



Uniwersytet im. A. Mickiewicza w Poznaniu

Wydział Matematyki i Informatyki

Praca magisterska

**Wykorzystanie metod uczenia
maszynowego w predykcji kursów akcji
na giełdzie papierów wartościowych.**

**The use of machine learning methods in predicting stock
prices on the stock exchange.**

Autor: Tomasz Łysiak

Numer albumu: 407259

Kierunek: Informatyka, Specjalność: Systemy inteligentne

Promotor:

prof. UAM dr hab. Krzysztof Jassem

Spis treści

Wstęp	8
1. Wprowadzenie do predykcji kursów akcji na giełdzie papierów wartościowych	10
1.1. Wstęp	10
1.2. Analiza fundamentalna	11
1.3. Analiza techniczna	12
1.4. Wybór strategii	13
1.5. Podsumowanie	13
2. Sposoby zdobywania wiedzy przez systemy uczenia maszynowego	15
2.1. Wstęp	15
2.2. Uczenie nadzorowane	16
2.3. Uczenie nienadzorowane	17
2.4. Uczenie częściowo nadzorowane	18
2.5. Uczenie przez wzmacnianie	19
2.6. Wybór sposobu nauki	21
2.7. Podsumowanie	21
3. Metody uczenia maszynowego	23
3.1. Wstęp	23
3.2. Regresja liniowa	24
3.3. Regresja wielomianowa	25
3.4. Regresja logistyczna	27
3.5. Algorytm k najbliższych sąsiadów	28
3.6. Maszyna wektorów nośnych	30
3.7. Drzewa decyzyjne i lasy losowe	31

3.8.	Sztuczne sieci neuronowe	33
3.8.1.	Podstawowe informacje o sztucznych sieciach neuronowych . . .	33
3.8.2.	Jednokierunkowe sieci neuronowe	36
3.8.3.	Rekurencyjne sieci neuronowe	38
3.9.	Wybór metody uczenia maszynowego	40
3.10.	Podsumowanie	41
4.	Wykorzystane narzędzia	42
4.1.	Wstęp	42
4.2.	Platformy programistyczne przeznaczone dla sztucznych sieci neuro- nowych	42
4.2.1.	Porównanie TensorFlow i PyTorch	42
4.2.2.	Wybór platformy	46
4.3.	Biblioteki pomocnicze	46
4.3.1.	Keras	46
4.3.2.	Scikit-learn	47
4.3.3.	Pandas	47
4.3.4.	Numpy	48
4.3.5.	Wybór bibliotek	49
4.4.	Podsumowanie	49
5.	Hipoteza błędzenia losowego oraz aktualne wyniki badań	50
5.1.	Wstęp	50
5.2.	Czy rynek jest nieprzewidywalny?	50
5.3.	Czy poprzez wykorzystanie metod uczenia maszynowego można osią- gnąć lepsze wyniki w wąskim horyzoncie czasowym niż w przypadku losowego wybierania spółek do portfolio?	51
5.4.	Jakie algorytmy są najefektywniejsze i jaka jest ich skuteczność? . . .	51
5.5.	Podsumowanie	53
6.	Przeprowadzone eksperymenty	54
6.1.	Wstęp	54
6.2.	Cel	54
6.3.	Założenia	55

6.4. Wyniki	63
6.5. Wnioski	70
6.6. Podsumowanie	72
7. Podsumowanie	73
Bibliografia	75
Spis ilustracji	78
Spis tabel	79

Poznań, dnia

Oświadczenie

Ja, niżej podpisany Tomasz Łysiak student Wydziału Matematyki i Informatyki Uniwersytetu im. Adama Mickiewicza w Poznaniu oświadczam, że przedkładaną pracę dyplomową pt: *"Wykorzystanie metod uczenia maszynowego w predykcji kursów akcji na giełdzie papierów wartościowych."* napisałem samodzielnie. Oznacza to, że przy pisaniu pracy, poza niezbędnymi konsultacjami, nie korzystałem z pomocy innych osób, a w szczególności nie zlecałem opracowania rozprawy lub jej części innym osobom, ani nie odpisywałem tej rozprawy lub jej części od innych osób.

Oświadczam również, że egzemplarz pracy dyplomowej w wersji drukowanej jest całkowicie zgodny z egzemplarzem pracy dyplomowej w wersji elektronicznej.

Jednocześnie przyjmuję do wiadomości, że przypisanie sobie, w pracy dyplomowej, autorstwa istotnego fragmentu lub innych elementów cudzego utworu lub ustalenia naukowego stanowi podstawę stwierdzenia nieważności postępowania w sprawie nadania tytułu zawodowego.

[]* - wyrażam zgodę na udostępnianie mojej pracy w czytelni Archiwum UAM

[]* - wyrażam zgodę na udostępnianie mojej pracy w zakresie koniecznym do ochrony mojego prawa do autorstwa lub praw osób trzecich

*Należy wpisać TAK w przypadku wyrażenia zgody na udostępnianie pracy w czytelni Archiwum UAM, NIE w przypadku braku zgody. Niewypełnienie pola oznacza brak zgody na udostępnianie pracy.

.....

(czytelny podpis studenta)

Streszczenie

Celem niniejszej pracy jest wykorzystanie metod uczenia maszynowego w predykcji kursów akcji na giełdzie papierów wartościowych. W pierwszej kolejności następuje wstęp do lokowania kapitału w akcje spółek. Następnie omówione zostają dwie najpopularniejsze analizy i strategie wspierające decyzje o zakupie lub sprzedaży instrumentów finansowych oraz wybór jednej z nich na potrzeby przeprowadzenia eksperymentów. W dalszej części pracy dokonuje się wyboru sposobu zdobywania wiedzy przez opracowywany algorytm oraz selekcja najskuteczniejszej metody uczenia maszynowego z perspektywy rozpatrywanego problemu. Ponadto następuje dobranie odpowiednich narzędzi umożliwiających zrealizowanie założeń projektowych. Po dokładnym zdefiniowaniu hipotez dokonuje się ich weryfikacji w oparciu o wyniki otrzymane w ramach przeprowadzonych badań oraz wyciągniętych na ich podstawie wniosków.

Słowa kluczowe: uczenie maszynowe, giełda papierów wartościowych, przewidywanie kursów akcji.

Abstract

The purpose of this work is to use machine learning methods to predict stock prices on the stock exchange. The thesis starts with an introduction to investing capital in company stocks. It then goes on to discuss the two most popular analyses, along with their associated strategies, that support financial instrument buying and selling decisions. From this discussion, the thesis identifies the most suitable analysis and associated strategy for experimentation. In the further part of the work a choice of the technique of acquiring knowledge by the algorithm is made and the optimal machine learning method from the perspective of the problem under consideration is chosen. During the next step the right tools are selected to achieve the project goals. After accurate formulation of the hypothesis they are verified in accordance to the results obtained as part of the research and the conclusions drawn from them.

Key words: machine learning, stock exchange, predicting stock prices.

Wstęp

Od wielu lat predykcja kursów akcji na giełdzie papierów wartościowych stanowi bardzo popularny i złożony temat, który stanowi źródło licznych kontrowersji. Pomimo zaangażowania niezliczonej ilości środków oraz czasu najwybitniejszych ekspertów na świecie problem nadal pozostaje nierozwiązany. Wciąż trwają ciągle dyskusje oraz spory dotyczące wyboru najlepszych i najefektywniejszych technik wykorzystywanych do przewidywania trendów na giełdzie. Jedni są zwolennikami analizy fundamentalnej i długoterminowego lokowania funduszy w aktywa, inni z kolei preferują analizę techniczną i spekulowanie w ramach tzw. „daytradingu”, czyli zakupu i sprzedaży walorów tego samego dnia działania giełdy. Nie brakuje również opinii, że rynek jest całkowicie nieprzewidywalny i jest to przypadek błędzenia losowego. Burton Gordon Malkiel w jednej z najwybitniejszych książek dotyczących gry na giełdzie pt. „Błądząc po Wall Street. Sprawdzona strategia skutecznego inwestowania” pokusił się nawet o następujące stwierdzenie:

“Małpa, rzucając lotkami, jest w stanie dobrać w przypadkowy sposób do portfela inwestycyjnego lepsze spółki, niż specjaliści po miesiącach analiz.” [13]

Czy rzeczywiście rynek jest nieprzewidywalny? Czy poprzez wykorzystanie metod uczenia maszynowego można osiągnąć lepsze wyniki w wąskim horyzoncie czasowym, niż w przypadku losowego wybierania spółek do portfolio? Jeżeli tak, to jakie metody uczenia maszynowego są najefektywniejsze i jaka jest ich skuteczność? Na wszystkie te pytania postaram się zawrzeć odpowiedź w niniejszej pracy. W pierwszej kolejności nastąpi wprowadzenie do problematyki lokowania kapitału na giełdzie papierów wartościowych oraz przegląd aktualnie stosowanych metod predykcji kursów akcji. W oparciu o bieżące badania nastąpi wybór odpowiedniej strategii inwestycyjnej.

Następnie zostanie dokonana analiza oraz wybór najatrakcyjniejszej metody uczenia maszynowego z punktu widzenia przyjętej strategii inwestycyjnej. W kolejnym kroku nastąpi selekcja narzędzi umożliwiających przeprowadzenie zaplanowanych wcześniej eksperymentów. W rozdziale pt. "Hipoteza błędzenia losowego oraz aktualne wyniki badań" zostaną szerzej omówione zagadnienia stanowiące cel przeprowadzanych eksperymentów. W końcowym etapie pracy nastąpi prezentacja otrzymanych rezultatów oraz zdefiniowanie na ich podstawie ostatecznych wniosków z przeprowadzonych badań.

Rozdział 1

Wprowadzenie do predykcji kursów akcji na giełdzie papierów wartościowych

Bieżący rozdział stanowi wstęp do problematyki lokowania kapitału na giełdzie papierów wartościowych. Omówione zostały w nim dwie najpopularniejsze strategie spekulacyjne oraz inwestycyjne umożliwiające wypracowanie zysku na podstawie zmian wartości instrumentów finansowych. Po dokonaniu analizy obu technik nastąpił wybór jednej z nich na potrzeby generowania predykcji notowań spółek zamknięcia kolejnego dnia działania giełdy w oparciu o metody uczenia maszynowego.

1.1. Wstęp

Giełda papierów wartościowych jest rynkiem finansowym umożliwiającym zakup i sprzedaż aktywów zgodnie z określonymi regulacjami. Lokowanie funduszy w instrumenty finansowe umożliwia osiągnięcie zysków związanych ze zmianami ich wartości przy jednoczesnej akceptacji ponadprzeciętnego ryzyka utraty kapitału. Generowanie zysków w dłuższej perspektywie czasu wymaga wypracowania odpowiedniej strategii inwestycyjnej. Do dzisiejszych czasów nie powstała metoda umożliwiająca osiąganie zysków bez ponoszenia ryzyka. Wybór odpowiednich technik ma jedynie na celu zmaksymalizowanie prawdopodobieństwa wygranej. Wśród najwybitniejszych ekspertów możemy wyróżnić podział na dwie grupy – zwolenników analizy

technicznej opierających się na spekulacji oraz wielbicieli analizy fundamentalnej lokujących kapitał w inwestycjach długoterminowych. Poniżej scharakteryzowano oba podejścia i dokonano wyboru rodzaju analizy danych na potrzeby wykorzystania metod uczenia maszynowego w predykcji kursów akcji.

1.2. Analiza fundamentalna

Analiza fundamentalna polega na inwestowaniu w wartość przedsiębiorstwa. Opiera się ona na badaniu spółki pod względem jej płynności finansowej, biznesplanu, wizji, marketingu, a także inflacji, sytuacji gospodarczej na świecie i wielu innych czynników wpływających na wartość przedsiębiorstwa w szerokim horyzoncie czasowym. Legendarny inwestor Warren Buffet nazywany przez wielu "wyrocznią z Omaha" jest wielkim zwolennikiem analizy fundamentalnej i długoterminowego lokowania funduszy w spółki. Podejście to nazywa "kup i trzymaj". Jego strategię idealnie odzwierciedlają słowa:

"Najlepszym przedziałem czasowym trzymania akcji jest wieczność." [2]

Ponadto wielokrotnie podkreśla, że:

"Nigdy nie próbuję zarabiać na giełdzie. Kupuję z założeniem, że następnego dnia mogą zamknąć giełdę i nie otworzą jej przez pięć lat." [2]

Cytaty te pozwalają nam zrozumieć, o jak szerokim horyzoncie czasowym należy myśleć podczas wykorzystywania analizy fundamentalnej do inwestowania na giełdzie papierów wartościowych. W najwybitniejszej książce Benjamina Grahama pt. "Inteligentny inwestor", o której Warren Buffett wypowiada się w sposób następujący:

„Pierwszy raz przeczytałam tę książkę na początku 1950 roku, gdy miałem 19 lat. Wydawało mi się wówczas, że to najlepsza książka o inwestowaniu, jaką kiedykolwiek napisano. Dzisiaj nadal tak uważam”. [9]

zostały zawarte trzy najważniejsze przesłanki dotyczące inwestowania opartego na analizie fundamentalnej:

- zanim kupisz akcje, musisz starannie przeanalizować firmę i zbadać solidność prowadzonej przez nią działalności,

- musisz świadomie zabezpieczać się przed poważnymi stratami,
- musisz starać się o uzyskiwanie "zadowalających" a nie nadzwyczajnych wyników.

1.3. Analiza techniczna

Analiza techniczna opiera się na badaniu historycznych notowań kursów akcji w celu przewidywania przyszłych trendów cenowych głównie w wąskim horyzoncie czasowym. Jest to agresywna strategia lokowania funduszy w aktywa, mająca na celu wypracowanie ponadprzeciętnych zysków przy jednoczesnym ponoszeniu ponadprzeciętnego ryzyka. Podejście to opiera się na trzech podstawowych przesłankach zawartych w książce Johna J. Murphy'ego pt. "Analiza techniczna rynków finansowych":

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

Zgodnie z tymi założeniami aktualna cena akcji odzwierciedla wszystkie kształtujące jej wartość czynniki, a notowania spółek przyjmują pewne okresowe kierunki wzrostu, spadku lub stagnacji. Istnieje większe prawdopodobieństwo kontynuacji trendu niż jego odwrócenia. Ponadto zachowanie rynku charakteryzuje określona powtarzalność. Zdaniem wybitnego analityka giełdowego Jessego Livermore'a:

"Nie ma nowości w spekulacji. To co wydarzyło się wcześniej, powtórzy się jutro i pojutrze, raz po raz. To efekt niezmiennych zachowań ludzi obecnych na rynku. Jestem tego pewien." [1]

Wykorzystując analizę techniczną do spekulacji na giełdzie papierów wartościowych mamy do wyboru wiele metod i strategii inwestycyjnych. Najpopularniejszymi z nich są:

- Transakcje w reakcji na sytuację rynkową — metoda opierająca się na podejmowaniu decyzji o nabywaniu lub zbywaniu instrumentów finansowych na podstawie sygnałów płynących z rynku. Podczas zwyżkowania kursów spółki następuje jej zakup i sprzedaż w momencie obrania kierunku spadkowego.

- Selektywność krótkoterminowa — technika polegająca na zakupie akcji o podwyższonych zyskach. Zgodnie z założeniem zwiększony popyt jest prognozą trendu wzrostowego.
- Selektywność długoterminowa - strategia wykorzystująca długotrwałe wykonywanie notowań spółki o przewidywanym dużym potencjale wzrostu w najbliższym okresie.

1.4. Wybór strategii

Zarówno analiza fundamentalna wraz z długoterminowym lokowaniem funduszy w spółki, jak i analiza techniczna oparta na kapitale spekulacyjnym, są z sukcesem wykorzystywane przez najwybitniejszych inwestorów na świecie. Nadszedł jednak czas na wybór jednej z nich na potrzeby zautomatyzowania procesu generowania predykcji przyszłych notowań spółek w oparciu o wykorzystanie metod uczenia maszynowego. Pomimo wielu atutów analizy fundamentalnej postanowiono porzucić tę strategię w dalszych badaniach. Założono, że podczas inwestowania w akcje spółek w dłuższym horyzoncie czasu, istnieje zbyt wiele czynników kształtujących ceny walorów, a bez dostępu do informacji, które ich dotyczą, eksperymenty z góry skazane są na niepowodzenie. Ponadto dzięki przyjęciu strategii krótkoterminowej możliwe jest szybkie zamknięcie pozycji¹ i zminimalizowanie poniesionych strat. Jest to niewątpliwie duży atut tego podejścia w porównaniu z analizą fundamentalną, gdzie czasami podczas bessy lub krachu większa część naszej łącznej alokacji kapitału w aktywa może zostać utracona.

1.5. Podsumowanie

Bieżący rozdział stanowił wstęp do lokowania kapitału na giełdzie papierów wartościowych. Omówione zostały dwie najpopularniejsze analizy oraz związane z nimi strategie dotyczące zakupu oraz sprzedaży akcji spółek. Analiza fundamentalna opiera się na dokładnym badaniu określonych przedsiębiorstw oraz wszystkich czynników mających wpływ na ich wartość w celu podjęcia długoterminowych in-

¹Zamykanie pozycji – proces polegający na sprzedaży dostępnych aktywów

westycji. Metoda ta charakteryzuje się obniżonym ryzykiem i niższą stopą zwrotu w porównaniu z analizą techniczną i spekulacją, która stanowi strategię agresywną umożliwiającą uzyskanie wysokich profitów w krótkim odstępie czasu. Wymaga ona jednak podjęcia ponadprzeciętnego ryzyka utraty kapitału. Po dokonaniu analizy obu podejść ostatecznie postanowiono w dalszych eksperymentach realizować strategię inwestycji krótkoterminowych opartych jedynie na historycznych notowaniach spółek. W związku z tym od teraz w dalszej części przyjęte zostały założenia analizy technicznej. Zgodnie z tym rynek dyskontuje wszystko, ceny podlegają trendom, a historia się powtarza. Na podstawie tej wiedzy postaramy się przewidzieć kursy akcji zamknięcia kolejnego dnia działania giełdy w oparciu o strategię transakcji w reakcji na sytuację rynkową.

Rozdział 2

Sposoby zdobywania wiedzy przez systemy uczenia maszynowego

W niniejszym rozdziale nastąpi wprowadzenie do wiedzy o uczeniu maszynowym oraz zostanie dokonana analiza najczęściej wykorzystywanych technik zdobywania wiedzy przez systemy informatyczne w oparciu o dostępne dane. Następnie omówiona zostanie próba znalezienia najlepszej metody z punktu widzenia generowania predykcji kursów akcji w oparciu o analizę techniczną oraz strategię transakcji w reakcji na sytuację rynkową.

2.1. Wstęp

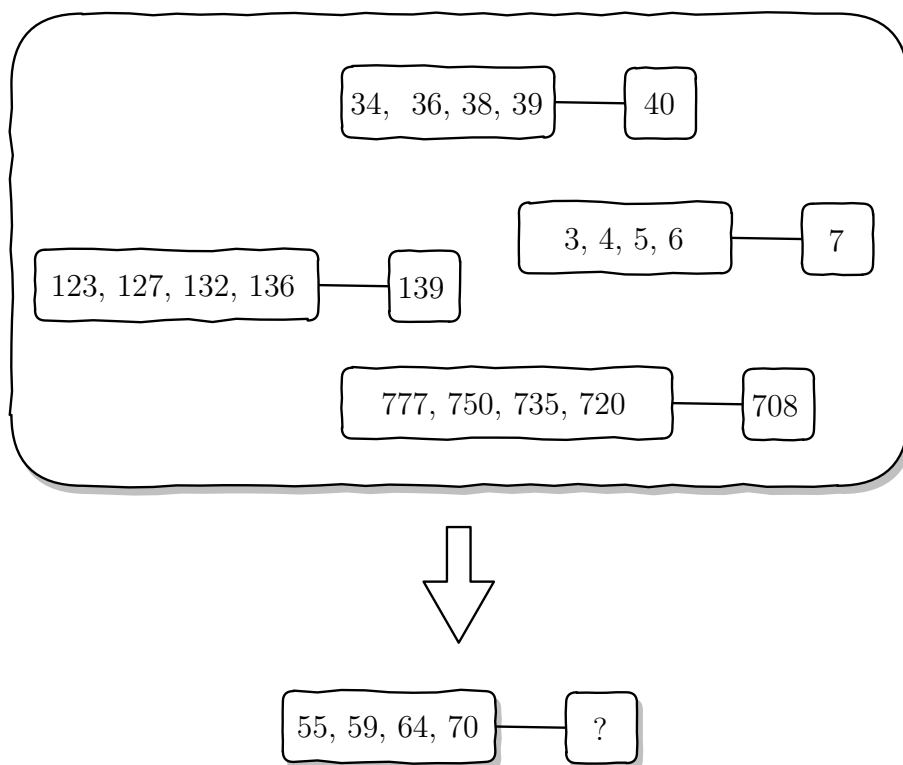
Uczenie maszynowe to metodologia umożliwiająca zdobywanie wiedzy przez systemy komputerowe na podstawie gromadzonych danych. Dokładniejsza definicja znajduje się w książce Toma M. Mitchella pt. "Machine Learning":

"Mówimy, że maszyna uczy się zadania T w oparciu o doświadczenie E i miarę jakości P , jeśli wraz z przyrostem doświadczenia E poprawia się jakość wykonywanego zadania T mierzona przez miarę P ." [14]

Systemy uczenia maszynowego mogą zdobywać wiedzę w oparciu o metody opisane w dalszej części bieżącego rozdziału.

2.2. Uczenie nadzorowane

Uczenie nadzorowane — do systemu przekazywane są dane wejściowe z oczekiwanymi rozwiązaniami nazywanymi etykietami. Celem systemu jest znalezienie związków zachodzących w danych i poprawne wygenerowanie predykcji na nieznanymi dotąd i nieoznaczonych próbkach. W przypadku przewidywania cen na giełdzie papierów wartościowych do systemu mogą zostać przekazane notowania spółek zamknięcia kolejnych dni działania giełdy z wybranego przedziału czasowego oraz oczekiwane wartości notowań do przewidzenia. Po procesie nauki na oznaczonych danych system może wygenerować predykcje kursu akcji zamknięcia kolejnej sesji. Schemat postępowania można prześledzić na rysunku 2.1.



Rys. 2.1. Poglądowy schemat przedstawiający ideę uczenia nadzorowanego.

Do najpopularniejszych metod uczenia nadzorowanego należą:

- Regresja liniowa¹
- Regresja wielomianowa²

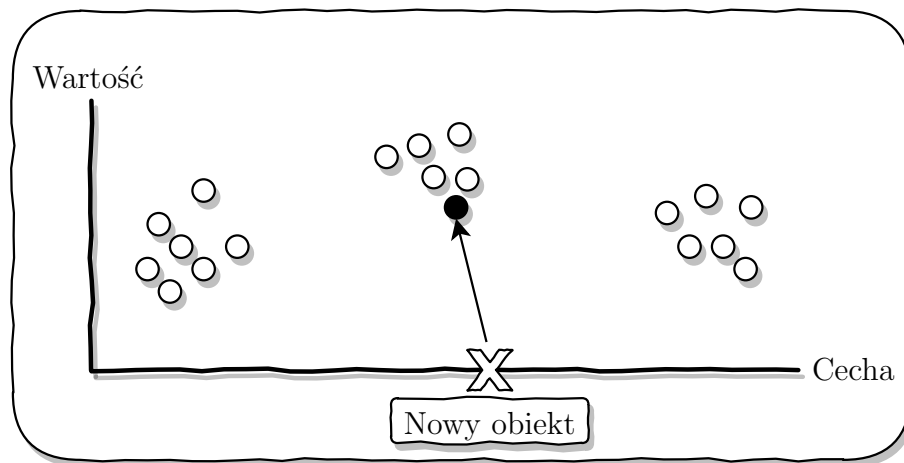
¹Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow", s. 106-107., 2017.

²Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow", s. 121-123., 2017.

- Regresja logistyczna³
- Algorytm k najbliższych sąsiadów⁴
- Maszyna wektorów nośnych⁵
- Drzewa decyzyjne i lasy losowe⁶
- Sztuczne sieci neuronowe⁷

2.3. Uczenie nienadzorowane

Uczenie nienadzorowane — do systemu przekazywane są jedynie dane nieoznaczone, a celem systemu jest znalezienie korelacji pomiędzy nimi oraz odpowiedniego wzorca. W przypadku przewidywania cen na giełdzie papierów wartościowych do systemu mogą zostać przekazane notowania spółek zamknięcia kolejnych dni działania giełdy z wybranego przedziału czasowego bez informacji o interesującym nas wyniku. Po udanym procesie nauki system powinien wykryć zależności pomiędzy wartościami spółek w kolejnych dniach. Schemat postępowania można prześledzić na rysunku 2.2.



Rys. 2.2. Poglądowy schemat przedstawiający ideę uczenia nienadzorowanego.

Do najpopularniejszych metod uczenia nienadzorowanego należą:

³Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow", s. 139-142., 2017.

⁴Altman, "An introduction to kernel and nearest-neighbor nonparametric regression", 1991.

⁵Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow", s. 145-165., 2017.

⁶Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow", s. 167-202., 2017.

⁷Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow", s. 253-311., 2017.

- Klasteryzacja
 - Algorytm centroidów⁸
 - Grupowanie hierarchiczne⁹
 - Algorytm oczekiwanie-maksymalizacja¹⁰

- Wizualizacja i redukcja wymiarowości
 - Analiza głównych składowych¹¹
 - Jądrowa analiza głównych składowych¹²
 - Zanurzanie lokalnie liniowe¹³
 - Stochastyczne zanurzanie sąsiadów przy użyciu rozkładu t¹⁴

- Algorytmy oparte o reguły asocjacyjne
 - algorytm Apriori¹⁵
 - algorytm Eclat¹⁶

2.4. Uczenie częściowo nadzorowane

Uczenie częściowo nadzorowane – do systemu przekazywane są zarówno dane oznaczone, jak i nieoznaczone a celem systemu jest odnalezienie wzorca pomiędzy danymi oznaczonymi. Po udanym procesie nauki system powinien wykryć zależności pomiędzy poszczególnymi próbami oraz być w stanie powiązać je z danymi nieoznaczonymi i oznaczyć je we własnym zakresie. W przypadku przewidywania cen na giełdzie papierów wartościowych do systemu mogą zostać przekazane notowania spółek zamknięcia kolejnych dni działania giełdy z wybranego przedziału czasowego bez oraz wraz z nazwami spółek. Po udanym procesie nauki system jest w stanie

⁸Hartigan, Wong, "A K-Means Clustering Algorithm", s. 100–108., 1979.

⁹Rokach, Lior, Maimon, "Clustering methods", s. 321-352., 2005.

¹⁰Borman, "The Expectation Maximization Algorithm: A short tutorial", 2004.

¹¹Krzanowski, "Principles of Multivariate Analysis: A User's Perspective", 2000.

¹²Hoffmann, Heiko, "Kernel PCA for Novelty Detection", 2007.

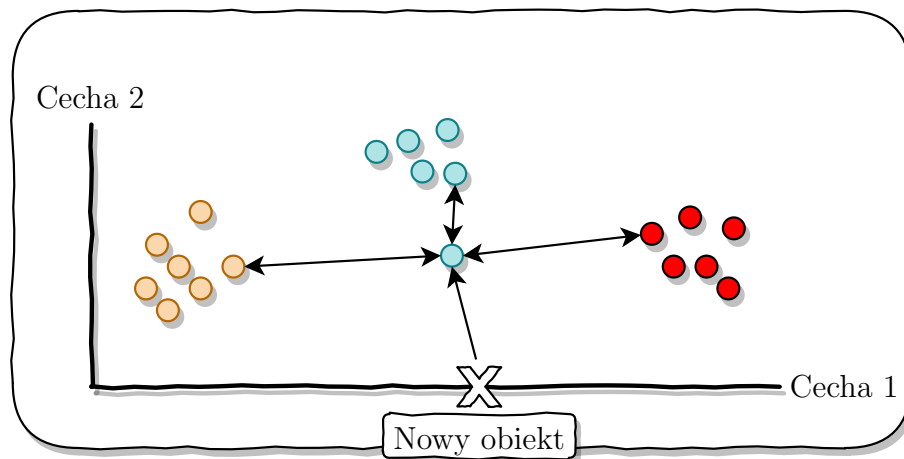
¹³Saul, Roweis, "An Introduction to Locally Linear Embedding", 2001.

¹⁴Maaten, Hinton, "Visualizing Data Using t-SNE", 2008.

¹⁵Agrawal, Srikant, "Fast algorithms for mining association rules", s. 487-499., 1994.

¹⁶Kaur, Grag, "ECLAT Algorithm for Frequent Itemsets Generation", 2014.

pogrupować dane należące do poszczególnych spółek bez znajomości ich nazw. Jeżeli w wybranym zbiorze znajduje się chociaż jedna próba oznaczona, to system automatycznie może przyporządkować nazwę przedsiębiorstwa do wszystkich pozostałych cen spółek znajdujących się w danej klasie. Schemat postępowania można prześledzić na rysunku 2.3.



Rys. 2.3. Poglądowy schemat przedstawiający ideę uczenia częściowo nadzorowanego.

Do uczenia częściowo nadzorowanego wykorzystuje się połączenie algorytmów uczenia nadzorowanego i nienadzorowanego. Do najpopularniejszych metod należą:

- Głębokie sieci przekonań¹⁷
- Maszyny wektorów nośnych¹⁸

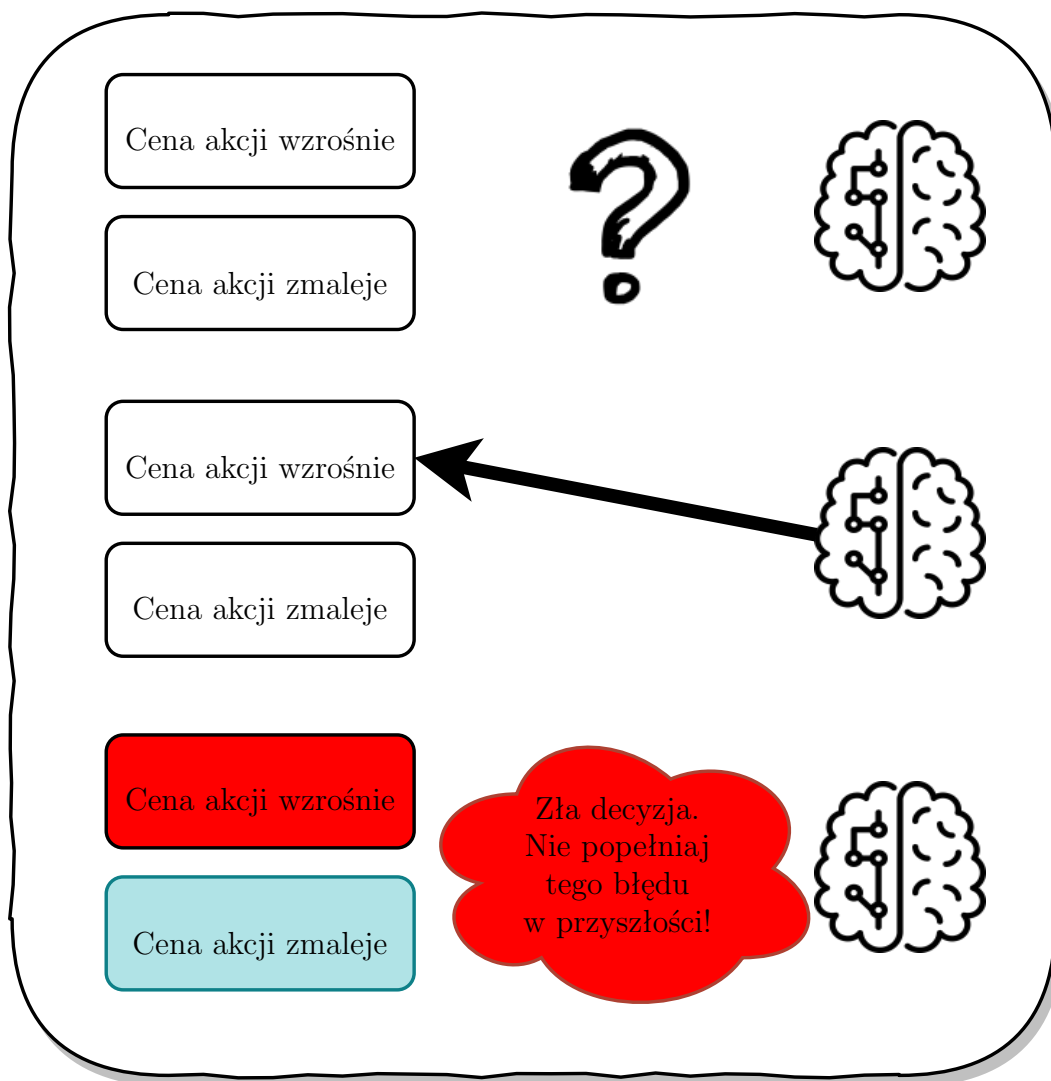
2.5. Uczenie przez wzmacnianie

Uczenie przez wzmacnianie — do systemu nie są przekazywane żadne dane w celu nauki. Podejście to oparte jest na metodzie prób i błędów. System nazywany agentem monitoruje ściśle określone środowisko i otrzymuje nagrody lub kary w oparciu o poprawnie lub błędnie podejmowane decyzje. Celem systemu jest wypracowanie odpowiedniej strategii nazywanej polityką umożliwiającą zdobycie jak największej

¹⁷Goodfellow, Bengio, Courville, "Deep learning : systemy uczące się", s. 671-674., 2018.

¹⁸Cristianini, Shawe-Taylor, "An Introduction to Support Vector Machines and other kernel-based learning methods", 2000.

nagrody. W przypadku przewidywania cen na giełdzie papierów wartościowych system może zostać wyuczony do podejmowania decyzji o zakupie, sprzedaży lub braku reakcji i trzymaniu akcji przedsiębiorstw w portfelu inwestycyjnym. Ponadto może on zostać wykorzystany do przewidywania ruchów giełdowych spółek następnego dnia działania giełdy względem poprzedniego. Po udanym procesie nauki w oparciu o obserwację notowań spółek system powinien być w stanie wypracowywać zysk związany z zakupem akcji i sprzedażą aktywów odbywającej się po wzroście ich wartości. Schemat postępowania można prześledzić na rysunku 2.4.



Rys. 2.4. Poglądowy schemat przedstawiający ideę uczenia częściowo nadzorowanego.

Do najpopularniejszych metod uczenia przez wzmacnianie należą:

- Algorytm SARSA¹⁹
- Algorytm Q-learning²⁰

2.6. Wybór sposobu nauki

Po dokonaniu przeglądu oraz analizy sposobów zdobywania wiedzy przez systemy uczenia maszynowego nadszedł czas na wybór najlepszej techniki z punktu widzenia zautomatyzowania procesu generowania predykcji przyszłych notowań spółek na giełdzie papierów wartościowych. Ze względu na posiadanie dostępu do wszystkich potrzebnych historycznych kursów spółek wchodzących w skład indeksu S&P 500²¹ – które mogą również stanowić dane wyjściowe, postanowiono zrezygnować w dalszej części pracy zarówno z podejścia opartego na uczeniu nienadzorowanym, jak i na uczeniu częściowo nadzorowanym. Ponadto, założeniem pracy magisterskiej jest predykcja przyszłych notowań spółek, a nie tworzenie autonomicznego systemu spekulującego na giełdzie bez ingerencji człowieka. W związku z tym na tym etapie prac odrzucono również podejście uczenia przez wzmacnianie. Najlepszą metodą nauki systemu po uwzględnieniu zdefiniowanych wcześniej założeń okazało się zatem podejście nadzorowane.

2.7. Podsumowanie

W bieżącym rozdziale nastąpiło wprowadzenie do wiedzy o systemach uczenia maszynowego. Omówione zostały różne techniki zdobywania wiedzy na podstawie gromadzonych i monitorowanych danych. Podejście oparte na uczeniu nadzorowanym polega na wyszukiwaniu wzorca na podstawie danych wejściowych i wyjściowych. W uczeniu nienadzorowanym system jest pozbawiony informacji o oczekiwanym wyniku i musi we własnym zakresie znaleźć korelacje jedynie w oparciu o dane nieoznaczone. Z kolei w uczeniu częściowo nadzorowanym system posiada dostęp zarówno

¹⁹Sutton, Barto, "Reinforcement Learning: An Introduction", rozdział 6.4, 2005.

²⁰Sutton, Barto, "Reinforcement Learning: An Introduction", rozdział 6.5, 2005.

²¹Indeks S&P 500 — indeks giełdowy zrzeszający 500 przedsiębiorstw o najwyższej kapitalizacji notowanych na New York Stock Exchange oraz NASDAQ.

do danych oznaczonych, jak i nieoznaczonych. Po procesie nauki system jest w stanie powiązać je z danymi nieoznaczonymi i nadać im etykiety według wyuczonego wcześniej schematu postępowania. Ostatnim omówionym podejściem jest uczenie przez wzmacnianie. W tej technice celem systemu jest wypracowanie odpowiedniej polityki w oparciu o monitorowanie określonego środowiska. Po dokonaniu analizy wszystkich dostępnych sposobów zdobywania wiedzy przez systemy uczenia maszynowego, postanowiono, że najlepszym podejściem z perspektywy generowania przyszłych notowań spółek będzie technika oparta na uczeniu nadzorowanym i to na niej zostanie skupiona uwaga w dalszej części pracy.

Rozdział 3

Metody uczenia maszynowego

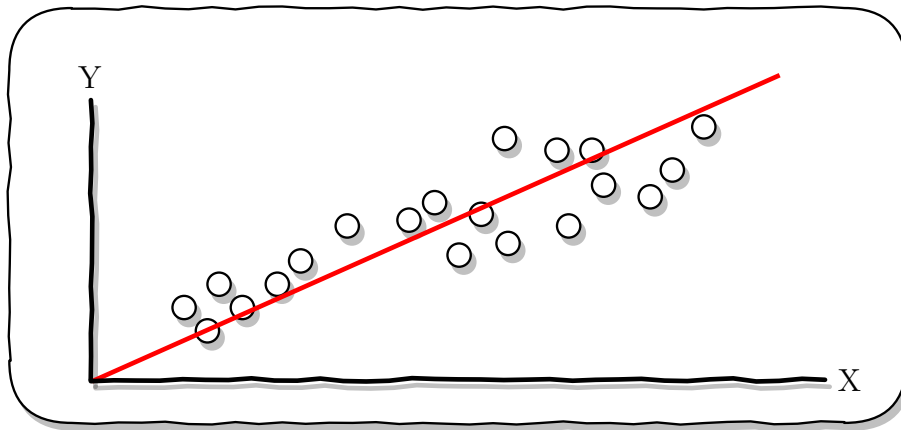
W niniejszym rozdziale zostaną przeanalizowane i omówione najpopularniejsze metody uczenia maszynowego wspierające podejście nadzorowane. Następnie w oparciu o aktualny stan wiedzy oraz odnalezione badania naukowe dokonany zostanie wybór jedynie jednej metody na potrzeby generowania predykcji przyszłych notowań spółek na giełdzie papierów wartościowych.

3.1. Wstęp

Metody uczenia maszynowego na przestrzeni ostatnich lat ewoluowały w bardzo dynamicznym tempie. Co roku powstają nowe wersje oparte o innowacyjne założenia oraz modele matematyczne, a także podejścia hybrydowe. Jaką w związku z tym wybrać metodę dla wybranego problemu, jakim w tym przypadku jest predykcja notowań spółek na giełdzie papierów wartościowych? Na to pytanie postaram się zamieścić odpowiedź w niniejszym rozdziale. Zgodnie ze wcześniejszymi założeniami w dalszej części pracy zostaną przeanalizowane i omówione najpopularniejsze metody uczenia maszynowego wspierające podejście nadzorowane oraz umożliwiające przyjęcie strategii inwestycyjnej wykorzystującej analizę techniczną. Następnie w oparciu o aktualny stan wiedzy oraz bieżące badania naukowe dokonany zostanie wybór metod na potrzeby generowania predykcji przyszłych notowań spółek.

3.2. Regresja liniowa

Regresja liniowa to metoda umożliwiająca wytrenowanie modelu reprezentującego liniową zależność pomiędzy danymi wejściowymi oraz wyjściowymi. Wykres przedstawiający przykładową predykcję modelu znajduje się na rysunku 3.1.



Rys. 3.1. Poglądowy wykres przedstawiający przykładową predykcję modelu regresji liniowej.

Podczas generowania predykcji zmiennych objaśnianych przyjmuje się następujące założenia:

- Niezależność i addytywność — każdy czynnik wpływający na proces doboru parametrów i predykcji wyników uznawany za niezależny, a w obliczeniu wartości zmiennej przewidywanej jest dodawany do pozostałych.
- Monotonniczność i liniowość — zmiana wartości zmiennych objaśniających nie powoduje utraty liniowości oraz zmiany kierunku prostej odwzorowującej generowane predykcje.

Regresja liniowa dokonuje prognozy wartości w oparciu o ważoną sumę cech dla danych wejściowych oraz wyrazu wolnego zgodnie ze wzorem:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

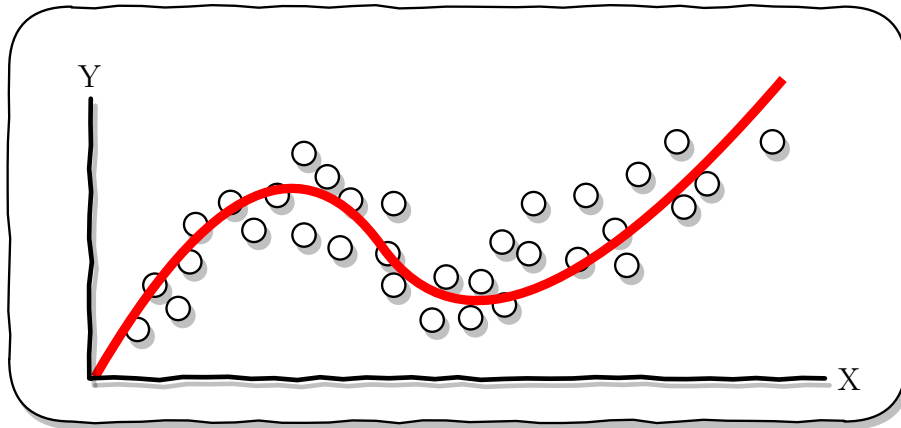
- y - zmienna objaśniana
- n - liczba cech

- x_i - wartość i-tej cechy
- b_j - wartość j-tego współczynnika regresji z uwzględnieniem wyrazu wolnego

W przypadku przewidywania notowań spółek na giełdzie papierów wartościowych do wyznaczenia najlepszych współczynników regresji oraz wyrazu wolnego mogłyby zostać wykorzystane oznaczone dane treningowe dotyczące historycznych kursów akcji. Umożliwiłoby to wygenerowanie predykcji przyszłych wartości waleń na podstawie stworzonego modelu oraz dobranych parametrów w procesie nauki. Dużym ograniczeniem związanym z wykorzystaniem tej metody modelowania zależności jest jej liniowa charakterystyka. Jest to istotna wada w kontekście przewidywania notowań spółek na giełdzie papierów wartościowych, gdzie ceny instrumentów finansowych w określonym oknie czasowym charakteryzują się nieliniowymi zależnościami.

3.3. Regresja wielomianowa

Regresja wielomianowa umożliwia wykrywanie nieliniowych zależności pomiędzy danymi wejściowymi i wyjściowymi. Wykorzystuje ona wielomian n-tego stopnia w celu odnalezienia korelacji pozwalających na wygenerowanie predykcji dla zmiennych objaśniających. Wykres przedstawiający przykładowy wzorzec wykryty przez model regresji wielomianowej znajduje się na rysunku 3.2.



Rys. 3.2. Poglądowy wykres przedstawiający przykładowy wzorzec wykryty przez model regresji wielomianowej.

Regresja wielomianowa dokonuje prognozy wartości w oparciu o ważoną sumę cech dla danych wejściowych podniesionych do odpowiedniej potęgi oraz wyrazu wolnego zgodnie ze wzorem:

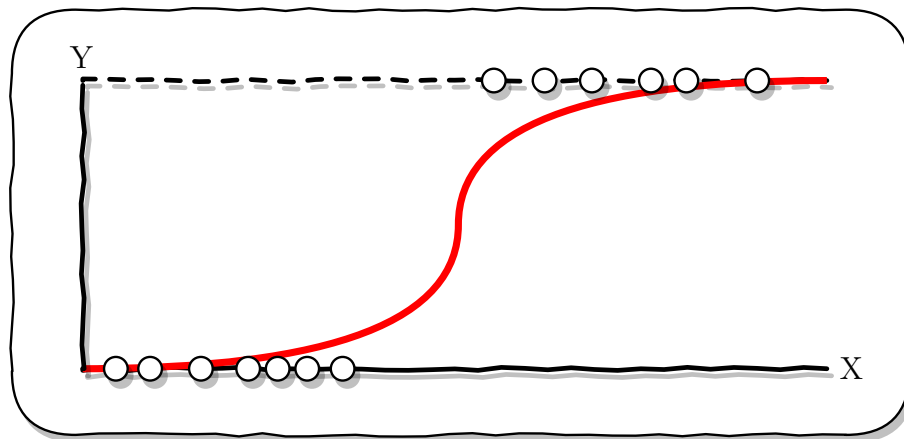
$$y = b_0 + b_1x + b_2x^2 + b_3x^3 + \dots + b_nx^n.$$

- y - zmienna objaśniana
- n - liczba cech
- x - wartość cechy
- b_j - wartość j -tego współczynnika regresji z uwzględnieniem wyrazu wolnego

Podobnie jak przypadku regresji liniowej i przewidywania notowań spółek na giełdzie papierów wartościowych do wyznaczenia najlepszych współczynników regresji oraz wyrazu wolnego mogłyby zostać wykorzystane oznaczone dane treningowe dotyczące historycznych kursów akcji. Umożliwiłyby to wygenerowanie predykcji przyszłych wartości walorów na podstawie stworzonego modelu oraz dobranych parametrów w procesie nauki. W odróżnieniu od regresji liniowej, w przypadku regresji wielomianowej istnieje możliwość dopasowania modelu do krzywoliniowej charakterystyki rozkładu wartości instrumentów w określonym oknie czasowym. Zgodnie z założeniami umożliwiłyby to wygenerowanie znacznie dokładniejszych predykcji dotyczących przyszłych notowań spółek.

3.4. Regresja logistyczna

Regresja logistyczna umożliwia rozwiązywanie problemów związanych z klasyfikacją binarną. Pozwala ona określić przynależność zmiennej wejściowej do jednej z dwóch zdefiniowanych wcześniej klas. Wykres przedstawiający przykładowy wzorec wykryty przez model regresji logistycznej znajduje się na rysunku 3.3.



Rys. 3.3. Poglądowy wykres przedstawiający wykryty wzorec przez model regresji logistycznej.

Regresja logistyczna, podobnie jak regresja liniowa, dokonuje prognozy wartości w oparciu o ważoną sumę cech dla danych wejściowych. Do szacowania prawdopodobieństwa wykorzystywany jest następujący wzór:

$$p(t) = \frac{1}{1 + e^{-t}}$$

- e - stała liczbowa Eulera
- t - zmienna objaśniana postaci $b_0 + b_1x + b_2x^2 + b_3x^3 + \dots + b_nx^n$
- n - liczba cech
- x - wartość cechy
- b_j - wartość j -tego współczynnika regresji z uwzględnieniem wyrazu wolnego

Regresja logistyczna w odróżnieniu od regresji liniowej nie zwraca bezpośrednio szacunkowego prawdopodobieństwa. W kolejnym kroku dokonywana jest klasyfikacja na przykład w oparciu o następujący model matematyczny:

$$y = \begin{cases} 0 & \text{dla } p(t) < 0.5, \\ 1 & \text{dla } p(t) \geq 0.5. \end{cases}$$

W związku z dychotomiczną charakterystyką regresji logistycznych niemożliwie jest wygenerowanie predykcji notowań spółek na giełdzie papierów wartościowych. Metoda ta mogłaby jednak znaleźć zastosowanie w przewidywaniu przyszłych ruchów kursów akcji. Zgodnie z przyjętymi założeniami analizy technicznej mogą one przyjmować kierunek wzrostowy, spadkowy lub horyzontalny. Z perspektywy wypracowywania zysku przez spekulanta stan horyzontalny ma mniejsze znaczenie i mógłby zostać pominięty na potrzeby dokonywania klasyfikacji binarnej. Historyczne notowania spółek umożliwiłyby dobranie odpowiednich parametrów w modelu matematycznym w celu przewidywania kierunku ruchu notowań akcji kolejnego dnia zamknięcia giełdy względem wartości z dnia poprzedniego.

3.5. Algorytm k najbliższych sąsiadów

Algorytm k najbliższych sąsiadów, w odróżnieniu od omówionych powyżej metod, nie opiera się na trenowaniu modelu w celu generowania predykcji zmiennych objaśnianych. Z tego powodu nazywany jest również często algorytmem leniwym. Idea jego działania polega na przyporządkowaniu wszystkim danym wejściowym zestawu cech oraz umieszczeniu ich w wielowymiarowej przestrzeni w oparciu o miarę podobieństwa. W przypadku przekazania do algorytmu próby nieoznaczonej następuje wyszukanie k najbliższych obiektów przy pomocy określonej metody. Najczęściej wykorzystywane są do tego celu następujące miary odległości:

- Euklidesowa:

$$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- Manhattan:

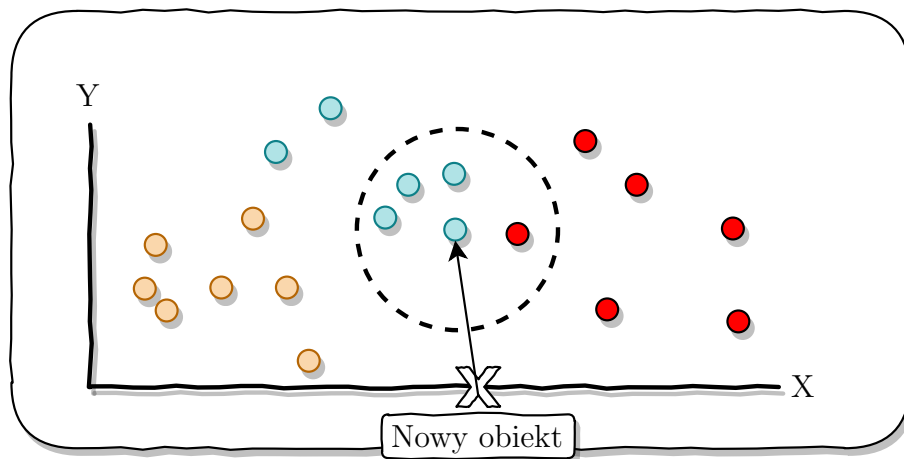
$$d = \sum_{i=1}^n |x_i - y_i|$$

- Minkowskiego:

$$d = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}}$$

- x_i i y_i - obserwacje służące do obliczania odległości,
- P - parametr przyjmujący wartość 1 dla odległości Manhattan oraz wartość 2 w przypadku odległości Euklidesowej.

W kolejnym kroku następuje zliczenie wystąpień wszystkich klas z wybranego zbioru najbliższych sąsiadów i przyporządkowanie etykiety najczęściej występującej grupy do zmiennej objaśnianej. Wykres przedstawiający przykładową predykcję modelu znajduje się na rysunku 3.4.

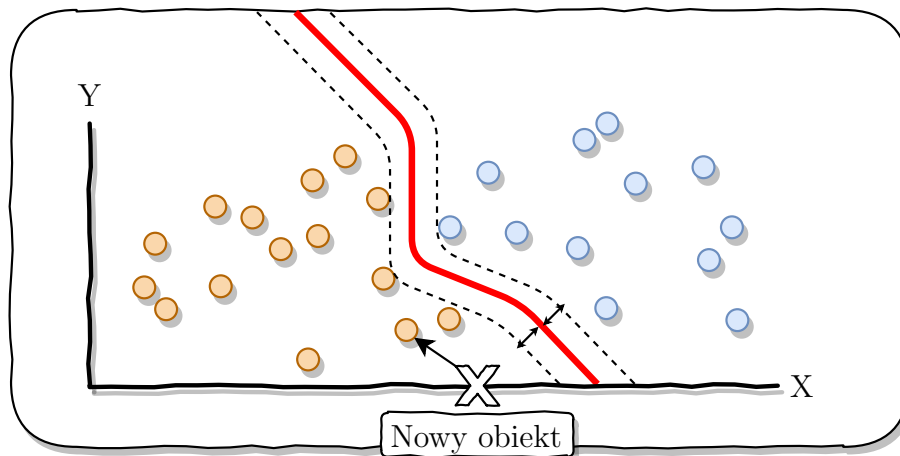


Rys. 3.4. Poglądowy wykres przedstawiający przykładową predykcję modelu algorytmu k najbliższych sąsiadów.

W przypadku przewidywania notowań spółek na giełdzie papierów wartościowych algorytm k najbliższych sąsiadów mógłby znaleźć zastosowanie w predykcji przyszłych ruchów kursów akcji. Na podstawie danych wejściowych oznaczających historyczne wartości spółek przydzielałby on je do jednej z trzech grup oznaczających ruch cen instrumentów finansowych w kierunku wzrostowym, spadkowym lub horyzontalnym w czasie kolejnego dnia działania giełdy.

3.6. Maszyna wektorów nośnych

Maszyna wektorów nośnych jest techniką wykorzystującą wielowymiarową przestrzeń w celu znalezienia maksymalnej hiperpłaszczyzny brzegowej umożliwiającej najbardziej precyzyjny sposób podziału danych na klasy. Wykres przedstawiający przykładową predykcję modelu znajduje się na rysunku 3.5.



Rys. 3.5. Poglądowy wykres przedstawiający przykładową predykcję modelu maszyny wektorów nośnych.

Idea działania maszyny wektorów nośnych opiera się na wyszukiwaniu maksymalnych odległości pomiędzy najbliższymi punktami nazywanymi marginesami zgodnie ze wzorem:

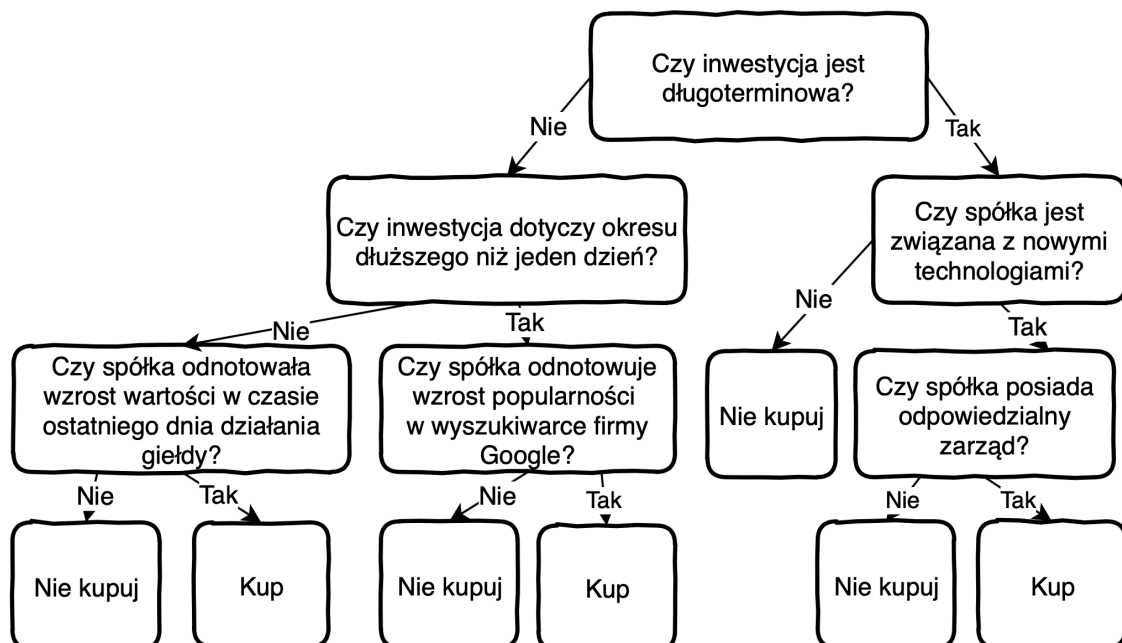
$$f(x) = \beta_0 + \sum_{i \in S} \alpha_i K(x_i, x_i')$$

- β_0 - wyraz wolny
- S - zbiór wszystkich obserwacji wektora nośnego
- α - parametry modelu przeznaczone do nauki
- (x_i, x_i') - pary obserwacji wektora nośnego
- K - funkcja nazywana kernelem porównująca podobieństwo pomiędzy x_i i x_i'

W przypadku przewidywania notowań spółek na giełdzie papierów wartościowych maszyna wektorów nośnych mogłaby zostać wykorzystana podobnie jak w przypadku regresji logistycznej oraz algorytmu k najbliższych sąsiadów w celu predykcji przyszłych ruchów kursów akcji. Algorytm na podstawie danych wejściowych oznaczających historyczne wartości spółek przewidywałby ruch wzrostowy, spadkowy lub horyzontalny na zakończenie kolejnego dnia działania giełdy względem poprzedniego. Oczywiście predykcje odbywałyby się przy wykorzystaniu wytrenowanego wcześniej modelu sekwencyjnie minimalizującego wybraną funkcję błędu w oparciu o próby uczące.

3.7. Drzewa decyzyjne i lasy losowe

Drzewa decyzyjne to metody wykorzystujące szereg zasad decyzyjnych do wytrenowania modelu w oparciu o zbiór uczący w celu generowania przyszłych predykcji na podstawie zmiennych objaśniających. Poniżej znajduje się schemat przedstawiający przykładową strukturę drzewa decyzyjnego (patrz rysunek 3.6) oraz opis jego najważniejszych składowych.



Rys. 3.6. Poglądowy schemat przedstawiający przykładową strukturę modelu drzewa decyzyjnego.

- Gałąź - krawędź łącząca pozostałe elementy drzewa
- Węzeł - wierzchołek łączący co najmniej jedną gałąź
- Korzeń - główny węzeł drzewa
- Liść - węzeł z którego nie wychodzi żadna gałąź

Głównym problemem w efektywnym korzystaniu z drzew decyzyjnych jest dobór odpowiedniej struktury dla rozważanego zagadnienia. W tym celu stosuje się algorytmy rekurencyjne umożliwiające maksymalizację zdobywania najistotniejszych informacji z punktu widzenia rozwiązywanego problemu podczas dokonywania decyzji oraz podziału w każdym węźle. W przypadku przewidywania cen na giełdzie papierów wartościowych wyznacznikiem dla reguły decyzyjnej mogłaby być wartość błędu średniokwadratowego postaci:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- y_i - rzeczywista wartość wyniku
- \hat{y}_i - przewidziana wartość

Algorytm rekurencyjny po znalezieniu najskuteczniejszych reguł decyzyjnych wykrytych na podstawie minimalizacji wartości błędu średniokwadratowego oraz zmiennych objaśniających mógłby generować przewidywania notowań spółek zamknięcia kolejnego dnia działa giełdy dla nieoznaczonych danych wejściowych. W celu przewidywania jeszcze dokładniejszych wyników mógłby zostać wykorzystany algorytm lasów losowych. Metoda ta polega na wykorzystaniu wielu drzew decyzyjnych w celu wygenerowania predykcji, a następnie wybrania z nich najczęściej występującego rezultatu.

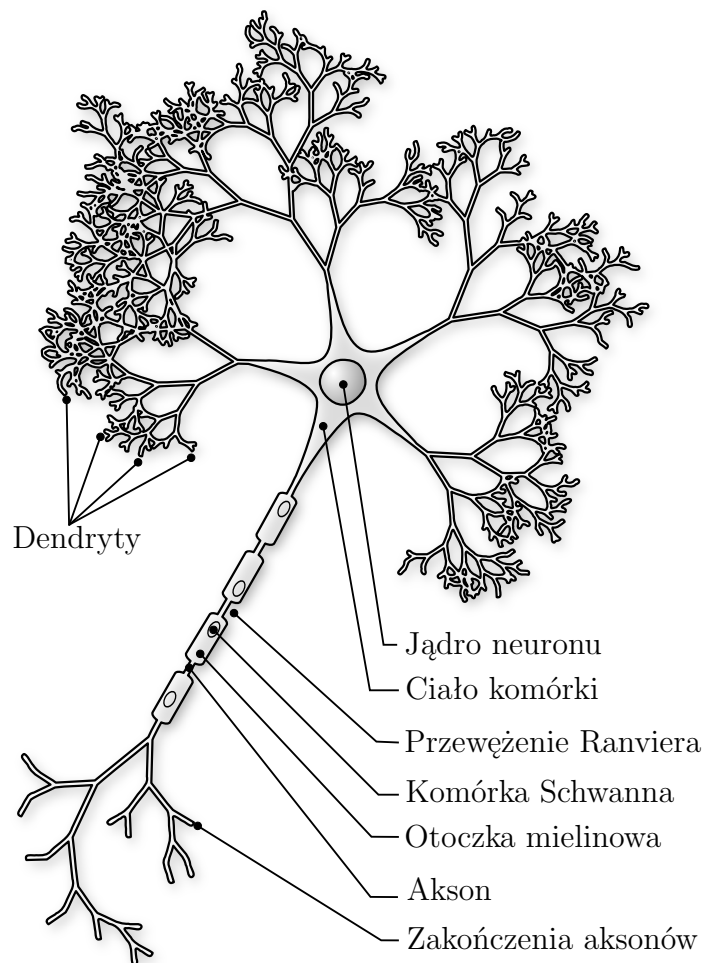
3.8. Sztuczne sieci neuronowe

3.8.1. Podstawowe informacje o sztucznych sieciach neuronowych

Sztuczne sieci neuronowe to modele matematyczne odpowiedzialne za przetwarzanie informacji w oparciu o ideę funkcjonowania ludzkiego mózgu. W książce pt. "Komputerowe metody analizy i przetwarzania danych." zostały one zdefiniowane w sposób następujący:

"Sieci neuronowe są metodami obliczeniowymi, których działanie inspirowane jest sposobem funkcjonowania układów nerwowych organizmów żywych. Podobnie jak układ nerwowy składa się z komórek nerwowych zwanych neuronami, tak również wykorzystywane w analizie danych sztuczne sieci neuronowe składają się z elementów przetwarzających nazywanych sztucznymi neuronami." [12]

W celu lepszego zrozumienia zasady działania wspomnianych sztucznych neuronów w pierwszej kolejności nastąpi przeanalizowanie struktury ich biologicznych odpowiedników w oparciu o schemat przedstawiony na rysunku 3.7.

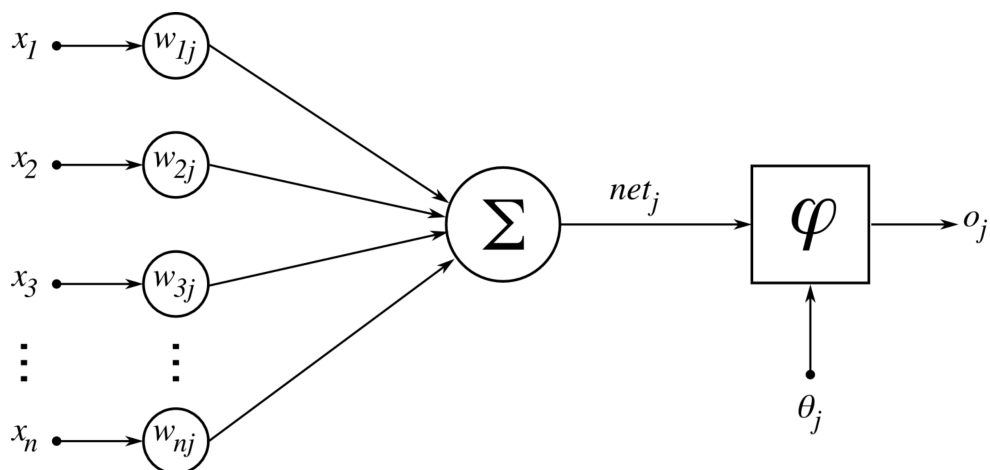


Rys. 3.7. Schemat przedstawiający uproszczoną budowę biologicznego neuronu. Źródło: <https://pl.wikipedia.org/wiki/Neuron>

Do najważniejszych elementów neuronu z punktu widzenia omawianych struktur matematycznych w kontekście predykcji przyszłych notowań spółek należą:

- Dendryty - rozgałęzione struktury umożliwiające zebranie impulsów elektrycznych od innych neuronów i dostarczenie ich do jądra komórki,
- Jądro - centrum obliczeniowe neuronu,
- Akson - włókna odpowiedzialne za przekazanie sygnałów wynikowych do innych neuronów,
- Synapsa - miejsce komunikacji aksonu z dendrytem innego neuronu.

Zgodnie z informacjami zamieszczonymi w publikacji pt. "Budowa aplikacji webowej umożliwiającej obrót instrumentami finansowymi na giełdzie." [11] w biologicznym neuronie proces przetwarzania informacji rozpoczyna się od odebrania przez dendryty impulsów elektrycznych i dostarczenia ich do jądra komórki. Następują tam procesy obliczeniowe oraz wygenerowanie sygnału wynikowego, który przesyłany jest od ciała komórki do synaps znajdujących się na zakończeniach aksonów. Okazuje się, że to mikroskopijnej wielkości miejsce komunikacji dwóch neuronów jest najważniejsze w całym procesie przetwarzania informacji. Podczas przesyłania sygnału elektrycznego przez synapsę uwalniane są specjalne substancje chemiczne zwane neuromediatorami. Proces uczenia się neuronu w skrajnym uproszczeniu polega na wygenerowaniu przez zakończenie aksonu odpowiedniej ilości neuroprzekaźnika w zależności od istotności przesyłanej informacji. Opisana idea działania została wykorzystana w modelu sztucznego neuronu zobrazowanego na rysunku 3.8. Na wejściu neuronu przekazywane są sygnały $x_1, x_2, x_3, \dots, x_n$. Następnie każdy z nich przemnażany jest przez odpowiednią wartość ze zbioru wag $W_{1j}, W_{2j}, W_{3j}, \dots, W_{nj}$. W dalszej kolejności zmodyfikowane sygnały trafiają do funkcji agregacji Σ sumującej wartości wejściowe. Ostatni etap opiera się na przekazaniu sygnału wynikowego net_j do charakterystyki neuronu φ . W funkcji aktywacji zostaje on przekształcony z uwzględnieniem funkcji progowej θ_j do sygnału wyjściowego neuronu o_j .



Rys. 3.8. Schemat modelu sztucznego neuronu. Źródło: https://en.wikibooks.org/wiki/Artificial_Neural_Networks

Poprzez zestawienie oraz połączenie ze sobą omówionych wyżej modeli matematycznych istnieje możliwość tworzenia bardziej złożonych struktur nazywanych sztucznymi sieciami neuronowymi. Wykazują one znacznie ciekawsze właściwości i są w stanie rozwiązywać bardziej skomplikowane oraz różnorodne problemy. Ze względu na odbywający się w nich przepływ informacji możemy dokonać podziału sieci neuronowych na jednokierunkowe i rekurencyjne. Obie klasy zostały scharakteryzowane poniżej.

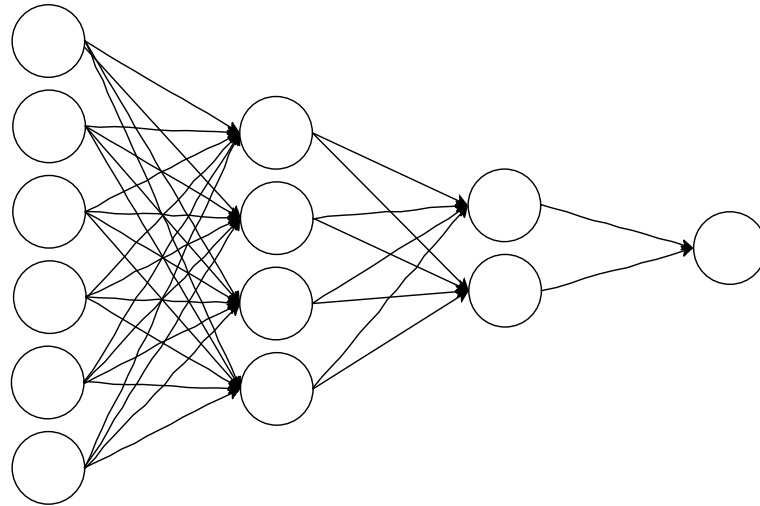
3.8.2. Jednokierunkowe sieci neuronowe

Jednokierunkowe sieci neuronowe opierają swoje działanie na ściśle określonym przepływie informacji od wejścia do wyjścia. W tej klasie sieci najczęściej wykorzystuje się model warstwowy polegający na zorganizowaniu struktury sieci w skupiska neuronów umożliwiające przekazywanie informacji pomiędzy elementami sąsiadujących warstw. Do najpopularniejszych modeli zaliczanych do klasy jednokierunkowych sieci neuronowych należą:

- Wielowarstwowy perceptron - sieć neuronowa zbudowana z perceptronów¹ posiadająca jedną warstwę wejściową, co najmniej jedną warstwę ukrytą² i jedną warstwę wyjściową. W tej topologii sieci połączenia występują między wszystkimi neuronami sąsiadujących warstw, a proces uczenia się najczęściej oparty jest na metodzie wstecznej propagacji błędów. Polega ona na modyfikacji wag dowolnych połączeń elementów przetwarzających w celu optymalizacji sygnałów wynikowych. Schemat przedstawiający przykładową strukturę modelu znajduje się na rysunku 3.9.

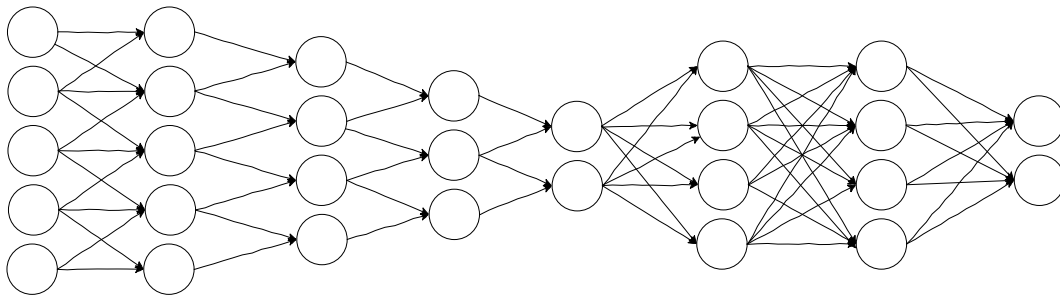
¹Perceptron – podstawowa struktura sztucznych sieci neuronowych składająca się z jednego lub wielu niezależnych neuronów.

²Warstwa ukryta – warstwa znajdująca się pomiędzy warstwą wejściową i warstwą wyjściową.



Rys. 3.9. Schemat przedstawiający przykładową strukturę modelu wielowarstwowego perceptronu.

- Konwolucyjna sieć neuronowa - zawiera jedną lub więcej warstw łączących się lub całkowicie połączonych i wykorzystuje odmianę wielowarstwowego perceptronu. Konwolucyjna sieć neuronowa charakteryzuje się współdzieleniem wag i wykorzystaniem operacji splotu umożliwiając tym samym przechodzenie w głąb sieci przy wykorzystaniu coraz mniejszej liczby parametrów. Schemat przedstawiający przykładową strukturę modelu znajduje się na rysunku 3.10.

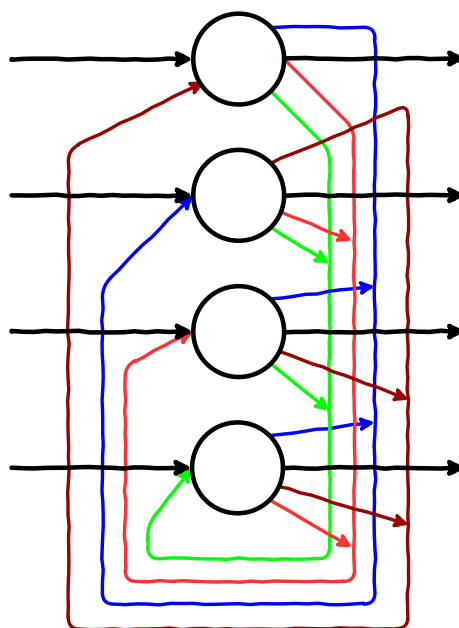


Rys. 3.10. Schemat przedstawiający przykładową strukturę konwolucyjnej sieci neuronowej.

3.8.3. Rekurencyjne sieci neuronowe

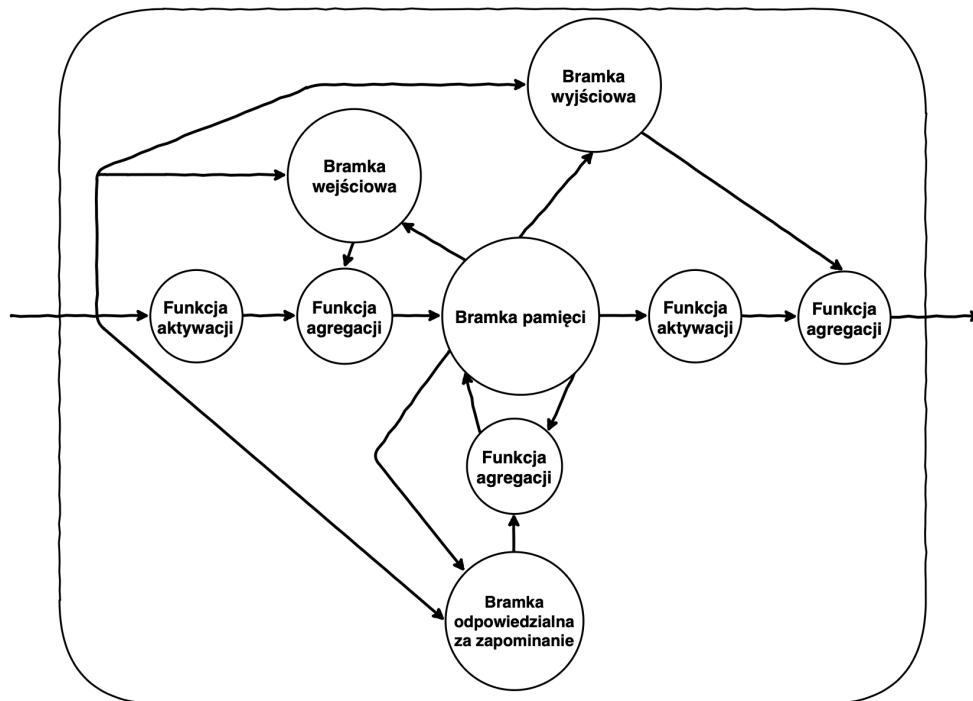
Rekurencyjna sieć neuronowa zawiera sprzężenia zwrotne, czyli połączenia dostarczające sygnały wyjściowe znajdujące się w dalszych warstwach sieci neuronowej do neuronów położonych w warstwie wejściowej lub w bliższych warstwach ukrytych. Umożliwia to wykonywanie bardziej złożonych obliczeń i uzyskiwanie dzięki temu bogatszych właściwości niż sieci posiadające jedynie jednokierunkowy przesył informacji. Do najpopularniejszych modeli zaliczanych do klasy rekurencyjnych sieci neuronowych należą:

- Sieć Hopfielda — idea sieci polega na połączeniu każdego neuronu w celu wytworzenia sprzężenia zwrotnego z każdym innym neuronem. W związku z tym wszystkie sygnały znajdujące się na wyjściach stanowią również wejścia i są to sygnały sprzężenia zwrotnego. Sieci Hopfielda charakteryzują się wysoką stabilnością procesów nauki ze względu na zastosowanie regularnej struktury, braku sprzężeń zwrotnych w obrębie pojedynczego neuronu oraz wykorzystywanie symetrycznych współczynników wagowych. Schemat przedstawiający przykładową strukturę modelu znajduje się na rysunku 3.11.



Rys. 3.11. Schemat przedstawiający przykładową strukturę sieci Hopfielda.

- Sieć neuronowa z pamięcią skojarzeniową – posiada zdolność do uczenia się okresowych zależności, nie wykorzystuje funkcji aktywacji³ w swoich sprzężeniach zwrotnych, zapisane wartości nie są modyfikowane, a gradient nie ma tendencji do zaniku⁴ podczas treningu. Liczba zapamiętywanych stanów⁵ ustalana jest przez architekta sieci. Najczęściej struktura sieci składa się z komórki pamięci⁶, bramki wejściowej⁷, bramki wyjściowej⁸ i bramki odpowiedzialnej za zapominanie⁹. Schemat przedstawiający przykładową strukturę modelu znajduje się na rysunku 3.12.



Rys. 3.12. Schemat przedstawiający przykładową strukturę rekurencyjnej sieci neuronowej z pamięcią skojarzeniową.

³Funkcja aktywacji – funkcja wykorzystywana do obliczania wartości wyjściowych sztucznych neuronów.

⁴Zanikanie gradientu – problem uniemożliwiający stabilny i efektywny proces nauki z powodu dokonywania zbyt nieznacznych modyfikacji wag neuronów modelu.

⁵Stany neuronów – modyfikowane wartości umożliwiające wykrywanie okresowych zależności.

⁶Komórka pamięci – składowa sieci odpowiedzialna za zachowywanie określonych stanów na różnych etapach czasowych.

⁷Bramka wejściowa – składowa sieci odpowiedzialna za kontrolowanie oraz wybór stanów wejściowych.

⁸Bramka wyjściowa – składowa sieci odpowiedzialna za kontrolowanie oraz wybór stanów wyjściowych.

⁹Bramka odpowiedzialna za zapominanie – składowa sieci odpowiedzialna za kontrolowanie oraz usuwanie wybranych stanów.

W przypadku przewidywania notowań spółek kolejnego dnia działania giełdy papierów wartościowych na wejście sztucznej sieci neuronowej mogłyby zostać wprowadzone historyczne kursy akcji. Model przy wykorzystaniu wybranego algorytmu uczenia oraz w oparciu o oczekiwany wynik dla każdego zestawu zmiennych objaśniających dostosowywałby wagi połączeń pomiędzy poszczególnymi elementami umożliwiając wygenerowanie predykcji na nieoznaczonych danych.

3.9. Wybór metody uczenia maszynowego

Po dokonaniu przeglądu oraz analizy nadzorowanych metod uczenia maszynowego nadszedł czas na wybór najlepszego algorytmu z punktu widzenia przewidywania przyszłych notowań spółek na giełdzie papierów wartościowych. Wszystkie omówione w bieżącym rozdziale techniki są bardzo różnorodne oraz umożliwiają generowanie prognoz z wykorzystaniem analizy technicznej i historycznych kursów akcji. Wysoce problematyczne jest jednoznaczne wskazanie najefektywniejszego algorytmu dla rozważanego problemu. W związku z tym w podjęciu decyzji postanowiono również uwzględnić aktualne wyniki badań oraz publikacje dotyczące wyżej wspomnianych metod w kontekście przewidywania kursów akcji na giełdzie papierów wartościowych. W większości odnalezionych eksperymentów najefektywniejszą metodą uczenia maszynowego dotyczącą przewidywania szeregów czasowych¹⁰ okazały się sztuczne sieci neuronowe.[8][16][17][18][5] Ponadto szczególnie wysoką skutecznością cechowały się ich rekurencyjne odmiany. Jak wszystkie metody uczenia maszynowego posiadają one również swoje wady. W wielu odnalezionych publikacjach poruszony został problem eksplozji gradientu¹¹ podczas przewidywania szeregów czasowych przy użyciu wspomnianej techniki.[4] Rekomendowanym rozwiązaniem dla omawianego problemu jest wykorzystanie rekurencyjnej sieci neuronowej z pamięcią skojarzeniową i to właśnie ten model uczenia maszynowego według opinii autora będzie najlepszym algorytmem dla generowania predykcji notowań spółek zamknięcia kolejnego dnia działa giełdy.

¹⁰Szereg czasowy – ciąg informacji uszeregowany według kolejności czasu zarejestrowania określonych obserwacji.

¹¹Eksplozja gradientu – problem uniemożliwiający stabilny i efektywny proces nauki z powodu dokonywania zbyt znacznych modyfikacji wag neuronów modelu.

3.10. Podsumowanie

W bieżącym rozdziale zostały omówione najpopularniejsze metody uczenia maszynowego wspierające podejście nadzorowane. Na początku poddano analizie regresję liniową. W przeciwieństwie do regresji wielomianowej umożliwia ona jedynie wyszukiwanie liniowych zależności pomiędzy danymi wejściowymi i wyjściowymi. Z kolei regresja logistyczna dokonuje klasyfikacji binarnej. Algorytm k najbliższych sąsiadów w odróżnieniu od omówionych metod nie opiera się na trenowaniu modelu. Sprawdza on, jaka jest najliczniejsza klasa w obrębie danej zmiennej objaśniającej i przyporządkowuje jej etykietę tej grupy. Omówiona w dalszej części bieżącego rozdziału maszyna wektorów nośnych klasyfikuje obiekty na podstawie odległości ich cech w wielowymiarowej przestrzeni. Następnie przeanalizowane zostały lasy losowe składające się z drzew decyzyjnych. Idea tego rozwiązania polega na wykorzystaniu szeregu reguł decyzyjnych w celu dostosowania modelu oraz wygenerowania predykcji. Ostatnią omówioną metodą były sztuczne sieci neuronowe oparte na idei funkcjonowania ludzkiego mózgu. Najpopularniejszymi modelami wykorzystującymi jednokierunkowy przepływ sygnałów są wielowarstwowy perceptron oraz konwolucyjna sieć neuronowa. Z kolei do rozwiązywania problemów przy użyciu sieci zawierających sprzężenia zwrotne najczęściej korzysta się z sieci Hopfielda lub sieci z pamięcią skojarzeniową. Po dokonaniu analizy wszystkich dostępnych metod uczenia maszynowego, postanowiono, że wybranym algorytmem z perspektywy generowania przyszłych notowań spółek będzie rekurencyjna sieć neuronowa z pamięcią skojarzeniową i to ona zostanie wykorzystana do przeprowadzenia eksperymentów.

Rozdział 4

Wykorzystane narzędzia

4.1. Wstęp

DO przeprowadzonych eksperymentów wybrano rekurencyjną sieć neuronową z pamięcią skojarzeniową. Aby możliwe było przeprowadzenie eksperymentów, w pierwszej kolejności musi zostać dokonana selekcja odpowiednich narzędzi. W niniejszym rozdziale znajduje się przegląd narzędzi dostępnych nieodpłatnie. Na początku uwaga skupiona zostanie na obecnie dwóch najpopularniejszych platformach programistycznych wspierających tworzenie i pracę ze sztucznymi sieciami neuronowymi. Następnie zostanie dokonany wybór jednej z nich. W drugiej części rozdziału nastąpi omówienie i wybór bibliotek pomocniczych.

4.2. Platformy programistyczne przeznaczone dla sztucznych sieci neuronowych

Poniżej znajduje się porównanie pod względem różnych kryteriów dwóch obecnie najpopularniejszych platform programistycznych przeznaczonych dla sztucznych sieci neuronowych oraz języka programowania Python - TensorFlow oraz PyTorch.

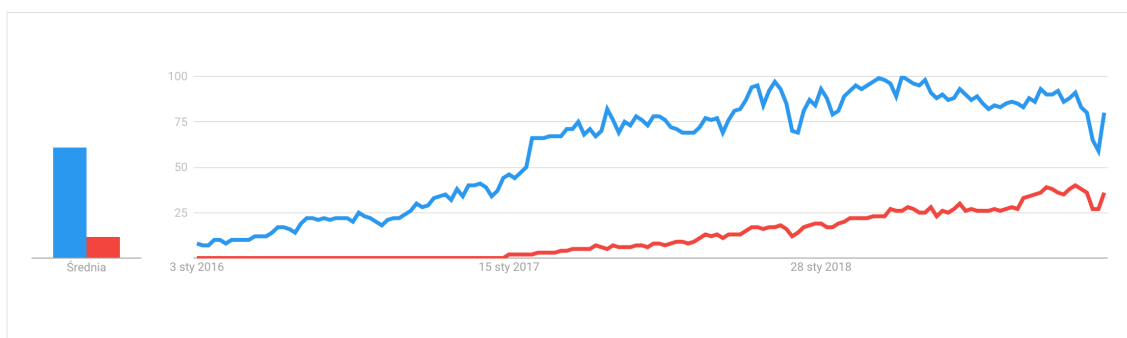
4.2.1. Porównanie TensorFlow i PyTorch

- Charakterystyka ogólna

TensorFlow to platforma stworzona przez firmę Google oraz upubliczniona na licencji Apache 2.0 w 2015 roku. Z kolei PyTorch to platforma opracowana przez firmę Facebook i wydana na licencji BSD w 2016 roku.

- Popularność

TensorFlow posiada znacznie liczniejszą społeczność niż PyTorch. Można zaobserwować tę zależność na rysunku 4.1. Na wykresie znajduje się porównanie platform w ujęciu czasowym pod względem popularności w wyszukiwarce firmy Google. Niebieski wykres dotyczy biblioteki TensorFlow, z kolei czerwony odzwierciedla zainteresowanie platformą PyTorch. Ponadto TensorFlow jest jednym z najpopularniejszych projektów w serwisie internetowym GitHub.



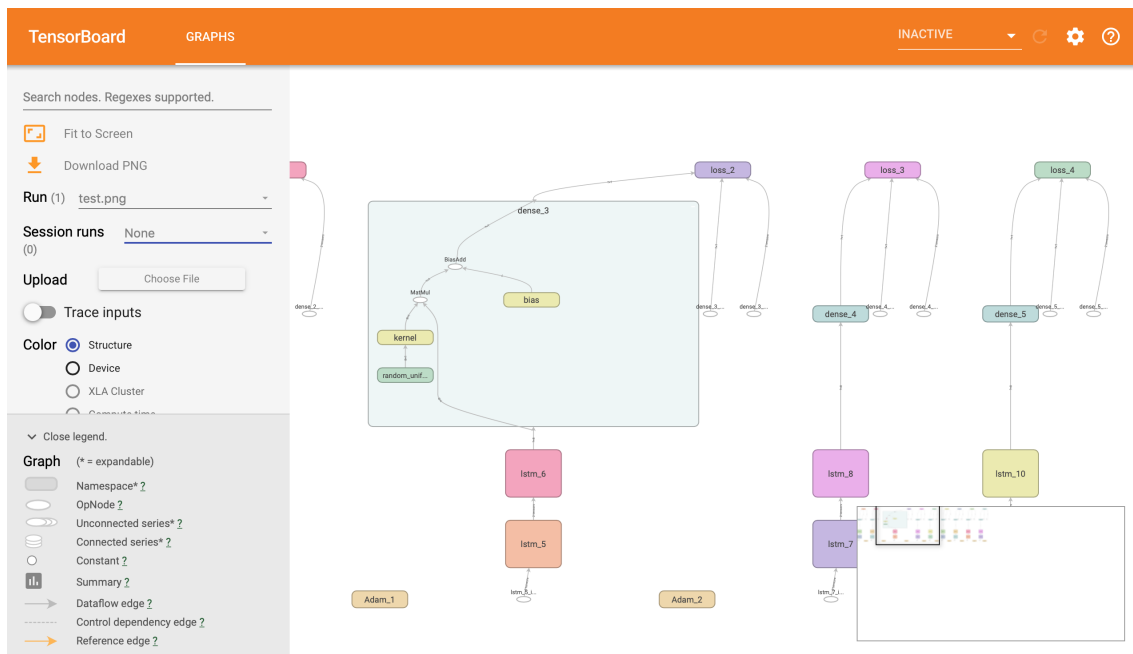
Rys. 4.1. Porównanie platform TensorFlow i PyTorch w ujęciu czasowym pod względem popularności w wyszukiwarce firmy Google. Źródło: <https://trends.google.pl/trends>

- Społeczność programistyczna

Wyższa popularność TensorFlow znajduje również odzwierciedlenie w znacznie liczniejszej społeczności programistów tworzących dostępne narzędzia. Oznacza to, że znalezienie rozwiązań lub zasobów do nauki jest mniej problematyczne niż w przypadku biblioteki PyTorch.

- Wizualizacja

TensorFlow posiada narzędzie wizualizacji danych o nazwie TensorBoard. Umożliwia ono przedstawienie graficzne, debugowanie, zrozumienie oraz optymalizację stworzonych modeli. Ponadto pozwala na wyświetlanie wykresu dotyczącego obliczeń, krzywych uczenia się i wielu innych zależności. Przykładowa wizualizacja znajduje się na rysunku 4.2. PyTorch nie udostępnia odpowiednika narzędzia TensorBoard.



Rys. 4.2. Przykładowa wizualizacja narzędzia TensorBoard.

- Wdrażanie

TensorFlow w odróżnieniu od PyTorch posiada elastyczny, intuicyjny i wysokowydajny system wdrażania modeli uczenia maszynowego dla środowisk produkcyjnych o nazwie TensorFlow Serving.

- Serializacja

TensorFlow w odróżnieniu od PyTorch umożliwia konwertowanie grafów do strumienia bajtów i wykorzystywanie ich w innych językach programowania.

- Obliczenia rozproszone

Obie platformy wspierają rozproszone wykonywanie, zarządzanie i optymalizowanie obliczeń przy wykorzystaniu procesorów, kart graficznych i układów scalonych.

- Dokumentacja

Obie platformy posiadają bardzo rozbudowane, dokładne i poparte wieloma praktycznymi przykładami dokumentacje.

- Integralność

TensorFlow i PyTorch są integralne z takimi platformami jak Windows, Linux i MacOS.

- Określanie struktury grafu

W TensorFlow definiowanie grafu obliczeniowego odbywa się w sposób statyczny. Zanim model zostanie uruchomiony, jego struktura musi być ściśle określona. Natomiast PyTorch wykorzystuje dynamiczne definiowanie struktury pozwalające na łatwe dostosowywanie i modyfikację sztucznej sieci neuronowej nawet podczas procesu jej trenowania.

- Debugowanie¹

TensorFlow w odróżnieniu od PyTorch posiada elastyczny system debugowania modeli uczenia maszynowego nazywany TensorFlow Debugger. Pozwala on między innymi na ewaluację modelu, przegląd tensorów i sprawdzanie operacji numerycznych.

¹Debugowanie - proces polegający na identyfikacji oraz usunięciu błędów z oprogramowania.

- Wsparcie zewnętrznych serwisów

TensorFlow i PyTorch ze względu na swoją popularność są wspierane niemalże we wszystkich platformach umożliwiających wykonywanie obliczeń w środowiskach chmurowych. Szczególne wsparcie TensorFlow posiada w platformie Google Cloud Platform rozwijanej przez jego firmę autorską.

4.2.2. Wybór platformy

Zarówno TensorFlow, jak i PyTorch są doskonałymi narzędziami umożliwiającymi interakcję i pracę ze sztucznymi sieciami neuronowymi. Zapewne obie platformy są w stanie sprostać wymaganiom stawianym im podczas generowania predykcji na giełdzie papierów wartościowych, jednak na potrzeby przeprowadzenia eksperymentów musi zostać dokonany wybór jednej z nich. Ostatecznie po dokonaniu analizy obu środowisk postanowiono w dalszych pracach wykorzystać pakiet TensorFlow. Według opinii autora jest to kompleksowe rozwiązanie i przewyższa w większości omówionych kryteriów platformę PyTorch. Pomimo braku możliwości dynamicznego definiowania grafów TensorFlow posiada znacznie liczniejszą grupę użytkowników oraz więcej źródeł pomocniczych. Co więcej pozwala on na korzystanie z wbudowanych narzędzi do wizualizacji, debugowania oraz wdrażania w systemach produkcyjnych sztucznych sieci neuronowych. Jest on również szerzej wspierany w platformie Google Cloud Platform, która zostanie wykorzystana do przeprowadzenia testów.

4.3. Biblioteki pomocnicze

Poniżej znajduje się opis bibliotek rozważanych do wykorzystania w eksperymentach dotyczących generowania predykcji notowań spółek na giełdzie papierów wartościowych.

4.3.1. Keras

Keras to biblioteka programistyczna o otwartym kodzie źródłowym przeznaczona do uczenia maszynowego. Została ona stworzona w języku programowania Python i stanowi interfejs wysokiego poziomu dla takich platform jak TensorFlow, CNTK

czy Theano. Keras charakteryzuje się następującymi właściwościami:

- Umożliwia dostęp do niemalże wszystkich funkcjonalności kompatybilnych platform.
- Pozwala na łatwe i szybkie tworzenie modeli uczenia maszynowego.
- Jest wykorzystywana w środowisku produkcyjnym.
- Umożliwia tworzenie modeli za pomocą gotowych modułów ze zdefiniowanymi parametrami. Moduły mogą stanowić na przykład abstrakcyjną reprezentację funkcji aktywacji, definiować funkcję kosztu czy poszczególne warstwy sieci neuronowej.

4.3.2. Scikit-learn

Scikit-learn to biblioteka programistyczna o otwartym kodzie źródłowym stworzona w języku programowania Python. Zawiera ona popularne algorytmy uczenia maszynowego. Projekt autorstwa Davida Cournapeau rozpoczął się w 2007 roku w ramach programu Google Summer of Code. Aktualnie biblioteka wydana jest na licencji BSD. Scikit-learn charakteryzuje się następującymi właściwościami:

- Wykorzystuje najpopularniejsze biblioteki związane z uczeniem maszynowym, takie jak NumPy, SciPy, Matplotlib, IPython, SamPy oraz Pandas.
- Jest elastyczna ze względu na kompatybilność z różnymi abstrakcyjnymi obiektami wchodzącymi w skład wykorzystywanych bibliotek. Umożliwia to na przykład dostarczenie danych do modelu zarówno w tablicach z pakietu Numpy, jak i w ramach danych pochodzących z biblioteki Pandas.

4.3.3. Pandas

Pandas to pakiet programistyczny zawierający dynamiczne, wydajne i elastyczne struktury danych. Jest on udostępniony na licencji BSD. Pandas charakteryzuje się następującymi właściwościami:

- Posiada szybkie i wydajne ramki danych służące do wykonywania operacji na zapisanych w nich obiektach.
- Oferuje dostęp do narzędzi umożliwiających odczyt, zapis i konwersję danych w różnych formatach.
- Automatyzuje procesy związane z uzupełnianiem brakujących danych i jest w stanie porządkować je według określonych reguł.
- Wspiera operacje na szeregach czasowych.
- Pozwala na elastyczne przekształcanie i manipulowanie zawartościami struktur danych.
- Umożliwia dynamiczną zmianę rozmiarów wierszy i kolumn oraz wykonywanie na nich określonych operacji.
- Etykietuje dane i rozdziela je na określone podzbiory.
- Osiąga wysoką wydajność wykonywanych obliczeń poprzez implementację kluczowych składowych w języku programowania Cython oraz C.
- Udostępnia wydajny mechanizm rozdzielania i łączenia zestawów danych.
- Potrafi operować na danych wielowymiarowych i skalować je do określonych potrzeb.

4.3.4. Numpy

Numpy to pakiet programistyczny o otwartym kodzie źródłowym przeznaczony do obliczeń naukowych. Udostępniony jest on na licencji BSD. Numpy charakteryzuje się następującymi właściwościami:

- Posiada wydajne tablice służące do wykonywania dużej ilości złożonych operacji matematycznych i logicznych na zapisanych w nich obiektach.
- Jest kompatybilny i wspierany przez wiele popularnych narzędzi przeznaczonych do obliczeń numerycznych wymagających wysokiej wydajności.

- Wykorzystuje statyczne tablice ze zdefiniowanymi rozmiarami. Zmiana ich wielkości lub wymiaru wymusza stworzenie nowej struktury.
- Przechowuje w tablicy jedynie dane tego samego typu.

4.3.5. Wybór bibliotek

Wszystkie omówione wyżej biblioteki pomocnicze stanowią doskonałe narzędzia do wykorzystania w eksperymentach związanych z przewidywaniem przyszłych notowań spółek na giełdzie papierów wartościowych. Zgodnie z założeniami autora biblioteka programistyczna Keras zostanie wykorzystana jako wysokopoziomowe API dla platformy TensorFlow. Z kolei biblioteki Scikit-learn, Pandas oraz Numpy będą podstawowymi narzędziami w całym procesie związanym z przetwarzaniem danych dotyczących historycznych kursów akcji oraz wygenerowanych na ich podstawie prognoz. W związku z tym wszystkie omówione wyżej narzędzia zostaną wykorzystane w dalszych badaniach.

4.4. Podsumowanie

W bieżącym rozdziale nastąpiło omówienie najpopularniejszych narzędzi umożliwiających przeprowadzenie eksperymentów dotyczących przewidywania notowań spółek. W pierwszej części zostały poddane analizie dwie najpopularniejsze platformy programistyczne przeznaczone dla sztucznych sieci neuronowych. Po dokładnym porównaniu obu środowisk postanowiono w dalszych badaniach wykorzystać TensorFlow. Następnie omówiona została biblioteka Keras pełniąca rolę wysokopoziomowego interfejsu dla TensorFlow. W kolejnym kroku dokonano analizy trzech najpopularniejszych bibliotek przeznaczonych dla języka Python i służących do przetwarzania danych - Scikit-learn, Pandas oraz Numpy. Ostatecznie uznano, że wszystkie opisane narzędzia posiadają unikalne funkcjonalności i zostaną wykorzystane na potrzeby przeprowadzenia eksperymentów.

Rozdział 5

Hipoteza błędzenia losowego oraz aktualne wyniki badań

5.1. Wstęp

We wstępie niniejszej pracy zostały zadane trzy następujące pytania:

- Czy rynek jest nieprzewidywalny?
- Czy poprzez wykorzystanie metod uczenia maszynowego można osiągnąć lepsze wyniki niż w przypadku losowego wybierania spółek do portfolio?
- Jakie metody uczenia maszynowego są najefektywniejsze i jaka jest ich skuteczność?

Znalezienie odpowiedzi na wyżej wspomniane pytania było celem przeprowadzonych eksperymentów. W związku z tym zagadnienia te zostały szerzej omówione w dalszej części bieżącego rozdziału.

5.2. Czy rynek jest nieprzewidywalny?

Jedną z najpopularniejszych teorii związanych z giełdą papierów wartościowych jest hipoteza Burtona Gordona Malkiela dotycząca błędzenia losowego. Zdefiniował on to pojęcie w książce wspomnianej we wstępie niniejszej pracy pt. "Błądząc

po Wall Street. Sprawdzona strategia skutecznego inwestowania” w następujący sposób:

”Spacer losowy polega na tym, że nie można przewidzieć przyszłych kroków lub kierunków na podstawie wcześniejszych działań. Zastosowanie tego terminu na giełdzie papierów wartościowych oznacza, że niemożliwe jest przewidzenie krótkoterminowych zmian cen akcji.” [13]

Zgodnie z powyższymi założeniami niemożliwe jest przewidzenie kursów akcji zamknięcia kolejnego dnia działania giełdy. W związku z tym przyjęte na potrzeby eksperymentów założenia dotyczące analizy technicznej w odniesieniu do wąskiego horyzontu czasowego powinny być całkowicie błędne. Hipoteza ta zostanie sprawdzona po przeprowadzeniu eksperymentów.

5.3. Czy poprzez wykorzystanie metod uczenia maszynowego można osiągnąć lepsze wyniki w wąskim horyzoncie czasowym niż w przypadku losowego wybierania spółek do portfolio?

W związku z omówioną hipotezą błędzenia losowego predykcje generowane przez sztuczne sieci neuronowe dotyczące ruchów notowań spółek zamknięcia kolejnego dnia działania giełdy papierów wartościowych powinny być co najwyżej tak samo skuteczne jak generator losowy. Hipoteza ta zostanie zweryfikowana po przeprowadzeniu eksperymentów.

5.4. Jakie algorytmy są najefektywniejsze i jaka jest ich skuteczność?

Obecnie można odnaleźć wiele badań dotyczących porównania skuteczności różnych modeli uczenia maszynowego w predykcji kursów akcji na giełdzie papierów wartościowych. Niestety niemożliwe jest określenie, jaka dokładnie metoda pozwala

na generowanie przewidywań z najwyższą skutecznością. Związane jest to z wieloma różnymi czynnikami. Jednym z nich jest brak możliwości przetestowania wszystkich dostępnych obecnie modeli i ich odmian. W dzisiejszych czasach istnieje niezliczona liczba algorytmów umożliwiających przewidywanie kursów akcji. Dodatkowo niektóre metody takie jak wybrane na potrzeby eksperymentów sztuczne sieci neuronowe mogą posiadać nieograniczoną liczbę możliwych struktur. W związku z tym jedynym rozwiązaniem wydaje się przegląd dokonanych eksperymentów oraz analiza dostępnych materiałów. Niestety o ile w ramach jednego eksperymentu wykonywanego w tych samych warunkach dla wszystkich metod można wskazać najskuteczniejsze algorytmy, to już w przypadku porównywania efektywności metod z różnych badań problem jest znacznie bardziej złożony. Najczęściej testy wykonywane są w różnych środowiskach. To co bardzo dobrze sprawdza się w przewidywaniu inwestycji długoterminowych, niekoniecznie może stanowić dobre narzędzie do generowania predykcji w ramach spekulacji. Ponadto nawet przy założeniu, że porównywane zostają techniki z różnych publikacji dotyczących generowania predykcji notowań spółek kolejnego dnia zamknięcia giełdy, to nie bierze się pod uwagę ciągłej zmienności rynku. Algorytm testowany w czasach bessy i małych zmian na rynku może dać skrajnie odmienne rezultaty w porównaniu z wynikami tego samego algorytmu testowanymi w czasach hossy i dużej zmienności gospodarczej. Dodatkowo w badaniach wykorzystuje się różne metody walidacji. Jedne eksperymenty i ich rozwiązania mogą być sprawdzane w oparciu o kursy jedynie wybranej spółki, z kolei w innych badaniach testy mogą odbywać się na podstawie wielu przedsiębiorstw. Oczywiście należy być również świadomym, że wiele opracowanych algorytmów przez duże przedsiębiorstwa jest poufne i ich skuteczność nie jest ujawniana publicznie. Przyniesione wyżej argumenty świadczą o braku możliwości wskazania najefektywniejszych algorytmów oraz ich skuteczności w generowaniu predykcji spółek. W związku z tym według opinii autora porównanie efektywności algorytmu opracowywanego w ramach niniejszej pracy ze skutecznością innych dostępnych metod nie byłoby miarodajne. Z tego powodu w ramach bieżącej pracy postanowiono jedynie wesprzeć się aktualnymi wynikami badań w wyborze metody uczenia maszynowego na potrzeby przeprowadzenia eksperymentów. Otrzymane wyniki nie będą jednak skonfrontowane z rezultatami osiąganymi przez inne algorytmy.

5.5. Podsumowanie

W bieżącym rozdziale nastąpiło omówienie hipotez stanowiących cel przeprowadzanych badań. W związku z tym w dalszej części pracy uwaga zostanie skupiona na eksperymentach umożliwiającym znalezienie odpowiedzi na pytania związane z hipotezą błędzenia losowego w kontekście przewidywania notowań spółek na giełdzie papierów wartościowych w krótkim horyzoncie czasowym.

Rozdział 6

Przeprowadzone eksperymenty

6.1. Wstęp

W pierwszym rozdziale niniejszej pracy został dokonany wybór strategii inwestycyjnej polegającej na analizie technicznej oraz podejmowaniu transakcji w reakcji na sytuację rynkową w krótkim horyzoncie czasowym. Następnie w oparciu o bieżący stan wiedzy oraz badania naukowe dokonano wyboru metody uczenia maszynowego na potrzeby generowania predykcji notowań spółek zamknięcia kolejnego dnia działania giełdy. Ostatecznie postanowiono w eksperymentach wykorzystać rekurencyjne sieci neuronowe z pamięcią skojarzeniową. W kolejnym kroku zostały wyselekcjonowane narzędzia umożliwiające realizację opisanych założeń. Ostatni etap polegał na dokładnym zdefiniowaniu problemów i hipotez mających na celu określenie idei niniejszej pracy oraz badań. W bieżącym rozdziale nadszedł czas na przeprowadzenie eksperymentów oraz interpretację uzyskanych wyników umożliwiających znalezienie odpowiedzi na postawione w poprzednim rozdziale pytania.

6.2. Cel

Celem przeprowadzenia niniejszych eksperymentów jest stworzenie algorytmu umożliwiającego generowanie predykcji kursów akcji zamknięcia kolejnego dnia działania giełdy dla wszystkich spółek wchodzących w skład indeksu S&P 500 przy wykorzystaniu możliwie najbardziej efektywnej struktury rekurencyjnej sieci neuronowej

z pamięcią skojarzeniową oraz historycznych notowań spółek.

6.3. Założenia

Celem eksperymentów jest wybranie najlepszej architektury dla sztucznych sieci neuronowych z punktu widzenia omawianego zagadnienia. W niniejszych eksperymentach zastosowano metodę prób i błędów. Pod uwagę wzięto szereg potencjalnie atrakcyjnych architektur z punktu widzenia efektywności generowania predykcji notowań spółek. Wyłonienie najskuteczniejszego rozwiązania nastąpiło w oparciu o określoną metodę walidacji. Wyselekcjonowany model został wykorzystany do generowania predykcji notowań spółek zamknięcia kolejnego dnia działania giełdy dla przedsiębiorstw wchodzących w skład indeksu S&P 500. W oparciu o przewidziane wartości następował wybór instrumentów finansowych umożliwiających uzyskanie najwyższych możliwych przychodów. Zgodnie z założeniem zakup walorów odbywał się pod koniec dnia działania giełdy, a jego sprzedaż następowała o tej samej porze w czasie kolejnej sesji. Cały proces został zrealizowany w poszczególnych etapach zgodnie z podaną kolejnością:

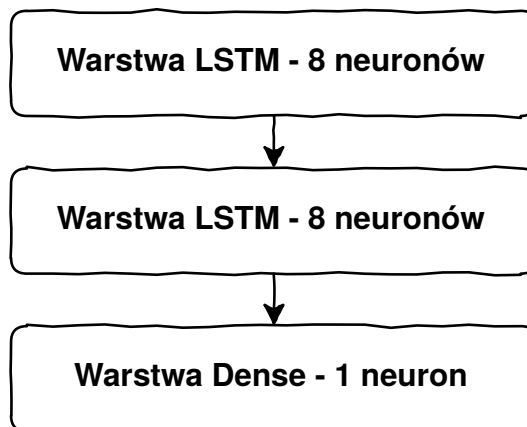
- Zgromadzenie danych - wszystkie dane niezbędne do przeprowadzenia eksperymentów zostały pobrane 25 sierpnia 2019 roku z zewnętrznej platformy. Notowania spółek wchodzących w skład indeksu S&P 500 dotyczą ostatnich 100 dni działania giełdy.
- Wstępne przetworzenie danych - zgromadzone dane zostały przekształcone do formatu umożliwiającego trenowanie oraz testowanie wszystkich modeli sztucznych sieci neuronowych z pamięcią skojarzeniową. Zgodnie z założeniami dotyczącymi analizy jedynie kursów zamknięcia, postanowiono we wstępnym przetwarzaniu danych odrzucić wszystkie pozostałe informacje dotyczące notowania otwarcia, najwyższej i najniższej wartości przedsiębiorstwa oraz liczby akcji, jaka zmieniła właściciela danego dnia. Dodatkowo sztuczne sieci neuronowe będą generowały kurs akcji zamknięcia kolejnej sesji w oparciu o dane historyczne z ostatnich 29 dni działania giełdy papierów wartościowych. W związku z tym wartości spółek przechowywane są w pliku zawierającym

30 kolumn. W każdym wierszu pierwsze 29 wartości stanowi dane wejściowe, a ostatnia wartość w 30 kolumnie jest oczekiwanym wynikiem. Ponadto wszystkie wiersze w pliku zostały poddane procesowi randomizacji.

- Podział na zbiór treningowy i testowy — w celu umożliwienia nauczenia oraz sprawdzenia skuteczności generowania predykcji przez sztuczne sieci neuronowe cały plik z danymi został podzielony losowo na zbiór treningowy składający się z 80% wszystkich wierszy (28003 linii) oraz zbiór testowy zawierający pozostałe 20% prób (7001 linii).
- Stworzenie instancji sztucznych sieci neuronowych z pamięcią skojarzeniową - do przeprowadzenia eksperymentów zostało wyselekcjonowanych 10 architektur o strukturach przedstawionych w dalszej części bieżącego podpunktu. Na rysunkach obrazujących modele przyjęto następujące oznaczenia:
 - Warstwa LSTM - warstwa neuronów rekurencyjnych
 - Warstwa Dense - warstwa neuronów nieposiadających sprzężeń zwrotnych

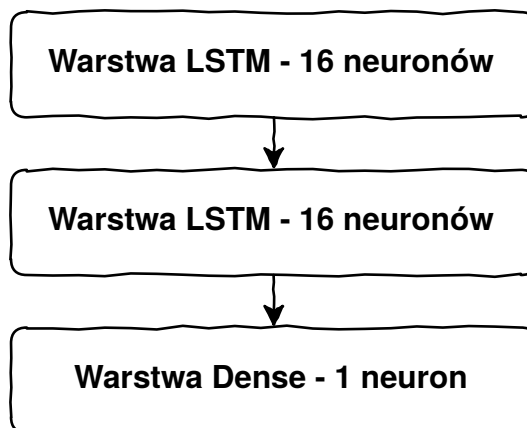
Wykorzystane w eksperymentach architektury sztucznych sieci neuronowych z pamięcią skojarzeniową prezentują się w sposób następujący:

- Schemat przedstawiający pogładową architekturę modelu numer 1 znajduje się na rysunku 6.1.



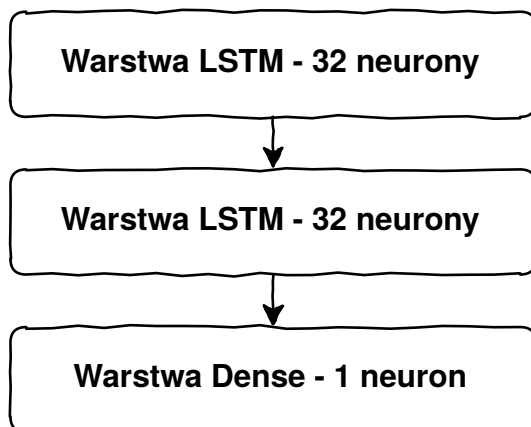
Rys. 6.1. Schemat architektury modelu numer 1.

- Schemat przedstawiający pogładową architekturę modelu numer 2 znajduje się na rysunku 6.2.



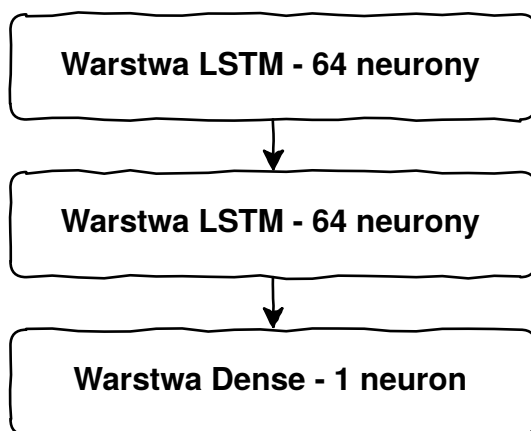
Rys. 6.2. Schemat architektury modelu numer 2.

- Schemat przedstawiający poglądową architekturę modelu numer 3 znajduje się na rysunku 6.3.



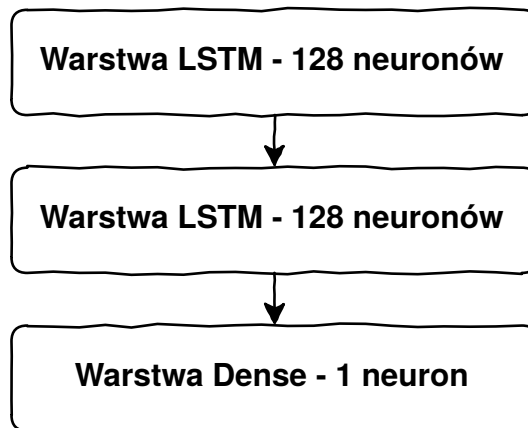
Rys. 6.3. Schemat architektury modelu numer 3.

- Schemat przedstawiający poglądową architekturę modelu numer 4 znajduje się na rysunku 6.4.



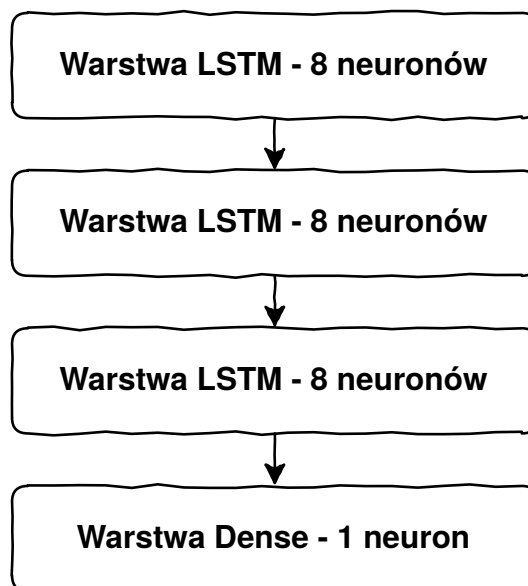
Rys. 6.4. Schemat architektury modelu numer 4.

- Schemat przedstawiający pogładową architekturę modelu numer 5 znajduje się na rysunku 6.5.



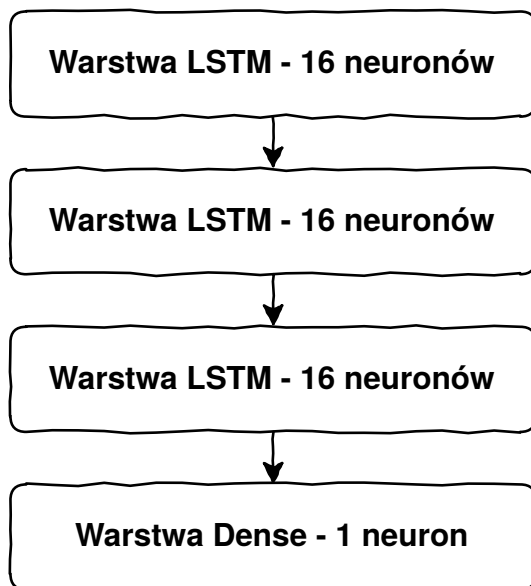
Rys. 6.5. Schemat architektury modelu numer 5.

- Schemat przedstawiający pogładową architekturę modelu numer 6 znajduje się na rysunku 6.6.



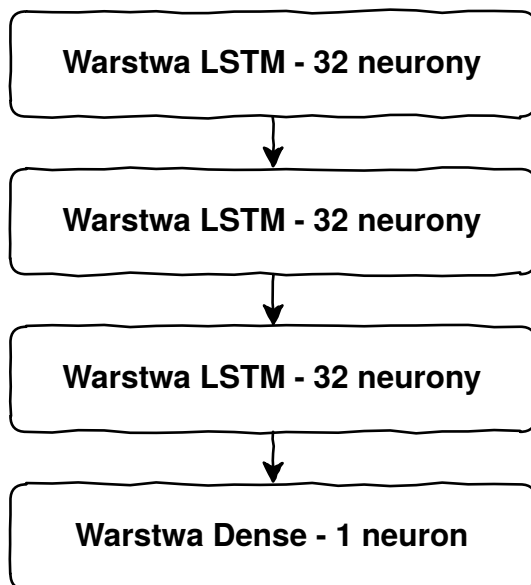
Rys. 6.6. Schemat architektury modelu numer 6.

- Schemat przedstawiający poglądową architekturę modelu numer 7 znajduje się na rysunku 6.7.



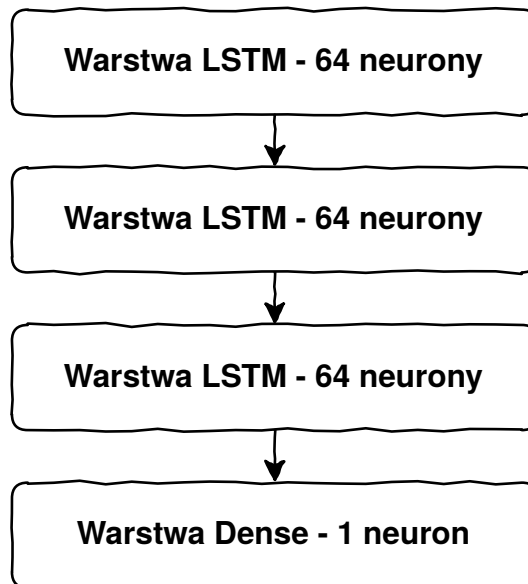
Rys. 6.7. Schemat architektury modelu numer 7.

- Schemat przedstawiający poglądową architekturę modelu numer 8 znajduje się na rysunku 6.8.



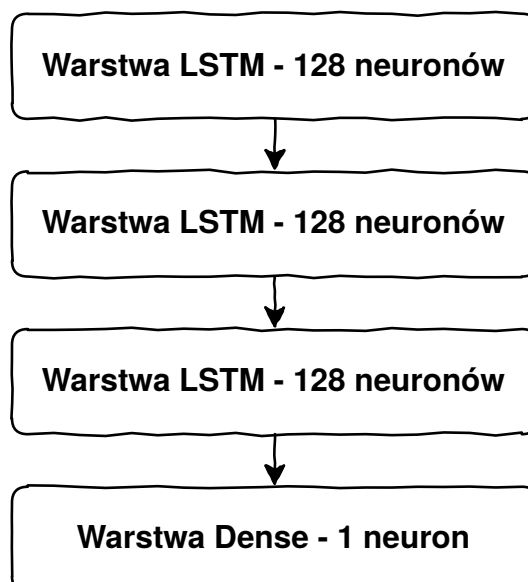
Rys. 6.8. Schemat architektury modelu numer 8.

- Schemat przedstawiający pogładową architekturę modelu numer 9 znajduje się na rysunku 6.9.



Rys. 6.9. Schemat architektury modelu numer 9.

- Schemat przedstawiający pogładową architekturę modelu numer 10 znajduje się na rysunku 6.10.



Rys. 6.10. Schemat architektury modelu numer 10.

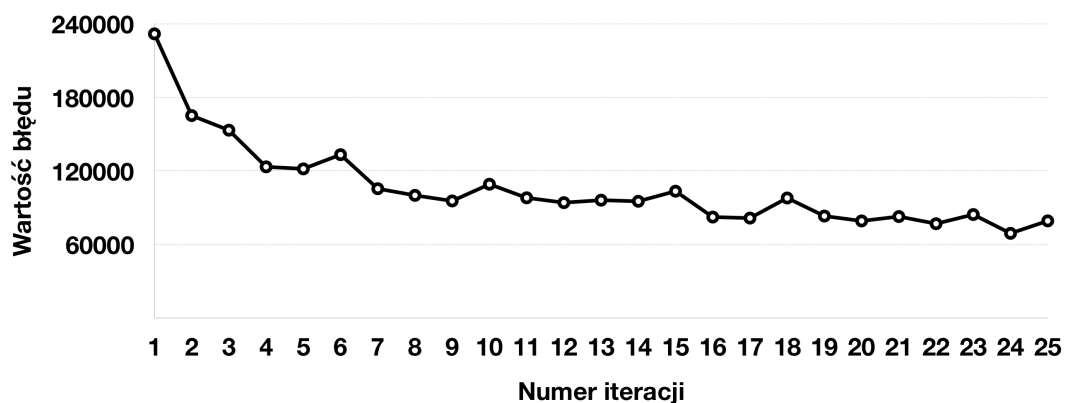
- Nauka sztucznych sieci neuronowych - etap ten opiera się na opracowanym we wcześniejszych krokach zbiorze trenującym umożliwiającym dostosowywanie odpowiednich wartości wag we wszystkich 10 architekturach sztucznych sieci neuronowych na podstawie danych wejściowych i wyjściowych.
- Walidacja sztucznych sieci neuronowych - etap ten następuje po każdej iteracji dotyczącej procesu nauki w oparciu o 28003 prób uczących. System wykorzystuje opracowany zbiór testowy do sprawdzenia skuteczności predykcji każdego z 10 modeli sztucznych sieci neuronowych. Odbywa się to poprzez obliczenie różnicy pomiędzy wygenerowanym wynikiem a oczekiwanym kursem dla każdej spośród 7001 prób testowych i zsumowaniem tych wartości odrębnie dla wszystkich 10 struktur. Niższa wartość błędu oznacza wyższą skuteczność generowanych predykcji przez opracowany model.
- Wybór najlepszego modelu - proces ten polega na sprawdzeniu wszystkich 10 wyników ostatniej iteracji i wybraniu najlepszego modelu posiadającego najmniejszą wartość błędu. Jeżeli wynik ten jest niższy od najlepszego wyniku poprzednich iteracji to następujące nadpisanie aktualnego modelu na strukturze poprzedniego. Dzięki temu w ostatecznym kroku zawsze otrzymany zostaje model z najlepszą ze wszystkich badanych architektur z najwyższą uzyskaną skutecznością we wszystkich przeprowadzonych iteracjach.
- Wybór najbardziej dochodowych spółek - wyselekcjonowany we wcześniejszym etapie najlepszy model rekurencyjnej sieci neuronowej z pamięcią skojarzeniową zostaje wykorzystany do wygenerowania predykcji kolejnego dnia działania giełdy dla spółek wchodzących w skład indeksu S&P 500. W kolejnym kroku następuje wyliczenie dla każdego przedsiębiorstwa różnicy pomiędzy ostatnim kursem zamknięcia dnia działania giełdy, a przewidzianą wartością zakończenia kolejnej sesji. Umożliwia to wyznaczenie oczekiwanej procentowej zmiany wartości dla każdego instrumentu finansowego względem dnia poprzedniego. Na podstawie otrzymanych wyników wyznaczane są walory umożliwiające z założenia osiągnięcie najwyższych możliwych zysków. W przeprowadzonych eksperymentach postanowiono dokonać dywersyfikacji całego dostępnego kapitału na 7 równych części. Zgodnie z tym założeniem każdego dnia działania giełdy

wybijamy i nabywamy akcje 7 spółek za tę samą kwotę oraz sprzedajemy pod koniec sesji papiery wartościowe 7 przedsiębiorstw nabytych w czasie poprzedniej sesji. Zabieg ten umożliwi zwiększenie stabilności oraz wiarygodności otrzymanych wyników przez opracowany algorytm, a uzyskane zyski lub straty nie będą dziełem przypadku.

6.4. Wyniki

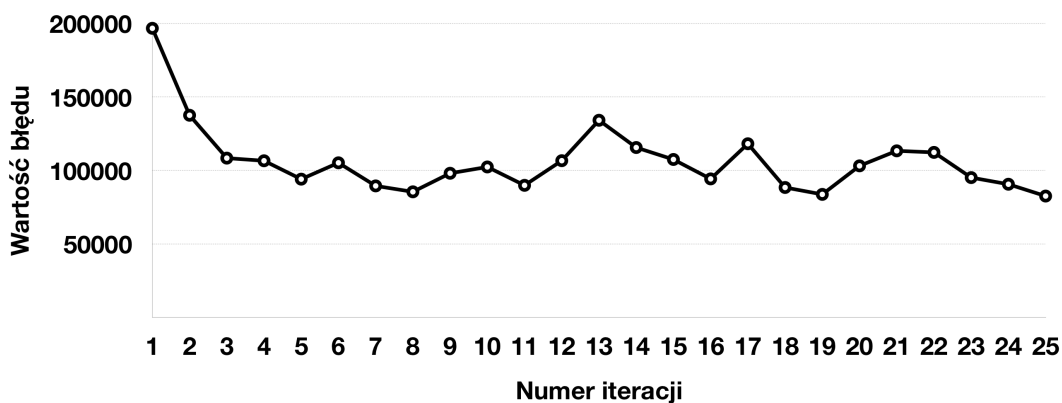
W dalszej części bieżącej sekcji znajdują się wyniki przeprowadzonych eksperymentów. W badaniach zostało wykonanych 25 pełnych iteracji przy wykorzystaniu całego zbioru treningowego. Zgodnie z przyjętą metodyką ewaluacji opisaną we wcześniejszym fragmencie niniejszego rozdziału wszystkie iteracje zakończone były procesem weryfikacji skuteczności każdej z architektur na podstawie zbioru testowego. Niższa uzyskiwana wartość błędu oznaczała wyższą skuteczność generowanych predykcji. W oparciu o uzyskane wyniki powstały poniższe wykresy odzwierciedlające cały proces nauki każdej opracowanej na potrzeby przeprowadzenia eksperymentów rekurencyjnej sieci neuronowej z pamięcią skojarzeniową.

- Wyniki uzyskane przez model numer 1 znajdują się na rysunku 6.11.



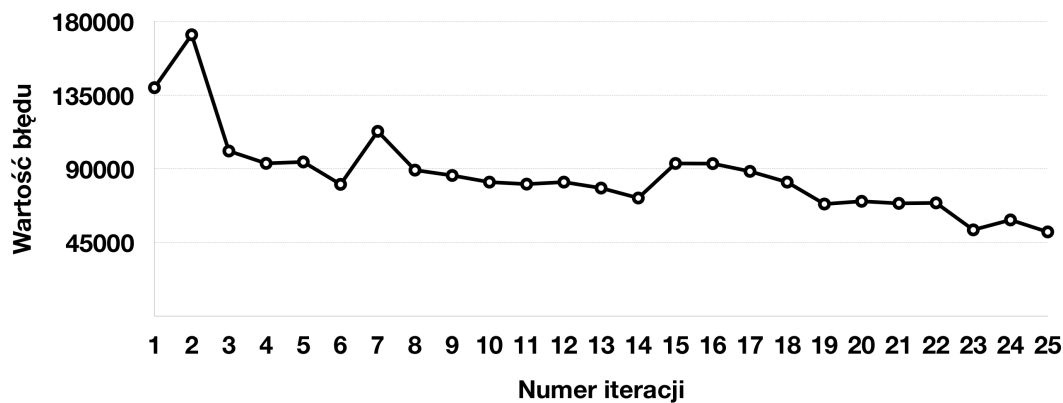
Rys. 6.11. Wykres przedstawiający zależność uzyskanych wyników przez model numer 1 od ilości iteracji.

- Wyniki uzyskane przez model numer 2 znajdują się na rysunku 6.12.



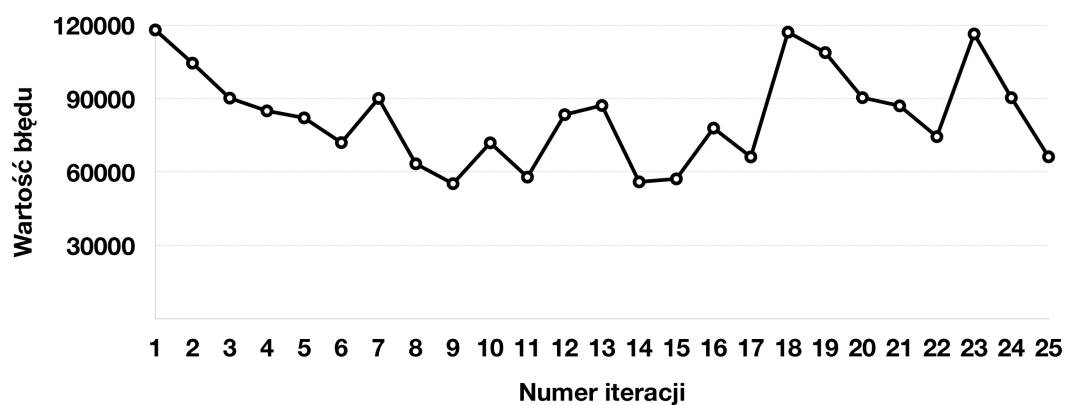
Rys. 6.12. Wykres przedstawiający zależność uzyskanych wyników przez model numer 2 od ilości iteracji.

- Wyniki uzyskane przez model numer 3 znajdują się na rysunku 6.13.



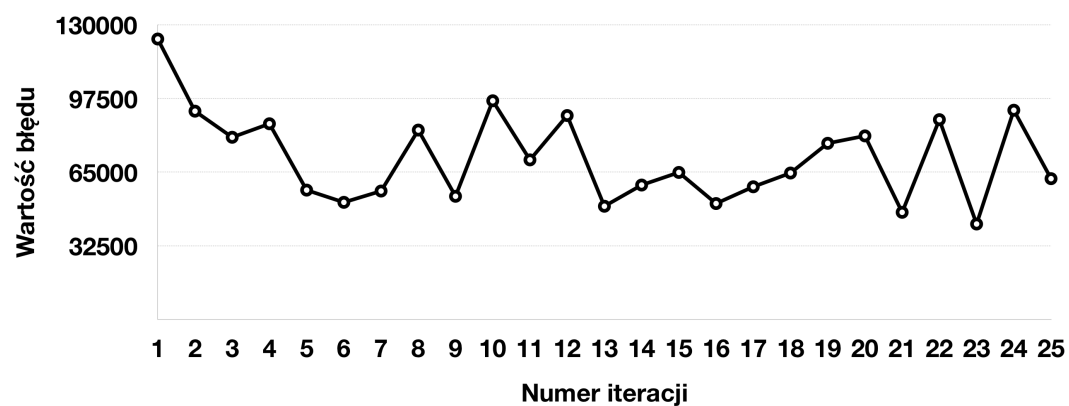
Rys. 6.13. Wykres przedstawiający zależność uzyskanych wyników przez model numer 3 od ilości iteracji.

- Wyniki uzyskane przez model numer 4 znajdują się na rysunku 6.14.



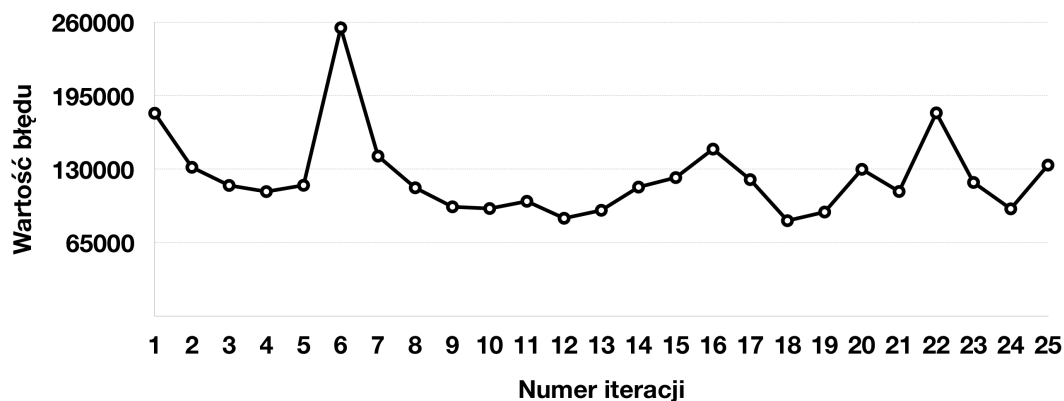
Rys. 6.14. Wykres przedstawiający zależność uzyskanych wyników przez model numer 4 od ilości iteracji.

- Wyniki uzyskane przez model numer 5 znajdują się na rysunku 6.15.



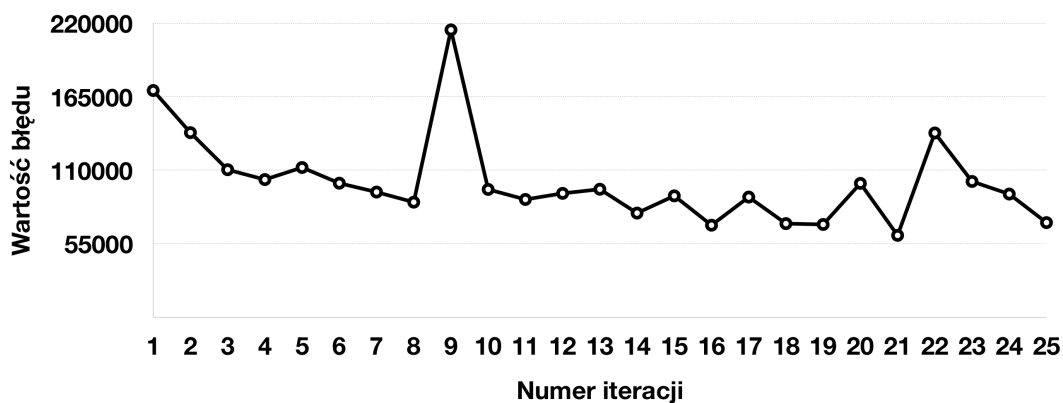
Rys. 6.15. Wykres przedstawiający zależność uzyskanych wyników przez model numer 5 od ilości iteracji.

- Wyniki uzyskane przez model numer 6 znajdują się na rysunku 6.16.



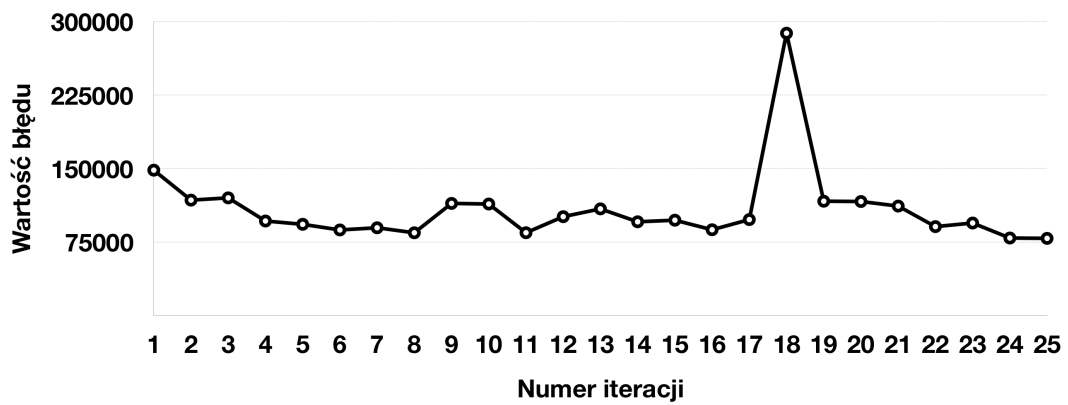
Rys. 6.16. Wykres przedstawiający zależność uzyskanych wyników przez model numer 6 od ilości iteracji.

- Wyniki uzyskane przez model numer 7 znajdują się na rysunku 6.17.



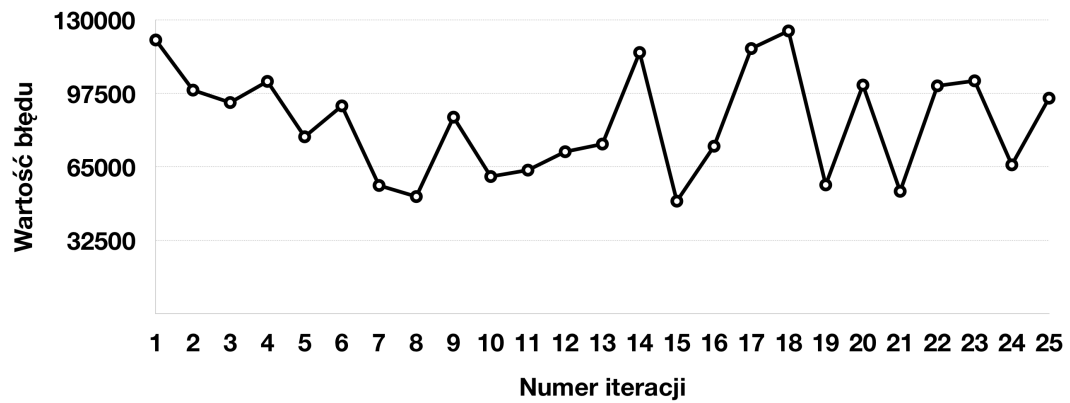
Rys. 6.17. Wykres przedstawiający zależność uzyskanych wyników przez model numer 7 od ilości iteracji.

- Wyniki uzyskane przez model numer 8 znajdują się na rysunku 6.18.



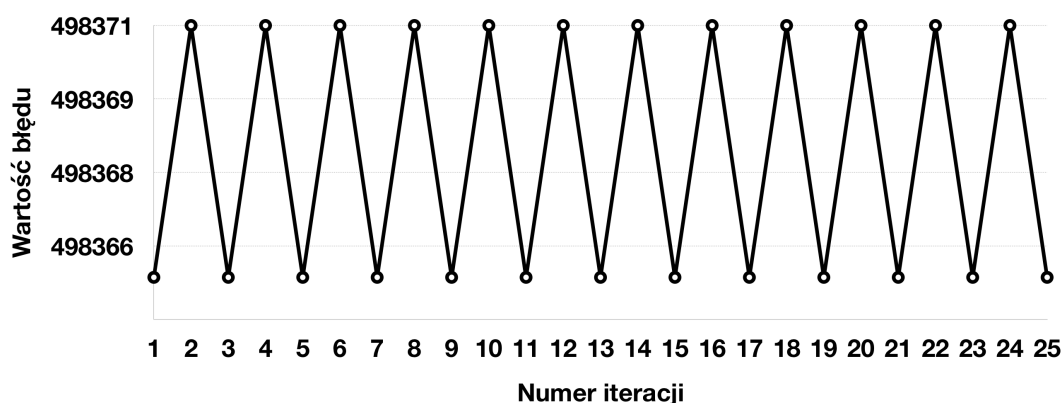
Rys. 6.18. Wykres przedstawiający zależność uzyskanych wyników przez model numer 8 od ilości iteracji.

- Wyniki uzyskane przez model numer 9 znajdują się na rysunku 6.19.



Rys. 6.19. Wykres przedstawiający zależność uzyskanych wyników przez model numer 9 od ilości iteracji.

- Wyniki uzyskane przez model numer 10 znajdują się na rysunku 6.20.



Rys. 6.20. Wykres przedstawiający zależność uzyskanych wyników przez model numer 10 od ilości iteracji.

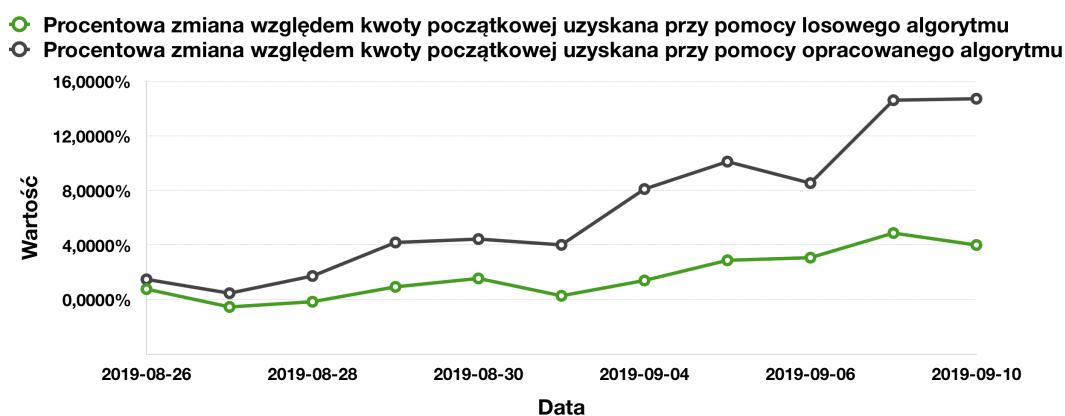
Zgodnie z wynikami zamieszczonymi na powyższych wykresach najskuteczniejsza w predykcji kursów akcji okazała się architektura modelu rekurencyjnej sieci neuronowej z pamięcią skojarzeniową o numerze 5. Zdecydowanie osiągała ona najlepsze wyniki spośród wszystkich testowanych rozwiązań niemalże we wszystkich przeprowadzonych testach. Najniższy wynik błędów został uzyskany podczas 23 iteracji i to właśnie ten model wyselekcjonowano do użycia w algorytmie wyboru spółek umożliwiających osiągnięcie najwyższych zysków. Aby możliwa była weryfikacja skuteczności działania systemu postanowiono na potrzeby przeprowadzenia testów opracować również losowy algorytm oparty o tę samą ideę działania. Jedynym wyjątkiem jest brak wykorzystania w nim rekurencyjnej sieci neuronowej. Selekcja 7 spółek wchodzących w skład indeksu S&P 500 każdego dnia odbywa się na podstawie wyboru funkcji pseudolosowej. Według opinii autora takie podejście jest najbardziej wiarygodnym i miarodajnym sposobem weryfikacji poprawności działania opracowanego rozwiązania oraz skuteczności generowanych przez niego predykcji. W związku z tym zgodnie z przyjętymi założeniami w przeprowadzonych eksperymentach na zakończenie każdego dnia działania giełdy następował zakup 7 akcji wyselekcjonowanych przez opracowany algorytm oraz 7 walorów wybranych na podstawie funkcji pseudolosowej. W czasie zamknięcia kolejnej sesji następowało zbywanie wszystkich 14 instrumentów finansowych i zakup kolejnych papierów wartościowych dotyczących ponownie wybranych w tym samym procesie postępowania spółek. Przeprowadzone badania trwały ponad dwa tygodnie i uwzględniły 11 dni działania gieł-

dy oraz 154 zrealizowane transakcje. Tabela przedstawiająca otrzymane rezultaty przez oba algorytmy znajduje się na rysunku 6.20.

DATA	PROCENTOWY ZMIANA WZGLĘDEM KWOTY POCZĄTKOWEJ UZYSKANA PRZY POMOCY LOSOWEGO ALGORYTMU	PROCENTOWA ZMIANA WZGLĘDEM KWOTY POCZĄTKOWEJ UZYSKANA PRZY POMOCY OPRACOWANEGO ALGORYTMU
2019-08-26	0,7630%	1,4764%
2019-08-27	-0,5453%	0,4622%
2019-08-28	-0,1617%	1,7221%
2019-08-29	0,9282%	4,1802%
2019-08-30	1,5453%	4,4310%
2019-09-03	0,2718%	4,0041%
2019-09-04	1,3949%	8,0956%
2019-09-05	2,8746%	10,1062%
2019-09-06	3,0670%	8,5307%
2019-09-09	4,8710%	14,6088%
2019-09-10	3,9947%	14,7159%

Tab. 6.20. Tabela przedstawiająca wyniki przeprowadzonych eksperymentów.

Zamieszczone wartości w kolumnach są uśrednionymi zyskami oraz stratami uzyskiwanymi na zakończenie każdej sesji oraz odzwierciedlają procentową zmianę kapitału względem jego stanu początkowego. Graficzna reprezentacja otrzymanych wyników w ujęciu czasowym została przedstawiona w formie wykresu na rysunku 6.21.



Rys. 6.21. Wykres przedstawiająca wyniki przeprowadzonych eksperymentów.

Zgodnie z informacjami zamieszczonymi powyżej stworzony algorytm w czasie przeprowadzania eksperymentów oraz 11 dni działania giełdy wypracował zysk równy 14,7159% względem wartości kapitału początkowego. Jest to rezultat zdecydowanie ponadprzeciętny i ponad trzykrotnie wyższy od wyniku wynoszącego 3,9947%,

który został uzyskany przez algorytm oparty na losowym wybieraniu spółek do portfela inwestycyjnego. Na wykresie można dostrzec odzwierciedlenie ogólnej tendencji panującej na rynku w czasie trwania eksperymentów, jednak niemalże każdego dnia opracowane rozwiązanie wybierało akcje spółek przynoszące znacznie wyższe zyski. Ponadto od pierwszego do ostatniego dnia trwania badań opracowany algorytm posiadał wyższą procentową wartość względem wartości kapitału początkowego.

6.5. Wnioski

Przeprowadzone eksperymenty oraz wyciągnięte na ich podstawie wnioski umożliwiły weryfikację hipotez oraz odnalezienie odpowiedzi na pytania stanowiące cel niniejszej pracy. Wszystkie konkluzje zostały zawarte oraz szerzej omówione w dalszej części bieżącego podrozdziału.

- Czy rynek jest nieprzewidywalny?

Zgodnie z założeniami hipotezy losowej niemożliwe jest przewidzenie kursów akcji zamknięcia kolejnego dnia działania giełdy w oparciu o historyczne notowania spółek. Wyniki przeprowadzonych eksperymentów zdają się zaprzeczać tej teorii. Niemalże wszystkie ze 308 zrealizowanych transakcji w oparciu o opracowany algorytm w ramach omówionych w pracy badań umożliwiło osiągnięcie nadzwyczajnie wysokich zysków w związku z poprawnym przewidzeniem trendu kolejnego dnia zamknięcia giełdy. Ponadprzeciętna skuteczność generowanych predykcji w tak wysokiej liczbie przeprowadzonych transakcji w żaden sposób nie wskazuje na dzieło przypadku w otrzymanych rezultatach. W związku z tym według opinii autora niniejszej publikacji hipoteza błędzenia losowego w kontekście spekulacji na giełdzie papierów wartościowych została sfalsyfikowana.

- Czy poprzez wykorzystanie metod uczenia maszynowego można osiągnąć lepsze wyniki niż w przypadku losowego wybierania spółek do portfolio?

Zgodnie z przyjętymi założeniami na potrzeby hipotezy błędzenia losowego predykcje generowane przez sztuczne sieci neuronowe dotyczące kursów akcji zamknięcia kolejnego dnia działania giełdy powinny przynosić niższe lub po-

dobne zyski w porównaniu do tych wypracowywanych przez algorytm losowy. Okazuje się jednak, że stworzone na potrzeby eksperymentów rozwiązanie w trwających ponad dwa tygodnie badaniach uzyskało ponadprzeciętny zysk przewyższający rezultat algorytmu losowego niemalże czterokrotnie. Ponadto w 9 z 11 badanych sesji podczas przeprowadzania eksperymentów uśredniona wartość zysku wszystkich zrealizowanych transakcji była wyższa w porównaniu do wyniku otrzymanego przez algorytm losowy. Według opinii autora przytoczone argumenty świadczą o nieprawdziwości hipotezy błędzenia losowego odnoszącej się spekulacji na giełdzie papierów wartościowych. Uzyskane wyniki jednoznacznie wskazują bowiem na możliwość przewidzenia przyszłych trendów oraz wartości kursów akcji w wąskim horyzoncie czasowym w oparciu o historyczne notowania spółek.

- Jakie metody uczenia maszynowego są najefektywniejsze i jaka jest ich skuteczność?

Zgodnie z informacjami zamieszczonymi w rozdziale pt. "Hipoteza błędzenia losowego oraz aktualne wyniki badań" według opinii autora niemożliwe jest w dzisiejszych czasach zbadanie skuteczności wszystkich możliwych metod uczenia maszynowego oraz wyłonienie najlepszej z nich w kontekście przewidywania kursów akcji na giełdzie papierów wartościowych. Jednak autor niniejszej pracy postanowił dokonać analizy wielu publikacji oraz przeprowadzonych badań w celu wyłonienia możliwie najbardziej perspektywistycznej techniki umożliwiającej przewidywanie szeregów czasowych z uwzględnieniem bieżących potrzeb i wymagań stawianych w ramach przeprowadzanych eksperymentów. Ostatecznie w dalszych testach postanowiono wykorzystać rekurencyjną sieć neuronową z pamięcią skojarzeniową. Ponadto zbadano jej szereg potencjalnie atrakcyjnych architektur z punktu widzenia efektywności generowania predykcji, a następnie nastąpiło wyłonienie najskuteczniejszego rozwiązania. Okazała się nim struktura zawierająca dwie warstwy LSTM. Każda z nich składa się ze 128 neuronów. Ponadto w architekturze zastosowano warstwę wyjściową Dense złożoną z jednego neuronu. Model ten umożliwił osiągnięcie zysku równego 14,7159% względem wartości kapitału początkowego w czasie 11 dni działania giełdy i to właśnie tę metodę uznano za najskutecz-

niejszą w kontekście generowania predykcji kursów akcji zamknięcia kolejnego dnia działa giełdy papierów wartościowych.

6.6. Podsumowanie

W niniejszym rozdziale nastąpiło omówienie wszystkich przeprowadzonych eksperymentów mających na celu predykcję kursów akcji kolejnego dnia działania giełdy dla spółek wchodzących w skład indeksu S&P 500. W pierwszej kolejności zdefiniowano cel badań. Następnie dokonano analizy wszystkich etapów prac zgodnie z kolejnością ich realizacji. W kolejnym kroku zostało przedstawionych 10 architektur rekurencyjnej sieci neuronowych z pamięcią skojarzeniową rozpatrywanych w ramach wyboru najskuteczniejszego modelu. Ostatecznie po przeprowadzeniu eksperymentów najlepsza okazała się struktura zawierająca dwie warstwy LSTM po 128 neuronów oraz jedną warstwę wyjściową Dense składającą się z 1 neuronu. W oparciu o opracowany algorytm wykorzystujący wyselekcjonowany model rekurencyjnej sieci neuronowej z pamięcią skojarzeniową zostały przeprowadzone badania trwające ponad dwa tygodnie i uwzględniające 11 dni działania giełdy. Po tym etapie nastąpiło wyciągnięcie wniosków na podstawie otrzymanych rezultatów w wyniku realizacji 308 transakcji kupna i sprzedaży akcji. Ostatecznie hipoteza błędzenia losowego została obalona w oparciu o uzyskany zysk równy 14,7159% przez opracowany algorytm względem wartości kapitału początkowego oraz w odniesieniu do zysku 3,9947% wypracowanego przez algorytm losowy.

Rozdział 7

Podsumowanie

Celem niniejszej pracy było wykorzystanie metod uczenia maszynowego w predykcji kursów akcji na giełdzie papierów wartościowych. W pierwszej kolejności nastąpił wstęp do lokowania kapitału w akcje spółek. Następnie omówione zostały dwie najpopularniejsze analizy i strategie wspierające decyzje o zakupie lub sprzedaży instrumentów finansowych. Ostatecznie na potrzeby przeprowadzania eksperymentów postanowiono porzucić analizę fundamentalną na rzecz analizy technicznej i strategii transakcji w reakcji na sytuację rynkową. W kolejnym rozdziale dokonano analizy wszystkich dostępnych sposobów zdobywania wiedzy przez systemy uczenia maszynowego. Ostatecznie postanowiono, że najbardziej obiecującym podejściem z perspektywy generowania przyszłych notowań spółek na giełdzie papierów wartościowych będzie technika oparta na uczeniu nadzorowanym. Po omówieniu metod uczenia maszynowego wspierających omówione podejście nauki zdecydowano się dalszych eksperymentach wykorzystać rekurencyjną sieć neuronową z pamięcią skojarzeniową. Ponadto po dokonaniu przeglądu najpopularniejszych narzędzi ostatecznie zdecydowano się zrealizować eksperymenty w oparciu o takie biblioteki jak TensorFlow, Keras, Scikit-Learn, Pandas oraz Numpy. W kolejnym rozdziale nastąpiło postawienie hipotez oraz określenie problemów stanowiących cel przeprowadzanych eksperymentów. Ostatecznie na podstawie uzyskanych wyników oraz płynących z nich wniosków sfalsyfikowano hipotezę błędzenia losowego w generowaniu przyszłych trendów oraz wartości notowań spółek w wąskim horyzoncie czasowym na giełdzie papierów wartościowych.

Bibliografia

- [1] 10 (nie)znanych cytatów jesse'ego livermore'a. <https://comparic.pl/10-nieznanych-cytatow-jesseego-livermorea/>, 2016.
- [2] Warren buffett - wikicytaty. https://pl.wikiquote.org/wiki/Warren_Buffett, 2017.
- [3] ALBON, C. *Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning*. 2018.
- [4] BENGIO, Y., SIMARD, P., AND FRASCONI, P. Learning long-term dependencies with gradient descent is difficult, 1994.
- [5] CHATZIS, S. P., SIAKOULIS, V., PETROPOULOS, A., STAVROULAKIS, E., AND VLACHOGIANNAKIS, N. "forecasting stock market crisis events using deep and statistical machine learning techniques", 2018.
- [6] CHOUDHURY, A. Tensorflow vs pytorch: Top 10 differences between the two ml libraries. <https://www.analyticsindiamag.com/tensorflow-vs-pytorch-top-10-differences-between-the-two-ml-libraries/>, 2019.
- [7] CONWAY, D., AND WHITE, J. M. *Machine Learning for Hackers*. 2012.
- [8] FISCHER, T., AND KRAUSS, C. Deep learning with long short-term memory networks for financial market predictions, 2017.
- [9] GRAHAM, B. *Inteligentny inwestor*. 2006.
- [10] GÉRON, A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. 2017.

-
- [11] LASKA, J., AND ŁYSIAK, T. Budowa aplikacji webowej umożliwiającej obrót instrumentami finansowymi na giełdzie, 2018.
- [12] LULA, P., MORAJDA, J., PALIWODA-PĘKOSZ, G., STAL, J., TADEUSIEWICZ, R., AND WILUSZ, W. *Komputerowe metody analizy i przetwarzania danych*. 2012.
- [13] MALKIEL, B. G. *A Random Walk Down Wall Street*. 1973.
- [14] MITCHELL, T. *Machine Learning*. 1997.
- [15] MURPHY, J. J. *Analiza techniczna rynków finansowych*. 2006.
- [16] NELSON, D. M. Q., DE OLIVEIRA, R. A., AND PEREIRA, A. C. M. Stock market's price movement prediction with lstm neural networks, 2017.
- [17] ROONDIWALA, M., PATEL, H. J., AND VARMA, S. Predicting stock prices using lstm, 2017.
- [18] SHAH, D., CAMPBELL, W., AND ZULKERNINE, F. H. A comparative study of lstm and dnn for stock market forecasting, 2018.

Spis rysunków

2.1. Poglądowy schemat przedstawiający ideę uczenia nadzorowanego.	16
2.2. Poglądowy schemat przedstawiający ideę uczenia nienadzorowanego.	17
2.3. Poglądowy schemat przedstawiający ideę uczenia częściowo nadzorowanego.	19
2.4. Poglądowy schemat przedstawiający ideę uczenia częściowo nadzorowanego.	20
3.1. Poglądowy wykres przedstawiający przykładową predykcję modelu regresji liniowej.	24
3.2. Poglądowy wykres przedstawiający przykładowy wzorzec wykryty przez model regresji wielomianowej.	26
3.3. Poglądowy wykres przedstawiający wykryty wzorzec przez model regresji logistycznej.	27
3.4. Poglądowy wykres przedstawiający przykładową predykcję modelu algorytmu k najbliższych sąsiadów.	29
3.5. Poglądowy wykres przedstawiający przykładową predykcję modelu maszyny wektorów nośnych.	30
3.6. Poglądowy schemat przedstawiający przykładową strukturę modelu drzewa decyzyjnego.	31
3.7. Schemat przedstawiający uproszczoną budowę biologicznego neuronu. Źródło: https://pl.wikipedia.org/wiki/Neuron	34
3.8. Schemat modelu sztucznego neuronu. Źródło: https://en.wikibooks.org/wiki/Artificial_Neural_Networks	35
3.9. Schemat przedstawiający przykładową strukturę modelu wielowarstwowego perceptronu.	37
3.10. Schemat przedstawiający przykładową strukturę konwolucyjnej sieci neuronowej.	37

3.11. Schemat przedstawiający przykładową strukturę sieci Hopfielda.	38
3.12. Schemat przedstawiający przykładową strukturę rekurencyjnej sieci neuronowej z pamięcią skojarzeniową.	39
4.1. Porównanie platform TensorFlow i PyTorch w ujęciu czasowym pod względem popularności w wyszukiwarce firmy Google. Źródło: https://trends.google.pl/trends	43
4.2. Przykładowa wizualizacja narzędzia TensorBoard.	44
6.1. Schemat architektury modelu numer 1.	57
6.2. Schemat architektury modelu numer 2.	57
6.3. Schemat architektury modelu numer 3.	58
6.4. Schemat architektury modelu numer 4.	58
6.5. Schemat architektury modelu numer 5.	59
6.6. Schemat architektury modelu numer 6.	59
6.7. Schemat architektury modelu numer 7.	60
6.8. Schemat architektury modelu numer 8.	60
6.9. Schemat architektury modelu numer 9.	61
6.10. Schemat architektury modelu numer 10.	61
6.11. Wykres przedstawiający zależność uzyskanych wyników przez model numer 1 od ilości iteracji.	63
6.12. Wykres przedstawiający zależność uzyskanych wyników przez model numer 2 od ilości iteracji.	64
6.13. Wykres przedstawiający zależność uzyskanych wyników przez model numer 3 od ilości iteracji.	64
6.14. Wykres przedstawiający zależność uzyskanych wyników przez model numer 4 od ilości iteracji.	65
6.15. Wykres przedstawiający zależność uzyskanych wyników przez model numer 5 od ilości iteracji.	65
6.16. Wykres przedstawiający zależność uzyskanych wyników przez model numer 6 od ilości iteracji.	66
6.17. Wykres przedstawiający zależność uzyskanych wyników przez model numer 7 od ilości iteracji.	66

6.18. Wykres przedstawiający zależność uzyskanych wyników przez model numer 8 od ilości iteracji.	67
6.19. Wykres przedstawiający zależność uzyskanych wyników przez model numer 9 od ilości iteracji.	67
6.20. Wykres przedstawiający zależność uzyskanych wyników przez model numer 10 od ilości iteracji.	68
6.21. Wykres przedstawiająca wyniki przeprowadzonych eksperymentów. . .	69

Spis tabel

6.20. Tabela przedstawiająca wyniki przeprowadzonych eksperymentów. . .	69
---	----