

Міністерство освіти і науки України
Львівський національний університет природокористування
Факультет механіки, енергетики та інформаційних технологій
Кафедра інформаційних технологій

КВАЛІФІКАЦІЙНА РОБОТА

першого (бакалаврського) рівня вищої освіти

на тему:

«Інтелектуальна інформаційна система прогнозування попиту на
сільськогосподарську продукцію»

Виконав: студент групи Іт-22сп
спеціальності 126
«Інформаційні системи та технології»

Юр Р.Ю.

Керівник:

Лиса О.В.

Рецензент:

Власовець В.М.

ЛЬВІВ-2023

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
 ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ПРИРОДОКОРИСТУВАННЯ
 ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
 КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Освітній ступінь «Бакалавр» за спеціальністю –
 126 – „Інформаційні системи та технології”

“ЗАТВЕРДЖУЮ”

Завідувач кафедри _____
 д.т.н., проф. А.М. Тригуба
 “ _____ ” _____ 202_ р.

ЗАВДАННЯ

на кваліфікаційну роботу студенту

Юр Руслан Юрійович

1. Тема роботи: «Інтелектуальна інформаційна система прогнозування попиту на сільськогосподарську продукцію»

Керівник роботи Лиса Ольга Володимирівна, к.т.н., доцент.

Затверджені наказом по університету від _____ лютого 2023 року № _____ /к-с.

2. Строк подання студентом роботи 2.06.2023 р.

3. Початкові дані до роботи: 1. Історичні дані біржової торгівлі
сільськогосподарськими товарами 2. Варіанти
інтелектуальних моделей для прогнозування

4. Зміст розрахунково-пояснювальної записки:

1. Аналіз предметної області

1.1. Інформаційні системи та їх класифікація

1.2. Інформаційні технології в сільському господарстві

2. Постановка задачі

2.1. Вибір методу дослідження ринку

2.2. Особливості аналізу часових фінансових рядів

2.3. Математична модель стратегії торгівлі

2.4. Вибір інструментарію та мови програмування

3. Розробка інформаційної системи

3.1. Розробка структурної схеми іс

3.2. Розвідувальний аналіз для інтелектуальної моделі

3.3. Розробка моделі машинного навчання та тестування

3.4. Розробка елементів графічного інтерфейсу користувача

4. Охорона праці

4.1. Основні положення охорони праці

4.2. Шкідливі та небезпечні виробничі фактори

4.3. Вимоги до персоналу при ремонті та обслуговуванні електрообладнання

Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень): Аналіз предметної області. Класифікація інформаційних систем. Вибір математичної моделі. Інструментарій вирішення задачі. Розвідувальний аналіз даних. Створення інтелектуальної моделі. Елементи графічного інтерфейсу користувача. Висновки.

6. Консультанти з розділів:

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1, 2, 3	<i>Лиса О.В., доцент кафедри інформаційних технологій</i>		
4	<i>Городецький І.М., доцент кафедри управління проектами та безпеки виробництва</i>		

7. Дата видачі завдання 19 лютого 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	<i>Написання першого розділу та означення головних завдань роботи</i>	19.02 - 21.03.23	
2	<i>Виконання другого розділу та формування початкових даних</i>	22.03 - 11.04.23	
3.	<i>Виконання третього розділу та узагальнення отриманих результатів роботи</i>	12.04 - 11.05.23	
4.	<i>Написання розділу: «Охорона праці»</i>	12.05 - 17.05.23	
5.	<i>Вартісне оцінення ефективності пропозицій роботи</i>	18.05 - 23.05.23	
6.	<i>Завершення роботи в цілому</i>	24.05 - 02.06.23	

Студент _____ Юр Р.Ю.
(підпис)

Керівник роботи _____ Лиса О.В.
(підпис)

УДК 004.8:338.43

Інтелектуальна інформаційна система прогнозування попиту на сільськогосподарську продукцію. Юр Р.Ю. – Кваліфікаційна робота бакалавра. Кафедра інформаційних технологій – Дубляни, ЛНУП, 2023.

50 сторінок текстової частини, 26 рисунків, 1 таблиця та 14 джерел інформації.

Текстова частина включає вступ, чотири розділи, висновки, список використаних джерел.

У вступі висвітлено актуальність обраної теми, а також важливість інформаційних систем в вирішенні задач в агропромисловому комплексі.

У першому розділі зроблено аналіз предметної області, подано варіанти класифікації інформаційних систем, зокрема, їх можливості в сільському господарстві, визначено тип і належність системи, яка розробляється.

У другому розділі обґрунтовано вибір відповідних для даної роботи: математичного апарату, інструментарію програмування та методів рішення.

У третьому розділі розроблена структурна схема інформаційної системи, реалізована інтелектуальна модель, створені елементи графічного інтерфейсу користувача.

У четвертому розділі розроблені заходи з охорони праці.

По результатам виконаної роботи зроблені відповідні висновки.

Ключові слова: інформаційна система, машинне навчання, аналіз часових рядів, прогнозування попиту, інтелектуальна модель.

Keywords: information system, machine learning, time series analysis, demand forecasting, intelligent model.

ЗМІСТ

ВСТУП.....	6
1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	8
1.1. Інформаційні системи та їх класифікація	8
1.2. Інформаційні технології в сільському господарстві	14
2. ПОСТАНОВКА ЗАДАЧІ.....	17
2.1. Вибір методу дослідження ринку.....	17
2.2. Особливості аналізу часових фінансових рядів.....	20
2.3. Математична модель стратегії торгівлі.....	23
2.4. Вибір інструментарію та мови програмування	28
3. РОЗРОБКА ІНФОРМАЦІЙНОЇ СИСТЕМИ	32
3.1. Розробка структурної схеми ІС	32
3.2. Розвідувальний аналіз для інтелектуальної моделі.....	33
3.3. Розробка моделі машинного навчання та тестування	39
3.4. Розробка елементів графічного інтерфейсу користувача	40
4. ОХОРОНА ПРАЦІ	44
4.1. Основні положення охорони праці.....	44
4.2. Шкідливі та небезпечні виробничі фактори	45
4.3. Вимоги до персоналу при ремонті та обслуговуванні електрообладнання .	46
ВИСНОВКИ	48
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	49

ВСТУП

Сьогодні інформаційні системи міцно увійшли до всіх сфер нашого життя. Машинний переклад, розпізнавання образів, генеративні інтелектуальні системи та багато іншого вже нікого не дивує. Однак ефективність впровадження нових інформаційних систем продовжує доводити, що ще є багато сфер людської діяльності, які потребують удосконалення та інновацій. Однією з таких сфер є агропромисловий комплекс та, зокрема, ринок сільськогосподарської продукції.

Розвиток агропромислового комплексу в сьогоднішніх умовах значною мірою залежить від інформаційного супроводу, тобто своєчасності, доступності та достовірності інформації про поточний стан сільського господарства, наявні проблеми та суперечності, передбачуваних шляхах їх вирішення, тісної взаємодії всіх галузей та рівнів управління, впровадження в процес управління та виробництва інформаційних систем, нових технологій, досягнень науки та техніки.

Звісно, сьогодні у зв'язку із ситуацією в країні складно говорити про довгострокові тенденції. Але ті компанії, хто розуміє важливість майбутніх змін і можливостей, не просто залишаються на екстенсивному шляху розвитку завдяки різниці цін на ціноутворюючі статті витрат, а знаходять сили інвестувати в інтенсифікацію та інформатизацію як у виробництво, так і в процеси управління.

Все це характеризує необхідність та наявність передумов для здійснення технічного та технологічного переоснащення шляхом впровадження інноваційних та інформаційних технологій.

Модернізація аграрного виробництва та торгівлі на основі впровадження інформаційних систем управління, використання нових технологій у процесі виробництва, управління та реалізації продукції сприятиме зростанню

аграрного сектору, зміцненню конкурентоспроможності та створення сталого стану на внутрішньому та світовому ринках.

Таким чином, створення та впровадження інформаційних систем, подібних до цієї роботи, є актуальним завданням і одним з найважливіших факторів розвитку сучасного агропромислового комплексу. Саме цим обумовлений вибір теми дослідження та визначається її актуальність.

Об'єктом роботи є створення інтелектуальних інформаційних систем в галузях, пов'язаних з сільським господарством.

Предметом дипломної роботи є розробка інтелектуальної інформаційної системи прогнозування попиту на сільськогосподарську продукцію.

Мета роботи – підвищення ефективності процесу реалізації сільськогосподарської продукції за рахунок розробки сучасної моделі із застосуванням елементів штучного інтелекту.

РОЗДІЛ 1

АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1. Інформаційні системи та їх класифікація

Поширений термін «інформаційна система» (ІС) не має суттєво уніфікованого понятійного визначення. Найчастіше це поняття трактується як «комплекс, що складається з інформаційної бази та процедур» для управління, оновлення, пошуку і завершення збору, зберігання, виправлення та спільного використання інформації. [1, 2]

Таке інтуїтивне визначення ІС пов'язане з усталеною і звичною, проте особливою формою цілеспрямованої діяльності людини - обробкою інформації як відомостей про щось, матеріально представлених на традиційних або машинних носіях і що забезпечують ефективність вирішення завдань її основної діяльності. Склад і структура ІС визначається, з вимог до рівня ефективності обслуговування інформаційних потреб кінцевих користувачів, передусім у частині перебування у накопичених масивах тих записів (документів), які імовірно містять потрібні відомості.

Таким чином, можна говорити, що інформаційною системою називають таку систему обробки інформації та такі інформаційні ресурси, які здатні забезпечити та поширити інформацію. До таких ресурсів належать людські, технічні, фінансові та інші. Інакше кажучи, до інформаційних систем відноситься така сукупність засобів, методів і персоналу, яка взаємопов'язана між собою і застосовується не тільки для зберігання інформації, але і для її обробки та видачі, досягаючи при цьому певної мети управління.

Інформаційний фонд системи, зазвичай, представляє базу чи сукупність баз даних (БД), створюваних як табличних, ієрархічних і мережевих структур. В інтернет-ресурсах широко використовується гіпертекст, здатний представляти модель організації інформаційного простору предметної галузі.

Тобто в цьому випадку не йдеться про класичні БД, а про деяку файлову, як правило, ієрархічну структуру.

У мережеских інформаційних системах використовують два способи взаємодії з кінцевими користувачами:

1. Розподіл часу, коли кожен учасник мережі якби користується своєю ЕОМ. Основне завдання розробників та адміністраторів мережі – захист даних від несанкціонованого доступу та забезпечення взаємної ізоляції учасників.

2. Забезпечення групових рішень, що передбачає організацію взаємодії користувачів у процесі прийняття рішень. Цей метод поєднує комунікаційну, обчислювальну технологію та технологію прийняття рішень для реалізації групою осіб складних неструктурованих завдань.

Будь-яка інформаційна система повинна мати такі властивості:

- функціональність, тобто будь-який її об'єкт повинен містити функціонально закінчену та максимально незалежну сукупність операцій з обробки даних, причому ззовні об'єкт представляє єдине ціле, а будь-яке поводження з використанням стандартизованого інтерфейсу відбувається до об'єкта в цілому;

- пов'язаність означає, що в об'єкті реалізується сукупність взаємозалежних функцій – методів, що працюють з тими самими даними, деякі з яких приховані для системи в цілому;

- маскуванню, тобто для системи доступні лише ті параметри об'єкта, які складають набори вхідного та вихідного інтерфейсів об'єкта, а функції та дані, що не беруть участь у взаємодії з системою, з іншими її об'єктами, приховані та поза об'єктом недоступні.

Так, наприклад, важливою властивістю ІС є мінімізація числа її інформаційних зв'язків, від яких залежать витрати на модифікацію системи при її функціонуванні в умовах, що змінилися, а також вирішенні нових функціональних завдань.

Єдина класифікація інформаційних систем ще не склалася остаточно, проте, певні розробки існують. Розглянемо найважливіші аспекти класифікації

ІС [3]. За призначенням функціонуючої інформації ІС поділяються на: державні, юридичні (законодавчі), ділові, фінансові, науково-технічні, навчальні, соціальні, розважальні та інші. При цьому, наприклад, фінансова інформація поділяється на: бухгалтерську, банківську, податкову та іншу, а медична (як і інші) може містити всі перелічені вище функції.

За рівнем управління виділяють стратегічні, тактичні та оперативні інформаційні системи.

За рівнем застосування технічних засобів інформаційні системи ділять на автоматизовані та неавтоматизовані. При цьому автоматизовані мають на увазі автоматизацію від окремих процесів та завдань до рівня автоматизації підприємств, установ та їх сукупності в масштабах території (регіону), тобто представляють клас систем, орієнтованих на автоматизацію окремих функцій чи процесів та клас інтегрованих систем та комплексів. Ступінь автоматизації залежить від різних причин. Вона передбачає електронну обробку та доставку даних, автоматизацію функцій та процесів управління, підтримку прийняття рішень та ін.

За типами інформації, а саме: документальної, фактографічної та документально-фактографічної пропонується виділити три типи ІС.

Документальні ІС включають інформаційно-пошукові системи (ІПС), інформаційно-логічні та інформаційно-семантичні системи.

Фактографічні ІС діляться на дві категорії: 1) системи обробки даних (СОД); 2) автоматизовані інформаційні системи (АІС) та автоматизовані системи управління (АСУ).

Документально-фактографічні ІС містять:

- 1) автоматизовані документально-фактографічні інформаційно-пошукові системи науково-технічної інформації (АДФІПС НТІ),
- 2) автоматизовані документально-фактографічні інформаційно-пошукові системи в автоматизованій системі нормативно-методичного забезпечення управління (АДФІПС в АСНМОУ).

Фахівці виділяють також інші ІС: бухгалтерські, банківські ІС, ІС ринку цінних паперів, інформаційні системи управління (ІСУ), системи підтримки прийняття рішень (СППР), експертні системи (ЕС), гібридні експертні системи (ГЕС), інформаційні системи моніторингу (ІСМ) та інші.

В АІС розміщують різні види інформації: бібліографічні дані (записи); фактографічні дані (записи); повнотекстові документи (записи); довідкові дані (у тому числі покажчики); математичні або чисельні (цифрові, табличні) дані; графічні дані; мультимедійні дані.

Відповідно, ІС можна класифікувати за видами оброблюваної інформації: текстові процесори та редактори (текст); графічні процесори та редактори (графіка); системи управління базами даних (СУБД), табличні процесори, алгоритмічні мови програмування (дані); експертні системи (знання), мультимедійні системи (об'єкти реального світу, що включають будь-які види інформації) та ін. Звичайно, така класифікація досить умовна. Так, сучасний текстовий процесор може забезпечувати присутність та взаємодію практично будь-яких видів інформації, гіпертексту та можливості комунікацій. Інша річ, наскільки він задовольнятиме відповідних користувачів. Наприклад, презентаційні матеріали доцільно створювати у спеціалізованому ПЗ.

Класифікація по сферах застосування має вигляд, представлений на рис.1.1. В свою чергу, кожен тип має свої класи систем, а саме:

- системи обробки транзакцій, які включають оперативну та пакетну обробку транзакцій;
- системи підтримки прийняття рішень, куди відносять системи оперативної аналітичної обробки та експертні системи;
- інформаційно-довідкові системи з такими класами, як системи електронної документації, географічні інформаційні системи та гіпертекстові системи;
- офісні інформаційні системи, які поділяють на документальні системи, системи автоматизації діловодства та управління документообігом.



Рисунок 1.1 - Класифікація ІС по сферах застосування

За способом організації інформаційних систем можна виділити такі їх типи (рис.1.2):

- системи на основі архітектури файл-сервер;
- системи на основі архітектури клієнт-сервер;
- системи на основі багаторівневої архітектури;
- системи на основі Інтернет-технологій.



Рисунок 1.2 - Класифікація ІС за способом організації

Забезпечення ІС поділяється на: інформаційне, технічне, математичне та програмне, методичне, лінгвістичне, правове та організаційне (рис. 1.3).

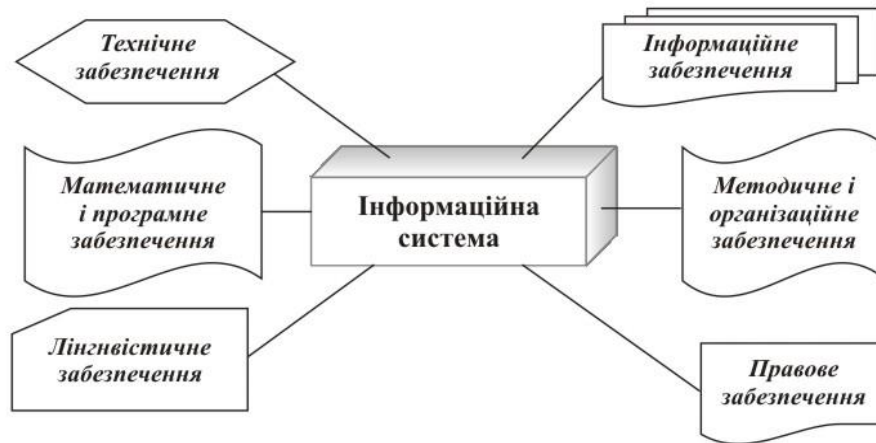


Рисунок 1.3 - Забезпечення інформаційної системи

Інформаційне забезпечення включає сукупність даних, методи побудови БД, а також проектних рішень щодо обсягів, розміщення, форм організації інформації, що циркулює в інформаційній системі організації.

Технічне забезпечення має на увазі комплекс технічних засобів, призначених для роботи ІС, документацію на ці засоби та технологічні процеси.

Математичне забезпечення представляє сукупність математичних методів, моделей, алгоритмів обробки інформації, що використовуються при вирішенні функціональних та проектних завдань у ІС.

Програмне забезпечення означає сукупність програм для реалізації цілей, завдань ІС та нормального функціонування як окремих, так і комплексу технічних засобів.

Методичне та організаційне забезпечення складають комплекс методів, засобів та документів, що регламентують взаємодію персоналу ІС з програмно-технічними засобами та між собою у процесі розробки та експлуатації ІС.

Лінгвістичне забезпечення має на увазі сукупність мов спілкування персоналу ІС та користувачів з програмно-технічним та інформаційним забезпеченням, а також суму термінів, що використовуються в ІС.

Правове забезпечення представляється правовими нормами та використовується для дотримання законності. До складу цього виду

забезпечення входять закони, укази, постанови державних органів влади, накази та інструкції вищих органів та керівників організації.

1.2. Інформаційні технології в сільському господарстві

Інформаційні технології (ІТ) не тільки стрімко увійшли до нашого життя, дозволяючи спростити вирішення складних технічних завдань, але й підвищити продуктивність праці, у тому числі й у галузях агропромислового комплексу.

Фахівці визначають ІТ як комбінації трьох ключових технологій: числові обчислення, зберігання інформації та трансляція числових сигналів по телекомунікаційних мережах. У вітчизняній літературі ІТ визначаються найчастіше, як технології, що використовують засоби мікроелектроніки для збирання, зберігання, обробки, передачі та подання даних, текстів, образів та звуків.

Сільське господарство – ідеальне середовище для застосування ІТ. У зв'язку з цим для ефективного і стійкого функціонування суб'єктів господарювання в нових умовах необхідно застосовувати передові інформаційні технології, що дозволяють виявити їх внутрішні резерви, залучити зовнішні вкладення, а також проводити реструктуризацію організаційних структур і виконувати реінжиніринг систем управління [4-7].

У сфері сільського господарства розвинених країн дедалі частіше з'являються умови та докладаються значних зусиль щодо впровадження інформаційних технологій. Найбільш відомі технології реалізовані у рамках прикладних комп'ютерних програм.

Це, насамперед, програми оптимізації розміщення сільськогосподарських культур у системах сівозміни та раціонів годівлі тварин; з розрахунку доз добрив; проведення комплексу землевпорядних робіт та управління земельними ресурсами; ведення державного земельного кадастру історії полів та розроблення технологічних карт вирощування сільськогосподарських культур; регулюванню режиму живлення рослин та мікроклімату в теплицях;

контролю процесу зберігання картоплі та овочів, якості вирощуваної продукції та кормів, забруднення ґрунтів; оцінки економічної ефективності виробництва; управлінню технологічними процесами в пташниках, виробничими процесами у переробці м'яса птиці та зберіганні продукції та багато іншого.

Сучасний ринок ІТ пропонує рішення практично для будь-якого виробництва, починаючи від вирощування пшениці та закінчуючи виведенням нових порід курей. Однак для кожного такого рішення є обмеження за мінімальним (втім, як і максимальним) розміром підприємства, в рамках яких впровадження буде ефективним.

Одним з актуальних напрямків використання ІТ в АПК стає точне землеробство, яке забезпечує стратегію управління врожайністю сільськогосподарських культур, що використовує глобальну систему позиціонування (GPS), географічні інформаційні системи (ГІС) та технології, та дані з множинних джерел про умови зростання та розвитку рослин та економічну ситуації кожної одиниці управління межах окремо взятого поля.

Одна з ознак застосування ІТ у господарствах – наявність комп'ютерів, а також їх з'єднання з Інтернетом. ІТ використовуються переважно для бухгалтерського обліку, автоматизації сільськогосподарських процесів.

Швидкість і обсяги інформації, що передається, через різні системи зв'язку забезпечать стабільне постачання виробників базами даних. Ці дані мають бути інтегровані до особливостей біологічних та фізичних систем для того, щоб отримати корисні знання про їхній поточний стан та прогнозувати результати можливих рішень. Впровадження наукових розробок шляхом використання Інтернету є надзвичайно важливим для розширення функціональних можливостей інформаційних систем.

Ведення сільського господарства в інформаційному суспільстві передбачає безперервне отримання інформації від зовнішніх джерел (через зовнішні мережі Інтернету) у час з будь-якої точки місцевості. Наприклад, постійне оновлення даних синоптиків може бути доступним протягом дня. Це дозволяє підвищити ефективність застосування хімічних засобів захисту

рослин, а також зменшує забруднення довкілля. Відомі розробки систем попередження про появу шкідників та хвороб рослин.

Сучасні ІТ дозволяють отримувати поради, рекомендації, незалежно від часу та місця їх розташування. Керівник може описувати свої проблеми через звичайну мову, ілюстровану фотографіями чи відеозаписами. При цьому час та розташування визначаються автоматично. Потім він може за допомогою електронної пошти надіслати свої матеріали підтримуючим службам ведення сільського господарства та отримати відповідь через деякий час, або він може вирішувати свою проблему в діалоговому режимі безпосередньо через Інтернет.

Розширення інформаційних баз даних – важлива, але недостатня умова для ефективного їх застосування у господарствах. Вихідна інформація має бути зручною для оцінки біологічних та фізичних систем з метою вироблення корисних знань про поточний стан господарств, а також прогнозування результатів під час реалізації різних сценаріїв. Знання, що накопичилися в сільськогосподарських дослідженнях, протягом багатьох років повинні бути застосовані для отримання практично корисної інформації шляхом обробки баз даних. Це означає, що ІТ – незамінне джерело для реалізації науково-дослідних розробок.

Підсумовуючи описане вище, можна класифікувати систему, яка розробляється в даній роботі, як систему підтримки прийняття рішень, а саме виконання оперативної аналітичної обробки. При цьому дані будуть зберігатись в базі даних, а щоб мати до неї доступ, треба буде використовувати клієнт-серверні технології.

РОЗДІЛ 2

ПОСТАНОВКА ЗАДАЧІ

2.1. Вибір методу дослідження ринку

Для того, щоб аналізувати інформацію про стан ринку, зокрема ринку сільськогосподарської продукції та попит на ньому, перш за все треба визначитись з джерелом такої інформації та методами її отримання.

Існує два основних типи способів збору інформації про ринок [8]: кількісні та якісні методи дослідження ринку (рис. 1.1).



Рисунок 2.1 - Методи дослідження ринку

Кількісні методи дослідження ринку надають структуровану та статистично опрацьовану інформацію про ринок. За допомогою кількісних методів можна отримати точні цифри, на підставі яких зробити прогноз продажу або оцінити обсяг ринку.

Результатом якісних методів аналізу ринку є гіпотези, ідеї, окремі думки, які не є структурованими і не піддаються статистичній оцінці, а можуть бути проаналізовані лише суб'єктивно.

Розрізняють п'ять основних методів дослідження ринку: опитування, фокус-групи, глибинні інтерв'ю, польові дослідження чи експерименти,

спостереження. Вибір методу дослідження залежить від бюджету та часових ресурсів. Короткий опис кожного з них наведено у табл. 1.1.

Таблиця 1.1 - Види дослідження ринку

Види дослідження ринку	Опис методів дослідження
Опитування (surveys)	Опитування цільового ринку за строго заданою анкетною. За розміром може бути як великими, і маленькими. В опитуванні важлива вибірка: чим вона більша, тим більше правильні та репрезентативний результат ви отримаєте. Це кількісний метод, який використовується, коли необхідно отримати точну цифру на конкретне питання.
Фокус-групи (focus groups)	Дискусія або круглий стіл, за яким відбувається обговорення заданої теми цільовою групою споживачів. Є модератор, який керує дискусією щодо заданого переліку питань. Якісний метод, ефективний розуміння причин поведінки, дослідження прихованих мотивів споживача, допомагає сформулювати гіпотези.
Глибинні інтерв'ю (personal interviews)	Бесіда з одним представником цільової аудиторії з певного переліку відкритих питань. Допомагають розібратися у питанні докладно та сформулювати гіпотези. Якісний метод дослідження.
Спостереження (observation)	Спостереження за представником цільової аудиторії у звичному для нього середовищі (наприклад: відеозйомка в магазині). Якісний метод дослідження.
Експерименти чи польові дослідження (experiments and field trials)	Кількісний спосіб дослідження ринку. Допомагає перевірити певні гіпотези та альтернативи на практиці.

Виходячи з вищенаведеного списку, можна з'ясувати, що перші чотири метода дослідження необхідно проводити з багаточисельною ринковою аудиторією, для чого треба мати в наявності власні людські ресурси. Дана

робота не націлена на це, і відтак було обрано метод кількісних досліджень. Однак для такого методу потрібні вхідні дані по ринку, котрі не завжди є в наявності у дослідника. Тому було прийнято рішення звернутись до даних біржового ринку сільськогосподарської продукції, які є у відкритому доступі. Логіка тут наступна.

Біржові ціни мають коливання в часі. Як правило, ринок намагається знайти баланс між попитом та пропозицією. Чим швидше зростає попит, а пропозиція не встигає за цим ростом - тим вище і ціна, і навпаки. Тобто ми спостерігаємо пряму кореляцію між попитом та ціною, і тому, якщо вдасться створити модель, яка прогнозує (хоча б з достатньою точністю) біржові ціни - тим самим ми можемо говорити про прогнозування попиту.

Тут треба зазначити, що прогнозування біржових цін не може здійснюватися з точністю, яку, наприклад, вже досягли моделі штучного інтелекту в розпізнаванні образів, текстів і т.д.

Справа в тому, що біржова ціна - це нестаціонарний процес у часі, через який, крім того, ринкові гравці намагаються дисконтувати майбутні ціни, причому кожен має власні цілі.

Спекулянти намагаються отримати прибуток, граючи на різниці цін. Виробники реалізують реальний товар та здійснюють спроби захисту від падіння ціни. Покупці товарів (або сировини, якщо річ йде про виробника) - навпаки, намагаються купити якомога дешевше, одночасно захищаючись від росту цін.

І все це в підсумку впливає на ціну, причому чим більшими об'ємами оперує окрема структура, тим суттєвіше її дії можуть скорегувати ціну.

В свою чергу, зміни ціни провокують інших гравців на відповідні дії. Тому ми маємо ще й множину зворотних зав'язків, які складно формалізувати. Тим не менш, досягти не стовідсоткової, але прийнятної точності в прогнозуванні біржових цін можливо.

І в цьому якраз допомагає той фактор, що ціни не є абсолютно випадковими, а мають деяку прив'язку до минулого. Тобто, аналізуючи історію

цін в минулому, особливо в недалекому, можна спробувати не тільки прогнозувати подальшу їх поведінку, але й планувати відповідні свої дії.

2.2. Особливості аналізу часових фінансових рядів

Аналіз часових фінансових рядів (іншими словами, технічний аналіз біржових цін) має свої особливості [9]. Технічний аналіз - це вивчення ціни в минулому для передбачення її в майбутньому за допомогою різних інструментів: індикаторів, графічних фігур та ін. Він не відповідає на питання «Чому змінилася ціна?», а лише показує найімовірнішу зміну ціни, виходячи з поведінки ринку в минулий період.

Ринок наочно ілюструє потоки пропозиції та попиту, які відображають надії та думки людей та формують їхню поведінку на ринку. Виходячи з цього, технічний аналіз повинен використовувати цінові коливання та сигнали індикаторів, маючи на увазі, що їхні рухи повторюватимуться в майбутньому.

Технічний аналіз ґрунтується на трьох базових постулатах.

1. Ринкова ціна враховує все. Згідно з цією аксіомою вся інформація, необхідна для ухвалення торгового рішення, вже врахована у графіку ціни та обсягах торгів. Передбачається, що ціна майже відразу реагує на всі події, які здатні вплинути на неї.

2. Рух цін підпорядкований тенденціям. На ринку періодично виникає тренд, який продовжується деякий час. Він може бути зростаючим або падаючим, залежно від співвідношення попиту та пропозиції на актив. За наявності тренда можуть бути періоди протилежного руху ціни (корекція). Тенденція може закінчитися будь-якої миті часу, тобто вона кінцева. При цьому на ринку можлива ситуація відсутності тренду, яка називається бічним рухом (боковик, флет). По суті, це горизонтально спрямований рух ціни протягом тривалого часу.

3. Історія повторюється. Якщо у минулому ціна змінювалася певним чином в певних ситуаціях, тобто є підстави вважати, що у майбутньому у таких самих ситуаціях ціна реагуватиме подібним чином.

Потік даних про біржові угоди може бути представлений у вигляді цінового графіка - лінії, кожна точка якої відображає інформацію про ціну та час здійснення кожної окремої угоди. Такий графік має звичайний лінійний вигляд і є тиковим за часовим інтервалом побудови (від англ. tick - цокання, мить). Тиковий графік відображає дані абсолютно про всі вчинені угоди, проте аналізувати його незручно саме через перевантаженість інформацією.

Для того, щоб оптимізувати процес графічного аналізу та отримати можливість прийняття рішень відповідно до обраного темпу торгівлі, дані про ціни та обсяг торгів зазвичай наносяться на графік не суцільним потоком, а відповідно до обраного часового інтервалу – таймфрейм.

Усі технічні індикатори та моделі базуються на 5 числах, що відображаються на графіку. Дані числа відповідають конкретному часовому інтервалу (наприклад, візьмемо 1 годину):

- ціна відкриття - ціна, з якої починаються торги в цю годину;
- ціна закриття – ціна останньої угоди за цю годину;
- максимальна ціна – найвища ціна операції протягом години;
- мінімальна ціна – найнижча ціна угоди протягом години;
- обсяг торгів – кількість угод за годину.

Технічний індикатор – це результат математичних розрахунків, зроблених на підставі даних про ціни та обсяги торгів. Індикаторний технічний аналіз дозволяє визначити стан ринку в поточний період та прийняти рішення про угоду на основі певних сигналів.

Різні індикатори мають різну природу та по-різному використовуються для аналізу та прогнозування цінового руху. Серед усіх індикаторів виділяються три основні групи:

- трендові індикатори;
- осцилятори;

- індикатори об'єму.

Трендові індикатори дають достовірні сигнали здійснення операцій на сильних трендових рухах. Осцилятори доцільніше використовуватиме торгівлі при послаблюванні тенденції чи в боковиках. Індикатори обсягу переважно використовуються як додаткові, незалежно від фази трендового руху, для підтвердження чи спростування поточного руху обсягом.

Використання індикаторів - це спроба описати поточний стан справ на ринку, зрозуміти розстановку сил продавців і покупців і на підставі цієї інформації зробити прогноз зміни ціни.

В зв'язку з тим, що ми розглядаємо можливості не спекулятивної торгівлі, а ринок з точку зору постачальника, тому нам найцікавішими будуть тенденції, тобто трендові індикатори.

Трендові індикатори є математичним усередненням ціни. На основі цього можна припустити напрямок руху ціни (тренд) у майбутньому.

Ковзна середня (рис. 2.2) – це найпростіший і найпоширеніший трендовий індикатор. Вона є середню ціну паперу за вибраний період і дозволяє трейдеру розглядати згладжені дані, а не фокусуватися на щоденних коливаннях цін. При цьому чим більший період розрахунку, тим більш згладженою щодо цінового графіка буде ковзна середня.



Рисунок 2.2 – Проста ковзна середня

2.3. Математична модель стратегії торгівлі

Ми прогнозуємо попит, тобто розглядаємо можливості ринку для постачальника або продавця товару (сільськогосподарської продукції). Якщо впродовж деякого періоду часу, наприклад року, постачальник кожен раз виходив би на ринок і рівномірно продавав товар, то ціна продажу мала б однаковий розподіл з ринковими цінами (приклад показаний на рис. 2.3)

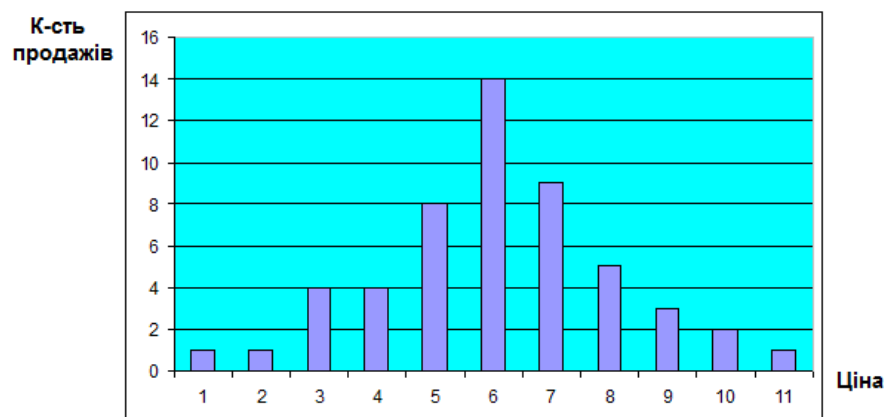


Рисунок 2.3 – Приклад розподілу ринкових цін

Для даного випадку середня ціна склала би 5,98 умовних одиниць.

Наприклад, розглянемо випадок при ідеальному прогнозі, коли ми знаємо наперед майбутній розподіл цін. Звісно, ми не можемо впливати на самі ціни, однак можемо міняти рівень своїх продаж у відповідності до прогнозування – збільшувати продажі у випадку високих цін і зменшувати у випадку низьких. Тоді за умовами обов’язкової хоча б мінімальної участі в продажах графік розподілу продаж, наприклад, міг би мати лінійний вигляд, показаний на рис.2.4.

Середня ціна продажів склала би 7,88, що перевищує середній рівень ринкових цін на 32%. Звісно, така ситуація цілком гіпотетична і в реальності неможлива.

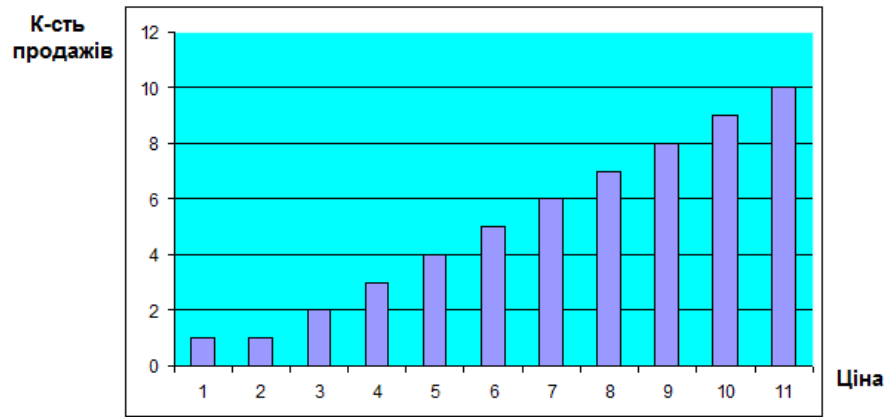


Рисунок 2.4 – Лінійна стратегія продажів

Якщо система має хоча б невелику здібність прогнозування, то розподіл продажу можна спробувати зсунути в бік більших цін і добитись, наприклад, такої умовної картини, яка показана на рис.2.5.

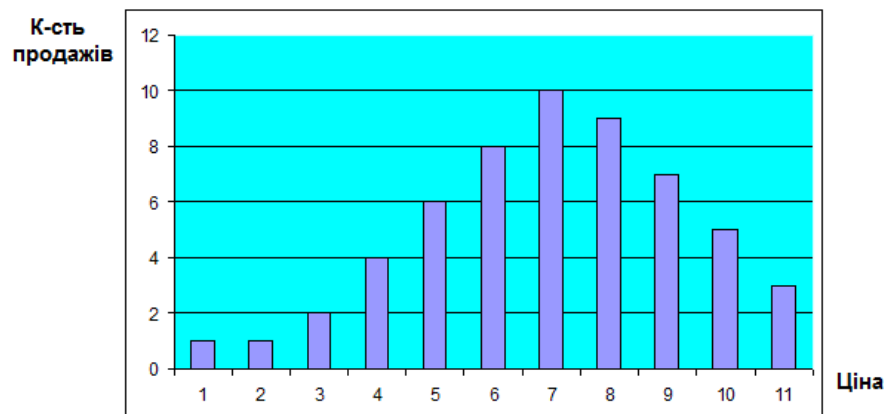


Рисунок 2.5 – Можливий реальний результат стратегії

Середня ціна продажів в такому випадку склала би 6,98 одиниць, що перевищує середній рівень ринкових цін майже на 17%, а це вже досить суттєве досягнення.

Для коректної роботи модель машинного навчання потребує первісної обробки вхідної інформації та виділення навчальної та тестової вибірок. Виходячи з того, що ми маємо справу з часовими рядами і в реальності майбутній розподіл цін невідомий, стандартний підхід випадкового розділення даних на навчальну і тестову вибірки тут не підходить. Тому скористаємось

підходом, аналогічним методу «часового вікна». На вхід моделі будуть подані тільки дані в минулому відносно прогнозованої частини, причому ці дані будуть обмежені в минулому часі. По мірі просування прогнозованої частини в майбутнє – також буде просуватись і «часове вікно» минулого.

В якості вхідних ознак можна використовувати як самі дані ринкових цін, так і результати їх обробки – індикатори. В якості останніх можуть бути, наприклад, нормалізовані відносно ціни ковзні середні з різним періодом.

Запишемо формальну математичну модель задачі.

Нехай відомі простір об'єктів X , які складаються з окремих ознак x , а також простір відповідей-прогнозів Y . Тоді x_n^k – відома k -та ознака в n -му «часовому вікні», $k=1\dots K$, $n=1\dots N$. Для навчальних вибірок також відомі показники $y_n^k \in Y$ – результат торгівлі в прогнозованому періоді. Тоді для вирішення задачі треба знайти такий алгоритм-функцію $f(x)$ сімейства функцій F , яка б на функціоналі помилки $L(f, X)$ алгоритма f на вибірці X приймала би мінімальне значення

$$f(x) = \operatorname{argmin}_{f \in F} L(f, X).$$

В якості функціоналу похибки (оцінки рішення) можна використовувати одну з поширених метрик, наприклад, середньоквадратичну помилку (mean squared error, MSE) або корінь із середньоквадратичної помилки (root mean squared error, RMSE).

Зазвичай лінійна регресія не може виявити нелінійні закономірності. Тому один з підходів – попередня обробка ознак таким чином, щоб зробити можливість виявити такі зв'язки. Нелінійність в модель можна внести, наприклад, перемножуючи ознаки між собою з різними степенями, в результаті отримуючи складні поліноміальні ознаки.

Після навчання модель треба перевірити та оцінити на тестовій вибірці. Крім прямої оцінки відповідності прогнозу і фактичним даним, можна спробувати промодельовати продажі та оцінити, чи буде кращим результат і

наскільки. Для моделювання можна також використовувати різні моделі, наприклад, спрощену лінійну, яка була показана на рис. 2.5. Ця модель перемножується з прогнозними показниками, в результаті чого знаходиться кількість товару, яку належить продати.

По підсумкам звітного періоду, наприклад року, можна порівняти середні ціни продажів такої стратегії і середні ринкові ціни та підрахувати, наскільки вигідним буде економічний результат.

Задача, яка стоїть перед нами, нагадує задачу класифікації (продавати чи ні) з тією різницею, що треба визначити додатково сам рівень продажів. Також відмінність від класичної логістичної регресії полягає в тому, що може існувати певний мінімальний рівень продажів, який компанія вимушена підтримувати, щоб покрити хоча б постійні витрати (наприклад, адміністративні).

Відомо, що одна з найпоширеніших функцій для задач класифікації - це так звана сигмоїда [10], яка має наступну формулу:

$$\sigma = \frac{1}{1+e^{-x}}$$

Діапазон її аргументів $]-\infty; +\infty[$, а значень, відповідно $[0; 1]$. Спробуємо додати параметри в цю функцію, щоб мати можливість отримати такі значення, які нас цікавлять. Це такі параметри:

- x_0 - початкова вхідна змінна;
- Δx - діапазон вхідної змінної;
- v_{\min} і v_{\max} - мінімальний і максимальний об'єм торгівлі за квант часу (день);
- s_{\min} та s_{\max} - мінімум та максимум степенів (зсув по горизонталі);
- k_e - коефіцієнт множення основи e (нахил або швидкість зростання).

Візуально вплив нових параметрів показаний на рис.2.6. Фактично, за допомогою цих параметрів ми можемо встановити, яким об'ємом потрібно торгувати, та в яких межах.

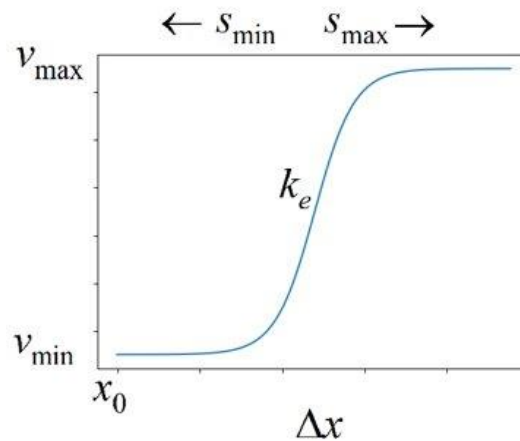


Рисунок 2.6 - Об'єм торгівлі

Враховуючи, що максимальний об'єм торгівлі залежить деяким чином від середнього плану торгівлі v_{av} та відповідного коефіцієнту збільшення k_v :

$$v_{\max} = k_v \times v_{av} + v_{\min}$$

можна вивести остаточну формулу визначення рекомендованого об'єму торгівлі:

$$v = v_{\min} + \frac{k_v \times v_{av}}{1 + (k_e \times e)^{-((S_{\max} - S_{\min}) \times (x - x_0) / (\Delta x + S_{\min}))}}$$

Маючи рекомендований об'єм та враховуючи деякі особливості (наприклад, ми мусимо піднімати мінімальний об'єм, якщо наближається кінець року, а у нас є велика кількість нереалізованої продукції), можна моделювати процес торгівлі. Її результат за рік ділиться на кількість продукції та встановлюється середня ціна, яку вже можна порівняти з фактичною середньою біржовою ціною. Ця різниця і є функція помилки, яку треба мінімізувати.

2.4. Вибір інструментарію та мови програмування

Насамперед для реалізації моделі необхідно вибрати мову програмування. Існує кілька мов, які використовуються під час роботи з даними. В результаті короткого дослідження вибір ліг на Python.

З одного боку, Python є мовою загального призначення і може бути успішно застосований для вирішення різних завдань [11]. З іншого боку, він, як і більшість проектів та бібліотек на ньому, є програмним забезпеченням з відкритим кодом, що робить його дуже популярним. Сама мова дуже проста і тому легка у вивченні. Тим більше, що для застосування в машинному навчанні не потрібно знати його тонкощів.

Якщо в традиційному програмуванні більшу частину часу програміст проводить у текстових редакторах чи інтегрованих середовищах розробки, то в обробці даних більшість коду пишеться в Jupyter Notebook.

Це простий та потужний інструмент для аналізу даних. Він дозволяє писати код Python, R та інших мовах, додавати текстові описи в Markdown, вбудовувати графіки та діаграми безпосередньо в інтерактивну веб-сторінку.

Плюс до всього Google випустив безкоштовний сервіс Google Colab, який надає хмарну версію Jupyter Notebook і дає можливість проводити обчислення на CPU і GPU.

Всі потрібні Python-бібліотеки вже встановлені, тому можна починати відразу і не потрібно встановлювати все локально. Часто це виявляється набагато зручнішим, оскільки немає необхідності встановлювати Jupyter на свій комп'ютер. Одним із базових пакетів для роботи з даними є бібліотека NumPy.

Основний функціонал бібліотеки NumPy полягає у підтримці багатовимірних масивів даних та швидких алгоритмів лінійної алгебри. Саме тому NumPy є ключовим компонентом інших бібліотек роботи з даними Scikit-learn, SciPy та Pandas.

Зазвичай NumPy використовують як допоміжну бібліотеку до виконання різних математичних операцій із структурами даних Pandas.

Аналіз та підготовка даних найчастіше займає більшу частину часу при вирішенні завдань машинного навчання (Machine Learning, ML). Дані можуть бути отримані в CSV, JSON, Excel або іншому структурованому (або неструктурованому) форматі і потрібно обробити їх для того, щоб використовувати в ML моделях.

Для цих цілей використовується бібліотека Pandas, яка служить в основному для отримання та підготовки даних. Це потужний інструмент, який дозволяє швидко аналізувати, модифікувати та готувати дані для подальшого використання в інших ML бібліотеках, таких як Scikit-learn, TensorFlow або PyTorch.

У Pandas можна завантажувати дані з різних джерел: SQL баз, CSV, Excel, JSON файлів та інших менш популярних форматів.

Коли дані завантажені у пам'ять, з ними можна виконувати безліч різних операцій для аналізу, трансформації, заповнення відсутніх значень та очищення набору даних. Pandas дозволяє виконувати безліч SQL-подібних операцій над наборами даних: об'єднання, угруповання, агрегування тощо. Також вона надає вбудований набір популярних статистичних функцій для базового аналізу. Jupyter Notebook також підтримує Pandas та реалізує гарну візуалізацію його структур даних.

Для побудови графіків та візуалізації даних використовуються пакети Matplotlib та Seaborn.

Matplotlib – це стандартний інструмент у наборі дата-інженера. Він дозволяє створювати різноманітні графіки та діаграми для візуалізації отриманих результатів. Графіки, створені Matplotlib, легко інтегруються в Jupyter Notebook. Це дозволяє візуалізувати дані та результати, отримані при обробці моделей.

Для цієї бібліотеки створено багато додаткових пакетів. Один із найпопулярніших — Seaborn.

Seaborn — це ефективна бібліотека Python для створення різних візуалізацій у Data Science, що використовує базову бібліотеку Matplotlib.

Однією з основних її особливостей є візуалізація даних, яка дозволяє побачити кореляцію там, де вона була не очевидна. Це дозволяє краще розуміти дані. Завдяки темам, що кастомізуються, і високорівневим інтерфейсам можна отримувати настільки якісні та репрезентативні візуалізації, що їх надалі можна показувати клієнтам.

Крім того, знадобиться бібліотека для машинного навчання. Серед інших було обрано Scikit-learn – одну з найпопулярніших бібліотек машинного навчання на сьогодні. Вона підтримує більшість алгоритмів навчання як з учителем, так і без: лінійна та логістична регресія, метод опорних векторів (SVM), Naive Bayes класифікатор, градієнтний бустинг, кластеризація, KNN, k-середні та багато інших.

Scikit-learn містить безліч корисних утиліт для підготовки даних та аналізу результатів. Ця бібліотека в основному призначена для класичних алгоритмів машинного навчання, тому її функціонал для нейронних мереж дуже обмежений, а для завдань глибокого навчання вона не може бути використана зовсім.

На додаток до дуже якісної документації Scikit-learn містить розділ, в якому показано, як працювати з бібліотекою, а також даються базові знання з машинного навчання.

В якості локального серверу було обрано збірку XAMPP – це сервер з повним набором модулів для повноцінної роботи, зокрема СУБД MySQL, сервер Apache і панель адміністрування phpMyAdmin. Переваги XAMPP:

- кросплатформеність, програма може встановлюватися на Windows, Linux та MacOS;
- додатково до PHP у збірці є Perl;
- безкоштовність та відкритий код;
- легка установка;
- доброзичливий інтерфейс.

Для задач побудови графічного інтерфейсу обраний популярну бібліотеку Tkinter - це пакет для Python, призначений для роботи з бібліотекою Tk.

Бібліотека Tk містить компоненти графічного інтерфейсу користувача (graphical user interface - GUI). Ця бібліотека написана на мові програмування Tcl. Під графічним інтерфейсом користувача (GUI) маються на увазі всі ті вікна, кнопки, текстові поля для введення, скролери, списки, радіокнопки, прапорці та ін., які користувач бачить на екрані, відкриваючи ту чи іншу програму. Через них користувач взаємодіє з програмою і керує нею. Всі ці елементи інтерфейсу називаються віджетами (widgets).

Як основний математичний метод вирішення завдання було обрано градієнтний спуск. Градієнтний спуск - метод знаходження мінімального значення функції втрат (існує безліч видів цієї функції). Мінімізація означає пошук найглибшої западини в цій функції, яка використовується для контролю помилки в прогнозах моделі машинного навчання. Пошук мінімуму означає отримання найменшої можливої помилки або підвищення точності моделі. Ми збільшуємо точність, перебираючи набір навчальних даних при налаштуванні параметрів нашої моделі (ваги та зміщення).

Суть алгоритму – покроковий процес отримання найменшого значення помилки. Це можна представити як спуск у западину у спробі знайти дно “ущелини” (найнижче значення помилки).

РОЗДІЛ 3

РОЗРОБКА ІНФОРМАЦІЙНОЇ СИСТЕМИ

3.1. Розробка структурної схеми ІС

Структура інформаційної системи мусить бути підкорена цілям, задачам і функціям, які виконує ця система. Була обрана клієнт-серверна архітектура, яка, взагалі, є збірним поняттям, що складається з двох компонентів, які взаємодоповнюють один одного: сервера і, власне, клієнта.

Клієнт – локальний комп'ютер на стороні віртуального користувача, який виконує надсилання запиту до сервера для надання даних або виконання певної групи системних дій.

Сервер – потужний комп'ютер або спеціальне системне обладнання, яке призначається для вирішення певного кола завдань процесу виконання програмних кодів. Він виконує роботи сервісного обслуговування за запитами клієнтів, надає користувачам доступ до певних системних ресурсів, зберігає дані або БД.

Особливості такої моделі полягають у тому, що користувач надсилає певний запит на сервер, де той системно обробляється та кінцевий результат надсилається клієнту. У можливості сервера входить одночасне обслуговування кількох клієнтів.

Якщо одночасно надходить більше одного запиту, такі запити встановлюються у певну чергу і сервером виконуються по черзі. Іноді запити можуть мати свої пріоритети. Частина запитів із вищими пріоритетами постійно виконуватимуться в першочерговому порядку.

Параметри, які можуть бути реалізовані на стороні сервера:

- зберігання, захист та доступ до даних;
- робота з клієнтськими запитами;
- процес надсилання відповіді клієнту.

Параметри, які можуть бути реалізовані на стороні клієнта:

- майданчик з надання користувальницького графічного інтерфейсу;
- формулювання запиту до сервера та його подальше відправлення;
- отримання підсумків запиту та надсилання додаткової групи команд (запити на додавання, оновлення інформації, видалення групи даних).

Архітектура системи клієнт-сервер формує принципи віртуального спілкування між локальними комп'ютерами, проте правила і принципи взаємодії перебувають усередині протоколу.

Крім перерахованого, треба пам'ятати, що в системі мусять бути додаткові підсистеми. Вони будуть відповідати за певні напрямки роботи, які прямо не стосуються роботи з БД, а виконують допоміжні функції. Це підсистема зв'язку з іншими інтернет-ресурсами, підсистема інтелектуальних розрахунків та підсистема внутрішнього зв'язку в межах компанії.

Спрощена структурна схема інформаційної системи представлена на рис.3.1.

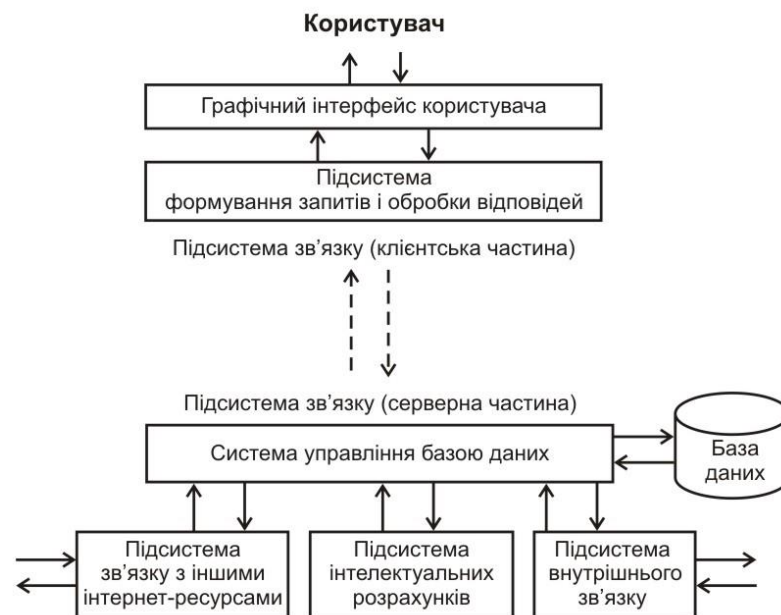


Рисунок 3.1 – Спрощена структура інформаційної системи

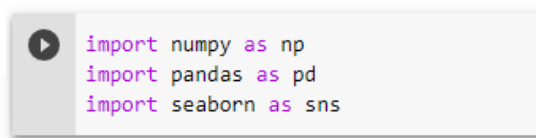
3.2. Розвідувальний аналіз для інтелектуальної моделі

Для якісного моделювання і роботи з даними треба визначитись з джерелом таких даних. Якщо брати дані по біржовим цінам - вільних джерел в

Інтернеті дуже багато. Автором було обрано один з популярних сайтів `investing.com`, а в якості безпосередніх даних - ціни на пшеницю, починаючи з січня 2001 року і закінчуючи груднем 2022 року [12].

Дані з сайту завантажуються в форматі `csv`, однак для безпосередньої передачі в програму обробки потребують невеличких правок (в основному - видалення зайвих символів), що і було зроблено. Для прискорення роботи файл відразу був записаний на `Google Drive`. Після підключення через відповідну бібліотеку він стане видимий для `Python` блокноту для завантаження вже туди.

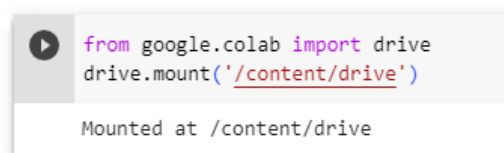
Робота в `Google Colab` починається зі створення нового файлу `Python` блокноту і надання йому імені. Перший крок розробки - імпорт основних бібліотек для безпосередньої роботи з даними (рис.3.2).



```
import numpy as np
import pandas as pd
import seaborn as sns
```

Рисунок 3.2 - Завантаження основних бібліотек

Далі - підключаємо `Google Drive` та даємо відповідну назву папці, щоб можна було бачити її з `Python` блокноту (рис.3.3).



```
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

Рисунок 3.3 - Підключення `Google Drive`

Зчитуємо файл `csv` і перетворюємо відразу його в формат `DataFrame` пакету `Pandas` (рис.3.4).

```

path = "/content/drive/MyDrive/Colab Notebooks/Ф'ючерс на пшоницю США.csv"
df = pd.read_csv(path, sep=';')
data = df[['date', 'price', 'change']][::-1]
data

```

	date	price	change
5754	01.01.2001	273.00	-2.33
5753	02.01.2001	276.25	1.19
5752	03.01.2001	283.25	2.53
5751	04.01.2001	284.25	0.35
5750	05.01.2001	283.00	-0.44
...
4	23.12.2022	776.00	1.80
3	27.12.2022	774.50	-0.19
2	28.12.2022	785.50	1.42
1	29.12.2022	774.00	-1.46
0	30.12.2022	792.00	2.33

5755 rows x 3 columns

Рисунок 3.4 - Вигляд даних

Бачимо, що у нас є 5755 записів, а зі стовпців нам потрібно тільки дати, ціни та зміни цін у відсотках. Також бажано переіндексувати дані, тому що на біржах зазвичай дати йдуть у зворотному напрямку (остання ціна - зверху).

Подивимось на деякі статистичні показники - мінімум, максимум і середнє значення для відсотків змін (рис. 3.5).

```

data.change.min(), data.change.max(), data.change.mean()

```

(-14.06, 22.43, 0.0422241529105126)

Рисунок 3.5 - Статистичні показники

Спробуємо візуалізувати дані, побудувавши їх розподіл (рис.3.6).

Бачимо досить рівну і майже симетричну криву, що говорить про можливі труднощі з пошуком ринкових неефективностей. У зв'язку з цим були прийняті наступні рішення: розділити дані на періоди (роки) і оцінювати кожен період окремо, а також додати нові ознаки - підрахунок середньої ціни з початку року та відсоток різниці між цією середньою та наявною ціною.

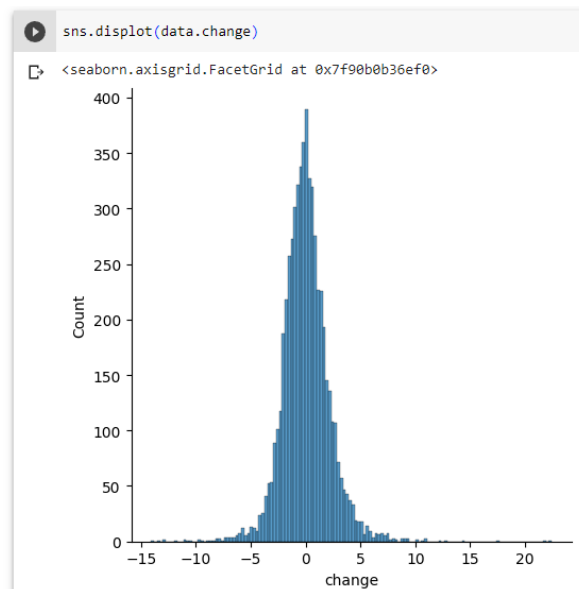


Рисунок 3.6 - Розподіл змін

Результат додавання наведено на рис.3.7.

```
data.reset_index(drop=True, inplace=True)
```

	date	price	change	av_st_year	av_year	differ
0	01.01.2001	273.00	-2.33	273.000000	272.169608	0.000000
1	02.01.2001	276.25	1.19	274.625000	272.169608	0.591716
2	03.01.2001	283.25	2.53	277.500000	272.169608	2.072072
3	04.01.2001	284.25	0.35	279.187500	272.169608	1.813298
4	05.01.2001	283.00	-0.44	279.950000	272.169608	1.089480
...
5750	23.12.2022	776.00	1.80	906.992103	905.031289	-14.442474
5751	27.12.2022	774.50	-0.19	906.468419	905.031289	-14.558524
5752	28.12.2022	785.50	1.42	905.992165	905.031289	-13.299471
5753	29.12.2022	774.00	-1.46	905.474549	905.031289	-14.519961
5754	30.12.2022	792.00	2.33	905.031289	905.031289	-12.489213

5755 rows x 6 columns

Рисунок 3.7 - Додавання нових ознак

Подивимось на розподіл різниці (рис.3.8) та її показники (рис.3.9).

Розподіл відразу показує асиметрію кривої, а середня різниця хоча і близька до нульового значення, але не настільки, як денні зміни. Все це дає основу для спроби побудови моделі.

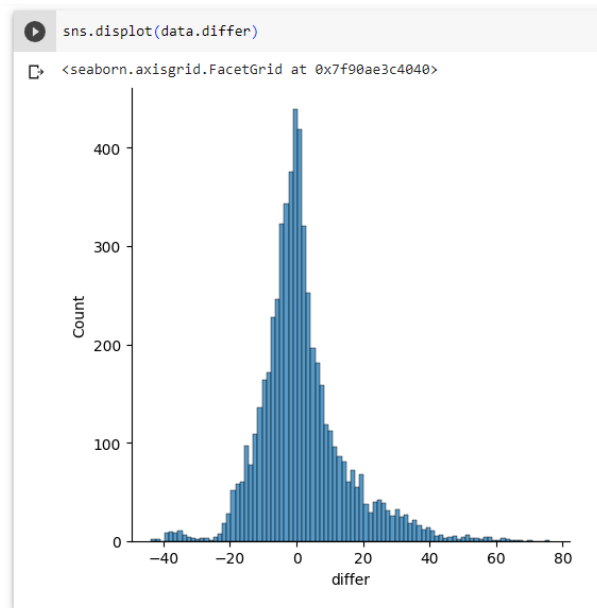


Рисунок 3.8 - Розподіл показника різниці

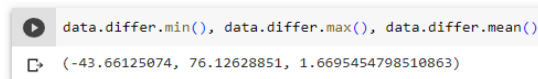


Рисунок 3.9 - Статистичні показники різниці

Подивимось перед цим, як виглядають графіки даних за різні роки (рис.3.10 - 3.12).

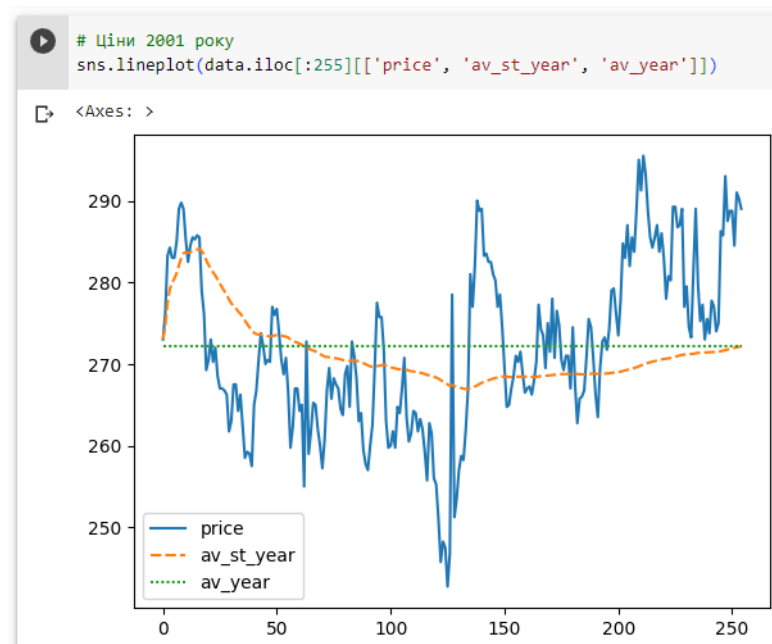


Рисунок 3.10 - Ціна на пшеницю впродовж 2001 року

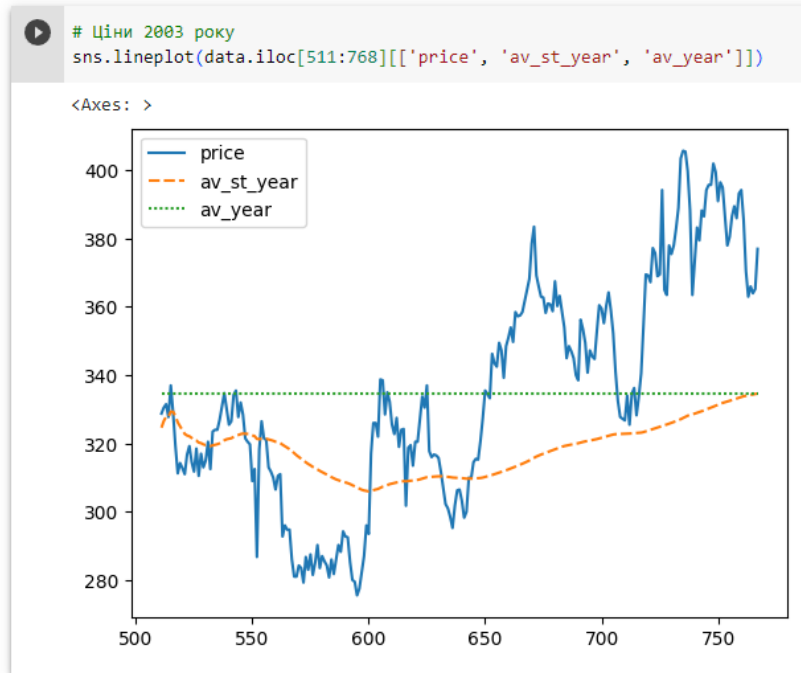


Рисунок 3.11 - Ціна на пшеницю впродовж 2003 року

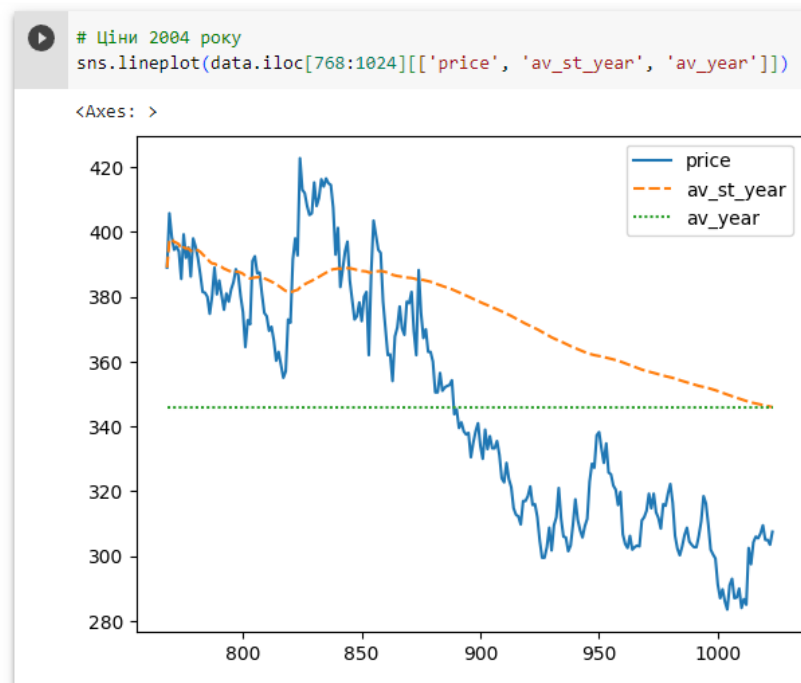


Рисунок 3.12 - Ціна на пшеницю впродовж 2004 року

Були обрані несхожі один на одного роки. В 2001 році ринок був відносно спокійним, а ціни коливались приблизно навколо середньої. В 2003-му році ми побачили невеличке падіння, а потім - тільки зростаючий ринок. В 2004-м все було дзеркально навпаки. Тобто, остаточне значення середньої на рік ми зможемо побачити тільки, коли сам рік закінчиться. Але навіть використання середньої з початку року вже може дати непогані орієнтири, які ціни є вигідними, а які не зовсім, для реалізації продукції. Перейдемо до визначення параметрів, які необхідні для остаточного моделювання.

3.3. Розробка моделі машинного навчання та тестування

В якості навчальної вибірки були обрані дані за 2001-2021 роки. На тестову вибірку вирішено відвести дані за 2022 рік.

Для пошуку рішення треба визначити межі області пошуку, або граничні значення, які її обмежують. В даному випадку були обрані такі значення:

- x_0 - від -10 до 10, відповідно, Δx від 0 до 20;
- v_{\min} і v_{\max} - від 1 до 21;
- s_{\min} та s_{\max} - від -10 до 10;
- k_e - від 1 до 3.

Початковий варіант рішення обирається випадково. Візуалізація одного з процесів рішення наведена на рис. 3.13.

В результаті пошуку були знайдені такі оптимальні параметри:

$$x_0 = 0,02; \Delta x = 10,09; v_{\min} = 1; v_{\max} = 21; s_{\min} = -10; s_{\max} = 10; k_e = 3.$$

Сумарно результат стратегії з оптимізованими параметрами на навчальній вибірці виявився краще на 34.16%, ніж стратегія рівномірного продажу.

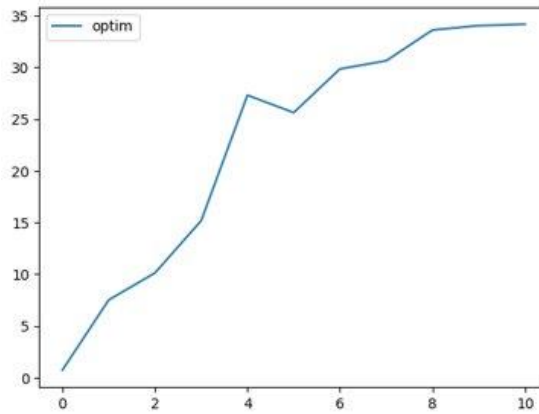


Рисунок 3.13 - Пошук рішення

Подивимось на результат роботи стратегії на тестовому періоді, який вона “не бачила”, а саме на 2022 році (рис.3.14).

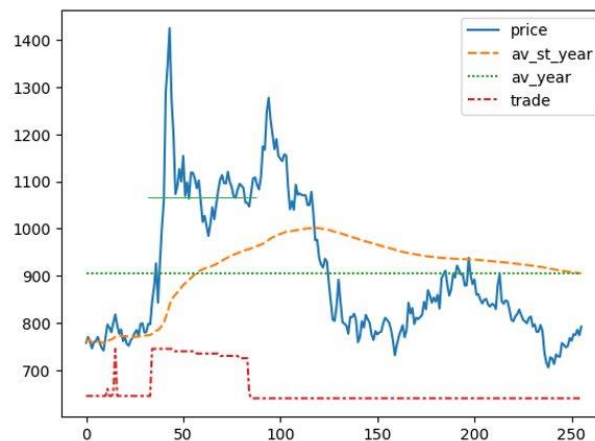


Рисунок 3.14 - Ціни та результат тестування роботи стратегії в 2022 році

Середня ціна продажів склала 1062 або на 17,4% більше, ніж середня річна ціна (905). На графіку додатково вказана активність системи (лінія trade) - ті періоди, коли система вела торгівлю.

3.4. Розробка елементів графічного інтерфейсу користувача

Програми графічного інтерфейсу користувача [13] є подієво-орієнтованими. Тобто певні частини програмного коду починають виконуватися лише тоді, коли відбуваються певні події.

Подієво-орієнтоване програмування базується на об'єктно-орієнтованому та структурному програмуванні. Навіть якщо ви не створюєте власні класи та об'єкти, вони все одно використовуються. Усі віджети є об'єктами, створеними вбудованими класами.

При цьому події можуть бути різними. Хтось клацнув мишею або натиснув Enter, почав вводити текст, поставив галочку на перемикач, прокрутив сторінку вниз тощо. Коли відбувається щось подібне і створюється відповідний обробник, частина програми запускається та видає результат.

Побудовані на основі єдиного інтерфейсу, форми додатків забезпечують покращений інтерфейс користувача для оптимізації роботи агента та підтримки контексту під час роботи з пов'язаними записами.

Редактори форм зазвичай використовуються для редагування форми та додавання чи зміни її елементів. Редактор форм дозволяє змінювати форми для всіх ваших програм на основі єдиного інтерфейсу. Форми використовуються для введення даних, створення запитів і виведення інформації після обробки. Джерелом даних форми може бути запис таблиці або запит.

Форми використовуються для надання наступних можливостей:

- введення і перегляду інформації в БД;
- зміни даних;
- друку;
- створення повідомлень;
- запуску додаткових програм.

Приклад форми наведений на рисунку 3.15. В даному випадку користувач може самостійно сформулювати запит до БД на вибірку необхідних даних і спробувати навчити на них нову модель. На рисунку якраз показаний приклад навчання тієї інтелектуальної моделі, яка наведена в даній роботі.

Крім того, форми використовуються для створення додаткових різноманітних запитів до БД.

The image shows a software window titled "Навчання моделей" (Model Training). It contains several input fields and buttons. The fields are arranged in a grid-like fashion. At the top, there are two dropdown menus: "Ринок" (Market) set to "США" and "Продукт" (Product) set to "Пшениця". Below these, there are two sections for time periods. The first section, "Період навчання" (Training Period), has "Початок" (Start) set to "01-01-2001" and "Закінчення" (End) set to "31-12-2021". The second section, "Період тестування" (Testing Period), has "Початок" set to "01-01-2022" and "Закінчення" set to "31-12-2022". At the bottom of the form, there are two more dropdown menus: "Алгоритми" (Algorithms) and "Метрики" (Metrics), both set to "Всі, що є в базі" (All in the database). At the very bottom, there are two buttons: "Start!" and "Cancel".

Рисунок 3.15 – Форма «Навчання моделей»

Запит є основним інструментом для отримання інформації з однієї або кількох таблиць БД. Ми можемо самостійно вносити зміни в базу даних за допомогою запитів. Запити служать джерелами даних для форм, звітів і сторінок доступу до даних. Результатом є нова електронна таблиця, яку можна переглядати, аналізувати, зберігати чи не зберігати.

Запити дозволяють вирішувати багато задач, не вдаючись до програмування. Наприклад, щоб відобразити дані в зведеній формі, виконати обчислення полів бази даних, згрупувати записи, використовувати різні статистичні функції для пошуку зведених значень для полів та багато іншого.

Формування запитів кінцевому користувачу зручно здійснювати також через форми.

Користувачеві достатньо ввести необхідні параметри, і після натискання кнопки «Переглянути звіт» програма замінить параметри на підготовлений шаблон запиту та відправить його в базу даних. Приклад такої форми запиту наведено на рисунку 3.16.

Рисунок 3.16 – Форма для запиту «Моделі для продукту»

Приклад звіту за цим запитом наведено на рисунку 3.17. Він показує, чи є в базі даних моделі, які були навчені користувачами, а також їх загальні параметри. Звіт формується як окрема html-сторінка, яку можна подивитись в браузері або в окремому вікні.

Звіт по знайденим моделям

Ринок: США

Продукт: Пшениця

Знайдено навчених моделей в базі даних: 2 шт.

Користувач	Період навчання		Період тестування	
	Початок	Закінчення	Початок	Закінчення
Коваленко С.І.	01.01.2001	31.12.2021	01.01.2022	31.12.2022
Степанов В.В.	01.01.2001	31.12.2019	01.01.2020	31.12.2021

Рисунок 3.17 – Звіт по запиті «Моделі для продукту»

РОЗДІЛ 4

ОХОРОНА ПРАЦІ

4.1. Основні положення охорони праці

Охорона праці - це система забезпечення безпеки життя та здоров'я працівників у процесі трудової діяльності, що включає правові, соціально-економічні, організаційні, технічні, психофізіологічні, санітарно-гігієнічні, лікувально-профілактичні, реабілітаційні та інші заходи та засоби. [14]

Безпека виробництва - це оптимальний баланс стану технологічного процесу, обладнання, робочих місць та поведінки людини, що обмежує вплив на працівника небезпечних та шкідливих виробничих факторів.

Робоче місце - зона, оснащена необхідними технічними засобами, в якій здійснюється трудова діяльність виконавця або групи виконавців, які спільно виконують одну роботу або операцію.

Небезпечний виробничий фактор - такий фактор, вплив якого на працюючого в певних умовах призводить до травми або іншого раптового різкого погіршення здоров'я або смерті.

Шкідливий виробничий фактор - такий фактор, вплив якого на працюючого в певних умовах може призвести до зниження працездатності та (або) захворювання (залежно від рівня та тривалості впливу шкідливий виробничий фактор може стати небезпечним).

Безпечні умови праці - умови праці, за яких виключено вплив на працівників шкідливих та (або) небезпечних виробничих факторів або рівні їх впливу не перевищують встановлених нормативів.

Засіб захисту на виробництві - засіб, призначений для запобігання ризику або зменшення його на працівника до допустимого рівня.

4.2. Шкідливі та небезпечні виробничі фактори

Розглянемо характеристику шкідливих та небезпечних виробничих факторів при технічній експлуатації електроустаткування.

До основних шкідливих та небезпечних виробничих факторів при технічній експлуатації електрообладнання належать:

- підвищене значення напруги в електричному ланцюзі, замикання якого може статися через тіло людини при наближенні на відстань менш допустимого до неізольованих струмопровідних частин та елементів обладнання, що знаходиться під напругою, а також при переміщенні та роботі в зонах розтікання струму на землю, впливу електричного поля та наведеної напруги;
- підвищена напруженість електричного та магнітного полів;
- електрична дуга, що виникла при перемиканнях в електроустановках або в аварійних ситуаціях;
- Недостатня освітленість робочої зони;
- Підвищена або знижена температура повітря робочої зони, а також поверхонь обладнання, матеріалів;
- підвищена або знижена вологість повітря, а також сильний вітер під час роботи поза приміщенням;
- рухомі машини та механізми; рухомі частини виробничого обладнання; вироби, що пересуваються, заготовлі, матеріали;
- конструкції, що руйнуються, та елементи обладнання в процесі виконання роботи та в аварійних ситуаціях;
- гострі кромки, задирки та шорсткості на поверхні заготовок, інструментів, обладнання;
- підвищена запиленість та загазованість повітря робочої зони.

4.3. Вимоги до персоналу при ремонті та обслуговуванні електрообладнання

До роботи електромонтером з ремонту та обслуговування електрообладнання (далі — електромонтер) допускаються особи не молодші 18 років, які мають професійну підготовку, пройшли медичний огляд, визнані придатними до роботи в діючих електроустановках та мають відповідну групу з електробезпеки.

Після навчання проводиться перевірка теоретичних знань електромонтера, результати якої оформлюються протоколом із присвоєнням відповідної групи з електробезпеки.

Електромонтеру видається посвідчення встановленої форми, до якого вносяться результати перевірки знань, і навіть записи про право виконання спеціальних робіт.

До спеціальних робіт відносяться:

- верхолазні роботи;
- роботи під напругою на струмоведучих частинах: чищення, обмив та заміна ізоляторів, ремонт проводів, мастило тросів;
- випробування електроустаткування з подачею підвищеної напруги від стороннього джерела;
- обслуговування щіткового апарату на працюючому генераторі та електродвигуні;
- роботи всередині баків силових трансформаторів (дугогасних реакторів);
- обслуговування акумуляторних батарей та зарядних пристроїв;
- роботи з імпульсним вимірювачем ліній електропередач;
- роботи з електровимірювальними кліщами та електроізолюючими штангами для проведення вимірювань.

Допуск електромонтера до самостійної роботи оформляється наказом чи розпорядженням.

Ті, хто працює в електроустановках, повинні бути забезпечені електрозахисними засобами відповідно до характеру роботи та зобов'язані правильно застосовувати їх під час роботи.

Працюючі, що помітили несправність електроустановки або електрозахисних засобів, зобов'язані негайно повідомити про це безпосереднього керівника, а за його відсутності — керівника, що стоїть вище.

У тих випадках, коли несправність в електроустановці становить явну небезпеку для працюючих або об'єктів, працівники, які її виявили, зобов'язані вжити заходів для унеможливлення наближення до електроустановки сторонніх осіб, потім повідомити про це безпосереднього керівника, а за його відсутності — вищого керівника.

Ті, хто працює в електроустановках, повинні бути навчені прийомам звільнення потерпілого від дії електричного струму та надання першої допомоги постраждалим при нещасних випадках.

ВИСНОВКИ

В ході виконання кваліфікаційної роботи була розроблена інтелектуальна інформаційна система прогнозування попиту на сільськогосподарську продукцію. Виконані такі задачі:

1. Проведений аналіз предметної області та визначені основні властивості системи, яка створюється.

2. Обрані відповідні для даної роботи математичний апарат, інструментарій програмування та метод рішення – градієнтний спуск.

3. Розроблена структурна схема інформаційної системи, реалізована інтелектуальна модель, розроблені елементи графічного інтерфейсу користувача.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Інформаційна система. URL: https://uk.wikipedia.org/wiki/Інформаційна_система (дата звернення 10.04.23).
2. «Знайомство з ІТ» / Матеріали Науково-технічної конференції студентства Навчально-наукового інституту Телекомунікацій та Інформатизації. Київ: ДУТ, 2016. 112 с.
3. Класифікація інформаційних систем. URL: <https://stboinf.wordpress.com/2011/05/28/класифікація-інформаційних-систем/> (дата звернення 10.04.23).
4. Матеріали IV Всеукраїнської науково-практичної інтернет-конференції «Управління ресурсним забезпеченням господарської діяльності підприємств реального сектору економіки». URL: <http://feb.tsatu.edu.ua/wp-content/uploads/2019/12/Materiali-IV-Vseukrayin.-internet-konfer.pdf> (дата звернення 10.04.23).
5. Нові інформаційні технології в сільському господарстві. URL: http://8ref.com/19/referat_196944.html (дата звернення 10.04.23).
6. Використання сучасних інформаційних технологій в сільському господарстві. URL: <https://ukrbukva.net/page,5,82592-Ispol-zovanie-sovremennyh-informacionnyh-tehnologiiy-v-sel-skom-hozyaiystve.html> (дата звернення 10.04.23).
7. Напрями досліджень в аграрній науці: стан та перспективи / Збірник наукових праць Всеукраїнської наукової конференції аспірантів, магістрів та студентів, 23 березня 2017 р. Вінниця: ВНАУ, 2017. 242 с.
8. Маркетинговий аналіз і дослідження ринку: методи, види і етапи. URL: <http://goldenkey.com.ua/index/2019/12/uk/vklucat-marketingovij-analiz-i-doslidzenna-rinku-gajd-metodi-vidi-i-etapi-powerbrandingru.php> (дата звернення 10.04.23).
9. Аналіз часових рядів. URL: https://uk.wikipedia.org/wiki/Аналіз_часових_рядів (дата звернення 10.04.23).

10. Сигмоїда. URL: <https://uk.wikipedia.org/wiki/Сигмоїда> (дата звернення 10.04.23).

11. З чого почати роботу з ML і DL. Огляд кращих бібліотек. URL: <https://www.seoblog.org.ua/5210/> (дата звернення 10.04.23).

12. US Wheat Futures Historical Data. URL: <https://www.investing.com/commodities/us-wheat-historical-data> (дата звернення 10.04.23).

13. Графічний інтерфейс користувача. URL: https://uk.wikipedia.org/wiki/Графічний_інтерфейс_користувача (дата звернення 10.04.23).

14. Нормативні акти про охорону праці. URL: <https://dnaop.com/398/2428/> (дата звернення 10.04.23).