

ПРОГРАМНА РЕАЛІЗАЦІЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ДЛЯ ВИРІШЕННЯ ПРОБЛЕМИ “ХОЛОДНОГО СТАРТУ”

Микола Баран¹, Вікторія Висоцька^{1,2}, Роман Голощук³

¹ Національний університет “Львівська політехніка”,
кафедра інформаційних систем та мереж, Львів, Україна

² Університет Оснабрюка, Інститут комп’ютерних наук,
Оснабрюк, Німеччина

³ Національний університет “Львівська політехніка”,
кафедра соціальних комунікацій та інформаційної діяльності, Львів, Україна

¹ E-mail: mykola.baran.mnitm.2021@lpnu.ua, ORCID: ORCID: 0000-0002-8199-0530,

² E-mail: Victoria.A.Vysotska@lpnu.ua, ORCID: 0000-0001-6417-3689

³ E-mail: Roman.O.Holoshchuk@lpnu.ua, ORCID: 0000-0002-1811-3025

© Баран М. І., Висоцька В. А., Голощук Р. О., 2023

За результатами дослідження описано та змодельовано один із підходів до побудови інтелектуальної інформаційної системи для рекомендації товарів користувачам із вирішенням проблеми “холодного старту”. В ході дослідження враховано переваги та недоліки методів, а також їхню сумісність під час їх комбінування, що є важливим фактором для забезпечення швидкодії системи та ефективності роботи алгоритму. Під час побудови інтелектуальної інформаційної системи здійснено реалізацію гібридного методу, а також тестування ефективності його роботи порівняно із класичним алгоритмом *k-means*. На основі результатів аналізу виконано практичне порівняння ефективності роботи інтелектуальної системи з базовим підходом до вирішення проблеми та гібридним.

Ключові слова: інтелектуальна інформаційна система; машинне навчання; холодний старт; *k-means*; Neural collaborative filtering.

Вступ. Загальна постановка проблеми

Фактично від самого початку історії людства так чи інакше існували силові структури. Спершу це була звичайна охорона, пізніше армії, а з розвитком державотворчих процесів з’явилися також спецслужби, органи правопорядку тощо. Для цих формувань як раніше, так і тепер важливу роль відіграє спорядження. Приблизно до 2000-х років все військове та наближене до нього спорядження виготовляли, закупували та поставляли в силові структури оборонні промисловості різних країн світу. Важливу нішу в процесах закупівлі спорядження для силових структур та армії займала Україна. З початком окупації та подальшої анексії Криму у березні 2014 р., пізніше окупації частини Донецької та Луганської областей у квітні 2014 р., а сьогодні війни із Росією, в Україні виник волонтерський рух, який має на меті допомагати українським бійцям в отриманні сучасного спорядження, амуніції, технічних засобів, транспорту тощо. Учасники більшості організацій не до кінця розуміли всі нюанси військового спорядження та речей, які необхідні для укомплектування армії, однак завдяки людям, які мали військовий досвід та необхідні знання, цю проблему вдалось вирішити за доволі короткий термін. Попри все, проблема закупівлі спорядження знову постала під час повномасштабного російського вторгнення в Україну. Оскільки потік людей був надто великим, волонтерські

організації не могли забезпечити всім необхідним кожного, хто вирішив стати на захист Батьківщини. Тому люди об'єднувались в групи та закуповували додаткове спорядження самостійно. Нерідко через те, що погляди людей ґрунтувались на кіно та літературі, вони підбирали неякісне спорядження. Сьогодні суспільство стало обізнанішим у цій сфері, однак у звичайних людей все ще виникають труднощі у пошуку тактичного, якісного, високофункціонального комплексу всього спорядження, зброї та техніки, яке гостро необхідне кожному військовослужбовцю, відповідатиме його роду діяльності, а головне – справді буде корисним та забезпечить перевагу під час виконання поставлених завдань. Отже, назріла необхідність надання людям засобу, який допоможе швидко і якісно вирішувати цю проблему. Зважаючи на розвиток інформаційних технологій та наявність у кожній людини засобу доступу до них, вирішити цю проблему зможуть усі потенційні користувачі, а засоби машинного навчання та штучного інтелекту нададуть максимально якісні рекомендації щодо товарів, необхідних їм.

Цілі цієї роботи – проектування та реалізація інтелектуальної інформаційної системи підбору тактичного спорядження для силових структур. Щоб досягти цієї мети, необхідно вирішити такі завдання:

- дослідити предметну область та проаналізувати подібні інформаційні технології, з'ясувати їхні переваги та недоліки, визначити особливості;
- здійснити системний аналіз предметної області та побудувати необхідні діаграми;
- розробити загальну типову архітектуру для подібних систем;
- проаналізувати доступні методи та засоби реалізації інформаційної системи;
- реалізувати та протестувати інтелектуальну інформаційну систему;
- виконати експериментальну апробацію розробленого прототипу системи.

Об'єкт дослідження – процес надання рекомендацій для користувачів. Предмет дослідження – засоби та методи надання рекомендацій. Практична цінність роботи полягає у реалізації інформаційної системи, яка забезпечить підбір максимально доцільного спорядження для користувача на основі даних про нього, зібраних за допомогою як явного, так і неявного підходу.

Аналіз останніх досліджень та публікацій

Сьогодні у мережі інтернет дуже багато інформації, кількість якої продовжує зростати та змінюватися з кожним днем. Для кожного користувача під час пошуку інформації виникає проблема її надлишковості та ізолюваності, оскільки пошукові системи, попри ефективне виконання завдання пошуку, дуже часто надають значно більше інформації, ніж користувач може опрацювати, або ж не можуть надати адекватної альтернативи відповідно до запитів користувача.

З огляду на зазначене вище, зростає актуальність випередження запиту користувача із пропонуванням йому потенційно цікавої інформації. Цю проблему вирішують, надаючи рекомендації. Основна відмінність алгоритмів надання рекомендацій від алгоритмів пошуку даних полягає у пропонуванні відповіді без явного запиту з боку користувача, тоді як пошукові алгоритми дають відповідь тільки на конкретний запит користувача. Безумовно, не можна стверджувати про перевагу того чи іншого підходу, тому що кожен із них доповнює інший. Пошук все ще має велике значення, однак його поєднання з рекомендаціями забезпечує набагато більше можливостей як для задоволення потреб користувача, так і для явної чи неявної взаємодії з ним. Рекомендації формують окремо для кожного користувача, спираючись на його попередні дії на основі минулої активності або на конкретному вебресурсі. Крім того, переважно значення має і поведінка попередніх учасників процесу. Концепція надання рекомендацій полягає у тому, що відповідні системи порівнюють схожі однотипні дані від різних об'єктів і виконують розрахунок списку рекомендацій для конкретного об'єкта.

Рекомендації є зручною альтернативою пошуковим алгоритмам, оскільки дають змогу виявити об'єкти, які не можуть знайти алгоритми пошуку. Варто зауважити, що системи, які ґрунтуються на рекомендаціях, часто використовують і пошукові машини.

Дуже часто під час побудови рекомендаційних систем, які генерують рекомендації певного вмісту, їх створюють на основі спрощеного підходу, наприклад, рекомендують товари тільки за попередніми вподобаннями користувача або ж за вподобаннями інших користувачів, які можуть мати зовсім інші інтереси, ніж цільовий користувач. Однак дуже часто за такого підходу якість рекомендацій надзвичайно низька, оскільки користувач може бути не зацікавлений у повторній рекомендації схожого товару або ж за великого переліку товарів чи послуг у різних категоріях. Те, що подобається одному типу користувачів, може бути зовсім недоречне для іншого типу. Такий підхід не просто знижує ефективність системи, але і зменшує зацікавленість користувачів у ній, відтак відбувається зменшення аудиторії та втрата репутації як рекомендаційної системи, так і компанії. Відповідно постає необхідність в застосуванні нових підходів для ефективнішої роботи системи та задоволення потреб цільової аудиторії.

Для вирішення цього завдання все частіше використовують інтелектуальні інформаційні системи, основані на рекомендаціях, які будують рейтинговий перелік об'єктів за певними критеріями і за допомогою яких кожен користувач може надати відповідну перевагу товару чи послугі. Перевага цих систем – залучення засобів машинного навчання та штучного інтелекту [1, 2].

Інтелектуальна інформаційна система (ІС) – один з видів автоматизованих інформаційних систем, яку інколи також називають системою, основаною на знаннях. Інтелектуальна інформаційна система – це комплекс програмних, логіко-математичних, а також лінгвістичних засобів і методів для реалізації основного завдання, яке полягає у здійсненні підтримки діяльності користувача і пошуку подальшої інформації за методом розширеного діалогу природною мовою. Інтелектуальні інформаційні системи вирішують такі завдання: діагностику; моніторинг; проектування; планування; прогнозування; навчання.

Нині ці системи демонструють, мабуть, найвідчутніший результат розвитку штучного інтелекту. Це проявляється в тому, що ми можемо зіткнутися з ними, просто зайшовши на сторінку одного з популярних інтернет-магазинів. Завдання таких систем полягає у тому, щоб за деякими ознаками, наприклад, списком товарів, які переглянув конкретний користувач, або часом перебування на сторінці з товаром визначити, які товари будуть найцікавіші й найдоречніші для користувача. Завдання пошуку доцільних товарів, як і завдання датамайнінгу, теж зводиться до навчання машини.

Загалом тепер можна виділити два чіткі види інтелектуальних систем з надання інформації, основаної на рекомендаціях. Перший принцип полягає в тому, що система основана на колаборативній фільтрації, тобто для того, щоб надати певну рекомендацію користувачу, вона переглядає наявну інформацію про те, як оцінювали потенційний об'єкт інші користувачі, за характеристиками схожі на цільового. Після цього, отримавши всю необхідну інформацію, система може спрогнозувати, наскільки високо цільовий користувач оцінить конкретний об'єкт (товар або послугу) та запропонувати його. Інший вид таких систем – системи-фільтри вмісту. Необхідною умовою для цих систем є наявність бази даних, яка повинна містити метрики всіх наявних об'єктів, та відповідно за потреби оновлювати їх. Після кількох дій користувача на сайті система уже здатна визначити, об'єкти приблизного якого типу або з якою ознакою подобаються користувачеві. Відповідно за наявності уже заздалегідь створених метрик система може підібрати нові об'єкти для рекомендування користувачеві, які будуть схожі на вже переглянуті раніше. Варто зазначити, що істотним недоліком таких систем є необхідність формування доволі великої бази даних з метриками, без яких нова система не зможе одразу розпочати надавати рекомендації користувачу, оскільки буде недостатньо початкової інформації [3]. Окрім цього, сама побудова метрики також може стати проблемою [4].

Проблема “холодного старту”. Одна із класичних проблем, яка виникає тоді, коли для роботи алгоритму ще не накопичено достатньо даних. Це доволі типова ситуація для нового або непопулярного об'єкта чи сайта загалом, який оцінила невелика кількість користувачів, або ж, навпаки, для неординарного користувача, який за вподобаннями дуже відрізняється від середньостатистичного користувача [5]. У таких випадках, якщо не застосовувати інші алгоритми, рейтинги доводиться

коригувати штучно. Наприклад, оцінку для об’єкта обчислюють не як середню за позицією, а як згладжену середню. Тобто за невеликої кількості оцінок рейтинг наблизатиметься до якоїсь “безпечної середньої”, але коли набирається достатньо реальних оцінок, то штучне усереднення не здійснюється.

Проблема “упередженості”. Іншою класичною проблемою є проблема упередженості, що виникає, коли неточно налаштовані алгоритми, закладені у якихось стереотипах, і навіть дії користувачів можуть вплинути на поточний аналіз та ранжування інформації. Окрім цього, дуже часто проблема може виникати з боку користувача, коли алгоритм не враховує ситуації, за якої користувач сам себе обмежує в категоріях, однак через це може пропустити певну значущу для нього інформацію, оскільки алгоритм чітко слідує його неявним даним (для прикладу, демографічним) та його явним вказівкам (для прикладу, фільтрам) [6].

Робіт, які стосуються розроблення та дослідження інтелектуальних інформаційних систем, оснований на рекомендаціях, способах побудови та їх проблемах, досить багато як у міжнародному просторі, так і у національному сегменті [7–18].

У дослідженні [12] розглянуто основні види рекомендаційних систем, наявні у мережі інтернет, оснований на методах контентної та колаборативної фільтрації, а також їх проблеми, із яких “до найпоширеніших проблем рекомендаційних систем можна віднести проблему холодного старту та бульбашки фільтрів” [12]. Ця проблема не нова та має багато способів вирішення. Як пишуть автори дослідження, “одним зі шляхів підвищення точності роботи рекомендаційних систем та вирішення проблем холодного старту є фільтри використання контексту, в якому відбувається фіксація вподобань користувача, та контексту, в якому відбувається формування рекомендацій” [12–18]. У цьому випадку під контекстом розуміють використання дати та часу формування рекомендацій, пристроїв, з яких користувач зайшов на вебсайт, його демографічні дані, геолокацію тощо. Завдяки прогресу, а також новим функціям, які забезпечив нам ES6+ та HTML5, отримати необхідні дані нескладно. Єдина перешкода, яка може виникнути, полягає в політиці конфіденційності браузерів, а тому деякі необхідні нам дані користувач може не надати або ця опція може бути взагалі вимкнена у нього за замовчуванням.

Ще одним зі знакових досліджень є стаття [13], автори якої наголошують, що серед алгоритмів найпоширенішими залишаються алгоритми колаборативної фільтрації та content-based filtering (CF і CBF), а також їх поєднання. Здебільшого ці алгоритми працюють на основі meta-інформації про новини. Проте в роботі відзначено їх недостатню якість безпосередньо стосовно рекомендацій новин. Останнім часом дослідники та розробники все більше звертають увагу на “навколишній контекст” та алгоритми глибокого навчання для покращення якості рекомендацій.

До “нових” підходів у розробленні NRS автори зараховують такі алгоритми:

- матрична факторизація і невід’ємна матриця факторизації;
- тензорна факторизація;
- імовірнісна матрична факторизація;
- байєсівський персоналізований рейтинг;
- загальне лінійне моделювання,

а також їх модифікації на основі нейронних мереж з використанням глибокого навчання, що особливо набуває популярності в останні роки.

У статті [14] автори розглянули попередні дослідження про теги до систем рекомендацій, щоб підвищити ефективність традиційних методів рекомендацій. В ході дослідження було визначено, що теги, які вказує користувач, зазвичай спричиняють багато проблем, таких як розрідженість, надмірність і неоднозначність. Щоб вирішити ці проблеми, автори запропонували один із нових підходів до рекомендацій на основі глибоких нейронних мереж. У запропонованому алгоритмі профілі користувачів спочатку представлені тегами, а потім модель глибокої нейронної мережі використовується для отримання глибинних характеристик із простору тегів шар за шаром. Завдяки цьому представлення

необроблених даних стають абстрактнішими та розширенішими. На їх основі профілі користувачів оновлюються та використовуються для надання рекомендацій. Експериментальні результати демонструють корисність запропонованого підходу та його вищу продуктивність, порівняно з класичними алгоритмами, що підтверджує ефективність комбінування класичних підходів до побудови та підходів, основаних на нейронних мережах і глибокому навчанні [15–18]. У всіх згаданих вище роботах особливо виділено потреби вирішення проблем холодного старту та бульбашки фільтрів. Окрім цього, варто додати ще одну поширену проблему, яка виникає як під час старту системи, так і в подальшій її роботі. Це проблема задіяння контексту та отримання даних користувача без його відома, що може, своєю чергою, створювати хибне уявлення про нього. Відповідно, у всіх роботах відзначено актуальність розроблення системи, яка враховуватиме всі наявні проблеми та буде побудована за принципом гібридності, що забезпечить масштабованість та надійність для системи.

Послуги – це не обов’язково прямий продаж запропонованого товару. Кожен сервіс, окрім надання послуг, може також отримувати комісійний дохід або просто підвищувати лояльність користувачів, яка відтак перетворюється на рекламні або інші прибутки. Відповідно, залежно від моделі бізнесу, пропозиції можуть бути як його основою (наприклад, у TripAdvisor), так і просто зручним сервісом, що є додатковим функціоналом (для прикладу, інтернет-магазині речей), який покликаний покращити Customer Experience та зробити навігацію у каталозі зручною. Враховуючи ці аспекти, розробники створили підходи, які відрізняються за поставленою метою.

Незважаючи на безліч алгоритмів побудови ІІС на основі підбору речей, всі вони зводяться до декількох фундаментальних підходів. Відповідно до найкласичніших можна зарахувати алгоритми:

- Summary-based (неперсоналізовані);
- Content-based (моделі, які ґрунтуються на описанні певного товару);
- Collaborative Filtering (колаборативна фільтрація);
- Knowledge-based filtering (фільтрація, основана на знаннях).

Фактично всі інші алгоритми є похідними від цих чотирьох та забезпечують покращену функціональність або із задіянням нових підходів та технологій, або комбінуванням цих алгоритмів.

Неперсоналізований підхід (Summary-based). Розпочнемо з неперсоналізованого підходу, оскільки він найпростіший у реалізації. За цим підходом потенційно можливий інтерес користувача встановлюють, визначаючи середній рейтинг товару: “Якщо усім подобається – відповідно сподобається і користувачеві”. За цим принципом зараз працює більшість сучасних бюджетних і не зовсім сервісів, реєстрація у системі не обов’язкова. Подаватись пропозиції можуть по-різному, наприклад, як певний банер біля опису товару або ж як фактичний результат запиту, що відсортований за певним параметром. Рейтинг для товару також може подаватись найрізноманітнішими методами. Це може бути різниця позитивних та негативних голосів (часто такий підхід реалізують на форумах), кількість вподобань, частка високих оцінок чи взагалі їх гістограма. Гістограми, звичайно, один із найінформативніших способів, але вони мають один значущий недолік, суть якого полягає у тому, що гістограми складно порівнювати між собою чи сортувати, коли необхідно подати товари списком відповідно до певних критеріїв. Саме за такого підходу і виникає поширена проблема “холодного старту”.

Персоналізований підхід (Content-based). Персоналізований підхід передбачає максимальне використання наявної інформації про користувача, насамперед про його попередні дії. Одним із перших у цій сфері з’явився підхід content-based filtering. У межах цього підходу опис товару (тобто content) зіставляється із інтересами користувача, встановленими з попередніх оцінок. Відповідно чим більше товар відповідає визначеним інтересам, тим вище оцінюється потенційна зацікавленість користувача. Очевидною вимогою є те, що у всіх товарів у каталозі повинен бути опис. Предметом Content-based найчастіше є товари з неструктурованим описом: книги, фільми, статті, ознаками є їх рецензії, текстові описи, список акторів тощо. Однак ніщо не заважає використовувати і звичайні

числа чи категорії. Опис неструктурованих ознак здійснюється типовим для тексту способом, а саме Vector-Space model. Кожен елемент вектора – це ознака, яка характеризує інтерес користувача. Відповідно продукт – це вектор, що міститься у тому самому просторі. Із урахуванням взаємодії користувача із системою векторні описи товарів, які придбав користувач, з’єднуються у єдиний вектор, що являє собою вектор його інтересів. Далі потрібно знайти товар, опис якого найближчий до вектора інтересів, тобто необхідно розв’язати задачу із пошуку n найближчих сусідів. Варто зауважити, що не всі елементи однаково значущі, тому щодо кількості елементів, які збігаються у двох векторах, всі виміри необхідно попередньо зважувати відповідно до їхньої значущості. Це завдання ефективно вирішує відоме в TextMining перетворення TF-IDF, що надає більшу вагу рідкіснішим інтересам. Збігу таких інтересів надається більше значення щодо близькості двох векторів, ніж просто збігу популярних.

Колаборативна фільтрація (Collaborative Filtering). Цей підхід широко застосовується переважно через простоту реалізації. Принцип роботи такої фільтрації насправді не є складним, хоча можна виділити два шляхи реалізації:

- Концепція, основана на зіставленні користувачів (user-based), ураховує подібність користувача щодо інших наявних користувачів.
- Концепція, що ґрунтується на зіставленні об’єктів (item-based), навпаки, виконує аналіз самих об’єктів і виявляє їх подібність на ті, які колись сподобалися користувачу.

Колаборативна фільтрація надає змогу користувачу отримати доволі доречні й точні рекомендації, основані на зіставленні та аналізуванні відмінностей у користувачів, поведінка яких схожа. Класична реалізація алгоритму ґрунтується на принципі k найближчих сусідів, який полягає в тому, що для кожного користувача здійснюється пошук найбільш схожих на нього користувачів і інформація цього користувача доповнюється, із урахуванням відомих даних про його сусідів. Отже, якщо відомо, що сусіди за інтересами захоплюються якимось фільмом, а користувач його чомусь ще не бачив, система вирішує, що це чудова нагода запропонувати йому цей фільм.

Фільтрація, основана на знаннях (knowledge-based filtering). Принцип цього підходу такий: якщо на сайті продаються будинки чи машини, покладатися на відгуки інших користувачів система не може: покупки штучні та рідкісні, а отже, інформації замало – немає статистичної значущості. Але ж ми хочемо, щоб користувач купив будинок своєї мрії у нас, а не у конкурентів? У цьому випадку можна додати фільтри: будинок у місті чи селі, кількість поверхів та квадратних метрів, матеріал стін. Після цього рекомендаційна система підбирає найвідповідніші будинки з каталогу. Такий тип рекомендацій точний: користувач вказує саме те, що він хоче бачити. Але алгоритми фільтрації, основані на знаннях, складніше продумати, тому що параметрів пошуку може бути дуже багато.

У 1979 р. Елейн Річ фактично несвідомо створила першу систему із надання рекомендацій Grundy. Вона шукала спосіб порекомендувати користувачеві книгу, яка йому може сподобатися. Її ідея полягала в тому, щоб створити систему, яка задає користувачеві конкретні питання і присвоює йому стереотипи залежно від його відповідей. Залежно від стереотипу користувача, потім йому рекомендують книгу, яка йому може сподобатися. Попри цікаву концепцію, цей підхід почали досліджувати і масово впроваджувати порівняно недавно, а саме з початком розвитку штучного інтелекту та машинного навчання. Протягом останніх кількох десятиліть, із розвитком Youtube, Amazon, Netflix та багатьох інших подібних вебсервісів, рекомендації товарів та послуг займають усе більше місця у нашому житті. Від електронної комерції (пропонувати покупцям товари, які можуть їх зацікавити) до онлайн-реклами (пропонувати користувачам правильний вміст, який відповідає їхнім уподобанням), системи рекомендацій сьогодні неминучі в наших щоденних онлайн-мандрівках. Окрім цього, ці сервіси спонукали вдосконалювати підходи до побудови із застосуванням штучного інтелекту та машинного навчання. Тепер на ринку можна знайти низку програм та інформаційних засобів зі схожими можливостями, але із різними підходами до реалізації. Для аналізу вибрано кілька найпопулярніших систем, а саме Hardware Inventory Management, Open-AudIT, Lansweeper.

Система “MollyUa” – переважно основана на власному виробництві товару. Завдяки співпраці із силовими структурами щодо функціоналу спорядження наявні якісні моделі відповідно до стандартів НАТО, якісна фурнітура та надійна тканина. Окрім цього, завдяки великим потужностям виробництва магазин характеризується оперативністю пошиття та доставки, що дуже важливо для специфіки цієї предметної області. Перевагою також є можливість пошиття речей на замовлення, що, звичайно, має як переваги, так і недоліки.

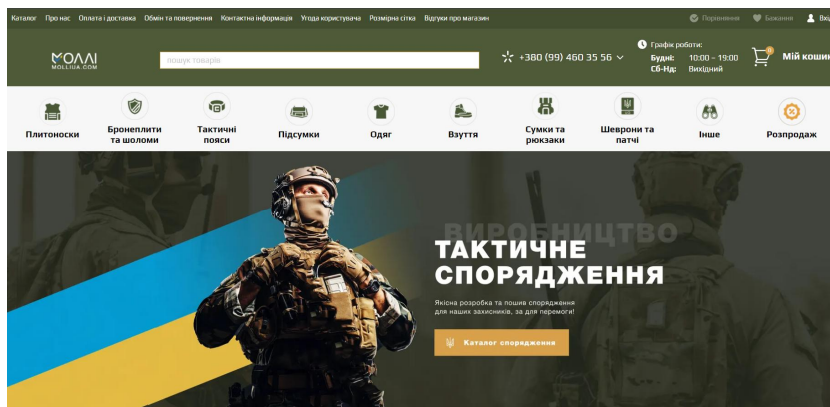


Рис. 1. Інтерфейс системи MollyUA

Переваги: система оптимізована та швидкодіюча; зручна навігація як для нових, так і для досвідчених користувачів; наявна техпідтримка та таблиці розмірів для спорядження. Недоліки: не вирішена проблема “холодного старту”; погано структурований опис товару; більшість продукції власного виробництва, що зменшує різноманітність вибору.

Militarist. Історія компанії ТМ “Мілітарист” почалася в 2005 р. з моменту відкриття першого в Україні магазину із тематики “Airsoft”. Тоді компанія пропонувала широкий асортимент страйкбольної зброї й аксесуарів до неї. Сьогодні ТМ “Мілітарист” – це мережа магазинів в Україні, яка пропонує широкий асортимент військово-тактичного спорядження і одягу від найвідоміших брендів. Також компанія представляє лінійку одягу в стилі “street wear”, “outdoor” і “casual” для людей, які ведуть активний спосіб життя. Варто зауважити, що магазин – ексклюзивний представник бренду “М-ТАС” в Україні, під яким випускають одяг, взуття, тактичне спорядження, різні аксесуари для військово-службовців, силовиків і просто любителів екстремального відпочинку. Переваги системи: хороша оптимізація та швидкодія; зручна навігація як для нових, так і для досвідчених користувачів; велика кількість товарів відомих брендів; чітко оформлений та поданий опис товару. Недоліки: за “холодного старту” товари подають тільки в одній категорії; таблиця розмірів є тільки у текстовому форматі.

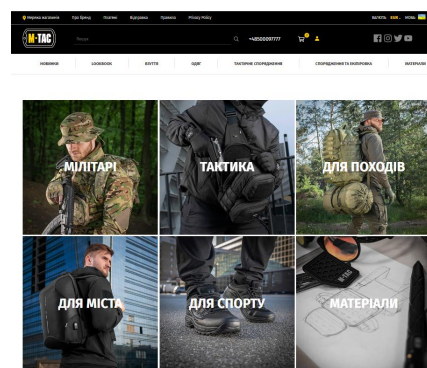
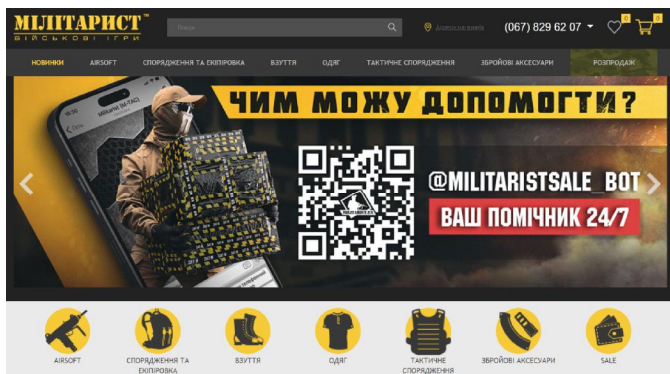


Рис. 2. Інтерфейс системи MILITARIST

М-Тас – це український виробник тактичного спорядження, взуття, одягу а також інших аксесуарів. У 2014 р. компанію заснувала команда досвідчених підприємців, які з 2005 р. Здійснювали торгівлю тактичним спорядженням та зброєю для страйкболу. У своєму виробництві компанія використовує високотехнологічні засоби та матеріали, а також сучасний дизайн. Завдяки здійсненню численних тестів і випробувань продукції учасниками спеціальних операцій і військовими інших спецпідрозділів компанія створює якісний та функціональний товар. Переваги системи: приємний та зручний інтерфейс; є функція зворотного зв’язку; чітко оформлений та поданий опис товару. Недоліки: не вирішена проблема “холодного старту”; дещо недостатній асортимент вибору; таблиця розмірів подана тільки у текстовому форматі.

Основні завдання дослідження та їх значення

Метою дослідження є визначення основних способів побудови алгоритмів із вирішенням проблеми “холодного старту”. Дослідження надасть засоби для побудови інтелектуальної інформаційної системи з урахуванням проблеми “холодного старту”. Для досягнення мети необхідно виконати такі основні завдання: визначити основні підходи до вирішення поставленої проблеми, на їх основі вибрати найдоцільніше рішення реалізації, виконавши детальний аналіз, переконатися у правильності тверджень та перевірити доцільність подальшого покращення алгоритму.

Основні результати досліджень

Згідно із загальноприйнятою методологією системного аналізу головною метою є створення інтелектуальної інформаційної системи підбору спорядження для силових структур. Це завдання можна виконати тільки після того, як будуть досягнуті всі наявні підцілі. Від вершини дерева цілей, що уособлює головну ціль, відходять три гілки, які ділять головну ціль на три підцілі, а саме:



Рис. 3. Дерево цілей інформаційної системи підбору товарів для силових структур

1. Аналіз та дослідження предметної області. Ця підціль відповідає за визначення особливостей предметної області, її аналіз з метою визначення основних переваг та недоліків, а також перешкод, що можуть виникнути під час подальшої реалізації. Ця підціль поділяється на:

- Збирання даних про інтелектуальні системи рекомендацій. Ця ціль дає змогу отримати всі необхідні дані про інтелектуальні системи, їхні основні завдання та цілі.
- Аналіз досліджень реалізації програмного та апаратного забезпечення. Надає можливість аналізувати інформацію щодо результатів роботи програмного та апаратного забезпечення у відомих дослідженнях та їх ефективності.

- Визначення основних аспектів розроблюваної системи дає можливість встановити основні вимоги, яким повинна відповідати інтелектуальна інформаційна система.

2. Проектування системи: виконання системного аналізу та формування загальної концепції системи на основі отриманих результатів дослідження. Передбачає:

- Проектування архітектури основних модулів системи. Дає змогу проаналізувати кожен із модулів системи, їхню доцільність та ефективність. Це дозволить проаналізувати всі ризики на етапі проектування та уникнути можливих проблем у майбутньому.

- Проектування механізмів комунікації компонентів. Уможливіє визначення оптимального механізму комунікацій для справної роботи та забезпечення надійної продуктивності системи;

- Проектування бази даних. Дає змогу виконати ефективне проектування, критично важливе для майбутнього отримання даних, їх структуризації та подальшої роботи з ними.

3. Програмна реалізація. Являє собою фінальну побудову та тестування системи. Цю підціль поділено, своєю чергою, на:

- Реалізацію програмного забезпечення інтелектуальної системи: створення основних компонентів, модулів та бази даних відповідно до попередньо визначених вимог, а також архітектури системи.

- Налаштування взаємодії компонентів у системі. Надає змогу здійснити налаштування і тестування модулів та засобів системи для належної роботи системи.

- Реалізацію графічного інтерфейсу системи. Уможливіє створення зрозумілого та зручного інтерфейсу для користувача з метою ефективної взаємодії із системою.

Для розроблення інтелектуальної інформаційної системи необхідно також додатково проаналізувати системи, що відповідають основним критеріям мети, та на основі отриманих даних вибрати відповідний тип системи. Критерії, на основі яких реалізуватиметься метод аналізу ієрархій (MAI), такі: надійність; актуальність; достовірність; продуктивність; функціональність; легкість у використанні. Окрім цього, для подальшого аналізу необхідно вибрати основні, найдоцільніші типи систем. Згідно із критеріями вибрано чотири основні типи системи, а саме:

- інформаційно-пошукова система. Основна функція системи полягає у проведенні пошуку інформації за певним сформованим запитом, який задав користувач;

- інформаційно-рекомендаційна система. Основна функція цієї системи – надання рекомендацій користувачу відповідно до його пріоритетів, із попереднім аналізом його активності, а також активності інших користувачів системи;

- інформаційно-довідкова система. Являє собою систему для збирання та подальшого опрацювання інформації, яка потім надається користувачеві у визначеному форматі;

- інформаційна система підтримки прийняття рішень. Ключовою особливістю системи є здатність здійснювати явний або неявний вплив на прийняття та ухвалення рішень, беручи за основу отримані дані.

Першим етапом є побудова ієрархії для проектованої системи, що зображена на рис. 4.



Рис. 4. Ієрархія MAI з вибору типу системи

Відповідно до підходу МАІ сформовано матрицю порівнянь. Ця матриця надає змогу зрозуміти важливість та взаємозалежність кожного критерію стосовно іншого. Результати порівняння із оцінками порівняння та векторами пріоритетів подано в табл. 1.

Таблиця 1

Матриця попарних порівнянь критеріїв

№	Критерії	1	2	3	4	5	6	Оцінка	Вектор пр.
1	Актуальність	1	1	0,33	2	1	0,5	0,83	0,12
2	Достовірність	1	1	0,33	0,5	1	0,5	0,66	0,1
3	Продуктивність	3	3	1	4	3,03	2	2,45	0,36
4	Функціональність	0,5	2	0,25	1	2	2	1	0,15
5	Надійність	1	1	0,33	0,5	1	0,5	0,66	0,1
6	Легкість у використанні	2	2	0,5	0,5	2	1	1,12	0,17

Наступний крок полягає в оцінюванні визначених типів інформаційних систем відповідно до зазначених ознак. Для цього будуємо матриці попарних порівнянь за кожним із критеріїв та формуємо подальші оцінки і вектори пріоритетів. Отримані результати наведено у відповідних табл. 2, 3.

Таблиця 2

Матриці попарних порівнянь

№	Тип системи	1	2	3	4	Оцінка	Вектор пріоритетів
за критерієм “Функціональність”							
1	Інформаційно-довідкова система	1	1	0,33	0,25	0,54	0,11
2	Інформаційно-пошукова система	1	1	0,33	0,25	0,54	0,11
3	Інформаційно-рекомендаційна система	3	3	1	0,75	1,61	0,33
4	Інформаційна система підтримки прийняття рішень	4	4	1,33	1	2,15	0,44
за критерієм “Актуальність”							
1	Інформаційно-довідкова система	1	0,33	0,17	0,5	0,41	0,08
2	Інформаційно-пошукова система	3	1	0,5	1,33	1,19	0,24
3	Інформаційно-рекомендаційна система	6	2	1	3,03	2,46	0,5
4	Інформаційна система підтримки прийняття рішень	2	0,75	0,33	1	0,84	0,17
за критерієм “Достовірність”							
1	Інформаційно-довідкова система	1	2	2	2	1,68	0,39
2	Інформаційно-пошукова система	0,5	1	0,5	1	0,71	0,17
3	Інформаційно-рекомендаційна система	0,5	2	1	2	1,19	0,28
4	Інформаційна система підтримки прийняття рішень	0,5	1	0,5	1	0,71	0,17
за критерієм “Легкість у використанні”							
1	Інформаційно-довідкова система	1	1	0,33	3,03	1	0,21
2	Інформаційно-пошукова система	1	1	0,5	2	1	0,21
3	Інформаційно-рекомендаційна система	3	2	1	4	2,21	0,47
4	Інформаційна система підтримки прийняття рішень	0,33	0,5	0,25	1	0,45	0,1
за критерієм “Надійність”							
1	Інформаційно-довідкова система	1	1,11	1,25	4	1,53	0,34
2	Інформаційно-пошукова система	0,9	1	1,33	3,33	1,41	0,31
3	Інформаційно-рекомендаційна система	0,8	0,75	1	4	1,24	0,27
4	Інформаційна система підтримки прийняття рішень	0,25	0,3	0,25	1	0,37	0,08
за критерієм “Продуктивність”							
1	Інформаційно-довідкова система	1	0,5	0,25	0,25	0,42	0,09
2	Інформаційно-пошукова система	2	1	0,5	0,33	0,76	0,16
3	Інформаційно-рекомендаційна система	4	2	1	0,5	1,41	0,29
4	Інформаційна система підтримки прийняття рішень	4	3	2	1	2,21	0,46

Матриця порівняння альтернатив

№	Критерії	Інформаційно-довідкова система	Інформаційно-пошукова система	Інформаційно-рекомендаційна	Інформаційна система ППР	
1	Актуальність	0,08	0,24	0,5	0,17	0,12
2	Достовірність	0,39	0,17	0,28	0,17	0,1
3	Продуктивність	0,09	0,16	0,29	0,46	0,36
4	Функціональність	0,11	0,11	0,33	0,44	0,15
5	Надійність	0,34	0,31	0,27	0,08	0,1
6	Легкість у використанні	0,21	0,21	0,47	0,1	0,17
	Узагальнені пріоритети	0,21	0,2	0,36	0,24	

Зважаючи на результати методу аналізу ієрархій, найбільшим показником є 0,36, що відповідає самій інформаційній системі рекомендацій, а отже, інтелектуальну інформаційну систему побудуємо на основі вибраного типу системи.

Основна мета створення інтелектуальної інформаційної системи підбору спорядження для силових структур полягає у наданні можливості ефективно підбирати спорядження для користувачів, завдяки забезпеченню здатності аналізувати користувачів, їхню поведінку, збирати явні й неявні дані, на цій підставі визначати групи користувачів та їхні особисті потреби, і відповідно до них надавати рекомендовані товари. Отримані рекомендації забезпечать користувача якісним та ефективним спорядженням відповідно до роду його діяльності, йому також рекомендуватимуть супутні товари. Призначення системи – аналіз дій користувача, опрацювання одержаних даних та надання на їх основі вичерпних рекомендацій із застосуванням машинного навчання щодо складання прогнозу доцільності товару та подальшої фільтрації і сортування. Такий підхід надає можливість ефективно та вичерпно рекомендувати товари, вилучати недоцільні та надалі пропонувати супутні, завдяки чому користувач матиме змогу підібрати товари, навіть не знаючи предметної області. Інтелектуальну систему підбору спорядження для силових структур спроектовано так, щоб її було зручно розгорнути та використовувати з будь-якого пристрою. Система моніторингу має змогу сканувати пристрої, щоб збирати неявні дані про користувача. Для повноцінної роботи системи необхідне підключення до мережі Internet.

Для розроблення та подальшого впровадження системи відповідно до результатів аналізу було вирішено використовувати нові підходи, основані на машинному навчанні та штучному інтелекті. Окрім цього, для калібрування результатів буде застосовано гібридний підхід, що дасть змогу продемонструвати високий ступінь гнучкості та можливості розширення. Під час реалізації системи буде створено інтуїтивний та зручний інтерфейс, орієнтований на цільового користувача. Цей інтерфейс буде розроблений із урахуванням роботи системи на багатьох браузерях та платформах. Перед впровадженням системи її буде розгорнуто на тестових пристроях для її тестування та усунення проблем, якщо в цьому виникне необхідність.

Після детального аналізу інтелектуальної інформаційної системи визначено певний перелік ефектів від використання системи, які варто проаналізувати після її реалізації та впровадження. До них можна зарахувати такі ефекти:

- економічний – дає змогу проаналізувати необхідність та доцільність залучення інвестицій у цей проєкт;
- фінансовий – відображає кінцевий прибуток реалізації проєкту;
- часовий – демонструє часові межі розроблення та впровадження системи, необхідні для створення системи;

- технічний – дає змогу проаналізувати можливість підвищення продуктивності;
- науково-технічний – дає можливість проаналізувати ступінь новизни впровадженої системи;
- соціальний – демонструє вплив на життя інших людей.

З метою успішного розроблення та впровадження інформаційної системи для вирішення поставленого завдання доцільно ретельно сформулювати завдання та їх формулювання. Також варто виконати аналіз та передбачити можливі труднощі в реалізації. Відповідно, щоб ефективно спроектувати систему із урахуванням всіх вимог та мінімальних затрат, було побудовано набір відповідних діаграм. На рис. 5 зображено подання діаграми класів, що складається із семи основних класів:

- клас “LogForm” описує модуль, який уможливує реєстрацію чи авторизацію користувача;
- клас “PersonalUserDataForm” описує модуль, який надає можливість редагувати дані користувача, а також здійснювати маніпуляції щодо видалення чи додавання нових макетів користувача;
- клас “TrackingManager” описує модуль, який сканує та аналізує дії користувача, стан системи, та збирає дані для подальшого оброблення, структуризації та їх передавання;
- клас “DataBase” описує базу даних, яка містить всі дані про користувачів, їх активності, а також про перелік товарів та інформацію щодо взаємодії з ними;
- клас “UserList” описує локальну базу даних, що містить останній аналіз користувача, які в майбутньому будуть передані в основну базу даних;
- клас “AI” описує основний логічний модуль системи, що виконує розрахунки для прогнозування доцільності рекомендацій товарів, їх фільтрацію, сортування, а також відповідає за кластеризацію користувачів.

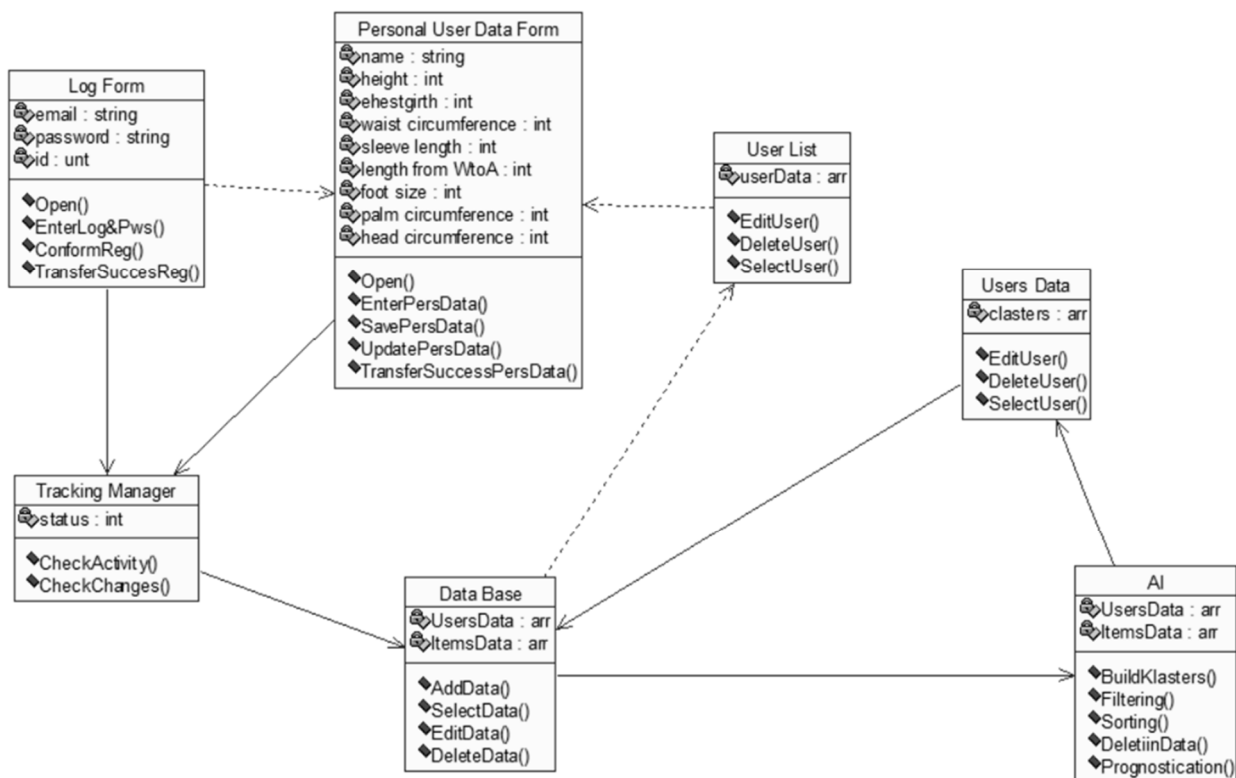


Рис. 5. Діаграма класів інтелектуальної інформаційної системи

На рис. 6 наведено діаграму варіантів використання для інтелектуальної інформаційної системи підбору товарів. На цій діаграмі три актори, тобто основні сутності:

- Користувач, який бажає підібрати один або декілька товарів та потребує певної рекомендації.
- Система, що авторизує користувача, збирає наявні дані, визначає кластер користувача, перевіряє отримані дані, здійснює сканування, аналізує та відображає результати у вигляді рекомендацій.

- База даних, що містить в собі всі наявні дані та має можливість здійснювати маніпуляції із даними, такі як: видалення, знаходження та редагування даних.

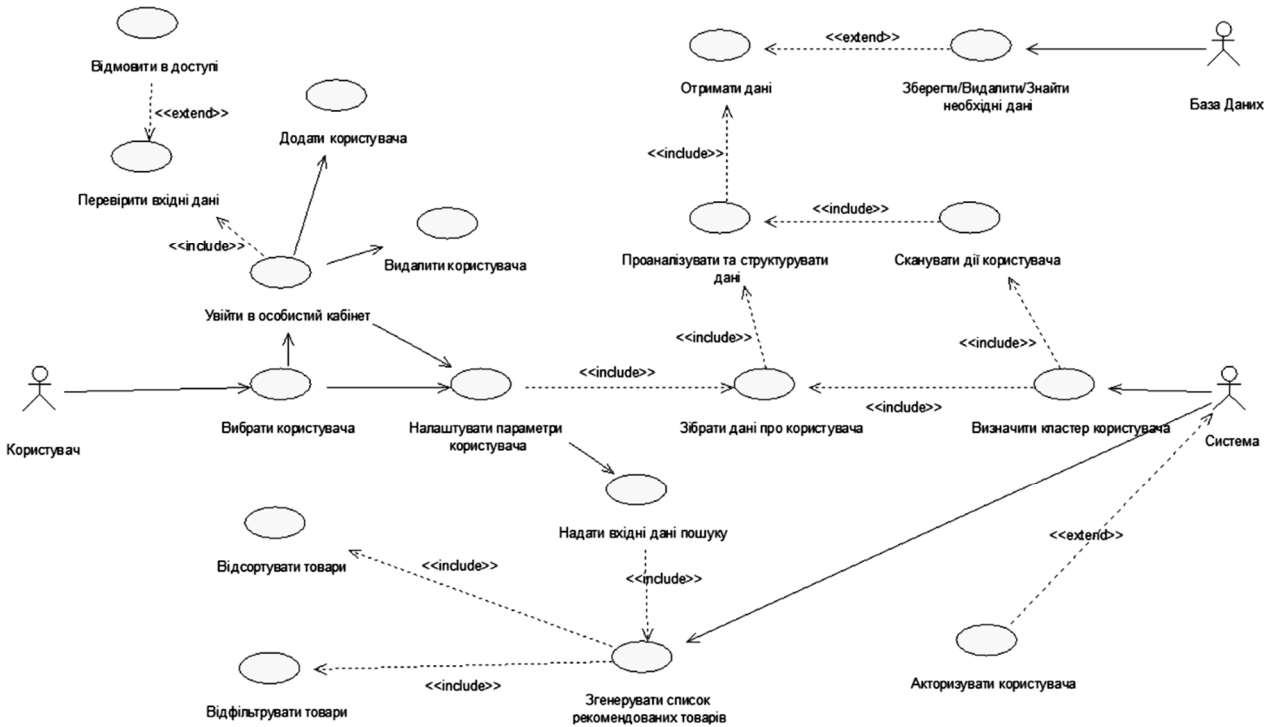


Рис. 6. Діаграма варіантів використання

Користувач системи налічує такі варіанти використання, як: “Вибрати користувача”, “Увійти в особистий кабінет”, “Додати користувача”, “Видалити користувача”, “Налаштувати параметри користувача”, “Надати вхідні дані пошуку”. Відповідно варіант використання “Увійти в особистий кабінет” пов’язаний відношенням <<include>> із варіантом використання “Перевірити вхідні дані”, який, своєю чергою, пов’язаний із варіантом використання “Відмовити в доступі” відношенням <<extend>>. Варіант використання “Налаштувати параметри користувача” пов’язаний відношенням <<include>> із варіантом використання “Зібрати дані про користувача”. Система ж налічує такі варіанти використання, як: “Згенерувати список рекомендованих товарів”, “Визначити кластер користувача”, а також “Авторизувати користувача”, який з’єднаний відношенням <<extend>>. Варіант використання “Згенерувати список рекомендованих товарів” пов’язаний із варіантами використання “Відсортувати товари” та “Відфільтрувати товари” за допомогою відношень <<include>>. Варіант використання “Визначити кластер користувача” пов’язаний із варіантами використання “Зібрати дані про користувача” та “Сканувати дії користувача” за допомогою відношень <<include>>. Своєю чергою ці варіанти використання пов’язані відношенням <<include>> із варіантом використання “Проаналізувати та структурувати дані”. Варіант використання “Проаналізувати та структурувати дані” пов’язаний відношенням <<include>> із варіантом використання “Отримати дані”, який, своєю чергою, пов’язаний із варіантом використання “Зберегти/Видалити/Знайти необхідні дані” відношенням <<extend>>. Сутність “База даних” має тільки один варіант використання, а саме “Зберегти/Видалити/Знайти необхідні дані”.

На рис. 7 відображено діаграму послідовності для розробленої інтелектуальної інформаційної системи. Під час запуску системи користувача основний сервіс завантажує графічний інтерфейс та форму для авторизації. Як тільки користувач введе вхідні дані для авторизації, їх буде відправлено на сервер та перевірено в базі даних, після чого буде дозволено доступ або відмовлено. Наступний етап – користувачу надається вибір підкористувачів, редагування даних, а також створення нового підкористувача.

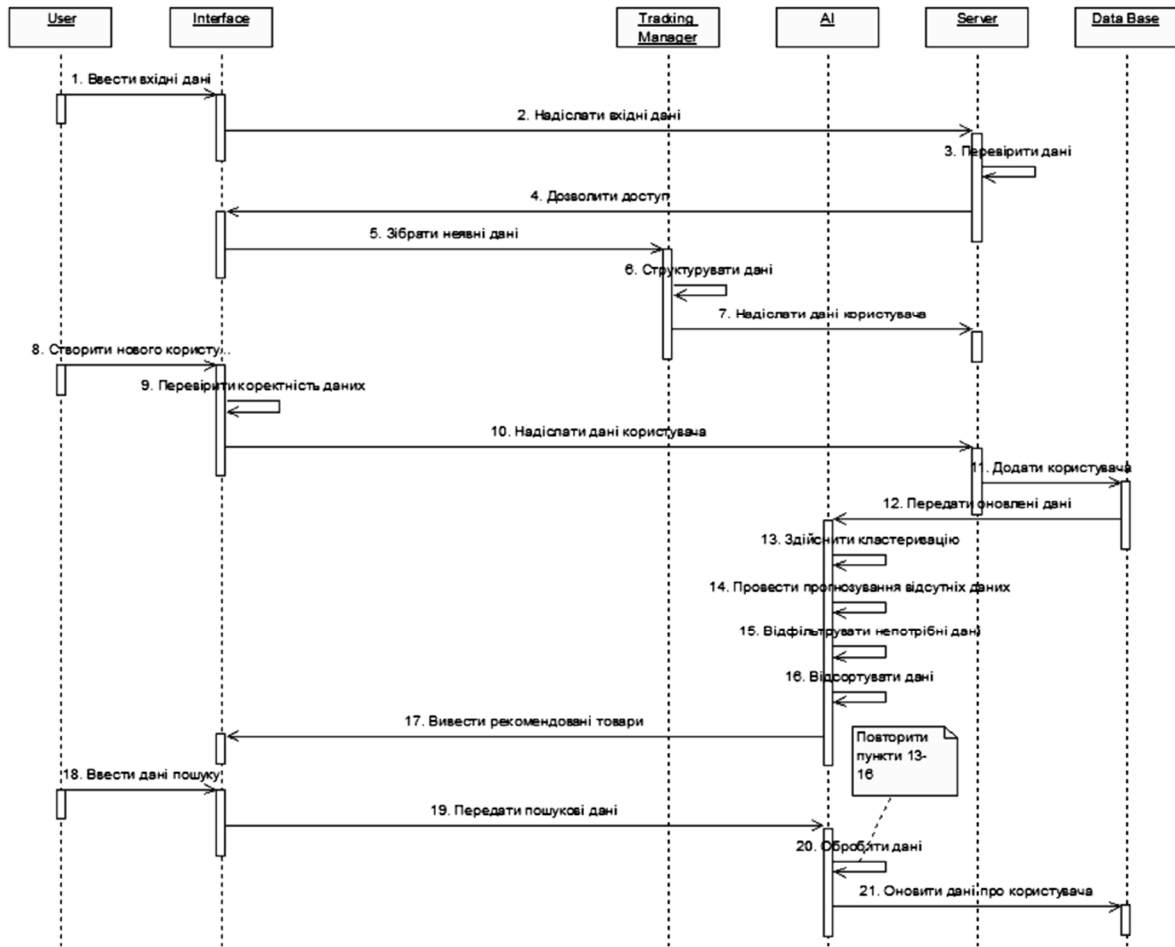


Рис. 7. Діаграма послідовності

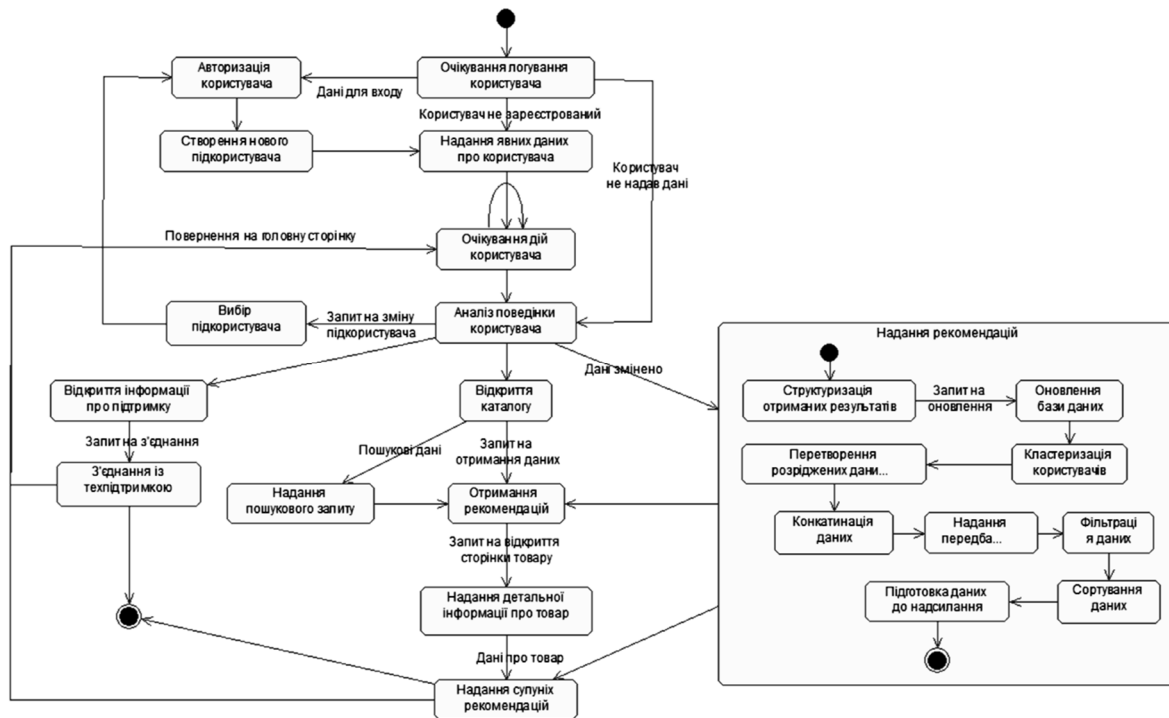


Рис. 8. Діаграма переходів станів

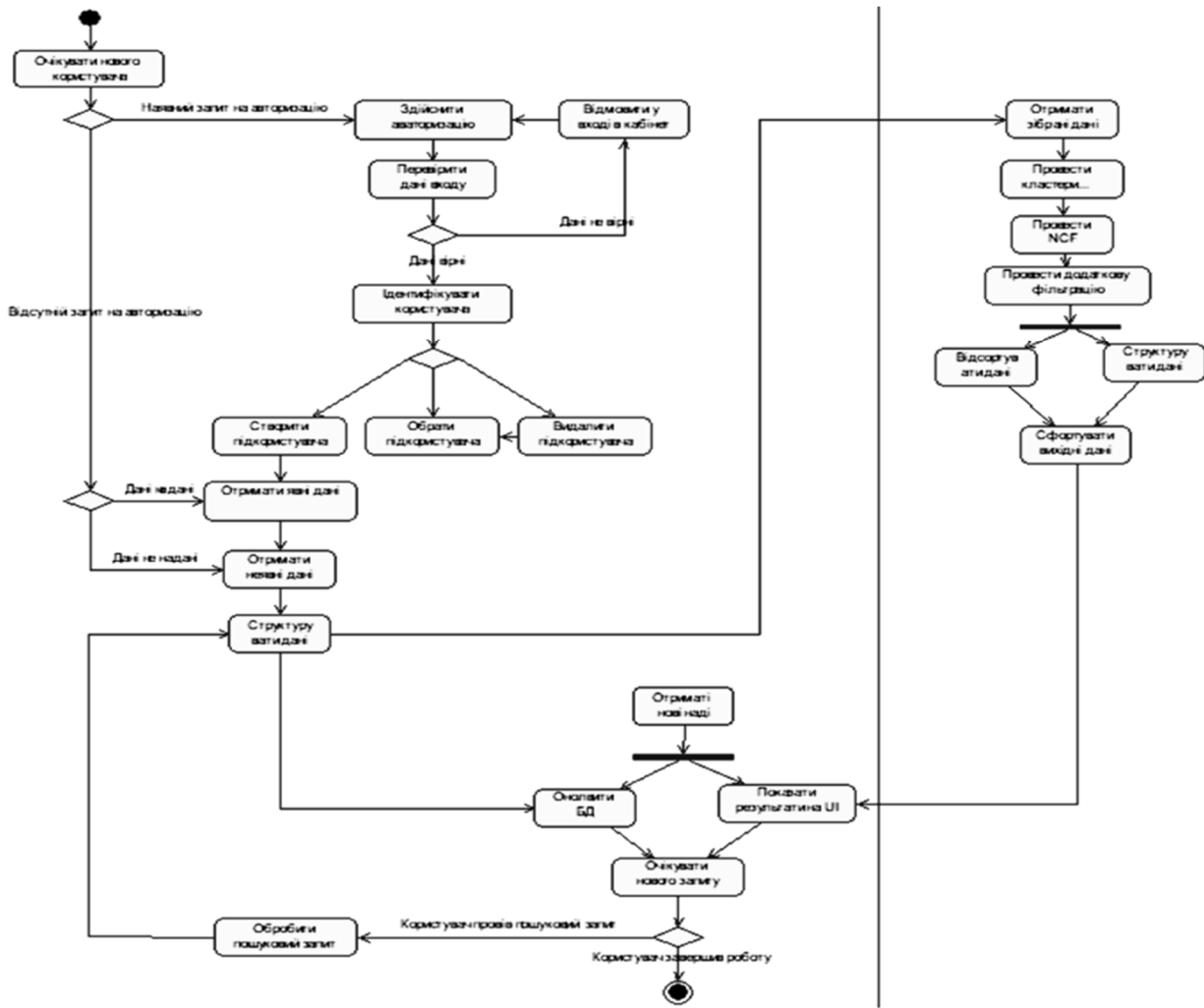


Рис. 9. Діаграма діяльності

Під час створення користувач вводить всі необхідні дані, які перевіряють на стороні клієнта, після чого їх відправляють на сервер, де відбувається їх подальше завантаження в базу даних. Після цього система проаналізує одержані явні та неявні дані, сформує кластери із користувачів з урахуванням останніх отриманих даних. Коли буде сформована вся необхідна інформація про користувача, система підбере рекомендації та виведе їх в інтерфейсі, в зручній для користувача формі. У разі додавання додаткових даних, таких як пошуковий запит, нова активність чи редагування явних даних про користувача, процес кластеризації та формування прогнозу рекомендацій буде повторено. На рис. 8 відповідно до завдань системного аналізу зображено реалізовану діаграму переходів станів для нашої розробленої інтелектуальної інформаційної системи підбору спорядження для силових структур. Для кращого розуміння роботи системи надання рекомендацій деталізовано за допомогою описання станів, що відбуваються у разі його застосування в системі. На рис. 9 наведена діаграма діяльності, на якій зображено основні стани системи. Сама діаграма відображає взаємозв'язок між різними компонентами системи та процеси, що виконуються у різних потоках. Початковим станом є дія "Очікування логування користувача". Окрім цього, діаграма містить декілька альтернативних переходів, у яких описано всі можливі випадки. Кінцевими ж діями системи є відображення даних на графічному інтерфейсі та у зручній для користувача формі. На рис. 10 наведено діаграму компонентів для інформаційної системи підбору спорядження для силових структур. На ній зображена основна програма та користувацький інтерфейс до неї, база даних, а також основні компоненти та модулі системи.

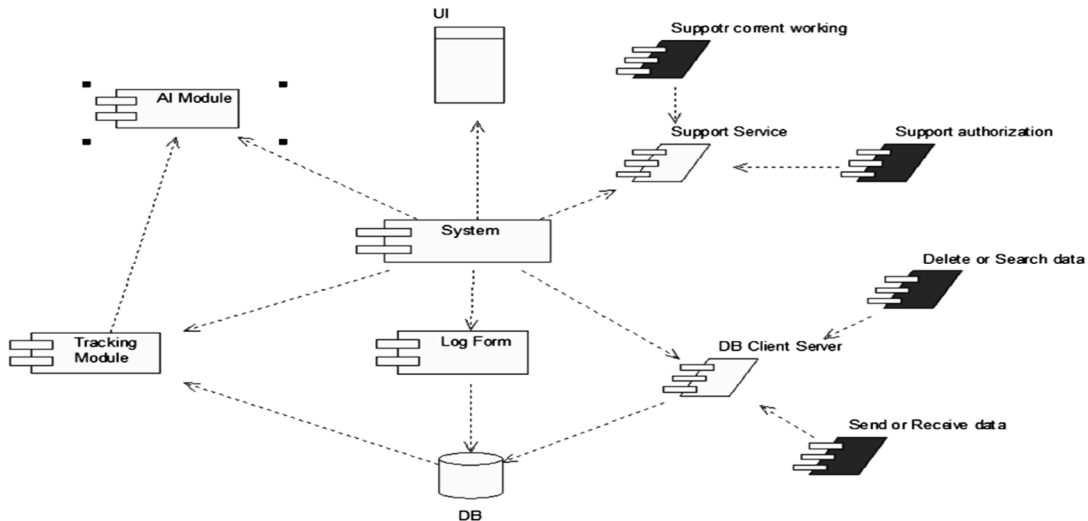


Рис. 10. Діаграма компонентів

У контексті цієї проблеми реалізація інтелектуальної інформаційної системи з урахуванням алгоритмів, що зможуть подолати проблему “холодного старту”, є доволі важливим процесом.

Загалом базове вирішення проблеми “холодного старту” для користувачів здійснюється на основі мінімальних даних про користувача, які так чи інакше можна отримати. Найпростіші дані – демографічні. В ідеалі, щоб отримати такі дані, користувачу пропонують пройти реєстрацію, під час якої він має змогу вказати певну інформацію про себе, однак, окрім явного отримання даних про користувача, можливе одержання і за допомогою API соціальних мереж, де більшість користувачів вказують свій вік, соціальний статус, рівень освіти тощо. Іншими даними для певного класу інтелектуальних інформаційних систем є фізичні дані користувача, як-от зріст, вага тощо.

Існують два основні підходи для визначення рекомендацій тих чи інших категорій користувачів на підставі отриманих даних:

- *експертно* – за допомогою визначення певних загальноприйнятих закономірностей із залученням певного експерта в конкретній сфері. Експерт сам визначає, що на початку роботи системи, коли вона перебуває на “холодному старті”, показувати кожній із категорій користувачів. Очевидним недоліком такого підходу є необхідність залучення експерта та його певна суб’єктивність.
- *автоматично* – категорії визначають, застосовуючи алгоритми, які можуть виявляти кластери користувачів зі схожими інтересами, після чого рекомендації будують на основі рейтингів, які визначили користувачі зі спільної категорії.

Оскільки проблему “холодного старту” ми вирішуємо для інтелектуальної інформаційної системи, другий підхід найдоцільніший для подальшого аналізу.

Для формування категорій користувачів найдоцільніше використовувати методи кластеризації, де користувачі будуть об’єктами кластеризації, а ознаками – необхідні дані про них. Серед цих методів виберемо метод k-середніх, тому що в цьому випадку кожен кластер визначається точкою свого центра, тому добре інтерпретується [7].

Оскільки, надаючи рекомендації, система повинна враховувати не тільки першочергові дані про користувача, але і його уже наявні оцінки чи оцінки інших користувачів, які до того ж можуть бути відсутніми, скористаємось алгоритмом Neural collaborative filtering (NCF), який спирається на колаборативні дані неявного фідбека (не лише відгук, це комплекс заходів зворотного зв’язку, тобто вплив результату функціонування будь-якої системи на її подальше функціонування) [8, 9]. Вхідні дані можуть бути різними, але спочатку використовують тільки колаборативні дані, закодовані за принципом one-hot encoding. Важливою перевагою цього підходу є те, що у випадку “холодного старту” на вхід можна подавати метадані користувачів і контенту. За вхідним шаром міститься повнозв’язний, який відповідає за перетворення розрідженого one-hot та представлення в щільні вкладення (embedding). Після отримання вкладення конкатинуються, а подальша архітектура власне

носить ім'я NCF і відповідає за перетворення вкладень пар користувач-об'єкт безпосередньо на передбачення. У цьому випадку стандартна середньоквадратична помилка може не відповідати виду реальних даних неявного фідбека, що представлені значеннями 0 і 1, тому використовується ймовірнісний підхід до навчання NCF. Тобто, якщо передбачене число розглядати як ймовірність значення 1 (релевантності об'єкта користувачеві) й обмежити вихід мережі діапазоном [0, 1], то відповідно можемо використовувати функцію втрат:

$$L = - \sum_{(u,i) \in \mathcal{Y}} \log \hat{y}_{ui} - \sum_{(u,j) \in \mathcal{Y}^-} \log(1 - \hat{y}_{uj})$$

$$= - \sum_{(u,i) \in \mathcal{Y} \cup \mathcal{Y}^-} y_{ui} \log \hat{y}_{ui} + (1 - y_{ui}) \log(1 - \hat{y}_{ui}),$$

де \mathcal{Y} – множина негативних прикладів, якими вибираємо певну підмножину взаємодій, які не спостерігаються (пар користувач-об'єкт, щодо яких немає даних). Негативні приклади відбирають рівномірно, але можна використовувати підходи, які враховують популярність, для підвищення продуктивності. Візуально це геометричний сенс сингулярного розкладання (рис. 11).

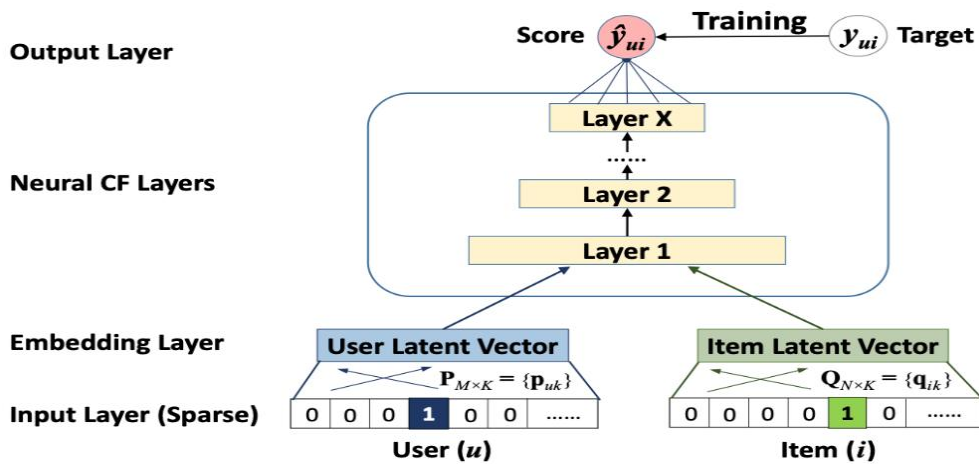


Рис. 11. Принцип роботи Neural collaborative filtering (NCF)

Отже, алгоритм працюватиме із урахуванням як подібності першочергових даних про користувача, так і явних і прогнозованих оцінок користувачів [10].

Здійснивши порівняння ефективності у разі застосування методу k -середніх із алгоритмом NCF та без нього, можемо спостерігати результати на рис. 12.

Ці метрики відображають рекомендації користувачам одягу із відношенням кількості переходів товару до кількості показів. Варто зауважити, що при цьому система була в умовах “холодного старту”, за яких всі об'єкти і користувачі вважались новими, а до старту алгоритмів не використовувалось жодної інформації щодо історії попередніх показів.

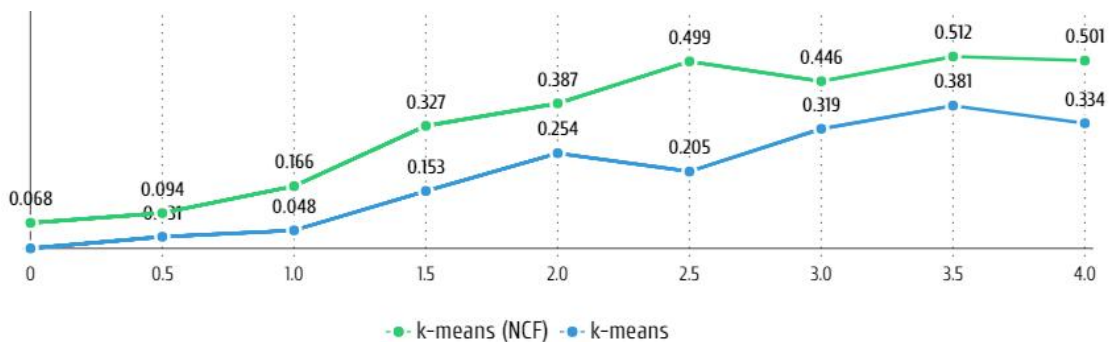


Рис. 12. Порівняння методу k -середніх із застосуванням NCF алгоритму та без нього

Звичайно, поліпшувати рекомендації в цьому випадку можна уже безліччю способів. Основний нюанс полягає в максимальній оптимізації алгоритму та усуненні прогалин і несумісності. Оскільки в нашому випадку за основу брали формування кластерів із користувачів, в цьому випадку доцільно поліпшити рекомендації із застосуванням групових рекомендацій, назва якого говорить сама за себе: ми підбираємо новому користувачеві такі рекомендації, які подобаються більшості користувачів його демографічної категорії. Існує низка різних стратегій, як агрегувати рейтинги різних користувачів у групову рекомендацію. Альтернативою до цього підходу є використання фільтр-ботів (filterbots), які генерують початкові рейтинги для нового користувача. Тобто під час реєстрації фільтр-боти автоматично генерують кілька базових рейтингів для користувача на основі його даних, зібраних явним або неявним методом, і які після цього використовують уже наявні алгоритми колаборативної фільтрації на “холодному старті”. Перевагою такого підходу є простота реалізації та відсутність необхідності змінювати наявні алгоритми. Варто зазначити, що фільтр-боти і групові рекомендації не конфліктують між собою, і відповідно цей фактор уможливило їх спільне використання. В такому випадку як початкові рейтинги фільтр-ботів беруть групові рейтинги [11].

Остаточний варіант концепції алгоритму із урахуванням зазначеного вище подано на рис. 13.

Принцип цих рекомендацій полягає в підборі новому користувачеві таких рекомендацій, які подобаються більшості користувачів його категорії (кластера). Варто зазначити, що цей підхід певною мірою нагадує принцип user-based, оснований на зіставленні користувачів, та бере до уваги подібність заданого користувача на інших користувачів, задіяних у системі. Відповідно до визначених вимог, архітектури, а також засобів реалізації побудовано інтелектуальну інформаційну систему підбору тактичного спорядження для силових структур. Ця система може працювати як на персональних комп’ютерах, так і на телефонах чи планшетах.

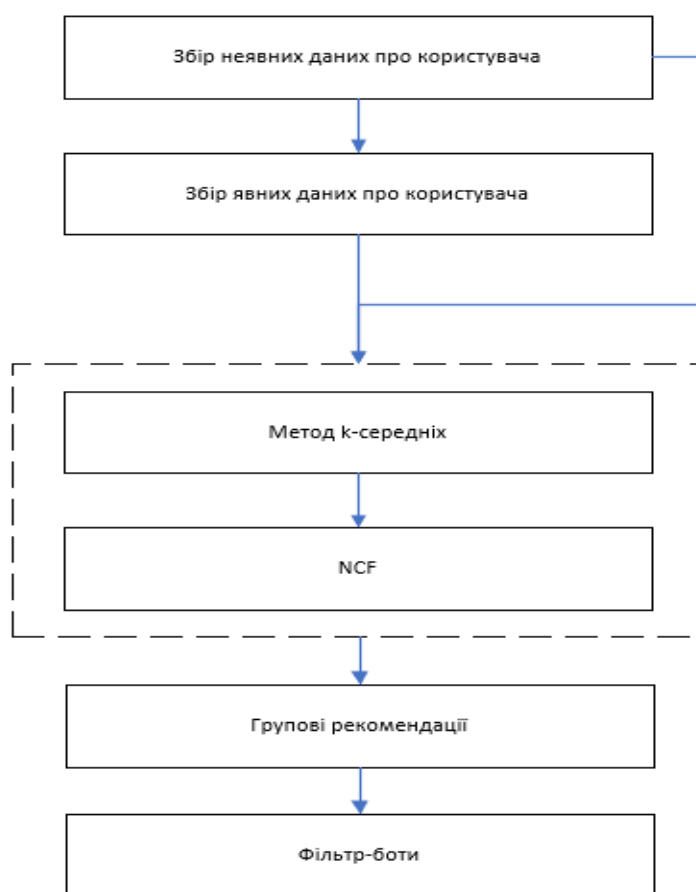


Рис. 13. Загальна схема алгоритму

Оскільки система являє собою вебсайт, користувачу не потрібно здійснювати додаткове розгортання системи, оскільки сам вебсайт міститься на хостингу разом із базою даних. База даних складається із кількох таблиць, пов'язаних відповідними відношенням. Варто зазначити, що для роботи системи необхідне інтернет-з'єднання, однак у разі тимчасового переривання цього з'єднання передбачено механізми локального збереження даних, з метою подальшого їх передавання після відновлення зв'язку із системою. Функціональне призначення системи полягає в підбиранні спорядження із застосуванням інтелектуальних засобів у поєднанні із класичними засобами рекомендацій. Окрім цього, користувачі матимуть змогу налаштовувати власні параметри, а тим користувачам, які з певних причин не бажатимуть проходити реєстрацію або ж заповнювати деталізовано пошукові фільтри, буде надано можливість пошуку спорядження у спрощеному алгоритмі, за допомогою вбудованої пошукової системи. Важлива взаємодія користувача із системою. У цьому випадку людина, користуючись системою, не бачитиме критичних візуальних відмінностей між типовими інтерфейсами і протоколами пошуку і підбору. Система побудована так, щоб мінімізувати кількість дій і зробити її максимально простою для будь-якого типу користувача. Загалом структуру проекту можна описати двома способами – логічно або файлово. Для кращого розуміння принципу функціонування розглянемо обидва.

Логічну структуру кожної інформаційної системи можна подати відповідно до чотирьох основних моделей: лінійної моделі, моделі “ґрати”, деревоподібної моделі та “павутини”. Варто зазначити, що існують додаткові комбінації на базі основних моделей, які надають змогу реалізувати будь-яку логічну структуру сайту. Оскільки, як зазначено вище, інтелектуальна інформаційна система доволі складна, зважаючи на її комбінацію алгоритмів, а розроблення такого ресурсу – клопітка робота, наявність оптимальної структури зведе до мінімуму можливість технічних помилок. Це стосується як помилок у коді, так і наявності або відсутності окремих сторінок і матеріалів на них. З огляду на ці фактори, ми вибрали логічну деревоподібну структуру. Логічна деревоподібна структура – найпопулярніша модель організації системи. Така структура дає змогу користувачам системи за бажанням керувати глибиною відвідування системи. Користувачі мають змогу входити тільки на сторінки найвищих рівнів або ж “спускатись” до нижчих. Наявні можливості вибору залежать від “ширини дерева”. Якщо користувачам для досягнення поставленої кінцевої мети необхідна занадто велика кількість клацань миші, то структура ієрархії системи може виявитися у цьому випадку занадто вузькою. Користувачів дратуватиме нескінченне “клацання”, оскільки їхні дії не даватимуть очікуваних результатів. Водночас дуже широке “дерево”, яке ґрунтується на доволі великій кількості варіантів вибору, змусить користувачів системи витратити чимало часу на вивчення різних варіантів. Це також не дасть позитивних результатів.

Зважаючи на ці фактори логічної структури системи, ми прагнули забезпечити оптимальний рівень глибини і ширини. Графічне зображення логічної деревоподібної структури системи подано на рис. 14.

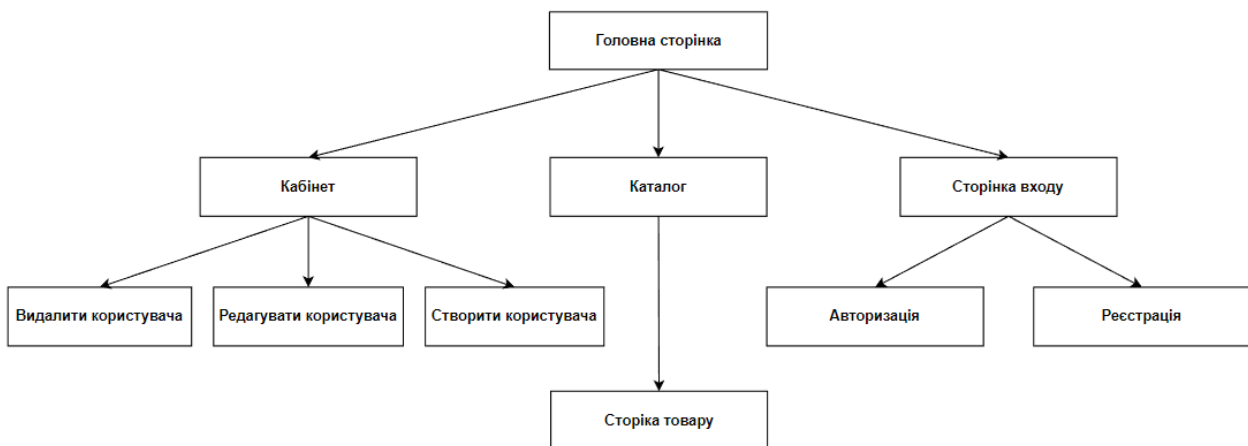


Рис. 14. Логічна структура проекту

Цей принцип дасть змогу зробити систему максимально швидкодіючою, наскільки це можливо, з мінімальною кількістю протоколів. Розроблена інтелектуальна інформаційна система являє собою набір компонентів та засобів, які формують чітку структуру та ієрархію. Більшість засобів відповідно до визначеної архітектури функціонують через середовище браузера, а наявні компоненти системи складаються із власне самої системи та бази даних.

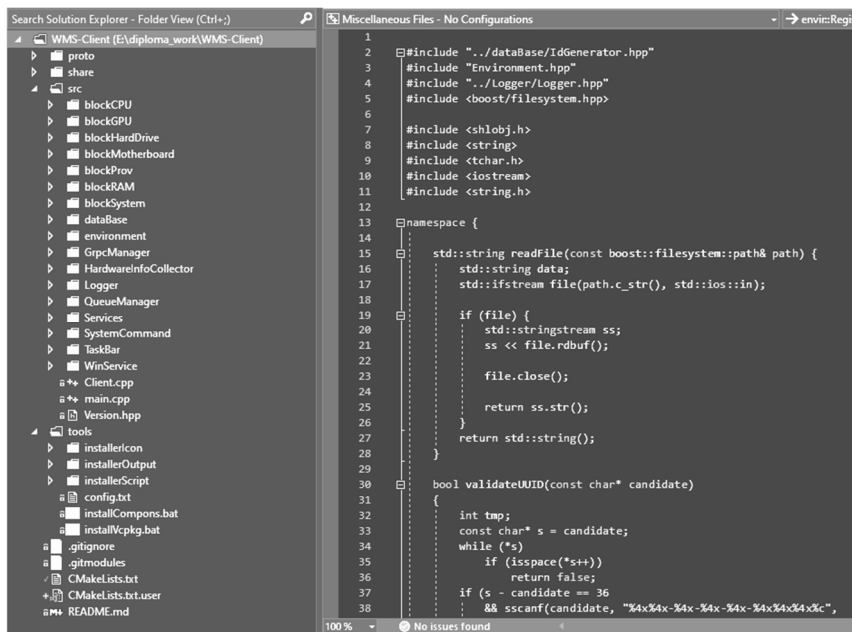


Рис. 15. Файлова структура проекту

Структура проекту складається із кількох підкаталогів та компонентів, які відповідають за відображення інтерфейсу користувача, розрахунки та весь функціонал. Основні підкаталоги:

- Images. Цей каталог містить зображення для ілюстрацій та побудови інтерфейсу, окрім цього, набір спрайтів іконок, необхідних для роботи;
- Src. Каталог містить файли із вихідним кодом програми;
- Node_modules. Каталог, в якому містяться всі модулі, необхідні для взаємодії елементів системи та її розгортання;
- Components. У цей каталог входять підкаталоги, що містять усі компоненти системи, а також UI бібліотека із стилістикою системи;
- Scss. У цьому каталозі містяться всі змінні, що відповідають за візуальну стилістику системи, міксіни, базове скидання стилів, а також використовувані шрифти.

З метою зберігання інформації про товари, користувача та його активність, а також роботу системи спроектовано базу даних. Принцип застосування бази даних полягає у її розміщенні на сервері. За необхідності за допомогою запитів система на стороні клієнта матиме можливість зв’язуватись з нею, її не потрібно розгортати на кожному пристрої користувача, що було б абсолютно недоцільним з погляду продуктивності та безпеки. У самій базі даних є кілька таблиць, які пов’язані відповідними зв’язками. Колонки в базі даних містять відповідну назву поля і його тип, відповідно вмістом поля може бути як конкретне значення, так і дані з іншої таблиці. Опишемо детальніше БД. Одна з основних таблиць в базі даних – таблиця “User”, яка містить всю наявну інформацію про певного користувача. За допомогою даних таблиці система може здійснювати ідентифікацію користувача та моніторити його активність у системі. Таблиця містить такі атрибути:

- id(uniqueidentifier) – ідентифікатор користувача;
- firstName(varchar) – ім’я користувача;

- lastName(varchar) – прізвище користувача;
- password(varchar) – пароль користувача;
- email(varchar) – електронну пошту для надсилання листів;
- createdAt(datetime) – дату реєстрації.

Наступна таблиця – “SessionsUser” – містить інформацію про сесію, яку веде поточний користувач. Ця таблиця допомагає в аналізі користувача, його кластеризації, а також у розумінні потреб користувачів та визначенні способів модернізації системи. Таблиця містить такі атрибути:

- id(uniqueidentifier) – ідентифікатор сесії;
- userId(uniqueidentifier) – ідентифікатор користувача;
- status(int) – статус роботи;
- generalClick(int) – кількість кліків під час роботи;
- generalBought(int) – кількість придбаних товарів;
- location(geography) – локація користувача;
- createdAt(datetime) – дата початку сесії;
- timeSpended(int) – час роботи сесії.

Кожен користувач матиме змогу сформувати декілька шаблонів підкористувачів (легенд), створена таблиця “Legends” міститиме інформацію про налаштування підкористувача та його основні відомості. Таблиця містить такі атрибути:

- id(uniqueidentifier) – ідентифікатор підкористувача;
- userId(uniqueidentifier) – ідентифікатор користувача;
- name(varchar) – ім'я підкористувача;
- type(varchar) – тип підкористувача;
- clusterType(int) – тип кластера підкористувача;
- height(float) – висота;
- chestGirth(float) – окружність грудей;
- waistCircumference(float) – окружність талії;
- sleeveLength(float) – довжина рукава;
- lengthFromWtoH(float) – довжина від талії до п'яти;
- footSize(float) – розмір ноги;
- palmCircumference(float) – окружність долоні;
- headCircumference(float) – окружність голови.

Оскільки система не просто підбирає товари, але і надає змогу їх придбати, відповідно необхідна таблиця “Orders” для їх оформлення та зберігання, якщо потрібно переглянути історію замовлень. Таблиця містить такі атрибути:

- id(uniqueidentifier) – ідентифікатор замовлення;
- legendId(uniqueidentifier) – ідентифікатор підкористувача;
- goodId(uniqueidentifier) – ідентифікатор товару;
- data(datetime) – дата купівлі;
- totalPrice(float) – загальна ціна;
- receiptPlace(varchar) – місце доставки;
- generalBoughtPerOrder(int) – кількість товарів.

Для моніторингу дій доцільно здійснювати аналіз не загального користувача акаунту, а підкористувачів, оскільки у кожного із них різні вимоги та параметри. Для цього передбачено таблицю “SessionsLegend”, яка містить такі атрибути:

- id(uniqueidentifier) – ідентифікатор сесії підкористувача;
- legendId(uniqueidentifier) – ідентифікатор підкористувача;
- generalClick(int) – кількість кліків під час роботи;

- generalBoughtPerSession(int) – кількість придбаних товарів;
- createdAt(datetime) – дата початку сесії;
- timeSpended(time) – час роботи сесії.

Із урахуванням сесій підкористувача здійснюються постійний моніторинг, збирання, аналіз та підсумовування інформації, а тому необхідна таблиця “Result”, яка міститиме відповідні дані, до яких будуть повторно за необхідності звертатись модулі. Таблиця містить такі атрибути:

- id(uniqueidentifier) – ідентифікатор поточного збирання даних;
- sessionLegendId(uniqueidentifier) – ідентифікатор сесії підкористувача;
- klusterType (float) – тип кластера підкористувача;
- priorityEquipment(float) – тип спорядження підкористувача;
- priorityVector(float) – пріоритет для підкористувача.

Останньою є таблиця “Catalog”, яка необхідна для збереження інформації про товари і містить такі атрибути:

- id(uniqueidentifier) – ідентифікатор товару;
- name(varchar) – ім’я товару;
- description(varchar) – опис товару;
- code(varchar) – код товару;
- type(varchar) – тип товару;
- quantity(int) – кількість товару;
- rating(float) – рейтинг товару;
- generalClick(int) – кількість кліків;
- generalSold(int) – кількість проданих товарів.

Графічне зображення ER-діаграми бази даних подано на рис. 16.

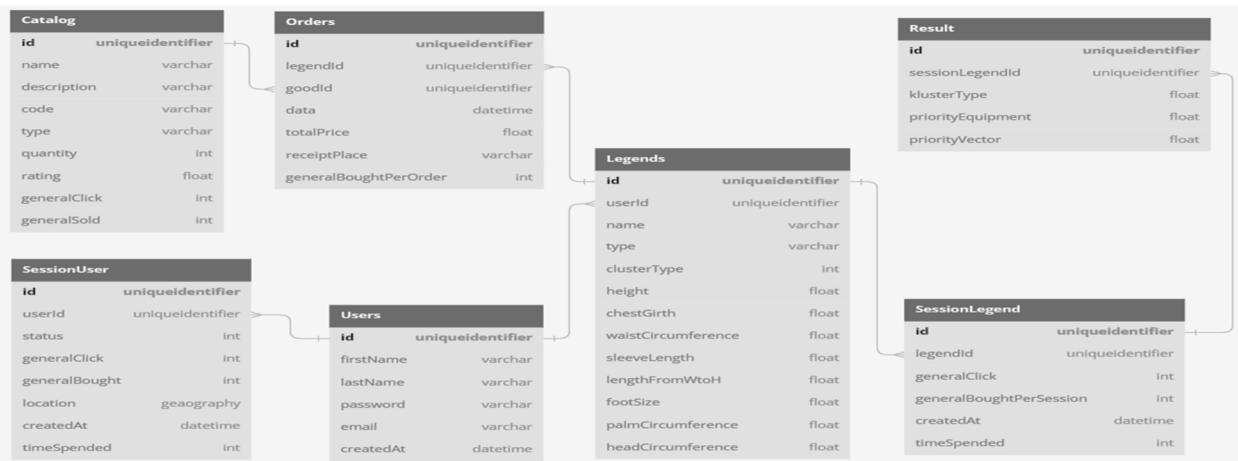


Рис. 16. Схема бази даних системи

Завершивши розроблення, ми порівняли ефективність роботи створеної інтелектуальної інформаційної системи із аналогічними базовими системами, що ґрунтуються на стандартних алгоритмах роботи. Результати наведено на рис. 17. Графік отриманих результатів оснований на відношенні кількості переходів до кількості показів рекомендованого товару.

Варто зазначити, що важливим фактором впливу на роботу та подальші результати системи були також реалізований інтерфейс та концепція початку роботи із системою для усунення поширених проблем. Разом з цим, порівняння ефективності не здійснювалось в умовах “холодного старту”, а реалізовано з моделюванням довготривалої роботи системи.

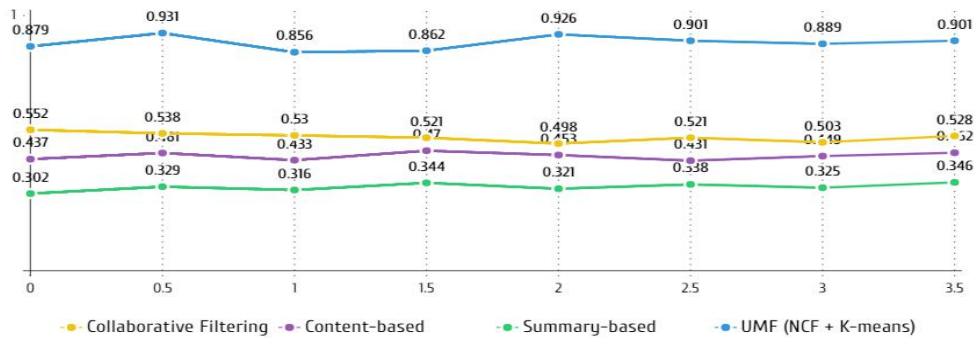


Рис. 17. Порівняння ефективності роботи системи

Висновок

В ході виконання роботи визначено основну мету реалізації інтелектуальної інформаційної системи підбору спорядження для силових структур, її новизну, актуальність та практичну цінність. У результаті реалізації створено типову інтелектуальну інформаційну систему, що являє собою SPA застосунок, із БД та AI частиною, яка відповідає за інтелектуальну складову.

Особливу увагу звернено на аналіз безпосередньо інтелектуальних інформаційних систем, їх архітектури та ускладнень, що можуть виникати під час їхньої реалізації. Під час проектування системи необхідно враховувати безліч чинників. На основі аналізу актуальних статей, наукових робіт та інших літературних джерел в міжнародному і національному сегментах ми визначили вимоги, типові проблеми та способи ефективної роботи системи, які надалі враховували під час проектування інтелектуальної інформаційної системи. Окрім цього, з метою визначення більшої кількості вимог, проаналізовано схожі за призначенням та функціональністю системи та виділено їхні переваги і недоліки, які допоможуть краще спроектувати власну інтелектуальну інформаційну систему. Виконано системний аналіз інтелектуальної інформаційної системи підбору товарів для силових структур. Цей аналіз охоплював першочергово побудову дерева цілей системи, визначення основних підцілей та критерії до системи, основного призначення системи, її середовища та місця використання, а також очікуваних ефектів від її реалізації та впровадження. З метою ефективного визначення архітектури системи і подальших деталей щодо її реалізації, реалізовано метод МАІ для заздалегідь визначених концепцій систем, які доцільно проаналізувати, із урахуванням попередніх вимог. На основі отриманих результатів здійснено побудову необхідних UML діаграм, які уможливають детальний аналіз структури системи, її поведінки під час роботи та взаємодії між елементами системи. Подано детальний опис кожної зі створених діаграм і відповідні коментарі. Здійснено детальний аналіз та обґрунтування методів і засобів, необхідних для реалізації інтелектуальної інформаційної системи, що могли б дати змогу виконати поставлені завдання та відповідати визначеним вимогам. Вибрані методи і засоби дають можливість реалізувати систему максимально ефективно, це виявляється у її швидкодії, потенціалі щодо масштабованості та відносній простоті роботи з нею. Окрім цього, за допомогою визначених засобів, система має змогу реалізовувати зручні інтерфейси користувача, легко взаємодіяти як із клієнтом, так із сервером, повністю використовувати всі наявні методи щодо реалізації AI модулів чи баз даних. Варто зауважити, що вибір всіх методів та засобів реалізації ґрунтувався на попередньому аналізованні актуальних наукових досліджень і проблем, а також на технічних висновках щодо доцільності взаємодії кожного із засобів чи методів у системі.

Наведено детальний опис створеної інтелектуальної інформаційної системи підбору спорядження для силових структур. Відповідно до попередніх вимог та архітектури, система являє собою SPA сайт, який містить БД для зберігання всієї необхідної інформації, а також AI частину для реалізації інтелектуальної складової системи. Окрім цього, здійснено опис як логічної, так і файлової

структури, структури бази, типів даних у робочих формах, а також вказано функціональне призначення системи. З метою надання детальної інструкції для користувача наведено описання типового алгоритму дій для нового користувача, із всіма вимогами щодо роботи із системою, здійснено аналіз інтерфейсу та зручності його використання. Для детальної перевірки роботи здійснено тестування та налагодження системи, а також аналіз контрольного прикладу, в якому продемонстровано результати роботи інтелектуальної інформаційної системи підбору спорядження для силових структур. Реалізована інтелектуальна інформаційна системи, як і всі інші новостворені системи, потребує подальшого тестування із більшою кількістю користувачів, щоб виявити можливі помилки чи способи оптимізації роботи.

У результаті дослідження описано та змодельовано один із підходів до побудови інтелектуальної інформаційної системи на основі рекомендації товарів користувачам із вирішенням проблеми “холодного старту”. У дослідженні враховано переваги та недоліки методів під час їх комбінування, а також їх сумісність, що є важливим фактором для швидкодії системи та ефективності роботи алгоритму. За результатами аналізу виконано практичне порівняння ефективності роботи системи із базовим підходом до вирішення проблеми та гібридним. Подальші дослідження будуть спрямовані на удосконалення алгоритму та створення інтелектуальної інформаційної системи із реалізацією підходів, визначених у результаті досліджень.

Список літератури

1. Grosan C., Abraham A. (2011). *Intelligent systems*, Vol. 17, 261–268. DOI: 10.1007/978-3-642-21004-4.
2. Ferré S., Ridoux O. (2004). Introduction to logical information systems. *Information Processing & Management*, 40(3), 383–419. DOI: 10.1016/S0306-4573(03)00018-9.
3. Injadat M., Moubayed A., Nassif A.B. et al. (2021). Machine learning towards intelligent systems: applications, challenges, and opportunities. *Artif. Intell. Rev.*, 54, 3299–3348. DOI: 10.1007/s10462-020-09948-w.
4. Bobadilla J., Ortega F., Hernando A., Bernal J. (2012). A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-based systems*, 26, 225–238. DOI: 10.1016/j.knosys.2011.07.021.
5. Lika B., Kolomvatsos K., Hadjiefthymiades S. (2014). Facing the cold start problem in recommender systems. *Expert systems with applications*, 41(4), 2065–2073. DOI: 10.1016/j.eswa.2013.09.005.
6. Vrontis D., Christofi M., Pereira V., Tarba S., Makrides A., Trichina E. (2022). Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review. *The International Journal of Human Resource Management*, 33(6), 1237–1266. DOI: 10.1080/09585192.2020.1871398.
7. Yadav N., Mundotiya R.K., Singh A.K., Pal S. (2021). Diversity in Recommendation System: A Cluster Based Approach. *Advances in Intelligent Systems and Computing*, Vol. 1179. DOI: 10.1007/978-3-030-49336-3_12.
8. Hsieh C. J., Natarajan N., Dhillon I. (2015). PU learning for matrix completion. In International conference on machine learning, PMLR, 2445–2453. URL: <http://proceedings.mlr.press/v37/hsiehb15.html>.
9. John MacKrell. (2000). Supporting Collaborative Product Definition via Scaleable, Web-Based PDM. Prepared by CIMdata, Inc.
10. Sammut C., Webb G. I. (2011). *Encyclopedia of machine learning*. Springer Science & Business Media. DOI: 10.1007/978-0-387-30164-8.
11. Aggarwal C. C. (2016). *Recommender systems*, Vol. 1. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-29659-3.
12. Meleshko E. V., Semenov S. G., Khokh V. D. (2018). Research on methods of building recommender systems on the Internet. DOI: 10.26906/SUNZ.2018.1.131.
13. Raza S., Ding C. (2022). News recommender system: a review of recent progress, challenges, and opportunities. *Artif. Intell. Rev.*, 55, 749–800. DOI: 10.1007/s10462-021-10043-x.
14. Zuo Y., Zeng J., Gong M., Jiao L. (2016). Tag-aware recommender systems based on deep neural networks. *Neurocomputing*, 204, 51–60. DOI: 10.1016/j.neucom.2015.10.134.
15. Lytvyn V., Vysotska V., Shatskykh, V., Kohut, I., Petruchenko O., Dzyubyk L., Bobrivets V., Panasyuk V., Sachenko S., Komar, M. (2019). Design of a recommendation system based on Collaborative Filtering and machine learning considering personal needs of the user. *Eastern-European Journal of Enterprise Technologies*, 4(2), 6–28. DOI: 10.15587/1729-4061.2019.175507.

16. Balush I., Vysotska V., Albota S. (2021). Recommendation System Development Based on Intelligent Search, NLP and Machine Learning Methods. *CEUR Workshop Proceedings*, 584–617.
17. Demchuk A., Lytvyn V., Vysotska V., Dilai M. (2020). Methods and Means of Web Content Personalization for Commercial Information Products Distribution. *Lecture Notes in Computational Intelligence and Decision Making*. Vol. 1020. Springer, Cham. DOI: 10.1007/978-3-030-26474-1_24.
18. Tulashvili Y., Turbal Y., Abd Alkaleg D., Pasichnyk V., Kunanets N. (2020). The Optimal Tour Problem in Smart Tourism Recommender Systems. In *IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT)*, Vol. 2, 246–250. DOI: 10.1109/CSIT49958.2020.9322043

References

1. Grosan C., Abraham A. (2011). *Intelligent systems*, Vol. 17, 261–268. DOI: 10.1007/978-3-642-21004-4.
2. Ferré S., Ridoux O. (2004). Introduction to logical information systems. *Information Processing & Management*, 40(3), 383–419. DOI: 10.1016/S0306-4573(03)00018-9.
3. Injadat M., Moubayed A., Nassif A.B. et al. (2021). Machine learning towards intelligent systems: applications, challenges, and opportunities. *Artif. Intell. Rev.*, 54, 3299–3348. DOI: 10.1007/s10462-020-09948-w.
4. Bobadilla J., Ortega F., Hernando A., Bernal J. (2012). A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-based systems*, 26, 225–238. DOI: 10.1016/j.knosys.2011.07.021.
5. Lika B., Kolomvatsos K., Hadjiefthymiades S. (2014). Facing the cold start problem in recommender systems. *Expert systems with applications*, 41(4), 2065–2073. DOI: 10.1016/j.eswa.2013.09.005.
6. Vrontis D., Christofi M., Pereira V., Tarba S., Makrides A., Trichina E. (2022). Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review. *The International Journal of Human Resource Management*, 33(6), 1237–1266. DOI: 10.1080/09585192.2020.1871398.
7. Yadav N., Mundotiya R.K., Singh A.K., Pal S. (2021). Diversity in Recommendation System: A Cluster Based Approach. *Advances in Intelligent Systems and Computing*, Vol. 1179. DOI: 10.1007/978-3-030-49336-3_12.
8. Hsieh C. J., Natarajan N., Dhillon I. (2015). PU learning for matrix completion. In *International conference on machine learning*, PMLR, 2445–2453. URL: <http://proceedings.mlr.press/v37/hsiehb15.html>.
9. John MacKrell. (2000). Supporting Collaborative Product Definition via Scaleable, Web-Based PDM. Prepared by CIMdata, Inc.
10. Sammut C., Webb G. I. (2011). *Encyclopedia of machine learning*. Springer Science & Business Media. DOI: 10.1007/978-0-387-30164-8.
11. Aggarwal C. C. (2016). *Recommender systems*, Vol. 1. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-29659-3.
12. Meleshko E. V., Semenov S. G., Khokh V. D. (2018). Research on methods of building recommender systems on the Internet. DOI: 10.26906/SUNZ.2018.1.131.
13. Raza S., Ding C. (2022). News recommender system: a review of recent progress, challenges, and opportunities. *Artif. Intell. Rev.*, 55, 749–800. DOI: 10.1007/s10462-021-10043-x.
14. Zuo Y., Zeng J., Gong M., Jiao L. (2016). Tag-aware recommender systems based on deep neural networks. *Neurocomputing*, 204, 51–60. DOI: 10.1016/j.neucom.2015.10.134.
15. Lytvyn V., Vysotska V., Shatskykh V., Kohut I., Petruchenko O., Dzyubyk L., Bobrivets V., Panasyuk V., Sachenko S., Komar, M. (2019). Design of a recommendation system based on Collaborative Filtering and machine learning considering personal needs of the user. *Eastern-European Journal of Enterprise Technologies*, 4(2), 6–28. DOI: 10.15587/1729-4061.2019.175507.
16. Balush I., Vysotska V., Albota S. (2021). Recommendation System Development Based on Intelligent Search, NLP and Machine Learning Methods. *CEUR Workshop Proceedings*, 584–617.
17. Demchuk A., Lytvyn V., Vysotska V., Dilai M. (2020). Methods and Means of Web Content Personalization for Commercial Information Products Distribution. *Lecture Notes in Computational Intelligence and Decision Making*. Vol. 1020. Springer, Cham. DOI: 10.1007/978-3-030-26474-1_24.
18. Tulashvili Y., Turbal Y., Abd Alkaleg D., Pasichnyk V., Kunanets N. (2020). The Optimal Tour Problem in Smart Tourism Recommender Systems. In *IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT)*, Vol. 2, 246–250. DOI: 10.1109/CSIT49958.2020.9322043

**SOFTWARE FOR THE IMPLEMENTATION OF AN INTELLIGENT SYSTEM
TO SOLVE THE PROBLEM OF “COLD START”**

Mykola Baran¹, Victoria Vysotska^{1,2}, Roman Holoshchuk³

¹ Lviv Polytechnic National University,

Information Systems and Networks Department, Lviv, Ukraine

² Osnabrück University, Institute of Computer Science, Osnabrück, Germany

³ Lviv Polytechnic National University,

Social Communications and Information Activities Department, Lviv, Ukraine

¹ E-mail: mykola.baran.mnitm.2021@lpnu.ua, ORCID: 0000-0002-8199-0530

² E-mail: Victoria.A.Vysotska@lpnu.ua, ORCID: 0000-0001-6417-3689

³ E-mail: Roman.O.Holoshchuk@lpnu.ua, ORCID: 0000-0002-1811-3025

© Baran M., Vysotska V., Holoshchuk R., 2023

As a result of the research, one of the approaches to building an intelligent information system based on the recommendation of products to users with a solution to the cold start problem is described and modeled. The conducted research takes into account the advantages and disadvantages of the methods, as well as their compatibility, when combining them, which is an important factor for the speed of the system and the efficiency of the algorithm. The implementation of the hybrid method for the construction of an intelligent information system, as well as its performance testing in comparison with the classical k-means algorithm, was carried out. Based on the received analysis, a practical comparison of the efficiency of the system with the basic approach to solving the problem and the hybrid one was carried out.

Key words: intelligent information system; machine learning; cold start; k-means; Neural collaborative filtering.