z tego samego dnia. W przypadku tygodniowej lub miesięcznej stopy zwrotu, kurs otwarcia rozumiemy jako kurs z pierwszego dnia notowań w tym okresie (zwykle pierwszy dzień roboczy), a kurs zamknięcia – z ostatniego dnia notowań.

W niniejszej analizie podstawą wyliczenia wartości zmiennej zależnej jest **miesięczna stopa zwrotu** z inwestycji w akcje spółki. Wyliczaliśmy ją dla każdego dnia oddzielnie, posługując się informacjami o kursie zamknięcia w danym dniu oraz o kursie otwarcia w dniu odległym o 21 dni roboczych (średnia liczba dni roboczych w miesiącu).

Przewidywanie absolutnej wartości zwrotu z inwestycji nie jest jednak zasadne. Inwestycja w akcje spółki, która notuje miesięczną stopę zwrotu np. rzędu 1 proc., nie zawsze jest tak samo korzystna. W okresie prosperity, gdy inne spółki odnotowują dużo większe wzrosty, zakup akcji takiej spółki byłby nieoptymalną alokacją zasobów. Inwestując te same środki w spółki szybciej rosnące, można by osiągnąć większy zysk. W warunkach dekoniunktury, gdy indeks całej giełdy jest ujemny, może to być jednak optymalny wybór. W takich warunkach nawet inwestycja w akcje spółki notującej niewielkie spadki (niższe niż inni emitenci) mogłaby być stosunkowo korzystna.

↘ Wykres 3. **Przeciętna wartość miesięcznej stopy zwrotu dla spółek z głównego parkietu GPW uwzględnionych w modelowaniu**



Źródło: opracowanie własne PIE.

Dobrze obrazuje to wykres 3, na którym pokazujemy zmiany przeciętnej stopy zwrotu dla analizowanych spółek notowanych na głównym rynku GPW. Gdyby rating przewidywał absolutną stopę zwrotu, proporcje spółek z oznaczeniami A, B i C byłyby bardzo mocno uzależnione od ogólnej kondycji gospodarki. W okresach wzrostu większość spółek miałaby rating A, w okresach spadku – najczęściej kategorię C. Naturalne jest unikanie takich zdarzeń. Poprawnie

stworzony rating powinien odzwierciedlać zasadność inwestycji w akcje spółki w danym czasie i przy zastanej kondycji gospodarki i giełdy.

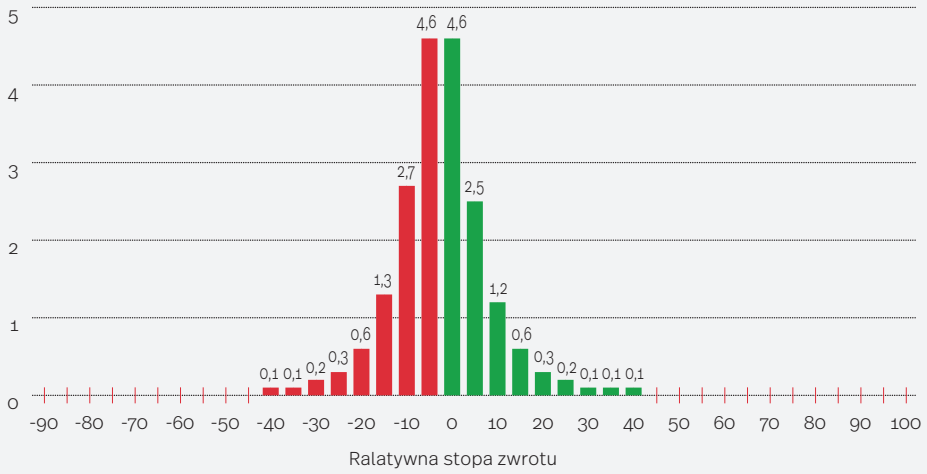
Dlatego **za zmienną przewidywaną obrano nie absolutną stopę zwrotu,** lecz relatywną stopę zwrotu, wyliczoną wg wzoru:

relatywna miesięczna stopa zwrotu = mies.st.zwrotu – mediana mies.st.zwrotu dla danego dnia

W ten sposób opisana wcześniej spółka z miesięczną stopą zwrotu równą 1 proc., w czasie w którym giełda rośnie przeciętnie w tempie 3 proc. na miesiąc, będzie miała relatywną miesięczną stopę zwrotu równą -2 proc. W okresie, gdy ceny akcji na giełdzie spadają o 3 proc. w okresie miesiąca, spółka ta przy tej samej dynamice wzrostu ma relatywną miesięczną stopę zwrotu równą +4 proc.

Na potrzeby algorytmów uczenia maszynowego wykorzystanych do stworzenia ratingu, relatywną miesięczną stopę zwrotu przekodowaliśmy na dwie wartości: z relatywną miesięczną stopą zwrotu powyżej i poniżej zera. Grupy te nazwaliśmy „powyżej mediany rynku" (spółki osiągające przeciętnie lepsze wyniki niż inni emitenci) i „poniżej mediany rynku" (przeciętnie gorsze wyniki). Ze względu na to, że relatywna miesięczna stopa zwrotu równa 0 jest równoważna miesięcznej stopie zwrotu na poziomie mediany, grupy te były bardzo zbliżone pod względem liczebności (49,9 proc. powyżej i 50,1 proc. poniżej mediany) (wykres 4).

↘ Wykres 4. **Rozkład wartości relatywnej miesięcznej stopy zwrotu**



Uwaga: kolorem czerwonym zaznaczyliśmy wartości poniżej mediany rynku, zielonym – powyżej mediany rynku.
Źródło: opracowanie własne PIE.

Jako zmienne niezależne (predyktory) wykorzystywaliśmy zmienne z mediów społecznościowych, których konstrukcję opisaliśmy wcześniej.

Model klasyfikacyjny

W podejściu opartym na klasyfikacji skupiliśmy się na budowie modelu przewidującego czy dana spółka w kolejnym okresie będzie dawała ponadprzeciętne zwroty z inwestycji. Jako zmienną zależną uwzględniliśmy binarną zmienną określająca czy miesięczna stopa zwrotu dla danej spółki była wyższa od przeciętnej, niższa bądź równa przeciętnej.

Do przewidywania wartości zmiennej zależnej wykorzystaliśmy tzw. lasy losowe (*random forest*). Ten model uczenia maszynowego należy do szerokiej klasy modeli wykorzystujących drzewa decyzyjne. Drzewa te powstają w wyniku sukcesywnych podziałów zbioru, wyznaczonych przez wartości zmiennych niezależnych, aż do uzyskania możliwie jak najwyższej zgodności między przypisaniem wskazanym przez drzewo decyzyjne (predykcją), a wartościami zmiennej zależnej. Lasy losowe od innych modeli wykorzystujących drzewa decyzyjne odróżnia fakt wykonywania dużej liczby (zwykle kilkuset) drzew – każdego na podstawie niepełnych danych. Każde drzewo powstaje na podstawie losowo wybranej części przypadków, a kolejne podziały wykorzystują tylko niewielką część obecnych w zbiorze zmiennych. Przewidywanie jest wynikiem agregacji decyzji podjętych przez wszystkie drzewa w „lesie". Decyzje o przynależności konkretnej obserwacji do grupy zapadają w wyniku głosowania większościowego, każde drzewo ma jeden głos. Zaletami lasów losowych, mimo ich dużej złożoności obliczeniowej, są odporność na losowe zakłócenia w danych oraz małe ryzyko przetrenowania modelu (*overfitting*), tj. nadmiernego dostosowania go do danych treningowych, kosztem jakości predykcji w zbiorze testowym.

Uwzględnienie w modelu wszystkich 706 potencjalnych predyktorów obecnych w zbiorze danych rodzi dwie trudności. Przede wszystkim wyliczanie modeli o tej skali trwa bardzo długo. Ponadto zalecane jest włączanie do lasów losowych mniejszej liczby zmiennych uprzednio wyselekcjonowanych. Dlatego model uwzględniający wszystkie predyktory (tzw. model nasycony, *saturated model*) wyliczyliśmy tylko raz, a przy wyliczaniu kolejnych modeli posługiwaliśmy się już skróconą listą 50 predyktorów, uzyskaną z modelu nasyconego. Do selekcji tych 50 predyktorów posłużyliśmy się miarą ważności zmiennych w modelu *Mean Decrease in Gini*. Miara ta, zwana też „ważnością Gini" (*Gini importance*), określa skuteczność zmiennej w podwyższaniu trafności predykcji we wszystkich drzewach tworzących las. Uwzględniona w ostatecznym modelu lista predyktorów powstała przez wybór 50 predyktorów z modelu nasyconego o najwyższej ważności Gini w tym modelu.

Działanie modelu uczenia maszynowego jest determinowane przez tzw. „hiperparametry". Dobór tych parametrów jest najtrudniejszym zadaniem przy konstruowaniu modelu. Ostatecznego wyboru dokonaliśmy wykonując wiele konkurencyjnych modeli i porównując osiągane przez nie wartości wskaźnika trafności przewidywania (*accuracy*). Zarazem jednak godziliśmy się na wartości wskaźników, które nieznacznie obniżały trafność predykcji, ale w zamian oferowały model przejrzystszy, oszczędniejszy, niebędący wynikiem przetrenowania.

W przypadku naszego modelu ostatecznie wybrane hiperparametry były następujące:

- liczba drzew w „lesie" – zwykle im jest większa, tym wyższa trafność przewidywania, jednak od pewnej liczby drzew zysk jest niewielki lub żaden, kompromis między trafnością a czasem obliczeń osiągnęliśmy przy liczbie drzew równej **500**;

- udział przypadków losowo usuwanych ze zbioru przy tworzeniu każdego kolejnego drzewa – zdecydowaliśmy się tu na powszechnie przyjęty poziom parametru: 1/e ~ **36,8 proc.**;
- liczba zmiennych branych pod uwagę przy kolejnych podziałach zbioru – wraz z jej wzrostem, rośnie trafność przewidywania, dlatego zdecydowaliśmy się na uwzględnienie przy każdym podziale wszystkich **50** predyktorów włączonych do modelu;
- maksymalna liczba końcowych węzłów drzewa (*nodes*) oraz minimalna liczebność grupy wyznaczonej przez końcowy węzeł, zwany też „liściem" – zmniejszanie pierwszego lub zwiększanie drugiego z tych parametrów jest równoznaczne z „przycięciem" gałęzi drzewa (*pruning*), ogranicza to wprawdzie trafność modelu, ale poprawia jego interpretowalność i pozwala wyciągać bardziej ogólne wnioski, dlatego do pewnego stopnia jest konieczne; zadowalające efekty uzyskaliśmy ograniczając liczebność końcowego węzła do **200** obserwacji.

Ostateczny model *random forest*, wyliczony z zastosowaniem tych wartości parametrów, osiągnął trafność predykcji na grupie testowej (*out-of-bag accuracy*) równą **83,3 proc.** Oznacza to, że składające się na las drzewa decyzyjne, oparte na części zbioru obejmującej losowo wybrane 63 proc. obserwacji, zastosowane do przewidywania zmiennej zależnej dla pozostałych obserwacji, myliły się tylko w co szóstym przypadku (*out-of-bag error rate* = 16,7 proc.). W zastosowaniu do całego zbioru, model był jeszcze skuteczniejszy: trafność przewidywania wynosiła aż 88,3 proc.

Kluczowymi zmiennymi dotyczącymi aktywności na stronach Facebook spółki i powiązanych z nią marek, które okazały się niezbędne do przewidzenia wyników giełdowych spółki, są:

- ogólne natężenie aktywności (wskaźnik INI – ogólnie i w przeliczeniu na post),
- łączna liczba polubień i reakcji „wow",
- liczba postów opublikowanych przez autora strony, szczególnie typu: status, zdjęcie, wideo,
- zmiana liczby fanów w czasie.

Wykres 5. **Macierz błędów predykcji – wszystkie obserwacje**

		Wartości rzeczywiste	
		poniżej mediany rynku	powyżej mediany rynku
Wartości przewidywane	poniżej mediany rynku	31 005 (44,5 proc.)	4 229 (6,1 proc.)
	powyżej mediany rynku	3 921 (5,6 proc.)	30 478 (43,8 proc.)

Źródło: opracowanie własne PIE.

Model predykcyjny wykazał bardzo wysoką, ponad 88-procentową trafność w przewidywaniu rzeczywistych wartości zmiennej zależnej dla całego zbioru danych. Trafność przewidywania była równie wysoka w przypadku obserwacji lokujących się powyżej i poniżej mediany rynku, co pokazuje macierz błędów (*confusion matrix*) (wykres 5).

Zadaliśmy sobie pytanie o stabilność modelu w czasie. Jak się okazuje, w kolejnych okresach trafność przewidywania nie zmieniała się, pozostając w wąskim przedziale między 86,7 proc. (10.2018 r.) a 90,8 proc. (02.2019 r.). Niezmienną skuteczność modelu w kolejnych okresach pokazujemy na wykresach 6 i 7.

↘ Wykres 6. **Macierz błędów predykcji w kolejnych okresach 3-miesięcznych**

Wartości przewidywane	Wartości rzeczywiste: poniżej mediany rynku	Wartości rzeczywiste: powyżej mediany rynku	Wartości rzeczywiste: poniżej mediany rynku	Wartości rzeczywiste: powyżej mediany rynku
	11.2018-01.2019		02.2019-04.2019	
poniżej mediany rynku	7 352 (44,3 proc.)	960 (5,8 proc.)	7 686 (45,3 proc.)	900 (5,3 proc.)
powyżej mediany rynku	981 (5,9 proc.)	7 315 (44,0 proc.)	820 (4,8 proc.)	7 579 (44,6 proc.)
	05.2018-07.2018		08.2018-10.2018	
poniżej mediany rynku	7 780 (44,0 proc.)	1 137 (6,4 proc.)	8 187 (44,6 proc.)	1 232 (6,7 proc.)
powyżej mediany rynku	1 080 (6,1 proc.)	7 666 (43,4 proc.)	1 040 (5,7 proc.)	7 918 (43,1 proc.)

Źródło: opracowanie własne PIE.

↘ Wykres 7. **Macierz błędów predykcji w kolejnych miesiącach**

Wartości przewidywane	Wartości rzeczywiste: poniżej mediany rynku	Wartości rzeczywiste: powyżej mediany rynku
05.2018		
poniżej mediany rynku	2 499 (44,7 proc.)	351 (6,3 proc.)
powyżej mediany rynku	301 (5,4 proc.)	2 435 (43,6 proc.)
06.2018		
poniżej mediany rynku	2 580 (43,8 proc.)	397 (6,7 proc.)
powyżej mediany rynku	373 (6,3 proc.)	2 543 (43,2 proc.)
07.2018		
poniżej mediany rynku	2 701 (43,7 proc.)	389 (6,3 proc.)
powyżej mediany rynku	406 (6,6 proc.)	2 688 (43,5 proc.)
08.2018		
poniżej mediany rynku	2 737 (44,1 proc.)	424 (6,8 proc.)
powyżej mediany rynku	382 (6,2 proc.)	2 665 (42,9 proc.)
09.2018		
poniżej mediany rynku	2 564 (45,3 proc.)	323 (5,7 proc.)
powyżej mediany rynku	278 (4,9 proc.)	2 495 (44,1 proc.)
10.2018		
poniżej mediany rynku	2 886 (44,3 proc.)	485 (7,5 proc.)
powyżej mediany rynku	380 (5,8 proc.)	2 758 (42,4 proc.)
11.2018		
poniżej mediany rynku	2 506 (44,4 proc.)	301 (5,3 proc.)
powyżej mediany rynku	316 (5,6 proc.)	2 519 (44,6 proc.)
12.2018		
poniżej mediany rynku	2 104 (43,9 proc.)	296 (6,2 proc.)
powyżej mediany rynku	299 (6,2 proc.)	2 095 (43,7 proc.
01.2019		
poniżej mediany rynku	2 742 (44,4 proc.)	363 (5,9 proc.)
powyżej mediany rynku	366 (5,9 proc.)	2 701 (43,8 proc.)
02.2019		
poniżej mediany rynku	2 534 (45,2 proc.)	246 (4,4 proc.)
powyżej mediany rynku	266 (4,8 proc.)	2 554 (45,6 proc.)
03.2019		
poniżej mediany rynku	2 616 (44,5 proc.)	324 (5,5 proc.)
powyżej mediany rynku	324 (5,5 proc.)	2 614 (44,5 proc.)
04.2019		
poniżej mediany rynku	2 536 (46,1 proc.)	330 (6,0 proc.)
powyżej mediany rynku	230 (4,2 proc.)	2 411 (43,8 proc.)

Źródło: opracowanie własne PIE.

Konstrukcja ratingu

Efektem zastosowania modelu predykcyjnego jest przypisanie do każdej spółki, w każdym dniu, jednej z dwóch wartości przewidywania: +1 (wyniki spółki powyżej mediany rynku) lub -1 (poniżej mediany rynku). Przewidywania te – podobnie jak rzeczywiste wartości zmiennej zależnej i podobnie jak kursy akcji – są nazbyt podatne na losowe wahania i dlatego nie należy na nich opierać algorytmu przypisującego oceny ratingowe. Rating, który zmieniałby się codziennie, nie byłby bowiem godny zaufania. Nie spełniałby również celów związanych z oceną atrakcyjności inwestycyjnej spółek. Stabilność może mu zapewnić tylko oparcie go na wynikach przewidywań z dłuższych okresów. Zbliża to metodę klasyfikacji do rzeczywistych ocen ratingowych wiarygodności kredytowej podmiotów gospodarczych, które zmieniane są raz na jakiś czas, a nie codziennie.

Zdecydowaliśmy się oprzeć rating spółek z Giełdy Papierów Wartościowych w Warszawie na agregacji pochodzących z modelu przewidywań dla okresów miesięcznych. Agregacja przewidywań dała dla każdej spółki i miesiąca liczbę dni (roboczych, w których prowadzone były notowania), w których dla tej spółki przewidywano, że jej wyniki będą lokowały się powyżej mediany rynku (dla uproszczenia nazwaliśmy je „dodatnimi"), i liczbę dni, w których przewidywano wyniki poniżej mediany rynku (przewidywania „ujemne"). Proporcje między tymi wartościami wyznaczają przydział do klasy:

- ocenę **A+** otrzymały spółki, dla których wszystkie lub prawie wszystkie przewidywania były dodatnie;
- ocenę **A** – spółki, dla których liczba przewidywań dodatnich była znacząco wyższa od liczby przewidywań ujemnych;
- ocenę **B** – spółki, które miały mniej więcej wyrównane proporcje przewidywań dodatnich i ujemnych;
- ocenę **C** – spółki, które miały wyraźnie większą liczbę przewidywań ujemnych;
- ocenę **C-** – spółki, które uzyskały wyłącznie lub prawie wyłącznie przewidywania ujemne.

Rating ten zmodyfikowaliśmy również do postaci uproszczonej, z podziałem na trzy grupy: A, B i C. Grupę A stanowiły połączone grupy A i A+, a grupę C – połączone grupy C i C-. W tej uproszczonej wersji, ocena A jest przynależna spółkom, których przewidywane wyniki lokują się konsekwentnie powyżej mediany rynku, a ocena C – spółkom poniżej mediany rynku. Klasa B zarezerwowana jest dla spółek, których wyniki giełdowe są przeciętne – równie często znajdują się powyżej, jak i poniżej mediany.

W dalszej części publikacji omówimy liczebności i udziały spółek z poszczególnych klas oraz zweryfikujemy różnice między spółkami różnych klas pod względem dynamiki wzrostu lub spadku cen akcji.

Częstości przypisania poszczególnych ocen atrakcyjności inwestycyjnej są różne, ale – jak widać w tabeli 1 – mimo nierównoliczności grup, ich rozkład odznacza się wysoką symetrią: ocena A+ jest równie częsta jak C-, a ocena A – równie częsta jak C.

Warto podkreślić, że od tego momentu jako jednostki obserwacji przyjmowaliśmy spółkomiesiące i dla nich wyliczaliśmy rating. Stąd łączna liczebność jest niższa niż liczebność zbioru danych surowych, pojedynczą obserwację stanowiły spółkodni, tj. iloczyn spółka × dzień.

Oprócz ratingu opartego na wartościach przewidywanych, pochodzących z modelu predykcyjnego, na potrzeby wykazania ich trafności

wyliczyliśmy też bliźniaczy rating, opierając się na rzeczywistych wynikach giełdowych spółek. Metoda tworzenia tego ratingu była identyczna jak w przypadku ratingu opartego na przewidywaniach, z tą różnicą, że posłużyliśmy się informacją o rzeczywistej dynamice zmian kursów akcji: czy miesięczna stopa zwrotu z inwestycji w akcje spółki lokowała się powyżej, czy poniżej mediany miesięcznych stóp zwrotu dla wszystkich spółek.

↘ Tabela 1. **Udziały liczebności i częstości przypisania różnych ocen atrakcyjności inwestycyjnej**

Ocena (5 klas)	N	Proc.	Ocena (3 klasy)	N	Proc.
A+	1 146	31,4	A	1 577	43,2
A	431	11,8			
B	438	12,0	B	438	12,0
C	438	12,0	C	1 633	44,8
C-	1 195	32,8			
RAZEM	3 648	100	RAZEM	3 648	100

Źródło: opracowanie własne PIE.

Macierz (wykres 8) zawiera prezentację trafności przypisania ocen. Udziały procentowe znajdujące się na głównej przekątnej oznaczają częstość sytuacji, w których trafnie przewidziano przypisanie do jednej z pięciu klas. Suma tych udziałów, równa **64,3 proc.**, może być uznana za **wskaźnik trafności ratingu**.

Szczegółowa analiza macierzy pozwala stwierdzić, że:

- przynależność do jednej z klas skrajnych jest wykrywana z bardzo wysoką trafnością, niezwykle rzadko zdarzają się sytuacje, gdy spółka – według rzeczywistych wyników giełdowych – powinna być zaliczona do klasy A+ lub C-, ale przewidywanie wskazuje inną klasę;
- przynależność do klasy środkowej (B), mimo jej stosunkowo niskiej liczebności, jest dość trafnie wykrywana; wśród spółek o rzeczywistej wartości ratingu równej B, ocena ta jest również najczęściej wskazywana przez model predykcyjny, jakkolwiek zdarzają się błędne przypisania do klas A i C;
- najczęstsze błędy polegają na przypisaniu do klasy skrajnej (A+ lub C-), podczas gdy spółce – na podstawie rzeczywistych danych – należałoby przypisać ocenę A lub C; jest to efekt, który trudno skorygować, nie obniżając ogólnej jakości modelu – wynika on przede wszystkim z tego, że model dokonuje pewnych uproszczeń i „wyostrza" obraz sytuacji, tam, gdzie spółka osiąga bardzo dobre wyniki, lecz od czasu do czasu zdarza jej się notować też spadki (relatywne w odniesieniu do mediany rynku), model upraszcza obraz sytuacji i przewiduje same wzrosty kursów.

↘ Wykres 8. Macierz błędów predykcji stopnia atrakcyjności inwestycyjnej (5 klas)

Wartości przewidywane \ Wartości rzeczywiste	A+	A	B	C	C-
A+	650 (19,3 proc.)	319 (9,5 proc.)	58 (1,7 proc.)	7 (0,2 proc.)	3 (0,1 proc.)
A	25 (0,7 proc.)	279 (8,3 proc.)	109 (3,2 proc.)	5 (0,1 proc.)	
B	6 (0,2 proc.)	52 (1,5 proc.)	291 (8,6 proc.)	55 (1,6 proc.)	2 (0,1 proc.)
C	2 (0,1 proc.)	7 (0,2 proc.)	135 (4,0 proc.)	266 (7,9 proc.)	15 (0,4 proc.)
C-	3 (0,1 proc.)	9 (0,3 proc.)	63 (1,9 proc.)	328 (9,7 proc.)	681 (20,2 proc.)

Źródło: opracowanie własne PIE.

Błędy w klasyfikacji, które następują między ocenami A i A+ oraz C i C-, skłaniają do zastosowania ratingu uproszczonego, w którym spółkom przypisywana jest jedna z trzech ocen: A, B lub C (wykres 9). W tym wariancie trafność przypisania ocen jest dużo wyższa i wynosi aż **84,7 proc.** (suma wartości procentowych na przekątnej poniższej macierzy). Zdecydowana większość błędów przypisania oceny polega na tym, że spółka, która wg rzeczywistych danych ma stopień atrakcyjności inwestycyjnej B, otrzymuje ocenę A lub B. Te sytuacje zdarzają się w ok. 11 proc. przypadków. Błędy następują dużo rzadziej w przypadku spółek, które powinny uzyskać oceny A lub C. Takie spółki niemal zawsze otrzymują prawidłową ocenę.

Najważniejszym sprawdzianem trafności ratingu jest jego przydatność do oceny rzeczywistej atrakcyjności inwestycyjnej spółki. Spodziewamy się, że spółki z grup A+ i A będą osiągały znacząco lepsze wyniki na giełdzie niż spółki z grupy B, a te z kolei będą miały wyższe stopy zwrotu niż spółki z grup C i C-. W tym celu porównaliśmy akcje spółek, które otrzymały różne oceny atrakcyjności inwestycyjnej.

↘ Wykres 9. Macierz błędów predykcji stopnia atrakcyjności inwestycyjnej (3 klasy)

Wartości przewidywane	Wartości rzeczywiste: A	B	C
A	1 273 (37,8 proc.)	167 (5,0 proc.)	15 (0,4 proc.)
B	58 (1,7 proc.)	291 (8,6 proc.)	57 (1,7 proc.)
C	21 (0,6 proc.)	198 (5,9 proc.)	1 290 (38,3 proc.)

Źródło: opracowanie własne PIE.

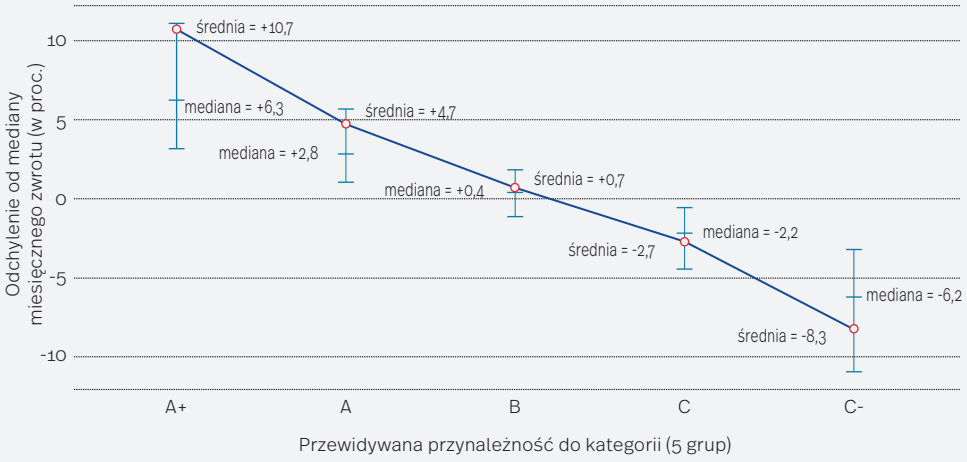
Na wykresach 10 i 11 pokazujemy różnice średnich i median wartości relatywnej miesięcznej stopy zwrotu z inwestycji między spółkami, którym przypisano poszczególne oceny, w wariancie oceny pięciostopniowej i trzystopniowej. Końce szarych pionowych odcinków na wykresach wyznaczają kwartyle rozkładów w podgrupach, zaś poziomy szary odcinek pośrodku lokalizuje medianę w podgrupie. Czerwone punkty połączone linią wyznaczają regresję średnich, czyli relatywne miesięczne stopy zwrotu w poszczególnych podgrupach. Jak widać różnice między tymi wartościami są bardzo duże.

W klasach A+ i A stopy zwrotu są znacząco wyższe od przeciętnych, natomiast w klasach C i szczególnie C- – znacząco niższe. Jednoczynnikowa analiza wariancji wykazała istotność statystyczną różnic między średnimi w podgrupach ($F = 125{,}2$ dla pięciu klas, $F = 223{,}9$ dla trzech, w obu przypadkach $p < 0{,}001$), co dowodzi, że wynik ten jest bardzo daleki od przypadkowości.

Uzyskane wyniki dowodzą, że tak stworzony rating z bardzo dużą trafnością pozwala przewidzieć czy miesięczna stopa zwrotu w kolejnym okresie będzie powyżej czy poniżej przeciętnej. Tym samym spełnia warunki bycia dobrym wskaźnikiem atrakcyjności inwestycyjnej.

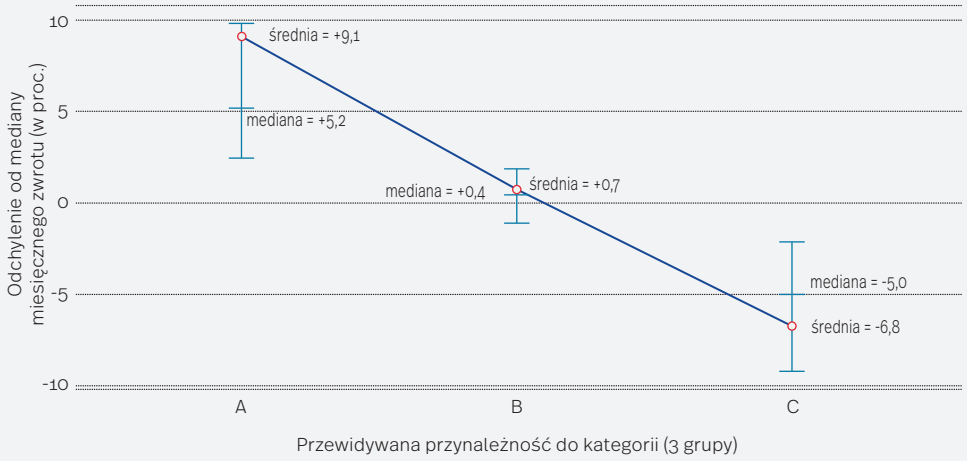
W kolejnym kroku oszacowaliśmy przypisanie analizowanych spółek do poszczególnych klas. Ratingi obejmują okres 12 kolejnych miesięcy. Przykład wyników zawiera tabela 2. Jak można zaobserwować widoczne są znaczne skoki pomiędzy klasami, w tym również pomiędzy skrajnymi klasami.

↘ **Wykres 10.** **Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (5 klas)**



Źródło: opracowanie własne PIE.

↘ **Wykres 11.** **Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (3 klasy)**



Źródło: opracowanie własne PIE.

↘ Tabela 2. **Fragment klasyfikacji spółek w podziale na trzy klasy dla dwunastu kolejnych miesięcy**

Nazwa	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12
11BIT	A	B	C	B	C	C	C	C	A	C	A	A
4FUNMEDIA	A	A	B	C	A	B	C	C	C	C	C	A
ABCDATA	A	A	C	B	A	C	C	C	A	B	A	C
ABPL	C	C	C	A	B	A	C	C	A	A	C	C
ACAUTOGAZ	A	A	A	A	A	A	A	A	A	A	A	A
ACTION	C	A	A	A	A	C	C	C	A	C	A	A
ADIUVO	A	A	C	C	A	A	A	C	C	A	A	C
AGORA	A	A	C	C	C	A	A	A	C	A	C	A
AILLERON	A	C	C	A	A	C	C	C	A	C	C	C
ALIOR	A	A	A	B	C	C	C	C	C	A	A	C
ALUMETAL	A	C	C	A	C	A	B	A	C	A	A	A
AMBRA	A	B	C	C	C	C	C	A	A	A	A	C
AMICA	C	C	B	A	C	C	B	A	A	A	A	B
APATOR	A	A	C	A	B	B	C	A	C	B	B	C
ARCHICOM	A	A	B	C	B	C	A	B	C	A	B	A

Źródło: opracowanie własne PIE.

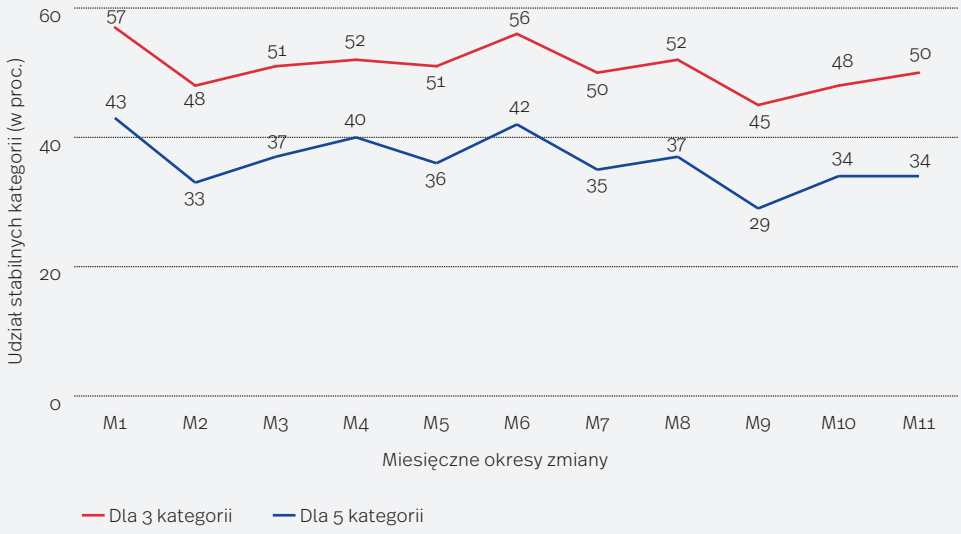
Duża zmienność klasyfikacji może być istotnym ograniczeniem metody, świadczy bowiem o tym, że w dłuższym okresie kondycja spółek jest niestabilna stabilna, co ma przełożenie na oceny. Warto w tym miejscu zaznaczyć, że nie tylko przewidywane wartości ratingu charakteryzują się dużą zmiennością. Wartości rzeczywiste również mocno „skaczą". Co prawda dzięki temu możliwe jest przewidywanie zwrotu z inwestycji w perspektywie miesiąca, ale już niekoniecznie w dłuższym okresie. Dlatego też, aby systematycznie sprawdzać stabilność wyników, przeprowadzaliśmy kolejne analizy.

W pierwszej kolejności ustaliliśmy poziom stabilności ratingu między kolejnymi okresami. Jak pokazaliśmy niżej, wartość rating waha się od 45 proc. do 57 proc. w przypadku podziału na 3 klasy i od 29 proc. do 43 proc. dla 5 klas (wykres 12).

W kolejnym kroku przyjrzeliśmy się stabilności wyników na poziomie spółek. Wyniki dla ratingu z podziałem na 3 i 5 klas przedstawiliśmy na wykresie 13.

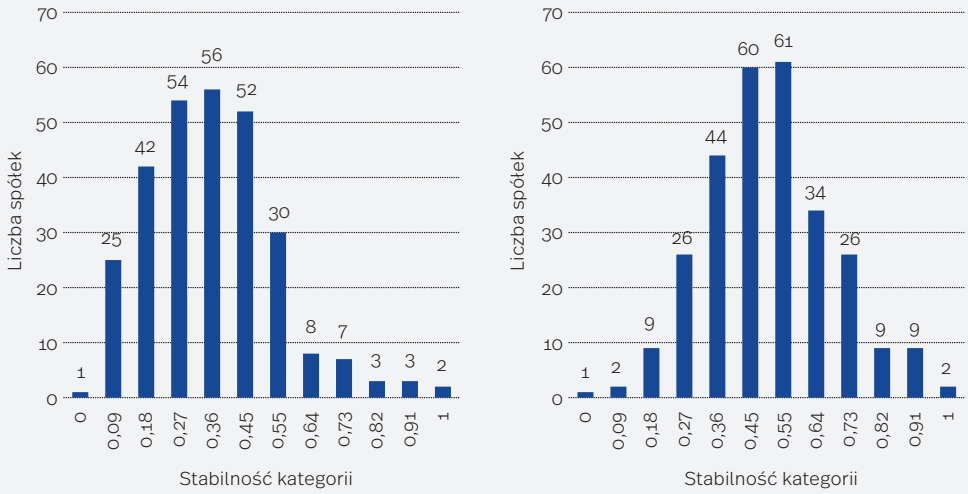
Dla ratingu opartego na 5 klasach stabilność przypisania spółek jest znacznie mniejsza. Rozwiązanie oparte na podziale na 3 klasy jest nie tylko znacznie bardziej precyzyjne, ale również charakteryzuje się mniejszą dynamiką, co w tym przypadku należy uznać za zaletę.

↘ **Wykres 12.** Poziom stabilności przypisania spółki do kategorii w czasie



Źródło: opracowanie własne PIE.

↘ **Wykres 13.** Rozkład stabilności spółek w okresie jednego roku z podziałem na 5 i 3 kategorie



Źródło: opracowanie własne PIE.

Walidacja metody na danych z giełdy wiedeńskiej

Przy ocenie jakości modelu predykcyjnego, a tym samym zaprezentowanego wyżej ratingu, liczą się dwa kryteria, które pozostają we wzajemnej opozycji. Pierwszym kryterium jest trafność przewidywania – to jak dokładnie model opisuje cechę, którą ma za zadanie opisywać. **Niniejszy rating atrakcyjności inwestycyjnej ma za zadanie odróżniać spółki notujące – relatywne w stosunku do ogółu rynku – wzrosty kursów akcji, od spółek notujących (relatywne) spadki.** Z tego zadania, jak się przekonaliśmy, wywiązuje się bardzo dobrze. Drugim kryterium jest tzw. *generalizowalność* modelu – to czy ma on zastosowanie w szerszym kontekście, poza tym wycinkiem rzeczywistości, którego bezpośrednio dotyczy.

Testem na ogólność zaprezentowanej tu metody budowy ratingu, dopasowanej do polskich danych, jest zastosowanie jej do innego, zagranicznego rynku papierów wartościowych. Wybór padł na rynek austriacki, ze względu na jego specyfikę, nieco odmienną od polskiej. Giełda wiedeńska przez większość inwestorów międzynarodowych jest uznawana za rynek rozwinięty. Emitenci to często korporacje międzynarodowe, rozpoznawalne banki, holdingi o dużo wyższym kapitale niż wielu spośród emitentów stosunkowo młodej giełdy warszawskiej. Wysoka szczegółowość i kompletność publikowanych na stronie internetowej http://www.wienerborse.at informacji na temat kursów notowanych na tej giełdzie spółek umożliwiła podjęcie analizy. W związku z tym zweryfikowaliśmy czy aktywność na stronach portali społecznościowych spółek austriackich będzie równie dobrym predyktorem kursów akcji, jak to ma miejsce w warunkach polskich.

Analizę danych austriackich przeprowadziliśmy w bardzo podobny sposób do analizy danych polskich:

- dla każdej spółki i każdego dnia w analizowanym okresie 01.2018-12.2019 odnaleźliśmy wszystkie przynależne do tych spółek strony na portalu Facebook (tj. strony spółki, spółek od niej zależnych lub posiadanych przez nią marek), spółki nieposiadające takich stron wykluczyliśmy z analizy;
- wykorzystując API portalu Facebook pobraliśmy dane na temat aktywności na tych stronach; wyliczyliśmy analogiczne wskaźniki jak dla spółek polskich, stosując do tego celu m.in. algorytmy przypisywania sentymentu (nacechowanie pozytywne lub negatywne) dla postów w języku niemieckim, angielskim i polskim;
- każdy wskaźnik poddaliśmy przekształceniom analitycznym, takim jak: rolowanie danych (uśrednianie wartości z ostatnich 30 dni), standaryzacja, normalizacja, logarytmizacja, binaryzacja i rekodowanie, łącząc je nieraz ze sobą dla uzyskania predyktorów o lepszych właściwościach; przy liczbie 50 spółek w bazie, ale dłuższym niż dla Polski okresie analizy, łączna liczba obserwacji wyniosła 19 012 spółkodni;
- sięgając po dostępne w internecie dane na temat kursów akcji spółek wyliczyliśmy miesięczną stopę zwrotu w 30-dniowym okresie następującym po danym dniu, medianę tej wartości dla danego dnia, a następnie – odejmując jedną od drugiej – uzyskaliśmy relatywną miesięczną stopę zwrotu, wyrażoną w wartościach procentowych;

- relatywną miesięczną stopę zwrotu poddaliśmy binaryzacji uzyskując zmienną dwuwartościową, którą potraktowaliśmy jako zmienną zależną w modelu predykcyjnym; predyktorami w tym modelu uczyniliśmy wszystkie, zarówno surowe, jak i przekształcone wskaźniki związane z aktywnością na stronach portalu Facebook;
- predyktory poddaliśmy wstępnej selekcji uzyskując listę 15 predyktorów najsilniej powiązanych z relatywną miesięczną stopą zwrotu;
- wykorzystaliśmy model uczenia maszynowego *random forest* do przewidywania wartości ww. zmiennej binarnej; hiperparametry modelu dobraliśmy w następstwie procesu „strojenia" modelu, kierując się zarówno maksymalizacją trafności, jak i troską o prostotę i ogólność modelu; kompromis osiągnęliśmy przy liczbie 2000 drzew, uwzględnieniu każdorazowo wszystkich 15 predyktorów na etapie podziału zbioru oraz przy maksymalnej liczbie węzłów końcowych („liści") drzewa ograniczonej do 400, oznacza to podział zbioru na części o średniej liczebności ok. 45-50 obserwacji, co oddala ryzyko „przetrenowania" modelu, tzw. *overfittingu*;
- ostateczny model predykcyjny osiągnął trafność przewidywania na zbiorze testowym (*out-of-bag accuracy*) równą **75,7 proc.**, z kolei model zbudowany na danych polskich osiągnął trafność 83,3 proc., ale wykorzystywał więcej predyktorów (50, nie 15)[1].

Trafność przewidywania zmiennej zależnej za pomocą modelu, w zastosowaniu do całego zbioru danych, była wyższa niż trafność przewidywania na zbiorze testowym. Wyniosła bowiem 81,6 proc. (dla modelu zbudowanego na danych z giełdy warszawskiej, było to nieco więcej: 88,3 proc.).

Wykres 14. Macierz błędów predykcji – wszystkie obserwacje

		Wartości rzeczywiste: poniżej mediany rynku	Wartości rzeczywiste: powyżej mediany rynku
Wartości przewidywane	poniżej mediany rynku	8 738 (46,0 proc.)	2 538 (13,3 proc.)
	powyżej mediany rynku	961 (5,1 proc.)	6 775 (35,6 proc.)

Źródło: opracowanie własne PIE.

[1] Mniejsza liczba predyktorów dla danych z giełdy wiedeńskiej wynikała ze znacznie mniejszej liczby notowanych spółek. Większa liczba predyktorów powodowałaby znaczące przeuczenie modelu.

Macierz błędów (wykres 14) ukazuje rozkład błędów przewidywań. Zarówno dla spółek znajdujących się ponad medianą rynku, jak i poniżej, trafność przewidywania była wysoka. W przypadku spółek osiągających ponadprzeciętne wartości stopy zwrotu model nieco częściej dokonywał błędnej klasyfikacji, ale za to trafniej identyfikował sytuacje, w których spółka osiągała rezultaty niższe od przeciętnych.

Kolejny zestaw macierzy błędów (wykres 15) pozwala prześledzić trafność predykcji w czasie. Powtarzający się w kolejnych kwartałach 2018 i 2019 r. wzór macierzy oraz niezmiennie wysokie wartości wskaźnika trafności przewidywania (77,5-86,8 proc.) upewniają nas, że model odznacza się wysoką jakością, niezależnie od przedziału czasu, do którego jest stosowany.

↘ Wykres 15. **Macierz błędów predykcji w kolejnych kwartałach**

		Wartości rzeczywiste							
		poniżej mediany rynku	powyżej mediany rynku	poniżej mediany rynku	powyżej mediany rynku	poniżej mediany rynku	powyżej mediany rynku	poniżej mediany rynku	powyżej mediany rynku
		I kw. 2018		Ii kw. 2018		III kw. 2018		IV kw. 2018	
Wartości przewidywane	poniżej mediany rynku	753 (47,4 proc.)	135 (8,5 proc.)	1 080 (46,4 proc.)	280 (12,0 proc.)	1 153 (47,4 proc.)	292 (12,0 proc.)	1 014 (45,3 proc.)	343 (34,6 proc.)
	powyżej mediany rynku	76 (4,8 proc.)	626 (39,4 proc.)	118 (5,1 proc.)	850 (36,5 proc.)	112 (4,6 proc.)	876 (36,0 proc.)	107 (4,8 proc.)	773 (34,6 proc.)
		I kw. 2019		II kw. 2019		III kw. 2019		IV kw. 2019	
	poniżej mediany rynku	1 120 (44,6 proc.)	410 (16,3 proc.)	1 163 (46,3 proc.)	359 (14,3 proc.)	1 282 (45,8 proc.)	350 (12,5 proc.)	1 173 (45,0 proc.)	369 (14,2 proc.)
	powyżej mediany rynku	156 (6,2 proc.)	826 (32,9 proc.)	133 (5,3 proc.)	855 (34,1 proc.)	138 (4,9 proc.)	1 028 (36,7 proc.)	121 (4,6 proc.)	941 (36,1 proc.)

Źródło: opracowanie własne PIE.

Ostatnim, a zarazem kluczowym etapem analizy danych z giełdy wiedeńskiej, była agregacja pochodzących z modelu wyników przewidywań do postaci ratingu przez wyliczenie ocen atrakcyjności inwestycyjnej dla każdej spółki i każdego miesiąca. Tak jak w przypadku danych z giełdy warszawskiej, spółkom przypisaliśmy pięć oznaczeń: A+, A, B, C i C-, a w wersji uproszczonej trzy: A, B i C. W tabeli 3 zaprezentowaliśmy proporcje poszczególnych klas. Jak widać udało się osiągnąć większą równowagę proporcji klas niż dla polskich danych (środkowa kategoria B jest liczniejsza). W zamian za to rating odznacza się lekką asymetrią, ciąży bowiem ku kategorii C- (28 proc. obserwacji), która jest w nim nadreprezentowana w stosunku do kategorii A+ i A (poniżej 16 proc.).

↘ Tabela 3. **Udziały liczebności i częstości przypisania różnych ocen atrakcyjności inwestycyjnej**

Ocena (5 klas)	N	Proc.
A+	138	14,0
A	156	15,8
B	213	21,6
C	203	20,6
C-	277	28,0
RAZEM	987	100, 0

Źródło: opracowanie własne PIE.

Trafność ratingu opartego na przewidywaniach w stosunku do klasyfikacji powstałej na podstawie rzeczywistych wyników giełdowych można odnaleźć w macierzy błędów (wykres 16). Niespodzianką jest fakt, że najczęstszy w przypadku klasyfikacji na spółkach polskich typ błędu przewidywania, polegający na uznaniu spółki z kategorii A lub C za reprezentanta kategorii skrajnej (odpowiednio A+ lub C-), występuje w przypadku ratingu spółek austriackich dużo rzadziej. W efekcie dla każdej „rzeczywistej" klasy najczęściej przypisywana jej reprezentantom ocena „przewidywana" jest zgodna z ich rzeczywistą kategoryzacją. Mimo nieco gorszych parametrów modelu predykcyjnego, sam rating przy podziale na 5 klas osiąga **60,2 proc.** trafności (rating spółek polskich: **64,3 proc.**), zaś rating w uproszczonym podziale na 3 klasy: **77,3 proc.** trafności (dla Polski: **84,7 proc.**). Należy stwierdzić, że są to wysokie wyniki, zważywszy na fakt, że do celów przewidywania posłużono się wyłącznie informacjami o aktywności na portalu Facebook, bez sięgania po jakiekolwiek informacje na temat osiąganych wyników finansowych.

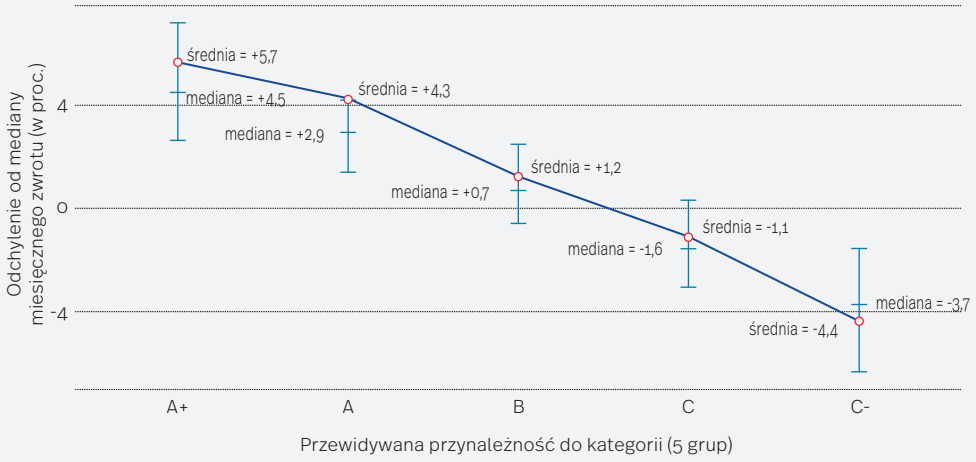
↘ Wykres 16. **Macierz błędów predykcji stopnia atrakcyjności inwestycyjnej (5 klas)**

Wartości przewidywane \ Wartości rzeczywiste	A+	A	B	C	C-
A+	84 (8,0 proc.)	44 (4,2 proc.)		4 (0,4 proc.)	
A	25 (2,4 proc.)	141 (13,5 proc.)	32 (3,1 proc.)	6 (0,6 proc.)	
B	10 (1,0 proc.)	33 (3,1 proc.)	122 (11,6 proc.)	17 (1,6 proc.)	
C	9 (0,9 proc.)	39 (3,7 proc.)	50 (4,8 proc.)	153 (14,6 proc.)	13 (1,2 proc.)
C-	3 (0,3 proc.)	14 (1,3 proc.)	21 (2,0 proc.)	97 (9,3 proc.)	131 (12,5 proc.)

Źródło: opracowanie własne PIE.

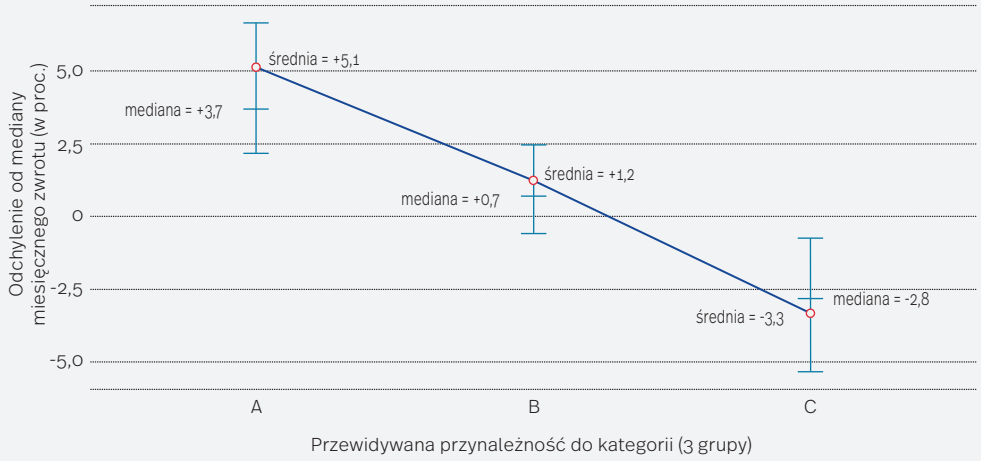
Na koniec badań pozostawiliśmy rozstrzygający test praktycznej przydatności modelu, którym jest porównanie wyników giełdowych spółek reprezentujących poszczególne stopnie atrakcyjności inwestycyjnej. Tak jak w wypadku spółek polskich, tak i tu różnice w wartościach relatywnej stopy zwrotu z inwestycji w akcje spółki są bardzo znaczące. Analiza wariancji ponownie dowiodła wysokiej istotności uzyskanych różnic ($F = 159,5$ dla pięciu klas, $F = 274,4$ dla trzech, $p < 0,001$) (wykres 17 i 18).

Wykres 17. Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (5 klas)



Źródło: opracowanie własne PIE.

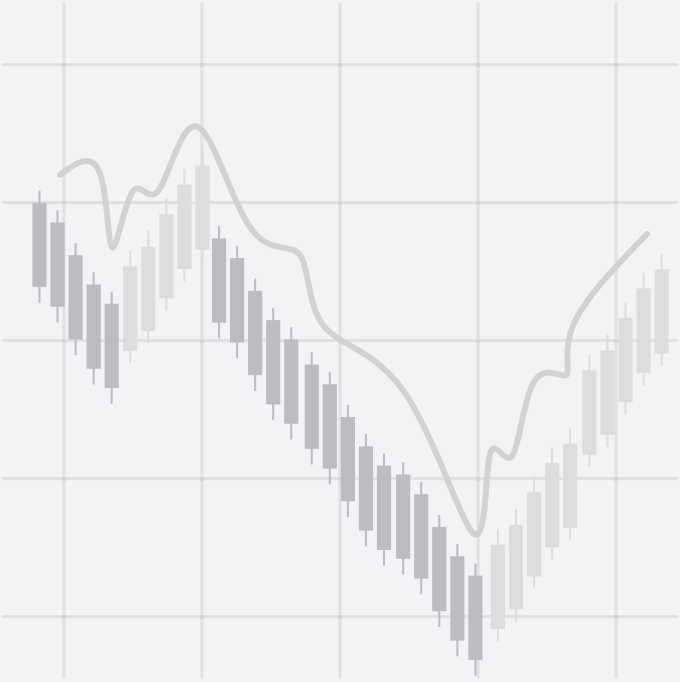
Wykres 18. Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (3 klasy)



Źródło: opracowanie własne PIE.

* * *

Reasumując, zbudowaliśmy skuteczny prototyp ratingu atrakcyjności inwestycyjnej, przypisujący emitentom z warszawskiej Giełdy Papierów Wartościowych, wyłącznie na podstawie danych o aktywnościach na stronach portali społecznościowych tych spółek, oceny silnie korespondujące z ich wynikami finansowymi. Metodę konstrukcji ratingu zastosowaliśmy także na danych pochodzących z giełdy wiedeńskiej i otrzymaliśmy równie zachęcający rezultat. Wierzymy, że skonstruowany w podobny sposób rating może stać się obiecującym instrumentem oceny atrakcyjności inwestycyjnej, który dzięki wykorzystaniu danych z mediów społecznościowych, może być stosowany również dla spółek spoza parkietu.



Bibliografia

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. (2014), *Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, https://doi.org/10.3115/v1/D14-1179.

Interaktywnie.com (2019), *Media społecznościowe. Raport*, https://interaktywnie.com/biznes/artykuly/raporty-interaktywnie-com/raport-media-spolecznosciowe-2019-259316 [dostęp: 01.12.2020].

Plan działania na rzecz tworzenia unii rynków kapitałowych (2015), Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów, COM(2015) 468 final.

Strategia Rozwoju Rynku Kapitałowego (2019), Załącznik do Uchwały Rady Ministrów nr 114 z 01.10.2019, Dz.U. 2019 poz. 1027.

https://fasttext.cc/ [dostęp: 06.12.2020].

http://www.wienerborse.at [dostęp: 01.12.2020].

Spis infografik, wykresów i tabel

SPIS INFOGRAFIK

- Infografika 1. Analizowane spółki i ich aktywność na Facebooku. 10
- Infografika 2. Rozkład aktywności internautów na posty firm ma Facebooku według miesięcy i dni tygodnia (w mln). 13
- Infografika 3. Rodzaje postów zamieszczanych przez firmy na Facebooku i reakcje internautów 16

SPIS WYKRESÓW

- Wykres 1. Rozkład liczby postów opublikowanych przez wszystkie profile na Facebooku przypisane do spółek z głównego parkietu GPW, uszeregowane od spółek, które publikowały najwięcej do spółek, które nie publikowały wcale 12
- Wykres 2. Rozkład liczby postów według poszczególnych rodzajów postów możliwych do opublikowania na Facebooku. 14
- Wykres 3. Przeciętna wartość miesięcznej stopy zwrotu dla spółek z głównego parkietu GPW uwzględnionych w modelowaniu 21
- Wykres 4. Rozkład wartości relatywnej miesięcznej stopy zwrotu 22
- Wykres 5. Macierz błędów predykcji – wszystkie obserwacje. 24
- Wykres 6. Macierz błędów predykcji w kolejnych okresach 3-miesięcznych 25
- Wykres 7. Macierz błędów predykcji w kolejnych miesiącach 26
- Wykres 8. Macierz błędów predykcji stopnia atrakcyjności inwestycyjnej (5 klas) 29
- Wykres 9. Macierz błędów predykcji stopnia atrakcyjności inwestycyjnej (3 klasy) 30
- Wykres 10. Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (5 klas). 31
- Wykres 11. Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (3 klasy). 31
- Wykres 12. Poziom stabilności przypisania spółki do kategorii w czasie 33
- Wykres 13. Rozkład stabilności spółek w okresie jednego roku z podziałem na 5 i 3 kategorie 33
- Wykres 14. Macierz błędów predykcji – wszystkie obserwacje. 35
- Wykres 15. Macierz błędów predykcji w kolejnych kwartałach 36
- Wykres 16. Macierz błędów predykcji stopnia atrakcyjności inwestycyjnej (5 klas) 38
- Wykres 17. Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (5 klas). 39
- Wykres 18. Relatywna miesięczna stopa zwrotu dla spółek o różnym stopniu atrakcyjności inwestycyjnej (3 klasy). 39

SPIS TABEL

- Tabela 1. Udziały liczebności i częstości przypisania różnych ocen atrakcyjności inwestycyjnej 28
- Tabela 2. Fragment klasyfikacji spółek w podziale na trzy klasy dla dwunastu kolejnych miesięcy. 32
- Tabela 3. Udziały liczebności i częstości przypisania różnych ocen atrakcyjności inwestycyjnej 37

Polski Instytut Ekonomiczny

Polski Instytut Ekonomiczny to publiczny *think tank* gospodarczy, którego historia sięga 1928 roku. Obszary badawcze Polskiego Instytutu Ekonomicznego to przede wszystkim handel zagraniczny, makroekonomia, energetyka i gospodarka cyfrowa oraz analizy strategiczne dotyczące kluczowych obszarów życia społecznego i publicznego Polski. Instytut zajmuje się dostarczaniem analiz i ekspertyz do realizacji Strategii na Rzecz Odpowiedzialnego Rozwoju, a także popularyzacją polskich badań naukowych z zakresu nauk ekonomicznych i społecznych w kraju oraz za granicą.

Polski
Instytut
Ekonomiczny