

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет кібербезпеки, комп'ютерної та програмної інженерії

Кафедра комп'ютерних інформаційних технологій

ДОПУСТИТИ ДО ЗАХИСТУ

Завідувач кафедри

_____ Савченко А.С.

“ _____ ” _____ 2020 р.

ДИПЛОМНА РОБОТА
(ПОЯСНЮВАЛЬНА ЗАПИСКА)

ВИПУСКНИКА ОСВІТНЬОГО СТУПЕНЯ
“МАГІСТРА”

ЗА СПЕЦІАЛІЗАЦІЄЮ “ІНФОРМАЦІЙНІ УПРАВЛЯЮЧІ СИСТЕМИ ТА
ТЕХНОЛОГІЇ (ЗА ГАЛУЗЯМИ)”

Тема: “Технологія застосування методів Data Mining при розробці
рекомендаційних експертних систем для підбору стилю”

Виконавець: Литовченко Олександра Вадимівна

Керівник: професор Воронін Альберт Миколайович

Нормоконтролер: _____ Райчев І.Е.

Київ 2020

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет кібербезпеки, комп'ютерної та програмної інженерії

Кафедра Комп'ютерних інформаційних технологій

Галузь знань, спеціальність, спеціалізація: 12 “Інформаційні технології”, 122 “Комп'ютерні науки”, “Інформаційні управляючі системи та технології (за галузями)”

ЗАТВЕРДЖУЮ

Завідувач кафедри

Савченко А.С.

“ ” 2020 р.

ЗАВДАННЯ

на виконання дипломної роботи студентки

Литовченко Олександри Вадимівни

1. Тема роботи: “Технологія застосування методів Data Mining при розробці рекомендаційних експертних систем для підбору стилю”

Затверджена наказом ректора від “ 02 ” жовтня 2020р за № 1891/ст .

2. Термін виконання роботи: з 05.10.2020р. до 31.12.2020р.

3. Вихідні дані до роботи: база правил та знань, рекомендаційна експертна система, задача підбору стилю.

4. Зміст пояснювальної записки: 1) аналіз предметної області, дослідження стану задачі підбору стилю, проведення огляду існуючих аналогів ПЗ, публікацій та досліджень, дослідження існуючих методів DM; 2) дослідження загальних рис та складових поняття ЕС; 3) проведення аналізу методів для проектування та розробки структури ЕС підбору стилю, аналіз їх переваг та недоліків; 4) розробка моделі експертної системи з використанням методів DM для поліпшення точності прийняття рішень; 5) розробка програмного застосунку експертної системи підбору стилю, отримання результатів та висновків.

5. Перелік обов'язкового ілюстративного матеріалу: рисунки, схеми, діаграми
слайди презентації доповіді у PowerPoint.

6. Календарний план-графік

<i>№ з/п</i>	<i>Завдання</i>	<i>Термін виконання</i>	<i>Підпис керівника</i>
1.	Визначення та узгодження теми МДР з дипломним керівником	05.10.20 – 06.10.20	
2	Складання плану на весь період виконання МДР та формування структури розділів	07.10.20 – 10.10.20	
3	Збір, аналіз та дослідження матеріалів для МДР	11.10.20 – 15.10.20	
4	Формування та оформлення першого розділу МДР	16.10.20 – 26.10.20	
5	Формування та оформлення другого розділу МДР	27.10.20 – 10.11.20	
6	Проектування та розробка програмного застосунку згідно із завданням МДР	11.11.20 – 25.11.20	
7	Формування та оформлення третього розділу МДР	25.11.20 – 31.11.20	
8	Формування та оформлення четвертого розділу МДР	31.11.20 – 08.12.20	
9	Оформлення та друк пояснювальної записки МДР	09.12.20 – 15.12.20	
10.	Створення доповіді та слайдів до неї. Підписання необхідних документів. Підготовка до захисту МДР	15.12.20 – 21.12.20	

7. Консультація з окремого(мих) розділу(ів) роботи:

Розділ	Консультант (посада, П.І.Б.)	Дата, підпис	
		Завдання видав	Завдання прийняв

8. Дата видачі завдання _____

Керівник дипломної роботи _____ **Воронін А.М.**

Завдання прийняв до виконання _____ **Литовченко О.В.**

(підпис випускника)

(ПІБ)

РЕФЕРАТ

Пояснювальна записка до дипломної роботи "Технологія застосування методів Data Mining при розробці рекомендаційних експертних систем для підбору стилю " містить: 88 сторінок з додатками, 41 рисунок, 33 бібліографічні посилання, 1 додаток 13 сторінок.

Актуальність магістерської роботи визначається необхідністю дослідження та розробки технології пошуку і накопичення знань для прийняття рішення в задачі підбору стилю, яка б дозволяла зручно зберігати і передавати інформацію користувачеві. Технологія має полегшувати вибір рішення та сприяла б більш точному результату. Створення такої технології обумовлене недостатньою кількістю досліджень і розробок в сфері підбору стилю.

Метою є дослідження застосування технології Data Mining при проектуванні експертних систем рекомендаційного типу та поліпшення процесу прийняття рішень в задачі підбору стилю на основі методів Data Mining.

Об'єктом є технологія Data Mining в задачах, що стосуються побудови експертних систем.

Предметом є методи сімейства Data Mining.

Методами дослідження є методи емпіричного дослідження (порівняння, вимірювання, моделювання) та формалізація.

Результати магістерської роботи: створено повністю працездатний програмний застосунок на основі задачі підбору стилю експертної системи із застосуванням методів Data Mining, на технологіях .NET, що містить тестування для прийняття рішення користувачем та графічне відображення динаміки результатів.

Ключові слова: ЕКСПЕРТНА СИСТЕМА, ПІДБОР СТИЛЮ, МОДЕЛЮВАННЯ, ПРОЕКТУВАННЯ, ДЕРЕВО РІШЕНЬ, DATA MINING, ПРИЙНЯТТЯ РІШЕНЬ.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ, ТЕРМІНІВ	7
ВСТУП.....	8
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	10
1.1. Аналіз стану задачі підбору стилю	10
1.2. Огляд існуючих аналогів ПЗ	11
1.3. Аналіз літератури, публікацій та досліджень	16
1.4. Аналіз існуючих методів та алгоритмів Data Mining.....	21
1.5. Постановка задачі	25
РОЗДІЛ 2. АНАЛІЗ МЕТОДІВ ТА ПІДХОДІВ ДЛЯ ПРОЕКТУВАННЯ ТА РОЗРОБКИ СТРУКТУРИ ЕС.....	27
2.1. Загальна характеристика поняття ЕС	27
2.2. Продукційний метод представлення знань	31
2.3. Проектування продукційної моделі бази знань прямого виводу	33
2.4. Класифікація стилю.....	35
РОЗДІЛ 3. ТЕХНОЛОГІЯ ЗАСТОСУВАННЯ МЕТОДІВ DATA MINING	44
3.1. Метод класифікації.....	44
3.2. Метод дерева рішень	45
3.3. Метод візуалізації даних	46
3.4. Застосування методів Data Mining при розробці ЕС підбору стилю	47
РОЗДІЛ 4. РОЗРОБКА ПРОГРАМНОГО ЗАСТОСУНКУ	53
4.1. Обґрунтування вибору платформи для розробки	53
4.2. Обґрунтування та вибір базових програмних засобів, технологій розробки	53
4.3. Опис програмної реалізації.....	54
4.4. Програмна реалізація методу дерева рішень	56
4.5. Алгоритм роботи експертної системи	61
4.6. Опис інтерфейсу користувача та результат роботи програми.....	62
ВИСНОВКИ	71
ПЕРЕЛІК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ	72
ДОДАТОК А	75

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ, ТЕРМІНІВ

БЗ – база знань;

БП – база правил;

ІЗ – інженер знань;

ІС – інформаційна система;

ІАД – інтелектуальний аналіз даних;

ІСАД – інтелектуальні системи аналізу даних;

ЕЗ – експертні знання;

ЕС – експертна система;

ОС – операційна система;

ПЗ – програмне забезпечення;

ПС – продукційні системи;

ІТ – інформаційні технології;

DM – Data Mining;

WPF – Windows Presentation Foundation.

ВСТУП

Життя мобільне та повне різноманітних подій, які потребують певного зовнішнього вигляду. На сьогоднішня відома невелика кількість подібного роду простих проте корисних систем.

Необхідно заповнити нішу та створити конкуренцію відомим брендам, які забрали велику частку світового ринку та диктують свої правила. Використання методів Data Mining в змозі створити умови для розробки ЕС підбору стилю з правильною базою знань.

Використання штучного інтелекту, нейронних мереж, наприклад генетичні алгоритми при самій системі не використовуються, оскільки вони необхідні для налаштування більш складних параметрів.

Якщо будуть застосовані такі алгоритми наприклад для створення бази правил, замість експертів, то це зруйнує саму ідею ЕС.

Автоматизована система генетичними алгоритмами дасть більш точні та кращі результати проте важко обрати грамотні параметри/функції/критерії корисності, а час буде витрачено.

У обраній предметній сфері не отримуються необхідні правила, оскільки існує велика залежність від людського фактору – власні смаки, переваги, досвід, почуття стилю.

«Машини» поки ще не розуміють творчість та емоційний стан. Наприклад в медичній сфері все по іншому, існують певні хвороби, симптоми та методи лікування, прогнозування, ця сфера потребує більш точних однозначних знань, аніж підбор стилю.

Актуальність магістерської роботи визначається необхідністю дослідження та розробки технології пошуку і накопичення знань для прийняття рішення в задачі підбору стилю, яка б дозволяла зручно зберігати і передавати інформацію користувачеві. Технологія має полегшувати вибір рішення та сприяла б більш точному результату. Створення такої технології обумовлене недостатньою кількістю досліджень і розробок в сфері підбору стилю.

Саме тому робота присвячена дослідженню та застосуванню методів Data Mining – дерева рішень та класифікації, для візуального представлення динаміки результатів рішення користувачеві на основі рекомендаційною ЕС підбору стилю.

Експертна система здатна запропонувати рішення або рекомендації щодо конкретних проблем в обраній предметній області.

Метою є дослідження застосування технології Data Mining при проектуванні експертних систем рекомендаційного типу та поліпшення процесу прийняття рішень в задачі підбору стилю на основі методів Data Mining.

Об'єктом є технологія Data Mining в задачах, що стосуються побудови експертних систем.

Предметом є методи сімейства Data Mining.

В ході виконання магістерської дипломної роботи були використані такі *методи*: метод прямого виводу, продукційний метод представлення знань (для розробки структури ЕС) та методи Data Mining (дерева рішень, візуалізації, класифікації для підвищення ефективності і точності роботи програмного засобу).

Для досягнення вказаної мети необхідно виконати *наступні завдання*:

- 1) провести аналіз предметної області, а саме дослідити стан задачі підбору стилю, провести огляд існуючих аналогів ПЗ, публікацій та досліджень, дослідити існуючі методи Data Mining;
- 2) дослідити загальні риси та складові поняття експертної системи;
- 3) провести аналіз методів та підходів для проектування та розробки структури ЕС підбору стилю, проаналізувати їх переваги та недоліки;
- 4) розробити модель експертної системи з використанням методів Data Mining для поліпшення точності прийняття рішень;
- 5) розробити програмний застосунок експертної системи підбору стилю, проаналізувати отримані результати роботи.

РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1. Аналіз стану задачі підбору стилю

Висока насиченість інформаційного середовища та ризик прийняття неправильних рішень в існуючих умовах підвищують актуальність інформаційних технологій як засобу підтримки прийняття рішень. Одяг швидко стає високотехнологічною промисловістю через швидкий прогрес в технологіях.

Одяг – невід’ємна частина сьогодення. Формування стилю людини забезпечує комфорт життя. Проте, не кожна людина наважується на зміни, та не кожна в змозі зробити такі зміни досконалими.

Одній жінці більш до вподоби підбори, а інша не вміє на них ходити. Тому вміння признати індивідуальність породжує легкість та комфорт. Треба вміти враховувати багато факторів як власні побажання, смаки, зовнішній вигляд, заплановані зустрічі, навіть настроїв.

У підбраному стилі можна підкреслити фігуру, показати себе з іншої сторони, показати свої досконалості. Стиль – гармонія зовнішнього вигляду з вашої унікальністю та характером.

Системи, які базуються на знаннях найбільше підходять для вирішення задач, де необхідні формальні розсуди – тобто рішенням такої проблеми є застосунок у вигляді рекомендаційної експертної системи, який представляє собою тестування і рекомендації стилю. Якщо людина не знає який образ їй личить в конкретній ситуації, ЕС підбору стилю рекомендуватиме декілька варіантів, які підходять найбільше.

Кафедра КІТ (47)				НАУ 20 13 79 000 ПЗ			
Виконав	Литовченко О.В.			Аналіз предметної області	Літера	Аркуш	Аркушів
Керівник	Воронін А.М.					10	17
Консульт							
Н.контр.	Райчев І.Е.					УС-211М	122
							10

Для удосконалення роботи такого застосунку необхідно використання методів Data Mining (наприклад: підходять такі методи як дерева рішень, методи класифікації, методи візуалізації даних, логічні методи), щоб результати були максимально точними, швидко отримані, порівняно з класичним підходом вирішення задачі.

1.2. Огляд існуючих аналогів ПЗ

При розробці програмного забезпечення варто спершу проаналізувати вже готові рішення. Це допоможе скласти уяву про те, як такий продукт повинен виглядати та якою має бути його зовнішня та внутрішня реалізація.

1) **Looklet.** Віртуальна студія стилю. На цьому сайті англомовного сервісу користувачу дають змогу одягнути одну з 36-ти моделей на свій смак. Налічуються сотні типів та видів одягу різних виробників, марок, від звичайних футболок, штанів до найсучасніших світшотів, светрів, худі в поєднанні з безліччю найрізноманітніших аксесуарів [1].

Looklet має на меті подолати розрив між технологіями та модою, а також змушує бізнес та споживачів переосмислити модні образи.

Looklet поєднує високоякісні динамічні зображення у простому та масштабованому процесі, ідеально підходить для нового світу онлайн-моди.

Після створення образу надається можливість його зберегти або використати сумісно з іншими користувачами веб-застосунку Looklet (див. рис.1.1).

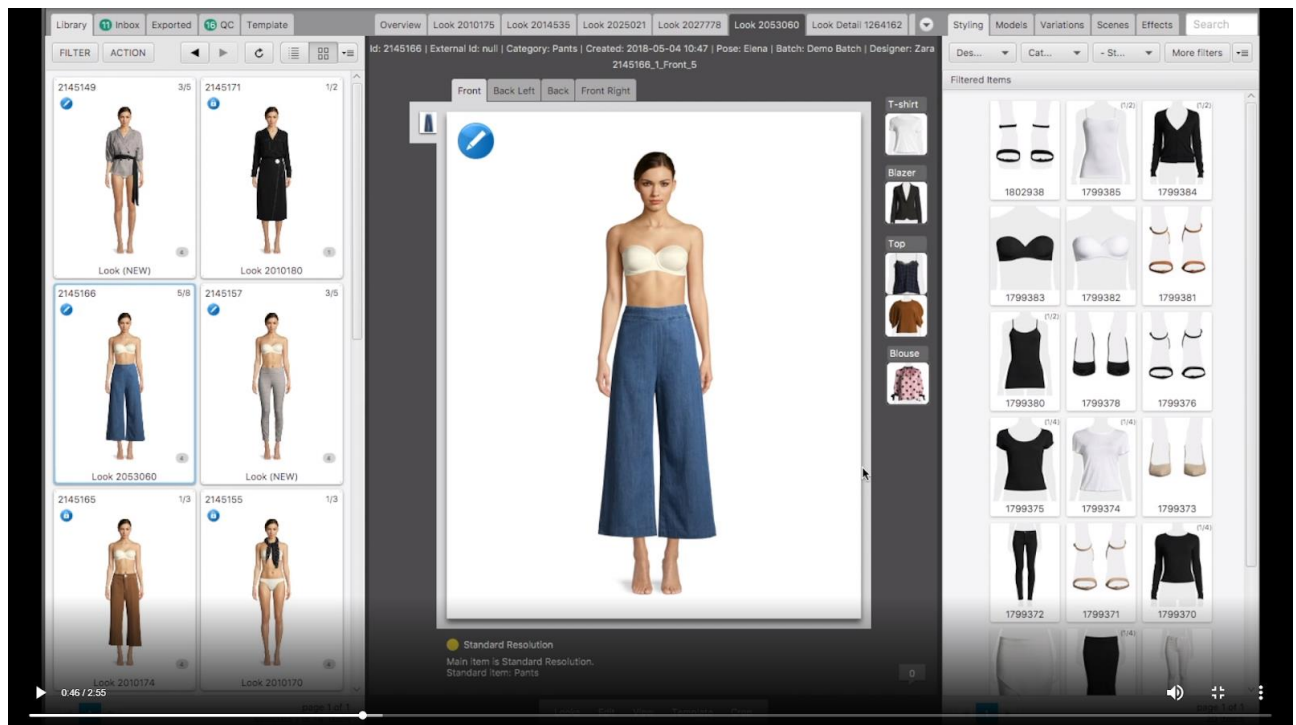


Рис. 1.1. Інтерфейс веб-застосунку «Looklet»

2) **Tesco**. Tesco запустила віртуальну примірювальну, яка дозволяє користувачам створювати 3D-версії різних нарядів «F&F» [2].

Tesco є першою мережею роздрібної торгівлі, яка використовує цю онлайн-технологію. Інтернет-примірювальна дозволяє клієнтам визначитися з підбором одягу. Клієнти потім зможуть легко додати те, що сподобалось в кошик і купити. На перший погляд звичайний інтернет-магазин. Можна поділитися посиланням одягу з друзями на Facebook, а також спробувати різні образи і зачіски (див. рис.1.2).

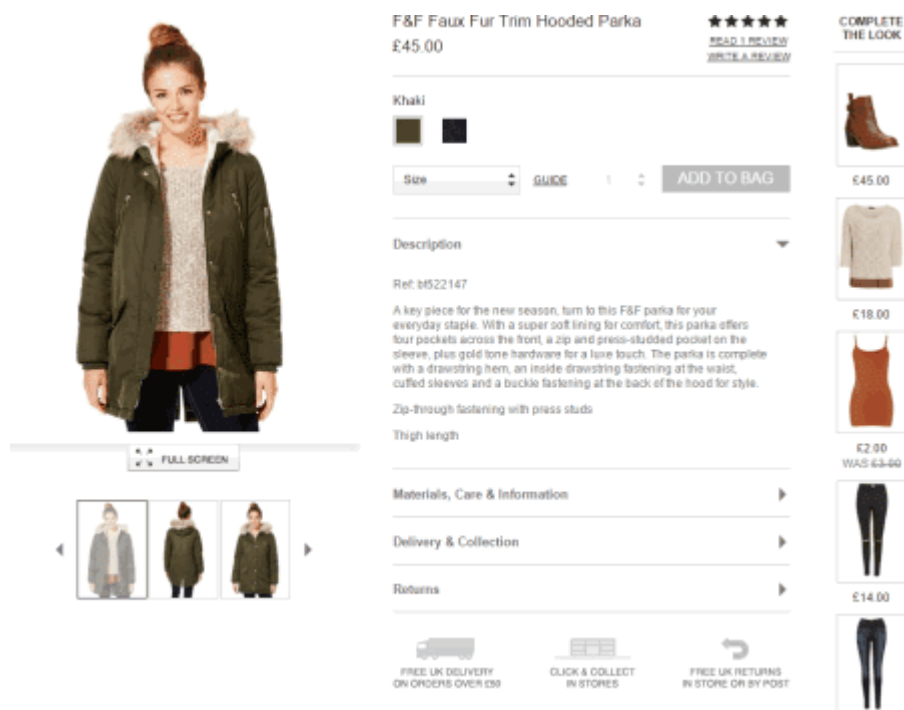


Рис. 1.2. Віртуальна студія одягу Tesco

3) **Fits.me.** У віртуальній примірювальній естонської компанії Fits.me можна підібрати одяг онлайн на віртуальних манекенах, ввівши особисті параметри фігури [3]. Сервіс допоможе підібрати розмір і повідомить про можливі проблеми (див. рис.1.3).



Рис. 1.3. Веб-застосунок «Fits.me»

4) **Polivore.com.** Сервіс, який відомий багатьом шанувальницям моди та стилю, який надає можливість швидко, за лічені хвилини створити неперевершений колаж [4].

За допомогою своєї технології та методів (особливо вузацізації), веб-сервіс настільки просто та швидко створює картинки як в популярних сучасних журналах.

Більш того, є можливість зберігати свої роботи та завантажувати їх, необхідно лише авторизуватися. Реєстрація безкоштовна.

У галереї ви знайдете одяг безлічі відомих брендів за категоріями, взуття, сумки, прикраси, косметику та ін.

Також можна скористатися для оформлення різними елементами декору: принти із зображенням людей, барвисті фони, текст, яскраві палітри, рамки, журнальні статті, символи, малюнки, ефекти, шаблони і навіть музика.

Галерея постійно поповнюється моделями зі свіжих колекцій.

Також можна поставити галочку в рядку «Quick share with friends» (Швидко поділитися з друзями), і ваші друзі в соціальних мережах побачать нове зображення, яке ви завантажили в Polyvore. (див. рис.1.4).



Рис. 1.4. Веб-сервіс «Polyvore.com»

5) «Майстерня Ксенії Штиль». Веб-застосунок у вигляді тестування для визначення стилю в одязі [5] (див. рис 1.5).

Поетапно виводяться питання з можливими варіантами відповіді користувача.

Онлайн-тест на определение стиля в одежде

УКАЖИТЕ ВАШЕ ИМЯ И ЕМЕЙЛ ТОЛЬКО ЕСЛИ ХОТИТЕ ПОЛУЧИТЬ РЕЗУЛЬТАТЫ ТЕСТА ПО ЭЛЕКТРОННОЙ ПОЧТЕ:

Ваше имя (по желанию) Ваш E-mail (по желанию)

1 2 3 4 5 6 7 8 9 10

1. КАКИЕ ЦВЕТА В ОДЕЖДЕ ВЫ ПРЕДПОЧИТАЕТЕ?

Люблю одежду ахроматических цветов, люблю контрастные сочетания в этих тонах. Это могут быть черный, белый, оттенки серого, бежевый.

Предпочитаю одежду немарких цветов, практичных в носке. Это могут быть оттенки серо-зеленого, серо-коричневого.

Отдаю предпочтение нежным, мягким, а также пастельным оттенкам. К ним могут относиться светло-розовый, нежно-голубой, нежно-фиолетовый, мятный.

Мне нравятся цвета и принты, присущие этно-культуре. Люблю элементы народных мотивов. Люблю сочетать 3-4 цвета в одном ансамбле, а также замысловатые орнаменты.

Рис.1.5. Веб-застосунок «Майстерня Ксенії Штиль»

б) **Metail**. Дозволяє клієнтам спробувати одяг у віртуальному приміщенні. Також представляє клієнтам імітаційну версію, яка враховує їх розмір і вагу (див. рис.1.6).

Вона має на меті зменшити кількість повернень, які клієнти здійснюють для онлайн– покупок, а також надати людям впевненість у пошуку предметів, які їх влаштовують і підходять, а також забезпечують загальний досвід клієнтів [6].

Metail має одну з найсучасніших технологій у цій галузі, і їх унікальний набір даних призводить до низки трансформаційних додатків для всього ланцюжка постачання одягу, що в кінцевому рахунку, веде до одягу, який краще підходить для клієнта (див. рис.1.7).

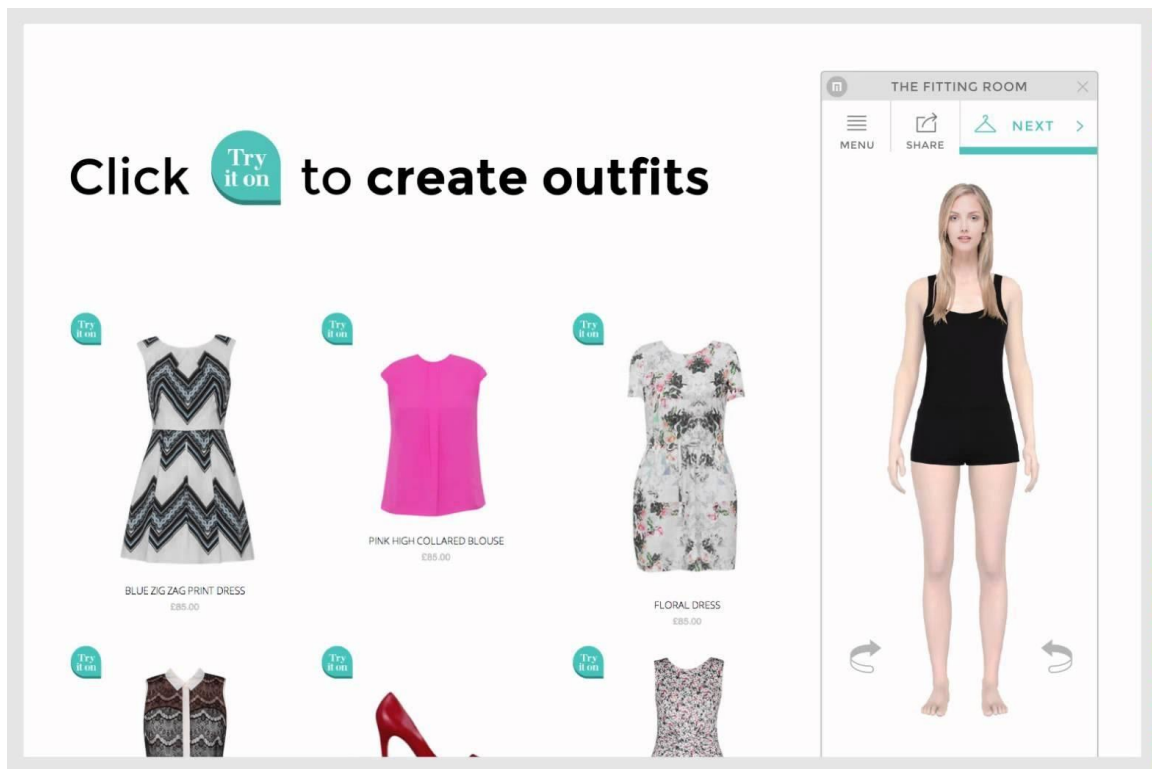


Рис. 1.6. Інтерфейс веб-застосунку «Metail»

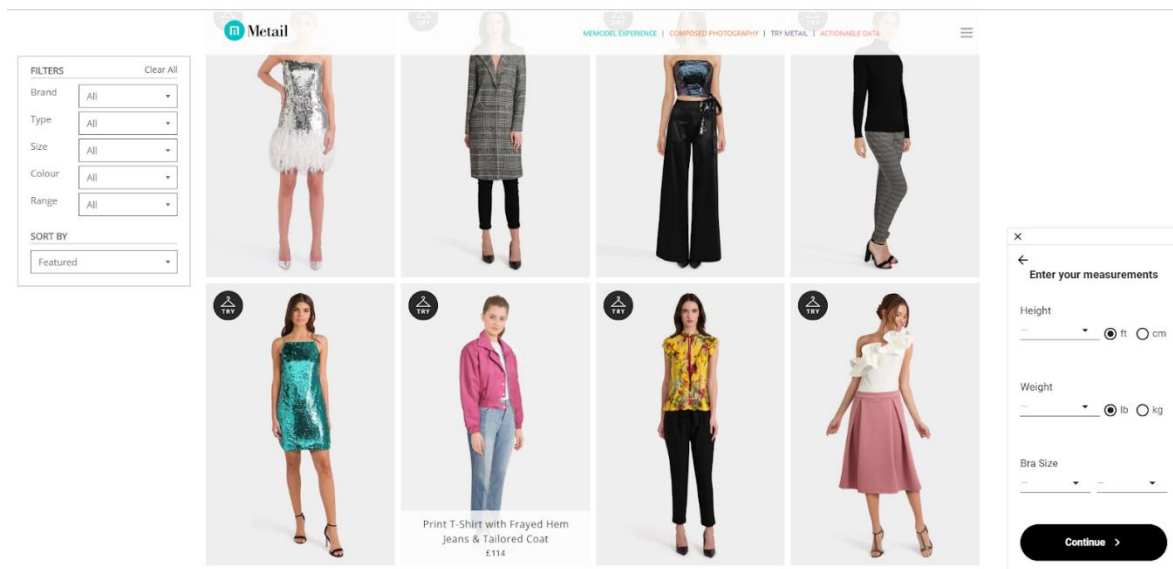


Рис.1.7. Веб-застосунок «Metail»

1.3. Аналіз літератури, публікацій та досліджень

Дослідженням питань ЕС займалися багато дослідників та розробників, нижче огляд деяких з них.

«Expert systems. Computers as sages» by Howard Rheingold

Edward Feigenbaum був один з перших людей, хто досліджував штучний інтелект та вирішив, в середині 1960-х років, що було важливо скільки комп'ютерна програма може знати, і як краще спробувати штучного експерта [7].

У пошуках необхідної області знань, Feigenbaum зіткнувся з Joshua Lederberg, який припустив, що хімікам-органікам вкрай необхідна допомога у визначенні молекулярної структури хімічних сполук. Разом з Bruce Buchanan, Lederberg, Feigenbaum почав роботу над DENDRAL, першої експертної системи, в 1965 році в Стенфордському університеті.

Звичайні комп'ютерні системи не забезпечили хіміків– органіків інструментом для прогнозування молекулярної структури. Хіміки знають, що можлива структура будь-якої хімічної сполуки залежить від ряду основних правил про те, як різні атоми можуть зв'язуватися один з одним.

Вони також знають багато фактів про різні атоми з відомими з'єднаннями. Коли вони роблять або відкривають раніше невідомі сполуки, вони можуть зібрати докази про з'єднання шляхом аналізу речовини з масою спектроскопа, який забезпечує багато даних, але не ключі до того, що все це означає. Побудова правильного виду програми «якщо-то», з достатньою гнучкістю, щоб використовувати правила – була лише перша серйозна проблема, яка повинна бути вирішена [8].

Коли ви думаєте, що ви створили структуру програми, здатну маніпулювати експертними знаннями, ви повинні помістити деякі знання в систему. Після «годування» комп'ютерних програм даними, творці DENDRAL опитали стільки експертів-хіміків: як вони могли б дізнатися, як вони зробили б свої рішення. На цьому етапі «придбання знань» має свої власні проблеми.

Коли його запитали, як вони знають і що вони знають, вони показали їм програму, яка приймає рішення і запитали їх, де в програмі неточності, і чому «Інженерія знань» – це мистецтво, ремесло і наука спостереження експертів, побудова моделей їх досвіду і вдосконалення моделі поки людські експерти не згодні, що це працює.

Одним з перших додаткових доходів від DENDRAL був Meta-DENDRAL, експертна система для тих людей, чий досвід полягає в побудові експертних систем. Buchanan був в змозі зробити інструмент для експертної системи будівельників.

В середині 1970-х років MYCIN був розроблений Edward-ом, лікарем і комп'ютерним вченим Стенфордської медичної школи.

Проблеми, пов'язані з діагностики певного класу інфекцій мозку була відповідна область для дослідження експертної системи і області особливо нагальною потребою людини, тому що перші двадцять чотири на дві доби мають вирішальне значення для лікування цих захворювань, щоб досягти успіху. З усіма його обіцянками, всі його лякали етичними наслідками, медицини, здається, одна з найбільш активних областей застосування для комерційної інженерії знань [9].

Механізм логічного виводу Mycin, відомий як E-MYCIN, був використаний дослідниками в Стенфорді і тихоокеанського медичного центру, щоб зробити Puff, експертну систему, яка допомагає діагностувати деякі захворювання легенів.

Навіть нова система, Кадуцей, тепер має більше бази знань, ніж будь-який фахівець з необроблених даних, що містить близько 80 відсотків медичної літератури в світі.

На сьогоднішній час, близько двох десятків корпорацій продажу експертних систем і послуг. Teknowledge, заснований Feigenbaum і сподвижниками в 1981 році, був першим. Intelli Genetics, мабуть, найекзотичніший, що спеціалізується в експертних системах для генної інженерії промисловості. Експертні системи в даний час в комерційних і дослідницьких цілях в ряді областей [10]:

- KAS (Knowledge Acquisition System – система збору знань) допомагають інженерам знання побудови експертних систем;
- ONCOCIN допомагає лікарям в управлінні складними схемами фармакотерапії для лікування онкологічних хворих;
- Molgen допомагає молекулярних біологів в плануванні експериментів ДНК;
- Guidon є експертом освіти, який вчить студентів, виправляючи відповіді на технічні питання;

- Genesis допомагає вченим у плануванні експериментів з клонування;
- TATR використовується BBC в плануванні нападу на противника авіабаз.

Yasser A. Nada у своїй статті «*Analysis, Design, and Implementation of Intelligent Expert System for Clothes Style Selection*» [11] поглиблено ознайомив з поняттями стилю, експертних систем, знань, методів, і як ці поняття поєднати в одну спільну робочу систему на CLIPS (C Language Integrated Production System) .

Oksana Zakharevich, Svetlana Kuleshova, Tetyana Zhylenko, Iryna Shuda «*Expert systems to solve informational issues in apparel design*». Основна мета дослідження досягається шляхом формування правил прийняття рішень, які мають на меті вирішення підзавдань швидкої зміни виробництва жіночого верхнього одягу, в тому числі вибору модних тканин і параметрів блоків шаблонів для конкретних видів одягу, вибору моделей готового одягу [12].

Верифікація результатів експерименту проводиться шляхом оцінки якості підгонки одягу, які були розроблені на основі роботи експертної системи.

Це підтверджує можливе застосування розробленої системи в розробці та виробництві одягу. Таким чином, створені необхідні умови для подальшого розвитку техніки штучного інтелекту в дизайні одягу.

Joseph C. Giarratano «*Expert systems principles and programming*»

В книгу включені відомості, що відносяться до двох основних напрямів: перший – викладається теорія експертних систем і показано, яке місце займають експертні системи в повному обсязі комп'ютерних наук [13].

Другий напрямок представлено відомостями з програмування за допомогою мови CLIPS. Ще одним новим засобом, описаним в даному виданні, є об'єктно-орієнтована мова програмування COOL. На початку книги міститься окреме введення в тематику штучного інтелекту, обсяг якого достатній для вивчення експертних систем.

Питанням розробки, проектування та дослідження експертних систем за допомогою методів Data Mining займалися такі науковці, нижче огляд деяких з них.

Alexandre Acácio De Andrade, Sergio Luiz Pereira, Eduardo Mario Dias, Caio Fernando Fontana у статті «*Using data mining techniques for development expert*

systems equipped with learning capabilities for use in automated industrial plants» розглянули питання застосування експертних систем, збільшення обсягу даних, що діють на сучасних заводах шляхом генерації автоматизованої розробки і тестування спеціалізованої системи, яка базується на Data Mining та володіє навчальною здатністю Artificial Intelligence (штучного інтелекту) [14].

Згадана система була названа NESISES. Вона працює в режимі реального часу з наглядом за промисловою автоматизацією систем, метою яких є мінімізувати часті процедури інженерії знань для постійного оновлення бази знань експертної системи.

Головна мета використання методів та технологій Data Mining була в пошуку знань у великих базах даних та автоматизації для подальшого повторного використання. Також ЕС працюють із системами контролю, що є потужним інструментом, який допомагає та оптимізує дії людських операторів на великих та складних автоматизованих промислових підприємствах.

Sholom M. Weiss, Stephen J. Buckley, Shubir Kapoor, and Soren Damgaard «Knowledge-Based Data Mining». У роботі спостерігається опис технології поєднання двох типів систем знань: експертного та машинного навчання. Обидві системи представляють інформацію шляхом виведення правил прийняття рішень або дерев [15].

Експертна система під назвою SEAS була побудована з метою виявлення потенційних клієнтів та рішень для продажу комп'ютерних продуктів. Система рекомендує продукти, які можуть покращити діяльність бізнесу. Серед потенційних переваг цього підходу є здатність визначати нові тенденції продажів.

Basilis Boutsinas «Accessing data mining rules through expert systems»

У цій роботі пропонується методологія доступу до правил Data Mining (видобутку знань), яка базується на використанні експертної системи. Представлено як різні типи правил видобутку даних можуть трансформуватися в знання доменів будь-якої експертної системи загального призначення; як певні значення атрибутів, надані користувачем як факти та/або цілі, можуть визначити за допомогою прямого та /або зворотного виведення, правила видобутку даних [16].

Zohreh Sedighian and Mahdi Javanmard «The effect of data mining on expert systems used for improving efficiency of correct speech E-learning systems». Аналіз та зберігання поведінки вчителів щодо правильної мовної освіти, а також пошук експертних підходів та пропозицій щодо корекції мовлення за допомогою експертної системи [17].

Знання, виявлені за допомогою Data Mining (інтелектуального аналізу даних) може використовуватися як додаткове джерело знань для експертної системи. Більше того, система інтелектуального аналізу даних, яка витягує знання, потребує експертного керівництва для подальшого використання прийняття рішень. У статті досліджується співпраця між інтелектуальним аналізом даних та експертом.

I.G.L. da Silva, B.P. Amorim, P. G. Campos, L.M. Brasil «Integration of Data Mining and Hybrid Expert System». В останні десятиліття значно зросла кількість використання/створення та збору даних в системах управління базами даних.

Однак традиційні методи генерують інформативні дані, проте не можуть аналізувати зміст цих даних, щоб вказати, які знання «цікавлять» найбільше. Ці труднощі сприяли виникненню інтелектуальних інструментів та методів Data Mining [18].

Дана робота полягає у застосуванні деяких з цих інструментів та методів для виявлення та закономірностей прогнозування медичної бази даних раку молочної залози в гібридній експертній системі, що допоможе в медичній діагностиці.

1.4. Аналіз існуючих методів та алгоритмів Data Mining

У всіх вище зазначених роботах прослідковується ріст та використання технології Data Mining у сенсі видобутку знань у великих масивах даних.

Згідно з деякими формами Data Mining можна встановлювати так звані "правила" або "обмеження", щоб підвищити рівень точності прогнозу події.

За допомогою інформаційних технологій, особливо розвитку в методах збору, обробки і зберігання можна отримати величезні масиви даних для подальшого аналізу [19].

Дані настільки величезні, що для обробки таких обсягів не вистачає можливостей експертів. Саме тому, в теперішній час відомий напрямок як інтелектуалізація методів обробки та аналізу даних.

Інтелектуальні системи аналізу даних (ІСАД) дозволяють вирішувати задачі прийняття рішення в процесі аналізу даних, також налаштування алгоритмів для аналізу, виявити закономірності, причинно-наслідкові зв'язки в системі для аналізу.

Data Mining поєднує широкий інструментарій. У технології об'єдналися кількісний і якісний аналіз даних, від статичного аналізу даних до методів кібернетики [20].

Основні методи Data Mining включають методи класифікації, прогнозування та моделювання. Також можна віднести статистичні методи, які працюють з невеликими обсягами даних.

Отже, саме тому для аналізу та дослідження слід розглянути статистичні методи як дерева рішень, методи класифікації, методи візуалізації даних, логічні методи.

Найважливіше призначення Data Mining – наочне подання результатів обчислень, що дозволяє використовувати інструментарій без спеціальної математичної підготовки. Проте при використанні статистичних методів допоможе володіння теорією ймовірності та статистики.

Знання, які добуваються методами Data Mining, прийнято представляти у вигляді моделей.



Рис.1.8. Моделі DM

Методи та алгоритми DM:

- дерева рішень;
- метод найближчого сусіда;
- метод k-найближчого сусіда;
- байєсівські мережі;
- кореляційно-регресійний аналіз;
- неієрархічні методи кластерного аналізу (алгоритми k-середніх і k-медіани);
- ієрархічні методи кластерного аналізу;
- лінійна регресія;
- еволюційне програмування;
- генетичні алгоритми;
- еволюційне програмування;
- генетичні алгоритми;
- нейронні мережі;
- візуалізація даних [21].

Data Mining використовує відомі математичні методи і алгоритми у поєднанні з новими можливостями технічних і програмних засобів.

Data Mining має сенс з роботою з «сирими даними» тобто необробленими, первинними, які були зібрані з джерела.

На теперішній час існують багато експертних систем для постановки медичних діагнозів. В їх основі правила, які описують симптоми захворювань. За допомогою такої системи та правил можна дізнатися спосіб лікування пацієнта. Правила можуть допомогти обрати найефективніший курс лікування, що включає курс лікування, протипоказання та прогнози [22].

Технології Data Mining дозволяють виявляти в медичних даних шаблони, що становлять основу зазначених правил. Data Mining аналіз допомагає в прийнятті рішень, тобто для чого факти не самі по собі, а знання про їх закономірності в процесах. Таким чином, Data Mining – це процес виявлення корисних знань.

Data Mining спрямований на вилучення знань (фактів, залежностей, угруповань) з великих обсягів інформації для виявлення взаємозв'язків, причин та наслідків.

Набори правил прийняття рішення під конкретну задачу зазвичай не формуються – частіше це набір розрізнених знань, які потрібно доповнювати отриманими з інших джерел, і таким чином при бажанні згодом побудувати класичну ЕС (на основі витягнутої за допомогою Data Mining інформації) [23].

Також Data Mining в можливе використання окремих статистичних процедур – наприклад визначення кластерів і віднесення до того чи іншого кластеру.

Доведено необхідність методів, які автоматично формують шаблони та логічні зв'язки з мінімальним втручанням користувача. Data Mining і є своєрідним найвідомішим методом, за допомогою якого відбувається розпізнавання корисних шаблонів всередині даних.

Метод надає користувачам та аналізаторам інформацію для прийняття найважливіших рішень на основі наданої інформації, оскільки продуктів аналізу даних недостатньо.

Для успішного аналізу даних необхідні професійні аналізатори та досвідчені люди, які здатні аналізувати та інтерпретувати отримані результати. Отже, обмеження видобутку даних пов'язані не з технологіями, з початковими даними або експертами.

1.5. Постановка задачі

Основною задачею дипломної роботи є дослідження застосування технології Data Mining при проектуванні та розробці програмного засобу експертної системи для поліпшення процесу прийняття рішень в задачі підбору стилю на основі методів Data Mining.

Актуальність магістерської роботи визначається необхідністю дослідження та розробки технології пошуку і накопичення знань для прийняття рішення в задачі підбору стилю, яка б дозволяла зручно зберігати і передавати інформацію користувачеві. Технологія має полегшувати вибір рішення та сприяти б більш точному результату. Створення такої технології обумовлене недостатньою кількістю досліджень і розробок в сфері підбору стилю.

Експертна система здатна запропонувати рішення або рекомендації щодо конкретних проблем в обраній предметній області.

Метою є дослідження застосування технології Data Mining при проектуванні експертних систем рекомендаційного типу та поліпшення процесу прийняття рішень в задачі підбору стилю на основі методів Data Mining.

В ході виконання магістерської дипломної роботи були використані такі *методи*: метод прямого виводу, продукційний метод представлення знань (для розробки структури ЕС) та методи Data Mining (дерева рішень, візуалізації, класифікації для підвищення ефективності і точності роботи програмного засобу).

Для досягнення вказаної мети необхідно виконати *наступні завдання*:

1) провести аналіз предметної області, а саме дослідити стан задачі підбору стилю, провести огляд існуючих аналогів ПЗ, публікацій та досліджень, дослідити існуючі методи Data Mining;

2) дослідити загальні риси та складові поняття експертної системи;

- 3) провести аналіз методів та підходів для проектування та розробки структури ЕС підбору стилю, проаналізувати їх переваги та недоліки;
- 4) розробити модель експертної системи з використанням методів Data Mining для поліпшення точності прийняття рішень;
- 5) розробити програмний застосунок експертної системи підбору стилю, проаналізувати отримані результати роботи.

РОЗДІЛ 2. АНАЛІЗ МЕТОДІВ ТА ПІДХОДІВ ДЛЯ ПРОЕКТУВАННЯ ТА РОЗРОБКИ СТРУКТУРИ ЕС

2.1. Загальна характеристика поняття ЕС

Для виконання дипломної роботи було обрано рекомендаційну експертну систему та задачу підбору стилю, в результаті якої надає рекомендації для прийняття рішення користувачем.

Експертні системи – системи, здатні запропонувати рішення або консультації щодо конкретних проблем в даній предметній області.

Експертна система – сукупність методів, моделей та даних.

Сукупність знань – основа для ЕС.

Для розробки експертних систем вимагається значний людський досвід та професіоналізм для вирішення проблем в області.

Експертні системи здатні вирішувати лише обмежені проблемні питання. Проте навіть у дуже обмежених областях експертні системи потребують великих обсягів знань порівнюючи з фахівцями; виконують операції, які зазвичай вимагають наявності значного людського досвіду.

Крім того, для цього потрібен експерт – людина, яка вміє знаходити вирішення проблем в конкретній предметній області.

Для створення складної ЕС, потрібні знання і практичні навички декількох фахівців.

Також не менш важливий інженер знань – людина, яка знає яким чином побудувати, змодельовати, спроектувати ЕС, як структурувати або організувати знання експерта [24].

Кафедра КІТ (47)				НАУ 20 13 79 000 ПЗ			
Виконав	Литовченко О.В.			Проектування та розробка структури ес за допомогою методів Data Mining	Літера	Аркуш	Аркушів
Керівник	Воронін А.М.					27	17
Консульт							
Н.контр.	Райчев І.Е.				УС-211М		122
							27

Для побудови ЕС необхідний програміст, який буде займатися розробкою ПЗ, використовуючи спеціальні або власні інструментальні засоби.

Класифікація ЕС за призначенням:

- інтерпретація та ідентифікація;
- прогнозування;
- діагностика;
- проектування;
- планування;
- моніторинг;
- налагодження та тестування;
- навчання;
- рекомендація;
- контроль [25].

Таким чином, відбувається поєднання людських знань з комп'ютером для вирішення проблеми.

Одним з найпотужніших атрибутів експертних систем є здатність пояснювати міркування та висновки.

Більш того, розробка експертної системи зазвичай проходить через кілька етапів, включаючи проблему відбору, набуття знань, представлення знань, програмування, тестування та оцінювання.

Основні компоненти ЕС

1. База знань.
2. Машина логічного виведення (висновку).
3. Інтерпретатор команд.
4. Інтерфейс (системи пояснення) (див. рис. 2.1) [26].

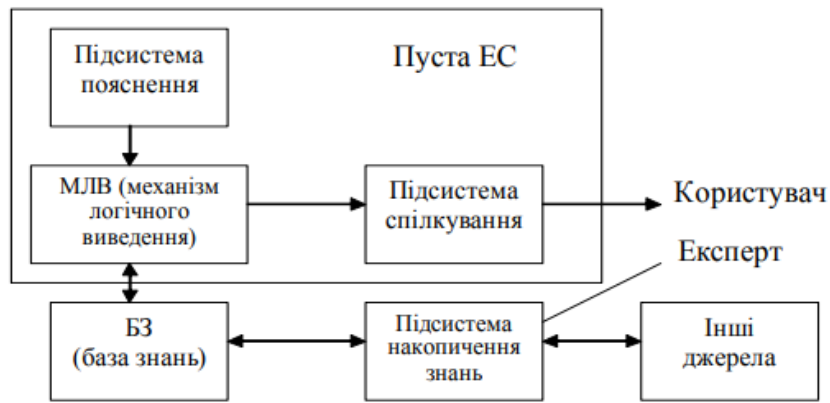


Рис. 2.1. Компоненти ЕС

БЗ та процедура логічного виведення – це ядро ЕС. Їх слід розглядати разом, оскільки знання, на основі яких не можна сформулювати результати та висновки, не мають будь-якого сенсу.

База знань

Знання – це сукупність понять про закономірності природи, суспільства і мислення, в процесі нагромадження діяльністю людства. Знання зберігаються в БЗ. БЗ – знання з предметної області. БЗ ЕС містить факти (дані) та правила (способи подання знань). Механізм логічного виведення складається з інтерпретатора, який визначає, як застосовувати правила для виводу нових знань [27].

База знань – це сукупність всіх знань, що формують ЕС.

Машина логічного виводу (висновку)

Існує прямий логічний вивід та зворотній логічний вивід. Перший – веде від даних до гіпотез, а другий – спроба знайти дані для доведення або спрощення певної гіпотези. Найбільш досконалі системи використовують комбінацію обох виводів [28].

Інтерпретатор команд

Інтерпретатор команд визначає як застосувати правила для виводу нових знань, встановлюючи порядок застосування цих правил.

Інтерфейс (система пояснення)

Спеціаліст використовує інтерфейс для введення інформації і команд в ЕС та одержання вихідної інформації з неї. Команди містять параметри, що спрямовують на процес обробки знань.

Інформація звичайно подається у формі певних значень, що присвоюються певним змінним.

Технологія ЕС забезпечує можливість одержання в якості вихідної інформації не тільки рішення, але й необхідні пояснення.

Етапи розробки ЕС:

- ідентифікація проблеми;
- вилучення знань;
- структуризація знань;
- формалізація;
- реалізація ЕС – створення одного або декількох прототипів ЕС, які вирішують поставлені задачі;
- тестування [28] (див. рис. 2.2).

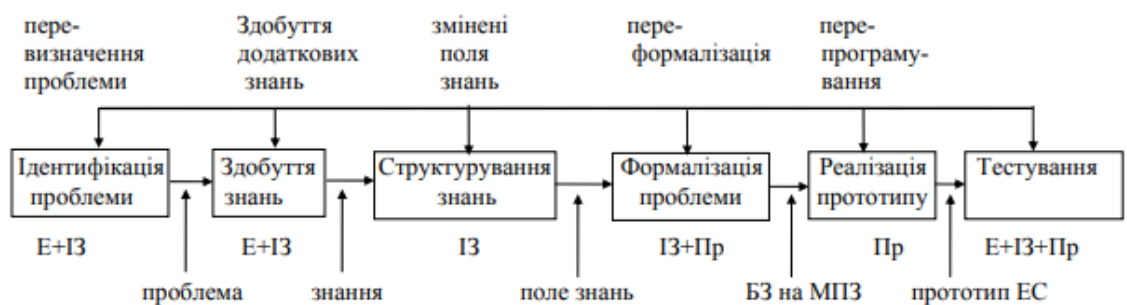


Рис. 2.2. Процес та етапи проектування ЕС

Головна відмінність ІС і ЕС від інших систем та програмних засобів – це наявність БЗ у якій знання зберігаються у формі, зрозумілій фахівцям предметної області і можуть бути змінені і доповнені також у зрозумілій формі.

2.2. Продукційний метод представлення знань

Продукційні правила – найбільш простий спосіб представлення знань (див. рис. 2.3). Він базується на представленні знань у формі правил, структурованих у відповідності до схеми *якщо-то*. Частина правила *якщо* називається умовою придатності (антецедент), а частина *то* – висновком (консеквент) [29].

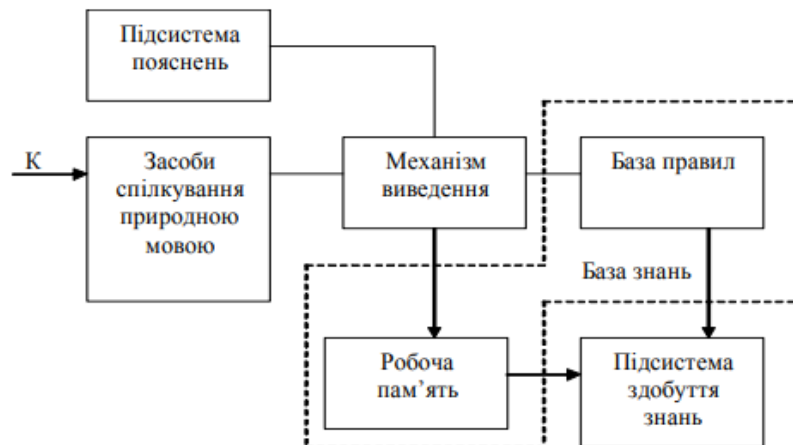


Рис. 2.3. Структура ЕС продукційного типу

Якщо умов придатності немає, тоді знання складаються лише з висновку і називаються фактами.

У продукційних системах використовуються два основні методи логічного виводу: прямий і зворотній.

За *прямого виводу* правила досліджуються одне за одним у певній послідовності. Виходячи з початкових умов (даних), введених користувачем, для кожного правила оцінюється істинність чи хибність його умови придатності.

Якщо умова – істина, то правило активізується, в іншому випадку – ні. Процедура виведення є ітеративною і може потребувати декількох повторів через усю низку правил, поки буде визначене певне значення цільової змінної.

За *зворотного виведення* припускається істина наслідку (тобто дії) деякого правила, після чого потрібно рухатись низкою правил у зворотній послідовності.

Проблеми

Продукційна модель знімає обмеження логіки, проте їй притаманні такі проблеми як нескінченні цикли, суперечність знань, непрозорість поведінки машини логічного виводу.

Нескінченні цикли виникають у тому випадку, коли машина виводу повертається до правил, які вже були переглянуті раніше.

Суперечливі знання з'являються тоді, коли додавання нових правил призводить до суперечності тим фактам, які можна було отримати раніше.

Непрозорість поведінки обумовлена хибним порядком виконання правил.

Внаслідок цього досить важко обробляти всі продукційні бази знань великого обсягу, оскільки навіть за умов правильності всіх наявних правил, хибний порядок їхнього виконання може призвести до помилок, які важко виявити.

Метод прямого логічного виводу

Машина логічного виводу – програма, яка робить логічні висновки з попередньо побудованої бази фактів.

Складається з двох компонентів:

- *компонент логічного виводу*, який реалізує власний дедуктивний висновок. Тобто, якщо в базі фактів є факт A , а в базі правил є правило $If A then B$, то робиться висновок про необхідність застосування дії B .
- *інтерпретатор правил*, який керує процесом перебирання фактів і застосування правил [29].

Схематично процес інтерпретації правил зображено на рисунку 2.4:

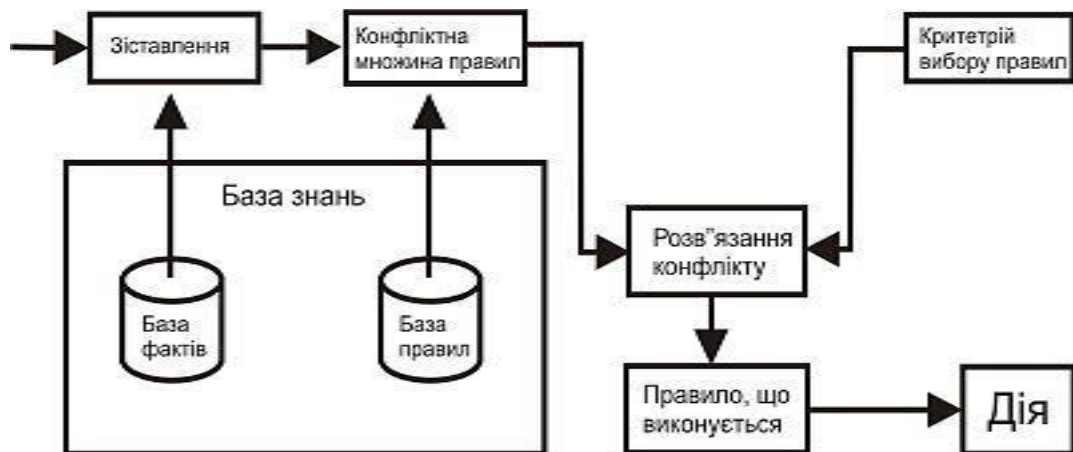


Рис. 2.4. Інтерпретації правил

Обґрунтування вибору продукційного методу та методу прямого логічного виводу

Для даної ЕС була обрана продукційна модель прямого виводу представлення знань, тому що вона є найбільш наочним засобом подання знань.

Вона близька до логічних моделей, що дозволяє організувати на її базі ефективні процедури виводу, і в той же час більш наочно (порівняно з класичними логічними моделями) відображає знання [33].

Продукційна модель найчастіше застосовується в більш легких ЕС, які базуються на простих формальних знаннях. Вона влучно підходить рекомендаційній ЕС із задачею підбору стилю з використанням методів Data Mining.

Також вона привертає розробників своєю наочністю, високою модульністю, легкістю змін і доповнень, простотою логічного виводу.

2.3. Проектування продукційної моделі бази знань прямого виводу

Продукційна модель знань представляє собою основу для проектування експертної системи. Її основні компоненти – це машина виводу, база правил та знань, також можна виділити робочу пам'ять.

Продукційну модель було обрано через формальну задачу підбору стилю на основі тестування та подачі рекомендацій. Вона не виконує складних операцій, рахує сумарні ймовірності та дає рекомендації, тому дана модель представлення знань є найбільш вдалою.

База знань поповнюється користувачем, відповідаючи на набір питань.

Існують такі вхідні дані задачі:

- 15 питань, для яких по 3–4 варіанти відповідей для кожного питання;
- кожний варіант відповіді має значення ймовірності від 0 до 1;

Вихідні дані задачі:

- результати у вигляді рекомендації підбору стилю та візуального представлення результату та його динаміки у вигляді дерева рішень;
- разом із результатами виводиться сумарна ймовірність для певного виду стилю, виводяться всі види, оскільки людина може обрати інший варіант з нижчою ймовірністю, що їй до вподоби.

Питання та відповіді, що складають базу знань див. в додатку А.

На основі цих даних створено БЗ продукційної моделі прямого виводу за допомогою простих конструкцій:

If (умова) and (умова) or (умова) ... then (дія)

If (умова) or (умова) and (умова) ... then (дія)

Порядок виконання операцій залежить від порядку проведення логічної операції в конструкції.

База правил. Приклад правил

Існує такий набір правил для моделі БЗ:

If (1 – 1), and (2 – 1), and (3 – 4), and (4 – 2), and (5 – 3), and (6 – 2), and (7 – 4), and (8 – 1), and (9 – 2), and (10 – 4), and (11 – 1), and (12 – 3), and (13 – 4), and (14 – 3), and (15 – 1), then (1 = 2,5), (2 = 0,5), (3 = 0,5), (4 = 0), (5 = 0,7), (6 = 0), (7 = 0,3), (8 = 2,5).

Примітка: в if перша цифра – номер питання, друга цифра – номер варіанта відповіді, у then перша цифра – номер рекомендації та виду стилю, друга цифра – значення ймовірності.

Отже, таким чином, можна побачити, що стиль 1 та 8 мають однакові значення отриманої ймовірності.

2.4. Класифікація стилю

Після аналізу інформації було розроблено власну класифікацію стилю на основі джерел публікацій та досліджень, оскільки мене обрано на роль експерта, інженера знань та програміста одночасно.

Інформації було накопичено достатньо, щоб представити основні компоненти класифікації стилю, що постає висновками (фактами) та результатами системи. До уваги бралися ресурси щодо накопичення знань в даній предметній області. Предметна область – задача підбору стилю.

Можна представити класифікацію таким чином:

1) Повсякденний стиль.

Опис: вам слід спробувати джинси у поєднанні зі заправленою кофтиною, футболкою, рубашкою, блузою. Якщо вам до вподоби плаття, спідниці, то вони повинні бути максимально не обтягуючими, типу трапеції або квадратної форми. Із взуття актуальні кеди та кросівки, стильно та комфортно. Прикраси, аксесуари та біжутерія будь-які за вашим смаком, проте головне правило – мінімалізм та відсутність вульгарності. Макіяж може бути як і нюдовим, так і більш ніжним чи яскравим. У залежності від подій та планів. Волосся як правила розпущене. Існує тенденція рівного гладенького волосся.

2) Спортивний стиль.

Опис: вам слід одягти спортивний костюм, який буде не лише максимально комфортним, слід звертати свою увагу на те, щоб надавати можливість тілу дихати. Класичний спортивний костюм складається зі штанів, футболки, спортивного взуття. Популярна тенденція: дихаючі легінси в поєднанні з топом, на який одягнута майка. Макіяж відсутній, але допускається з використанням легких водостійких косметичних засобів натуральних кольорів. З аксесуарів пасують розумні годинники, відсутність біжутерії для комфорту. Волосся повинно бути зібраним за смаком: пучок, кінський хвіст.

3) Урочистий стиль.

Опис: такий стиль є найбільш яскравим, відкритим та може мати будь– які індивідуальні побажання. Він використовується на урочистих подіях (весілля, випускний, день народження, вечірки), вимагає максимальної святковості та урочистості. Макіяж та образ можуть бути максимально яскравими. Як правила образ складається із сукні (як довгої чи пишної так і короткої), взуття на підборах. Із волосся краще зробити зачіску в салоні, щоб та не розпалася.

4) Базовий офіційно– діловий стиль.

Опис: такий стиль повинен бути максимально офіційним, без високих тенденцій моди та складатися з базових частин офіційно– ділового стилю. Як правило, він необхідний людям, які працюють в певній галузі, де є дрес– код та необхідний офіційний вигляд. Проте завжди можна виглядати максимально привабливо та стильно. Елементи модної тенденції можна комбінувати з офіційно– діловими костюмами, спідниця може бути строгою проте мати якусь особливість. Штани, піджак можуть бути підкороченими, завуженими, наприклад знизу. Аксесуар як сумка повинен бути не яскравих кольорів проте може виділятися. Волосся як правило потрібно бути зібраним.

5) Класичний стиль.

Опис: такий стиль може мати елементи офіційності, проте зазвичай складається з класики, тих тенденцій які існують завжди. Як правило більшість людей, яким до вподоби класика не мають інтересу або жаги до модних тенденцій. Тому слід одягатися звичайно, мінімалістично. Такий стиль може більше підходити людям, які люблять темні кольори, або мають проблеми з фігурою, щоб якнайкраще прикрити недоліки. Такий стиль є універсальним, зазвичай мінімалістичним

б) Повсякденний діловий стиль.

Опис: такий стиль використовують для навчання, роботи. Як правило, не має чітких обмежень, проте необхідно виглядати по дрес-коду. Вбрання повинно вимагати простоти та комфорту, проте можна використовувати деякі частини повсякденного стилю, як наприклад джинси з блузою та піджаком, та кросівки. Це

залежить від установи та події. Як правило занадта строгість відсутня, но її елементи повністю взаємозалежні. Аксесуари, біжутерію краще використовувати мінімально. Зачіска звичайна, як іноді розпущена так і підібрана для комфорту та офіційності.

7) Урочистий повсякденний стиль.

Опис: такий стиль змішує елементи урочистого та повсякденного. Більш підходить для вечірніх прогулянок, походу до кінотеатру, якогось звичайного свята тобто подія, яка потребує якоїсь невеличкої урочистості та жіночності. Вам слід звернути увагу на сумісні костюми, з відкритими плечима або спиною, якогось незвичного для вас кольору. Також плаття ніжних світлих кольорів, в яких комфортно та не виглядає вульгарно. Можна поєднати із взуттям на підборах, платформі. Аксесуари, біжутерія за смаком повинні бути. Макіяж більш святковий, додати деякі елементи вечірнього, проте більш ніжною гамою. Волосся може бути накрученим, так і прямим.

8) Повсякденний сучасний стиль.

Опис: вам слід найбільше звернути увагу на модні тенденції. Як правило, такий стиль може бути максимально неординарним, використання тенденцій одягу over– size, або протилежність – обтяжність. Краще всього поєднувати ці два компоненти, наприклад широка худі з короткими обтягуючими джинсами або спідницею. Плаття в стилі розпущеної довгої рубашки. Зі взуття оптимальні низькі підбори, лодочки, мокасини, кросівки, кеди. Аксесуари та взуття яскравих різних кольорів. У біжутерії та аксесуарах обмежень немає. Волосся повинно виглядати гладеньким здоровим, проте ефект 'розтріпаного волосся' ніхто не відміняв.

Отже, до уваги було обрано лише ті стилі, які найчастіше зустрічаються у повсякденному житті.

2.4. Метод дерева рішень

Дерева рішень

Дерево має багато аналогій у реальному житті, і виявляється, що воно вплинуло на широку область машинного навчання, охоплюючи як класифікацію, так і регресію [30].

При аналізі рішень дерево рішень може використовуватися для візуального та явного представлення рішень та прийняття рішень. Як випливає з назви, вона використовує деревоподібну модель рішень. Незважаючи на те, що інструмент інтелектуального аналізу даних часто використовується для вироблення стратегії досягнення певної мети.

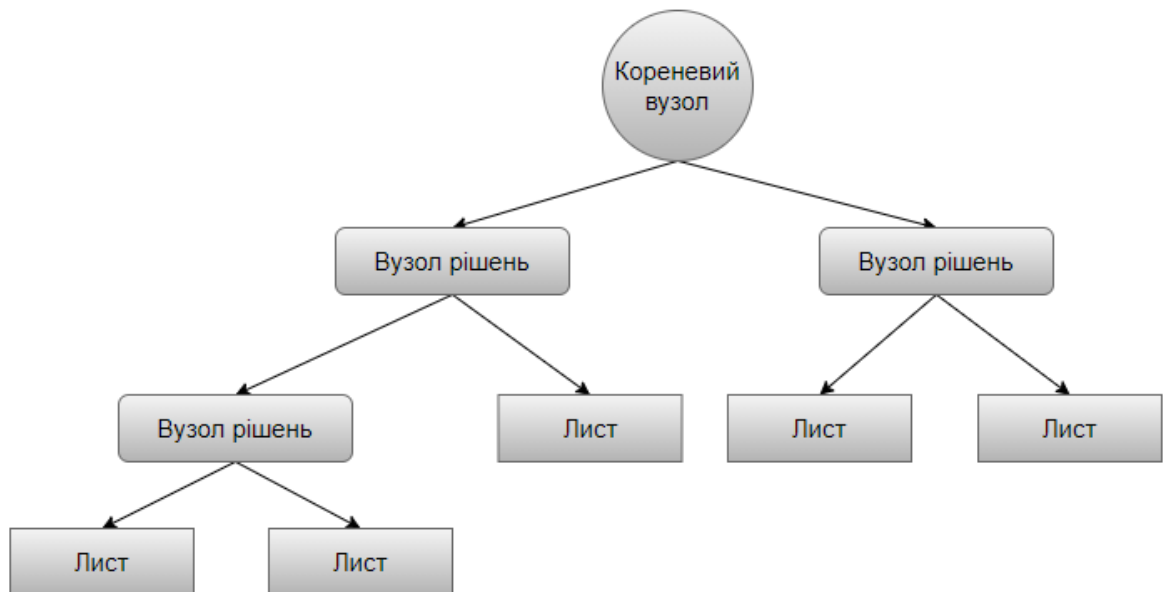


Рис. 2.5. Основні компоненти дерева рішень

Представлення алгоритма у вигляді дерева

Для цього треба розглянути дуже базовий приклад, який використовує набір даних для прогнозування виживання пасажирів чи ні. Нижче модель використовує 3 ознаки/атрибути/стовпці з набору даних, а саме стать, вік та супутні (кількість подружжя чи дітей разом).

Дерево рішень малюється догори дном, а корінь зверху. На зображенні зліва жирний текст чорним кольором представляє умову/внутрішній вузол, на основі

якого дерево розбивається на гілки/межі. Кінець гілки, яка більше не ділиться – це рішення/лист, в даному випадку чи загинув пасажир, чи вижив, як червоний та зелений текст відповідно.

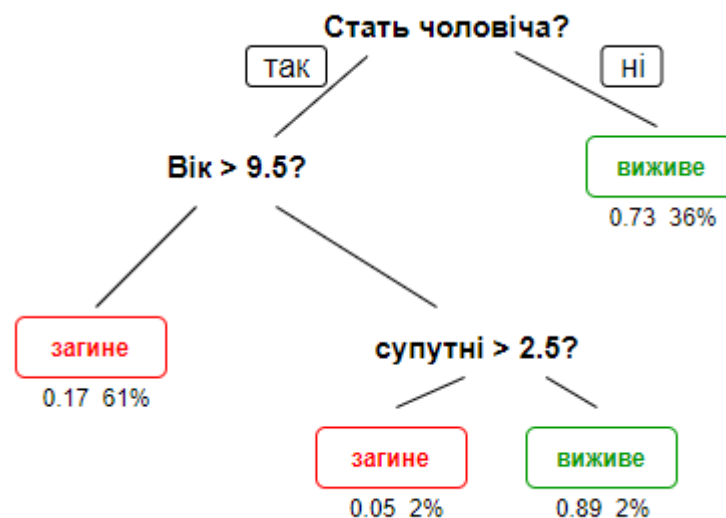


Рис. 2.6. Приклад дерева рішень

Хоча справжній набір даних матиме набагато більше можливостей, і це буде просто гілкою набагато більшого дерева, але не можна ігнорувати простоту цього алгоритму. Важливість функції зрозуміла, і відносини можна легко переглянути. Ця методологія більш відома як дерево прийняття рішень за даними, і вище дерево називається деревом класифікації, оскільки метою є класифікація пасажирів як вижилих або загиблих. Древа регресії представлені однаково, просто вони передбачають безперервні значення, такі як ціна будинку. Загалом алгоритми дерева рішень називаються CART або класифікаційними та регресійними деревами.

“Вирощування” дерева передбачає прийняття рішення про те, які особливості вибрати і які умови використовувати для розбиття поряд із знанням часу зупинки. Оскільки дерево зазвичай росте довільно, його необхідно буде обрізати, щоб воно виглядало так як потрібно. Саме тому існують різні загальні техніки, які використовуються для розбиття.

Дерево рішень – один із керованих алгоритмів машинного навчання. Цей алгоритм може бути використаний для задач регресії та класифікації, але в основному використовується для задач класифікації. Дерево рішень слідує за набором умов if - else, щоб візуалізувати дані та класифікувати їх відповідно до умов.



Рис. 2.7. Приклад дерева рішень

Метод дерева рішень – це метод аналізу величезних масивів даних.

Слід також зазначити, що метод дерева рішень використовує ієрархічну структуру.

Основні складові алгоритму дерева рішень

Кореневий вузол

Цей атрибут використовується для розподілу даних на два або більше наборів. Атрибут об'єкта в цьому вузлі обирається на основі методів вибору атрибутів.

Гілка або піддерево

Частина всього дерева рішень називається гілкою або піддеревом.

Розбиття

Ділення вузла на два або більше підвузлів на основі умов if– else.

Вузол прийняття рішень

Після поділу підвузлів на подальші підвузли він називається вузлом прийняття рішень.

Лист або термінальний вузол

Кінець дерева рішень, де його не можна розділити на подальші підвузли.

Обрізання

Видалення підвузла з дерева називається обрізанням.

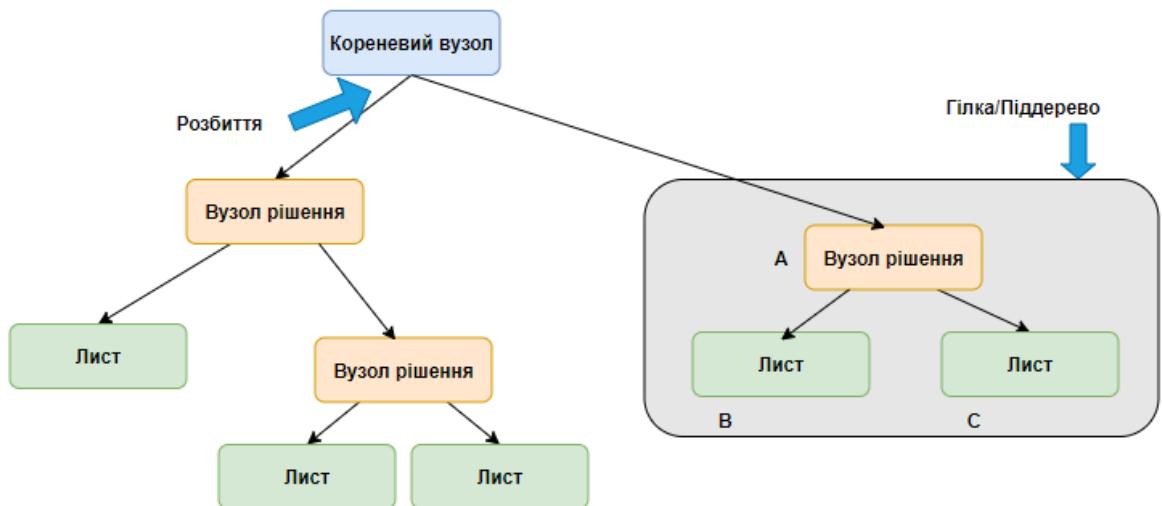


Рис. 2.8. Основні складові алгоритму дерева рішень

Робота дерева рішень

Функція кореневого вузла обирається на основі результатів вимірювання вибору атрибутів (ASM).

ASM повторюється доти, доки кінцевий вузол або кінцевий вузол не можна розділити на підвузли.

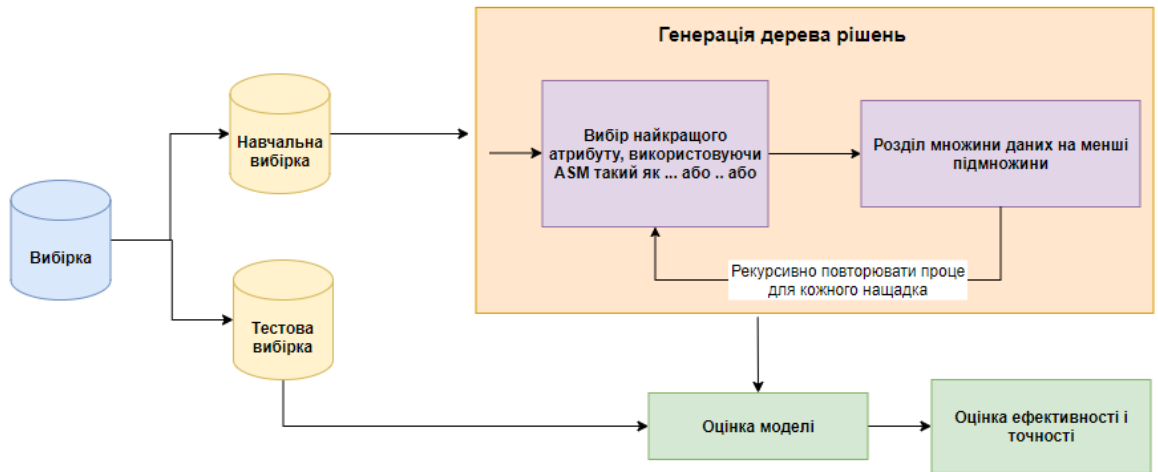


Рис. 2.9. Приклад генерації дерева рішень

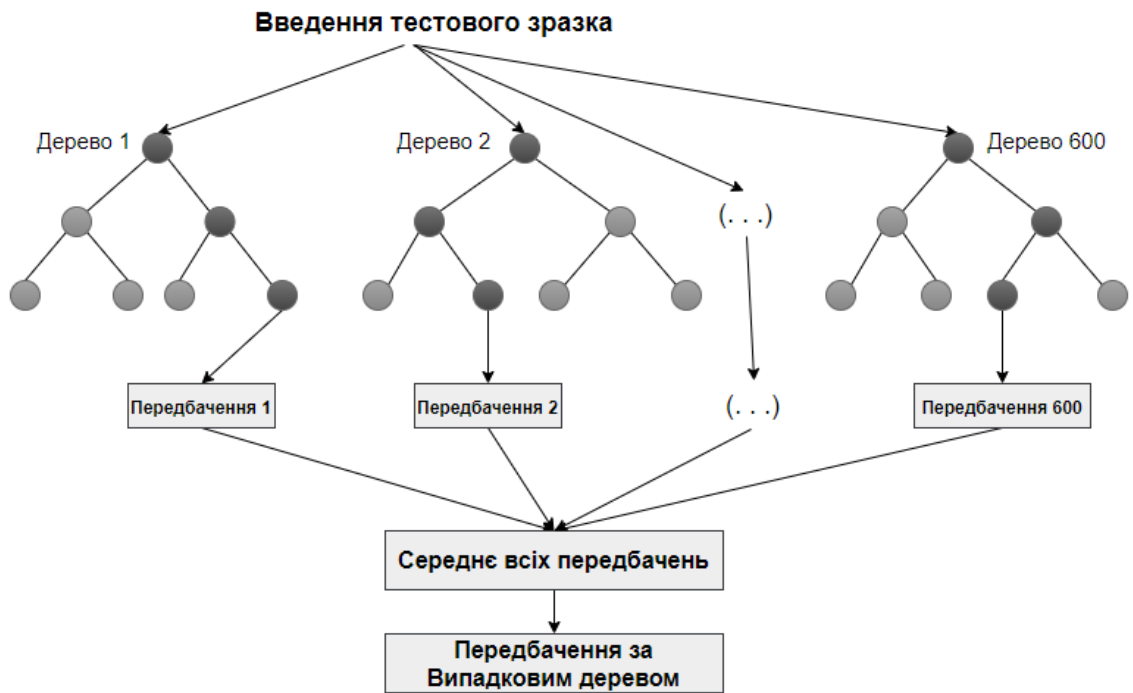


Рис. 2.10. Алгоритм побудови дерева рішень

Модель дерева рішень

Модель дерева рішень представляє собою модель розрахунків, в якій алгоритмом є дерево рішень тобто послідовність розбиття об'єктів на основі порівняння.

Просте дерево рішень

Модель, в якій кожне рішення з'являється за допомогою порівняння двох чисел. Застосовується для сортування та пошуку.

Лінійне дерево рішень

Це просте дерево рішень, проте рішення з'являється за допомогою сукупності деяких значень вхідних даних та мають три вихідні гілки рішень.

Використання методу дерева рішень:

- зберігання інформації в необ'ємній і зручній формі, яка містить точні описи об'єктів;
- застосування завдання класифікації – віднесення об'єктів до класів;
- застосування регресії – визначення залежності змінної від незалежних вхідних змінних у випадку недостовірних значень змінної;
- покращення якості та точності прийняття рішень;
- графічне відображення та візуалізація послідовності рішень.

РОЗДІЛ 3. ТЕХНОЛОГІЯ ЗАСТОСУВАННЯ МЕТОДІВ DATA MINING

3.1. Метод класифікації

Класифікація є сферою досліджень, що розширюється, особливо у відносно недавньому контексті аналізу даних [32].

Класифікація використовує рішення для класифікації даних. Кожне рішення встановлюється за запитом, пов'язаним з однією із вхідних змінних. На основі підтверджень екземпляр даних класифікується.

Кілька добре охарактеризованих класів, як правило, забезпечують ефективний синтез сукупності об'єктів. Таким чином, класифікація є потужним інструментом для дослідження даних.

Класифікаційний аналіз – це завдання Data Mining, яке ідентифікує та присвоює категорії до набору даних для більш точного аналізу. Класифікаційний аналіз можна використовувати для опитування, прийняття рішення чи прогнозування поведінки за допомогою алгоритмів.

Класифікація складається з алгоритму класифікації, в основі якого логічні правила класифікації, такі ж як в правилах експертних систем. Саме тому класифікація суміжна з ЕС та деревом рішень та підходить для вирішення поставленої мети та задач. Дерева рішень відносяться до алгоритмів та методів класифікації. В сукупності ці алгоритми та методи збільшують точність розрахунків та покращують результат прийняття рішень.

Класифікація – одна з найбільш часто використовуваних технік, коли мова йде про класифікацію великих наборів даних.

Цей метод аналізу даних включає алгоритми контрольованого навчання, адаптовані до якості даних.

Кафедра КІТ (47)				НАУ 20 13 79 000 ПЗ			
Виконав	<i>Литовченко О.В.</i>			Технологія застосування методів data mining	<i>Літера</i>	<i>Аркуш</i>	<i>Аркушів</i>
Керівник	<i>Воронін А.М.</i>					44	9
Консульт							
Н.контр.	<i>Райчев І.Е.</i>					УС-211М	122
							44

Метод класифікації використовує такі алгоритми, як дерево рішень, для отримання корисної інформації. Компанії використовують цей підхід, щоб дізнатись про поведінку та уподобання своїх клієнтів.

За допомогою класифікації можна розрізнити дані, які корисні для мети, та дані, які не мають значення.

3.2. Метод дерева рішень

Алгоритм побудови дерева рішень

Задача вибору стилю може розглядатись як задача класифікації на основі дерев рішень як методу її реалізації. При цьому, чітко прослідковується статична ієрархія компонентів обраної структури, з фіксованою кількістю рівнів.

Компонентами верхнього рівня є власне стилі, що представляють собою альтернативні рішення E_i , які пропонуються людині, що приймає рішення.

Ступенем оптимальності кожного альтернативного рішення отримується за допомогою агрегації тих чинників C_j , що мають певний вплив на нього [33].

Такими чинниками у запропонованій задачі виступають відповіді людини, що приймає рішення, на питання-фактори F_{jk} .

Питання виступають вузлами найнижчого рівня і слугують радше меті об'єднати окремі чинники у логічні групи. Власне, відповіді є компонентами проміжного рівня.

Отримана ймовірність з бази знань експертної системи може розглядатись як вага для дерева рішень.

Зв'язки між окремими чинниками та альтернативними рішеннями несуть у собі інформацію про рівень впливу на вирішальний ступінь оптимальності – ваги.

При цьому, кожен з чинників може впливати на одне чи більше рішень із різним рівнем впливу. Вага може знаходитись на діапазоні $[0, 1]$.

Розрахунок оптимальності альтернативних рішень виконується у два основні етапи:

- агрегація чинників, що відповідають даному рішенню;
- акумуляція ваг чинників для цього рішення.

Власне акумуляція може виконуватись із використанням різних математичних операторів, як сума чи множення.

У випадку із запропонованою задачею, застосовується операція додавання. Даного оператора цілком достатньо для простої моделі акумуляції ваг чинників. Наведемо формулу акумуляції:

$$E_i = \sum_{r=1}^m w_r \cdot a_r,$$

де w_r відповідає вазі чинника F_{jk} попередньо агрегованого для рішення E_i ,

r відповідає індексу чинника F_{jk} ,

a_r – коефіцієнт активації чинника F_{jk} , що приймає значення 0 або 1,

m – кількість чинників, що відповідають

Отже, дерева рішень допомагають проаналізувати, які частини бази даних дійсно корисні або яка частина містить рішення вашої проблеми.

Це інструмент підтримки, який використовує схему прийняття рішень або модель та її можливі наслідки.

З точки зору рішення, дерево рішень – це найменша кількість питань, які слід визнати, щоб оцінити ймовірність прийняття точного рішення.

Переглядаючи прогнози чи значення для кожного поділу на дереві, можна уявити деякі ідеї або знайти відповіді на запитання.

3.3. Метод візуалізації даних

Візуалізація інформації та даних допомагає уявити та відобразити результат роботи у різноманітних формах і різноманітними способами.

Вона є невід’ємною частиною при роботі з методами не лише Data Mining, а й взагалі.

Найкращі прийоми та способи візуалізації даних для малих та великих даних:

- гістограми;
- діаграми (кругові, пончикові);
- діаграми гістограми;
- графіки (функцій);

- кластери;
- мережеві діаграми;
- матриці;
- дерева (рішень).

Отже, таким чином у дипломній роботі велика увага була виділена саме візуалізації не лише результату експертної системи, а й відображенні методів, за допомогою яких система працює та отримує результат.

У процесі роботи візуалізація була застосована до методу класифікації та дерева рішень і відображає динаміку вибору, рішення проблеми предметної області. У розділі 4 можна детально ознайомитись з візуалізацією методів програмного застосування.

3.4. Застосування методів Data Mining при розробці ЕС підбору стилю

Тестування роботи експертної системи із зазначеними методами Data Mining

Існують 8 груп стилів, які позначаються *g*, наприклад *style id="g1"* та відносять відповідь до того чи іншого стилю.

Набір груп

g1 = "Повсякденний стиль."

g2 = "Спортивний стиль."

g3 = "Урочистий стиль."

g4 = "Базовий офіційно– діловий стиль."

g5 = "Класичний стиль."

g6 = "Повсякденний діловий стиль."

g7 = "Урочистий повсякденний стиль."

g8 = "Повсякденний сучасний стиль."

CaseНомер – відповідь на питання.

FactorНомер – запитання.

Rate = "*Число*" – ймовірність стилю, яка проставляється та сумується для кожного варіанту відповіді відповідно до виду стилю.

Style id = "НомерГрупи" – група стилю.

Результат подається у вигляді лінійного рівняння, лінійної функції виду

$$case1 + \dots caseN = style1$$

$$y = x1*k1 + \dots xN*kN$$

Алгоритм виконання

1. Беремо стиль, у якому зацікавлені.
2. Дивимось його зв'язки.
3. Далі задаємо вхідні дані.
4. Все, що не входить в ці вхідні дані – 0, усе інше сумується.
5. Отримуємо оцінку стилю:

$$case1 + \dots caseN = style1$$

Рівняння можна записати у вигляді:

$$y = x1 + \dots + xN$$

$$y = x1*k1 + \dots xN*kN$$

kN – коефіцієнт впливу N -ої відповіді на стиль

xN – може бути лише 1 або 0 в залежності від активованої відповіді

$$y1 = 0*0.1 + 1*0.3 \dots$$

У майбутньому систему можна розширювати, наприклад, шляхом додавання рівня впевненості у відповіді, і тоді буде $xN \in [0, 1]$.

Приклад

Який ваш настрій?

- поганий (0.1);
- нейтральний (0.7);
- хороший (0.2).

Припустимо, що обрано другий варіант та ще якийсь.

$$0 + 0.3 + \dots 0.7 = 1.8$$

Вхідні дані

Існують 15 запитань та 3–4 варіанти відповідей.

Жовтим кольором виділені обрані варіанти відповідей.

factor1="Чи займаєтесь ви спортом?"

case1 ="Так, регулярно"

style id="g2" rate="0.8"

case2 ="Ні"

style id="g1" rate="0.2"

case3="Інколи"

style id="g2" rate="0.4"

case4 ="Декілька разів на тиждень"

style id="g2" rate="0.6"

factor2="Чим ви займаєтесь взагалі?"

case1="Навчання"

style id="g5" rate="0.7"

case2 ="Робота"

style id="g6" rate="0.5"

case3="Жоден з варіантів"

style id="g1" rate="0.1"

factor3="Якщо ви працюєте, який тип роботи у вас?"

case1="Вимагає дрес– код та класики"

style id="g4" rate="0.5"

case2="Вимагає простоти та комфорту"

style id="g6" rate="0.4"

case3="Вимагає спортивного вбрання"

style id="g2" rate="0.4"

case4="Не працюю"

style id="g1" rate="0.1"

factor4="Який макіяж вам подобається найбільше?"

case1="Вечірній"

style id="g3" rate="0.4"

case2="Нюдодовий– натуральний"

style id="g1" rate="0.2"

case3="Яскравий"

style id="g7" rate="0.2"
case4="Не використовую косметику взагалі"
style id="g2" rate="0.0"
factor5="Яка подія очікує вас найближчим часом?"
case1="Урочисте (весілля, випускний, день народження)"
style id="g3" rate="0.6"
case2="Похід до кінотеатру/ресторану"
style id="g7" rate="0.6"
case3="Звичайна прогулянка"
style id="g1" rate="0.3"
case4="Відсутнє"
style id="g1" rate="0.0"
factor6="Який стиль ви оберете або маєте особисту перевагу?"
style id="g2" rate="0.8"
case2="Повсякденний"
style id="g1" rate="0.8"
case3="Урочистий"
style id="g3" rate="0.8"
case4="Діловий"
style id="g4" rate="0.8"
factor7="Які кольори вам найбільше подобаються?"
case1="Яскраві"
style id="g3" rate="0.3"
case2="Ніжні"
style id="g7" rate="0.2"
case3="Темні"
style id="g5" rate="0.3"
case4="Світлі"
style id="g8" rate="0.2"
factor8="Як ви вважаєте, чи маєте 'жагу до стилю'?"

case1" name="Так"
style id="g8" rate="0.5"

case2" name="Частково"

style id="g5" rate="0.4"

case3" name="Ні"
style id="g1" rate="0.2"

factor9="Який тип фігури ви маєте?"

style id="g1" rate="0.3"

case2="Середній/стрункий"

style id="g8" rate="0.4"

case3" name="Зайва вага"

style id="g5" rate="0.2"

factor10="Чи вдягаєте спідниці/плаття або маєте перевагу до
штанів/джинсів?"

case1="Перевага до спідниць"

style id="g3" rate="0.5"

case2="Перевага до плаття"

style id="g7" rate="0.4"

case3="Перевага до штанів"

style id="g2" rate="0.3"

case4="Перевага до джинсів"

style id="g1" rate="0.6"

factor11="Чи подобаються вам аксесуари?"

case1="Так"

style id="g8" rate="0.7"

case2" name="Ні"

style id="g5" rate="0.2"

case3="Інколи використовую"

style id="g3" rate="0.5"

factor12="Чи подобається вам біжутерія/прикраси?"

case1="Так"

style id="g8" rate="0.7"

case2="Ні"

style id="g5" rate="0.2"

case3="Інколи використовую"

style id="g3" rate="0.5"

factor="Ви зазвичай вдягаєте облягаючий одяг?"

case1="Так"

style id="g3" rate="0.4"

case2="Ні"

style id="g5" rate="0.1"

case3="Інколи"

style id="g1" rate="0.2"

case4="Перевага до мішкуватого та over-size"

style id="g8" rate="0.7"

Вихідні дані

$case1 + case1 + case4 + case2 + case4 + case2 + case2 + case2 + case2 + case4 + case3 + case3 + case4 + case4 + case1 = 0 + 0 + 0.1 + 0.2 + 0.0 + 0.8 + 0 + 0 + 0 + 0.6 + 0 + 0 + 0 + 0 + 0.5 = 2.2$

$y = 0.8*0 + 0.7*0 + 0.1*1 + 0.2*1 + 0.0*1 + 0.8*1 + 0.2*0 + 0.4*0 + 0.4*0 + 0.6*1 + 0.5*0 + 0.5*0 + 0.7*0 + 0.2*0 + 0.5*1 = 2.2$

Отже, експертна система рекомендує стиль №1 – *повсякденний*, про що свідчить найбільший коефіцієнт та варіанти відповіді, які відносять саме до цього стилю.

Таким чином, в процесі тестування роботи були отримані результати, що свідчать про підвищення якості та точності роботи задачі підбору стилю за допомогою алгоритмів класифікації, застосування методів дерева рішень та візуалізації; отримано рекомендаційну ЕС, яка в результаті надає точні рекомендації для прийняття рішення користувачами.

РОЗДІЛ 4. РОЗРОБКА ПРОГРАМНОГО ЗАСТОСУНКУ

4.1. Обґрунтування вибору платформи для розробки

.NET – це безкоштовна крос-платформа для розробників з відкритим кодом для створення багатьох різних типів програмних засобів.

За допомогою .NET можна використовувати декілька мов, редакторів та бібліотек для веб, мобільних пристроїв, настільних комп'ютерів, ігор.

Потужний інструментарій

Середовище бібліотеки базових класів містить необхідний функціонал для вирішення будь-яких задач будь-якої складності.

Поділ коду

В .NET використовуються збірки, які є способом поділу коду між програмними продуктами. Більш того, .NET – єдина платформа під різні потреби та мови.

Visual Studio

.NET поставляється разом із середовищем розробки Visual Studio, що надає необхідний інструментарій для ефективною та швидкою розробки застосунків із простим та легким графічним інтерфейсом. Visual Studio – основне середовище розробки.

4.2. Обґрунтування та вибір базових програмних засобів, технологій розробки

Під час планування розробки програмного продукту, було вирішено використати наступні технології та засоби:

- 1) C# – основна мова програмування.
- 2) .NET 5.

Кафедра КІТ (47)				НАУ 20 13 79 000 ПЗ			
Виконав	Литовченко О.В.			Розробка програмного застосунку	Літера	Аркуш	Аркушів
Керівник	Воронін А.М.					53	18
Консульт							
Н.контр.	Райчев І.Е.					УС-211М	122
							53

3) WPF – для розробки інтерфейсу користувача.

4) Microsoft Windows .NET 5 Core Runtime.

Мова C# була обрана з декількох причин. По-перше, для виконання даного індивідуального завдання ця мова є основною при розробці програмних застосунків.

По-друге, C# є актуальним на ринку праці серед технологій розробки; активно розвивається та націлена на різноманітні цілі використання.

Вибір другої технології був покликаний тим, що вона дозволяє створити бібліотеки класів, котрі можуть бути використані на всіх реалізаціях .NET.

Технологія WPF (Windows Presentation Foundation) є частина системи платформи .NET 5 і являє собою підсистему для побудови графічних інтерфейсів. На сьогоднішній час WPF є основним найбільш функціональним та продуманим фреймворком для створення десктопних застосунків.

Як альтернативу цій технології, можна було обрати Windows Forms, але ця платформа вважається застарілою через неактуальні засоби та набір інструментів, та не є такою актуальною як WPF.

Microsoft Windows .NET 5 Core Runtime дозволяє запускати існуючий застосунок на будь-якому пристрої Windows не потребуючи додаткових завантажень, програм та IDE.

При збірці релізу застосунків збирається цілком компактно (7-9 файлів, половина з яких стандартні бібліотеки, які входять безпосередньо в застосунок, інші в Runtime).

4.3. Опис програмної реалізації

Структура застосунку

Програмний застосунок – це один проект з двома представленнями.

Перше представлення – клієнтський інтерфейс, друге – інтерфейс візуалізації (дерева рішень). Перше представлення представляє собою застосунок для виконання задачі підбору стилю користувача та прийняття рішення. Вона

складається з вікна головної сторінки, вікна питань та відповідей, перехід до наступних питань та відповідей, що відбувається через оновлення методом.

Потім можна регулювати питання назад– вперед, останні обрані дані зберігаються. Далі слідує останнє питання та останні варіанти відповідей, та результат – діалогове вікно з класифікацією видів основних стилів та рекомендації до кожного, з розрахованим коефіцієнтом.

Друге представлення є допоміжним для візуалізації задачі підбору стилю з використанням методів Data Mining як дерево рішень. Воно складається з вікна головної сторінки, яка передбачає огляд системи, при натисканні якої з’являються усі питання та можливі відповіді, при тестуванні яких з правої частини вікна з’являється візуальне представлення методу дерева рішень, динаміка кожного вибору на подальше рішення та отримання рекомендації стилю.

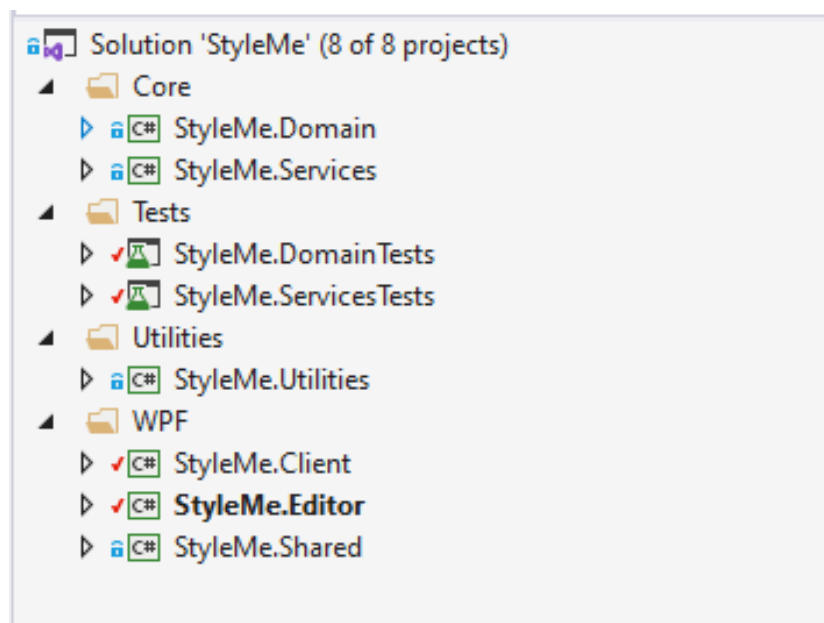


Рис. 4.1. Загальна структура проекту

Той, що має найбільший коефіцієнт – підходить найбільше. Робота експертної системи базується на автоматичному завантаженні та зчитуванні xml-файлу «styling», який доданий як ресурс.

4.4. Програмна реалізація методу дерева рішень

Велика частина – це бібліотечні методи і класи, які входять в стандартну бібліотеку .NET.

Із зовнішніх майже нічого немає – xUnit для тестів. Вони, по– перше, створюються на старті розробки. Тобто, слідуючи TDD – Test Driven Design, спочатку пишуться тести, а потім вже реалізуються методи для яких вони написані.

Юніт тести, крім цього, ще корисні при підтримці і зміні. Тобто, вносити в правку, запускати тест, якщо його статус fail – значить є помилки. Раз є помилки – треба правити. При чому, це відбувається до того, як запускається сам додаток. Юніт тести необхідні для перевірки, щоб все правильно працювало та розраховувало. Не потрібно запускати додаток та власноруч тестувати.

Клієнт потрібен для кінцевого користувача.

Редактор виконує роль візуалізатора у плані роботи.

Вони обидва використовують одні і ті ж сервіси і моделі, які визначені у відповідних бібліотеках Domain і Services.

У першій – визначено сутність-моделі.

У другій – сервіси для читання файлів, записи файлів, розрахунку оптимального стилю.

Принцип побудови метода дерева рішень

Терміни: питання – фактор, відповідь – кейс, стиль і є стиль.

Завантажується модель з файлу.

Стилі перетворюються в окрему колекцію. Фактори з кейсами – в окрему, з вкладенням (кожен фактор так само є колекцією, з кейсами). Для кожного кейса знаходяться стилі, з чого отримуємо пари ID. За допомогою цих пар формується третя колекція – зв'язок. Кожний зв'язок знає про вплив кейса, до якого він прив'язаний, на стиль. Коли активується кейс (ми відповіли на питання) – стиль додає до свого значення істинності. Якщо кейс деактивовано – віднімається.

На інтерфейсі це все представлено головною панеллю: двома вкладеними в панель блоками і одним проміжним шаром для зв'язків.

При завантаженні кожного представлення зв'язку (лінія), керуюча панель шукає відповідне представлення шару і кейса (блок, він же вузол)

Якщо один з них ще не завантажений – очікуємо.

Коли будь-якої вузол завантажується, панель перевіряє чи немає зв'язків, які очікують на його появу.

Якщо є, виконуємо прив'язку.

Прив'язка проста – у обох блоків є свої точки якірні для цього, у кейса – знизу по центру, у стилю – зверху по центру.

При будь-якому переміщенні або зміні розміру, точка оновлюється.

Так як якірна точка оголошена як Dependency Property, вона при зміні своїх значень автоматично повідомляє про це всіх, хто до неї прив'язався. В тому числі, нашу лінію зв'язку.

Так що лінію досить один раз прив'язати до кейсу та стилю, а далі вона сама буде міняти свою відрисовку, якщо раптом розмітка панелей зміниться.

Сам User Control, який представляє лінію займає повністю всю головну панель, так що йому абсолютно не важливо, де знаходяться точки прив'язки.

Метод сервісу, який виконує головні розрахунки

```
public IReadOnlyCollection<StyleEvaluationResult> Evaluate(StyleDefinitions definitions, IEnumerable<StyleFactorCaseActivation> caseActivations)
{
    var styles = definitions.Styles.ToDictionary(s => s.Id, s => 0.0);
    var cases = definitions.StyleFactors.SelectMany(sfs => sfs.Cases.Select(c => (factor: sfs.Id, @case: c))).ToDictionary(t => $"{t.factor}^{t.@case.Id}", t => t.@case);
    var activationGroups = GroupCaseActivationsByStyles(caseActivations, cases);

    foreach (var styleId in styles.Keys.ToArray())
    {
        if (activationGroups.TryGetValue(styleId, out var activatedCases))
        {
            styles[styleId] += activatedCases.Sum(c => c.AffectedStyles.First(s => s.Id == styleId).Rate);
        }
    }

    return styles.Select(kvp => new StyleEvaluationResult(kvp.Key, kvp.Value)).ToArray();
}
```

Він виконує агрегацію даних, і розраховує якому стилю скільки привласнити степінь значимості.

Допоміжний метод сервісу, який відповідає за розрахунки

```
private static IReadOnlyDictionary<string, IReadOnlyCollection<StyleFactorCase>> GroupCaseActivationsByStyles(IEnumerable<StyleFactorCaseActivation> caseActivations, IReadOnlyDictionary<string, StyleFactorCase> cases)
{
    var groups = new Dictionary<string, ICollection<StyleFactorCase>>();
    foreach (var caseActivation in caseActivations)
    {
        if (cases.TryGetValue(caseActivation.UniqueId, out var styleFactorCase))
        {
            foreach (var affectRate in styleFactorCase.AffectedStyles)
            {
                if (groups.TryGetValue(affectRate.Id, out var collection))
                {
                    collection.Add(styleFactorCase);
                }
                else
                {
                    groups.Add(affectRate.Id, styleFactorCase.AsEnumerable().ToList());
                }
            }
        }
    }
    return new ReadOnlyDictionary<string, IReadOnlyCollection<StyleFactorCase>>(groups.ToDictionary(kvp => kvp.Key, kvp => (IReadOnlyCollection<StyleFactorCase>) kvp.Value.ToArray()));
}
```

На вхід подаємо власне наші визначення факторів і стилів, з їх зв'язками (перший параметр), там же і вказані ваги зв'язків.

Другий параметр – відповіді на питання, вони ж – активовані кейси.

Зіставляємо відповіді з їх ваговими коефіцієнтами по кожному пов'язаному стилю.

У результаті отримуємо оцінку "правдивості", тобто коефіцієнт по кожному стилю, по тим класам, які працюють з представленням, все розбите на дрібні складові, навіть умовний корінь компонування невеликий.

Код XML-ресурсу «styling»

```
<?xml version="1.0"?>
<definition xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <styles>
    <style id="style1" name="Style 1" description="Here goes style 1 description"/>
    <style id="style2" name="Style 2" description="Here goes style 1 description"/>
  </styles>

  <factors>
    <factor id="factor1" name="Question or statement 1">
      <cases>
        <case id="case1" name="Possible answer 1">
          <affects>
            <style id="style1" rate="0.8" />
          </affects>
        </case>
        <case id="case2" name="Possible answer 2">
          <affects>
            <style id="style1" rate="0.2" />
            <style id="style2" rate="0.4" />
          </affects>
        </case>
      </cases>
    </factor>
    <factor id="factor2" name="Question or statement 2">
      <cases>
        <case id="case1" name="Possible answer 2">
          <affects>
            <style id="style1" rate="0.7" />
          </affects>
        </case>
        <case id="case2" name="Possible answer 2">
          <affects>
            <style id="style2" rate="0.5" />
          </affects>
        </case>
      </cases>
    </factor>
  </factors>
</definition>
```

Кожен фактор (він же питання) дозволяє задати набір можливих кейсів (відповідей), і кожен кейс може впливати на кілька стилів, для кожного з різним ступенем.

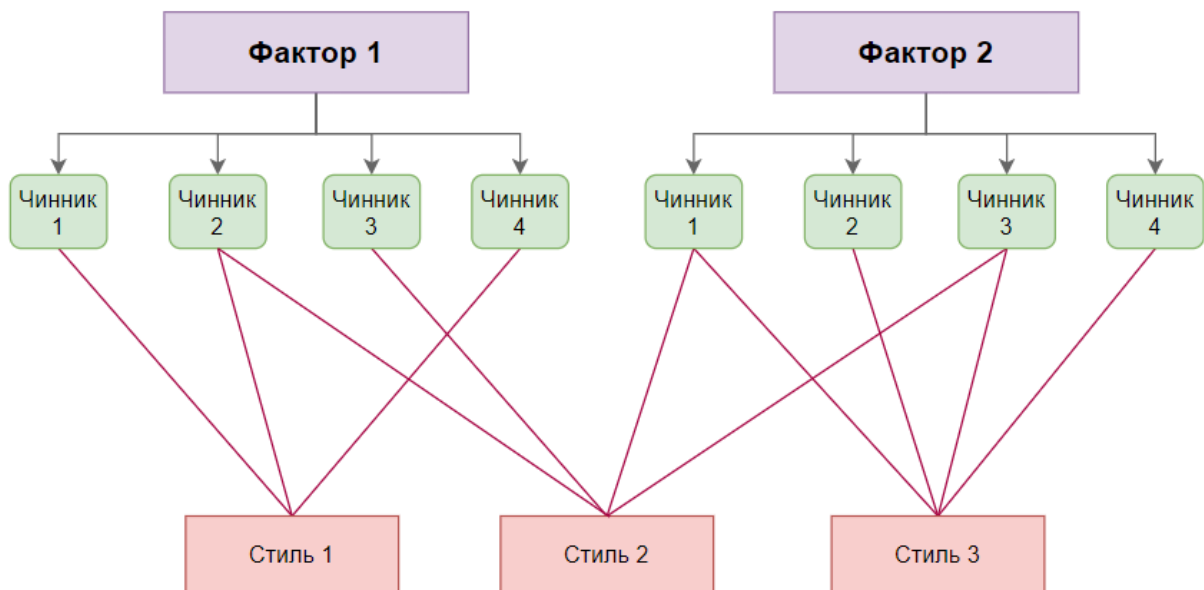


Рис. 4.2. Дерево рішень

Методи завантаження/зчитування файлу з питаннями, відповідями та запису рекомендацій підбору стилю

```
public StyleDefinitions LoadFrom(Stream stream)
{
    var serializer = new XmlSerializer(typeof(StyleDefinitions));
    return (StyleDefinitions) serializer.Deserialize(stream);
}
```

Простір імен стандартної бібліотеки .NET надає готовий сервіс System.Xml.Serialization.XmlSerializer, який виконує всю роботу,

залишається тільки розмістити серіалізовані класи потрібними атрибутами.

Майже всі моделі, які представлені в програмі, успадковуються від основних базових класів.

```

[XmlAttribute("style", Namespace = "element")]
12 references
public sealed class Style : NamedEntity, IEquatable<Style>
{
    4 references
    public Style() : base(string.Empty, string.Empty)
    {
        Description = string.Empty;
    }

    4 references
    public Style(string id, string name, string description) : base(id, name)
    {
        Description = description;
    }

    [XmlAttribute("description")]
    12 references
    public string Description { get; set; }

    4 references
    public override string ToString() => $"Style [{Id}] - {Name}. {Description}";

    4 references
    public override bool Equals(object obj)
    {
        return obj is Style other && Equals(other);
    }

    10 references
    public bool Equals(Style other)
    {
        return Id == other.Id && Name == other.Name && Description == other.Description;
    }

    23 references
    public override int GetHashCode()
    {
        var hash = 13;

        hash = (hash * 7) + Id.GetHashCode();
        hash = (hash * 4) + Name.GetHashCode();
        hash = (hash * 3) + Description.GetHashCode();

        return hash;
    }
}

```

```

public abstract class NamedEntity : Entity
{
    0 references
    public NamedEntity() : base(string.Empty)
    {
        Name = string.Empty;
    }

    6 references
    public NamedEntity(string id, string name) : base(id)
    {
        Name = name;
    }

    [XmlAttribute("name")]
    38 references
    public string Name { get; set; }
}

```

```

public abstract class Entity
{
    public Entity()
    {
        Id = string.Empty;
    }

    -references
    public Entity(string id)
    {
        Id = id;
    }

    [XmlAttribute("id")]
    -references
    public string Id { get; set; }
}

```

4.5. Алгоритм роботи експертної системи

При ініціалізації системи завантажується файл з питаннями styling.xml. При старті тестування програма по черзі перебирає питання з файлу виводячи вміст на екран. Питання розділені на групи. Кожна відповідь у свою чергу має значення ймовірності "value" зі значенням від 0 до 1, яке накопичуючись, визначає подальший вибір групи – тобто виду стилю та рекомендацію підбору до нього.

Паралельно цьому з лівої сторони застосунку будується дерево рішень, де можна прослідкувати кожен крок та етап, динаміку прийняття власного рішення. На інтерфейсі це все представлено головною панеллю: двома вкладеними в панель блоками і одним проміжним шаром для зв'язків.

Після того як користувач відповість на питання, програма на підставі накопичених балів вибере відповідну групу і таким чином відбувається поява рекомендацій стилю з прорахованими проставленими коефіцієнтами та побудоване дерево рішень. Відповідаючи на питання, користувач додає або віднімає значення на користь того чи іншого варіанта. Ці значення також грають роль коефіцієнта на екрані результатів.

4.6. Опис інтерфейсу користувача та результат роботи програми

При старті системи користувач бачить повідомлення з привітанням, у нижньому кутку кнопка «Почати»/«Огляд системи», по натисканню на яку розпочинається тестування.

На наступному кроці у вікні програми з xml-файлу «styling» (база знань) виводиться питання та список доступних варіантів відповіді.

Після вибору програма переходить до наступного питання, до тих пір, поки список не дійде до кінця. Після чого система відображає результати тестування у форматі переліку відповідних видів стилю з урахуванням коефіцієнтів. Стили виводяться всі, щоб користувач міг обирати на будь-який смак з нижчим коефіцієнтом.

Для переходу до наступного питання використовується кнопка «Далі». У кінці тестування на останньому питанні після натискання кнопки «Далі», виведуться результати тестування.

Результат роботи в представленні інтерфейсі користувача містить лише текстовий варіант рекомендації, а візуалізатор вже відображає дерево рішень для отримання та візуального відображення користувачеві найбільш точного рішення і динаміки вибору під час тестування. Можна обирати представлення в залежності від потреби.

Результати роботи програмного застосунку щодо визначення та підбору стилю та його тестування див. рис.4.3. – 4.25.

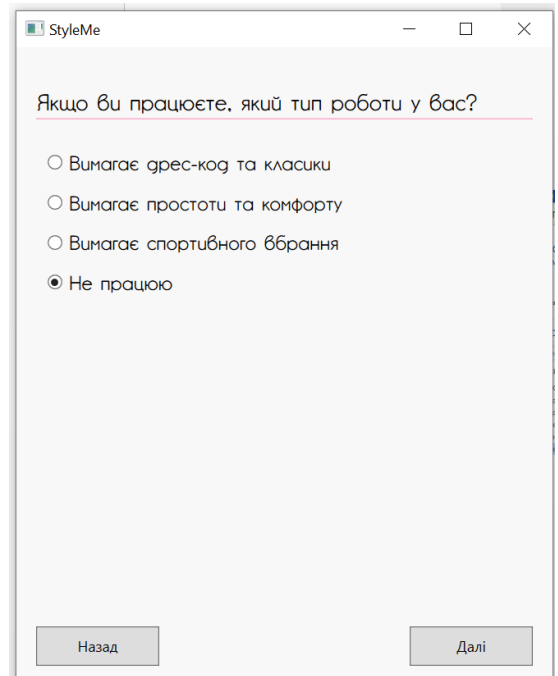
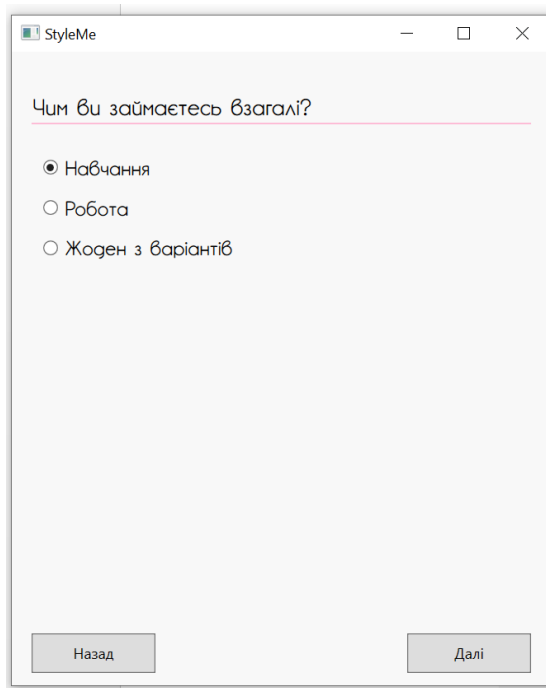
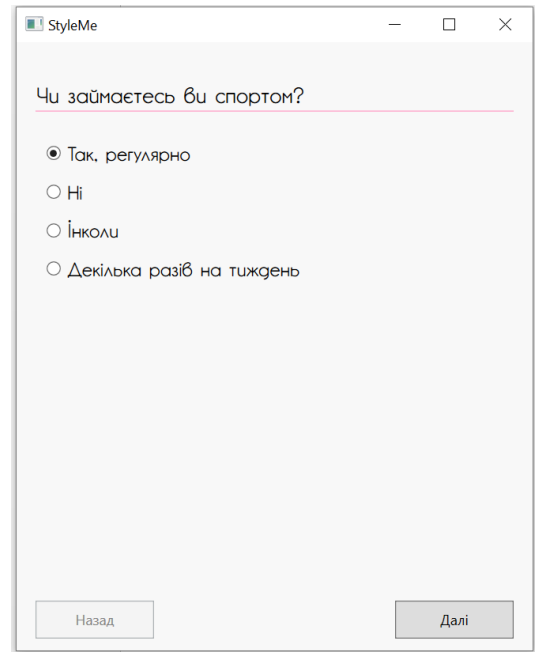
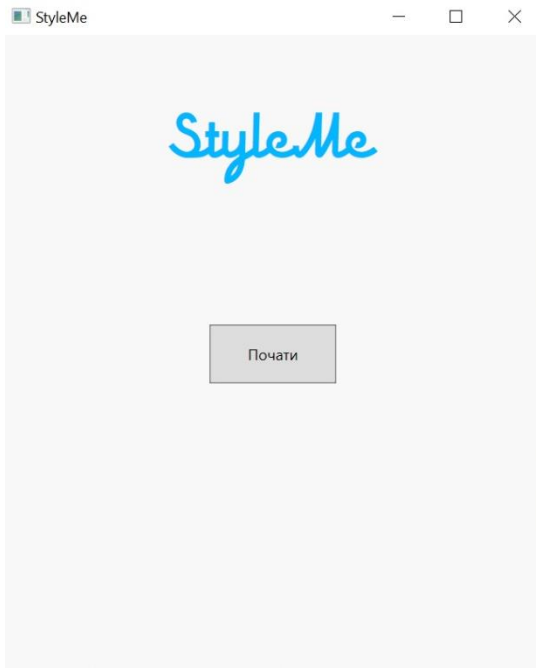


Рис. 4.3. - 4.6. Результат роботи застосунку у вигляді тестування

StyleMe

Яка подія очікує вас найближчим часом?

- Урочисте(весілля, випускний, день народження)
- Похід до кінотеатру/ресторану
- Звичайна прогулянка
- Відсутнє

Назад Далі

StyleMe

Як ви вважаєте, чи маєте 'жагу до стилю'?

- Так
- Частково
- Ні

Назад Далі

StyleMe

Який стиль ви оберете або маєте особисту перевагу?

- Спортивний
- Повсякденний
- Урочистий
- Діловий

Назад Далі

StyleMe

Який тип фігури ви маєте?

- Худощавий
- Середній/стрункий
- Зайва вага

Назад Далі

Рис. 4.6. - 4.10. Результат роботи застосунку у вигляді тестування

StyleMe

Яку зачіску вважаєте найкомфортнішою, або віддаєте найбільшу перевагу?

- Зібране волосся, наприклад у пучок або конський
- Коротка зачіска, наприклад каре
- Розпущене волосся
- Неохайна натуральна зачіска

Назад Далі

StyleMe

Ви зазвичай вдягаєте облягаючий одяг?

- Так
- Ні
- Іноколи
- Перевага до мішкуватого та over-size

Назад Далі

StyleMe

Чи вдягаєте спідниці/плаття або маєте перевагу до штанів/джинсів?

- Перевага до спідниць
- Перевага до плаття
- Перевага до штанів
- Перевага до джинсів

Назад Далі

StyleMe

Які кольори вам найбільше подобаються?

- Яскраві
- Ніжні
- Темні
- Світлі

Назад Далі

Рис. 4.10. - 4.14. Результат роботи застосунку у вигляді тестування

StyleMe

Який у вас настрій?

- Нейтральний
- Поганий
- Гарний, веселий

Назад Далі Завершити

StyleMe

Який макіяж вам подобається найбільше?

- Вечірній
- Нюдовий-натуральний
- Яскравий
- Не використовую косметику взагалі

Назад Далі

StyleMe

Чи подобається вам біжутерія/прикраси?

- Так
- Ні
- Іноколи використовую

Назад Далі

StyleMe

Чи подобаються вам аксесуари?

- Так
- Ні
- Іноколи використовую

Назад Далі

Рис. 4.14. - 4.18. Результат роботи застосунку у вигляді тестування

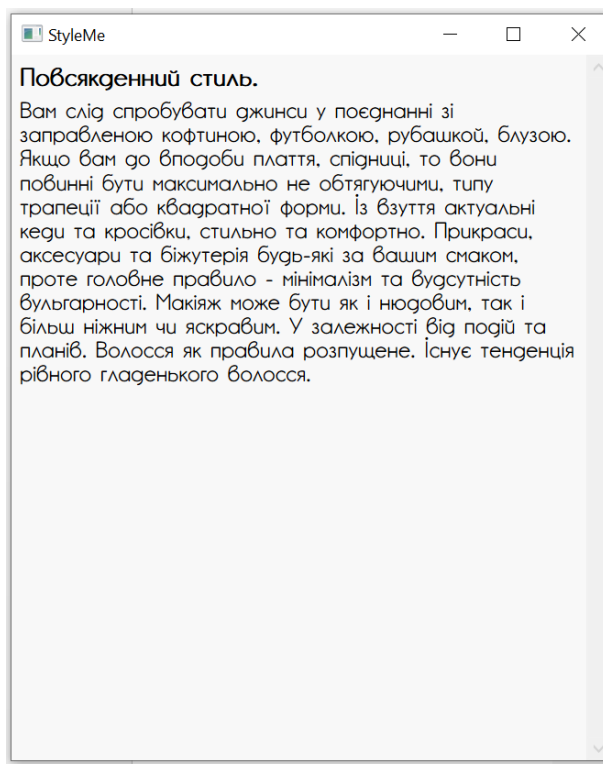


Рис. 4.19. Клієнт застосунку. Результат роботи програми у вигляді рекомендації

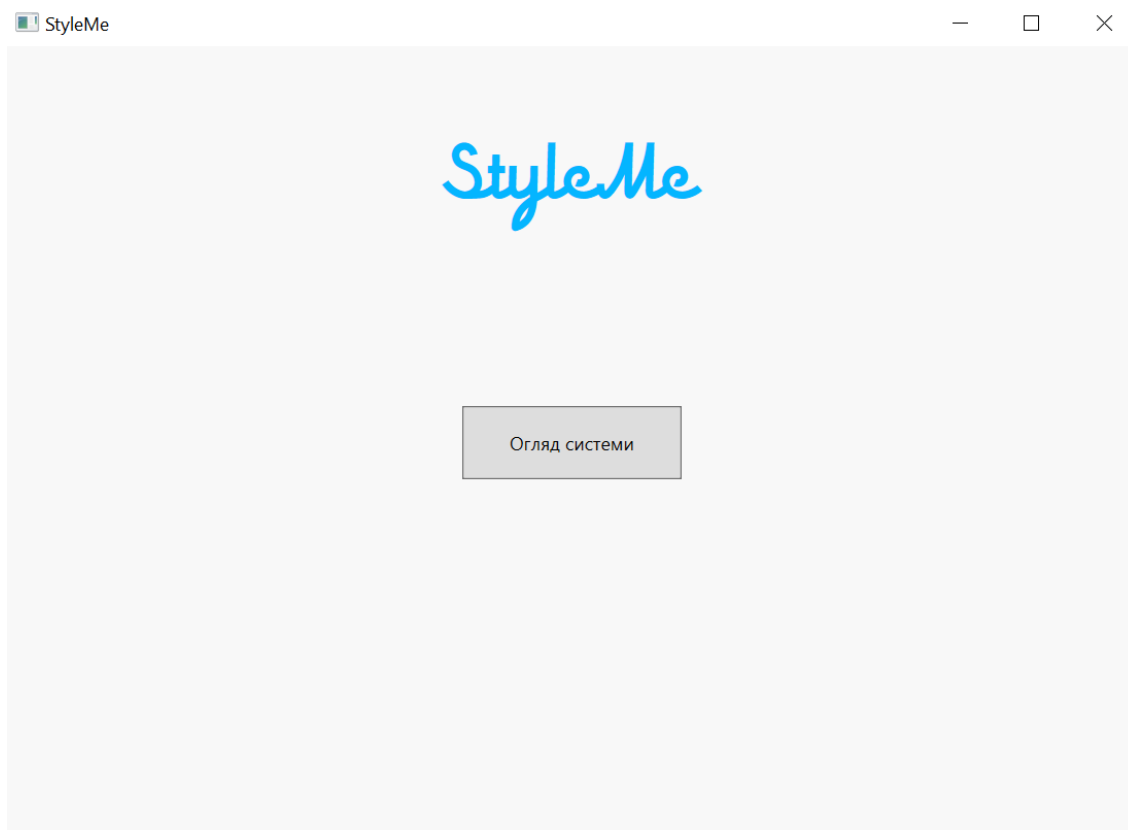


Рис. 4.20. Візуалізатор застосунку. Головне меню

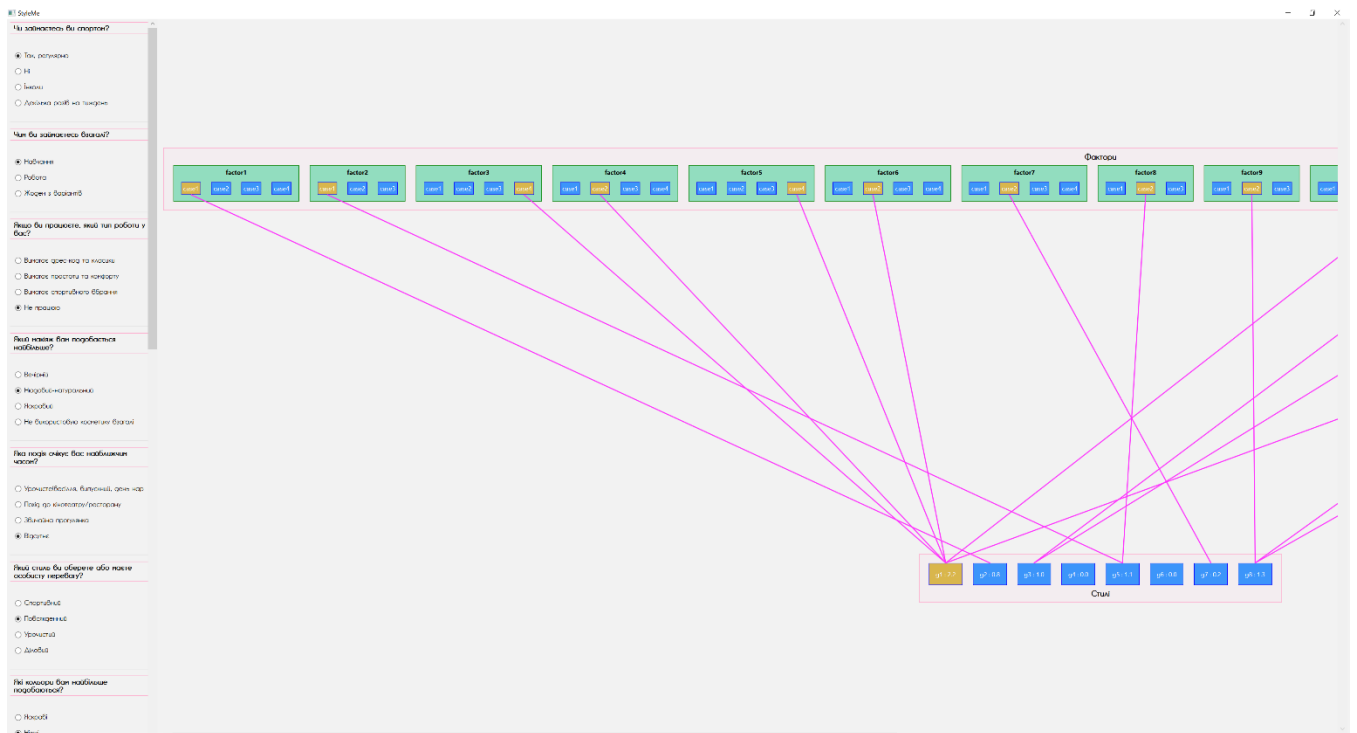


Рис. 4.21. Візуалізатор застосунку. Результат роботи програми

При наведенні курсором миші по фактору (питанню) у значенні кейсу (відповідь) або фактору (запитання) можна побачити текстовий варіант для зручності див. рис. 4.22.

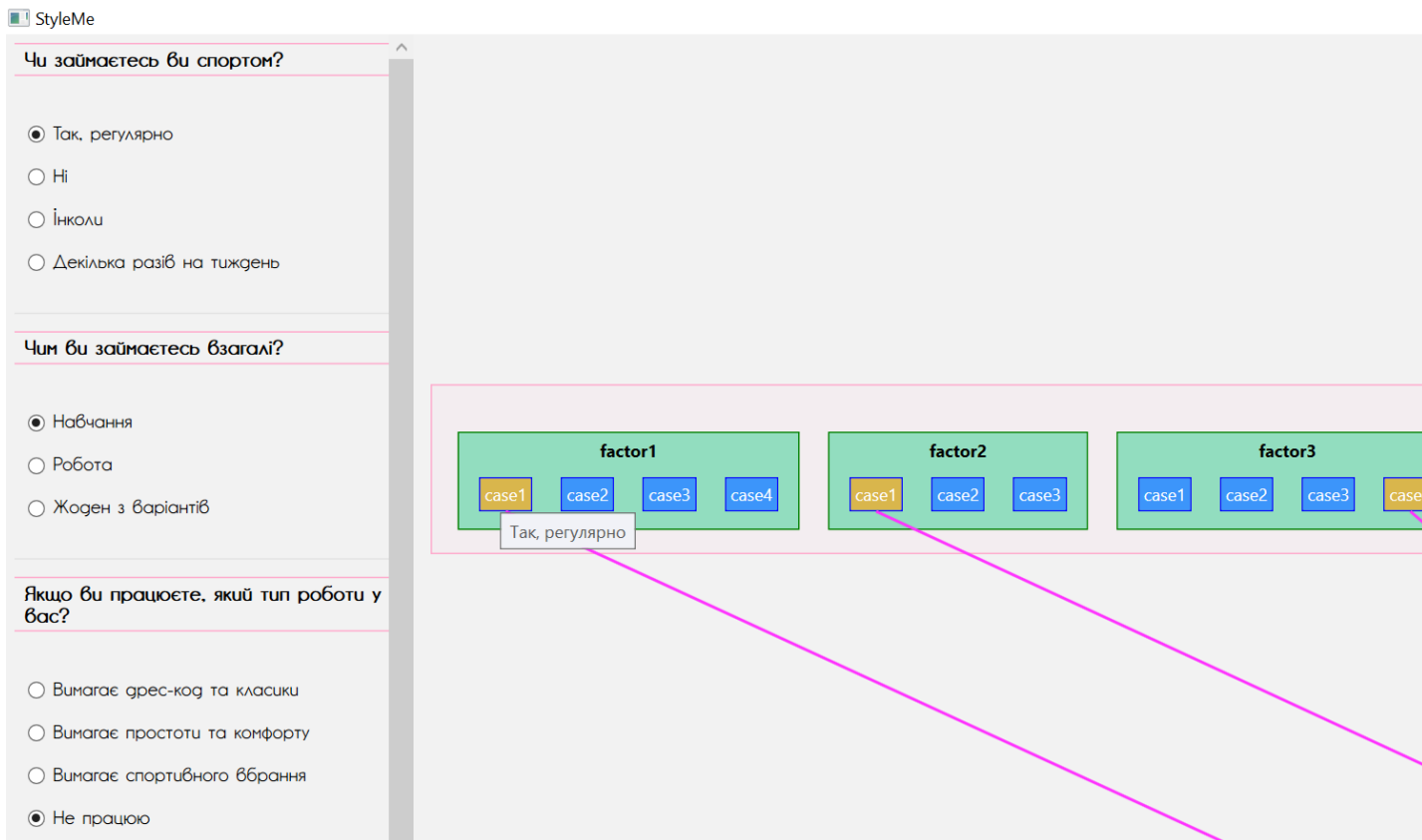


Рис. 4.22. Візуалізатор застосунку

При наведенні курсором миші по стилю у значенні варіанту g можна побачити текстовий варіант рекомендації підбору стилю для зручності див. рис. 4.23.

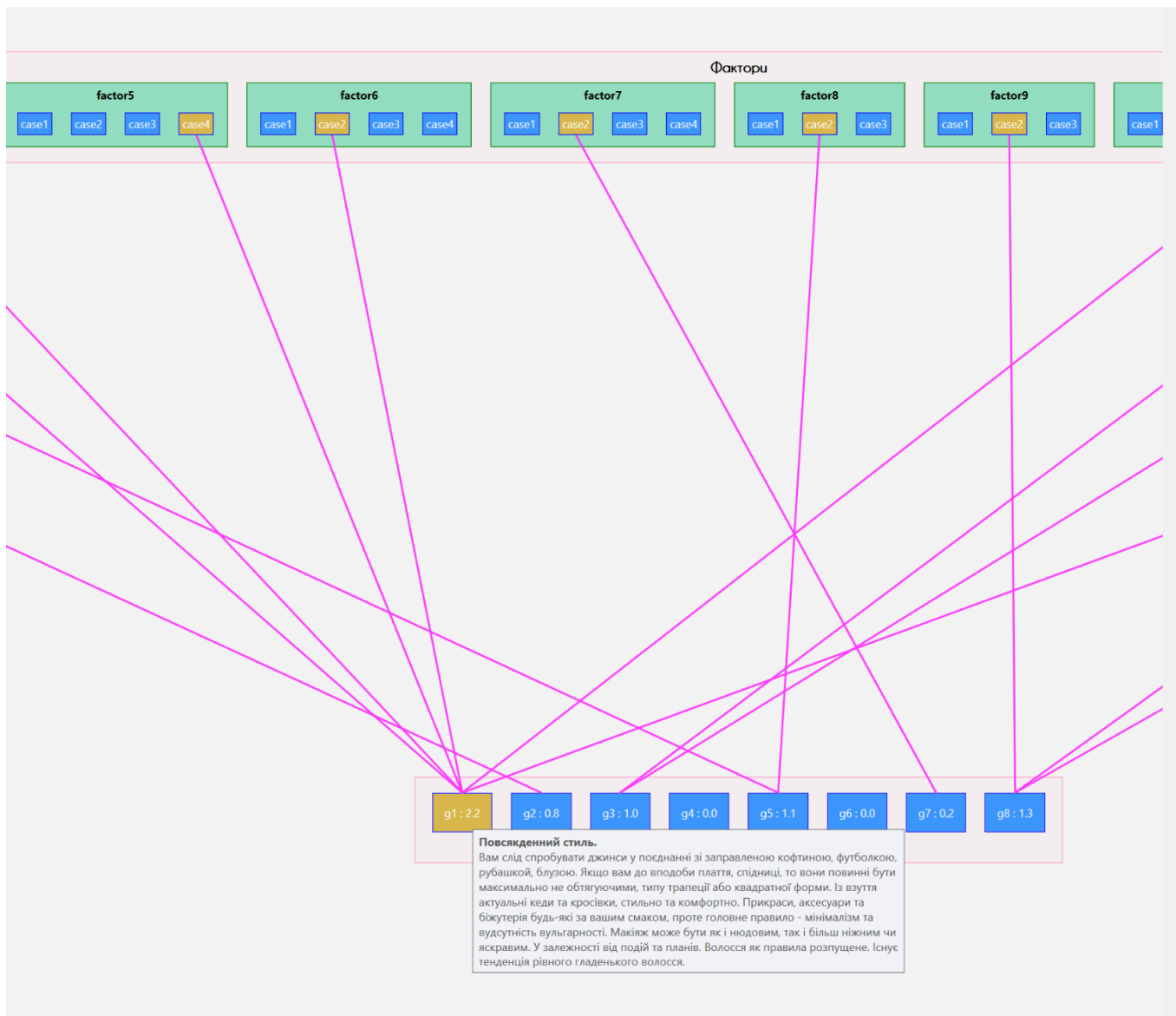


Рис. 4.23. Візуалізатор застосунку

ВИСНОВКИ

У даній магістерській науковій роботі було досліджено застосування технології Data Mining при проектуванні експертних систем рекомендаційного типу та поліпшення процесу прийняття рішень в задачі підбору стилю на основі методів Data Mining.

У першому розділі було проведено аналіз предметної області, а саме дослідження стану задачі підбору стилю, проведення огляду існуючих аналогів програмних засобів, публікацій та досліджень вітчизняного та іноземного походження, дослідження існуючих методів, алгоритмів та підходів Data Mining для використання в роботі.

У другому розділі було досліджено загальні риси та складові поняття ЕС, а саме її компоненти, класифікацію, методи представлення знань; також проведено аналіз методів та підходів для проектування та розробки структури експертної системи підбору стилю, проаналізовано їх переваги та недоліки, розроблено власну систему з базою правил та знань.

У третьому розділі було розроблено модель ЕС з використанням методів Data Mining як дерева рішень, візуалізація, класифікація, для поліпшення точності прийняття рішень. Більш того, визначено основні складові та алгоритми роботи вказаних методів. Також було проведено тестування роботи системи за допомогою вхідних даних задля отримання і порівняння результатів із програмним застосунком.

У четвертому розділі було розроблено програмний застосунок ЕС підбору стилю із застосуванням методів DM, проаналізувано отримані результати роботи.

У результаті виконання дипломної роботи був набутий досвід самостійної науково-дослідної роботи для вирішення поставлених завдань, згідно теми «Технологія застосування методів Data Mining при розробці рекомендаційних експертних систем для підбору стилю».

Отже, таким чином були визначені основні позитивні риси роботи як покращення працездатності, ефективності та точності роботи ЕС підбору стилю шляхом застосування методів Data Mining, визначені такі недоліки роботи: необхідне доопрацювання бази знань та структури ЕС, застосування більш складного алгоритму розрахунків для прийняття рішень та побудови дерева рішень.

ПЕРЕЛІК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ

1. Chouprina N.V.: Characteristics of «fast fashion» concept in fashion industry, *Vlakna a textil* 21(1), 2014, pp. 31– 36.
2. Zakharkevich O.V., Pochuprin A.V.: Development of prototype of expert system for rapid change in production of women's outerwear, *Easter– European Journal of Enterprise Technologies* 2/2(68), 2014, pp. 50– 55, <https://doi.org/10.15587/1729–4061.2014.23327>.
3. Nada Y.A., Meshref H.: Analysis, design and implementation of intelligent expert system for clothes style selection, *International Journal of Computer Applications* 105(4), 2014, pp. 15– 20.
4. Santos M.: An expert system to support clothing design process, *ACM Digital Library*, 2007, Retrieved from <http://dl.acm.org/citation.cfm?id=1784393>.
5. Kuleshova S.G., Zakharkevich O.V., Koshevko J.V., Ditkovska O.A.: Development of expert system based on Kansei Engineering to support clothing design process, *Vlakna a textil* 24(3), 2017, pp. 30– 41.
6. Turban, Efraim. Ramesh Shard, and Dursun Delen. "Decision Support System and Business Intelligent Systems", Prentice Hall. 9th Edition., 2011.
7. Keith D., "The Essence of Expert Systems", Prentice Hall, 2000.
8. Gruber, T. "Toward principles for the design of ontologies used for knowledge sharing". In *Formal Analysis in Conceptual Analysis and Knowledge Representation*. Kluwer 1993.
9. Giarratano, Joseph. Riley, Gary., "Expert Systems Principles and Programming", International Thompson Publishing. 3rd Edition, 1994.
10. Дж. Ф. Люгер. Искусственный интеллект: стратегии и методы решения сложных проблем. 4-е издание: Пер. с англ. – М.: ИД.
11. Практическое введение в технологию искусственного интеллекта и экспертных систем с иллюстрациями на Бэйсике / Р.Левин, Д.Дранг, Б.Эдельсон: Пер.с англ. – М.: Финансы и статистика, 1990 – 239 с.
12. П. Джексон. Экспертные системы. – М.: ИД «Вильямс», 2001. – 609 с.
13. Ситник В.Ф. та ін. Основи інформаційних систем: Навчальний

посібник. – К.: КНЕУ, 2001. – 420 с.

14. Алексеев А.А., Костіна Н.І., Кононець О.Я. Фінансово– економічні експертні системи. Навчальний посібник // За ред. Н.І. Костіної. – К.: Видавничий дім „Скарби”, 2004. – 208 с.

15. Уотермен Д. Руководство по экспертным системам. – М.: Мир, 1989.

16. Петров В.М., Яблонский А.И. Математика и социальные процессы. – М.: Знание, 1980.

17. Представление и использование знаний. / Под. ред. Х. Уэно, М. Исидзука: Пер. с японск. – М.: Мир, 1989. – 220 с.

18. Базы знаний интеллектуальных систем. / Т.А. Гаврилова, В.Ф. Хорошевский – СПб: Питер, 2000. – 384 с.

19. Базы и банки знаний. / Под ред. В.Н. Четвирикова. – Москва: Высшая школа, 1992. – 348 с.

20. Нейлор К. Как построить свою экспертную систему: Пер. с англ. – М.: Энергоатомиздат, 1991. – 286 с.

21. Дж. Элти, М. Кумбс. Экспертные системы: концепции и примеры / Пер. с англ. и предисл. Б.И. Шитикова. – М.: Финансы и статистика, 1987. – 191 с.

22. Разработка и применение экспертно– обучающих систем: Сб. научн. трудов. – М.: НИИВШ, 1989. – 154 с.

23. Таунсенд К., Фохт Д. Проектирование и программная реализация экспертных систем на персональных ЭВМ: Пер. с англ. – М.: Финансы и статистика, 1990.

24. Кнут Д. Искусство программирования для ЭВМ: В 3 т. – М.: Мир. – 1978.

25. Джонс М.Т. Программирование искусственного интеллекта в приложениях. Пер. с англ. Осипов А.И. – М.: ДМК Пресс, 2004. – 312 с.

26. Clancey W.J. Heuristic Classification. – Artificial Intelligence, 1985. – p. 289– 350.

27. Alexander et al. Knowledge level engineering: ontological analysis. – Proc. of National Conference on Artificial Intelligence, 1986. – p. 963– 968.

28. Winston P.H. Artificial Intelligence. – MA, Addison– Wesley, 1984.
29. Robinson J.A. Logic: Form and Function. – Edinburgh, Edinburgh Press., 1979.
30. Бойко В.В., Савинков В.П. Проектирование базы данных информационных систем. – Москва: Финансы и статистика, 1989. – 347 с.
31. Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. eds. 1996. Advanced in Knowledge Discovery and Data Mining. Cambridge, MA: AAAI/MIT Press.
32. Hui_Chuan Chu, Min_Juliao, 2011. learning case adaptation for problem-oriented e-learning on mathematics. Teaching for students with mild disabilities, journal of expert systems with applications, 38(7): 1269-1281.
33. Fernando Alonso, Loic Martinez, Aurora Perez, P. Juan Valente, 2012. Cooperation between expert knowledge and data mining discovered knowledge: Lesson learned. journal of expert systems with applications, 39: 7524-7535.

ДОДАТОК А

Styling.xml

```
<?xml version="1.0"?>
```

```
<definition xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
```

```
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
```

```
<styles>
```

```
<style id="g1" name="Повсякденний стиль." description="Вам слід спробувати джинси у поєднанні зі заправленою кофтиною, футболкою, рубашкой, блузою. Якщо вам до вподоби плаття, спідниці, то вони повинні бути максимально не обтягуючими, типу трапеції або квадратної форми. Із взуття актуальні кеди та кросівки, стильно та комфортно. Прикраси, аксесуари та біжутерія будь-які за вашим смаком, проте головне правило - мінімалізм та вудсутність вульгарності. Макіяж може бути як і нюдовим, так і більш ніжним чи яскравим. У залежності від подій та планів. Волосся як правила розпущене. Існує тенденція рівного гладенького волосся."/>
```

```
<style id="g2" name="Спортивний стиль." description="Вам слід одягти спортивний костюм, який буде не лише максимально комфортним, слід звертати свою увагу на те, щоб надавати можливість тілу дихати. Класичний спортивний костюм складається зі штанів, футболки, спортивного взуття. Популярна тенденція: дихаючі легінси в поєднанні з топом, на який одягнута майка. Макіяж відсутній, але допускається з використанням легких водостійких косметичних засобів натуральних кольорів. З аксесуарів пасують розумні годинники, відсутність біжутерії для комфорту. Волосся повинно бути зібраним за смаком: пучок, конський хвіст."/>
```

```
<style id="g3" name="Урочистий стиль." description="Такий стиль є найбільш яскравим, відкритим та може мати будь-які індивідуальні побажання. Він використовується на урочистих подіях (весілля, випускний, день народження, вечірки), вимагає максимальної святковості та урочистості. Макіяж та образ можуть бути максимально яскравими. Як правила образ складається із сукні(як
```

довгої чи пишної так і короткої) , взуття на підборах. Із волосся краще зробити зачіску в салоні, щоб та не розпалася."/>

<style id="g4" name="Базовий офіційно-діловий стиль." description="Такий стиль повинен бути максимально офіційним, без високих тенденцій моди та складатися з базових частин офіційно-ділового стилю. Як правило, він необхідний людям, які працюють в певній галузі, де є дрес-код та необхідний офіційний вигляд. Проте завжди можна виглядати максимально привабливо та стильно. Елементи модної тенденції можна комбінувати з офіційно-діловими костюмами, спідниця може бути строгою проте мати якусь особливість. Штани, піджак можуть бути підкороченими, завуженими, наприклад знизу. Аксесуар як сумка повинен бути не яскравих кольорів проте може виділятися. Волосся як правило потрібно бути зібраним."/>

<style id="g5" name="Класичний стиль." description="Такий стиль може мати елементи офіційності, проте зазвичай складається з класики, тих тенденцій які існують завжди. Як правило більшість людей, яким до вподоби класика не мають інтересу або жаги до модних тенденцій. Тому слід одягатися звичайно, мінімалістично. Такий стиль може більше підходити людям, які люблять темні кольори, або мають проблеми з фігурою, щоб якнайкраще прикрити недоліки. Такий стиль є універсальним, зазвичай мінімалістичним."/>

<style id="g6" name="Повсякденний діловий стиль." description="Такий стиль використовують для навчання, роботи. Як правило, не має чітких обмежень, проте необхідно виглядати по дрес-коду. Вбрання повинно вимагати простоти та комфорту, проте можна використовувати деякі частини повсякденного стилю, як наприклад джинси з блузою та піджаком, та кросівки. Це залежить від установи та події. Як правило занадта строгість відсутня, но її елементи повністю взаємозалежні. Аксесуари, біжутерію краще використовувати мінімально. Зачіска звичайна, як іноді розпущена так і підібрана для комфорту та офіційності."/>

<style id="g7" name="Урочистий повсякденний стиль." description="Такий стиль змішує елементи урочистого та повсякденного. Більш підходить для вечірніх прогулянок, походу до кінотеатру, якогось звичайного свята тобто подія, яка

потребує якоїсь невеличкої урочистості та жіночності. Вам слід звернути увагу на сумісні костюми, з відкритими плечима або спиною, якогось незвичного для вас кольору. Також плаття нижніх світлих кольорів, в яких комфортно та не виглядає вульгарно. Можна поєднати із взуттям на підборах, платформі. Аксесуари, біжутерія за смаком повинні бути. Макіяж більш святковий, додати деякі елементи вечірнього, проте більш ніжної гами. Волосся може бути накрученим, так і прямим."/>

<style id="g8" name="Повсякденний сучасний стиль." description="Вам слід найбільше звернути увагу на модні тенденції. Як правило, такий стиль може бути максимально неординарним, використання тенденцій одягу over-size, або протилежність - обтяжність. Краще всього поєднувати ці два компоненти, наприклад широка худі з короткими обтягуючими джинсами або спідницею. Плаття в стилі розпущеної довгої рубашки. Зі взуття оптимальні низькі підбори, лодочки, мокасіни, кросівки, кеди. Аксесуари та взуття яскравих різних кольорів. У біжутерії та аксесуарах обмежень немає. Волосся повинно виглядати гладеньким здоровим, проте ефект 'розтріпаного волосся' ніхто не відміняв."/>

</styles>

<factors>

<factor id="factor1" name="Чи займаєтесь ви спортом?">

<cases>

<case id="case1" name="Так, регулярно">

<affects>

<style id="g2" rate="0.8" />

</affects>

</case>

<case id="case2" name="Ні">

<affects>

<style id="g1" rate="0.2" />

</affects>

</case>

```
<case id="case3" name="Інколи">
  <affects>
    <style id="g2" rate="0.4" />
  </affects>
</case>
<case id="case4" name="Декілька разів на тиждень">
  <affects>
    <style id="g2" rate="0.6" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor2" name="Чим ви займаєтесь взагалі?">
  <cases>
    <case id="case1" name="Навчання">
      <affects>
        <style id="g5" rate="0.7" />
      </affects>
    </case>
    <case id="case2" name="Робота">
      <affects>
        <style id="g6" rate="0.5" />
      </affects>
    </case>
    <case id="case3" name="Жоден з варіантів">
      <affects>
        <style id="g1" rate="0.1" />
      </affects>
    </case>
  </cases>
```

```
</factor>
<factor id="factor3" name="Якщо ви працюєте, який тип роботи у вас?">
  <cases>
    <case id="case1" name="Вимагає дрес-код та класики">
      <affects>
        <style id="g4" rate="0.5" />
      </affects>
    </case>
    <case id="case2" name="Вимагає простоти та комфорту">
      <affects>
        <style id="g6" rate="0.4" />
      </affects>
    </case>
    <case id="case3" name="Вимагає спортивного вбрання">
      <affects>
        <style id="g2" rate="0.4" />
      </affects>
    </case>
    <case id="case4" name="Не працюю">
      <affects>
        <style id="g1" rate="0.1" />
      </affects>
    </case>
  </cases>
</factor>
<factor id="factor4" name="Який макіяж вам подобається найбільше?">
  <cases>
    <case id="case1" name="Вечірній">
      <affects>
        <style id="g3" rate="0.4" />
      </affects>
    </case>
  </cases>
</factor>
```

```

    </affects>
</case>
<case id="case2" name="Нюдодвий-натуральний">
  <affects>
    <style id="g1" rate="0.2" />
  </affects>
</case>
<case id="case3" name="Яскравий">
  <affects>
    <style id="g7" rate="0.2" />
  </affects>
</case>
<case id="case4" name="Не використовую косметику взагалі">
  <affects>
    <style id="g2" rate="0.0" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor5" name="Яка подія очікує вас найближчим часом?">
  <cases>
    <case id="case1" name="Урочисте(весілля, випускний, день народження)">
      <affects>
        <style id="g3" rate="0.6" />
      </affects>
    </case>
    <case id="case2" name="Похід до кінотеатру/ресторану">
      <affects>
        <style id="g7" rate="0.6" />
      </affects>

```



```
</case>
<case id="case3" name="Звичайна прогулянка">
  <affects>
    <style id="g1" rate="0.3" />
  </affects>
</case>
<case id="case4" name="Відсутнє">
  <affects>
    <style id="g1" rate="0.0" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor6" name="Який стиль ви оберете або маєте особисту
перевагу?">
  <cases>
    <case id="case1" name="Спортивний">
      <affects>
        <style id="g2" rate="0.8" />
      </affects>
    </case>
    <case id="case2" name="Повсякденний">
      <affects>
        <style id="g1" rate="0.8" />
      </affects>
    </case>
    <case id="case3" name="Урочистий">
      <affects>
        <style id="g3" rate="0.8" />
      </affects>
    </case>
  </cases>
</factor>
```

```
</case>
<case id="case4" name="Діловий">
  <affects>
    <style id="g4" rate="0.8" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor7" name="Які кольори вам найбільше подобаються?">
<cases>
  <case id="case1" name="Яскраві">
    <affects>
      <style id="g3" rate="0.3" />
    </affects>
  </case>
  <case id="case2" name="Ніжні">
    <affects>
      <style id="g7" rate="0.2" />
    </affects>
  </case>
  <case id="case3" name="Темні">
    <affects>
      <style id="g5" rate="0.3" />
    </affects>
  </case>
  <case id="case4" name="Світлі">
    <affects>
      <style id="g8" rate="0.2" />
    </affects>
  </case>
```

```
</cases>
</factor>
<factor id="factor8" name="Як ви вважаєте, чи маєте 'жагу до стилю?'">
<cases>
  <case id="case1" name="Так">
    <affects>
      <style id="g8" rate="0.5" />
    </affects>
  </case>
  <case id="case2" name="Частково">
    <affects>
      <style id="g5" rate="0.4" />
    </affects>
  </case>
  <case id="case3" name="Ні">
    <affects>
      <style id="g1" rate="0.2" />
    </affects>
  </case>
</cases>
</factor>
<factor id="factor9" name="Який тип фігури ви маєте?">
<cases>
  <case id="case1" name="Худощавий">
    <affects>
      <style id="g1" rate="0.3" />
    </affects>
  </case>
  <case id="case2" name="Середній/стрункий">
    <affects>
```

```
<style id="g8" rate="0.4" />
</affects>
</case>
<case id="case3" name="Зайва вага">
  <affects>
    <style id="g5" rate="0.2" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor10" name="Чи вдягаєте спідниці/плаття або маєте перевагу до
штанів/джинсів?">
  <cases>
    <case id="case1" name="Перевага до спідниць">
      <affects>
        <style id="g3" rate="0.5" />
      </affects>
    </case>
    <case id="case2" name="Перевага до плаття">
      <affects>
        <style id="g7" rate="0.4" />
      </affects>
    </case>
    <case id="case3" name="Перевага до штанів">
      <affects>
        <style id="g2" rate="0.3" />
      </affects>
    </case>
    <case id="case4" name="Перевага до джинсів">
      <affects>
```

```

    <style id="g1" rate="0.6" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor11" name="Чи подобаються вам аксесуари?">
  <cases>
    <case id="case1" name="Так">
      <affects>
        <style id="g8" rate="0.7" />
      </affects>
    </case>
    <case id="case2" name="Ні">
      <affects>
        <style id="g5" rate="0.2" />
      </affects>
    </case>
    <case id="case3" name="Інколи використовую">
      <affects>
        <style id="g3" rate="0.5" />
      </affects>
    </case>
  </cases>
</factor>
<factor id="factor12" name="Чи подобається вам біжутерія/прикраси?">
  <cases>
    <case id="case1" name="Так">
      <affects>
        <style id="g8" rate="0.7" />
      </affects>

```

```
</case>
<case id="case2" name="Hi">
  <affects>
    <style id="g5" rate="0.2" />
  </affects>
</case>
<case id="case3" name="Інколи використовую">
  <affects>
    <style id="g3" rate="0.5" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor13" name="Ви зазвичай вдягаєте облягаючий одяг?">
  <cases>
    <case id="case1" name="Так">
      <affects>
        <style id="g3" rate="0.4" />
      </affects>
    </case>
    <case id="case2" name="Hi">
      <affects>
        <style id="g5" rate="0.1" />
      </affects>
    </case>
    <case id="case3" name="Інколи">
      <affects>
        <style id="g1" rate="0.2" />
      </affects>
    </case>
  </cases>
</factor>
```

```

<case id="case4" name="Перевага до мішкுவатого та over-size">
  <affects>
    <style id="g8" rate="0.7" />
  </affects>
</case>
</cases>
</factor>
<factor id="factor14" name="Яку зачіску вважаєте найкомфортнішою, або
віддаєте найбільшу перевагу?">
  <cases>
    <case id="case1" name="Зібране волосся, наприклад у пучок або конський
хвіст">
      <affects>
        <style id="g4" rate="0.5" />
      </affects>
    </case>
    <case id="case2" name="Коротка зачіска, наприклад каре">
      <affects>
        <style id="g6" rate="0.4" />
      </affects>
    </case>
    <case id="case3" name="Розпущене волосся">
      <affects>
        <style id="g7" rate="0.3" />
      </affects>
    </case>
    <case id="case4" name="Неохайна натуральна зачіска">
      <affects>
        <style id="g8" rate="0.2" />
      </affects>

```

```
</case>
</cases>
</factor>
<factor id="factor15" name="Який у вас настрій?">
  <cases>
    <case id="case1" name="Нейтральний">
      <affects>
        <style id="g1" rate="0.5" />
      </affects>
    </case>
    <case id="case2" name="Поганий">
      <affects>
        <style id="g4" rate="0.1" />
      </affects>
    </case>
    <case id="case3" name="Гарний, веселий">
      <affects>
        <style id="g3" rate="0.5" />
      </affects>
    </case>
  </cases>
</factor>
</factors>
</definition>
```