

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Системотехніки
(повна назва)

АТЕСТАЦІЙНА РОБОТА Пояснювальна записка

другий (магістерський)
(освітньо-кваліфікаційний рівень)

ГЮІК.506120.012 ПЗ
(позначення документа)

Розробка методів CRM на базі генерації асоціативних правил та
їх використання в інформаційній системі продажу квітів
(тема роботи)

Виконав: студент II курсу, групи ІТІМ-19-1
спеціальності 122 - «Комп'ютерні науки та
інформаційні технології»
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні технології
проекткування
(повна назва освітньої програми)

Шубіна Н.А.
(прізвище, ініціали)

Керівник проф. Ситніков Д.Е.
(прізвище, ініціали)

Допускається до захисту

Зав. кафедри СТ _____
(підпис)

проф. Гребеннік І.В.
(прізвище, ініціали)

2020 р.

Харківський національний університет радіоелектроніки

(назва вищого навчального закладу)

Факультет Комп'ютерних наук

(повна назва)

Кафедра Системотехніки

(повна назва)

Рівень вищої освіти другий (магістерський)

Спеціальність 122 «Комп'ютерні науки та інформаційні технології»

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні технології проектування

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри СТ _____

(підпис)

" _____ " _____ 2019
р.

ЗАВДАННЯ

НА АТЕСТАЦІЙНУ РОБОТУ

студентові Шубіній Наталії Андріївні

(прізвище, ім'я, по батькові)

1. Тема роботи: Розробка методів CRM на базі генерації асоціативних правил та їх використання в інформаційній системі продажу квітів

затверджена наказом по університету від « 02 » 11 2020 р. № 1517 Ст.

2. Термін подання студентом роботи 12 грудня 2020 р.

3. Вихідні дані до роботи Розробити методи для CRM-системи на основі генерації асоціативних правил. Система повинна являти собою проект клієнтського web-додатка з інтерфейсом доступу до бази даних. Перелік використовуваних програмних засобів: ОС Windows 10, PyCharm, IntelliJ Idea Community 2018 Технічне забезпечення: IBM-сумісний ПК з МП Core 2 Duo та вище.

4. Зміст пояснювальної записки (перелік питань, що потрібно розробити) 4.1 Вступ. 4.2 Визначення сфери застосування CRM-системи в електронній комерції. 4.3 Аналіз аналогічних програмних систем. 4.4 Постановка задачі. 4.5 Дослідження методів та алгоритмів вирішення задачі. 4.6 Впровадження алгоритму асоціативних правил Apriori в інформаційну систему 4.7 Розробка методу CRM з урахуванням предметної області. 4.8 Проектування системи. 4.9 Розробка інформаційного та програмного забезпечення. 4.10 Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, плакатів). 5.1 Контекстна діаграма IDEF0 інформаційної системи (1 аркуш формату А4). 5.2 IDEF0-діаграма декомпозиції (1 аркуш формату А4). 5.3 Діаграма варіантів використання. (1 аркуш формату А4). 5.5 Модель бази даних системи (1 аркуш формату А4). 5.6 Алгоритм Apriori (1 аркуш формату А4).

6. Консультанти з роботи із зазначенням розділів роботи, що їх стосуються

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		(підпис)	(дата)
Аналіз предметної області	проф. Ситніков Д.Е.		
Дослідження методів та алгоритмів вирішення задачі	проф. Ситніков Д.Е.		

КАЛЕНДАРНИЙ ПЛАН

Пор. №	Назва етапів атестаційної роботи	Термін виконання етапів роботи	Примітка
1.	Отримання завдання атестаційної роботи	02.11.2020	
2.	Аналіз завдання, літератури та аналогів з теми атестаційної роботи	02.11-12.11.2020	
3.	Формування постановки задачі	13.11 - 15.11.2020	
4.	Дослідження методів та алгоритмів вирішення задачі	16.04 - 20.11.2020	
5.	Розробка методу CRM системи на основі асоціативних правил	21.11 - 25.11.2020	
6.	Проектування та розробка інформаційної системи	26.11 - 03.12.2020	
7.	Аналіз результатів, отриманих за допомогою програмного засобу	04.12 - 05.12.2020	
8.	Оформлення пояснювальної записки та програмної документації	06.12 - 09.12.2020	
9.	Оформлення графічної частини та презентаційних матеріалів комп'ютерного захисту	10.12 - 12.06.2020	
10.	Представлення на рецензування	12.12.2020	
11.	Подання атестаційної роботи до ЕК	13.12.2020	

Дата видачі завдання 02 листопада 2020 р.

Студент



(підпис)

Шубіна Н.А.

Керівник роботи



(підпис)

проф. Ситніков Д.Е.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка до атестаційної роботи містить 80 аркушів, 23 рисунків, 1 таблицю, 3 додатків, 19 джерел. Графічна частина атестаційної роботи містить 6 плакатів.

ЕЛЕКТРОННА КОМЕРЦІЯ, CRM-СИСТЕМА, CRM-МЕТОДИ, АСОЦІАТИВНІ ПРАВИЛА, APRIORI БАЗА ДАНИХ, СИСТЕМА УПРАВЛІННЯ БАЗОЮ ДАНИХ, WEB-СЕРВІС, РЕКОМЕНДАЦІЙНІ СИСТЕМИ, РЕКОМЕНДАЦІЇ

Об'єктом дослідження атестаційної роботи є методи CRM-системи побудовані основі генерації асоціативних правил та їх використання в інформаційній системі з продажу квітів.

Предметом дослідження є алгоритм пошуку асоціативних правил, та методи його застосування для створення рекомендацій, котрі будуть пропонуватися кінцевим користувачам системи.

Метою роботи є дослідження та аналіз існуючих методів побудови рекомендацій (зокрема основаних на генерації асоціативних правил) на предмет можливості їх застосування для розробки методу CRM системи.

Методи дослідження – системний підхід, методи структурного аналізу і моделювання

Результати роботи – розроблені методи CRM-системи на основі генерації асоціативних правил, а також веб-сервіс продажу квітів з вбудованим до нього рекомендаційного методу.

ABSTRACT

Explanatory note contains 80 sheets, 25 figures, 1 table, 3 annexes, 19 sources; graphic material – 6 p.

ELECTRONIC COMMERCE, CRM-SYSTEM, CRM-METHODS, ASSOCIATED RULES, APRIORI, DATABASE, DATA-BASED MANAGEMENT SYSTEM, WEB-SERVICE, RECOMENDER SYSTEM, RECOMMENDATIONS

The object of the study are the methods of CRM-system based on the generation of associative rules and their use in the information system for the sale of flowers.

The subject of the study is the algorithm for finding associative rules, and methods of its application to create recommendations that will be offered to end users of the system.

The purchase of the attestation – is to study and analyze the existing methods of constructing recommendations (in particular based on the generation of associative rules) for the possibility of their application to develop a method of CRM system.

Research methods - a systematic approach, methods of structural analysis and modeling

The results of the attestation work - developed methods of CRM-system based on the generation of associative rules, as well as a web service for selling flowers with a built-in recommendation method.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	8
ВСТУП.....	9
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	11
1.1 Визначення сфери застосування CRM-системи в електронній комерції.....	13
1.2 Актуальність роботи	16
1.3 Аналіз аналогічних програмних систем	17
1.4 Постановка задачі	21
2 ДОСЛІДЖЕННЯ МЕТОДІВ ТА АЛГОРИТМІВ РІШЕННЯ ЗАДАЧІ.....	27
2.1 Фільтрація на основі контенту (content based)	27
2.2 Колаборативні методи фільтрації;	29
2.3 Гібридні методи	32
2.4 Рекомендаційні механізми на основі асоціативних правил.....	33
2.5 Впровадження алгоритму асоціативних правил Аргіогі в інформаційну систему	36
2.6 Розробка методу CRM з урахуванням предметної області.....	40
3 ПРОЕКТУВАННЯ СИСТЕМИ	44
3.1 Розробка системних вимог до CRM-системи.....	44
3.2 Визначення функціональних вимог	46
3.3 Розробка моделі потоків даних компонентів CRM-системи	55
3.4 UML-проекткування	63
4 РОЗРОБКА ІНФОРМАЦІЙНОГО ТА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ.....	68
4.1 Обґрунтування вибору СУБД.....	68
4.2 Проектування БД	69
4.3 Вибір мови програмування.....	76
ВИСНОВКИ	78
ПЕРЕЛІК ПОСИЛАНЬ	79
ДОДАТОК А Графічні матеріали.....	81
ДОДАТОК Б Текст програми.....	86
ДОДАТОК В Відомість атестаційної роботи	91

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

CRM – customer relationship management;

ІС – інформаційна система;

АІС – автоматизована інформаційна система;

БД – база даних;

СУБД – система управління базами даних;

ПЗ – програмне забезпечення;

IS – Information system;

SQL – Structured Query Language;

URL – Uniform Resource Locator;

UML – Unified Modeling Language;

РС – рекомендаційна система;

Collaborative filtering – колаборативна або спільна фільтрація;

Content-based filtering – фільтрація на основі вмісту або контентна фільтрація;

Associated rules – асоціативні правила

ВСТУП

На сьогодні, відбувається значне поширення нових інформаційних технологій в різні сфери діяльності людини, а також їх постійні зміни під впливом різноманітних аспектів повсякденного життя. Все це раніше було неможливим для людини.

Процес збору та збереження даних став необхідною частиною фактично всіх областей людської діяльності. Бізнес, торгівля, електронна комерція, навчання, банківська справа, медицина – це ті сфери, що накопичують велику кількість даних в процесі своєї діяльності. До них можна віднести різноманітні дані про клієнтів/користувачів, їх вподобання, дані про покупки та замовлення, що ними здійснюються, тощо.

В умовах, коли потоки даних все більш зростають, сховища даних наповнюються величезними обсягами даних. З часом, відбувається їх накопичення та переповнення. В таких випадках, гостро стає проблема обробки та аналізу даних. Це пояснюється тим, що відбувається значний дефіцит інформації та знань, що зберігаються в необроблених («сирих») даних. У зв'язку з цією проблемою виникає інтерес до методів, що дозволяють проаналізувати та виявити знання з величезних масивів даних.

Виявлення закономірностей в базах даних для генерації рекомендацій користувачам інформаційної системи являється актуальною задачею. Серед величезної кількості товарів, що є в асортименті інформаційної системи квітів, клієнтам доведеться самостійно шукати квіти чи букети певного типу, а також схожі на них. Часто трапляється так, що в системі дуже складна навігація, та без рекомендаційних блоків, користувачу буде важко орієнтуватися, через що він просто скористається послугами іншої інформаційної системи.

Завдяки відстежуванню покупок користувачів та використанню й аналізу зібраної інформації, інформаційна система здатна буде пропонувати клієнту товари.

Метою даної роботи є аналіз предметної області та дослідження існуючих методів (зокрема основаних на генерації асоціативних правил) на предмет можливості їх застосування для розробки методу CRM системи сервісу з продажу квітів; проаналізувати різні джерела інформації та існуючі методи для створення рекомендацій, що базуються на прогнозуванні уподобань користувачів системи продажу квітів.

Враховуючи мету роботи, були поставлені наступні завдання для дослідження:

- необхідно провести дослідження й аналіз існуючих методів та оцінок побудови рекомендацій;
- провести дослідження предметної області (магазину квітів) та виявити основні бізнес процеси, що в ній відбуваються;
- розробити метод, що базується на генерації асоціативних правил з урахуванням особливостей предметної області;
- розробити інформаційну систему для магазину квітів;
- використати розроблений метод в інформаційній системі;

Об'єктом дослідження є інформаційна система, що орієнтується на продаж квітів, надає доступ кінцевим користувачам до товарів, що містяться в системі та має дуже низьку частоту переглядів товарів.

Предметом дослідження є алгоритм пошуку асоціативних правил, та методи його застосування для створення рекомендацій, котрі будуть пропонуватися кінцевим користувачам системи.

Методи дослідження. Досягнення мети атестаційної роботи ґрунтується на генерації асоціативних правил для розробки методів CRM для інформаційної системи квітів

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

Початок XXI сторіччя відзначився розвитком електронної індустрії. З розвитком Інтернету та ринкових відносин електронна комерція отримала широке поширення серед малого й середнього бізнесу. Це пояснюється тим, що за допомогою глобальної мережі стала можливою співпраця з існуючими та потенційними клієнтами через віртуальний ринок. Він дає можливість значно зменшити витрати бізнесу та стрімко розширювати клієнтську базу. Внаслідок цього з'являється багато сервісів, котрі направлені на задоволення запитів користувачів.

Квітковий бізнес завжди був і буде користуватися популярністю. Раніше, щоб придбати незвичайні квіти доводилось витратити багато часу та докласти значних зусиль на пошуки. Це пояснювалось тим, що не всі квіткові магазини мали в асортименті необхідні квіти, проте зараз, покупець може не виходячи з дому знайти, порівняти та купити квіти, що йому необхідні.

Великою популярністю користуються інформаційні системи для продажу квітів, оскільки вони значно економлять час клієнтів та дозволяють підібрати необхідні товари по оптимальній ціні. Це пов'язано з тим, що перехід на використання таких систем дозволяє заощаджувати значну кількість грошей на персонал та оренду кіоску, а тому і ціни в них є нижче, чим у квіткових крамницях. Таким чином, перехід магазину на простір мережі Інтернет є необхідним.

Автоматизовані технології, що використовуються в інтернет-магазині дозволяють уникнути паперової роботи, зменшити витрати часу на обробку і накопичення інформації та знизити фінансові витрати на обслуговування. Завдяки автоматизації з прийому замовлень співробітники компанії будуть здатні швидко отримувати та аналізувати інформацію про товари, що купують клієнти. За допомогою цієї інформації буде зрозуміло на які товари більший попит, на який менше, а це, в свою чергу, дозволить запровадити рекомендаційні механізми.

В рамках даної атестаційної роботи для автоматизації буде розглядатись діяльність компанії «Flowers Store». Основний профіль його діяльності - здійснення оптово-роздрібною торгівлі квітами, а основний бізнес-процес – це продаж квітів, букетів та супутніх до них товарів. Цільова аудиторія не обмежена певною віковою категорією чи статтю. Користуватися послугами компанії може хто завгодно, в тому числі й приватні компанії.

Щоб придбати квіти, покупець повинен завітати до квіткового кіоску компанії, вибрати товар та оплатити його. Продавець має зафіксувати в паперовому журналі всю інформацію про проданий товар, його кількість, поточну дату та суму, що була сплачена покупцем за замовлення. Також, клієнт має змогу замовити товар через телефон, що також фіксується в паперовому журналі.

Компанія займається пошуком та фіксуванням контактів та зв'язків з постачальниками квітів, цукерок, іграшок та інших супутніх товарів. Всі поточні постачальники фіксуються в паперових журналах.

Всі товари, котрі надходять від постачальників, зберігаються на складі в умовах, необхідних для збереження товарного вигляду квітів та супутніх товарів. Велика частина товарів зі складу розташовується у квіткових крамницях компанії, інша ж – зберігається на складі.

Організаційна структура фірми представлена на рис. 1.1.

Діяльність компанії здійснюється на підставі законів України від 27.03.1991 р. «Про підприємництво» (з усіма змінами та доповненнями) [1], від 07.02.1991 р. «Про власність» (з усіма змінами та доповненнями) [2], від 14.10.1992 р. «Про охорону праці» [3], від 24.03.1995 р. «Про оплату праці» [4].



Рисунок 1.1 – Організаційна структура компанії «Flowers Store»

Враховуючи вказані функції, можна прийти до висновку, що облік та продаж квітів й супутніх товарів – це основний бізнес-процес підприємства. В компанії відсутня автоматизація процесів, оскільки ведеться паперовий документообіг (журнали обліку постачальників, квітів, клієнтів, ведення звітів). Всі форми документообороту компанії необхідно автоматизувати, аби збільшити кількість клієнтів та прибутків компанії.

1.1 Визначення сфери застосування CRM-системи в електронній комерції

Електронна комерція була однією з перших сфер людської діяльності, що успішно почала використовувати всі можливості, котрі надає мережа Інтернет.

Під терміном «електронна комерція» заведено вважати, перш за все, підприємницьку діяльність, направлену на надання, поширення та рекламування товарів і послуг через глобальну мережу Інтернет [5].

Для того щоб робота систем електронної комерції стала ефективнішою в них впроваджується концепція CRM (Customer Relationship Management). Вона забезпечує автоматизоване керування взаємодії з клієнтами, та створення механізмів в котрих потреби клієнта є пріоритетними.

CRM-система – це база даних та програмне забезпечення, котрі призначені щоб автоматизувати стратегій взаємодії системи з користувачами, оптимізації маркетингу та покращення обслуговування клієнтів [6].

Завданням CRM-систем є збереження значних обсягів інформації про клієнтів (наприклад, контактну інформацію, історію транзакцій) аби мати змогу керувати зв'язками з користувачами системи, що опираються на результат аналітичної обробки даних БД системи.

CRM-системи мають широкий спектр сфер використання. Вони є незамінними в розумінні потреб користувача, збору інформації та застосуванню її для залучення більшої кількості користувачів.

Основними компонентами/складовими CRM-систем є (рис. 1.2):

- інтерфейс – надає доступ до БД CRM-системи;
- база даних – зберігає всі дані взаємодії користувача з системою і не тільки;
- аналітична – збирання, обробка та підготовка даних про клієнтів для прийняття організаційних рішень (залучення нових користувачів, збільшення обсягів продажу, пошук мір для підняття середнього чека);
- операційна – авторизація операцій, забезпечення звітності;
- фронтальна – застосування великої кількості каналів взаємодії з кінцевими користувачами (електронна пошта, дзвінки за телефоном, тощо).

Основною метою впровадження CRM-системи є необхідність створення системи, котра має єдину базу контактів (клієнтів та контрагентів) з повною історією взаємин з ними, для ефективного здійснення контролю якості продажів, отримання

статистики та аналітики роботи, керування взаємовідношеннями з клієнтами, планування подальшої роботи та розробки програм лояльності, і т. д.

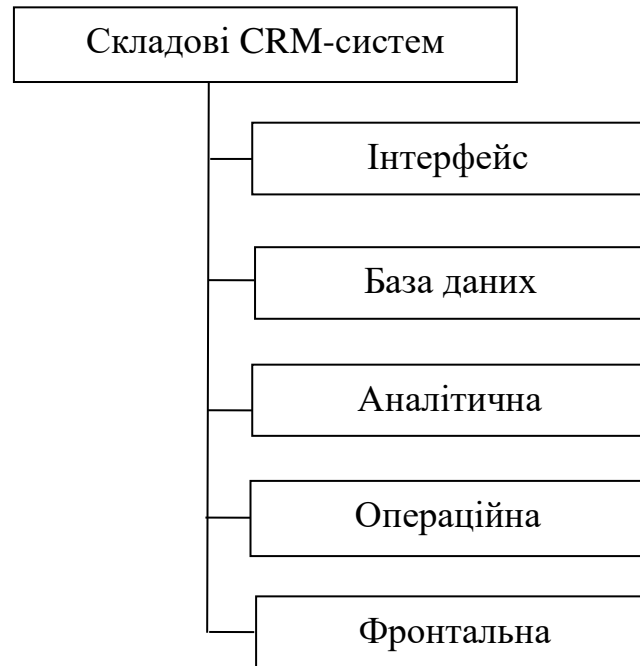


Рисунок 1.2 – Складові CRM-систем

Властивості та функції CRM-систем:

- управління продажами;
- об'єднання різноманітних бізнес-процесів в єдиній платформі;
- автоматизація документообороту;
- управління контактами (забезпечую збереження контактної інформації та доступу до неї);
- відстеження взаємодії з користувачами системи;
- збір відомостей про клієнтів;
- аналіз цільової аудиторії;
- контроль діяльності продавців та менеджерів;
- прогнозування продажів.

Інформаційна система стає більш ефективною, коли в ній впроваджуються методи CRM-систем на основі генерації асоціативних правил: підвищуються продажі, покращується зв'язок з клієнтами, удосконалюються бізнес-процеси системи. Всі ці дії направлені на залучення до системи все більшої кількості користувачів та підняття середнього чека продажів.

1.2 Актуальність роботи

Останнім часом підвищився інтерес до методів, що дозволяють знаходити знання в величезних обсягах даних. Розміри баз даних постійно збільшуються, а пошук зв'язків між даними за допомогою лише одних людських зусиль стає практично неможливим.

Знаходження знань в величезних обсягах даних є актуальною проблемою в області електронної комерції, зокрема і для інформаційної системи з продажу квітів. Користувачі інформаційної системи знаходяться в процесі вибору серед великої кількості товарів. Досить часто, клієнт заходить до інформаційної системи з метою ознайомитися з пропозиціями та каталогом товарів, його необхідно зацікавити, тому доцільним є використання рекомендаційних блоків. Коли ця функція присутня в системі – вона створює для клієнта позитивний досвід взаємодії з системою, оскільки користувач економить свій час. Це є основними факторами для прийняття рішення відносно покупки товару і, відповідно, сприяє зростанню прибутку компанії.

Проблема залучення все більшої кількості клієнтів та утримання уже наявних, в рамках сучасної конкуренції, нині є актуальною. Рекомендаційні механізми є одним зі способів її вирішення. Саме з їх допомогою можна зробити наступне:

- збільшити обсяги продажів;
- збільшити тривалості сесії та глибини перегляду сторінок;
- спростити навігацію;
- підвищити частоту покупок;

- ненав'язливо просувати різні асортиментні групи;
- збільшити середній чек.

За допомогою рекомендаційних механізмів, окрім реалізації бізнес-цілей, також буде підвищуватися розуміння потреб користувача, що грає дуже важливу роль для будь-якого проекту.

1.3 Аналіз аналогічних програмних систем

На сьогоднішні виникає все більше і більше різноманітних інформаційних систем, котрі допомагають підприємствам взаємодіяти з користувачами дистанційно, або, інакше кажучи – віртуально.

Огляд аналогів інформаційних систем, котрі орієнтуються на торгівлю різноманітними товарами являється дуже важливим процесом. Метою цього аналізу є порівняння схожих систем для виявлення їх переваг та недоліків, щоб впровадити в інформаційну систему ті функції, що будуть найбільш корисними, функціональними та зручними, а також проаналізувати те, як застосовуються в них рекомендаційні механізми. До розгляду будуть взяті харківські інформаційні системи з продажу квітів, оскільки саме вони будуть основними конкурентами розроблюваної системи.

На сайті flogina.kh продемонстрований великий асортимент товарів. Він розбитий по категоріях. Клієнт може побачити товари, що представлені на головній сторінці в різних категоріях та ознайомитися з ними. На сайті також продемонстрований блок з акціями, що є чинними на певні товари. Цей блок дає змогу системі підняти інтерес до знижок й на цьому підвищити чек від покупок.

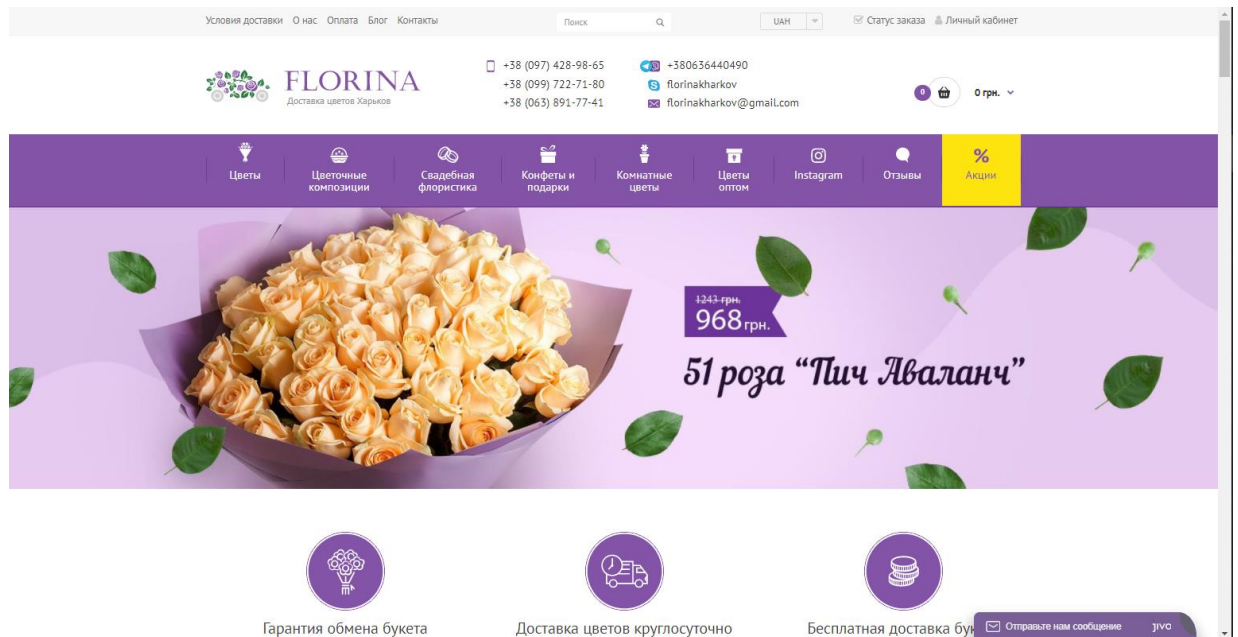


Рисунок 1.3 – Сайт florina.kh.ua

В даній інформаційній системі не реалізована функція відстежування замовлення, проте є можливість перегляду його статусу.

Варто зазначити одним з недоліків сайту – відсутність будь-якого блоку рекомендацій (як для зареєстрованих так і не для зареєстрованих користувачів), наприклад: «Популярні товари» або «Товари, що часто купують». Навіть коли клієнт буде користуватися сайтом та переглядати товари – інформація, що збирається, не буде застосовуватися на допомогу сайту та монетизації. Система ніяк не аналізує її.

На рисунку 1.3 зображена сторінка з деталями по певному товару. Зазвичай користувач звик бачити якісь пропозиції від системи, наприклад: «Ви вже розглядали цей товар» чи «Схожий товар», проте, на даному сайті такі рекомендаційні механізми не застосовуються.

The screenshot shows the Florina website interface. At the top, there is a navigation bar with icons for 'Цветы', 'Цветочные композиции', 'Свадебная флористика', 'Конфеты и подарки', 'Комнатные цветы', 'Цветы отпуску', 'Instagram', 'Отзывы', and 'Акции'. The main content area features a large image of a bouquet of chrysanthemums wrapped in purple and orange paper. To the right of the image, the product name 'Хризантемы "Бархатная осень"' is displayed, along with a star rating of 4.9/5 and a price of 490 грн. Below the price is a quantity selector set to 1, with a note that the minimum order quantity is 1. Two buttons are present: 'ЗАКАЗАТЬ' and 'КУПИТЬ В 1 КЛИК'. A detailed composition list is provided below the buttons, listing various chrysanthemum varieties and their quantities. At the bottom of the page, there is a section for 'Описание' and 'Отзывов (4)'. A small text block at the very bottom describes the delivery service in Kharkiv.

Florina
Бесплатная доставка цветов

+38 (097) 428-98-65
+38 (099) 722-71-80
+38 (063) 891-77-41

+380636440490
florinakerkov
florinakerkov@gmail.com

0 грн.

Цветы
Цветочные композиции
Свадебная флористика
Конфеты и подарки
Комнатные цветы
Цветы отпуску
Instagram
Отзывы
Акции

Главная / Услуги / Цветы / Подарки / Хризантемы "Бархатная осень"

Хризантемы "Бархатная осень"
★★★★★ 4 отзыва
490 грн.

Количество: Минимальное количество заказа этого товара: 1

ЗАКАЗАТЬ КУПИТЬ В 1 КЛИК

Состав:

Хризантема "Бакарди жёлтая"	1 шт.
Хризантема "Зембла белая"	1 шт.
Хризантема "Филин Грин"	2 шт.
Хризантема "Тренд Гинк"	2 шт.
Зелень "Солидаго голд"	2 шт.
Хризантема "Микс"	1 шт.
Упаковка	3 шт.
Лента	1 шт.

Описание Отзывов (4)

Доставка цветов Хризантемы "Бархатная осень" в Харьков.

Вообще, Хризантемы можно подарить в любое время года и это будет подарок без промаха. Но посмотрите на этот букет, сразу можно вспомнить осень и разноцветные листья, которые украшают обочины тротуара или предают деревьям той праздной атмосферы карнавала цветов. И хотя Хризантемы «Бархатная осень» - это не листья, но они способны передать ту красоту красок, на которую способен природа. Подарок учительнице в школе поднимет ей настроение и украсит лицо приятной улыбкой. А время цвета ожидает в памяти только самые приятные воспоминания. [Доставка цветов](#) Харьков сделает ваш лучший день незабываемым.

Рисунок 1.4 – Сторінка з деталями товару

Сайт є інтуїтивно зрозумілий, не переповнений безліччю інформації, з присмним дизайном. Проте, серед основних недоліків є ігнорування будь-яких рекомендаційних механізмів для покращення роботи з клієнтами. Їх відсутність є менш цікавою для кінцевого користувача та не підвищує конверсію з сайту.

Наступна інформаційна система, котра була розглянута – це svetkof.kh.ua. Вона також орієнтується на продаж квітів, букетів та подарункових товарів у вигляді цукерок та м'яких іграшок.

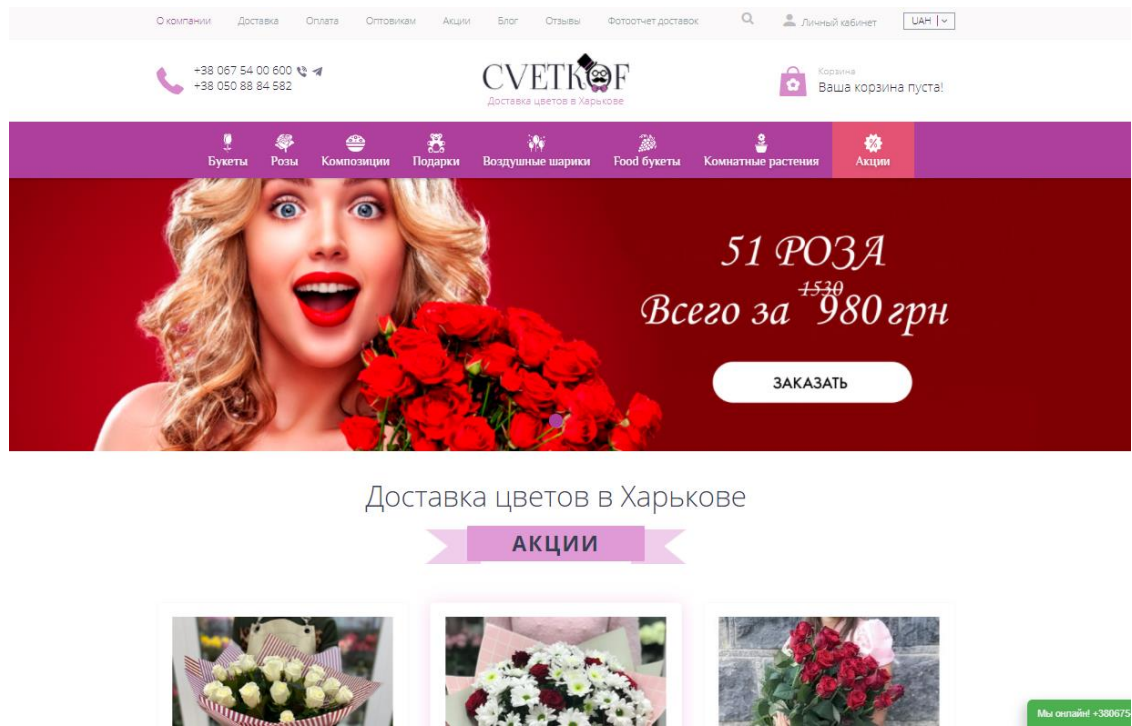


Рисунок 1.5– Інформаційна система svetkof.kh.ua

Даний сервіс має незначну схожість з попереднім (flogina.kh). проте в ньому реалізований більш цікавий підхід для передачі інформації та ознайомлення клієнта з товарами, що є на сайті.

Всі товари та категорії розташовані на сайті дуже органічно. Система дає змогу клієнту ознайомитися з різними категоріями. При переході на сторінку з деталями про товар, варто звернути увагу на те, що система пропонує йому квіти в блоці, котрий називається «Рекомендуемые товары». Якщо прогорнути нижче – то можна побачити також блок під назвою «Дополните свой заказ», основна задача котрого полягає в тому, щоб разом з одним товаром, користувач захотів би придбати ще щось із рекомендацій, тим самим підвищити чек продажу для компанії (рис. 1.6).

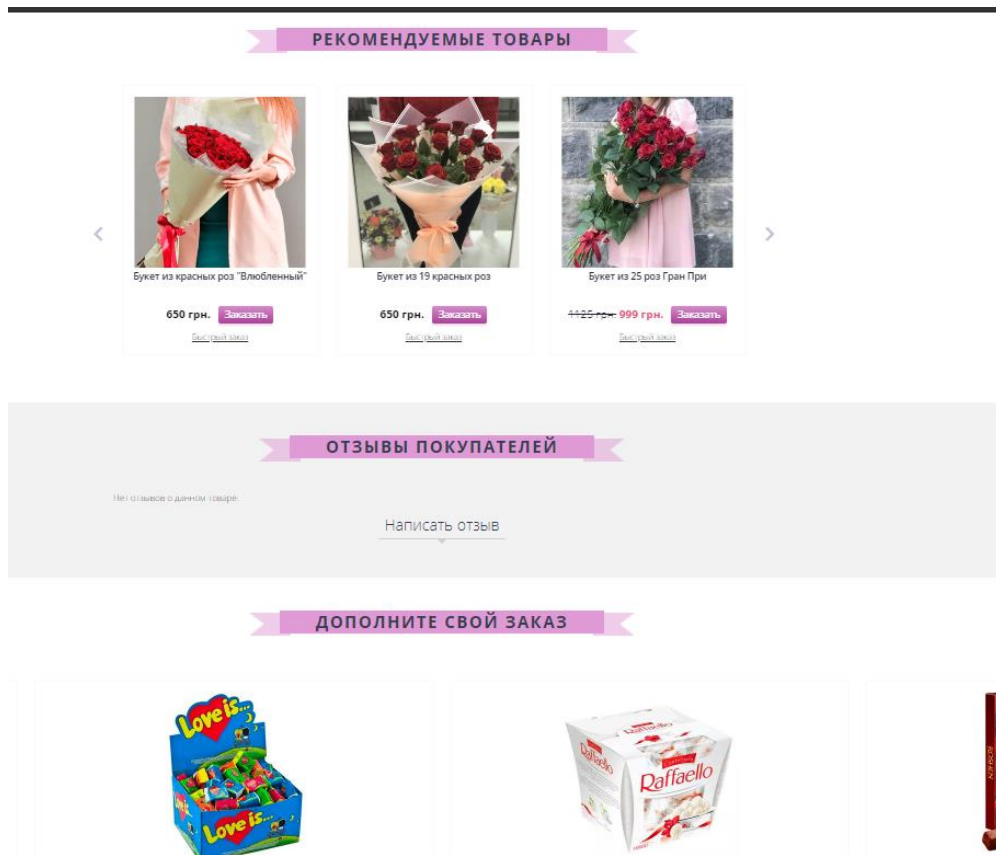


Рисунок 1.6 – Рекомендаційні блоки сервісу svetkof.kh.ua.

Даний сайт використовує рекомендаційні механізми, що здатні підвищити продажі в системі та спростити користувачу пошук і навігацію по сайту, саме це виділяє його з-поміж інших сайтів.

1.4 Постановка задачі

Об'єктом дослідження атестаційної роботи є методи CRM-системи побудовані основі генерації асоціативних правил та їх використання в інформаційній системі з продажу квітів.

1.4.1 Формалізована постановка задачі

До атестаційної роботи ставляться наступні задачі:

- проведення дослідження й аналіз існуючих методів та оцінок побудови рекомендацій;
- провести дослідження предметної області (магазину квітів) та виявити основні бізнес-процеси, що в ній відбуваються.
- обґрунтувати доцільність вибору методу генерації асоціативних правил для предметної області, що розглядається;
- розробити метод, що базується на генерації асоціативних правил з урахуванням особливостей заданої предметної області;
- впровадити розроблений метод в інформаційну систему продажу квітів;
- провести дослідження комерційних перспектив (збільшення середнього чека, частоти покупок, просування та продаж різних товарів, збільшення конверсії).

1.4.2 Характеристика предметної області

Задача розробки методів CRM-системи на основі генерації асоціативних правил, буде розглядатися й виконуватися на прикладі сфери з продажу квітів.

На сьогоднішній день існує безліч сервісів, що пропонують свої послуги та товари для клієнтів. До найбільш поширених серед них відносять:

- перегляд каталогу товарів;
- пошук товарів за категоріями;
- перегляд детальних характеристик про товари;
- реєстрація та авторизація в системі;
- додавання товарів до кошика користувача

- оформлення замовлення на товари
- доставлення квітів.

Система, котра буде розроблюватися в даній атестаційній роботі буде реалізовувати описані вище стандартні можливості систем-аналогів та спеціалізуватися на продажі квітів, букетів та супутніх до них товарів.

1.4.3 Перелік задач необхідних для вирішення

Необхідно буде вирішити комплекс задач, призначених для створення методів CRM-системи з впровадженням в неї алгоритмом генерації асоціативних правил. Дана система має відповідати наступним вимогам:

- відсутність проблеми «холодного старту» (пов'язана з відсутністю вподобань користувачів на початку роботи з системою);
- відсутність високого порогу входження;
- висока точність рекомендацій;
- відсутність побудови прогнозів на оцінках користувачів;
- можливість зацікавити нового користувача пропозиціями з перших кроків;
- зменшення «вузьких місць» для придбання знань з вибраної предметної області;
- врахувати особливості предметної області (низький термін придатності товару) при розробці методу.

1.4.3 Вхідна та вихідна інформація

Вхідною інформацією для даної задачі є:

- minsup - значення порогу мінімальної підтримки;
- minconf - значення порогу мінімальної вірогідності;

- minlift - значення порогу для мінімального підйому;
- база даних транзакцій (або двійковий контекст), тобто набір операцій;
- presc – кількість рекомендацій, що будуть виділені для однієї транзакції;
- metric – значення метрики, що використовується при видачі рекомендації.

Вихідними даними системи є:

Вихідні дані будуть являти собою список товарів рекомендацій, що були відібрані алгоритмом та список рекомендацій для введеного набору транзакцій. Результатом цього алгоритму буде набір усіх правил, що задовольняють умовам.

1.4.4 Математична постановка завдання

Для опису математичної задачі в загальному вигляді, перш за все необхідно визначити об'єкти, що складаються в досліджувальні набори наступною множиною I (1.1):

$$I = \{i_1, i_2, \dots, i_j, \dots, i_n\}, T \in I, \quad (1.1)$$

де i_j – це об'єкт, котрий входить до набору, що аналізується, n – потужність набору (загальна кількість об'єктів).

Всі об'єкти, котрі зберігаються в множині I називають транзакціями. Кожна транзакція T являє собою сукупність конкретних елементів набору (підмножин) множини I .

Набір транзакцій у яких інформація доступна для аналізу буде позначатися наступним чином (1.2):

$$D = \{T_1, T_2, \dots, T_r, \dots, T_m\}, \quad (1.2)$$

де m – кількість доступних транзакцій для аналізу.

Множина транзакцій, в котру входить об'єкт i_k , буде позначений наступним чином (1.3):

$$D_{ik} = \{T_r \mid i_j \in T_r; j = 1..n; r = 1..m\} \subseteq D \quad (1.3)$$

Довільний об'єкт транзакцій буде позначений наступним чином (1.4):

$$F = \{i_j \mid i_j \in I, j = 1..n\} \quad (1.4)$$

Множина транзакцій до якої входить набір F , буде позначатися наступним чином (1.5):

$$D_F = \{T_r \mid F \subseteq T_r; r = 1..m\} \subseteq D \quad (1.5)$$

Відношення кількості транзакцій, в яке входить набір F , до загальної кількості транзакцій називається підтримкою набору F та позначається $Supp(F)$ (1.6):

$$Supp(F) = \frac{|D_F|}{|D|}. \quad (1.6)$$

Мінімальне значення підтримки правила ($Supp_{min}$), вводиться користувачем для пошуку набору певних елементів. Частим варто вважати той набір, у якого значення підтримки буде більшим ніж значення $Supp_{min}$, що ввів користувач (1.7).

$$L = \{F \mid Supp(F) > Supp_{min}\} \quad (1.7)$$

Таким чином, основна математична постановка завдання зводиться до пошуку множини всіх частих наборів елементів.

2 ДОСЛІДЖЕННЯ МЕТОДІВ ТА АЛГОРИТМІВ РІШЕННЯ ЗАДАЧІ

Щоб інформаційна система з продажу квітів користувалася популярністю серед користувачів, вона повинна відповідати їх потребам. Проблема знаходження релевантної інформації на сьогодні є актуальною. Для її вирішення було створено багато різних алгоритмів, методів та підходів. Серед них будуть відібрані та розглянуті ті, котрі найбільш підходять для вирішення поставленої задачі, а саме: фільтрація на основі вмісту, колаборативна фільтрація, гібридний підхід та асоціативні правила.

2.1 Фільтрація на основі контенту (content based)

Фільтрація на основі контенту аналізує характеристики елементів, для того щоб ідентифікувати серед них ті, котрі мають особливий інтерес для користувача. В рамках даного підходу опис товару (квітів, іграшок, подарунків) зіставляється з інтересами користувача, отриманими з його попередніх оцінок. Чим більше товар цим інтересам відповідає, тим вище оцінюється потенційна зацікавленість користувача.

Для даного підходу необхідно створити профіль користувача та профіль елемента. Опираючись на параметри елементів системи можна буде зробити висновки про належність якогось конкретного елемента, якомусь конкретному користувачу. Для опису елементів системи та створення їх профілю рекомендаційні системи ставлять у відповідність кожному елементу певний набір ключових слів.

Так склалось, що об'єктом рекомендацій на основі контенту, частіше за все є товари з неструктурованим описом: музика, книги, статті. Такими ознаками можуть бути, наприклад, текстові описи, рецензії, склад акторів та інше. Однак ніщо не заважає використовувати і звичайні числові або категоріальні ознаки. Цей тип

системи рекомендацій може використовуватися в різних сферах, таких як рекомендації в інтернет-магазинах, новинних статей і контенту в соціальних мережах.

Основною метою контент-орієнтованого підходу є створення «профілю» для кожного користувача та елемента

При виконанні даного підходу аналізується набір характеристик предметів, які оцінюються цільовим користувачем і формують профіль інтересів цього користувача на основі характеристик предметів. Процес рекомендації полягає в зіставленні атрибутів профілю користувача з набором властивостей вмісту товару. Після цього, на основі параметрів елементів системи, можна зробити висновок щодо відповідності конкретного елемента конкретному користувачу.

Одна з поширених технік для цієї стратегії полягає в зберіганні ключових слів документів в n -мірних векторах. Кожному слову призначається своя вага. Для визначення ваги можуть бути використані техніки TD-IDF, відстань Окарі, Кульбака-Лейблера, тощо. Після того, як побудовані вектори ключових слів, необхідно вчислити переходи між користувачами та предметами. Пропозиції повинні бути рекомендовані користувачу, якщо вони мають найбільшу взаємодію з іншими користувачами або елементами, прочитаними користувачем. Один із найбільш популярних способів досягти цього є метод косинусної схожості. Чим менше кут, тим ближче документ до профілю користувача [7]. Формула косинусного коефіцієнта має наступний вигляд (2.1):

$$\text{cosine}(x, y) = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (2.1)$$

, де x – вектор профілю користувача, y – вектор документа

Переваги підходу:

- можливість зацікавити нового користувача пропозиціями з перших кроків.

Для цього не потрібно мати велику історію взаємодій клієнта з сайтом.

- присутня можливість рекомендувати навіть ті об'єкти, що не мають оцінку інших користувачів;
- є досить універсальним, тому швидко приносить результати;
- не обов'язково мати детальну інформацію про продукти, товари, тощо.

Недоліки підходу:

- сильна залежність від предметної області;
- обмежений аналіз контенту;
- точність рекомендацій дуже низька;
- обмежена корисність рекомендацій;
- проблема холодного старту (відсутність вподобань користувачів на початку роботи з системою).
- Профіль користувача та елементів (товарів), має містити в собі однаковий набір характеристик, для того, щоб їх легше було порівнювати.

2.2 Колаборативні методи фільтрації;

Колаборативна фільтрація – це метод, що дозволяє прогнозувати невідомі вподобання одного конкретного користувача інтернет-ресурсу, порівнюючи його інтереси з інтересами інших користувачів. Тобто, користувачі, котрі однаково оцінили будь-які елементи системи – можуть однаково оцінювати і інші елементи системи. Алгоритми даного методу основані на машинному навчанні.

Даний метод будується виключно на основі поведінки користувача чи групи користувачів. Принцип дії колаборативної фільтрації досить простий: якщо користувач зробив покупки, чи просто переходив на сторінку товару, то система знаходить інших користувачів зі схожими запитамі. Після цього система буде рекомендувати користувачу ті продукти, котрими цікавились інші клієнти, а користувач ще ні.

Методи колаборативної фільтрації прогнозують цінність різних елементів для певного користувача, базуючись на попередніх оцінках, що були зроблені іншими користувачами.

User-based підхід пропонує товари, котрі будуть придбані схожими користувачами: будуть усереднені рейтинги товару, проставлені іншими користувачами з вагами по ступені схожості користувачів.

Підходи на основі колаборативної фільтрації, на сьогодні, є більш популярними, чим підходи на основі фільтрації вмісту, оскільки являють собою зображення практичного досвіду.

Міра схожості $sim(u, u')$ обчислюється за матрицею оцінок R . Найбільш загальноживана метрика схожості – кореляція Пірсона і косинусна відстань рядків (стовпців) матриці [8].

Коефіцієнт Пірсона (2.2):

$$s_{i,j} = \frac{\sum_{u \in U} (r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{u \in U} (r_{u,i} - \bar{r}_i)^2} \sqrt{\sum_{u \in U} (r_{u,j} - \bar{r}_j)^2}}, \quad (2.2)$$

де $U = U_i \cup U_j$ – це множина користувачів, які оцінили об'єкти i чи j .

Косинус кута між двома векторами r_i і r_j (2.3):

$$s_{i,j} = \cos(r_i, r_j) = \frac{r_i \cdot r_j}{|r_i| \cdot |r_j|} \quad (2.3)$$

Проводиться формування кінцевої множини об'єктів S , найбільш близьких до об'єкта o . Обчислення рейтингу об'єкта o здійснюється за формулою (2.4)

$$\hat{r}_{u,o} = \frac{\sum_{j \in S} s_{o,j} r_{u,j}}{\sum_{j \in S} |s_{o,j}|} \quad (2.4)$$

Популярний підхід до формування багатьох рекомендацій - це упорядкування всіх об'єктів за критерієм схожості і вибірці деякого фіксованого кількості об'єктів з максимальним рейтингом [9]. В якості запобіжної схожості двох об'єктів виступає косинус кута між n-мірними векторами.

Переваги:

- користувач має змогу ознайомитися з повним спектром товарів ресурсу;
- висока точність рекомендацій.

Недоліки:

- нові елементи в системі повинні оцінюватися користувачами до використання в процесі рекомендації. Колаборативні системи керуються тільки вподобаннями користувачів, а нові товари не можуть бути рекомендовані, доки не наберуть необхідну кількість оцінок;
- високий поріг входження – не знаючи нічого про користувача, рекомендації є марними та не точними (проблема «холодного старту»);
- розрідженість рейтингів. Користувачі оцінюють схожі товари (предмети, елементи), що приводить до того, що тільки кілька товарів отримали оцінку користувача. Тобто, кількість оцінок, що необхідно показати набагато перевищує кількість даних оцінок.
- кількість оцінок, що треба показати набагато перевищує кількість даних оцінок.
- «білі ворони». Тип користувачів, думка та вподобання котрих не підходять під більшість оцінок
- синонімія. Схожі товари мають різне найменування.

2.3 Гібридні методи

Гібридні методи розроблені для отримання кращих результатів прогнозування, шляхом поєднання в собі декількох рекомендаційних алгоритмів. Завдяки цьому поєднанню вирішуються ті проблеми, що виникають при використанні лише одного з цих методів.

Алгоритми колаборативної фільтрації, на практиці, показують високі показники ефективності, саме тому вони часто фігурують в гібридних методах в парі з іншими алгоритмами фільтрації. Щоб гібридні методи працювали ефективно, необхідно правильно підбирати вагові коефіцієнти, інакше, використання методу не принесе жаданих результатів [10].

Наприклад, у нас є дані про опис продуктів, профіль користувачів і історію його покупок. Ми можемо покращити систему прогнозування рекомендацій шляхом об'єднання методів колаборативної фільтрації й алгоритмів фільтрації за змістом. В разі появи нового користувача в системі, про який немає історії покупок, ми зможемо вибирати, котрий з методів доцільно застосовувати. Основним недоліком алгоритму колаборації є проблема «холодного» старту (недостатність вхідної інформації для точного прогнозування), в той час, як фільтрація на основі змісту здатна запропонувати користувачу якийсь товар з самого початку його взаємодії з системою. Таким чином, гібридний метод застосує фільтрацію на основі змісту для побудування рекомендацій новим користувачам, а в разі великого обсягу статистичних даних будувати більш точний прогноз, використовуючи методи колаборативної фільтрації.

Коли користувач інформаційної системи дає свій відгук на різні об'єкти, рекомендаційна система формує матрицю оцінок переваг R . З часом, до матриці переваг будуть застосовуватися алгоритми колаборативної фільтрації.

Перетворення зворотного зв'язку користувача в числове значення його переваг – це не проста задача. Вона виконує важливе завдання для налаштування рекомендаційної системи.

Існують наступні види зворотного зв'язку

- з явним зворотним зв'язком;
- з неявним зворотним зв'язком.

Часто можна зустріти в інформаційних системах використання оцінок товарів з зображенням п'яти зірок. Ця шкала є одним зі способів отримання явного зворотного зв'язку. Неявний зворотній же зв'язок будується на основі покупок чи відвідування сторінок користувачем, тощо.

Неявний зворотний зв'язок важко визначити, оскільки не зрозуміло, позитивно чи негативно відноситься до товару користувач, переглядаючи його. Виникає невизначеність зворотного зв'язку на ступінь переваги.

В гібридному методі є багато різних моделей та підходів. Однією з них є єднальна модель. Поєднання різноманітних джерел отримання інформації здатне виправити проблеми використання кожного з методів, а також значно покращити результати. Від доступності даних у системі будуть залежати значення вагових коефіцієнтів. У випадку, коли отриманої інформації буде достатньо для роботи цих підходів, вони будуть робити однаковий внесок до рейтингів.

Важливою перевагою даних систем над іншими є те, що вони комбінують в собі рекомендації, що дозволяють покращити якість й ефективність системи в цілому (підняти якість прогнозування рекомендацій, тощо).

Не зважаючи на те, що гібридні системи допомагають боротися з недоліками описаних раніше методів та є найбільш ефективними серед них, вони все ж дуже складні для реалізації. Особливо, ця складність відчутна при адаптації алгоритмів до конкретної сфери застосування рекомендаційної системи.

2.4 Рекомендаційні механізми на основі асоціативних правил

Алгоритм пошуку асоціативних правил є одним з найпопулярніших серед методів інтелектуального аналізу даних. Даний метод призначений для виявлення

закономірностей між пов'язаними фактами та подіями, інформація про які зберігається в базі даних [11]. Наприклад, можна виявити закономірність придбання троянд та супровідних товарів у вигляді іграшкових ведмедів до дня всіх закоханих.

Асоціативні правила являють собою механізм знаходження логічних закономірностей (правил) між пов'язаними елементами (подіями чи об'єктами). Знання таких наборів дозволяє оптимізувати товарний асортимент та підвищити ефективність продажів. На прикладі інформаційної системи квітів можна продемонструвати асоціативне правило наступним чином: коли клієнт намагається додати товар до кошика користувача, знизу буде розташовано рекомендаційний блок, в якому система пропонує йому придбати ще ті товари, що за статистикою, купуються разом з букетом (наприклад, коробка цукерок, чи м'які іграшки).

Щоб знайти асоціативні правила, необхідно провести пошук таких наборів елементів, котрі відповідають умовам мінімальної підтримки (MinSupport). Асоціативні правила будуть генеруватися з тих наборів елементів, котрі були знайдені з урахуванням того, що їх вірогідність відповідає умовам мінімального порогу (MinConfidence).

Асоціативні правила інтуїтивно зрозумілі. На практиці їх можна позначити як: якщо (умова) то (результат). Наприклад, асоціативним правилом є: якщо (Квіти, букети) то (цукерки) означає, що якщо користувач придбає квіти чи букети, то також є можливість того, що він придбає цукерки.

Правила поділяються на:

- корисні
- тривіальні
- незрозумілі

Корисними правилами вважаються ті, що відкривають дійсну, раніше невідому інформацію, що має логічне пояснення. До тривіальних відносять інформацію, що показує вже відомі закономірності, тому вони не несуть практичної користі.

Незрозумілі правила дають такі результати, що складно інтерпретувати, та не може бути застосована без додаткового аналізу.

Одним з популярних алгоритмів, для здобування даних є Apriori. Він є масштабованим алгоритмом для генерації асоціативних правил. Apriori здатен обробляти величезні набори даних та вирішувати поставлену задачу за оптимальний час.

Щоб можна було оцінювати продуктивність та корисність правил, застосовуються різні частотні критерії, котрі аналізують появу елемента в масиві експериментальних даних. До них відносять підтримку (support), вірогідність (confidence).

Для оцінки корисності та продуктивності правил, що перебираються – використовуються різні частотні критерії, що аналізують появу елемента в масиві експериментальних даних. Найважливішими з них є підтримка (support) та Вірогідність (confidence). Експерт визначає значення цих параметрів правил [12].

Підтримкою прийнято вважати кількість транзакцій, що містять в собі певний набір даних. Підтримка визначається за наступною формулою (2.5):

$$supp(x_1 \cup x_2) = \frac{\sigma(x_1 \cup x_2)}{|T|}, \quad (2.5)$$

де σ – це кількість транзакцій, в котрих є x_1 та x_2 , T – кількість транзакцій

Вірогідність – показник того, як часто правило спрацьовує для всього набору даних. Визначається за наступною формулою (2.6):

$$conf(x_1 \cup x_2) = \frac{supp(x_1 \cup x_2)}{supp(x_1)}, \quad (2.6)$$

2.5 Впровадження алгоритму асоціативних правил Apriori в інформаційну систему

Для розробки методів CRM в інформаційній системі, буде застосовуватися алгоритм генерації асоціативних правил. Він опирається на аналіз транзакцій (кошика по купця). Метод дозволяє визначити те, які товари були придбані разом. Наприклад, якщо в результаті аналізу було визначено, що типовим шаблоном покупки є букет квітів з коробкою цукерок, то вони можуть розміщуватися в одному блоці на сайті, щоб можна було «спровокувати» користувача придбати їх сумісно один з одним.

Рекомендації, побудовані на основі асоціативних правил дозволяють зменшити запаси товарів на складі, збільшити прибуток від їх збуту, а також мати можливість рекомендувати перш за все ті товари, термін придатності котрих сплине раніше ніж в інших.

Впровадження асоціативних правил в інформаційну систему зумовлено необхідністю рекомендувати користувачам товари, котрі найчастіше зустрічаються в кошиках покупців. Алгоритм повинен виявляти нестандартні закономірності серед товарів, що були придбані споживачами одночасно.

Існує багато реалізацій алгоритму генерації асоціативних правил. В якості алгоритму для розробки асоціативних правил, пошуку шаблонів та виявлення залежностей між товарами був застосований самонавчальний алгоритм Apriori. Він являє собою масштабований алгоритм пошуку асоціативних правил та здатен обробляти величезні об'єми даних за прийнятний час

Серед його основних переваг є простота, інтуїтивна зрозумілість правил, вичерпність та швидкість пошуку правил.

При розробці алгоритму були задані наступні критерії для пошуку:

- мінімальний рівень підтримки – 0,3;
- мінімальний рівень вірогідності – 0,6.

Значення для параметрів мінімальної вірогідності та підтримки були вибрані таким чином, щоб обмежити кількість знайдених правил. При використанні низького значення підтримки – буде відбуватися пошук цікавих, унікальних та не тривіальних товарів. Значення мінімальної підтримки було встановлено 0.3, як оптимально низький рівень підтримки, оскільки, якщо він буде нище – можлива генерація статистично не аргументованих правил.

Рівень вірогідності був виставлений рівним 0,6. Це пояснюється тим, що вірогідність не має бути занадто малою, оскільки цінність правила, в такому випадку, буде викликати багато сумнівів. Якщо вірогідність занадто велика, то може статися так, що асоціації можуть бути не знайденими. Саме тому, указана вірогідність є оптимальним показником.

Для використання алгоритму, всі дані, що будуть надходити на його вхід, необхідно привести до нормалізованого бінарного вигляду та змінити їх структуру (якщо елемент є в транзакції – буде виставлена одиниця, якщо відсутній – нуль).

В роботі алгоритму Apriori є підтримка властивості антимонотонності, що дозволяє скоротити простір пошуку. В рамках її властивостей виконується наступна нерівність (при розгляді множини товарів $\{A, B, C\}$) (2.7):

$$supp(A \cup B \cup C) \leq supp(A \cup B) \leq supp(A) \quad (2.7)$$

Антимонотонність дозволяє уникнути «жадібного» перебору даних та проводити швидко обробку величезних масивів даних.

Принцип роботи алгоритма Apriori полягає в наступному: спочатку, знаходяться набори елементів, що зустрічаються найчастіше, а потім з них вилучаються асоціативні правила. Правила, рівень вірогідності яких перевищує заданий поріг вважаються вірогідними і можуть використовуватись для рекомендації користувачу. Той факт, що квіти – це товар з коротким терміном зберігання необхідно врахувати в процесі рекомендацій. Для цього обираються товари, термін придатності, яких сплине

у наступні декілька днів та відсортовуються за зменшенням спливу терміну. Після цього, порівнюються товари в правилах з відсортованими товарами, на предмет входження останніх до правил. Якщо ми знаходимо співпадіння – будемо використовувати це правило для рекомендацій, якщо жодний з товарів не знаходиться в правилах – то пропонувати той, що найшвидше зіпсується у зв'язці з товарами, котрі знаходяться в найвірогіднішому правилі.

Алгоритм асоціативних правил зображено на рис. 6.1

На першому кроці алгоритму відбувається зчитування кожного елемента транзакції та подальший підрахунок одноелементних наборів, котрі часто зустрічаються.

На другому кроці необхідно сформувати кандидатів (часті набори елементів) та визначити підтримку для кожного елемента з набору (кандидату), шляхом сканування множини транзакцій. Цей етап є дуже ресурсоемним

На третьому кроці відбувається порівняння підтримки кожного кандидата з вказаною мінімальною підтримкою ($minsupp$). Після порівняння відкидаються кандидати, підтримка яких менше вказаного мінімуму ($S \geq minsupp$), якщо ж вона задовольняє умову – одноелементний кандидат зберігається в певній множині (L1).

Четвертий крок – це формування нових кандидатів. Вони складаються з двох елементів, котрі найчастіше зустрічаються в одноелементних кандидатах.

П'ятий крок – зчитування транзакцій з бази даних, щоб визначити підтримку S для кожного кандидату набору k -елементів.

Шостий крок – відкидання/видалення кандидатів, підтримка яких менше вказаного мінімуму ($S \geq minsupp$), якщо підтримка кандидата задовольняє умову – необхідно додати його до частих наборів k -елементів, що зберігаються в певній множині.

Сьомий крок – формування нових кандидатів, що складаються з трьох елементів, котрі найчастіше зустрічаються в двоелементних кандидатах.

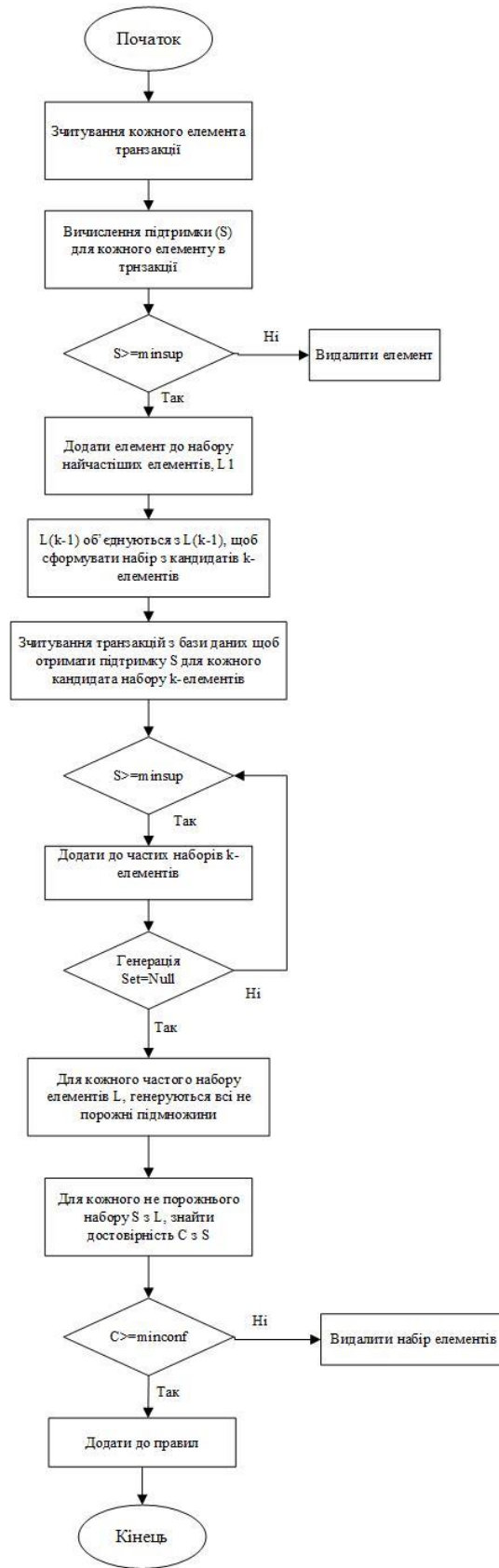


Рисунок 2.1 – Алгоритм асоціативних правил

Восьмий крок – перевірити, чи значення підтримки всіх кандидатів не перевищують значень заданого мінімального порогу підтримки, якщо ні – то повторюється крок 6-7, якщо так – то всі множини товарів, що найчастіше зустрічаються – знайдені, вони формуються в не порожні підмножини, а алгоритм переходить до другого етапу – пошуку асоціативних правил.

Дев'ятий крок - взяти один елемент з множини товарів, що часто зустрічаються, в вигляді умови, та один товар в вигляді наслідку, якщо мінімальна впевненість більше ніж у набору елементів, якщо ні – додати до правил.

Алгоритм асоціативних правил представляє список товарів (з урахуванням терміну придатності квітів), що входять до рекомендацій кінцевим користувачам. В результаті впровадження алгоритму збільшився середній чек продаж, конверсія та зменшилась кількість товарів, що зіпсувалися й були не придатні до продажу.

2.6 Розробка методу CRM з урахуванням предметної області

В рамках виконання даної атестаційної роботи розглядається інформаційна система з продажу квітів. Квіти – це товар, що дуже швидко псується та втрачає свою продовольчу спроможність. Якщо вчасно їх не продати – компанія зазнає збитків.

Специфіка даної предметної області в тому, що товари не завжди можуть бути доступними (в рамках сезонності), вони можуть бути не сумісними один з одним та мають дуже короткий термін придатності. При розробці CRM методу, ці фактори необхідно брати до уваги. Враховуючи це, було прийнято рішення впровадити в інформаційну систему метод генерації асоціативних правил.

Беручи до уваги особливості предметної області метод CRM буде розроблятися з використанням алгоритму Apriori. В рамках проведеного аналізу предметної області було виявлено, що раніше даний алгоритм не застосовувався в системах з продажу квітів.

Багато рекомендаційних методів гарно працюють лише тоді, коли є числові оцінки об'єктів користувачами. В цій особливості є одна з їх найбільших проблем. Для того, щоб такі рекомендаційні методи користувались популярністю потрібні активні дії від користувачів, через це не завжди рекомендації та прогнози будуть точними.

Системи рекомендацій, що базуються на таких методах як метод колаборативної фільтрації, фільтрації за змістом, гібридні методи – є яскравими представниками методів прогнозування побудованих на нейронних мережах.

Нейронні мережі – це потужний метод моделювання, котрий дозволяє відтворювати надзвичайно складні залежності, з метою автоматизувати, прискорити та допомагати людині в роботі з системою, покращити роботу програми, тощо [13].

Нейронні мережі широко застосовуються для вирішення різноманітних задач. Їх часто застосовують саме в області прогнозування, оскільки з їх допомогою можна передбачити об'єми продажів товару, рекомендувати товари, що найбільш відповідають інтересам клієнта, тощо. Основною перевагою для нейронних мереж є їх здатність в самостійному навчанні, проте, їх доцільно застосовувати в тому випадку, коли не можна розв'язати проблему за допомогою чіткого алгоритму.

Методи на основі нейронних мереж не зовсім підходять до вибраної предметної області. Однією з причин, є неоднозначність рішень. Тобто, результатом роботи буде набір елементів (товарів), в котрих не має однозначного вирішення поставленої задачі – рекомендації товарів (або товарів буде занадто багато, або мало, або не буде взагалі). Впровадження в систему методу асоціативних правил дозволить отримати однозначні рекомендації з урахуванням таких особливостей – як короткий термін зберігання. До того ж, спочатку в рекомендаціях будуть показуватися ті товари, у котрих термін зберігання буде підходити до кінця раніше, ніж в інших.

Використання методів побудованих на нейронних мережах потребує великих зусиль на те щоб правильно організувати навчальну вибірку, проте це все одно може привести до тупикових ситуацій, крім того поведінка навченої нейронної мережі не завжди може бути однозначно вірогідною.

Головним недоліком рекомендаційних методів, побудованих на нейронних мережах є те, що вони побудовані по принципу «чорний ящик». Неможливо побачити та зрозуміти, що відбувається всередині мережі. Це дуже ускладнює процес інтерпретації результатів.

На відміну від нейронних мереж, метод генерації асоціативних правил є «білим ящиком». Використовуючи його можна отримати результати в вигляді логічних правил, що можна легко інтерпретувати.

Більшість алгоритмів машинного навчання працюють з числовими наборами даних, а тому мають тенденцію бути математичними. Проте застосування аналізу асоціативних правил – підходить, як для числових, так і для не числових даних. Правила асоціацій корисні для аналізу та прогнозування кошику користувача, їх переваг. Враховуючи ці фактори, створюються логічні правила, котрі здатні враховувати термін придатності товарів (квітів, букетів, супутніх товарів) та їх сумісність один з одним.

Методи колаборативної фільтрації, фільтрації за змістом та гібридної фільтрації опираються на заповнені профілі користувача з великою історією взаємодій та їх вподобань зіштовхуються з рядом проблем. Однією з головних є проблема «холодного» старту. Ця проблема є критичною, для інформаційної системи, що розробляється. Продаж квітів не відноситься до товарів першої необхідності (наприклад, їжа, медикаменти, тощо). Розроблювана система буде характеризуватися не регулярним використанням, тому використання вищеописаних алгоритмів не є доцільним. В методі генерації асоціативних правил, даної проблеми не буде, оскільки буде використовуватися вже існуюча база даних транзакцій.

Таким чином, для розробки CRM методів був вибраний метод генерації асоціативних правил, оскільки він краще підходить до специфіки предметної області та подібних систем взагалі.

Розроблюваний CRM метод буде застосовувати алгоритм Аргіогі для того, щоб будувати рекомендації групі користувачів не залежно від приватних інтересів,

вподобань та оцінок товарів користувачами системи (для побудови рейтингів). Ефективність алгоритму асоціативних правил аналізується на великій вибірці транзакцій.

3 ПРОЕКТУВАННЯ СИСТЕМИ

3.1 Розробка системних вимог до CRM-системи

Предметною областю для проектування CRM системи була вибрана сфера з продажу квітів. Алгоритм генерації асоціативних правил буде впроваджений до CRM системи, як метод рекомендацій. Розроблювана система має виконати завдання автоматизації роботи компанії з продажу квітів та мати впроваджений в неї рекомендаційний метод.

Завдяки використанню та впровадженню алгоритму асоціативних правил в інформаційній системі будуть пропонуватися товари, котрі можуть зацікавити кінцевих користувачів.

Інформаційна система призначена для продажу квітів та ведення комерційного бізнесу в мережі Інтернет, для підвищення прибутку компанії. Виходячи з описаної інформації та цілей розробки даної системи були поставлені наступні вимоги до її проектування:

В системі мають бути присутніми два логічних блока: серверна частина і клієнтська частина.

Серверна частина повинна реалізовувати внутрішню роботу всієї системи, реалізовувати алгоритм реалізації рекомендацій в системі, відповідати на запити клієнтської частини та надавати їй інтерфейс для керування інформаційною системою. База даних серверної частини (СУБД) повинна реалізовувати внутрішні процедури збереження й доступу до даних.

Клієнтська ж система вирішує питання взаємодії системи з користувачами. Вона взаємодіє з серверною частиною, відправляючи на сторону сервера запити та отримує від нього необхідні дані. Ці дані вона має показувати кінцевому користувачу в

зручному для його розуміння вигляді, приховуючи технічні деталі функціонування інформаційної системи.

Інформаційна система, повинна бути масштабованою та відмовостійкою, тому що в майбутньому, кількість користувачів системи буде зростати.

Системні вимоги до системи автоматизації:

а) серверна частина інформаційної системи має бути розроблена з використанням платформи СУБД PostgreSQL;

б) клієнтська частина інформаційної системи має бути реалізована в вигляді з використанням мови програмування Java та його фреймворку Spark;

В інформаційній системі клієнтська частина реалізується в вигляді додатку для локальної мережі (web-сторінки глобальної мережі Інтернет). Необхідно уточнити вимоги до інтерфейсу клієнтської частини інформаційної CRM-системи з продажу квітів.

Інтерфейс web-сторінок доступу до даних системи має забезпечувати виконання наступних аспектів:

- відкрити повний доступ адміністратору до всіх функцій системи;
- обмежити доступ до деяких функцій системи для незареєстрованого користувача. Він не повинен мати змогу покласти товар до кошика покупця і, відповідно, здійснювати оформлення замовлень;
- відкрити доступ до таких функцій системи, як кошик покупця та оформлення замовлення в системі, для зареєстрованого та авторизованого користувача.

3.2 Визначення функціональних вимог

Щоб визначити функціональні вимоги до CRM системи, була розроблена її функціональна модель з використанням стандарту IDEF0. Цей стандарт дозволяє виділити основні бізнес-процеси в системі.

Стандарт IDEF0 – це методологія функціонального моделювання, використовуючи яку, система, що вивчається, постає у вигляді набору взаємопов’язаних функцій (функціональних блоків) [14].

Будь-яка схема IDEF0 розпочинається з концептуальної діаграми. Ця діаграма описує систему в узагальненому вигляді.

На рис. 3.1 зображені вхідні дані. Після надходження до входу системи, вони обробляються для отримання кінцевого результату. На виході контекстної діаграми зображується дані, що є результатом діяльності функцій системи. Контекстна діаграма також містить документи та інструкції, що необхідні для правильного управління процесами, а також механізми, що представляють собою ресурс/ресурси, для виконання основної бізнес-функції. Головною функцією системи є облік замовлень товарів.

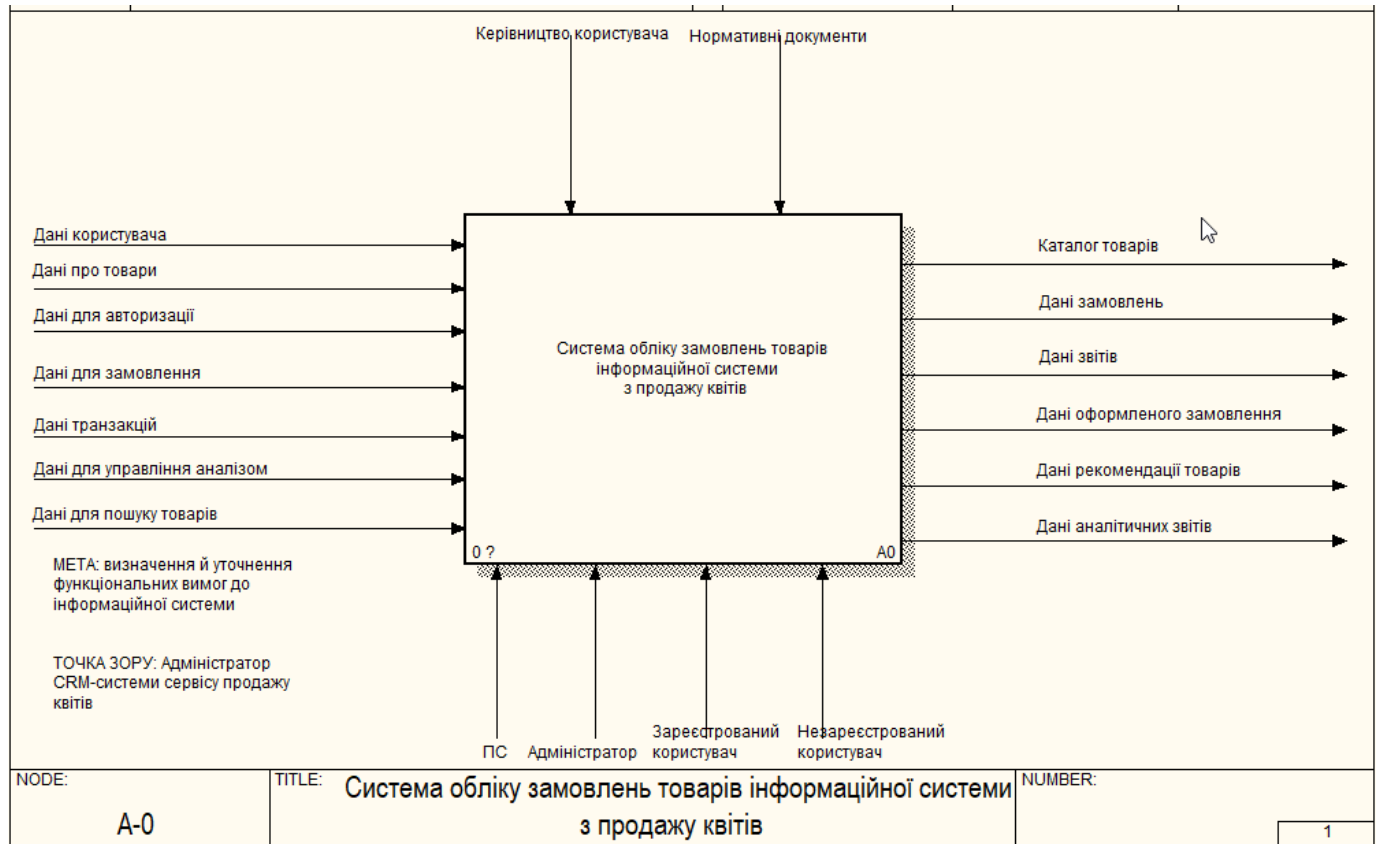


Рисунок 3.1 – Концептуальна діаграма системи електронної комерції

Розроблена контекстна діаграма відображає погляд на функції системи з точки зору адміністратора інформаційної системи з продажу квітів. Для того щоб більш детально зрозуміти, на які функції буде розбиватися основний бізнес-процес проводиться функціональна декомпозиція.

На рис. 3.2 зображена декомпозиція концептуальної діаграми. З її допомогою відбувається уточнення основної функції розробки CRM-системи для сервісу з продажу квітів.

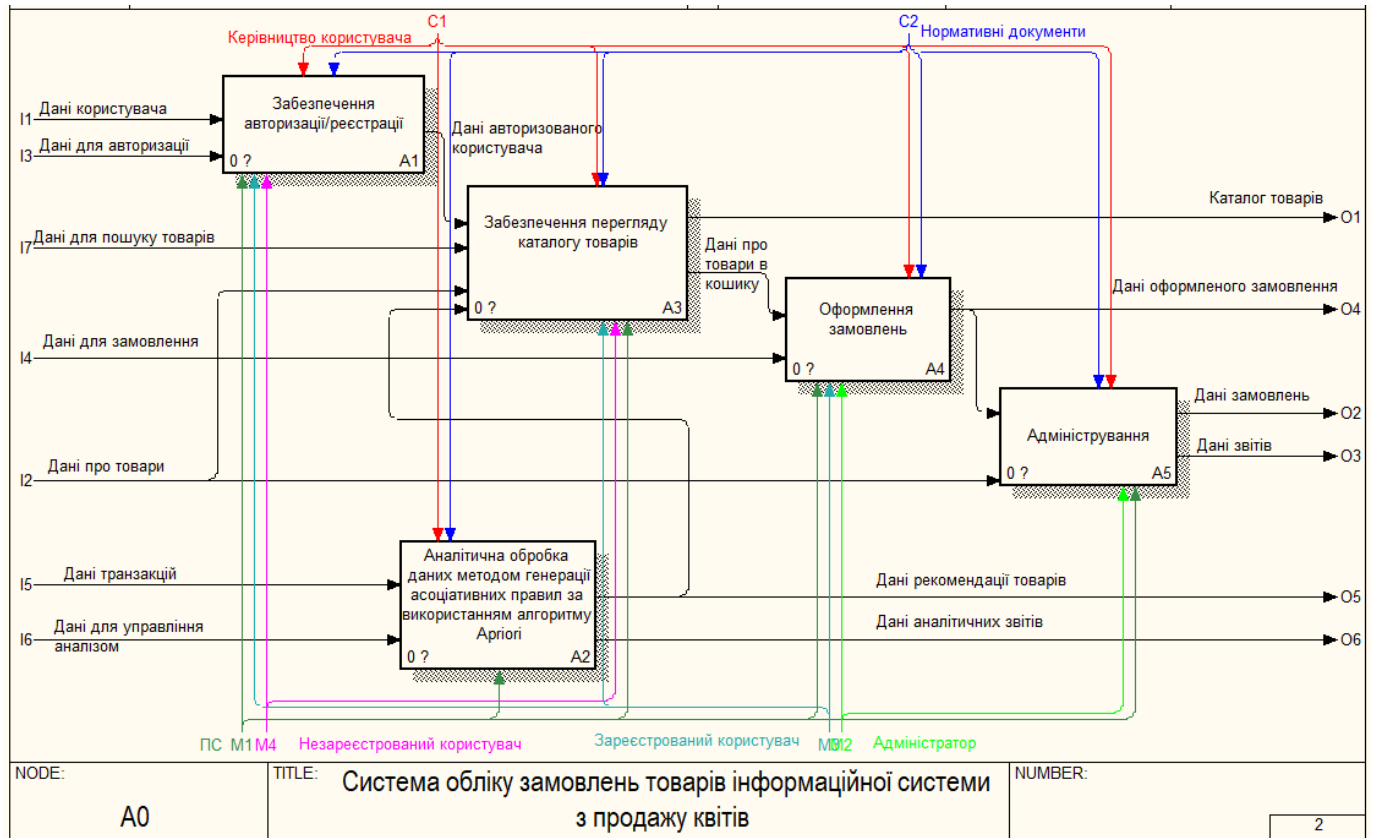


Рисунок 3.2 – Діаграма декомпозиції, що представляє функції системи

До декомпозиції увійшли наступні підфункції, що реалізуються системою:

- забезпечення авторизації/реєстрації;
- забезпечення перегляду каталогів квітів;
- оформлення замовлень;
- адміністрування;
- аналітична обробка даних.

Функція «Забезпечення авторизації/реєстрації» потребує для свого виконання дані користувача (для реєстрації) або дані для авторизації (для входу в систему). Після її виконання на вхід функції «Забезпечення перегляду каталогів квітів» будуть надходити дані про вже авторизованого користувача. Окрім того, на вхід функції будуть надходити дані про товари та рекомендаційні товари, а також дані для пошуку товарів. В результаті виконання даної функції, на виході буде отримано каталог

товарів та дані про товари в кошику. Для функції «Оформлення замовлень» на вхід надходять дані про замовлення (тобто дані, що будуть вводитися в форму замовлень) а також дані про товари в кошику користувача, результатом її роботи будуть дані оформленого замовлення. Функція «Адміністрування» отримує на вхід дані про товари та дані оформленого замовлення, а на виході виводить дані замовлень та дані звітів. На вхід функції «Аналітична обробка даних» надходять дані транзакцій та дані, що необхідні для управління аналізом. Результатом виконання будуть дані рекомендації товарів та дані аналітичних звітів.

На рис. 3.3 зображена декомпозиція функції системи «Забезпечення авторизації/реєстрації користувача». З її допомогою відображаються ті основні процеси, що відповідають за забезпечення реєстрації нових користувачів системи та авторизації користувачів, що вже зареєстровані в системі.

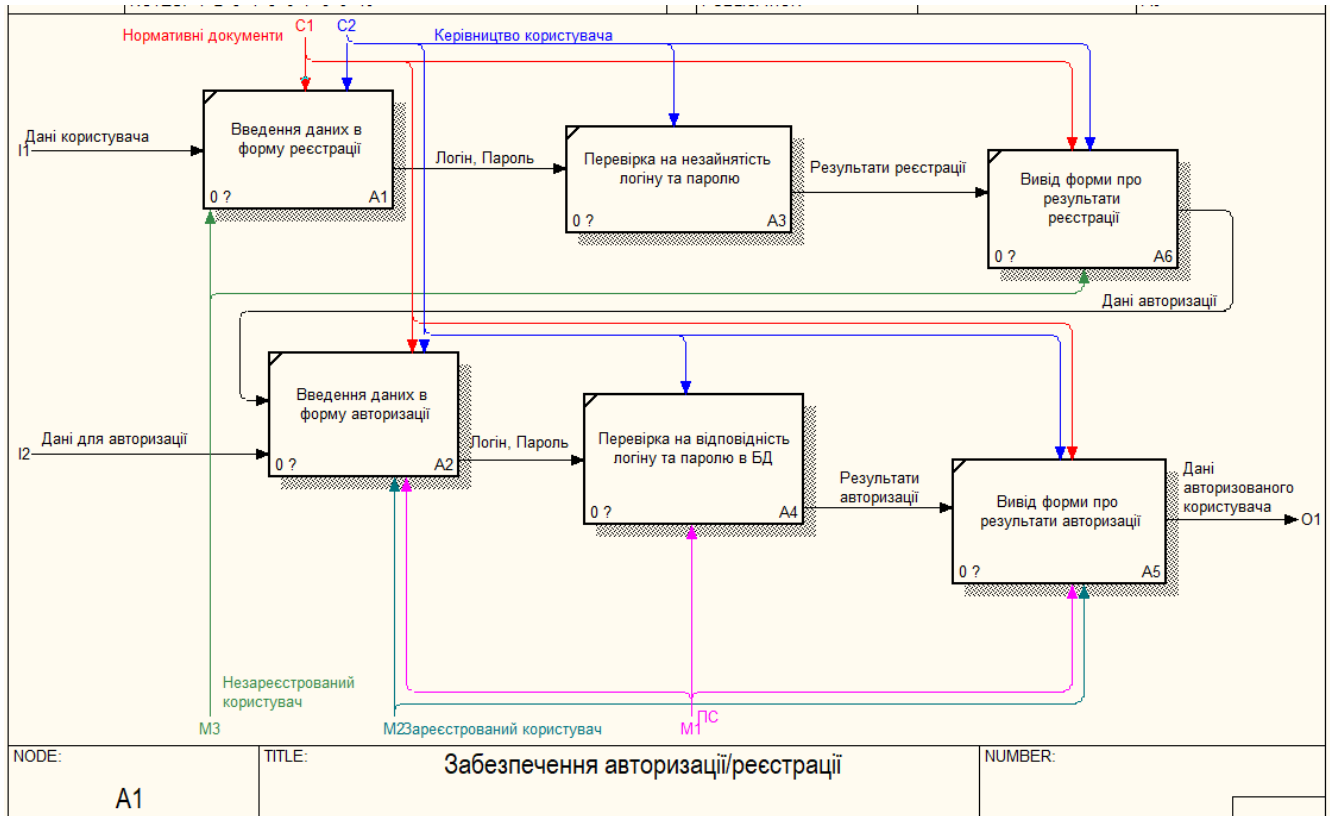


Рисунок 3.3 – Діаграма декомпозиції функції «Забезпечення авторизації/реєстрації користувача» системи

Функція розбивається на наступні підфункції:

- введення даних в форму реєстрації;
- перевірка на незайнятість логіну та паролю;
- вивід форми про результати реєстрації;
- введення даних в форму авторизації;
- перевірка на відповідність логіну та паролю в БД;
- вивід форми, про результати авторизації.

На рис. 3.4 зображена декомпозиція функції «Забезпечення перегляду каталогу». Дана функція відповідає за дії системи при перегляді асортименту товарів, як авторизованими, так і не авторизованими користувачами. Крім того, демонструє коли будуть застосовуватися результати роботи рекомендаційного алгоритму, що буде розроблюватися для системи. Рекомендації будуть відображатися при перегляді усього асортименту товарів, а також на сторінці перегляду інформації про товар.

Функція розбивається на наступні підфункції:

- перегляд товарів;
- вибрати товар;
- покласти товар до кошику користувача.

Для виконання функції перегляду товарів необхідно отримати на її вхід дані про товари, рекомендаційні товари, та, при необхідності, дані для пошуку товарів, в результаті виконання функції, буде отриманий каталог товарів. Він буде надходити до входу функції «Вибрати товар», на виході будуть отримані дані вибраного товару та надходити до функції «Покласти товар до кошика користувача» разом з даними авторизованого користувача, оскільки в системі цю функцію може виконати лише зареєстрований користувач. В результаті виконання функції, будуть отримані дані про товари в кошику.

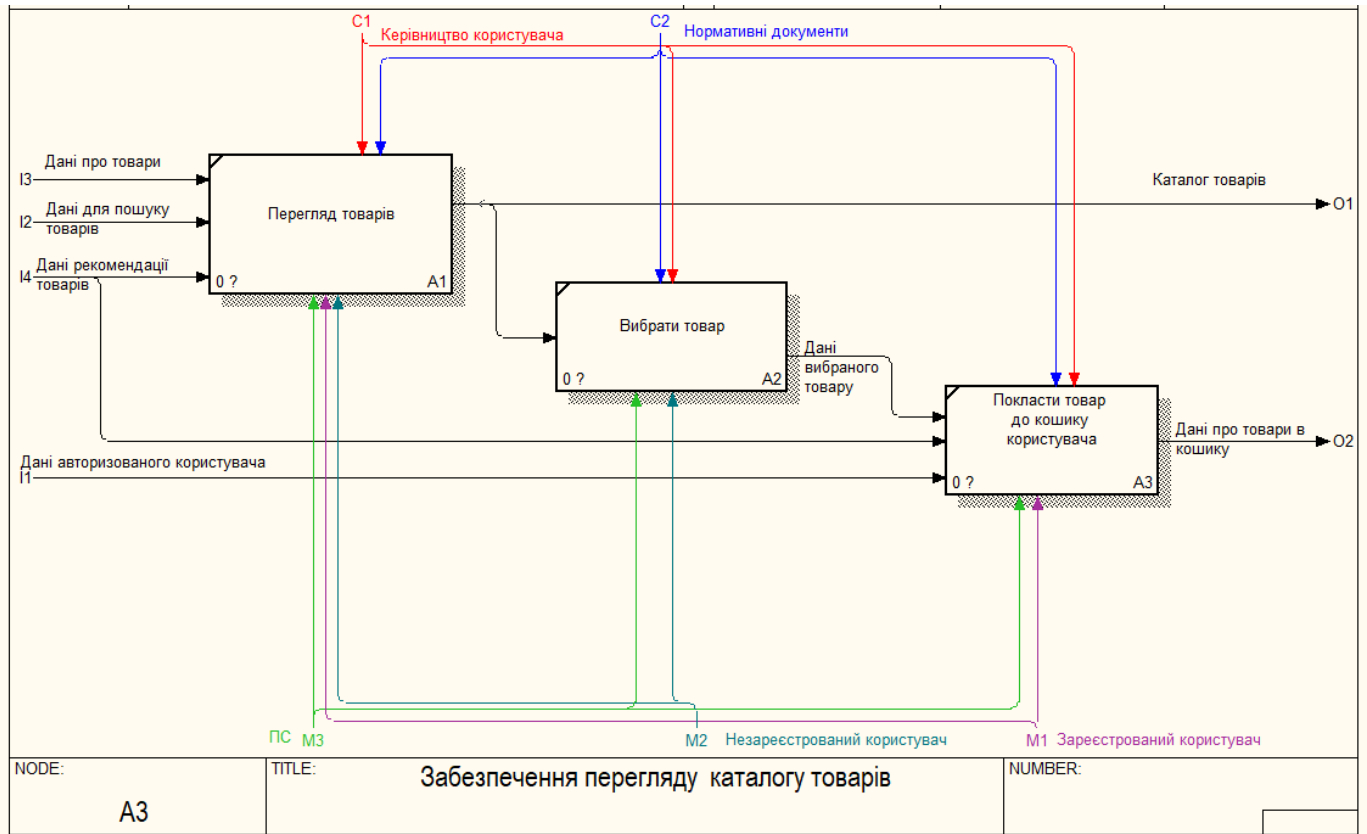


Рисунок 3.4 – Діаграма декомпозиції функції «Забезпечення перегляду каталогу» системи

На рис. 3.5 представлена декомпозиція функції «Оформлення замовлень». Функція відображає один з найголовніших процесів в системі – оформлення замовлення. На діаграмі продемонстрована послідовність виконання процесів в системі, що необхідні для того, щоб забезпечити оформлення замовлення на товар.

Функція розбивається на ряд підфункцій:

- введення даних в форму для створення замовлення;
- виведення форми, що замовлення успішно оформлене;
- обробка замовлення.

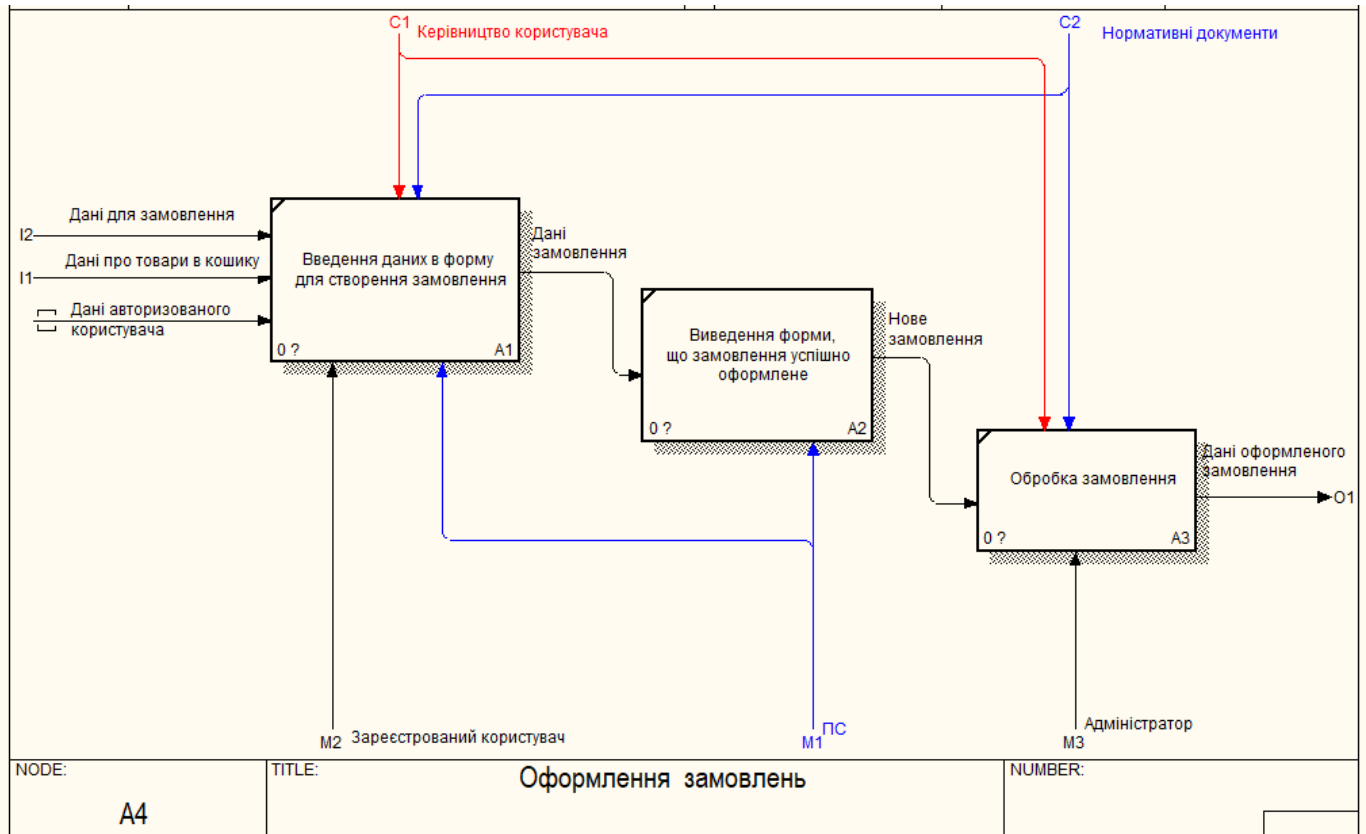


Рисунок 3.5 – Діаграма декомпозиції функції «Оформлення замовлень» системи

Для виконання функції «Введення даних в форму для створення замовлення» на вхід надходять дані для замовлення, про товари в кошику користувача, та, безпосередньо дані про самого авторизованого користувача. Після виконання функції на виході отримуються дані замовлення. Коли вони надходять на вхід функції «Виведення форми, що замовлення успішно оформлене» система отримує нове замовлення, що йде на вхід до функції «Обробка замовлення» Результатом роботи цієї функції є дані оформленого замовлення.

На рис. 3.6 представлена декомпозиція функції «Адміністрування» системи. Дана функція демонструє процес управління адміністратором інформаційною системою. До процесів керування системою відноситься: оновлення асортименту товарів, публікування актуальної інформації та актуалізація каталогу товарів;

створення звітів з продажу квітів, опираючись на товари, що часто замовляють та тих, котрі замовляють рідко.

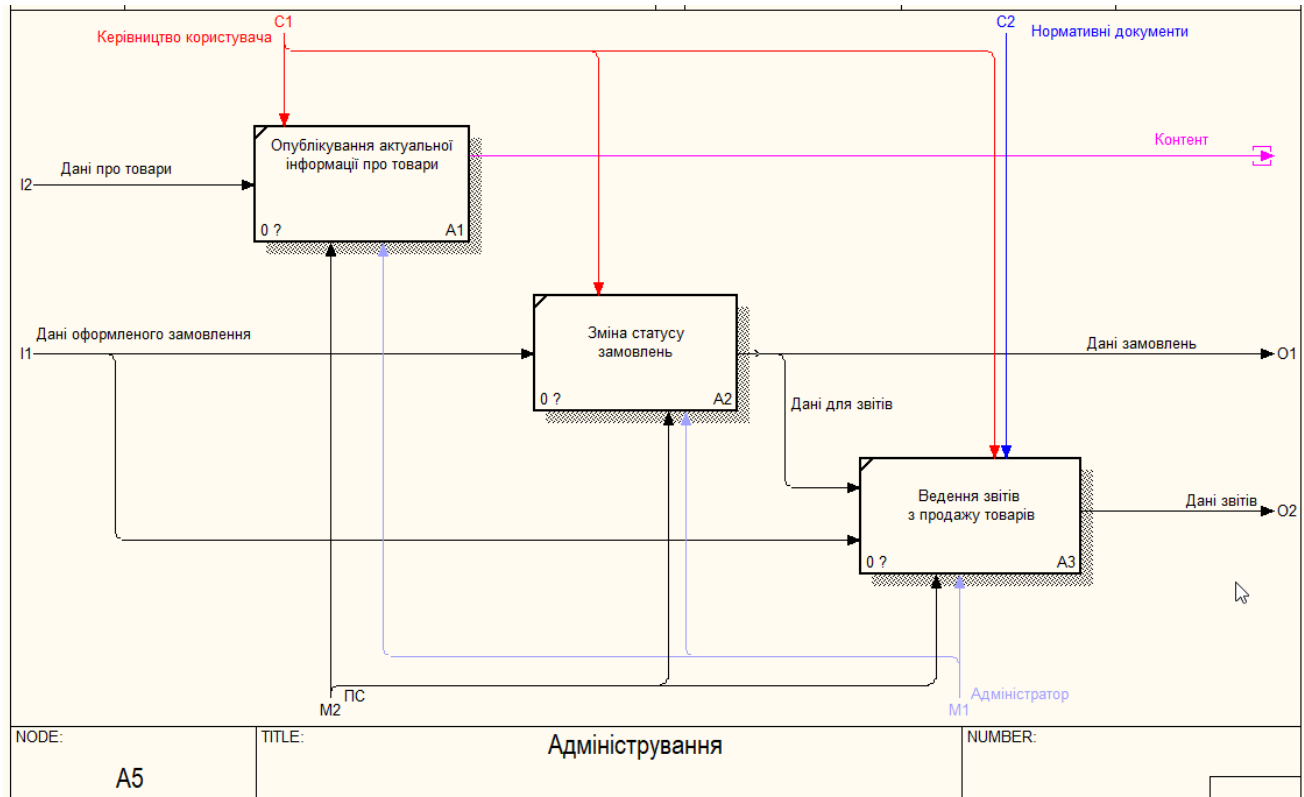


Рисунок 3.6 – Діаграма декомпозиції функції «Адміністрування» системи

Дана функція розбивається на наступні функції:

- опублікування актуальної інформації про товари
- зміна статусу замовлень
- ведення звітів з продажу товарів.

Дані про товари надходять до функції «Опублікування актуальної інформації про товари» результатом виконання даної функції є контент, що буде демонструватися в інформаційній системі. Для виконання функції «Зміна статусу замовлення» на вхід повинні надходити дані оформленого замовлення, а результатом роботи даної функції будуть дані замовлень та дані для звітів. Функція «ведення звітів

з продажу товарів» має на своєму вході дані для звітів та дані оформленого замовлення. На виході даної функції будуть дані звітів.

На рис. 3.7 представлена декомпозиція функції «Аналітична обробка даних». Вона демонструє те, як система оброблює дані транзакцій, щоб відібрати рекомендації товарів. Спочатку дані транзакції обробляються до необхідного вигляду, а далі використовується алгоритм генерації асоціативних правил для отримання рекомендацій товарів.

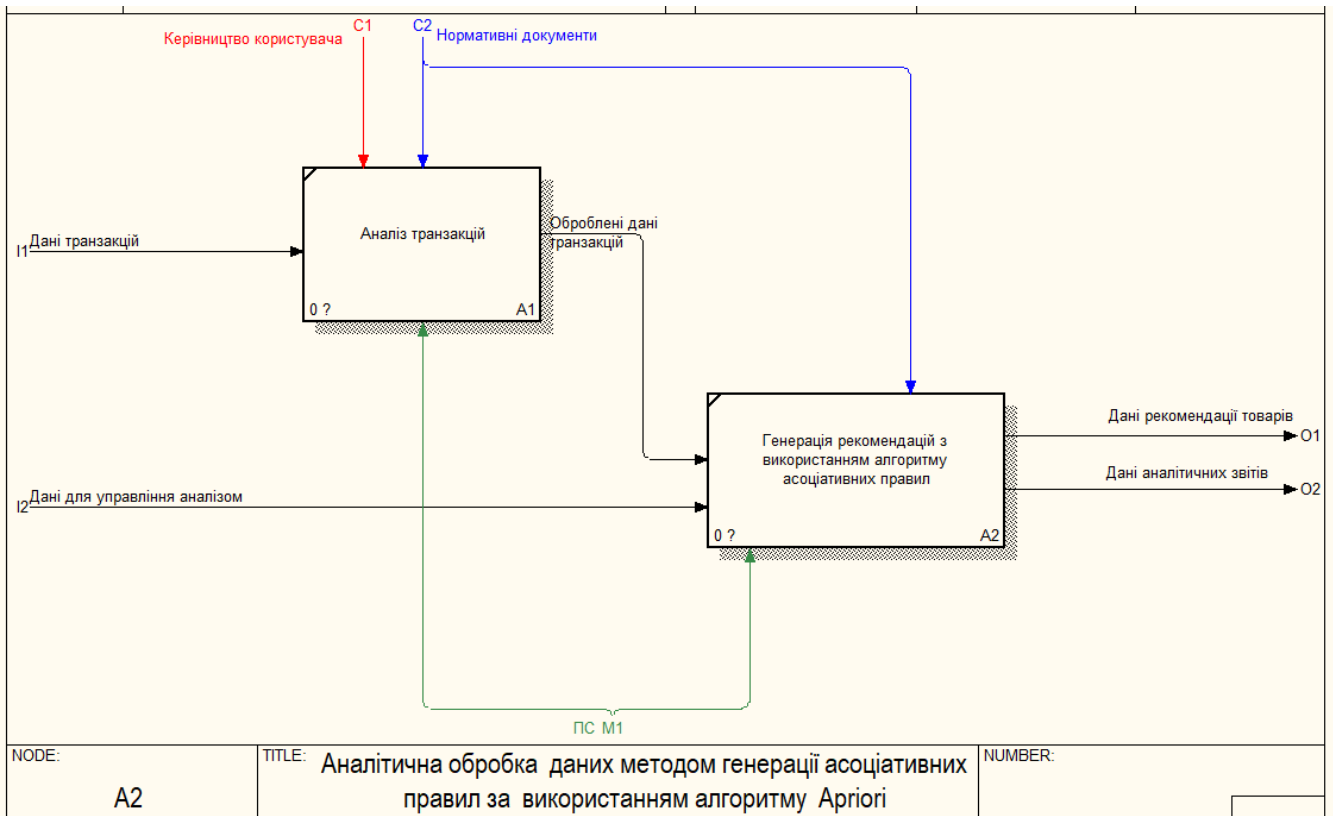


Рисунок 3.7 Діаграма декомпозиції функції «Аналітична обробка даних» системи

Функція «Аналітична обробка даних» розбивається на декілька функцій:

- аналіз даних;
- генерація рекомендацій з використанням алгоритму асоціативних правил.

Дані транзакцій, що надходять до входу функції «Аналіз даних» оброблюються в рамках даної функції, формуються дані транзакцій, а на виході отримуються оброблені дані транзакцій, що необхідні для застосування функцією «Генерація рекомендацій з використанням алгоритму асоціативних правил». До її входу також надходять й дані для управління аналізом. На виході даної функції будуть дані про рекомендовані товари та дані аналітичних звітів.

В рамках проведеного функціонального моделювання були уточнені функціональні вимоги до розроблювальної CRM-системи.

Функціональні вимоги до розроблюваної системи наступні:

- забезпечення авторизації;
- забезпечення реєстрації нових користувачів;
- перегляд асортименту товарів;
- редагувати інформації про товари;
- надання рекомендацій;
- забезпечення оформлення замовлень;
- забезпечення обробки замовлень;
- забезпечення можливості адміністрування;
- забезпечення публікування актуальної інформації;
- мати можливість змінювати статус замовлення.

3.3 Розробка моделі потоків даних компонентів CRM-системи

Для більш точного визначення вимог до інформаційної системи є створення діаграми потоків даних DFD.

DFD-діаграми забезпечують зручний спосіб опису потоків інформації, як між частинами розроблювальної системи, так і між системою та її зовнішнім середовищем [15].

Концептуальна DFD-діаграма відображає взаємодію системи з зовнішнім середовищем, тобто інформаційних потоків між системою і зовнішніми сутностями (такими як користувач системи, чи адміністратор), з котрими вона має бути зв'язана. На рисунку 3.8 зображена концептуальна діаграма, що відображає погляд на функції системи з боку адміністратора.

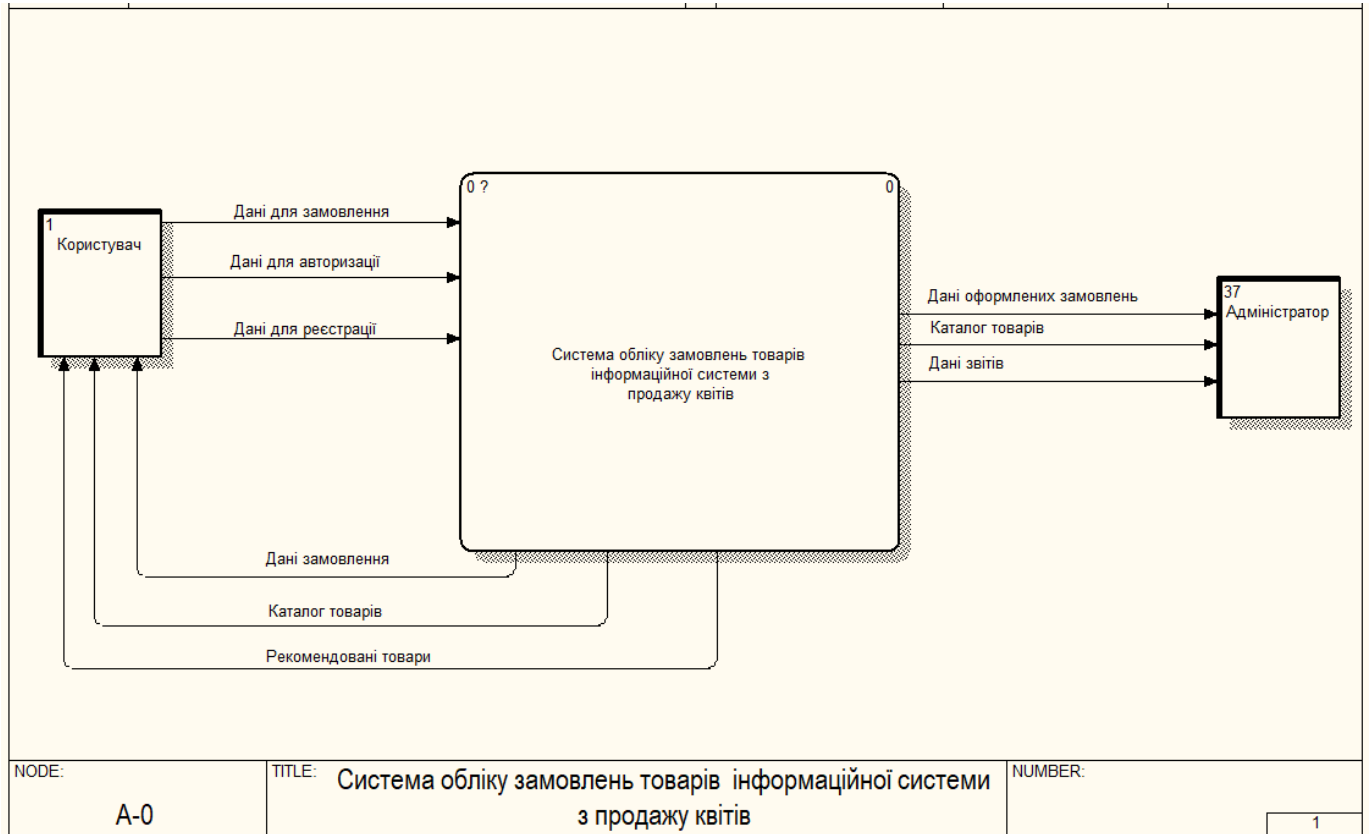


Рисунок 3.8 – Концептуальна діаграма DFD інформаційної системи

На контекстній діаграмі демонструється зв'язок інформаційної системи з зовнішніми сутностями «Користувач» та «Адміністратор». Зовнішня сутність «Користувач» взаємодіє з системою шляхом надання на її входи таких потоків даних як: дані для авторизації, реєстрації та замовлення. В результаті виконання внутрішніх процесів в інформаційній системі, на виході користувач отримує дані замовлення, каталог товарів та товари, що рекомендує система. Адміністратор системи отримує

дані оформленого замовлення та дані, що необхідні для звітів, а також каталог товарів, що необхідний для редагування, видалення та додання нових товарів.

На рис. 3.9 представлена декомпозиція концептуальної діаграми DFD. Вона дозволяє декомпонувати зовнішні потоки інформації концептуальної діаграми та зв'язати їх з внутрішніми потоками даних, що потрібні для функціонування системи.

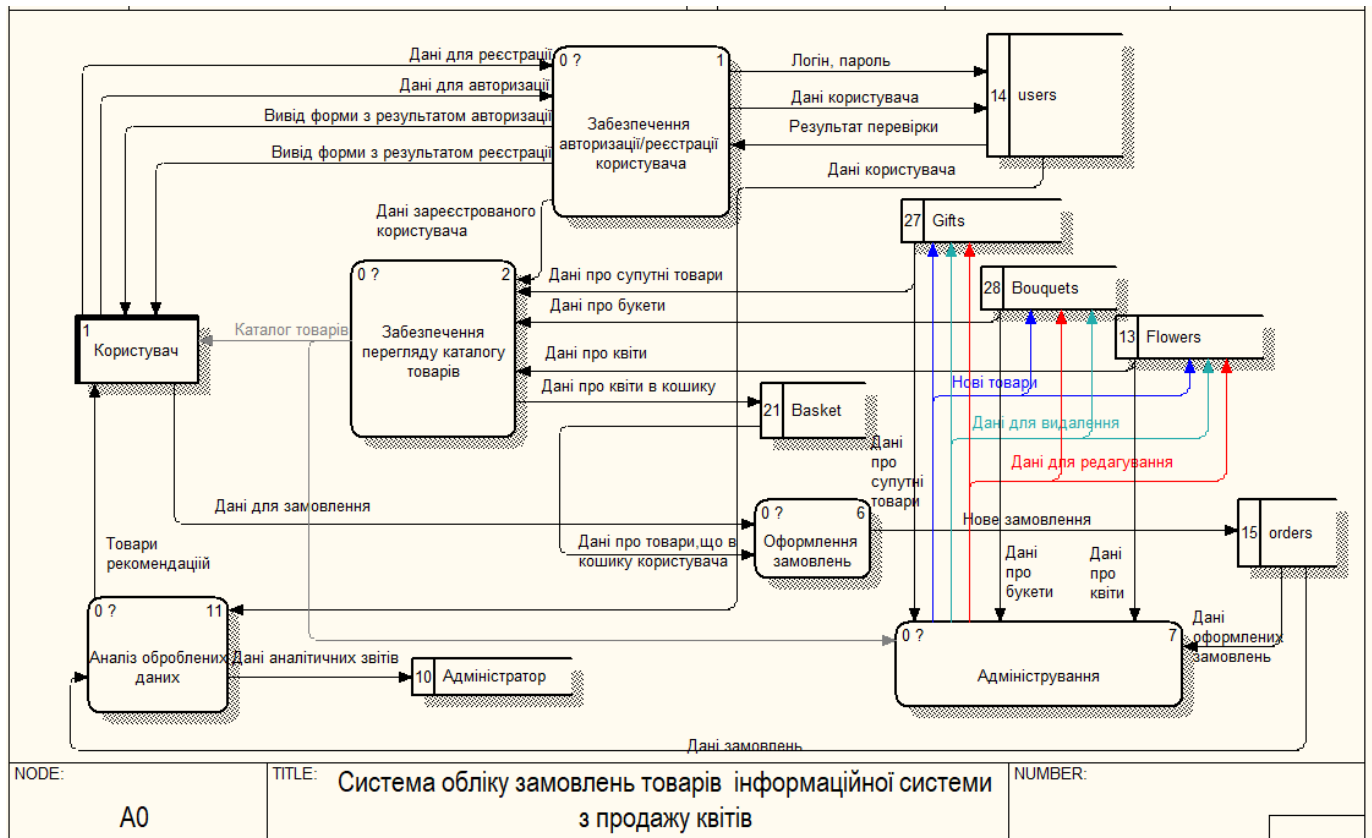


Рисунок 3.9 – Діаграма потоків даних DFD декомпозиції

Основний процес розбивається на наступні роботи/процеси:

- забезпечення авторизації/реєстрації користувача;
- забезпечення перегляду каталогу товарів;
- оформлення замовлення;
- адміністрування;
- аналіз оброблених даних.

На рис. 3.10 представлена декомпозиція функції «Забезпечення авторизації/реєстрації користувача» системи.

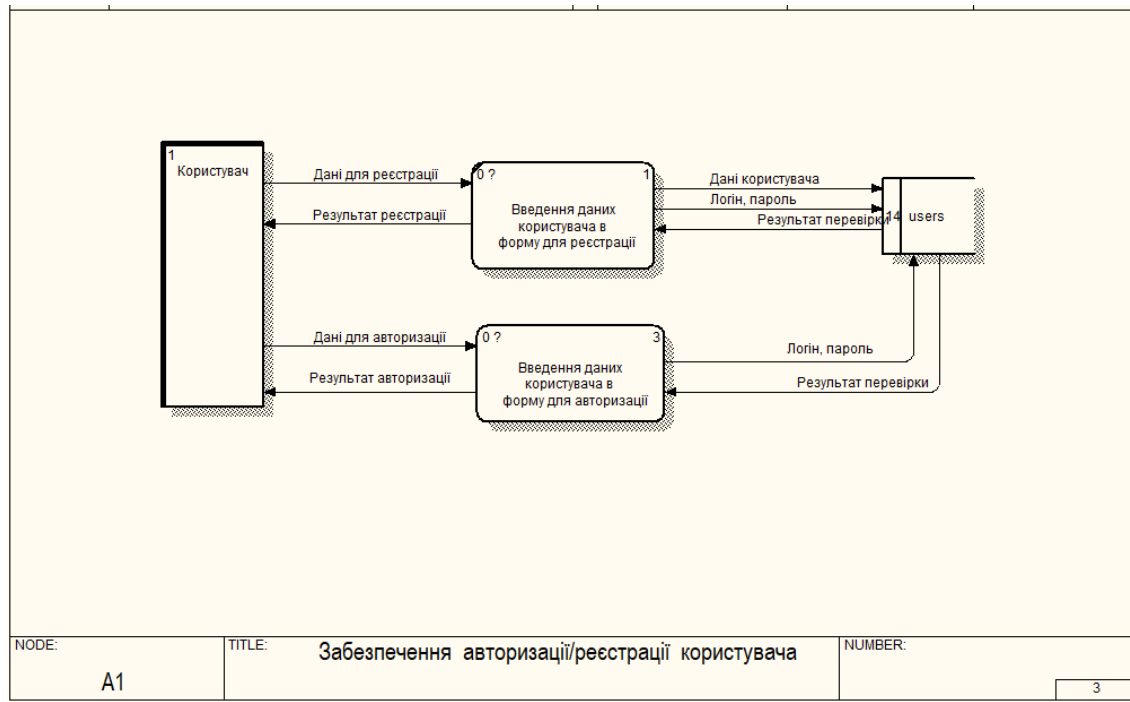


Рисунок 3.10 – Діаграма декомпозиції функції «Забезпечення авторизації/реєстрації користувача» системи

Для виконання функції «Забезпечення авторизації/реєстрації користувача» необхідні сховища даних (Data Store):

- сутність «users» з атрибутами: «First Name», «Surname», «Birthday», «Login», «Password».

Використання сутності «users» дає можливість зберігати інформацію зареєстрованого клієнта та звіряти співпадіння логіну та паролю з тими, що вже зберігаються в системі.

На рис. 3.11 представлена декомпозиція функції «Забезпечення перегляду каталогу товарів» системи.

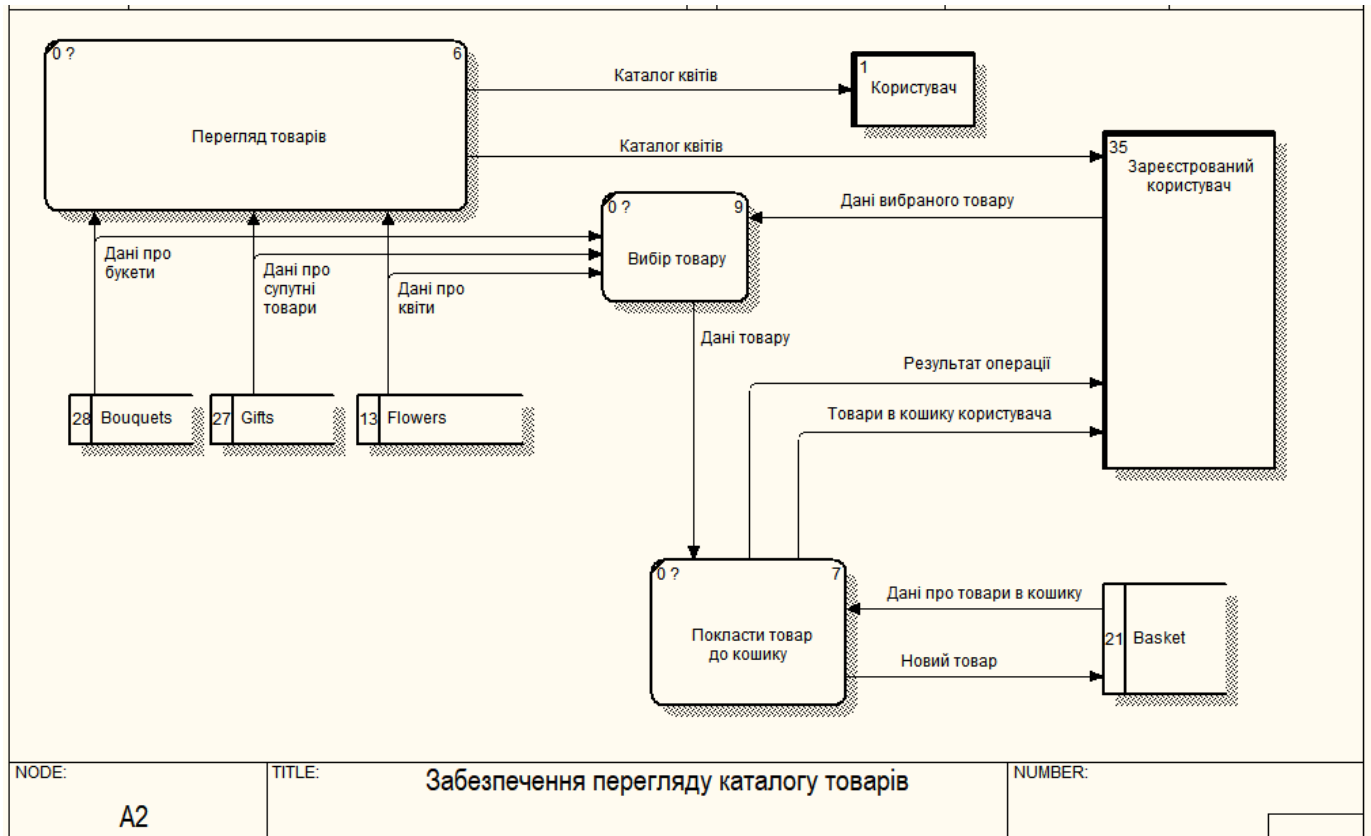


Рисунок 3.11 – Діаграма декомпозиції функції «Забезпечення перегляду каталогу товарів» системи

Для виконання функції «Забезпечення перегляду каталогу товарів» необхідні наступні сховища даних (Data Stores):

- сутність «Basket» атрибути якої (idBasket, idFlowers, idOrders, Kolichestvo, Price, idService) дають можливість додавати товари до кошика, для можливості їх подальшого придбання, а також переглядати інформацію про товар в кошику.
- сутність «Flowers» атрибути якої (idFlowers, Name, Description, Price, Quantity, Date_coming, idCategory) дають можливість зчитувати інформацію про квіти, котрі мають в системі
- сутність «Bouquets» атрибути якої (idBouquet, Name, Description, Price, Quantity, FKidFlBouq, idCategory) дають можливість зчитувати інформацію про букети, їх кількість котрі мають в системі

- сутність «Gifts» атрибути якої (idGifts, Name, Description, Price, Quantity, Date_coming, idCategory) дають можливість зчитувати інформацію про супутні товари, котрі маютья в системі

На рис. 3.12 представлена декомпозиція функції «Оформлення замовлення» системи.

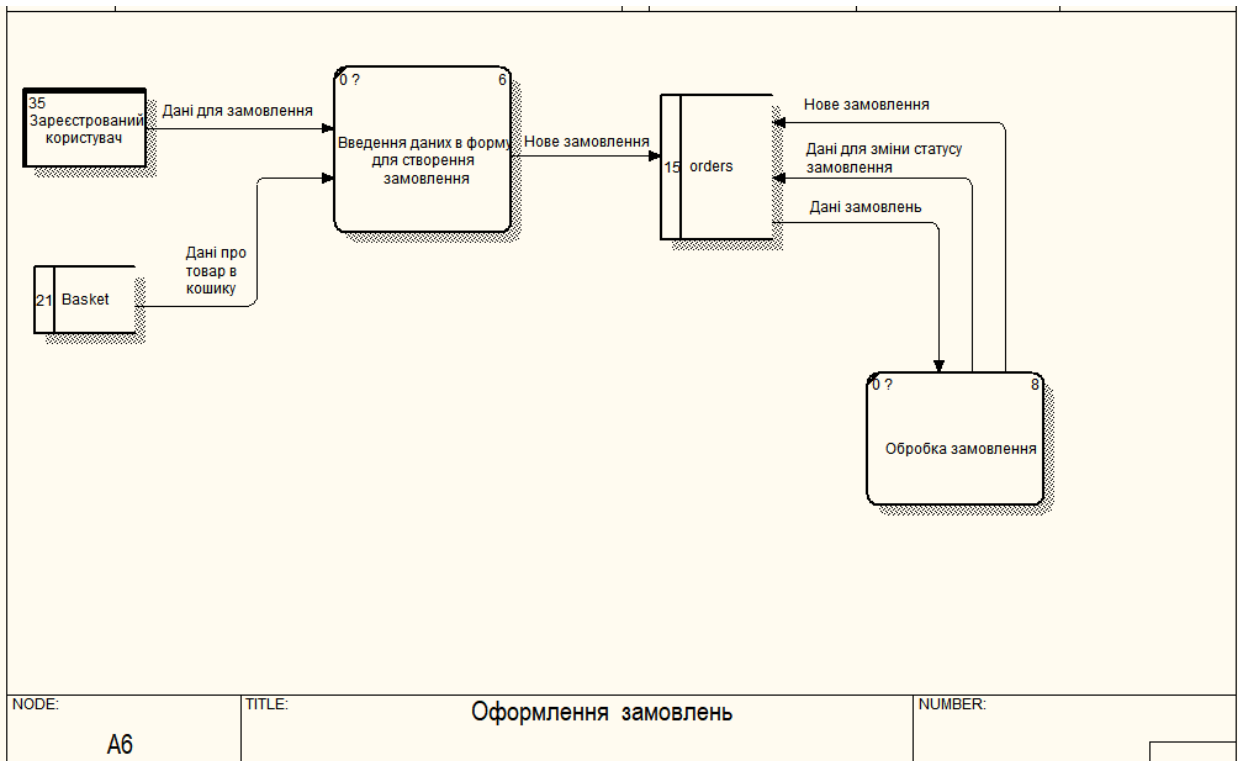


Рисунок 3.12 – Діаграма декомпозиції функції «Оформлення замовлень» системи

Для виконання функції «Оформлення замовлення» необхідні наступні сховища даних (Data Stores):

- сутність «orders» атрибути якої (idOrders, idUser, Date, Price, Address, Telephone, idDostavka, idStatusOrder) дають можливість читати інформацію з таблиць БД та використовувати її з метою аналізу даних;

- сутність «Basket» атрибути якої (idBasket, idFlowers, idOrders, Kolichestvo, Price, idService) дають можливість додавати товари до корзини, для можливості їх подальшого придбання, а також переглядати інформацію про товар в кошику.

На рис. 3.13 представлена декомпозиція функції «Адміністрування» системи.

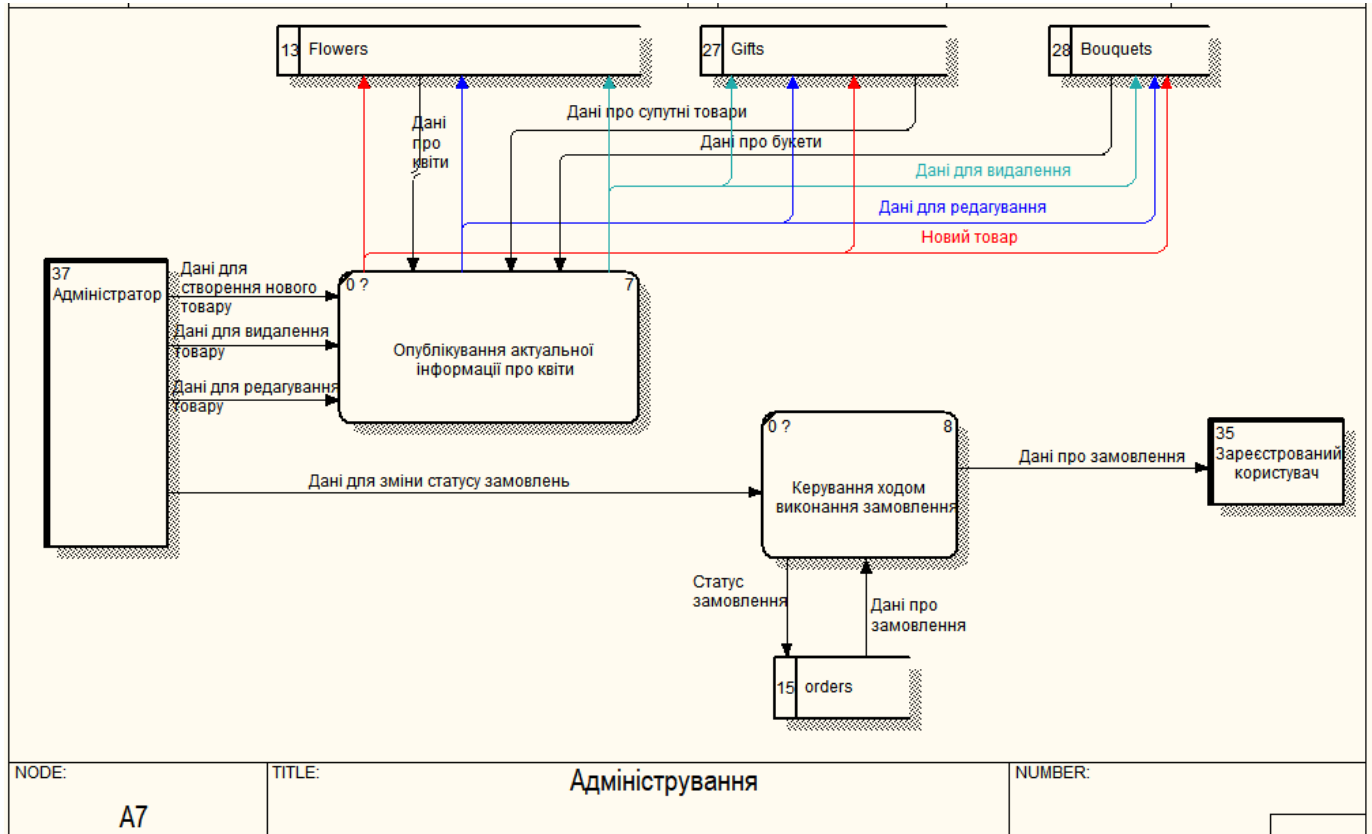


Рисунок 3.13 – Діаграма декомпозиції функції «Адміністрування» системи

Для виконання функції необхідні наступні сховища даних (Data Stores):

- сутність «Flowers» атрибути якої (idFlowers, Name, Description, Price, Quantity, Date_coming, idCategory) дають можливість зчитувати інформацію про квіти, котрі маютья в системі

- сутність «Bouquets» атрибути якої (idBouquet, Name, Description, Price, Quantity, FKidFlBouq, idCategory) дають можливість зчитувати інформацію про букети, їх кількість котрі маються в системі
- сутність «Gifts» атрибути якої (idGifts, Name, Description, Price, Quantity, Date_coming, idCategory) дають можливість зчитувати інформацію про супутні товари, котрі маються в системі
- сутність «orders» атрибути якої (idOrders, idUser, Date, Price, Adress, Telephone, idDostavka, idStatusOrder) дають можливість читати інформацію з таблиць БД, та змінювати її ;

На рис. 3.14 представлена декомпозиція функції «Аналіз даних» системи.

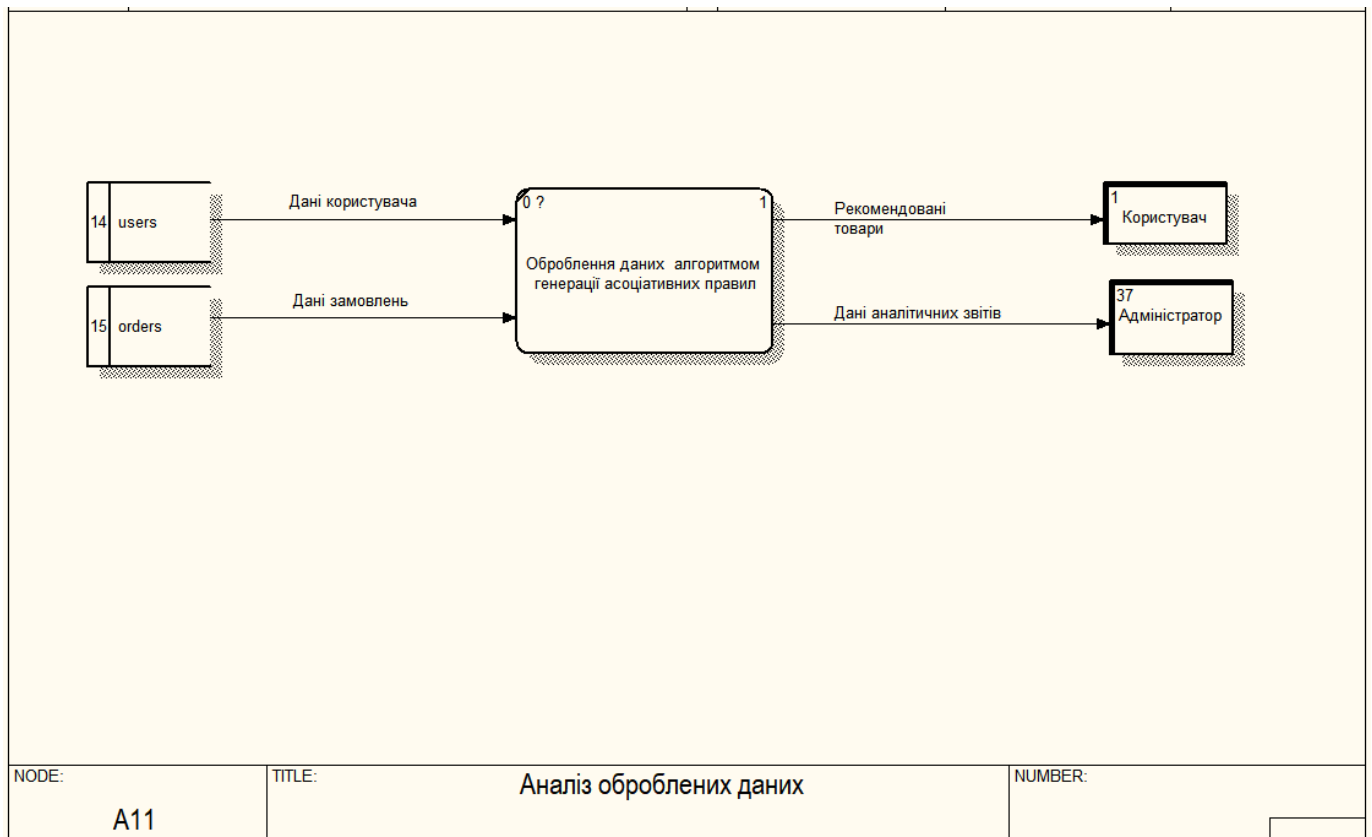


Рисунок 3.13 - Діаграма декомпозиції функції «Аналіз даних» системи

Для виконання функції необхідні наступні сховища даних (Data Store):

- сутність «users», атрибути, якої («Surname», «Name», «Patronymic», «Birthday», «FKidGender» «Login», «Password») дають можливість отримати інформацію про користувача, та використовувати її для аналізу.
- сутність «orders» атрибути якої (idOrders, idUser, Date, Price, Adress, Telephone, idDostavka, idStatusOrder) дають можливість читати інформацію з таблиць БД та використовувати її з метою аналізу даних;

В результаті проведення моделювання потоків даних, були уточнені функціональні вимоги до системи, що розроблюється, а також визначені сутності, що будуть використовуватися в структурі бази даних.

3.4 UML-проектування

Уніфікована мова моделювання (UML) – це уніфікована графічна мова проектування для опису, візуалізації, проектування та документування об’єктно орієнтованих систем [16].

На сьогоднішній день UML є одним з найбільш розповсюджених та популярних методів моделювання бізнес-процесів. Він оснований на схематичних зображеннях різних компонентів програмного забезпечення з метою визначення того, як була спроектована інформаційна система. Саме завдяки графічному представленню людина здатна краще сприймати та розуміти можливі недоліки чи помилки в програмному забезпеченні або бізнес-процесах.

UML знайшла своє призначення в документації бізнес-процесів чи робочих процесів, оскільки вона надає найбільш стандартизований спосіб їх моделювання та широкий набір функцій для підвищення ефективності та зручності в читанні.

3.4.1 Діаграма варіантів використання CRM-системи для визначення вимог до її інтерфейсу

Діаграма варіантів використання (Use Case Diagram) – це графічне представлення всіх акторів, прецедентів та їх взаємодій у системі. У кожній системі, як правило, є головна діаграма прецедентів, котра відображає границі системи (акторів) та основну функціональну поведінку системи (прецеденти) [17].

На рис. 3.1 зображена діаграма варіантів використання розроблюваної інформаційної системи. Вона являє собою певну послідовність дій, котра виконується системою у відповідь на подію, що ініціюється зовнішнім об'єктом (діючим актором).

Варіанти використання, зазвичай, описують типову взаємодію між системою та користувачем (актором). Кожен актор відображає певну роль, котру виконує користувач по відношенню до інформаційної системи.

На діаграмі варіантів використання (рис. 4.1) зображено три діючих актора: незареєстрований користувач, зареєстрований користувач, адміністратор.

Актор «Незареєстрований користувач» – це користувач системи, котрий має обмеження при взаємодії з системою. Він може здійснювати тільки наступні дії: перегляд каталогу (пошук за категорією, перегляд рекомендаційних товарів), авторизація та реєстрація.

Актор «Зареєстрований користувач» – це користувач системи, що має можливість користуватися всіма сервісами системи. Він має дещо розширені можливості користування системою, на відміну від актора «Незареєстрований користувач».

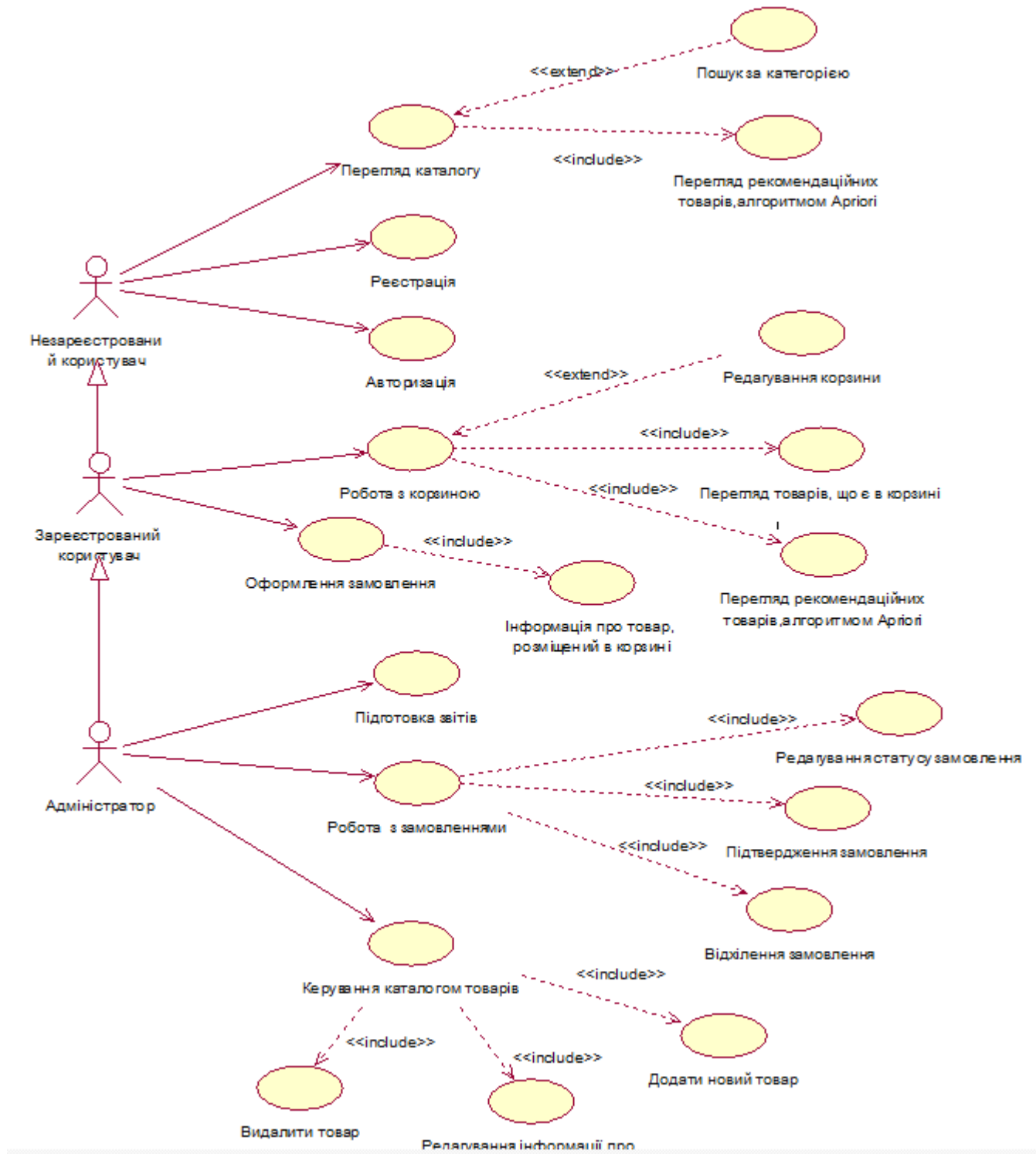


Рисунок 3.15 - Діаграма варіантів використання

Зареєстрований користувач може виконувати ті ж дії з системою, що й незареєстрований користувач, але, додатково : «Робота з корзиною» (додавання, видалення, перегляд товарів), а також «Оформлення замовлення».

Актор «Адміністратор» – це користувач системи, котрий має найбільш повні та розширені права взаємодії з системою. Він може здійснювати підготовку звітів, роботу з замовленнями, керування каталогами товарів

Варіант використання «Перегляд каталогу» – відповідає за функцію ознайомлення користувачів з товарами, котрі містяться в системі для продажу.

Варіант використання: «Перегляд рекомендованих товарів» - відповідає за функцію ознайомлення користувачів системи з найбільш популярними товарами, котрі часто опиняються в кошиках покупок інших користувачів.

Варіант використання «Пошук за категорією» - відповідає за функцію пошуку товарів за певною категорією. Функція використовується користувачем в разі його бажання звужити спектр пошуку товарів, до тих, що належать вибраній категорії

Варіант використання «Реєстрація» – відповідає за функцію реєстрації в системі. Реєстрація дозволить користувачу мати розширений доступ до її сервісів (міг покласти товар до кошика, переглядати кошик, оформляти замовлення).

Варіант використання «Авторизація» – відповідає за процес вводу даних авторизації для того, щоб користувач авторизувався в системі.

Варіант використання «Робота з кошиком» – відповідає за функцію виконання певних операцій зареєстрованого користувача з товаром, що знаходиться в кошику

Варіант використання «Видалення товарів з кошика» - відповідає за функцію видалення товару зареєстрованим користувачем з кошику. Включається, як зіставний компонент до прецеденту «Робота з кошиком»

Варіант використання «Перегляд товарів, що є в кошику» - відповідає за функцію перегляду доданих товарів до кошика. Включається, як зіставний компонент прецеденту «Робота з кошиком».

Варіант використання «Оформлення замовлення» – відповідає за виконання основної бізнес-функції системи, котра відповідає за оформлення замовлення на певний товар.

Варіант використання «Робота з замовленням» – відповідає за виконання функції обробки та управління замовленнями адміністратором системи.

Варіант використання «Підтвердження замовлення» – відповідає за функцію прийняття замовлення. В рамках виконання даного прецеденту відбувається зміна статусу замовлення. Цей варіант використання включається в базову послідовність дій прецедента «Робота з замовленням»

Варіант використання «Відхилення замовлення» – відповідає за функцію відхилення замовлення, в разі, коли воно не підтверджено користувачем. Прецедент входить в базову послідовність дій прецедента «Робота з замовленням».

Варіант використання «Керування каталогом товарів» - відповідає за функцію управління контентом системи. Адміністратор повинен публікувати та оновлювати інформацію до актуальної.

Варіант використання «Додавати новий товар» – відповідає за функцію додання нового товару та супровідної інформації про нього. Прецедент включається в базову послідовність дій прецедента «Керування каталогом товарів».

Варіант використання «Редагувати інформації про товар» – відповідає за функцію редагування інформації про товар. Прецедент включається в базову послідовність дій прецедента «Керування каталогом товарів».

Варіант використання «Видалити товар» - відповідає за функцію видалення товару, якого більше не буде в продажу. Прецедент включається в базову послідовність дій прецедента «Керування каталогом товарів».

Варіант використання «Перегляд рекомендаційних товарів» - відповідає за функцію перегляду користувачем популярних товарів, що виводяться в блоці з рекомендаціями.

4 РОЗРОБКА ІНФОРМАЦІЙНОГО ТА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

4.1 Обґрунтування вибору СУБД

В наш час існують велика кількість систем управління базами даних (СУБД). Серед них найбільш популярними є такі, як: MS SQL Server, MySQL, Postgres, MongoDB.

При виборі СУБД для інформаційної системи, особлива увага приділялась до здатностей масштабування, надійності, продуктивності, легкості інтегрування, та підтримуваних сучасних мов програмування (Java, Python і т.д.).

Враховуючи основні критерії до СУБД була вибрана PostgreSQL. Вона підходить до вибраної мови програмування та технологій, що будуть застосовуватись для розробки проекту.

PostgreSQL – це об'єктно-реляційна система управління базами даних (ORDBMS) з відкритим вихідним кодом [18]. Вона підходить як для великих та високонавантажених, так і для малих проектів, пропонує велику кількість різноманітних інструментів розробки, має багатий функціонал.

Postgres включає такі функції, як успадкування таблиці та перевантаження функцій. СУБД також реалізує мультиверсійний паралельний контроль (MVCC) без блокування читання, підтримує паралельні плани запитів, які можуть використовувати декілька процесорів, ядер.

До основних переваг PostgreSQL відносять:

- відкритий вихідний код (СУБД в вільному доступі);
- масштабованість (здатність працювати з великими обсягами даних);
- легкість застосування;
- надійність (повністю відповідає принципам ACID);
- можливість розширення;

- продуктивність (швидкість обробки запитів).
- безпека даних;
- підтримка JSON (дозволяє з'єднуватись з іншими джерелами даних, наприклад NoSQL).

4.2 Проектування БД

База даних, котра буде проектуватися, повинна задовольняти всім поставленим до неї вимогам та дотримуватися правил нормалізації. Розробка БД розпочинається, в першу чергу, з концептуального проектування. Таке проектування являє собою аналіз та збір вимог до даних, що досягається шляхом дослідження предметної області та її інформаційної структури; виділення основних сутностей системи та зв'язків між ними; опис обмежень цілісності, тобто вимог для допустимих значень даних; моделювання та інтеграція всіх представлень.

Концептуальна база даних проектується не орієнтуючись на певну СУБД, вона описує модель даних, опираючись на представлення предметної області кінцевим користувачем. Після створення концептуальної моделі БД відбувається логічне (створення схеми БД на основі реляційної моделі даних) та фізичне проектування (створення схеми БД для СУБД PostgreSQL).

Проектування бази даних CRM системи з продажу квітів проводилося з застосуванням інструментального CASE-засобу All Fusion ErWin Data Modeler. На рисунку 5.1 зображена ER-діаграма розроблюваної бази даних.

Предметною областю системи являється діяльність фірми по продажу квітів.

Спроектowana схема бази даних містить наступні сутності системи:

- сутність «Flowers» – містить інформацію про назву квітки латинською мовою, сорт, сімейство до якої вона належить, висоту стебла та дату його зрізу;

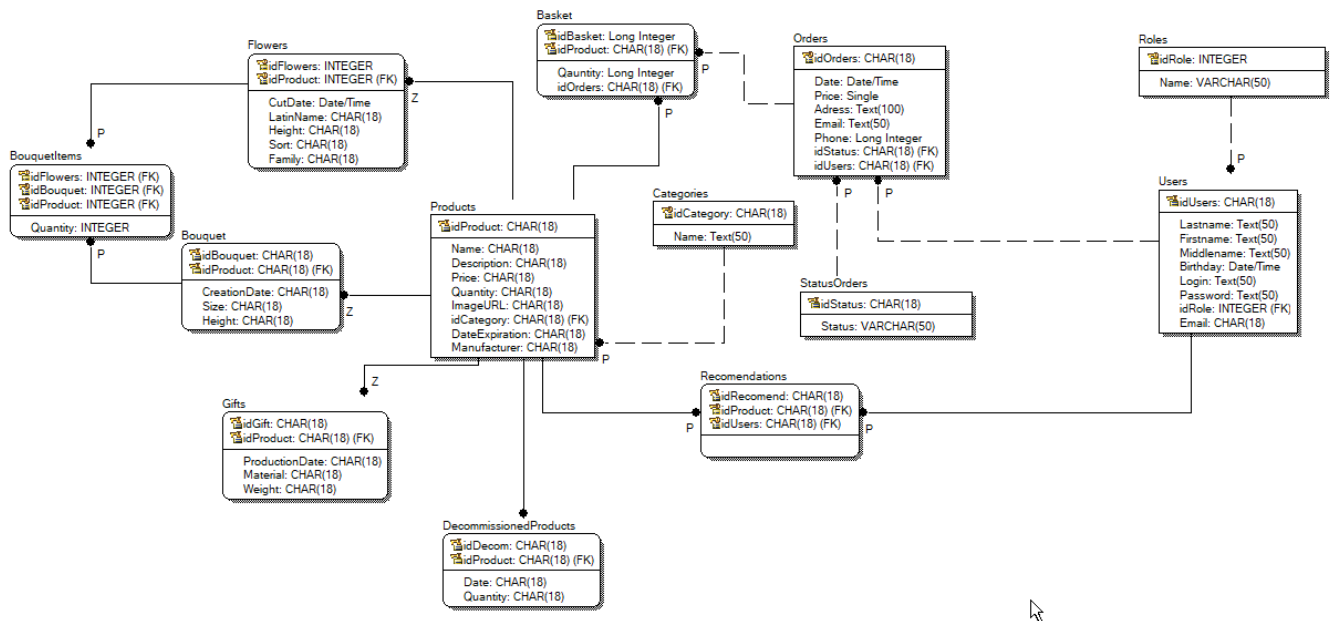


Рисунок 5.1 – ER-діаграма бази даних

- сутність «Bouquets» – містить інформацію про дату створення букету, його розмір (малий, середній чи великий) та висоту всієї квіткової композиції;
- сутність «BouquetItems» – містить в собі інформацію про квіти, що містяться в букеті, та їх кількість;
- сутність «Gifts» – містить в собі інформацію про дату створення товарів, матеріал (іграшок), вагу;
- сутність «Products» – містить в собі узагальнену інформацію, що є спільною для сутностей «Flowers», «Bouquet» та «Gifts», а саме: ім'я, опис, ціна, кількість товарів, URL зображення, термін придатності, виробник та категорія товару;
- сутність «Category» - містить в собі інформацію про всі існуючі категорії в системі;

- сутність «Basket» - містить в собі інформацію про товари, що знаходяться в ній та їх кількість;
- сутність «Orders» - містить в собі інформацію про замовлення, його статус, коштовність, адресу доставки, електронну пошту, користувача, дату створення, телефонний номер;
- сутність «StatusOrders» - містить в собі інформацію про статус в якому знаходиться замовлення (в процесі створення, в обробці, прийнятий, відхилений, закритий);
- сутність «Users» - містить в собі інформацію про авторизованих користувачів системи;
- сутність «Roles» - містить в собі інформацію про ролі користувача системи (адміністратор чи користувач);
- сутність «Recommendations» - містить в собі інформацію про товари, що будуть запропоновані користувачам системи;
- сутність «DecomissionedProducts» - містить в собі інформацію про товари, в котрих вийшов термін придатності. В ньому міститься інформація про дату потрапляння до утилізації, кількість товарів, що треба видалити, та можливість написати коментар до

Таблиця 5.1 – Сутності логічної моделі даних системи

№	Найменування сутності	Найменування атрибута	Тип даних (домен)	Призначення
1.	Products	idProduct	Лічильник, цілочисловий	Первинний ключ
		Name	Текст, довжиною в 100 символів	Найменування товару

Продовження таблиці 5.1

№	Найменування сутності	Найменування атрибута	Тип даних (домен)	Призначення
2.		Description	Текст, довжиною в 200 символів	Опис товару
		Price	Десятичний тип даних з фіксованою крапкою	Ціна за одиницю товару
		Quantity	Цілочисловий	Кількість товару (шт.)
		ImageURL	Текст	URL адреса знаходження картинки
		idCategory	Цілочисловий	Зовнішній ключ
		DateExpiration	Дата	Термін придатності
		Manufacturer	Текст, довжиною в 200 символів	Виробник
3.	Flowers	idFlower	Лічильник, цілочисловий	Первинний ключ
		idProduct	Цілочисловий	Зіставний первинний ключ
		CutDate	Дата	Дата зрізу стебла
		LatinName	Текст, довжиною в 200 символів	Ім'я латинською
		Height	Десятичний тип даних з фіксованою крапкою	Висота стебла

Продовження таблиці 5.1

№	Найменування сутності	Найменування атрибута	Тип даних (домен)	Призначення
		Sort	Текст, довжиною в 200 символів	Сорт
		Family	Текст, довжиною в 150 символів	Сімейство
4.	Buckets	idBouquet	Лічильник, цілочисловий	Первинний ключ
		idProduct	Цілочисловий	Зіставний ключ
		CreationDate	Дата	Дата створення букету
		Size	Текст, довжиною в 150 символів	Розмір букету
		Height	Десятичний тип даних з фіксованою крапкою	Висота букету
5.	Gifts	idGift	Лічильник, цілочисловий	Первинний ключ
		idProduct	Цілочисловий	Зіставний ключ
		ProductionDate	Дата	Дата виробництва
		Material	Текст, довжиною в 100 символів	Матеріал, з якого створено
		Weight	Текст, довжиною в 100 символів	Вага товару
6.	BouquetItems	idFlower	Цілочисловий	Зіставний ключ
		idBouquet	Цілочисловий	Зіставний ключ

Продовження таблиці 5.1

№	Найменування сутності	Найменування атрибута	Тип даних (домен)	Призначення
		idProduct	Цілочисловий	Зіставний ключ
7.	Users	idUser	Лічильник, цілочисловий	Первинний ключ
		FirstName	Текст, довжиною в 50 символів	Прізвище користувача
		LastName	Текст, довжиною 50 символів	Ім'я користувача
		MiddleName	Текст, довжиною 50 символів	По-батькові користувача
		Birthday	Дата	День народження користувача
		Login	Текст, довжиною в 15 символів	Логін для входу на сайт
		Password	Текст, довжиною в 20 символів	Пароль для входу на сайт
		idRole	Текст, довжиною в 15 символів	Статус користувача
		Email	Текст довжиною в 50 символів	Адрес електронної пошти
8.	Category	idCategory	Лічильник, цілочисловий	Первинний ключ
		Name	Текст, довжиною в 50 символів	Назва категорії

Продовження таблиці 5.1

№	Найменування сутності	Найменування атрибута	Тип даних (домен)	Призначення
9.	Role	idRole	Лічильник, цілочисловий	Первинний ключ
		Name	Текст довжиною в 50 символів	Роль користувача при авторизації
10.	StatusOrders	idStatus	Лічильник, цілочисловий	Первинний ключ
		Status	Текст, довжиною в 35 символів	Статус замовлення
11.	Basket	idBasket	Лічильник, цілочисловий	Первинний ключ
		idProduct	Цілочисловий	Зіставний первинний ключ
		idOrders	Цілочисловий	Зовнішній ключ
		Quantity	Цілочисловий	Кількість одиниць товару
12.	Orders	idOrders	Лічильник, цілочисловий	Первинний ключ
		Date	Дата	Дата оформлення замовлення
		Price	Десятичний тип	Ціна замовлення
		Address	Текст, довжиною в 50 символів	Адреса доставки замовлення

№	Найменування сутності	Найменування атрибута	Тип даних (домен)	Призначення
		Email	Текст, довжиною в 40 символів	Електронна адреса користувача
		Phone	Цілочисловий	Телефон користувача
		idStatus	Цілочисловий	Зовнішній ключ
		idUser	Цілочисловий	Зовнішній ключ
13.	Recomendations	idRecomend	Лічильник, цілочисловий	Первинний ключ
		idProduct	Цілочисловий	Зіставний ключ
		idUser	Цілочисловий	Зіставний ключ
14.	DecommissionedProducts	idDecom	Лічильник, цілочисловий	Первинний ключ
		idProduct	Цілочисловий	Зіставний ключ
		Date	Дата	Дата потрапляння товару на утилізацію
		Quantity	Цілочисловий	Кількість одиниць товару

4.3 Вибір мови програмування

Інформаційна система була розроблена на мові програмування Java з використанням фреймворку Apache Spark.

Java – це об’єктно-орієнтована мова програмування, що має сувору типізацію, необхідну для написання різних програмних продуктів та найбільший ступінь переносимості програм серед усіх існуючих мов програмування [19].

На відміну від C#, Java є мовою програмування загального призначення, що є паралельним, об’єктно-орієнтованим і розробленим спеціально, для того щоб було як можна менше залежностей реалізації, а також відзначаються стабільністю в порівнянні з програмами на інших мовах програмування.

Особливістю Java в порівнянні з іншими мовами програмування є забезпечення високої продуктивності програмування, а також можливості створювати програмні продукти, що будуть здатні працювати на різних платформах.

Основними перевагами Java є:

- незалежність від платформи й архітектури;
- об’єктно-орієнтованість;
- гнучкість, надійність, безпечність;
- можливість розробки через тестування (TDD);
- підтримка багатопоточності;
- потужні стандартні бібліотеки;
- багатопоточність.

Мова Java здатна забезпечувати ефективну підтримку відкритих та комерційних баз даних, в таких СУБД, як: MS SQL Server, Oracle, MySQL, PostgreSQL.

Розробка алгоритму асоціативних правил була реалізована за використання мови програмування Python та інтегрування його до інформаційної системи, що написана на Java.

Засоби, що також використовувались під час розробки:

- набір бібліотек Anaconda 4.3.3;
- середовище розробки PyCharm (Community Edition);

ВИСНОВКИ

В рамках виконання атестаційної роботи був розроблений та впроваджений CRM метод на основі генерації асоціативних правил для інформаційної системи з продажу квітів.

Під час аналізу предметної області була визначена сфера застосування CRM систем в електронній комерції, проведено аналіз вже існуючих інформаційних систем, а також визначено специфічні особливості для предметної області, що розглядається. В результаті проведеного аналізу була розроблена постановка задачі.

Для визначення того, який підхід найбільше підходить для розробки методу CRM, було проведено аналіз та дослідження наступних існуючих методів побудови рекомендацій: фільтрації на основі контексту/вмісту, колаборативна фільтрація, гібридна фільтрація та метод генерації асоціативних правил. В результаті дослідження було визначено, що використання методу генерації асоціативних правил є найбільш доцільним для специфіки розглянутої предметної області.

Для розробки методів CRM системи з урахуванням специфіки предметної області був застосований алгоритм асоціативних правил Argioi. Розроблений метод опирається на особливості предметної області. Він враховує специфіку товару (низький термін придатності) та уникає рекомендацій з несумісними товарами. Після розробки методу були описані основні бізнес-процеси та визначені функціональні вимоги до інформаційної завдяки методології IDEF0.

Для визначення того, які інформаційні потоки будуть проходити між системою та зовнішніми сутностями була розроблена модель потоків даних CRM системи, а для визначення взаємодії користувачів з системою – Use Case діаграма.

Під час розробки інформаційного забезпечення був проведений обґрунтований аналіз вибору СУБД та спроектована база даних. Метод CRM системи був розроблений на мові програмування мова Java з використанням фреймворку Spark.

ПЕРЕЛІК ПОСИЛАНЬ

1. Закон України "Про підприємництво" [Електронний ресурс].– Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/show/698-12>.
2. Закон України «Про власність» [Електронний ресурс] – Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/main/697-12>.
3. Закон України "Про охорону праці" [Електронний ресурс] – Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/main/2694-12>.
4. Закон України "Про оплату праці" [Електронний ресурс] – Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/main/108/95-%D0%B2%D1%80>.
5. Що таке електронна комерція? [Електронний ресурс]. – 2020. – Режим доступу до ресурсу: <https://www.interkassa.com/ua/blog/chto-takoe-elektronnaya-kommerciya-e-commerce-dlya-nachinayushchih/>.
6. CRM система, её компоненты, классификация [Електронний ресурс] – Режим доступу до ресурсу: <http://websekretar.chizh.ua/?p=548>.
7. Метрики сходства и расстояния для науки о данных и машинного обучения [Електронний ресурс] – Режим доступу до ресурсу: <https://www.machinelearningmastery.ru/similarity-and-distance-metrics-for-data-science-and-machine-learning-e5121b3956f8/>.
8. Item-based collaborative filtering словами [Електронний ресурс] – [http://www.cs.carleton.edu/cs_comps/0607/recommend/recommender/item based.html](http://www.cs.carleton.edu/cs_comps/0607/recommend/recommender/item%20based.html)
9. Su X., Khoshgoftaar T.M. A Survey of Collaborative Filtering Techniques; Advances in Artificial Intelligence, 2009.
10. Гончаров М. Data Mining: Рекомендательные системы [Электронный документ] (http://kek.ksu.ru/EOS/WM/50_132-670.pdf)
11. Н. Шубіна. Використання асоціативних правил для аналізу кошика покупок магазину квітів // V Всеукраїнській науково-практична конференція «Перспективні напрямки сучасної електроніки, інформаційних і комп'ютерних

систем». Зб. матеріалів конференції. С. 1. «Перспективні напрямки сучасної електроніки, інформаційних і комп'ютерних систем» – Дніпро: ДНУ. 2020. – С. 55-56.

12. Sitnikov D. Assessment of extended aggregated association rules / Sitnikov D., Ryabov O., Titova O., Kovalenko A. // Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies, DESSERT'2018, 24 27 May, 2018, Kyiv, Ukraine, page 95 99. (IEEE Xplore, Scopus)

13. Нейронные сети [Електронний ресурс] – Режим доступу до ресурсу: <https://future2day.ru/nejronnye-seti/>.

14. Похилько А. Ф. CASE-технология моделирования процессов с использованием средств BPWin и ERWin учебное пособие / А. Ф. Похилько, И. В. Горбачев. – Ульяновск: УЛГТУ, 2008. – 120 с.

15. Стандарти IDEF. Методології SADT, DFD, BPMN, ARIS, UML. Порівняльний аналіз методологій моделювання. [Електронний ресурс] – Режим доступу до ресурсу: <http://um.co.ua/8/8-16/8-162794.html>.

16. Общая характеристика языка UML [Електронний ресурс] – Режим доступу до ресурсу: <http://www.informicus.ru/default.aspx?SECTION=6&id=73&id=2>.

17. Київська К. І. Побудова діаграми прецедентів (варіантів використання) [Електронний ресурс] / К. І. Київська. – 2018. – Режим доступу до ресурсу: http://org2.knuba.edu.ua/pluginfile.php/58191/mod_resource/content/1/2018metoda_labr_1.html.

18. PostgreSQL [Електронний ресурс]. – 1. – Режим доступу до ресурсу: <https://coderlessons.com/tutorials/bazy-dannykh/uchebnik-postgresql/1-chto-takoe-postgresql>.

19. Объектно-ориентированное программирование в Java и Python [Електронний ресурс] – Режим доступу до ресурсу: <https://habr.com/ru/post/455796/>