

Politechnika Wrocławska
Wydział Informatyki i Telekomunikacji

Kierunek: **Informatyczne Systemy Automatyki (ISA)**
Specjalność: **Zastosowania Technologii Informacyjnych (IZT)**

PRACA DYPLOMOWA
MAGISTERSKA

**Głębokie sieci neuronowe w zadaniach predykcji cen
akcji oraz sygnałów transakcyjnych**

Predicting stock prices and providing recommendations using
deep neural networks

AUTOR

Piotr Kozub

OPIEKUN PRACY

Dr inż. Piotr Ciskowski

WROCLAW 2023

STRESZCZENIE

Celem pracy jest zbadanie przydatności głębokich sieci neuronowych w systemach transakcyjnych wykorzystywanych podczas gry na giełdzie. W pierwszej kolejności przedstawiono zagadnienia związane z rynkami finansowymi, a następnie opisano najważniejsze elementy sztucznych sieci neuronowych. Wstęp teoretyczny zakończył się przeglądem artykułów, gdzie przeprowadzono eksperymenty zbliżone do tematyki projektu.

Główną ideą wykonanych badań było porównanie podejścia polegającego na klasyfikacji trendów giełdowych z predykcją ceny w następnych dniach. W tym celu wykorzystano historyczne dane pary walutowej EUR/USD, wzbogacone o wskaźniki techniczne i odpowiednie etykiety. Następnie przystąpiono do implementacji oraz uczenia rekurencyjnych sieci neuronowych, a najlepsze model posłużyły do stworzenia kilku strategii inwestycyjnych. Zaproponowane rozwiązania zostały przetestowane za pomocą handlu symulowanego, w którym wyznaczono roczną stopę zwrotu, informującą o wartości wypracowanego przychodu. Końcowa analiza uzyskanych wyników pozwoliła na wybranie najkorzystniejszego systemu pod względem opłacalności oraz ryzyka inwestycji.

Słowa kluczowe:

giełda papierów wartościowych, sieci neuronowe, LSTM, sygnały transakcyjne, predykcja cen akcji

ABSTRACT

The purpose of this paper is to investigate the usefulness of deep neural networks in trading systems used when playing the stock market. Firstly introduced the issues related to financial markets, followed by a description of the most important elements of artificial neural networks. The theoretical part ended with a review of articles, where experiments similar to the topic of the project were realised.

The main idea of the performed experiments was to compare the stock market trend classification approach with next-day price prediction. For this purpose, historical data of the EUR/USD currency pair was used, enhanced with technical indicators and relevant labels. The implementation and learning of recurrent neural networks was then undertaken, and the best model was used to create several investment strategies. The suggested solutions were tested by means of simulated trading, in which the yearly rate of return was determined, giving information on the value of the generated revenue. A final analysis of the results allowed the selection of the most beneficial system in terms of profitability and investment risk.

Key words:

stock market, neural networks, LSTM, trading signals, stock price prediction

SPIS TREŚCI

1. Wprowadzenie.....	1
1.1. Opis problemu	1
1.2. Cel pracy.....	1
1.3. Zarys pracy	1
2. Wstęp.....	2
2.1. Giełda.....	2
2.2. Instrumenty finansowe.....	2
2.2.1. Akcje	2
2.2.2. Obligacje	4
2.2.3. Instrumenty pochodne	4
2.2.4. ETF.....	4
2.3. Forex	5
2.4. Systemy wspomagania decyzji inwestycyjnych	6
2.4.1. Analiza fundamentalna.....	7
2.4.2. Analiza techniczna	7
3. Sztuczne sieci neuronowe	10
3.1. Budowa neuronu.....	10
3.2. Rozwój sieci neuronowych.....	13
3.2.1. Perceptron Rosenblatta.....	13
3.2.2. Sieci wielowarstwowe i algorytm propagacji wstecznej.....	13
3.2.3. Głębokie sieci neuronowe	15
3.3. Rekurencyjne sieci neuronowe	16
3.3.1. Struktura i budowa	16
3.3.2. LSTM	17
3.3.3. GRU	19
4. Przegląd literatury	20
5. Przygotowanie zbioru danych.....	22
5.1. Pobranie danych	22
5.2. Dobór wskaźników	23
5.3. Przygotowanie danych treningowych i testowych	25

6. Uczenie sieci neuronowych	27
6.1. Przygotowanie modeli	27
6.1.1. Klasyfikacja	27
6.1.2. Predykcja.....	29
6.2. Analiza wyników uczenia.....	30
6.3. Wybór najlepszego modelu	34
7. Implementacja systemów transakcyjnych	36
7.1. Strategie oparte na klasyfikacji	36
7.1.1. Pojedynczy sygnał	36
7.1.2. Trzy sygnały.....	37
7.2. Strategie oparte na predykcji.....	37
7.2.1. Wzrost względny	37
7.2.2. Wzrost bezwzględny	38
8. Symulacja systemów na danych historycznych	39
8.1. Pojedynczy sygnał	39
8.2. Trzy sygnały	43
8.3. Wzrost względny	45
8.4. Wzrost bezwzględny	49
9. Podsumowanie	52
9.1. Wybór najlepszego systemu	52
9.2. Wnioski.....	53
Bibliografia	55

1. WPROWADZENIE

1.1. OPIS PROBLEMU

Praca magisterska porusza temat zastosowania algorytmów sztucznej inteligencji w systemach transakcyjnych stosowanych na giełdzie papierów wartościowych. Znalezienie skutecznego algorytmu, przynoszącego realny zysk w trakcie inwestycji środków, pozwoliłoby na uzyskanie pasywnego przychodu dla potencjalnego inwestora. Wobec tego, poszukiwane są strategie, bazujące na modelach głębokich sieci neuronowych, które udowodnią swoją skuteczność w trakcie badań przeprowadzonych na danych historycznych.

1.2. CEL PRACY

Celem pracy jest sprawdzenie przydatności głębokich sieci neuronowych w systemach transakcyjnych wykorzystywanych podczas gry na giełdzie, poprzez porównanie podejścia związanego z predykcją oraz klasyfikacją szeregów czasowych. Wykonanie badań wiąże się z zebraniem oraz przygotowaniem odpowiednich danych uczących, które pozwolą na wytrenowanie wybranych modeli głębokich sieci neuronowych. Uzyskane źródło wiedzy ma posłużyć do zaimplementowania strategii inwestycyjnych, które umożliwią obliczenie potencjalnego zysku przetestowanego podczas symulacji na zbiorach testowych.

1.3. ZARYS PRACY

Merytoryczna część pracy rozpoczyna się w drugim rozdziale, gdzie przedstawione zostały podstawowe zagadnienia związane z giełdą, akcjami oraz światem inwestycji. Następnie opisano historie prac nad rozwojem sztucznych sieci neuronowych, zaczynając od porównania biologicznego i sztucznego neuronu, kończąc na uczeniu głębokim. Część teoretyczna kończy się przeglądem artykułów naukowych, w których przeprowadzono eksperymenty poruszające zbliżoną tematykę.

Część projektowa rozpoczyna się w rozdziale piątym, gdzie opisano pozyskanie oraz przygotowanie danych, które następnie służą do wytrenowania rekurencyjnych sieci neuronowych. Rozdział szósty został poświęcony badaniom na temat zaprojektowania najkorzystniejszej architektury sieci. W tym punkcie wybrane zostają najskuteczniejsze z przygotowanych modeli do predykcji oraz klasyfikacji szeregów czasowych.

W rozdziale siódmym opracowano różne strategie inwestycyjne, oparte na wiedzy stworzonych sieci neuronowych. Następnie szczegółowo opisano przebieg badań, sprawdzający ich skuteczność podczas symulacji na danych historycznych w wybranym środowisku testowym. Ostatecznie wykonano analizę otrzymanych wyników, pozwalającą na wyłonienie najlepszej strategii inwestycyjnej z wszystkich przygotowanych rozwiązań. Na podstawie przeprowadzonych badań wyciągnięto idące za nimi wnioski.

2. WSTĘP

2.1. GIEŁDA

Giełda Papierów Wartościowych (GPW) to miejsce transakcji, w którym kupujący i sprzedający mogą handlować papierami wartościowymi oraz instrumentami pochodnymi. Funkcjonuje jako otwarty, transparentny system umożliwiający inwestorom przeprowadzanie transakcji zgodnie z zasadą popytu i podaży. Kluczową rolę w trakcie transakcji odgrywa makler, który otrzymuje i weryfikuje zlecenia przed przekazaniem ich na giełdę. W skład procesu transakcyjnego na GPW wchodzi zbieranie zleceń od domów maklerskich i łączenie stron zainteresowanych realizacją transakcji. Oznacza to, że do uzyskanie pomyślnego zakończenia, oferta kupna musi znaleźć odpowiednie zlecenie sprzedaży.

Czas otwarcia giełdy jest podzielony na tak zwane sesje, które są wyznaczane w dni robocze od poniedziałku do piątku, w godzinach zależnych od konkretnego rynku. Inwestorzy mają możliwość monitorowania zmian cen akcji na bieżąco, dzięki dostępowi do arkusza zleceń.

Każda osoba dysponująca odpowiednim kapitałem oraz posiadająca otwarty rachunek inwestycyjny u wybranego maklera może inwestować na giełdzie. Transakcje kupna i sprzedaży są realizowane w ramach dwóch różnych systemów notowań, ciągłych lub jednolitych, które odpowiadają za ustalanie kursu akcji. System notowań ciągłych jest najczęściej stosowany dla dużych firm o wysokiej płynności. W tej metodzie kurs akcji może ulegać zmianie przy każdej transakcji. GPW oferuje również notowania jednolite, przeznaczone dla firm o niższej płynności. Ten system jest mniej popularny, a ceny akcji zmieniają się znacznie rzadziej - zazwyczaj dwa razy dziennie [1, 2].

2.2. INSTRUMENTY FINANSOWE

Rynki finansowe oferują szeroki wachlarz instrumentów, które umożliwiają przyjęcie różnorodnych strategii inwestycyjnych. Inwestorzy mają do wyboru nie tylko akcje, ale również obligacje, instrumenty pochodne, fundusze indeksowe i wiele innych. Wybór odpowiedniego instrumentu finansowego zależy od indywidualnych preferencji inwestora, jego apetytu na ryzyko oraz celów inwestycyjnych [1, 2].

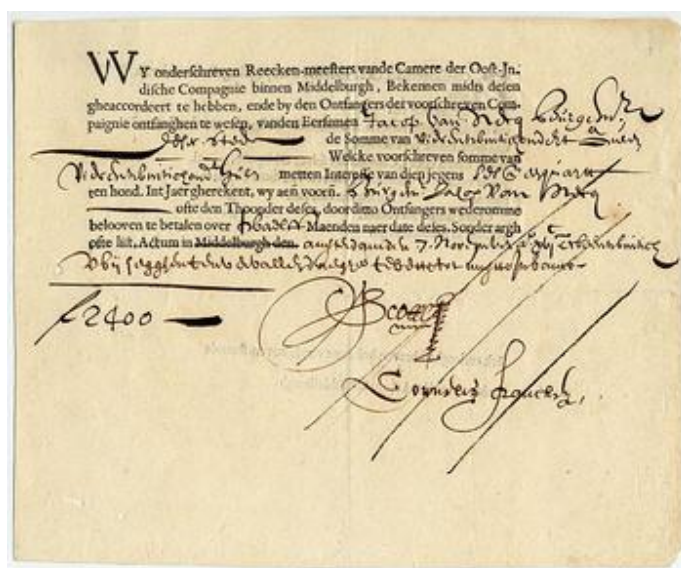
2.2.1. AKCJE

Akcje to podstawowe instrumenty finansowe oferowane na giełdach papierów wartościowych. Łączą prawa majątkowe oraz niemajątkowe, które wynikają z uczestnictwa ich właścicieli, zwanych akcjonariuszami, w spółce akcyjnej. Akcję należą do firmy przez którą zostały wyemitowane. Oznacza to, że ich kupno jest równoznaczne z nabyciem udziału w kapitale spółki [3].

Historia akcji jako instrumentów finansowych sięga początków XVII wieku, kiedy to Holenderska Kompania Wschodnioindyjska (VOC) jako pierwsza spółka w historii zdecydowała się na publiczną emisję akcji. VOC została utworzona przez rząd Niderlandów, aby prowadzić handel morski z Azją, przede wszystkim z Indiami i Indonezją, a później również z Japonią i Chinami. Pierwsze wyprawy były finansowane przez grupy handlowców, które tworzyły tzw. kompanie konsorcjalne. Każda z nich była oddzielnym przedsięwzięciem, na które zbierano fundusze, jednak nie było to rozwiązanie efektywne. Podróże wymagały środków na budowę statków, zatrudnienie załóg oraz zebrania dużej

ilości zapasów, co wiązało się z wysokimi kosztami. Zdobywanie takiego kapitału nie było wówczas łatwym zadaniem. Niewielka ilość potencjalnych inwestorów obawiała się niepowodzenia misji, przez co wyprawy opóźniały się lub nie dochodziły do skutku.

Holenderska Kompania Wschodnioindyjska znalazła rozwiązanie tego problemu poprzez emisję akcji. To nowatorskie podejście pozwoliło na dywersyfikację ryzyka - inwestorzy ponosili tylko część niebezpieczeństwa związanego z każdą wyprawą, a potencjalne straty z jednej ekspedycji mogły być zrekompensowane zyskami z innych. Od tego momentu VOC nie musiało szukać majątnych inwestorów, gdyż pieniądze dużej grupy średniozamożnych osób pozwalały na zebranie odpowiedniej ilości środków. Zgromadzony kapitał pozwolił na zbudowanie floty statków i prowadzenie stałej działalności handlowej w Azji Wschodniej, co przekształciło handel morski w bardziej stabilny i przewidywalny biznes. W ten sposób, VOC stała się pierwszą na świecie spółką akcyjną i wyznaczyła kierunek dla przyszłych innowacji finansowych [4].



Rys. 2.1. Akcja wydana przez VOC w 1623 roku [5]

Koncepcja emisji akcji rozpoczęła nową erę finansów, umożliwiając firmom pozyskiwanie kapitału od szerokiego grona inwestorów, jednocześnie znacznie przyspieszając rozwój gospodarczy. Akcjonariusze mogą liczyć na następujące benefity:

- **Dywidendy** – to część zysków firmy, która jest wypłacana właścicielom akcji. Wielkość dywidendy jest zależna od zysków firmy i decyzji zarządu o jej dystrybucji.
- **Zyski kapitałowe** – jeżeli wartość firmy rośnie, rośnie również wartość akcji. Akcjonariusze mogą sprzedać swoje akcje po wyższej cenie, niż ta za którą je kupili, co generuje zysk kapitałowy.
- **Prawo głosu** – inwestorzy mają prawo do głosowania na walnym zgromadzeniu akcjonariuszy. Mogą w ten sposób wpływać na kierunek rozwoju firmy, głosując na przykład w sprawach nominacji do zarządu czy dywidendy.
- **Prawo do informacji** – akcjonariusze mają prawo do regularnych informacji o działalności firmy, które pomagają im śledzić kondycję finansową i operacyjną spółki. Dostęp do tych danych, pozwala akcjonariuszom na podjęcie świadomych decyzji o kupnie, sprzedaży lub utrzymaniu akcji. [6]

2.2.2. OBLIGACJE

Obligacje są instrumentami dłużnymi, które reprezentują pożyczkę udzieloną przez inwestora, będącego kupcem, emitentowi obligacji. Kupując obligację, inwestor staje się wierzycielem emitenta i otrzymuje prawo do regularnych odsetek oraz zwrotu nominalnej wartości obligacji po określonym czasie, zwanym terminem zapadalności. W zależności od rodzaju podmiotu emitującego wyróżniamy:

- **Obligacje skarbowe** – są emitowane przez rząd, dla którego jest to sposób na finansowanie deficytu budżetowego. Charakteryzują się one niskim ryzykiem oraz terminem zapadalności większym niż rok.
- **Obligacje korporacyjne** – emitują je duże przedsiębiorstwa w celu zebrania funduszy na nowe inwestycje.
- **Obligacje komunalne** – ich emitentem są poszczególne jednostki terytorialne, takie jak powiaty lub gminy.

Zakup obligacji jest możliwy bezpośrednio od emitenta, na rynku wtórnym poprzez Giełdę Papierów Wartościowych lub w ramach funduszu inwestycyjnego. Cena ich nabycia może być zarówno niższa jak i wyższa od ceny za którą zostaną odkupione, jednak zazwyczaj są to odchylenia znacznie mniejsze niż w przypadku akcji. Główne ryzyko związane z inwestowaniem w obligacje zależy od wiarygodności kredytowej emitenta, przez co na ogół są postrzegane jako bezpieczne lokum funduszy. Co istotne, w przeciwieństwie do akcji, inwestorzy nie zyskują prawa do dywidendy lub współwłasności dającej prawo głosu na walnych zgromadzeniach [7].

2.2.3. INSTRUMENTY POCHODNE

Instrumenty pochodne, takie jak kontrakty terminowe, opcje czy kontrakty na różnicę kursową (CFD), są bardziej skomplikowanymi instrumentami finansowymi, które bazują na wartości innego aktywa (aktywa bazowego), takiego jak akcje, obligacje, indeksy giełdowe, surowce czy waluty. Wartość instrumentu pochodnego jest zatem "pochodną" wartości innego aktywa. Instrumenty pochodne są często wykorzystywane do zabezpieczenia przed ryzykiem związanym ze zmianą wartości aktywów bazowych, ale mogą być również wykorzystywane do spekulacji, co wiąże się z wyższym ryzykiem [9, 10].

2.2.4. ETF

ETF-y, czyli fundusze indeksowe, są instrumentami pozwalającymi inwestorom na zdywersyfikowanie swojego portfela poprzez inwestycję w koszyk różnych akcji, który odzwierciedla dany indeks giełdowy, np. WIG20 czy S&P500. ETF-y są notowane na giełdzie oraz mogą być kupowane i sprzedawane podobnie jak akcje. To nic innego jak wskaźniki tworzone na podstawie kursów wybranych aktywów notowanych na danym rynku. Oznacza to, że ich wartość stara się naśladować ogólną kondycję spółek, które wchodzą w skład ETF-a.

Zarządzający funduszami starają się jak najdokładniej odwzorowywać obecną sytuację rynku poprzez proces zwany replikacją. Polega ona na zachowaniu odpowiednich proporcji wszystkich indeksów giełdy. W przypadku bardziej rozbudowanych wskaźników brana jest pod uwagę jedynie ich część, ponieważ jednoczesna kontrola tysięcy spółek byłaby ciężka do opanowania i w konsekwencji prowadziła do nieuzasadnionych wzrostów lub spadków notowania funduszu [8].

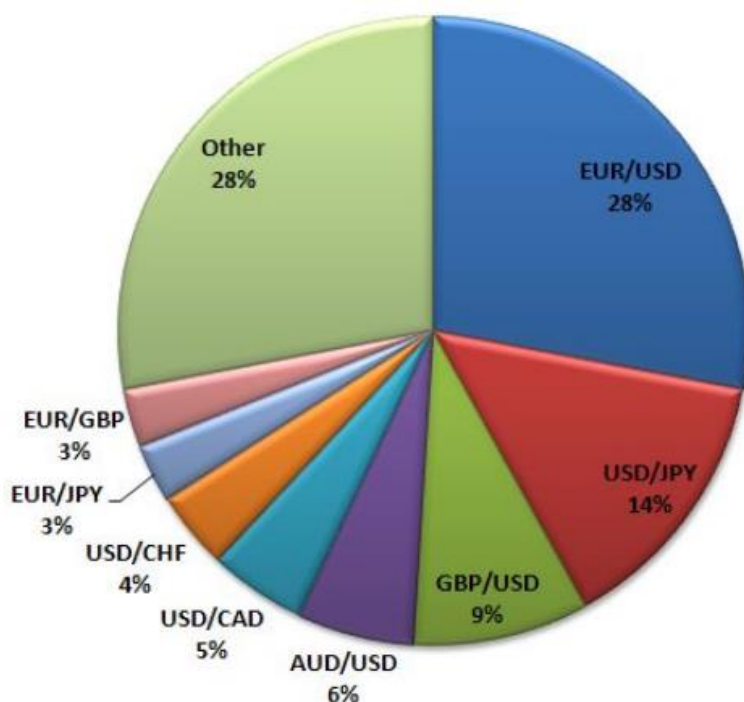
2.3. FOREX

Forex (ang. Foreign Exchange Market) stanowi największy na świecie, międzynarodowy rynek walutowy, w ramach którego rządy, banki centralne, wielkie korporacje oraz inwestorzy z całego świata dokonują operacji wymiany walut. Forex to zdecentralizowany rynek pozagiełdowy, co oznacza, że transakcje odbywają się bezpośrednio pomiędzy kontrahentami, w odróżnieniu od scentralizowanej giełdy, na której oferty kupna i sprzedaży muszą zostać połączone przez pośrednika. Ponadto Forex charakteryzuje się największym i najbardziej płynnym rynkiem na świecie, z dziennym obrotem przekraczającym 6 bilionów dolarów [9].

Handel na rynku Forex otwiera się w niedzielę o godzinie 23:00 czasu polskiego, a zamyka w piątek o 22:00. Oznacza to, że jest czynny we wszystkie dni robocze przez niemal 24 godziny na dobę. Dzieje się tak, ponieważ nie jest on zlokalizowany w jednym konkretnym miejscu – Forex to sieć komputerowa łącząca uczestników z całego świata [9].

Właściwym przedmiotem handlu na rynku Forex są pary walutowe. Konstrukcja każdej z nich polega na ustaleniu stosunku wartości wybranej waluty do drugiej. Pierwsza waluta w parze nazywana jest walutą bazową, zaś druga kwotowaną. W ten sposób cena pary walutowanej określa, ile jednostek waluty kwotowanej jesteśmy w stanie kupić za jedną jednostkę waluty bazowej. Jeśli cena pary USD/PLN wynosi 4,2, oznacza to, że za jednego dolara możemy kupić za 4 złote i 20 groszy [9, 10]. Pary dzielą się również na kilka grup:

- **Majors (główne)** – składają się z najpopularniejszych walut światowych, do których zaliczają się m.in. EUR/USD, USD/CHF, GBP/USD, USD/JPY i inne.
- **Minors (drugorzędne)** – należą do nich waluty o mniejszym globalnym znaczeniu, np. USD/CAD, AUD/USD itp.
- **Exotic (egzotyczne)** – to pary walutowe, gdzie jedna z walut jest mało popularna na świecie i należy do grona wciąż rozwijających się krajów, np. USD/PLN. [10]



Rys. 2.2. Ranking najpopularniejszych par walutowych na rynku Forex [11]

Handel na rynku Forex został upodobany przez inwestorów z całego świata, chcących powiększyć swój portfel na wahaniami cen. Kursy walut są zależne od wielu czynników, takich jak stopy procentowe, inflacja, polityka rządów, dane ekonomiczne i ogólne nastroje rynkowe. Jest to rynek zmienny, który potrafi odwracać swoje trendy kilka razy dziennie. Ponadto Forex posiada największą płynność ze wszystkich rynków na świecie, co oznacza, że zawsze znajdzie się wystarczająco dużo kupujących i sprzedających, aby przeprowadzić nawet bardzo wysokie transakcje. To ważny aspekt dla graczy, którzy w szczególności skupiają się na krótkoterminowych inwestycjach, takich jak day-trading [9, 10].

Zmiany cen tradycyjnych walut posiadają wysoką częstotliwość, jednak ich amplituda jest stosunkowo niska w porównaniu ze skokami innych instrumentów finansowych, takich jak akcje lub kryptowaluty. Z tego względu popularną praktyką na rynku Forex jest korzystanie z dźwigni finansowej. To narzędzie pozwala na zawieranie kontraktów, których wartość jest znacznie wyższa niż możliwości finansowe inwestora. W ten sposób gracze kupują dużą ilość walut, posiadając jedynie kapitał zabezpieczający małą część ich wartości. Reszta to pożyczka zaciągnięta u brokera, która wynosi np. dziesięciokrotność wkładu własnego. Taka forma inwestycji oznacza, że ewentualny wzrost ceny wybranej waluty przyniesie zysk proporcjonalny do całkowitej wartości kontraktu. Należy jednak pamiętać, że nietrafiona decyzja spowoduje adekwatną utratę pieniędzy [9, 10].

W związku z tym, inwestowanie na rynku Forex wymaga dogłębnej wiedzy na temat rynków finansowych, analizy technicznej i fundamentalnej, a także świadomości ryzyka związanego z inwestowaniem.

2.4. SYSTEMY WSPOMAGANIA DECYZJI INWESTYCYJNYCH

Decyzje inwestycyjne zwykle pociągają za sobą wysokie ryzyko. Ich sukcesy oraz porażki mają istotne konsekwencje z punktu widzenia osób grających na giełdzie, które dążą do finansowego dobrobytu. Z tego powodu inwestorzy często szukają metod i narzędzi, które mogą im pomóc w dokonaniu właściwego wyboru. W tym celu chętnie korzystają z systemów wspomagających decyzje kupna.

Systemy wspomagania decyzji to narzędzia wykorzystywane w celu optymalizacji procesu podejmowania decyzji. W kontekście inwestycji, systemy te pomagają inwestorom w analizie danych rynkowych, prognozowaniu trendów i ocenie ryzyka. Mają na celu usprawnienie procesu kupna i sprzedaży poprzez ograniczenie błędów oraz zwiększenie prawdopodobieństwa sukcesu inwestycji przy jednoczesnym zachowaniu ostatecznej decyzji w rękach osoby wykorzystującej dane narzędzie.

Inwestorzy wykorzystują systemy wspomagania decyzji, ponieważ pomagają one zredukować skomplikowany i często niejednoznaczny świat inwestowania do uporządkowanych oraz zrozumiałych zestawów danych i analiz. W ten sposób umożliwiają skupienie się wyłącznie na kluczowych informacjach, zamiast analizowania ogromu dostępnych danych niosących wiadomości niepowiązane z aktualnym obiektem badań. W praktyce, inwestorzy mogą korzystać z różnych narzędzi i technik, które poszerzą ich wiedzę na temat danego rynku. Niektóre z nich mogą polegać na prostych wykresach i wskaźnikach, inne mogą być bardziej zaawansowane, wykorzystując algorytmy do przewidywania trendów cenowych. Bez względu na konkretną formę pomagają inwestorom skoncentrować się na tym, co najważniejsze dla ich decyzji inwestycyjnej.

W kontekście giełdy, główne metody wykorzystywane w systemach wspomagania decyzji to analiza fundamentalna i techniczna. Oba te podejścia mają na celu przewidywanie przyszłych ruchów cen, przy czym biorą pod uwagę odmienne obserwacje i wskaźniki.

2.4.1. ANALIZA FUNDAMENTALNA

Analiza fundamentalna to metoda oceny wartości aktywów oparta na głębokim zrozumieniu czynników, które wpływają na ich rzeczywistą wartość. Zdaniem zwolenników tego podejścia, rynek może nie odzwierciedlać prawdziwej wartości aktywów w krótkim okresie, jednakże w dłuższym czasie ceny mają tendencję do osiągnięcia ich rzeczywistej wartości. Celem analizy fundamentalnej jest identyfikacja tak zwanych "wartości inwestycyjnych" – akcji firm, które są dobrze zarządzane, mają solidne podstawy finansowe i działają w korzystnych warunkach gospodarczych, ale których cena na rynku jest niższa, niż sugerowałaby ich rzeczywista wartość [12, 13].

Analiza fundamentalna skupia się na badaniu wszelkiego rodzaju danych, które mogą mieć wpływ na wartość firmy. Dane te, obejmują informacje finansowe, takie jak dochody, koszty, zyski, przepływy pieniężne, dług, aktywa, a także szereg innych miar wartościowania, takich jak wskaźnik ceny do zysku (P/E), wskaźnik ceny do wartości księgowej (P/B) lub wskaźnik dywidendy. Nie ogranicza się jednak tylko do danych finansowych. Analiza fundamentalna bierze także pod uwagę szeroki zakres czynników zewnętrznych, takich jak warunki gospodarcze, stopy procentowe, inflacja, polityka rządowa, trendy branżowe i wiele innych [13].

To podejście często wymaga dogłębnego zrozumienia biznesu, branży, w której działa firma, a także kontekstu makroekonomicznego. Analiza fundamentalna zakłada, że posiadanie takiego zrozumienia pozwoli inwestorowi zidentyfikować akcje, które są niedocenione lub przewartościowane przez rynek, co w efekcie przełoży się na trafną decyzję.

Przykładem zastosowania analizy fundamentalnej może być inwestor, który analizuje firmę produkującą samochody. Inwestor ten mógłby zbadać sprawozdania finansowe firmy, aby zrozumieć jej kondycję finansową, a także przeprowadzić analizę branżową, aby zrozumieć konkurencję, popyt na samochody, trendy technologiczne i inne czynniki, które mogą wpłynąć na przyszłość firmy.

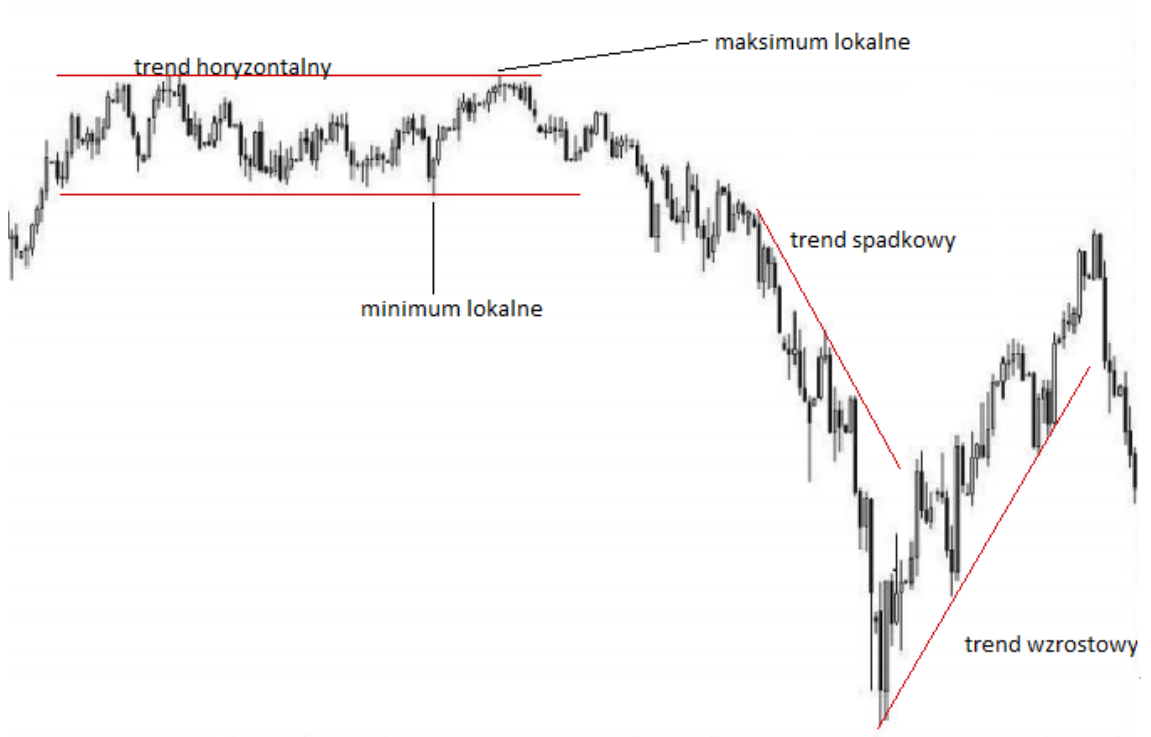
Należy pamiętać, że choć analiza fundamentalna dostarcza inwestorom użytecznych narzędzi do oceny aktywów, to nie jest wolna od ryzyka. Wyniki analizy są często zależne od założeń i prognoz, które mogą okazać się błędne, a nawet jeśli są prawidłowe, rynek może nie docenić danego aktywa. Ponadto zmiany cen na giełdzie są uwarunkowane wieloma ogólnymi czynnikami związanymi z aktualną sytuacją na świecie. Nawet najlepsze analizy nie są w stanie przewidzieć nagłych zdarzeń, które potrafią całkowicie zmienić stan gospodarki, tym samym doprowadzając do wysokich strat. Z tego powodu, choć analiza fundamentalna jest cennym narzędziem, nie powinna być jedynym kryterium decyzji inwestycyjnej [12, 13].

2.4.2. ANALIZA TECHNICZNA

Analiza techniczna to metoda prognozowania ruchów cen na podstawie badania przeszłych danych rynkowych. Różni się od analizy fundamentalnej, ponieważ zamiast skupiać się na realnych wartościach firmy, takich jak zysk lub przychody, koncentruje się na badaniu wzorców historycznych cen i wolumenów handlu. Głównym założeniem analizy technicznej jest koncepcja, że ceny aktywów finansowych odzwierciedlają wszystkie dostępne informacje i składają się z szeregu powtarzalnych wzorców, które są uznawane za wskaźniki przyszłych ruchów cen. Innymi słowy, dotychczasowe zmiany w cenach mogą dawać sygnały do przyszłych zmian [14, 15].

Jednym z najważniejszych aspektów analizy technicznej jest identyfikacja trendów, czyli kierunków, w których porusza się cena. W trakcie analizy wykresów można zauważyć, że w pewnych okresach przebiegi poszczególnych notowań potrafią podążać w jednym kierunku. W ten sposób wyróżnia się trzy główne rodzaje trendów:

- **Trend wzrostowy** – pojawia się w momencie, gdy linia łącząca co najmniej trzy minima lokalne jest prostą skierowaną w górę. Trend wzrostowy powstaje na skutek przewagi osób kupujących nad sprzedającymi. Utworzona prosta jest nazywana „linią wsparcia” [15].
- **Trend spadkowy** – to sytuacja odwrotna, która pojawia się w momencie, gdy linia łącząca co najmniej trzy maksima lokalne jest prostą malejącą. Powstaje na skutek przewagi podaży nad popytem. Linia ta nazywana jest „linią oporu” [15].
- **Trend horyzontalny** – występuje w sytuacji, gdy na wykresie nie można dostrzec żadnej wyraźnej sekwencji, a cena cyklicznie oscyluje w górę oraz dół w ograniczonym zakresie [15].



Rys. 2.3. Rodzaje trendów [16]

W procesie analizy technicznej używa się wiele różnych narzędzi i technik, z których każda ma swoje zastosowanie w zależności od sytuacji rynkowej. Najpopularniejsze z nich obejmują:

- **Wskaźniki pędu (momentum)** – wskaźniki pędu, takie jak oscylator stochastyczny lub wskaźnik siły względnej (RSI), pomagają mierzyć prędkość, z jaką zmieniają się ceny. Mogą one pomóc inwestorom zidentyfikować nadchodzące punkty zwrotne, sugerując kiedy aktywa są "przekupione" lub "przesprzedane" [14].
- **Wskaźniki wolumenu** – wskaźniki wolumenu, takie jak wskaźnik przepływów pieniężnych Chaikina, są używane do mierzenia siły trendu na podstawie poziomu aktywności handlowej. Wskaźniki te pomagają zrozumieć, jak duży jest udział inwestorów w trendzie cenowym [14].

- **Wskaźniki średnich kroczących** – średnie kroczące, takie jak prosta średnia krocząca (SMA) i wykładnicza średnia krocząca (EMA), są używane do identyfikacji ogólnych trendów rynkowych. Średnie te pomagają wygładzić krótkoterminowe oscylacje cen, umożliwiając inwestorom skupienie się na długoterminowym trendzie [14].
- **Wskaźniki formacji cenowych** – wskaźniki te identyfikują specyficzne wzorce na wykresach cenowych, które mogą sugerować przyszłe ruchy cen. Przykłady takich formacji to flagi, trójkąty, podwójne szczyty i podwójne dołki [14].

Ponadto, analiza techniczna nie jest ograniczona tylko do pojedynczych aktywów. Inwestorzy i analitycy często stosują analizę techniczną do oceny trendów na większych rynkach, takich jak indeksy giełdowe, rynki towarowe lub walutowe, tym samym biorąc pod uwagę wiele indeksów na raz.

Zaawansowani inwestorzy często łączą analizę techniczną z innymi formami analizy, takimi jak analiza fundamentalna lub analiza sentymentu, aby uzyskać pełniejszy obraz rynku. Coraz popularniejszą techniką wykorzystywaną w ramach analizy technicznej jest uczenie maszynowe. Jego szczególnym przypadkiem są sieci neuronowe, które nauczone na podstawie danych historycznych starają się odwzorować przyszłe trendy.

Ostatecznie, pomimo skomplikowanej natury analizy technicznej, jej głównym celem jest pomoc inwestorom w zrozumieniu rynków i przewidywaniu przyszłych trendów cen. Choć nie jest to nigdy gwarancją sukcesu, dobrze przeprowadzona może pomóc w zrozumieniu świata inwestycji, identyfikować atrakcyjne możliwości inwestycyjne oraz ograniczać ryzyko z nimi związane.

3. SZTUCZNE SIECI NEURONOWE

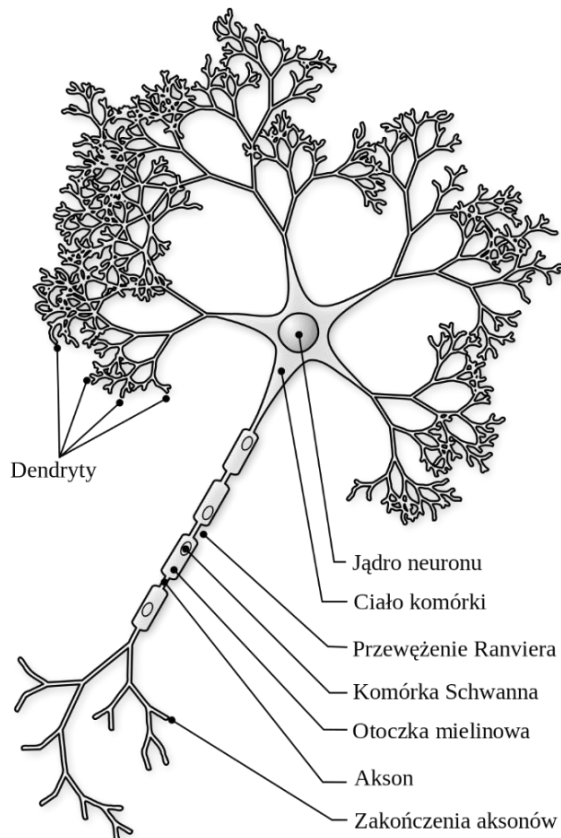
Zgodnie z raportem IDC (International Global Corporation) ilość danych na świecie wzrasta o 23% rocznie, a w 2025 roku osiągnie rozmiar 180 zettabajtów (10^{21} bajtów). Przy tak dużym napływie informacji, zdolność do ich efektywnego przetwarzania staje się coraz bardziej istotna. W tym kontekście, sztuczne sieci neuronowe, część większego obszaru nazywanego uczeniem maszynowym, zdobyły sobie uznanie jako potężne narzędzie umożliwiające odkrywanie ukrytych wzorców i zależności w dużych zbiorach danych.

Sztuczne sieci neuronowe, są zainspirowane biologicznymi sieciami neuronowymi obecnymi w ludzkim mózgu. Każdy "neuron" w takiej sieci otrzymuje dane wejściowe, przetwarza je, a następnie przekazuje wynikowe dane na wyjście. Neurony są połączone w sieci, a siła tych połączeń, nazywana wagami, jest kluczem do uczenia się i adaptacji sieci.

Pomysł na sztuczne sieci neuronowe sięga lat 40. XX wieku, ale prawdziwy przełom nastąpił wraz z rozwojem technologii komputerowych i dostępności dużych zbiorów danych. Dzięki temu, sieci neuronowe stały się praktycznym narzędziem do rozwiązywania problemów w różnych dziedzinach, począwszy od analizy obrazu i rozpoznawania mowy, po przewidywanie trendów na giełdzie [17].

3.1. BUDOWA NEURONU

Inspiracją do budowy pierwszego sztucznego neuronu był pomysł na stworzenie matematycznego opisu biologicznego neuronu. Idea ta nie polegała na odwzorowaniu budowy naturalnego układu, lecz próbie naśladowania sposobu przetwarzania oraz przesyłania informacji.



Rys. 3.1. Schemat budowy biologicznego neuronu [18]

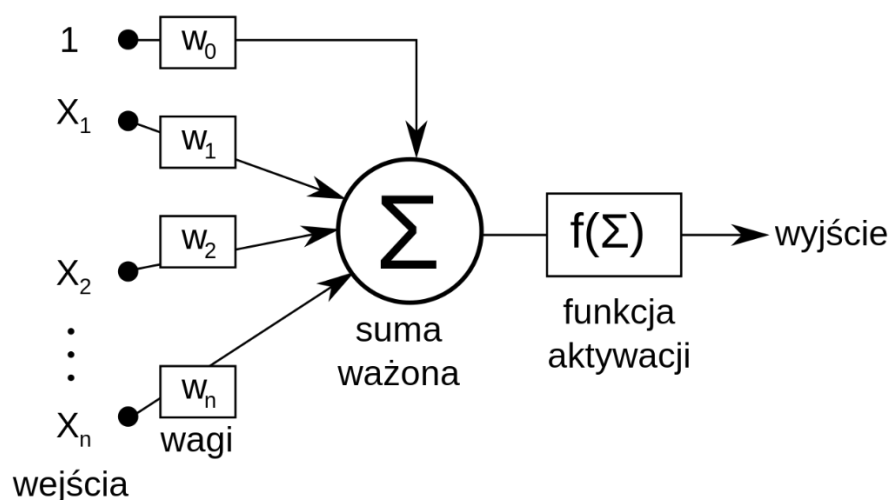
Na rysunku 3.1 przedstawiony został model budowy biologicznego neuronu, w którego skład wchodzi wiele elementów. Do jego najważniejszych członów odpowiedzialnych za proces przekazywania informacji należą:

- **Dendryty** – liczne wypustki odbierające sygnały zewnętrzne i przekazujące je do jądra neuronu,
- **Jądro** – miejsce obliczeń,
- **Akson** – wyjście neuronu zakończone wieloma rozgałęzieniami.

Proces transmisji sygnałów w układzie nerwowym rozpoczyna się od odebrania przez dendryty wielu bodźców zewnętrznych i przekazaniu ich do ciała komórki. Wówczas na podstawie otrzymanych informacji jądro decyduje o ewentualnym wysłaniu sygnału wyjściowego według zasady „wszystko albo nic”. Zasada ta mówi o tym, że neurony są w stanie emitować wyłącznie jeden rodzaj informacji. Jeśli otrzymane bodźce są wystarczająco silne, wysyłany zostaje sygnał elektryczny o amplitudzie około 100 mV i czasie trwania 1-2 ms. Silniejsza stymulacja dendrytów może jedynie zwiększyć ilość i częstotliwość generowanych impulsów [18, 19].

Wytworzony sygnał przechodzi przez zakończenia aksonów, które łączą się z dendrytami innych neuronów lub organami takimi jak mięśnie, tworząc **synapsy**. Większość spośród wszystkich synaps to synapsy chemiczne, które działają na zasadzie przekształcania dochodzącego do zakończenia aksonu sygnału elektrycznego (impuls nerwowy) do chemicznej reprezentacji w postaci wyprodukowanego neuroprzekaźnika. Tak skonstruowane połączenia stanowią miejsce komunikacji pomiędzy wszystkimi neuronami, w efekcie czego tworzą sieć neuronową odpowiedzialną za procesy neurologiczne organizmów żywych. Z kolei uczenie biologicznej sieci neuronowej polega w dużym uproszczeniu na zwiększaniu sił połączenia synaps poprzez dostosowanie ilości produkowanego neuroprzekaźnika przez poszczególne neurony, w zależności od rodzaju transmitowanej informacji [18, 19].

W 1943 roku, podczas próby zrozumienia i naśladowania działania biologicznego neuronu, Warren McCulloch i Walter Pitts zaproponowali model matematyczny, który nazwano sztucznym neuronem. Ich praca, zatytułowana "A Logical Calculus of the Ideas Immanent in Nervous Activity", zdefiniowała podstawy dla tego, co znamy dziś jako sieci neuronowe [19÷21].



Rys. 3.2. Model neuronu McCullocha-Pittsa [20]

Przedstawiony na rysunku 3.2 model neuronu McCullocha i Pittsa jest uproszczonym odzwierciedleniem jego biologicznego odpowiednika. Działa na zasadach prostych reguł logicznych i matematycznych, a jego struktura odtwarza podstawowe mechanizmy przetwarzania informacji w biologicznym neuronie.

W skład modelu McCullocha i Pittsa wchodzi następujące elementy:

- **Wejścia** – to pierwszy składnik sztucznego neuronu który jest odpowiedzialny za odbieranie sygnałów z danych wejściowych lub innych neuronów. Każde wejście odbiera sygnał w sposób analogiczny do odbierania impulsów nerwowych przez dendryty od innych neuronów układu nerwowego (poprzez synapsy).
- **Wagi** – każde wejście ma przypisaną wagę, która mnoży początkową wartość, tym samym modyfikując jej poziom istotności. Odpowiada to mechanizmowi działania sił synaptycznych, które regulują przekazywanie sygnałów między neuronami w biologicznym neuronie.
- **Suma ważona** – jest miejscem agregacji wszystkich ważonych sygnałów wejściowych, podobnie jak jądro neuronu sumuje i przetwarza informacje otrzymane od dendrytów.
- **Funkcja aktywacji** – ta funkcja decyduje, czy neuron powinien być aktywowany oraz wylicza ostateczną wartość wyjścia, na podstawie wyniku z funkcji sumowania. Jest to rozwiązanie analogiczne do właściwości biologicznego neuronu, w którym sygnał musi przekroczyć pewien próg, aby wyemitować impuls elektryczny (zasada „wszystko albo nic”).
- **Wyjście (odpowiednik aksonu)** – jest to miejsce, skąd sygnał jest przekazywany do innych neuronów lub do środowiska. Podobnie jak w biologicznym neuronie, gdzie sygnał jest przekazywany do innych neuronów lub do mięśni przez zakończenia aksonu. [19, 21]

Mimo że model neuronu McCullocha i Pittsa odzwierciedla podstawowe mechanizmy działania biologicznego neuronu, posiada on jednak istotne ograniczenia, które wpływają na jego funkcjonowanie. Przede wszystkim, model ten jest deterministyczny i statyczny, co oznacza, że zawsze dostarcza te same wyniki dla tych samych danych wejściowych. Nie posiada zdolności do zmieniania się w czasie lub adaptowania do nowych informacji, co jest cechą charakterystyczną dla biologicznych neuronów, które wykazują dużą dynamikę i zmienność w swoim działaniu [19, 21].

Model McCullocha i Pittsa jest w stanie symulować prostą logikę binarną i wykonywać podstawowe operacje, takie jak AND, OR i NOT. Jednak niezdolność do nauki i adaptacji znacznie ogranicza jego zastosowanie. Mimo uproszczonej struktury, obliczenia wykonywane przez model McCullocha i Pittsa są procesem złożonym i czasochłonnym. Wynika to zarówno z natury operacji logicznych jakie powinien wykonywać, jak i z konieczności dokładnej kontroli wszystkich elementów procesu. Bez współczesnych narzędzi i technologii, takie jak sieci neuronowe czy algorytmy uczące, model ten jest daleki od efektywności biologicznego neuronu [22].

W konsekwencji, pomimo swojego fundamentalnego znaczenia dla historii rozwoju sztucznej inteligencji, model neuronu McCullocha i Pittsa jest teraz przede wszystkim historycznym przypomnieniem o początkach tej dziedziny. Został zastąpiony przez nowocześniejsze modele, które są w stanie naśladować złożoność i dynamikę działania biologicznych neuronów oraz zachować dużo większą dokładność predykcji.

3.2. ROZWÓJ SIECI NEURONOWYCH

Po początkowym okresie prób odwzorowania zasad działania biologicznych neuronów, nadszedł czas na dalszy rozwój idei sztucznych sieci neuronowych. Początek tej ery datuje się na rok 1957, kiedy to Frank Rosenblatt, psycholog i pionier w dziedzinie uczenia maszynowego, wprowadził swój model o nazwie **perceptron** [23, 24].

3.2.1. PERCEPTRON ROSENBLATTA

Rosenblatt rozwinął neuron McCullocha i Pittsa, nadając mu zdolność do nauki oraz adaptacji, co oznaczało ogromny postęp w porównaniu do dotychczasowego statycznego i deterministycznego modelu. Perceptron, nazywany również jednokierunkową siecią neuronową, opierał się na pojedynczym neuronie, który mógł się uczyć poprzez dostosowywanie wag wejść w odpowiedzi na błędy w przewidywaniach [23, 24].

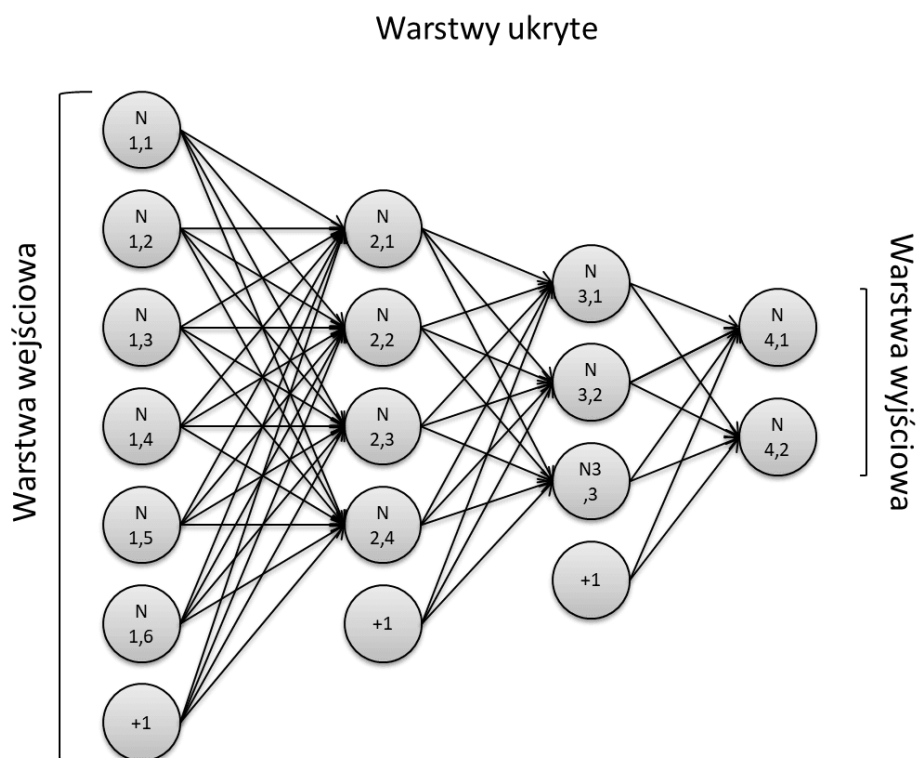
Mechanizm działania perceptronu był dość prosty. Na początku, wagi były inicjowane losowo. Następnie, dla każdego przykładu treningowego, perceptron dokonywał przewidywania i porównywał je z rzeczywistymi etykietami, a uzyskana skuteczność była oceniana za pomocą funkcji błędu. Jeśli udzielona odpowiedź była nieprawidłowa, to przed kolejnym podaniem przykładów ze zbioru treningowego aktualizowano wagi. Proces aktualizacji wag polegał na dodaniu do nich iloczynu błędu i wartości wejściowych, pomnożonego przez tzw. współczynnik uczenia. Współczynnik uczenia był stałą wartością, która decydowała, jak duże korekty były wprowadzane do wag. Wielokrotne powtarzanie opisanego schematu powodowało zminimalizowanie wartości funkcji błędu, które oznaczało polepszenie skuteczności modelu, a tym samym jego uczenie się [23, 24].

Choć perceptron Rosenblatta był znacznie bardziej zaawansowany niż model McCullocha i Pittsa, posiadał on pewne ograniczenia. Najważniejszym z nich było to, że mógł on prawidłowo klasyfikować tylko te zestawy danych, które były liniowo separowalne. Oznacza to, że musiała istnieć prosta lub hiperpłaszczyzna, która mogła oddzielić różne klasy danych bez żadnych błędów. Jeżeli takiej linii nie dało się wyznaczyć, perceptron nie był w stanie nauczyć się poprawnej klasyfikacji. W praktyce nadawał się on do rozwiązywania problemów wymagających wyboru pomiędzy jedną z dwóch klas [23].

3.2.2. SIECI WIELOWARSTWOWE I ALGORYTM PROPAGACJI WSTECZNEJ

Rozwiązaniem problemu liniowej separowalności, który ograniczał pierwsze perceptrony, okazały się sieci wielowarstwowe, zwane również sieciami gęsto połączonymi lub perceptronem wielowarstwowym (MLP). Ich implementacja wykorzystuje dokładnie te same neurony McCullocha i Pittsa, jednak wyróżniają się one odmienną architekturą. Ich budowa polega na pogrupowaniu pewnej liczby neuronów jeden obok drugiego, w taki sposób, by zostały wydzielone co najmniej trzy warstwy: warstwa wejściowa, jedna lub więcej warstw ukrytych oraz warstwa wyjściowa [24].

W sieciach jednowarstwowych, takich jak perceptron, wszystkie wejścia węzłów źródłowych (sygnały zewnętrzne) są połączone bezpośrednio z jedną warstwą neuronów wyjściowych. Natomiast w sieciach wielowarstwowych, neurony warstwy wejściowej są połączone z neuronami pierwszej warstwy ukrytej, a te z neuronami warstwy wyjściowej. Neurony z jednej warstwy przekazują swoje wartości do wszystkich neuronów warstwy kolejnej, ale nie wymieniają informacji w obrębie tych samych lub wcześniejszych grup. Z kolei każde powstałe połączenie ma przypisaną wagę modyfikującą przekazywaną wartość. Schemat takiego modelu został przedstawiony na rysunku 3.3 [23, 24].



Rys. 3.3. Schemat wielowarstwowej sieci neuronowej [25]

W sieciach MLP, warstwy ukryte wprowadzają dodatkowy stopień złożoności, który umożliwia modelowanie skomplikowanych zależności między danymi wejściowymi i wyjściowymi. Dodatkowe warstwy pozwalają na tworzenie nowych cech na wyższym poziomie abstrakcji, które nie były pierwotnie dostępne w danych wejściowych. Dzięki temu, MLP mogą modelować zależności nieliniowe, czego nie potrafiły perceptrony jednowarstwowe. Oznacza to, że są w stanie rozwiązywać problemy liniowo nieseparowalne [23, 24].

Problemem sieci wielowarstwowych było jednak to, że nie były one efektywne w użyciu. Obliczanie wag dla wielu warstw, zwłaszcza w dużych sieciach, było zadaniem niezwykle złożonym. Istniało również ryzyko utknięcia algorytmu w lokalnym minimum podczas optymalizacji, co mogło prowadzić do suboptymalnych rozwiązań.

Wszystko zmieniło się w 1975 roku, kiedy Paul Werbos przedstawił algorytm propagacji wstecznej (ang. Backpropagation). Algorytm ten umożliwił efektywne trenowanie sieci neuronowych z wieloma warstwami [24, 26].

Algorytm propagacji wstecznej działa w dwóch etapach: propagacji do przodu i propagacji wstecznej. W fazie propagacji do przodu, dane wejściowe są przekazywane przez sieć od warstwy wejściowej do warstwy wyjściowej. Każdy neuron odbiera dane wejściowe, mnoży je przez swoje wagi, sumuje i przepuszcza przez funkcję aktywacji, zanim przekaże wynik do kolejnej warstwy [24, 26].

Po zakończeniu fazy do przodu, sieć oblicza błąd przewidywania, porównując wyjście sieci z oczekiwanym wynikiem. Błąd ten jest następnie przekazywany z powrotem przez sieć, od warstwy wyjściowej do warstwy wejściowej, stąd nazwa "propagacja wsteczna". Podczas propagacji wstecznej, algorytm oblicza gradient funkcji błędu względem każdej wagi w sieci, co pozwala na aktualizację wag w sposób, który minimalizuje błąd. W praktyce oznacza to, że algorytm analizuje popełniane błędy, dostosowując wagi w sieci tak, aby

z czasem zwiększyć skuteczność przewidywań. Backpropagation to rewolucyjny wynalazek, który pozwolił na efektywne trenowanie dużych sieci wielowarstwowych, otwierając drogę do rozwoju głębokich sieci neuronowych [24, 26].

3.2.3. GŁĘBOKIE SIECI NEURONOWE

Wraz z upływem lat, rozwój techniki nieustannie zwiększał moce obliczeniowe dostępnych komputerów, co doprowadziło do znacznego przyspieszenia procesu trenowania sieci neuronowych oraz wzrostu ilości dostępnych danych. Dzięki temu sieci wielowarstwowe pręźnie się rozwijały z czasem przeradzając się w bardziej złożone struktury, które znane są dziś jako głębokie sieci neuronowe (ang. deep neural networks).

Uczenie głębokie bazuje na rozbudowanych sieciach neuronowych, składających się z wielu warstw ukrytych, które mogą przetwarzać informacje na coraz bardziej złożone poziomy abstrakcji. W efekcie, sieć jest w stanie uczyć się złożonych zależności i wzorców w danych, co czyni ją niezwykle potężnym narzędziem do przetwarzania i analizowania informacji, niejednokrotnie przewyższając ludzkie możliwości. Są one skutecznie wykorzystywane w dziedzinach takich jak:

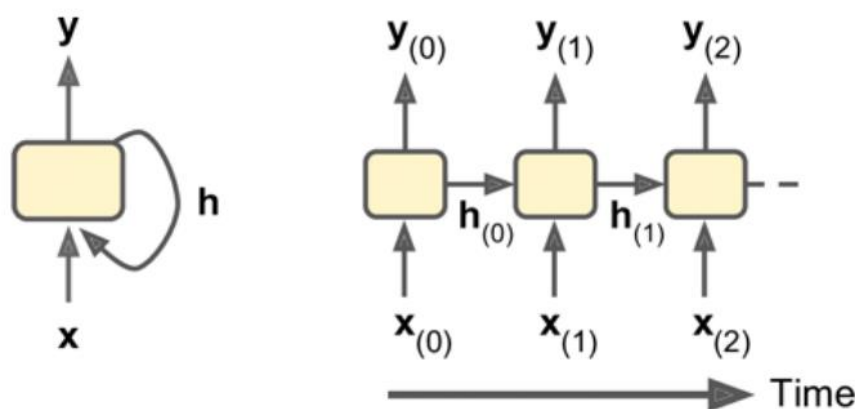
- **Rozpoznawanie obrazów i wideo** – głębokie sieci neuronowe są wykorzystywane do detekcji oraz klasyfikacji obiektów na zdjęciach lub filmach, śledzenia ruchów lub rozpoznawania twarzy, które pozwoliło na stworzenie systemów autoryzacji aplikacji i urządzeń za pomocą twarzy.
- **Tworzenie i edycja obrazów** – ponadto potrafią tworzyć w pełni realistyczne, sztuczne obrazy, które w rzeczywistości nigdy nie powstały. Są w stanie generować zdjęcia osób, które nie istnieją, zmieniać warunki pogodowe w tle, a nawet przenosić style artystyczne z jednego obrazu na inny.
- **Rozpoznawanie mowy** – sieci te pomagają również w transkrypcji mowy na tekst, dzięki czemu przyczyniły się do powstania asystentów głosowych taki jak Asystent Google lub Siri.
- **Przetwarzanie języka naturalnego (NLP)** – głębokie sieci neuronowe umożliwiają analizę i generowanie języka naturalnego. Umożliwiło to zbudowanie zaawansowanych komunikatorów i chatów, które charakteryzują się niezmiernie obszerną wiedzą oraz inteligentnym zrozumieniem użytkowników.
- **Medycyna** – są wykorzystywane do analizy obrazów medycznych, detekcji chorób na podstawie objawów czy genetycznych markerów, a także do przewidywania efektów zażycia leków.
- **Pojazdy autonomiczne** – głębokie sieci neuronowe to także kluczowe narzędzie w rozwoju pojazdów autonomicznych, gdzie są wykorzystywane do analizy danych z czujników i podejmowania decyzji o wykonywanych manewrach.
- **Rekomendacje** – wiele systemów rekomendacji, takich jak te używane przez Netflix czy Amazon, opiera się na głębokich sieciach neuronowych do analizy preferencji użytkowników i generowania spersonalizowanych podpowiedzi lub reklam.
- **Gry strategiczne** – dzięki możliwością analizy skomplikowanych kombinacji, głębokie sieci neuronowe stały się podstawą niepokonanych algorytmów w grach takich jak szachy, warcaby lub Go, które z łatwością ogrywają najlepszych graczy na świecie. Są w stanie tworzyć innowacyjne oraz zaskakujące strategie, które były uprzednio nie odkryte przez człowieka.
- **Finanse** – głębokie uczenie znajduje swoje zastosowanie także w świecie finansów, gdzie generuje podpowiedzi dotyczące strategii inwestycyjnych. [27, 28]

3.3. REKURENCYJNE SIECI NEURONOWE

Rekurencyjne sieci neuronowe (ang. Recurrent Neural Networks, RNN) to specyficzny typ sieci neuronowych, które wykorzystują sekwencyjność danych wejściowych w celu wygenerowania ostatecznej odpowiedzi. Dzięki temu doskonale sprawdzają się w przetwarzaniu i analizie szeregów czasowych, co czyni je szczególnie użytecznymi w kontekście przewidywania cen akcji i innych instrumentów finansowych. Używa się ich również do przetwarzania języka naturalnego, generowania tekstu, rozpoznawania mowy, tłumaczenia maszynowego i wielu innych zadań, które wymagają analizy danych zmiennych w czasie [29÷31].

3.3.1. STRUKTURA I BUDOWA

Rekurencyjne sieci neuronowe wyróżniają się możliwością uwzględnienia dodatkowego wymiaru danych jakim jest czas. W tradycyjnej sieci neuronowej, wszystkie dane wejściowe są traktowane niezależnie, co oznacza że nie ma znaczenia w jakiej kolejności zostaną one podane na wejście sieci. RNN biorą pod uwagę sekwencyjny charakter danych, przez co informacje z przeszłych kroków są wykorzystywane do generowania przewidywań dla aktualnego wejścia [29, 30].



Rys. 3.4. Schemat neuronu rekurencyjnego [29]

Podstawowa struktura rekurencyjnej sieci neuronowej składa się z pojedynczego neuronu, na którego wejściu pojawiają się dwa rodzaje danych:

- $x(t)$: standardowa dana wejściowa w kroku czasowym t ,
- $h(t)$: komórka pamięci zwana stanem ukrytym. W najprostszej postaci jej wartość odpowiada informacji na wyjściu neuronu $y(t)$ dla kroku czasowego $t = t - 1$. [29, 30]

Schemat działania rekurencyjnej sieci neuronowej został przedstawiony na rysunku 3.4. W pierwszej kolejności neuron otrzymuje dane wejściowe $x(t)$ oraz stan ukryty $h(t)$ obliczony w poprzednim kroku. Dodatkowo zarówno informacje wejściowe jak i komórka pamięci posiadają swoje osobne wagi Wx oraz Wh , które są aktualizowane w procesie uczenia sieci. Suma wejścia pomnożonego przez wagę Wx i stanu ukrytego pomnożonego przez wagę Wh , zostaje przekazana do funkcji aktywacji, której wartość jest jednocześnie wartością wyjścia $y(t)$ i stanu ukrytego $h(t)$ dla następnego kroku czasowego $t + 1$ [29, 30].

Ten prosty mechanizm umożliwił rekurencyjnym sieciom neuronowym znacznie większą efektywność w przetwarzaniu danych sekwencyjnych, jednak jego podstawowa wersja borykała się z kluczowymi problemami. Zastosowanie komórki pamięci obejmującej jedynie ostatni krok nie było szczególnie skutecznym rozwiązaniem. Taki model traci wiedzę na temat starszych informacji zaraz po przejściu do następnej iteracji, przez co nie radzi sobie z zadaniami takimi jak analiza sentymentu czy predykcja cen akcji. Próby rozbudowania funkcji stanu ukrytego w taki sposób, by przechowywał on wartości z bardziej oddalonych kroków czasowych również nie przyniosły pozytywnych rezultatów. Głównym problemem takiego podejścia okazało się zjawisko zanikającego gradientu [29, 30].

W trakcie procesu uczenia złożonych sieci neuronowych dąży się do znalezienia odpowiednich wag dla wszystkich połączeń pomiędzy neuronami za pomocą metody propagacji wstecznej. Jeśli zauważymy, że w trakcie tego procesu, gradient błędu dla wczesnych warstw uzyskuje się poprzez pomnożenie gradientów późniejszych warstw, to przy ich wartościach mniejszych niż jeden, wartość pochodnej dla początkowych warstw zaczyna spadać w bardzo szybkim tempie. Efektem tego zjawiska jest zdecydowanie dłuższy czas potrzebny na wytrenowanie sieci, a w skrajnych przypadkach, gdy gradient jest bliski zeru, wagi całkowicie przestają się aktualizować. W ten sposób dalsze uczenie staje się niemożliwe [29, 30].

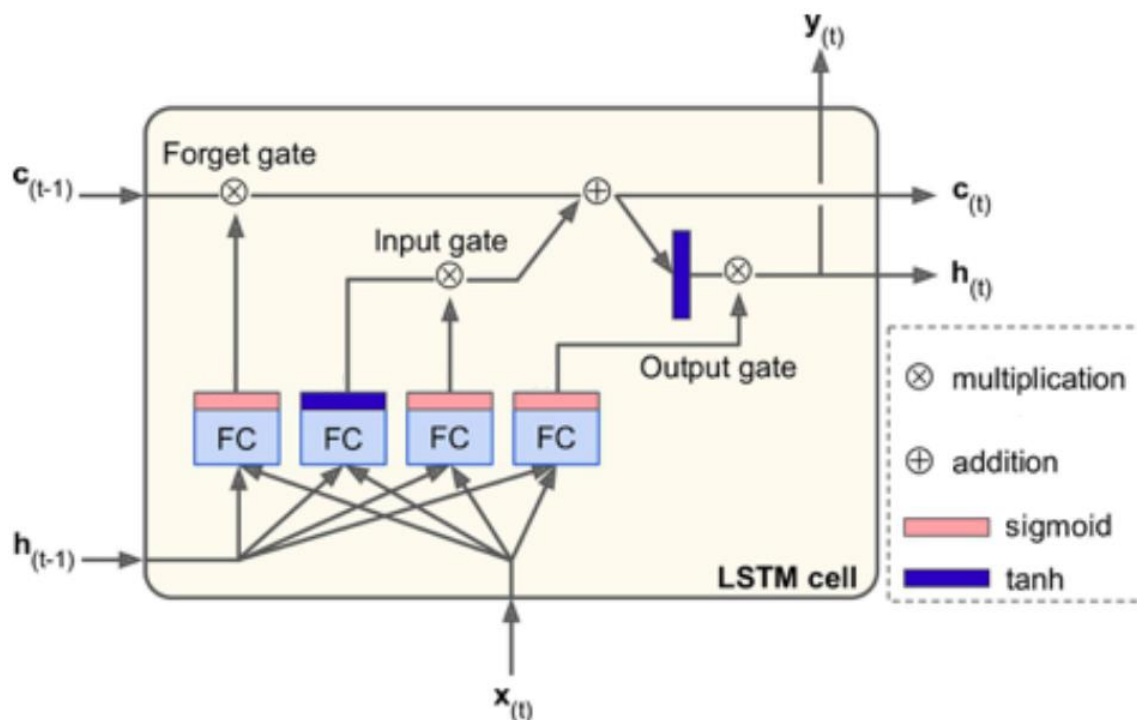
Ten problem okazuje się być szczególnie dotkliwy w kontekście sieci rekurencyjnych. Zapamiętywanie większej ilości informacji wiąże się z uzyskaniem osobnych wag dla każdego stanu ukrytego, co zdecydowanie powiększa ilość optymalizowanych parametrów. Podczas uczenia RNN, gradienty są przenoszone wstecz przez wszystkie kroki czasowe. Oznacza to, że jeśli sieć ma na celu nauczenie się zależności pomiędzy znacznie oddalonymi elementami, gradienty te mogą zanikać do momentu, gdy stają się zbyt małe, aby skutecznie zaktualizować wagi sieci [29, 30].

3.3.2. LSTM

Sieci neuronowe typu LSTM (ang. Long Short-Term Memory) to specjalny rodzaj rekurencyjnych sieci neuronowych, które zostały zaprojektowane, aby skutecznie radzić sobie z problemami zanikającego gradientu oraz szybkiego gubienia informacji występujących w podstawowych modelach RNN. Choć sieci LSTM zostały wprowadzone pod koniec zeszłego wieku, do dzisiaj są jednymi z najbardziej efektywnych rozwiązań dla przetwarzania danych sekwencyjnych [29, 30].

Podobnie jak w przypadku standardowych RNN, podstawowy element sieci LSTM to neuron, niemniej jednak jego struktura jest zdecydowanie bardziej złożona. Składa się ona ze znacznie większej liczby elementów, których głównym zadaniem jest zarządzanie dwoma rodzajami pamięci:

- **krótkoterminowej $h(t)$** , która jest odwzorowaniem stanu ukrytego mającego na celu przekazywanie informacji pomiędzy kolejnymi krokami oraz przechowuje informacje potrzebne do krótkoterminowych zadań,
- **długoterminowej $c(t)$** , która pozwala sieci na przechowywanie i odzyskiwanie informacji przez dłuższy czas. W trakcie uczenia jest każdorazowo aktualizowana, a jej aktualna wartość jest nazywana komórką stanu. [29, 30]



Rys. 3.5. Budowa komórki LSTM [29]

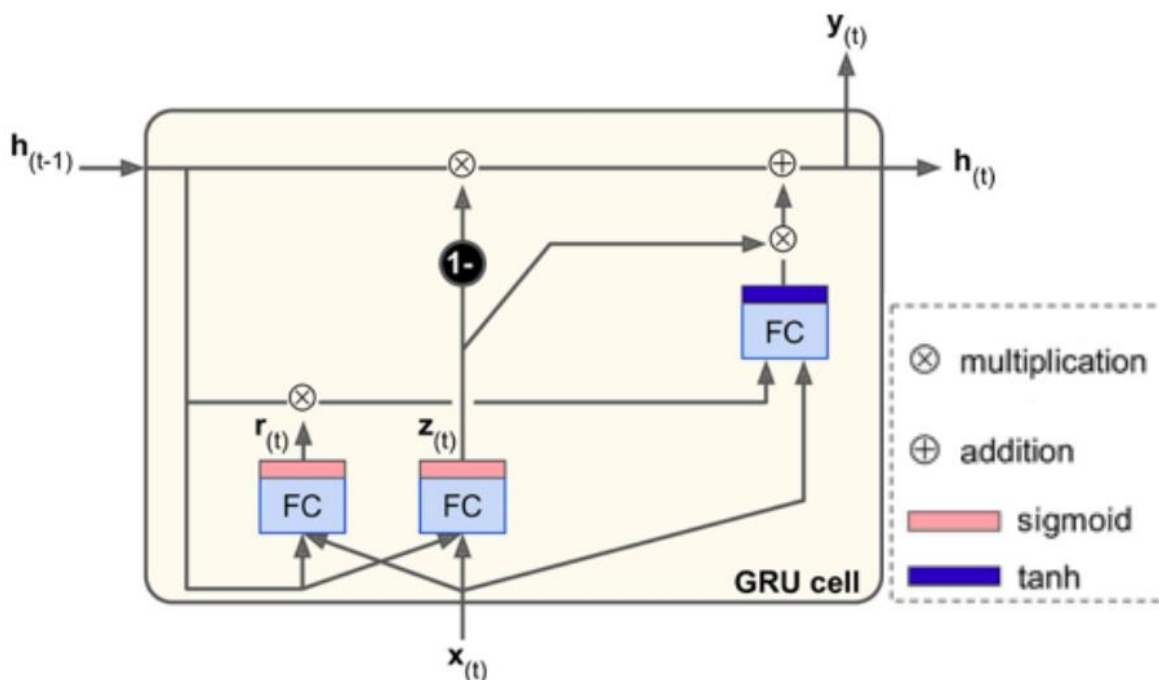
W celu zarządzania dwoma rodzajami pamięci, komórka LSTM posiada trzy rodzaje bramek, odpowiedzialnych za koordynację przepływu informacji w trakcie procesu uczenia:

- **Bramka Zapomnienia (Forget Gate)** – Bramka zapomnienia decyduje, które informacje zostaną odrzucone z komórki stanu. W pierwszym etapie otrzymuje ona informacje z aktualnego wejścia sieci oraz poprzedniego stanu ukrytego, a następnie przechodzi przez funkcję sigmoidalną. Wyjście z tej funkcji to wartość z zakresu od 0 do 1, która jest używana do mnożenia z aktualnym stanem długoterminowym. Im bardziej zbliża się ona do 0, tym większy jest stopień zapomnienia długoterminowych informacji. W ten sposób brama zapomnienia usuwa niechciane informacje z pamięci komórki [29÷31].
- **Bramka Wejściowa (Input Gate)** – Bramka wejściowa decyduje, które informacje zostaną zaktualizowane lub dodane do komórki stanu. Tak jak brama zapomnienia, otrzymuje ona aktualne wejście oraz poprzedni stan ukryty, ale przekazuje je przez dwie różne funkcje: sigmoidalną i tangens hiperboliczny. Wartość tangensu to nowy kandydat na stan komórki i jest on mnożony przez analogiczne wyjście z funkcji sigmoidalnej, tak jak w poprzedniej bramce. Ostateczna wartość jest dodawana do komórki stanu, która od teraz stanowi końcową postać pamięci długoterminowej $c(t)$ dla aktualnego kroku czasowego [29÷31].
- **Bramka Wyjściowa (Output Gate)** – Bramka wyjściowa decyduje, jaki będzie następny stan ukryty. Ponownie wyjście z funkcji sigmoidalnej uzyskiwane jest na podstawie aktualnego wejścia i poprzedniego stanu ukrytego. Tym razem wartość ta jest mnożona z pamięcią długoterminową przekazaną do funkcji tangensu hiperbolicznego. Otrzymany iloczyn wyznacza wartość pamięci krótkoterminowej $h(t)$ oraz wyjścia $y(t)$ [29÷31].

Dzięki zastosowaniu powyższych mechanizmów, sieci LSTM są w stanie przechowywać i przetwarzać informacje przez dłuższe okresy, co czyni je idealnym wyborem dla złożonych zadań związanych z danymi sekwencyjnymi.

3.3.3. GRU

Sieci GRU (ang. Gated Recurrent Unit) są obok LSTM drugim, najpopularniejszym rodzajem rekurencyjnych sieci neuronowych stosowanym do analizy szeregów czasowych. Zostały wprowadzone w 2014 roku, jako uproszczona alternatywa dla sieci LSTM, która oferuje podobną zdolność do modelowania długoterminowych zależności, ale jest mniej skomplikowana i szybsza do trenowania [29, 30].



Rys. 3.6. Budowa komórki GRU [29]

Podobnie jak LSTM, sieci GRU mają mechanizm bramek, który pomaga im z łatwością przechowywać i zapominać informacje, jednak robią to w bardziej skondensowany sposób. GRU posiadają jedynie jeden rodzaj pamięci, który jest zarządzany przez dwa rodzaje bramek:

- **Bramka aktualizacji (Update Gate)** – Bramka aktualizacji decyduje, jak wiele informacji z poprzedniego stanu ukrytego ma zostać przeniesione do obecnego stanu. Gdy jej wartość jest bliska 1, stan ukryty jest zaktualizowany tylko nieznacznie, a większość informacji z poprzedniego stanu zostaje zachowana. W przeciwnym razie stan ukryty ma mniejsze znaczenie, co pozwala na zapomnienie starszych danych [30, 31].
- **Bramka resetowania (Reset Gate)** – Bramka resetowania kontroluje, jak wiele informacji z poprzedniego kroku powinno być użyte przy obliczaniu kandydata na nowy stan ukryty. Gdy wartość bramki resetowania jest bliska 1, nowa wartość pamięci będzie pokrywać się z poprzednim stanem ukrytym. W odwrotnej sytuacji stan ukryty bazuje w głównej mierze na nowej wartości wejściowej [30, 31].

Dzięki temu, że sieci GRU mają mniej bramek oraz nie posiadają oddzielnej komórki pamięci, są znacznie szybsze do trenowania, co poprawia ich skuteczność przy niewielkiej ilości danych. W praktyce, sieci GRU i LSTM wykazują często podobną skuteczność, a wybór pomiędzy nimi zależy w głównej mierze od specyfiki problemu i dostępnych zasobów obliczeniowych.

4. PRZEGLĄD LITERATURY

W artykule „Stock Market Trend Prediction and Investment Strategy by Deep Neural Networks”, Shi, M. oraz Zhao, Q. postanowili zbadać możliwość rozpoznawania trendów giełdowych za pomocą głębokich sieci neuronowych [32]. Główna idea pracy polega na porzuceniu rozwiązania wykorzystującego prognozowanie cen akcji ze względu na ich nierugalmą skalę wahań. W zamian tego, autorzy proponują znalezienie sygnałów rozpoznających rodzaj nadchodzącego trendu.

Do wykonania tego zadania, na wykresie notowania S&P 500 w latach 1995 – 2020, odnalezione zostały wszystkie wystąpienia dwóch formacji technicznych, jakimi są złoty krzyż oraz krzyż śmierci. W celu znalezienia obu wskaźników, wyznaczane są dwie średnie kroczące, gdzie jedna z nich to średnia długoterminowa, zaś druga krótkoterminowa. Złoty krzyż pojawia się, gdy średnia o krótszym interwale przecina się powyżej średniej długoterminowej. Taka sytuacja może sygnalizować początek długotrwałego trendu wzrostowego. Z kolei krzyż śmierci powstaje podczas przeciwnego przecięcia się obu wskaźników i może oznaczać początek spadku wartości akcji.

Następnie za pomocą metody obliczania punktu zwrotnego, podzielono zebrane dane na krzyże prawdziwe, które poprawnie zwiastowały nadajęcie określonego trendu oraz fałszywe, czyli takie po których nie zaobserwowano spodziewanej reakcji rynku. Na podstawie przygotowanego zbioru danych, dokonano klasyfikacji za pomocą głębokich sieci neuronowych MLP. Wejście sieci zawierało informacje giełdowe z ostatnich 60 dni przed wystąpieniem wyznaczonej formacji, zaś wyjście miało określić czy krzyż jest prawdziwy lub fałszywy.

Rezultat przeprowadzonego uczenia początkowo nie przyniósł dobrego rezultatu. Co prawda model poprawnie sklasyfikował 80.75% sygnałów wejściowych, jednak ze względu na wysokie niebalansowanie klas, wynoszące 79.46 % do 20.54%, uzyskany wynik nie dawał realnych korzyści. Dopiero zastosowanie prognozy akceptującej odpowiedzi powyżej 90% dokładności predykcji, pozwoliło na znaczne zwiększenie skuteczności. W takiej sytuacji 93% wszystkich krzyży zostało sklasyfikowane poprawnie. Oczywiście odbyło się to kosztem pomijania części danych, jednak transakcje przeprowadzone na bazie tak wyselekcjonowanych predykcji, znacznie zwiększają szanse na dokonanie właściwej inwestycji. Zdaniem autorów, dobranie prognozy o wartości 0.9 równoważy ryzyko oraz zysk podczas transakcji giełdowych. [32]

W kolejnym artykule, Thomas Fisher oraz Christopher Krauss zdecydowali się stworzyć system transakcyjny oparty na prognozowaniu kursu notowania S&P 500 w latach 1992 – 2015 [33]. Głównym celem pracy jest wykorzystanie rekurencyjnych sieci neuronowych LSTM oraz porównanie ich z trzema innymi metodami uczenia maszynowego: drzewami decyzyjnymi, standardowymi głębokimi sieciami neuronowymi i regresją logistyczną.

W pierwszej kolejności przygotowano dane uczące w taki sposób, by wejścia zawierały 240 kroków czasowych. Następnie zaimplementowano architekturę każdego z czterech testowanych modeli. Budowa sieci LSTM składała się z jednej warstwy ukrytej, liczącej 25 komórek oraz 10 procentowego dropoutu. Po przeprowadzeniu procesu uczenia, autorzy stworzyli strategię, mającą na celu przekształcanie uzyskanych informacji na sygnały transakcyjne. Za miarę efektywności zaproponowanych rozwiązań wybrana została dzienna stopa zwrotu. Jej uśredniona wartość pozwoli wybrać metodę, która może przynieść największy przychód. [33]

Przeprowadzone badania składały się ze 100 000 symulacji, sprawdzających osiągnięte zyski. Okazało się, że sieć LSTM uzyskała najwyższą średnią stopę zwrotu na poziomie 0.46%. Po odjęciu kosztów transakcji, model nadal był w stanie zarabiać 0.26% dziennie. Niewiele gorzej sprawdziła się regresja logistyczna, która zakończyła testy z wynikiem 0.43%. Z kolei klasyczna głęboka sieć neuronowa uzyskała znacznie gorszy rezultat, w postaci dziennej stopy zwrotu o wartości 0.32%. Badanie udowadnia, że zastosowanie modeli rekurencyjnych podczas analizy szeregów czasowych przynosi wyraźne korzyści.

W dalszej części artykułu, sprawdzono także w jaki sposób strategie radzą sobie w poszczególnych okresach zbioru testowego. Okazuje się, że opłacalność stosowania metod uczenia maszynowego skończyła się w 2010 roku. W pozostałych 5 latach średnia stopa zwrotu dla modelu LSTM oscylowała w okolicach zera, zaś pozostałe podejścia odnotowały ujemne wyniki. Taka sytuacja ma związek z kryzysem ekonomicznym, który długotrwale odwrócił trendy na amerykańskich rynkach finansowych. Należy więc pamiętać, że wykorzystanie sztucznej inteligencji podczas podejmowania decyzji inwestycyjnych nie zabezpiecza przed nagłymi zmianami gospodarczymi.

Na podstawie zebranych informacji oraz wiedzy pochodzącej z dotychczas przeprowadzonych badań, wykonany zostanie projekt magisterski, w którym zaproponowane będą własne propozycje zastosowania głębokich sieci neuronowych podczas wspomagania decyzji inwestycyjnych. W kolejnych rozdziałach pracy znajdzie się szczegółowy opis poszczególnych etapów badań wraz z analizą otrzymanych wyników oraz wyciągnięciem wniosków.

5. PRZYGOTOWANIE ZBIORU DANYCH

Pierwszym etapem przygotowującym bazę do wykonania całego projektu jest zebranie oraz przygotowanie odpowiednich danych. Do dalszych badań wykorzystana zostanie para walutowa EUR/USD, która reprezentuje ilość dolarów amerykańskich (USD) jakie można otrzymać za jedno euro (EUR). Jest to najpopularniejsza para na rynku Forex, która charakteryzuje się wysoką płynnością oraz stabilnością ceny, dzięki czemu jest chętnie wybierana przez osoby zajmujące się tradingiem. Doskonale sprawdzi się również jako baza do badań nad strategiami, gdyż jej cena zmienia się bardzo często, ale utrzymuje stały poziom, oscylujący w granicach nieznacznie przekraczających jednego dolara. Para EUR/USD nie spowoduje zatem znacznych przekłamań wynikających z ciągłego wzrostu lub spadku kursu, tym samym umożliwiając wykonanie rzetelnych badań.

5.1. POBRANIE DANYCH

Dane historyczne dla pary walutowej EUR/USD zostały pobrane za pomocą platformy MetaTrader 4, która jest jednym z najpopularniejszych narzędzi dla inwestorów na rynku Forex. Platforma oferuje szereg zaawansowanych funkcji, w tym możliwość pobierania szczegółowych danych historycznych dla instrumentów finansowych od brokerów z całego świata. Niezwykle przydatna funkcja, która charakteryzuje MetaTrader, to dostępność wielu wskaźników, zaczynając od najbardziej podstawowych, a kończąc na profesjonalnych formułach, które pozwalają na wykonanie zaawansowanej analizy technicznej. Dodatkowo platforma oferuje możliwość pisania własnych wskaźników w języku MQL.

Pobrane dane obejmują okres od początku roku 2000 i są przedstawione w formie świec japońskich z interwałem dziennym. Każda świeca reprezentuje jednodniowy okres handlu i składa się z następujących elementów:

- **Open:** cena, po której rozpoczął się handel w danym dniu,
- **High:** najwyższa cena w ciągu dnia,
- **Low:** najniższa cena w ciągu dnia,
- **Close:** cena, po której zakończył się handel w danym dniu.

Tab. 5.1. Pobrane dane EUR/USD

	date	open	high	low	close
0	2000-01-02	1.00730	1.01410	1.00730	1.01370
1	2000-01-03	1.01400	1.02780	1.00540	1.02620
2	2000-01-04	1.02600	1.03400	1.02130	1.02920
3	2000-01-05	1.02910	1.04020	1.02840	1.03230
4	2000-01-06	1.03250	1.04150	1.02720	1.03050
...
7149	2022-12-07	1.04670	1.05493	1.04427	1.05114
7150	2022-12-08	1.05115	1.05643	1.04892	1.05551
7151	2022-12-09	1.05549	1.05878	1.05032	1.05325
7152	2022-12-11	1.05286	1.05371	1.05176	1.05204
7153	2022-12-12	1.05204	1.05795	1.05054	1.05407

Przebieg ceny na zamknięciu w latach 2000-2022 został przedstawiony na rysunku 5.1. Można na nim zaobserwować, że na przestrzeni lat cena ma tendencję do częstych zmian, ale widoczne są również charakterystyczne trendy wzrostowe i spadkowe z niewielką ilością trendów bocznych.



Rys. 5.1. Przebieg ceny close EUR/USD w latach 2000 - 2022

5.2. DOBÓR WSKAŹNIKÓW

Pozyskany zbiór danych zawiera jedynie surowe ceny z poszczególnych dni. W celu wzbogacenia go o dodatkowe dane, które mogą poprawić skuteczność sieci neuronowych, dodane zostanie kilka popularnych i powszechnie wykorzystywanych wskaźników. Większa ilość danych wejściowych może również usprawnić proces trenowania i zmniejszyć ryzyko wystąpienia przeuczenia. Z tego względu dodane zostaną:

- **Return rate** – stopa zwrotu przedstawia stosunek ceny na zamknięciu z dnia n do tej samej wartości z dnia $n - 1$. Ten najprostszy wskaźnik informuje czy inwestycja w danym okresie została zakończona zyskiem lub stratą.
- **RSI (Relative Strength Index)** – wskaźnik siły względnej, który określa siłę aktualnego trendu. RSI przyjmuje wartości z zakresu od 0 do 100 oraz jest dany wzorem:

$$RSI = 100 - \frac{100}{1 + \frac{a}{b}} \quad (1)$$

,gdzie:

a – to zmienna określająca średnią wartość zysku z ostatnich n dni,

b – to zmienna określająca średnią wartość strat z ostatnich n dni.

Wszystkie zyski oraz straty są określane na podstawie stopy zwrotu z poszczególnych dni. Jeśli w danym okresie nie ma żadnych strat, to iloraz a oraz b będzie dążył do nieskończoności, co w efekcie da wartość wskaźnika RSI równą 100.

W sytuacji odwrotnej RSI będzie maleć do zera. Wysokie wartości wskaźnika mogą wskazywać na szansę do odwrócenia trendu na spadkowy, natomiast niskie do powrotu na trend wzrostowy. W pracy wykorzystane zostały dwa rodzaje wskaźnika RSI: z 7 oraz 14 dni.

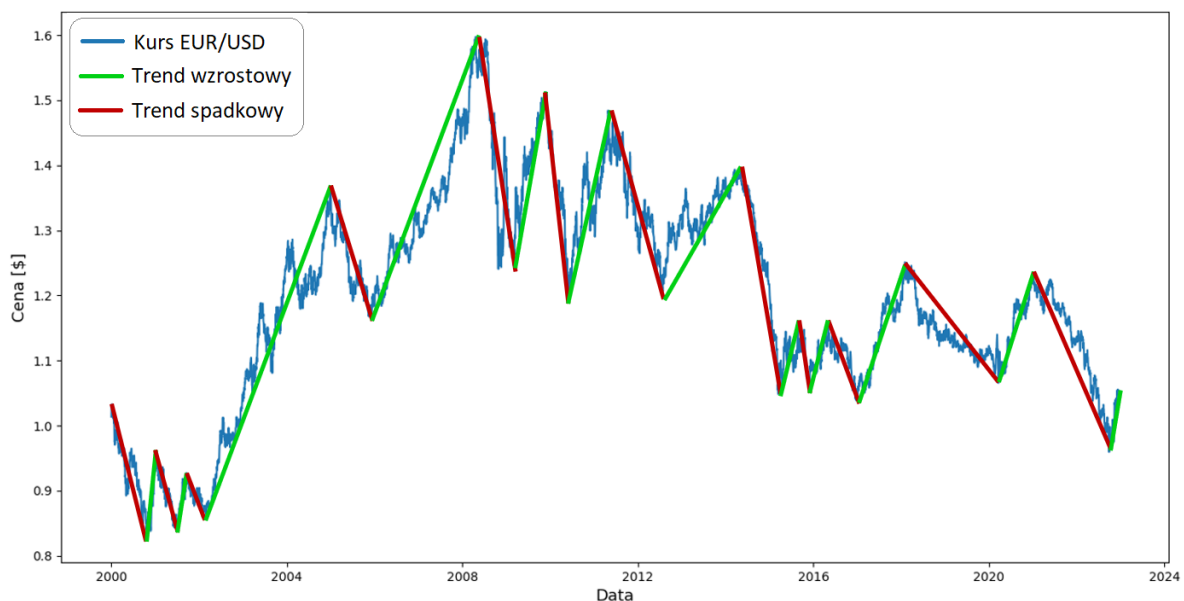
- **MA (Moving average)** - średnia krocząca lub inaczej średnia ruchoma, to wskaźnik monitorujący trendy na podstawie średniej wyciągniętej z wartości danego instrumentu w ostatnich n dniach. Podobnie jak w przypadku RSI, wskaźnik obliczony będzie na bazie ostatnich 7 oraz 14 dni.

Ostatnim niezbędnym elementem zbioru danych, dzięki któremu możliwe będzie wykonanie klasyfikacji szeregu czasowego, jest dodanie odpowiedniej etykiety. W tym przypadku założenie projektowe przewiduje znalezienie miejsc, w których następuje zmiana trendu. Niestety dodanie etykiet jedynie w miejscach przełamania, spowodowałoby bardzo wysokie niezrównoważenie klas, które uniemożliwia poprawne wytrenowanie sieci neuronowych. W związku z tym efektywniejszym rozwiązaniem będzie dodanie informacji o aktualnym kierunku zmiany ceny w każdym wierszu zbioru danych.

W tym celu wykorzystany zostanie wskaźnik ZigZag, dostępny na platformie MetaTrader. Składa się on z linii, które łączą lokalne ekstrema cen wykresie, dzięki czemu odwzorowuje historyczne trendy. Wykreślone proste łączą powstałe szczyty i doliny, przez co kształtem przypominają literę „Z”, skąd przyjęła się nazwa ZigZag.

Zasada działania wskaźnika polega na zastosowaniu procentowego kryterium, które analizuje przebieg kursu danego instrumentu. Jego wartość zastosowana w projekcie wynosi 5%. Oznacza to, że każde nowe ekstremum jest wyznaczane w momencie gdy po zmianie trendu cena wzrosła lub spadła o co najmniej 5%.

Na rysunku 5.2 przedstawiony został poglądowy wykres wskaźnika ZigZag. Wszystkie zielone linie oznaczają, że w danym okresie cena miała tendencję rosnącą, a linie czerwone odpowiadają kierunkowi spadkowemu. W rzeczywistości liczba odnalezionych trendów jest zdecydowanie większa, jednak znacznie zmniejszyłaby czytelność wykresu, stąd zdjęcie ma charakter jedynie schematyczny.



Rys. 5.2. Schemat działania wskaźnika ZigZag

Końcowa wersja przygotowanego zbioru danych została przedstawiona w tabeli 5.2. Ostatecznie składa się on z daty, która jest jednocześnie indeksem zbioru, 9 kolumn z cenami i wybranymi wskaźnikami oraz kolumny ZigZag. Wartość tego wskaźnika zapisano zgodnie z poniższą konwencją:

- **0** – trend wzrostowy,
- **1** – trend spadkowy.

W tabeli można zauważyć, że w dniu 31 stycznia 2000 roku cena Close osiąga najniższą wartość po serii dni zakończonych coraz niższym kursem. Od tego momentu cena zaczyna konsekwentnie rosnąć. W związku z tym wartość wskaźnika ZigZag zmienia się z 1 na 0, co oznacza wyznaczenie nowego minimum lokalnego i zmianę trendu na wzrostowy.

Tab. 5.2. Końcowa wersja zbioru danych

Date	Open	High	Low	Close	Return rate	RSI_7	RSI_14	MA_7	MA_14	ZigZag
24.01.2000	1.0043	1.0077	0.9997	1.0057	1.0014	32.3532	36.8229	1.0104	1.0171	1
25.01.2000	1.0063	1.0096	0.9992	1.0014	0.9957	27.2200	33.2253	1.0090	1.0152	1
26.01.2000	1.0012	1.0055	0.9989	0.9999	0.9985	25.5690	32.0490	1.0069	1.0133	1
27.01.2000	0.9997	1.0030	0.9868	0.9870	0.9871	15.8955	24.1353	1.0034	1.0101	1
28.01.2000	0.9873	0.9892	0.9736	0.9749	0.9877	11.2414	19.3170	0.9974	1.0061	1
30.01.2000	0.9773	0.9804	0.9772	0.9796	1.0048	21.6387	25.5355	0.9933	1.0029	1
31.01.2000	0.9796	0.9829	0.9667	0.9705	0.9907	17.1110	22.0000	0.9884	0.9999	0
01.02.2000	0.9707	0.9782	0.9667	0.9717	1.0012	19.6961	23.5041	0.9836	0.9970	0
02.02.2000	0.9718	0.9808	0.9694	0.9766	1.0050	30.0836	29.4836	0.9800	0.9945	0
03.02.2000	0.9767	0.9947	0.9710	0.9891	1.0128	49.5181	41.9496	0.9785	0.9927	0
04.02.2000	0.9890	0.9940	0.9747	0.9840	0.9948	43.7318	38.9260	0.9781	0.9907	0
06.02.2000	0.9832	0.9832	0.9817	0.9818	0.9978	41.3029	37.6649	0.9790	0.9882	0

5.3. PRZYGOTOWANIE DANYCH TRENINGOWYCH I TESTOWYCH

I. NORMALIZACJA

Po dodaniu odpowiednich wskaźników do zbioru danych, przyszła pora na ich odpowiednie przygotowanie do procesu uczenia. W związku z tym, dane zostały przeskalowane do zakresu od -1 do 1. Taki zabieg jest często stosowany w procesie uczenia maszynowego, ponieważ przekształca wszystkie cechy do jednolitego zakresu. Zastosowana normalizacja danych ułatwia proces uczenia, skracając czas potrzebny na osiągnięcie satysfakcjonujących wyników.

II. OKNO CZASOWE

Kolejnym niezbędnym elementem podczas procesu predykcji oraz klasyfikacji cen akcji jest zastosowanie **okna czasowego**. To technika często stosowana w analizie szeregów czasowych, która polega na grupowaniu danych w zależności od długości ustalonego okresu. Za pomocą okna czasowego wyznacza się ilość wierszy, które będą brane pod uwagę jako pojedyncze wejście sieci oraz na podstawie którego wygenerowana zostanie przewidywana odpowiedź. Jeśli okno posiada długość n , to obliczenie pojedynczej odpowiedzi sieci neuronowej nastąpi po podaniu wszystkich danych z n dni poprzedzających przewidywany dzień.

III. WYDZIELENIE ETYKIET

Ze względu na zastosowanie dwóch różnych podejść do uczenia nadzorowanego, dane wejściowe dla klasyfikacji oraz predykcji muszą się od siebie różnić.

W przypadku klasyfikacji, etykietą jest wcześniej opisywany wskaźnik ZigZag, który przyjmuje dwie wartości 0 oraz 1. Kolumna ta musi zostać wydzielona do osobnej tabeli i będzie wykorzystywana wyłącznie do procesu uczenia. Istotną informację stanowi fakt, że po podziale danych na dwie grupy, otrzymano wysoki poziom zrównoważenia klas, wynoszący 51.2% do 48.8% z przewagą dla trendu spadkowego.

Z kolei do zadań predykcji, etykietą jest cena na zamknięciu w kolejnym dniu zbioru danych. Istnieje również możliwość przewidywania cen nie w jednym, a większej ilości następujących dni. Oznacza to, że po podaniu np. ostatnich 30 dni, sieć zwróci swoje przewidywania w postaci cen w kilku następnych dniach. W projekcie przewidywane będą ceny w pięciu kolejnych dniach. Zastosowanie wydłużonego okresu wiąże się z ryzykiem zwiększenia błędu predykcji. Jednakże posiadanie wiedzy na temat bardziej odległej przyszłości stworzy lepszy obraz aktualnej sytuacji rynkowej, które może wpłynąć na podjęcie korzystniejszej decyzji.

IV. PODZIAŁ NA DANE TRENINGOWE, TESTOWE I WALIDACYJNE

Ostatni krok przed przystąpieniem do implementacji sieci neuronowych to podział danych na zbiór treningowy oraz testowy. Często spotykanym podejściem jest podzielenie zbiorów w stosunku 70/30 lub 80/20 z przewagą dla zbioru treningowego. Taki zabieg dostarcza wystarczającą ilość danych do rzetelnego sprawdzenia stworzonego modelu.

W dalszej części pracy modele będą sprawdzane również podczas symulacji giełdowej, na podstawie której zostanie wyznaczona roczna stopa zwrotu, stanowiąca kluczowy element przeprowadzanych badań. W związku z tym korzystnym rozwiązaniem jest wyznaczenie długości zbioru testowego na jeden rok kalendarzowy. Każdy z testowanych okresów będzie posiadał oddzielnie wytrenowany model, na podstawie danych z 10 lat bezpośrednio go poprzedzających. Oznacza to, że przykładowy model testowany na zbiorze z 2021 roku, będzie wytrenowany na danych z okresu 2011 – 2020. Takie podejście zapewnia przekazanie sieciom najświeższych danych, występujących bezpośrednio przed przewidywanym okresem, zabezpieczy przed przypadkowym wykorzystaniem zbioru testowego do nauki oraz ułatwi wykonanie symulacji.

6. UCZENIE SIECI NEURONOWYCH

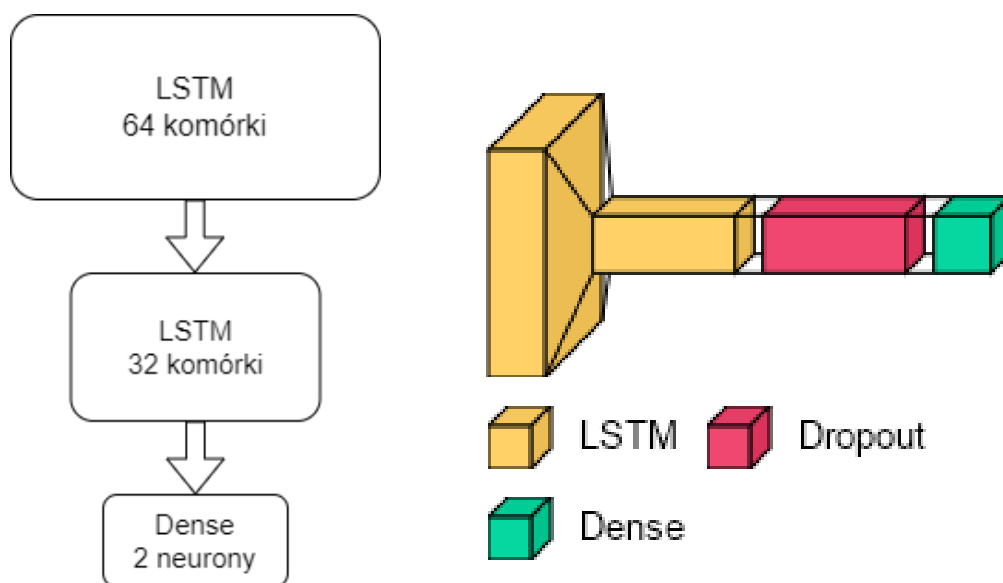
Po przygotowaniu wszystkich niezbędnych danych możliwe jest przystąpienie do przygotowania modeli głębokich sieci neuronowych. W tym celu wykorzystana zostanie biblioteka TensorFlow, która jest najpopularniejszym frameworkiem języka Python wykorzystywanym w uczeniu maszynowym. Dodatkowo będzie ona wzbogacona o bibliotekę Keras, będącą wysokopoziomowym interfejsem użytkownika. Keras jest znany ze swojej intuicyjnej i łatwej do zrozumienia składni, która znacznie upraszcza proces tworzenia złożonych modeli.

6.1. PRZYGOTOWANIE MODELI

W projekcie wykorzystano dwa rodzaje rekurencyjnych sieci neuronowych LSTM oraz GRU, które zostały szczegółowo opisane w drugim rozdziale pracy. W trakcie ich implementacji sprawdzone zostało wiele różnych architektur różniących się liczbą warstw ukrytych, neuronów, zastosowanym dropoutem lub współczynnikiem uczenia. W dalszej części pracy przedstawione zostaną badania przeprowadzone jedynie na modelach, które odznaczyły się najlepszą skutecznością i stabilnością w trakcie procesu uczenia. Bazując na ich wynikach wybrana zostanie najkorzystniejsza szerokość okna czasowego spośród 10, 20 i 30 dni.

6.1.1. KLASYFIKACJA

Schemat modelu LSTM przygotowanego do zadania klasyfikacji szeregu czasowego został przedstawiony na rysunku 6.1. Składa się on z warstwy wejściowej o rozmiarze 64 komórek LSTM/GRU, która przyjmuje wejście w formie 9 wymiarowych danych o długości zależnej od testowanego okna czasowego. Jest ona połączona z pojedynczą warstwą ukrytą, zawierającą 32 komórki LSTM/GRU. Mniejsza wielkość warstwy ukrytej od warstwy wejściowej na ogół prowadziła do większej stabilności podczas procesu uczenia.



Rys. 6.1. Schemat architektury sieci LSTM

Kolejnym elementem modelu jest warstwa Dropout o współczynniku 0.2. Jej działanie polega na losowym wyłączeniu 20% neuronów warstwy ukrytej podczas treningu, w celu zapobiegnięcia nadmiernemu dopasowaniu do danych treningowych. Wyjście sieci zostało zdeterminowane przez rodzaj problemu, zatem składa się ono z warstwy dwóch klasycznych neuronów. Odpowiadają one poleceniom kup oraz sprzedaj zwracanym przez przegotowany model.

Pełna implementacja sieci na przykładzie LSTM została przedstawiona na rysunku 6.2. Można w niej zauważyć kilka dodatkowych informacji takich jak:

- **Funkcja aktywacji** – w warstwie wyjściowej została, zastosowana funkcja Softmax, która jest klasycznym rozwiązaniem w problemach klasyfikacji. Odpowiada ona za przekształcenie wektora liczb rzeczywistych do wektora prawdopodobieństw, który sumuje się do jednego. Oznacza to, że każdy neuron z wyjściem typu Softmax zwraca prawdopodobieństwo przynależności do danej klasy.
- **Funkcja strat** – model wykorzystuje funkcję strat zwaną Categorical Crossentropy, która jest podstawowym wyborem w zadaniach klasyfikacji. Oblicza ona stratę między prawdziwymi etykietami, a odpowiedziami prognozowanymi przez model.
- **Optymalizator** – we wszystkich modelach wykorzystany został optymalizator Adam, który jest odpowiedzialny za proces aktualizacji wag, oparty na metodzie propagacji wstecznej.
- **Batch Size** – określa liczbę próbek jednocześnie podawanych na wejście sieci, po której następuje aktualizacja wag. W trakcie trenowania sieci najlepiej spisywał się wsad o rozmiarze 100 próbek.
- **Liczba epok** – informuje ile razy algorytm przechodzi przez cały zbiór treningowy w trakcie procesu uczenia. Liczba przejść była często zmieniana w trakcie realizacji projektu na rzecz analizy zachowania sieci, jednak zazwyczaj wystarczyło kilka epok do osiągnięcia maksymalnej skuteczności.

```
def LSTM_Classification(input_dim, input_feature_size):
    model = models.Sequential()
    model.add(LSTM(units=64, return_sequences=True,
                    input_shape=(input_dim, input_feature_size)))
    model.add(LSTM(units=32))
    model.add(Dropout(0.2))
    model.add(Dense(units=2, activation='softmax'))

    model.compile(optimizer='adam',
                  loss='categorical_crossentropy',
                  metrics=['accuracy'])

    history = model.fit(X_train, y_train, epochs=N_EPOCH,
                        batch_size=BATCH_SIZE, verbose=2)

    return model, history
```

Rys. 6.2. Kod źródłowy z implementacją sieci LSTM przeznaczonej do klasyfikacji

6.1.2. PREDYKCJA

Dla zachowania jak największej rzetelności porównania obu podejść, modele przeznaczone do predykcji przyszłej ceny posiadają architekturę jak najbardziej zbliżoną do sieci wykorzystywanych w klasyfikacji. Główną różnicą w budowie jest warstwa wyjściowa, która w tym przypadku składa się z pięciu neuronów. Różnica wynika z założeń projektowych, w których przyszła strategia ma bazować na przewidywanych cenach z pięciu kolejnych dni.

```
def LSTM_Prediction(input_dim, input_feature_size):
    model = Sequential()
    model.add(LSTM(units=64, return_sequences=True,
                    input_shape=(input_dim, input_feature_size)))
    model.add(LSTM(units=32))
    model.add(Dropout(0.2))
    model.add(Dense(units=5))

    model.compile(optimizer=Adam(lr = LR), loss="mse")

    history = model.fit(X_train, y_train, epochs=N_EPOCH,
                        batch_size=BATCH_SIZE)

    return model, history
```

Rys. 6.3. Kod źródłowy z implementacją sieci LSTM przeznaczonej do predykcji

Zadanie predykcji wymaga zastosowania również kilku innych parametrów, ze względu na przewidywanie dokładnej ceny. Wobec tego sieć posiada następujące zmiany:

- **Funkcja aktywacji** – wyjście sieci musi przyjmować dowolne wartości, w zależności od przewidywanej ceny, z tego względu wykorzystana została liniowa funkcja aktywacji, będąca domyślnym ustawieniem w bibliotece TensorFlow.
- **Funkcja strat** – podczas predykcji szeregu czasowego wykorzystuje się błąd średniokwadratowy MSE, którego zadaniem jest zmierzenie różnicy między wartościami rzeczywistymi, a przewidywanymi przez model. Dany jest on wzorem:

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

,gdzie:

n – liczba obserwacji,

y_i – wartości rzeczywiste dla i -tej obserwacji,

\bar{y} – wartość przewidywana dla i -tej obserwacji.

Ponadto często wykorzystywaną funkcją przy ocenie predykcji szeregów czasowych jest pierwiastek błędu średniokwadratowego **RMSE**. Jego zaletą jest likwidacja kwadratu powstałego przy obliczaniu MSE, przez co zachowana zostaje ta sama jednostka. Ułatwia to zrozumienie wielkości powstałego błędu, gdyż jego wartość może być bezpośrednio porównana z ceną rzeczywistą.

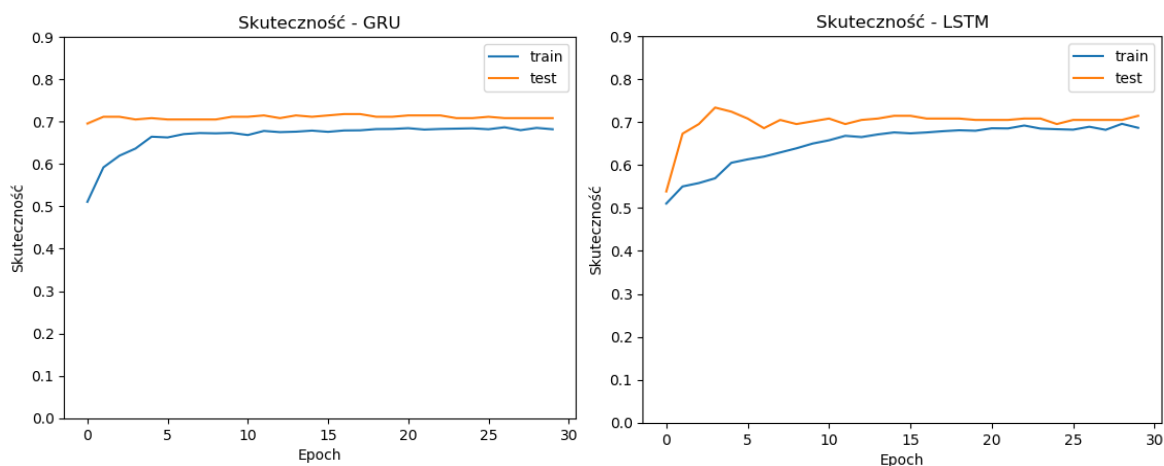
- **Współczynnik uczenia** – to hiperparametr określający tempo uczenia sieci neuronowych. W praktyce aktualizacje wag są wprost proporcjonalne do jego wartości. Współczynnik uczenia optymalizatora Adam jest domyślnie ustawiony na 0,001. Taka wartość nie sprawiała problemów w trakcie klasyfikacji, jednak podczas predykcji, uzyskiwany błąd sieci był często niestabilny i oscylował wokół pewnej wielkości. Z tego względu wartość współczynnika została zmniejszona dwukrotnie na 0,0005.

6.2. ANALIZA WYNIKÓW UCZENIA

W trakcie badań wytrenowane zostały osobne modele dla każdego roku z przedziału 2016 – 2022. Do treningu wykorzystano dane z 10 lat poprzedzających wybrany okres, zgodnie z przyjętymi założeniami. Szczegółowo opisywane wyniki pochodzą z modelu dla 2021 roku, zarówno dla klasyfikacji jak i predykcji, a długość zastosowanego okna czasowego wynosiła 20 dni. W ocenie efektów uczenia wzięto pod uwagę następujące kryteria:

- Skuteczność
- Najmniejsza wartość błędu
- Stabilność
- Powtarzalność
- Prędkość uczenia

Na rysunku 6.4 znajduje się przykładowy przebieg uczenia sieci klasyfikującej, który składa się ze skuteczności osiągniętej na zbiorze treningowym i testowym. Model GRU znajdujący się po lewej stronie, charakteryzuje się wysoką stabilnością i bardzo szybkim tempem nauki. Skuteczność na zbiorze testowym osiąga maksymalny poziom 72% już po dwóch epokach i stabilnie utrzymuje się w raz z kolejnymi przejściami. Z kolei zbiór treningowy wymagał nieznacznie dłuższego czasu na osiągnięcie maksymalnej skuteczności, gdyż stabilizowała się ona po około 5 epokach. Ponadto można zauważyć powtarzającą się zależność niższej skuteczności na zbiorze treningowym, która może być spowodowana znacznie większą ilością danych. W przypadku modeli z innych lat, sieć GRU uzyskiwała rezultaty z przedziału 63% – 74% i z reguły zapewniała wysoką powtarzalność. Przy opisywanej architekturze nie zaobserwowano efektów przeuczenia lub oscylacji.

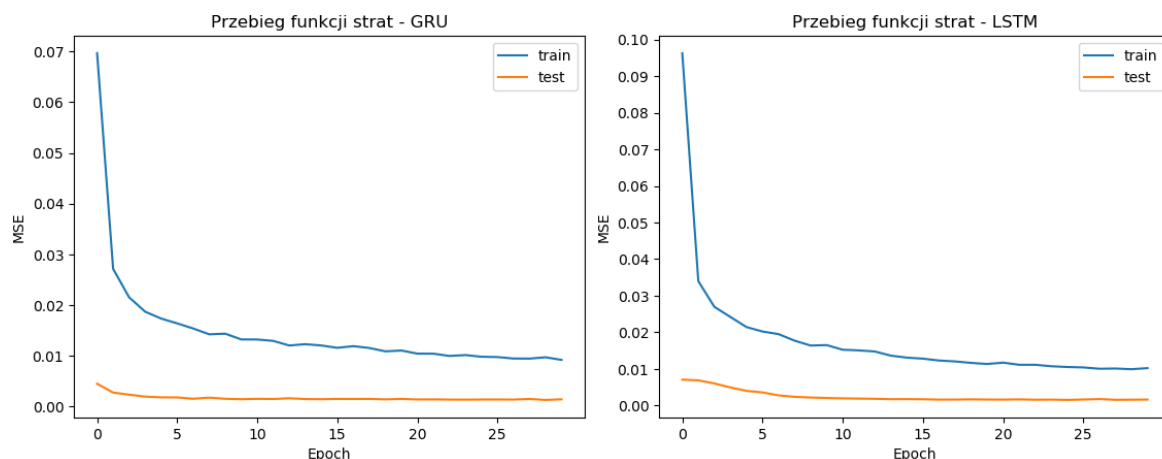


Rys. 6.4. Przebieg uczenia modeli GRU i LSTM – klasyfikacja

Przebieg uczenia sieci LSTM widoczny po prawej stronie rysunku 6.4 znacząco różni się od swojego poprzednika. Na wykresie możemy zauważyć, że skuteczność na zbiorze testowym jest zdecydowanie mniej stabilna. W początkowym etapie obserwujemy szybki wzrost, który kończy się szczytem na poziomie 74%. Po następnych kilku epokach skuteczność stabilizuje się i przewyższa tą osiągniętą przez GRU o około 1 punkt procentowy. Opanowanie zbioru treningowego w przypadku sieci LSTM trwa znacznie dłużej niż w modelu GRU. Skuteczność stale rośnie przez około 20 epok, po czym zatrzymuje się na poziomie tylko minimalnie niższym, niż ten uzyskany na zbiorze testowym. Wydłużenie czasu jest spowodowane dodatkowym zadaniem dobrania odpowiednich wag dla pamięci długoterminowej, której sieć GRU nie posiada.

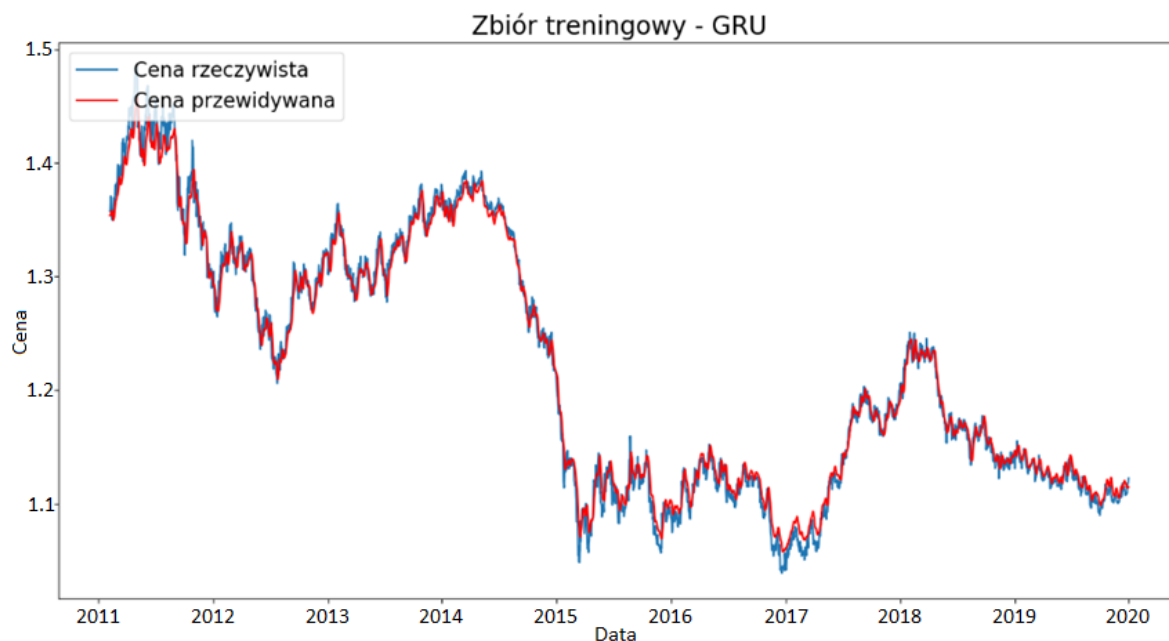
Podsumowując porównanie sieci LSTM i GRU w zadaniu klasyfikacji szeregu czasowego, można stwierdzić że pierwsza z nich jest w stanie osiągnąć nieco wyższą skuteczność zarówno dla zbioru treningowego jak i testowego, natomiast druga cechuje się większą stabilnością i szybszym czasem uczenia.

Przechodząc do porównania sieci zaprojektowanych do predykcji ceny, które zostały wytrenowane na tych samych danych, analizowana będzie wartość funkcji strat, którą stanowi błąd średniokwadratowy MSE. Jego przebiegi w trakcie procesu uczenia zostały przedstawione na rysunku 6.5. W tym przypadku okazuje się, że różnice pomiędzy dwoma rodzajami modeli są znacznie mniej zauważalne. Obie sieci zachowywały podobną stabilność i brak oscylacji, które zdarzały się w początkowej fazie badań, podczas sprawdzania innych architektur sieci (szczególnie dla LSTM). Można zauważyć, że ponownie to sieć LSTM potrzebowała o kilka epok więcej na osiągnięcie najniższego błędu. Po raz kolejny lepszy efekt został osiągnięty na zbiorze testowym, a korzyść na rzecz sieci z dwoma rodzajami pamięci była marginalna.



Rys. 6.5. Przebieg uczenia modeli GRU i LSTM – predykcja

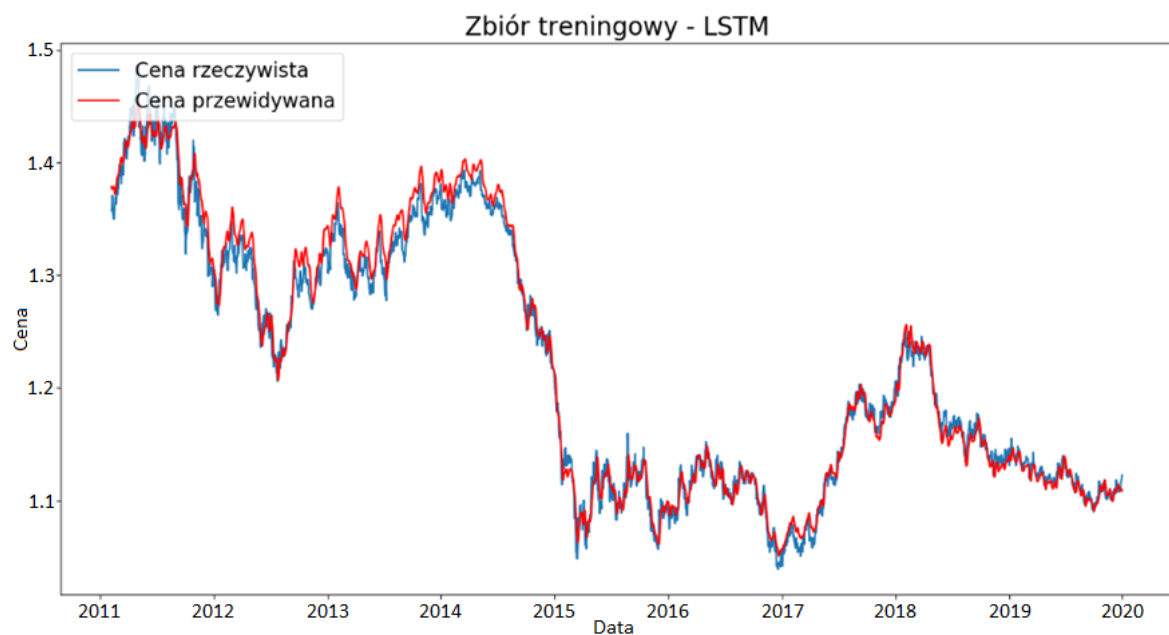
Dodatkowo różnice pomiędzy analizowanymi modelami można porównać poprzez bezpośrednie zestawienie przebiegu rzeczywistej ceny EUR/USD z przewidywaniami wygenerowanymi przez sieci. W celu zachowania czytelności zostały one rozdzielone na dwa osobne rysunki – 6.6 i 6.7. Na pierwszym z nich obserwujemy efekt uczenia sieci GRU. Przewidywana cena bardzo dobrze dopasowuje się do danych uczących w całym przedziale czasu. Minimalne rozbieżności są widocznie głównie w najbardziej ekstremalnych miejscach, gdzie dochodzi do przełamania trendu. Pozostała część wykresu prezentuje bardzo dobre efekty.



Rys. 6.6. Przebieg ceny przewidywanej przez model GRU – zbiór treningowy

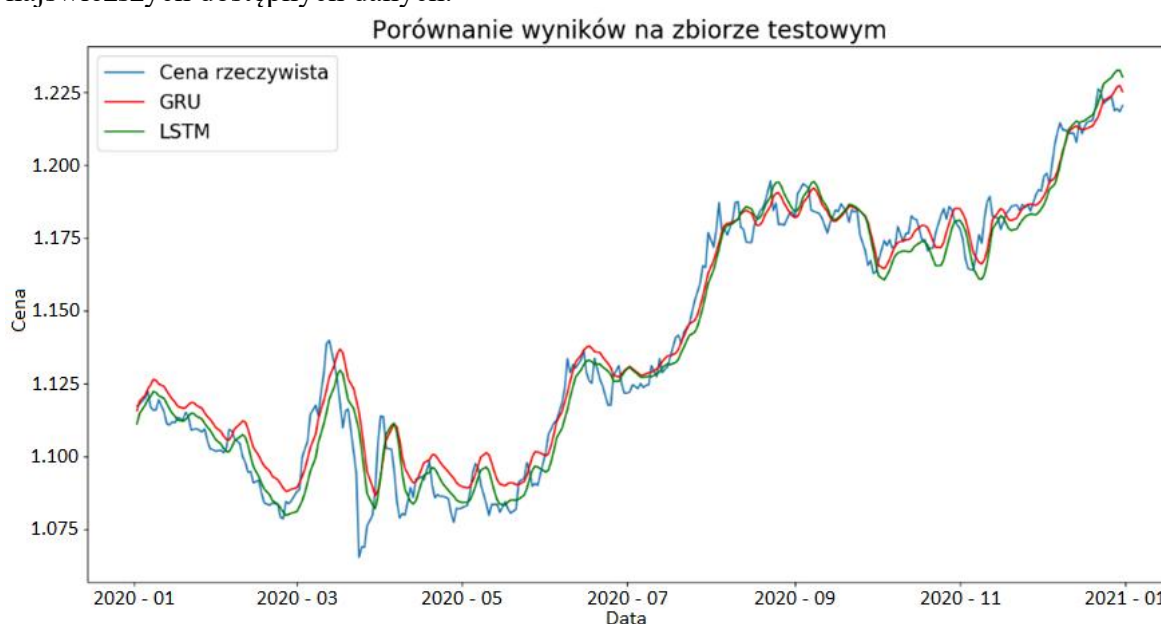
Z kolei w przypadku modelu LSTM, możemy zaobserwować większe rozbieżności w pierwszej części zbioru treningowego. Do roku 2014 sieć ma tendencję do zawyżania oczekiwanej ceny. Po tym czasie sytuacja odwraca się i to właśnie sieć LSTM może pochwalić się lepszym dopasowaniem. Nawet w miejscach nagłych zmian trendu, przebieg ceny przewidywanej pokrywa się z rzeczywistymi wartościami.

Rozbieżność w pierwszej części wykresu może być spowodowana dostosowaniem wag dla pamięci długoterminowej w komórkach LSTM. W trakcie procesu uczenia dane podawana są chronologicznie, w sposób zgodny z przypisaną im datą. Jest to niezbędny element uczenia sieci rekurencyjnych, dzięki czemu mogą one prawidłowo korzystać ze swoich stanów ukrytych. W zaistniałej sytuacji błąd modelu LSTM na zbiorze treningowym może być zawyżony.



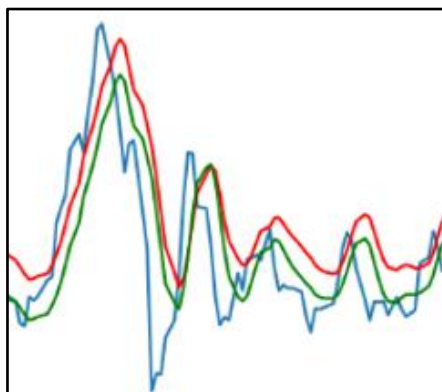
Rys. 6.7. Przebieg ceny przewidywanej przez model LSTM – zbiór treningowy

Na rysunku 6.8 przedstawione zostało porównanie ceny rzeczywistej oraz przewidywań modeli GRU i LSTM na zbiorze testowym, którym jest rok 2021. Należy pamiętać, że stosunkowo dobre dopasowanie jest efektem generowania ceny dla każdego kolejnego dnia w formie odpowiedzi na podanie okna czasowego z pełną informacją na temat 20 poprzedzających go dni. Modele wiedzą więc dokładnie na jakiej cenie zakończył się ostatni zanotowany dzień. Taki sposób symulacji nie jest w stanie wygenerować danych sięgających w dalszą przyszłość. Byłoby to możliwe przy uzyskiwaniu kolejnych odpowiedzi poprzez bazowanie na uprzednio obliczonych przewidywaniach. Wygenerowałyby to jednak dużo większe rozbieżności, które nie są potrzebne przy tym rodzaju zadania. Tworząc system transakcyjny, chcemy uzyskać informacje, czy w danym momencie występuje dogodny czas na kupno lub sprzedaż. W tej sytuacji zawsze możliwe jest podanie na wejście sieci najświeższych dostępnych danych.



Rys. 6.8. Przebieg uczenia modeli GRU i LSTM – predykcja

Przechodząc do dalszej analizy wykresu, obserwujemy, że mniejsza skala uwypukla główny problem obu testowanych modeli. Przy większości zmian trendów sieć zaczyna zauważać zaistniałą sytuację z lekkim opóźnieniem. Dopiero po jednym lub kilku dniach przebiegi zaczynają podążać w stronę zgodną z nowym kierunkiem zmian ceny. W ten sposób sprawiają wrażenie powtarzania danych wejściowych z przesunięciem w czasie. Efekt jest dobrze widoczny na rysunku 6.9, gdzie w większym przybliżeniu przedstawiono wybrany fragment omawianego wykresu.



Rys. 6.9. Porównanie przebiegów cen w przybliżeniu

6.3. WYBÓR NAJLEPSZEGO MODELU

Trenując modele na 7 zbiorach danych, które odpowiadają wszystkim testowanym okresom, nie zawsze ten sam rodzaj sieci lub ta sama długość okna czasowego okazała się najlepsza. Błędem byłoby jednak zdecydowanie się na wybór różnych architektur dla poszczególnych lat, gdyż w rzeczywistości nie jesteśmy w stanie stwierdzić, który z modeli będzie najskuteczniejszy w danym momencie. Z tego względu należy wybrać sieć, która sumarycznie poradziła sobie najlepiej na wszystkich testowanych zbiorach i w przyszłości posłuży do zaprojektowania strategii systemu decyzyjnego.

W tabeli 6.1 przedstawione zostały efekty uczenia dla wszystkich modeli służących do klasyfikacji pary walutowej EUR/USD. Dla każdego testowanego roku przedstawiono osiągnięty poziom skuteczności przez sieć LSTM oraz GRU przy wykorzystaniu okna czasowego o długości 10, 20 i 30 dni. W efekcie uzyskano bezpośrednie zestawienie 6 wartości, spośród których najwyższa z nich została zaznaczona kolorem zielonym.

Na podstawie poniższych wyników można zaobserwować, że tylko jeden raz udało się uzyskać wyższą skuteczność przy wykorzystaniu sieci GRU. Taka sytuacja ma miejsce w roku 2019 i tym samym przewyższa swojego konkurenta o niespełna 1 punkt procentowy. W pozostałych przypadkach to sieć LSTM popełniała mniej błędów, co potwierdza wcześniejsze spostrzeżenia.

Z kolei wpływ długości okna na uzyskiwane wyniki nie był aż tak różnicujący jak rodzaj wybranej sieci. Dwukrotnie najskuteczniejsze okazały się przedziały 10 lub 20 dniowe, lecz ich przewaga była bardzo niewielka. Okno 30 dniowe wygrało trzykrotnie, przy czym w roku 2020 i 2021 przewyższyło pozostałe wyniki o 2 - 3 punkty procentowe.

Tab. 6.1. Zestawienie wyników modeli klasyfikujących.

	Skuteczność [%]					
	10 Dni		20 Dni		30 Dni	
	LSTM	GRU	LSTM	GRU	LSTM	GRU
2022	67.23	66.22	64.15	62.85	66.55	63.18
2021	71.08	70.76	72.72	72.08	74.54	73.72
2020	70.1	70.74	72.03	71.38	74.92	73.95
2019	63.78	64.74	63.78	62.82	62.82	62.82
2018	71.06	70.42	68.49	68.17	71.84	70.1
2017	70.01	69.01	71.86	70.94	71.34	71.66
2016	66.45	65.81	67.1	64.42	65.48	66.77
Średnia	68.82	68.24	68.73	67.52	69.64	68.89

Ostatecznie jako najlepszy model, który uzyskał największą średnią skuteczność na przestrzeni wszystkich zbiorów testowych, wybrana zostaje sieć z warstwami komórek LSTM przy wykorzystaniu okna czasowego o długości 30 dni. Nieznacznie większe trudności w trenowaniu tej architektury są zrekompensowane najwyższą skutecznością, dzięki czemu ma ona potencjał na zapewnienie najkorzystniejszych decyzji inwestycyjnych.

Analogiczne zestawienie wyników uzyskanych dla modeli zajmujących się predykcją cen zostało przedstawione w tabeli 6.2. Tym razem porównana zostanie wartość pierwiastka błędu średniokwadratowego RMSE, którą obliczono dla wszystkich zbiorów testowych. Opisywany błąd określa o jaką średnią wartość różnią się ceny przewidywane przez sieć w porównaniu do całego zbioru testowego. W tym przypadku RMSE przedstawia różnicę ceny za jedno euro wyrażonej w dolarach.

Uzyskane pierwiastki błędu średniokwadratowego dla wszystkich zbiorów i modeli mieszczą się w przedziale 0.03 – 0.055, co oznacza zakres od 3 do ponad 5 centów. Tym razem aż w 4 testowanych latach lepszy wynik uzyskała architektura oparta o komórki GRU, zaś LSTM przewyższało ją w pozostałych trzech przypadkach. Ponadto 30 dniowe okno czasowe dawało realne korzyści, szczególnie w porównaniu z najkrótszym, 10 dniowym przedziałem.

Tab. 6.2. Zestawienie wyników modeli predykcji

	RMSE [\$]					
	10 Dni		20 Dni		30 Dni	
	LSTM	GRU	LSTM	GRU	LSTM	GRU
2022	0.049	0.046	0.055	0.054	0.049	0.053
2021	0.044	0.046	0.043	0.041	0.038	0.037
2020	0.038	0.039	0.035	0.035	0.033	0.035
2019	0.055	0.054	0.053	0.049	0.051	0.052
2018	0.042	0.044	0.041	0.039	0.035	0.038
2017	0.044	0.043	0.038	0.040	0.039	0.041
2016	0.053	0.055	0.050	0.049	0.045	0.042
Średnia	0.046	0.047	0.045	0.044	0.041	0.043

Spoglądając na uśrednione wyniki ze wszystkich zbiorów, okazuje się że najmniejsze odchylenie od cen rzeczywistych uzyskał model LSTM z 30 dniowym oknem czasowym. Taka sama sytuacja miała miejsce również w przypadku klasyfikacji.

Bazując na przeprowadzonych badaniach, w dalszej części projektu wykorzystany zostanie model sieci LSTM z 64 komórkami w warstwie wejściowej, 32 w warstwie ukrytej i dropoutem na poziomie 0.2. Dodatkowo będzie on trenowany na danych wydzielonych przez 30 dniowe okno czasowe.

7. IMPLEMENTACJA SYSTEMÓW TRANSAKCYJNYCH

Dane uzyskiwane z modeli głębokich sieci neuronowych stanowią jedynie zaplecze wiedzy, którą należy umiejętnie wykorzystać podczas implementacji systemu decyzyjnego do gry na giełdzie. Za wygenerowanie odpowiedniego sygnału transakcyjnego bezpośrednio odpowiada przygotowana strategia inwestycyjna. Jej głównym celem jest zarządzanie portfelem inwestora, poprzez skupowanie i sprzedaż instrumentów finansowych we właściwym momencie.

W świecie giełdowym, istnieje wiele rodzajów strategii inwestycyjnych. Jedne z nich bazują na wiedzy jaką niesie analiza fundamentalna, inne korzystają z technik analizy technicznej, zaś niektóre bazują wyłącznie na sztucznej inteligencji. Zadaniem do wykonania w projekcie jest znalezienie podejścia, które w najskuteczniejszy sposób wykorzysta przygotowane modele neuronowe, tak by zmaksymalizować zysk inwestora.



Rys. 7.1. Schemat sygnałów transakcyjnych [34]

Zaawansowane strategie giełdowe oprócz wysyłania sygnału transakcyjnego w odpowiednim momencie, regulują również stawkę za jaką wchodzi do gry lub ilość akcji jakie sprzedają. Dla uproszczenia symulacji i jak największym skupieniu się na zbadaniu przydatności sieci neuronowych podczas gry na giełdzie, przygotowane strategie nie będą zarządzać wielkością otwieranych pozycji. Po każdorazowym otrzymaniu sygnału transakcyjnego, cała zawartość portfela zostanie wydana lub wszystkie posiadane aktywa będą sprzedane.

7.1. STRATEGIE OPARTE NA KLASYFIKACJI

7.1.1. POJEDYNCZY SYGNAŁ

Strategie oparte na klasyfikacji trendów pary walutowej EUR/USD nie są w stanie wnieść szczególnie rozbudowanych taktyk, gdyż sama sieć została zaprojektowana w taki sposób, by rozpoznawać i przewidywać panujące trendy. Z tego względu pierwsza strategia będzie zarządzać wysyłanymi sygnałami zgodnie z przewidywaniami modelu. Transakcja dojdzie do skutku jeśli w portfelu inwestora znajdują się jakiekolwiek środki, a sieć da sygnał kupna. Jeśli następne sygnały również będą wskazywały na dobry moment do zainwestowania pieniędzy, strategia będzie czekać. Dopiero po pojawieniu się pierwszego sygnału informującego, że nadszedł czas na sprzedaż, wszystkie posiadane aktywa zostaną sprzedane, a pieniądze wrócą do portfela, aż do czasu następnego kupna.

W trakcie badań regulowana będzie również pewność modelu co do danej decyzji. Transakcje dojdą do skutku jeśli prawdopodobieństwo prawdziwości wygenerowanej informacji na wyjściu sieci jest większe od 70 procent. W ten sposób ograniczone zostanie ryzyko związane z niepoprawną interpretacją aktualnej sytuacji rynkowej. Modele nie są także w stanie wykrywać trendów bocznych, przez co w takich sytuacjach mogą generować sygnały z niską pewnością. Ustawienie 70 procentowego progu ma na celu wykluczenie transakcji bazujących na tego rodzaju sygnałach.

7.1.2. TRZY SYGNAŁY

Strategia trzech sygnałów opiera się na podobnej metodzie skupowania i sprzedaży aktywów. Główną różnicą w porównaniu do „pojedynczego sygnału” jest wykorzystanie trzech kolejnych informacji od modelu sieci neuronowej. Strategia wysyła sygnał kupna lub sprzedaży w momencie gdy trzy kolejne informacje sugerują wykonanie tego samego ruchu.

Wykorzystanie większej liczby dni niesie za sobą pewne konsekwencje. Po pierwsze podnosi to pewność podjętej decyzji, przez co ogranicza ryzyko utraty kapitału. Mniejsze jest prawdopodobieństwo wykonania nietrafnej transakcji, gdy w trzy dni z rzędu przygotowany model ocenił sytuację rynkową w ten sam sposób. W teorii zwiększa się szansa na uniknięcie błędów popełnianych przez sieć neuronową.

Z drugiej strony czas reakcji na zmianę trendu zostaje znacznie zwiększony. Nawet jeśli sieć precyzyjnie wykryje odwrócenie kierunku zmian ceny, potrzebne są co najmniej kolejne dwa dni zanim transakcja dojdzie do skutku. Może to prowadzić do zbyt długiego czasu oczekiwania i przegapienia odpowiedniego momentu na kupno lub sprzedaż. Jeśli w szybkim czasie trend odwróci się ponownie, portfel inwestora uszczupli się.

7.2. STRATEGIE OPARTE NA PREDYKCJI

W trakcie prac nad strategiami bazującymi na prognozowanej cenie pary EUR/USD, sygnały transakcyjne nie są otrzymywane wprost od modelu sieci neuronowej. W związku z tym niezbędne jest opracowanie zasady, która zdefiniuje momenty wysłania odpowiedniego sygnału transakcyjnego.

7.2.1. WZROST WZGLĘDNY

Pierwsza strategia opiera się na porównaniu średniej wartości prognozowanych cen z poprzedzającą ją ceną zanotowaną na giełdzie. Gdy model dostarcza swoje predykcje w postaci przewidywań kursu w następnych 5 dniach, obliczana jest ich średnia arytmetyczna. Wówczas można ją porównać z ostatnią, rzeczywistą, zanotowaną ceną. Jeśli uzyskana średnia jest od niej większa, strategia ma zadanie dokonać transakcji kupna. W przeciwnym razie posiadana waluta zostaje sprzedana.

Kluczowym elementem tego rozwiązania jest zastosowanie odpowiedniego progu, który wyznacza o jaką kwotę średnia cena musi być wyższa lub niższa by transakcja doszła do skutku. Jego brak lub zbyt niska wartość doprowadzi do ciągłego skupowania i sprzedawania, nawet w sytuacjach drobnych korekt ceny na rynku. Grozi to dokonaniem nietrafnej decyzji, która może doprowadzić do straty pieniędzy. Z drugiej strony zbyt wysoki próg spowoduje brak jakichkolwiek transakcji lub dokonywanie ich ze zbyt dużym opóźnieniem. Aby uniknąć obu opisanych przypadków należy wyznaczyć jego najkorzystniejszą wartość w trakcie badań. Początkowo dobrany zostanie próg w postaci 1 procenta ostatniej ceny na zamknięciu, co zazwyczaj nieznacznie przekroczy 1 cent.

7.2.2. WZROST BEZWZGLĘDNY

Ostatnia strategia opiera się na nieco innej metodzie wykorzystania otrzymywanych prognoz przyszłej ceny pary EUR/USD. Tym razem uzyskane wartości nie będą odnosiły się do ostatniej, znanej ceny na zamknięciu. W zamian za to strategia sprawdzi, jaki kierunek zmian obierze kurs w obrębie pięciu kolejnych dni. Jeśli prognozowane wartości będą rosły, wysłany zostanie sygnał kupna, zaś w sytuacji odwrotnej – sygnał sprzedaży.

W odróżnieniu od swojego poprzednika, strategia nie wykorzystuje odniesienia do historycznej ceny w trakcie generowania sygnałów transakcyjnych. W ten sposób ograniczany jest efekt niedokładności modelu przy obliczaniu przyszłego kursu. Wystarczy, że prognozowany trend okaże się trafny, by wykonana transakcja zakończyła się sukcesem.

W trakcie symulacji, obliczane będą dwie średnie arytmetyczne. Jedna z pierwszych dwóch dni, a następna z pozostałych trzech. Jeśli wartość drugiej średniej przewyższy pierwszą, przekazana zostanie informacja o możliwym kupnie instrumentu. Gdy z biegiem czasu wartości zaczną stawać się coraz niższe, nastąpi sprzedaż. Podobnie jak w przypadku wzrostu względnego, strategia będzie posiadała próg o jaki dwie średnie muszą się różnić by doszło do transakcji. Z racji, że różnice wartości w obrębie pięciu prognozowanych dni są znacznie mniejsze, zostanie on ustawiony na 0.002 dolara.

8. SYMULACJA SYSTEMÓW NA DANYCH HISTORYCZNYCH

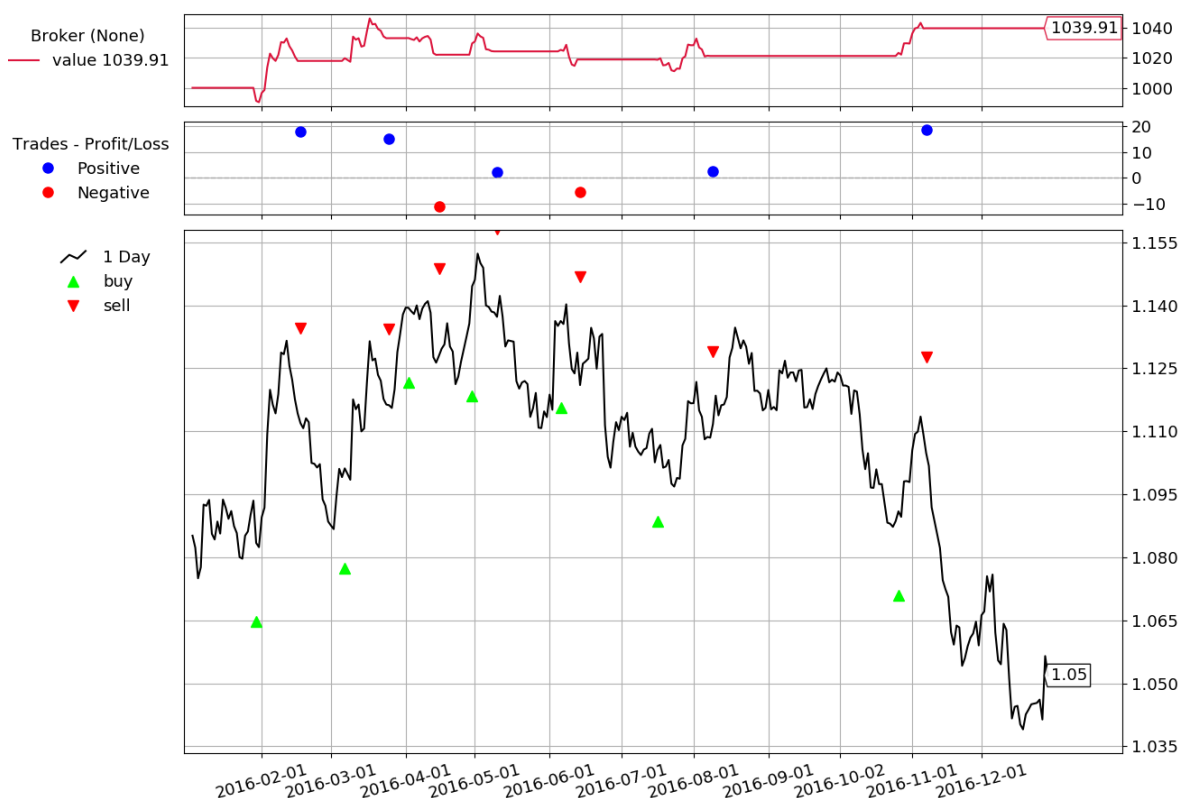
Po przygotowaniu wszystkich danych, wytrenowaniu modeli sieci neuronowych oraz zaimplementowaniu strategii giełdowych, możliwe jest przeprowadzenie symulacji sprawdzających realny potencjał zarobkowy przygotowanego systemu. W tym celu wykorzystana zostanie otwartoźródłowa biblioteka Pythona – BackTrader. Umożliwia ona tworzenie i testowanie strategii handlowych, poprzez wykonywanie wirtualnych transakcji bazujących na otrzymywanych sygnałach transakcyjnych.

Na początku każdej symulacji wartość środków przeznaczonych na grę będzie wynosiła 1000 dolarów. Wraz z upływem czasu wykonywane będą kolejne transakcje, a rentowność każdej z nich zostanie zarejestrowana. Jednocześnie monitorowana będzie całkowita wartość portfela. Na podstawie takiego zapisu istnieje możliwość przeanalizowania zachowania danej strategii, a końcowy stan konta pokaże, czy zakończony rok przyniósł oczekiwany zysk.

W kolejnych punktach opisane zostaną wybrane okresy dla poszczególnych strategii. Przedstawienie i porównanie wyników dla wszystkich przeprowadzonych symulacji znajdzie się w następnym rozdziale.

8.1. POJEDYNCZY SYGNAŁ

Na rysunku 8.1 przedstawiony został wykres z wykonaną symulacją na danych z roku 2016 przy wykorzystaniu strategii „pojedynczy sygnał”. W dolnym segmencie wykresu znajduje się przebieg ceny ze wszystkimi transakcjami zaznaczonymi za pomocą trójkątów. Zielony wskaźnik oznacza kupno, zaś czerwony – sprzedaż.



Rys. 8.1. Strategia „pojedynczy sygnał” – rok 2016

Rentowność wszystkich inwestycji została zaznaczona na środkowym segmencie wykresu. Po każdej sprzedaży obliczana jest różnica pomiędzy wartością waluty w momencie jej spieniężenia oraz kupna. Wszystkie transakcje zakończone zyskiem przedstawiono w postaci niebieskiej kropki, a straty przyjęły kolor czerwony. Dodatkowo można odczytać jaki przychód wygenerowała każda z nich. Z kolei najwyższy segment wykresu stale monitoruje wartość portfela inwestora.

Jak możemy zauważyć strategia nie najgorzej poradziła sobie w trakcie gry w roku 2016. Tylko dwie z siedmiu dokonanych transakcji zakończyły się niewielką stratą. W przedstawionym okresie kurs przeważnie oscylował wokół stałej stawki, a jedynie w jego końcowej części zaobserwowano większy spadek. Pomimo to strategia zakończyła rok z około 4 procentowym zyskiem.

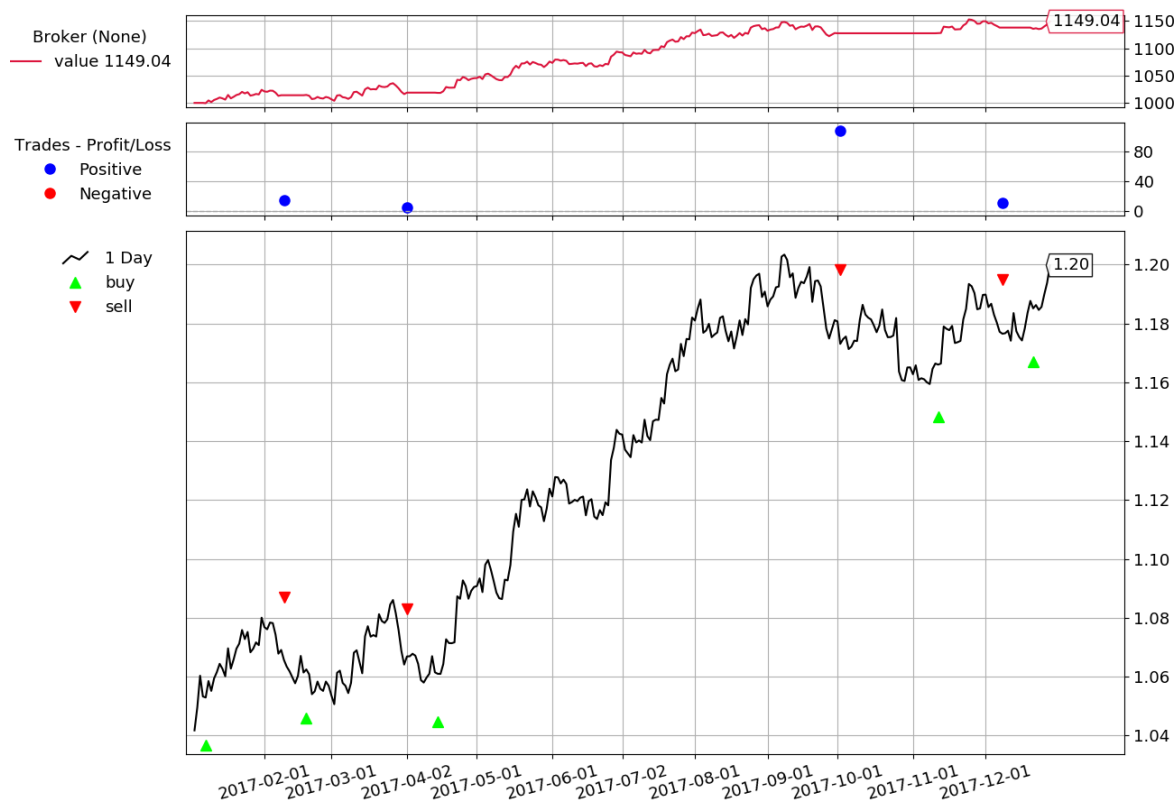
Większość wystawionych sygnałów transakcyjnych jest słuszna. Zawsze gdy cena zaczyna wyraźnie rosnąć, obserwujemy kupno waluty, zaś przy jej spadku – sprzedaż. Można jednak zauważyć, że zysk byłby znacznie większy, gdyby nie delikatne opóźnienie w podejmowaniu wszystkich decyzji. Jest to szczególnie widoczne na przybliżonym fragmencie przedstawionym na rysunku 8.2. Już pierwsza transakcja sprzedaży omija wierzchołek zmiany trendu. Wobec tego sprzedaż odbywa się po potencjalnie niższym kursie. Z kolei druga transakcja kupna całkowicie przegapia nagły wzrost ceny, przez co sprzedaż zakończyła się stratą.



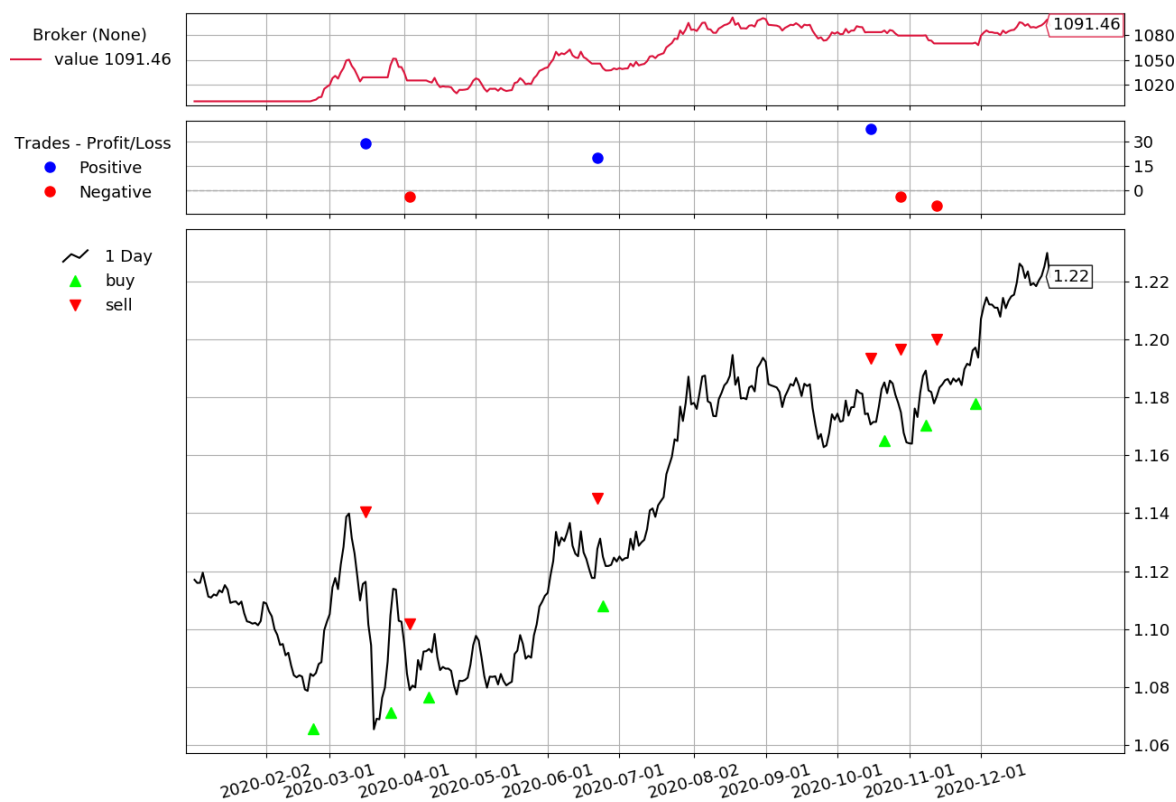
Rys. 8.2. Fragment symulacji z wykorzystaniem strategii „pojedynczy sygnał” – rok 2016

Mimo wszystko wynik strategii w 2016 roku można uznać za mały sukces. Podejmowanie głównie prawidłowych decyzji oraz końcowy zysk z symulacji w roku z niewielką tendencją do spadku ceny z pewnością mieści się w oczekiwaniach.

Znacznie większy zysk odnotowano w latach 2017 i 2020 zaprezentowanych na rysunku 8.3 i 8.4. Oba okresy są zestawione obok siebie ze względu na tą samą tendencję do długotrwałego wzrostu wartości kursu EUR/USD. W pierwszym przypadku strategia wykonała niewiele transakcji, jednak wszystkie z nich zakończyły się sukcesem. Warty uwagi jest fakt precyzyjnego wykrycia kilkumiesięcznego trendu wzrostowego, na którym udało się zarobić ponad 10 procent zainwestowanych środków. Ze względu na wysoki wzrost ceny, wykorzystanie strategii w 2017 roku nie przynosi zysku wyższego od klasycznego ulokowania pieniędzy w parę EUR/USD, jednak dużym plusem jest odnalezienie panującego trendu i niezmarowanie dobrej okazji na zarobek.



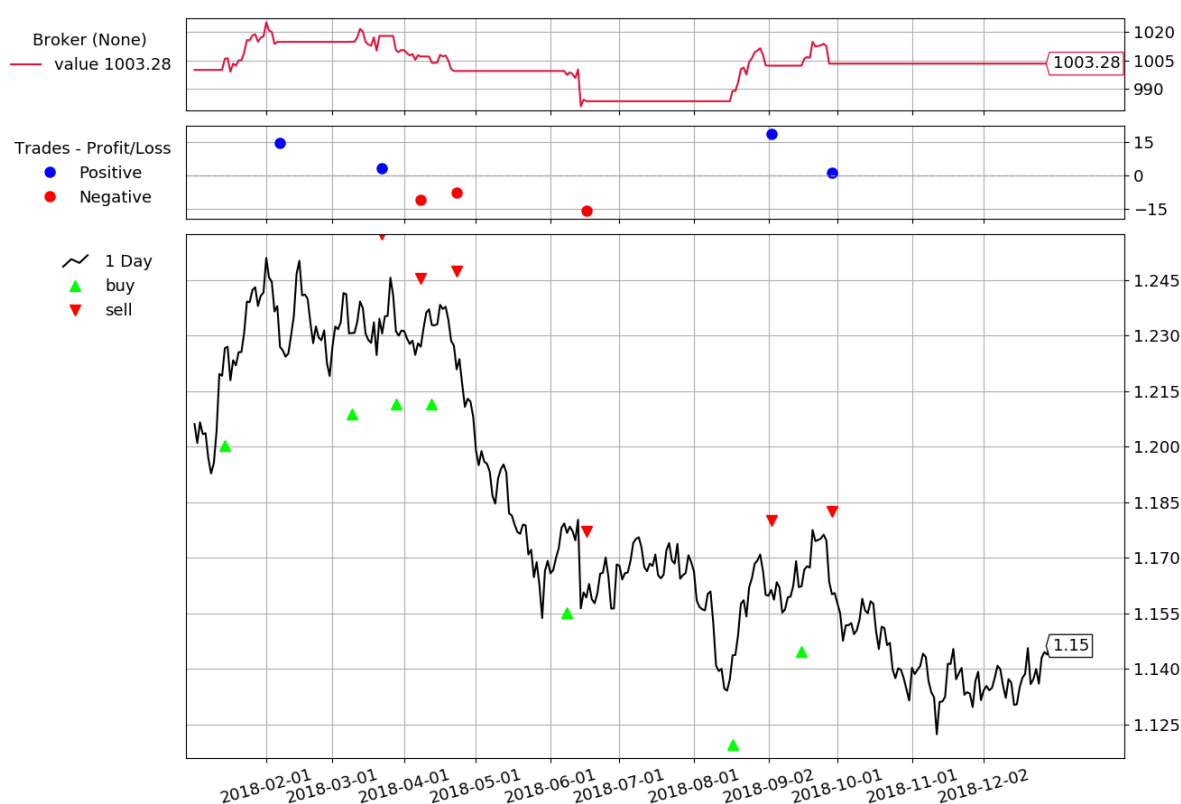
Rys. 8.3. Strategia „pojedynczy sygnał” – rok 2017



Rys. 8.4. Strategia „pojedynczy sygnał” – rok 2020

W przypadku roku 2020 odnotowano przychód niższy o ponad połowę. Strategia kilkakrotnie wygenerowała błędne sygnały zakończone minimalną stratą, jednak tak samo jak poprzednio, odnaleziono zostały kluczowe trendy. W efekcie portfel inwestora ponownie powiększył swoją wartość.

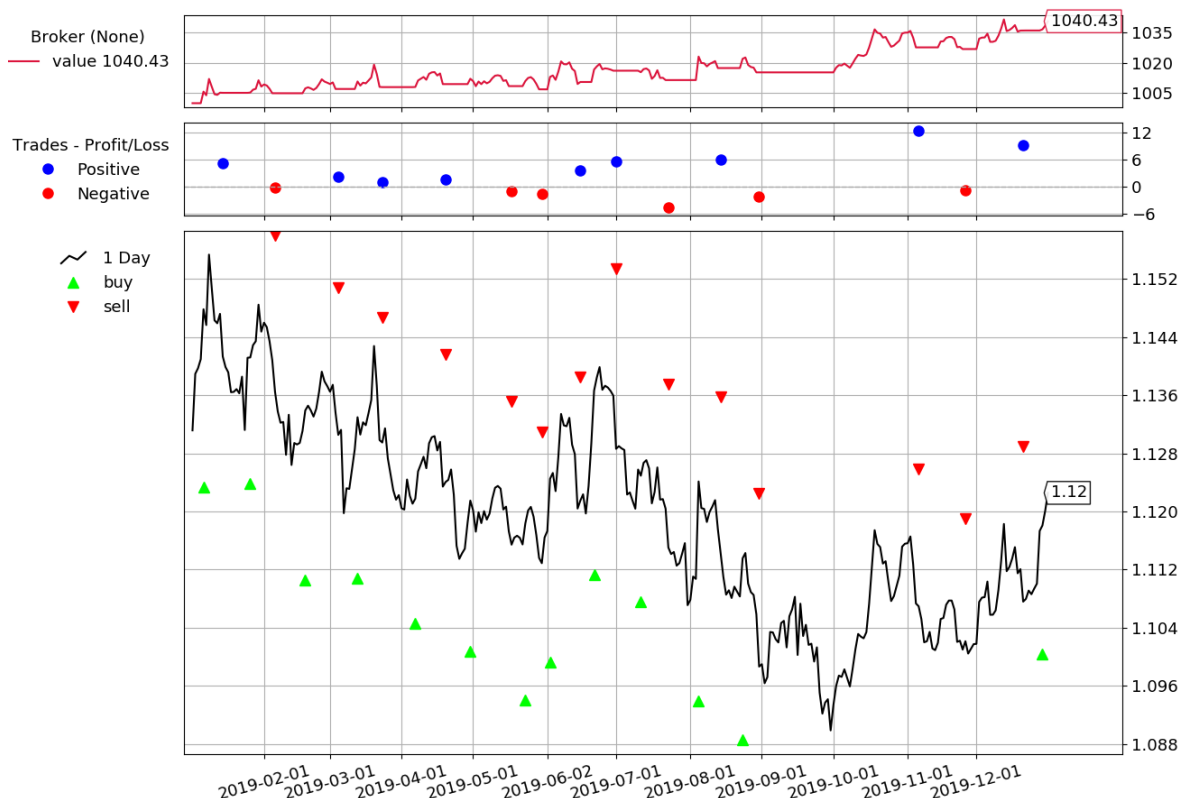
Przeciwieństwem do lat 2017 i 2020 jest sytuacja w roku 2018 przedstawiona na rysunku 8.5. Na wykresie widać znaczny spadek wartości euro do dolara na przestrzeni całego roku. Istnieje jedynie kilka momentów pozwalających na niewielki zysk, które strategia wykorzystwała prawidłowo. Głównym sukcesem w przedstawionym okresie jest uniknięcie wysokich strat na największych trendach spadkowych. Poza kilkoma nieudanymi decyzjami, które w większym stopniu nie nadszarpnęły portfela inwestora, strategia potrafiła przeczekać najmniej korzystne momenty. W ten sposób rok udało zakończyć się mniej więcej na wejściowym poziomie pieniężnym, co można uznać za względny sukces, biorąc pod uwagę bardzo niski potencjał zarobkowym w tym czasie.



Rys. 8.5. Strategia „pojedynczy sygnał” – rok 2018

Ciekawą sytuację przedstawia zapis symulacji z roku 2019 widoczny na rysunku 8.6. W tym okresie istnieje wiele dogodnych momentów na zainwestowanie środków. Można zaobserwować naprzemienną zmianę trendów ze wzrostowego na spadkowy. W takiej sytuacji wymagana jest dobra strategia, która precyzyjnie zareaguje na szybkie zmiany cen.

Wykorzystanie pojedynczego sygnału połowicznie poradziło sobie z tym zadaniem. Rok został zakończony 4 procentowym zyskiem oraz przewagą pozytywnych transakcji. Szybkie zmiany trendów spowodowały jednak niepełne wykorzystanie potencjału zarobkowego. Tak jak poprzednio, większość transakcji jest wykonywana z opóźnieniem, przez co odnotowano tylko niewielki zysk. Gdyby transakcje zostały przeprowadzone o 2 – 3 dni wcześniej, strategia mogłaby zarobić nawet trzykrotnie więcej. Niemniej jednak wygenerowane sygnały mieszczą się w przedziałach, które pozwoliły na końcowy zysk z inwestycji.



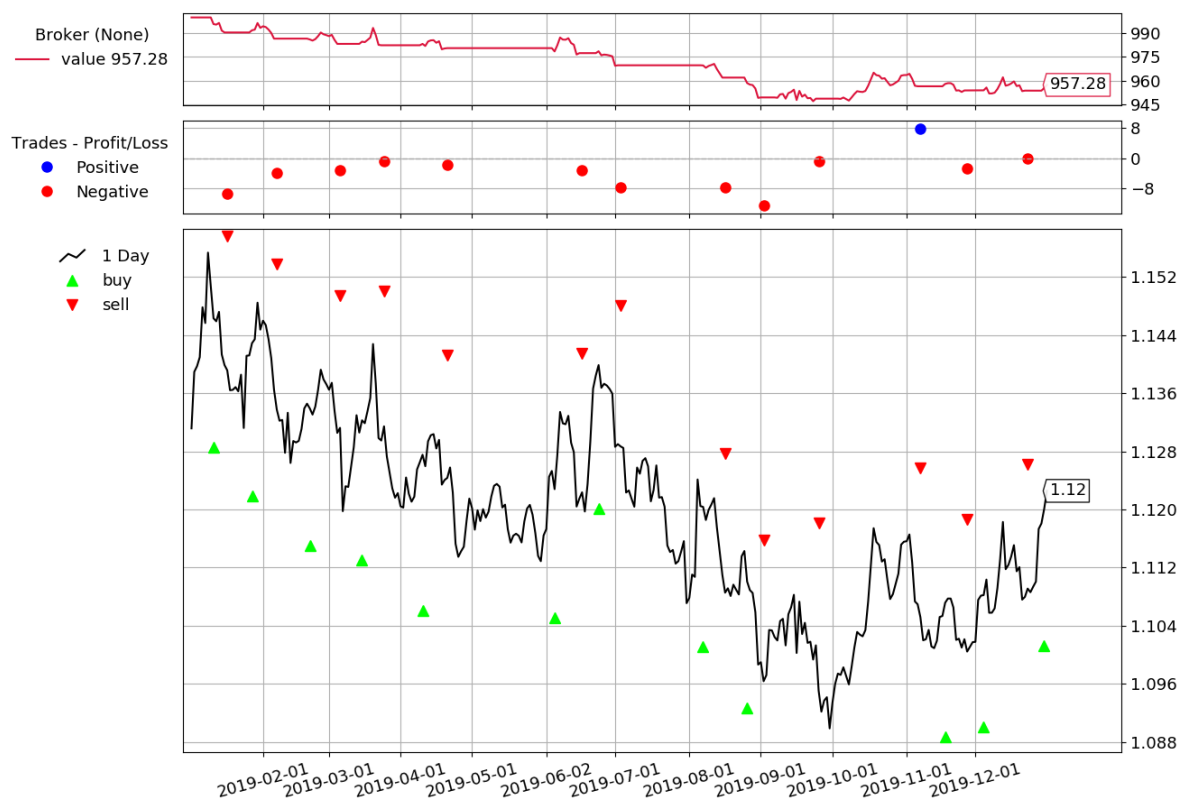
Rys. 8.6. Strategia „pojedynczy sygnał” – rok 2019

8.2. TRZY SYGNAŁY

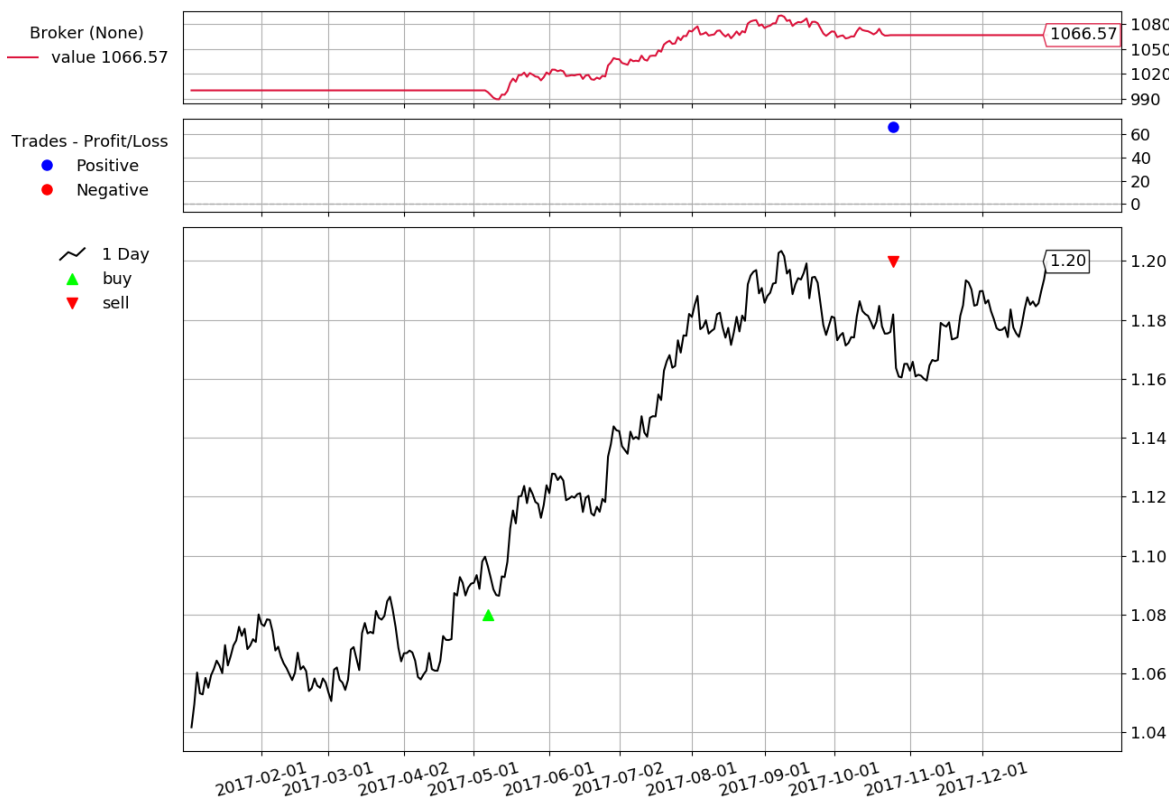
W trakcie symulacji gry na giełdzie znacznie gorzej poradziła sobie strategia poszukująca trzech występujących po sobie sygnałów kupna lub sprzedaży generowanych przez sztuczną inteligencję. Przykładowym dowodem całkowitego niepowodzenia jest między innymi rok 2019, którego zapis został przedstawiony na rysunku 8.7. W tym okresie, z wyłączeniem jednej pozytywnej transakcji, wszystkie sygnały wygenerowane przez algorytm zakończyły się utratą pieniędzy.

Przyglądając się dokładnie momentom dokonywania transakcji przez strategię „trzy sygnały”, można zauważyć że są one wykonywane w podobnym czasie jak w przypadku pojedynczego sygnału. Główną różnicą jest przesunięcie ich w czasie o kilka dni. W ten sposób przegapiane są wszystkie ekstrema lokalne, co prowadzi do handlu po niekorzystnych cenach. W efekcie rok został zakończony ponad 4% stratą, co znacznie przegrywa z poprzednią strategią. Powodem opisywanego opóźnienia jest zbyt długi czas oczekiwania na otrzymanie trzech kolejnych sygnałów sugerujących wykonanie tego samego ruchu. Założeniem strategii było uzyskanie dużej pewności z odpowiedzi modelu sieci neuronowej, które zabezpieczyłoby przed popełnieniem błędu. Niestety wydłużenie czasu potrzebnego na zebranie takich danych spowodowało kompletną bezużyteczność zaproponowanego rozwiązania.

Inny przykładem potwierdzającym słabą efektywność strategii jest rok 2017. Na rysunku 8.8 można zauważyć, że algorytm wykonał tylko jedną transakcję. Co prawda zakończyła się ona zyskiem, jednak w roku ciągłego wzrostu ceny, osiągnięty przychód daje niesatysfakcjonujący efekt. Dla porównania poprzednia strategia zarobiła ponad dwukrotnie więcej.



Rys. 8.7. Strategia „trzy sygnały” – rok 2019

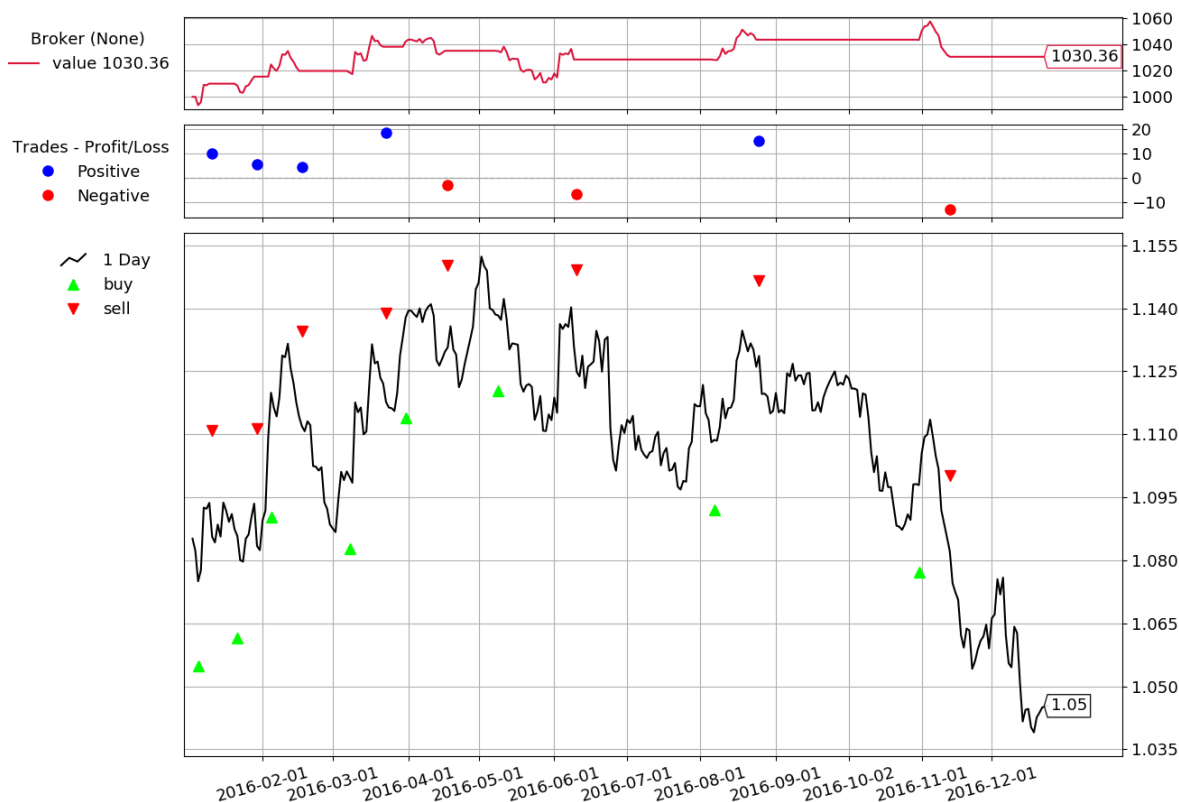


Rys. 8.8. Strategia „trzy sygnały” – rok 2017

8.3. WZROST WZGLĘDNY

Strategia badająca wzrost średniej w odniesieniu do ostatniej ceny na zamknięciu to pierwsza ze strategii opartych na predykcji ceny. Opis jej zachowania w trakcie symulacji rozpocznie się od roku 2016 przedstawionego na rysunku 8.9. Momenty wystawiania sygnałów transakcyjnych wyraźnie różnią się od podejścia bazującego na klasyfikacji. „wzrost średniej” zanotował kilka bardzo dobrych ruchów na początku roku, gdzie udało mu się wypracować widoczny zysk. Transakcje te są dotychczasowo najlepsze spośród wszystkich omawianych strategii, gdyż doskonale wpasowują się w przełamania trendów.

Z kolei w dalszej części roku poza jedną skuteczną reakcją, strategia wyraźnie spóźniała się z momentami kupna i sprzedaży. Jest to szczególnie widoczne w ostatniej z przeprowadzonych inwestycji. Zakup następuje na przełomie października i listopada, gdzie cena praktycznie osiągnęła już swoje lokalne maksimum. Dopiero po około tygodniu od przełamania, waluta zostaje sprzedana. Zbyt długie opóźnienie spowodowało stratę przy dobrej okazji na powiększenie wartości portfela.



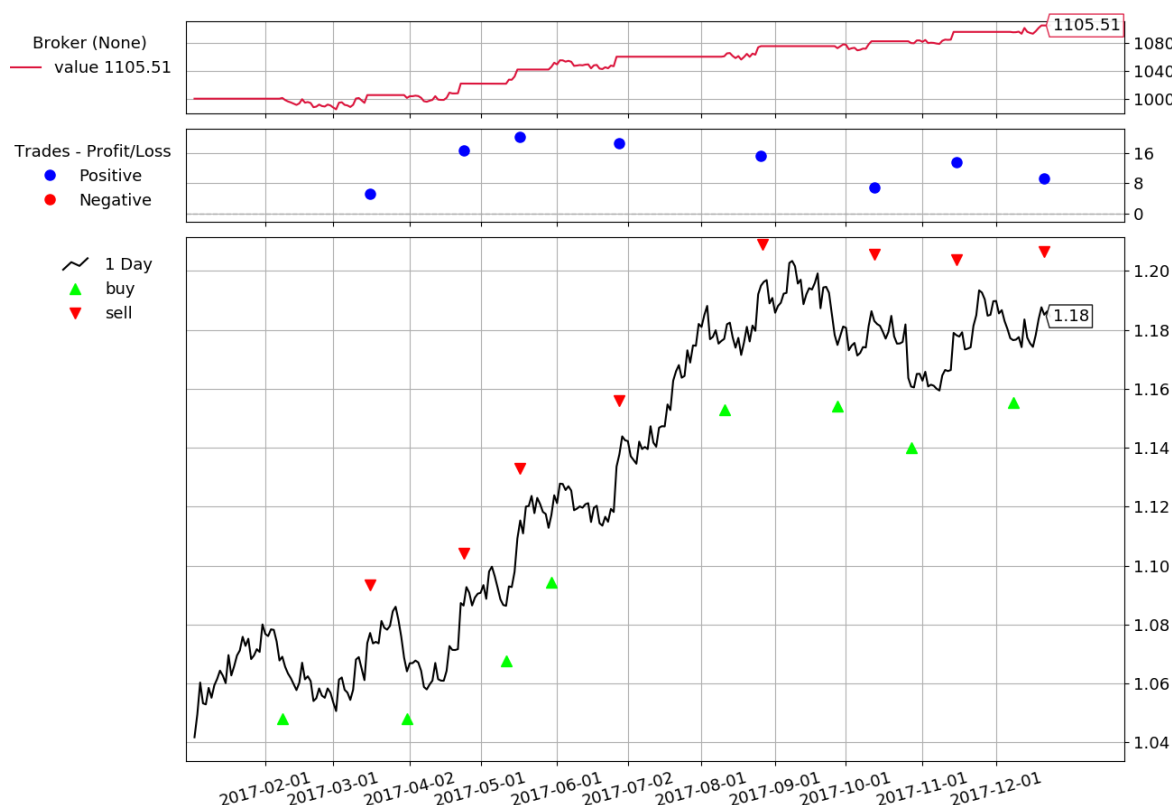
Rys. 8.9. Strategia „wzrost względny” – rok 2016

Ostatecznie strategii udało się uzyskać 3% przychód, co jest wynikiem nieznacznie gorszym od „pojedynczego sygnału”. Czynnikiem decydującym o końcowym sukcesie było dobranie odpowiedniego progu o jaki prognozowana cena musi przewyższać ostatnią, zanotowaną wartość na zamknięciu. Początkowo ustalony 1 cent powodował zbyt małą liczbę transakcji, które było zbyt mocno opóźnione. Za niskie wartości sprawiały zaś, że strategia naprzemiennie skupowała i sprzedawała w krótkich odstępach czasu. Optymalny z punktu widzenia tej oraz kolejnych symulacji okazał się próg o wartości 0.006 dolara. W miarę możliwości wykluczył on generowania niewłaściwych sygnałów transakcyjnych, a za razem nadmiernie nie opóźnił w czasie pozostałych z nich.

Innym wartym uwagi przykładem dobrze obrazującym charakterystykę omawianej strategii jest rok 2017. Prezentowany przebieg symulacji doskonale pokazuje różnicę pomiędzy wykorzystaniem modelu zaprojektowanego do predykcji ceny, a podejściem klasyfikującym. Na rysunku 8.10 możemy zauważyć, że w tym przypadku udało się wypracować ponad 10 procentowy zysk. Jest on jednak wyraźnie mniejszy niż poprzednio, ponieważ w okresie długotrwałego wzrostu ceny, najlepsza decyzja to zainwestowanie wszystkich środków na cały czas trwania trendu. Przedstawiana strategia zdecydowała, że w terminie od marca do października wystawi aż 8 sygnałów transakcyjnych. Dla porównania, „pojedynczy sygnał” nie sprzedał waluty aż do momentu zatrzymania wzrostu.

Większa liczba transakcji jest spowodowana ciągłą analizą zmieniającej się ceny. Model klasyfikujący został nauczony na danych, które jasno opisywały aktualnie panujące trendy. Oznacza to, że przedziały czasowe podobne do opisywanego, były oznaczone w zbiorze treningowym jako dobre momenty na kupno od początku do końca trwania tendencji wzrostowej. Model został więc nauczony, że po rozpoczęciu trendu nie ma potrzeby sprzedaży posiadanych instrumentów, aż do jego zakończenia.

Z kolei strategia badająca prognozowaną cenę sprzedawała zakupione euro w momentach krótkotrwałego zahamowania wzrostu. W takiej sytuacji pomimo zauważalnego zysku, wypracowany przychód jest niższy niż potencjalny zarobek przy mniejszej ilości transakcji.

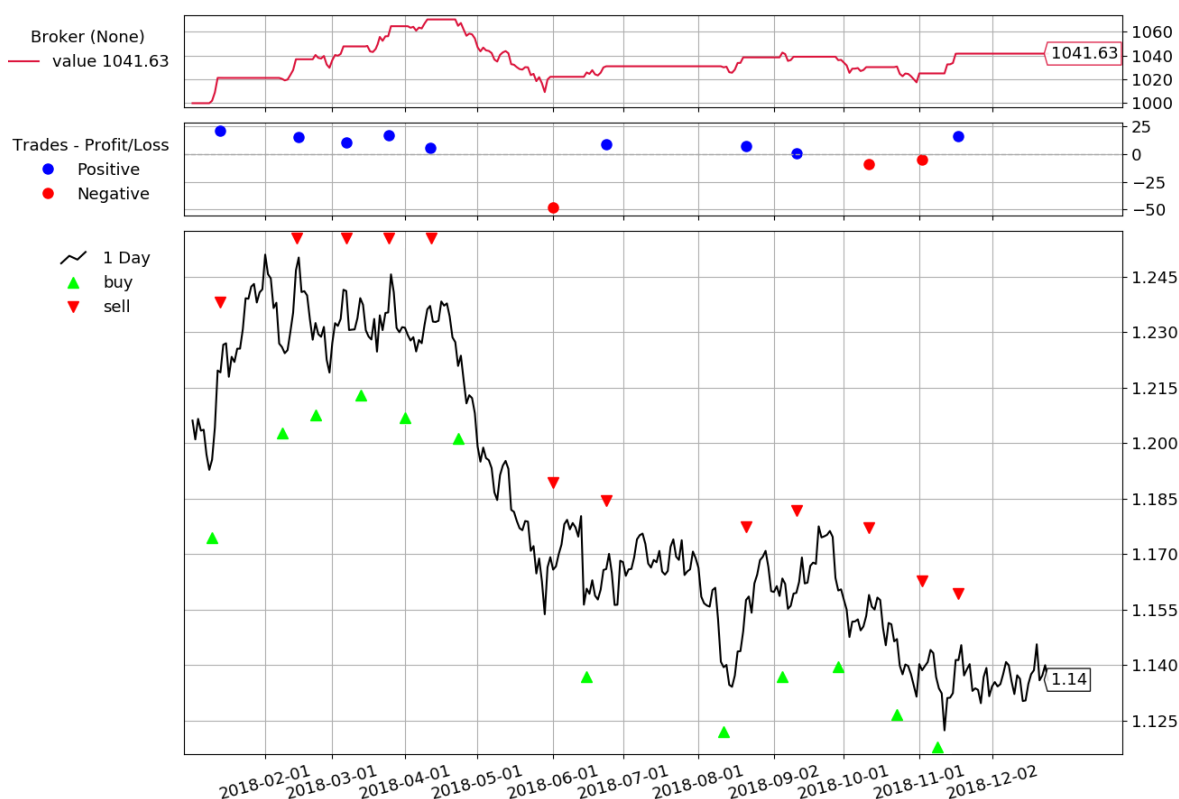


Rys. 8.10. Strategia „wzrost względny” – rok 2017

Największym zaskoczeniem w trakcie testowania strategii wzrostu względnego jest wynik uzyskany w roku 2018. W okresie przeważającego trendu spadkowego, algorytm był w stanie wypracować 4 procentowy przychód, co daje najwyższy zysk ze wszystkich dotychczas testowanych strategii. Zastanawiający jest jednak sposób w jaki udało się uzyskać wspomniany wynik.

Na rysunku 8.11 można zaobserwować, że w okresie od stycznia do maja 2018 roku, strategia wykonała 5 skutecznych transakcji, które przyniosły przychód przekraczający 60 dolarów. Biorąc pod uwagę przebieg zmian ceny, jaki miał miejsce w tym czasie, taki wynik jest wyjątkowo skutecznym rezultatem. Kurs EUR/USD charakteryzował się wówczas gwałtownymi przejściami pomiędzy kierunkiem wzrostowym i spadkowym. Dodatkowo, zmiany te oscylowały w bardzo niewielkim przedziale pomiędzy 1.22, a 1.25 dolara. Wykrycie tak krótkich i stromych zbroczy, oznaczałoby niesamowitą precyzję zastosowanego modelu neuronowego. Ponadto wysłane sygnały transakcyjne trafiają idealnie w ekstrema lokalne, gdyż transakcje zostały przeprowadzone niemal dokładnie w trakcie przełamania trendów.

Elementem podważającym tak wysoką skuteczność strategii jest następujący po omawianym okresie, długotrwały trend spadkowy. Okazuje się, że w tym przypadku, podjęta decyzja była całkowicie nie trafna. Po zainwestowaniu środków na koniec kwietnia 2018 roku, algorytm całkowicie nie wykrył dalszego spadku ceny, przez co stracił aż 50 dolarów.



Rys. 8.11. Strategia „wzrost względny” – rok 2018

W celu znalezienia odpowiedzi na pytanie, w jaki sposób strategia tak precyzyjnie odnalazła momenty zmian trendów w pewnej części okresu symulacyjnego, sprawdzony został przebieg ceny prognozowanej przez wykorzystany model. Na rysunku 8.12 przedstawiono analizowany fragment symulacji z dodatkowym sygnałem pokazującym informacje na temat przewidywanego kursu instrumentu.

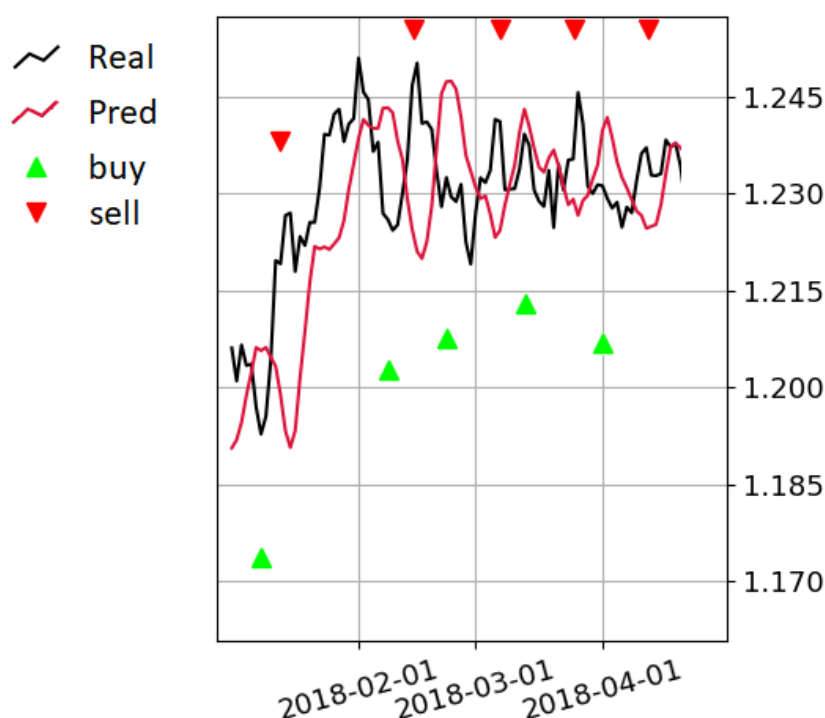
Okazuje się, że seria skutecznych decyzji nie była wynikiem prawidłowych predykcji modelu. Na ostateczny rezultat wpłynęły dwa główne czynniki:

- Opóźnienie w przewidywaniu przyszłych trendów
- Niedokładność prognozowanej ceny

Szczegółowa analiza momentów wykonywania poszczególnych transakcji pokazuje, że różnicę pomiędzy średnią z prognoz, a ceną rzeczywistą, osiągnęły poziom wybranego progu dokładnie w miejscach przełamania trendu. Przede wszystkim przewidywane zmiany kierunków przebiegu notowania wystąpiły zdecydowanie za późno. Ponadto prognozowana cena nie pokrywała się idealnie z danymi historycznymi. W efekcie wystawienie sygnałów transakcyjnych nie było spowodowane perfekcyjnym funkcjonowaniem systemu, a jedynie przypadkowym sukcesem w trakcie popełnienia szeregu błędów.

Szybkie zmiany trendów, które nie zostały wykryte na czas, spowodowały, że cena na zamknięciu była wystarczająco niższa od przewidywanego kursu w ekstremach lokalnych analizowanego przebiegu. Niestety nie były to te same ekstrema. Gdy prawdziwa cena przechodziła przez swoje minimum lokalne, model stwierdzał, że znajduje się w lokalnym maksimum. Wówczas różnica obu cen miała na tyle wysoką wartość, że algorytm postanowił wysłać sygnał kupna.

Ta nadzwyczajnie skuteczna decyzja została osiągnięta przez przypadkowe złożenie się kilku czynników, wpływających na końcowy rezultat. Hipotetyczny inwestor korzystający z przedstawianej strategii, faktycznie osiągnąłby satysfakcjonujący przychód w danym okresie. Jednak gdyby kurs powędrował w innym kierunku, doświadczyłby on utraty pieniędzy. Przykładem takiego zachowania jest wcześniej opisywana transakcja, w której stracono 50 dolarów.



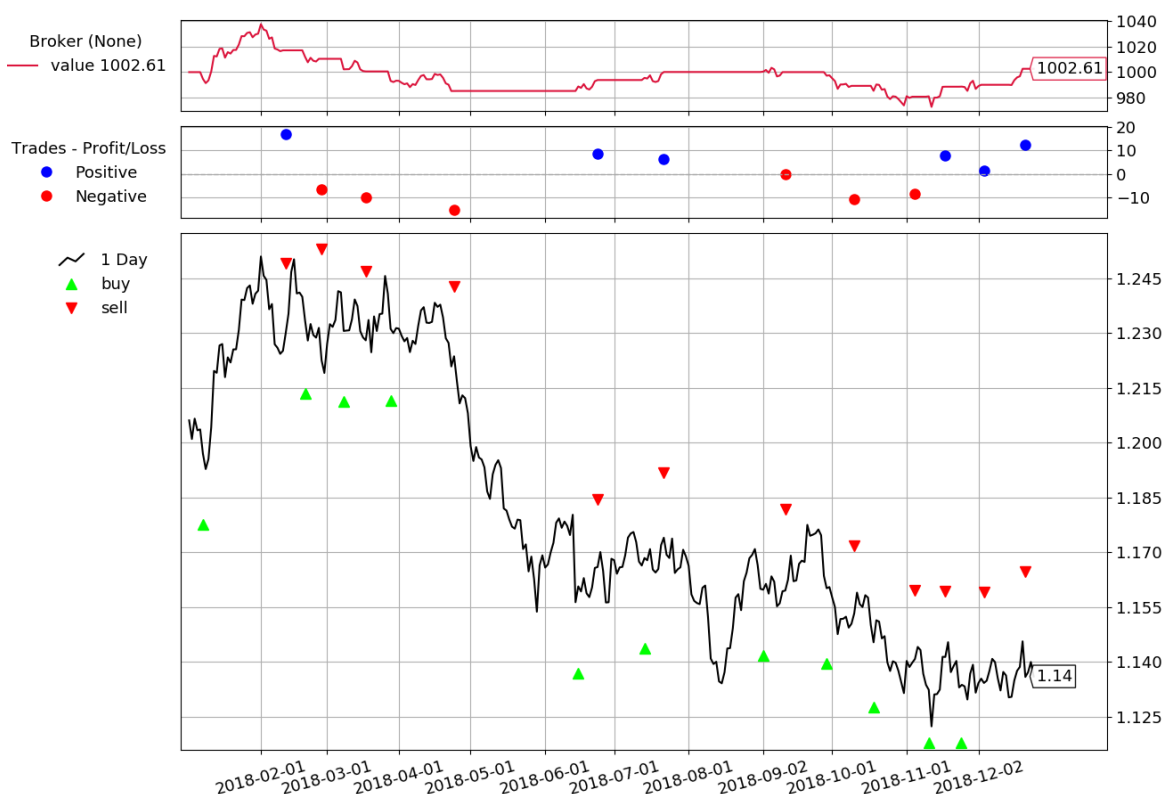
Rys. 8.12. Strategia „wzrost względny” – analiza prognozowanej ceny

Strategia została sprawdzona w ten sam sposób również na innych zbiorach testowych. W wielu sytuacjach podejmowane decyzje były przeprowadzane na tej samej podstawie jak przykład opisany powyżej. Pozytywne rezultaty często wynikały z błędów jakie popełniał przygotowany model, które przypadkowo okazywały się skuteczne. Z tego powodu strategia „wzrost względny” nie jest systemem polecanym do wykorzystania podczas inwestycji prawdziwych pieniędzy w czasie rzeczywistym.

8.4. WZROST BEZWZGLĘDNY

Gwarancję uniknięcia tych samych błędów symulacyjnych gwarantuję strategia badająca wzrost bezwzględny, która sprawdza zmiany jedynie wewnątrz prognozowanych wartości. W celu porównania zasady wysyłania sygnałów transakcyjnych, w pierwszej kolejności pod uwagę wzięty zostanie ten sam zbiór testowy, którym jest rok 2018. Całość symulacji przedstawiono na rysunku 8.13.

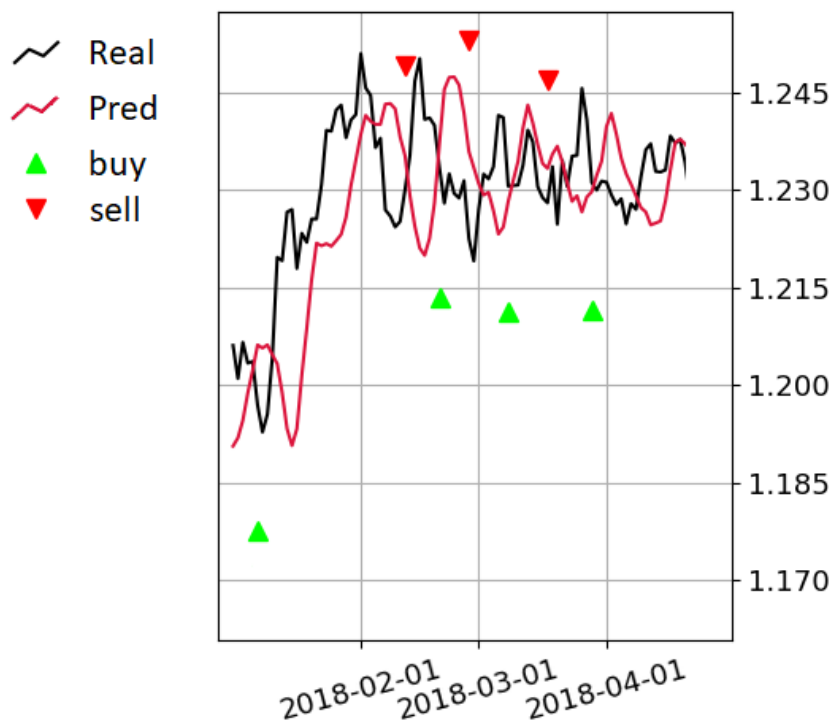
Na wstępie widoczny jest znacznie mniejszy przychód niż w przypadku „wzrostu względnego”. Tym razem rok został zakończony z mniej więcej zerowym bilansem portfela. Warto jednak zauważyć, że strategia wykonała zdecydowanie inne transakcje. Jej plusem jest dobre wykrycie najdłuższych trendów, na których stan konta może zmienić się najmocniej. Z kolei przy mniejszych zmianach, odnotowano większą ilość strat.



Rys. 8.13. Strategia „wzrost bezwzględny” – rok 2018

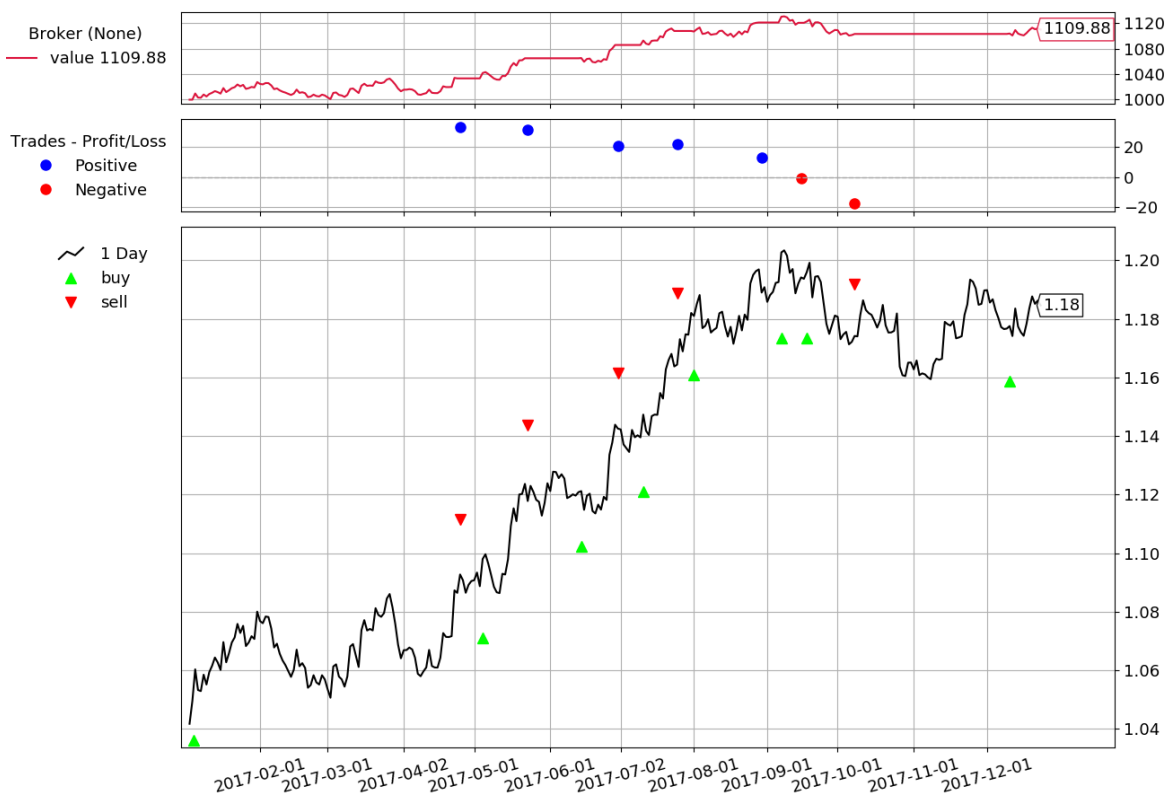
Szczegółowa analiza pierwszej części roku została przedstawiona na rysunku 8.14. Pomimo tego samego przebiegu prognozowanej ceny, sygnały transakcyjne są generowane w zdecydowanie innych momentach. Na początku strategia zauważyła wzrost ceny, więc inwestuje pieniądze, aż do czasu pierwszego wyraźnego spadku. Sprzedaż nie nastąpiła już w tak idealnym momencie, jakim jest znajdujące się bardzo blisko maksimum lokalne, jednak tym razem nie wynika ona z przypadkowego ustawienia ceny rzeczywistej i prognozowanej.

W dalszej części następuje seria niekorzystnych inwestycji zakończonych utratą kapitału. Dzieje się tak przez błędne przewidywanie ceny, które charakteryzuje się opóźnieniem w stosunku do historycznego zapisu. Tym razem widać jednak, że podejmowane decyzje wynikają bezpośrednio z danych otrzymywanych od modelu neuronowego, a ryzyko otrzymania przypadkowych zysków, błędnie wpływających na ocenę strategii jest znacznie zmniejszone.



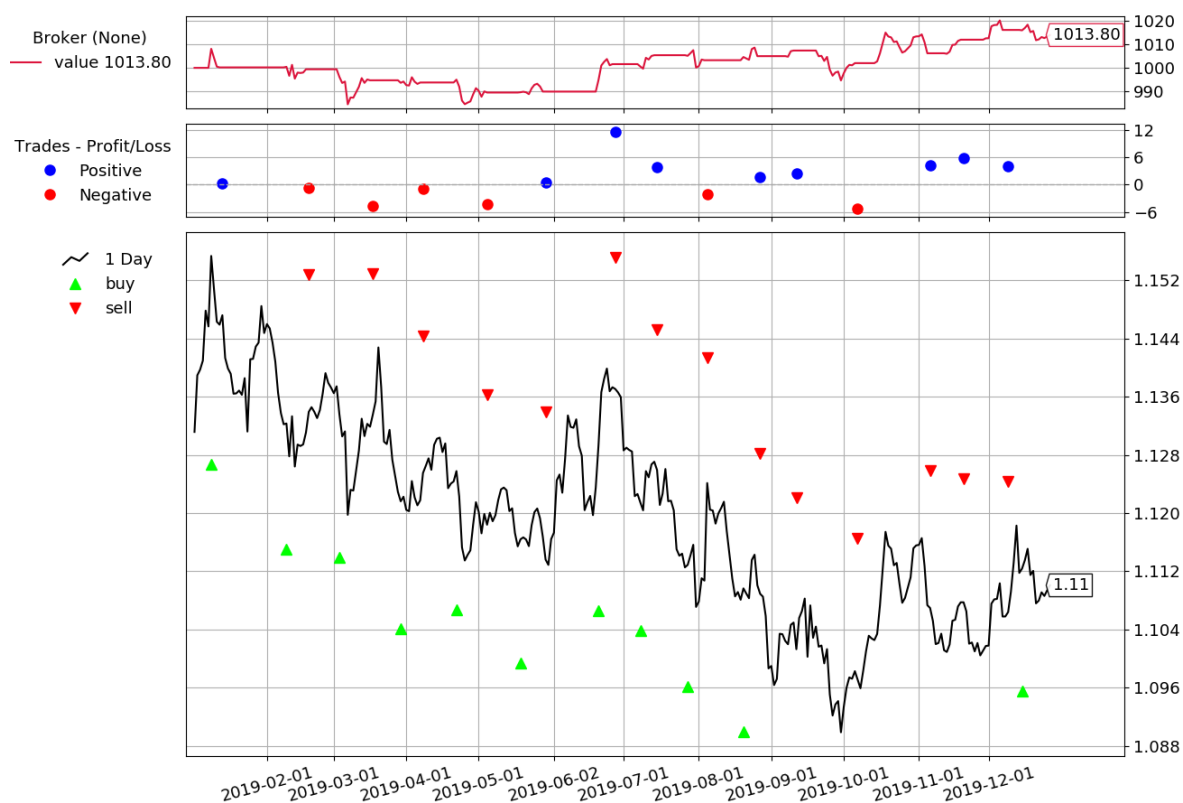
Rys. 8.14. Strategia „wzrost bezwzględny” – analiza prognozowanej ceny

Na rysunkach 8.15 i 8.16 przedstawiono wyniki strategii na innych zbiorach testowych. W pierwszym z nich można zaobserwować, że przy rosnącym kursie, udało się wypracować całkiem wysoki zysk. Ze względu na ciągłą analizę prognozowanej ceny oraz błędy modelu, tylko mała część transakcji zakończyła się stratą.



Rys. 8.15. Strategia „wzrost bezwzględny” – rok 2017

W ciężkim z inwestycyjnego punktu widzenia okresie, jakim jest rok 2019, także udało wypracować się przychód. Jest on jednak niewielki i mniejszy od strategii wykorzystującej klasyfikację szeregu czasowego. Zyski są widoczne głównie w trakcie największych trendów wzrostowych. W przypadku szybkich zmian, model generuje dane ze zbyt dużym przesunięciem, by inwestycje mogły zakończyć się powodzeniem. W efekcie strategia wykonała kilka negatywnych transakcji, które nie były jednak w stanie doprowadzić do utraty pieniędzy.



Rys. 8.16. Strategia „wzrost bezwzględny” – rok 2019

9. PODSUMOWANIE

9.1. WYBÓR NAJLEPSZEGO SYSTEMU

Przeprowadzenie symulacji gry giełdowej na przygotowanych zbiorach testowych wraz z analizą uzyskanych wyników, dostarcza wystarczającą ilość wiedzy do świadomego wyboru najlepszego ze wszystkich zaproponowanych systemów inwestycyjnych. Ostateczna decyzja zostanie podjęta nie tylko na podstawie spostrzeżeń z poprzednich rozdziałów pracy, ale również zestawienia wszystkich zebranych danych będących efektem wykonanych badań.

Tabela 9.1 przedstawia roczne stopy zwrotów osiągnięte przez testowane strategie we wszystkich symulowanych latach. Ponadto w ostatniej kolumnie znajdują się rzeczywista zmiana ceny, obliczona na podstawie ilorazu kursu początkowego i końcowego w danym roku. Pozwoli ona na określenie czy wypracowany przychód nie wynikał jedynie z historycznego wzrostu ceny pary walutowej. Dodatkowo w ostatnim wierszu tabeli zapisano średnią stopę zwrotu uzyskaną podczas wszystkich symulacji.

Tab. 9.1. Zestawienie wszystkich zebranych wyników symulacji

Roczna stopa zwrotu					
Rok	Pojedynczy sygnał	Trzy Sygnały	Wzrost względny	Wzrost bezwzględny	Rzeczywista zmiana ceny
2022	-1.9%	-9.8%	-4.3%	-2.5%	-7.7%
2021	0.1%	-6%	-0.4%	-1%	-7.4%
2020	9.1%	1.2%	5%	6.2%	9.3%
2019	4%	-4.3%	2%	1.3%	-0.9%
2018	0.3%	-8.5%	4.1%	0.2%	-5%
2017	14.9%	6.6%	10.5%	11%	15.3%
2016	4%	-3.8%	3%	3.2%	-2.8%
Średnia	4.35%	-3.50%	2.84%	2.62%	0.1%

W pierwszej kolejności odrzucona zostanie strategia „trzy sygnały”. Wyniki przez nią osiągnięte w głównej mierze zakończyły się utratą środków. Zaproponowane podejście okazało się być kompletnie nieskuteczne przez zbyt duże opóźnienie w trakcie wystawiania sygnałów transakcyjnych.

Porównywalne przychody osiągnęły obie strategie oparte na modelach prognozujących przyszłą cenę instrumentu finansowego. Zarówno „wzrost względny” jak i „wzrost bezwzględny” zarobiły średnio ponad 2.5% w skali roku. Biorąc jednak pod uwagę analizę przeprowadzoną w poprzednim rozdziale pracy, algorytm odnoszący się do ostatniej zanotowanej ceny na zamknięciu, wykonywał wiele transakcji opierających się na przypadkowym układzie dwóch kursów. Przedstawione przykłady pokazały, że część decyzji zakończyła się sukcesem jedynie przez szczęśliwe wystąpienie poszczególnych trendów w przyszłości. Z tego względu rozsądniejszym wyborem jest wykorzystanie strategii „wzrost bezwzględny”, która przy skromnym zysku bazowała na realnej wiedzy dostarczanej przez przygotowany model neuronowy.

Ostatecznie okazuje się, że największy przychód osiągnęła najprostsza ze strategii, w której bazą do wygenerowania sygnału transakcyjnego była pojedyncza informacja otrzymywana od modelu klasyfikującego szereg czasowy. W efekcie otrzymano średni zysk na poziomie 4.35% każdego roku. Warto zauważyć, że w latach największego wzrostu ceny, strategia była w stanie najefektywniej wykorzystać długotrwałe trendy wzrostowe. Z kolei przy zdecydowanie mniej atrakcyjnych okresach, zazwyczaj odnotowywano stopę zwrotu o około 5% większą niż roczna zmiana ceny EUR/USD w tym czasie.

Na podstawie przeprowadzonych badań można więc stwierdzić, że przygotowany system transakcyjny oparty na klasyfikacji szeregu czasowego jest rozwiązaniem skuteczniejszym niż konkurencyjne podejście bazujące na prognozowaniu ceny.

9.2. WNIOSKI

Podsumowując efekty badań nad zastosowaniem głębokich sieci neuronowych w zadaniach predykcji cen akcji oraz sygnałów transakcyjnych, można stwierdzić, że cel pracy magisterskiej został w pełni zrealizowany. Przygotowane dane oraz modele pozwoliły na zaprojektowanie różnych strategii, które w trakcie symulacji określiły skuteczność tego typu rozwiązań.

Wyniki pracy pokazują, że przewidywanie przyszłych zmian cen na giełdzie papierów wartościowych jest zadaniem niezwykle trudnym i ciężkim pod względem uzyskania realnego przychodu. Wiedza pozyskana z przeprowadzonych eksperymentów skłania do wyciągnięcia następujących wniosków:

- Głębokie sieci neuronowe nie są szczególnie efektywne podczas predykcji ceny w następnych dniach. Ich wysoka dokładność, wykazywana przez niski pierwiastek błędu średniokwadratowego, jest tylko pozornie skuteczna w kontekście przewidywania przyszłości. Przyglądając się dokładnie otrzymywanym odpowiedziom, za każdym razem można zaobserwować pewne opóźnienie w czasie względem prawdziwych wartości. W ten sposób modele osiągają wysokie wyniki statystyczne, jednak opracowane na ich podstawie strategie, nie spisują się równie dobrze w trakcie symulacji gry na giełdzie.
- Rozwiązanie poszukujące odpowiedzi na pytanie w jakim kierunku podąży cena w przyszłości, przyniosło całkiem dobre rezultaty. Najlepsze modele osiągnęły skuteczność na poziomie niespełna 75%, która daje nadzieje na osiągnięcie sukcesu inwestycyjnego. Tak jak w poprzednim przypadku, największym problemem są momenty przełamania trendów, gdyż to w tych miejscach pojawia się najwięcej błędów. Z kolei strategia oparta na klasyfikacji przyniosła najlepsze przychody ze względu na lepsze zachowanie w trakcie największych wzrostów i spadków cen. Dzięki temu zaprojektowany system został uznany jako najlepsza z testowanych propozycji.
- Możliwość wykorzystania głębokich sieci neuronowych podczas inwestycji prawdziwych pieniędzy na giełdzie budzi wiele wątpliwości. Sygnały transakcyjne wysyłane w oparciu jedynie o wiedzę wytrenowanych modeli są obarczone bardzo dużym ryzykiem straty. W trakcie badań przeprowadzono setki symulacji, które miały na celu dostroić wszystkie parametry, takie jak progi cenowe lub prawdopodobieństwo prawdziwości na wyjściu sieci. Dzięki temu przy najlepszym rozwiązaniu udało się uzyskać średni przychód przekraczający 4% w skali roku.

Istnieje jednak wysoka szansa na to, że przy kolejnych zbiorach testowych stworzonych na podstawie innych okresów lub instrumentów finansowych, wyniki badań nie powtórzyłyby się. Przed realnym wykorzystaniem systemów należałoby więc zwiększyć ilość symulacji testowych wraz z ich szczegółową analizą.

- Biorąc pod uwagę charakterystykę wybranego zbioru danych, osiągnięty przychód nie jest wynikiem kompletnie niesatysfakcjonującym. Wysoka płynność oraz niewielkie zmiany cen pary walutowej EUR/USD, stanowią ciężkie wyzwanie dla inwestorów przy próbach osiągnięcia zysku. Ponadto zastosowanie systemu transakcyjnego stanowi źródło pasywnego dochodu, które nie wymaga stałego monitorowania sytuacji rynkowej.

Z drugiej strony ryzyko ewentualnej straty jest na tyle wysokie, że 4 procentowa stopa zwrotu nie rekompensuje niebezpieczeństwa związanego z inwestycją. Szczególnie w czasach wysokiej inflacji, istnieją inne możliwości ulokowania środków, które przynoszą taki sam lub większy zysk, przy jednoczesnym zachowaniu pełnego bezpieczeństwa.

BIBLIOGRAFIA

- [1] Redakcja A.O, "Giełda Papierów wartościowych: definicje i nie tylko!", 16.05.2023 <https://wzr.pl/ekonomia/gielda-papierow-wartosciowych-definicja-i-nie-tylko/>
- [2] Krzysztof Szymański, "Jak działa giełda? Jakie są rodzaje giełd i w co możemy na nich inwestować?", 16.05.2022, <https://direct.money.pl/artykuly/porady/jak-dziala-gielda>
- [3] Główny Urząd Statystyczny - Akcja
- [4] Dr Wojciech Świder, "Pierwsza spółka akcyjna - Holenderska Kompania Wschodnioindyjska", 19.11.2022
- [5] https://en.wikipedia.org/wiki/File:Vereinigtes_Ostindische_Companie_bond.jpg
- [6] Karina K. "Na czym polega akcja? Jak jest jej znaczenie w inwestowaniu?" 14.06.2023 <https://www.gowork.pl/poradnik/17/zostan-pracodawca/na-czym-polega-akcja-jakie-jest-jej-znaczenie-w-inwestowaniu/>
- [7] "Co to są obligacje i czy warto w nie inwestować?" <https://www.skarbiec.pl/news/co-to-sa-obligacje-i-czy-warto-w-nie-inwestowac/>
- [8] Marcin Kowalczyk, "ETF - co to jest i jak w nie inwestować", 20.12.2022 <https://www.najlepszekonto.pl/etf-co-to-jest-i-jak-w-nie-inwestowac>
- [9] <https://www.cmcmarkets.com/pl-pl/edukacja-forex/czym-sa-transakcje-forex>
- [10] Krzysztof Szymański, "Jak działa Forex? Jak inwestować na nim w waluty?", 27.07.2022, <https://direct.money.pl/artykuly/porady/jak-dziala-forex-jak-inwestowac-na-nim-w-waluty>
- [11] <https://www.tradersdna.com/switch-different-currency-pair/>
- [12] Szymon Machniewski, "Analiza fundamentalna - do czego jest wykorzystywana?", 21.08.2022,
- [13] Enrique Valls, "Co to jest analiza fundamentalna?", 20.06.2023, <https://rankia.pl/analizy-gieldowe/czym-jest-analiza-fundamentalna/>
- [14] Szymon Machniewski, "Wskaźniki analizy technicznej - czym są, podział skuteczność", 02.10.2022
- [15] <https://www.xtb.com/pl/edukacja/analiza-techniczna>
- [16] Encyklopedia Zarządzania, "Analiza trendu", https://mfiles.pl/pl/index.php/Analiza_trendu

- [17] Naveen Joshi, "The evolution of neural networks", 16.12.2018
<https://www.allerin.com/blog/the-evolution-of-neural-networks>
- [18] <https://pl.wikipedia.org/wiki/Neuron>
- [19] W. Zamora-Cárdenas, M. Zumbado, Trejos-Zelaya, "McCulloch-Pitts, Artificial Neuron and Rosenblatt's Perceptron: An abstract specification in Z", 01.03.2020
- [20] https://pl.wikipedia.org/wiki/Neuron_McCullocha-Pittsa
- [21] Richard Nagyfi, "The differences between Artificial and Biological Neural Networks", 04.09.2018
- [22] Akshay L Chandra, "McCulloch-Pitts Neuron - Mankind's First Mathematical Model Of A Biological Neuron", 24.07.2018
- [23] Jean-Christophe B. Loiseau, "Rosenblatt's perceptron, the first modern neural network", 11.03.2019
- [24] Widrow, B., & Lehr, M. A. (1990). 30 years of adaptive neural networks: perceptron, Madaline, and backpropagation. *Proceedings of the IEEE*, 78(9), 1415–1442. doi:10.1109/5.58323
- [25] Kamil Ciomcia, "Czym jest deep learning i sieci neuronowe", 03.02.2020
- [26] HECHT-NIELSEN, R. (1992). Theory of the Backpropagation Neural Network. *Proceeding in the International Joint Conference on Neural Networks* 1, 593–611, June 1989. © 1989 IEEE. 65–93. doi:10.1016/b978-0-12-741252-8.50010-8
- [27] "Głębokie sieci neuronowe — czym są i jakie mają zastosowanie?", 15.12.2022
<https://mindboxgroup.com/pl/glebokie-sieci-neuronowe-czym-sa-i-jakie-maja-zastosowanie/>
- [28] "Top 25 Deep Learning Applications Used Across Industries". 21.07.2023
<https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-applications>
- [29] Agnieszka Ławrynowicz, "Rekurencyjne sieci neuronowe, LSTM i GRU", Wydział Informatyki Politechniki Poznańskiej, 08.05.2018
https://www.cs.put.poznan.pl/alawrynowicz/PJN_5.pdf
- [30] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, 1–36. doi:10.1162/neco_a_01199
- [31] <https://blog.lukaszogan.com/informatyka/rekurencyjne-sieci-neuronowe-generuja-muzyke/>
- [32] Shi, M., & Zhao, Q. (2020). Stock Market Trend Prediction and Investment Strategy by Deep Neural Networks. 2020 11th International Conference on Awareness Science and Technology (iCAST). doi:10.1109/icast51195.2020.9319488

- [33] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669. doi:10.1016/j.ejor.2017.11.054
- [34] <https://www.ig.com/en-ch/trading-strategies/the-complete-guide-to-trading-strategies-and-styles-190709>