

Одеський національний університет імені І. І. Мечникова
Факультет математики, фізики та інформаційних технологій
Кафедра оптимального керування та економічної кібернетики

Кваліфікаційна робота

на здобуття ступеня вищої освіти «магістр»

**«Вейвлетний аналіз та прогнозування фінансових
часових рядів»**

«Wavelet analysis and financial time series forecasting»

Виконала: здобувачка денної форми навчання

спеціальності 113 Прикладна математика

Освітня програма «Прикладна математика»

Міловська Карина Миколаївна

Керівник: канд. техн. наук, доц. Мороз В. В. _____

Рецензент: канд. фіз.-мат. наук, доц. Страхов Є. М.

Рекомендовано до захисту:

Протокол засідання кафедри

№ ____ від _____ 2022 р.

Завідувач кафедри

Захищено на засіданні ЕК № _____

Протокол № ____ від _____ 2022 р.

Оцінка _____ / _____ / _____

Голова ЕК

Одеса — 2022 р.

ЗМІСТ

Вступ	4
1 Аналіз методів прогнозу курсу криптовалют	7
1.1 Основні поняття	7
1.2 Прогнозування курсу криптовалют	8
1.3 Прогнозування часових рядів	9
1.3.1 Метод використання інтегрованої моделі авторегресії	10
1.3.2 Метод використання штучних нейронних мереж . . .	11
1.3.3 Метод використання рекурентної мережі з довгою короткостроковою пам'яттю	12
1.3.4 Метод використання згорткової нейронної мережі . .	13
1.4 Використання вейвлет-аналізу у фінансах	14
1.4.1 Огляд підходів	14
1.4.2 Вейвлети в прогнозуванні криптовалют	16
1.5 Висновки до розділу	18
2 Нейронні мережі та вейвлетний аналіз	19
2.1 Визначення та характеристики часового ряду	19
2.2 Прогнозування часових рядів	20
2.3 Штучна нейронна мережа	24
2.4 Рекурентна нейронна мережа	26
2.5 Згорткова нейронна мережа	28
2.6 Вейвлет-аналіз	29
2.7 Дискретне вейвлет-перетворення	31
2.7.1 Вейвлет Хаара	31
2.7.2 Вейвлет Добеші	32
2.8 Дискретне вейвлет-перетворення максимального перекриття	33
2.9 Висновки до розділу	35
3 Алгоритм прогнозування WA-LSTM	36
3.1 Запропонована схема прогнозування	36

3.2	Емпіричні дослідження	36
3.2.1	Опис вхідних даних	36
3.2.2	Етап 1: Декомпозиція курсу криптовалюти Bitcoin . .	38
3.2.3	Етап 2: Прогнозування та реконструкція	42
3.3	Оцінка ефективності прогнозу	43
3.4	Висновки до розділу	44
4	Практична реалізація програмного продукту	45
4.1	Вибір мови програмування та бібліотек	45
4.2	Вхідні дані	46
4.2.1	Data Exploration – дослідження даних	46
4.2.2	Data Preprocessing — попередня обробка даних	48
4.3	Побудова моделі LSTM	49
4.4	Результати прогнозу на основі "чистої" моделі LSTM	51
4.5	Використання вейвлет-перетворення та результати дослідження	52
4.5.1	DWT (тип: Хаар) + LSTM	52
4.5.2	DWT (тип: Добеші) + LSTM	54
4.5.3	MODWT (рівень розкладання 1) + LSTM	55
4.5.4	MODWT (рівень розкладання 2) + LSTM	56
4.6	Оцінка ефективності моделі	57
4.7	Висновки до розділу	57
	Висновки	59
	Список літератури	61

ВСТУП

Протягом останніх десятиліть у світі активно відбувався процес інформатизації економіки, набуло популярності поняття віртуальної валюти, унаслідок чого найцікавішим питанням сучасних фінансів є криптовалюта та відповідний їй клас активів. Криптовалюта — різновид цифрової валюти, створення і контроль за якою базуються на криптографічних методах, за своєю природою вона є унікальним активом, який має власні властивості. Термін закріпився після статті «Cryptocurrency», яка була опублікована у 2011 році у журналі «Forbes» [9].

Криптовалюта у сучасній системі міжнародних валютно-фінансових та кредитних відносин є зручною формою електронних розрахунків і перспективною для здійснення інвестицій. На сьогодні «CoinMarketCap» включає 9167 цифрових активів. Мають попит серед споживачів ті криптовалюти, які мають високу ліквідність, постійність курсових показників, перспективи подальшого розвитку і позитивну репутацію засновників.

Найбільш популярними трендами сфери цифрових грошових одиниць є:

- Біткоїн (Bitcoin, BTC) — лідер серед віртуальних грошей, на основі блокчейна постійно створюються різноманітні технічні рішення, а розробники постійно працюють над їх модернізацією;
- Ефіріум (Ethereum, ETH) — стимулює появу децентралізованих стартапів на основі блокчейна, що базуються на смарт-контрактах;
- Ripple (XRP) — найшвидша криптовалюта;

У результаті виявлених трендів розвитку криптовалюти зростає увага до криптовалютних торгів, які являють собою постійно зростаючий і перспективний фінансовий ринок. Проблема прогнозування на ринку криптовалют є актуальною та цікавою задачею з точки зору його високої волатильності, невизначеності та сильними коливаннями цін з часом, а самі криптовалюти знаходяться на перехідній стадії, через що є нестабільними.

Крім того, багато фінансових змінних тепер можна вимірювати на хвилинній або навіть секундній основі, ця зростаюча складність вимагає інструментів обробки даних. На сьогодні вейвлетні технології є досить новим підходом для дослідження різноманітних систем, але їх використання

вже досягло широкого розмаху в економіці та фінансах. Основна сфера застосування вейвлет-аналізу — обробка нестаціонарних (у часі) або неоднорідних (у просторі) сигналів різних типів. Разом з тих, досить часто почали використовувати вейвлет-аналіз, щоб покращити вхідні дані, тобто з метою усунення шумів, що є великою проблемою у ході дослідження динаміки курсу, і декомпозиції даних, що у комбінуванні з нейронними мережами значно підвищує точність і зменшує величину помилок прогнозу.

Актуальність дослідження: інноваційність та великі перспективи робить криптовалюту цінним активом, який варто прогнозувати, тому питання формування та прогнозування на ринку криптовалют є надзвичайно цікавим та актуальним.

Метою дослідження є дослідження прогнозування курсу криптовалют за допомогою вейвлет-аналізу та нейронної мережі, порівняння отриманих результатів з прогнозом без використання вейвлет-перетворення й визначення ефективності створеного рішення.

Поставлена мета зумовила необхідність вирішення таких завдань:

- Аналіз і дослідження предметної області;
- Огляд і аналіз існуючих методів прогнозування курсів криптовалют;
- Визначення ефективності наявних методів прогнозування;
- Прийняття рішення про використання вейвлетного аналізу та методу машинного навчання для прогнозування курсу криптовалют;
- Написання програмного коду;
- Порівняння отриманих результатів з прогнозом без попередньої обробки даних;
- Визначення ефективності створеного рішення.

Об'єктом дослідження є ефективність використання комбінованих моделей, які поєднують вейвлет-перетворення та модель глибокого навчання для прогнозування криптовалют.

Предметом дослідження є вейвлет-аналіз та методи машинного навчання для прогнозування криптовалют.

У 2020 році дана робота пройшла апробацію на 20-й міжнародній науково-технічній конференції «Проблеми інформатики та моделювання»,

яка проводилася на базі національного технічного університету «Харківський політехнічний інститут». Тези було опубліковано у збірнику [27]. У 2021 році робота пройшла апробацію результатів дослідження на 77-й звітній студентській науковій конференції Одеського національного університету імені І.І.Мечникова. Тези опубліковано у збірнику [28].

У 2022 році основні положення та висновки роботи пройшли апробацію на 22-й міжнародній науково-технічній конференції «Проблеми інформатики та моделювання», яка проводилася на базі національного технічного університету «Харківський політехнічний інститут». Тези було опубліковано у збірнику [29]. За матеріалами магістерської роботи була опублікована наукова стаття [30].

РОЗДІЛ 1

АНАЛІЗ МЕТОДІВ ПРОГНОЗУ КУРСУ КРИПТОВАЛЮТ

1.1. Основні поняття

Поява та функціонування криптовалют у світі створило нову течію досліджень у науковій сфері. Обіг криптовалют набирає обертів, що зумовлює інтерес науковців до їх сутності, аналіз якого викладено у працях А. Бейкверді [3], Г. Хілемана, М. Раухс [11], Д. Вахрушева, О. Железова. Попри таку велику кількість робіт, тематика криптоіндустрії та обігу криптовалют в науковому контексті є недостатньо розробленою, зокрема, більшість опублікованих матеріалів носить здебільшого описовий характер.

Криптовалюта — різновид цифрової валюти, створення і контроль за якою базуються на криптографічних методах. Вона має надійний захист від підробки та може використовуватись у якості засобу інвестування. Це підтверджується створенням та успішним функціонуванням бірж криптовалют, інвестиційних проектів.

Початкові ідеї створення криптовалют почали виникати в кінці 90-х років минулого століття. Ідеологи криптовалют в першу чергу прагнули забезпечити повну незалежність нової валюти від держави, щоб гарантувати бажану анонімність для учасників торгівлі криптовалютою. У 2008 році світ побачив першу повноцінну криптовалюту. Сатоші Накамото (за деякими версіями йдеться не про одну особу, а про групу вчених під псевдонімом) запропонував концепцію цього виду валюти, розробив протокол криптовалют Bitcoin і використав його, створивши першу версію програмного забезпечення за даним протоколом [32]. Вже в 2009 була запущена мережа Bitcoin. Успіх її застосування, наявність протоколу та відкритого програмного коду сприяли стрімкому зростанню кількості криптовалют та учасників, що зацікавлені в їх існуванні.

1.2. Прогнозування курсу криптовалют

Проблема прогнозування розвитку фондового або страхового ринків, на сьогодні є достатньо вивченою. Натомість прогнозування на ринку криптовалют є актуальною та цікавою задачею. Це пов'язано з тим, що ринок криптовалют характеризується високим ступенем волатильності, невизначеності та сильними коливаннями цін з часом, а самі криптовалюти знаходяться на перехідній стадії, через що є нестабільними. На сьогодні, криптовалюта — один з найновіших способів інвестування фінансів, криптовалютні торги являють собою постійно зростаючий і перспективний фінансовий ринок. Щоб отримати прибуток у торгівлі криптовалютою, ви повинні вміти правильно прогнозувати майбутні зміни цін.

Розглядаючи тему прогнозування криптовалют згадаємо Сословського В.Г. [43], який пропонує використовувати системний підхід у дослідженні криптовалют. Він визначає ринок криптовалют як відкриту та складну стохастичну систему. Вона знаходиться в активній фазі формування, взаємодіючи з існуючими економічними суб'єктами у зовнішньому середовищі, виступаючи компонентом фінансової системи та надалі впливаючи на її розвиток.

Рішення проблеми прогнозування курсу криптовалюти на ринку є, але зазвичай вони представлені у вигляді комерційних послуг, а деталі реалізації не розголошуються. Серед факторів, що впливають на курси криптовалют, дослідники зосередилися на поширеності самих цифрових валют. Останнім часом біткойн зазнав широкого просування у всіх ЗМІ по всьому світу, що безсумнівно сприяло зростанню курсу.

Також пост із згадкою відповідної криптовалюти в соціальній мережі Twitter може вплинути на обсяги торгів BTC наступного дня [41]. Крім того, [17] встановив, що існує вплив коментарів користувачів на онлайн-платформах на коливання цін криптовалют, а саме BTC особливо корелює з числом позитивних коментарів у соціальних мережах.

Як вже було зазначено вище, однією з головних проблем криптовалют є висока волатильність, але саме через це задача прогнозування їх курсу є дуже важливою. Розглядаючи більш детально криптовалюту Bitcoin (BTC),

можемо сказати, що ціни BTC демонструють нестационарну поведінку, де статистичний розподіл даних змінюється з часом. На малюнку 1.1 можна побачити динаміку курсу BTC за весь період:

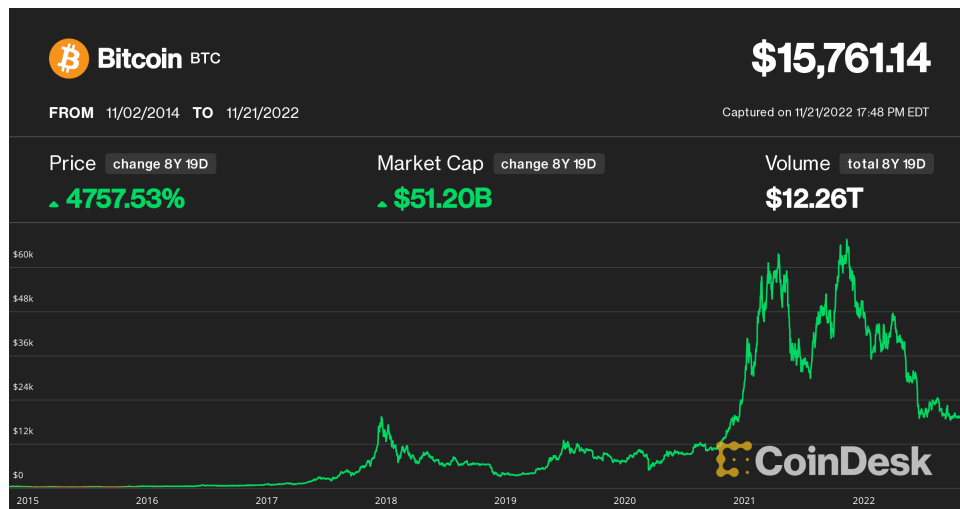


Рис. 1.1. Динаміка курсу Bitcoin за всю історію

Ціни на BTC в цей період були вкрай нестабільними. До 2013 року інтерес до біткоіну, його використання в віртуальних транзакціях і його ціни були низькими. А вже далі криптовалюти почали поширюватися, що викликало інтерес у людей, у результаті ціни збільшувалися. Але попри те, що ціни на BTC демонструють надзвичайну волатильність, BTC, як цифровий актив, досить стійкий, оскільки може відновити свою вартість після значних падінь, і навіть коли на ринку висока невизначеність, наприклад, під час пандемії COVID-19 [42].

1.3. Прогнозування часових рядів

Проблема прогнозування цін належить до аналізу часових рядів. Тому розглядаючи прогнозування курсу криптовалют з боку прогнозування часових рядів, зазначимо, що існують різні підходи, що базуються на моделях для прогнозування часових рядів відповідно до [14].

1.3.1. Метод використання інтегрованої моделі авторегресії

Інтегрована модель авторегресії — змінного середнього (Autoregressive Integrated Moving Average — ARIMA), мабуть, найпопулярніший метод, коли справа стосується прогнозування часових рядів. Вперше системний підхід до побудови моделей ARIMA опубліковано Г.Е.П. Боксом і Г.М. Дженкінсом у 1976 році [4]. Популярність використання цієї моделі пояснюється тим, що ARIMA здатна забезпечити подібну та іноді навіть кращу точність порівняно з іншими методами машинного навчання [?]. До того ж моделі ARIMA легше пояснити, зокрема, порівняно з альтернативами глибокого навчання.

У дослідженні [1] описуються результати використання ARIMA моделі для прогнозування цін на Bitcoin (BTC), Ripple (XRP) та Ethereum (ETH). Експеримент повторили для щоденних, щотижневих та щомісячних наборів даних шляхом застосування ARIMA з різними параметрами. На рисунку 1.2 узагальнено результати помилок на навчальній та тестовій вибірках.

Test.	Bitcoin			XRP			Ethereum		
	D.	W.	M.	D.	W.	M.	D.	W.	M.
MAE.	313.894	764.24	1937.78	0.041	0.125	0.254	12.949	31.249	94.05
MSE	294560	1389560	10012856	0.0097	0.0681	0.1822	410.01	1712	14818
RMSE	542.735	1178.8	3164.31	0.096	0.261	0.427	20.25	41.380	121.72

Рис. 1.2. Експериментальні результати для помилок MAE, MSE та RMSE

Дослідники помітили, що використання моделі ARIMA на щоденній повторній вибірці перевершила інші моделі з точки зору MAE, MSE та RMSE для прогнозування ціни на BTC, XRP та ETH. Модель ARIMA працювала краще на щотижневих наборах даних, ніж на щомісячних значеннях для BTC та XRP з точки зору MAE, MSE та RMSE, тоді як помилка тієї ж моделі в ETH гірша через менший розмір набору даних.

Перевагами моделей прогнозування ARIMA є їхня простота та прозорість моделювання, а також їх чітка математична та статистична основа. Однак самостійне створення моделі ARIMA вимагає багато ресурсів і часу. Недоліки включають відсутність адаптивності — коли надходять нові дані,

модель необхідно періодично переоцінювати та, можливо, повторно ідентифікувати. Такі моделі рекомендовані для короткострокового прогнозування, але можуть бути використані для середньострокового прогнозування, але не підходять для довгострокового прогнозування, що було доведено у статті [1].

1.3.2. Метод використання штучних нейронних мереж

Враховуючи специфіку ринку біткойнів, використання методів машинного навчання може забезпечити об'єктивні та обґрунтовані результати з огляду на їхню ефективність у подібних сферах. Низка досліджень показує збільшення використання штучних нейронних мереж (англ. Artificial Neural Network — ANN) для аналізу часових рядів і прогнозування, особливо у фінансовому секторі.

У роботі [12] показано, що ШНМ має кращі передбачувані властивості, ніж моделі часових рядів та інші алгоритми ML для прогнозування фінансової ефективності. Водночас останні дослідження показали, що моделі нейронних мереж є перспективними через неканонічний характер обробки інформації. Це забезпечує більш ефективну роботу всієї системи прогнозування. Якщо навчання пройшло успішно, нейронна мережа повертає прийнятні результати на основі даних, які були відсутні у навчальній вибірці, неповних та зашумлених даних [39]. Однією з головних переваг нейронних мереж перед традиційними алгоритмами — є їх здатність до навчання. Під час навчання нейронні мережі можуть розпізнавати складні залежності між вхідними даними та вихідними.

Застосування ШНМ для прогнозування ціни біткойна створює мережу прогнозування тенденцій. Метою цієї мережі є визначення ймовірностей трьох станів: ціна буде рости, падати або залишатися незмінною в наступні періоди [26].

Основні переваги нейронних мереж включають ефективність прогнозування, здатність відтворювати складні нелінійні залежності, адаптивність до надходження нових даних і здатність вивчати моделі. Недоліками є складність алгоритму навчання, ресурсомісткість процесу навчання, відсу-

тність доступу до проміжних результатів прогнозування, посттренінговий потенціал.

1.3.3. Метод використання рекурентної мережі з довгою короткостроковою пам'яттю

Одним з найперспективніших сучасних технологій машинного навчання є застосування глибинних нейронних мереж, в основі яких лежить застосування глибокого навчання (англ. Deep Learning).

Рекурентна нейронна мережа (Recurrent Neural Network, RNN) — мережа із циклів, які можуть зберігати інформацію. Така нейронна мережа здатна знаходити не тільки прості взаємозв'язки, а й взаємозв'язки між взаємозв'язками. У моделі RNN внутрішній шар служить для запам'ятовування інформації з попередніх спостережень. Головною проблемою є запам'ятовування невеликої кількості попередніх спостережень, що не підходить для довгих (фінансових, економічних) періодів.

Зупинімося докладніше на LSTM мережі (Long Short-Term Memory Units) — архітектурі рекурентних нейронних мереж, запропонованою 1997 року Зеппом Хохрайтером та Юргеном Шмідгубером [13].

В останні роки з'явилося багато досліджень щодо використання LSTM для прогнозування курсу криптовалют. Наприклад, у [31] порівнюються LSTM і ARIMA моделі для прогнозування протягом одного, семи, тридцяти та дев'яноста днів. З результатів видно, що моделі машинного навчання (LSTM) потребують набагато більше часу для прогнозування через складні розрахунки, ніж традиційні моделі. Час компіляції моделі LSTM становить 61 мілісекунду, а для моделі ARIMA — 4 мілісекунди. Але розрахунки помилки RMSE показали, що модель машинного навчання LSTM перевершує традиційну модель ARIMA. Дослідження фокусується лише на ціні біткоїна для розробки прогнозної моделі та не бере до уваги інші економічні фактори, такі як новини про біткоїни, державну політику та ринкові настрої, з якими можна набагато точніше передбачити ціну. Отже, доведено, що моделі, розроблені з використанням LSTM, мають більшу точність, ніж традиційні моделі, тому вони є ефективними у питанні прогнозування біткоїнів.

1.3.4. Метод використання згорткової нейронної мережі

Здатність згорткових нейронних мереж (англ. Convolutional Neural Network — CNN) вивчати й автоматично отримувати ознаки з необроблених вхідних даних може застосовуватися до завдань прогнозування часових рядів. Послідовність спостережень можна розглядати як одновимірне зображення, яке модель CNN може зчитувати і розкласти на основні елементи [5]. Порівняно з іншими нейронними мережами, літератури з прогнозування часових рядів за допомогою згорткових нейромереж справді мало, оскільки ці типи мережі набагато частіше застосовуються в проблемах класифікації.

Розглянемо останню роботу, в яких дослідники займалися питанням використання CNN для прогнозування на фондовому ринку та ринку криптовалют. У [38] порівнюються архітектури LSTM, RNN та CNN для короткострокового прогнозування майбутніх цін на акції. У цьому підході архітектура CNN перевершує LSTM і RNN. Це пов'язано з раптовими змінами, що відбуваються на фондових ринках. Ці зміни можуть не завжди мати регулярний характер або не завжди йти за одним і тим же циклом. Залежно від компаній й секторів існування тенденцій, період їх існування будуть відрізнятися. Аналіз цих типів тенденцій і циклів дасть більший прибуток інвесторам. Тому дослідники вважають, що для аналізу такої інформації повинні використовувати мережі, такі як CNN, оскільки вони використовують поточну інформацію, надану в конкретний момент та здатні фіксувати різкі зміни в системі.

Головною перевагою згорткових нейронних мереж є те, що використовуються згорткові шари, щоб виявити ознаки мережі, що дозволяє тренувати нейронну мережу без складної попередньої обробки, оскільки корисні функції будуть вивчені під час навчання. Разом з тим вони потребують великої кількості даних, що у результаті дає точний прогноз.

1.4. Використання вейвлет-аналізу у фінансах

На сьогодні вейвлети є досить новим підходом для дослідження різноманітних систем, але уже широко застосовуються в астрономії (при аналізі супутникових зображень планет), в медицині (аналіз кардіограм, рентгенограм, а також райдужної оболонки ока), у згортці великих обсягів інформації (наприклад, відбитків пальців), для розпізнавання образів.

Зокрема, набуває популярності використання вейвлет-технології в бізнесі та фінансах. Останні кілька десятиліть були епохою великих даних, особливо для фінансового сектора. Це пояснюється тим, що багато фінансових змінних, таких як ціни на акції, зараз дуже часто можна виміряти «за хвилину чи секунду». Із цією підвищеною складністю виникає потреба в інструментах обробки даних як ніколи.

Вейвлет-аналіз — це математична процедура, яка дозволяє нам зменшувати дані на різних частотних компонентах і досліджувати кожен компонент із роздільною здатністю, залежно від його масштабу. Крім того, це різновид лінійного перетворення, яке має нести інформацію про конкретні локальні координати, в яких з'являються його характерні частоти.

Основна сфера застосування вейвлет-аналізу — обробка різного роду нестационарних (у часі) або неоднорідних (у просторі) сигналів. Також вейвлет-аналіз використовують для видалення шуму із сигналу, оскільки саме зашумленість є великою проблемою у ході дослідження фінансових часових рядів, насамперед динаміки курсу.

1.4.1. Огляд підходів

Найперша робота, пов'язана з вейвлет-аналізом, належить угорському математику Альфреду Хаару і була написана у 1910 році [10]. Вейвлет Хаара — один з найпростіших вейвлетів, що заснований на ортогональній системі функцій. З часу появи роботи Хаара було зроблено багато інших важливих внесків у галузі вейвлет-аналізу. Наприклад, у 1976 році було відкрито безперервне вейвлет-перетворення (CWT) [50], а з часом бельгійська науковиця з фізики та математики Інгрід Добеші продовжила розвивати

теорію ортогональних вейвлетів, але вже використовувала функції, які обчислюють ітераційним шляхом [6].

Починаючи з 2000-х років, вейвлет-декомпозиція поєднується з моделями часових рядів як метод попередньої обробки. Вейвлет-декомпозиція (або вейвлет-перетворення) розкладає дані часових рядів на компоненти наближення та деталізації (approximation and details), щоб до кожного компонента можна було застосувати різні моделі прогнозування. Ця властивість може підвищити ефективність прогнозування. Справедливість цього підходу була доведена різними дослідженнями. Прикладом слугує наукова робота Юсефі та ін. [49], в якій використовувався вейвлет-рдекомпозиція для перевірки ефективності ф'ючерсних ринків. Ф'ючерси NYMEX від одного до чотирьох місяців порівнювали з цінами спотового ринку West Texas Intermediate, прогнозованими моделями, що використовують вейвлет-розклад і моделі числового аналізу. Порівняння отриманих коефіцієнтів кореляції показало вищу кореляцію прогнозованих значень, ніж ф'ючерсних цін.

За даними Тана та ін. [46], результати прогнозування ціни на електроенергію значно покращилися, коли вони побудували модель, використовуючи вейвлет-декомпозицію, а також моделі ARIMA та GARCH. У статті [33] використовували DWT у поєднанні з XGBoost та аналіз головних компонентів (PCA), щоб проаналізувати та спрогнозувати тенденції фондового ринку до ефективної покупки або продажу акцій. Порівняно з деякими найновішими методами прогнозування цін, запропонований метод показав значне покращення точності прогнозування. Дослідження [36] запропонувало модель прогнозування, засновану на DWT і алгоритмі регресії опорного вектора (SVR) для прогнозування цін на золото. У дослідженні використовувалося дискретне вейвлет-перетворення з максимальним перекриттям (MODWT) і фільтр Хаара. Результати підтвердили перевагу моделі прогнозування DWT-SVR над Random Forest і лінійною регресією (OLS) як за точністю регресії, так і за економічною прибутковістю. Крім того, використовуючи DWT з фільтрами Добеші (db) різної довжини, а саме фільтри db1, db2, db3 і db4, і модель прогнозування на основі SVR [16], було спрогнозовано чотири індекси фондового ринку. Вони використовували сплайни багатовимірної адаптивної регресії (MARS) для вибору значущих підсерій DWT і показали, що їхній підхід до декомпозиції, відбору та прогно-

зування перевершує інші моделі, включаючи ARIMA та адаптивну мережу на основі системи нечіткого виводу (ANFIS) відповідно до статистичної точності.

Автори статті [19] у своїй роботі описують аналіз декомпозиції фінансових часових рядів індексів BSE та NSE за допомогою дискретного вейвлет-перетворення максимального перекриття (MODWT), де вони дійшли висновку, що аналіз декомпозиції шкали часу на основі MODWT дає кращі результати, ніж спектральний аналіз на основі перетворення Фур'є. А в роботі [45] представлено новий алгоритм усунення шумів на основі вейвлетів (а саме MODWT) для високочастотного аналізу фінансових даних.

Розглянемо більш детально випадки, коли вейвлет-декомпозицію комбінували зі штучними нейронними мережами. Наприклад, Моаззамі та ін. [24] застосували вейвлет-декомпозицію та штучні нейронні мережі до сезонних даних для національної електромережі Ірану, щоб спрогнозувати пікове навантаження на день вперед. Використовувалися три типи алгоритмів штучних нейромереж. Модель, яка поєднувала нейромережу прямого розповсюдження з вейвлет-декомпозицією, дала найкращі результати. Шафієха [40] застосував вейвлет-перетворення, ARIMA та нейронні мережі радіально базисних функцій (RBFN) для прогнозування ціни на електроенергію в Іспанії, розглядаючи поведінку ціни як нелінійну функцію. Він розклав дані за допомогою вейвлет-перетворення, і ARIMA розпізнала чотири розкладені ряди. Після застосування ARIMA було виконано зворотне вейвлет-перетворення, а потім мережа RBFN використовувалася для виправлення помилок предиктора вейвлет-ARIMA. Інший дослідник, який використовував вейвлети, об'єднав їх із RNN і методом ройового інтелекту (штучної бджолиної колонії (ABC)), щоб передбачити ціну акцій і максимізувати прибуток. DWT використовувався для зменшення шуму в кожній функції, що подається в модель [15].

1.4.2. Вейвлети в прогнозуванні криптовалют

Вейвлети локалізують процес у часі та частоті. Це робиться за допомогою функції вейвлет-перетворення. Його роль полягає у захопленні

низько- та високочастотних характеристик у часовому ряді на послідовних рівнях розкладання. Коли вейвлет-перетворення має високу частоту, воно фіксує детальні події. Коли він має низьку частоту, він фіксує тривалі події, що робить його ідеальним для аналізу нестационарних часових рядів.

Розглянемо та зробимо аналіз останніх наукових робіт, в яких досліджувалося використання вейвлетів саме для прогнозування курсу криптовалют.

У цій сфері застосовувалося емпіричне перетворення вейвлетів (EWT) [2], вперше запропоноване у 2013 році [7], яке полягає у вилученні різних внутрішніх режимів часового ряду шляхом створення адаптивних вейвлетів. Алгоритм поєднав EWT з LSTM для прогнозування вартості криптовалюти та Cuckoo Search (CS) для оптимізації навчання LSTM. Ця конкретна робота показує, що вейвлет-декомпозиція в емпіричних вейвлетах призвела до значного зменшення величин помилок прогнозу. У роботі [44] представлено швидке вейвлет-перетворення, яке полягає у використанні вейвлетів Добеші, а потім використання зворотного перетворення для прогнозування цін криптовалюти.

Автори статті [34] досліджували кілька глобальних ринків у пошуках предиктора прибутку BTC. У зв'язку з цим, запропонували часові ряди цін на золото, нафту, SP500, ETH, XRP і BTC, а також часові ряди VIX і USDI як потенційні предиктори для дослідження їх прогностичної сили в прогнозуванні ціни BTC на крок вперед. По-перше, було застосовано DWT з фільтром db3, щоб усунути шум часового ряду. Далі була створена модель прогнозування на основі нейронних мереж LSTM і спрогнозовано BTC на крок вперед, використовуючи обидва підходи однофакторної регресії та багатофакторної регресії.

Вищезазначені дослідження показують, що вейвлети є ефективним методом усунення шумів і декомпозиції даних, а їх застосування у фінансових даних різко підвищило точність і зменшило величину помилок прогнозу навчених моделей.

1.5. Висновки до розділу

Висвітлено загальні відомості про сферу криптовалют та їх прогнозування. Показана актуальність даної задачі, проведено огляд поточного стану проблемної області й огляд існуючих рішень. Крім того, був розглянутий новий підхід до прогнозування з використанням вейвлет-аналізу.

Аналіз застосування традиційних моделей прогнозування показує, що вони спираються на припущення щодо лінійного характеру залежності між показниками та пристосовані для даних, в яких можна виділити тенденції, сезонність та шум, а криптовалюти характеризуються низькою сезонністю. Крім того, оскільки часові ряди криптовалют характеризуються високим рівнем нелінійності й нестационарності, ці методи не дуже ефективні для прогнозування цін криптовалют. За результатами порівняння методів прогнозування курсу криптовалют були виділені їх переваги та недоліки.

Таким чином, перспективним і потужним підходом до прогнозування стають методи обчислювального інтелекту, до числа яких, в першу чергу, слід віднести нейронні мережі. Здійснений аналіз дає змогу підтвердити перспективність та доцільність їх використання для прогнозування. Разом з тим, набуває популярності використання вейвлет-аналізу для покращення вхідних даних, тобто з метою усунення шумів і декомпозиції даних, що у комбінуванні з нейронними мережами значно підвищує точність і зменшує величину помилок прогнозу.

РОЗДІЛ 2

НЕЙРОННІ МЕРЕЖІ ТА ВВЕЙВЛЕТНИЙ АНАЛІЗ

2.1. Визначення та характеристики часового ряду

Означення 2.1. [8] Часовий ряд — це ряд спостережень за значеннями $y_1, y_2, \dots, y_T$ деякого показника (ознаки), упорядкований у хронологічній послідовності, тобто у порядку зростання змінного t -часового параметра. Будемо позначати часовий ряд з T елементами Y_T .

Окремі спостереження часового ряду називають його рівнями. Якщо кожному моменту часу відповідають значення лише одного показника, тобто, якщо ряд містить ряд містить записи однієї змінної, він є одновимірним. В іншому випадку, коли розглядаються розглядати записи більш ніж однієї змінної — ряд є багатовимірним.

Часові ряди широко застосовуються в різних галузях, таких як статистика, розпізнавання зображень, обробка сигналів, прогнозування погоди, економетрика, прогнозування землетрусів, електроенцефалограма, автоматичне керування, астрономія та комунікаційні технології, та в будь-якій області прикладної науки, яка включає часові вимірювання.

Компоненти часового ряду

Вважається, що на весь часовий ряд впливають чотири основні компоненти: тенденція, сезонність, циклічність і випадкова складова.

- Тренд — системна складова в довгостроковому розвитку досліджуваного показника. Трендом називають гладку зміну, яка проявляється протягом тривалого періоду часу і є результатом дії довготривалих факторів.
- Сезонна компонента характеризує циклічні коливання досліджуваного показника навколо трендової компоненти протягом року, протягом

сезону.

- Циклічна компонента описує середньострокові зміни, спричинені повторюваними випадками.
- Нерегулярна компонента — випадкова, яка може формуватися як під впливом непередбачених природних катаклізмів (землетруси, пожежі, епідемії), так і поточних чинників — помилки вимірювання.

Через наявність випадкової складової неможливо безпомилково передбачити значення часового ряду. Але будь-який фактичний процес містить випадкову складову.

Таким чином, модель часового ряду можна описати як

$$y_t = u_t + s_t + c_t + \varepsilon_t \quad (2.1)$$

— адитивна модель (2.1) чи

$$y_t = u_t \cdot s_t \cdot c_t \cdot \varepsilon_t \quad (2.2)$$

— мультиплікативна модель (2.2),

де y_t — спостереження; u_t — тренд; s_t — сезонність; c_t — циклічність; ε_t — випадкова компонента.

Слід зазначити, що не завжди можна описати кожен компонент окремо. Іноді краще робити прогнози щодо всієї моделі, ніж намагатися відокремити кожен компонент.

2.2. Прогнозування часових рядів

В результаті появи нових технологій, зросла довіра до методів, що включають складну техніку обробки даних. Комп'ютери, в сукупності з кількісними методами розрахунків, які завдяки їм стали загальнодоступними, для сучасних організацій є вже не просто зручним інструментом, а фактично їх невід'ємною частиною.

Метою прогнозування є аналіз того, що може відбутися при реалізації можливих прогнозів у майбутньому та до яких наслідків це призведе.

Прогнозування відіграє важливу роль у багатьох сферах, тому сучасні організації потребують короткострокових, середньострокових, довгострокових прогнозів, залежно від конкретного застосування.

Розглянемо, в яких випадках використовуються ті чи інші прогнози на ринку криптовалют:

- Короткострокові.

Вони складаються на один-три дні й містять максимально точну інформацію про найближчі зміни вартості електронної валюти. Такі прогнози складаються для трейдерів, які щодня проводять валютні операції, з метою своєчасного повідомлення про можливі коливання курсу.

- Середньострокові.
- Довгострокові.

Такі прогнози дозволяють дізнатися загальну динаміку зміни курсу електронної валюти на найближчий місяць. Їх головна мета — своєчасно попередити про можливу появу факторів, що ведуть до тривалого зниження вартості електронної валюти.

Метрики оцінки точності прогнозів

Перед тим як використовувати модель для складання реальних прогнозів, важливо оцінити точність прогнозу. Це можна визначити лише з урахуванням того, наскільки добре працює модель з новими даними, які не використовувались для навчання моделі. Тобто наявні дані розділяють на дві частини: навчальний та тестовий набори, де останній повинен надати вірогідну вказівку на те, наскільки вдало модель прогнозує нові дані.

Зважаючи на фундаментальну важливість прогнозування часових рядів у багатьох реальних ситуаціях, слід бути дуже обережним при виборі конкретної моделі. Внаслідок чого, запропоновано різні показники ефективності для оцінки прогнозів і порівняння моделей.

Основним показником міри точності прогнозу є його помилка (2.3) —

різниця між прогнозним і фактичним значенням досліджуваної змінної:

$$e_t = y_t - \hat{y}_t, \quad (2.3)$$

де y_t — фактичне значення змінної для t -го виміру; $\hat{y}_t$ — прогнозне значення досліджуваної змінної.

Для метрики оцінки точності прогнозів частіше використовуються наступні показники: R^2 , MSE, RMSE, MAPE, MAE, що передбачає їх розрахунок за n кроків:

- коефіцієнт детермінації (2.4):

$$R^2 = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{\hat{y}})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (2.4)$$

де $\bar{\hat{y}}$ — середнє арифметичне оцінених значень y за моделлю, y — середнє арифметичне часового ряду. Коефіцієнт детермінації змінюється від 0 до 1. Чим ближче R^2 до 1, тим точніша модель.

- середня квадратична похибка (2.5):

$$MSE = \frac{1}{n} \sum_t (y_t - \hat{y}_t)^2 \quad (2.5)$$

MSE чутливо реагує на зміну масштабу і трансформацію даних.

- корінь із середньоквадратичної похибки (2.6):

$$RMSE = \sqrt{\frac{1}{n} \sum_t (y_t - \hat{y}_t)^2} \quad (2.6)$$

Оскільки RMSE — це квадратний корінь обчисленого MSE, то властивості MSE зберігаються і для RMSE.

- середня абсолютна похибка (2.7):

$$MAE = \frac{1}{n} \sum_t |y_t - \hat{y}_t| \quad (2.7)$$

MAE залежить від масштабу вимірювань і перетворення даних, а також

отримане значення MAE має бути якомога меншим для хорошого прогнозу.

- середня відносна похибка (2.8):

$$MAPE = \frac{100}{n} \sum_t \left(\frac{y_t - \hat{y}_t}{y_t} \right)^2 \quad (2.8)$$

MAPE не залежить від масштабу вимірювання, але на нього впливають перетворення даних.

MAPE	Точність прогнозу
менше 10%	Висока
10% - 20%	Добра
20% - 40%	Задовільна
40% - 50%	Погана
більше 50%	Незадовільна

Розглянемо ще одну метрику — показник подібності динамічного викривлення часу (англ. Dynamic Time Warping, DTW). DTW — це міра подібності між часовими рядами, яка була введена незалежно в літературі [48] і [37], в обох випадках для мовних додатків.

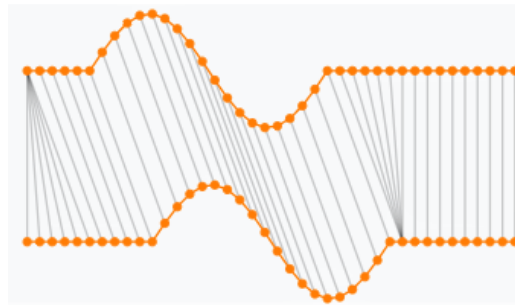


Рис. 2.1. Dynamic Time Warping

DTW між x і y формулюється як така задача оптимізації (2.9):

$$DTW(x, y) = \min_{\pi} \sqrt{\sum_{(i, j) \in \pi} d(x_i, y_j)^2} \quad (2.9)$$

де $\pi = [\pi_0, \dots, \pi_K]$ — шлях, який задовольняє такі властивості:

- 1) це список пар індексів $\pi_k = (i_k, j_k)$ з $0 \leq i_k < n$ і $0 \leq j_k < m$;
- 2) $\pi_0 = (0, 0)$ та $\pi_K = (n - 1, m - 1)$;

3) для всіх $k > 0$, $\pi_k = (i_k, j_k)$ пов'язане з $\pi_{k-1} = (i_{k-1}, j_{k-1})$ таким чином (2.10):

$$\begin{aligned} i_{k-1} &\leq i_k \leq i_{k-1} + 1; \\ j_{k-1} &\leq j_k \leq j_{k-1} + 1 \end{aligned} \quad (2.10)$$

Тут шлях можна розглядати як часове вирівнювання часових рядів таким чином, що евклідова відстань між вирівняними (тобто перевибраними) часовими рядами є мінімальною.

Зазначимо, що кожен з цих методів має унікальні властивості, відмінні від інших, тому в експериментах краще розглядати більш ніж один критерій ефективності для винесення суджень.

2.3. Штучна нейронна мережа

Глибинне навчання (англ. Deep Learning, DL) належить до потужних алгоритмів машинного навчання, які спеціалізуються на вирішенні нелінійних та складних проблем. Прогнозування ціни на криптовалюту за своєю природою є такою проблемою, оскільки її значення з часом мають дуже великі коливання внаслідок майже хаотичної та непередбачуваної поведінки. Тому методи глибинного навчання можуть становити належну методологію розв'язання цієї проблеми.

Штучні нейрони — це прототипи біологічних нейронів, які визначають вузли штучної нейронної мережі, ШНМ. Нейрон приймає певну множину сигналів і генерує з них вихідний сигнал, який потім подає або на вихід, або на вхід наступних нейронів. У другому варіанті коли нейрони між собою з'єднані, вони утворюють ШНМ. Першою спробою змодельовати ці процеси були нейронні мережі У.Маккалок і У.Піттса [25].

ШНМ — система з'єднаних і взаємодіючих між собою штучних нейронів. Кожен процесор такої мережі обробляє лише сигнали, які він періодично отримує, і сигнали, які він надсилає іншим процесорам. На рис. 2.2 представлено структуру штучного нейрона.

Модель роботи нейрона має такий алгоритм. Сигнали x_i , надходять на

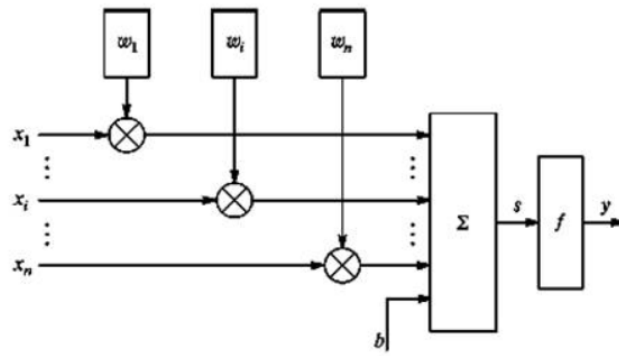


Рис. 2.2. Структура штучної нейронної мережі

вхід, де реалізується функція синапсів, де кожен має свій ваговий коефіцієнт w_i . Після цього сигнали стають зваженими й надходять до лінійного суматора (сигнали додаються), далі переходить до блоку активаційної функції f . Потім відбувається відповідна обробка і тоді параметр подається на вихід у вигляді сигналу y . Зазвичай функція активації обмежує сигнал нейрона у певному діапазоні, $[0; 1]$ або $[-1; 1]$. Також сигнал нейрона змінюється від початкового зрушення b , який подається на вхідний сигнал блоку функції активації.

Математично модель нейрона можна описати так (2.11):

$$y = f \left(\sum_{i=1}^n w_i x_i + b \right) \quad (2.11)$$

де y — вихідний сигнал нейрона; f — активаційна функція нейрона; x_i — вхідні сигнали нейрона; w_i — синаптична вага нейрона; b — зміщення нейрона.

Функція активації показує залежність вихідного сигналу нейрона від суми зважених вхідних сигналів. Є багато видів функцій активації: порогова бінарна функція, лінійна обмежена функція, гіперболічний тангенс, сигмоїдальна, випрямлена лінійна (Rectified linear unit, ReLU).

Важлива особливість ШНМ стосовно застосування до проблем прогнозування часових рядів полягає в здатності нейронних мереж до нелінійного моделювання, без будь-якого припущення про статистичний розподіл часового ряду. Кожна модель адаптивно навчається на основі даних, тому мережа має самоадаптивний характер. Але існує складність вибору алгоритму навчання та вибору архітектури.

2.4. Рекурентна нейронна мережа

Рекурентна нейронна мережа (англ. Recurrent Neural Network, RNN) — мережа із циклів, які можуть зберігати інформацію. У рекурентній нейронній мережі є деякі входи, декілька прихованих шарів і декілька виходів. Це зображено на рис. 2.3:

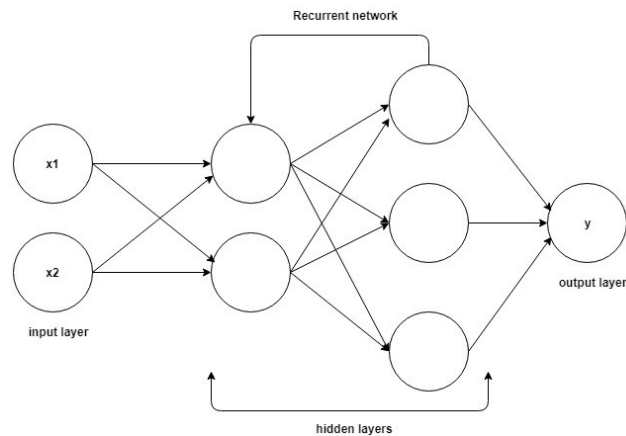


Рис. 2.3. Базова схема рекурентної нейронної мережі

Сигнал з вихідних нейронів або нейронів схованого шару частково передається назад на входи нейронів вхідного шару. Як будь-яка система, що має зворотний зв'язок, рекурентна мережа прагне до стійкого стану. Така нейронна мережа здатна знаходити не тільки прості взаємозв'язки, а й взаємозв'язки між взаємозв'язками. У моделі RNN внутрішній шар служить для запам'ятовування інформації з попередніх спостережень. Головною проблемою є запам'ятовування невеликої кількості попередніх спостережень, що не підходить для довгих (фінансових, економічних) періодів.

Модель довгої короткострокової пам'яті

LSTM мережі (Long Short-Term Memory Units) — архітектура рекурентних нейронних мереж, запропонована 1997 року Зеппом Хохрайтером та Юргеном Шмідгубером [13].

Метою розробки модулів LSTM було формування більш стабільної мережевої пам'яті, що значно спрощує пошук довгих контекстних залежностей. Спільним для всіх рекурентних мереж є те, що вони складаються з

ланцюга повторюваних блоків. Зазвичай цей блок має дуже просту структуру. Структура LSTM схожа на ланцюг із модулями, що містять чотири шари, і ці шари взаємодіють особливим чином. Цю комірку зображено на рис. 2.4:

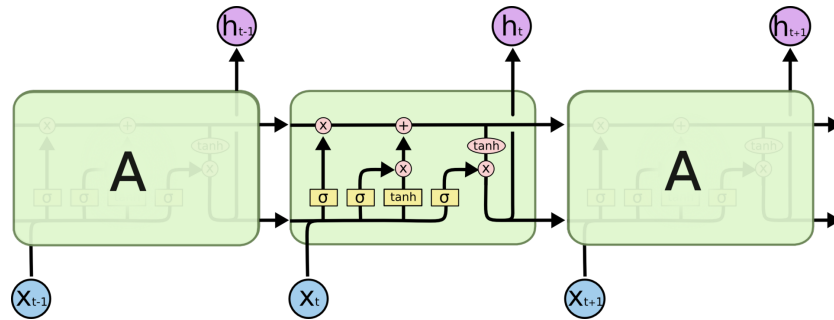


Рис. 2.4. Архітектура LSTM

Основним компонентом LSTM є стан комірки, горизонтальна лінія, яка проходить уздовж верхніх ділянок схеми. Інформацію, яку він містить, можна зберігати в незмінному вигляді, записувати або читати, подібно до даних у пам'яті комп'ютера. Модуль LSTM потенційно дуже корисний для роботи з рядками з довгостроковими залежностями: можливість додавати та видаляти інформацію про стан комірки. Водночас цей процес регулюється "фільтрами" — регуляторами, що відповідають за процес забування або збереження інформації. Комірка пам'яті LSTM зберігає значення протягом тривалого та короткого періодів часу. Це досягається завдяки функції активації кожної комірки пам'яті.

Комірка визначає, що вона зберігає та коли її можна читати, писати та стерти. На відміну від комп'ютерної цифрової пам'яті, ці гейти є аналоговими і реалізовані шляхом елементарного множення на сигмоїди, в діапазоні від 0 до 1. Ці ворота діють на отриманий сигнал, блокуючи або передаючи інформацію, яка підлягає фільтрації, за допомогою власних ваг, подібно до вузлів нейронної мережі. Іншими словами, комірка дізнається, коли вводити, виходити чи видаляти дані, через ітеративний процес створення припущень, зворотного поширення помилок і коригування ваг за допомогою градієнтного спуску.

2.5. Згорткова нейронна мережа

Згорткова нейронна мережа — спеціальна архітектура нейронних мереж, від самого початку націлена на ефективне розпізнавання зображень. Ця мережа за основу схему з'єднання нейронів зорової кори тварин [23]. Згорткова нейронна мережа зазвичай складається з чергування згорткових шарів (convolution layers, C-layers), шарів субдискретизації (subsampling layers, S-layers) і за наявності повнозв'язних шарів (fully-connected layer, F-layers) на виході. Всі три види шарів можуть чергуватися в довільному порядку (рис. 2.5).

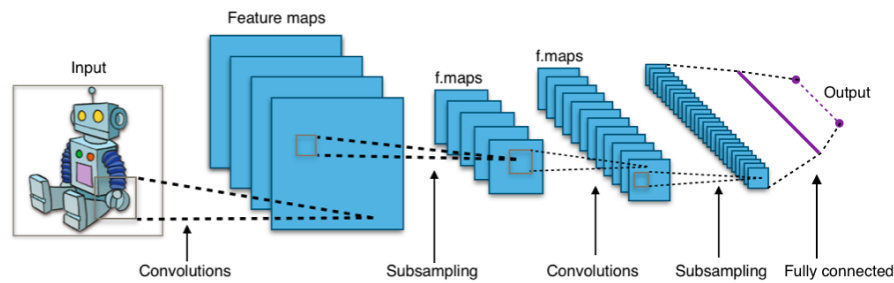


Рис. 2.5. Архітектура згорткової нейронної мережі

Згортка — операція над парою матриць A і B , результатом якої є матриця $C = A \cdot B$. Кожен елемент обчислюється як скалярний добуток матриці B і деякої підматриці A такого ж розміру. Операція згортки (рис. 2.6) визначається таким виразом (2.12):

$$C_{i,j} = \sum_{u=0}^{m_x-1} \sum_{v=0}^{m_y-1} A_{i+u,j+v} B_{u,v} \quad (2.12)$$

де $B_{u,v}$ — значення елемента ядра згортки (u, v) , $C_{i,j}$ — значення пікселя зображення, що отримаємо, $A_{i+u,j+v}$ — значення пікселя вхідного зображення, $m_x - 1$, $m_y - 1$ — розмір ядра згортки.

Шари субдискретизації виконують зменшення розмірності, найчастіше для цього використовується метод вибору максимального елемента (max-pooling). Вся матриця ознак поділяється на квадрати, з яких вибираються найбільші за значенням елементи (рис. 2.7).

Основною перевагою згорткових нейронних мереж є використання згорткових шарів для виявлення характеристик мережі. Це дозволяє нам

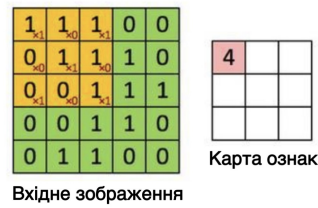


Рис. 2.6. Алгоритм згортки

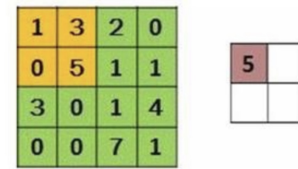


Рис. 2.7. Алгоритм субдискретизації

навчати нейронні мережі без складної попередньої обробки, оскільки корисні функції вивчаються під час навчання. Згорткові нейронні мережі можна використовувати для обробки складних нелінійних моделей. Однак це вимагає великої кількості даних, які ведуть до точних прогнозів.

2.6. Вейвлет-аналіз

В останні роки було виявлено, що звичайний метод представлення довільних функцій і сигналів рядами Фур'є неефективний для функцій з локальними особливостями, особливо для імпульсних і цифрових сигналів. Це пов'язано з тим, що базисна функція рядів Фур'є визначена від $-\infty$ до $+\infty$ і за своєю природою є гладкою і періодичною. Унаслідок чого поступово набуває поширення використання вейвлетів в економічній та фінансовій сферах.

Цей тип перетворення сигналу використовується, коли аналіз конкретного сигналу (фінансового часового ряду) повинен містити інформацію не тільки про його характерних частотах, але і про конкретні локальні координати, в яких ці частоти знаходяться.

З основних переваг вейвлет-аналізу над традиційними інструментами аналізу даних можемо назвати наступні:

- 1) є гнучким і не вимагає серйозних припущень щодо процесу генерування даних, тобто має здатність представляти дуже складні дані без необхідності знати їхню функціональну форму;
- 2) не вимагає припущення стаціонарності даних, а більшість економічних і фінансових часових рядів не є стаціонарними;
- 3) надає інформацію як у часовій, так і в частотній областях.

Назва вейвлет походить від англійського слова "wavelet" що перекладається як "маленька хвиля", або "хвилі, що йдуть одна за одною". Це короточасне хвилеподібне коливання, яке починається з нульової амплітуди, зростає та повертається до нуля. Вейвлети створюються за допомогою спеціальних базових функцій (прототипів), які визначають вигляд і властивості, які задовольняють багатьом конкретним умовам. Вейвлет-декомпозиція поєднує всі функції, зміщуючи та розтягуючи їх уздовж осі часу.

Варто зазначити, що вейвлет-перетворення поділяють на:

- дискретне вейвлет-перетворення (DWT),
- вейвлет-перетворення пакетів (WPT),
- безперервне вейвлет-перетворення (CWT).

Різниця між ними полягає в тому, як реалізуються ці перетворення. Наприклад, DWT можна реалізувати шляхом поділу сигналу коефіцієнти наближення (низькочастотні) та деталізації (високочастотні). Після першого рівня, апроксимація розкладається на вищий рівень. Для WPT, після першого рівня деталізація та наближення розкладаються на наступний рівень. А CWT працює у будь-якому масштабі. Під час обчислення вейвлет масштабується та зміщується в межах області аналізованого сигналу [21].

Основними перевагами вейвлет-перетворення є:

- 1) великий набір базисних функцій;
- 2) гарна локалізація за частотою та часом;
- 3) доступні функції різного ступеня гладкості.

До недоліків можна віднести можливість спотворення даних, наприклад, підвищення рівня декомпозиції може призвести до значного згладжування даних і втрати інформації [20]. До того ж порівняти результати вейвлет-аналізу, отримані під час роботи в різних, досить складно.

У роботі було використано дискретне вейвлет-перетворення та його модифікація, тому у наступному розділі більш детально розглянемо цей тип.

2.7. Дискретне вейвлет-перетворення

Дискретне вейвлет-перетворення (ДВП) (англ. Discrete wavelet transform, DWT) — це вейвлет-перетворення, у якому останні представлені як дискретний сигнал (вибірок фрагментів).

В загальному вигляді вейвлет-перетворення функції $f(t)$ визначається виразами (2.13, 2.14):

$$W_f(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}^*(t)dt \quad (2.13)$$

де $\psi_{a,b}^*(t)$ — комплексно-спряжені функції сімейства вейвлет-функцій $\psi_{a,b}(t)$ виведених з материнського вейвлета $\psi(t)$:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{2}}\psi\left(\frac{t-b}{a}\right) \quad (2.14)$$

де a — параметр розтягування (масштабу), а b — параметр зсуву.

Ідея вейвлет-перетворень полягає в тому, що саме перетворення дозволяє лише зміни на осі часу(t), але форма функції залишається незмінною. Це залежить від базисних функцій, яких існує велика кількість, що відрізняються гладкістю та іншими показниками.

Зупинимось докладніше на деяких вейвлетах, що використовуються за материнські у дискретному вейвлет-перетворенні. Слід сказати, що термін "материнський" означає, що функції з різною шириною носія, що використовуються у перетворенні, генеруються однією базовою функцією — материнським вейвлетом. Тобто материнський вейвлет є прототипом для всіх віконних функцій [18].

2.7.1. Вейвлет Хаара

Це система базисних ортогональних функцій, що володіє наступними властивостями: має компактний носій, локальну область визначення, але не є гладкими.

Функція перетворення — материнський вейвлет $\psi(x)$, має нульове зна-

чення інтеграла, що визначає особливості сигналу, задається виразом (2.15):

$$\psi(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2} \\ -1, & \frac{1}{2} \leq x < 1 \\ 0, & x \notin [0,1) \end{cases} \quad (2.15)$$

а функція масштабування $\phi(x)$ має одиничне значення інтеграла та визначає грубе наближення (апроксимацію) сигналу (2.16):

$$\phi(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & x \notin [0,1) \end{cases} \quad (2.16)$$

На рис. 2.8 можемо бачити графік функції материнського вейвлету:

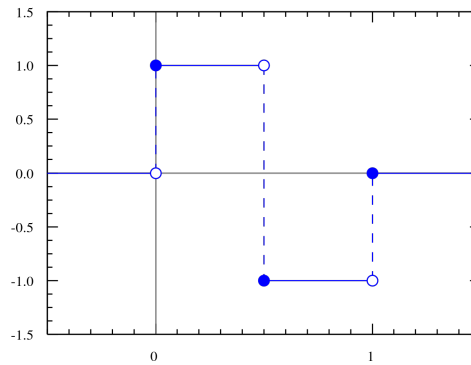


Рис. 2.8. Графік функції материнського вейвлету

2.7.2. Вейвлет Добеші

Вейвлет Добеші (див. рис. 2.9) обчислюється ітераційно, і його форма залежить від ступеня полінома та кількості обчислених коефіцієнтів.

Вейвлети цього типу характеризуються двома функціями: вейвлет-функцією ψ і масштабуючою функцією ϕ , які задаються наступними рівняннями (2.17, 2.18):

$$\varphi(t) = \sqrt{2} \sum_k h_k \varphi(2t - k) \quad (2.17)$$

$$\psi(t) = \sqrt{2} \sum_k g_k \varphi(2t - k) \quad (2.18)$$

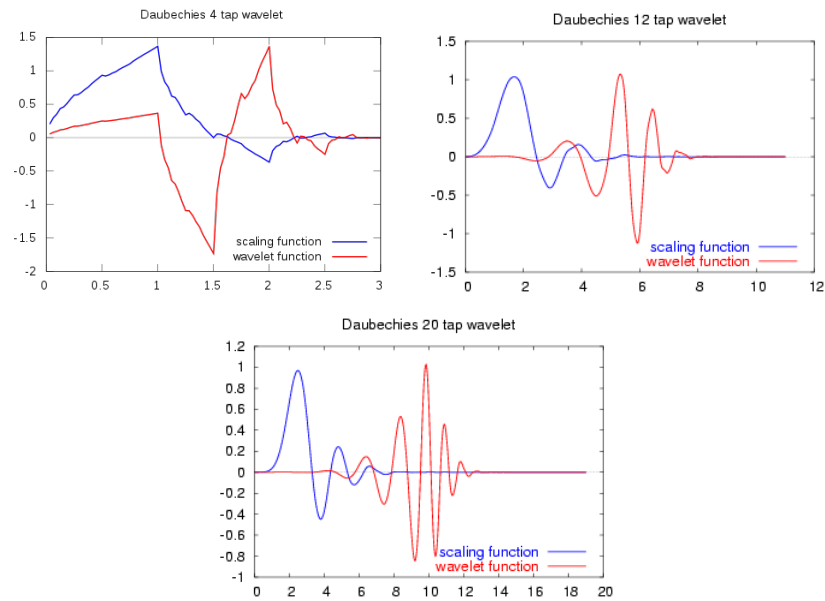


Рис. 2.9. Графік вейвлету Добеші та масштабуючої функції

Підсумуємо основні властивості вейвлетів Добеші:

- 1) є ортогональними, але не симетричними та мають кінцеву кількість фільтруючих коефіцієнтів;
- 2) добре локалізовані в частотній області;
- 3) здатні виконувати перетворення за допомогою алгоритму швидкого вейвлет-перетворення;
- 4) мають можливість реконструкції сигналів.

2.8. Дискретне вейвлет-перетворення максимального перекриття

Дискретне вейвлет-перетворення максимального перекриття (англ. Maximal Overlap Discrete Wavelet Transform, MODWT) — один з найсучасніших методів попередньої обробки фінансових часових рядів [35], модифікація звичайного дискретного вейвлет-перетворення.

У звичайного ДВП є деякі обмеження, а саме: вимога, щоб довжина набору даних була діадичною, а також вихідні дані, створені ДВП, сильно залежать від походження ряду, що аналізується. Невеликий зсув може

впливати на генеровані результати, через це важко узгодити перетворені сигнали з часом. Щоб подолати два вищезазначені обмеження, почали використовувати модифікацію, яка втрачає ортогональність, але набуває атрибутів, придатних для економічних досліджень, таких як плавність і можливість для аналізу недіадичних процесів (процесів, не кратних двом).

MODWT — це операція лінійного фільтрування, яка використовується для перевірки поведінки сигналу (часового ряду), що залежить від масштабу. MODWT покращує узгодження розкладених вейвлетів і коефіцієнтів масштабування на кожному рівні з вихідним часовим рядом. Таким чином, легко проаналізувати локалізовану зміну сигналу щодо масштабу та часу.

Функція $f(x)$ є лінійною комбінацією функції масштабування та вейвлет-функції, де j_0 є кількістю рівнів розкладання (2.19). Результатом є комбінація коефіцієнтів деталізації, які створюються перетворенням на кожному рівні, та масштабування, які створюються тільки на кінцевому рівні.

$$f(x) = \sum_{K=0}^{N-1} c_k 2^{-\frac{j_0}{2}} \odot (2^{-j_0} x - k) + \sum_{j=1}^{j_0} f_j(x) \quad (2.19)$$

$$f_j(x) = \sum_{k=0}^{N-1} d_{j,k} 2^{-\frac{j}{2}} \psi(2^{-j} x - k) \quad (2.20)$$

MODWT повертає N -багато коефіцієнтів $\{c_k\}$ і $(j_0 * N)$ багато коефіцієнтів деталізації $\{d_{j,k}\}$.

$$\|x\|^2 = \sum_{j=1}^{j_0} \|w_j\|^2 + \|v_{j_0}\|^2 \quad (2.21)$$

Де x — це вхідні дані, w_j — коефіцієнти деталізації в масштабі j і v_{j_0} — коефіцієнти масштабування кінцевого рівня.

Дискретне перетворення максимального перекриття фільтрує вхідний сигнал на основі кількості рівнів, де на кожному рівні перетворення здатне видалити коефіцієнти шуму, присутні в сигналі.

2.9. Висновки до розділу

Проведений аналіз предметної області, висвітлені загальні відомості про прогнозування часових рядів та перераховано основні міри точності прогнозів. Розглянуто ARIMA моделі, апарат штучної нейронної мережі та нейронні мережі, в основі яких лежить застосування глибокого навчання: LSTM та CNN, виділені їх переваги та недоліки. Розглянуто вейвлет-аналіз, який використовується для попередньої обробки даних. Висвітлені загальні відомості про вейвлет та деякі типи із сімейства вейвлетів: Хаара та Добеші, а також дискретне вейвлет-перетворення. А також розглянуто перетворення, яке найчастіше використовується для обробки фінансових часових рядів — ДВП максимального перекриття.

РОЗДІЛ 3

АЛГОРИТМ ПРОГНОЗУВАННЯ WA-LSTM

3.1. Запропонована схема прогнозування

У роботі було запропоновано гібридні моделі DWT-LSTM та MODWT-LSTM для прогнозування курсу Bitcoin на 1 крок вперед. Ця модель об'єднує переваги моделі декомпозиції (а саме DWT/MODWT) і моделі машинного навчання (LSTM). Розглянемо кроки цієї схеми (див. Рис. 3.1):

- 1) Вхідний ряд розкладається на набір різних підрядів (також відомих як вейвлети) за допомогою DWT та MODWT (використовуючи різні типи вейвлету за материнський).
- 2) Кожна підсерія прогнозується окремо з використанням LSTM.
- 3) Прогнозовані підряди повторно комбінуються для отримання остаточного прогнозного значення (вейвлет-реконструкція).
- 4) Графічне відтворення отриманих результатів прогнозування.
- 5) Показники ефективності/помилки обчислюються шляхом порівняння остаточного прогнозованого значення з вхідним рядом.

3.2. Емпіричні дослідження

3.2.1. Опис вхідних даних

Досить часто можна зіткнутися з ситуацією, коли даних, необхідних для моделювання і прогнозування, недостатньо, або ці статистичні дані спотворені. В результаті, при застосуванні прогнозних моделей, може статися збій в програмі або кінцеві дані можуть показати неприйнятний, для ситуації, що розглядається, результат.

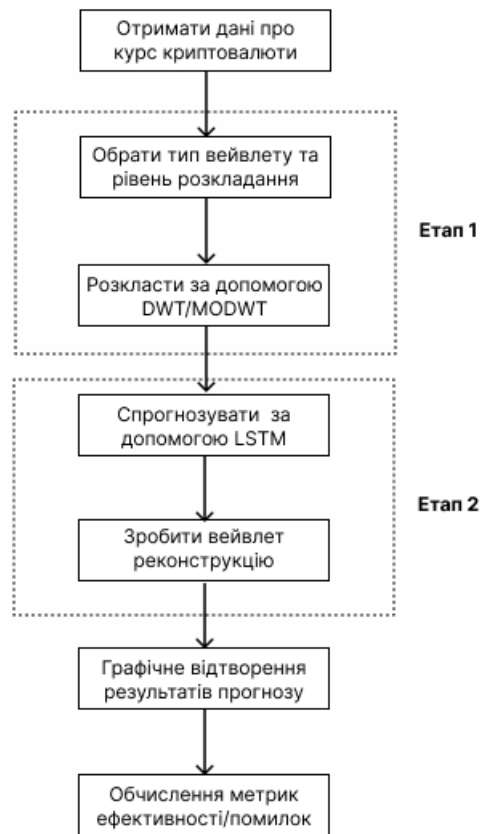


Рис. 3.1. Блок-схема запропонованої гібридної моделі

Під історичними даними ми будемо розуміти деякий набір даних, обмежених як за кількістю (за показниками), так і за датою. Іншими словами, дослідник в даній ситуації має у своєму розпорядженні не повний набір всієї необхідної інформації, а лише деякі дані, обмежені деяким проміжком часу.

Криптовалютні набори даних відрізняються повнотою опису торгів криптовалюти, які відбулися. Залежно від часових проміжків фіксації стану торгів і наявності даних за заявками виділяють кілька типів даних: Intraday data, Tick data, Time and sales data. Розглянемо перший тип: Intraday data — загальна характеристика даних, що показує стан торгів на біржі протягом дня. На відміну від класичних бірж, торги на криптовалютних біржах не зупиняються на ніч. Залежно від того, наскільки часто ми будемо знімати дані про стан торгів ми будемо отримувати різний ступінь повноти картини того, що відбувається на біржі. Від повноти картини залежить правильність настройки алгоритмів та їх похибка.

У даній роботі будуть розглядатися дані, які були зібрані з біржі криптовалют, у період з 1 червня 2022 року 00:15 по 16 червня 2022 року 00:00, тобто у нас є дані про курс біткоіна за 2 тижні з відліком 15 хвилин. Вибірка склала 1440 значень.

Наші дані — це OHLCV, вони характеризуються тим, що показують котирування на момент відкриття (open) торгового дня (початок дня), найвищий (high) рівень цін досягнутий за день, найменший (low) рівень цін досягнутий за день, рівень цін на момент закриття (close) торгового дня (на кінець дня), а також обсяг торгів (volume) за день. Такі дані вважаються не зовсім повними, оскільки це дані торгів, що відбулися на біржі. Однак ці дані не включають заявки (orders) на покупку (bid orders) і продаж (sell orders), які не відбулися, а це також важливо знати. Наприклад, може складатися ситуація, коли трейдер виставив на продаж біткоіни на 200 мільйонів доларів, але продав лише на 70 мільйонів доларів. 130 мільйонів залишилися невиконаними. Зазначимо, що це важлива інформація для розуміння подальшої можливої зміни курсу. Для відображення таких ситуацій використовують іншу групу типів даних про торги на біржі.

3.2.2. Етап 1: Декомпозиція курсу криптовалюти Bitcoin

3.2.2.1. Тип вейвлету та рівень декомпозиції

Для декомпозиції вхідного ряду використовується DWT та його модифікація — MODWT. Такий вибір зумовлено тим, що дискретне вейвлет-перетворення не вимагає припущення стаціонарності даних та є економним з точки зору кількості операцій і необхідної пам'яті, водночас надаючи достатньо інформації як для аналізу ряду, так і для його синтезу [18].

Щодо рівнів розкладання слід зазначити, що вибраний рівень декомпозиції залежить від кількості доступних точок даних. Згідно з [47], уся прихована інформація в даних виявляється, коли рівень декомпозиції дорівнює 6. Однак зі збільшенням рівня декомпозиції, дані згладжуються більшою мірою, що може призвести до втрати інформації [20]. Таким чином,

рівень декомпозиції, обраний для експериментів, дорівнює 1 чи 2. Для DWT було обрано 2 типи вейвлетів:

- Вейвлет Хаара (рівень декомпозиції 1)
- Вейвлет Добеші порядку 2 (рівень 2)

Різниця між цими вейвлетами у тому, що вейвлет Добеші має довший базис, що значно покращує функціональність цих нових перетворень.

3.2.2.2. Декомпозиція

DWT та MODWT використовувалися для декомпозиції курсу криптовалюти Bitcoin на різні підсерії, також відомі як вейвлети.

- DWT (тип: Хаар)

Було отримано один вейвлет-коефіцієнт (коефіцієнт деталізації) $W1$ і один коефіцієнт масштабування $V1$. Зазначимо, що коефіцієнти (ваги), пов'язані з функцією масштабування, звані коефіцієнтами апроксимації, фіксують низькочастотну інформацію, тоді як коефіцієнти, пов'язані з вейвлет-функцією, звані коефіцієнтами деталізації, фіксують високочастотну інформацію.

Розкладання ряду DWT (Хаар) показано на рисунку 3.2, де на першому рядку коефіцієнт $V1$, це згладжений компонент, малюнок якого дуже схожий на оригінальний часовий ряд (на останньому рядку). Коефіцієнт $V1$ відображає тенденцію руху курсу криптовалюти. Другий рядок представляє вейвлет-коефіцієнт $W1$, тобто високочастотну складову.

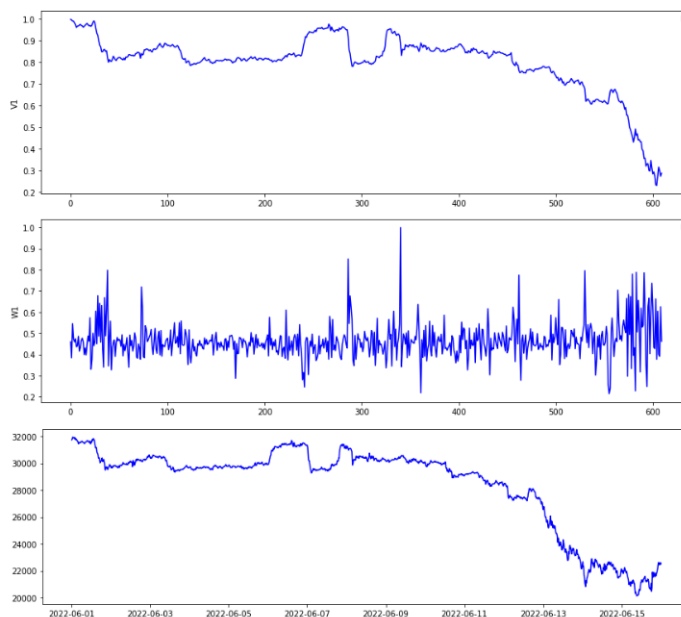


Рис. 3.2. Декомпозиція DWT (Хаар)

- DWT (тип: Добеші)

Розкладання часового ряду показано на рисунку 3.3. Отримано коефіцієнт масштабування $V2$ на першому рядку та два коефіцієнти деталізації $W2$, $W1$ на другому та третьому рядках відповідно. Останній рядок представляє початковий часовий ряд.

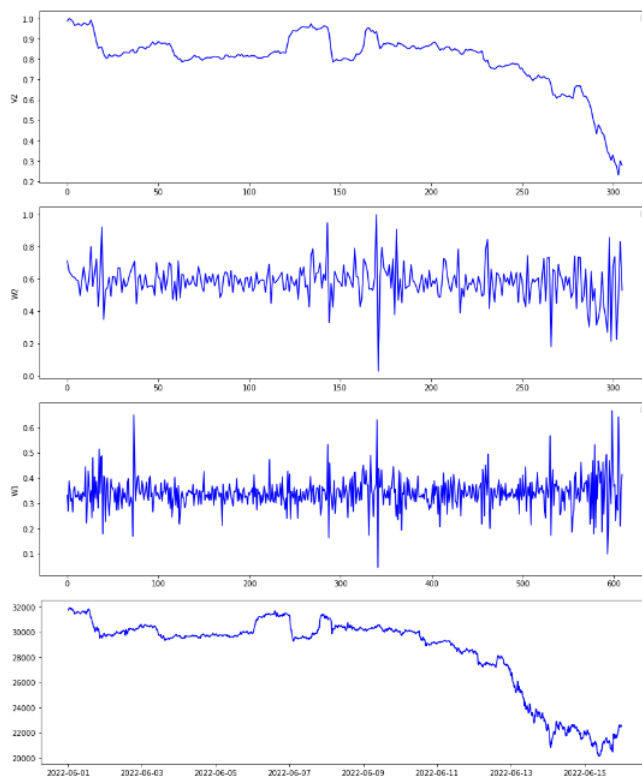


Рис. 3.3. Декомпозиція DWT (Добеші)

- MODWT (тип: Добеші, рівень розкладання 1)

На рис. 3.4 представлена декомпозиція ряду за допомогою MODWT з рівнем розкладання 1. На першому рядку коефіцієнт масштабування $V1$, а на другому — високочастотна складова $W1$.

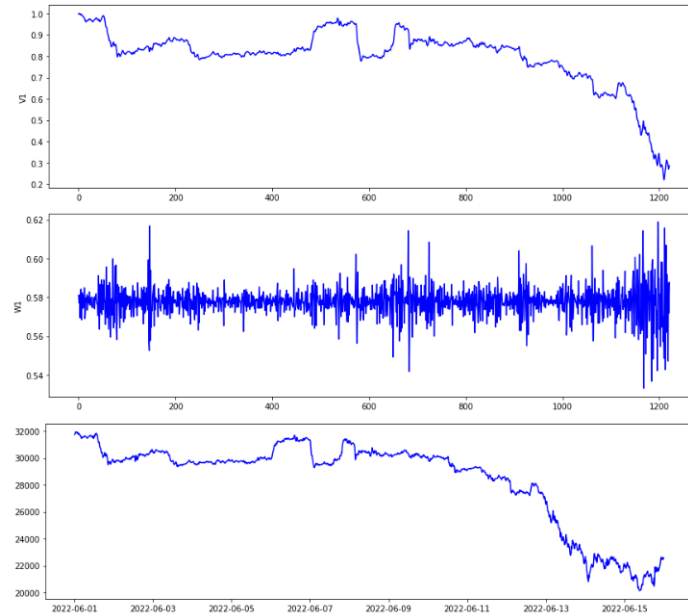


Рис. 3.4. Декомпозиція MODWT (рівень декомпозиції 1)

- MODWT (тип: Добеші, рівень розкладання 2)

Оскільки рівень розкладання 2, то було отримано два коефіцієнти деталізації $W2, W1$, які можна побачити на рис. 3.5 на 2 та 3 рядках відповідно, та один коефіцієнт масштабування $V2$ (перший рядок).

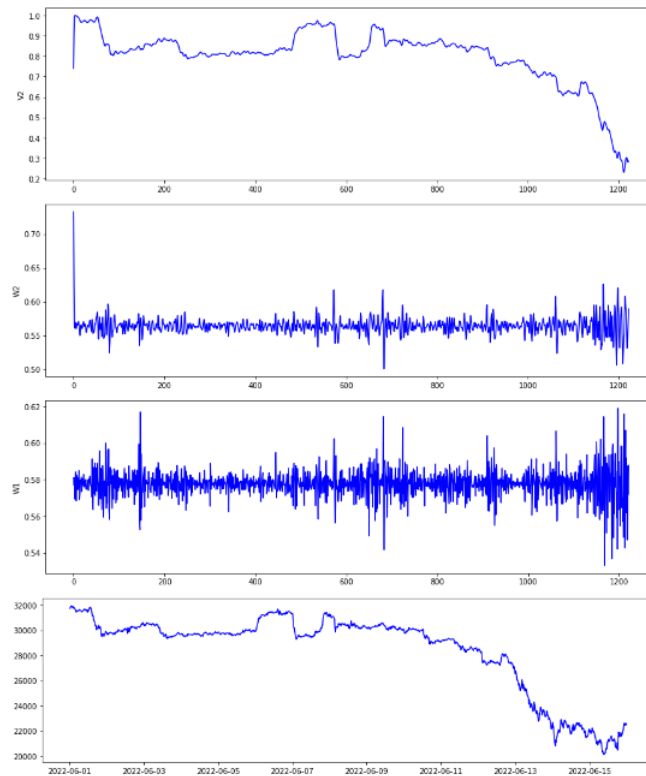


Рис. 3.5. Декомпозиція MODWT (рівень декомпозиції 2)

3.2.3. Етап 2: Прогнозування та реконструкція

На цьому етапі кожна отримана підсерія прогнозується за допомогою рекурентної нейронної мережі з довгою короткостроковою пам'яттю (LSTM).

3.2.3.1. Прогноз підсерій

Прогноз робиться 1 крок вперед для кожної підсерії, використовуючи при цьому $look_back = 1,2$ попередніх часових кроків для прогнозування наступного. Така модель називається одноетапною моделлю прогнозування. Перш ніж перейти до побудови рекурентної мережі LSTM, необхідно попередньо обробити вхідні дані. Спочатку проведено їх нормування, метою якого є зміна значень числових стовпців у наборі даних до загальної шкали, не спотворюючи відмінностей в діапазонах значень. Також важливо центрування даних — наявність у даних середнього значення і середньоквадратичного відхилення, тобто треба обчислити і застосувати до вибірки ці параметри. Далі дані були розділені на тренувальний та тестовий набори, у відношенні

85% та 15% відповідно. Після підготовки даних, настав час побудови моделі. Параметри моделі LSTM наведено у табл. 4.1.

3.2.3.2. Вейвлет-реконструкція

Після того як використали дискретне вейвлет-перетворення (та його модифікація) для декомпозиції фінансового часового ряду, необхідно зібрати компоненти назад у вихідний сигнал без втрати інформації. Цей процес називається реконструкцією, або синтезом. Математична маніпуляція, яка впливає на синтез, називається зворотним дискретним вейвлет-перетворенням (IDWT). Як відбувається цей процес?

- 1) Коефіцієнти DWT спочатку підвищуються (коефіцієнти апроксимації та деталізації обробляються окремо) шляхом розміщення нулів між кожними коефіцієнтами, фактично подвоюючи довжину кожного.
- 2) Потім вони згортаються за допомогою фільтра масштабування реконструкції для коефіцієнтів апроксимації (фільтр масштабування реконструкції — це просто вихідний фільтр масштабування, який було перевернуто зліва направо) і вейвлет-фільтра реконструкції для коефіцієнтів деталізації.
- 3) Потім ці результати додаються разом, щоб отримати вихідний сигнал.

3.3. Оцінка ефективності прогнозу

Отримані остаточні прогнози потім порівнюються з початковими даними для обчислення похибки. Крім того, щоб проаналізувати ефективність запропонованого підходу (гібридних моделей DWT-LSTM, MODWT-LSTM), оригінальний ряд (тобто ряд без декомпозиції) моделюється та прогнозується за допомогою LSTM.

Для оцінки ефективності прогнозування використовуються наступні показники: коефіцієнт детермінації R^2 , середня абсолютна похибка (MAE), відсоткова похибка (MAPE), корінь з середньоквадратичної похибки (RMSE) та метрика Dynamic Time Warping (DTW). Зазначимо, що значення MAE та RMSE розраховуються в тих самих одиницях, в яких представлені вхідні

дані (у нашому випадку, USD). Наприклад, у контексті ціни Bitcoin, MAE 125 означає, що прогнозована ціна становить ± 125 доларів США від фактичної ціни. Модель, яка іноді передбачає помилкові значення, буде мати вище значення RMSE, хоча у неї все ж може бути нижче MAE. Найкращим вважається випадок з найнижчими значеннями MAE, MAPE, RMSE, DTW та найвищим значенням R^2 . Таким чином, моделі слід оцінювати за всіма показниками.

3.4. Висновки до розділу

Розглянуто криптовалютні набори даних, а саме один з його типів — Intraday data. Один з прикладів таких даних — є OHLCV (Open, High, Low, Close, Volume), саме такий набір даних буде використовуватися для прогнозування курсу криптовалют.

Крім того, описано запропонований підхід прогнозування з використанням вейвлет-перетворення, наведено блок-схему. Далі більш детально розглянуто кожен крок, наведено рисунки отриманих декомпозицій ряду. Обрано параметри моделі нейронної мережі з комірками LSTM, а також описано вейвлет-реконструкцію. А також описано, які метрики використовуються для оцінки ефективності прогнозу.

РОЗДІЛ 4

ПРАКТИЧНА РЕАЛІЗАЦІЯ ПРОГРАМНОГО ПРОДУКТУ

4.1. Вибір мови програмування та бібліотек

У цьому розділі буде наведено обчислювальні результати, пов'язані з навчальними процесами, а також детальний опис реалізації дослідження та архітектури програмного продукту.

В якості мови програмування була обрана мова Python. Перевагами саме такого вибору стали наявність широкого вибору бібліотек та фреймворків для роботи з даними, стислість та читабельність коду (що особливо доцільно для задач машинного навчання).

Для прогнозування курсу криптовалюти біткоїн на 1 крок вперед було обрано метод глибинного навчання, а саме: рекурентна нейрона мережа з довгою короткостроковою пам'яттю (LSTM).

Для реалізації мережі LSTM було застосовано бібліотеку методів глибинного навчання Keras, тому що там шари додаються поступово, а не визначають всю мережу відразу. Так ми можемо швидко змінювати кількість і тип шарів, оптимізуючи нейронну мережу.

Щоб зробити вейвлет-декомпозицію та реконструкцію часового ряду, використовувалося програмне забезпечення PyWavelets. Це бібліотека для вейвлет-перетворення з відкритим кодом, яка є дуже простою у використанні та поєднує в собі простий інтерфейс високого рівня з продуктивністю C і Cython [22]. Даний пакет реалізує кілька n-вимірних дискретних вейвлет-перетворень, крім того, надається широкий вибір попередньо визначених вейвлетів. Усі багатовимірні перетворення реалізовано в Python за допомогою окремого застосування одновимірних перетворень. API для PyWavelets був розроблений таким чином, щоб бути подібним до інструментарію вейвлетів Matlab, і такі функції, як 1D, 2D і 3D дискретні вейвлет-перетворення, перевіряються на точність порівняно з їхніми аналогами Matlab. PyWavelets

має додаткову функціональність, незвичайну для інших наборів інструментів вейвлетів, такі як підтримка розмірності $n > 3$, та підтримка як дійсних, так і комплексних даних. Зазначимо, що PyWavelets був розроблений для використання вченими, які працюють у низці сфер, включаючи аналіз часових рядів, обробку сигналів, обробку зображень і медичну візуалізацію.

Зауважимо, що наш аналіз буде зосереджений лише на інформації про ціни, залишаючи осторонь будь-який фактор, який може вплинути на ціну біткоіна, як, наприклад, новини, які можуть відігравати дуже важливе правило. З цієї причини, треба сказати, що відповідний проєкт призначений лише для навчальних цілей і не доречно використовувати його у трейдерстві. Це надто спрощена модель, яка допоможе зрозуміти прогнозування часових рядів з використанням Python та рекурентних нейронних мереж, а також методу онлайн-навчання та проаналізувати вплив довжини вибірки та частоти дискретизації на точність прогнозу.

4.2. Вхідні дані

Існує досить багато ресурсів, які ми можемо використовувати для отримання історичних даних про ціни на біткоіни. Деякі з цих ресурсів дозволяють користувачам завантажувати файли CSV вручну, інші пропонують API, який можна приєднати до його коду.

Як вже було зазначено у попередньому розділі, у даній роботі будуть розглядатися дані Intraday data — OHLCV, які були зібрані з біржі криптовалют, у період з 1 червня 2022 року 00:15 по 16 червня 2022 року 00:00, тобто у нас є дані про курс біткоіна за 2 тижні з відліком 15 хвилин. Вибірка склала 1440 значень.

4.2.1. Data Exploration — дослідження даних

Першим чином, завантажимо наші набори даних за допомогою бібліотеки Pandas, на малюнку нижче можемо побачити, як виглядають дані з 15-хвилинним інтервалом.

	<DATE>	<TIME>	<OPEN>	<HIGH>	<LOW>	<CLOSE>	<VOL>
1361	150622	04:30	21207.78	21251.01	20915.00	20978.42	958
1173	130622	05:30	25696.04	25822.10	25600.17	25635.38	223
1169	130622	04:30	25798.75	26229.91	25733.31	26069.71	524
1318	140622	17:45	22618.66	22672.76	22322.05	22383.89	455
1223	130622	18:00	23519.72	23578.49	23398.47	23512.60	611
1357	150622	03:30	21817.38	21827.20	21590.82	21681.11	1153
89	10622	22:30	29528.95	29682.01	29507.01	29680.40	177

Рис. 4.1. Вхідні дані з 15-хвилинними інтервалом

Після цього видалимо зайві дані та переназвемо деякі з них. Крім того, перетворимо інформацію про дату та час на `pandas.TimeStamp`. Отримаємо:

	values	ts
923	29536.05	2022-06-10 15:00:00
735	30455.16	2022-06-08 16:00:00
1092	27493.98	2022-06-12 09:15:00
461	29962.26	2022-06-05 19:30:00
512	31337.43	2022-06-06 08:15:00
638	29821.80	2022-06-07 15:45:00
283	29703.28	2022-06-03 23:00:00

Рис. 4.2. Оновлені дані

Для поставленої мети нас цікавить стовпець `Close`, який містить значення цін біткоїнів на кінець дня. Нижче наведено графік (див. рис. 4.3), який ілюструє зміну ціни біткоїна протягом обраного періоду.

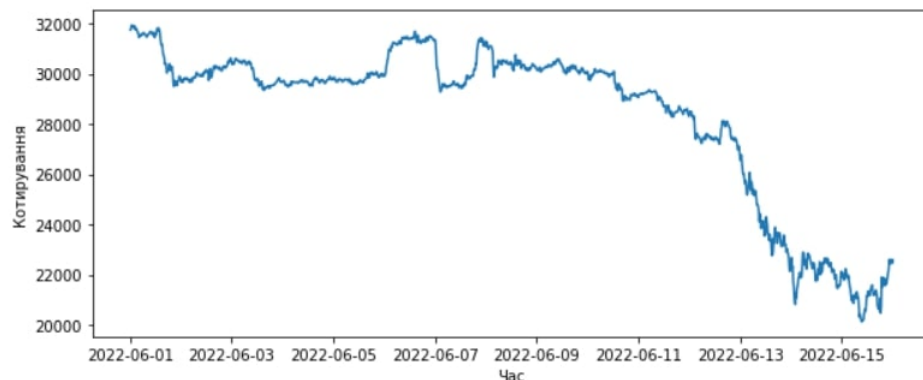


Рис. 4.3. Зміна курсу Bitcoin протягом 2 тижнів червня 2022 року

Найвища ціна за цей період — 31928.34\$, 01.06.2022 о 00:45.

Найнижча ціна за цей період — 20134.03\$, 15.06.2022 о 09:15.

Тепер перевіримо, чи містяться у стопчику *values* будь-які нульові значення, нульові значення нам не корисні, і ми повинні їх замінити, якщо такі є. Зробимо це за допомогою функції *.info()*, яка використовується для отримання короткого огляду набору даних. Це дуже зручно при дослідницькому аналізі даних. Перевіrivши, отримали, що нульових значень немає, тому продовжуємо працювати з даними.

4.2.2. Data Preprocessing — попередня обробка даних

- Крок 1: Нормування

Ціни криптовалют, що використовуються в обчисленнях, часто збільшуються з часом, що призводить до виходу більшості значень в тестовому наборі за допустимі межі. Виходить так, що мережа повинна передбачити значення, які вона ніколи не бачила раніше. Очевидно, що рано чи пізно модель перестає поводитися адекватно. Метою нормування є зміна значень числових стовпців у наборі даних до загальної шкали, не спотворюючи відмінностей в діапазонах значень.

Тому в першу чергу для адекватної роботи мережі в цілому, потрібно унормувати дані. Для цього завдання з бібліотеки *scikit-learn* був викликаний клас *MinMaxScaler*. Для кожного значення в датасеті *MinMaxScaler* віднімає мінімальне значення з датасета і ділить на діапазон — різницю між вихідним максимумом і мінімумом (він становить від 0 до 1).

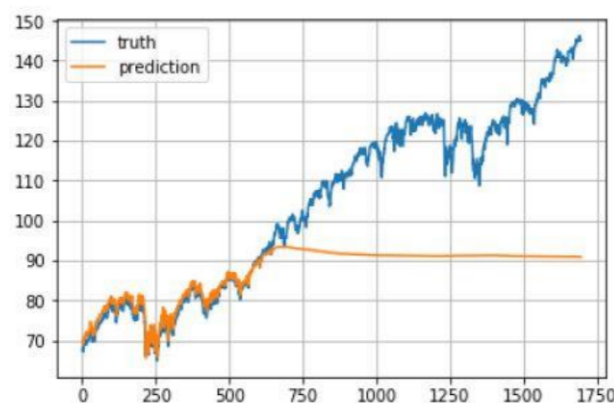


Рис. 4.4. Приклад використання ненормованих даних

Для адекватної роботи мережі також важливо центрування даних — наявність у даних середнього значення і середньоквадратичного відхилення.

Відбувається це наступним чином: віднімається середнє значення, а потім цей результат ділиться на стандартне відхилення. Функція `fit_transform()` з бібліотеки `scikit-learn`, використовувана в цій роботі, допомагає зробити ці дві операції одночасно (обчислити й застосувати до вибірки ці параметри).

- Крок 2: Розбиття даних

На цьому кроці потрібно розділити наш набір даних на тренувальний та тестовий набори. Тренувальний набір даних використовується для навчання моделі, тоді як тестовий набір використовується для забезпечення об'єктивної оцінки кінцевої моделі. Це дуже важливо, оскільки потрібно переконатись, що наші прогнози мають сенс, але ми не можемо перевіряти ті самі дані, якими ми навчаємо нашу мережу, оскільки ми можемо зіткнутися з ризиком перенавчання. Для тренувальних даних візьмемо 85%, для тестових – 15%.

- Крок 3: Формування даних

Завершальним етапом попередньої обробки даних є переформування вхідного масиву, щоб він був сумісним із шарами LSTM.

4.3. Побудова моделі LSTM

Після підготовки наших даних настав час для побудови мережі, яку ми згодом будемо навчати. Для цього, в першу чергу, необхідно завантажити пакети даних з бібліотеки Keras:

- Sequential використовується в якості моделі послідовної мережі.
- Dense застосовується для створення прихованих (повнозв'язних) шарів мережі.
- LSTM призначений для створення шарів з довгою короткостроковою пам'яттю для довготривалого зберігання інформації.

Створення першого LSTM-шару мережі відбувалося наступним чином: аргументом `units` задавалося число внутрішніх блоків LSTM-мережі. Налаштування `return_sequences` визначає, чи повертати останній висновок як результат роботи програми, або "відправити цей висновок далі" для цього аргументу був обраний параметр `True`, бо передбачається подальше включення

додаткових шарів LSTM-типу.

Далі було обрано оптимізатор *Adam* який є методом стохастичного градієнтного спуску, заснований на оцінці моментів першого і другого порядків. Цей метод ефективний в обчисленнях, має невелику потребу в пам'яті. Також потрібно обрати функцію втрат, мета якої обчислити помилку, яку модель повинна прагнути мінімізувати в процесі навчання. В рамках даного дослідження була обрана середньоквадратична помилка моделі (*mean squared error*).

Використовуючи *fit* можна почати тренування мережі, попередньо обравши параметри *batch_size*, підмножина навчальної вибірки фіксованого розміру (кількість прикладів), і *epochs*. Епохою називається одна ітерація в процесі навчання мережі. Оскільки неможливо пропустити всі дані разом, то його ділять на *batches* — маленькі партії. Дуже важливо коректно оцінити необхідну кількість епох, адже їх мала кількість призводить до недонавчання, а надто велика — до перенавчання.

Тепер можна приступити безпосередньо до прогнозування, скориставшись *predict()*. Використавши *inverse_transform()*, результати прогнозування були отримані в абсолютних одиницях (долар США).

Усі параметри цієї мережі зведемо до таблиці 4.1:

Таблиця 4.1

Параметри моделі

Кількість шарів LSTM	4
Кількість блоків шару LSTM (units)	50
Кількість повнозв'язних шарів (Dense)	1
Оптимізатор	Adam
Функція втрат	Mean Squared Error
Batch-size	32
Кількість епох	100

4.4. Результати прогнозу на основі "чистої" моделі LSTM

На рис. 4.5 можемо бачити графік порівняння тестових даних та прогнозу, зробленого за допомогою цієї мережі без використання вейвлет-перетворення до вхідних даних. Прогноз був зроблений на 1 крок вперед, використовуючи при цьому 20 попередніх часових кроків для прогнозування наступного. Така модель називається одноетапною моделлю прогнозування.

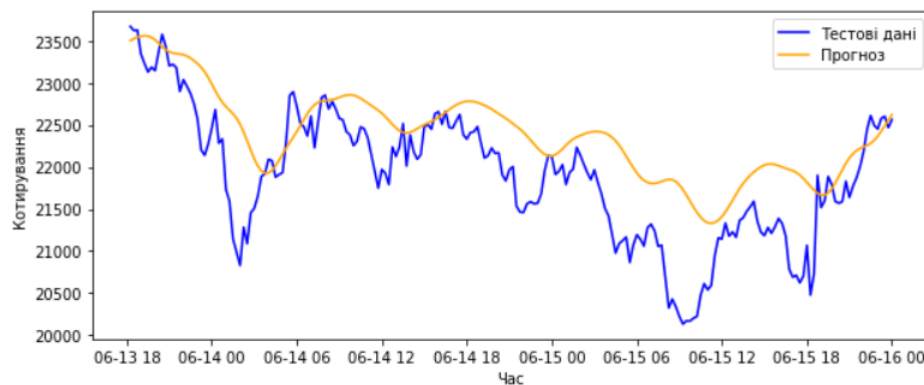


Рис. 4.5. Порівняння тестових даних та створеного прогнозу

Тепер порахуємо показники ефективності моделі, для цього скористаємося модулем *sklearn.metrics*, який реалізує функції, що оцінюють помилку прогнозу. Також завантажимо модуль *tslearn.metrics*, що забезпечує специфічні показники часових рядів. Скористаємося функцією *dtw()*, яка обчислює показник подібності динамічного викривлення часу (DTW) між часовими рядами.

- $R^2 = 0.66$;
- $MAE = 361.29$;
- $MAPE = 17\%$;
- $RMSE = 446.16$;
- $DTW = 2459.79$.

Дивлячись на вищенаведений графік та отримані значення показників, можемо зробити висновок, що створений прогноз не є точним. Тому скористаємося запропонованою схемою: спочатку застосувати вейвлет-декомпозицію, потім зробити прогноз за допомогою LSTM.

4.5. Використання вейвлет-перетворення та результати дослідження

Для того, щоб довести ефективність використання вейвлет-перетворень в якості попередньої обробки даних, будемо проводити декілька експериментів з використанням різних типів вейвлетів для дискретного вейвлет-перетворення, а також візьмемо два рівні розкладання для ДВП максимального перекриття.

4.5.1. DWT (тип: Хаар) + LSTM

Почнемо наше дослідження з однорівневого дискретного вейвлет-перетворення з застосуванням вейвлету Хаара у якості функції перетворення. Для декомпозиції будемо використовувати функцію *dwt()*, у якій наступні параметри:

- *data* — вхідний ряд;
- *wavelet* — тип вейвлету;
- *mode* — режим розширення сигналу (ряду);
- *axis* — вісь, по якій обчислюється DWT; якщо не вказано, використовується остання вісь.

Зупинимося докладніше на параметрі *mode*. Оскільки найпоширенішим і практичним способом представлення цифрових сигналів є скінченні масиви значень, необхідно виконати деяку екстраполяцію вхідних даних, щоб розширити сигнал перед обчисленням дискретного вейвлет-перетворення. Залежно від методу екстраполяції, під час цього процесу на границях сигналу можуть бути внесені значні артефакти (дефект, шум у сигналі), що може призвести до неточних обчислень DWT на кінцях сигналу. PyWavelets надає кілька методів екстраполяції сигналу, які можна використовувати для мінімізації цього негативного ефекту. Будемо використовувати симетричне заповнення — *symmetric*. Тобто сигнал розширюється за допомогою дзеркальних семплів. Отримаємо таку функцію:

$$A1, D1 = \text{pywt.dwt}(df['values'].values, 'haar', mode = 'symmetric')$$

Після цього будемо ділити дані на тренувальні та тестові. Показник *look_back*, тобто кількість попередніх часових кроків для прогнозування наступного, дорівнює 2. Низьке значення показника вказує на те, що LSTM не є оптимальною моделлю для створення прогнозу, при вищих значеннях відбувається згладження результуючої кривої. Отримали наступні набори: Тренувальні точки: 612. Тестові точки: 108. Тренувальні точки: 612. Тестові точки: 108.

Далі робимо прогноз окремо для кожної підсерії за допомогою LSTM та вейвлет-реконструкцію, використовуючи функцію *idwt()*. Необхідні параметри:

- *sArray_like* — коефіцієнти апроксимації;
- *sDarray_like* — коефіцієнти деталізації;
- *wavelet* — тип вейвлету;
- *mode* — режим розширення сигналу (ряду);
- *axis* — вісь, по якій обчислюється зворотний DWT.

В результаті маємо однорівневу реконструкцію сигналу за заданими коефіцієнтами, яка і буде отриманим прогнозом. Далі побудуємо графік порівняння тестових даних та створеного прогнозу (див. рис. 4.6), а також розрахуємо значення показників ефективності моделі.

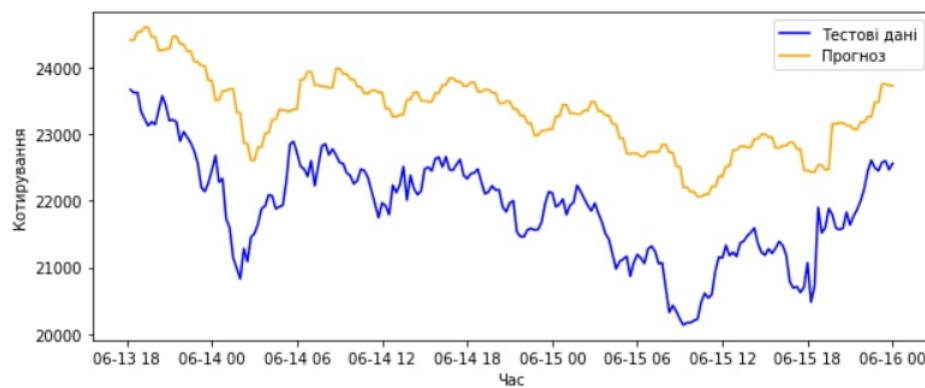


Рис. 4.6. Порівняння тестових даних та створеного прогнозу

Показники ефективності моделі:

- $R^2 = 0.67$;
- $MAE = 355.92$;
- $MAPE = 16\%$;
- $RMSE = 444.42$;

- $DTW = 3150.45$.

4.5.2. DWT (тип: Добеші) + LSTM

Далі розглянемо дворівневе ДВП з застосуванням вейвлету Добеші порядку 2. Для багаторівневої декомпозиції використовується інша функція, а саме *wavedec()*.

$A2, D2, D1 = pywt.wavedec(df['values'].values, 'db2', mode='smooth', level=2)$

Smooth — сигнал розширюється за першими похідними, обчисленими на ребрах.

Після дворівневої декомпозиції ряду, поділимо дані на тренувальний та тестовий набори:

Тренувальні точки: 310. Тестові точки: 55.

Тренувальні точки: 310. Тестові точки: 55.

Тренувальні точки: 614. Тестові точки: 109.

Поділивши отримані робимо прогноз окремо для кожної підсерії за допомогою LSTM та вейвлет-реконструкцію, використовуючи функцію *waverec()*. У результаті отримаємо прогноз, який треба зобразити на графіку (див. рис. 4.7), а також розрахуємо значення показників ефективності моделі.

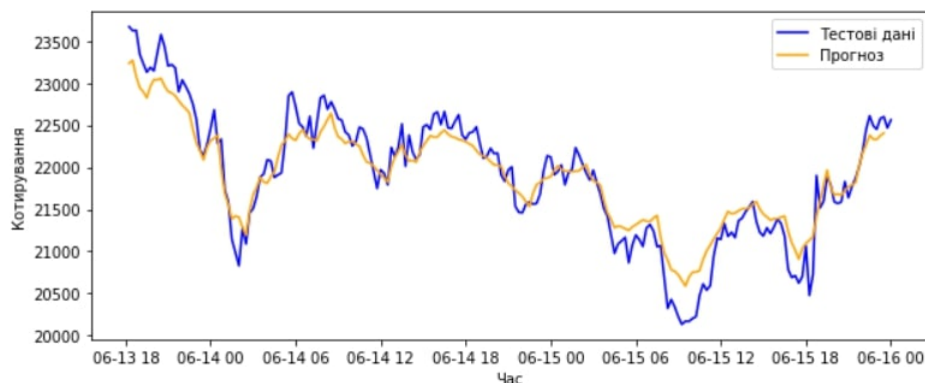


Рис. 4.7. Порівняння тестових даних та створеного прогнозу

Показники ефективності моделі:

- $R^2 = 0.90$;

- $MAE = 195.20$;
- $MAPE = 9\%$;
- $RMSE = 244.58$;
- $DTW = 2432.62$.

4.5.3. MODWT (рівень розкладання 1) + LSTM

Тепер скористаємося сучасним методом попередньої обробки даних — MODWT. Для реалізації цієї декомпозиції будемо використовувати функцію *swt()*. Зазначимо, що Stationary Wavelet Transform це синонім MODWT. Ця функція має наступні параметри:

- *data* — вхідний сигнал (ряд);
- *wavelet* — тип вейвлету;
- *level* — кількість рівнів декомпозиції;
- *start_level* (необов'язково) — рівень, на якому почнеться декомпозиція (за замовчуванням: 0);
- *axis* (необов'язково) — вісь, по якій обчислюється SWT;
- *trim_approx* (необов'язково) — якщо *True*, коефіцієнти апроксимації на кінцевому рівні зберігаються;
- *norm* (необов'язково) — якщо *True*, перетворення нормалізується так, що енергія коефіцієнтів дорівнюватиме енергії даних.

Слідуючи документації [22] при використанні з *norm = True* це перетворення є MODWT. Тому наша функція має вид:

```
A1, D1 = pywt.swt(df['values'].values, 'db2', level = 1, trim_approx = True,
                  norm = True)
```

Прогноз робимо окремо для кожної підсерії за допомогою LSTM та потім застосовуємо вейвлет-реконструкцію *iswt()*. Порівняння отриманого прогнозу та реальних даних можемо побачити на графіку (див. рис. 4.8).

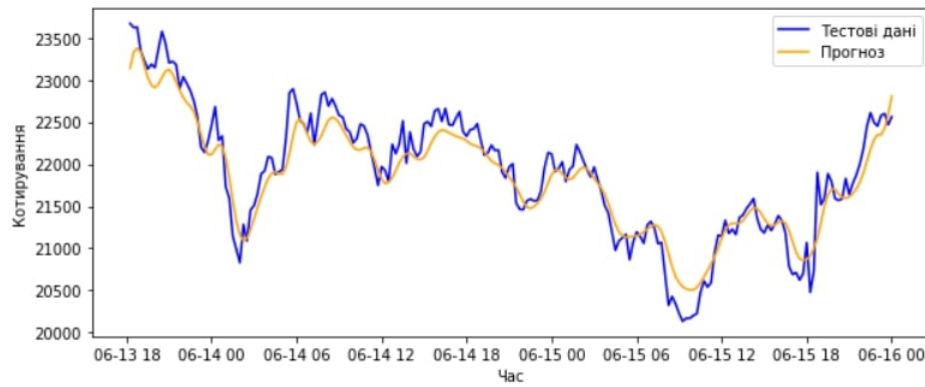


Рис. 4.8. Порівняння тестових даних та створеного прогнозу

- $R^2 = 0.88$;
- $MAE = 200.16$;
- $MAPE = 9\%$;
- $RMSE = 262.14$;
- $DTW = 2337.39$.

4.5.4. MODWT (рівень розкладання 2) + LSTM

Тепер подивимося, як зміниться результат прогнозу, якщо вкажемо рівень розкладання 2. Наведемо графік порівняння прогнозних значень та реальних (див. рис. 4.9) та розрахуємо показники ефективності моделі.

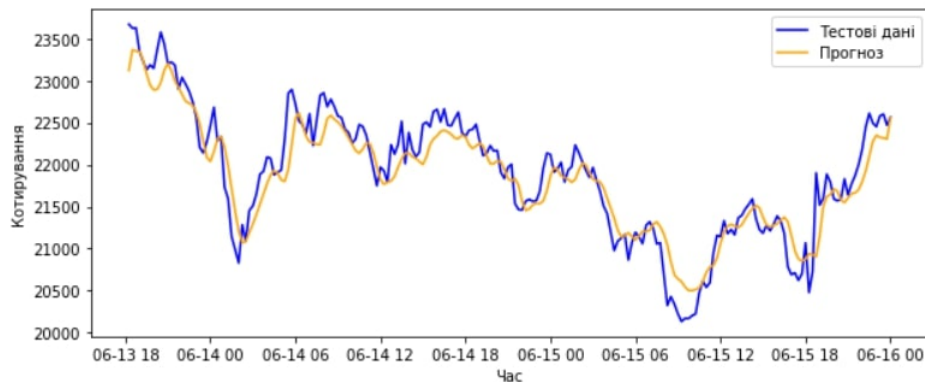


Рис. 4.9. Порівняння тестових даних та створеного прогнозу

- $R^2 = 0.88$;
- $MAE = 200.16$;
- $MAPE = 9\%$;
- $RMSE = 262.14$;
- $DTW = 2337.39$.

4.6. Оцінка ефективності моделі

Для більш зручного порівняння отриманих результатів, зведемо усі показники в таблицю 4.2.

Таблиця 4.2

Оцінка точності прогнозування

Модель	R^2	MAE	MAPE	RMSE	DTW
LSTM	0.66	361.29	17%	446.16	2459.79
DWT (Хаар) + LSTM	0.67	355.92	16%	444.42	3150.45
DWT (Добеші) + LSTM	0.90	195.20	9%	244.58	2432.62
LSTM + MODWT (рівень 1)	0.88	200.16	9%	262.14	2337.39
LSTM + MODWT (рівень 2)	0.91	175.11	8%	226.15	2160.28

Отримані результати дають змогу дійти висновку, що попередня обробка даних за допомогою вейвлет-перетворень значно покращує точність прогнозування, а використання MODWT з рівнем розкладання 2 найкраще підходить для роботи на основі запропонованих даних.

Варто зазначити, що добрий результат було отримано і з вейвлетом Добеші, і з MODWT з рівнем розкладання 1. Разом з тим, було виявлено, що різні материнські вейвлети, які застосовуються для аналізу того ж часового ряду, дають різні результати. Унаслідок чого, існує складність у використанні такого підходу, яка полягає у правильному виборі вейвлет-перетворення, самого материнського вейвлету та рівня розкладання.

4.7. Висновки до розділу

Проаналізовано інструменти, що підходять для розв'язання поставленої задачі та обрано мову програмування Python, бібліотеку машинного навчання Keras та бібліотеку PyWavelets. Проведено попередню обробку даних та створено тренувальний та тестовий набори. Наступним кроком був

вибір параметрів моделі нейронної мережі з комірками LSTM та проведене її навчання на створених наборах. Після цього зроблено прогноз на основі "чистої" моделі.

Отримані результати були не точними, тому також використовували вейвлет-перетворення у якості попередньої обробки даних. Проведено декілька експериментів, де змінювали тип материнського вейвлету та рівні розкладання. Наведено графіки порівняння створеного прогнозу та реальних даних, а також розраховано показники точності прогнозування. У результаті проведеної роботи, було доведено ефективність використання обраного підходу.

ВИСНОВКИ

У результаті проведеної роботи була досягнута поставлена мета, а саме було досліджено прогнозування курсу криптовалют за допомогою вейвлет-аналізу та нейронної мережі, порівняно отримані результати з прогнозом на основі "чистої" моделі LSTM та доведено ефективність використання обраного підходу.

Поставлена мета зумовила необхідність вирішення таких завдань: було розглянуто основні поняття у сфері криптовалют та виконано аналіз актуальності задачі прогнозування курсу криптовалют, яка підтвердила високу популярність з боку трейдерів. Виявлено основні фактори, що впливають на зміну курсу криптовалют. Проведено огляд поточного стану прогнозування на ринку криптовалют й огляд існуючих досліджень та рішень.

Після аналізу існуючих підходів прогнозування часових рядів щодо зручності, часозатрат та ефективності прогнозу, було виділено 5 основних з них:

- Метод використання інтегрованої моделі авторегресії (ARIMA)
- Метод використання штучної нейронної мережі (ANN)
- Метод використання рекурентної мережі з довгою короткостроковою пам'яттю (LSTM)
- Метод використання згорткової нейронної мережі (CNN)
- Метод використання комбінованих нейронних моделей

У результаті аналізу, для прогнозування було обрано підхід, який базується на використанні рекурентної мережі з довгою короткостроковою пам'яттю (LSTM). Такий вибір зумовлено тим, що така мережа здатна обробляти цілі послідовності даних, тому добре підходить для обробки і побудови прогнозів на основі часових рядів.

Крім того, було проаналізовано популярність використання вейвлетних технологій у фінансовій сфері. У роботі було використано дискретне вейвлет-перетворення (DWT), з різними типами материнських вейвлетів, а також дискретне вейвлет-перетворення максимального перекриття (MODWT), який є одним з найсучасніших методів попередньої обробки фінансових часових рядів. Такий вибір зумовлено тим, що дискретне вейвлет-перетворення

не вимагає припущення стаціонарності даних та є економним з точки зору кількості операцій і необхідної пам'яті, водночас надаючи достатньо інформації як для аналізу ряду, так і для його синтезу.

Проведені експерименти дають змогу дійти висновку, що попередня обробка даних за допомогою вейвлет-перетворень значно покращує точність прогнозування, а використання MODWT з рівнем розкладання 2 найкраще підходить для роботи на основі запропонованих даних. Варто зазначити, що добрий результат було отримано і з вейвлетом Добеші, і з MODWT з рівнем розкладання 1.

Запропонована модель є корисною для прогнозування курсу, оскільки вона базується на декомпозиції фінансового часового ряду. Це допомагає розкласти часовий ряд на різні частотні компоненти. Окремі компоненти легше передбачити, ніж нестационарні та нелінійні вхідні дані в цілому. Разом з тим, було виявлено, що різні материнські вейвлети, які застосовуються для аналізу того ж часового ряду, дають різні результати. Унаслідок чого, існує складність у використанні такого підходу, яка полягає у правильному виборі вейвлет-перетворення, самого материнського вейвлету та рівня розкладання.

Оскільки в роботі використовувалася тільки інформація про ціну криптовалюти, без урахування факторів, які можуть впливати на зміну курсу, в майбутніх дослідженнях планується побудова точнішої моделі з використанням зовнішніх факторів, що впливають на курс.

СПИСОК ЛІТЕРАТУРИ

1. Alahmari, S. (2019). Using Machine Learning ARIMA to Predict the Price of Cryptocurrencies.
2. Altan, A., Karasu, S., & Bekiros, S. (2019). Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques.
3. Beikverdi, A. (2015). Is China the Primary Driving Force in Bitcoin?. Cointelegraph.
4. Box, G., Reinsel, G., & Jenkins, G. (1976). Time Series Analysis: Forecasting and Control.
5. Brownlee, J. (2018). Deep Learning for Time Series Forecasting. Predict the Future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery.
6. Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets.
7. Gilles, J. (2013). Empirical wavelet transform.
8. Gusarov, V. M. (2001). Statistika: Uchebnoe posobie dlja vuzov. [Statistics: Textbook for Universities]. [in Russian]
9. Greenberg, A. (2011). Crypto Currency. Retrieved from <https://www.forbes.com/forbes/2011/0509/technology-psilocybin-bitcoins-gavin-andresen-crypto-currency.html#6110e19c353e>.
10. Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. Math. Ann. 69, 331–371.
11. Hileman, G., & Rauchs, M. (2017). Global Cryptocurrency Benchmarking Study. United Kingdom, Cambridge: Cambridge Centre For Alternative Finance.
12. Hitam, N., & Ismail, A. (2018). Comparative Performance of Machine Learning Algorithms for Cryptocurrency Forecasting. Indonesian Journal Of Electrical Engineering And Computer Science, 11(3), 1121.
13. Hochreiter, S. & Schmidhuber, J. (1997). Long Short-term Memory. Neural Computation, 9(8), 1735–1780.
14. Hyndman, R., & Athanasopoulos, G. (2018). Forecasting: principles and practice (2nd ed.). OTexts, Melbourne, Australia.

15. Hsieh, T.-J., Hsiao, H.-F., & Yeh, W.-C. (2011). Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm.
16. Kao, L. J., Chiu, C.C., Lu, C.J., & Chang, C.H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting.
17. Kim, Y., Kim, J., Kim, W., Im, J., Kim, T., Kang, S., & Kim, C. (2016). Predicting Fluctuations in Cryptocurrency Transactions Based on User Comments and Replies.
18. Kolomiiets, R. O. (2020). Osnovy veivlet-peretvorennia.
19. Kumar, A., & Joshi, LK. (2011). MODWT based time scale decomposition analysis of BSE and NSE indexes financial time series. *International Journal of Mathematical Analysis* 5(27):1343-1352
20. Lahmiri, S. (2014). Wavelet low- and high-frequency components as features for predicting stock prices with backpropagation neural networks, *Journal of King Saud University - Computer and Information Sciences*, Vol. 26, Issue 2, pp. 218-227.
21. Leavey, C., James, M. N., Summerscales, J., & Sutton, R. (2003). An introduction to wavelet transforms: A tutorial approach. *Insight*, 45, 344-353.
22. Lee, G.R., Gommers, R., Wasilewski, F., Wohlfahrt, K., & O’Leary A. (2019). PyWavelets: A Python package for wavelet analysis. *Journal of Open Source Software*, 4(36), 1237.
23. Matsugu, M., Mori, K., Mitari, Y., & Kaneda, Y. (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*, 16(5-6), 555-559.
24. Moazzami, M., Khodabakhshian, A., & Hooshmand, R. (2013). A new hybrid day-ahead peak load forecasting method for Iran’s National Grid. *Applied Energy*.
25. McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–133.
26. McNally, S. (2016). Predicting the price of Bitcoin using Machine Learning.

- National College of Ireland.
27. Milovska, K. M, & Moroz V.V. (2020). Neiromerezhevyi pidkhid do prohnozuvannia chasovykh riadiv pry obmezhenni istorychnykh danykh. In: The National Technical University «Kharkiv Polytechnic Institute», Problems of informatics and modeling, Proceedings of the XX International Science and Technology Conference, Karolino-Bugaz, pp. 52.
 28. Milovska, K. M, & Moroz V.V. (2021). Odesa I. I. Mechnikov National University. 77th reporting student scientific conference of Odesa I.I. Mechnikov National University. Odesa, pp. 104.
 29. Milovska, K. M, & Moroz V.V. (2022). Veivletnyi analiz chasovykh riadiv ta prohnoz z zastosuvanniam neiromerezhi LSTM. In: The National Technical University «Kharkiv Polytechnic Institute», Problems of informatics and modeling, Proceedings of the XXII International Science and Technology Conference, Kharkiv, pp. 59.
 30. Milovska, K. M, & Moroz V.V. (2022). Veivletnyi analiz ta prohnoz finansovykh chasovykh riadiv. Herald of the National Technical University "KhPI". Series of "Informatics and Modeling"[accepted for publication]
 31. Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. Neural Computing And Applications.
 32. Nakamoto, S. (2009). Bitcoin: A Peer-to-Peer Electronic Cash System.
 33. Nobre, J. & Neves, R. F. (2019). Combining principal component analysis, discrete wavelet transform and xgboost to trade in the financial markets.
 34. Parvini, N., Abdollahi, M., Seifollahi, S., & Ahmadian, D. (2022). Forecasting Bitcoin returns with long short-term memory networks and wavelet decomposition: A comparison of several market determinants.
 35. Percival, D., & Walden, A. (2000). Wavelet Methods for Time Series Analysis Cambridge Series in Statistical and Probabilistic Mathematics.
 36. Risse, M. (2019). Combining wavelet decomposition with machine learning to forecast gold returns.
 37. Sakoe, H. & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. IEEE Transactions on Acoustics, Speech, and Signal Processing 26: 159-165.

38. Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V., & Soman, K. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017 International Conference On Advances In Computing, Communications And Informatics (ICACCI).
39. Schmidhuber, J. (2015) Deep Learning in Neural Networks: An Overview. *Neural Networks*, ISSN: 0893-6080, Vol: 61, pp. 85-117.
40. Shafie-khah, M., Moghaddam, MP., & Sheikh-El-Eslami, MK. (2011). Price forecasting of day-ahead electricity markets using a hybrid forecast method.
41. Shen, D., Urquhart, A., & Wang, P. (2019). Does twitter predict Bitcoin?. *Economics Letters*, 174, 118-122.
42. Shukla, S., & Dave, S. (2020). Bitcoin beats coronavirus blues. *The Economic Times*.
43. Soslovskij, V., & Kosovskij, I. (2016). Rinok kriptovaljut jak sistema. [Cryptocurrency market as a system]. *Visnik Universitetu Bankivs'koï Spravi — Bulletin of the University of Banking*, 236-246.
44. Stocchi, M., & Marchesi, M. (2018). Fast wavelet transform assisted predictors of streaming time series.
45. Sun, EW., & Meinl, T. (2012). A new wavelet-based denoising algorithm for high-frequency financial data mining. *European Journal of Operational Research* 217(3):589-599.
46. Tan, Z., Zhang, J., Wang, J., & Xu, J. (2010). Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. *Applied Energy*.
47. Tsai, D.-M., & Chiang, C.-H. (2003). Automatic band selection for wavelet reconstruction in the application of defect detection. *Image and Vision Computing*.
48. Vintsyuk, T. K. (1968). Speech discrimination by dynamic programming. *Cybern Syst Anal* 4, 52–57.
49. Yousefi, S., Weinreich, I., & Reinartz, D. (2005) Wavelet-based prediction of oil prices. *Chaos, Solitons Fractals*.
50. Zweig, G. (1976). Basilar membrane motion. *Cold Spring Harbor Symposia on Quantitative Biology*, 40:619–633.