



Uniwersytet Ekonomiczny
we Wrocławiu

Ekonomia

Hubert Gładysz

Nr albumu 182908

PRACA DYPLOMOWA LICENCJACKA

**„Wpływ uczenia maszynowego i sztucznej
inteligencji na rozwój rynku usług bankowości
inwestycyjnej”**

WROCŁAW 2023

Spis treści

Wstęp	3
Rozdział 1: Historia zastosowania sztucznej inteligencji i uczenia maszynowego w finansach.....	5
1.1 Komputerowe modele matematyczne	5
1.2 Zastosowanie algorytmów w analizie danych.....	8
1.3 Zaawansowane systemy eksperckie oraz początek handlu algorytmicznego	10
1.4 Wykorzystanie modeli uczenia maszynowego w handlu algorytmicznym	13
Rozdział 2: Machine Learning Pipeline.....	17
2.1 Pozyskiwanie danych.....	17
2.2 Czyszczenie danych	17
2.3 Przygotowanie danych	18
2.4 Podział na zestawy danych.....	20
2.5 Trenowanie modelu	21
2.6 Testowanie, wdrażanie i monitorowanie modelu.....	24
2.7 Data Science Pipeline	25
Rozdział 3: Praktyczne zastosowanie algorytmów i matematyki finansowej do analizy rynku.....	27
3.1 Techniki analizy rynku	27
3.2 Modele statystyczne	28
3.3 Modele stochastyczne.....	29
3.4 Algorytm regresji liniowej	33
3.5 Deep Learning	35
3.6 Drzewa regresyjne oraz las losowy (Random Forest)	38
3.7 Reinforcement Learning.....	39
3.8 High Frequency Trading (HFT).....	40
Rozdział 4: Analiza SWOT i podsumowanie	44
4.1 Zalety.....	44
4.2 Wady	46
4.3 Szanse.....	46
4.4 Zagrożenia	47
Podsumowanie.....	49
Bibliografia	51
Spis rysunków.....	52

Wstęp

Wraz z dynamicznym rozwojem Internetu i technologii komputerowej, tradycyjne podejście do eksplorowania oraz analizy danych również ewoluuje. Predykcje analityków nie muszą już bazować jedynie na podejściu statystycznym, ponieważ można je łączyć z różnymi dodatkowymi technologiami przetwarzania informacji takimi jak uczenie maszynowe.

Inwestorzy od zawsze chcieli odkryć zasady funkcjonowania rynku, które mogły dać im przewagę nad konkurencją. Analitycy próbują stosować różne metody analizy inwestycyjnej i ilościowej analizy danych, aby przewidzieć trend rynkowy lub wyliczyć potencjalną cenę, przy której trend może się zmienić. Biorąc pod uwagę, że na wycenę instrumentów finansowych wpływa wiele czynników rynkowych, jak i pozarynkowych, trudno jest stworzyć dokładny model, który pozwoli chociaż częściowo przewidzieć zachowanie cen na giełdzie. Zanim komputery miały możliwość przechowywania dużej ilości danych w pamięci, było to wręcz niemożliwe. Obecnie uczenie maszynowe jako model predykcyjny jest coraz częściej stosowane w przewidywaniu cen instrumentów finansowych.

Wraz z zwiększaniem mocy obliczeniowej komputerów, coraz bardziej skomplikowane algorytmy sztucznej inteligencji mogą pozwolić na bardziej precyzyjne wyliczenie ceny instrumentu, ryzyka związanego z inwestycją lub optymalizacji portfela inwestycyjnego poprzez identyfikację najlepszych aktywów i realokację kapitału.

Celem niniejszej pracy jest przeprowadzenie szczegółowej analizy porównawczej pomiędzy metodami inwestycyjnymi zastosowanymi przez fundusze inwestycyjne, które wykorzystują sztuczną inteligencję (AI) oraz technologie uczenia maszynowego (ML), a ich tradycyjnymi odpowiednikami. Celem analizy jest zbadanie aktualnych strategii inwestycyjnych stosowanych przez fundusze hedgingowe, oraz identyfikacja najbardziej efektywnych. Analiza ta ma również na celu scharakteryzowanie potencjalnej ewolucji finansów inwestycyjnych, biorąc pod uwagę zmieniające się trendy i innowacje technologiczne. W szczególności, badanie skupi się na kwestii, czy klasyczne podejście do inwestowania gdzie decyzje są podejmowane na podstawie wiedzy ekspertów i intuicji pozostanie dominującym modelem, czy też środowisko finansowe skieruje się ku modelowi gdzie procesy decyzyjne są w całości sterowane przez zaawansowane algorytmy.

Praca ta obejmuje szczegółowe omówienie ewolucji sztucznej inteligencji oraz precyzyjny opis procesu działania algorytmów wykorzystywanych w finansach, pozwalając na

głębsze zrozumienie mechanizmów kształtujących współczesne metody inwestycyjne. Zawiera opis i szczegółową analizę strategii inwestycyjnych oraz metod analizy danych.

Rozdział 1: Historia zastosowania sztucznej inteligencji i uczenia maszynowego w finansach.

1.1 Komputerowe modele matematyczne

Początki dziedziny sztucznej inteligencji sięgają lat 40. i 50. XX wieku. Pierwsze maszyny AI to proste programy, które wykorzystywały reguły logiczne do rozwiązywania prostych, powtarzalnych problemów.

Jednym z pierwszych programów był „Logic Theorist” stworzony przez Allena Newella i Herberta Simona w 1955r. Program pomagał badaczom w udowodnieniu twierdzeń matematycznych korzystając z zaprogramowanego zestawu reguł i aksjomatów. Logic Theorist był napisany językiem IPL¹ (ang. Interactive Problem Language), czyli w pierwszym języku programowania wysokiego poziomu, który został opracowany w latach 60. XX wieku przez Johna McCarthy'ego. Jego celem było ułatwienie tworzenia programów do rozwiązywania różnorodnych problemów, w tym problemów matematycznych, fizycznych i inżynierskich. IPL cechował się prostotą i możliwością interaktywnego programowania, a jego składnia opierała się na języku naturalnym. Pomimo swojej wczesnej historii, IPL odegrał ważną rolę w rozwoju sztucznej inteligencji, mimo że nie jest już powszechnie stosowany. Na jego podstawie rozwinęły się inne programy i języki programowania takie jak APL, MATLAB i Mathematica.

Sztuczna inteligencja po raz pierwszy zaistniała w świecie naukowym w 1956 r. podczas konferencji na Uniwersytecie Dartmouth w Hanover w USA. Na sześciotygodniowej konferencji grupa badaczy z dziedziny informatyki, matematyki i filozofii rozważała, czy można zbudować maszynę, która będzie w stanie symulować myślenie i wnioskowanie podobne do ludzkiego. Na konferencji rozważano zastosowanie sztucznej inteligencji w wielu branżach np. planowanie, logistyka, przetwarzanie języka naturalnego. Wielu uczestników przewidywało, że maszyny osiągną poziom inteligencji równy z ludzkim w ciągu zaledwie kilku lat.

W konferencji udział wzięli m.in:

- John McCarthy – doktor matematyki, jest on autorem terminu „sztuczna inteligencja”, który sformułował na konferencji w Dartmouth. W roku 1960 stworzył język programowania Lisp,

¹ P. Ceruzzi (1998) “A History of Modern Computing”

który był aktualizowany i komercyjnie używany aż do roku 1994. W roku 1971 za wkład w dziedzinie sztucznej inteligencji został laureatem nagrody Turinga.

- Claude E. Shannon – matematyk i inżynier, jeden z pierwszych naukowców, którzy dostrzegli wartość kodu binarnego. W 1956 roku stworzył najpotężniejszy komputer swoich czasów, MANIAC 1 (Mathematical Analyzer, Numerator, Integrator, and Computer). Komputer rozegrał jedną z pierwszych partii szachowych między komputerem, a człowiekiem, wybierał najkorzystniejszą strategię na podstawie koncepcji teorii gier Minimax². C. Shannon zapoczątkował również teorię informacji która stanowi fundamentalne narzędzie w dziedzinie naukowych i inżynierskich dyscyplin zajmujących się przesyłaniem, przetwarzaniem oraz analizą informacji. Podstawowe założenia teorii informacji obejmują pojęcia takie jak entropia informacyjna, która jest miarą niepewności lub niespodziewanego charakteru informacji, oraz redundancja, która odnosi się do stopnia, w jakim informacja jest przewidywalna.

- Marvin Minsky – doktor matematyki, współpracował razem z Johnem McCarthy. W 1951 r. zbudował Snarc, czyli pierwszą w historii maszynę wykorzystującą sieci neuronowe³.

Sieć neuronowa to model obliczeniowy inspirowany strukturą i funkcjonowaniem ludzkiego mózgu. Składa się z warstw neuronów, które są elementami przetwarzającymi. Neurony są uproszczonymi modelami biologicznych neuronów i są odpowiedzialne za obliczenia i przekazywanie informacji. Celem takiego modelu jest uczenie się wzorców, rozpoznawanie i klasyfikacja danych wejściowych oraz przewidywanie wyników na podstawie nauczonych wzorców. Sieci neuronowe dzielą się na trzy główne warstwy: warstwę wejściową, warstwę wyjściową i jedną lub więcej warstw ukrytych.

1. Warstwa wejściowa: Odbiera dane wejściowe i przekazuje je do warstw ukrytych. Ilość neuronów w tej warstwie zwykle odpowiada wymiarowi danych wejściowych.

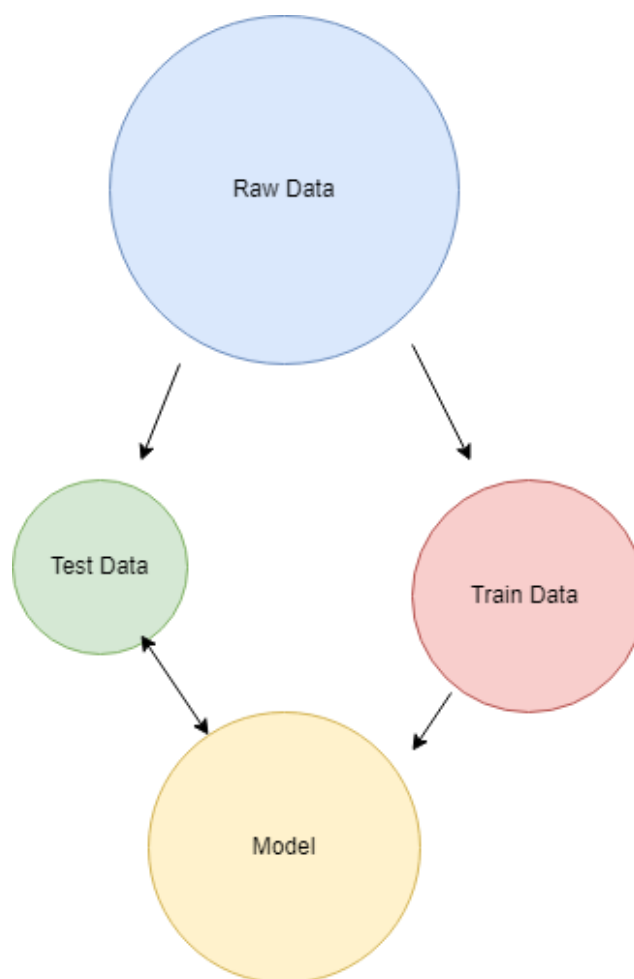
2. Warstwy ukryte: Są to warstwy pośrednie między warstwą wejściową, a wyjściową gdzie odbywa się faktyczne przetwarzanie i uczenie algorytmu. Neurony w tych warstwach są połączone z neuronami z sąsiednich warstw za pomocą wag, które są aktualizowane podczas procesu uczenia.

² Minimax to strategia podejmowania decyzji w grach o sumie zerowej z dwoma graczami.

³ Goodfellow, I., Bengio, Y., & Courville, A. (2016). "Deep Learning"

3. Warstwa wyjściowa: Generuje końcowy wynik sieci neuronowej, który może być interpretowany jako klasyfikacja, prognoza lub wartość numeryczna. Liczba neuronów w warstwie wyjściowej zależy od problemu, który próbujemy rozwiązać.

Proces uczenia sieci neuronowej polega na dostosowywaniu wag pomiędzy neuronami w oparciu o dane uczące⁴. Uczenie sieci neuronowej odbywa się za pomocą algorytmów takich jak wsteczna propagacja błędu (backpropagation), który pozwala na dostosowanie wag w celu zmniejszenia błędu między prognozami sieci a rzeczywistymi danymi. Wagi są aktualizowane stopniowo, wielokrotnie, aż osiągnięty zostanie ustalony poziom dokładności.



Rysunek 1. Uproszczona wizualizacja techniki szkolenia algorytmu,

Źródło: opracowanie własne

⁴ Dane uczące (ang. training data) to zbiór danych, na podstawie którego algorytm przeprowadza proces uczenia, zbiór ten musi być częścią danych bazowych które dzielimy na zbiór testowy i uczący - Aureliena Geron (2019) "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow"

Ilustracja przedstawiona powyżej reprezentuje zasadniczą strukturę działania algorytmu w kontekście uczenia nadzorowanego⁵. W modelu tym, inicjalny zbiór danych dzielony jest na dwie odrębne grupy: zestaw treningowy (przeważnie obejmujący około 70% danych) oraz zestaw testowy (zbierający zazwyczaj pozostałe 30%). Algorytm poddaje analizie zestaw danych treningowych i na tej podstawie dostosowuje swoje parametry - wagi - tak, aby zminimalizować błąd pomiędzy wartościami przewidywanymi przez model, a rzeczywistymi etykietami danych. Proces kwantyfikacji błędu może opierać się na różnych metrykach, przy czym najbardziej podstawowe z nich wykorzystują średnią arytmetyczną i odchylenie standardowe. Głównym celem powyższego procesu jest stworzenie modelu zdolnego do efektywnej predykcji na podstawie nowych, nieznanych mu wcześniej danych. W tym kontekście, kluczowe jest aby struktura i kategorie nowych danych odpowiadały tym, które wykorzystano w procesie trenowania modelu.

Sieci neuronowe mają wiele zastosowań, takich jak rozpoznawanie obrazów, analiza tekstu, tłumaczenie maszynowe, rozpoznawanie mowy, systemy rekomendacyjne i wiele innych. Są one podstawą dla wielu zaawansowanych technologii sztucznej inteligencji takich jak głębokie uczenie (deep learning) i sieci konwolucyjne (CNN).

1.2 Zastosowanie algorytmów w analizie danych

W latach 60. i 70. XX wieku, dzięki wzrostowi mocy obliczeniowej komputerów, zaczęto tworzyć bardziej zaawansowane systemy. W latach 60. badacze pracowali nad rozwojem systemów eksperckich, które mogły udzielać odpowiedzi na pytania ze skomplikowanych dziedzin, takich jak medycyna, finanse czy inżynieria. Systemy eksperckie, podobnie jak program "Logic Theorist", opierały się na zbiorze zapisanych reguł, na podstawie których wyciągały wnioski i podejmowały decyzje. W finansach systemy eksperckie miały jedynie pośrednie zastosowanie ze względu na charakter branży. Przykładem takiego systemu jest program "PROSPECTOR"- opracowany przez Daniela Bobrowa. Program ten analizował dane geologiczne i pomagał w poszukiwaniu złóż minerałów. Dzięki programowi inwestorzy mogli zastosować zupełnie nową strategię inwestycyjną i szacować wartość surowców na podstawie danych stworzonych przez algorytm.

Dopiero wykorzystanie sieci neuronowych pozwoliło na opracowanie systemu "STOCKS" do prognozowania zmian cen na rynku akcji. Program analizował dane i szukał

⁵ Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press

zależności pomiędzy różnymi czynnikami rynkowymi takimi jak zmienność rynku, wolumen, cena akcji, a nawet okres wypłaty dywidendy. Następnie przetwarzał dane wejściowe i każdej z danych nadawał odpowiednią wagę - im większa waga, tym większy wpływ czynnika na zmianę ceny. Mimo że "Stocks" był przełomowym programem wykorzystującym AI do analizy rynku, jego wyniki były jedynie niewiele lepsze od wyników uzyskiwanych przez tradycyjne modele statystyczne.

Modele statystyczne były jednak o wiele prostsze, przez co tańsze i łatwiejsze w obsłudze. Opierały się na bazowych miarach statystycznych, takich jak średnia geometryczna, najczęściej używana w analizie cen i wolumenu, odchyleniu standardowym i wariancji, które były traktowane jako współczynniki określające ryzyko, a współczynnik korelacji pomagał w dywersyfikacji portfela.

Mimo tego, że modele statystyczne wydawały się w tamtym momencie bardziej efektywnym rozwiązaniem, firmy wciąż wierzyły w rozwój AI. Program Stocks był ważnym krokiem do rozwoju sztucznej inteligencji w finansach, ponieważ pokazał, że maszyny mogą być wykorzystane do przewidywania zmian rynkowych na podstawie danych historycznych. Jedynym ograniczeniem dla naukowców jest moc obliczeniowa i ilość pamięci komputerów.

Kolejnym programem, który powstał w tym okresie był „BACON” (Bank of America Computerized Order eNtry” stworzony i używany przez Bank of America. Zasada działania systemu była bardzo zbliżona do Stocks, jednak Bacon dodatkowo analizował wyniki swoich poprzednich rekomendacji i na podstawie tych danych oceniał ryzyko rynkowe i uwzględniał je w swoich kolejnych rekomendacjach. W przeciwieństwie do poprzedników Bacon okazał się bardzo skutecznym narzędziem dla Bank of America i przyczynił się do poprawy wyników inwestycyjnych, a zatem zysków banku.

System Bacon był pionierski jeśli chodzi o zastosowane rozwiązania logiczne. W przeciwieństwie do poprzedników zostały w nim wykorzystane elementy logiki rozmytej zamiast klasycznej logiki dwuwartościowej. Rozwiązanie to było stałym punktem wykorzystywanym w budowie kolejnych systemów i do dziś stanowi podstawę systemów inteligentnych. Logika rozmyta⁶ różni się od klasycznej liczbą wartości. W dwuwartościowej wyróżniamy tylko prawdę (1) lub fałsz (0), w logice rozmytej natomiast pojawia się szereg wartości pośrednich zwanych również stopniami prawdy, które mogą przyjmować dowolne

⁶ Nguyen, H. T., & Walker, E. A. (2006). A First Course in Fuzzy Logic

wartości w zakresie od 0 do 1. Przed zastosowaniem w sieciach neuronowych logika rozmyta miała zastosowanie w inżynierii i mechanice.

Sztuczna inteligencja znajduje zastosowanie w finansach i to nie tylko w zakresie inwestycji, ale również w bankowości komercyjnej. "LIMID" to jeden z pierwszych programów sztucznej inteligencji, który wykorzystywał algorytmy uczenia maszynowego i przetwarzania języka naturalnego w celu analizy dokumentów finansowych i wyliczania zdolności kredytowej klientów. LIMID był w stanie ustalić zdolność kredytową na podstawie formularzy podatkowych i wyciągów bankowych, analizował wysokość dochodów, historię kredytową, wydatki i inne czynniki związane z sytuacją finansową klienta. Program był dokładniejszy i szybszy w ocenie zdolności kredytowej niż człowiek, wyeliminował błąd ludzki, a decyzje były podejmowane jedynie na podstawie faktów.

1.3 Zaawansowane systemy eksperckie oraz początek handlu algorytmicznego

W latach 80. i 90. nastąpiła zmiana podejścia do sztucznej inteligencji, zamiast próbować stworzyć systemy o pełnej świadomości i myśleniu analogicznym do ludzkiego, skoncentrowano się na tworzeniu systemów o węższym zakresie funkcjonalnym.

W 1983 roku firma Intellicorp stworzyła program KEE (Knowledge Engineering Environment). Program ten służył do budowy baz danych (knowledge bases), które pozwalały przechować niespotykaną wcześniej ilość danych i reguł. Dzięki temu systemy opracowane przy użyciu KEE mogły analizować więcej danych, przez co były znacznie dokładniejsze niż poprzednie systemy.

W późniejszych latach KEE stał się inspiracją do stworzenia takich narzędzi jak CLIPS (C Language Integrated Production System) czy JESS (Java Expert System Shell). CLIPS i JESS to jedne z najbardziej popularnych języków programowania używanych obecnie. Narzędzia te służą do tworzenia zaawansowanych systemów na podstawie reguł, funkcji, definicji i konkretnych warunków, które mają zostać wykonane przez komputer. Zarówno CLIPS, jak i JESS, znajdują obecnie zastosowanie w wielu dziedzinach związanych z zarządzaniem i analizą danych.

W 1994 roku firma Mead Data Central opracowała LexisNexis - system do przetwarzania języka naturalnego (NLP - Neuro-Linguistic Programming). Początkowo dane były używane jedynie przez prawników, ponieważ dokumentacja generowana przez program ułatwiała pracę. LexisNexis zbierał i przechowywał duże ilości informacji tekstowych takich

jak artykuły prasowe, publikacje naukowe, raporty finansowe, sprawozdania rządowe, a po powstaniu Internetu nawet wpisy na blogach i forach.

Po pobraniu danych, system filtrował i indeksował tekst, który później był udostępniany dla użytkowników. Użytkownik miał możliwość wyszukiwania informacji na podstawie słów kluczowych, fraz i innych kryteriów, a także tworzenia niestandardowych raportów i analiz. Analizy tworzone przez program szybko znalazły zastosowanie jako dodatkowe narzędzie służące do badania sentymentu rynkowego. Dzięki zastosowaniu narzędzi NLP i analizy sentymentu, LexisNexis mógł automatycznie analizować tysiące artykułów i publikacji na temat konkretnej firmy, branży lub rynku, aby określić, jakie trendy i wydarzenia wpływają na ceny akcji lub kursy walut. System mógł również wykryć pozytywny lub negatywny sentyment w artykułach na temat konkretnych firm lub rynków, co pomagało inwestorom w podejmowaniu decyzji inwestycyjnych.

Po komercjalizacji Internetu w 1990 roku, zaczęły powstawać systemy do analizy informacji w mediach, np. WiseWindow, NewsEdge, Inxight, ClearForest. Systemy te działały na tej samej zasadzie co LexisNexis, jednak miały dużo większą pamięć, przez co potrafiły gromadzić więcej danych i były szybsze ze względu na większą moc obliczeniową.

Po udostępnieniu Internetu dla szerokiego grona odbiorców, analiza trendów stała się bardzo popularna i wszystkie firmy inwestycyjne korzystały z usług firm świadczących takie usługi, lub opracowywały własne systemy. Aktualnie, systemy NLP są bardzo popularne w wielu branżach, np. w finansach, branży marketingowej. AI stale analizuje dane z raportów finansowych, przekazów z konferencji prasowych, telewizji, mediów społecznościowych, historii przeglądania użytkownika, aby wygenerować raport, lub w czasie rzeczywistym dopasować treści do odbiorcy.

W latach 90. Rozwój technologiczny i komercjalizacja Internetu pozwoliły na wprowadzenie handlu algorytmicznego⁷ na rynkach finansowych. Algo trading to zastosowanie zaawansowanych matematycznych modeli i algorytmów komputerowych do automatyzacji procesów inwestycyjnych na rynkach finansowych. Handel algorytmiczny ma na celu zautomatyzowanie decyzji inwestycyjnych, eliminując potrzebę ręcznego wprowadzania zleceń przez inwestorów.

⁷ Chlistalla, M. (2011). Algorithmic Trading. Deutsche Bank Research

Oprócz lepszej analizy rynku algorytmy pozwoliły na poprawienie jakości handlu na giełdzie. Do największych zalet należą:

- a) Poprawa jakości transakcji: Inwestor mógł bardziej precyzyjnie określić różne aspekty transakcji takie jak: czas transakcji, cena, wolumen, zastosowanie zlecenia stop loss⁸.
- b) Redukcja kosztów transakcyjnych: Handel algorytmiczny pozwala na automatyzację procesu inwestycyjnego, co prowadzi do zmniejszenia kosztów związanych z realizacją transakcji. Koszty te obejmują prowizje brokerskie, podatki oraz spread (różnica między ceną kupna, a sprzedaży). Algorytmy mogą zoptymalizować strategie handlowe, aby zminimalizować te koszty, co przekłada się na większą efektywność inwestycji.
- c) Zwiększenie szybkości transakcji: Algorytmy działają znacznie szybciej niż ludzie, co pozwala na natychmiastowe reagowanie na zmieniające się warunki rynkowe i wykorzystanie krótkotrwałych okazji inwestycyjnych. W handlu algorytmicznym, zlecenia są składane i realizowane w ułamkach sekundy, co pozwala na lepsze wykorzystanie informacji i szybsze adaptowanie się do zmieniających się sytuacji na rynku. Możliwość generowania dużych ilości transakcji w bardzo szybkim czasie pozwoliła na powstanie nowych strategii inwestycyjnych, które polegały na wykorzystaniu zaawansowanych modeli statystycznych i matematycznych do identyfikacji niewielkich, krótkotrwałych anomalii cenowych na rynku. Wspomniana strategia została nazwana handlem kwantowym (Tłumaczenie z Quantitative Trading, handel kwantowy jest również nazywany kwantytatywnym lub ilościowym)
- d) Minimalizowanie czynnika ludzkiego. Ludzkie emocje, takie jak strach czy chciwość, mogą wpłynąć na decyzje inwestycyjne i prowadzić do nieoptymalnych wyników. Algorytmy są pozbawione emocji i podejmują decyzje oparte wyłącznie na analizie danych, co eliminuje wpływ emocji na proces inwestycyjny.

Jim Simons⁹, założyciel funduszu hedgingowego Renaissance Technologies, jest znanym zwolennikiem handlu algorytmicznego i kwantowego. Simons jest słynny ze swojego stanowiska – uważa on, że emocje i subiektywne interpretacje danych często prowadzą do błędów w procesie inwestycyjnym. Jego filozofia polega na zredukowaniu, a najlepiej

⁸ Stop loss to rodzaj zlecenia handlowego stosowanego przez inwestorów na rynkach finansowych, mające na celu ograniczenie potencjalnych strat. Zlecenie stop loss jest aktywowane, gdy cena instrumentu finansowego osiąga określony poziom. - K. Jajuga, T. Jajuga (2020) Inwestycje

⁹ Gregory Zuckerman (2019) "Człowiek, który rozwiązał rynek: Jak Jim Simons zapoczątkował rewolucję quantów"

wyeliminowaniu, roli człowieka w podejmowaniu decyzji inwestycyjnych, zastępując go algorytmami, które podejmują decyzje oparte na twardych danych i obiektywnej analizie.

Dla osiągnięcia tego celu, Simons zebrał zespół naukowców – matematyków, fizyków i informatyków – którzy stworzyli i wdrożyli algorytmy do automatycznego handlu. Zamiast polegać na intuicji czy odczuciach, algorytmy funduszu Renaissance Technologies analizowały ogromne ilości danych, szukając wzorców i trendów, które mogłyby wskazywać na potencjalne okazje inwestycyjne. Inne firmy, które obecnie wykorzystują podobne techniki to Two Sigma, DE Shaw & Co, Citadel i AQR Capital Management.

1.4 Wykorzystanie modeli ucznia maszynowego w handlu algorytmicznym

Z początkiem nowego stulecia, rosnąca ilość danych dostępnych do analizy stała się kluczowym czynnikiem wpływającym na rozwój uczenia maszynowego i sztucznej inteligencji w bankowości inwestycyjnej. Dane te, często nazywane "Big Data"¹⁰, termin używany do opisanie ogromnej ilości danych, które są zbyt duże lub zbyt złożone, aby można je było przetwarzać za pomocą tradycyjnych metod. Big Data obejmuje zarówno dane strukturalne, takie jak liczby i daty, jak i dane niestukturalne, takie jak tekst, obrazy i dźwięk.

Jednym z kluczowych aspektów Big Data jest to, że zawierają one wiele różnych typów informacji, które gdy są ze sobą powiązane, mogą dostarczyć głębokich i wartościowych wglądów. Na przykład - analiza sentymentu z mediów społecznościowych może być skorelowana z danymi o kursach akcji, aby przewidzieć przyszłe ruchy cen. Warto również dodać, że w latach 2000 zaczęto używać zaawansowanych technik uczenia maszynowego takich jak głębokie sieci neuronowe. Głębokość¹¹ sieci neuronowej odnosi się do liczby warstw ukrytych pomiędzy warstwą wejściową, a wyjściową w sieci. Każda warstwa w sieci neuronowej składa się z zestawu neuronów, które przetwarzają informacje od poprzedniej warstwy i przekazują je do następnej.

W odpowiedzi na kryzys finansowy z 2008 roku, banki inwestycyjne i instytucje finansowe na całym świecie zaczęły inwestować w rozwój zaawansowanych ilościowych systemów wczesnego ostrzegania (QEWAS).

¹⁰ Bholat, D., Brookes, J., Cai, C., Grundy, K., & Lund, J. (2020). Machine learning explainability in finance: an application to default risk analysis. Bank of England working papers.

¹¹ Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

Nowe metody, które zostały wdrożone, obejmują techniki uczenia maszynowego takie jak: sieci neuronowe, algorytmy drzew decyzyjnych¹², maszyny wektorów nośnych¹³, a także bardziej tradycyjne metody statystyczne takie jak analiza regresji¹⁴. Drzewo decyzyjne to rodzaj modelu predykcyjnego. Każdy węzeł w drzewie reprezentuje cechę (lub atrybut), każda gałąź reprezentuje regułę decyzyjną, a każdy liść reprezentuje wynik. Główną zaletą drzew decyzyjnych jest ich prostota i łatwość interpretacji. Maszyny Wektorów Nośnych (Support Vector Machines, SVM): to model uczenia maszynowego używany głównie w problemach klasyfikacji i regresji. SVM jest najbardziej skuteczny, gdy wymiarowość¹⁵ danych jest wysoka, a dane są liniowo separowalne¹⁶. W kontekście uczenia maszynowego i analizy danych, wysoka wymiarowość może prowadzić do problemów, takich jak przetrenowanie modelu i trudności w interpretacji wyników. SVM nie reguluje wymiarowości natomiast znajduje hiperpłaszczyznę w przestrzeni wielowymiarowej, która najlepiej oddziela różne klasy danych.

Ilościowe systemy wczesnego ostrzegania (QEWAS)¹⁷ działają poprzez monitorowanie szeregów wskaźników finansowych i makroekonomicznych takich jak: stopy procentowe, wskaźniki inflacji, poziomy długu, ceny akcji i wiele innych. Gdy te wskaźniki osiągną określone poziomy lub zaczną wykazywać pewne wzorce, systemy mają za zadanie wysłać ostrzeżenie, wskazując na potencjalne zagrożenie dla stabilności finansowej.

Obecnie wiele firm korzysta z usług firm specjalizujących się w systemach QEWAS. Te firmy takie jak RiskMetrics Group, MSCI, Moody's Analytics i wielu innych, oferują zaawansowane narzędzia do monitorowania i zarządzania ryzykiem dla banków inwestycyjnych i innych instytucji finansowych.

¹² Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3), 660-674

¹³ Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.

¹⁴ Analiza regresji to metoda statystyczna, która pozwala zrozumieć, jak wartość zależnej zmiennej zmienia się w zależności od wartości jednej lub więcej zmiennych niezależnych - James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*

¹⁵ Wymiarowość danych odnosi się do liczby atrybutów lub cech, które opisują każdy obiekt w zestawie danych. Każdy atrybut reprezentuje jeden wymiar w przestrzeni danych - Bellman, R. (2015). *Adaptive control processes: a guided tour*. Princeton university press.

¹⁶ Liniowo separowalne" to termin używany w kontekście klasyfikacji w uczeniu maszynowym, który odnosi się do możliwości oddzielenia dwóch lub więcej klas danych. Na przykład zbiór danych z dwiema klasami, które można rozdzielić tak, że wszystkie punkty z jednej klasy należą do jednej kategorii a wszystkie punkty z drugiej klasy należą do drugiej.

¹⁷ Fendoglu, S. (2017). Credit cycles and capital flows: Effectiveness of the macroprudential policy framework in emerging market economies. *Journal of Banking & Finance*

Dane są podstawą działania każdego dobrego algorytmu. Wraz z rozwojem nowych technologii i nagłego wzrostu ilości generowanych danych zaczęto wprowadzać nowe rozwiązania do bezpiecznego przechowywania informacji.

Jedną z takich technologii jest - blockchain¹⁸, czyli podstawa kryptowalut. Blockchain został wprowadzona po raz pierwszy w 2009 roku przez anonimowego twórcę o pseudonimie Satoshi Nakamoto. Od tamtej pory technologia przeszła długą drogę i stała się fundamentalnym elementem dla wielu innowacyjnych projektów, zwłaszcza w sektorze finansowym. Blockchain umożliwia bezpieczne przechowywanie i udostępnianie danych, co jest kluczowe dla procesów uczenia maszynowego. Ponadto, blockchain pozwala na tworzenie zdecentralizowanych sieci, gdzie różne podmioty mogą współpracować i dzielić się zasobami bez konieczności polegania na jednym centralnym organie.

Podobną do blockchain technologią jest zapisywanie w chmurze¹⁹, obie technologie mają swoje unikalne zalety. Istnieją jednak aspekty które mogą sprawić, że zapisywanie w blockchain jest lepsze przy pracy nad niektórymi projektami. Do zalet blockchain należą:

1. Niezmiennność: W przeciwieństwie do chmur, gdzie dane mogą być modyfikowane lub usuwane, dane zapisane na blockchainie są niezmiennie. Raz zapisane, nie mogą być zmienione ani usunięte. To zapewnia wysoki poziom bezpieczeństwa i sprawia, że blockchain jest idealny do przechowywania ważnych danych, które nie mogą ulec zmianie.
2. Bezpieczeństwo: Blockchain korzysta z zaawansowanych technik kryptografii, co zapewnia wysoki poziom bezpieczeństwa. Ponieważ jest to sieć rozproszona, nie ma pojedynczego punktu awarii, co oznacza, że jest trudniejsza do zhakowania niż tradycyjne chmury.
3. Zaufanie: W przypadku blockchajna, nie jest potrzebne zaufanie do trzeciej strony (takiej jak dostawca usług chmurowych) do przechowywania i zarządzania danymi.

¹⁸ Tapscott, D., & Tapscott, A. (2016). "Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World"

Mougayar, W. (2016). "The Business Blockchain: Promise, Practice, and Application of the Next Internet Technology"

¹⁹ Chmura (cloud computing) to model dostarczania usług IT za pośrednictwem Internetu. W kontekście przechowywania danych, chmura umożliwia użytkownikom zapisywanie i pobieranie danych z serwerów zdalnych zarządzanych przez dostawcę usług chmurowych, zamiast przechowywać te dane na lokalnych serwerach lub na osobistych komputerach. - "What is cloud storage?" - Microsoft Azure

Wszystkie transakcje są w pełni transparentne i można je sprawdzić niezależnie, co przyczynia się do budowania zaufania.

Blockchain to stosunkowo nowa technologia i wszystkie jej zastosowania nie są jeszcze znane. Wiadomo jednak, że ze względu na powyższe cechy zaufanie do usług AI ciągle rośnie, a to przyczynia się do dalszego rozwoju tej technologii. Istnieją już firmy które wykorzystują blockchain do przechowywania informacji na przykład firmy SingularityNET i Ocean Protocol. Są to platformy które umożliwiają udostępnianie i monetyzację usług opartych na AI. Użytkownicy mogą publikować zestawy danych, które są później używane do trenowania algorytmów ML.

Rozdział 2: Machine Learning Pipeline

Procesy ML i Data Science Pipeline składają się na niezbędne podstawy tworzenia skutecznych i dokładnych modeli, które są kluczowe dla procesu podejmowania decyzji inwestycyjnych. Przedstawiony tekst oferuje dogłębne zrozumienie tych procesów. Machine Learning Pipeline stanowi podstawowy schemat, a koncept Data Science Pipeline jest rozszerzeniem tego schematu, uwzględnia nie tylko budowę modelu, ale również analizę, wizualizację i interpretację wyników.

2.1 Pozyskiwanie danych

To proces zbierania danych z różnych źródeł, które mogą obejmować bazy danych za pomocą SQL²⁰, pliki w różnych formatach między innymi CSV²¹, JSON²², Parquet, HDF5²³ i strumienie danych w czasie rzeczywistym²⁴.

Szczególną uwagę należy zwrócić na strumienie danych w czasie rzeczywistym, ponieważ ten format pozwala na ciągłe przesyłanie danych do modelu, bez konieczności uczenia go na danych statycznych. Dzięki temu dane nie muszą być gromadzone i przechowywane, bo model analizuje je na bieżąco. Taki format jest najczęściej wykorzystywany przy modelach działających na danych giełdowych. Strumienie danych w czasie rzeczywistym są obsługiwane przez specjalistyczne technologie, takie jak Apache Kafka, Amazon Kinesis czy Google Cloud Pub/Sub. Po zebraniu danych, należy je przekształcić na dowolny, ale taki sam format dla całego zbioru.

2.2 Czyszczenie danych

Dane często zawierają błędy, nieprawidłowości, brakujące wartości lub są w różnych formatach. Czyszczenie danych to proces usunięcia tych problemów, aby dane były spójne i gotowe do dalszego przetwarzania. Na tym etapie szczególną uwagę należy zwrócić na uzupełnienie brakujących danych, w wielu przypadkach dane mogą być niekompletne i będą musiały zostać uzupełnione przez analityków. Niewłaściwie uzupełnienie danych np. wypełnienie brakujących wierszy zerami może prowadzić do manipulacji algorytmu. Najprostszym rozwiązaniem jest uzupełnienie brakujących danych średnią, medianą lub wartością dominującą w danym zbiorze.

Podczas gdy uzupełnianie brakujących wartości na podstawie parametrów statystycznych, może okazać się skuteczna dla danych o rozkładzie normalnym. Nie działa w bardziej skomplikowanych zbiorach np. szeregów czasowych, w takich sytuacjach stosuje się interpolację, która polega na estymowaniu brakujących wartości, na podstawie wartości sąsiednich. Metody interpolacji mogą obejmować interpolację liniową, wielomianową lub

²⁰ SQL (Structured Query Language): Język zapytań służący do ekstraktowania danych z relacyjnych baz danych.

²¹ CSV (Comma Separated Values): Bardzo prosty i popularny format danych, służy do przechowywania małej ilości informacji w formie tekstu podzielonego na tabele.

²² JSON (JavaScript Object Notation): Jest to format danych, który jest często używany w kontekście danych strukturyzowanych i semi-strukturyzowanych. JSON jest bardziej elastyczny niż CSV, ponieważ pozwala na zagnieżdżanie struktur danych i obsługuje większą różnorodność typów danych.

²³ Parquet, HDF5: to formaty plików kolumnowych, które są zoptymalizowane do pracy z dużymi zbiorami danych. Są szczególnie efektywne ponieważ pozwalają na szybkie odczytywanie tylko wybranych kolumn.

²⁴ Strumień danych w czasie rzeczywistym odnosi się do procesu ciągłego gromadzenia i przesyłania danych, które są natychmiast przetwarzane i analizowane

szeregów harmonicznych, w zależności od charakterystyki danych. Innym podejściem, które również wykorzystuje relacje pomiędzy najbliższymi wartościami, jest metoda k-najbliższych sąsiadów (k-nearest neighbors imputation)²⁵. W tej technice, wszystkie wartości są grupowane na podstawie określonych parametrów statystycznych. Każde brakujące wartości są wtedy zastępowane wartością, która jest najbliższa, lub najbardziej reprezentatywna dla grupy, do której należy. W efekcie, uzupełnione wartości są zgodne z dominującą strukturą danych w ich bezpośrednim otoczeniu.

Rozszerzeniem statycznego uzupełniania danych jest modelowanie statystyczne, w tej technice wykorzystuje się modele regresji aby znaleźć ukryte wzorce strukturalne w zbiorze danych. Przy pomocy takich modeli możliwe jest obliczenie brakujących wartości na podstawie wartości innych, skorelowanych zmiennych.

2.3 Przygotowanie danych

Surowe dane, które są zbierane z różnych źródeł, rzadko są idealnie dostosowane do bezpośredniego użycia w modelach uczenia maszynowego. Przygotowanie danych to szereg technik i procedur mających na celu przekształcenie surowych danych wejściowych w formę, która jest optymalna dla konkretnego algorytmu lub modelu. Podstawową techniką jest kodowanie danych katagorycznych²⁶, czyli danych które najczęściej są nie liczbowe, nie mają żadnej kategorii, ani skali. Przykłady obejmują kolory (czerwony, zielony, niebieski) lub płeć (męska, żeńska). Wiele algorytmów uczenia maszynowego wymaga jednak numerycznej reprezentacji danych, dlatego dane katagoryczne muszą być zakodowane do numerycznej formy. Najpopularniejsza technika to kodowanie porządkowe (ordinal encoding) w której każdej kategorii przypisuje się unikalną liczbę.

Następna technika konieczną dla optymalnego uczenia algorytmu jest normalizacja²⁷ czyli proces przekształcania wartości numerycznych cech w zestawie danych na wspólną skalę, bez utraty informacji. Wspólna skala przyspiesza i zwiększa efektywność uczenia algorytmu. Wyróżniamy dwie najbardziej popularne metody normalizacji. Pierwsza to skalowanie mini-max czyli przeskalowanie wartości na liczby w zakresie od 0 do 1 w praktyce wykonuje się to za pomocą wzoru:

$$n' = \frac{n - n.min}{n.max - n.min}$$

Gdzie:

n' to nowa, znormalizowana wartość

n to oryginalna wartość obserwacji

n.min to minimalna wartość w zestawie danych

n.max to maksymalna wartość w zestawie danych

²⁵ Kevin P. Murphy (2020) „Machine Learning: A Probabilistic Perspective”

²⁶ Max Kuhn, Kjell Johnson (2013) „Applied Predictive Modeling”

²⁷ Normalizacja to proces przekształcania wartości numerycznych w zestawie danych na wspólną skalę, bez zakłócania różnic w zakresach wartości lub utraty informacji.

Kolejną metodą normalizacji jest standaryzacja (z-score normalization). W tej metodzie wartości są przeskalowane aby średnia w zbiorze danych była równa zero a odchylenie standardowe 1. Po standaryzacji obserwacje będą się mieścić w zakresie od -3 do 3 zgodnie z regułą rozkładu normalnego, czyli 68% w zakresie -1:1, 95% w zakresie -2:2, 99,7% w zakresie -3:3. Standaryzowaną wartość wylicza się ze wzoru:

$$n' = \frac{n - \mu}{\sigma}$$

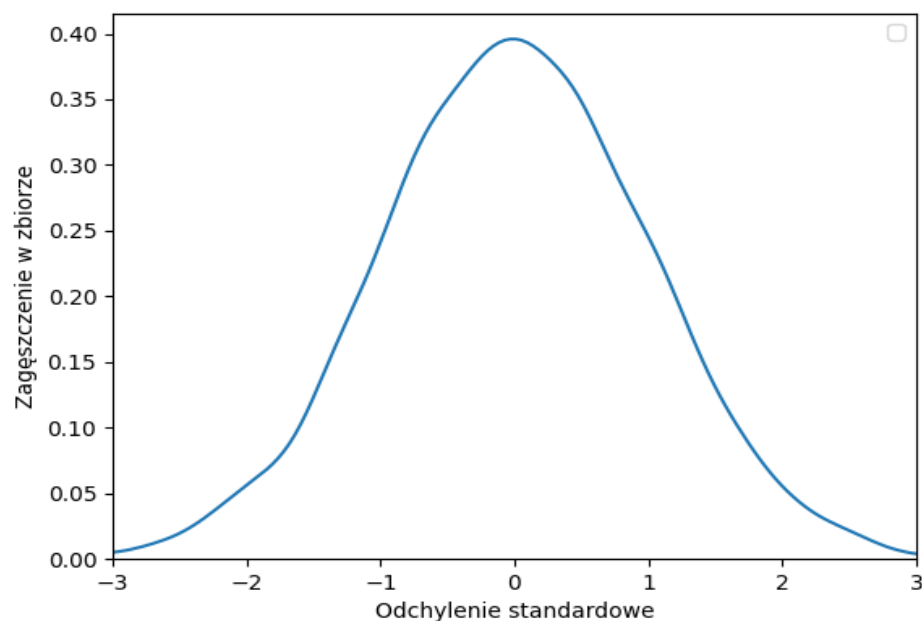
Gdzie:

n' to standaryzowana wartość

n to oryginalna wartość obserwacji

μ to średnia wartości w zestawie danych,

σ to odchylenie standardowe wartości w zestawie danych.



Rysunek 2 Wizualizacja rozkładu normalnego, opracowanie własne.

Jeżeli dane dystrybuują się zgodnie z zasadami rozkładu normalnego, należy przeprowadzić proces selekcji cech (feature selection)²⁸ - predyktorów²⁹. Selekcja cech polega na identyfikacji i wyborze zmiennych wejściowych które mają największy wpływ na zmienną zależną. Wybór to proces optymalizacyjny, upraszcza modele, aby te były łatwiejsze do

²⁸ Guyon, I., & Elisseeff, A. (2003), "An Introduction to Variable and Feature Selection"

²⁹ Predyktor to wartość, która jest używana do przewidywania wartości zmiennej zależnej. Model uczy się związków między predyktorami a zmienną zależną, aby móc prognozować wartości zmiennej zależnej dla nowych danych. Na przykład, w modelu przewidującym ceny domów, predyktory mogą obejmować powierzchnię domu, ilość sypialni itp.

interpretacji, skraca czasy szkolenia, zwiększa generalizację³⁰ poprzez redukcję overfittingu³¹ (przetrenowania), a także poprawia dokładność poprzez redukcję szumów³². Techniki wyboru cech obejmują:

- metody filtracyjne (filter methods), pozwalają wybrać odpowiednie cechy na podstawie miar statystycznych np. korelacji Pearsona (r).
- metody opakowujące (wrapper methods), automatycznie testują różne kombinacji cech i sprawdzają ich wpływ na wydajność modelu.

Oprócz wspomnianych metod istnieją jeszcze metody osadzone, lecz kwalifikują się jako część procesu uczenia modelu i są unikalne dla danego algorytmu.

2.4 Podział na zestawy danych

Dane są losowo dzielone na zestaw treningowy, walidacyjny i testowy. Ten podział ma na celu zapewnić, że model jest w stanie generalizować³³ dobrze na nowych, nieznanych danych, a nie tylko zapamiętywać wyniki dla danych treningowe.

Zestaw treningowy to największy podzbiór danych (zwykle 70-80%) używany do "uczenia" modelu, co oznacza dostosowywanie jego parametrów w celu minimalizacji funkcji kosztu³⁴. Funkcja kosztu, znana również jako funkcja straty, funkcja błędu lub funkcja celu jest miarą jak dobrze model uczenia maszynowego przewiduje wynik w porównaniu do rzeczywistego wyniku. Podczas trenowania modelu, celem jest minimalizacja wartości funkcji kosztu, czyli znalezienie takich parametrów modelu, które minimalizują błąd pomiędzy przewidywanym a rzeczywistym wynikiem. Różne typy modeli mogą używać różnych funkcji kosztu na przykład w regresji liniowej używa się MSE (Mean Squared Error), który jest sumą kwadratów różnic między wartościami wygenerowanymi przez model, a rzeczywistymi. Funkcje kosztu są kluczowym elementem algorytmów uczenia maszynowego, ponieważ definiują "cel", który model stara się osiągnąć podczas procesu trenowania.

³⁰ Generalizacja czyli zdolność przetrenowanego modelu do przewidywania wyników dla nowych, nieznanych danych.

³¹ Overfitting może wystąpić gdy model jest zbyt skomplikowany (zbyt głęboki) w porównaniu do ilości danych treningowych. W rezultacie model doskonale dopasowuje się do danych treningowych, jednak słabo przewiduje wartości dla nowych danych. - James, Gareth (2013), "An Introduction to Statistical Learning: with Applications in R"

³² Szum to zwykle niepożądane fluktuacje lub zakłócenia w danych, które mogą wprowadzać błędy lub nieprawidłowości w analizie danych lub procesie uczenia maszynowego.

³³ Generalizacja odnosi się do zdolności modelu do efektywnego przewidywania lub klasyfikowania nowych, nieznanych wcześniej danych.

³⁴ James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). "An Introduction to Statistical Learning"

Zestaw walidacyjny jest używany do dostrojenia hiperparametrów³⁵ modelu i oceny jego wydajności podczas procesu trenowania. Dostrojenie hiperparametrów polega na zmianie parametrów sterujących procesem uczenia algorytmu i decydują o wydajności modelu. Nie wszystkie modele będą posiadały takie same hiperparametry ze względu na swoją charakterystykę, ale podstawowe hiperparametry to:

- a) Tempo uczenia: Określa, jak szybko model uczy się z danych. Dosłownie małe tempo oznacza mniejszą amplitudę aktualizacji wag między neuronami w sieci, model robi małe kroki w kierunku minimalizacji błędu, co może prowadzić do wolniejszego, ale potencjalnie dokładniejszego uczenia. Zbyt duże tempo uczenia może prowadzić do niestabilnego procesu uczenia, ale grozi „przeskoczeniem” minimum funkcji kosztu, czyli wyniku który chcemy osiągnąć.
- b) Liczba epok: To liczba pełnych cykli, które algorytm uczenia maszynowego wykonuje na całym zestawie danych treningowych. Epoki są nazywane również iteracjami. Optymalna liczba iteracji jest bardzo ważna ponieważ zbyt mało prowadzi do niskiej wydajności modelu, a zbyt dużo do przetrenowania (overfitting).
- c) Wielkość próbki: Wpływa na to, jak wiele przykładów treningowych jest używanych jednocześnie do aktualizacji wag modelu. Zbyt małe próbki prowadzą do słabego przyuczenia modelu więc w przypadku większości modeli zalecane jest używanie jak największej próbki danych treningowych, ograniczeniem może być jedynie ilość pamięci do przechowywania danych.
- d) Liczba warstw: Zwana również „głębokością” modelu odpowiada za poziom komplikacji. Głębokość reguluje się empirycznie metodą prób i błędów, ponieważ jest bardzo zależna od danych na jakich pracuje model. Podobnie jak przy liczbie epok, jeśli model będzie miał zbyt dużą liczbę warstw dojdzie do przeuczenia.

Ostatnim zbiorem jest zestaw testowy który nie jest używany ani do treningu, ani do walidacji modelu. Po wytrenowaniu i dostrojeniu modelu, zestaw testowy jest używany do oceny ostatecznej wydajności modelu. Wyniki z zestawu testowego dają najlepszą ocenę, pokazując jak model prawdopodobnie poradzi sobie na nowych, nieznanych danych.

2.5 Trenowanie modelu

Trenowanie modelu to kluczowy etap uczenia maszynowego, polega na dostosowaniu parametrów modelu tak, aby jak najlepiej przywidywał wyniki dla danych treningowych. Cały

³⁵ Hiperparametr – nazwa używana do określenia parametrów, które sterują procesem uczenia algorytmu.

proces uczenia jest bardzo złożony i zależy od wielu czynników, poniżej jedynie przybliżę proces, a dokładniejszy opis będzie w kolejnej części pracy.

Pierwszym i najważniejszym krokiem jest wybór algorytmu uczenia zależnie od problemu który model ma rozwiązać. Najważniejsze cechy które odróżniają od siebie modele to: uczenie nadzorowane i nienadzorowane, problemy regresji lub klasyfikacji oraz linearne lub nieliniarne typy zależności cech.

W uczeniu nadzorowanym, model uczy się na podstawie zestawu danych treningowych, które zawierają zarówno cechy, jak i odpowiednie etykiety lub wyniki. Na przykład, jeśli model ma przewidzieć ceny domów na podstawie różnych cech, takich jak wielkość, lokalizacja, liczba sypialni itd., to cena domu byłaby etykietą, którą model próbuje oszacować. Najpopularniejsze algorytmy uczenia nadzorowanego to m.in. regresja liniowa, regresja logistyczna, maszyny wektorów nośnych (SVM), sieci neuronowe i drzewa decyzyjne. W uczeniu nienadzorowanym, model uczy się na podstawie zestawu danych, które nie zawierają etykiet. Celem jest znalezienie pewnych wzorców lub struktur w danych. Przykładowe zastosowania uczenia nienadzorowanego obejmują grupowanie (clustering), w którym model próbuje pogrupować dane w różne klastry³⁶ na podstawie podobieństwa cech, oraz redukcję wymiarowości, w której model próbuje zredukować liczbę cech, tracąc jak najmniej informacji. Przykłady algorytmów uczenia nienadzorowanego to k-means, hierarchiczne grupowanie, PCA (Principal Component Analysis), autoenkodery³⁷.

Algorytmy regresji polegają na przewidywaniu wyniku (zmiennej zależnej) na podstawie wielu zmiennych niezależnych. Na przykład, przewidywanie ceny domu na podstawie jego cech (wielkość, liczba sypialni itp.) to problem regresji. Klasyfikacja polega na przewidywaniu, do której z określonych kategorii należy dany obiekt. Na przykład, przewidywanie, czy dany e-mail jest spamem.

Modele liniowe, takie jak regresja liniowa, zakładają, że wynik jest liniową kombinacją cech wejściowych. Czyli, jeśli zmienimy wartość jednej cechy, przewidywany wynik zmieni się o stałą wartość, niezależnie od wartości innych cech. Modele nieliniarne, takie jak sieci neuronowe, są w stanie modelować złożone nieliniowe zależności między cechami wejściowymi a wynikiem. Czyli, zmiana wartości jednej cechy może wpływać na

³⁶ Klaster odnosi się do grupy danych, które są podobne do siebie według określonych kryteriów. To podstawowa technika grupowania, która jest częścią uczenia nienadzorowanego

³⁷ Autoenkoder to specjalny rodzaj sieci neuronowej używany do uczenia się skompresowanych, efektywnych reprezentacji danych, często zwanych kodami.

przewidywany wynik w zależności od wartości innych cech. Modele nieliniarne są zazwyczaj bardziej złożone i trudniejsze do interpretacji, ale mogą również obliczać bardziej skomplikowane wzorce w danych.

Kolejnym krokiem jest dostrojenie parametrów modelu, które zostały wybrane podczas podziału danych na 3 próbki – treningową, testową i walidacyjną. Funkcja straty pozwala oceniać jak dobrze model przewiduje wyniki w porównaniu do wartości danych testowych, jest to metryka którą algorytm stara się zminimalizować podczas treningu. W zależności od problemu i rodzaju modelu, mogą być używane różne funkcje kosztu. Na przykład, w problemach klasyfikacji, często stosuje się entropię krzyżową, dla problemów regresji błąd średniokwadratowy. Warto zauważyć, że wybór odpowiedniej funkcji kosztu jest kluczowy dla procesu trenowania modelu, ponieważ to ta funkcja jest optymalizowana podczas procesu uczenia. Wybór nieodpowiedniej funkcji kosztu może prowadzić do gorszych wyników lub utrudnić proces uczenia.

Po dopasowaniu najlepszej funkcji kosztu, model wykorzystuje technik matematyczne takie jak spadek gradientu do zmiany parametrów modelu (np. wagi w sieci neuronowej), aby zredukować wartość funkcji kosztu. Przy prostych algorytmach opartych na mało skomplikowanych danych, wystarczy oszacować najlepsze parametry metodą empiryczną, jednak przy bardziej skomplikowanych modelach, szacowanie zajmie zdecydowanie za dużo czasu, lub będzie niemożliwe. Spadek gradientu³⁸ (Gradient Descent) pokazuje kierunek najszybszego wzrostu funkcji kosztu, aby znaleźć minimum funkcji wystarczy zmienić znak przy wektorze kierunkowym. Aby skutecznie oszacować minimum funkcji należy powtarzalnie odejmować wartość gradientu od bieżących wartości parametrów, wielkość kroku jest kontrolowana przez parametr tempa uczenia. Analogicznie jak przy tempie uczenia algorytmu należy wybrać optymalną wartość, aby nie ominąć minimum funkcji. Krok odejmowania gradientu od parametrów jest powtarzany, aż do osiągnięcia minimalnej lub bardzo małej funkcji kosztu. Wyróżnia się wiele wariantów techniki spadku gradientu³⁹:

- Batch Gradient Descent: oblicza błąd dla każdego przykładu (przykład w tym kontekście oznacza różnicę między oszacowaną wartością, a wartością w zestawie walidacyjnym) w zestawie danych, ale aktualizuje model dopiero po przeprowadzeniu obliczeń dla wszystkich przykładów. Jest efektywny pod kątem zasobów

³⁸ <http://cs231n.github.io/optimization-1/> - Stanford University

³⁹ T. Hastie, R. Tibsharini, J. Friedman (2017) "The Elements of Statistical learning"

komputerowych, ponieważ nie wymaga aktualizacji wag modelu dla każdego przykładu treningowego. Ta metoda może być jednak bardzo czasochłonna dla dużych zestawów danych.

- Stochastic Gradient Descent (SGD): Zamiast przeprowadzać obliczenia na całym zestawie danych, SGD aktualizuje wagi modelu na podstawie jednego przykładu na raz. Jest to znacznie szybsze dla dużych zestawów danych, ale może prowadzić do większej zmienności w wynikach, ponieważ każda aktualizacja jest oparta na jednym przykładzie, a nie na całym zestawie danych.
- Mini-Batch Gradient Descent: Kompromis między Batch Gradient Descent a Stochastic Gradient Descent. Mini-Batch Gradient Descent aktualizuje wagi modelu po przetworzeniu niewielkiej grupy przykładów, zwykle od 10 do 1000. Pozwala to na szybsze przetwarzanie niż Batch Gradient Descent, ale jest stabilniejsze niż SGD, ponieważ aktualizacje są oparte na większym zestawie przykładów. W praktyce najczęściej używa się Mini-Batch Gradient Descent.

Po optymalizacji modelu następuje proces iteracyjny, zwykle polegający na wielokrotnym przetwarzaniu danych treningowych i dostosowywaniu parametrów modelu w celu zminimalizowania funkcji straty. To, kiedy proces zostanie zatrzymany zależy od wcześniej ustalonych parametrów np. gdy zostanie osiągnięta określona liczba epok, gdy błąd na zbiorze walidacyjnym przestaje się zmniejszać, gdy zmiana funkcji straty pomiędzy kolejnymi epokami przestaje się znacząco zmniejszać.

2.6 Testowanie, wdrażanie i monitorowanie modelu

Po kalibracji modelu na zestawie treningowym i walidacyjnym model jest testowany na zestawie testowym. Jest to zestaw danych, który nie był używany podczas treningu ani walidacji. Taki test daje rzetelną ocenę, jak dobrze model jest w stanie generalizować swoje predykcje na całkiem nowych danych. W praktyce, metryki wydajności są obliczane dla predykcji modelu na zestawie testowym i służą jako finalna ocena jakości modelu.

Jeśli wydajność modelu jest zadowalająca, następnym krokiem jest wdrożenie, czyli proces integracji modelu z bazami danych i systemami w organizacji tak, by model mógł być używany do przewidywania wyników na nowych danych w czasie rzeczywistym. Model będzie działał prawidłowo jeśli nowe dane źródłowe będą miały taką samą budowę co dane treningowe. Po wdrożeniu, ważne jest regularne monitorowanie wydajności przeprowadzając regularne testy zbliżone do tych wykonywanych podczas walidacji. W miarę jak nowe dane napływają,

wydajność modelu może się zmieniać. Ponadto, mogą wystąpić zmiany w danych wejściowych, które mogą wpływać na wydajność modelu.

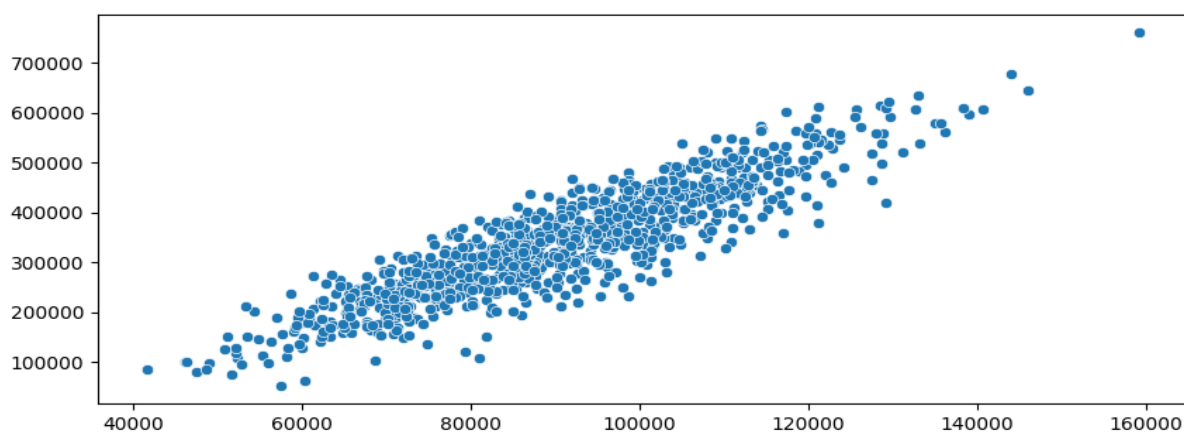
Na przykład, mogą pojawić się nowe kategorie w zmiennej kategorycznej, które nie były obecne w danych treningowych. Regularne monitorowanie modelu pozwala szybko wykryć takie problemy i podjąć odpowiednie działania. Jeśli model przestaje być precyzyjny należy go zaktualizować. Aktualizacja może obejmować dostosowanie hiperparametrów, ponowny trening na nowych danych, a nawet całkowite przeprojektowanie modelu.

2.7 Data Science Pipeline

Oprócz wyżej wymienionych procesów, które są podstawą każdego algorytmu, szczególnie w bankowości popularne jest dodanie dodatkowych analiz - Data Science (DS.). Data Science łączy statystykę, analizę danych, uczenie maszynowe i ich powiązane metody, aby zrozumieć i analizować rzeczywiste zjawiska za pomocą danych.

W praktyce, DS często angażuje się w procesie tworzenia decyzji i strategii biznesowych, pracując na przecięciu biznesu, statystyki i informatyki. Data Science pomaga organizacjom podejmować decyzje oparte na danych, identyfikować trendy i wzorce, które mogą pomóc firmom, oraz tworzyć modele, które mogą przewidzieć przyszłe zdarzenia.

Podstawową techniką DS jest eksploracyjna analiza danych (EDA) pozwala na zrozumienie i wizualizację ogólnego kształtu, struktury oraz charakterystyki danych. EDA może być wykonywana zarówno w początkowym jak i końcowym etapie przygotowywania danych, lub po otrzymaniu wyników w celu sformułowania hipotez. Przed załadowaniem danych pomaga zrozumieć naturę danych, ich strukturę, zidentyfikować potencjalne problemy (takie jak brakujące dane), odkryć wzorce i zależności między zmiennymi. Dodatkowo może służyć do wizualizacji funkcji kosztu lub metryk służących do jak najlepszej kalibracji modelu. Po załadowaniu danych do modelu, ponowna EDA może wykryć anomalie lub zidentyfikować obszary, gdzie wyniki modelu znacznie odbiegają od średniej. Takie anomalie mogą być łatwe do przeoczenia na dużych zbiorach danych i dopiero wizualizacja używając wykresu typu scatter-plot pozwoli je wykryć.



Rysunek 3 Poglądowy rozkład danych w modelu na wykresie punktowym,
Źródło: opracowanie własne.

Oprócz wizualizacji i analizy danych która ma pomóc w doborze i rozumieniu danych przetwarzanych przez algorytm analitycy DS. wykorzystują inżynierię cech (Feature Engineering) która poprawia efektywność modelu. Feature Engineering to proces tworzenia nowych cech (zmiennych, atrybutów) z istniejących podczas procesu przygotowywania danych do modelu. Podstawową różnicą między inżynierią cech a czyszczeniem i przygotowywaniem danych jest to, że Feature Engineering tworzy zupełnie nowe kategorie cech które są później używane przy uczeniu modelu. Nowe cechy mogą być stworzone z kilku istniejących gdy relacja między cechami i zmienną docelową jest nieliniowa lub złożona. Przykładem wykorzystania interakcji jest połączenie wolumenu i wzrostu ceny danej akcji przy trendzie wzrostowym. Taka nowa cecha w modelu może służyć jako wskaźnik czy mimo rosnącej ceny inwestorzy nadal są chętni do kupna akcji. Oprócz tworzenia cech, dane mogą być również grupowane w bardziej ogólne grupy. Na przykład wiek może zostać zgrupowany w przedziały 10-letnie. Tworzenie przedziałów upraszcza model i redukuje funkcje kary.

Po trenowaniu modelu, wyniki są interpretowane i omawiane w kontekście problemu biznesowego. Mogą to być wykresy, statystyki, wyniki feature importance⁴⁰, itp.

⁴⁰ Feature importance (ważność cech) to technika używana w uczeniu maszynowym do określenia, które cechy mają największy wpływ na przewidywane wyniki modelu.

Rozdział 3: Praktyczne zastosowanie algorytmów i matematyki finansowej do analizy rynku.

W skomplikowanym i dynamicznie zmieniającym się rynku finansowym, tradycyjne strategie inwestycyjne często nie wystarczają. Zasady inwestycyjne, które kiedyś przynosiły stabilne zyski, mogą stać się przestarzałe w obliczu nowych trendów rynkowych, zmienności czy nawet globalnych zmian ekonomicznych. Rola uczenia maszynowego i sztucznej inteligencji koncentruje się na przewidywaniu trendów rynkowych, optymalizacji portfela inwestycyjnego, zarządzaniu ryzykiem i wielu innych zadaniach, które były wcześniej niemożliwe do wykonania na dużą skalę. W tym rozdziale dokładnie opisze zasady działania algorytmów które są najczęściej wykorzystywane przy tworzeniu rozwiązań dla finansów oraz przybliżę strategie inwestycyjne które wykorzystują algorytmy.

3.1 Techniki analizy rynku

Analiza rynku jest kluczowym elementem podejmowania świadomych decyzji inwestycyjnych. Różnorodność metod analitycznych umożliwia inwestorom dogłębne zrozumienie dynamiki rynku i identyfikację potencjalnych możliwości inwestycyjnych. Dwa główne podejścia do analizy rynku, które są przedmiotem tego rozdziału, to analiza techniczna i analiza fundamentalna.

Analiza fundamentalna jest procesem oceny rzeczywistej wartości firmy, biorąc pod uwagę jej zdolność do generowania zysków, stabilność finansową, kondycję rynku, zarządzanie i inne czynniki, które mogą wpłynąć na jej przyszłą rentowność. Ten rodzaj analizy jest często stosowany przez inwestorów długoterminowych, którzy starają się zidentyfikować instrumenty finansowe które są niedocenione przez rynek.

Analiza techniczna polega na studiowaniu wzorców cenowych i trendów rynkowych za pomocą danych historycznych i statystycznych. Ta metoda często wykorzystuje wykresy i różnego rodzaju wskaźniki techniczne do prognozowania przyszłych ruchów cen. Stosowana jest głównie przez krótkoterminowych traderów, takich jak spekulanci i day traderzy, którzy starają się wykorzystać fluktuacje cen na rynku. Analiza techniczna jest agresywną strategią mającą na celu wypracowanie ponadprzeciętnych zysków przy wyższej tolerancji na ryzyko. John J. Murphy w książce „Analiza techniczna rynków finansowych” określa takie podejście do inwestowania w trzech najważniejszych zasadach:

- Rynek dyskontuje wszystko
- Ceny podlegają trendom
- Historia się powtarza

Zgodnie z tymi zasadami inwestorzy starają się określić okresowe trendy, szukać wzorców cenowych, historycznych analogii oraz szacować szansę na zmianę trendu.

Metody uczenia maszynowego, ze względu na swoją zdolność do przetwarzania dużej ilości danych i odkrywania złożonych wzorców świetnie adaptują do obu tych technik. Uczenie maszynowe wykorzystuje zarówno techniki analizy technicznej, jak i fundamentalnej, integrując je w zaawansowane modele predykcyjne. Ta unikalna kombinacja pozwala na generowanie dokładniejszych i bardziej dynamicznych prognoz, które mogą znacznie poprawić proces decyzyjny inwestora. W tym rozdziale zostanie opisane, jak różne techniki wycen

instrumentów finansowych oparte o statystykę, matematykę finansową i uczenie maszynowe mogą być stosowane w kontekście tych dwóch podejść do analizy rynku.

3.2 Modele statystyczne

Modele statystyczne to matematyczne przedstawienie danych liczbowych. Jednym z popularnych modeli statystycznych jest model autoregresyjny średniej zwrotnej (Autoregressive Mean Reversion - ARM). Ten model zakłada, że cena aktywów będzie dążyć do swojej średniej historycznej, a odchylenia od tej średniej będą miały tendencję do zmniejszania się z czasem. To jest podstawą dla strategii inwestycyjnej mean reversion (średnia zwrotna). W praktyce, strategia ta polega na identyfikacji sytuacji, w których cena aktywów odchyła się znacznie od swojej średniej, co jest interpretowane jako sygnał, że cena prawdopodobnie powróci do swojej średniej w przyszłości. Strategia średniej⁴¹ została pierwszy raz zastosowana w latach 80 przez J. Simons'a do handlu na giełdach towarowych i walutowych, ponieważ te instrumenty wykazywały tendencję do powrotu do średniej ceny. Szybko jednak strategia przestała działać i została określona jako „low-hanging fruit” czyli była zbyt prosta i zbyt popularna aby przynosiła długoterminowe zyski. Obecnie algorytmy mogą wykorzystywać wskaźnik SMA⁴² lub EMA⁴³ jako jeden z parametrów do identyfikacji trendów rynkowych.

Bardziej zaawansowanym modelem jest analiza szeregów czasowych która koncentruje się na analizie danych zbieranych w regularnych odstępach czasu. Celem analizy jest identyfikacja istotnych wzorców, trendów i cykli w danych, które mogą być wykorzystane do prognozowania przyszłych obserwacji. Jednym z kluczowych narzędzi w analizie szeregów czasowych jest model autoregresyjny z zintegrowanym średnim ruchem, znany jako ARIMA. Model jest specyficzny ponieważ pozwala na modelowanie i prognozowanie danych, które mogą nie być stacjonarne⁴⁴. ARIMA działa w następujący sposób:

- Autoregresja⁴⁵ (AR): Ta część modelu szuka i oblicza zależności między obecną i poprzednią lub kilkoma poprzednimi obserwacjami.
- Zintegrowanie danych (I): Ten element różniczuje dane w szeregu aby zniwelować trend tj. sprawić aby szereg był stacjonarny
- Średnia krocząca (MA): Wykorzystuje zależności między obserwacją, a błędem z poprzedniej prognozy. Komponent MA porównuje wartości wyliczone przez model z wartościami w zbiorze testowym co pozwala modelowi na „uczenie się” i dostosowanie prognoz w celu minimalizacji błędu.

⁴¹ Gregory Zuckerman (2020) „Człowiek, który rozwiązał rynek: Jak Jim Simons zapoczątkował rewolucję quantów”

⁴² SMA(Simple Moving Average, Prosta Średnia Krocząca) Popularna technika analizy technicznej jest wyliczana poprzez sumowanie n ostatnich obserwacji (np. cen zamknięcia) i następnie podzielenie tej sumy przez n.

⁴³ EMA (Exponential Moving Average, Wykładnicza Średnia Krocząca) – Bardziej zaawansowana technika, podobnie jak SMA liczy średnią z n obserwacji, ale dodatkowo nadaje wyższą wagę dla najnowszych danych.

⁴⁴ Stacjonarność danych odnosi się do właściwości szeregów czasowych, w których statystyczne właściwości, takie jak średnia i wariancja, nie zmieniają się w czasie. Średnia roczna temperatura powietrza może być stacjonarnym szeregiem danych, wraz z kolejnym rokiem wartość oczekiwana (średnia) nie wiele różni się od poprzedniego roku.

⁴⁵ Autoregresja to specjalny przypadek regresji liniowej, który jest stosowany w analizie szeregów czasowych. W modelu autoregresji, zmienna zależna (y) jest modelowana jako funkcja jej własnych poprzednich wartości.

Modele szeregów czasowych, opierają się na kilku założeniach, które często nie są spełnione w rzeczywistości rynków finansowych, co może wpływać na ich zdolność do efektywnego prognozowania cen akcji. Zakładają, że zależności w danych są liniowe i stacjonarne. Jednakże, na rynkach finansowych, dynamika cen często jest złożona i nieliniowa. Zmienność cen może być skomplikowana, a zależności między zmiennymi finansowymi często zmieniają się w czasie. Dodatkowo zakładają, że dane są autokorelowane⁴⁶. Choć istnieją dowody na pewną formę autokorelacji w danych finansowych w krótkim terminie, nie jest to zasada. Eugen Fama, twórca teorii Efektywnego Modelu Rynku argumentuje, że ceny akcji są efektem "błądzenia losowego", gdzie przyszłe zmiany są niezależne od przeszłych z ewentualnymi krótkimi okresami trendu lub powrotu do średniej. Abstrahując od „losowości” cen modele szeregów czasowych nie uwzględniają wpływu zdarzeń zewnętrznych, takich jak decyzje polityczne, wydarzenia gospodarcze czy zmiany w preferencjach rynkowych, które nie są możliwe do przewidzenia analizując poprzednie wyceny, a mogą gwałtownie wpłynąć na cenę akcji.

Podsumowując, klasyczne modele statystyczne, mimo że mają swoje ograniczenia, nadal mogą być użyteczne jako narzędzia do wstępnej analizy danych finansowych. Choć mogą nie być w stanie dokładnie przewidzieć przyszłych cen akcji z powodu złożoności i dynamiki rynków finansowych, mogą dostarczyć użytecznych informacji o trendach i wzorcach w danych.

3.3 Modele stochastyczne

Modele stochastyczne⁴⁷ to matematyczne modele, które uwzględniają element losowości. Są one wykorzystywane w ekonomii i finansach, do analizy i prognozowania zjawisk, które są podatne na losowe perturbacje. W latach 60. i 70. XX wieku pojawiły się pierwsze modele stochastyczne wykorzystywane do wyceny instrumentów pochodnych. Od tamtej pory modele stochastyczne stały się nieodzowną częścią analizy finansowej i inwestycyjnej. Są one wykorzystywane do oceny ryzyka, wyceny różnych instrumentów finansowych i modelowania portfeli inwestycyjnych. Ich zaawansowanie i złożoność rosną wraz z postępem technologicznym.

W modelach stochastycznych, zamiast traktować zjawiska jako deterministyczne⁴⁸, uwzględnia się wpływ czynników losowych. Analiza stochastyczna w kontraście do deterministycznej, nie koncentruje się na szczegółowej analizie indywidualnych czynników wpływających na proces analizowany np. wycena instrumentu finansowego, a raczej, skupia się na kwantyfikacji niepewności. Istotą takiego podejścia jest generowanie wielu możliwych scenariuszy i obliczanie ich prawdopodobieństwa wystąpienia. Podstawowym narzędziem stosowanym w tym celu są procesy losowe, które pozwalają na uwzględnienie wpływu różnorodnych czynników ryzyka na modelowane zjawisko, bez konieczności dokładnego opisu ich struktury i relacji. W efekcie, pomimo braku pełnej informacji na temat dynamiki systemu, możliwe jest oszacowanie dystrybucji prawdopodobieństwa wyników, co daje podstawy do oceny ryzyka i podejmowania decyzji w warunkach niepewności.

⁴⁶ Autokorelacja, znana również jako korelacja serialna, to statystyczna korelacja między danymi szeregów czasowych dla kolejnych okresów.

⁴⁷ Mark S. Joshi (2003) „The Concepts and Practice of Mathematical Finance”

⁴⁸ Zjawisko deterministyczne to takie, w którym przyszłe stany lub wyniki są w pełni określone przez aktualne stany i warunki systemu, a także znane prawa rządzące tym systemem.

Najbardziej znanym modelem jest model Blacka-Scholesa, który opisuje ruch cen jako proces stochastyczny, znany jako ruch Browna. Model jest stosowany do wyceny instrumentów pochodnych, głównie opcji europejskich⁴⁹, został wymyślony przez ekonomistów Fischera Blacka, Myrona Scholesa i Roberta Mertona w 1973r.

Model Blacka-Scholesa opiera się na kilku założeniach. Zakłada między innymi, że:

- Cena instrumentu bazowego (np. akcji) zmienia się według procesu stochastycznego z ciągłym czasem.
- Nie ma kosztów transakcji ani podatku.
- Stopa wolna od ryzyka jest stała i znana.
- Opcje są europejskie, co oznacza, że mogą być wykonywane tylko w dniu wygaśnięcia.
- Nie ma możliwości arbitrażu.

Proces stochastyczny w kontekście modeli wyceny jest używany do opisanie zmiany cen aktywów finansowych. Kluczową cechą procesu stochastycznego jest to, że przyszłe zmiany są nieprzewidywalne i zależą od oszacowanego prawdopodobieństwa. Termin "w ciągłym czasie" oznacza, że zmienna losowa może zmienić wartość w dowolnym momencie, a nie tylko w określonych punktach czasowych. W praktyce jednak dane finansowe są często obserwowane w dyskretnych punktach czasowych (np. codziennie, co godzinę). Dla niektórych modeli wymagane jest aby zmienna losowa była obserwowana w czasie ciągłym. Dlatego istnieją techniki takie jak metoda Monte Carlo, które pozwalają symulować ścieżki procesów stochastycznych w czasie ciągłym na podstawie dyskretnych danych.

Drift w modelu stochastycznym jest terminem używanym do opisanie średniego trendu wzrostu procesu, a w modelach wyceny jako oczekiwany zwrot z inwestycji. Volatylność jest miarą rozproszenia zmiennych losowych w procesie. Wysoka volatylność oznacza, że cena akcji może znacznie fluktuować w krótkim okresie czasu, co zwiększa ryzyko. Volatylność jest kluczowym składnikiem w modelach wyceny opcji.

Podstawowym elementem modelu Blacka-Scholesa jest tzw. równanie różniczkowe Blacka-Scholesa. Równanie opisuje, jak cena opcji (lub ogólniej, pochodnej finansowej) zmienia się w czasie w odpowiedzi na zmiany ceny aktywa bazowego.

$$C = S_0 * N(d_1) - X * e^{-rT} * N(d_2)$$

gdzie:

C – wartość opcji,

S₀ - aktualna cena aktywa bazowego,

X – cena wykonania opcji (Strike price),

r - bezpieczna stopa procentowa⁵⁰,

⁴⁹ Opcja europejska to typ kontraktu opcji, który daje posiadaczowi prawo do kupna lub sprzedaży określonego aktywa po określonej cenie w dniu wygaśnięcia opcji.

⁵⁰ Bezpieczna stopa procentowa to teoretyczna stopa zwrotu inwestycji, która nie niesie ze sobą ryzyka np. obligacji rządowej.

T - czas do wygaśnięcia opcji,

e^{51} – liczba Eulera

N() - dystrybuanta standardowego rozkładu normalnego,

σ - odchylenie standardowe aktywa bazowego, czyli miara volatylności.

LN - Logarytm naturalny z ilorazu S0 i X daje miarę relacji między obecną ceną aktywa a ceną wykonania opcji

$$d1 = (\ln(S0 / X) + (r + \frac{\sigma^2}{2}) * T) / (\sigma * \sqrt{T}),$$

$$d2 = d1 - \sigma * \sqrt{T},$$

Cena, którą oblicza model Blacka-Scholesa, nie jest prognozą przyszłej ceny aktywa bazowego, ale raczej wyceną samej opcji, której wartość zależy od przyszłej ceny aktywa. Dużym ograniczeniem modelu jest założenie, że zmienność jest stała i niezależna od czasu co jest rzadkością na rynkach finansowych. Dlatego powstały inne modele które służą do wyceny opcji.

Kolejnym bardzo popularnym modelem służącym do wyceny opcji jest model dwumianowy, zwany też drzewem dwumianowym. Głównym założeniem modelu jest to, że cena aktywa może poruszać się tylko w dwóch kierunkach: w górę lub w dół, z pewnymi prawdopodobieństwami, przez pewien określony okres czasu. W przeciwieństwie do modelu Blacka-Scholesa dane finansowe mogą być obserwowane w dyskretnych punktach czasowych.

Przy założeniu, że cena aktywa obecnie wynosi S, to w kolejnym kroku czasowym cena ta może wzrosnąć do uS (gdzie $u > 1$) lub spaść do dS (gdzie $0 < d < 1$). Te mnożniki u (up) i d (down) są zazwyczaj ustalane tak, aby odzwierciedlały oczekiwaną zmienność aktywa. Dla opcji na akcje, cena opcji w modelu binominalnym jest wyznaczana poprzez modelowanie przyszłych możliwych cen akcji, następnie wyliczanie wartości opcji dla tych możliwych cen i dyskontowanie tych wartości z powrotem do dnia dzisiejszego.

Główne równanie modelu binominalnego dla wyceny opcji Europejskich wygląda następująco:

$$C = e^{(-rT)} * (pC_u + (1-p)C_d)$$

gdzie:

C wartość opcji,

r to bezpieczna stopa procentowa,

T to czas do wygaśnięcia opcji,

p to prawdopodobieństwo wzrostu ceny,

⁵¹ e - reprezentuje podstawę logarytmów naturalnych, czyli liczbę Eulera. Jest to stała matematyczna, która jest irracjonalna (nie można jej wyrazić jako dokładny iloraz dwóch liczb całkowitych) i transcendentna (nie jest rozwiązaniem żadnego niewielkiego wielomianu o współczynnikach całkowitych). W przybliżeniu $e = 2.71828$.

C_u to wartość opcji, gdy cena aktywa wzrasta,

C_d to wartość opcji, gdy cena aktywa spada.

Prawdopodobieństwo wzrostu ceny (p) jest wyliczane za pomocą wzoru:

$$p = (e^{r\Delta t} - d) / (u - d)$$

Drzewo dwumianowe jest konstruowane od końca (od daty wygaśnięcia opcji) co sprawia, że model może być przydatny również do wyceny opcji Amerykańskich, które mogą być wykonane przed terminem wygaśnięcia. Model mimo tego, że jest dosyć prosty daje dość precyzyjne wyniki szczególnie dla opcji z krótkim czasem wygaśnięcia.

Bardziej zaawansowane modele wykorzystują dwa inne procesy stochastyczne które są fundamentalne w matematyce finansowej – Proces Wienera i Proces Poissona. Wiele modeli wykorzystuje, te procesy ale pojedynczo są zbyt proste aby służyły do kwantyfikacji dynamiki rynku.

Proces Wienera, często nazywany ruchem Browna, to proces stochastyczny, który ma niezależne przyrosty o normalnym rozkładzie. Czyli liczba zdarzeń (lub "skoków") w jednym okresie czasu nie wpływa na liczbę zdarzeń w innym okresie. Właściwość ta jest znana jako brak pamięci. Proces Wienera definiuje wzór, a otrzymany wynik to funkcja rozkładu „skoków” w czasie:

$$W(t) = \sigma Z \sqrt{t}$$

gdzie:

t to czas,

σ to odchylenie standardowe (zmiennosc),

Z ⁵² to zmienna losowa o standardowym rozkładzie normalnym

Proces Poissona modeluje liczbę zdarzeń występujących w określonym okresie czasu, jest zdefiniowany wzorem:

$$N(t) = \text{Poisson}(\lambda t)$$

gdzie:

t to czas,

λ to średnia liczba zdarzeń na jednostkę czasu (zwana intensywnością),

$\text{Poisson}(\lambda t)$ oznacza, że liczba zdarzeń w czasie t ma rozkład Poissona o średniej λt . λ reprezentuje średnią liczbę "skoków" (lub zdarzeń) na jednostkę czasu. W kontekście modelowania finansowego, jednostką czasu może być dzień, tydzień, miesiąc, itd., a "skok" może być zdefiniowany na różne sposoby, w zależności od konkretnego zastosowania modelu.

⁵² Z w praktyce jest generowane przez komputerowy generator liczb losowych, który jest zaprojektowany tak, aby generować liczby, które są niezależne i identycznie rozłożone zgodnie z rozkładem normalnym. Generatory wykorzystują różne techniki np. metoda Monte Carlo.

Na przykład, może to być określone jako gwałtowna zmiana ceny instrumentu, która przekracza określony próg. W takim przypadku, λ byłoby średnią liczbą takich gwałtownych zmian na jednostkę czasu.

Kolejnym rodzajem modeli służących do wyceny opcji są modele zmienności skokowej np. model Mertona. Kluczową cechą tych modeli jest uwzględnienie "skoków" w cenie, co może modelować nagłe, nieprzewidywalne zmiany na rynku, takie jak gwałtowne spadki cen akcji.

Model Mertona wyrażany jest wzorem:

$$dS = \mu S dt + \sigma S dW + S dJ$$

Gdzie:

S jest wartością aktywa.

dS to zmiana wartości aktywa.

μ to drift

σ to zmienność aktywa

W to proces Wienera, reprezentujący losowe ruchy cen aktywów.

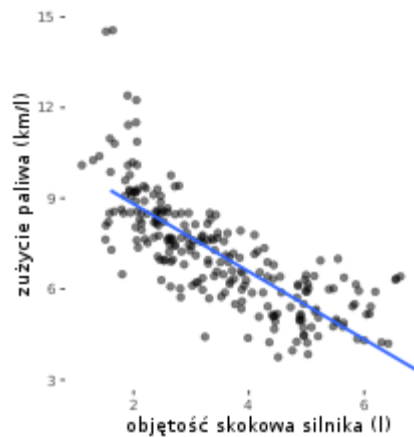
J to proces Poissona

dt to zmiana czasu.

Można to interpretować jako dodanie składnika "skoku" do standardowego modelu Blacka-Scholesa, który zakłada, że ceny aktywów zmieniają się płynnie w czasie. W efekcie model jest bardziej elastyczny i dokładniejszy. Można go rozszerzyć o dodatkowe parametry takie jak średnia wielkość skoku (m) oraz zmienność skoków (v).

3.4 Algorytm regresji liniowej

Model regresji liniowej stanowi fundamentalne narzędzie w statystyce i analizie danych, które ma na celu zidentyfikowanie liniowej zależności pomiędzy zmienną zależną (target) a jedną lub wieloma zmiennymi niezależnymi (cechy). Głównym zadaniem tego modelu jest minimalizacja sumy kwadratów różnic między wartościami obserwowanymi a wartościami estymowanymi przez model. Na poniższym rysunku punkty to cechy, a niebieska linia to najbardziej optymalny model regresji liniowej.



Rysunek 4 – Wizualizacja regresji liniowej

Źródło: pl.wikipedia.org/wiki/Regresja_liniowa

Podstawowy model regresji liniowej próbuje znaleźć najlepszą prostą (w przypadku jednej zmiennej) lub płaszczyznę/hyperpłaszczyznę (w przypadku wielu zmiennych). Czyli szuka relacji liniowej pomiędzy zmienną zależną a cechami. Matematyczny model regresji liniowej określany jest wzorem:

$$Y = aX_1 + bX_2 + \dots + zX_n + c$$

Gdzie:

Y jest zmienną zależną (target),

X_1, X_2, \dots, X_n to zmienne niezależne (cechy),

a, b, ..., z to wagi przypisane do zmiennych niezależnych,

c to bias (przesunięcie).

Wagi są dobierane w taki sposób, aby minimalizować różnicę między przewidywanymi a rzeczywistymi wartościami zmiennej zależnej. Bias odnosi się do stałej wartości, która jest dodawana do sumy cech modelu. Jest to jedna z dwóch podstawowych składowych modelu, obok wag (które mnożą cechy). Bias umożliwia przesunięcie linii regresji, dzięki czemu model staje się bardziej elastyczny i może lepiej dopasować się do danych.

Choć algorytm regresji liniowej jest stosunkowo prosty, ma praktyczne zastosowanie w strategii APT (Arbitrage pricing theory). APT wykorzystuje modele czynnikowe np. Fama-French Three Factor czyli klasyczne techniki wyceny instrumentów na podstawie analizy fundamentalnej. W modelu Fama-French realna wartość spółki jest wyliczana na podstawie sumy znormalizowanych wartości: beta(współczynnik ryzyka indeksowego)⁵³, kapitalizacji i współczynnika ceny do wartości księgowej⁵⁴. Następnie ta realna wartość jest porównywana ze wzrostem cen akcji na giełdzie. Jeśli wyliczona wartość jest wysoka w porównaniu do innych firm z podobną kapitalizacją lub wartość w modelu rośnie, a cena akcji nie, oznacza to, że

⁵³ Beta reprezentuje „wpływ” szerszego rynku, np. indeksu giełdowego, na zmiany ceny akcji. Na przykład jeśli beta dla aktywa wynosi 1 oznacza to, że przewidyuje się, że aktywo to zareaguje na ruchy rynku w taki sam sposób jak rynek. Czyli, jeśli rynek wzrośnie o 10%, oczekuje się, że aktywo również wzrośnie o 10%.

⁵⁴ C/WK to popularny indyktor wykorzystywany w analizie fundamentalnej. Jest obliczany poprzez podzielenie ceny akcji przez aktywa firmy minus zobowiązania.

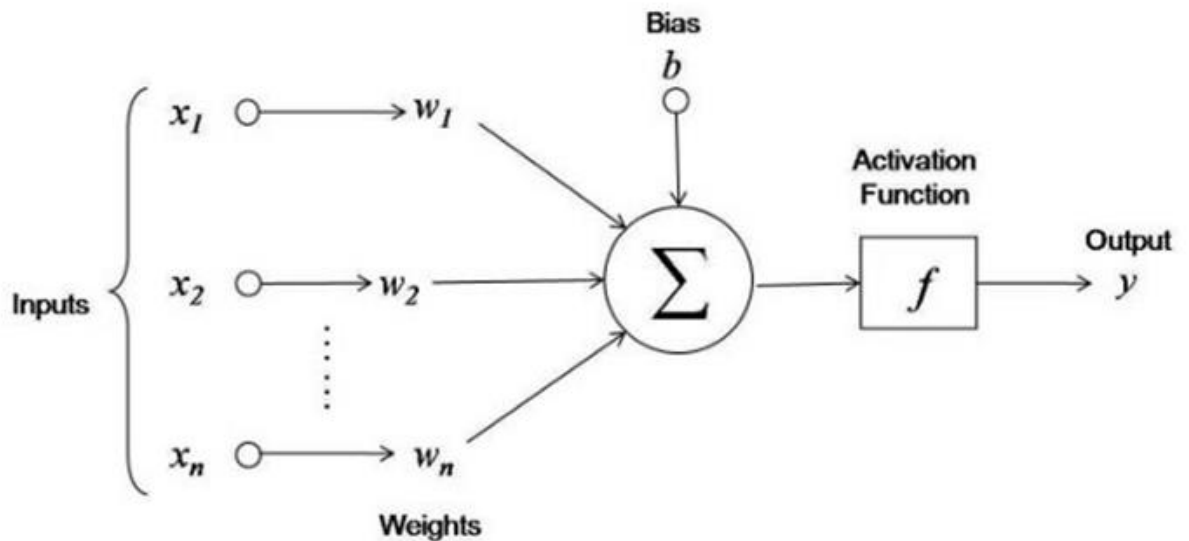
aktywo jest chwilowo wycenione poniżej wartości. Taka wycena pozwala inwestorowi ocenić wiele spółek, wybrać te najniżej wycenione i wykorzystać przewagę. Arbitraż w znaczeniu tej strategii nie odnosi się do wyceny na innych giełdach, tylko na różnicy wycen między wartością ceny akcji, a faktyczną wartością wyliczoną przez model, w literaturze określa się tą technikę jako arbitraż statystyczny. Model Fama-French jest bardzo prosty i obecnie raczej nie używany na szeroką skalę w klasycznej formie.

Algorytmy regresji liniowej znacząco rozszerzają możliwości implementacji modeli czynnikowych. Po pierwsze, algorytmy te nie ograniczają liczby czynników analizowanych w modelu, co umożliwia dodanie dowolnej liczby zmiennych które mają potencjalny wpływ na wycenę giełdową akcji. Po drugie, regresja liniowa pozwala na przypisanie indywidualnych wag dla każdego czynnika, co zwiększa precyzyjność estymacji oraz dokładność prognoz. W rezultacie, analiza ta umożliwia bardziej szczegółowe zrozumienie dynamiki rynkowej oraz skuteczniejsze prognozowanie przyszłych tendencji długoterminowych.

Strategie APT i arbitraż statystyczny jest stosowany na szeroką skalę przy budowie portfeli inwestycyjnych. Dodatkowo regresja liniowa jest wykorzystywana w wstępnej fazie zaawansowanej analizy danych np. przy tworzeniu nowych cech.

3.5 Deep Learning

Sieci neuronowe konstruowane są z szeregu warstw neuronów, tworzących strukturę połączoną. Każdy neuron w warstwie przetwarza dane wejściowe przez przyznawanie wag dla poszczególnych połączeń, a następnie przekazuje informacje dalej do kolejnej warstwy. Przetworzone sygnały są agregowane za pomocą funkcji sumującej, po czym podlegają operacji określonej przez wybraną funkcję aktywacji. Ten proces jest kontynuowany od warstwy wejściowej, poprzez liczne warstwy ukryte, aż do osiągnięcia warstwy wyjściowej. Istnieje wiele różnych funkcji aktywacji, a ich wybór zależy bezpośrednio od konkretnego zadania, które ma rozwiązać sieć. W kontekście problemów regresji, jak prognozowanie przyszłej ceny instrumentów finansowych, często wybieraną funkcją jest funkcja liniowa (identity function), która nie wprowadza żadnej zmiany do wartości wyjściowej. Głębokość sieci neuronowej zależy od ilości warstw ukrytych między warstwą wejściową a wyjściową. Im więcej warstw tym bardziej sieć jest skomplikowana. W skomplikowanych problemach może to zminimalizować funkcje celu, jednak gdy warstw będzie za dużo model będzie przeuczony i nie będzie działał na nieznanymi danych.



Rysunek 5 Uproszczony model działania sieci neuronowej

Źródło: <https://pl.wikipedia.org/wiki/Perceptron>

W trakcie procesu uczenia, sieci neuronowe dostosowują wartości wag występujących w połączeniach pomiędzy neuronami. Celem jest minimalizacja różnicy pomiędzy prognozowanymi wartościami a obserwacjami empirycznymi. Procedura ta jest najczęściej realizowana za pomocą techniki propagacji wstecznej⁵⁵ (ang. backpropagation) i optymalizacji gradientowej. Propagacja wsteczna jest kluczowym mechanizmem stosowanym do efektywnego trenowania sieci neuronowych, a jej zasada działania jest następująca:

1. Forward Pass: Dane wejściowe są przekazywane przez wszystkie warstwy sieci aż do warstwy wyjściowej.
2. Obliczanie błędu: Za pomocą funkcji straty przeliczany jest błąd przewidywań sieci zestawu uczącego w porównaniu do zestawu walidacyjnego.
3. Backward Pass: W ostatnim kroku algorytm zaczyna aktualizować wagi w celu minimalizacji błędu.
4. Kroki od 1 do 3 są powtarzane, aż do momentu kiedy funkcja kary osiągnie ustaloną wartość

Propagacja wsteczna używa pochodnych cząstkowych i łańcuchowej reguły różniczkowania do obliczenia wag, dzięki temu sieć „uczy się” z błędów popełnianych w poprzednich iteracjach i minimalizuje funkcję kary.

Sieci neuronowe sprawdzają się najlepiej w problemach, które wymagają identyfikacji skomplikowanych wzorców i nieliniowych zależności. Wymagają dużych ilości danych do efektywnego trenowania oraz charakteryzują się niską interpretowalnością⁵⁶. Głębokie sieci neuronowe są kluczowym elementem w strategiach Regime Switching, czyli wykorzystywane są do wykrywania krótkoterminowej zmiany trendu. Strategia wykrycia zmiany trendu może

⁵⁵ Ł. Wordliczek (2021) “Wykorzystanie sztucznych sieci neuronowych”

⁵⁶ Interpretowalność modelu uczenia maszynowego odnosi się do łatwości, z jaką człowiek może zrozumieć, dlaczego model dokonał określonego przewidywania. Model jest uważany za wysoce interpretowalny, jeżeli jego działanie i decyzje są łatwe do zrozumienia i wyjaśnienia.

być stosowana na wielu instrumentach, ale najczęściej jest używana na indeksach, towarach oraz walutach. Sieci neuronowe przetwarzają obszerne zbiory danych w celu identyfikacji momentów, w których trend rynkowy ulegał krótkoterminowej zmianie, zazwyczaj po znacznych fluktuacjach cenowych. Sieci neuronowe analizują czynniki takie jak:

- częstotliwość występowania znaczących skoków cenowych, którym towarzyszyła krótkoterminowa zmiana trendu
- wymagana wielkość skoku cenowego, aby mógł on być potencjalnym sygnałem zmiany trendu
- optymalny okres utrzymania „nowego trendu”

Po wykonaniu początkowej analizy, mającej na celu ustalenie, czy na danym instrumencie występują zmiany trendu po gwałtownym wzroście ceny, przeprowadzany jest dalszy proces analizy. Głęboka sieć neuronowa analizuje zmienność cen w różnych stanach rynku⁵⁷, dążąc do znalezienia punktu, w którym gwałtowny wzrost ceny najprawdopodobniej sygnalizuje krótkoterminową zmianę trendu. Okres w którym utrzymuje się „nowy” trend, jest średnim okresem dla poszczególnych stanów rynku. Tradycyjnie stany rynku to:

- Hossa - okres, w którym ceny rosną lub mają tendencję do wzrostu
- Bessa - Okres, w którym ceny spadają lub mają tendencję do spadku
- Trend boczny - Okresy bez wyraźnej tendencji wzrostowej lub spadkowej

Rozwój sztucznej inteligencji pozwolił jednak na wykrycie innych stanów rynku, ich rozpoznanie jest kluczowe dla prawidłowego działania strategii inwestycyjnych.

- Rynek wykupiony (Overbought Market) - Rynek, który, według kryteriów technicznych, osiągnął stan, w którym cena jest wyższa niż uzasadnia to jej wartość.
- Rynek wyprzedany (Oversold Market) - Rynek, który, według kryteriów technicznych, osiągnął stan, w którym cena jest niższa niż uzasadnia to jej wartość.
- Rynek o wysokiej zmienności (High Volatility Market) - Ceny gwałtownie rosną i spadają, taki stan jest bardzo częsty przed i po ogłoszeniu ważnych danych makroekonomicznych
- Okresy wysokiego przepływu kapitału - Okresy, kiedy duża ilość pieniędzy jest inwestowana lub wycofywana z rynku.
- Okresy niskiego przepływu kapitału - Okresy, kiedy przepływ kapitału na rynek jest minimalny.
- Stany związane z niskimi lub wysokimi stopami procentowymi.

Sieci neuronowe, jako zaawansowane metody uczenia maszynowego, odgrywają kluczową rolę w wielu strategiach inwestycyjnych, które poszukują krótkoterminowych wzorców cenowych, sezonowości, cykliczności lub niejawnych korelacji. Implementacja tych strategii na rynku finansowym jest często określana jako handel algorytmiczny, który polega na systematycznym wykorzystywaniu zidentyfikowanych przez algorytmy zależności do generowania sygnałów kupna i sprzedaży. Stosowanie handlu algorytmicznego pozwala na automatyzację procesu inwestycyjnego i zwiększa jego efektywność, pozwalając na skupienie się na zarządzaniu ryzykiem i optymalizacji portfela.

⁵⁷ "Market States" (1996) L. Chan

Współcześnie, parametry algorytmów służących do handlu są traktowane jako poufne informacje handlowe, z uwagi na ich potencjał generowania zysków. Ale w przeszłości, algorytmy wykorzystywane przez te instytucje umożliwiły identyfikację specyficznych sekwencji rynkowych, te wzorce mogą wyglądać na bardzo proste, jednak stały się podstawą obecnych które są bardziej skomplikowane.

Przykładowo, anomalie w dynamice waluty niemieckiej Marki, obecnie nieistniejącej, w relacji do dolara amerykańskiego, stały się obszarem szczególnego zainteresowania. Zidentyfikowano znaczącą korelację jednodniową, zgodnie z którą istniały poważne prawdopodobieństwo kontynuacji trendu na kolejnej sesji - to jest, jeżeli kurs obu walut w danym dniu spadał, było wysokie prawdopodobieństwo dalszego spadku w dniu następnym (lub analogicznie wzrost po wzroście). Algorytmy były także w stanie zaobserwować sytuacje, w których ceny niektórych aktywów obniżały się bezpośrednio przed publikacją istotnych danych makroekonomicznych, a następnie rosły bezpośrednio po ich prezentacji niezależnie od tego czy dane były pozytywne czy negatywne dla rynku. Ponadto, niektóre fundusze inwestycyjne implementowały strategię oparte na analizie otwarcia rynku w stosunku do zamknięcia poprzedniej sesji. Jeżeli otwarcie rynku nastąpiło znacząco poniżej zamknięcia poprzedniej sesji, fundusz często decydował się na zakup kontraktów. Analogicznie, jeżeli otwarcie rynku nastąpiło niespodziewanie wysoko ponad zamknięciem poprzedniej sesji, kontrakty były sprzedawane. W analizie technicznej taka strategia jest potocznie określana jako „wypełnianie luki cenowej”

3.6 Drzewa regresyjne oraz las losowy (Random Forest)

Drzewo decyzyjne to algorytm służący do klasyfikacji danych na podstawie serii zasad decyzyjnych. Model dzieli dane na podgrupy (węzły) i ustala wartość progową dla każdego węzła. Drzewo decyzyjne nie modyfikuje wag poszczególnych cech, a jedynie ustanawia minimalny próg wejścia. Uczenie drzewa polega na minimalizacji funkcji kary, poprzez zmianę progu poszczególnych cech. Taka metoda klasyfikacji pozwala na wydzielenie z dużej bazy danych najbardziej wartościowych cech. Funkcje kary wyliczamy na podstawie wyciągnięcia średniej z tych najbardziej wartościowych cech i porównanie ich do średniej z zbioru treningowego.

Drzewa decyzyjne mogą być używane do regresji ale są bardzo skłonne do przeuczenia. Z tego powodu w finansach wykorzystuje się bardziej skomplikowany algorytm - las losowy. Random Forest tworzy wiele drzew decyzyjnych i łączy je w celu uzyskania bardziej precyzyjnych i stabilnych prognoz. Każde drzewo jest trenowane na różnych, losowo wybranych próbkach z oryginalnego zestawu danych, co jest techniką nazywaną bootstrap (niezależny zestaw danych). Bootstrap oznacza, że jeśli nasz oryginalny zestaw danych zawiera 100 obserwacji, to stworzymy nową próbkę o rozmiarze 100, wybierając losowo obserwacje z oryginalnego zestawu danych z możliwością powtórzeń. Niektóre obserwacje mogą się powtarzać w naszej próbce, a niektóre mogą nie wystąpić wcale taka metoda zapewnia, że każde drzewo stworzone przez algorytm będzie unikalne.

Dla każdej próbki bootstrap, algorytm tworzy drzewo decyzyjne i szuka najlepszego podziału tylko wśród cech próbki. Ten proces jest powtarzany wielokrotnie, tworząc "las" drzew decyzyjnych, stąd nazwa "Random Forest". Gdy wszystkie drzewa decyzyjne są utworzone, model łączy ich prognozy. Dla problemów regresji, zwykle używa się średniej prognoz od wszystkich drzew. Najważniejszym atutem Random Forest jest odporność na

przeuczenie ze względu na to, że każde drzewo jest trenowane osobno. Nawet gdyby przeuczenie wystąpiło w części drzew zostałoby zredukowane przez średnią.

W praktycznej implementacji, algorytmy typu Random Forest często odgrywają istotną rolę w procesie selekcji cech. Ten proces polega na identyfikacji najważniejszych zmiennych wejściowych, które najskuteczniej przyczyniają się do skuteczności modelu predykcyjnego. Na przykład, w kontekście modeli regresji liniowej wykorzystywanych w strategii Arbitrage Pricing Theory (APT), Random Forest może zostać wykorzystany do identyfikacji najbardziej istotnych czynników wpływających na cenę akcji. Dodatkowo, algorytmy Random Forest mogą być wykorzystywane w procesie klasyfikacji zaawansowanych stanów rynkowych, takich jak "rynek o wysokiej zmienności" czy "rynek przekupiony". Tego rodzaju informacje pozwalają na dostosowanie strategii inwestycyjnej do panujących warunków na rynku, zwiększając tym samym jej skuteczność.

3.7 Reinforcement Learning

Uczenie przez wzmacnianie (Reinforcement Learning, RL)⁵⁸ jest gałęzią uczenia maszynowego, która rozwinęła się już kilka dekad temu, ale jej popularność znacznie wzrosła w ostatnich latach dzięki postępom w dziedzinie sztucznej inteligencji i uczenia głębokiego. Wielkie sukcesy, takie jak wykorzystanie RL przez firmę DeepMind (część Google) do stworzenia algorytmu AlphaGo, który pokonał mistrzów świata w grze Go, przyciągnęły uwagę do potencjału tej technologii. Od zawsze uważano, że Go jest zbyt złożoną grą aby algorytm mógł pokonać człowieka.

RL różni się od typowych technik klasyfikacji i regresji. Główny cel polega na uczeniu agenta (modelu), jak maksymalizować sumę nagród (funkcja celu) w dynamicznym środowisku⁵⁹, gdzie decyzje agenta wpływają na nagrodę. Celem nie jest bezpośrednio przewidywanie konkretnej wartości (jak w regresji) ani klasyfikowanie obserwacji do określonej klasy (jak w klasyfikacji), zamiast tego agent może podejmować decyzje o kupnie, sprzedaży lub zatrzymaniu akcji, a jego nagrodą może być zysk wynikający z tych decyzji.

Sposób działania algorytmu RL zaczyna się od zainicjowania środowiska w którym model będzie podejmował decyzje. Takim środowiskiem może być np. plansza gry w Go wtedy akcjami są możliwe ruchy. W strategiach inwestycyjnych model RL jest stosowany w środowisku inwestycyjnym np. na danym instrumencie. W takim wypadku środowiskiem są wszystkie dane które dodamy do algorytmu np. indykatory giełdowe, dodatkowe modele które śledzą i oceniają nastroje inwestorów, właściwie do środowiska można dodać wszystko co według osoby nadzorującej algorytm ma wpływ na cenę instrumentu.

Następnym krokiem jest interakcja modelu ze środowiskiem, agent⁶⁰ podejmuje decyzje według polityki decyzyjnej i uczy się reakcji. Każda akcja prowadzi do nowego stanu środowiska i generuje nagrodę lub karę. Początkowo polityka decyzyjna może być losowa, ale z czasem ewoluuje, aby maksymalizować nagrodę oznaczaną symbolem Q.

⁵⁸ A. Barto R. S. Sutton (1992) "Reinforcement Learning: An Introduction"

⁵⁹ Środowisko odnosi się do specyficznego kontekstu modelu RL, w którym agent podejmuje decyzje, a nie bazy danych tak jak w przypadku poprzednich modeli.

⁶⁰ <https://www.nature.com/articles/nature14236>

W strategiach inwestycyjnych używa się dodatkowych komplikacji RL. Do najpopularniejszych należą:

Deep Q Networks (DQN) – dodatkowo integruje głębokie sieci neuronowe do szybszej i dokładniejszej estymacji funkcji Q.

A3C/A2C (Asynchronous Advantage Actor-Critic): które wykorzystują równoczesne uczenie z wielu kopii środowiska, co przyspiesza proces uczenia.

W metodologii sztucznej inteligencji, uczenie przez wzmacnianie (Reinforcement Learning, RL) samo w sobie, nie prezentuje bezpośredniego fundamentu dla strategii inwestycyjnych. Potencjalnie może być jednak znakomitym mechanizmem uzupełniającym, szczególnie w kontekście operacji giełdowych.

Jednym z popularnych modeli, który wykorzystuje inną technikę uczenia, jest Chat-GPT 3. Jest to model generacji języka opracowany przez OpenAI, który został wytrenowany na ogromnej ilości danych tekstowych pochodzących z internetu. Jego zadaniem jest prognozowanie kolejnego słowa w sekwencji na podstawie poprzednich słów. W rezultacie, model jest zdolny do generowania płynnych i konkretnych odpowiedzi, sprawiając wrażenie dialogu z osobą posiadającą rozległą wiedzę. Zgodnie z najnowszymi badaniami, w ramach ewolucji algorytmu, GPT-4, proces ten został rozbudowany o dodatkowy mechanizm oparty na uczeniu przez wzmacnianie. W praktycznym zastosowaniu, użytkownik ma możliwość oceniać odpowiedzi generowane przez model językowy. Ta interakcja umożliwia modelowi dokładniejsze zrozumienie preferencji użytkownika w zakresie oczekiwanych odpowiedzi na konkretne pytania.

Analogicznie takie podejście może być zastosowane jako fundament dla strategii inwestycyjnej. Sieć neuronowa o zaawansowanej strukturze może być skonfigurowana do procesu uczenia się na podstawie danych rynkowych, zgodnie z specyficzną strategią inwestycyjną, określoną przez użytkownika. Po przeprowadzeniu procesu walidacji modelu na zestawie danych testowych, istnieje możliwość integracji modelu z rzeczywistym rynkiem i implementacji mechanizmu uczenia przez wzmacnianie na danych rynkowych w czasie rzeczywistym.

Taki system stanowi dynamiczną platformę adaptacyjną, zdolną do samodzielnego dostosowywania się do zmian warunków rynkowych i korygowania swoich hiperparametrów w oparciu o uzyskiwane wyniki. Ta ewolucja w czasie rzeczywistym pozwala na nieprzerwane optymalizowanie strategii inwestycyjnej, co może przynosić znaczne korzyści, szczególnie w kontekście dynamicznie zmieniających się warunków rynkowych.

3.8 High Frequency Trading (HFT)

HFT stanowi specyficzną odmianę strategii algorytmicznych, który zasadniczo różni się od tradycyjnych metod analizy rynkowej. Podczas gdy konwencjonalne strategie polegają na analizie makroekonomicznej, identyfikacji powtarzających się wzorców oraz na długoterminowym prognozowaniu, HFT kieruje swe działania na wykorzystanie krótkotrwałych fluktuacji cen, które pojawiają się na arkuszu zleceń. Strategie tego typu wykorzystują błyskawiczną prędkość przetwarzania danych i decyzji, umożliwiającą przez zaawansowane technologie komputerowe. Algorytmy HFT są zaprojektowane do błyskawicznego analizowania arkusza zleceń, identyfikowania niewielkich, krótkotrwałych

różnic cen i realizowania transakcji na podstawie tych różnic. Dzięki temu mogą one przeprowadzać tysiące transakcji w ułamku sekundy.

HFT budzi również pewne kontrowersje. Ze względu na skomplikowanie i szybkość tych operacji, HFT może zwiększać ryzyko destabilizacji rynków finansowych, co prowadzi do dyskusji o potrzebie dalszej regulacji tego rodzaju działalności. Spoofing i layering to strategie handlowe stosowane w kontekście High Frequency Trading (HFT), jednak ich stosowanie zostało zakazane w wielu jurysdykcjach ze względu na manipulacyjną naturę.

Spoofing jest techniką, w której operator zamieszcza zlecenie kupna lub sprzedaży znacznej ilości instrumentu finansowego np. akcji, nie mając zamiaru realizacji tego zlecenia. Celem jest manipulacja percepcją rynkową, wywołując wrażenie, pojawienia się ważnej informacji związanej ze spółką. Może to skłonić innych traderów do kupna tych akcji w oczekiwaniu na wzrost ich ceny. Gdy cena akcji wzrośnie, trader anuluje swoje pierwotne zlecenie i sprzedaje akcje po wyższej cenie. Layering polega na składaniu wielu zleceń kupna lub sprzedaży instrumentu finansowego po różnych cenach, ale bez rzeczywistego zamiaru ich realizacji. Cel jest taki sam - manipulowanie percepcją rynkową, tym razem poprzez stworzenie sztucznego obrazu zwiększenia się wolumenu na różnych poziomach cenowych. Obie strategie mają działać szczególnie przeciw inwestorom wykorzystującym algorytmy, nagłe ruchy mają spowodować błędny sygnał kupna lub sprzedaży odbierany przez algorytm, który zmienia się w ułamku sekund tak szybko jak inwestor zakupi pakiet akcji. Po zakupie zlecenie jest anulowane, a akcje inwestora wykorzystującego nielegalną strategię sprzedawane.

Obie te praktyki były wykorzystywane na rynkach i przez wiele lat przynosiły systematyczne zyski nieuczciwym inwestorom. Zostały zakazane w USA w 2010r. kiedy weszła w życie ustawa Dodd-Frank Wall Street Reform. Ustawa była następstwem pojawienia się na rynku zupełnie nowego zjawiska znanego jako „Flash crash”⁶¹. Termin "flash crash" odnosi się do nagłych, gwałtownych spadków wartości instrumentów finansowych, które zwykle są związane z błędami lub manipulacjami w handlu algorytmicznym. Fenomen ten stanowi potężne wyzwanie dla stabilności rynkowych, ponieważ nawet drobne zmiany w cenach instrumentów finansowych mogą wywołać wieloskładnikowe i nieprzewidywalne reakcje w systemie rynkowym. W strukturze rynkowej opartej na handlu algorytmicznym, pojedynczy sygnał sprzedaży, wywołany przez jeden algorytm, może wywołać serię innych sygnałów sprzedaży z innych algorytmów. Ta kaskadowa reakcja przypomina efekt domina, gdzie upadek jednego elementu prowadzi do upadku kolejnego, co prowadzi do niewspółmiernej do skali początkowego zdarzenia, zmiany na rynku. Wynikające stąd gwałtowne i często niewspółmierne do początkowych przesłanek zmiany cenowe stwarzają znaczne ryzyko dla uczestników rynku, narażając ich na potencjalnie duże straty.

W kontekście historii amerykańskich giełd, najbardziej znaczące zjawisko typu "flash crash" miało miejsce na indeksie Dow Jones Industrial Average w 2010r⁶². W kilka minut, indeks ten zanotował dramatyczny spadek wartości o ponad 1000 punktów, co stanowiło około 9% całkowitej wartości indeksu, aby w przeciągu zaledwie 5 minut odzyskać większość swojej straconej wartości. Wydarzenie to zostało przypisane brytyjskiemu traderowi, który

⁶¹ A. Kirilenko, A. Kyle (2017) "The Flash Crash: High Frequency Trading in an Electronic Market"

⁶² <https://www.bbc.com/news/explainers-51265169>

został skazany w 2015 roku za manipulację rynkową. Trader zastosował technikę znana jako "layering", przy pomocy algorytmu, składając szereg dużych, lecz niewykonanych zleceń sprzedaży. Sprawca krachu giełdowego został skazany dopiero po 5 lata co stanowi przykład nie tylko złożoności procesu identyfikacji takich działań, ale także podkreśla potrzebę ciągłego doskonalenia mechanizmów nadzoru i regulacji, aby przeciwdziałać potencjalnym manipulacjom na rynkach finansowych.

Kolejne zdarzenie oznaczone mianem "flash crash" miało miejsce w 2011⁶³ roku, na rynku surowców, konkretnie dotycząc srebra. W przeciągu zaledwie 11 minut, cena tego kruszcu zanotowała gwałtowny spadek o ponad 12%, po czym równie szybko powróciła do swojej pierwotnej wartości. Mimo analiz i dochodzeń, źródło tego zjawiska nie zostało jednoznacznie zidentyfikowane, a ewentualni sprawcy nie zostali ukarani.

Następnie, w 2016⁶⁴ roku, doszło do kolejnego gwałtownego załamania cen, tym razem na rynku walutowym, dotycząc funta brytyjskiego. Pomimo faktu, że rynek Forex jest uważany za jeden z najbardziej płynnych rynków finansowych na świecie, kurs funta spadł o ponad 6% w krótkim okresie kilku minut. W tym przypadku, istnieją podejrzenia, że oprócz potencjalnej manipulacji za pomocą nielegalnych technik handlowych, przyczyną tego gwałtownego spadku mogła być również niepewność związana z wynikiem referendum dotyczącego wyjścia Wielkiej Brytanii z Unii Europejskiej, znanego jako Brexit. Niemniej jednak, podobnie jak w przypadku flash crash z 2011 roku, nie udało się zidentyfikować i ukarać sprawców tej anomalii rynkowej.

Strategie wykorzystujące HFT nie ograniczają się wyłącznie do podejść o kwestionowanej legalności, ale obejmują również prawidłowe i dozwolone taktyki, które mają na celu generowanie zysków. Jedną z takich strategii jest arbitraż cenowy, który polega na wykorzystaniu krótkotrwałych różnic w cenach tych samych instrumentów finansowych występujących na różnych giełdach. Z uwagi na rosnącą szybkość i efektywność rynków, tradycyjny arbitraż między parami instrumentów staje się coraz mniej popularny. To jest efektem coraz większej korelacji między giełdami oraz rosnącą konkurencją między inwestorami HFT, którzy rywalizują o te same, coraz rzadsze możliwości arbitrażu.

Arbitraż między parami instrumentów, mimo że coraz rzadziej stosowany na dużych rynkach finansowych, może odgrywać kluczową rolę na rynkach nowych i rozwijających się instrumentów. Wyrazem tego są wydarzenia z lat 2016 i 2017, kiedy to rosnąca popularność kryptowalut otworzyła nowe możliwości dla strategii arbitrażowych. Podczas tej burzliwej fazy ekspansji rynku kryptowalut, różnice cenowe tych samych kryptowalut między różnymi giełdami były często znaczne, tworząc atrakcyjne okazje do arbitrażu. Inwestorzy wykorzystali te różnice do osiągnięcia zysków, kupując kryptowaluty na giełdach, na których były one notowane po niższych cenach, a następnie sprzedając je na giełdach, na których ceny były wyższe.

Równocześnie na płynnych rynkach, takich jak rynek walutowy, inwestorzy stosują strategię arbitrażu trójkątnego. Na przykład, przyjmując trzy waluty - frank szwajcarski (CHF), euro (EUR) i dolar amerykański (USD), arbitraż trójkątny mógłby przebiegać następująco:

⁶³ <https://www.cnbc.com/2011/04/26/buyers-beware-silver-crashed-11-in-24-hours.html>

⁶⁴ <https://www.bis.org/publ/mkctc10.htm>

- Inwestor wymienia początkowo posiadaną walutę, na przykład CHF, na EUR na giełdzie, gdzie kurs CHF/EUR jest najkorzystniejszy.
- Następnie, inwestor wymienia te EUR na USD na giełdzie, gdzie kurs EUR/USD jest najkorzystniejszy.
- Na koniec, inwestor wymienia te USD z powrotem na CHF na giełdzie, gdzie kurs USD/CHF jest najkorzystniejszy.

Jeśli występują odpowiednie różnice kursów na poszczególnych giełdach, takie transakcje mogą skutkować netto zyskiem dla inwestora, pomimo że z założenia transakcja zaczynała się i kończyła na tej samej walucie. Jest to doskonały przykład, jak strategie arbitrażu ewoluują, aby dostosować się do zmieniających się warunków rynkowych.

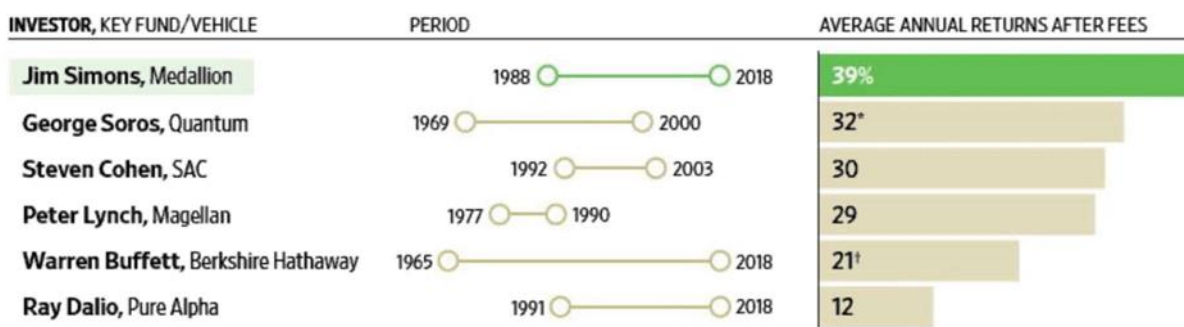
Rozdział 4: Analiza SWOT i podsumowanie

W kontekście dynamicznie rozwijającego się sektora finansowego, gdzie coraz bardziej zaawansowane technologie stają się kluczowym elementem strategii inwestycyjnych, fundusze hedgingowe oparte na sztucznej inteligencji (AI) zyskują na znaczeniu. Rozwiązania te, znane również jako fundusze algorytmiczne, wykorzystują zaawansowane techniki uczenia maszynowego i przetwarzania danych w celu analizy szerokiego zakresu informacji rynkowych i podejmowania decyzji inwestycyjnych. Analiza SWOT skupia się na unikalnych mocnych i słabych stronach tych funduszy, a także na szansach i zagrożeniach, które wynikają z ich działalności w sektorze finansowym. Celem jest zrozumienie, jak fundusze hedgingowe oparte na AI mogą przyczynić się do przyszłego kształtu przemysłu inwestycyjnego i jakie wyzwania stoją przed nimi w kontekście rosnącej konkurencji, regulacji i ewolucji technologicznej.

4.1 Zalety

Algorytmiczne fundusze inwestycyjne, reprezentowane przez takie podmioty jak Renaissance Technologies i Two Sigma, historycznie wykazały znaczącą przewagę pod względem stóp zwrotu w porównaniu do ich tradycyjnych odpowiedników. Ta różnica staje się jeszcze bardziej widoczna, gdy analizujemy wyniki na przestrzeni wielu lat. Na przykład, Medallion Fund, który jest zarządzany przez Renaissance Technologies, jest powszechnie uważany za jeden z najbardziej dochodowych funduszy hedgingowych w historii finansów. Od swojego powstania w 1988 roku do 2018 roku, fundusz generował średnią roczną stopę zwrotu wynoszącą 66%. Po odjęciu kosztów i honorarium dla funduszu średni zysk netto dla klienta Renaissance Technologies wynosił 39%.

Dla porównania, przeciętny fundusz hedgingowy według BarclayHedge⁶⁵ generuje średnio 7.2% rocznej stopy zwrotu, a TOP50 funduszy 15,5% co jest znacznie poniżej wyników osiąganych przez algorytmiczne fundusze inwestycyjne. Ponadto, wyniki osiągane przez tradycyjne fundusze inwestycyjne często są bardziej zmienne i podatne na błędy ludzkie, co może prowadzić do nieprzewidywalności i zwiększonego ryzyka dla inwestorów.



Rysunek 6 Porównanie średniego zwrotu rocznego netto największych funduszy inwestycyjnych

Źródło: <https://dnarynkow.pl/jim-simons-najskuteczniejszy-trader-w-historii-i-srednioroczna-stopa-zwrotu-rowna-66-procent/>

⁶⁵ <https://www.forbes.com/sites/jacobwolinsky/2022/08/30/heres-how-the-top-50-hedge-funds-generate-consistent-returns/?sh=4f02a2083faf>

W świetle teorii⁶⁶ Andre Kostolany'ego, która głosi, że 90% rynku giełdowego opiera się na psychologii, eliminacja czynnika ludzkiego staje się kluczowym atutem funduszy algorytmicznych, przyczyniającym się do lepszych decyzji inwestycyjnych. W kontekście naukowym, wpływ emocji i subiektywnych uprzedzeń na podejmowanie decyzji inwestycyjnych jest dobrze udokumentowany. Tzw. behawioralne skłonności, takie jak przewartościowywanie ostatnich informacji (ang. recency bias), nadmierne przekonanie we własną rację (ang. overconfidence bias) czy niechęć do realizacji strat (ang. loss aversion), mogą prowadzić do niestandardowych i często nieoptymalnych decyzji inwestycyjnych nawet przez najbardziej doświadczonych inwestorów. Poniższe punkty stanowią bezpośrednie behawioralne ryzyka inwestycyjne w których sztuczna inteligencja ma przewagę nad człowiekiem:

- **Szybkość Decyzji:** Sztuczna inteligencja jest zdolna do podejmowania decyzji w ułamkach sekund, co staje się nieocenionym atutem w dynamicznie zmieniającym się środowisku rynkowym. W dobie automatyzacji i high-frequency trading (HFT), zdolność do szybkiego reagowania na zmiany rynkowe jest kluczowa dla utrzymania konkurencyjności.
- **Bezstronność i obiektywizm:** Sztuczna inteligencja jest odporna na wpływy emocjonalne, które często wpływają na decyzje inwestycyjne podjęte przez ludzi. Decyzje podejmowane przez AI są w pełni oparte na danych i algorytmach, co eliminuje wpływ strachu, chciwości lub intuicji, które mogą skomplikować proces decyzyjny.
- **Optymalizacja:** Sztuczna inteligencja nie wymaga przerw w działaniu. Co umożliwia jej ciągłe dostosowanie się do zmieniających się warunków rynkowych i optymalizację strategii w czasie rzeczywistym.

Fundusze hedgingowe oparte na algorytmach mają zwykle szerszą dywersyfikację niż ich tradycyjne odpowiedniki, ponieważ mogą analizować i inwestować w wiele różnych instrumentów finansowych jednocześnie. Dzięki temu, mają zdolność do dynamicznego zarządzania portfelem i często nie utrzymują poszczególnych pozycji przez długi okres czasu. Algorytm pozwala na stosowanie strategii, w których instrumenty finansowe, które nie reagują zgodnie z oczekiwaniami krótko po zakupie, są sprzedawane z minimalną stratą. Daje to możliwość ograniczenia potencjalnych strat i zwiększenia stopnia dywersyfikacji. W przeciwieństwie do tradycyjnych funduszy, które często wykorzystują strategię typu Long Equity, algorytmiczne fundusze hedgingowe mają większą elastyczność i dywersyfikację, co zazwyczaj prowadzi do mniejszego ryzyka. Przykładem ryzyka związanego z brakiem dywersyfikacji jest sytuacja funduszu White Square Capital, który zbankrutował w wyniku tzw. short squeeze na akcjach firmy GameStop, przeprowadzony przez skoordynowane działania wielu inwestorów indywidualnych.

Oprócz umożliwienia ograniczenia ryzyka, algorytmiczne strategie inwestycyjne charakteryzują się możliwością przeprowadzenia tzw. backtestu. Backtesting to proces oceny wydajności strategii inwestycyjnej poprzez aplikację jej do historycznych danych rynkowych. Ta metoda analizy jest w pewnym sensie analogiczna do testowania modelu w kontekście nauk o danych, z tą różnicą, że wynik nie jest oceniany za pomocą funkcji straty, lecz jest interpretowany jako hipotetyczny zysk, który mógłby zostać uzyskany, gdyby dana strategia była zastosowana w przeszłości. Chociaż wyniki backtestingu nie są idealne i nie zawsze

⁶⁶ Andre Kostolany (1999) „Psychologia giełdy”

dokładnie odzwierciedlają rzeczywiste wyniki, dają one inwestorom i zarządzającym funduszami ważny obraz potencjalnej skuteczności strategii.

4.2 Wady

Podczas gdy algorytmy są bardzo skuteczne w analizie i przetwarzaniu dużych ilości danych, mogą mieć trudności z interpretacją niuansów i kontekstu, które mogą być istotne dla podejmowania decyzji inwestycyjnych. Fundusze w których decyzje inwestycyjne są podejmowane przez ludzi mogą lepiej interpretować i rozumieć znaczenie takich czynników jak zmiany polityczne, nastroje rynkowe czy globalne kryzysy.

Programowanie i implementacja algorytmów zdolnych do efektywnego prognozowania trendów rynkowych stanowi zadanie o wysokiej stopniu skomplikowania. Jest to zagadnienie na tyle złożone, że tylko nielicznym podmiotom udało się z powodzeniem wdrożyć taki model biznesowy i zaistnieć na rynku. Dominujące na rynku korporacje bankowe, dzięki swej długiej historii sukcesów w dziedzinie doradztwa i zarządzania kapitałem, posiadały wystarczającą podstawę do utworzenia i skutecznego zarządzania funduszami opartymi na algorytmach. Dla nowych firm, pragnących zarządzać kapitałem w taki sam sposób, wejście na rynek jest procesem zdecydowanie trudniejszym. Algorytm musi podlegać intensywnemu testowaniu pod kątem efektywności w różnych okresach zmienności rynkowej. Wymaga to nie tylko wykorzystania historycznych danych testowych, ale przede wszystkim rzeczywistych. Proces ten jest czasochłonny i generuje znaczne koszty, a dodatkowo dopiero po potwierdzeniu skuteczności modelu na podstawie realnych danych, można przystąpić do poszukiwania klientów. Zastosowanie takiego podejścia pozwala na zminimalizowanie ryzyka, ale jednocześnie generuje istotne wyzwania związane z koniecznością początkowego nakładu finansowego oraz zyskania zaufania potencjalnych inwestorów.

Działania algorytmów mogą być trudne do zrozumienia dla inwestorów, zwłaszcza jeśli używane są skomplikowane techniki, takie jak głębokie uczenie się. To może prowadzić do braku zaufania lub niezrozumienia decyzji inwestycyjnych podejmowanych przez algorytmy. Nie wpływa to oczywiście na wynik inwestycyjny, jedynie na biznesowe powodzenie funduszu.

Proces wdrażania i utrzymania skomplikowanych systemów algorytmicznych może implikować znaczne nakłady kapitałowe. Może to obejmować koszty związane z akwizycją i utrzymaniem specjalistycznej infrastruktury technologicznej, angażowaniem ekspertów w dziedzinie danych i algorytmów, a także kontynuowanym nadzorem i modernizacją algorytmów. Takie koszty mogą znacznie przekroczyć wydatki związane z tradycyjnym zarządzaniem funduszem. Z tego powodu struktura kosztowa funduszy opartych na algorytmach znacząco odbiega od tradycyjnych modeli opłat, takich jak prowizja 1% od całego obracanego kapitału oraz 5% z zysku. Przykładowo, Fundusz Medallion, mimo generowania 66% zysku w skali roku, był w stanie wypłacić inwestorom jedynie 39% tego zysku. Aż 27% stanowiły koszty operacyjne, które obejmowały oczywiście zarówno koszty funkcjonowania funduszu, jak i generowany przez niego zysk. Taka struktura kosztów może stanowić istotne wyzwanie dla rentowności netto inwestycji w fundusze algorytmiczne.

4.3 Szanse

Stałe innowacje w sferze mocy obliczeniowej, przechowywania danych oraz ulepszania technologii algorytmicznej stwarzają istotne potencjał dla dalszej optymalizacji efektywności funduszy algorytmicznych. Wzrost możliwości obliczeniowych, zarówno w kontekście

przetwarzania w chmurze, jak i w stacjonarnych systemach komputerowych, umożliwia szybszą i dokładniejszą analizę danych. Rozwój technologii komputerów kwantowych stanowi kolejne pole możliwości dla funduszy algorytmicznych. Komputery kwantowe, które operują na poziomie cząsteczek subatomowych, mogą oferować ogromne przyspieszenie obliczeń, co może przyczynić się do budowy głębszych i bardziej złożonych modeli uczenia maszynowego. Adaptacja nowatorskich algorytmów z różnych dziedzin nauki i technologii stanowi istotny potencjał rozwoju dla funduszy algorytmicznych. Na przykład techniki uczenia przez wzmocnienie (reinforcement learning), które pierwotnie były stosowane w obszarach takich jak przetwarzanie języka naturalnego (NLP) i gry komputerowe, zyskują obecnie na znaczeniu w kontekście modelowania strategii inwestycyjnych.

Wzrost stopnia cyfryzacji społeczeństwa prowadzi do generowania i gromadzenia znacznej ilości danych, które stanowią główne źródło informacji dla algorytmów. Algorytmy zyskują coraz szerszy dostęp do zasobów informacyjnych, umożliwiając precyzyjne modelowanie zjawisk rynkowych, prognozowanie trendów i optymalizację strategii inwestycyjnych. Zmiany w regulacjach finansowych mające na celu zwiększenie przejrzystości rynków finansowych dostarczają kolejnego rodzaju cennych danych. Zgodnie z zasadami przejrzystości, spółki giełdowe oraz instytucje finansowe są zobligowane do udostępniania coraz większej ilości informacji, co tworzy dodatkowy zasób danych giełdowych, które są wykorzystywane w opracowywaniu strategii inwestycyjnych.

4.4 Zagrożenia

Gwałtowny postęp w dziedzinie sztucznej inteligencji zaowocował pojawieniem się grupy badaczy dążących do wprowadzenia regulacji ograniczających ten rozwój. Ta inicjatywa wynika z konieczności zarządzania możliwym ryzykiem związanym z niewłaściwym lub niekontrolowanym wykorzystaniem AI. Wśród rozważanych restrykcji znalazłyby się propozycje takie jak zakaz podłączania systemów opartych na sztucznej inteligencji do Internetu lub restrykcje w zakresie szkolenia AI w dziedzinie psychologii ludzkiej. Takie regulacje mogą mieć istotne implikacje dla funduszy algorytmicznych. Na przykład, restrykcje dotyczące dostępu do Internetu mogą ograniczać zdolność algorytmów do gromadzenia i analizowania informacji z szerokiego spektrum źródeł, co jest kluczowe dla tworzenia skutecznych strategii inwestycyjnych. Z kolei ograniczenia w zakresie edukacji AI w dziedzinie psychologii ludzkiej mogą utrudniać modelowanie i prognozowanie zachowań rynkowych, które często są napędzane przez czynniki psychologiczne.

Algorytmiczne fundusze są ściśle powiązane z postępem technologicznym, szczególnie w dziedzinie sztucznej inteligencji. To powiązanie wpływa na dynamikę rozwoju funduszy - jeżeli rozwój AI zostanie opóźniony lub spowolniony, wpłynie to na zdolność do innowacji i poprawy ich wydajności.

Fundusze algorytmiczne ze względu na istotne uzależnienie od technologii są podatne na ataki cybernetyczne. Takie naruszenia mogą prowadzić do poważnych konsekwencji na przykład utrata danych, nieautoryzowany dostęp do systemów lub manipulacja algorytmami. Kluczowe znaczenie w tej kwestii ma ochrona kodu algorytmu, który jest najcenniejszym aktywem dla funduszy algorytmicznych. W przeciwieństwie do klasycznych funduszy, kompetencji osoby zarządzającej nie da się "ukraść", kod algorytmu jest podatny na wyniesienie i przekazanie do konkurencji. Jeżeli konkurencja pozyska wiedzę na temat strategii inwestycyjnych funduszu, może to wykorzystać na swoją korzyść, narażając tym

samym fundusz na znaczne straty finansowe. Takie zagrożenie podkreśla nie tylko konieczność stosowania najnowszych standardów cyber-bezpieczeństwa, ale także budowanie kultury bezpieczeństwa wśród pracowników i partnerów funduszu. Zabezpieczenie tajemnic handlowych i innych cennych aktywów intelektualnych jest kluczowym elementem zarządzania ryzykiem w kontekście funduszy algorytmicznych.

Konkurencja na rynku funduszy algorytmicznych nasila się, co jest wynikiem rosnącej liczby podmiotów z sektora finansowego i technologicznego inwestujących w rozwój własnych systemów algorytmicznych. Ta zwiększona konkurencja wywiera presję na marże i może oznaczać mniejszą pulę dostępnych zasobów przeznaczanych na innowacje, co potencjalnie może utrudnić utrzymanie przewagi konkurencyjnej. Wraz z wzrostem ilości transakcji generowanych przez algorytmy rynek staje się coraz bardziej skomplikowany, charakteryzując się większą zmiennością i nagłymi skokami cen. Co doprowadziło do pojawienia się zupełnie nowego zjawiska na rynkach zwanego „Flash Crash”

Warto również zwrócić uwagę na potencjalne zagrożenia związane z nielegalnymi działaniami na rynkach finansowych, takimi jak te związane z wysokoczęstotliwościowym handlem. Mimo że te praktyki są nielegalne, mogą one nadal występować i wpływać na działanie zautomatyzowanych funduszy hedgingowych.

Podsumowanie

Rozwój technologiczny przyczynia się do wzrostu wyników inwestycyjnych osiąganych przez fundusze inwestycyjne które stosują strategie inwestycyjne bazujące na algorytmach. Postęp w dziedzinie przetwarzania danych, przepustowości sieci oraz moc obliczeniowa procesorów mają bezpośredni wpływ na zdolność algorytmów do przetwarzania, analizy i reagowania na zmienność rynkową w czasie rzeczywistym. Rozwój technologii umożliwia przetwarzanie danych w skali, która była nieosiągalna zaledwie kilka lat temu. Możliwość przechowywania i analizy ogromnych ilości danych w chmurze, a następnie szybkiego dostępu do nich, pozwala algorytmom na bardziej efektywne identyfikowanie trendów, anomalii i wzorców, co zwiększa ich efektywność. Nieprzerwany rozwój technologiczny jest kluczowym czynnikiem, który stanowi o przewadze strategii inwestycyjnych opartych na algorytmach od klasycznych odpowiedników takich jak analiza fundamentalna lub modele statystyczne.

Fundusze algorytmiczne mimo wielu zalet, generują potencjalne ryzyko destabilizacji rynków finansowych. Gwałtowne zmiany w cenach aktywów, które mogą wynikać z nagłych i masowych decyzji transakcyjnych podejmowanych przez algorytmy, prowadzą do zjawisk takich jak „flash crash”. Zdolność algorytmów do reagowania w ekstremalnie krótkim czasie, połączona z potencjałem dla wielkoskalowej koordynacji decyzji inwestycyjnych, może prowadzić do bardzo wysokiej zmienności rynkowej uniemożliwiającej inwestowanie na giełdzie dla tradycyjnych inwestorów. Ponadto, przewaga technologiczna funduszy algorytmicznych może prowadzić do zjawiska, w którym bariera wejścia na rynek staje się zbyt wysoka dla nowych funduszy hedgingowych. Wymóg dysponowania zaawansowaną infrastrukturą technologiczną i specjalistyczną wiedzą może ograniczyć konkurencję, skupiając kontrolę nad rynkiem w rękach nielicznych podmiotów, które dysponują niezbędnymi zasobami.

Na podstawie bieżących badań i trendów technologicznych, nie ulega wątpliwości, że sztuczna inteligencja zyska na znaczeniu w kontekście analizy rynkowej. Modele algorytmiczne przynoszą wiele korzyści, są dokładniejsze, powtarzalne, bardziej dopasowane do obecnej dynamiki rynkowej. Przy czym nie przynoszą takich zagrożeń jak agresywne strategie HFT lub wyszukiwanie wzorców cenowych. Biorąc pod uwagę te przewagi, przyszłość analizy rynkowej zdecydowanie zależy od rozwoju sztucznej inteligencji.

Pod znakiem zapytania stoją jedynie fundusze algorytmiczne które w pełni są zarządzane przez algorytmy ,w tym kontekście, regulacje mogą odegrać kluczową rolę w zarządzaniu ryzykiem związanym z rozwojem sztucznej inteligencji. Pytania dotyczące etyki, bezpieczeństwa i prywatności, które wiążą się z wykorzystaniem AI, mogą prowadzić do wprowadzenia nowych ograniczeń i regulacji. Zastosowanie AI w handlu instrumentami finansowymi jest obszarem szczególnego zainteresowania dla regulatorów, którzy będą dążyć do zminimalizowania potencjalnych negatywnych skutków dla stabilności rynków i uczciwości konkurencji. Równoważenie dynamiki innowacji i dostosowywania prawa do najnowszych technologii będzie kluczowym wyzwaniem dla regulatorów. Wprowadzanie zbyt rygorystycznych ograniczeń może całkowicie zatrzymać innowacje i konkurencyjność, podczas gdy zbyt luźne regulacje mogą stwarzać potencjalne zagrożenia dla stabilności rynkowej i integralności systemu finansowego.

Podsumowując, perspektywa przyszłości algorytmicznych funduszy hedgingowych zdaje się istotnie zależeć od globalnego nastawienia do sztucznej inteligencji. Z tego punktu widzenia, decyzje regulatorów zdecydowanie ukształtują krajobraz finansowy w nadchodzących latach. Jeżeli regulacje przesuną się w stronę liberalizmu, a więc przyjęcia podejścia skrajnie wolnorynkowego, prawdopodobne jest, że większość decyzji inwestycyjnych w ciągu kilku lat będzie podejmowana przez zaawansowane algorytmy. Tym samym, konsekwencją takiej sytuacji byłaby dominacja algorytmicznych funduszy hedgingowych na rynku finansowym.

Alternatywnie, w przypadku, gdy regulacje ograniczą stosowanie algorytmów w handlu instrumentami finansowymi np. tylko do analizy, rynek może utrzymać swój obecny stan równowagi. Takie podejście regulatorów pozwoliłoby na koegzystencję różnych form inwestycji: indywidualnej, tradycyjnych funduszy, jak również tych opartych na algorytmach. Podkreślić należy, że ewolucja sektora finansowego jest nieodłącznie związana z rosnącym wpływem technologii informacyjnych i komunikacyjnych, w tym sztucznej inteligencji. Niezależnie od tego, jakie regulacje zostaną wprowadzone, sektor finansowy będzie nadal ewoluować w kierunku większego wykorzystania technologii. Dlatego niezbędne jest przeprowadzenie dalszych badań na temat optymalnej roli algorytmów w handlu finansowym, aby zminimalizować ryzyko i maksymalizować korzyści dla wszystkich uczestników rynku.

Bibliografia

- 1) P. Ceruzzi (1998) "A History of Modern Computing" - The MIT Press
- 2) Goodfellow, I., Bengio, Y., & Courville, A. (2016). „Deep Learning" - The MIT Press
- 3) Aureliena Geron (2019) "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" - O'Reilly Media
- 4) Nguyen, H. T., & Walker, E. A. (2006) "A First Course in Fuzzy Logic" CRC Press
- 5) Chlistalla, M. (2011) "Algorithmic Trading" Deutsche Bank Research
- 6) K. Jajuga, T. Jajuga (2020) „Inwestycje Instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa" Wydawnictwo Naukowe PWN
- 7) Gregory Zuckerman (2019) „Człowiek, który rozwiązał rynek: Jak Jim Simons zapoczątkował rewolucję quantów" MT Biznes
- 8) Bholat, D., Brookes, J., Cai, C., Grundy, K., & Lund, J. (2020) "Machine learning explainability in finance: an application to default risk analysis" Bank of England working papers.
- 9) Safavian, S. R., & Landgrebe, D. (1991) "A survey of decision tree classifier methodology. IEEE transactions on systems, man, and cybernetics" IOP Publishing
- 10) Cortes, C., & Vapnik, V. (1995) "Support-vector networks and Machine learning" IOP Publishing
- 11) James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013) "An introduction to statistical learning" Springer
- 12) Bellman, R. (2015) "Adaptive control processes: a guided tour" Princeton university press.
- 13) Fendoglu, S. (2017) "Credit cycles and capital flows: Effectiveness of the macroprudential policy framework in emerging market economies" Journal of Banking & Finance
- 14) Tapscott, D., & Tapscott, A. (2016). "Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World" Penguin Publishing Group
- 15) Mougayar, W. (2016). "The Business Blockchain: Promise, Practice, and Application of the Next Internet Technology" Wiley
- 16) Kevin P. Murphy (2020) „Machine Learning: A Probabilistic Perspective" The MIT Press
- 17) Max Kuhn, Kjell Johnson (2013) „Applied Predictive Modeling" Springer Science & Business Media
- 18) Guyon, I., & Elisseeff, A. (2003), "An Introduction to Variable and Feature Selection"
- 19) Stanford University "Convolutional neural Networks for Visual recognition"
<http://cs231n.github.io/optimization-1/>
- 20) T. Hastie, R. Tibshirani, J. Friedman (2017) "The Elements of Statistical learning" Springer
- 21) Mark S. Joshi (2003) „The Concepts and Practice of Mathematical Finance" Cambridge University Press
- 22) Ł. Wordliczek (2021) "Wykorzystanie sztucznych sieci neuronowych" Uniwersytet Jagielloński w Krakowie

- 23) L. Chan (1996) "Market States" Journal of Empirical Finance
- 24) Barto R. S. Sutton (1992) "Reinforcement Learning: An Introduction" The MIT Press
- 25) Kirilenko, A. Kyle (2017) "The Flash Crash: High Frequency Trading in an Electronic Market" The Journal of Finance
- 26) <https://www.bbc.com/news/explainers-51265169>
- 27) <https://www.cnbc.com/2011/04/26/buyers-beware-silver-crashed-11-in-24-hours.html>
- 28) <https://www.bis.org/publ/mktc10.htm>
- 29) <https://www.forbes.com/sites/jacobwolinsky/2022/08/30/heres-how-the-top-50-hedge-funds-generate-consistent-returns/?sh=4f02a2083faf>
- 30) Andre Kostolany (1999) „Psychologia giełdy” Europa

Spis rysunków

Rysunek 1: Uproszczona wizualizacja techniki szkolenia algorytmu, opracowanie własne.

Rysunek 2: Wizualizacja rozkładu normalnego, opracowanie własne.

Rysunek 3: Poglądowy rozkład danych w modelu, opracowanie własne.

Rysunek 4: Wizualizacja regresji liniowej pl.wikipedia.org/wiki/Regresja_liniowa

Rysunek 5: Uproszczony model działania sieci neuronowej
<https://pl.wikipedia.org/wiki/Perceptron>

Rysunek 6: Porównanie średniego zwrotu rocznego netto największych funduszy inwestycyjnych <https://dnarynkow.pl/jim-simons-najskuteczniejszy-trader-w-historii-i-srednioroczna-stopa-zwrotu-rowna-66-procent/>