МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ХАРЧОВИХ ТЕХНОЛОГІЙ

Інститут (факультет) _Автоматизації і комп'ютерних систем_____
Кафедра _Інформаційних технологій, штучного інтелекту і кібербезпеки_

«До захисту в ЕК»
Директор інституту(декан факультету)
Андрій ФОРСЮК
(підпис) (ім'я та прізвище)
13 грудня 2024р.

«До захисту допущено»
Завідувач кафедри
Сергій ГРИБКОВ
(підпис) (ім'я та прізвище)
13 грудня 2024р.

КВАЛІФІКАЦІЙНА РОБОТА
НА ЗДОБУТТЯ ОСВІТНЬОГО СТУПЕНЯ МАГІСТРА

зі спеціальності 122 «Комп'ютерні науки»
(код та назва спеціальності)
освітньо-професійної програми _Управління інформацією та аналітика даних

на тему: Дослідження та аналіз ринку молочної продукції України з використанням методів штучного інтелекту

Виконав: здобувач _2_ курсу, групи КН-2-4М
Бойко Ілля Олександрович (підпис)
(прізвище, ім'я, по батькові повністю)

Керівник Харкянен Олена Валеріївна (підпис)
(прізвище , ім'я та по батькові повністю)

Консультанти _____ _____
(ім'я та прізвище) (підпис)
(ім'я та прізвище) (підпис)
(ім'я та прізвище) (підпис)

Рецензент Ольга КРАВЧЕНКО (підпис)
(ім'я та прізвище)

Я як здобувач(ка) Національного університету харчових технологій розумію і підтримую політику університету з академічної доброчесності. Я не надавав(-ла) і не одержував(-ла) недозволеної допомоги під час підготовки цієї роботи. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

Здобувач _____
(підпис)

Київ - 2024р.

2

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ХАРЧОВИХ ТЕХНОЛОГІЙ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ХАРЧОВИХ ТЕХНОЛОГІЙ

Інститут (факультет) Автоматизації і комп'ютерних систем
Кафедра Інформаційних технологій, штучного інтелекту і кібербезпеки
Освітній ступінь магістр
Спеціальність 122 «Комп'ютерні науки»
(код і назва)
Освітньо-професійна програма Управління інформацією і аналітика даних
(назва)

ЗАТВЕРДЖУЮ
Завідувач
кафедри Інформаційних технологій, штучного інтелекту і кібербезпеки
Грибков С.В.
"07" жовтня 2024року

З А В Д А Н Н Я

НА КВАЛІФІКАЦІЙНУ РОБОТУ ЗДОБУВАЧА

Бойка Іллі Олександровича
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження та аналіз ринку молочної продукції України з використанням методів штучного інтелекту
керівник роботи Харкянен Олена Валеріївна, к. т. н., доцент,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)
затверджені наказом закладу вищої освіти від 7 жовтня 2024 року №884-кс
2. Строк подання здобувачем роботи 6 листопада 2024
3. Вихідні дані до роботи
1) Статистичні дані про виробництво молочної продукції в Україні
2) Ринкові дані
4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)
Розділ 1. Дослідження методів машинного навчання
Розділ 2. Опис математичної моделі для прогнозування споживання молочної продукції
Розділ 3. Аналіз та моделювання прогнозів доданої вартості у виробництві молочної продукції
5. Перелік графічного матеріалу
1) Алгоритми роботи методів штучного інтелекту
2) Приклади коду розробленого ПЗ

3

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1	Харкянен О.В., доцент	07.10.2024	
2	Харкянен О.В., доцент	07.10.2024	
3	Харкянен О.В. доцент	07.10.2024	

7. Дата видачі завдання 7 жовтня 2024 року

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів виконання кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Дослідження методів збору і аналізу даних, обґрунтування їхнього вибору	08.10.2024-14.10.2024	Виконано
2	Постановка задач щодо виконання збору і аналізу відгуків	15.10.2024-20.10.2024	Виконано
3	Дослідження та вибір математичної моделі	21.10.2024-22.10.2024	Виконано
4	Розробка та налаштування методу машинного навчання для аналізу та прогнозування	23.10.2024-14.11.2024	Виконано
5	Оформлення автореферату	15.11.2024-28.11.2024	Виконано
6	Оформлення презентації	29.11.2024-06.12.2024	Виконано

Здобувач _____ (підпис) Бойко І.О. (прізвище та ініціали)

Керівник роботи _____ (підпис) Харкянен О.В. (прізвище та ініціали)

АНОТАЦІЯ

Бойко Ілля Олександрович - Дослідження та аналіз ринку молочної продукції України з використанням методів штучного інтелекту.

Кваліфікаційна робота: 66 сторінок, 29 рисунків, 2 додатки, 13 джерел.

У кваліфікаційній роботі проведено дослідження застосування методів машинного навчання для прогнозування тенденцій споживання та виробництва м’ясо-молочної продукції, зокрема із використанням алгоритму LSTM.

Результатом роботи є програмне забезпечення, розроблене мовою Python, яке використовує алгоритм LSTM для аналізу даних і генерації прогнозів споживання та виробництва молочної продукції. Даний інструмент може бути використаний підприємствами для оптимізації виробничих процесів та розробки стратегій у динамічних ринкових умовах.

Ключові слова: ПРОГНОЗУВАННЯ, АНАЛІЗ, НЕЙРОННІ МЕРЕЖІ, МАШИННЕ НАВЧАННЯ, LSTM, МОЛОЧНА ПРОДУКЦІЯ.

ABSTRACT

Boiko Illia Oleksandrovych, Research and Analysis of the Ukrainian Dairy Market Using Artificial Intelligence Methods.

Diploma thesis: 66 pages, 29 figures, 2 appendices, 13 references.

In the qualification work, a study was conducted on the application of machine learning methods for forecasting consumption and production trends of meat and dairy products, specifically using the LSTM algorithm.

The outcome of the work is software developed in Python that utilizes the LSTM algorithm for data analysis and generating forecasts of dairy product consumption and production. This tool can be used by enterprises to optimize production processes and develop strategies in dynamic market conditions.

Keywords: FORECASTING, ANALYSIS, NEURAL NETWORKS, MACHINE LEARNING, LSTM, DAIRY PRODUCTS.

ЗМІСТ

ВСТУП 7
РОЗДІЛ 1. ДОСЛІДЖЕННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ 9
1.1. Актуальність дослідження 9
1.2. Історія машинного навчання в промисловості 10
1.3. Постанова проблеми 11
1.4. Методи та підходи дослідження 13
1.5. Висновки до першого розділу 19
РОЗДІЛ 2. ОПИС МАТЕМАТИЧНОЇ МОДЕЛІ ДЛЯ ПРОГНОЗУВАННЯ СПОЖИВАННЯ МОЛОЧНОЇ ПРОДУКЦІЇ 20
2.1. Основні методи та інструменти побудови моделі прогнозування споживання молочної продукції 20
2.2. Вибір математичної моделі для прогнозування 21
2.3. Висновки до другого розділу 34
РОЗДІЛ 3. АНАЛІЗ ТА МОДЕЛЮВАННЯ ПРОГНОЗІВ ДОДАНОЇ ВАРТОСТІ У ВИРОБНИЦТВІ МОЛОЧНОЇ ПРОДУКЦІЇ 35
3.1. Опис вхідних даних для аналізу і прогнозування 35
3.2. Збереження даних та забезпечення їх доступності 37
3.3. Розробка та навчання моделі 38
3.4. Результати аналізу ринку молочної продукції та побудова графіків 42
3.5. Висновки до третього розділу 52
ВИСНОВКИ 53
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 54
ДОДАТКИ 57

ВСТУП

Актуальності теми. Машинне навчання вже давно стало невід'ємною частиною сучасного життя, впливаючи на різні сфери діяльності людини. Його технології впроваджуються там, де потрібно обробляти великі обсяги даних, прогнозувати тенденції та знаходити оптимальні рішення. Сфера виробництва активно використовує здобутки таких дисциплін, як аналіз часових рядів, аналіз даних, теорія ймовірностей та Data Science. Ці інструменти дозволяють підприємствам ефективніше планувати ресурси, оптимізувати процеси та підвищувати продуктивність.

Ринок молочної продукції України також потребує впровадження і використання інструментів на основі технології обробки та аналізу даних. У сучасних умовах, коли обсяги виробництва зростають, а конкуренція на ринку посилюється, здатність прогнозувати попит, обсяги виробництва та тенденції споживання стає ключовим фактором успіху. Використання машинного навчання у цій галузі дозволяє автоматизувати процеси аналізу та створювати точні прогнози, що є особливо актуальним в умовах динамічних змін ринку та коливань споживчого попиту.

Зв'язок роботи з науковими програмами, планами, темами кафедри, університету.

Кваліфікаційна робота виконана в межах наукової тематики кафедри інформаційних технологій, штучного інтелекту і кібербезпеки НУХТ – «Дослідження та використання сучасних інформаційних технологій для виконання функцій та завдань виробничого і організаційного управління підприємств харчової галузі».

Завдання дослідження.

У межах роботи виконати аналіз сучасного стану ринку молочної продукції України, дослідити основні тенденції у виробництві та споживанні. Провести огляд сучасних методів і машинного навчання, орієнтованих на роботу з часовими рядами, зокрема алгоритму LSTM, та оцінити їх ефективність для розв'язання задач прогнозування в економічній галузі.

Мета дослідження є вивчення можливостей застосування сучасних методів машинного навчання для аналізу та прогнозування тенденцій на ринку молочної продукції України. Основна увага приділена використанню алгоритму LSTM для роботи з часовими рядами з метою створення інструменту, який дозволить автоматизувати аналіз великих обсягів даних, здійснювати точні прогнози та сприяти підвищенню ефективності управлінських процесів у молочній галузі.

Об'єкт дослідження. Споживання та виробництво молочної продукції в Україні.

Предмет дослідження. Використання методів машинного навчання, а саме алгоритм LSTM, для аналізу даних і прогнозування тенденцій на ринку молочної продукції.

Мета роботи. Дослідити можливості використання сучасних методів штучного інтелекту для аналізу ринку молочної продукції України.

Кваліфікаційна робота присвячена дослідженню ринку молочної продукції України із застосуванням методів штучного інтелекту, зокрема алгоритмів аналізу часових рядів, таких як LSTM. Метою роботи є створення програмного забезпечення, яке дозволить оцінити поточні тенденції та прогнозувати обсяги виробництва й споживання. Розроблений інструмент стане корисним як для підприємств молочної галузі, так і для аналітиків, допомагаючи приймати зважені рішення та адаптувати стратегії розвитку відповідно до змін на ринку.

Наукова новизна одержаних результатів роботи полягає в застосуванні алгоритму LSTM для прогнозування виробництва та споживання молочних продуктів в Україні, що забезпечує високу точність та ефективність у порівнянні з традиційними методами.

Практичне значення одержаних результатів полягає у створенні інструменту для планування виробничих ресурсів і розробки бізнес-стратегій. Розроблений продукт допоможе підприємствам та аналітикам точніше прогнозувати тенденції ринку та адаптувати свої стратегії до ринкових змін.

РОЗДІЛ 1. ДОСЛІДЖЕННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ

1.1. Актуальність дослідження

Агропромисловий комплекс (АПК) є ключовим сектором економіки України та багатьох країн світу, забезпечуючи населення як продуктами харчування, так і сировиною для різноманітних галузей промисловості. Особливу роль у цьому процесі відіграє м'ясо-молочна продукція, яка не лише задовольняє базові потреби людей, але й є важливим елементом національної продовольчої безпеки. Ефективне управління виробничими процесами в агропромислових компаніях значно впливає на стабільність цін на харчові продукти, а також на доступність якісної сировини для кінцевих споживачів. У таких умовах здатність адаптуватися до динамічних змін ринку стає вирішальним фактором для розвитку бізнесу та збільшення доходів підприємств.

Сучасний етап глобалізації, швидкий розвиток технологій та посилення конкуренції вимагають від виробників активного впровадження інноваційних рішень, зокрема технологій машинного навчання та аналізу даних. Ці інструменти дозволяють оптимізувати виробничі процеси, знижувати ризики та витрати, а також підвищувати якість продукції. Завдяки машинному навчанню підприємства можуть ідентифікувати приховані закономірності у великих масивах даних, оперативно приймати обґрунтовані управлінські рішення і швидше реагувати на зміни ринку.

М'ясні та молочні продукти традиційно займають важливе місце в раціоні людей і відіграють суттєву роль у структурі продовольчого ринку. Прогнозування попиту та споживання допомагає підприємствам забезпечувати оптимальний рівень виробництва, уникати дефіциту або надлишку продукції, що в кінцевому результаті сприяє стабільності цін та підвищенню прибутковості. До того ж, оптимізація виробничих і логістичних

процесів дає змогу знизити витрати на транспортування і забезпечити ефективне управління ресурсами.

1.2. Історія машинного навчання в промисловості

Машинне навчання бере свій початок у середині XX століття, коли вчені почали активно досліджувати можливості навчання машин на основі наявних даних. Одним із перших проривів у цій галузі стало створення перцептрону у 1950-х роках. Ця модель була розроблена для імітації роботи людського мозку. Значний внесок у її розвиток зробив Френк Розенблатт, який заклав фундамент для створення сучасних нейронних мереж. Перцептрон став ключовим етапом у розвитку технологій, на основі яких згодом було побудовано багатошарові нейронні моделі, що активно використовуються сьогодні.

У 1970–1980-х роках акцент у розвитку машинного навчання змістився на розробку алгоритмів, таких як дерева рішень, методи кластеризації та регресії. Особливе місце зайняв метод зворотного поширення помилки, який дозволив ефективно навчати багатошарові нейронні мережі. Цей підхід суттєво підвищив точність моделей, відкривши шлях для їх широкого використання у різних сферах діяльності.

Починаючи з 1990-х років, машинне навчання почало знаходити практичне застосування в різних галузях, включаючи агропромисловий комплекс. Зокрема, воно використовувалося для автоматизації виробничих процесів, оптимізації логістичних систем і управління запасами. Завдяки цим інструментам стало можливим ефективніше прогнозувати попит, оптимізувати технічне обслуговування обладнання та створювати системи підтримки прийняття рішень. Такі системи значно покращили економічну ефективність і конкурентоспроможність підприємств.

На початку XXI століття розвиток машинного навчання зіткнувся з новими викликами, пов’язаними з обробкою великих обсягів даних. Водночас суттєве зростання обчислювальних потужностей серверів і комп’ютерів

дозволило впоратися з цією проблемою. Використання великих масивів даних дало змогу створювати складніші архітектури нейронних мереж, підвищуючи їхню точність та ефективність. У результаті машинне навчання стало невід'ємною частиною таких галузей, як фінанси, охорона здоров'я, енергетика та агропромисловість.

Сьогодні машинне навчання є основою Індустрії 4.0 — четвертої промислової революції, яка передбачає повну автоматизацію та цифровізацію виробничих процесів. Ці технології забезпечують аналіз та прогнозування в реальному часі, сприяючи підвищенню ефективності виробництва, зменшенню витрат і покращенню якості продукції. Машинне навчання стає незамінним інструментом у сучасному світі, допомагаючи підприємствам залишатися конкурентоспроможними та адаптивними до стрімких змін ринку.

1.3. Постанова проблеми

Процес виробництва та реалізації молочної продукції супроводжується низкою значних викликів, які вимагають комплексного підходу та впровадження інноваційних технологій для їх ефективного подолання.

Основна мета цього дослідження — аналіз сучасних викликів у сфері виробництва та реалізації молочної продукції та об'єктивна оцінка можливостей їх подолання за допомогою інформаційних технологій. У рамках дослідження розглядаються основні проблеми, що впливають на ефективність виробництва молочної продукції, включаючи захворювання тварин та залежність від погодних умов. Також досліджується вплив сезонних і культурних коливань попиту на планування та логістику, аналізуються логістичні й управлінські аспекти, пов'язані з обмеженим терміном придатності продукції. Особливу увагу приділено забезпеченню високої якості продукції на всіх етапах виробничого процесу. Важливим елементом є оцінка ролі сучасних інформаційних систем, що допомагають автоматизувати процеси, аналізувати дані й прогнозувати попит і виробничі обсяги.

Однією з основних проблем є вразливість сільськогосподарських тварин до хвороб та епідемій, які можуть спричинити значні втрати у виробництві. До цього додається залежність виробничих процесів від погодних умов, які складно передбачити чи контролювати. Ігнорування цих чинників може призвести до зниження обсягів виробництва, що негативно позначається на постачанні продукції та стабільності ринку.

Ще одним важливим викликом є сезонні та культурні коливання попиту і цін. Наприклад, у США попит на індичку зростає під час Дня подяки, а в Україні в літній період спостерігається збільшення попиту на свинину для шашликів. Такі зміни у вподобаннях споживачів ускладнюють планування виробництва та логістики, особливо коли спостерігаються раптові зрушення в попиті на певні види продукції.

Ефективність логістики та управління ланцюгами постачання також залишається важливою складовою у виробництві молочної продукції. Через обмежений термін придатності та специфічні умови зберігання продукції потрібне ретельне планування всіх логістичних операцій. Збої в постачанні можуть спричинити фінансові втрати, затримки та зниження якості продукції, що вимагає оптимізації маршрутів, оновлення транспортних засобів і впровадження автоматизованих рішень у логістиці.

Забезпечення високої якості продукції є ключовим фактором для збереження конкурентоспроможності. Контроль якості на кожному етапі виробничого процесу та реалізації продукції є вирішальним завданням, адже саме він визначає довіру споживачів і стабільний попит.

У сучасних умовах подолання зазначених викликів стає можливим завдяки використанню сучасних інформаційних систем. Такі системи дають змогу автоматизувати збір і аналіз даних, проводити моніторинг стану тварин, контролювати кліматичні умови та враховувати інші важливі параметри, які впливають на ефективність виробництва. Застосування алгоритмів машинного навчання дозволяє аналізувати великі обсяги інформації, виявляти приховані

закономірності та будувати прогнози, що сприяє прийняттю обґрунтованих управлінських рішень.

Інформаційні системи також надають можливість створення моделей прогнозування врожайності, виробничих обсягів та попиту. Це допомагає підприємствам більш ефективно розподіляти ресурси, управляти запасами та оптимізувати виробничі процеси. Завдяки цьому компанії можуть зменшувати витрати, забезпечувати стабільність ринку та підвищувати свою конкурентоспроможність.

Впровадження сучасних інформаційних систем для аналізу та прогнозування у сфері виробництва і реалізації молочної продукції є критично важливим для успішного управління підприємствами. Інноваційні технології допомагають адаптуватися до динамічних умов ринку, підвищувати якість продукції та забезпечувати сталий розвиток і зростання бізнесу.

1.4. Методи та підходи дослідження

У сучасних умовах ринок молочної продукції України є складною та багатогранною системою, яка зазнає впливу економічних, соціальних та культурних чинників. Для забезпечення конкурентоспроможності виробники мають постійно адаптуватися до змін попиту, зважаючи на уподобання споживачів, ринкові тенденції та економічну ситуацію в країні. У цьому контексті важливим є розробка ефективних підходів до аналізу попиту та прогнозування його змін.

У сфері аналізу показників та метрик для дослідження попиту та виробництва існує велика кількість методів, які зазвичай використовуються маркетинговими відділами компаній. Для дослідження ринку молочної продукції в Україні застосовуються три основні методи: опитування, спостереження та експеримент. Розглянемо кожен з них і визначимо ефективність їхнього використання у межах даної роботи.

Опитування є одним із найцінніших методів збору інформації, оскільки дозволяє отримати різноманітні дані про вподобання клієнтів та їхні наміри. Цей метод особливо ефективний у стабільних ринкових умовах, де покупці демонструють сталість у своїх уподобаннях та здатні чітко визначити свої наміри на тривалий період. Виробництво молочної продукції є прикладом, де опитування є доцільним, оскільки більшість людей споживають молочні продукти протягом тривалого часу, і їхні уподобання змінюються рідко, у порівнянні з більш волатильними ринками, такими як ринок дитячих іграшок.

Проте, незважаючи на переваги, опитування може мати свої обмеження. Один із них — критична суб'єктивність відповідей, яка залежить від різних факторів, таких як регіональні особливості, етнічні та релігійні переконання, а також рівень матеріального достатку. Усі ці фактори можуть вплинути на точність даних, що збираються, особливо при проведенні опитувань у великих масштабах. Крім того, на достовірність результатів може вплинути бажання респондентів показати себе в кращому світлі, що може спотворити інформацію.

Метод спостереження передбачає збір та аналіз статистичних даних, що стосуються купівлі та реалізації товарів у минулому. Цей метод є дуже корисним у разі, якщо виробництво веде облік обсягу виробленої та реалізованої продукції. Спостереження дозволяє створити надійну статистичну базу, яка є основою для подальшого аналізу та прогнозування попиту.

У контексті цієї роботи метод спостереження є найбільш підходящим, оскільки він забезпечує об'єктивні дані про попит і пропозицію на ринку молочної продукції. Важливим аспектом є використання відповідних інструментів та підходів для збору, обробки та аналізу статистичних даних, які будуть розглянуті у наступних розділах.

Метод експерименту передбачає зміну окремих аспектів стратегії реалізації, реклами чи виробництва та збір даних для подальшого аналізу. Хоча експеримент може бути корисним для перевірки гіпотез і зміни трендів,

його застосування в даній роботі є недоцільним. Це зумовлено тим, що дослідження орієнтоване на прогнозування трендів без втручання в сам процес, що не відповідає основній меті експерименту.

Додатково для моделювання та прогнозування попиту застосовуються методи математичного та машинного навчання. Серед них:

— лінійна регресія;

— нелінійна регресія.

Лінійна регресія є одним із найбільш поширених і простих методів для прогнозування залежностей між змінними. Вона використовується для моделювання ситуацій, де одна змінна є незалежною, а інша залежною від неї. У рамках цієї роботи, лінійна регресія розглядалася як перший варіант для створення прогностичних моделей. Однак, незважаючи на свою простоту та ефективність у певних випадках, лінійна регресія не є найкращим вибором для цієї задачі.

Лінійна регресія є ефективною для створення простих моделей, але не підходить для складних функцій із багатьма змінними, особливо коли йдеться про прогнозування складних взаємозв'язків. Навіть множинна регресія, що розширює лінійну модель шляхом додавання декількох незалежних змінних, не вирішує проблему прогнозування складних залежностей. Важливим недоліком лінійної регресії є її обмеження у врахуванні нелінійних зв'язків між змінними.

Нелінійна регресія є більш підходящою для моделювання складних залежностей між змінними. Цей метод дозволяє враховувати нелінійні зв'язки, що значно покращує точність прогнозів у порівнянні з лінійною регресією. Зокрема, поліноміальна регресія може бути використана для моделювання змін у попиті на молочну продукцію, що враховують різні фактори впливу.

Однак, використання нелінійної регресії має свої обмеження. Перш за все, вона вимагає великої кількості даних для ефективного навчання, що є критичним аспектом у разі обмеженого доступу до даних. В умовах недостатньої кількості спостережень модель має високий ризик перенавчання,

що може призвести до надмірної чутливості до коливань у даних і погіршення якості прогнозування.

Крім того, вибір правильного функціонального виду залежності між змінними є важливою задачею. Неправильний вибір може призвести до зниження точності моделі, а також до труднощів у її інтерпретації та використанні для подальших прогнозів.

Для досягнення високої точності прогнозування попиту на молочну продукцію були розглянуті різні методи машинного навчання, які дозволяють будувати складні прогностичні моделі та враховувати багатозмінні взаємозв'язки. Зокрема, серед найбільш ефективних методів для цієї роботи можна відзначити наступні:

— рішення дерева;

— випадкові ліси;

— градієнтний бустинг;

— нейронні мережі.

Рішення дерева є одним із базових алгоритмів машинного навчання, що використовується для задач класифікації та регресії. Метод побудови моделі на основі рішень дерева передбачає розбиття даних на підмножини за допомогою умов на основі значень вхідних змінних. Кожна гілка дерева представляє собою вирішення певної умови, а кінцеві вузли (листи) показують кінцевий прогноз.

Переваги використання рішень дерева:

— простота інтерпретації: модель легко інтерпретується, що дозволяє зрозуміти, як змінні впливають на результат;

— гнучкість: здатність працювати з як числовими, так і категоріальними даними.

Недоліки:

— перенавчання: великі дерева схильні до перенавчання, що може знижувати їхню здатність до узагальнення на нових даних.

Випадкові ліси є вдосконаленням рішень дерева і представляють собою ансамблевий метод, який об'єднує результати декількох рішень дерев. Кожне дерево в лісі будується на випадковій підмножині даних, що дозволяє знижувати кореляцію між деревами та покращувати загальну ефективність моделі.

Переваги використання випадкових лісів:

— стійкість до перенавчання: завдяки ансамблевій природі метод має знижений ризик перенавчання;

— висока точність: забезпечує добрі результати навіть при великій кількості змінних;

— важливість змінних: можливість оцінки важливості кожної змінної для прогностичної моделі, що може допомогти у виборі ключових показників для аналізу.

Недоліки:

— обчислювальні витрати: при великій кількості дерев модель може вимагати значних обчислювальних ресурсів;

— інтерпретовані результати: хоч і кожне дерево інтерпретується окремо, загальна модель у вигляді випадкового лісу не є такою прозорою.

Градієнтний бустинг є потужним методом ансамблевого навчання, який створює послідовність слабких моделей (зазвичай рішень дерев), де кожна наступна модель намагається виправити помилки попередньої. Цей підхід дозволяє поступово покращувати точність моделі.

Переваги градієнтного бустингу:

— висока точність: цей метод забезпечує відмінні результати при використанні на реальних даних;

— гнучкість у налаштуванні: можливість налаштування різних параметрів моделі для досягнення оптимальних результатів;

— важливість змінних: як і у випадку зі случайними лісами, градієнтний бустинг дозволяє оцінити важливість змінних.

Недоліки:

— вразливість до перенавчання: хоча метод ефективний, він може перенавчитися при великій кількості параметрів, якщо не здійснюється належна регуляризація;

— час навчання: навчання градієнтного бустингу може займати більше часу через послідовний процес побудови моделей.

Нейронні мережі є одним із найпотужніших інструментів у машинному навчанні. Вони імітують роботу людського мозку, використовуючи багатошарові структури для виявлення складних залежностей у даних. У контексті прогнозування попиту на молочну продукцію нейронні мережі можуть бути надзвичайно корисними завдяки здатності працювати з великим обсягом даних і багатозмінними взаємозв'язками.

Типи нейронних мереж:

— згорткові нейронні мережі (CNN): застосовуються здебільшого для аналізу зображень, але можуть бути корисними у випадках, коли дані містять геопросторові особливості;

— рекурентні нейронні мережі (RNN): ідеально підходять для аналізу часових рядів, оскільки враховують залежності між даними у часі;

— глибокі нейронні мережі (DNN): дозволяють аналізувати багатовимірні та складні дані, що включають велику кількість змінних.

Переваги нейронних мереж:

— висока точність: здатність досягати вищих результатів у порівнянні з іншими методами;

— адаптивність: можливість роботи з даними різної природи, включаючи текст, числа та зображення;

— автоматичне визначення особливостей: мережі можуть автоматично визначати найважливіші ознаки даних.

Недоліки нейронних мереж:

— висока вимогливість до обчислювальних ресурсів: моделі потребують потужного обладнання для навчання;

— складність налаштування: багато параметрів, які потрібно налаштовувати;

— чутливість до обсягу даних: для навчання якісних моделей потрібна велика кількість якісних даних.

1.5. Висновки до першого розділу

Застосування різних методів та підходів до аналізу ринку молочної продукції в Україні дозволяє отримати комплексне уявлення про динаміку попиту. У цьому дослідженні акцент зроблено на використанні спостереження, нелінійної регресії та сучасних алгоритмів машинного навчання. Такі підходи сприяють більш точному прогнозуванню змін попиту та підтримують прийняття стратегічних рішень для підвищення ефективності виробництва.

РОЗДІЛ 2. ОПИС МАТЕМАТИЧНОЇ МОДЕЛІ ДЛЯ ПРОГНОЗУВАННЯ СПОЖИВАННЯ МОЛОЧНОЇ ПРОДУКЦІЇ

2.1. Основні методи та інструменти побудови моделі прогнозування споживання молочної продукції

У даному розділі розглядаються основні інструменти, що були застосовані для побудови математичної моделі прогнозування споживання молочної продукції на основі часових рядів. Описано обрані методи аналізу даних, а також програмні засоби, які використовувались для реалізації поставленої задачі.

Метою дослідження є побудова математичної моделі, що дозволяє прогнозувати обсяг споживання молочної продукції на основі історичних даних. Для цього використовується метод машинного навчання, здатний враховувати сезонні коливання, тренди та інші фактори, що можуть впливати на попит. Оскільки дані мають часову складову, особлива увага приділяється методам, що працюють з часовими рядами і дозволяють урахувати залежності між значеннями на різних етапах часу.

Модель була реалізована за допомогою мови програмування Python, яка є одним із найпоширеніших інструментів для роботи з даними та машинного навчання. Python надає широкий спектр бібліотек та інструментів для аналізу даних, побудови моделей машинного навчання та візуалізації результатів. Його переваги включають простий синтаксис, багатофункціональність та активну спільноту розробників.

Для реалізації проекту були використані наступні бібліотеки:

1. pandas: бібліотека для обробки та аналізу структурованих даних. Використовувалася для завантаження, очищення та попередньої обробки даних. Pandas забезпечує зручну роботу з табличними даними та дозволяє виконувати складні трансформації з мінімальними зусиллями;

2. numPy: забезпечувала ефективну роботу з числовими масивами, необхідними для створення послідовностей та маніпуляцій із даними. Її функціонал включає швидкі операції над багатовимірними масивами та матрицями, що критично важливо для роботи з великими наборами даних;

3. scikit-learn: використовувалася для масштабування даних за допомогою Min-Max Scaler, що забезпечило стабільну роботу моделі. Scikit-learn також підтримує інші методи попередньої обробки та аналізу даних;

4. keras: високорівнева бібліотека для роботи з нейронними мережами, що дозволила швидко створити й налаштувати архітектуру LSTM. Keras надає зручний інтерфейс для побудови складних моделей та інтеграцію з TensorFlow;

5. matplotlib: використовувалася для візуалізації результатів прогнозування та аналізу точності моделі. Бібліотека дозволяє створювати високоякісні графіки, що сприяє кращому розумінню даних.

2.2. Вибір математичної моделі для прогнозування

У цьому розділі розглядаються основні інструменти та методи, які були застосовані для побудови математичної моделі прогнозування споживання молочної продукції на основі часових рядів. Описано обрані методи аналізу даних, зокрема використання рекурентних нейронних мереж (РНН), а також програмне забезпечення, яке використовувалось для реалізації поставленої задачі.

Метою дослідження є побудова математичної моделі, здатної прогнозувати обсяг споживання молочної продукції на основі історичних даних. Для цього застосовується метод машинного навчання, зокрема рекурентні нейронні мережі (RNN), які є ефективними для роботи з часовими рядами. RNN дозволяють моделювати залежності між значеннями на різних етапах часу, враховуючи як короткострокові, так і довгострокові тренди,

сезонні коливання та інші фактори, що впливають на попит на молочну продукцію.

Рекурентні нейронні мережі є потужним інструментом у сфері обробки послідовних даних, таких як текст, аудіо та відео. Основною особливістю RNN є їх здатність зберігати інформацію про попередні етапи, що дає змогу ефективно працювати з даними, де важливий контекст. Ця властивість робить RNN незамінними для задач, де порядок елементів у послідовності має велике значення.

Основи роботи з RNN.

У звичайних нейронних мережах кожен елемент вхідних даних передається окремо через шар. Водночас у RNN кожен вхід обробляється послідовно, і результат обчислень на поточному етапі залежить від стану попереднього етапу, що дозволяє мережам зберігати інформацію про контекст.

Переваги використання RNN.

RNN мають багато переваг, зокрема здатність працювати з послідовними даними та враховувати контекст. Це робить їх ефективними для застосування в задачах, таких як машинний переклад, генерація тексту, розпізнавання мови та аналіз часових рядів. Вони можуть використовуватись для прогнозування фінансових показників, прогнозування погоди, аналізу попиту на продукцію тощо.

Проблеми використання RNN.

Проте звичайні RNN мають суттєві обмеження. Основними з них є:

— вибух градієнта: під час зворотного поширення помилки градієнти можуть швидко ставати занадто великими, що призводить до нестабільності навчання;

— згасання градієнта: градієнти можуть зменшуватися до дуже малих значень, що ускладнює навчання, особливо при обробці довгих послідовностей;

— обмежена пам'ять: звичайні RNN не можуть ефективно зберігати інформацію про довгострокові залежності через зазначені проблеми.

Модифікації RNN.

Для подолання проблем звичайних RNN були розроблені удосконалені версії, зокрема:

— LSTM (Long Short-Term Memory) — це спеціальний тип рекурентної мережі, який використовує "ворота" (gates) для контролювання потоку інформації. Вони дозволяють LSTM ефективно зберігати інформацію на довгий час і обробляти довгострокові залежності.

— GRU (Gated Recurrent Unit) — спрощена версія LSTM з двома ворітьми (обновлення та скидання). GRU має менш складну структуру, але при цьому здатна досягати подібних результатів, що й LSTM, з меншою обчислювальною витратою.

Застосування RNN.

RNN використовуються в різних галузях, таких як:

— обробка природної мови: RNN є основою для машинного перекладу, створення чат-ботів, аналізу настроїв і генерації тексту;

— аналіз часових рядів: вони знаходять застосування у прогнозуванні фінансових ринків, прогнозуванні погоди та інших випадках, де важлива послідовність подій;

— обробка аудіо: RNN використовуються для розпізнавання мови, генерації музики та обробки звукових сигналів.

Порівняння RNN з іншими архітектурами.

Хоча LSTM і GRU значно покращують ефективність RNN у роботі з довгостроковими залежностями, сучасні моделі, такі як трансформери, стали новим стандартом у обробці послідовностей. Завдяки механізму уваги (attention mechanism) трансформери забезпечують паралельну обробку даних, що робить їх дуже ефективними для великих обсягів інформації.

LTSM як вибрана модель для прогнозування.

Для прогнозування споживання молочної продукції було розглянуто модель довготривалої короткочасної пам'яті (Long Short-Term Memory, LSTM), яка є розширенням рекурентних нейронних мереж (RNN). Ця модель

дозволяє ефективно працювати з часовими рядами, зберігаючи інформацію про довгострокові та короткострокові залежності в даних. Використання LSTM виправдане задачами прогнозування, де важливо враховувати тренди та сезонні коливання, що притаманні ринку молочної продукції.

Основи роботи моделі LSTM.

LSTM була розроблена для подолання обмежень класичних RNN, зокрема проблеми зникнення або вибуху градієнтів при навчанні довгих послідовностей. Основним елементом моделі є осередок пам'яті (*cell state*), який зберігає інформацію протягом тривалого часу та динамічно оновлюється через три ключові механізми.

На рисунку 2.1 наведено схему обробки нейронною мережею RNN послідовності даних.

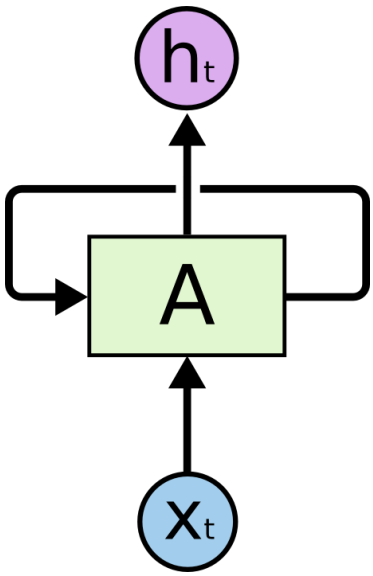

Рисунок 2.1 — Фрагмент нейронної мережі LSTM

x_t —отримана інформація, що надходить у блок А, який є процесором, що її обробляє. Для обробки інформації блок А бере до уваги не лише отримані дані, а й ті, що вже були отримані та збережені в h_{t-1} до цього.

Блок А об’єднує нову інформацію x_t з попередньою пам’яттю h_{t-1} і створює оновлену пам’ять h_t. Ця пам’ять зберігає контекст історії, який потім передає далі, щоб обробляти наступну інформацію або фрагмент даних.

Таким чином, RNN дозволяє «згадувати» інформацію, що була отримана до останньої операції, і використовувати її для розуміння або прогнозування майбутнього, аналізу тексту, прогнозування продажів чи обробки мови.

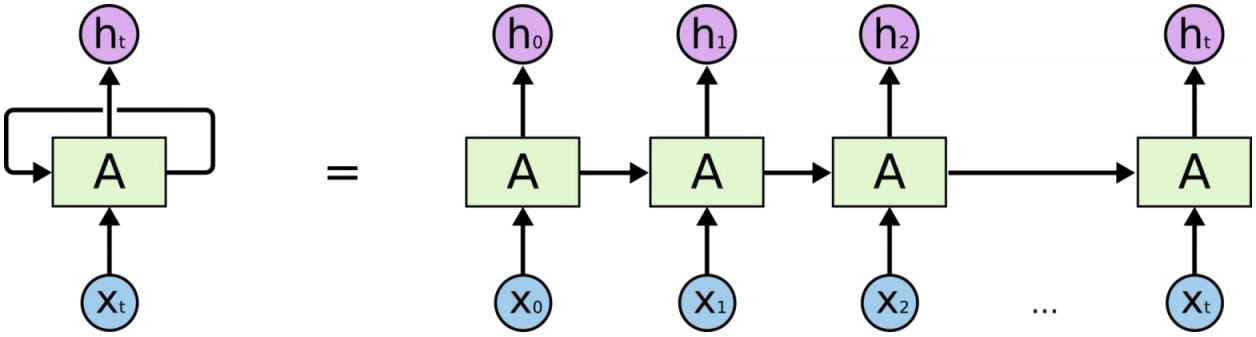

Рисунок 2.2 — Рекурентна нейронна мережа в розгортці

У звичайних рекурентних мережах виникає проблема обробки довгострокових залежностей. Однак, мережі типу LSTM ефективно вирішують цю проблему.

Будь-яка рекурентна нейронна мережа має структуру у вигляді ланцюга, що продемонстрований на рисунку 2.2, і складається з однакових модулів. У звичайних рекурентних мережах ці модулі мають досить просту будову та часто складаються з одного шару з функцією активації tanh.

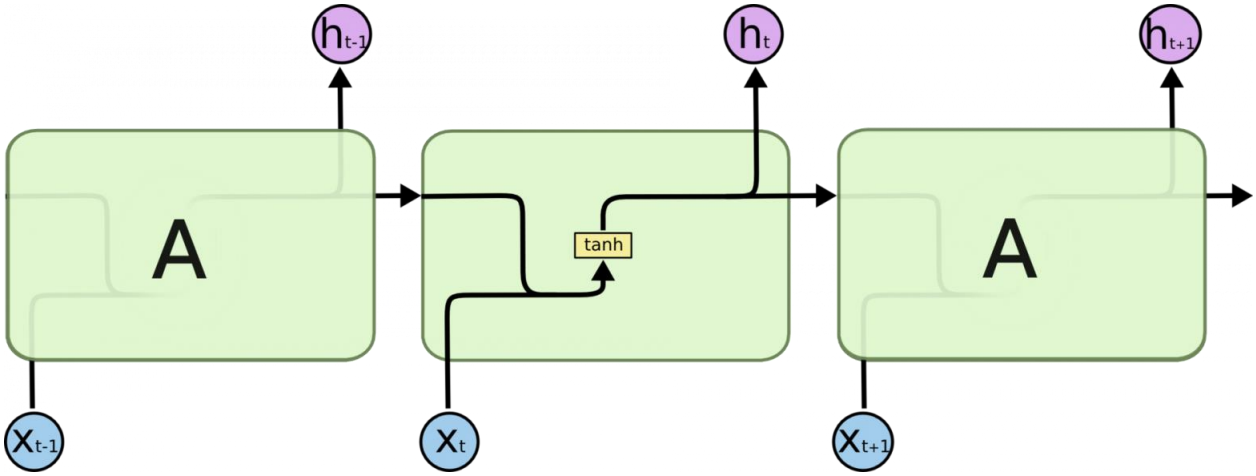

Рисунок 2.3 — Проста архітектура рекурентної нейронної мережі

Однак, LSTM мережа має чітку структуру, яка представлена на рисунку 2.3, що замість одного шару має 4, що взаємодіють особливим чином.

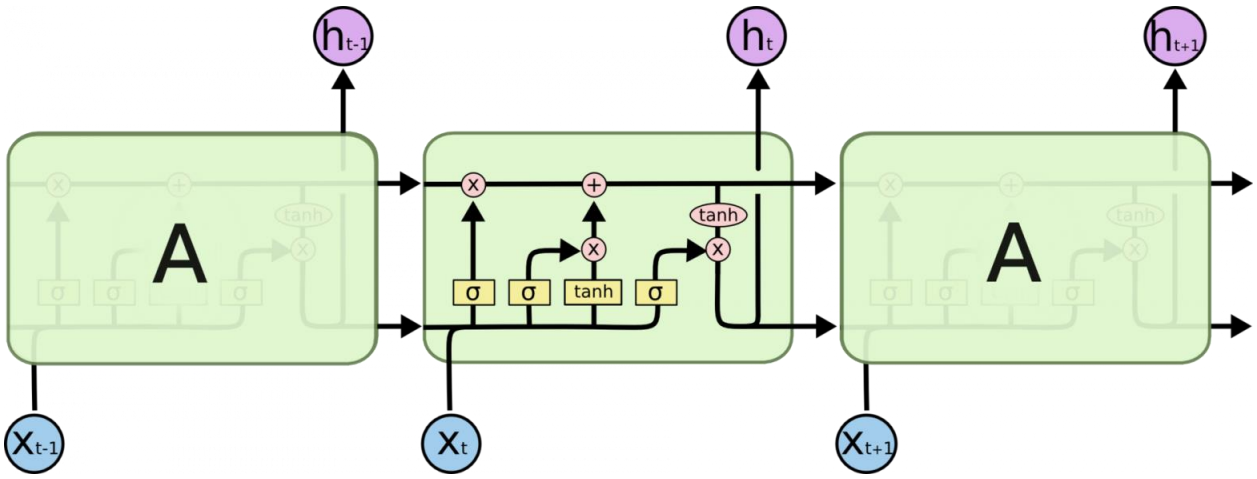

Рисунок 2.4 — Схематичне зображення архітектури LSTM мережі

Головна концепція LSTM полягає в cell state, шо проходить напряму через весь ланцюг та використовується лише в декількох лінійних перетвореннях, передаючи інформацію без її змін, як зображено на рисунку 2.4.

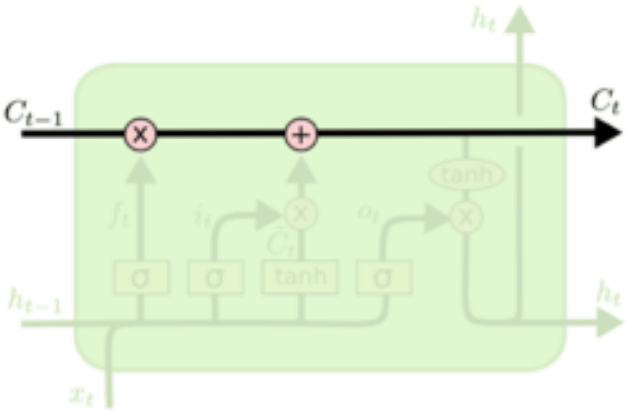

Рисунок 2.5 — Cell state в архітектурі LSTM

Інформація у таких мережах потребує модифікації, навіть враховуючи описану вище архітектуру. Цей процес контролюється спеціальними фільтрами, які називаються воротами (gates). Ворота працюють на основі певних умов, пропускаючи або блокуючи інформацію. Їхня структура включає шар сігмоїдальної нейронної мережі та операцію множення.

Сігмоїдальний шар генерує значення в діапазоні від 0 до 1, яке визначає, яка частина інформації буде передана до стану комірки (cell state). Значення 1 означає, що інформація пропускається повністю (100%), а 0 — що вона повністю блокується (0%).Загалом, LSTM має 3 таких фільтри, що дозволяють захищати стан клітинки від непотрібної інформації.

Отже, першим кроком в даній архітектурі є необхідність вирішити які дані нейронна мережа має пропустити, а які зберегти для подальших операцій. Це рішення приймає сігмоїдальний шар, зображений на рисунку 2.5, який також називається forget gate layer.

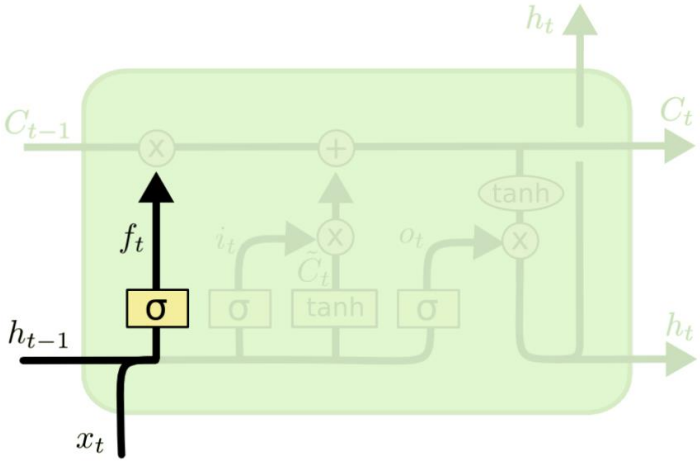

Рисунок 2.6 — Перший крок, фільтрація

Можна побачити, що під час даного кроку фільтрації використовується 2 вхідних джерела інформації, саме на основі h_{t-1} та x_t *forget gate layer* приймає рішення з допомогою чисел із інтервалу (0, 1) скільки інформації має перейти в даний сегмент. Загалом, цей процес можна описати наступною формулою (2.1):

$$f_t = \sigma\left(W_t \cdot [h_{t-1}, x_t] + b_f\right) \quad (2.1)$$

Наступним кроком є рішення щодо інформації, яка має зберігатись в даному вузлі, це відбувається у 2 етапи, як представлено на рисунку 2.6:

1. Input layer gate визначає які значення необхідно оновити (2.2);
2. шар з функцією активаці tanh будує вектор нових значень-кандидатів $\tilde{C}_t$ (2.3).

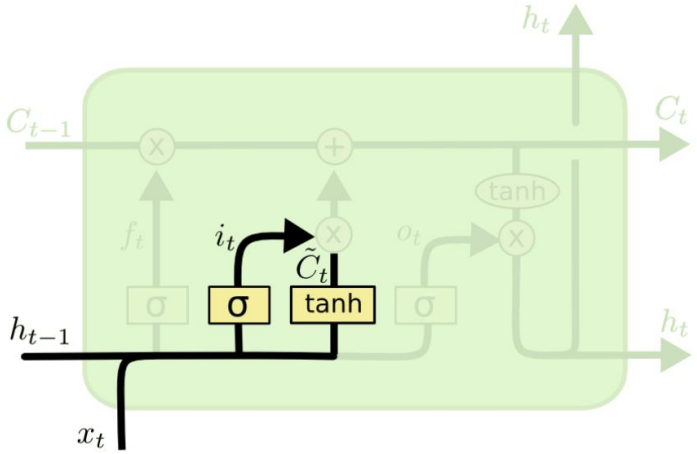

Рисунок 2.7 — Крок оновлення даних вузла

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2.2}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{2.3}$$

На рисунку 2.7 зображено, як відбувається заміна інформації з C_{t-1} на C_t. Для цього попереднє значення множиться на певний коефіцієнт та додається нова інформація (2.4).

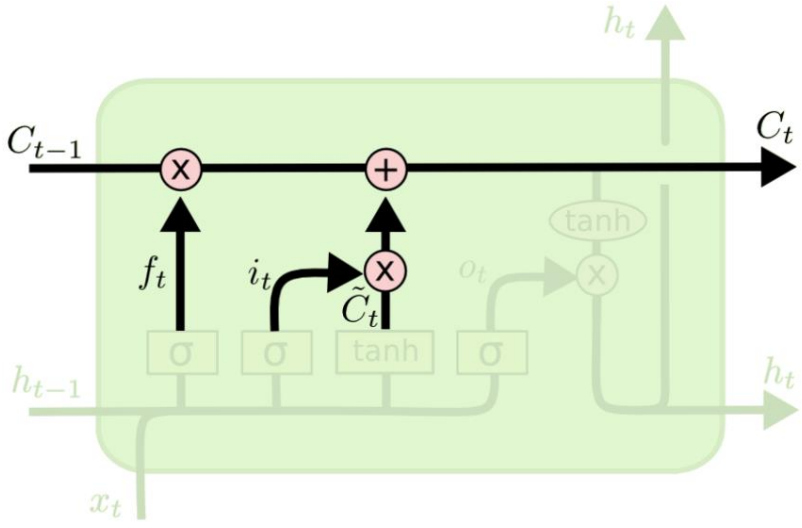

Рисунок 2.8 — Збереження інформації в вузол

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{2.4}$$

В останньому етапі, який можна побачити на рисунку 2.8 та на рисунку 2.9, відбувається передача інформації з вузла до наступного. Для цього визначається, які саме дані мають бути на виході, базуючись на стані вузла та з використанням фільтра. Спершу застосовується сигмоїдальний шар, який визначає, яку інформацію слід передати далі (2.5). Потім отримані дані проходять через шар tanh, що перетворює їх у діапазон (-1, 1), і перемножуються із результатами сигмоїдального шару (2.6).

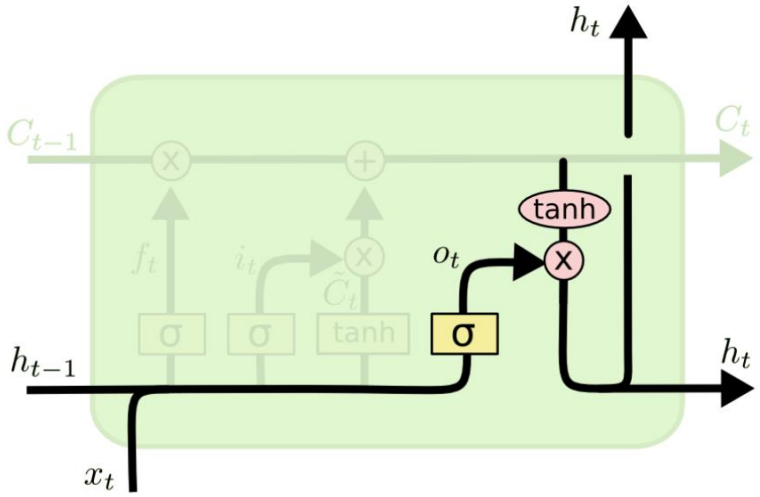

Рисунок 2.9 — Фінальний крок обробки даних вузлом

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{2.5}$$

$$h_t = o_t * \tanh(C_t) \tag{2.6}$$

Альтернативні архітектури для прогнозування

Окрім LSTM, для аналізу часових рядів можуть застосовуватися такі архітектури:

- автокодери;
- ARIMA;
- 1D-CNN;
- Transformer.

Автокодери — тип штучної нейронної мережі, який використовується для навчання ефективного кодування даних. Вони працюють у рамках без

використання розмічених даних, і призначені для зменшення розмірності, відновлення даних або виявлення ключових особливостей у вхідній інформації. Основна ідея автокодера полягає в тому, щоб стискати дані у компактне представлення (кодування) і потім відновлювати їх якомога ближче до початкових.

Автокодер складається з двох основних компонентів:

1. енкодер (Encoder):

– перетворює вхідні дані у компактний вектор фіксованої розмірності, який називається латентним простором або кодуванням;

– ця частина автокодера поступово зменшує розмірність вхідних даних шляхом застосування лінійних і нелінійних перетворень.

2. декодер (Decoder):

– відновлює вихідні дані з латентного простору. Мета полягає в тому, щоб отриманий вихід був максимально схожим на оригінальні дані;

– у декодері виконується обернена операція енкодера.

Автокодери підходять для виявлення прихованих структур у даних, але мають обмеження в роботі зі складними залежностями у часових рядах.

ARIMA — це статистичний метод для моделювання та прогнозування часових рядів. Він є одним із найбільш популярних підходів у сфері аналізу часових рядів завдяки своїй гнучкості та простоті використання. ARIMA поєднує три основні компоненти, які дозволяють ефективно моделювати та прогнозувати дані:

Основні компоненти ARIMA:

1. AR (Autoregressive, авторегресія): цей компонент моделі використовується для вираження залежності поточного значення часового ряду від попередніх значень того самого ряду. У моделі авторегресії використовується коефіцієнти для створення лінійної комбінації попередніх значень;

2. I (Integrated) - це компонент, який використовується для стабілізації часу ряду. Інтеграція означає взяття різниці між послідовними значеннями часового ряду, щоб зробити його стаціонарним (тобто щоб його статистичні властивості не змінювались із часом). Це необхідно, оскільки багато статистичних моделей, включаючи ARIMA, припускають, що вхідні дані є стаціонарними;

3. MA (Moving Average): цей компонент моделі враховує вплив випадкових помилок (шуму) на поточне значення часового ряду. Ковзаюче середнє виражає залежність поточного значення від середнього значення попередніх помилок.

ARIMA потужний інструмент для прогнозування часових рядів і має багато застосувань у різних галузях. Однак ефективність залежить від правильного вибору параметрів і передумов, таких як стаціонарність даних.

1D-CNN — це тип згорткових нейронних мереж, які спеціально призначені для роботи з одновимірними даними. Дані мережі широко використовуються для обробки послідовних даних, таких як часові ряди, текст або будь-які інші типи даних, які мають структуровану послідовність.

Основні компоненти 1D-CNN:

1. згортковий шар: основний компонент, який застосовує фільтри до вхідних даних і створює нові ознаки. Фільтри, зазвичай, мають малий розмір, але можуть мати різну глибину;

2. шар активації: використовується для додавання нелінійності в модель;

3. шар підсумовування (Pooling layer): зменшує розміри ознак (часто застосовується max-pooling), щоб зменшити обчислювальні витрати і зберегти найважливіші ознаки;

4. повнозв’язний шар (Dense layer): застосовується для інтеграції та обробки інформації перед виходом моделі;

5. шар виходу: формує кінцевий результат, який може бути ймовірністю для класифікації, прогнозованим значенням тощо.

1D-CNN забезпечує ефективний аналіз локальних залежностей у даних, однак не можуть враховувати довгострокові тренди без комбінації з іншими моделями.

Transformer — це архітектура нейронної мережі, яка була представлена у роботі "Attention is All You Need" у 2017 році і відтоді значно змінила підхід до обробки послідовних даних у сферах обробки природної мови (NLP) та інших галузях штучного інтелекту.

Transformer складається з двох основних компонентів:

1. кодувальник (Encoder): обробляє вхідну послідовність і створює контекстуальне представлення для кожного елементу;
2. декодувальник (Decoder): використовується для генерації вихідної послідовності на основі представлення, отриманого від кодувальника.

Кожен з блоків складається з кількох однакових шарів, що включають:

— механізм уваги (Attention Mechanism): дозволяє моделі фокусуватися на різних частинах входу при обробці кожного елементу. В основі лежить механізм self-attention, який дозволяє кожному слову в послідовності звертатися до інших слів, що дає змогу моделі враховувати важливість різних слів у контексті;

— шари нормалізації і додавання (Layer Normalization & Additive Layer): підвищують стабільність навчання;

— мережі прямого зв'язку (Feed-Forward Networks): використовуються для обробки представлень на кожному етапі.

Transformer сучасна архітектура, що використовує механізми уваги (*attention*), дозволяючи ефективно аналізувати довгі послідовності.

LSTM було обрано через її здатність працювати з нелінійними залежностями, ефективно враховувати довгострокові тренди та адаптуватися до змін у даних, що робить її більш гнучкою порівняно з ARIMA чи CNN.

Архітектура моделі

Для реалізації прогнозування було обрано архітектуру, що складається з:

— вхідного шару, який приймає нормалізовані значення часового ряду;

— LSTM-шару з 50 нейронами, який виконує аналіз послідовностей даних;

— щільного шару (*Dense layer*), що забезпечує остаточний прогноз.

Функція втрат моделі — середньоквадратична похибка (MSE), яка мінімізує різницю між реальними та прогнозованими значеннями. Для оптимізації ваг використовується алгоритм Adam, який забезпечує швидку й стабільну збіжність.

Порівняння з альтернативними методами.

- ARIMA: хоча цей метод забезпечує хороші результати для лінійних стаціонарних рядів, він не враховує нелінійності та довгострокові залежності, що обмежує його застосування для прогнозування споживання молочної продукції.
- 1D-CNN: ефективний для виявлення локальних залежностей, але для аналізу довгострокових трендів вимагає комбінації з іншими моделями, такими як RNN чи LSTM.
- Transformer: забезпечує високу точність і гнучкість, однак його використання потребує значних обчислювальних ресурсів, що може бути обмеженням у реальних умовах.

Вибір Python як мови програмування для реалізації моделі обумовлений широкими можливостями для роботи з даними, числовими масивами та машинним навчанням, що забезпечується через бібліотеки, такі як pandas, numpy, scikit-learn, keras і matplotlib. Завдяки їм вдалося створити ефективну модель, яка надає точні прогнози, враховуючи всі основні фактори, що впливають на попит на молочну продукцію.

2.3. Висновки до другого розділу

У другому розділі побудована математична модель на основі алгоритму LSTM, що є потужним інструментом для прогнозування споживання молочної продукції, здатним забезпечити точні прогнози та ефективно враховувати залежності між значеннями в часових рядах.

Застосування даного підходу дозволяє харчовим підприємствам приймати більш обґрунтовані рішення щодо планування виробництва молочної продукції, оптимізувати стратегій збуту на ринку молочних продуктів України.

РОЗДІЛ 3. АНАЛІЗ ТА МОДЕЛЮВАННЯ ПРОГНОЗІВ ДОДАНОЇ ВАРТОСТІ У ВИРОБНИЦТВІ МОЛОЧНОЇ ПРОДУКЦІЇ

3.1. Опис вхідних даних для аналізу і прогнозування

Для аналізу та прогнозування доданої вартості у сфері перероблення молока та виробництва масла і сиру були використані дані, зібрані з офіційних джерел, доступних у відкритому доступі. Оскільки інформація про детальні фінансові показники для провідних молочних підприємств часто є конфіденційною, використано дані з публічних ресурсів, що надають офіційну статистику по регіонах України.

Дані були отримані переважно з веб-ресурсу Державної служби статистики України (Держстат), який є основним джерелом економічної статистики в Україні[2].

У роботі використані дані за період 2014–2023 років, що охоплюють показники доданої вартості в регіонах України у сфері перероблення молока та виробництва масла і сиру. Дані включають загальний обсяг доданої вартості по Україні та по окремих областях[2].

Я вважаю ці дані найбільш інформативними та релевантними для аналізу, оскільки вони дають повне уявлення про зміну економічної діяльності в даній сфері за різні роки та допомагають зробити точні прогнози та аналіз тенденцій. На рисунку 3.1 наведено приклад частини даних у вигляді таблиці, яка ілюструє показники доданої вартості за роками та регіонами.

	Регіон	Додана вартість за витратами виробництва суб'єктів господарювання/ *Value added of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
виробництво молочних продуктів	Україна	5988741.7	5666211.7	7599253.1	10488803.9	12062249.2	13179487.8	13886398.5	13482159.1	14835865.2	17384882.5
	Вінницька область	675849.1	554061.7	1074815.9	1100397.4	1413093.8	1788138.6	1819450.2	1093179.9	2042182.5	1825340.9
	Волинська область	10209.3	6431.4	2180.2	к/с	54802.4	2311.2	к/с	к/с	34290.9	104058.9
	Дніпропетровська область	314795.7	238051.9	309661.8	к/с	550925.8	1681532.4	к/с	916868.6	к/с	к/с
	Донецька область	433273.9	80696.1	68199.1	к/с	к/с	72450.1	77141.4	82842.5	7787.3	к/с
	Житомирська область	516783.7	446829.4	425552.1	817426.2	883449.2	875225.6	799411.7	961984.7	к/с	1549008.3
	Закарпатська область	к/с	698.0	1236.8	1782.5	к/с	к/с	к/с	2268.5	к/с	2560.2
	Запорізька область	98628.2	34399.5	57849.3	69660.6	57826.6	81112.1	75343.7	68673.1	к/с	51260.5
	Івано - Франківська область	51721.5	16165.5	34814.3	135216.1	к/с	к/с	103026.5	141892.8	128219.4	245531.2
	Київська область	1417082.9	1723607.4	1303364.2	1104031.0	1345026.2	1388510.9	к/с	1577394.7	1328757.0	2282034.7
	Кіровоградська область	к/с	26236.0	к/с	к/с	к/с	к/с	250759.7	237857.5	к/с	100092.4
	Луганська область	20213.2	к/с	к/с	26157.4	33715.1	к/с	к/с	к/с	к/с	к/с
	Львівська область	178995.6	123939.9	270518.4	363649.2	455252.1	к/с	к/с	574725.4	к/с	к/с
	Миколаївська область	к/с	186672.6	285365.6	772504.1	398867.8	412412.9	734543.5	908439.8	к/с	к/с
	Одеська область	43531.8	34305.7	68862.5	76439.5	84468.7	111377.4	179482.6	147295.4	к/с	251912.8
	Полтавська область	484791.4	559044.7	668280.0	988789.5	645976.8	387774.7	443606.9	681646.9	886402.8	793460.5
	Рівненська область	216325.3	245955.2	341188.0	439327.3	к/с	к/с	к/с	743360.5	759935.5	938398.1
	Сумська область	50503.5	125816.5	108128.0	210707.6	149878.0	186151.7	188010.6	67105.9	к/с	к/с
	Тернопільська область	249754.6	176736.6	275460.8	402577.7	445348.9	505055.1	606310.8	537895.1	к/с	к/с
	Харківська область	159838.5	225812.8	255339.7	596506.0	473287.4	618825.5	879297.9	1384698.7	1297588.5	618970.3
	Херсонська область	141709.7	к/с	272941.5	к/с	к/с	781705.2	к/с	к/с	6661.0	12338.4
	Хмельницька область	-152641.2	-150234.2	-37988.9	-497502.1	753002.0	792268.2	1071868.5	333969.1	к/с	к/с
	Черкаська область	124002.5	91021.8	430746.0	797633.3	770257.9	421597.8	611385.1	383251.3	840367.6	998346.6
	Чернівецька область	-1376.2	138.0	-2258.6	к/с	778.5	к/с	7950.2	10282.2	к/с	53372.8
	Чернігівська область	270633.2	196615.7	486750.3	569581.8	382569.0	139845.8	407670.5	271482.4	621329.5	272488.0
	м. Київ	492866.1	596473.7	785941.8	1255117.8	1204475.4	1578091.9	1248648.5	1511622.3	2244454.2	2799084.4

Рисунок 3.1 — Масив даних, який містить інформацію щодо виготовлення молочної продукції в Україні

Розглянемо структуру масиву даних який містить інформацію щодо виготовлення молочної продукції в Україні:

— регіон: назва області України або загальний показник для країни;

— рік: рік, для якого представлені дані;

— додана вартість (тис. грн): обсяг доданої вартості, розрахованої на основі витрат виробництва підприємств у відповідному регіоні та році.

```
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Year    10 non-null     int32
 1   Value   10 non-null     float64
dtypes: float64(1), int32(1)
memory usage: 252.0 bytes
```

Рисунок 3.2 — Структура даних

Основні характеристики:

— охоплення: дані включають показники по всіх областях України, а також для м. Київ та загальноукраїнські показники;

— часовий інтервал: показники за періоди з 2014 по 2023 рік, що дозволяє аналізувати зміни протягом 10 років;

— формат даних: дані представлені у вигляді таблиці з числовими показниками для кожного року та регіону;

— якість даних: дані зібрані на основі офіційної статистики, що гарантує їх достовірність, проте у деяких випадках інформація відсутня або відображена як "к/с" (не доступно).

3.2. Збереження даних та забезпечення їх доступності

Збереження даних є важливим етапом у процесі роботи з інформацією, особливо при використанні їх у подальшому аналізі або моделюванні. Для цього застосовуються різноманітні формати файлів, які забезпечують зручність використання та надійність зберігання. Одним із найпопулярніших форматів є Excel (формат файлів .xlsx), який забезпечує структуроване представлення даних у вигляді таблиць.

Розглянемо вибір формату для збереження даних.

Формат Excel (.xlsx) є оптимальним для багатьох задач завдяки наступним перевагам:

— зручність у використанні: файли Excel можна легко відкривати, редагувати та переглядати за допомогою програмного забезпечення, такого як Microsoft Excel, Google Sheets або LibreOffice Calc.

— підтримка структурованих даних: табличний формат дозволяє впорядковувати дані за рядками та стовпцями, що сприяє легкому доступу до потрібної інформації.

— можливість збереження метаінформації: формат Excel дозволяє додавати заголовки, примітки та формули, які можуть бути корисними під час роботи з даними.

— сумісність із популярними бібліотеками: формат підтримується багатьма мовами програмування, включаючи Python (через бібліотеки, такі як pandas і openpyxl).

3.3. Розробка та навчання моделі

Для початку створення моделі необхідно імпортувати необхідні інструменти, які знадобляться нам у процесі розробки.

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import MinMaxScaler
from keras.models import Sequential
from keras.layers import LSTM, Dense
import matplotlib.pyplot as plt
```

Рисунок 3.3 — Імпорт бібліотек

На рисунку 3.3 показано використання pandas та numpy: для роботи з таблицями, числами, та виконання базових обчислень. MinMaxScaler - інструмент для приведення чисел у діапазон від 0 до 1. Це потрібно, щоб модель швидше і точніше навчалася.

Keras – бібліотека яка була використана для створення та навчання нейронних мереж.

Sequential - дозволяє створювати нейронну мережу, додаючи шари послідовно. LSTM - спеціальний тип шару нейронної мережі, який добре працює із часовими рядами. Dense - звичайний шар нейронів, що прогнозує одне значення. matplotlib.pyplot – бібліотека що використовується для створення графіків.

```
file_path = 'C:/diplom/statistic.xlsx'
data = pd.ExcelFile(file_path)


sheet1_data = data.parse('ДВМП')
```

Рисунок 3.4 — Завантаження даних

На рисунку 3.4 показано, що дані зберігаються у файлі Excel за адресою «C:/diplom/statistic.xlsx». pd.ExcelFile дозволяє працювати з таблицями Excel. У нашому випадку вибрано лист із назвою ДВМП.

```
ukraine_data = sheet1_data.iloc[2, 2:].reset_index(drop=True)
years = sheet1_data.iloc[0, 2:].reset_index(drop=True)
```

Рисунок 3.5 — Попередня обробка даних

Код, представлений на рисунку 3.5, виконує операції з даними, зокрема обробку та вибір певних рядків з об'єкта sheet1_data, отриманого з Excel-файлу. У результаті утворюється новий DataFrame під назвою ukraine_data.

```
ukraine_df = pd.DataFrame({
    "Year": years,
    "Value": ukraine_data
})
```

Рисунок 3.6 — Формування таблиці та очищення

Наведений на рисунку 3.6 фрагмент коду відповідає за формування таблиці з даних які були сформовані у списки для подальшого очищення та використання.

```
ukraine_df = ukraine_df.dropna()
ukraine_df["Year"] = ukraine_df["Year"].astype(int)
ukraine_df["Value"] = pd.to_numeric(ukraine_df["Value"], errors='coerce')
ukraine_df = ukraine_df.dropna().reset_index(drop=True)
```

Рисунок 3.7 — Процес очищення даних

Зображений на рисунку 3.7 код відповідає за очищення даних, видаляються пусті або некоректні рядки. Також він типізує дані, «Year» - перетворює на цілі числа, а «Value» - стає числом за плаваючою точкою. У результаті ми отримуємо таблицю з підготовленими для аналізу даними.

```
scaler = MinMaxScaler(feature_range=(0, 1))
ukraine_df["Scaled_Value"] = scaler.fit_transform(ukraine_df[["Value"]])


print("До масштабування:")
print(ukraine_df["Value"].describe())
print("Після масштабування:")
print(ukraine_df["Scaled_Value"].describe())
```

Рисунок 3.8 — Масштабування даних

Мета коду представленого на рисунку 3.8 - привести значення до діапазону [0, 1] за допомогою функції MinMaxScaler. Це потрібно для того, щоб нейронна мережа ефективніше працювала, оскільки у нас маштабні дані. У результаті ми отримуємо нову колонку «Scaled_Value», із нормалізованими значеннями.

```
def create_sequences(data, sequence_length=3):
    x, y = [], []
    for i in range(len(data) - sequence_length):
        x.append(data[i:i + sequence_length])
        y.append(data[i + sequence_length])
    return np.array(x), np.array(y)

sequence_length = 3
data = ukraine_df["Scaled_Value"].values
X, y = create_sequences(data, sequence_length)
```

Рисунок 3.9 — Формування послідовностей

Фрагмент коду зображений на рисунку 3.9 відповідає за створення послідовностей для навчання моделі. «X» – вводить послідовність із трьох років, «Y» – наступне значення після цих трьох років які є ціллю. У результаті ми отримуємо набір входів і відповідних цільових значень для навчання моделі.

```
X = X.reshape((X.shape[0], X.shape[1], 1))


model = Sequential([
    LSTM(50, activation='relu', input_shape=(X.shape[1], X.shape[2])),
    Dense(1)
])


model.compile(optimizer='adam', loss='mse')


print("Навчання моделі...")
history = model.fit(X, y, epochs=200, verbose=1)
```

Рисунок 3.10 — Побудова та навчання моделі

Представлений на рисунку 3.10 код взаємодіє з даними «X» які мають форму: (кількість прикладів, довжина послідовності) та перетворює їх у потрібний для роботи з LSTM: (кількість прикладів, довжина послідовності, кількість ознак). В результаті отримаємо (n_samples, sequence_lenght, 1). Наступним кроком є створення моделі LSTM, вона здатна запам`ятовувати довготривалі залежності в даних завдяки спеціальним коміркам пам`яті. Модель має 50 нейронів та вхідний формат: (довжина послідовності, кількість ознак). Також використовується Dense-шар для передбачення наступного числа. В результаті проводиться навчання моделі, відбувається 200 ітерацій під час яких модель бачить весь набір даних, поступово модель оновлює дані та результати, намагаючись зменшити втрати(MSE).

```
predicted = model.predict(X)


predicted_rescaled = scaler.inverse_transform(predicted)


result_df = ukraine_df.iloc[sequence_length:].copy()
result_df["Predicted"] = predicted_rescaled


future_years = list(range(ukraine_df["Year"].iloc[-1] + 1, 2026))
last_sequence = data[-sequence_length:]

future_predictions = []
for year in future_years:

    last_sequence_reshaped = last_sequence.reshape(1, sequence_length, 1)
    next_pred = model.predict(last_sequence_reshaped)
    future_predictions.append(next_pred[0, 0])


    last_sequence = np.append(last_sequence[1:], next_pred)


future_predictions_rescaled = scaler.inverse_transform(np.array(future_predictions).reshape(-1, 1))
```

Рисунок 3.11 — Прогноз на майбутні роки

Фрагмент коду зображений на рисунку 3.11 створює список років, для яких потрібен прогноз, вибирається остання відома послідовність, та створюється прогноз. Прогнозування відбувається таким чином, що остання послідовність перетворюється у формат (1, sequence_lenght, 1), потім модель передбачає наступне значення, останнє значення видаляється і додається передбачене значення, усі передбачення зберігаються у «future_predictions». Після всього алгоритму прогнози повертаються у початковий формат.

```
final_df = pd.concat([result_df, future_df], ignore_index=True)


print("Останні записи фінального DataFrame:")
print(final_df.tail(10))


plt.figure(figsize=(10, 6))
plt.plot(final_df["Year"], final_df["Value"], label="Реальні дані", marker="o")
plt.plot(final_df["Year"], final_df["Predicted"], label="Прогноз LSTM", linestyle="--")
plt.xlabel("Рік")
plt.ylabel("Значення")
plt.title("Прогноз виробництва молочної продукції до 2025 року")
plt.legend()
plt.grid(True)
plt.show()
```

Рисунок 3.12 — Об’єднання даних та візуалізація

Код наведений на рисунку 3.12 об`єднує реальні дані та прогнози у одну таблицю, після чого відбувається побудова графіку. На графіку відображаються реальні дані (суцільна лінія з точками) та прогноз моделі (пунктирна лінія). Це дозволяє візуально оцінити якість прогнозу та тенденцію для майбутніх років.

3.4. Результати аналізу ринку молочної продукції та побудова графіків

У цьому розділі наведено результати аналізу ринку молочної продукції України з використанням методів прогнозування та візуалізації даних. Результати представлені у вигляді графіків, кожен із яких містить реальні дані за попередні роки та прогноз до 2025 року, виконаний за допомогою моделі LSTM. Після кожного графіка наведено короткий опис його змісту та ключові результати аналізу.

Розглянемо рисунок 3.13 де наведено прогноз доданої вартості за витратами виробництва молочної продукції.

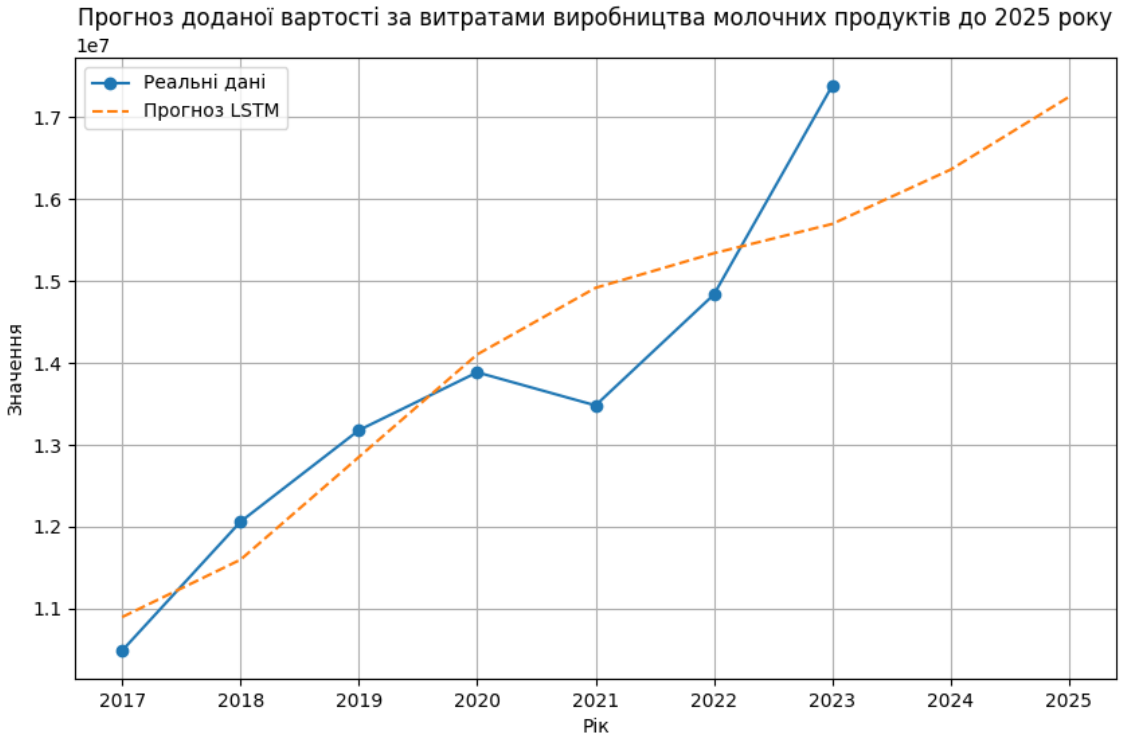

Рисунок 3.13 — Прогноз доданої вартості за витратами виробництва молочної продукції

Аналізуючи графік, представлений на рисунку 3.13 стає зрозуміло, що з 2017 до 2023 року спостерігається поступове зростання доданої вартості з певними періодами уповільнення у 2021 році. Прогноз на 2025 рік демонструє сталу позитивну динаміку, яка досягає пікових значень. Це можна пояснити кількома ключовими факторами. По-перше, відзначається зростання попиту на молочну продукцію як на внутрішньому, так і на зовнішньому ринках. По-друге, технологічні інновації сприяють підвищенню продуктивності й якості виробництва. Також важливим чинником є розширення експортних можливостей, що відкривають нові ринки збуту. Крім того, державна підтримка, спрямована на розвиток аграрного сектору, позитивно впливає на розвиток галузі.

Розглянемо рисунок 3.14 де наведено прогноз доданої вартості за витратами виробництва масла та сиру.

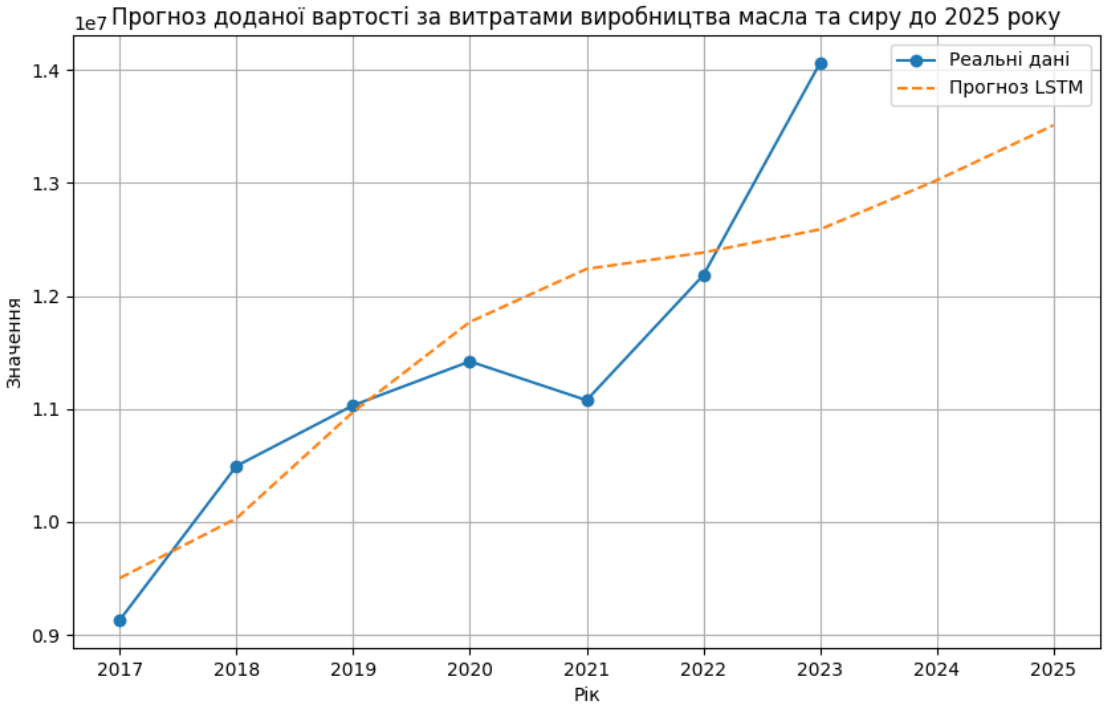

Рисунок 3.14 — Прогноз доданої вартості за витратами виробництва масла та сиру

Розглянувши графік, зображений на рисунку 3.14, ми спостерігаємо, що у період з 2017 до 2023 року спостерігається зростання доданої вартості зі значними коливаннями, проте загальна тенденція до підвищення зберігається. Прогноз на 2025 рік свідчить про сталий розвиток, хоча темпи зростання дещо уповільнюються. Це може бути зумовлено такими факторами, як розширення технологій переробки молока, зокрема для виробництва високоякісного масла та сиру, активна експортна політика у відповідному сегменті ринку, а також підвищення попиту на сирні продукти преміум-класу.

Розглянемо рисунок 3.15 де наведено прогноз доданої вартості за витратами виробництва морозива.

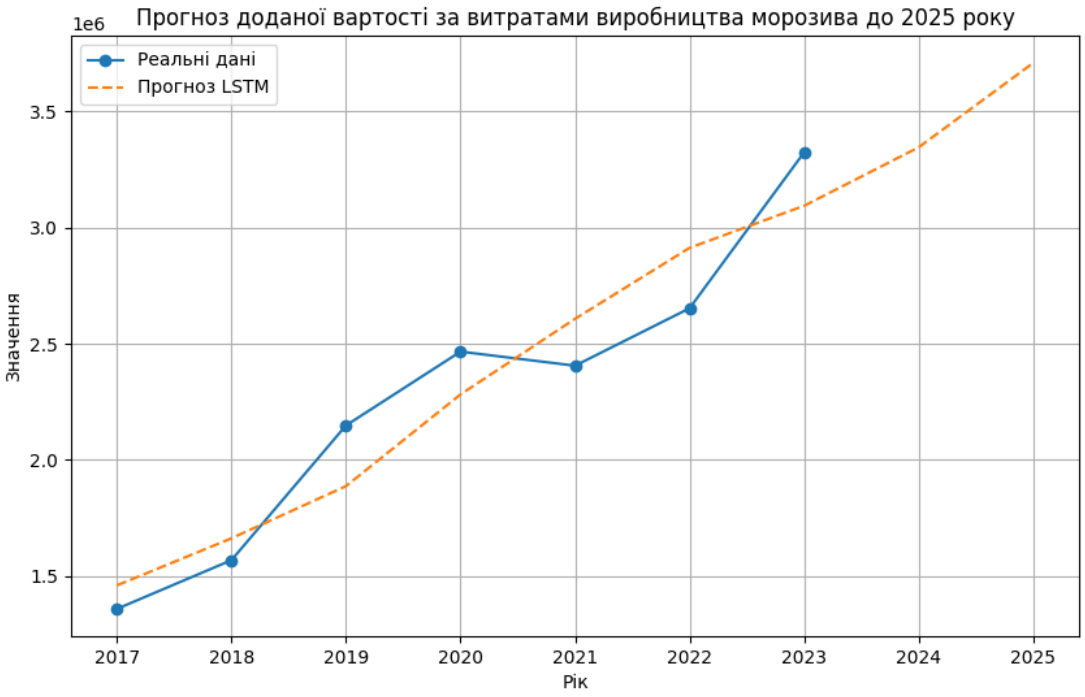

Рисунок 3.15 — Прогноз доданої вартості за витратами виробництва морозива

Динаміка зростання доданої вартості у виробництві морозива, яку ми спостерігаємо на рисунку 3.15, демонструє стійке підвищення, починаючи з 2017 року. Незважаючи на уповільнення у 2021 році, прогноз на 2025 рік свідчить про повернення до стабільного зростання. Така тенденція може пояснюватися кількома чинниками. Зокрема, зростаючий попит на морозиво у літні періоди стимулює розвиток ринку. Крім того, розширення асортименту продукції, що включає використання натуральних інгредієнтів, відповідає сучасним споживчим запитам. Інновації у пакуванні, які забезпечують довготривале зберігання продукту, також сприяють підвищенню конкурентоспроможності виробників морозива.

Розглянемо рисунок 3.16 де наведено прогноз обсягу реалізованої молочної продукції.

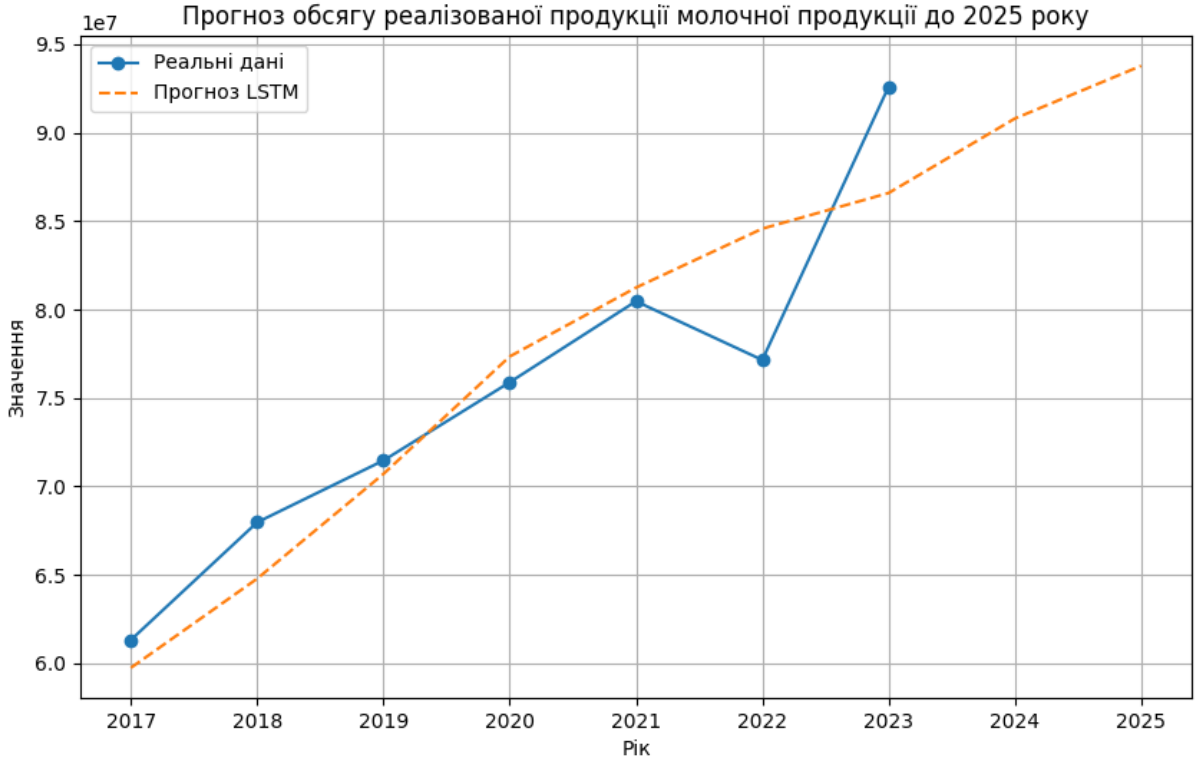

Рисунок 3.16 — Прогноз обсягу реалізованої молочної продукції

Аналізуючи графік, зображений на рисунку 3.16, ми спостерігаємо, що динаміка зростання доданої вартості у виробництві молочної продукції демонструє стабільне підвищення з 2017 року, хоча у 2021 році було зафіксовано незначне зниження. Проте вже у 2022–2023 роках спостерігалося різке зростання, а прогнози до 2025 року свідчать про подальше поступове збільшення. Це зумовлено кількома факторами. Постійний попит на молочну продукцію пояснюється її високою поживною цінністю, що забезпечує стабільний ринок збуту. Виробники активно розширюють асортимент, враховуючи сучасні потреби споживачів, що сприяє підвищенню конкурентоспроможності. Інновації у виробничих процесах дозволяють знижувати витрати і водночас покращувати якість продукції, що позитивно впливає на додану вартість. Окрім того, вагомим фактором залишається державна підтримка молочної галузі, яка стимулює розвиток сектору та забезпечує сприятливі умови для зростання.

Розглянемо рисунок 3.17 де наведено прогноз обсягу реалізованої продукції масла та сиру.

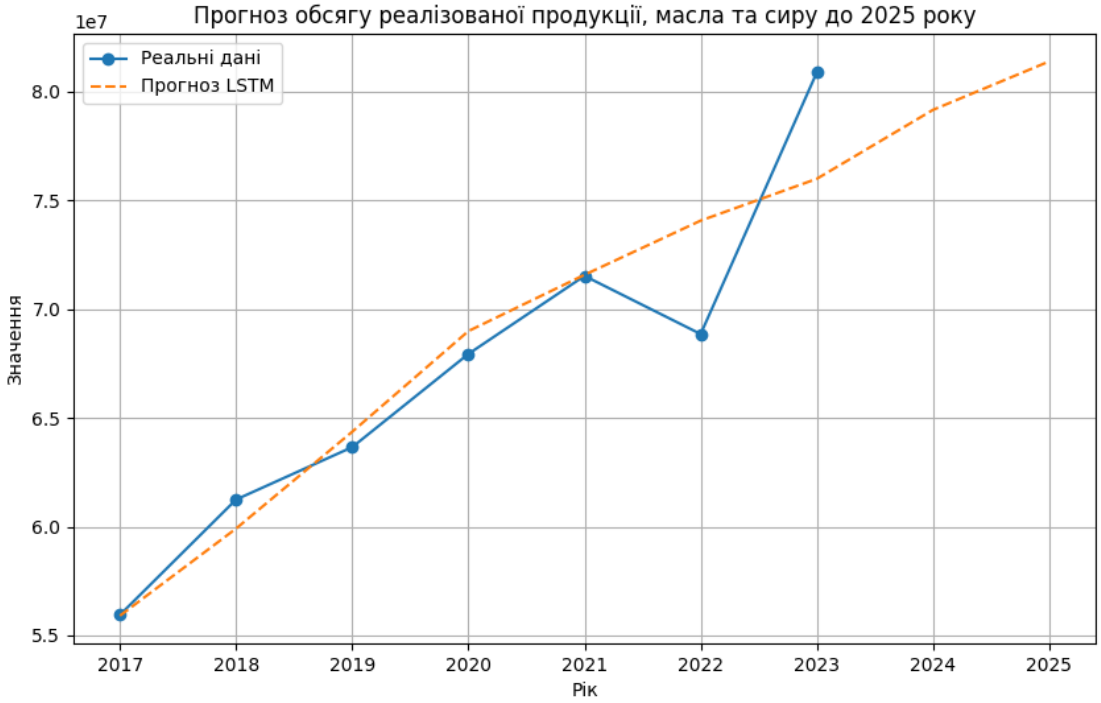

Рисунок 3.17 — Прогноз обсягу реалізованої продукції, масла та сиру

Розглянувши графік, зображений на рисунку 3.17, стає зрозуміло, що динаміка зростання доданої вартості у виробництві масла та сиру демонструє стабільне підвищення з 2017 року, незважаючи на певні коливання у 2021 році. Загальний тренд залишається позитивним, а прогноз до 2025 року свідчить про сталий розвиток галузі. Це обумовлено кількома ключовими чинниками. Підвищення попиту на натуральне вершкове масло та сири преміум-класу відображає зростаючу зацікавленість споживачів у якісній молочній продукції. Впровадження нових технологій у виробництво сприяє зниженню витрат, збільшенню маржі та підвищенню конкурентоспроможності продукції. Крім того, розширення експорту до країн із високим рівнем споживання молочних продуктів відкриває нові ринки збуту, що забезпечує додатковий імпульс для розвитку виробництва масла та сиру.

Розглянемо рисунок 3.18 де наведено прогноз обсягу реалізованої продукції, морозива.

Рисунок 3.18 — Прогноз обсягу реалізованої продукції, морозива

Динаміка зростання доданої вартості у виробництві морозива, яку ми спостерігаємо на рисунку 3.18, демонструє стабільний приріст із 2017 року, хоча у 2021 році спостерігалося певне уповільнення. Прогнози до 2025 року свідчать про впевнене зростання, що відкриває позитивні перспективи для галузі. Це зумовлено кількома ключовими факторами. Попит на морозиво, особливо в літній сезон, продовжує зростати, забезпечуючи стабільний ринок збуту. Виробники активно розширюють асортимент, впроваджуючи інноваційні смаки з використанням натуральних інгредієнтів, що відповідає сучасним споживчим трендам. Удосконалення методів пакування дозволяє подовжувати термін придатності продукції, зберігаючи її якість, що підвищує привабливість для споживачів. Крім того, маркетингові акції та адаптація до змін у споживчих вподобаннях, наприклад, створення низькокалорійного морозива, сприяють розширенню аудиторії та зміцненню позицій виробників на ринку.

Розглянемо рисунок 3.19 де наведено прогноз обсягу виробленої молочної продукції.

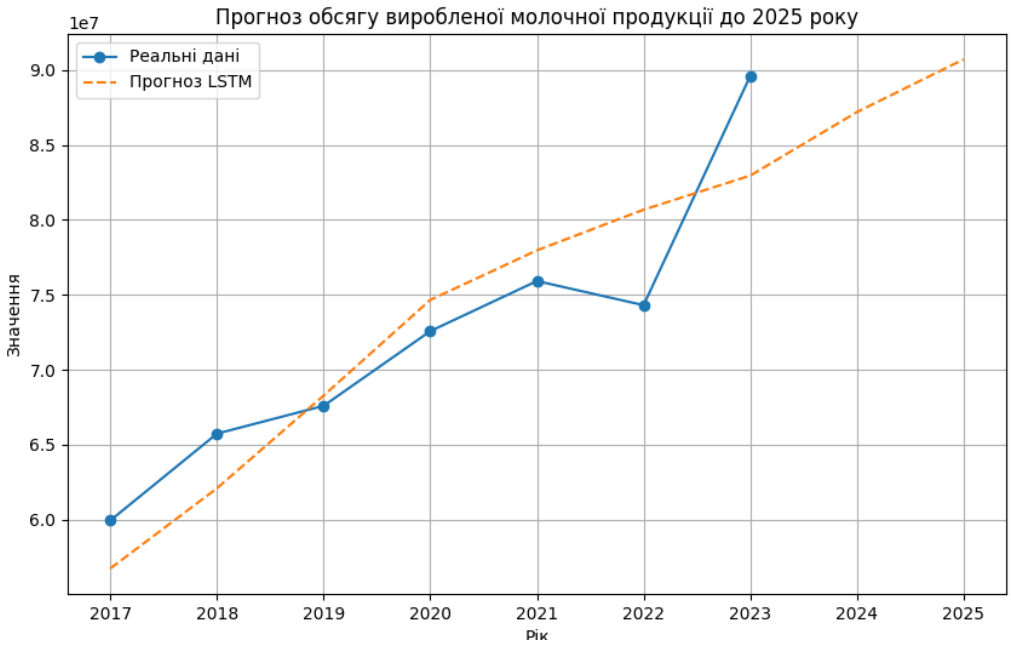

Рисунок 3.19 — Прогноз обсягу виробленої молочної продукції

Аналізуючи графік, зображений на рисунку 3.19, ми спостерігаємо, що динаміка зростання у виробництві молочних продуктів демонструє стабільне підвищення обсягів із 2017 року, зокрема помітне зростання у 2022 році. Це може бути зумовлено як збільшенням попиту, так і впровадженням новітніх технологій у виробництві. Прогноз до 2025 року свідчить про подальше впевнене зростання, що підтверджує стабільний розвиток галузі. До основних причин цього зростання належить розширення ринків збуту, оскільки попит на якісні молочні продукти як у межах країни, так і на експорт, продовжує зростати. Впровадження сучасних технологій переробки сприяє підвищенню ефективності виробничих процесів і зниженню витрат. Значну роль відіграють державні програми підтримки сільськогосподарського сектору, що стимулюють виробників до модернізації та розширення виробництва. Окрім того, зростає популярність органічних молочних продуктів, які дедалі більше відповідають уподобанням сучасних споживачів.

Розглянемо рисунок 3.20 де наведено прогноз обсягу виробленої продукції, масла та сиру.

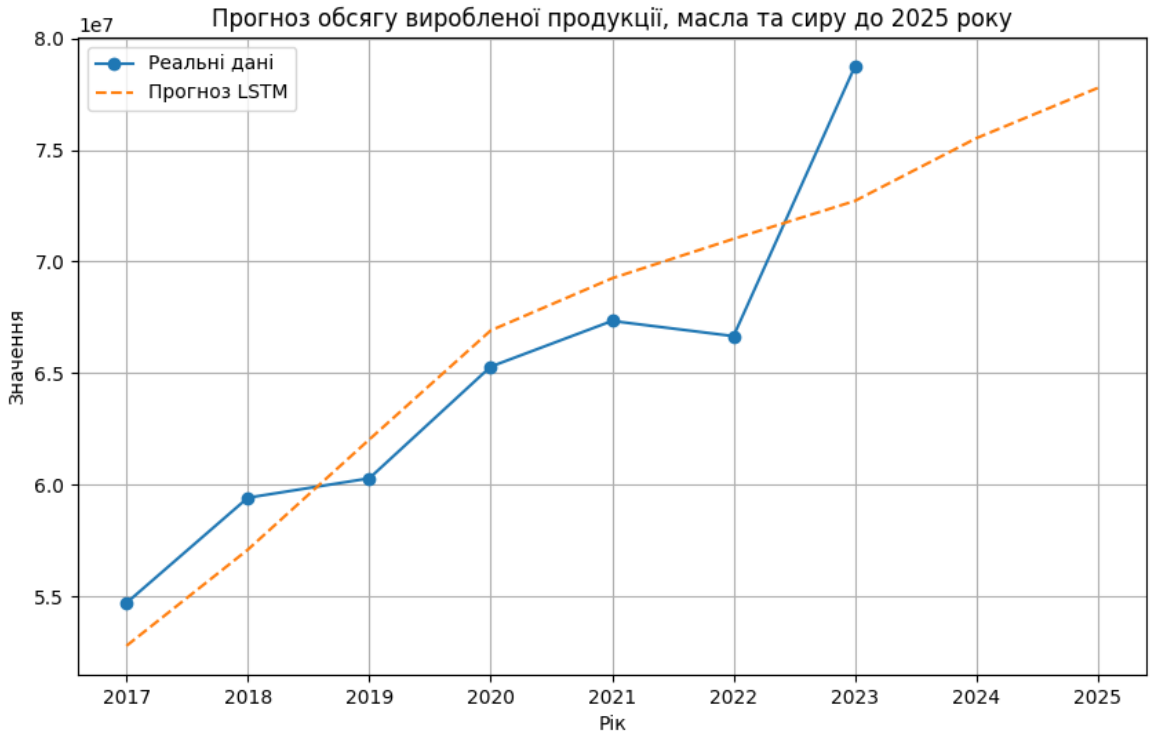

Рисунок 3.20 — Прогноз обсягу виробленої продукції, масла та сиру

Розглянувши графік, зображений на рисунку 3.20, стає зрозуміло, що виробництво масла та сиру демонструє стабільну позитивну динаміку з незначним уповільненням у 2021 році та різким зростанням у 2023 році. Прогноз до 2025 року підтверджує подальше збереження цієї тенденції, що свідчить про перспективний розвиток галузі. Зростання обсягів виробництва зумовлене кількома ключовими факторами. Попит на високоякісну молочну продукцію залишається стабільним, що стимулює виробників до збільшення пропозиції. Значну роль відіграє розвиток експорту, адже масло та сири мають високий попит у країнах із традиційно високим рівнем споживання молочних продуктів. Окрім того, зростає інтерес до традиційних сирів і вершкового масла як натуральних і корисних продуктів, що посилює конкурентоспроможність виробників на внутрішньому та зовнішньому ринках.

Розглянемо рисунок 3.21 де наведено прогноз обсягу виробленої продукції, морозива.

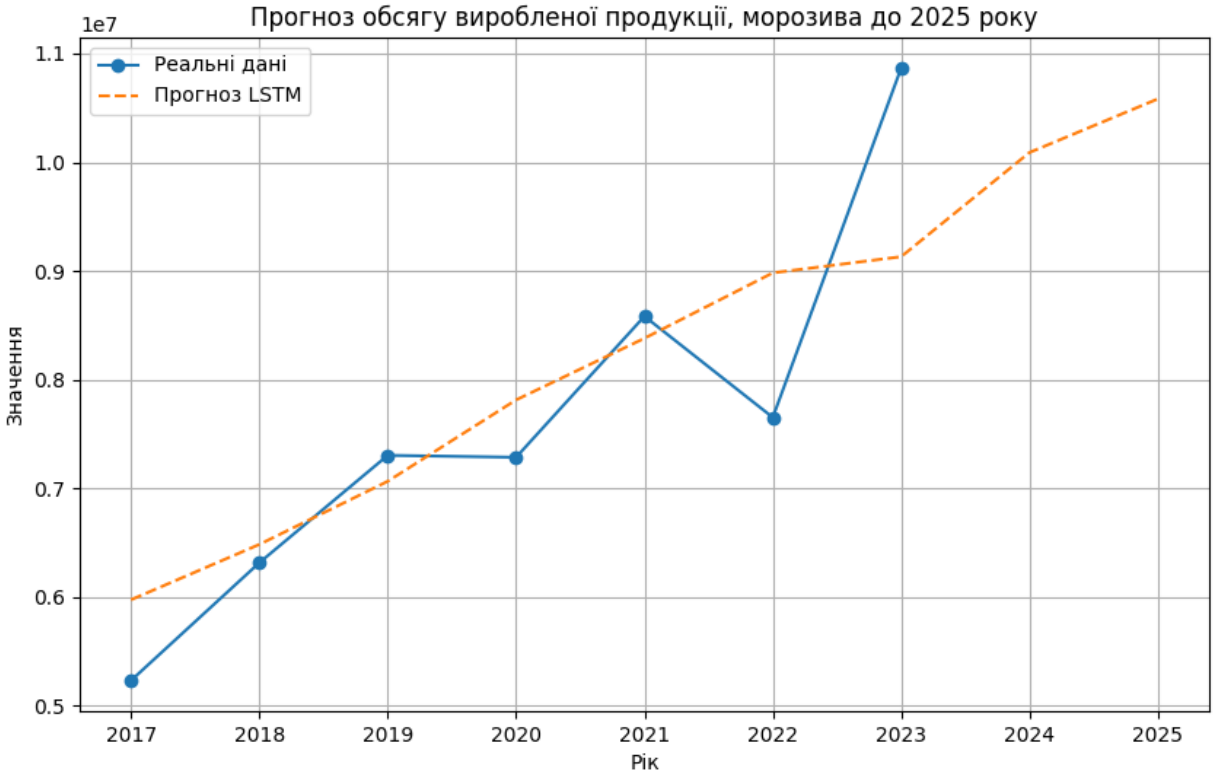

Рисунок 3.21 — Прогноз обсягу виробленої продукції, морозива

Як ми можемо спостерігати на рисунку 3.21, виробництво морозива демонструє стабільну позитивну тенденцію з 2017 року, незважаючи на тимчасове уповільнення в 2021 році. У 2023 році обсяги досягли нового піку, а прогноз до 2025 року вказує на подальше впевнене зростання. Основними чинниками цього є постійне зростання попиту, зумовлене розширенням асортименту, що враховує різноманітні уподобання споживачів. Важливу роль відіграють ефективні маркетингові кампанії та вдосконалення методів реклами, які сприяють зміцненню позицій продукту на ринку. Окрім того, популяризація морозива, адаптованого до сучасних споживчих трендів, таких як низькокалорійні чи веганські варіанти, розширює аудиторію та сприяє збільшенню обсягів виробництва.

3.5. Висновки до третього розділу

Загалом проаналізувавши графіки та результати прогнозування ми можемо спостерігати, що стабільне зростання доданої вартості у виробництві молочних продуктів, масла, сиру та морозива відображає загальну позитивну динаміку розвитку галузі. Це зростання зумовлене підвищенням попиту на якісну продукцію, впровадженням сучасних технологій, а також розширенням асортименту, який відповідає актуальним споживчим трендам. Важливу роль у стимулюванні розвитку відіграють державні програми підтримки, що сприяють залученню додаткових ресурсів і модернізації виробництва.

Прогноз, побудований на основі моделі LSTM, демонструє впевнене зростання до 2025 року, що підтверджує перспективність цього сектору. Модель успішно відображає сезонні коливання, структуру споживання та ключові фактори впливу, дозволяючи підприємствам галузі приймати обґрунтовані управлінські рішення. Вона показує високу точність і кореляцію з реальними даними, що підтверджує її практичну цінність для аналізу та прогнозування ринку молочної продукції.

Розроблений інструмент, інтегрований із сучасними методами аналізу даних, допоможе підприємствам не лише підвищити продуктивність, але й зменшити ризики, пов’язані з коливаннями споживчого попиту. Для подальшого розвитку галузі рекомендується орієнтуватися на інноваційні продукти, такі як органічна та функціональна продукція, впроваджувати новітні технології виробництва й пакування, розвивати експортний потенціал і вдосконалювати маркетингові стратегії. Це сприятиме зміцненню позицій української молочної продукції як на внутрішньому, так і на міжнародному ринках.

ВИСНОВКИ

Робота спрямована на підвищення ефективності аналізу та прогнозування тенденцій ринку молочної продукції України за допомогою сучасних методів штучного інтелекту. Основна мета полягала у створенні інструменту, що дозволяє підприємствам галузі краще адаптуватися до динамічних змін ринку, забезпечуючи точність управлінських рішень.

Проведений аналіз ринку виявив основні тенденції виробництва та споживання, що створюють основу для прогнозування майбутніх змін. Огляд сучасних методів машинного навчання зосередив увагу на алгоритмі LSTM, який довів свою ефективність у задачах роботи з часовими рядами.

Розроблене програмне забезпечення для аналізу та прогнозування допомагає ідентифікувати важливі ринкові закономірності, підвищує якість управлінських процесів та мінімізує ризики, пов’язані з коливаннями попиту. Практична перевірка на основі реальних даних підтвердила високу точність прогнозів, демонструючи перспективність використання LSTM для задач у галузі.

Результати роботи засвідчують потенціал впровадження методів штучного інтелекту в управління підприємствами молочної галузі. Подальший розвиток може включати інтеграцію нових джерел даних і застосування сучасних архітектур нейронних мереж для підвищення точності аналізу та розширення можливостей інструменту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Habr (2024) 'Градієнтний бустінг', Habr. [Електронний ресурс] Режим доступу: https://habr.com/ua/companies/sberbank/articles/672022/#id4 (дата звернення: 17.11.2024).
2. Державна служба статистики України (2024), [Електронний ресурс] Режим доступу: https://www.ukrstat.gov.ua/ (дата звернення: 28.11.2024).
3. Ekmair (2024) 'Градієнтний бустінг (Gradient boosting)', Ekmair. [Електронний ресурс] Режим доступу: https://ekmair.ukma.edu.ua/server/api/core/bitstreams/c315b222-4b7e-44b6-a306-06ebb90b7d5b/content#:~:text=Градієнтний%20бустінг%20(Gradient%20boosting)%20-,на%20відповідних%20залишках%20від%20минулих (дата звернення: 30.11.2024).
4. Habr (2024) 'Machine Learning в ODS', Habr. [Електронний ресурс] Режим доступу: https://habr.com/ua/companies/ods/articles/327250/ (дата звернення: 30.11.2024).
5. Foxminded (2024) 'Реляційні бази даних', Foxminded. [Електронний ресурс] Режим доступу: https://foxminded.ua/reliatsiini-bazy-danykh/ (дата звернення: 28.11.2024).
6. Habr (2024) 'Основи Data Science', Habr. [Електронний ресурс] Режим доступу: https://habr.com/ua/articles/658153/ (дата звернення: 27.11.2024).
7. Neurohive (2024) 'Градієнтний бустінг', Neurohive. [Електронний ресурс] Режим доступу: https://neurohive.io/ru/osnovy-data-science/gradientyj-busting/ (дата звернення: 27.11.2024).
8. Habr (2024) 'Стаття від Wunderfund', Habr. [Електронний ресурс] Режим доступу: https://habr.com/ua/companies/wunderfund/articles/331310/ (дата звернення: 28.11.2024).
9. Bishop, C.M. (2006) *Pattern Recognition and Machine Learning*. New York: Springer.

10. Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep Learning*. Cambridge, MA: MIT Press.
11. Murphy, K.P. (2012) *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: MIT Press.
12. Provost, F. and Fawcett, T. (2013) *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. Sebastopol, CA: O'Reilly Media.
13. Shalev-Shwartz, S. and Ben-David, S. (2014) *Understanding Machine Learning: From Theory to Algorithms*. New York: Cambridge University Press.
14. Khan Academy (2024) 'Introduction to Machine Learning', Khan Academy. Режим доступу: https://www.khanacademy.org/computing/computer-science/algorithms.
15. Coursera (2024) 'Machine Learning', Coursera. [Електронний ресурс] Режим доступу: https://www.coursera.org/learn/machine-learning (дата звернення: 30.11.2024).
16. GeeksforGeeks (2024) 'Gradient Boosting Algorithm', GeeksforGeeks. [Електронний ресурс] Режим доступу: https://www.geeksforgeeks.org/gradient-boosting-algorithm/ (дата звернення: 28.11.2024).
17. Domingos, P. (2015) *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books.
18. Flach, P. (2012) *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge: Cambridge University Press.
19. Mitchell, T.M. (1997) *Machine Learning*. New York: McGraw-Hill.
20. Chollet, F. and Allaire, J.J. (2018) *Deep Learning with R*. Shelter Island, NY: Manning Publications.
21. Aggarwal, C.C. (2015) *Data Mining: The Textbook*. New York: Springer.

22. Mullainathan, S. and Spiess, J. (2017) 'Machine Learning: An Applied Econometric Approach', *Journal of Economic Perspectives*, 31(2), pp. 87-106.

23. McKinney, W. (2017) *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. 2nd edn. Sebastopol, CA: O'Reilly Media.

24. Fisher, M. (2023) 'The Role of Agriculture in Global Economics', Journal of Agricultural Studies, 15(4), pp. 245-260.

25. Robinson, T. and Clark, R. (2018) *Smart Farming and Data Analytics*. New York: Academic Press.

26. Patel, V. (2017) 'Globalization and Its Effects on Agricultural Production', *Global Economic Review*, 23(2), pp. 74-89.

ДОДАТКИ

ДОДАТОК А. Ілюстративний матеріал

	Регіон	Додана вартість за витратами виробництва суб'єктів господарювання/ *Value added of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
виробництво молочних продуктів	Україна	5988741.7	5666211.7	7599253.1	10488803.9	12062249.2	13179487.8	13886398.5	13482159.1	14835865.2	17384882.5
	Вінницька область	675849.1	554061.7	1074815.9	1100397.4	1413093.8	1788138.6	1819450.2	1093179.9	2042182.5	1825340.9
	Волинська область	10209.3	6431.4	2180.2	к/с	54802.4	2311.2	к/с	к/с	34290.9	104058.9
	Дніпропетровська область	314795.7	238051.9	309661.8	к/с	550925.8	1681532.4	к/с	916868.6	к/с	к/с
	Донецька область	433273.9	80696.1	68199.1	к/с	к/с	72450.1	77141.4	82842.5	7787.3	к/с
	Житомирська область	516783.7	446829.4	425552.1	817426.2	883449.2	875225.6	799411.7	961984.7	к/с	1549008.3
	Закарпатська область	к/с	698.0	1236.8	1782.5	к/с	к/с	к/с	2268.5	к/с	2560.2
	Запорізька область	98628.2	34399.5	57849.3	69660.6	57826.6	81112.1	75343.7	68673.1	к/с	51260.5
	Івано - Франківська область	51721.5	16165.5	34814.3	135216.1	к/с	к/с	103026.5	141892.8	128219.4	245531.2
	Київська область	1417082.9	1723607.4	1303364.2	1104031.0	1345026.2	1388510.9	к/с	1577394.7	1328757.0	2282034.7
	Кіровоградська область	к/с	26236.0	к/с	к/с	к/с	к/с	250759.7	237857.5	к/с	100092.4
	Луганська область	20213.2	к/с	к/с	26157.4	33715.1	к/с	к/с	к/с	к/с	к/с
	Львівська область	178995.6	123939.9	270518.4	363649.2	455252.1	к/с	к/с	574725.4	к/с	к/с
	Миколаївська область	к/с	186672.6	285365.6	772504.1	398867.8	412412.9	734543.5	908439.8	к/с	к/с
	Одеська область	43531.8	34305.7	68862.5	76439.5	84468.7	111377.4	179482.6	147295.4	к/с	251912.8
	Полтавська область	484791.4	559044.7	668280.0	988789.5	645976.8	387774.7	443606.9	681646.9	886402.8	793460.5
	Рівненська область	216325.3	245955.2	341188.0	439327.3	к/с	к/с	к/с	743360.5	759935.5	938398.1
	Сумська область	50503.5	125816.5	108128.0	210707.6	149878.0	186151.7	188010.6	67105.9	к/с	к/с
	Тернопільська область	249754.6	176736.6	275460.8	402577.7	445348.9	505055.1	606310.8	537895.1	к/с	к/с
	Харківська область	159838.5	225812.8	255339.7	596506.0	473287.4	618825.5	879297.9	1384698.7	1297588.5	618970.3
	Херсонська область	141709.7	к/с	272941.5	к/с	к/с	781705.2	к/с	к/с	6661.0	12338.4
	Хмельницька область	-152641.2	-150234.2	-37988.9	-497502.1	753002.0	792268.2	1071868.5	333969.1	к/с	к/с
	Черкаська область	124002.5	91021.8	430746.0	797633.3	770257.9	421597.8	611385.1	383251.3	840367.6	998346.6
	Чернівецька область	-1376.2	138.0	-2258.6	к/с	778.5	к/с	7950.2	10282.2	к/с	53372.8
	Чернігівська область	270633.2	196615.7	486750.3	569581.8	382569.0	139845.8	407670.5	271482.4	621329.5	272488.0
	м. Київ	492866.1	596473.7	785941.8	1255117.8	1204475.4	1578091.9	1248648.5	1511622.3	2244454.2	2799084.4

Рисунок А.1 — Зберігання даних в таблиці ДВМП

	Регіон	Додана вартість за витратами виробництва суб'єктів господарювання/ *Value added of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
перероблення молока, виробництво масла та сиру	Україна	5200572.1	4709759.1	6755670.5	9131020.3	10494665.2	11031512.3	11420361.9	11076800.3	12183007.3	14060463.5
	Вінницька область	к/с	к/с	к/с	к/с	1413075.3	к/с	к/с	1091227.0	2041806.1	1824726.7
	Волинська область	к/с	6431.4	1453.8	к/с	54767.8	2311.2	к/с	к/с	к/с	к/с
	Дніпропетровська область	к/с	к/с	к/с	359248.1	к/с	к/с	к/с	573491.5	856649.0	1080233.0
	Донецька область	408925.1	67861.3	к/с	к/с	305629.6	к/с	к/с	к/с	7787.3	к/с
	Житомирська область	160686.9	156149.6	214043.7	к/с	458040.8	399499.1	407739.1	601433.7	к/с	к/с
	Закарпатська область	к/с	к/с	819.1	1115.3	к/с	к/с	к/с	к/с	к/с	к/с
	Запорізька область	93517.7	33871.4	57179.6	к/с	к/с	к/с	к/с	67363.1	к/с	51260.5
	Івано - Франківська область	к/с	к/с	к/с	к/с	185745.1	91462.6	100439.9	138915.2	к/с	243979.2
	Київська область	к/с	1723057.2	1302641.6	1103202.9	1344085.1	к/с	к/с	1575932.0	1326835.6	2275234.6
	Кіровоградська область	1425.7	к/с	5960.8	к/с	47112.4	26813.8	к/с	148676.5	к/с	24897.7
	Луганська область	к/с	к/с	к/с	26157.4	33715.1	к/с	к/с	к/с	к/с	к/с
	Львівська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Миколаївська область	205252.7	185968.8	к/с	к/с	к/с	411955.0	к/с	к/с	к/с	257016.1
	Одеська область	к/с	к/с	55544.6	к/с	45881.8	74834.7	119972.0	73069.0	60868.1	к/с
	Полтавська область	475720.4	к/с	654913.9	к/с	629623.5	к/с	429701.8	629448.4	к/с	к/с
	Рівненська область	к/с	к/с	332136.1	к/с	к/с	592641.4	662063.9	725340.1	к/с	897843.8
	Сумська область	к/с	к/с	к/с	210707.6	149878.0	186151.7	188010.6	к/с	к/с	к/с
	Тернопільська область	к/с	к/с	к/с	к/с	к/с	424524.5	495226.9	к/с	к/с	к/с
	Харківська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	713302.9	696850.5	216983.8
	Херсонська область	141451.3	к/с	272643.1	к/с	к/с	к/с	879265.7	769900.5	6661.0	к/с
	Хмельницька область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	333846.8	Ы	к/с
	Черкаська область	к/с	к/с	к/с	к/с	к/с	к/с	610278.6	381868.1	к/с	к/с
	Чернівецька область	-1376.2	138.0	-2258.6	к/с	к/с	-146.9	к/с	к/с	33683.1	46113.8
	Чернігівська область	к/с	к/с	486750.3	569581.8	к/с	к/с	к/с	271482.4	621329.5	к/с
	м. Київ	к/с	к/с	631920.3	к/с	к/с	к/с	к/с	1153186.5	1772426.0	2168861.7

Рисунок А.2 — Зберігання даних в таблиці ДВПМП

	Регіон	Додана вартість за витратами виробництва суб'єктів господарювання/ *Value added of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
виробництво морозива	Україна	788169.6	956452.6	843582.6	#####	1567584.0	2147975.5	2466036.6	2405358.8	2652857.9	3324419.0
	Вінницька область	к/с	к/с	к/с	к/с	18.5	к/с	к/с	1952.9	376.4	614.2
	Волинська область	к/с	–	к/с	к/с	к/с	–	–	–	к/с	к/с
	Дніпропетровська область	к/с	к/с	к/с	к/с	к/с	к/с	518578.5	343377.1	353193.8	310155.5
	Донецька область	к/с	12834.8	к/с	к/с	33.8	к/с	к/с	к/с	–	к/с
	Житомирська область	356096.8	290679.8	211508.4	к/с	425408.4	475726.5	391672.6	360551.0	567898.6	к/с
	Закарпатська область	к/с	к/с	417.7	667.2	835.4	893.3	445.4	к/с	311.8	к/с
	Запорізька область	5110.5	528.1	669.7	к/с	к/с	к/с	к/с	1310.0	к/с	0.0
	Івано - Франківська область	к/с	к/с	к/с	к/с	к/с	к/с	2586.6	2977.6	к/с	1552.0
	Київська область	к/с	550.2	722.6	828.1	941.1	к/с	372.0	1462.7	1921.4	6800.1
	Кіровоградська область	к/с	к/с	к/с	к/с	89207.4	к/с	к/с	к/с	к/с	к/с
	Луганська область	к/с	–	к/с	–	–	–	–	–	–	–
	Львівська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Миколаївська область	к/с	703.8	к/с	к/с	к/с	457.9	к/с	к/с	к/с	к/с
	Одеська область	к/с	к/с	13317.9	к/с	38586.9	36542.7	59510.6	74226.4	к/с	к/с
	Полтавська область	к/с	к/с	13366.1	к/с	16353.3	к/с	13905.1	52198.5	к/с	к/с
	Рівненська область	к/с	к/с	9051.9	к/с	16168.5	к/с	к/с	к/с	к/с	к/с
	Сумська область	к/с	к/с	к/с	–	–	–	–	к/с	к/с	к/с
	Тернопільська область	к/с	к/с	к/с	к/с	к/с	80530.6	111083.9	к/с	к/с	к/с
	Харківська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	671395.8	600738.0	401986.5
	Херсонська область	258.4	278.2	298.4	381.3	600.1	к/с	к/с	к/с	–	к/с
	Хмельницька область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Черкаська область	к/с	к/с	к/с	к/с	к/с	к/с	1106.5	1383.2	к/с	к/с
	Чернівецька область	–	–	–	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Чернігівська область	к/с	к/с	–	–	к/с	к/с	к/с	–	–	к/с
	м. Київ	к/с	к/с	154021.5	к/с	к/с	к/с	к/с	358435.8	472028.2	630222.7

Рисунок А.3 — Зберігання даних в таблиці ДВМ

	Регіон	Обсяг реалізованої продукції (товарів, послуг) суб'єктів господарювання/ *Turnover of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
виробництво молочних продуктів	Україна	35360251.1	41024860.3	45979487.8	61320528.1	67969754.0	71464972.5	75888393.0	80468245.7	77149522.6	92590958.7
	Вінницька область	5107316.6	7081734.3	7908582.3	10850708.7	11659049.3	12987007.1	12933409.0	9159870.1	10451759.4	11262454.6
	Волинська область	120216.4	35067.8	9854.0	к/с	631719.6	9142.7	к/с	к/с	137941.6	234759.7
	Дніпропетровська область	1740861.7	1940496.1	2313029.9	к/с	3864070.4	4446343.6	к/с	5084313.7	к/с	к/с
	Донецька область	1136537.3	542149.7	470046.9	к/с	к/с	358268.4	313194.9	323481.3	53705.4	к/с
	Житомирська область	1944839.8	2501799.0	3175151.4	4304121.1	5185276.4	5049336.6	5098411.4	5887136.0	к/с	7815858.9
	Закарпатська область	к/с	3048.0	4512.6	5260.1	к/с	к/с	к/с	8551.2	к/с	10608.2
	Запорізька область	833251.8	305815.5	403440.5	480948.6	354938.9	497185.7	468776.1	457844.3	к/с	356317.4
	Івано - Франківська область	268634.1	96809.3	232845.7	460435.8	к/с	к/с	521854.8	671650.9	712615.1	1227484.9
	Київська область	6544585.8	5972840.2	4616697.3	4199979.1	4857780.9	4968077.3	к/с	4499410.2	4838017.5	7342441.3
	Кіровоградська область	к/с	299078.2	к/с	к/с	к/с	к/с	1013408.5	1012193.4	к/с	1199283.5
	Луганська область	139891.7	к/с	к/с	188464.3	212261.3	к/с	к/с	к/с	к/с	к/с
	Львівська область	1081911.8	1256392.5	1306681.0	1835657.1	2123642.9	к/с	к/с	3380996.6	к/с	к/с
	Миколаївська область	к/с	1951559.9	1926379.0	2611855.2	3618874.1	3988952.2	3862490.1	4391350.7	к/с	к/с
	Одеська область	301952.7	301453.5	389395.8	531846.4	576672.5	646594.4	636069.6	742734.0	к/с	982127.5
	Полтавська область	4259107.8	5062067.3	5650599.0	6033466.6	5937055.3	6091916.3	5837388.2	5910304.9	6827468.2	7351377.8
	Рівненська область	1025367.6	1401996.7	1852892.8	2664908.0	к/с	к/с	к/с	3689120.7	4445866.5	4911282.7
	Сумська область	709132.1	958831.0	1231333.8	2051295.5	2228778.6	1647365.2	1424728.4	1397299.8	к/с	к/с
	Тернопільська область	1133395.4	1414764.4	1781711.6	2390743.7	2657996.9	2677466.3	2816276.7	3296645.3	к/с	к/с
	Харківська область	1072877.6	1515135.3	1862650.3	2693579.7	2818435.1	3135867.2	3756741.9	5041275.0	2790623.6	4110671.3
	Херсонська область	983362.5	к/с	3307461.0	к/с	к/с	4180836.4	к/с	к/с	55472.1	56628.0
	Хмельницька область	1132637.7	1272494.6	1623108.0	2410116.9	2734354.4	3204189.3	3548191.1	1959945.7	к/с	к/с
	Черкаська область	1831400.7	2051088.3	2370407.3	3454130.8	3832758.6	3837200.9	4042562.4	4910326.5	5517249.8	6619047.6
	Чернівецька область	5833.2	9869.5	19998.6	к/с	2110.6	к/с	26514.6	41288.6	к/с	70987.3
	Чернігівська область	1091281.4	1424706.6	1552698.1	1920557.1	1996385.3	1998592.8	1570673.9	2034491.1	2157038.3	1925180.6
	м. Київ	864487.6	1339962.4	1376406.5	3813920.3	3727272.8	4031460.6	7988952.8	12004939.8	9569174.1	10951268.5

Рисунок А.4 — Зберігання даних в таблиці РПМП

	Регіон	Обсяг реалізованої продукції (товарів, послуг) суб'єктів господарювання/ *Turnover of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
переоброблення молока, виробництво масла та сиру	Україна	32504798.7	37306088.0	41471384.7	55949616.1	61249617.3	63659746.3	67943347.8	71531768.9	68858385.0	80886797.4
	Вінницька область	к/с	к/с	к/с	к/с	11659025.2	к/с	к/с	9150446.7	10451759.4	11262449.5
	Волинська область	к/с	35067.8	7976.3	к/с	631674.6	9142.7	к/с	к/с	к/с	к/с
	Дніпропетровська область	к/с	к/с	к/с	2166637.5	к/с	к/с	к/с	3405038.2	4065020.3	5535005.8
	Донецька область	1088861.7	514383.3	к/с	к/с	430025.3	к/с	к/с	к/с	53705.4	к/с
	Житомирська область	757155.9	970959.8	1399557.9	к/с	2975733.0	2831637.9	2875271.2	3422439.8	к/с	к/с
	Закарпатська область	к/с	к/с	3282.4	3616.7	к/с	к/с	к/с	к/с	к/с	к/с
	Запорізька область	804323.2	303690.6	401236.3	к/с	к/с	к/с	к/с	454046.4	к/с	356317.4
	Івано - Франківська область	к/с	к/с	к/с	к/с	523518.5	544426.4	514809.8	663905.2	к/с	1222420.7
	Київська область	к/с	5970434.8	4614315.3	4197302.9	4854956.0	к/с	к/с	4495936.7	4834854.1	7324644.7
	Кіровоградська область	6850.4	к/с	21244.4	к/с	172119.6	371871.4	к/с	471276.0	к/с	322681.2
	Луганська область	к/с	к/с	к/с	188464.3	212261.3	к/с	к/с	к/с	к/с	к/с
	Львівська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Миколаївська область	1785497.0	1948372.6	к/с	к/с	к/с	3986833.0	к/с	к/с	к/с	3085839.4
	Одеська область	к/с	к/с	318807.2	к/с	437301.0	467815.0	397227.9	441132.9	433560.6	к/с
	Полтавська область	4218980.4	к/с	5568357.8	к/с	5822684.4	к/с	5702967.0	5753812.4	к/с	к/с
	Рівненська область	к/с	к/с	1821273.3	к/с	к/с	3023637.8	3263869.4	3638232.1	к/с	4792725.4
	Сумська область	к/с	к/с	к/с	2051295.5	2228778.6	1647365.2	1424728.4	к/с	к/с	к/с
	Тернопільська область	к/с	к/с	к/с	к/с	к/с	2435417.0	2501727.4	к/с	к/с	к/с
	Харківська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	3666923.4	1849012.6	2372456.1
	Херсонська область	982266.6	к/с	3306432.8	к/с	к/с	к/с	4190832.8	4317387.0	55472.1	к/с
	Хмельницька область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	1959461.0	к/с	к/с
	Черкаська область	к/с	к/с	к/с	к/с	к/с	к/с	4039466.5	4905874.8	к/с	к/с
	Чернівецька область	5833.2	9869.5	19998.6	к/с	к/с	28076.0	к/с	к/с	55243.6	55746.5
	Чернігівська область	к/с	к/с	1552698.1	1920557.1	к/с	к/с	к/с	2034491.1	2157038.3	к/с
	м. Київ	к/с	к/с	1080286.0	к/с	к/с	к/с	к/с	10858057.5	8563438.6	9533639.0

Рисунок А.5 — Зберігання даних в таблиці РППМП

	Регіон	Обсяг реалізованої продукції (товарів, послуг) суб'єктів господарювання/ *Turnover of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
виробництво морозива	Україна	2855452.4	3718772.3	4508103.1	5370912.0	6720136.7	7805226.2	7945045.2	8936476.8	8291137.6	11704161.3
	Вінницька область	к/с	к/с	к/с	к/с	24.1	к/с	к/с	9423.4	0.0	5.1
	Волинська область	к/с	–	к/с	к/с	к/с	–	–	–	к/с	к/с
	Дніпропетровська область	к/с	к/с	к/с	к/с	к/с	к/с	1620573.5	1679275.5	1450669.9	2141154.9
	Донецька область	к/с	27766.4	к/с	к/с	к/с	к/с	к/с	к/с	–	к/с
	Житомирська область	1187683.9	1530839.2	1775593.5	к/с	2209543.4	2217698.7	2223140.2	2464696.2	2277808.5	к/с
	Закарпатська область	к/с	к/с	1230.2	1643.4	2161.7	2953.6	988.4	к/с	1090.2	к/с
	Запорізька область	28928.6	2124.9	2204.2	к/с	к/с	к/с	к/с	3797.9	к/с	0.0
	Івано - Франківська область	к/с	к/с	к/с	к/с	к/с	к/с	7045.0	7745.7	к/с	5064.2
	Київська область	к/с	2405.4	2382.0	2676.2	2824.9	к/с	879.8	3473.5	3163.4	17796.6
	Кіровоградська область	к/с	к/с	к/с	к/с	692977.0	к/с	к/с	к/с	к/с	к/с
	Луганська область	к/с	–	к/с	–	–	–	–	–	–	–
	Львівська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Миколаївська область	к/с	3187.3	к/с	к/с	к/с	2119.2	к/с	к/с	к/с	к/с
	Одеська область	к/с	к/с	70588.6	к/с	139371.5	178779.4	238841.7	301601.1	к/с	к/с
	Полтавська область	к/с	к/с	82241.2	к/с	к/с	к/с	134421.2	156492.5	к/с	к/с
	Рівненська область	к/с	к/с	31619.5	к/с	45145.5	к/с	к/с	к/с	к/с	к/с
	Сумська область	к/с	к/с	к/с	–	–	–	–	к/с	к/с	к/с
	Тернопільська область	к/с	к/с	к/с	к/с	к/с	242049.3	314549.3	к/с	к/с	к/с
	Харківська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	1374351.6	941611.0	1738215.2
	Херсонська область	1095.9	1353.8	1028.2	1141.8	1505.5	к/с	к/с	к/с	–	к/с
	Хмельницька область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Черкаська область	к/с	к/с	к/с	к/с	к/с	к/с	3095.9	4451.7	к/с	к/с
	Чернівецька область	–	–	–	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Чернігівська область	к/с	к/с	–	–	к/с	к/с	к/с	–	–	к/с
	м. Київ	к/с	к/с	296120.5	к/с	к/с	к/с	к/с	1146882.3	1005735.5	1417629.5

Рисунок А.6 — Зберігання даних в таблиці РПМ

	Регіон	Обсяг виробленої продукції (товарів, послуг) суб'єктів господарювання/ *Value of output of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ total, thsd. UAH	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
виробництво молочних продуктів	Україна	33452816.6	38538472.9	43390769.9	59941685.5	65737360.6	67593294.9	72577826.7	75921802.8	74311302.4	89603947.0
	Вінницька область	4977505.7	6684481.6	7652994.5	10744919.5	11621089.3	12975529.3	12505559.9	9181510.1	10488689.7	11234464.9
	Волинська область	91352.3	30496.2	10676.0	к/с	630596.1	8437.7	к/с	к/с	138092.1	236173.8
	Дніпропетровська область	1753753.5	1856773.4	2164112.2	к/с	3855876.4	4451082.9	к/с	4986409.7	к/с	к/с
	Донецька область	1144059.3	377294.7	261534.2	к/с	к/с	250616.7	274430.9	302190.1	45866.1	к/с
	Житомирська область	1789889.4	2235165.4	2958798.5	4173904.3	4801981.0	4833704.7	4683345.8	5356857.1	к/с	7302501.9
	Закарпатська область	к/с	2348.5	3938.4	4968.8	к/с	к/с	к/с	8028.6	к/с	10723.5
	Запорізька область	803275.2	263902.6	345096.3	451459.3	337090.4	492511.9	452787.3	438971.7	к/с	348382.2
	Івано - Франківська область	262312.6	93642.9	231975.3	503835.7	к/с	к/с	535840.3	702008.3	675641.3	1154983.9
	Київська область	5740979.9	6254729.1	4729100.7	4232375.3	4823245.3	4873068.2	к/с	4448394.2	4810886.1	7416571.1
	Кіровоградська область	к/с	266891.0	к/с	к/с	к/с	к/с	1011961.0	1010410.5	к/с	1181119.0
	Луганська область	137656.6	к/с	к/с	180551.8	206665.5	к/с	к/с	к/с	к/с	к/с
	Львівська область	1104217.2	1164146.3	1284227.6	1845851.7	2176697.1	к/с	к/с	3216846.6	к/с	к/с
	Миколаївська область	к/с	1672888.1	1801043.5	2689154.8	3363081.0	3410808.3	3493154.8	3937584.1	к/с	к/с
	Одеська область	295793.9	277639.3	394108.3	515865.1	568233.2	647475.5	671042.2	749734.5	к/с	996693.8
	Полтавська область	3783210.3	4637755.1	5120619.0	5814024.4	5789989.3	5632282.8	5563984.0	5699543.8	6440468.6	7154316.1
	Рівненська область	1011413.7	1405008.4	1894325.3	2727884.1	к/с	к/с	к/с	3625509.1	4397295.3	4908542.5
	Сумська область	651157.6	899585.4	1150386.9	2027451.6	2139007.5	1612670.9	1377404.3	1341535.7	к/с	к/с
	Тернопільська область	1138530.6	1329893.3	1735164.7	2390168.8	2661117.5	2722000.4	2871662.0	3292794.8	к/с	к/с
	Харківська область	1074573.3	1372418.2	1713515.7	2830018.3	2956960.6	3167500.0	3541973.8	5161689.4	2948963.9	3934521.6
	Херсонська область	985222.5	к/с	1965549.0	к/с	к/с	2700697.5	к/с	к/с	47668.8	56999.8
	Хмельницька область	1093479.4	1320623.8	1763502.7	2787981.8	2703895.9	2871363.6	3522174.8	1873795.5	к/с	к/с
	Черкаська область	1548147.0	1830302.4	2305004.4	3200228.1	3697397.9	3489063.5	4034717.1	4122057.7	5054046.2	6050126.4
	Чернівецька область	20370.0	9180.6	13741.4	к/с	2053.5	к/с	32314.6	42521.6	к/с	73097.5
	Чернігівська область	1159502.8	1400008.9	1757776.6	1996363.5	2035972.0	2035737.8	1625685.7	1948130.8	2152167.5	1963746.5
	м. Київ	1029170.0	1696465.9	1551479.9	3645178.6	3551874.7	3905034.1	7983159.6	11589570.4	9145382.0	9916441.6

Рисунок А.7 — Зберігання даних в таблиці ВПМП

	Регіон	Обсяг виробленої продукції (товарів, послуг) суб'єктів господарювання/ *Value of output of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ total, thsd. UAH	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
перероблення молока, виробництво масла та сиру	Україна	30710290.4	34904610.4	39213993.7	54707324.2	59419905.2	60288489.9	65291090.6	67338396.7	66657802.0	78738498.8
	Вінницька область	к/с	к/с	к/с	к/с	11621065.2	к/с	к/с	9172137.2	10488313.3	11233849.5
	Волинська область	к/с	30496.2	8841.2	к/с	630551.1	8437.7	к/с	к/с	к/с	к/с
	Дніпропетровська область	к/с	к/с	к/с	2093075.3	к/с	к/с	к/с	3321420.2	3983288.0	5461284.8
	Донецька область	1094271.0	349707.7	к/с	к/с	494397.6	к/с	к/с	к/с	45866.1	к/с
	Житомирська область	785776.2	960778.4	1468753.6	к/с	3002881.5	2953301.7	2894859.8	3260544.8	к/с	к/с
	Закарпатська область	к/с	к/с	2851.3	3380.2	к/с	к/с	к/с	к/с	к/с	к/с
	Запорізька область	777269.5	262290.1	343288.8	к/с	к/с	к/с	к/с	435539.2	к/с	348382.2
	Івано - Франківська область	к/с	к/с	к/с	к/с	528728.4	511148.1	529173.9	694777.3	к/с	1150147.7
	Київська область	к/с	6252942.3	4726990.9	4229887.3	4820615.9	к/с	к/с	4445001.0	4807517.9	7399397.3
	Кіровоградська область	5194.5	к/с	20816.0	к/с	189299.4	265237.5	к/с	481581.4	к/с	320793.3
	Луганська область	к/с	к/с	к/с	180551.8	206665.5	к/с	к/с	к/с	к/с	к/с
	Львівська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Миколаївська область	1631715.3	1670026.8	к/с	к/с	к/с	3408801.1	к/с	к/с	к/с	2790434.4
	Одеська область	к/с	к/с	321349.4	к/с	426208.3	467814.1	433213.2	445172.2	432910.1	к/с
	Полтавська область	3741412.6	к/с	5041798.2	к/с	5683889.1	к/с	5447613.8	5549755.8	к/с	к/с
	Рівненська область	к/с	к/с	1868721.9	к/с	к/с	3029152.0	3260449.3	3579878.0	к/с	4815552.6
	Сумська область	к/с	к/с	к/с	2027451.6	2139007.5	1612670.9	1377404.3	к/с	к/с	к/с
	Тернопільська область	к/с	к/с	к/с	к/с	к/с	2473794.4	2564158.1	к/с	к/с	к/с
	Харківська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	3655916.1	1920278.9	2482419.6
	Херсонська область	984313.1	к/с	1964719.8	к/с	к/с	к/с	2628211.7	2638344.9	47668.8	к/с
	Хмельницька область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	1873310.8	к/с	к/с
	Черкаська область	к/с	к/с	к/с	к/с	к/с	к/с	4031785.8	4117867.8	к/с	к/с
	Чернівецька область	20370.0	9180.6	13741.4	к/с	к/с	28888.8	к/с	к/с	83442.1	56591.8
	Чернігівська область	к/с	к/с	1757776.6	1996363.5	к/с	к/с	к/с	1948130.8	2152167.5	к/с
	м. Київ	к/с	к/с	1197580.4	к/с	к/с	к/с	к/с	10538556.8	8228021.6	8607356.6

Рисунок А.8 — Зберігання даних в таблиці ВППМП

	Регіон	Обсяг виробленої продукції (товарів, послуг) суб'єктів господарювання/ *Value of output of business entities*									
		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
		усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ total, thsd. UAH	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*	усього, тис.грн/ *total, thsd. UAH*
виробництво морозива	Україна	2742526.2	3633862.5	4176776.2	5234361.3	6317455.4	7304805.0	7286736.1	8583406.1	7653500.4	10865448.2
	Вінницька область	к/с	к/с	к/с	к/с	24.1	к/с	к/с	9372.9	376.4	615.4
	Волинська область	к/с	–	к/с	к/с	к/с	–	–	–	к/с	к/с
	Дніпропетровська область	к/с	к/с	к/с	к/с	к/с	к/с	1621775.6	1664989.5	1265118.9	2132898.8
	Донецька область	к/с	27587.0	к/с	к/с	33.8	к/с	к/с	к/с	–	к/с
	Житомирська область	1004113.2	1274387.0	1490044.9	к/с	1799099.5	1880403.0	1788486.0	2096312.3	1920761.2	к/с
	Закарпатська область	к/с	к/с	1087.1	1588.6	2047.5	2763.2	988.4	к/с	937.5	к/с
	Запорізька область	26005.7	1612.5	1807.5	к/с	к/с	к/с	к/с	3432.5	к/с	0.0
	Івано - Франківська область	к/с	к/с	к/с	к/с	к/с	к/с	6666.4	7231.0	к/с	4836.2
	Київська область	к/с	1786.8	2109.8	2488.0	2629.4	к/с	944.0	3393.2	3368.2	17173.8
	Кіровоградська область	к/с	к/с	к/с	к/с	667082.6	к/с	к/с	к/с	к/с	к/с
	Луганська область	к/с	–	к/с	–	–	–	–	–	–	–
	Львівська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Миколаївська область	к/с	2861.3	к/с	к/с	к/с	2007.2	к/с	к/с	к/с	к/с
	Одеська область	к/с	к/с	72758.9	к/с	142024.9	179661.4	237829.0	304562.3	к/с	к/с
	Полтавська область	к/с	к/с	78820.8	к/с	106100.2	к/с	116370.2	149788.0	к/с	к/с
	Рівненська область	к/с	к/с	25603.4	к/с	40448.0	к/с	к/с	к/с	к/с	к/с
	Сумська область	к/с	к/с	к/с	–	–	–	–	к/с	к/с	к/с
	Тернопільська область	к/с	к/с	к/с	к/с	к/с	248206.0	307503.9	к/с	к/с	к/с
	Харківська область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	1505773.3	1028685.0	1452102.0
	Херсонська область	909.4	1007.0	829.2	1052.3	1465.8	к/с	к/с	к/с	–	к/с
	Хмельницька область	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Черкаська область	к/с	к/с	к/с	к/с	к/с	к/с	2931.3	4189.9	к/с	к/с
	Чернівецька область	–	–	–	к/с	к/с	к/с	к/с	к/с	к/с	к/с
	Чернігівська область	к/с	к/с	–	–	к/с	к/с	к/с	–	–	к/с
	м. Київ	к/с	к/с	353899.5	к/с	к/с	к/с	к/с	1051013.6	917360.4	1309085.0

Рисунок А.9 — Зберігання даних в таблиці ВПМ

	Year	Value	Scaled_Value	Predicted
0	2017	5234361.3	0.306766	5976783.5
1	2018	6317455.4	0.440104	6485650.5
2	2019	7304805.0	0.561655	7066185.0
3	2020	7286736.1	0.559430	7814413.5
4	2021	8583406.1	0.719061	8381227.5
5	2022	7653500.4	0.604582	8982820.0
6	2023	10865448.2	1.000000	9131335.0
7	2024	NaN	NaN	10089708.0
8	2025	NaN	NaN	10582493.0

Рисунок А.10 — Вигляд таблиці з масштабованими даними

```
Після масштабування:
count    10.000000
mean      0.447790
std       0.304931
min       0.000000
25%       0.209118
50%       0.499767
75%       0.593850
max       1.000000
Name: Scaled Value, dtype: float64
```

Рисунок А.11 — Вигляд даних після масштабування

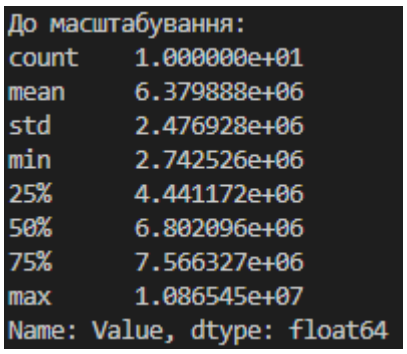

Рисунок А.12 — Вигляд даних до масштабування

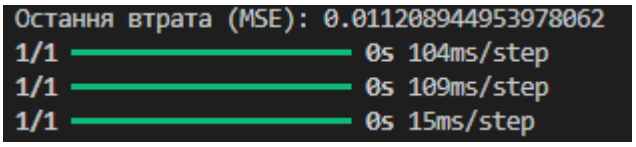

Рисунок А.13 — Втрата MSE

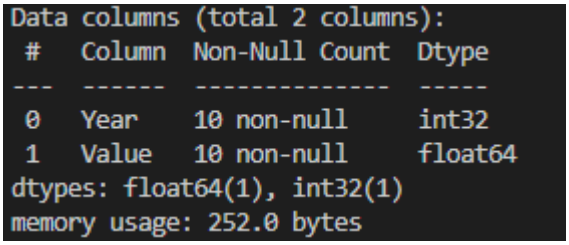

Рисунок А.14 — Вигляд таблиці для роботи з даними

ДОДАДОК Б. Лістинг програми

App.py

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import MinMaxScaler
from keras.models import Sequential
from keras.layers import LSTM, Dense
import matplotlib.pyplot as plt

file_path = 'C:/diplom/statistic.xlsx'
data = pd.ExcelFile(file_path)

sheet1_data = data.parse('ВПМ')

print("Перші кілька рядків таблиці:")
print(sheet1_data.head())

ukraine_data = sheet1_data.iloc[2, 2:].reset_index(drop=True)
years = sheet1_data.iloc[0, 2:].reset_index(drop=True)

print("Роки:", years.head())
print("Значення для України:", ukraine_data.head())

ukraine_df = pd.DataFrame({
    "Year": years,
    "Value": ukraine_data
})

ukraine_df = ukraine_df.dropna()
ukraine_df["Year"] = ukraine_df["Year"].astype(int)
```

```
ukraine_df["Value"] = pd.to_numeric(ukraine_df["Value"], errors='coerce')
ukraine_df = ukraine_df.dropna().reset_index(drop=True)

print("Інформація про очищені дані:")
print(ukraine_df.info())
print(ukraine_df.head())

scaler = MinMaxScaler(feature_range=(0, 1))
ukraine_df["Scaled_Value"] = scaler.fit_transform(ukraine_df[["Value"]])

print("До масштабування:")
print(ukraine_df["Value"].describe())
print("Після масштабування:")
print(ukraine_df["Scaled_Value"].describe())

def create_sequences(data, sequence_length=3):
    x, y = [], []
    for i in range(len(data) - sequence_length):
        x.append(data[i:i + sequence_length])
        y.append(data[i + sequence_length])
    return np.array(x), np.array(y)

sequence_length = 3
data = ukraine_df["Scaled_Value"].values
X, y = create_sequences(data, sequence_length)

print("Перша послідовність (X):", X[0])
print("Перше цільове значення (y):", y[0])
print("Форма X:", X.shape)
print("Форма y:", y.shape)
```

```
X = X.reshape((X.shape[0], X.shape[1], 1))

model = Sequential([
    LSTM(50, activation='relu', input_shape=(X.shape[1], X.shape[2])),
    Dense(1)
])

model.compile(optimizer='adam', loss='mse')

print("Навчання моделі...")
history = model.fit(X, y, epochs=200, verbose=1)

print("Остання втрата (MSE):", history.history['loss'][-1])

predicted = model.predict(X)

predicted_rescaled = scaler.inverse_transform(predicted)

result_df = ukraine_df.iloc[sequence_length:].copy()
result_df["Predicted"] = predicted_rescaled

future_years = list(range(ukraine_df["Year"].iloc[-1] + 1, 2026))
last_sequence = data[-sequence_length:]

future_predictions = []
for year in future_years:

    last_sequence_reshaped = last_sequence.reshape(1, sequence_length, 1)
    next_pred = model.predict(last_sequence_reshaped)
```

```
            future_predictions.append(next_pred[0, 0])


            last_sequence = np.append(last_sequence[1:], next_pred)

        future_predictions_rescaled =
scaler.inverse_transform(np.array(future_predictions).reshape(-1, 1))

        future_df = pd.DataFrame({
            "Year": future_years,
            "Value": np.nan,
            "Predicted": future_predictions_rescaled.flatten()
        })

        final_df = pd.concat([result_df, future_df], ignore_index=True)

        print("Останні записи фінального DataFrame:")
        print(final_df.tail(10))

        plt.figure(figsize=(10, 6))
        plt.plot(final_df["Year"], final_df["Value"], label="Реальні дані", marker="o")
        plt.plot(final_df["Year"], final_df["Predicted"], label="Прогноз LSTM",
linestyle="--")
        plt.xlabel("Рік")
        plt.ylabel("Значення")
        plt.title("Прогноз обсягу виробленої продукції, морозива до 2025 року")
        plt.legend()
        plt.grid(True)
        plt.show()
```