

Uniwersytet Śląski w Katowicach
Wydział Matematyki, Fizyki i Chemii

Paweł Lewandowski

Nr albumu 298807

Ocena ryzyka upadłości spółek na
podstawie własności multifraktałnych
notowań akcji z użyciem metody maszyn
wektorów nośnych - budowa modelu,
zastosowanie w handlu algorytmicznym.

PRACA DYPLMOWA
MAGISTERSKA

promotor:
dr Łukasz Machura

Katowice, 2015

Słowa kluczowe: **akcje, ryzyko, MF-DFA, SVM, handel algorytmiczny**

Oświadczenie autora pracy

Ja, niżej podpisany:

imiona i nazwisko: **Paweł Piotr Lewandowski**

autor pracy dyplomowej pt.: *Ocena ryzyka upadłości spółek na podstawie własności multifraktałnych notowań akcji z użyciem metody maszyn wektorów nośnych - budowa modelu, zastosowanie w handlu algorytmicznym,*

Nr albumu: **298807**,

student Uniwersytetu Śląskiego w Katowicach, Wydziału Matematyki, Fizyki i Chemii, kierunku studiów **Ekonofizyka** oświadczam, że ww. praca dyplomowa:

- została przygotowana przeze mnie samodzielnie¹,
- nie narusza praw autorskich w rozumieniu ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (tekst jednolity Dz. U. z 2006 r. Nr 90, poz. 631, z późn. zm.) oraz dóbr osobistych chronionych prawem cywilnym,
- nie zawiera danych i informacji, które uzyskałem/łam w sposób niedozwolony,
- nie była podstawą nadania dyplomu uczelni wyższej lub tytułu zawodowego ani mnie, ani innej osobie.

Oświadczam również, że treść pracy dyplomowej zamieszczonej przeze mnie w Archiwum Prac Dyplomowych jest identyczna z treścią zawartą w wydrukowanej wersji pracy.

Jestem świadoma/y odpowiedzialności karnej za złożenie fałszywego oświadczenia.

.....

Data

.....

Podpis autora pracy

¹1 uwzględniając merytoryczny wkład promotora (w ramach prowadzonego seminarium dyplomowego)

Spis treści

1	Wstęp	1
1.1	Istota ryzyka na rynku finansowym	2
1.2	Metody analizy i prognozowania ryzyka	6
1.3	Prognozowanie niewypłacalności przedsiębiorstw	9
1.4	Algorytmiczny obrót instrumentami finansowymi	13
1.4.1	Wprowadzenie do handlu algorytmicznego	13
1.4.2	GPW S.A. - kształtowanie specyfiki handlu w wyniku zmian technicznych i technologicznych	19
2	Metodologia i metodyka przeprowadzonych badań	22
2.1	Badanie własności multifraktałnych rynków finansowych	22
2.1.1	FMH - Hipoteza rynku fraktalnego	24
2.1.2	Multifraktałna analiza szeregów czasowych	27
2.1.3	Metoda MF-DFA	29
2.2	Analiza dyskryminacyjna	31
2.2.1	Metoda maszyn wektorów nośnych - SVM	34
2.3	Koncepcja algorytmu MF-DFA - SVM oceny ryzyka inwestycyj- nego pod kątem niewypłacalności emitenta	38

3 Budowa i zastosowanie modelu MF DFA-SVM w predykcji upadłości spółek notowanych na GPW S.A.	42
3.1 Dobór walorów oraz zakresu danych giełdowych do badań	43
3.2 Przeprowadzenie procedury MF DFA, utworzenie wektorów obserwacji	48
3.3 Wstępna analiza charakterystyk fraktalnych	53
3.4 SVM - budowa i walidacja modelu	60
3.5 Zastosowanie modelu MF DFA-SVM na przykładzie spółki CD Projekt S.A.	64
3.5.1 Ocena ryzyka upadłości spółki	67
3.5.2 Symulacja działania prymitywnego algorytmu generowania sygnałów kupna/sprzedaży	70
3.6 Podsumowanie badań	73
3.7 Wnioski	75

Streszczenie

Praca prezentuje alternatywne dla dotychczas stosowanych podejście do kwantyfikacji ryzyka upadłości spółek, których akcje są notowane na giełdzie.

Zgodnie z hipotezą rynku fraktalnego przyjęto, że określenie własności multifrakalnych notowań może stanowić zbiorcze ujęcie złożonej rynkowej oceny kondycji spółki.

Analizę przeprowadzono dla spółek, w przypadku których miało miejsce złożenie wniosku o ogłoszenie upadłości, oraz przedsiębiorstw o dobrej kondycji finansowej. Podstawę analizy utworzyły szeregi czasowe cen zamknięcia i dziennej liczby transakcji. Celem określenia natury zmienności poszczególnych walorów użyto metody multifraktalnej analizy detrendowanych fluktuacji (MF-DFA), na bazie której utworzono 4-elementową charakterystykę własności multifrakalnych.

Na podstawie uzyskanych charakterystyk skonstruowano z wykorzystaniem metody maszyn wektorów nośnych (SVM) model dyskryminacyjny. Model zastosowano w okresowym określaniu ryzyka upadłości spółki CD Projekt S.A. (dawniej Optimus S.A.) dla notowań w latach 2004-2015.

W ostatnim etapie badań dokonano symulacji działania prymitywnego algorytmu inwestycyjnego opartego o utworzony model.

Rozdział 1

Wstęp

*If we knew what it was we were
doing, it would not be called
research, would it?*

Albert Einstein

Dynamiczny rozwój technologiczny doprowadził do silnego rozwoju narzędzi stosowanych w analizie rynków finansowych. Identyfikacja i kwantyfikacja ryzyka stanowi kluczowy element w podejmowaniu decyzji inwestycyjnych. Jednym z aspektów ryzyka inwestycyjnego jest ewentualność ogłoszenia upadłości przez emitenta. Częstotliwość publikacji informacji o stanie finansów spółek giełdowych często bywa niższa niż ta, która koresponduje z horyzontem inwestycyjnym ewentualnego akcjonariusza.

Pojawia się zatem przestrzeń dla metod abstrahujących od analizy fundamentalnej a opartych na ocenie rynku. Według hipotez rynku (patrz podrozdział 2.1.1), wartość tej oceny wyrażana jest w cenie akcji emitenta. Techniki oferowane w ramach ekonometrycznego podejścia do analizy jej zmienności bazują, przynajmniej w pewnej mierze, na hipotezie rynku efektywnego.

W obliczu wątpliwości dla tej koncepcji oraz powstania fraktalnej hipotezy rynku, w niniejszej pracy proponuję metodę prognozowania upadłości spółek notowanych na giełdzie w oparciu o multifraktalną charakterystykę zmienności kursów ich akcji.

W celu jak najlepszego zobrazowania ogółu zamysłu dla propozycji zawartej w tej pracy, pozwoliłem sobie na krótkie omówienie istoty oraz sposobów analizy ryzyka. Nawiązałem również do metod oceny niewypłacalności przedsiębiorstw opartych o analizę fundamentalną. Dalej, z dwojakich względów, postanowiłem omówić krótko specyfikę oraz możliwości handlu algorytmicznego, w szczególności w kontekście GPW S.A.. Powodami są: 1) wzgląd na istotny udział algorytmów w generowaniu obrotu na rynkach oraz kształtowaniu poziomu ich efektywności, 2) upatrywanie ewentualnego wykorzystania tej propozycji w ramach rozwijanych przez różnego rodzaju podmioty algorytmów inwestycyjnych.

Obraną przeze mnie metodą badania notowań jest multifraktalna analiza detrendowanych fluktuacji (*Multifractal Detrended Fluctuation Analysis* - MF-DFA), a klasyfikatorem uzyskanych w ten sposób charakterystyk metoda maszyn wektorów nośnych (*Support Vector Machines* - SVM).

Celem niniejszej pracy jest przedstawienie koncepcji, zaproponowanie i wyznaczenie wartości zmiennych charakteryzujących fraktalne własności szeregów czasowych akcji, utworzenie na ich bazie modelu dyskryminacyjnego oraz podjęcie próby weryfikacji zasadności dla złożonego w ten sposób algorytmu w wyłanianiu spółek, dla których natura zmienności kursów akcji wskazuje na istotne ryzyko ogłoszenia upadłości emitenta.

1.1 Istota ryzyka na rynku finansowym

Rynek finansowy w sposób pośredni przenika codzienność praktycznie każdej cywilizowanej jednostki ludzkiej. Wiele osób również bezpośrednio podejmuje obecności na rynkach walutowych, kapitałowych, dłużnych czy instrumentów pochodnych.

Dostępność rachunków maklerskich, jak i rachunków dla kontraktów terminowych stała się powszechna, a legislacja w tym zakresie bardzo liberalna. Istnieje również wiele platform transakcyjnych dla rynków walutowych, a jedynym ograniczeniem dla dostępności pozostaje posiadanie określonej minimalnej ilości pieniądza. Wiele z tych rynków podlega wysokiemu lewarowaniu, co skutecznie przyciąga kapitał spekulacyjny. Inwestorzy indywidualni alokują nadwyżki kapitału również w instytucjach wspólnego inwestowania, najczęściej w

postaci uczestnictwa w funduszach inwestycyjnych.

Z drugiej strony mamy przedsiębiorstwa, gdzie wzrost wolumenu wymiany transgranicznej towarów, kontraktowanie dostaw (dla różnych horyzontów czasowych) lub sprzedaży surowców skłaniają je do zawierania kontraktów futures, wchodzenia w strategie opcyjne czy umowy wymiany (*swap*).

Niezależnie od celu wszystkie wspomniane strony decydując się na obecność na rynku podejmują **ryzyko**. Podstawowym celem każdego z inwestorów jest przynajmniej ochrona kapitału. Zgodnie z klasycznym podejściem łączącym modele: CAPM, pocisk Markowitza oraz teorię użyteczności, racjonalny inwestor winien obierać portfel aktywów maksymalizujący oczekiwaną stopę zwrotu przy minimalnym ryzyku (awersja do ryzyka)[1][3][4].

Wielość koncepcji oraz miar ryzyka sprawiają jednak, że jego definicja staje się bardzo szeroką. W najprostszym ujęciu można by przyjąć równoznaczność tego terminu ze słowem niepewność, jednak ogranicza ono postrzeganie zjawiska. Rozważania w tym zakresie przyczyniły się do rozwoju dwóch głównych koncepcji ryzyka - negatywnej i neutralnej oraz dwóch ujęć w modelowaniu ryzyka - jako miary zmienności albo jako miary wrażliwości na czynniki ryzyka[1].

Negatywna koncepcja ryzyka spaja się z potocznym rozumowaniem, gdzie ponoszenie ryzyka tożsame jest z możliwością poniesienia straty. Przykładem takiego podejścia może być choćby koncepcja poziomu bezpieczeństwa, z której wywodzi się rodzina modeli wartości zagrożonej (*Value at Risk* - **VaR**). W ramach koncepcji VaR istotnego znaczenia nabrał model wartości oczekiwanej dla straty (*Expected Shortfall, Conditional Value at Risk* - **CVaR**)[2].

Modele VaR oparte są o poziom aspiracji, będący graniczną wartością akceptowalnego poziomu stopy zwrotu. Wartość zagrożona oraz wartość oczekiwana ewentualnej straty wyznaczane są z rozkładu stóp zwrotu waloru będącego podstawą inwestycji, a tym samym rozkładu zmienności cen rozważanego waloru. Ponieważ mówimy tutaj o ewentualnej stracie, istotny w modelowaniu jest lewy ogon tego rozkładu. Właściwe jego oszacowanie jest bardzo trudne, zwłaszcza w obliczu niestacjonarnych szeregów stóp zwrotu.

Neutralna koncepcja ryzyka wydaje się być znacznie bliższa osobom zaangażowanym na rynkach finansowych z pobudek innych niż zabezpieczenie wartości kapitału. Ryzyko w takim ujęciu jest interpretowane jako możliwość uzyskania stopy zwrotu innej niż oczekiwana. Może ona zatem być zarówno niższa (jak w koncepcji negatywnej), ale również przynieść korzyści wyższe od spodziewanych.

Ryzyko jako miara zmienności jest koncepcją bezpośrednio i wyłącznie nawiązującą do obserwowanych zmian cen waloru. Wszelkie wskaźniki budowane w nawiązaniu do tej koncepcji opierają się tylko na szeregu czasowym cen waloru, analizując jedynie łączny wpływ ewentualnych czynników na zmienność cen. Powszechnie przyjętą podstawową miarą ryzyka inwestycji jest poziom fluktuacji cen rozumiany jako odchylenie standardowe stopy zwrotu obliczone dla stopy zwrotu z waloru lub z portfela¹. Miara ta jest wykorzystywana przez klasyczny model wyceny aktywów kapitałowych (*Capital Asset Pricing Model - CAPM*)[3], z którą koresponduje koncepcja pocisku Markowitza (*efficient frontier*), prowadząca do wyznaczenia maksymalnego stosunku oczekiwanej wartości stopy zwrotu do zmienności portfela (koncepcja linii rynku kapitałowego, *Capital Market Line - CML*)[4].

Klasyczna ocena efektywności zarządzania portfelem również wykorzystuje ryzyko rozumiane jako zmienność cen (w tym **ryzyko systematyczne**), czego dobrymi przykładami są popularne wskaźniki Sharpe'a, Treynora, Blacka - Treynora, Jensena, Information Ratio, czy Sortino Ratio[5].

Jednak, jak zauważył Mandelbrot, następstwami niskich (wysokich) poziomów zmienności, są zazwyczaj niskie (wysokie) poziomy zmienności - mamy do czynienia ze swego rodzaju gromadzeniem zmienności (*volatility clustering*). Z ekonomicznego punktu widzenia okresy, w których zmienność utrzymuje się na wysokim poziomie, wiązać można w następujący sposób[16]:

- wysoka zmienność oznacza zwiększenie ryzyka, którego wynikiem jest obniżenie cen aktywów (efekt ryzyka);
- niższe ceny redukują poziom majątności, co prowadzi do wzrostu awersji do ryzyka, czego skutkiem jest zwiększenie dywersyfikacji (efekt majątności);

¹Niekiedy używane jest w analizach również odchylenie przeciętne, odchylenie ćwiartkowe lub semiodchylenie stopy zwrotu.

- w trakcie spadków cen na całym rynku wzrasta zmienność, co zmniejsza możliwości dywersyfikacji (efekt dywersyfikacji);
- w okresach wysokich zmienności zmniejsza się płynność rynku, co wymusza na sprzedających godzenie się na zawieranie transakcji mimo znacznej redukcji ceny (efekt płynności).

Spostrzeżenia Mandelbrota wprowadzają element wątpliwości w zakresie trafności modelowania i oceny efektywności z użyciem wspomnianych wyżej metod klasycznych.

Koncepcja ryzyka jako miary wrażliwości powszechna jest wśród analityków instrumentów pochodnych, w szczególności opcji, a także analityków instrumentów dłużnych[1].

W przypadku opcji, wycena modelem Blacka - Scholesa - Mertona (**BSM**) (lub jakimkolwiek modelem wywodzącym się z BSM, np. modelem Garmana - Kohlhagena, czy chociażby modelem Blacka - Karasińskiego) uwzględnia co najmniej wpływ: ceny instrumentu podstawowego, czasu do wygaśnięcia opcji, zmiany stóp procentowych, zmienności cen. Zmienność ceny opcji względem zmienności wymienionych czynników powszechnie oznaczana jest literami greckimi, a współczynniki nazywane są wspólnie grekami. Ponieważ wyrażają one zależność zmian ceny opcji od zmian czynników ryzyka, są zatem miarami wrażliwości ceny opcji[6].

Podobnie rzecz się ma w przypadku instrumentów dłużnych. Najczęściej wyliczaną miarą wrażliwości ceny obligacji jest jej zależność od zmienności stopy procentowej i nazywana jest czasem trwania (*duration*). Miara została wyznaczona przez Macaulaya i popularnie zwana jest **duration Macaulaya**, dla odróżnienia od wielu jej pochodnych, jak np. *modified duration*, *key rate duration*, *dollar duration*, czy *effective duration*[7].

Duration jest pochodną pierwszego stopnia, zatem uwzględnia wpływ czynnika ryzyka w sposób liniowy, jednak przy znacznych zmianach stopy procentowej istotny staje się również wpływ czynników wyższego stopnia. Z tego względu analiza obligacji uwzględnia również drugą składową rozwinięcia funkcji ceny od stopy procentowej w szereg Taylora (druga pochodna ceny względem zmiany stopy procentowej) popularnie nazywaną **wypukłością** (*convexity*).

Uogólniając, powyższe ujęcie sugeruje, że ryzyko pojmowane jako miara wrażliwości może być wyrażone przez k -tą pochodną cząstkową funkcji ceny po i -tym czynniku ryzyka:

$$\frac{\partial^k P}{\partial x_i^k}. \quad (1.1)$$

W większości przypadków analizowane są liniowe wpływy czynników. Pojawiają się jednak również zmienności stopnia drugiego, jak wspomniana wypukłość - w przypadku obligacji, czy chociażby **współczynnik gamma** - wyrażający wpływ zmienności ceny opcji w zależności od zmiany zmienności instrumentu podstawowego dla tej opcji.

Podsumowując, ryzyko jest dla profesjonalnych uczestników rynku oraz inwestorów indywidualnych kluczowym problemem, a jego identyfikacja i kwantyfikacja stała się szczególnie istotna. Nie bez powodu zatem, istotną rolę w podejmowaniu decyzji przez podmioty rynkowe zaczyna pełnić zarządzanie ryzykiem (*Risk Management* - RM), w którym silne stały się takie obszary jak zarządzanie ryzykiem finansowym (*Financial Risk Management* - FRM), zarządzanie ryzykiem przedsiębiorstwa (*Enterprise Risk Management* - ERM), zarządzanie ryzykiem energetycznym (*Energy Risk Management*). W niniejszej pracy będziemy koncentrować się na obszarze FRM, jednak wymieniane niżej koncepcje analizy i prognozowania ryzyka w znacznej mierze przekrywają się w całości RM.

1.2 Metody analizy i prognozowania ryzyka

Modelowanie rozkładów stóp zwrotu (zmienności cen) jest bez wątpienia najistotniejszym z elementów prognozowania w świecie finansów, a zarazem jednym z najtrudniejszych zadań stojących przed analitykami oraz zarządzającymi portfelami inwestycyjnymi.

Ocena poziomu ryzyka ma wspomagać procesy decyzyjne - stanowi punkt wyjścia dla wyznaczania oczekiwanych zysków, wartości zagrożonej, oczekiwanej wartości ewentualnej straty, które wyznaczane mogą być zarówno dla pojedynczego waloru, jak i całego portfela. Praktyka w tym zakresie jest szeroka i obejmuje m.in. podejścia oparte o:

1. zmienność historyczną (wykorzystuje historyczne szeregi czasowe zmienności cen w celu określenia kształtu rozkładu warunkowego), gdzie wyróżnia się podejścia:
 - parametryczne (zakłada określony rodzaj rozkładu),
 - nieparametryczne (wykorzystuje rozkład otrzymany na podstawie danych historycznych),
 - hybrydowe (połączenie obu powyższych);
2. zmienność implikowaną (wyznaczana przy użyciu modeli wyceny instrumentów pochodnych oraz aktualnych cen)[8];
3. kombinację zmienności historycznej oraz zmienności implikowanej[9].

W ramach pierwszej linii analityki stosowane są przykładowo modele: auto-regresywnej średniej ruchomej (*autoregressive moving average* - **ARMA**), eksponencjalnie ważonej średniej ruchomej (*exponentialy weighted moving average* - **EWMA**²) oraz jej uogólnienia - rodziny modeli **GARCH** (*Generalized Auto-Regressive Conditional Heteroskedasticity model*), symulacje historyczne (*historical simulation* - **HS**) oparte o określone okna czasowe, ważne ekspercko wielowymiarowe estymacje rozkładu gęstości (*multivariate density estimation* - **MDE**), albo modele hybrydowe, które łączą podejścia RiskMetrics oraz HS.

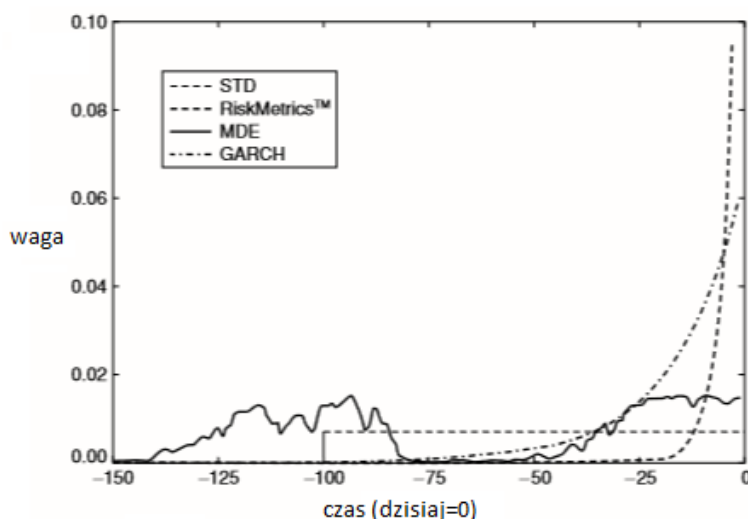
Z kolei przykładem modeli łączących zmienność implikowaną oraz wynikającą z danych historycznych jest model korygujący historyczną średnią krocząca indeksem zmienności (*istorical moving average - volatility index adjustment model* - **HMA-VIX**)[14].

Wszystkie te sposoby traktują ryzyko jako miarę zmienności cen i są powszechnie stosowane w modelach rodziny VaR. Wykorzystują one historyczne poziomy zmienności, przypisując ryzyku rejestrowanemu w poszczególnych chwilach czasu określone wagi (patrz Rysunek 1.1).

Zmienność implikowana jest wyznaczana na podstawie modelu wyceny danego rodzaju aktywa³. Tym samym, można wiązać ją z koncepcją wrażliwości. Najczęściej pojmowana jest jako wynik nieefektywności rynku, albo czysto

²Miara zaproponowana została w wywodzącym się z JP Morgan RiskMetrics, stąd często nazywana jest modelem **RiskMetrics**.

³Często spotykaną konsekwencją stosowania metod zmienności implikowanej jest występowanie zjawiska uśmiechu zmienności (*volatility smile*).



Rysunek 1.1: Rozkład współczynników wagowych przypisywanych poziomom ryzyka w poszczególnych chwilach czasu w modelowaniu ryzyka z użyciem RiskMetrics, MDE, GARCH (źródło: [8]).

stochastycznej zmienności. W ramach takiego rozumowania przypisuje się jej często składniki **zmienności rzeczywistej** oraz premii wynikającej ze **zmienności stochastycznej**.

Błędnie oszacowany poziom zmienności, albo szerzej - niewłaściwa ocena natury zmienności, może być dla podmiotów bazujących na modelach, a w szczególności posiadających ekspozycje w postaci silnie lewarowanych pozycji, katastrofalna. Trudność w modelowaniu zmienności dla rynków finansowych wynika, przede wszystkim, z braku stacjonarności jej rozkładu w czasie. Natura zmienności rozkładu stóp zwrotu jest zazwyczaj trudna, albo wręcz niemożliwa do zapisu modelem matematycznym.

Innego rodzaju podkreśleniem nietrywialności problemu jest fakt, że w ostatnich trzydziestu latach obserwowany jest silny rozwój modeli wykorzystujących złożone narzędzia analityczne, a instytucje finansowe działające w skali globalnej chętnie zatrudniają w roli analityków ilościowych (*quantitative analysts*, powszechnie *quants*) naukowców zajmujących się analizami układów złożonych, w szczególności przewidywaniem pogody, aktywnością sejsmiczną,

dynamiką sieci społecznych, itp.

Nurty metod analizy ryzyka rozlewają się w wielu kierunkach, w szczególności ze względu na zmiany techniczne i technologiczne, które zaszły w otoczeniu giełd oraz na samych parkietach. W wielu instytucjach intensywnie rozwija się złożone algorytmy decyzyjne, które dzięki wyłączeniu czynnika ludzkiego, mogą osiągać znacznie lepsze rezultaty oraz zarządzać pozycjami w czasie ułamka sekundy. Handel algorytmiczny, o którym szerzej będzie mowa w dalszej części pracy, wzbudza dziś wiele kontrowersji, jednocześnie elektryzuje rynek złożonością i domniemaną efektywnością.

1.3 Prognozowanie niewypłacalności przedsiębiorstw

Jednym z podstawowych rodzajów ryzyka, jaki towarzyszy inwestowaniu w określoną spółkę, jest możliwość ogłoszenia jej upadłości. Akcjonariusze są jednymi z ostatnich w szeregu uprawnionych do masy upadłościowej przedsiębiorstwa[46].

Upadłość przedsiębiorstwa nie jest zjawiskiem występującym w sposób nagły. Powstanie zagrożenia bankructwem poprzedza zwykle długotrwały kryzys w firmie, który w rozumieniu definicji proponowanej przez Zimniewicz[47], składa się z ciągu zdarzeń stanowiących zagrożenie dla egzystencji przedsiębiorstwa lub jej uniemożliwienia. Zaistnienie sytuacji kryzysowej można w ogólności ująć jako nieprawidłowe dostosowanie wewnętrznych zasobów przedsiębiorstwa do zmieniającego się w sposób ciągły otoczenia.

Ze względu na długość i zmienność dynamiki procesu, jakim jest kryzys przedsiębiorstwa, Zimniewicz wyróżnia cztery jego fazy[47]:

1. kryzysu potencjalnego,
2. kryzysu ukrytego,
3. kryzysu ostrego,
4. końcowa – bankructwo.

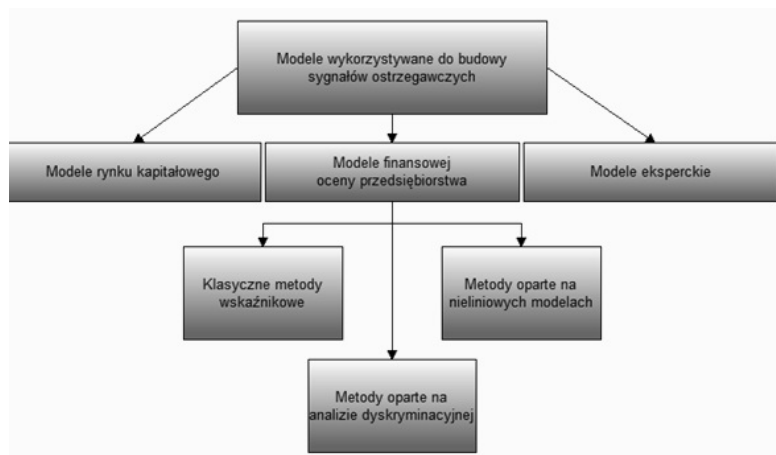
Często niewykrywalnymi okazują się różnice w funkcjonowaniu przedsiębiorstwa, które miałyby świadczyć o wystąpieniu fazy pierwszej kryzysu. Za-

zwyczaj są to odchylenia jakościowe lub ilościowe, którym, w przypadku wykrycia, kierownictwo może zaradzić. Ze względu na wielkość odchyień ich wykrycie stanowi pewną trudność, w skutek czego mogą pozostać one niedostrzeżone w odpowiednim momencie. Z drugiej strony niewłaściwe ustosunkowanie się do zaobserwowanych nieprawidłowości, może również mieć negatywne skutki, co może wprowadzić przedsiębiorstwo w fazę kryzysu ukrytego.

W drugiej fazie istnieje większe prawdopodobieństwo wykrycia anomalii w funkcjonowaniu firmy. Nieumiejętne lub zbyt powolne reagowanie na zaistniałą sytuację, kieruje przedsiębiorstwo ku fazie kryzysu ostrego.

Trzecia faza objawia się kumulacją trudności w prowadzeniu przedsiębiorstwa, którym towarzyszyć może: wzrost zapasów wyrobów, wzrost kosztów, malejąca sprzedaż, trawienie rynków zbytu, niezrealizowanie zadań finansowych. Zaistniałe czynniki w sposób znaczący ograniczają możliwości zmian w przedsiębiorstwie. Nieprzezwyciężenie kryzysu w fazie trzeciej prowadzi do bankructwa firmy[48].

Istnieje co najmniej kilka różnych typów modeli wczesnego ostrzegania, opracowanych na potrzeby analizy finansowej do generowania sygnałów o istotnym ryzyku upadłości. Klasyfikacja zaproponowana przez Siedleckiego[42] wyróżnia trzy podstawowe ich rodzaje (patrz Rysunek 1.2).



Rysunek 1.2: Rodzaje modeli wykorzystywanych w tworzeniu systemów wczesnego ostrzegania (źródło: [42]).

Modele oparte na rynku kapitałowym to przede wszystkim modele wyceny akcji i opcji oraz analizy ryzyka. Jednymi z najpopularniejszych są: model **EDF** (*Expected Default Frequency*; oparty na koncepcji wyceny opcji i obligacji – najczęściej wykorzystujący model BSM) oraz VaR. Modele te, jak wcześniej opisywano, wykorzystują implikowaną zmienność cen, albo prognozy ryzyka opracowane na podstawie metod zmienności historycznych.

Z kolei modele eksperckie to głównie metody oparte na subiektywnej ocenie specjalisty, z czego najpopularniejszymi są metody punktowe.

Wśród metod opartych na analizie finansowej przedsiębiorstwa, Siedlecki wyróżnił trzy rodzaje metod: wskaźnikowe, analizy dyskryminacyjnej oraz nieliniowe.

Podstawowym warunkiem zastosowania klasycznej **metody wskaźnikowej**, jako metody prognostycznej, jest poprzedzająca ją analiza strategiczna przedsiębiorstwa. Istotna jest tutaj właściwa ocena dalszego i bliższego otoczenia analizowanego podmiotu. Dzięki temu możliwe jest dokonanie wyboru właściwych wskaźników finansowych, sygnalizujących pogarszającą się sytuację firmy. Ocena wartości uzyskiwanych dla poszczególnych wskaźników, w kontekście zagrożenia upadłością, powinna uwzględniać aktualną fazę cyklu życia analizowanego podmiotu.

W tego rodzaju metodach wykorzystuje się wskaźniki odnoszące się do wszystkich finansowych aspektów działalności przedsiębiorstwa, obejmując: płynność, rentowność, zadłużenie i zasilanie finansowe, sprawność działania, wskaźniki rynku kapitałowego. Najczęściej wykorzystywanymi są: rentowność kapitałów całkowitych, poziom ogólnego zadłużenia, wskaźnik kapitału obrotowego oraz wskaźniki rotacji (zapasów, należności, zobowiązań).

W ramach klasycznych metod wskaźnikowych stworzone zostały również metody opierające się na zagregowanych wskaźnikach finansowych, do których należą m.in.: analiza Du Ponta, test szybki (quick test), model stopnia płynności względnej (SPW). Modele te mogą wykorzystywać kombinacje pojedynczych wskaźników (model Du Ponta, model SPW) lub ujmować w sposób syntetyczny pojedyncze oceny poszczególnych wskaźników, sumując punkty przez nie uzyskane (test szybki)[42].

W ostatnich latach dużym powodzeniem cieszyły się metody oceny przedsiębiorstw oparte na metodzie sztucznych sieci neuronowych (SSN). Koncepcja

ta stanowi bardzo uproszczony model mózgu. SSN, podobnie jak biologiczny organ, składa się z pewnej ilości elementów (neuronów) przetwarzających informacje. Każdy z neuronów przetwarza dane wejściowe poprzez funkcję aktywacji (z reguły funkcja ta jest postaci funkcji: liniowej, tangensa hiperbolicznego, logistycznej, Heaviside'a, gaussowskiej). Istotnym elementem są połączenia sieci, poprzez które następuje transmisja informacji z wyjścia jednego neuronu do wejścia kolejnego. Siła poszczególnych połączeń zmienia się w trakcie procesu uczenia sieci tak, aby osiągnąć określone wartości parametrów modelowanego zjawiska[42].

Koncepcja sztucznych sieci neuronowych znalazła zastosowanie również w prognozowaniu upadłości przedsiębiorstw. Jeden z pierwszych artykułów poświęconych temu zastosowaniu SSN[49], porównuje wyniki stworzonego przez autorów artykułu modelu sieci z wynikami modelu dyskryminacyjnego Z-score Altmana⁴. Procentowe ujęcie trafności prognoz dla przebadanych 128 przedsiębiorstw amerykańskich, wskazało na model SSN jako efektywniejszy. Jednak nie wszyscy autorzy porównań uznają wyższość koncepcji SSN nad liniowymi modelami funkcji dyskryminacyjnych. Trzy spośród pięciu prac badawczych występujących w zestawieniu zaprezentowanym przez Siedleckiego[42], wskazują na wyższą skuteczność analizy dyskryminacyjnej.

Inną, wartą wspomnienia, propozycją prognozowania kondycji przedsiębiorstw jest podejście zaproponowane przez Klobucnika i Sieversa[50], które wykorzystuje modelowanie procesów stochastycznych (*bankruptcy prediction based on stochastic processes* - BPSP). Koncepcja ta wymaga jednak zaprojektowania wielu procesów stochastycznych (modelujących m.in. poziomy wartości: sprzedaży, kosztów, czy kapitału pracującego netto), co sprawić może w praktyce, że metodyka będzie swoista dla każdej spółki. Z tego względu uzasadnione wydaje się być wykorzystanie tej propozycji jedynie w prognozowaniu wewnątrz organizacji lub w przypadku dużych pozycji inwestycyjnych.

Bardzo istotną słabością powyżej przedstawionych koncepcji opartych o analizę fundamentalną jest względnie długi dystans czasu między kolejnymi datami sprawozdawczymi. Obrót giełdowy dokonywany jest praktycznie każdego dnia. Pojawiające się pogłoski oraz transakcje insiderów wyceniane są bezwzględnie przez rynek. Reakcja następująca po publikacji wyników jest zazwy-

⁴patrz podrozdział 3.2.

czaj jedynie korektą zbiorczych oczekiwań.

W przypadku stosowanych obecnie modeli opartych na analizie stochastycznej, ocenia się głównie poziom zmienności. Braknie oceny natury obserwowanych zmian cen. W założeniu niniejszej pracy, podobnie jak i hipotezy rynku fraktalnego⁵, w obliczu znacznych niepewności co do wyników spółki, również taka informacja jest zawarta w cenie waloru. Poprawne scharakteryzowanie zachowań zmienności ceny waloru może prowadzić do wyłonienia powyższych informacji z samego szeregu cen.

1.4 Algorytmiczny obrót instrumentami finansowymi

1.4.1 Wprowadzenie do handlu algorytmicznego

Początkowo algorytmy były wykorzystywane w handlu instrumentami finansowymi do optymalizacji momentu wprowadzenia zlecenia na rynek, ewentualnej jego fragmentacji, itp. Nie podejmowały one decyzji inwestycyjnej, a jedynie realizowały polecenie człowieka. Nazywano je systemami automatycznej realizacji (*Automated Execution System* - **AES**).

W 2001 roku Amerykańska Komisja Papierów Wartościowych i Giełd (*Securities and Exchange Commission* - SEC) wprowadziła decymalizację. Z obowiązujących dotychczas szesnastych części dolara przeskoku cenowego, rynek zaczął pracować na minimalnym spreadzie w wysokości 0,01 dolara.

Handel algorytmiczny zaczął ewoluować w kierunku samodzielnego podejmowania decyzji przez maszyny. Transakcje były przeprowadzane znacznie sprawniej (głównie dzięki eliminacji czynnika ludzkiego). Zwiększenie szybkości w podejmowaniu decyzji inwestycyjnych oraz ich realizacji w warunkach niskiego spreadu, zwiększyło ilość zawieranych transakcji. Doprowadziło to do znacznego zwiększenia konkurencji w zakresie obniżania kosztów transakcyjnych. Oba czynniki wpłynęły z kolei na zmniejszenie przeciętnej wielkości zlecenia, co również miało pozytywny wpływ na dalszy rozwój zawierania transakcji w oparciu o algorytmy. Dzięki automatyzacji procesów, w tym analizy pod

⁵patrz podrozdział 3.2.

kątem strategii inwestycyjnych, przeciętny czas utrzymania pozycji zaczął spadać. Mniej kaloryczne transakcje były obserwowane coraz częściej, co, *ceteris paribus*, zwiększyło płynność poszczególnych walorów.

Większa płynność jest wiązana teoretycznie ze zmniejszeniem zmienności instrumentu, co oznacza zmniejszenie ryzyka inwestycyjnego. W ogólności, operowanie na niższym ryzyku systematycznym wiąże się ze zmniejszeniem oczekiwanych stóp zwrotu z portfela efektywnego. Taka sytuacja prowadzi z kolei do ponownego podwyższenia konkurencyjności między instytucjami obecnymi na rynku, które z natury chcą osiągać ponadprzeciętne stopy zwrotu[39].

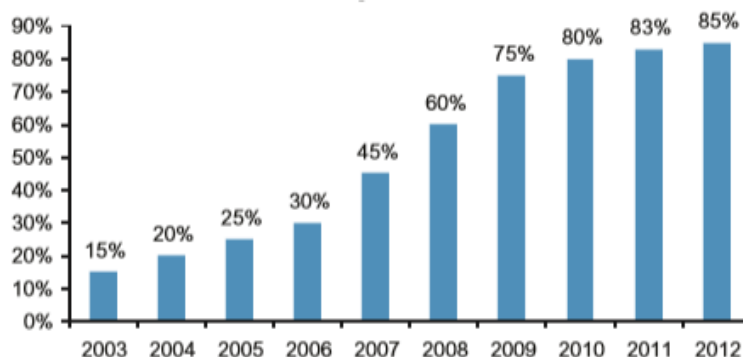
W warunkach wyższej konkurencyjności często poszukiwane są nowe, efektywne rozwiązania. Ścieżka rozumowania wraca w tym momencie do handlu algorytmicznego⁶, w ramach którego:

1. wszelkie decyzje inwestycyjne mogą być podejmowane przy minimalnym opóźnieniu (dokonywanie transakcji w oparciu o nieustannie zbierane parametry stanu rynku dzieje się przy eliminacji czynnika ludzkiego);
2. dobór optymalnej strategii może być na bieżąco monitorowany;
3. możliwe są nieustanne testy dostępnych strategii oraz optymalizacja w ramach porównań typu *champion-challenger*.

W taki oto sposób udział handlu algorytmicznego dynamicznie rósł w ostatnich latach - szacuje się, że na NYSE wyniósł on w 2012r. 85% (patrz Rysunek 1.3).

Warto wspomnieć, że zazwyczaj maszyna egzekwująca algorytmy (*Black Box* - **BB**) nie obiera jednej optymalnej strategii, ale dokonuje transakcji na wiele z dostępnych mu sposobów[55]. Strategie dające lepsze wyniki są wykorzystywane częściej, sumarycznie operując większym wolumenem. Mniej efektywne strategie niekoniecznie od razu są porzucane, ponieważ również mogą nieść istotną wartość poznawczą dla BB. Najczęściej są to bowiem inteligentne maszyny wykorzystujące statystyczne techniki uczenia i samodzielnie tworzące testowane strategie, dla których czynnikiem ludzkim jest jedynie pewien zbiór wartości skrajnych dla parametrów przez nie wykorzystywanych.

⁶Innym istotnym czynnikiem, który przyczynił się do dalszego rozwoju handlu algorytmicznego, a który w niniejszym opracowaniu został pominięty, jest zaistnienie alternatywnych systemów generowania obrotu w ramach tzw. ukrytej płynności (*dark pools, dark liquidity*).



Rysunek 1.3: Udział handlu algorytmicznego w ogóle wolumenu obrotu NYSE w poszczególnych latach (źródło: [14]).

Dla lepszego zrozumienia zarysu sposobu działania handlu algorytmicznego rozważmy przykładowy algorytm prymitywny (*dummy algo*).

Jednym z podstawowych parametrów rynku/instrumentu, którym mogą sugerować się algorytmy jest średnia cena ważona wolumenem (*Volume-Weighted Average Price - VWAP*)[14][16][55]. Prymitywny algorytm traktujący ten wskaźnik jako sygnał kupna albo sprzedaży mógłby egzekwować następujące działania:

1. jeżeli w arkuszu dostępna jest oferta sprzedaży instrumentu X w cenie mniejszej lub równej wartości oczekiwanej rozkładu śród-dziennego (*intraday*) VWAP ⁷ ($E(iVWAP)$) pomniejszonej ⁸:
 - o ustaloną kwotę, albo
 - ustalonym mnożnikiem z zakresu $[0;1)$, albo
 - określonym percentylem rozkładu $iVWAP$, albo

⁷W ogólności, rodzaj rozkładu VWAP (śród-dzienny, dzienny, tygodniowy, itd.) winien odzwierciedlać przyjęty horyzont inwestycyjny. W tym wypadku posłużono się horyzontem w wysokości jednego dnia obrotu, dla którego rozkładem powiązany jest śród-dzienny rozkład VWAP

⁸Wartość $E(iVWAP)$ po pomniejszeniu oznaczmy jako $E^-(iVWAP)$.

- inaczej określona wartość poniżej $E(iVWAP)$,

dokonaj transakcji kupna.

2. Jeżeli minimalna cena sprzedaży w arkuszu zleceń instrumentu X znajduje się powyżej $E^-(iVWAP)$, ale maksymalna cena ofert kupna znajduje się poniżej $E^-(iVWAP)$, umieść zlecenie kupna w cenie wyższej od obecnie proponowanej maksymalnej oferty kupna, jednak nie wyższej niż cena $E^-(iVWAP)$.
3. Jeżeli w arkuszu zleceń maksymalna oferta kupna przekracza wartość $E^+(iVWAP)$ (określona jako wartość $E(iVWAP)$ powiększoną w sposób analogiczny do sposobu pomniejszenia $E(iVWAP)$ do $E^-(iVWAP)$), wystaw zlecenie sprzedaży w tej cenie⁹.
4. Jeżeli w arkuszu zleceń minimalna oferta sprzedaży znajduje się powyżej ceny $E^+(iVWAP)$, wystaw zlecenie sprzedaży w cenie niższej od obecnego minimum, ale nie niższej niż $E^+(iVWAP)$.

Założeniem powyższego algorytmu jest tendencja powrotu ceny, w ramach rozważanego horyzontu czasowego, do wartości średniej. W warunkach utrzymywania się ceny instrumentu w trendzie bocznym (inaczej: poziomy kanał zmienności, ograniczona fluktuacja wokół stałej wartości średniej), strategia oparta o VWAP winna przynosić ponadprzeciętne zyski, ponieważ inwestor dokonywałby zakupu poniżej ceny przeciętnej oraz sprzedaży powyżej tej ceny.

Innymi przykładami parametrów obrotu instrumentami są: średnia cena ważona czasem (*Time-Weighted Average Price* - TWAP), czy chociażby algorytm przyływu informacji w mediach (*newsflow algo*), czy zmienności natężenia określonego rodzaju ruchu w Internecie (np. częstotliwość wyszukiwania informacji na temat określonej spółki).

Strategie algorytmiczne uwzględniają również wpływ potencjalnej transakcji na rynek, a także modelują koszty transakcyjne. Sprawdzany jest także ewentualny sposób wykonania, co szczególnie istotne jest dla bardziej kalorycznych zleceń. Zlecenie może zostać rozłożone przez system giełdowy na kilka

⁹W przypadku braku posiadania stosownej liczby jednostek instrumentu X lub braku posiadania instrumentu X, w ogóle wystawione może zostać zlecenie krótkiej sprzedaży o ile dla tego instrumentu zoszło tego rodzaju działanie dopuszczone przez giełdę.

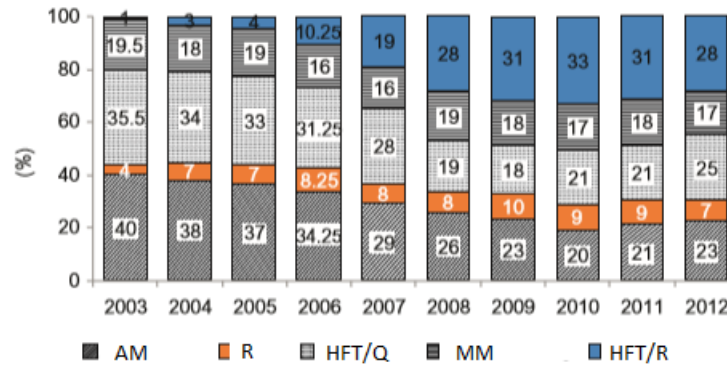
mniejszych, jeśli po przeciwnej stronie arkusza nie znajduje się zlecenie powodujące całkowitą realizację zlecenia wstępującego. W takiej sytuacji możliwe jest, że część przesłanego zlecenia będzie oczekiwała na realizację (czas oczekiwania może stanowić zmienną wyrażającą czynnik ryzyka).

W przypadku fragmentacji i oczekiwania przez część zlecenia na realizację, istnieje również ryzyko, że ostatecznie osiągnięta przeciętna cena realizacji zlecenia będzie różniła się od pierwotnie zakładanej. Wynika to z faktu, że poszczególne części zlecenia mogą być realizowane po różnych kwotach, w zależności od składanych do arkusza przeciwstawnych zleceń.

Wśród typów handlu opartego o maszyny, szczególnie istotnym jest ilościowy handel algorytmiczny (*Quantitative Algorithmic Trading*), opierający się głównie o modele statystyczne, których celem jest wykonanie zlecenia przy minimalizacji zidentyfikowanych rodzajów ryzyka. Jednocześnie, jako że tego rodzaju modele zarządzają pozycją, a nawet portfelem, winny one charakteryzować się wysokim stopniem immunizacji na zmienność rynku. Może to być osiągnięte dzięki minimalizacji wartości zmiennych ryzyka, albo jego mitygacji dzięki dywersyfikacji lub zabezpieczeniu pozycji instrumentami pochodnymi (*hedge*).

W ostatnich latach dużym zainteresowaniem cieszą się algorytmy dokonujące obrotu na poziomie wysokich częstotliwości (*High Frequency Trading - HFT*)[14](patrz Rysunek 1.4). Horyzont czasowy przypisywany pozycjom przez nie zajmowanym waha się od milisekund do kilku godzin. Ze względu na znaczną ilość obrotu, jaka jest przez nie realizowana, algorytmy HFT uważane są za jedno ze źródeł wspomaganie osiągania efektywności rynku. Metodyki BB realizujących HFT podlegają ścisłej tajemnicy, zarówno ze względu na zawarty w nich know-how, jak i kontrowersyjny sposób ich działania[40].

Algorytmom HFT przypisuje się wykorzystywanie strategii typu *front-running*, polegającej na zajęciu pozycji przez algorytm przed inwestorem wcześniej wysyłającym zlecenie. Wy tłumaczenia tego zjawiska dopatrywać można w zróżnicowaniu szybkości przesyłu informacji oraz opóźnień w jej przetwarzaniu. Wprost ujmując, algorytmy HFT wyłapują zlecenia kierowane na rynek i są na nim wcześniej. Dzięki temu, to one są stroną sprzedającą, mogą zatem na transakcji zarobić. Zysk z takiej pojedynczej transakcji jest najprawdopodobniej minimalny, jednak horyzont czasowy pozycji jest rzędu milisekund, co oznacza, że dziennie możliwe jest przeprowadzenie setek lub tysięcy tego rodzaju



Rysunek 1.4: Udział HFT oraz innych rodzajów handlu algorytmicznego w ogóle obrotów generowanych tym sposobem w poszczególnych latach. (źródło: [14]).

transakcji¹⁰.

Do szerokiej publiczności informacja o istnieniu tego rodzaju maszyn dobiegła 6 maja 2010 roku, kiedy to o godzinie 2:45 miało miejsce gwałtowne załamanie się cen spółek DJIA, powodujące zmianę indeksu o -10%, po którym w ciągu minuty nastąpiło wybicie niwelujące spadek. Niektóre ze spółek, jak np. P&G rejestrowały spadki cen przekraczające -30% (patrz Rysunek 1.5). Zdarzenie zostało określone mianem Krachu 2:45 (*The Crash of 2:45*, często nazywany również *2010 Flash Crash*). Powszechnie uważa się, że zdarzenie było wynikiem działania HFT - trudno określić, czy wynikającym z błędów algorytmów, czy też wytworzenia nierównowagi rynkowej przez jedną maszynę lub ich grupę¹¹.

¹⁰W ostatnich latach zaczęła rozwijać się giełda **IEX**, której wyróżnikiem jest gwarancja stałej wartości opóźnienia czasowego przyjmowanych zleceń. Dzięki temu zabiegowi strategię typu *front-running* mogą być skutecznie eliminowane.

¹¹Raport SEC wskazał na złożenie błędnego zlecenia o dużej wartości przez jeden z funduszy, jednak w ocenie przynajmniej niektórych analityków jest to niewiarygodne.



Rysunek 1.5: Zmiany cen PG oraz Accenture w dniu Krachu 2:45 (źródło: [58]).

1.4.2 GPW S.A. - kształtowanie specyfiki handlu w wyniku zmian technicznych i technologicznych

System zawierania transakcji na GPW S.A. zmienił się istotnie we wrześniu 2013 - momencie, w którym Giełda przeszła z WARSETu na UTP – system o innej specyfice, wielokrotnie zwiększonej mocy zawierania transakcji w stosunku do poprzednika.

W ramach systemu WARSET, wdrożonego 17 listopada 2000 roku, możliwe było zawieranie odpowiednio 40 lub 45 zleceń na sekundę w ramach systemu notowań ciągłych w jego fazie głównej lub fazie sesji przed otwarciem, co stanowiło około 10 – krotne zwiększenie mocy w tym aspekcie w stosunku do poprzedniego systemu[13] opracowanego przez Société de Bourse Français oraz spółkę Euronext.

W czerwcu 2007 roku wdrożono zmianę umożliwiającą zawieranie 180 zleceń na sekundę podczas fazy notowań ciągłych[30]. Kolejna zmiana zwiększyła

ten parametr do wartości 850[31]. Dziś UTP przetwarza na polskim parkiecie 20 000 zleceń na sekundę. WARSET był także systemem stosunkowo zawodnym, co, *ceateris paribus*, podnosiło ryzyko operacyjne związane z efektywnością inwestycji krótkoterminowych na rodzimym rynku kapitałowym.

W obrocie opartym o WARSET nie istniała możliwość składania dostępnych w UTP zleceń typu PEG, WNF, WNZ (VFC), WDC, cross[32]. Inna była specyfika działania członków giełdy wykonujących działania o charakterze animatora emitenta, czy animatora rynku, która została w ramach UTP przekształcona do osobnego systemu zawierania składania zleceń i zawierania transakcji.

Wraz z UTP pojawiła się techniczna możliwość zastosowania strategii inwestycyjnych opartych o handel algorytmiczny. Dla efektywnego działania algorytmów przesunięcia czasowe (**latencje**) wynikające z przesyłu, przetworzenia i responsu na zadane zlecenie mają kluczowe znaczenie[33]. Ciekawym i ważnym tego przykładem jest Project Exspress[34], budowany za 400mln USD ma zmniejszyć czas przesunięcia czasowego między Nowym Jorkiem a Londynem o 5ms[31]. Podobnych do tego przykładów jest oczywiście więcej. Połączenia między Tokio, Nowym Jorkiem, Szanghajem, czy wyspecjalizowanym w derywatach Seulem – bo każdorazowe wyprzedzenie konkurencji o choćby jedną milisekundę, pozwala zarobić. Odległość Warszawy od Londynu, czy Frankfurtu nie jest aż tak znaczne, natomiast oprócz szybkiego przesyłu informacji, istotna jest szybkość jej przetworzenia.

Przesyłana informacja musi zostać odpowiednio przetworzona przez system implementujący strategię inwestora, by ostatecznie zlecenie trafiło na giełdę. Dzieje się to za pośrednictwem omawianych wcześniej maszyn algorytmów, w szczególności BB. Przyjmijmy, że *Black Box* zbiera informacje z parkietu, na którym działa oraz z innych parkietów; jego połączenia, w dużym uproszczeniu, to dwie linie wejścia (z parkietem, na którym operuje oraz z otoczeniem tego parkietu¹²) oraz jedna wyjściowa (z parkietem, na którym operuje).

¹²Za jedną z funkcji Black Box'ów uważane jest również wytwarzanie szumu w otoczeniu giełdy, na której działa. Polega ono przykładowo na algorytmicznym wytwarzaniu sugestywnych artykułów dotyczących poszczególnych spółek. Artykuły te, wyłapane przez inne Black Boxy mogłyby wprowadzić je w błąd, co do oceny otoczenia. W przypadku udanego zaszumienia otoczenia, Black Box generujący szum może czerpać korzyści z nierównowagi w dostępie do informacji.

Celem jest minimalizacja przesunięcia czasowego, zatem logiczne jest, aby BB znalazł się możliwie blisko swojego parkietu transakcyjnego – w **kolokacji** z nim. Na GPW stało się to możliwe po przejściu na UTP, dzięki usłudze HPA (High Performance Access)[35]. Efektywność odbioru informacji giełdowej została podniesiona, w wyniku wprowadzenia usługi XDP (Exchange Data Publisher)[36]. Możliwe stało się także przeniesienie części ciężaru obliczeniowego, związanego z ograniczaniem niektórych parametrów ryzyka, poprzez usługę RMA (Risk Management Access)[37].

Wszelkie implementacje algorytmów HFT, ale także i efektywne działanie maszyn niskich częstotliwości, byłyby niemożliwe na GPW opartej o WARSET. Jak pokazują szacunki, w 2012 roku handel algorytmiczny stanowił 85% obrotu na giełdzie nowojorskiej[14]. Zmiany te są również istotne przy spojrzeniu przez pryzmat założeń hipotezy fraktalnej rynku. Istnieje bowiem pogląd zbliżony z intuicją, że obecność algorytmów na rynku podnosi jego efektywność. Wpływają one zatem na specyfikę obrotu - niepodobna zatem, by badać zachowania cen bez powzięcia powyższego.

Rozdział 2

Metodologia i metodyka przeprowadzonych badań

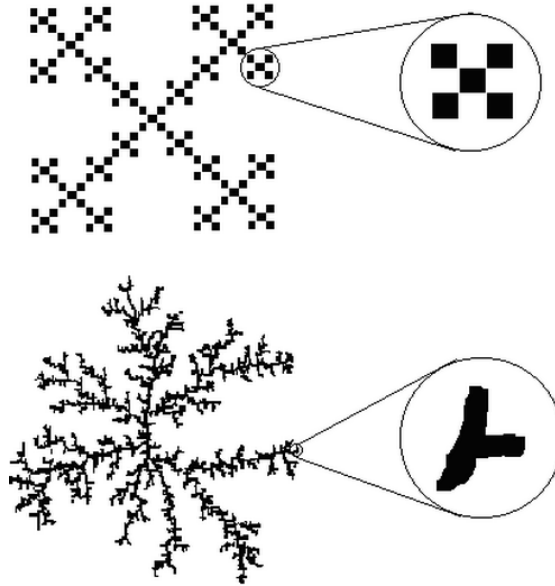
*Prediction is very difficult,
especially about the future.*

Niels Bohr

2.1 Badanie własności multifraktałnych rynków finansowych

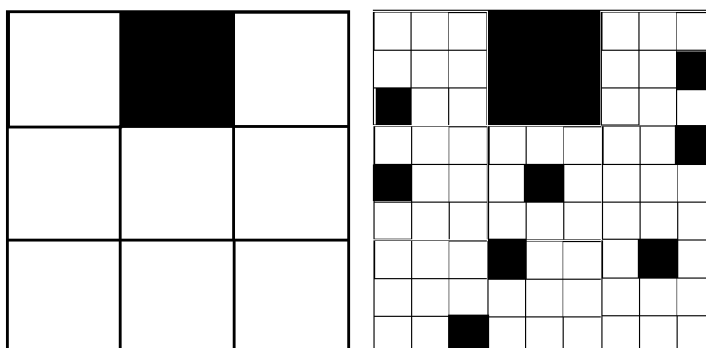
Między matematycznym wyobrażeniem fraktala, a tym obserwowanym w przyrodzie, jak i na rynkach finansowych, istnieje pewna różnica. W jakościowym rozumieniu, modele matematyczne fraktali (**fraktale matematyczne**) stanowią obiekty, dla których w każdej kolejnej skali obserwowane jest idealne powielenie poprzedniego układu (**samopodobieństwo**). **Fraktale fizyczne** mogą zawierać jednak pewne zaburzenia w powieleniu poprzednich skal. Na rynkach finansowych własność ta jest jeszcze silniej zniekształcana szumem. Możliwe jest jednak określenie własności skalowania dla takich obiektów, a fraktale tego rodzaju nazywane są **fraktalami statystycznymi**[51](patrz Rysunek 2.1).

Oczywiście istnieje możliwość wprowadzenia zaburzeń struktury fraktalnej dla modeli matematycznych, czego przykładem jest zdefektowany dywan Sierpińskiego (patrz Rysunek 2.2), czy też przypadkowa krzywa Kocha (patrz Rysunek 2.3). Mimo zaburzenia, wymiar fraktalny pozostał niezmienny. Innymi słowy, obiekt jest samopodobny w sensie statystycznym.



Rysunek 2.1: Zestawienie przykładów fraktali: matematycznego oraz statystycznego (źródło: [52]).

W przypadku fraktali samopodobnych transformacja, z jaką mamy do czynienia jest izotropowa. Istnieją jednak fraktale nie wykazujące równomiernego skalowania we wszystkich wymiarach. W tym wypadku mamy do czynienia z szerszą grupą transformacji, których wynikiem są obiekty samodzielnne, inaczej zwane **samoafinicznymi** (*self-affine*). Przykładowo, skalowanie w dwóch wymiarach generujące obiekty samoafiniczne może dokonywać w jednym z wymiarów powiększania dwukrotnego, a w drugim czterokrotnego[53]. Relacja skalowania samodzielnnego wykorzystuje eksponentę Hursta, o której mowa będzie w dalszych rozdziałach.



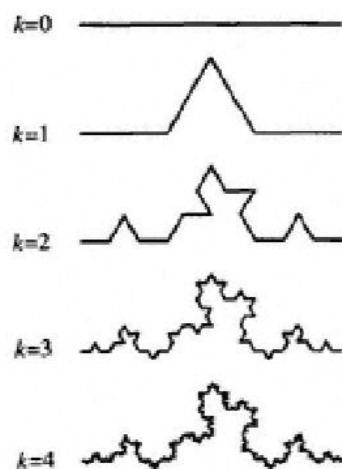
Rysunek 2.2: Przykład dwóch kroków generowania zdefektowanego dywanu Sierpińskiego.

2.1.1 FMH - Hipoteza rynku fraktalnego

Badania w zakresie fraktalnych właściwości rynków finansowych doprowadziły do powstania Hipotezy Rynku Fraktalnego (*Fractal Market Hypothesis - FMH*) [11], która w odróżnieniu od Hipotezy Rynku Efektywnego (*Efficient Market Hypothesis - EMH*) uwzględnia różnorodność horyzontów czasowych poszczególnych inwestorów, a tym samym brak jednorodności w zakresie oceny ryzyka.

Wcześniej zaproponowana hipoteza miała kilka wersji: silnej, pół-silnej oraz słabej EMH. Pierwsza z nich abstrahowała od jakiegokolwiek analizy czynników zewnętrznych w kształtowaniu się cen, zakładając, że ich wpływ jest już uwzględniony w cenie (dotyczy to również informacji, które nie zostały podane do publicznej wiadomości). Silna EMH wymagała, aby obserwacje były niezależne lub charakteryzowały się co najwyżej krótką pamięcią. Innymi słowy, w ramach takich założeń przyszła zmienność cen nie mogła zostać określona na podstawie zmienności historycznej (proces o własnościach martyngału niskiego rzędu). Rozumowanie takiego podejścia można uzasadnić następująco: nie jest możliwe przewidzenie nowej informacji dotyczącej rynku/waloru, lecz w momencie, gdy ona pojawi się, ceny bezzwłocznie zostaną skorygowane do poziomu odzwierciedlającego nowy stan rynku/waloru.

Hipotezy pół-silna oraz słaba EMH ograniczały ilość informacji zawartej w aktualnej cenie instrumentu. Pierwsza z nich zakładała, że cena odzwierciedla



Rysunek 2.3: Sposób tworzenia przypadkowej krzywej Kocha (źródło: [54]).

wszystkie przeszłe oraz aktualne informacje dostępne do publicznej wiadomości. Natomiast ostatnia, że w cenie są uwzględnione wszelkie informacje przeszłe. Tym samym, nawet najslabsza z hipotez EMH, celuje w bezzasadność analizy technicznej.

Według EMH zmienność cen powinna być dobrze reprezentowana przez rozkład normalny. W rzeczywistości, krzywa gaussowska nie jest zbyt trafnym przybliżeniem tego rozkładu. Pojawiają się zdarzenia ekstremalne, charakterystyczne, co prawda, dla rozkładów z rodziny Levy'ego, jednak nie rozkładu normalnego, a rozkładów o grubych ogonach. Rejestrowane na rynkach ekstremalne wartości zmienności noszą nazwę przelotów Levy'ego (*Levy flights*) i, poza finansami, są obserwowane w wielu obszarach badań m.in. w analizach dynamiki sieci społecznych[45].

Zaproponowana przez Petersa Fraktalna Hipoteza rynku wprowadza następujące podejście[11]:

1. Rynek jest **stabilny**, jeśli składa się z inwestorów charakteryzujących się zróżnicowanymi horyzontami czasowymi (zapewnienie płynności).
2. Wpływ poszczególnych determinant cen jest zależny od horyzontu czasowego. Czynniki techniczne oraz sentyment rynku przejawiają silniejszy

wpływ na cenę w krótkim okresie. Dla dłuższych okresów dominujące mogą okazać się czynniki fundamentalne. Zatem zmienność cen odzwierciedla w dominującym stopniu te czynniki, które są istotne dla dystansu czasowego będącego podstawą określenia zmienności.

3. Jeżeli pojawia się zdarzenie wprowadzające wątpliwość co do zasadności informacji fundamentalnej, inwestorzy długookresowi przestają uczestniczyć w rynku lub przechodzą na krótkookresowy obrót instrumentem. Dodatkowo, jeśli ogólny horyzont czasowy inwestorów zaczyna ujednocilać się, rynek staje się **niestabilny** - struktura fraktalna załamuje się.
4. Ceny odzwierciedlają kombinację krótkoterminowego, technicznego obrotu oraz wyceny długoterminowej. Tym samym, zmienność krótkoterminowa jest wyższa (bardziej zaszumiona) od długoterminowej.
5. Jeśli instrument nie charakteryzuje się powiązaniem z cyklami ekonomicznymi, długoterminowe trendy nie powinny być obserwowane. W takim wypadku dominować powinna płynność i informacja krótkookresowa.

W przeciwieństwie do EMH, hipoteza fraktalna nie musi uwzględniać całej dostępnej informacji o walorze, a jedynie informację istotną¹. Co więcej, według Petersa dyfuzja informacji jest nierównomierna, co wynika z założenia drugiego. Może ona jednak zmieniać się - zgodnie z założeniem trzecim. Wpływ pojawiających się informacji może powodować zatem zjawiska perturbacji cen znacznie bardziej złożone niż korygowanie jej według EMH.

Klasyczne metody cechują się tendencją do niedowważania ryzyka[28]. Istnieją przykłady spektakularnych bankructw wynikających z niewłaściwej oceny ryzyka (LTCM (1998), upadłości będące pokłosiem kryzysu z roku 2008).

Fraktalne podejście do rynku finansowego, zarządzania aktywami oraz ryzykiem mogą w przyszłości stać się silnie wykorzystywanymi koncepcjami². Wyciągnąć można bowiem z założeń FMH, że ocena zmienności wymaga złożonych narzędzi analitycznych ujmujących różnorodność jej natury. Wspomniane wcześniej metody GARCH, RiskMetrics, itp. nie mają takich zdolności.

¹Miejmy na uwadze, że nie rozróżnia się charakteru czy dostępności tej informacji, a jedynie to czy jest ona istotna

²Ważnym elementem fraktalnego podejścia do rynków finansowych jest zaproponowany przez Mandelbrota multifraktalny model stóp zwrotu z aktywów (*Multifractal Model of Asset Returns* - **MMAR**)

Zmienność rozumiana jako wrażliwość nabiera, w ramach koncepcji FMH, nowego wymiaru. Sposób kształtowania się ceny instrumentu należy wiązać nie tylko z poziomem zmienności, ale całym zbiorem parametrów jej fraktalnej charakterystyki.

Przyjmując, że taka charakterystyka transformuje informację istotną w sposób poprawny, możliwe winno być prognozowanie na jej bazie dalszego zachowania waloru, jak również ocena prawdopodobieństwa upublicznienia istotnej informacji dotyczącej emitenta. W szczególności, informacje o znacznym pogorszeniu się sytuacji finansowej podmiotu dostępne insiderom, jak również zawarte w cenie reakcje na pogłoski o pogorszeniu się sytuacji określonej spółki, powinny być możliwe do wyekstrahowania z szeregu cen.

2.1.2 Multifraktalna analiza szeregów czasowych

Fraktalna lub multifraktalna charakterystyka może być wykorzystana do modelowania szeregów czasowych, sporządzania prognoz dotyczących zdarzeń ekstremalnych, modelowania reakcji układu na czynniki zewnętrzne w ramach zadanych warunków brzegowych. W celu obserwacji fraktalnych i multifraktalnych własności skalowania w szeregach czasowych stworzono kilka narzędzi, pierwotnie zakładających stacjonarność danych, a w późniejszym czasie również dla szeregów niestacjonarnych. Analiza układów złożonych jest trudnym zadaniem, głównie ze względu na brak możliwości podziału na grupy składowych systemu przy jednoczesnym zachowaniu własności jego dynamiki.

O ile proste układy mogą charakteryzować się łatwo rozpoznawalną periodycznością ograniczoną do pojedynczej częstości, układy złożone z reguły charakteryzują się fluktuacjami w wielu skalach czasowych. W takich przypadkach dynamika może być określona przy pomocy praw skali (*scaling laws*) i oznaczana jako fraktalna lub multifraktalna, zależnie od tego, czy taką charakterystykę oprócz można o pojedynczą eksponentę skalującą, czy też o wiele wartości takich eksponent.

Zmiany stanu skupienia (przemiany fazowe) układów złożonych często wiążą się ze zmianami w ich dynamice fraktalnej, pozwalając na wykrycie takiej zmiany poprzez analizę fraktalną. Rynkową analogią przemiany fazowej jest reakcja na pojawienie się nowej, istotnej informacji.

Samoafiniczność nie determinuje fraktalności układu w ścisłym rozumieniu, niemniej w szerokim ujęciu można uważać układy za fraktalne wszędzie tam, gdzie **eksponenta Hursta** może być w sposób uzasadniony zdefiniowana. Oś czasu oraz oś wartości szeregu czasowego nie są równe, zatem przeskalowanie osi czasu czynnikiem a może wymagać przeskalowania osi wartości innym czynnikiem a^H w celu otrzymania statystycznie podobnego obrazu:

$$x(t) \rightarrow a^H x(at) \quad (2.1)$$

gdzie H - eksponenta Hursta - charakteryzuje typ samoafiniczności. Dla błędzenia przypadkowego $H = 0,5$ - reskalowanie czasu przez 4 oznacza reskalowanie wartości przez 2. W przypadku wartości $H > 0,5$ szereg wykazuje **persystencję**, a $H < 0,5$ oznacza **antypersystencję**³.

Dla danych stacjonarnych, funkcja autokowariancji przyrostów w szeregu może być badana w celu określenia stopnia persystencji. Dla korelacji długozasięgowych, zatem persystentnych, które można zdefiniować jako

$$C(s) \propto s^{-\gamma}, \quad (2.2)$$

gdzie własność skalowania określona jest przez $H = 1 - \gamma/2$. Dla korelacji krótkozasięgowych autokowariancja $C(s)$ słabnie eksponencjalnie.

Istotnym zagadnieniem są również pojawiające się przekrywania (*crossovers*) w skalowaniu. Mogą wynikać one z różnych mechanizmów regulacji dla szybkich oraz wolnych skal czasu. Przekrywania mogą także wynikać z niestacjonarności. Dobrym przykładem danych niestacjonarnych jest zbiór składający się z segmentów silnych fluktuacji występujących na przemian z obszarami o niskiej zmienności poziomu sygnału.

Prawdziwa multifraktalność charakteryzuje się różnorodnymi własnościami skalowania na przestrzeni całej rozpiętości jednostek czasu.

Wyróżnia się dwa główne typy multifraktalności w szeregach czasowych:

1. wynikający z szerokiego rozkładu prawdopodobieństwa (gdzie nie można wyeliminować multifraktalności poprzez przetasowywanie danych z szeregu),

³Należy mieć na uwadze fakt, że eksponenta Hursta praktycznie zawsze jest estymowana. Nie ma całkowitej pewności, czy wyznaczona wartość pokrywa się z wartością rzeczywistą.

2. wynikająca z różnych korelacji długo-zasięgowych zależnie od wielkości fluktuacji.

W drugim przypadku rozkład prawdopodobieństwa wartości szeregu może charakteryzować się skończonymi momentami i być np. rozkładem gaussowskim. W takim przypadku przedstawienie danych w szeregu niszczy długozasięgowe korelacje, a tym samym własności multifraktalne szeregu. Jeśli oba typy multifraktalności są obecne w szeregu, przetasowanie danych osłabia własności multifraktalne.

Z uwagi na fakt, że finansowe szeregi czasowe często wykazują trendy, które mogą być traktowane jako przejaw niestacjonarności, silne trendy mogą prowadzić do niepoprawnej oceny pamięci szeregu⁴. Metodą pozwalającą na systematyczną eliminację trendów jest MF-DFA[17].

2.1.3 Metoda MF-DFA

Procedura MF-DFA składa się z pięciu kroków. Dla danego szeregu czasowego $X_{k=1}^N$ dokonujemy kolejno:

1. scałkowania sygnału zmienności skorygowanego o średnią wartość X_{avg} , otrzymując szereg

$$Y_i = \sum_{k=1}^i [X_k - X_{avg}] \text{ dla } i = 1, 2, \dots, N \quad (2.3)$$

2. Scałkowany szereg Y dzielimy na N_s nieprzekrywających się przedziałów o równej długości równej s . (Długość szeregu nie zawsze jest wielokrotnością s , z tego względu dzielenia szeregu Y można dokonać poczynając od najstarszej oraz najmłodszej chwili czasu, otrzymując tym samym $2N_s$ przedziałów. W ten sposób uwzględnione w analizie pozostają wszystkie punkty szeregu.)
3. Dla każdego z N_s (albo $2N_s$) przedziałów wyznaczamy trend. Zwykle przyjmuje się metodę najmniejszych kwadratów (*Ordinary Least Squares* - OLS) do wyznaczania funkcji trendu. Często wyznaczanym trendem jest

⁴Skoro silne trendy mogą prowadzić do niewłaściwej oceny długo-zasięgowej persystencji, trendy i fluktuacje winny być poddawane osobnym analizom.

liniowy, jednak możliwe są zastosowania funkcji wyższych rzędów. W zależności od rzędu zastosowanej funkcji, nazwa metody przyjmuje znacznik stając się: MF-DFA1 - dla funkcji liniowych, MF-DFA2 - dla dopasowań wielomianem drugiego stopnia, itd. W ramach poszczególnych przedziałów ν , od wartości szeregu Y odejmujemy wartości wynikające z wyznaczonej funkcji trendu, tworząc funkcję fluktuacji $F^2(s, \nu)$:

$$F^2(s, \nu) = \frac{1}{s} \sum_{i=1}^s \{Y[(\nu - 1)s + i] - y_\nu(i)\}^2 \quad (2.4)$$

a w przypadku, gdy mamy $2N_s$ przedziałów, fluktuacje przedziałów N_s do $2N_s$ wyliczamy ze wzoru:

$$F^2(s, \nu) = \frac{1}{s} \sum_{i=1}^s \{Y[(\nu - N_s - 1)s + i + l_s] - y_\nu(i)\}^2 \quad (2.5)$$

4. Uśredniamy po wszystkich przedziałach, otrzymując funkcję fluktuacji q -tego rzędu⁵:

$$F_q^2(s) = \left\{ \frac{1}{2N_s} \sum_{i=1}^{2N_s} [F^2(s, \nu)]^{q/2} \right\}^{1/q} \quad (2.6)$$

5. Określamy własności skalowania funkcji fluktuacji poprzez analizę wykresu logarytmu funkcji fluktuacji $F_q^2(s)$ od logarytmu wielkości przedziałów s . Wynika to z prawa skalowania:

$$F_q^2(s) \propto s^{h(q)}. \quad (2.7)$$

Należy przy tym mieć na uwadze, że dla skal $s > N/4$, funkcja fluktuacji przestaje być wiarygodna statystycznie ze względu na niewielką liczbę segmentów, dlatego należy takie skale wykluczyć z analizy.

Następnie, na podstawie otrzymanych wartości $h(q)$ obliczamy wartości charakteryzujące własności multifraktalne:

1. eksponenty Renyi $\tau(q)$ [56]:

$$\tau(q) = qh(q) - 1 \quad (2.8)$$

⁵W podejściu monofrakalnym rozważane są fluktuacje rzędu drugiego - przyjmuje się $q = 2$.

2. eksponenty Lifszycy-Höldera⁶ α zwanej również siłą osobliwości⁷:

$$\alpha = h(q) + qh'(q) \quad (2.9)$$

3. spektrum osobliwości $f(\alpha)$ wyrażającego wymiar podzbioru szeregu charakteryzowanego przez α :

$$f(\alpha) = q[\alpha - h(q)] + 1 \quad (2.10)$$

Uogólniony wymiar fraktalnych $D(q)$ powiązany jest z powyższymi zależnościami:

$$D(q) = \frac{\tau(q)}{q-1} \quad (2.11)$$

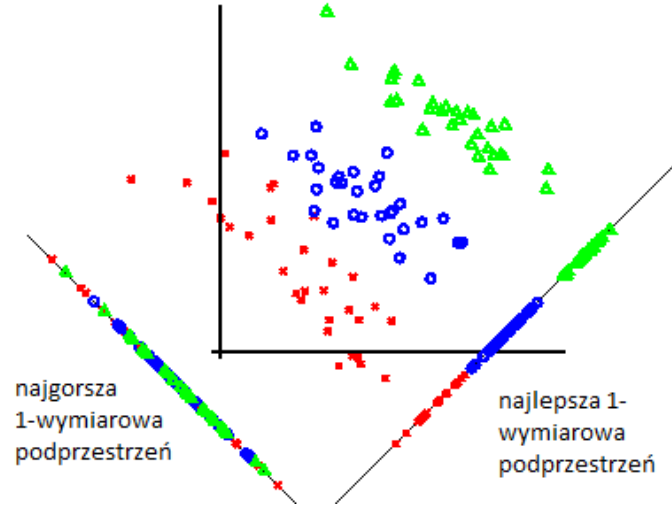
2.2 Analiza dyskryminacyjna

Analiza dyskryminacyjna w ogólnym ujęciu polega na określeniu relacji między zmienną kategoryzującą a grupą zmiennych powiązanych określonych dla poszczególnych obiektów podlegających rozróżnieniu (inaczej dyskryminacji), a następnie wykorzystaniu tej relacji w określaniu obiektów, dla których zmienna kategoryzująca nie jest znana. Z natury, wartości zmiennej kategoryzującej określają jednoznacznie przypisanie obiektów do poszczególnych kategorii, klas czy grup obiektów[18]. W najprostszym ujęciu rozważać można zmienną dychotomiczną, określającą przynależność do wyróżnionej klasy lub brak takiej przynależności, albo też przynależność do jednej z dwóch grup. Problematyka funkcji dyskryminacyjnych została zapoczątkowana w 1936 roku przez R. Fishera, który przy pomocy funkcji 4 zmiennych maksymalizującej stosunek zmienności międzygrupowej od zmienności wewnątrz grupowej, dokonał rozróżnienia dwóch gatunków roślin[19]. Modele klasyfikujące obiekty pochodzące z dwóch grup są chyba najczęściej wykorzystywane, przy czym często używaną jest liniowa analiza dyskryminacyjna (*Linear Discriminant Analysis* - **LDA**), będąca uogólnieniem funkcji Fishera.

W ramach metody poszukuje się takiego przekształcenia liniowego przestrzeni zmiennych, które zoptymalizuje stosunek pierwotnie wyznaczany przez Fishera. Idea przekształcenia została przedstawiona na Rysunku 2.4.

⁶Stanowi ona wynik transformaty Legendre dla eksponenty Renyi: $\alpha = \frac{d}{dq} \tau(q)$

⁷w odróżnieniu od siły multifraktalności wyrażanej wzorem $\Delta\alpha = \alpha_{max} - \alpha_{min}$



Rysunek 2.4: Idea przekształcenia liniowego wykorzystywanego w metodzie LDA (źródło: [60]).

Analiza próby uczącej w ramach LDA ma w szczególności prowadzić do wyłonienia takiej kombinacji współczynników dyskryminacyjnych, która dokona wyżej wspomnianego przekształcenia. Wynikiem jest macierz współczynników:

$$a^T = [a_1 \dots a_k]^T, \quad (2.12)$$

która zostaje podana do funkcji dyskryminacyjnej o postaci:

$$Z(x) = a^T x + a_0. \quad (2.13)$$

Znanym i powszechnie wykorzystywanym przykładem implementacji LDA jest funkcja dyskryminacyjna **Z-score Altmana**, której zadaniem jest klasyfikacja kondycji finansowej przedsiębiorstw, w szczególności wyodrębnienie firm o dużym prawdopodobieństwie bankructwa w przeciągu roku od zaklasyfikowania. Klasyfikacja dokonywana jest na podstawie zmiennych wywodzących się z analizy fundamentalnej, w szczególności analizy wskaźnikowej.

Pierwszy model Altmana powstał w 1968 roku na próbie 66 przedsiębiorstw amerykańskich. Zbiór ten stanowiły przedsiębiorstwa produkcyjne, przy

czym liczba upadłych równała się połowie badanych. Postać funkcji przedstawia się następująco:

$$Z = 1,2 \cdot W_1 + 1,4 \cdot W_2 + 3,3 \cdot W_3 + 0,6 \cdot W_4^{-1} + 1,0 \cdot W_5 \quad (2.14)$$

Autor do budowy funkcji wykorzystał 5 wskaźników: wskaźnik udziału kapitału obrotowego w aktywach (W_1), wskaźnik stosunku zysku skumulowanego do aktywów ogółem (W_2), wskaźnik stosunku zysku przed opodatkowaniem i zapłatą odsetek ($EBIT$) do kapitału całkowitego (W_3), wskaźnik stosunku wartości rynkowej kapitału akcyjnego do wartości księgowej kapitału obcego (W_4^{-1}) oraz wskaźnik rotacji aktywów (W_5). Dla tak zbudowanej funkcji zostały wyznaczone trzy przedziały wartości Z : $(-\infty; 1,8)$, $(1,8; 3)$, $(3; +\infty)$. Dla wartości poniżej 1,8 ryzyko upadłości określono jako bardzo wysokie. Dla przedziału wartości od 1,8 do 3 to ryzyko oznaczono jako nieokreślone. W przypadku wartości Z powyżej 3 ryzyko bankructwa było niewielkie.

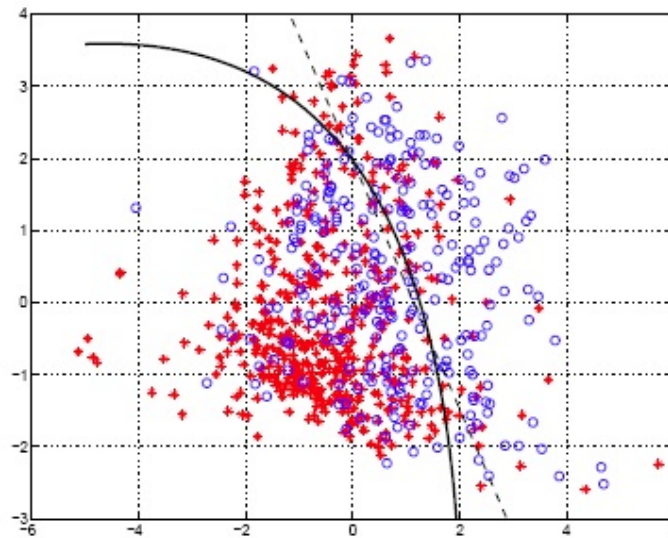
W późniejszym okresie powstały jeszcze dwie propozycje Altmana dla funkcji dyskryminacyjnych, jednak spotkały się z krytyką w obliczu braku uzasadnienia wyboru wskaźników jednowymiarowych oraz braków w przedstawieniu metod przypisania poszczególnych parametrów funkcji. W szczególności wymienia się tutaj wpływ czasu istnienia przedsiębiorstwa na wartości niektórych wskaźników tworzących model, a zatem i wartości uzyskiwane przez funkcję. Najsilniej zależnym od długości czasu istnienia podmiotu jest wskaźnik W_2 informujący o stosunku zysku skumulowanego do aktywów ogółem, gdzie jako zysk skumulowany rozumie się zysk zatrzymany na potrzeby przedsiębiorstwa w ciągu całego okresu jego egzystencji. Zatem oczekiwane jest otrzymywanie wyższych wskazań w tym wymiarze dla przedsiębiorstw dłużej utrzymujących się na rynku. Nawet sam Altman obchodził się z rezerwą do stworzonych przez siebie funkcji, podkreślając ich niską reprezentatywność dla konglomeratów, przedsiębiorstw handlowych i usługowych oraz firm spoza Stanów Zjednoczonych⁸.

W analogii do propozycji Altmana powstały przykłady modeli wyznaczonych dla polskich przedsiębiorstw autorstwa m.in. Hadasik, Mączyńskiej,

⁸W zastosowaniu funkcji Z dla polskich przedsiębiorstw mogą pojawić się trudności w wyznaczeniu wielkości zysku skumulowanego, ponieważ w naszej sprawozdawczości nie jest on wyodrębniany. W uproszczeniu przyjęć można za wynik skumulowany sumę kapitału zapasowego i rezerwowego oraz nie podzielonego wyniku finansowego[43].

Gajdki i Stosa, Hołdy. Co warto podkreślić, wszystkie bazują na klasyfikatorze LDA.

Wracając do szerszego spojrzenia na zagadnienie dyskryminacji, istnieją również metody zupełnie analogiczne do LDA, ale rzędów wyższych. Przykładem jest choćby kwadratowa analiza dyskryminacyjna (*Quadratic Discrimination Analysis - QDA*), która po dokonaniu nieliniowego przekształcenia przestrzeni klasyfikuje obiekty w osi/płaszczyźnie prostopadłej do krzywej lub płaszczyzny przekształcenia (patrz Rysunek 2.5).



Rysunek 2.5: Przykład dyskryminacji w dwuwymiarowej przestrzeni zmiennych uczących z wykorzystaniem LDA (linia przerywana) oraz QDA (linia ciągła) (źródło: [59]).

W ostatnich latach pojawiły się także inne metody analizy dyskryminacyjnej, w szczególności metoda wektorów nośnych, które wykazują wyższość w zakresie skuteczności klasyfikacyjnej nad LDA, QDA, czy też metodami opartymi o sieci neuronowe.

2.2.1 Metoda maszyn wektorów nośnych - SVM

Metoda maszyn wektorów nośnych (*Support Vector Machines - SVM*) należy do grupy metod opartych na funkcjach jądra. W ostatnich latach prze-

żywała intensywny rozwój poparty utworzeniem wielu modeli oraz wykorzystaniem w różnorodnych obszarach badań.

W odróżnieniu od metod opartych o sieci neuronowe, w SVM granice klasyfikacji są determinowane bezpośrednio na zbiorze uczącym w taki sposób, aby maksymalizować marginesy tych granic w wielowymiarowej przestrzeni cech (*feature space*). Dzięki temu możliwa jest efektywna minimalizacja błędów klasyfikacji na zbiorze uczącym oraz docelowym[27].

Logika metody w przypadku liniowej separowalności klas jest następująca[20]:

1. Dany jest zbiór uczący składający się z N obserwacji o postaci

$$\{(\mathbf{x}^1, y^1), \dots, (\mathbf{x}^N, y^N)\}, \quad (2.15)$$

gdzie \mathbf{x}^i jest d -wymiarowym wektorem zmiennych przestrzeni dyskryminacji a y^i wartością zmiennej kategoryzującej.

2. Poszukiwane są rozwiązania równania

$$\mathbf{A} \cdot \mathbf{x} + A_0 = 0 \quad (2.16)$$

dla wektora \mathbf{A} w przestrzeni \mathbf{R}^{d+1} .

3. Pośród rozwiązań \mathbf{A} dokonuje się wyboru optymalnego rozwiązania. Osiągnięcie takiego rozwiązania oznacza znalezienie hiperpłaszczyzny, dla której odległość (margines) od punktów obu klas jest maksymalna. Innymi słowy, zadaniem optymalizacyjnym jest maksymalizacja średnicy marginesu przy nałożonym warunku dyskryminacji:

$$\begin{cases} \max 2r, \\ y^i \frac{1}{\|\mathbf{A}\|} (\mathbf{A} \cdot \mathbf{x}^i + A_0) \geq r. \end{cases} \quad (2.17)$$

Przyjmując $\|\mathbf{A}\| = \frac{1}{r}$ otrzymujemy:

$$\begin{cases} \min \frac{1}{2} \|\mathbf{A}\|^2, \\ y^i (\mathbf{A} \cdot \mathbf{x}^i + A_0) \geq 1, \end{cases} \quad (2.18)$$

przy czym, dla drugiej linii, osiągnięcie równości jest równoznaczne ze znalezieniem się punktu na granicy marginesu. Warto zaznaczyć, że mamy tutaj do czynienia z zagadnieniem optymalizacji wypukłej z kwadratową

funkcją celu (stąd użycie programowania kwadratowego), co ujęte w funkcji Lagrange'a przyjmuje postać:

$$L_p(\mathbf{A}, A_0, \alpha) = \frac{1}{2} \|\mathbf{A}^2\| + \sum_{i=1}^N \alpha_i [y^i (\mathbf{A} \cdot \mathbf{x}^i + A_0) - 1]. \quad (2.19)$$

Podstawienie wartości rozwiązań funkcji dla warunku maksimum (rozwiązania równania $\frac{\partial L_p}{\partial A_i} = 0$) do L_p i rozwiązanie tak otrzymanego równania (przy warunku nieujemności współczynników Lagrange'a: $\alpha_i \geq 0$) prowadzi do jednoznacznego rozwiązania:

$$\mathbf{A} = \sum_{i=1}^N \alpha_i y^i \mathbf{x}^i, \quad (2.20)$$

gdzie wektor \mathbf{x}^i nazywamy wektorem nośnym, czyli obserwacją istotną dla dyskryminacji.

4. Wartość parametru A_0 jest wyznaczana dla dowolnie ustalonego wektora nośnego.
5. Funkcję dyskryminacyjną stanowi:

$$\hat{G}(\mathbf{x}) = [\hat{\mathbf{A}} \cdot \mathbf{x} + A_0] \quad (2.21)$$

W przypadku, gdy nie jest możliwa liniowa separacja klas, SVM wykorzystuje przekształcenie zbioru uczącego:

$$\{(\mathbf{x}^1, y^1), \dots, (\mathbf{x}^N, y^N)\} \rightarrow \{(\varphi(\mathbf{x}^1), y^1), \dots, (\varphi(\mathbf{x}^N), y^N)\}, \quad (2.22)$$

a margines jest redukowany lub wręcz pozwala się na błędy klasyfikacji. W takim przypadku, wnikięcie w głąb marginesu klasyfikacji określone przez ξ_i w zadaniu optymalizacyjnym przyjmuje:

$$\begin{cases} \min \frac{1}{2} \|\mathbf{A}\|^2 + C \sum_{i=1}^N \xi_i, \\ y^i (\mathbf{A} \cdot \mathbf{x}^i + A_0) \geq 1 - \xi_i, \xi_i \geq 0, \end{cases} \quad (2.23)$$

gdzie C stanowi parametr mechanizmu regularyzacji (współczynnik dokładności dopasowania). Modele tego typu nazywane są modelami miękkiego marginesu L1 (*L1 Soft-Margin Support Vector Machines - L1 SVM*)[27].

Postać funkcji $\varphi(\cdot)$ w ogólności nie musi być znana. Istotny jest natomiast iloczyn skalarny w przestrzeni, do której przekształcenie mapuje \mathbf{x}^i ⁹. Najczęściej

⁹wynika z twierdzenia Mercera[20]

wykorzystywanymi funkcjami definiującymi iloczyn (funkcjami jądrowym) są: liniowa, wielomianowa, Gaussa (funkcja o podstawie radialnej), oparta o sieci neuronowe, Mahalanobisa oraz funkcja sigmoidalna.

Początkowo, SVM była wykorzystywana do rozwiązywania dychotomicznych problemów klasyfikacji i taką klasyfikację rozważamy w niniejszym opracowaniu. Warto jednak wspomnieć, w szczególności w perspektywie dalszych badań, że z czasem powstały modele pozwalające na dyskryminację wielu grup (*multiclass SVM* - **mSVM**). W ogólności modele te można podzielić na:

1. "jeden przeciw wszystkim" (*One-against-All* - OAA),
2. "jeden na jednego" (*One-against-One* - OAO),
3. "wszyscy razem" (*All-at-Once* - AAO).

Podstawową różnicą pomiędzy metodami jest liczba wykorzystywanych przez nie funkcji dyskryminacyjnych - w przypadku pierwszych dwóch typów wyznaczane jest ich wiele, natomiast w ostatnim przypadku poszukiwana jest jedna funkcja. OAA traktuje zbiór uczący w taki sposób, że pojedyncza klasa dyskryminowana jest względem wszystkich pozostałych obserwacji. Procedura wykonywana jest dla każdej z klas, a otrzymane funkcje łączy się w jeden model. Z kolei, metoda "jeden na jednego" dokonuje dyskryminacji pomiędzy poszczególnymi kombinacjami par klas. Obie metody *de facto* wykorzystują binarną dyskryminację, której wielokrotne wykorzystanie zostaje zsumowane tworząc ostateczny model[20].

W obszarze ryzyka finansowego metoda SVM została wykorzystana m.in. do:

- predykcji zachowań arkusza zleceń w warunkach handlu wysokich częstotliwości [23];
- tworzenia systemu handlu akcjami [24] opartego o *okienkową* teorię akcji (*box theory*);
- tworzenia modeli scoringowych oceny ryzyka kredytowego[21].

Istnieją jednak pewne niedogodności w stosowaniu metody - m.in. względnie długi czas uczenia, co wynika z wykorzystania programowania kwadratowego, gdzie rozwiązywane zagadnienie składa się ze zmiennych, których ilość

jest równa mocy zbioru uczącego¹⁰. Niemniej, w prognozowaniu finansowych szeregów czasowych *czysta* wersja metody SVM została wskazana jako prowadząca do lepszych wyników niż te oparte o sieci neuronowe[25].

2.3 Koncepcja algorytmu MF-DFA - SVM oceny ryzyka inwestycyjnego pod kątem niewyplacalności emitenta

Powstanie IEX daje impuls dla nowych trendów w sposobach organizacji giełd. W warunkach stałej, wysokiej latencji wykorzystanie strategii typu *front running* wydaje się bezprzedmiotowe. Algorytmy, w szczególności HFT, winny obrać inne sposoby osiągania przewagi. Jedną z możliwych ścieżek, zarówno dla HFT, jak i obrotu niskich częstotliwości (*Low Frequency Trading* - LFT), jest dalsza penetracja informacji ukrytej w szeregach cen, wolumenów, statystykach arkuszy zleceń. Wielokrotnie wykorzystywana do analizy zachowań cen metoda MF-DFA oraz, szerzej, koncepcje FMH i MMAR wydają się być ciekawym kierunkiem rozwoju.

Proponuję koncepcję prognozowania ryzyka upadłościowego spółek notowanych na giełdzie w oparciu o charakterystykę multifraktałną jej notowań. W celu weryfikacji proponowanego podejścia wykonany zostanie eksperyment składający się z następujących etapów:

1. określenie zmiennych tworzących charakterystykę MF-DFA,
2. określenie sposobu wyznaczania wartości zmiennej celu,
3. określenie zakresu zbioru uczącego,
4. obliczenie wartości zmiennych MF-DFA oraz przypisanie im zmiennej celu,
5. przeprowadzenie proces uczenia maszynowego - dokonanie klasyfikacji poszczególnych obserwacji wyników MF-DFA w oparciu o wartości zmiennej celu,

¹⁰Dla przypadków dużych ilości danych istnieją techniki łączące SVM z metodami drzew decyzyjnych czy też sieci neuronowych[27].

6. weryfikacja skuteczności klasyfikacji na zbiorze testowym.

Metodą klasyfikacji obraną w niniejszej pracy jest SVM, ponieważ w innych badaniach wykazywała ona wyższą poprawność klasyfikacji niż LDA, QDA, jak również metody oparte o drzewa klasyfikacyjne, w szczególności dla małych prób uczących[41].

Wykorzystanie połączenia MF-DFA i SVM do prognozowania zachowania określonego instrumentu finansowego, kondycji emitenta, w szczególności określenia prawdopodobieństwa bankructwa nie zostało przez autora napotkane. Istnieje natomiast opracowanie dotyczące rynków finansowych łączące MF-DFA z metodami analizy dyskryminacyjnej (w omawianym przypadku jest to LDA), gdzie weryfikowano skuteczność zaklasyfikowania walorów do określonych rynków obrotu (NYSE, NASDAQ, AMEX lub TEX) na podstawie własności multifrakalnych tych instrumentów. Wyniki eksperymentu wskazały na wysokie podobieństwo w zakresie własności multifrakalnych walorów obracanych na tym samym rynku[17].

W zakresie innych dziedzin badawczych, klasyfikowanie wyników analizy multifraktalnej z użyciem metody SVM zostało wykorzystane m.in. do detekcji błędów w obwodach analogowych[22], gdzie parametry spektrum zostały ujęte w postaci wektora:

$$F = [\alpha_{min}, \alpha_{max}, \Delta\alpha, f(\alpha_{min}), f(\alpha_{max}), \Delta f]. \quad (2.24)$$

Jednocześnie, zastosowanym typem metody SVM był OAA, wykorzystujący radialną funkcję jądra.

Innym przykładem wykorzystania połączenia metod MF-DFA oraz SVM jest próba stworzenia systemu rozpoznawania emocji na bazie wyników elektroencefalogramu (EEG)[26]. Analogicznie do wcześniejszego przykładu, wyniki badania EEG poddano analizie MF-DFA, a następnie zdyskryminowano metodą SVM opartą o gaussowską funkcję jądra. W badaniu, konkurencyjnie do SVM, posłużono się również LDA oraz QDA, jednak wyniki metodą wektorów nośnych wykazały lepszą zdolność prognostyczną.

Finalnym przykładem jest wykorzystanie MF-DFA i SVM do klasyfikacji obiektów powietrznych[29]. Metoda MF-DFA została tutaj zaimplementowana do analizy struktury sygnału zwrotnego rejestrowanego przez radary niskich rozdzielczości. Klasyfikacja polegała na wyodrębnieniu sygnałów pochodzących od

trzech rodzajów statków powietrznych - dwóch typów samolotów (o napędzie śmigłowym oraz odrzutowym) oraz helikopterów. Wektor obserwacji użyty w dyskryminacji - parametry charakterystyki multifraktalnej - składa się z następujących zmiennych:

1. stopień symetrii indeksu masowego wyrażany jako:

$$R_\tau = \left| \frac{\max \tau(q)}{\min \tau(q)} \right| \quad (2.25)$$

2. szerokość spektrum:

$$\Delta\alpha = \alpha_{\max} - \alpha_{\min} \quad (2.26)$$

3. różnica między wartościami wymiaru fraktalnego dla maksimum oraz minimum podzbioru prawdopodobieństwa:

$$\Delta f = |f(\alpha_{\min}) - f(\alpha_{\max})| \quad (2.27)$$

4. indeks asymetrii:

$$R_\alpha = \frac{\Delta\alpha_L - \Delta\alpha_R}{\Delta\alpha_L + \Delta\alpha_R}, \quad (2.28)$$

gdzie

$$\Delta\alpha_L = \alpha_0 - \alpha_{\min}, \quad (2.29)$$

$$\Delta\alpha_R = \alpha_{\max} - \alpha_0. \quad (2.30)$$

Powyższe przypadki stanowią analogię poniższych badań. Obierane jest wiele zmiennych charakteryzujących własności multifraktalne analizowanego szeregu cen, nie pojedynczy uogólniony wykładnik Hursta. Zakłada się powiązanie natury zmienności notowań akcji z oceną rynku co do ewentualnej upadłości spółki emitenta - zmienne stanowią charakterystykę spółki. Ich wartości mierzone są na dwuletnim oknie czasowym, a wartość zmiennej celu określana jest bezpośrednio po ostatnim dniu analizowanego okna. Zmienna celu jest w niniejszym przypadku binarna i wyraża zaistnienie zdarzenia złożenia wniosku o ogłoszenie upadłości spółki.

Zbiór uczący stanowią spółki, które były notowane na giełdzie na przestrzeni zadanego okna czasowego, w przypadku których 1) w dniu następującym po ostatnim rozważanym dniu sesyjnym miało miejsce złożenie wniosku o ogłoszenie upadłości emitenta, lub które 2) cechowały się dobrą sytuacją finansową na przestrzeni okna.

Zbiór charakterystyk multifraktałnych stanowiących zmienne pochodzące z pojedynczej charakterystyki określony zostaje w postaci wektora:

$$\vec{E} = [R_\tau, \Delta\alpha, \Delta f, R_\alpha] \quad (2.31)$$

Wektor \vec{E} zostanie wyznaczony dla szeregów:

1. kursu akcji na zamknięciu notowań (*cz*),
2. dziennej liczby transakcji (*lt*).

W ten sposób dla poszczególnych instrumentów otrzymujemy układ wektorów $\vec{E}(x_i)$, gdzie $x_i = \{cz, lt\}$, który tworzy wektor wyższego rzędu \vec{U} :

$$\vec{U} = [\vec{E}(cz), \vec{E}(lt)] \quad (2.32)$$

Wektor \vec{U} wspólnie z wartością zmiennej celu tworzy pojedynczą obserwację. Analiza klasyfikacyjna wykonana zostanie przy użyciu pakietu SVM zbudowanego w R[57], którego biblioteki dostępne są w SageMath.

Rozdział 3

Budowa i zastosowanie modelu MF DFA-SVM w predykcji upadłości spółek notowanych na GPW S.A.

*An approximate answer to the
right problem is worth a good deal
more than an exact answer to an
approximate problem.*

John Tukey

3.1 Dobór walorów oraz zakresu danych giełdowych do badań

W badaniach uwzględniono 34 spółki których akcje były notowane na rynku głównym GPW S.A., z czego 25 stanowiły spółki, względem których miało miejsce złożenie wniosku o ogłoszenie upadłości (zestawienie podmiotów wraz z datą wniosku zawarto w Tabeli 3.1). Pozostałe 9 spółek to firmy o dobrej kondycji (patrz Tabela 3.3), dla których historycznie nie zanotowano złożenia wniosku¹.

Rocznie składanych jest zazwyczaj kilka wniosków o ogłoszenie upadłości emitentów rynku głównego GPW. Mając na celu stworzenie możliwie efektywnego modelu dyskryminacyjnego, wybrano spółki, dla których złożenie wniosku miało miejsce w latach 2009-2014, przy czym analizie podlegały notowania z dwóch lat bezpośrednio poprzedzających datę złożenia wniosku.

Co ważne, w analizie nie są uwzględniane bezpośrednio skutki prawne złożenia wniosku o ogłoszenie upadłości. Nie weryfikowano czy jego następstwem była upadłość, jej rodzaju, czy spółka przeszła restrukturyzację, albo czy wniosek został uznany przez Sąd za bezpodstawny. Przyjęto natomiast, arbitralnie, że jakikolwiek wniosek tego rodzaju świadczy o słabości przedmiotowej spółki.

Analogiczny do spółek z wnioskiem sześcioletni okres analizy dotyczył grupy spółek o dobrej kondycji finansowej, jednak w tym wypadku brano pod uwagę 3 dwuletnie okresy notowań - w latach 2009-2010, 2011-2012, 2013-2014. Celem 3-krotnego pomiaru na spółkach bez wniosku było ograniczenie aspektu specyfiki okresu zbierania danych, w tym m.in. zmniejszenie uzależnienia od stanu całego rynku akcji w określonym przedziale dwuletnim. W ten sposób otrzymano 27 obserwacji w grupie, co stanowi grupę porównywalną do 25 obserwacji dla spółek z wnioskiem.

Podstawą wyłonienia spółek do grupy bez wniosku było wykazanie dobrej kondycji według dyskryminacyjnej analizy fundamentalnej (Zestawienie zawarto w Tabeli 3.3). W celu wyłonienia tych spółek posłużono się ratingiem Altmana dla rynków wschodzących. Został on wyznaczony na podstawie scoringu opar-

¹W dalszej części rozdziału będziemy je nazywać niekiedy jako spółki *bez wniosku*, a drugą grupę jako spółki *z wnioskiem*.

Walor	Data wniosku
KROSNO	2009-01-30
SFINKS POLSKA	2009-02-17
DUDA	2009-03-24
MONNARI TRADE	2009-05-08
TECHMEX	2009-10-05
JAGO	2011-11-02
PBG	2012-06-04
ABM SOLID	2012-06-29
BOMI	2012-07-11
WILBO	2012-07-19
POLSKIE JADŁO	2012-09-28
IDEON	2013-04-03
FOTA	2013-06-23
MEDIATEL	2013-07-16
GANT	2013-10-16
GREENECO	2013-10-30
RAFAKO	2013-12-02
CASH FLOW	2013-12-17
PBO ANIOŁA	2014-02-25
IDM	2014-05-21
ONE-2-ONE	2014-11-06
INSTAL KRAKÓW	2015-01-29
DSS	2015-02-19
E-KANCELARIA	2015-03-13
MSX	2015-05-27

Tablica 3.1: Zestawienie instrumentów wykorzystanych w analizie, dla których względem spółki emitenta złożono wnioski o ogłoszenie upadłości.

tego o model dyskryminacyjny **Z-score**²[62]. W modelu wykorzystywane są wartości wskaźników:

1. kapitału pracującego³ w ogóle aktywów (X_1),
2. rentowności skumulowanego zysku zatrzymanego (zyski zatrzymane / aktywa) (X_2),
3. rentowności operacyjnej aktywów (EBIT / aktywa) (X_3),
4. stopień pokrycia długu kapitałem własnym (X_4).

Funkcja dyskryminacyjna przyjmuje postać:

$$Z'' = 6,56 \cdot X_1 + 3,26 \cdot X_2 + 6,72 \cdot X_3 + 1,05 \cdot X_4 + 3,25 \quad (3.1)$$

Dla określonych zakresów wartości wyznaczanych z użyciem funkcji Z'' Altman przypisał rating rozpisany w Tabeli 3.2.

Na podstawie określonego w powyższy sposób ratingu dokonano losowego wyboru 9 spółek notowanych na rynku głównym GPW spośród tych, które należały do strefy bezpiecznej na dzień pomiaru (dane o ratingach pobrano z [62]).

²Niekiedy zwany jest on również EM-score (*Emerging Markets score*).

³Kapitał pracujący definiowany jest jako różnica między aktywami obrotowymi a zobowiązaniami bieżącymi.

Strefa	Rating Altmana	Wartość Z-score
bezpieczna	AAA	8,15
	AA+	7,6
	AA	7,3
	AA-	7
	A+	6,85
	A	6,65
	A-	6,4
	BBB+	6,25
	BBB	5,85
niepewna	BBB-	5,65
	BB+	5,25
	BB	4,95
	BB-	4,75
	B+	4,5
	B	4,15
	zagrożona	B-
CCC+		3,2
CCC		2,5
CCC-		1,75
D		0

Tablica 3.2: Wartości graniczne dla poszczególnych ratingów modelu Z-score (źródło: [62]).

Kod waloru	rating Altmana	Okresy analizy
ATLANTA	AAA	okres I: 2009-2010, okres II: 2011-2012, okres III: 2013-2014
COMARCH	AAA	
GRAAL	AA	
KGHM	AA	
LPP	A-	
NETIA	AA	
PULAWY	AAA	
WAWEL	AAA	
ZETKAMA	AAA	

Tablica 3.3: Zestawienie spółek wykorzystanych w analizie, dla których nie miało miejsca złożenie wniosku o ogłoszenie upadłości, a które cechowały wysokie ratingi Altmana (źródło: [62]).

Dla poszczególnych spółek pobrano dane o możliwie najmniejszej dostępnej jednostce czasowej - dzienne (źródło: [63]). Spośród dostępnych danych, tj.: ceny otwarcia, ceny maksymalnej, ceny minimalnej, ceny zamknięcia, wartości obrotu, liczby transakcji, w analizie wykorzystano **cenę zamknięcia** oraz **liczbę transakcji**.

Wybór ten wynika z faktu, iż dla walorów o niższej płynności lub cechujących się znaczną zmiennością cen, często nie ma miejsca wyznaczenie ceny otwarcia. W przypadku braku obrotu również nie ustala się dziennej ceny minimalnej oraz maksymalnej. Znacznie rzadziej natomiast zdarza się brak ustalenia ceny zamknięcia, co, podobnie jak wcześniejsze, wynika z kształtu uregulowań obowiązujących na GPW (w szczególności Regulaminu Giełdy oraz Szczegółowych Zasad Obrotu Giełdowego).

W sytuacjach opisanych powyżej, zastosowanie metody MF-DFA, w szczególności zbudowanie profilu, staje się niewykonalne dla tych punktów (dni), gdzie nie ustalono wspomnianych wartości. Możliwe jest zastosowanie w poszczególnych punktach wartości obranych arbitralnie - np. określenie ceny otwarcia równej ostatnio ustalonej cenie zamknięcia⁴. Z punktu widzenia niniejszej

⁴Przykłady zastosowania takiego przypisania znajdziemy również wśród polskich serwisów [61].

pracy, uznano, że tego typu manipulacja danymi spowodowałaby, że posługiwanie się ceną otwarcia, minimalną czy maksymalną, dawałoby nikły wkład do poprawienia efektywności modelu ze względu na powielanie danych.

W przypadku braku obrotu w danym dniu, wyznaczenie zmienności względnej w analizie multifraktalnej dla liczby transakcji wymagało również pewnej ingerencji w dane, jednak w znacznie bardziej ograniczonym zakresie. Problem polegał na tym, że w przypadku braku obrotu w dniu poprzedzającym dany dzień (na którym zarejestrowano transakcje), zmienność względna sięgała nieskończoności. Aby ten efekt wyeliminować dodano arbitralnie 1 do liczby transakcji dla każdego analizowanego dnia obrotu.

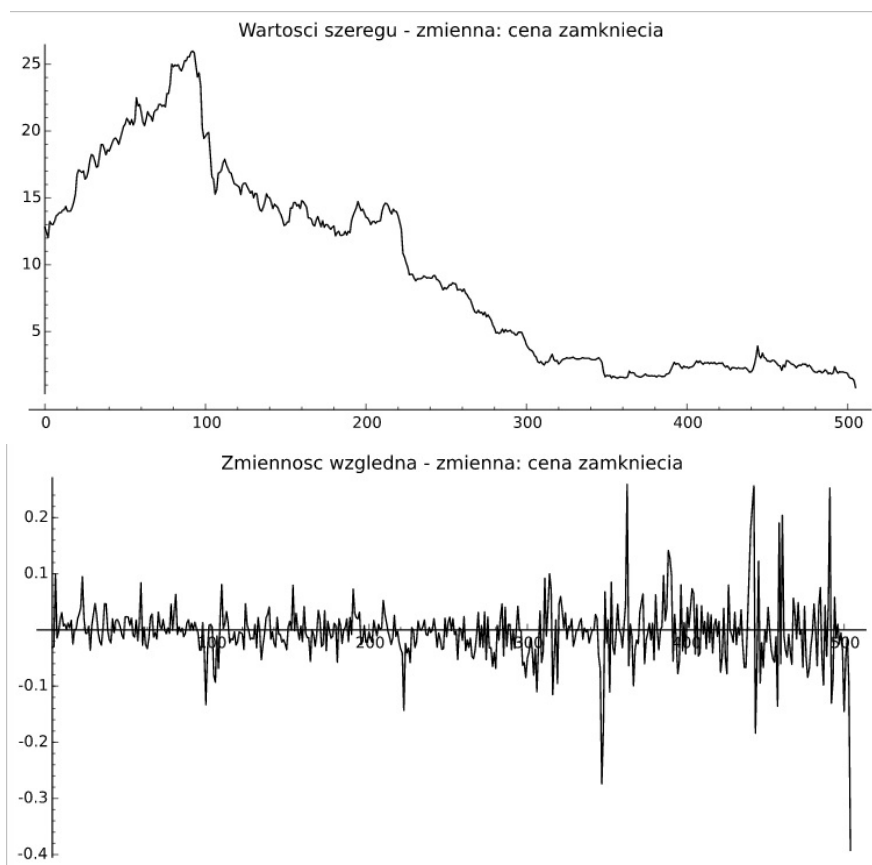
Przykładową analizę MF-DFA dla ceny zamknięcia akcji spółki ABM SOLID S.A. oraz analizę wyników charakterystyki multifraktalnej w całości badanej próby prezentuje kolejny podrozdział.

3.2 Przeprowadzenie procedury MF DFA, utworzenie wektorów obserwacji

Dla 52 obserwacji dokonano analizy MF-DFA dwóch zmiennych: cena zamknięcia oraz liczba transakcji. Na tej podstawie zostały wyznaczone charakterystyki multifraktalne - po dwie dla każdej z obserwacji. Każda z charakterystyk zawiera 4 wartości: stopień symetrii indeksu masowego, szerokość spektrum osobliwości, różnica wartości funkcji f dla otrzymanych skrajnych argumentów oraz indeks asymetrii. Zatem, dla poszczególnych obserwacji wyznaczono po 8 zmiennych wykorzystanych w dalszej analizie jako wyjaśniające, które zostały zestawione ze zmienną celu w postaci binarnej, przyjmującą wartość 1 dla spółek, względem których złożono wnioski o ogłoszenie upadłości.

Proces przeprowadzenia obliczeń metodą MF-DFA został przykładowo zaprezentowany dla cen zamknięcia notowań akcji spółki ABM SOLID S.A. Kolejne kroki przedstawiono na Rysunkach 3.1, 3.2, 3.3. W pierwszej kolejności zaprezentowano wykres kształtowania się ceny zamknięcia oraz jej zmienność w kolejnych dniach 2-letniego okresu bezpośrednio poprzedzającego złożenie wniosku o ogłoszenie upadłości spółki.

Następnie, po scałkowaniu zmienności otrzymano profil, który wielokrotnie



Rysunek 3.1: Szeregi czasowe ceny zamknięcia oraz jej zmienności względnej dla waloru ABM SOLID.

detrendowano przedziałami z zastosowaniem krzywych najlepszego dopasowania trzeciego stopnia. Przykład dopasowania przedstawia czerwona krzywa na górnym diagramie Rysunku 3.2. Zdetrendowana w ramach przedziału zmienność podlega przeliczeniu na poziom fluktuacji w ramach przedziału. Poziomy fluktuacji w poszczególnych przedziałach zostały przedstawione w postaci łamanej na środkowym wykresie Rysunku 3.2. Proces detrendowania, wyznaczania poziomów fluktuacji oraz ich uśredniania wykonywany jest wielokrotnie dla różnych długości przedziałów s (różnych wielkości skal).

Dalej, fluktuacje podlegają uśrednieniu, zgodnie ze wzorem 2.6. Uśrednienie wykonywane jest dla wartości rzędu q z zakresu $(-20;20)$ z wyłączeniem 0. Zależność funkcji fluktuacji F dla poszczególnych wartości q od wartości s przedstawia dolny z wykresów na Rysunku 3.2, na którym nałożono proste najlepszego dopasowania dla poszczególnych $F(s)$ przy określonym rzędzie q .

Współczynnik kierunkowy prostej $F(s)$ wyznacza wartość wykładnika Hursta. Wyznaczany jest on zatem dla poszczególnych rzędów q . Otrzymaną zależność $H(q)$ prezentuje górny wykres Rysunku 3.3. W przypadku notowań cen zamknięcia akcji spółki ABM SOLID S.A., podobnie jak i w pozostałych, dostrzegalna jest zmiana wartości H przy zmianie rzędu q , co potwierdza multifrakalną naturę badanej zmiennej⁵.

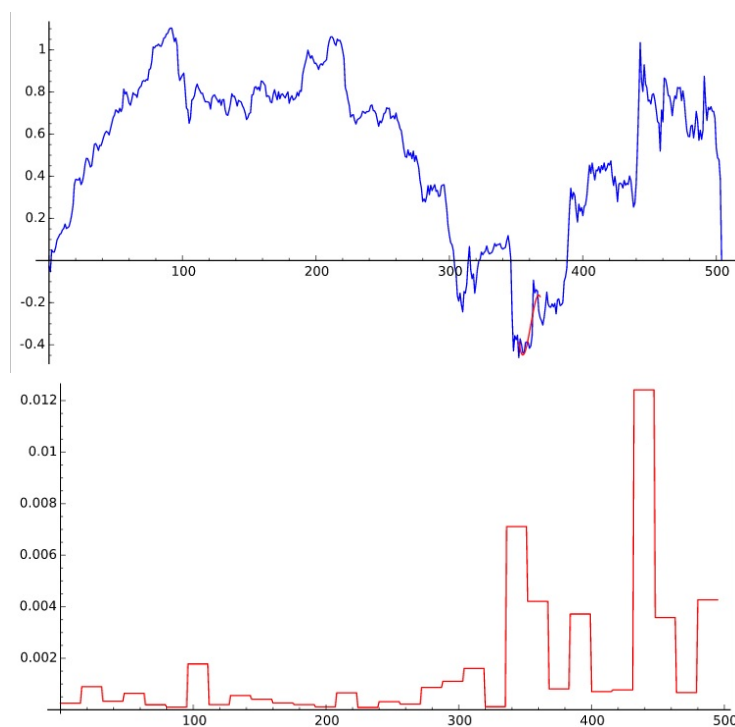
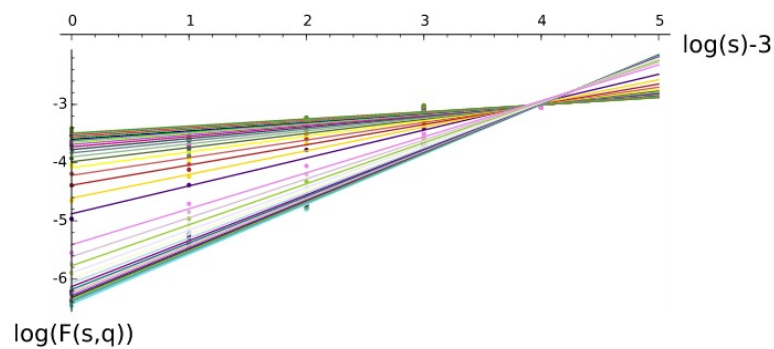
Na podstawie $H(q)$ obliczane są, zgodnie ze wzorem 2.8, wartości funkcji wymiaru Renyi co prezentuje kolejny z wykresów Rysunku 3.3. W wyniku wykonania ostatniego kroku procedury MF-DFA otrzymano wykres spektrum osobliwości $f(a)$ (dolny na Rysunku 3.3).

Na podstawie otrzymanych zależności $f(a)$ obliczono elementy charakterystyk multifrakalnych:

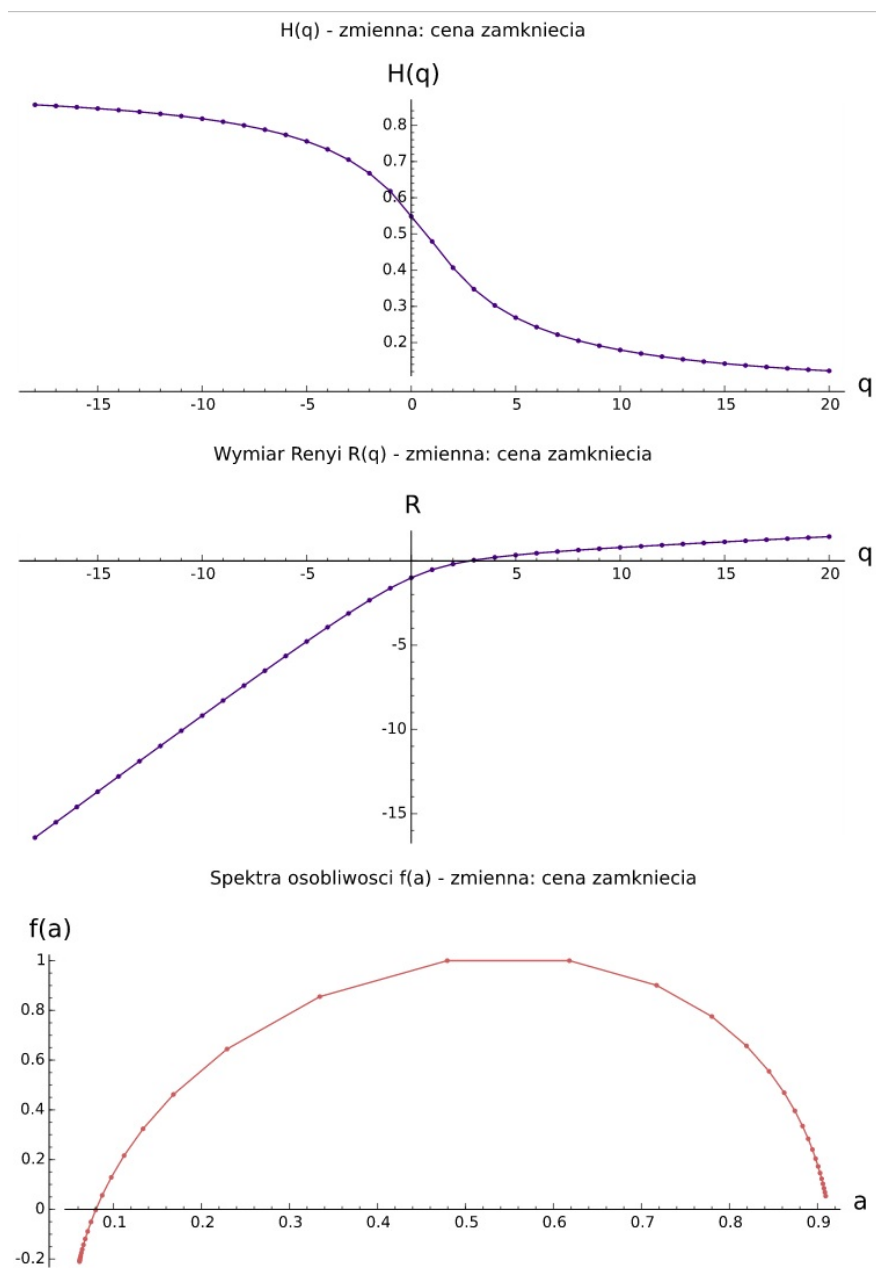
- R_τ - stopień symetrii indeksu masowego,
- $\Delta\alpha$ - szerokość spektrum osobliwości,
- Δf - różnica wartości $f(a)$ przyjmowanych dla minimalnego oraz maksymalnego argumentu,
- R_α - indeks asymetrii⁶.

⁵W przypadku sygnału monofrakalnego $H(q)$ jest funkcją stałą.

⁶Formuły oraz opis wyznaczania poszczególnych wartości został opisany w podrozdziale 2.3.

Skalowanie funkcji fluktuacji $F(s,q)$ w zależności od wielkości okna s - zmienna: cena zamknięcia

Rysunek 3.2: Profil z naniesioną przykładowo funkcją detrendującą, wykres fluktuacji lokalnych, oraz zależności funkcji fluktuacji od wielkości okna dla wszystkich badanych rzędów q .



Rysunek 3.3: Zależność wykładnika Hursta od rzędu q , wymiaru Renyi $R(q)$, oraz spektrum osobliwości $f(a)$.

Wyznaczenia powyższych wartości dokonano dla wszystkich 52 obserwacji, osobno dla spektrów ceny zamknięcia oraz liczby transakcji.

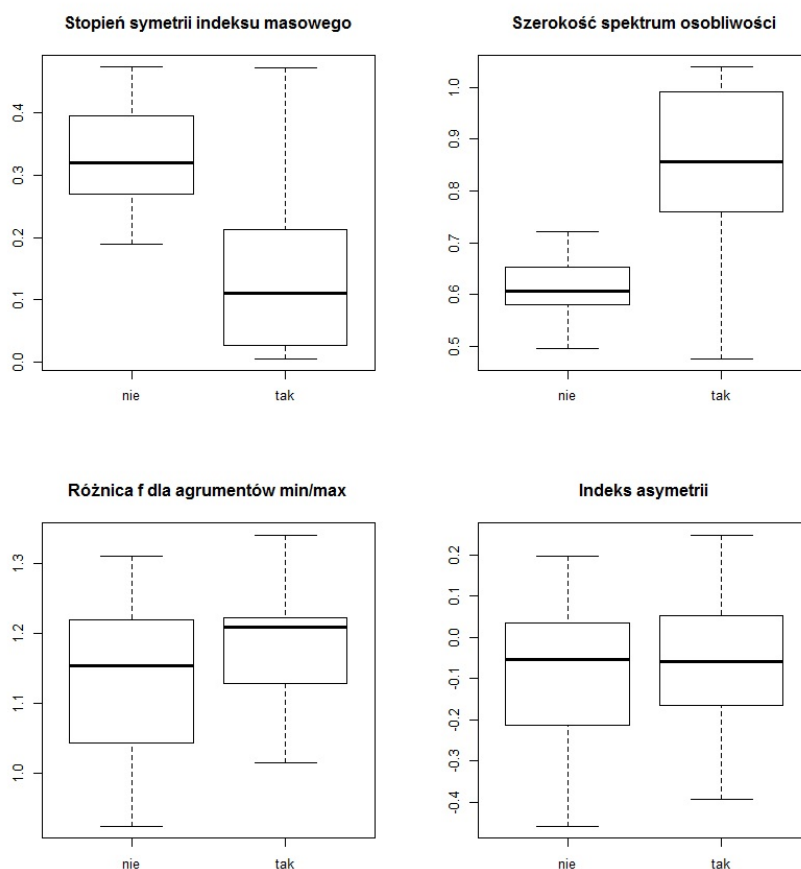
Każdej z obserwacji przypisano 9-elementowy wektor obserwacji. Pierwszy z elementów stanowi zmienna binarna określająca przynależność do grupy spółek *z wnioskiem* albo *bez wniosku*. Kolejne 4 wartości to wynik charakterystyki multifraktalnej dla ceny zamknięcia, a ostatnie 4 - dla liczby transakcji. W przypadku spółki ABM SOLID otrzymano wektor zmiennych objaśniających, który został rozpisany w Tabeli 3.4.

Zmienna	Wartość
Grupa	1
R_τ - cena zamknięcia	0,08773
$\Delta\alpha$ - cena zamknięcia	0,84771
Δf - cena zamknięcia	1,21012
R_α - cena zamknięcia	-0,01421
R_τ - liczba transakcji	0,25164
$\Delta\alpha$ - liczba transakcji	0,61108
Δf - liczba transakcji	1,20423
R_α - liczba transakcji	-0,16615

Tablica 3.4: Zestawienie wyników otrzymanej charakterystyki multifraktalnej dla zmiennych: cena zamknięcia oraz liczba transakcji, dla spółki ABM SOLID S.A.

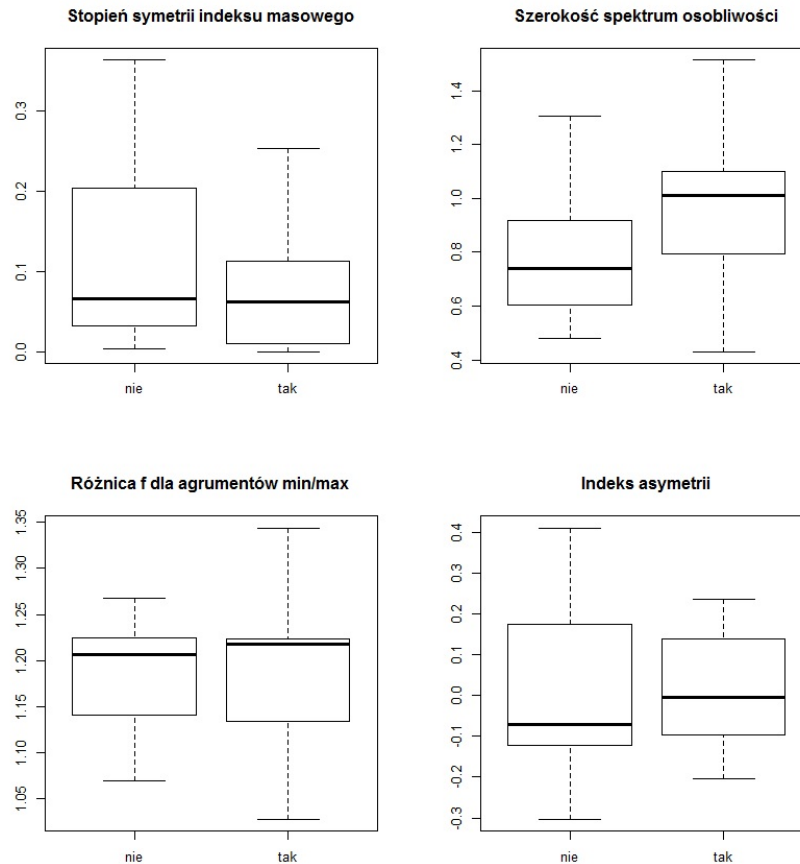
3.3 Wstępna analiza charakterystyk fraktalnych

Komplet wektorów obserwacji poddano obróbce tworząc wykresy pudełkowe umieszczone na Rysunkach 3.4, 3.5. Jakościowy ogląd wskazuje, że największe różnice pomiędzy podpopulacjami *z wnioskiem* i *bez wniosku* obserwuje się w przypadku rozkładów zmiennych R_τ dla ceny zamknięcia oraz $\Delta\alpha$ dla ceny zamknięcia i liczby transakcji.



Rysunek 3.4: Zestawienie wykresów pudełkowych zmiennych objaśniających dotyczących ceny zamknięcia dla spółek: z wnioskiem - tak, bez wniosku - nie.

W przypadku zmiennej R_τ dla ceny zamknięcia blisko 75% obserwacji spośród spółek z wnioskiem leży poniżej wartości 0,2, podczas gdy w drugiej grupie takie wartości notują jedynie 3 z 27 przypadków (11%). Dla szerokości spektrum osobliwości ceny zamknięcia 96% spółek bez wniosku notuje wartości mniejsze od 0,72, natomiast ponad 75% obserwacji dla emitentów z wnioskiem znajduje się powyżej tej wartości.



Rysunek 3.5: Zestawienie wykresów pudełkowych zmiennych objaśniających dotyczących liczby transakcji dla spółek: z wnioskiem - tak, bez wniosku - nie.

Wśród zmiennych dotyczących liczby transakcji największe różnice w rozkładach są obserwowane dla szerokości spektrum osobliwości. W tym wypadku spółki *bez wniosku* wykazują nieco większą koncentrację od drugiej próby, przy czym mediany wynoszą odpowiednio 0,73 oraz 1,02. W pozostałych przypadkach mediany kształtują się na zbliżonym poziomie. Położenia i szerokości drugiego i trzeciego kwartyłu są najbardziej zróżnicowane w przypadku stopnia symetrii indeksu masowego. Tutaj jakościowy ogląd wskazuje, że spółki *bez wniosku* winien charakteryzować rozkład o większej wartości skośności.

W celu dokładniejszego zbadania obserwowanych różnic w rozkładach, a dokładniej wpływu przynależności do jednej z dwóch grup na wartości poszczególnych zmiennych, wykonano **jednoczynnikową analizę wariancji (ANOVA)**. Wyniki testu znajdują się w Tabeli 3.5. Dla stopnia symetrii indeksu masowego dla ceny zamknięcia niezależnie od poziomu istotności przynależność do grupy *z wnioskiem* lub *bez wniosku* ma wpływ na uzyskiwane wartości. Dla 5% poziomu istotności takie stwierdzenie można wysnuć dla jeszcze dwóch zmiennych dotyczących ceny zamknięcia - Δf oraz $\Delta\alpha$. W przypadku zmiennych opartych o liczbę transakcji, różnice są istotne na poziomie 10% dla R_τ oraz $\Delta\alpha$. W przypadku pozostałych zmiennych, ANOVA wskazuje na brak wpływu zmiennej niezależnej.

zmienna	Sum Sq	Mean Sq	F value	Pr(>F)
R_τ - c.z.	0.4463	0.4463	35.21	2.77e-07
$\Delta\alpha$ -c.z.	5.31	5.309	4.864	0.032
Δf - c.z.	0.0545	0.05452	6.173	0.0164
R_α - c.z.	0.0166	0.01665	0.334	0.566
R_τ - l.tr.	0.0406	0.04060	3.348	0.0733
$\Delta\alpha$ - l.tr.	2.43	2.4338	3.115	0.0837
Δf - l.tr.	0.0001	0.000062	0.008	0.928
R_α - l.tr.	0.0007	0.00073	0.019	0.892

Tablica 3.5: Zestawienie wyników jednoczynnikowej analizy wariancji - ANOVA dla zależności poszczególnych zmiennych od zmiennej celu.

Zbadanie samego istnienia wpływu zmiennej objaśniającej na objaśnianą nie informuje o potencjale skuteczności dyskryminacji względem danej zmiennej objaśnianej. W celu zbadania mocy dyskryminacyjnej stojącej za poszczególnymi zmiennymi, obliczono wartości **uproszczonego⁷ współczynnika Gini** rozumianego jako stosunek pola zawartego pomiędzy dystrybuantami empirycznymi wyznaczonymi dla obu wartości zmiennej niezależnej do całości pola diagramu. Jakościowo obrazują je zakreskowane obszary na Rysunkach 3.6 , 3.7. Zestawienie wartości współczynnika dla wszystkich zmiennych zawarto w Tabeli 3.6.

⁷Niedokładność wynika tutaj z oparcia dystrybuant o szeregi rozdzielcze rozkładów empirycznych, a nie bezpośrednio o rozkłady.

Współczynnik Gini przyjmuje wartości z przedziału $[0;100\%]$, przy czym 100% oznacza maksymalną zdolność dyskryminacyjną. Należy pamiętać, że nie jest wielkością addytywną - Gini dla modelu wielu zmiennych nie stanowi sumy wartości współczynników zmiennych tworzących go, co wynika istnienia niezerowych komomentów⁸. Wynik modelu zawarty jest między największą wartością Gini spośród wszystkich wykorzystanych zmiennych, a wspomnianą sumą, jednak nie większy niż 100%. Dokładanie kolejnych zmiennych do modelu nie powinno zmniejszyć zdolności dyskryminacyjnej, niezależnie od stopnia jej korelowania z pozostałymi zmiennymi.

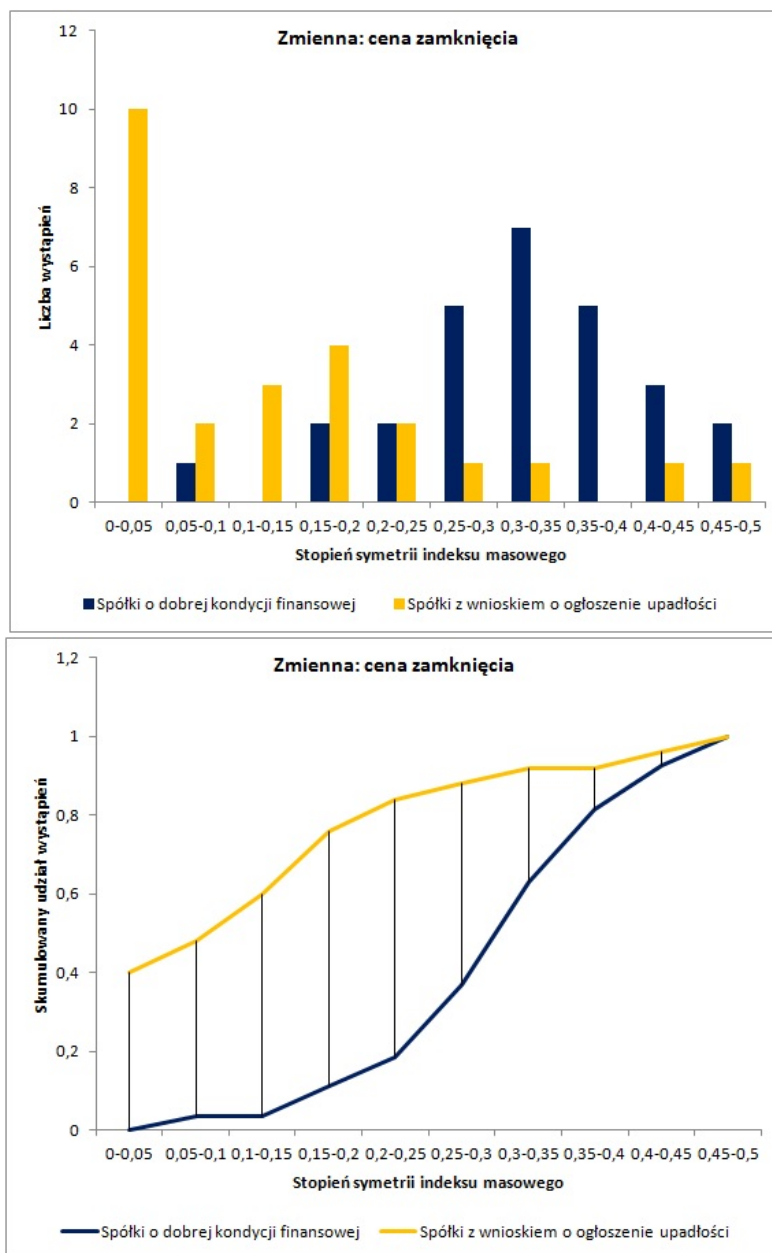
zmienna	uproszczony Gini [%]
R_τ - cena zamknięcia	36,49
$\Delta\alpha$ - cena zamknięcia	29,54
Δf - cena zamknięcia	15,80
R_α - cena zamknięcia	6,00
R_τ - liczba transakcji	10,67
$\Delta\alpha$ - liczba transakcji	14,85
Δf - liczba transakcji	4,81
R_α - liczba transakcji	8,90

Tablica 3.6: Wartości współczynnika Gini dla poszczególnych elementów charakterystyki multifraktalnej badanej próby.

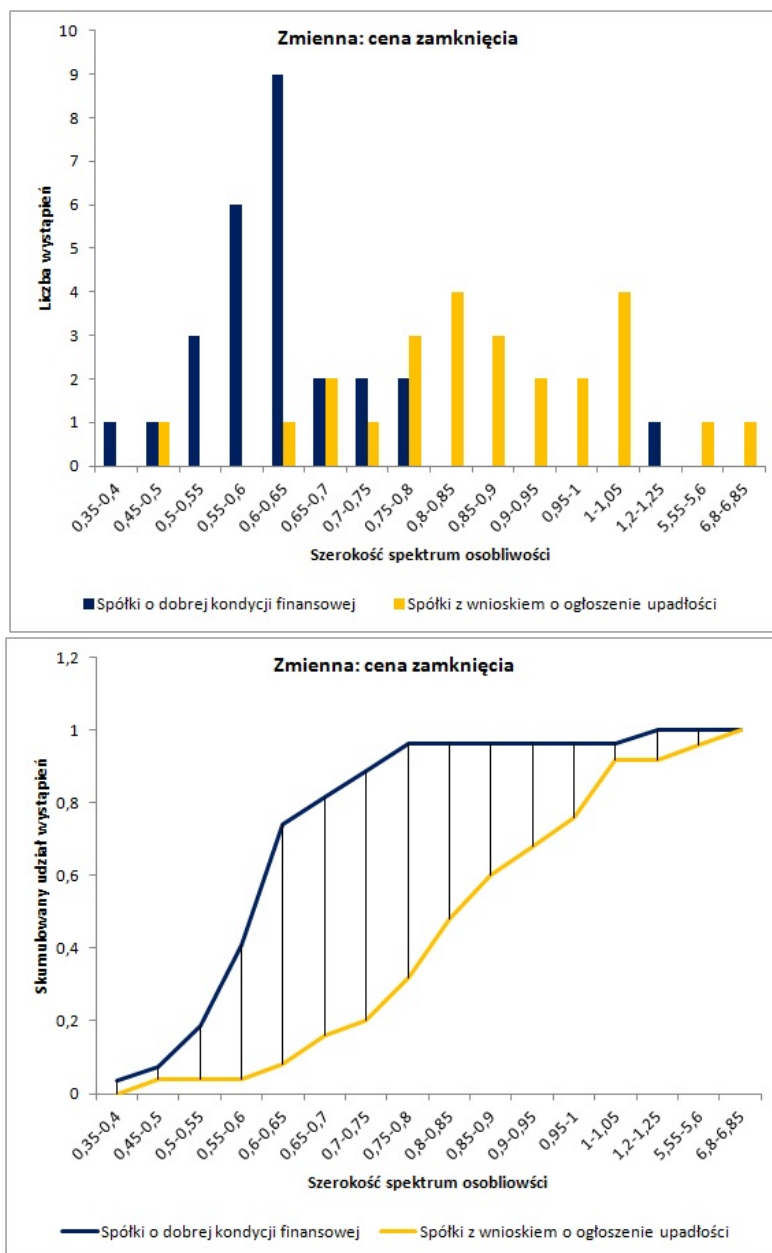
Największy potencjał wniesienia wkładu do budowanego w kolejnym podrozdziale modelu dyskryminacyjnego przejawia stopień symetrii indeksu masowego R_τ dla ceny zamknięcia (Gini na poziomie 36,5%). Blisko 30% wartość wskaźnika wykazuje szerokość spektrum osobliwości $\Delta\alpha$, również dla ceny zamknięcia. Szeregi rozdzielcze oraz dystrybuanty próby dla obu zmiennych pokazano na Rysunkach 3.6, 3.7, gdzie zakresowane obszary prezentują pole stanowiące licznik współczynnika Gini.

3 z 4 zmiennych dotyczących ceny zamknięcia wykazują wyższą zdolność dyskryminacyjną niż dla liczby transakcji. Wyjątkiem jest indeks asymetrii R_α , tutaj jednak w obu przypadkach Gini nie przekracza 10%.

⁸Zwykle analizowana jest kowariancja, jednak istotne mogą okazać się również komomentary wyższych rzędów, jak np. koskość czy kokurtoza.



Rysunek 3.6: Uzyskane rozkłady oraz dystrybuanty wyników stopnia symetrii indeksu masowego dla ceny zamknięcia.



Rysunek 3.7: Uzyskane rozkłady oraz dystrybuanty wyników szerokości spektrum osobliwości dla ceny zamknięcia.

3.4 SVM - budowa i walidacja modelu

Wektory obserwacji wykorzystujące otrzymane uprzednio wyniki charakterystyk multifrakalnych poddano dyskryminacji metodą maszyn wektorów nośnych. Modelowanie wykonano w środowisku R z wykorzystaniem pakietu `e1071`. Obraną metodą walidacji otrzymanych modeli został sprawdzian krzyżowy N -krotnej walidacji na próbie N -krotnej polegający na sprawdzeniu poprawności przypisania pojedynczej obserwacji na podstawie modelu otrzymanego z reszty próby (*leave-one-out cross-validation* - **LOOCV**) wykonany kolejno dla wszystkich obserwacji. Zatem, w tym przypadku LOOCV polegała na utworzeniu modelu z 51 obserwacji i walidacji 1 dla wszystkich 52 kombinacji.

W celu optymalizacji zdolności dyskryminacyjnej, modelowanie przeprowadzono dla zależności zmiennej objaśnianej od wszystkich pozostałych elementów wektora obserwacji dla kilku funkcji jądrowych. Dla każdego z typów modeli wykonano LOOCV, której wyniki zestawiono w Tabeli 3.7. Najwyższą skuteczność modeli otrzymano dla wykorzystujących sigmoidalną funkcję jądrową i taką też funkcję obrano ostatecznie w modelu. Pełną listę wyników sprawdzaniu krzyżowego LOOCV obranego modelu SVM zawarto w Tabeli 3.8.

funkcja jądrowa	Poprawność LOOCV [%]
liniowa	82,69
wielomianowa	59,62
sigmoidalna	86,54
radialna	75,00

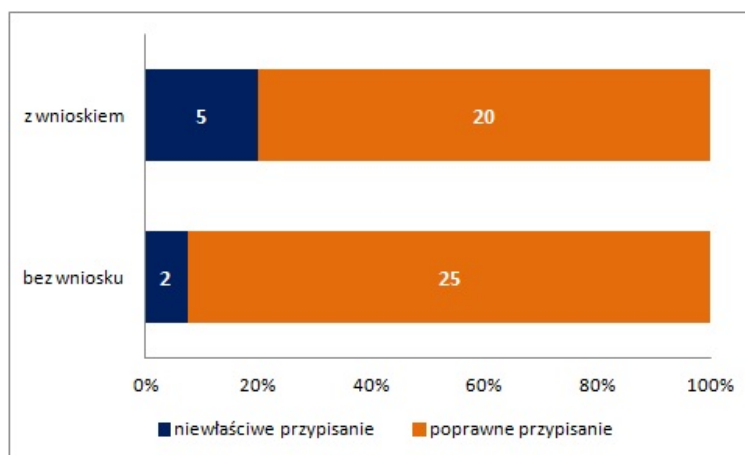
Tablica 3.7: Zestawienie wyników sprawdzianu krzyżowego LOOCV w zależności od rodzaju wykorzystywanej funkcji jądrowej w modelu SVM.

Walidacja sprawdzianem krzyżowym LOOCV wskazała skuteczność modeli cząstkowych na poziomie 86,5%. Błędnie przypisanych zostało 7 spośród 52 obserwacji, z czego 5 pochodziło z grupy *z wnioskiem* co dało 80% skuteczność w tej grupie (patrz Rysunek 3.8). Wśród spółek *z wnioskiem* jedynie 7,4% walorów zaklasyfikowano błędnie.

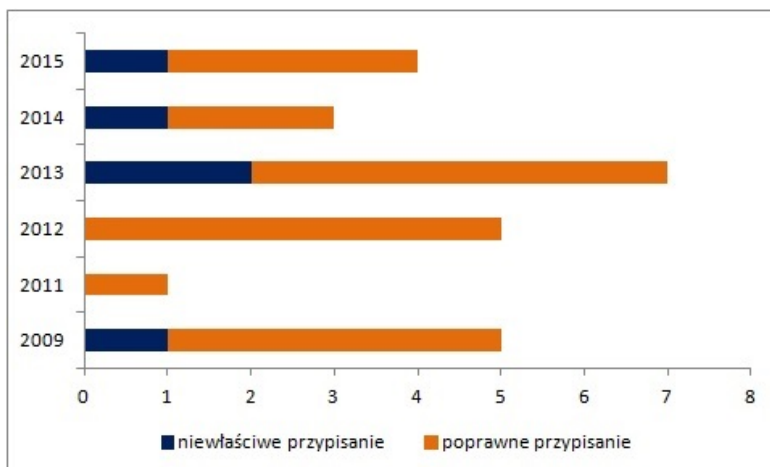
Rozłożenie niewłaściwych przypisań nie wykazuje koncentracji względem czasu - w obu grupach błędy pojawiają się jednostkowo w czasie, za wyjątkiem roku 2013 i grupy *z wnioskiem* gdzie pojawiły się 2 przypadki (patrz Rysunki 3.9, 3.10).

Walor	nr okresu (dotyczy spółek <i>bez wniosku</i>)	Czy wniosek o ogłoszenie upadłości?
PULAWY	III	NIE
COMARCH	II	NIE
KROSNO		TAK
INSTAL KRK		TAK
IDM		TAK
CASH FLOW		TAK
ANTI		TAK

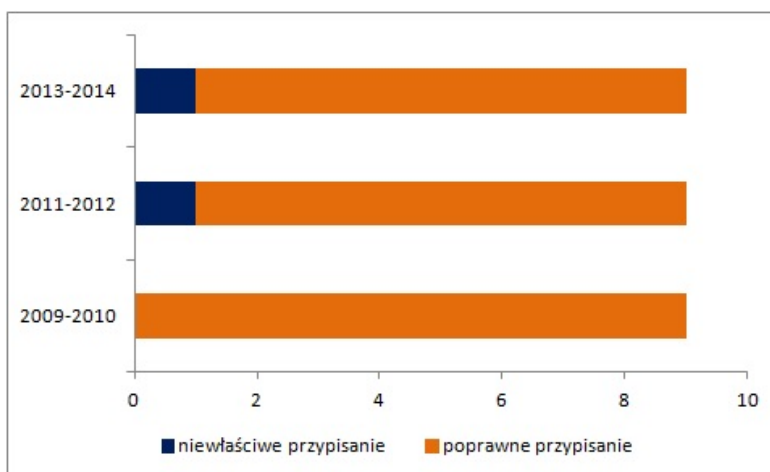
Tablica 3.8: Zestawienie błędnie przypisanych walorów w wyniku sprawdzianu krzyżowego LOOCV.



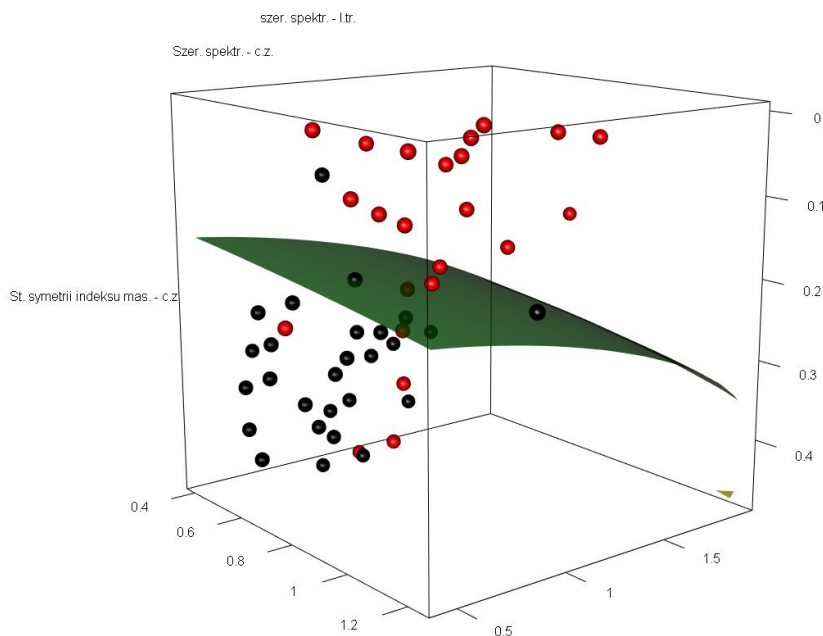
Rysunek 3.8: Poprawność sprawdzianu krzyżowego LOOCV dla całości próby w podziale na spółki z wnioskiem i bez wniosku.



Rysunek 3.9: Poprawność sprawdzianu krzyżowego LOOCV dla walorów spółek, dla których złożono wnioski o ogłoszenie upadłości w zależności od roku złożenia wniosku.



Rysunek 3.10: Poprawność sprawdzianu krzyżowego LOOCV dla walorów spółek o wysokim ratingu Altmana (bez wniosku) w zależności od lat notowań.



Rysunek 3.11: Wizualizacja dyskryminacji dla 3 zmiennych objaśniających wraz z płaszczyzną podziału otrzymaną przy zastosowaniu SVM. Na czerwono zaznaczono spółki, dla których miało miejsce złożenie wniosku o ogłoszenie upadłości.

Ideą funkcji dyskryminacyjnych jest utworzenie hiperpłaszczyzny skutecznie rozdzielającej populację względem zmiennej niezależnej. W przypadku modeli wielu zmiennych (>3) wizualizacja pełnego modelu staje się niemożliwa. Z tego względu zdecydowano się na zaprezentowanie płaszczyzny podziału w przestrzeni 3 zmiennych wykazujących wysokie zdolności dyskryminacyjne, tj. R_{τ} dla ceny zamknięcia, $\Delta\alpha$ dla ceny zamknięcia oraz ilości transakcji (patrz Rysunek 3.11).

3.5 Zastosowanie modelu MF DFA-SVM na przykładzie spółki CD Projekt S.A.

Uzyskany model zastosowano w kwantyfikacji ryzyka upadłości spółki CD Projekt S.A., a następnie w stworzeniu prymitywnego algorytmu kupna / sprzedaży akcji emitenta. Wybór padł na tą spółkę ze względu na jej bogatą historię, zawierającą m.in. złożenie przez emitenta wniosku o ogłoszenie upadłości z możliwością zawarcia układu.

Dzisiaj CD Projekt zajmuje się, przede wszystkim, produkcją i dystrybucją gier komputerowych, jednak zarówno nazwa, jak i kluczowe obszary działalności zmieniały się na przestrzeni czasu. Pierwotną nazwą wprowadzonej w 1994 roku na giełdę spółki była Optimus S.A. - powszechnie kojarzona z dostawcą komputerów, osiągającym blisko 15% udziału i pozycję lidera w rodzimym rynku pod koniec ubiegłego wieku. Spółka, zarządzana wówczas przez Romana Kluskę - twórcę przedsięwzięcia, miała swój udział w informatyzacji polskich szkół. W ramach jej działalności stworzono Onet.pl - jeden z czołowych portali horyzontalnych, sprzedany później grupie ITI.

W 2000 roku pojawiły się problemy firmy wynikające z nałożonych w wyniku kontroli skarbowej kar na podmiot. Powodem była wspomniana już umowa Optimusa z Ministerstwem Edukacji dotycząca dostaw komputerów, w ramach której spółka eksportowała sprzęt na Słowację celem późniejszego importu przez Ministerstwo. Operacja powstała po stronie Ministerstwa i umożliwiła mu zakup towaru z zerową stawką VAT. Sprawa została przekazana przez fiskusa do prokuratury i ciągnęła się kilka lat by ostatecznie zakończyć się wyrokiem Naczelnego Sądu Administracyjnego, całkowicie oczyszczającego Optimus S.A. ze stawianych zarzutów. Utraconych korzyści, renomy, zaufania ze strony kontrahentów czy instytucji kredytowych, jak również wielu zwolnionych pracowników tworzących sukces spółki nie udało się odzyskać.

Problemy dotknęły spółkę również w późniejszych latach. W sierpniu 2006 roku Komisja Papierów Wartościowych i Giełd (KPWiG, dziś funkcjonująca w ramach Komisji Nadzoru Finansowego) podjęła względem waloru decyzję wykluczeniu go z obrotu i nałożyła na spółkę karę pieniężną. Argumentami dla decyzji były:

- celowe ukrycie informacji o podwyższeniu kapitału zakładowego w wyniku prywatnej emisji akcji;
- braki w informowaniu o umowie znaczącej (sprzedaż akcji spółki Zatra);
- nieterminowe przekazywanie informacji (podpisanie aneksu do umowy kredytowej z Bankiem BPH);
- brak przekazania listy akcjonariuszy uprawnionych do głosowania na NWZA⁹ z dnia 10 lipca 2006.

Optimus złożył wniosek o ponowne rozpatrzenie sprawy - ostatecznie walory pozostały w obrocie, jednak negatywna ocena inwestorów zaczęła pogłębiać się.

W latach 2007 i 2008 dokonywano zatrzymań członków zarządu emitenta: Piotra Lewandowskiego i Michała Lorenca przez Centralne Biuro Śledcze w ramach stawianych zarzutów nieprawego podwyższenia kapitału w spółce, a także Piotra Lewandowskiego przez Centralne Biuro Antykorupcyjne z zarzutem wyprowadzenia kapitału ze spółki poprzez zawarcie fikcyjnej umowy z firmą windykacyjną.

W styczniu 2008 roku emitent złożył do Sądu wnioski o ogłoszenie upadłości, który później wycofał ratując spółkę finansowaniem środkami pozyskanymi w wyniku przeprowadzenia publicznej emisji akcji. W tym samym roku firma Zarta - wierzyciel Optimusa - złożyła kolejny taki wniosek, jednak został on oddalony przez Sąd.

Kolejny rok przyniósł poprawę nastrojów. Wśród inwestorów zaczęły krążyć pogłoski o pojawieniu się inwestora zagranicznego (w późniejszym terminie emitent podał informację o bezowocnych negocjacjach z Sincere Equity Limited), jak również o połączeniu z lokalnym podmiotem.

⁹Nadzwyczajne Walne Zgromadzenie Akcjonariuszy

Rok później przystąpiono do połączenia Optimusa z CD Projekt Investment w ramach przejęcia odwrotnego (*reverse takeover*). Działanie tego typu ma na celu umożliwienie nienotowanej na parkiecie spółce, jaką w tym wypadku był CD Projekt, względnie szybkie wejście na giełdę. Połączony podmiot obrał nową nazwę - CD Projekt RED S.A. - a jego dominującym obszarem działalności stała się produkcja i dystrybucja gier komputerowych, którą do tej pory prowadził CD Projekt. Sztandarowym produktem emitenta stała się popularna w Polsce, jak i na świecie, seria gier *Wiedźmin*.

W roku 2013 dokonano sprzedaży marki Optimus spółce AB S.A.

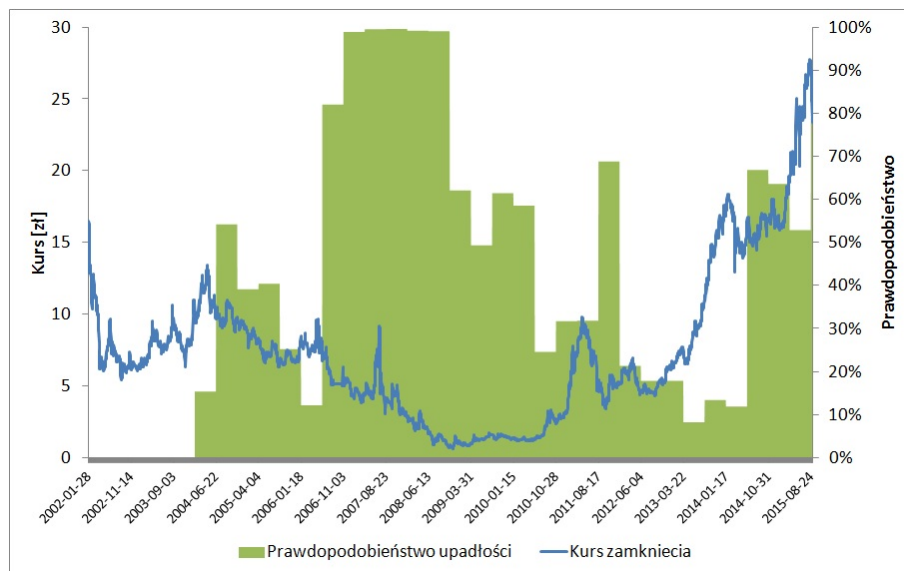
Rok	Zdarzenie
1988	założenie firmy Optimus przez Romana Kluskę
1994	wprowadzenie spółki Optimus S.A. na GPW
2000	sprzedaż akcji przez Romana Kluskę; zawiadomienie prokuratury przez organy skarbowe o możliwości popełnienia przestępstwa; wszczęcie śledztwa przeciw byłemu prezesowi
2002	złożenie przez spółkę skargi do NSA; zwolnienia grupowe
2003	NSA - decyzje fiskusa bezpodstawne; umorzenie śledztwa przeciw R. Klusce
2006	pozew o odszkodowanie; decyzja o wykluczeniu waloru z obrotu (nie weszła w życie)
2007	zatrzymania członków zarządu przez CBS
2008	zawieszenie działalności; 2 wnioski o ogłoszenie upadłości; zatrzymanie członka zarządu przez CBA
2009	wyrok NSA; pogłoski na temat potencjalnego inwestora
2010	przedstawienie planu połączenia z CD Projekt Investment
2011	połączenie; zmiana nazwy na CD Projekt RED S.A.; wprowadzenie do sprzedaży gry Wiedźmin II
2012	zmiana nazwy na CD Projekt S.A.
2013	sprzedaż marki Optimus
2014	spór sądowy z Empik S.A
2015	premiera gry Wiedźmin III

Tablica 3.9: Wybrane zdarzenia z historii spółki CD Projekt S.A. (opracowanie własne na podstawie [64], [65], [66]).

3.5.1 Ocena ryzyka upadłości spółki

Na wstępie należy dokonać wyjaśnienia sposobu użycia określenia *ryzyko upadłości*. W istocie, zadaniem utworzonego modelu jest dyskryminowanie spółek względem których złożono wnioski o ogłoszenie upadłości od tych, dla których rzecz taka nie miała miejsca. *De facto* ocenie podlega prawdopodobieństwo złożenia wspomnianego wniosku, nie zaś ostateczna decyzja sądu. Na potrzeby niniejszej pracy przyjęto arbitralnie, że w rozumieniu hipotetycznego inwestora pojawienie się informacji o złożeniu do sądu wniosku o ogłoszenie upadłości emitenta stanowi ostateczne potwierdzenie złej kondycji spółki. W takich warunkach rozważany inwestor wycofuje się z partycypacji w spółce.

Pojedyncze spektrum multifraktalne określano na próbie 500 dni sesyjnych. Przeliczenia dokonywano co 100 dni sesyjnych począwszy od 27 stycznia 2004.

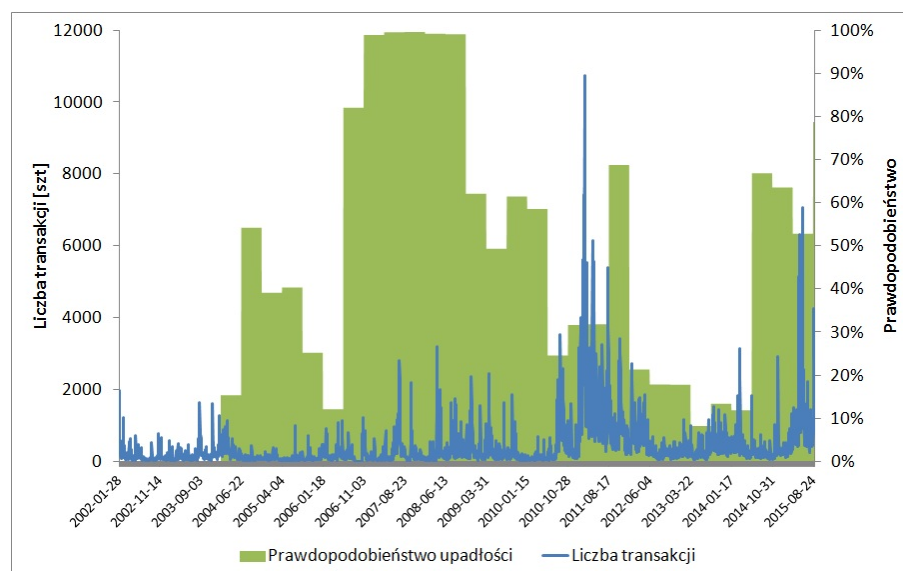


Rysunek 3.12: Kształtowanie się ceny zamknięcia dla notowań spółki CD Projekt S.A. z nałożonymi poziomami wyznaczonego ryzyka upadłości emitenta.

W ocenie wykorzystano dostępną w pakiecie e1071 środowiska R funkcję przypisującą wartość prawdopodobieństwa przynależności obserwacji do określonej grupy. Określono prawdopodobieństwo pochodzenia z subpopulacji spółek

z wnioskiem. Wyniki znajdują się w Tabeli 3.10, a ich zobrazowanie wraz z naniesioną ceną zamknięcia i liczbą transakcji przedstawiają odpowiednio Rysunki 3.12 oraz 3.13.

Wyznaczone poziomy ryzyka korespondują w znacznej mierze ze zdarzeniami z historii spółki. Najwyższe wskazania otrzymano w latach 2006 - 2008, gdzie zanotowano 99% wartości prawdopodobieństwa złożenia wniosku o ogłoszenie upadłości spółki. W rzeczywistości tak też się stało - na początku 2008 roku emitent złożył tego rodzaju wniosek do sądu. Ponowne złożenie wniosku, tym razem przez spółkę Zatra, miało miejsce we wrześniu 2008. Jak opisano we wcześniejszym podrozdziale, lata 2006-2007 również dawały inwestorom powody do negatywnej oceny spółki (patrz Tabela 3.9).



Rysunek 3.13: Kształtowanie się liczby transakcji dla notowań spółki CD Projekt S.A. z nałożonymi poziomami wyznaczonego ryzyka upadłości emitenta.

Poczynając od drugiej połowy roku 2008 do roku 2014 ryzyko wykazuje tendencję spadkową, co koresponduje z poprawą oceny procesów zachodzących w spółce przez akcjonariat. W tym czasie nastąpiło przejście odwrotne Optimusa przez CD Projekt Investment, reorganizacja, zmiana kluczowych obszarów działalności, sukcesy sprzedażowe *Wiedźmina*, rozwój dystrybucji cyfrowej gier komputerowych przez GOG.com.

Data końca okresu	Prawdopodobieństwo upadłości
2004-01-27	15,3%
2004-06-21	54,1%
2004-11-09	39,0%
2005-04-01	40,3%
2005-08-25	25,1%
2006-01-17	12,1%
2006-06-12	82,0%
2006-11-02	98,8%
2007-03-27	99,5%
2007-08-22	99,5%
2008-01-17	99,1%
2008-06-12	99,0%
2008-10-31	62,0%
2009-03-30	49,3%
2009-08-21	61,4%
2010-01-14	58,5%
2010-06-09	24,5%
2010-10-27	31,6%
2011-03-22	31,7%
2011-08-16	68,7%
2012-01-09	21,3%
2012-06-01	17,8%
2012-10-24	17,7%
2013-03-21	8,1%
2013-08-19	13,3%
2014-01-16	11,8%
2014-06-10	66,7%
2014-10-30	63,5%
2015-03-30	52,8%

Tablica 3.10: Zestawienie wyników oceny ryzyka upadłości spółki CD Projekt S.A. z użyciem modelu MF DFA-SVM.

W roku 2014 silny trend wzrostowy kursu akcji spółki ulega zatrzymaniu i fluktuuje na poziomie 15 zł. Na ocenę inwestorów mogły mieć wpływ w tym okresie: sprawa sądowa wytoczona przez emitenta spółce Empik, a także słabe wyniki finansowe [67]. Ryzyko oszacowane dla tego okresu kształtuje się na poziomie 65% i spada w kolejnym do 53%.

3.5.2 Symulacja działania prymitywnego algorytmu generowania sygnałów kupna/sprzedaży

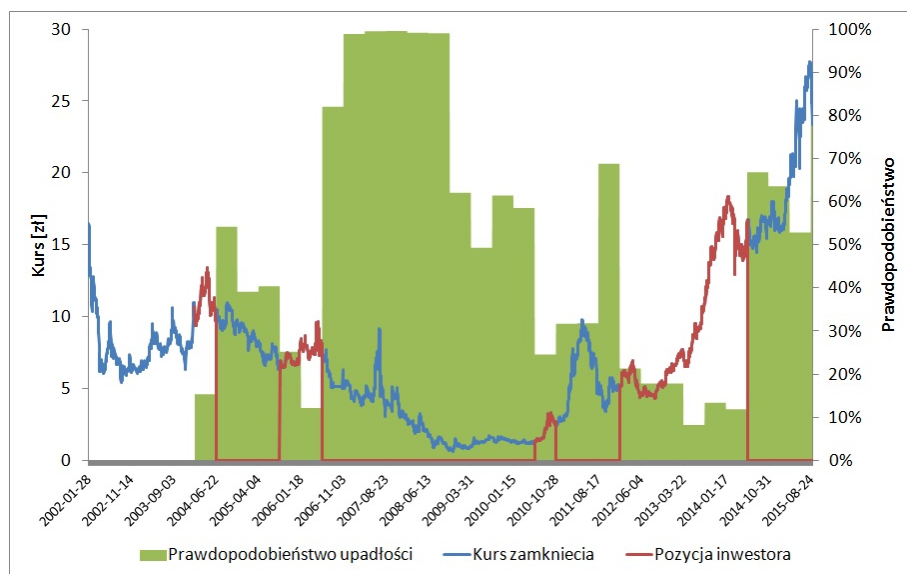
Wykorzystując opisane powyżej wyniki sporządzono symulację dla prostego algorytmu inwestycyjnego opartego o wynik modelu MF DFA-SVM.

Algorytm nie uwzględnia stopniowej akumulacji czy redukcji pozycji inwestycyjnej, a jedynie kupno lub sprzedaż waloru. Pominięte zostało także zakładanie zleceń typu *stop loss* oraz jakichkolwiek prostych elementów strategii wykorzystujących inne dostępne typy zleceń giełdowych.

Kolejnym uproszczeniem jest określenie ceny kupna i sprzedaży akcji na poziomie ceny zamknięcia, co w praktyce sprowadzałoby się najczęściej do składania zleceń na fixingu. Niekiedy w rzeczywistości mogłoby jednak nie dojść do zawarcia transakcji ze względu na brak zleceń przeciwnych albo zbyt późne złożenie zlecenia przez rozważanego inwestora w warunkach dysproporcji liczby oferowanych instrumentów po obu stronach arkusza zleceń.

Poziom akceptowalnego przez inwestora maksymalnego ryzyka określono arbitralnie jako 30%. Strategia działania zakłada dokonywanie kupna i sprzedaży waloru jedynie w dniach pomiaru ryzyka. Zmiana posiadania akcji ma miejsce w wypadku przekroczenia progu dla badanej wartości prawdopodobieństwa. *Ceteris paribus* oznacza to, że w przypadku posiadania przez inwestora akcji, określenie ryzyka na poziomie wyższym niż 30% skutkuje sprzedażą. W przypadku przeciwnym, inwestor, który dotychczas nie był w posiadaniu waloru dokonuje zakupu, a jeśli posiadał już akcje - nie wykonuje żadnych działań.

Zobrazowaniem stanu pozycji portfela inwestora jest naniesiona na uprzednio pokazywany wykres czerwona linia zaznaczona na Rysunku 3.14. Z kolei dokładne dane na temat poszczególnych transakcji, prostej stopy zwrotu jaka z nich wynikała oraz skumulowanej prostej stopy zwrotu z wszystkich zawartych transakcji zawarto w Tabeli 3.11.



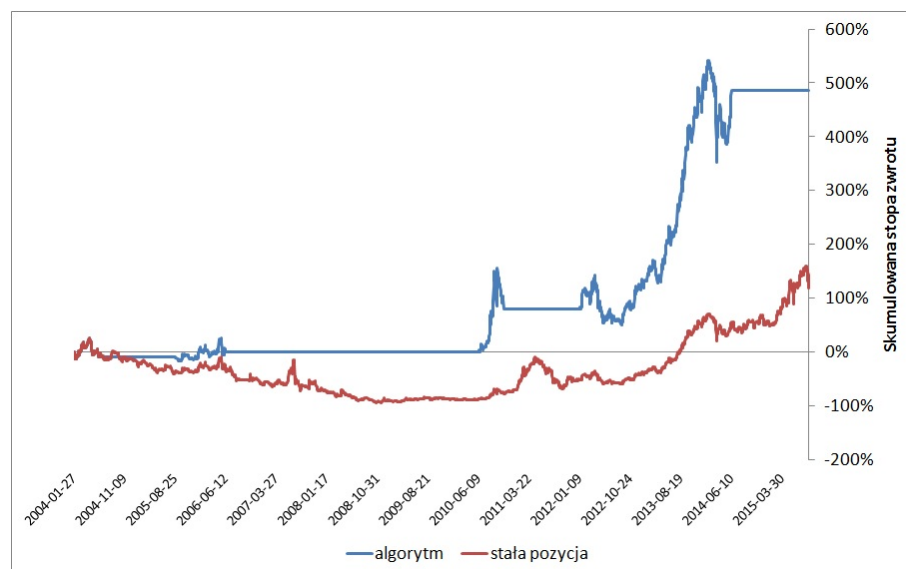
Rysunek 3.14: Kształtowanie się ceny zamknięcia dla notowań spółki CD Projekt S.A. z naniesioną linią obrazującą działanie algorytmu inwestycyjnego opartego o wyznaczone poziomy prawdopodobieństwa upadłości dla progu 30%.

Symulacja zastosowania algorytmu dla analizowanego okresu od stycznia 2004¹⁰ do sierpnia 2015 daje skumulowany zysk na poziomie 485%, podczas gdy dokonanie zakupu na początku tego okresu i sprzedaży u jego końca dałaby zwrot w wysokości 127% (patrz Rysunek 3.15). Jedynie w przypadku pierwszej pary transakcji K/S (kupna/sprzedaży) zanotowano stratę. Najwyższy zysk (227%), a zarazem najdłuższe utrzymanie waloru w posiadaniu osiągnięto w latach 2012-2014. Wysokie wartości ryzyka otrzymane w drugiej połowie 2014 roku oraz w roku 2015 wyeliminowały ewentualność osiągnięcia dodatkowej korzyści w wyniku kontynuacji zapoczątkowanego w 2012 roku silnego trendu wzrostowego. Z drugiej strony, hipotetyczny inwestor praktycznie nie partycypował w trendach spadkowych z roku 2005 oraz z lat 2007-2009.

¹⁰Pierwszym dniem zastosowania algorytmu jest dzień sesyjny, w którym po raz pierwszy wyznaczono ryzyko upadłości spółki.

Data	Typ zlecenia	Kurs [zł]	Stopa zwrotu [%]	Skumulowana stopa zwrotu [%]
2004-01-27	K	10,7		
2004-06-18	S	9,65	-9,8	-9,8
2005-08-25	K	6,95		
2006-06-09	S	7,65	10,1	-0,7
2010-06-09	K	1,3		
2010-10-26	S	2,35	80,8	79,5
2012-01-09	K	5,13		
2014-06-09	S	16,75	226,5	485,9

Tablica 3.11: Podsumowanie wyników na symulowanych transakcjach zawartych z użyciem algorytmu inwestycyjnego opartego o wyznaczone poziomy prawdopodobieństwa upadłości dla progu 30%.



Rysunek 3.15: Porównanie skumulowanej stopy zwrotu uzyskanej w wyniku algorytmu inwestycyjnego opartego o wyznaczone poziomy prawdopodobieństwa upadłości dla progu 30% z pasywną strategią zakupu i sprzedaży w krańcowych momentach okresu dla spółki CD Projekt S.A.

3.6 Podsumowanie badań

Niniejsza praca ukazała możliwość praktycznego zastosowania własności multifraktałnych notowań akcji oraz koncepcji zawartej we fraktalnej hipotezie rynku. Jej celem, w szczególności, było stworzenie modelu oceny ryzyka upadłości, który różnicowałby spółki na podstawie informacji zawartych w spektrach osobliwości ich notowań.

Metodą wyznaczania spektrum była multifraktałna analiza detrendowanych fluktuacji (MF-DFA), natomiast dyskryminację przeprowadzono przy użyciu metody maszyn wektorów nośnych (SVM).

W sumie badaniu poddano szeregi czasowe dziennych cen zamknięcia oraz liczby transakcji 35 spółek notowanych na GPW S.A. w latach 2009-2014, z czego w przypadku 26 spółek miało miejsce złożenie wniosku o ogłoszenie upadłości (spółki *z wnioskiem*). Pozostałe 9 to spółki o dobrej kondycji finansowej (spółki *bez wniosku*), tj. cechujące się wysokimi ratingami Altmana opartymi o model Z-score utworzony dla rynków wschodzących.

Notowania dla 25 spółek¹¹ *z wnioskiem* analizowane były w 2-letnim okresie bezpośrednio poprzedzającym złożenie wniosku. Walory *bez wniosku* poddano badaniu na 3 nieprzekrywających się 2-letnich oknach czasowych. Otrzymano w ten sposób 52 obserwacje, z których każda opierała się o 2 spektra osobliwości.

Uzyskane metodą MF-DFA spektra osobliwości zostały opisane charakterystyką składającą się z 4 elementów zawierającą: stopień symetrii indeksu masowego, szerokość spektrum, różnicę w wartościach spektrum dla skrajnych argumentów, indeks asymetrii spektrum.

Otrzymane charakterystyki dla ceny zamknięcia przejawiały większą zdolność skutecznej dyskryminacji niż te oparte o liczbę transakcji. Analiza ANOVA dla ceny zamknięcia wskazuje na 5% poziomie istotności na brak podstaw do odrzucenia hipotezy o wpływie przynależności spółki do grupy *z wnioskiem* lub *bez wniosku* na wartości 3 z 4 badanych elementów charakterystyki multifraktałnej. 2 z nich - szerokość spektrum oraz stopień symetrii indeksu masowego - uzyskały odpowiednio 30% i 36% zdolność dyskryminacyjną metodą współczynnika Gini.

¹¹CD Projekt S.A. zostało wyłączone z badań na tym etapie.

Wartości charakterystyk wspólnie ze zmienną binarną określającą przynależność do grupy z *wnioskiem* utworzyły 9-elementowe wektory obserwacji, które poddano analizie SVM. Modele SVM utworzono dla kilku funkcji jądrowych. Zdolność dyskryminacyjną modeli oceniano sprawdzianem krzyżowym wielokrotnego jednoelementowego wyłączenia (LOOCV). Najlepszą z propozycji okazał się model wykorzystujący sigmoidalną funkcję jądrową, dla którego osiągnięto blisko 87% skuteczność walidacji.

Uzyskany model zastosowano w ocenie ryzyka upadłości spółki CD Projekt S.A. Notowania jej akcji poddano badaniu na 29 częściowo przekrywających się oknach czasowych o długości 500 dni sesyjnych. Analizę MF-DFA wykonywano co 100 dni sesyjnych, a jej charakterystyki poddawano ocenie modelem SVM. Pomiar dokonywane w czasie słabej kondycji spółki, w szczególności w okolicach daty złożenia wniosku o upadłość spółki, wskazywały na ponad 95% prawdopodobieństwo zaistnienia tego zdarzenia. Jednocześnie stopniowa poprawa jej sytuacji zbiegała się w czasie ze znacznym i prawie systematycznym spadkiem wartości oceny ryzyka.

Wyznaczone poziomy ryzyka posłużyły do przeprowadzenia symulacji strategii inwestycyjnej, w której o zakupie, trzymaniu, sprzedaży akcji decydowało przekroczenia 30% progu prawdopodobieństwa ryzyka złożenia wniosku. W wyniku zastosowania tego prostego algorytmu dla spółki CD Projekt S.A. otrzymano blisko czterokrotnie wyższą skumulowaną stopę zwrotu nad zyskiem osiąganym poprzez zakup akcji na początku obranego okresu i sprzedaż u jego końca. Całkowity zysk w ujęciu procentowym wyniósł 485%, co dało ponad 16% przeciętną roczną stopę zwrotu.

3.7 Wnioski

Niewątpliwie najistotniejszymi z ogólnych wniosków płynących z przeprowadzonych badań są:

- I **Multifraktalna natura zmienności szeregów czasowych notowań zależy od kondycji spółki.**
- II **Istnieje możliwość dokonania pomiaru rynkowej oceny ryzyka stwarzanego z walorem z wykorzystaniem koncepcji rynku fraktalnego.**

W szczególności, zarówno w aspekcie poznawczym, jak i utylitarnym wnioskować można następująco:

1. **Charakterystyki multifraktalne spółek, względem których złożono wniosek o ogłoszenie upadłości różnią się od charakterystyk spółek o dobrej kondycji finansowej.**

Analiza ANOVA 4-elementowych charakterystyk opartych o szeregi cen zamknięcia wykazała na 5% poziomie istotności wpływ przynależności do grupy spółek, dla których dokonano złożenia wniosku o ogłoszenie upadłości na 3 elementy charakterystyki. W przypadku szeregów liczby transakcji ustalono taki wpływ na 2 elementy na 90% poziomie ufności.

2. **Różnice w charakterystykach pozwalają na utworzenie relatywnie silnego (w odniesieniu do ilości użytych obserwacji) modelu dyskryminacyjnego.**

Model SVM oparty o sigmoidalną funkcję jądrową wykazał podczas wielokrotnego sprawdzaniu krzyżowego z jednoelementowym wyłączeniem skuteczność na poziomie 87%.

3. **Ocena ryzyka dokonana przy zastosowaniu modelu MF DFA-SVM pozytywnie koreluje z sytuacją finansową spółki.**

W czasach trudności finansowych Optimus S.A. szacowane ryzyko było bliskie jedności, natomiast w okresie poprawy sytuacji spółki obserwowany był znaczny jego spadek.

4. **Wykorzystanie modelu może prowadzić do osiągnięcia ponadprzeciętnych stóp zwrotu w średnim i długim horyzoncie inwestycyjnym.**

Prymitywny algorytm oparty o utworzony model zyskał w rozważanym okresie blisko czterokrotnie więcej niż strategia pasywna.

Spis rysunków

1.1	Rozkład współczynników wagowych przypisywanych poziomom ryzyka w poszczególnych chwilach czasu w modelowaniu ryzyka z użyciem RiskMetrics, MDE, GARCH(źródło: [8]).	8
1.2	Rodzaje modeli wykorzystywanych w tworzeniu systemów wczesnego ostrzegania (źródło: [42]).	10
1.3	Udział handlu algorytmicznego w ogóle wolumenu obrotu NYSE w poszczególnych latach (źródło: [14]).	15
1.4	Udział HFT oraz innych rodzajów handlu algorytmicznego w ogóle obrotów generowanych tym sposobem w poszczególnych latach. (źródło: [14]).	18
1.5	Zmiany cen PG oraz Accenture w dniu Krachu 2:45 (źródło: [58]).	19
2.1	Zestawienie przykładów fraktali: matematycznego oraz statystycznego (źródło: [52]).	23
2.2	Przykład dwóch kroków generowania zdefektowanego dywanu Siepińskiego.	24
2.3	Sposób tworzenia przypadkowej krzywej Kocha (źródło: [54]). . .	25
2.4	Idea przekształcenia liniowego wykorzystywanego w metodzie LDA (źródło: [60]).	32
2.5	Przykład dyskryminacji w dwuwymiarowej przestrzeni zmiennych uczących z wykorzystaniem LDA (linia przerywana) oraz QDA (linia ciągła) (źródło: [59]).	34

3.1 Szeregi czasowe ceny zamknięcia oraz jej zmienności względnej dla waloru ABM SOLID.	49
3.2 Profil z naniesioną przykładowo funkcją detrendującą, wykres fluktuacji lokalnych, oraz zależności funkcji fluktuacji od wielkości okna dla wszystkich badanych rzędów q	51
3.3 Zależność wykładnika Hursta od rzędu q , wymiaru Renyi $R(q)$, oraz spektrum osobliwości $f(a)$	52
3.4 Zestawienie wykresów pudełkowych zmiennych objaśniających dotyczących ceny zamknięcia dla spółek: <i>z wnioskiem</i> - tak, <i>bez wniosku</i> - nie.	54
3.5 Zestawienie wykresów pudełkowych zmiennych objaśniających dotyczących liczby transakcji dla spółek: <i>z wnioskiem</i> - tak, <i>bez wniosku</i> - nie.	55
3.6 Uzyskane rozkłady oraz dystrybuanty wyników stopnia symetrii indeksu masowego dla ceny zamknięcia.	58
3.7 Uzyskane rozkłady oraz dystrybuanty wyników szerokości spektrum osobliwości dla ceny zamknięcia.	59
3.8 Poprawność sprawdzianu krzyżowego LOOCV dla całości próby w podziale na spółki z wnioskiem i bez wniosku.	61
3.9 Poprawność sprawdzianu krzyżowego LOOCV dla walorów spółek, dla których złożono wnioski o ogłoszenie upadłości w zależności od roku złożenia wniosku.	62
3.10 Poprawność sprawdzianu krzyżowego LOOCV dla walorów spółek o wysokim ratingu Altmana (bez wniosku) w zależności od lat notowań.	62
3.11 Wizualizacja dyskryminacji dla 3 zmiennych objaśniających wraz z płaszczyzną podziału otrzymaną przy zastosowaniu SVM. Na czerwono zaznaczono spółki, dla których miało miejsce złożenie wniosku o ogłoszenie upadłości.	63
3.12 Kształtowanie się ceny zamknięcia dla notowań spółki CD Projekt S.A. z nałożonymi poziomami wyznaczonego ryzyka upadłości emitenta.	67

- 3.13 Kształtowanie się liczby transakcji dla notowań spółki CD Projekt S.A. z nałożonymi poziomami wyznaczonego ryzyka upadłości emitenta. 68
- 3.14 Kształtowanie się ceny zamknięcia dla notowań spółki CD Projekt S.A. z naniesioną linią obrazującą działanie algorytmu inwestycyjnego opartego o wyznaczone poziomy prawdopodobieństwa upadłości dla progu 30%. 71
- 3.15 Porównanie skumulowanej stopy zwrotu uzyskanej w wyniku algorytmu inwestycyjnego opartego o wyznaczone poziomy prawdopodobieństwa upadłości dla progu 30% z pasywną strategią zakupu i sprzedaży w krańcowych momentach okresu dla spółki CD Projekt S.A. 72

Spis tablic

3.1	Zestawienie instrumentów wykorzystanych w analizie, dla których względem spółki emitenta złożono wnioski o ogłoszenie upadłości.	44
3.2	Wartości graniczne dla poszczególnych ratingów modelu Z-score (źródło: [62]).	46
3.3	Zestawienie spółek wykorzystanych w analizie, dla których nie miało miejsca złożenie wniosku o ogłoszenie upadłości, a które cechowały wysokie ratingi Altmana (źródło: [62]).	47
3.4	Zestawienie wyników otrzymanej charakterystyki multifraktalnej dla zmiennych: cena zamknięcia oraz liczba transakcji, dla spółki ABM SOLID S.A.	53
3.5	Zestawienie wyników jednoczynnikowej analizy wariancji - ANOVA dla zależności poszczególnych zmiennych od zmiennej celu.	56
3.6	Wartości współczynnika Gini dla poszczególnych elementów charakterystyki multifraktalnej badanej próby.	57
3.7	Zestawienie wyników sprawdzianu krzyżowego LOOCV w zależności od rodzaju wykorzystywanej funkcji jądrowej w modelu SVM.	60
3.8	Zestawienie błędnie przypisanych walorów w wyniku sprawdzianu krzyżowego LOOCV.	61
3.9	Wybrane zdarzenia z historii spółki CD Projekt S.A. (opracowanie własne na podstawie [64], [65], [66]).	66
3.10	Zestawienie wyników oceny ryzyka upadłości spółki CD Projekt S.A. z użyciem modelu MFDFA-SVM.	69

3.11 Podsumowanie wyników na symulowanych transakcjach zawartych z użyciem algorytmu inwestycyjnego opartego o wyznaczone poziomy prawdopodobieństwa upadłości dla progu 30%.	72
---	----

Bibliografia

- [1] K. Jajuga, T. Jajuga, *Inwestycje - instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa*, PWN, Warszawa, 2011.
- [2] M. Miller, *Mathematics and Statistics for Financial Risk Management*, John Wiley Sons, Inc., Hoboken, 2014.
- [3] E. Elton, M. Gruber, S. Brown, W. Goetzmann, *Modern Portfolio Theory and Investment Analysis*, John Wiley Sons Inc., Hoboken, 2010.
- [4] R. Haugen, *Teoria Nowoczesnego Inwestowania*, WIG Press, Warszawa, 1996.
- [5] N. Amenc and V. Le Sourd, *Portfolio Theory and Performance Analysis*, John Wiley Sons Ltd., Southern Gate, 2003.
- [6] J. Hull, *Options, Futures and other derivatives*, Pearson Education, Inc., Upper Saddle River, 2006.
- [7] F. Fabozzi, *Bond Markets, Analysis and Strategies*, Prentice-Hall Inc., Upper Saddle River, 2000.
- [8] L. Allen, J. Boudoukh and A. Saunders, *Understanding Market, Credit, and Operational Risk (the Value at Risk approach)*, Blackwell Publishing Ltd., Malden, 2004.
- [9] D. Grech and Z. Mazur, *Can One Make Any Crash Prediction in Finance Using the Local Hurst Exponent Idea?*, Physica A 336, pp. 133-145, 2004.
- [10] A. Schmidt, *Quantitative Finance for Physicists - an Introduction*, Elsevier Inc., Burlington, 2005.

- [11] E. E. Peters, *Fractal Market Analysis - Applying Chaos Theory to Investment and Economics*, John Wiley Sons Inc., New York, 1994.
- [12] J. Kantelhardt, *Fractal and Multifractal Time Series*, 2008.
- [13] W. Dębski, *Rynek Finansowy i jego mechanizmy*, PWN, Warszawa, 2000.
- [14] R. Kissell, *The Science of Algorithmic Trading and Portfolio Management*, Elsevier Inc., San Diego, 2014.
- [15] M. Glantz, R. Kissell, *Multi-Asset Risk Modeling Techniques for a Global Economy in an Electronic and Algorithmic Trading Era*, Elsevier Inc., 2014.
- [16] K. Kendall, *Electronic and Algorithmic Trading Technology The Complete Guide*, Elsevier Inc., 2007.
- [17] O. Snguanyat, *Stochastic modelling of financial time series with memory and multifractal scaling*, PhD Thesis at Queensland University of Technology, Brisbane, 2009.
- [18] G. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley Sons, Inc., Hoboken, 2004.
- [19] R. Fisher, *The Use of Multiple Measurements in Taxonomic Problems*, Annale of Eugenics, 7: 179-188, 1936.
- [20] M. Trzęsiok, *Metody i modele analiz ilościowych w ekonomii i zarządzaniu*, Prace Naukowe AE w Katowicach, Katowice, 2009.
- [21] T. Van Gestel, B. Baesens, J. Garcia, P. Van Dijke, *A Support Vector Machine Approach to Credit Scoring*, 2003.
- [22] J. Tang, J. Chen, C. Zhang, *Nonlinear Analog Circuit Fault Diagnosis Based on MFDFA Method*, Applied Mechanics and Materials Vols 263-266 (2013) pp 108-113, 2013.
- [23] Y. Zhang, *Modeling High-Frequency Order Book Dynamics with Support Vector Machines*. Phd Thesis, 2013.
- [24] Q. Wen, Z. Yang, Y. Song, P. Jia, *Intelligent Stock Trading System based on SVM Algorithm and Oscillation Box Prediction*, roceedings of International Joint Conference on Neural Networks, Atlanta, 2009.

- [25] L. Cao, E. Tay, *Support Vector Machine With Adaptive Parameters in Financial Time Series Forecasting*, IEEE Transactions On Neural Networks, vol. 14, No. 6, November 2003.
- [26] P. Sanada, M. Ankita, G. Poulami, D. Tibarewala, G. Vimalarani, *EEG Based Emotion Recognition System using MFDDFA as Feature Extractor*, International Conference on Robotics, Automation, Control and Embedded Systems – RACE 2015, Chennai, 2015.
- [27] S. Abe, *Support Vector Machines for Pattern Classification*, Springer-Verlag Ltd., London, 2010.
- [28] Y. Hayek Kobeissi, *Multifractal Financial Markets - An Alternative Approach to Asset and Risk Management*, Springer, NY, 2013.
- [29] Q. Li, W. Xie, *Target Classification with Low-Resolution Surveillance Radars Based on Multifractal Features*, Progress In Electromagnetics Research B, Vol. 45, 291–308, 2012.
- [30] [http : //pl.wikipedia.org/wiki/Warset](http://pl.wikipedia.org/wiki/Warset), stan na 2015-01-20.
- [31] T. Świderek, Gra o milisekundy, [https : //www.obserwatorfinansowy.pl/tematyka/rynki – finansowe/grao – milisekundy/](https://www.obserwatorfinansowy.pl/tematyka/rynki-finansowe/grao-milisekundy/), stan na 2015-01-20.
- [32] http://www.gpw.pl/pub/files/PDF/foldery/UTP_Warset.pdf, stan na 2015-01-20.
- [33] M. Meerman, *Backlight: Money and Speed, inside the black box*, VPRO, 2013, <https://www.youtube.com/watch?v=6DsKfR3hAfo>, stan na 2015-01-20.
- [34] Project Express, <http://submarinenetworks.com/systems/transatlantic/project-express>, stan na 2015-01-20.
- [35] [http : //www.gpw.pl/kolokacja](http://www.gpw.pl/kolokacja).
- [36] [http : //static.gpw.pl/pub/files/PDF/uslugi/new/Opis_uslugi_Market_Data_Gateway.pdf](http://static.gpw.pl/pub/files/PDF/uslugi/new/Opis_uslugi_Market_Data_Gateway.pdf), stan na 2015-04-01.
- [37] [http : //static.gpw.pl/pub/files/PDF/uslugi/Opis_uslugi_RMA.pdf](http://static.gpw.pl/pub/files/PDF/uslugi/Opis_uslugi_RMA.pdf), stan na 2015-04-01.

- [38] M. Dacorogna, R. Gencay, U. Muller, R. Olsen, *An Introduction to High-Frequency Finance* Academic Press, San Diego, 2001.
- [39] F. Reilly, K. Brown, *Investment Analysis and Portfolio Management*, South-Western Cengage Learning, Mason, 2011.
- [40] Friederich, S Payne, R 2011, *Computer based trading, liquidity and trading costs*, Government Office for Science, 2011, [https : //www.gov.uk/government/uploads/system/uploads/attachment_data /file/289011/11 – 1240 – dr5 – computer – based – trading – liquidity – and – trading – costs.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/289011/11-1240-dr5-computer-based-trading-liquidity-and-trading-costs.pdf), stan na 2015-07-01.
- [41] J.Huang, J. Lu, C. Ling, *Comparing Naive Bayes, Decision Trees, and SVM with AUC and Accuracy*, The Third IEEE International Conference on Data Mining, 2003, [http : //citeseerx.ist.psu.edu/viewdoc/download?doi = 10.1.1.141.751rep = rep1type = pdf](http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.141.751rep=rep1type=pdf), stan na 2015-07-01.
- [42] R. Siedlecki, *Finansowe sygnały ostrzegawcze w cyklu życia przedsiębiorstwa*, C.H. Beck, Warszawa, 2007.
- [43] T. Stasiewski, *Z-Score – indeks przewidywalnego upadku przedsiębiorstwa*, *Rachunkowość* nr 12/96, 1996.
- [44] B. Mandelbrot, *Self-affinity and fractal dimension*, *Physica Scripta*: 32, 257-260, 1985.
- [45] N. Christakis, J. Fowler, *W sieci. Jak sieci społeczne kształtują nasze życie*, Smak Słowa ,Sopot, 2011.
- [46] K. Flaga-Gieruszyńska, *Prawo upadłościowe i naprawcze*, Wydawnictwo C.H. Beck, Warszawa, 2009.
- [47] K. Zimmewicz, *Nauka o organizacji i zarządzaniu*, PWN, Warszawa-Poznań, 1990.
- [48] B. Prusak, *Uwarunkowania upadłości przedsiębiorstw*, [http : //www.zie.pg.gda.pl/ pb/wup.pdf](http://www.zie.pg.gda.pl/pb/wup.pdf), stan na 2015-07-01.
- [49] M. Odom, R. Sharda, *A Neutral Network Model for Bankruptcy Prediction*, materiały konferencyjne, San Diego, 1990.

- [50] J. Klobucnik, S. Sievers, *Bankruptcy prediction based on stochastic processes: a new model class to predict corporate bankruptcies?*, [http : //arw – suisse.ch/wp – content/uploads/2013/05/BPSP.pdf](http://arw-suisse.ch/wp-content/uploads/2013/05/BPSP.pdf), stan na 2015-07-01.
- [51] R. Kutner, *Wprowadzenie do Ekonofizyki: niegaussowskie procesy stochastyczne oraz niedebey'owska relaksacja w realu. Elementy teorii ryzyka rynkowego wraz z elementami teorii zdarzeń ekstremalnych*, [https : //www.fuw.edu.pl/tl_files/studia/materialy/ef/bloweier.pdf](https://www.fuw.edu.pl/tl_files/studia/materialy/ef/bloweier.pdf), 2014, stan na 2015-07-01.
- [52] L. Alves, [http : //www.intechopen.com/books/applied – fracture – mechanics/foundations – of – measurement – fractal – theory – for – the – fracture – mechanics](http://www.intechopen.com/books/applied-fracture-mechanics/foundations-of-measurement-fractal-theory-for-the-fracture-mechanics), stan na 2015-07-01.
- [53] A. Barabasi, E. Stanley, *Fractal Concepts in Surface Growth*, Cambridge University Press, 2002.
- [54] P. Addison, *Fractals and Chaos An Illustrated Course*, IOP Publishing Ltd, London, 1997.
- [55] R. Narang, *Inside the Black Box - The Simplest Truth About Quantitative Trading*, John Wiley Sons Inc., Hoboken, 2009.
- [56] D. Lee, C. You, J. Jung, K. Chang, S. Kim, K. Kim, *Multifractal Structures in Temperature and Humidity*, Journal of the Korean Physical Society, Vol. 57, No. 2, August 2010, pp. 296299.
- [57] [ttp : //cran.r – project.org/web/packages/e1071/vignettes/svmdoc.pdf](http://cran.r-project.org/web/packages/e1071/vignettes/svmdoc.pdf), stan na 2015-07-01.
- [58] [http : //blogs.wsj.com/moneybeat/2013/05/06/flash – crash – three – years – later – what – have – we – learned/](http://blogs.wsj.com/moneybeat/2013/05/06/flash-crash-three-years-later-what-have-we-learned/), stan na 2015-07-01.
- [59] [https : //onlinecourses.science.psu.edu/stat557/sites/onlinecourses.science.psu.edu.stat557/files/lesson04/image_08.gif](https://onlinecourses.science.psu.edu/stat557/sites/onlinecourses.science.psu.edu/stat557/files/lesson04/image_08.gif), stan na 2015-07-01.
- [60] [http : //courses.ee.sun.ac.za/Pattern_Recognition13/lectures/lecture01/node6.html](http://courses.ee.sun.ac.za/Pattern_Recognition13/lectures/lecture01/node6.html), stan na 2015-07-01.
- [61] [http : //www.gpwin.fostrefa.pl](http://www.gpwin.fostrefa.pl) , stan na 2015-07-01.

- [62] <http://wiadomosci.stockwatch.pl/kompendium-wskaznika-altmana-i-ratingu-kredytowego-w-stockwatch-pl,ratingi-i-upadlosci-spolek-gieldowych>, 37570, stan na 2015-07-01.
- [63] <http://www.gpwinfostrefa.pl/>, stan na 2015-07-01.
- [64] <http://pwp.pb.pl/2525966,5761,optimus-mial-szanse-byc-globalnym-mocarzem>, stan na 2015-08-25.
- [65] https://pl.wikipedia.org/wiki/CD_ProjektOptimus_SA_.281988.E2.80.932011.29, stan na 2015-08-25.
- [66] http://wyborcza.biz/biznes/1,101558,15684120,CD_Projekt_walczy_w_sadzie_z_Empikiem_o_10 mln_zl.html, stan na 2015-08-25.
- [67] <http://biznes.pl/gielda/wiadomosci/strata-netto-cd-projekt-to-2-92-mln-zl-w-iii-kw-2014-wobec-zysku-rok-wczesniej/qggr5>, stan na 2015-08-25.