

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Системотехніки
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

(позначення документа)

Розробка методів оцінювання ІТ-компаній
з використанням анонімних відгуків
(тема)

Виконав: здобувач групи ІТПм-21-2

спеціальності 122 Комп'ютерні науки

(код і повна назва спеціальності)

освітньої програми Інформаційні технології
проектування

(повна назва освітньої програми)

Білоус І.С.

(прізвище, ініціали)

Керівник к.т.н. Ситніков Д.Е.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри системотехніки

(підпис)

Гребеннік І.В.

(прізвище, ініціали)

2022 р.

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____
Кафедра _____ Системотехніки _____
Рівень вищої освіти _____ другий(магістерський) _____
Спеціальність _____ 122 Комп'ютерні науки _____
(код і повна назва)
Тип програми _____ освітньо-професійна _____
Освітня програма _____ Інформаційні технології проектування _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » _____ 20 ____ р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові _____ Білоусу Івану Сергійовичу _____
(прізвище, ім'я, по батькові)

1. Тема роботи Розробка методів оцінювання ІТ-компаній з використанням анонімних відгуків затверджена наказом університету від 03 жовтня 2022 р. № 1263СТ
2. Термін подання студентом роботи до екзаменаційної комісії 21 грудня 2022 р.
3. Вихідні дані до роботи Дослідити та розробити метод оцінювання ІТ-компаній з використанням анонімних відгуків з метою пришвидшення процесу пошуку підходящої компанії. Перелік використаних програмних засобів: ОС Ubuntu/MacOS, Webstorm, PostgreSQL, PgAdmin. Мова програмування додатку – JavaScript, з використанням фреймворків Nest.js та Next.js і бібліотеки React.
4. Перелік питань, що потрібно опрацювати в роботі Провести дослідження інформаційних систем; провести аналіз предметної області та існуючих досліджень в предметній області; проаналізувати існуючі аналоги системи, їх переваги та недоліки; оглянути підходи до побудови системи рекомендацій; сформулювати постановку задачі дослідження; розробити алгоритмічне та математичне забезпечення; розробити функціональні та нефункціональні вимоги до програмного забезпечення, створити діаграму варіантів використання; описати архітектуру системи; розробити модель даних; розробити веб додаток; проаналізувати роботу програмного забезпечення і навести приклади його використання.
5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Скріншоти, діаграма варіантів використання, діаграма розгортання, модель даних, діаграма послідовності, фотографія.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1.	Отримання завдання атестаційної роботи	03.10.22	Виконано
2.	Огляд існуючих досліджень в предметній області	04.10.22 – 10.10.22	Виконано
3.	Аналіз існуючих систем	13.10.22	Виконано
4.	Огляд підходів до побудови систем рекомендацій	19.10.22 – 22.10.22	Виконано
5.	Постановка задачі дослідження	23.10.22 – 25.10.22	Виконано
6.	Розробка алгоритмічного та математичного забезпечення	27.11.22 – 07.11.22	Виконано
5.	Створення основних програмних специфікацій	10.11.22 – 13.11.22	Виконано
6.	Проектування системи	15.11.22 – 18.11.22	Виконано
7.	Створення бази даних	20.11.22 – 23.11.22	Виконано
8.	Розробка програмного забезпечення	25.11.22 – 10.12.22	Виконано
11.	Оформлення пояснювальної записки	11.12.22 – 17.12.22	Виконано
12.	Опис графічних матеріалів та програмної документації	17.12.22 – 20.12.22	Виконано
13.	Представлення на рецензування	17.12.22	Виконано
14.	Представлення роботи до ДЕК	21.12.22	Виконано

Дата видачі завдання 3 жовтня 2022 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис)

к.т.н. Ситніков Д.Е.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка до магістерської кваліфікаційної роботи: 50 с., 9 табл., 17 рис., 25 джерел інформації.

ІНФОРМАЦІЙНА СИСТЕМА, ІТ-КОМПАНІЇ, ВІДГУКИ, ВАКАНСІЇ, СИСТЕМА РЕКОМЕНДАЦІЙ, ФІЛЬТРАЦІЯ НА ОСНОВІ ВМІСТУ, ВЕКТОРНА МОДЕЛЬ, TF-IDF ЗВАЖУВАННЯ

Об'єктом дослідження є процес підтримки створення анонімних відгуків на ІТ-компанії з використанням інтелектуальних систем.

Предметом дослідження є математичне, алгоритмічне та програмне забезпечення для створення анонімних відгуків на ІТ-компанії.

Мета досліджень – розробка методів оцінювання ІТ-компаній з використанням анонімних відгуків з метою пришвидшення процесу пошуку підходящої компанії.

У роботі проведено аналіз предметної області, в рамках якого встановлено актуальність проблеми, розглянуто проведені дослідження, існуючі продукти на ринку, здійснено ознайомлення з принципами систем рекомендацій та доведено доцільність створення платформи анонімних відгуків на ІТ-компанії. На основі отриманих під час аналізу предметної області знань створено метод оцінювання ІТ-компаній та розроблено відповідне алгоритмічне, математичне і програмне забезпечення. Наведена робота користувача з програмним забезпеченням і аналіз результатів.

Галузь застосування – ІТ-спільнота України. Розроблене ПЗ і метод оцінювання ІТ-компаній прискорять процес пошуку підходящої компанії для кандидата, який шукає працедавця.

ABSTRACT

Explanatory note for Masters diploma work: 50 p., 9 tables., 17 figures., 25 sources.

INFORMATION SYSTEM, IT-COMPANIES, REVIEWS, VACANCIES, RECOMMENDER SYSTEM, CONTENT-BASED FILTERING, VECTOR MODEL, TF-IDF WEIGHTING

The object of this study of this work is a process of creation of anonymous reviews for IT-companies using intelligent systems.

The subject of study is mathematical and algorithmic methods, and software for creating anonymous reviews for IT-companies.

The purpose of this work is to create methods of evaluation of IT-companies using anonymous reviews in order to speed up the process of searching for suitable companies.

In scope of this work the analysis of the domain area has been conducted. In scope of domain area analysis the relevance of the problem has been stated, existing researches have been reviewed, existing products on market have been examined, approaches in recommender systems have been reviewed and the importance of platform for anonymous IT-reviews has been proven. Using the knowledge received from domain area analysis the method of evaluation of IT-companies and corresponding mathematical and algorithmic methods, and software have been developed. The example of interaction between software and the end user and analysis of results are demonstrated in this work.

Field of application is the IT-community of Ukraine. Developed software and method of evaluation of IT-companies will speed up the process of searching for suitable companies.

ЗМІСТ

ПЕРЕЛІК ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ	8
ВСТУП	9
1 ОСНОВНІ ПРОБЛЕМИ, ЩО ПОВ’ЯЗАНІ З ОЦІНЮВАННЯМ ІТ-КОМПАНІЙ	11
1.1 Платформа анонімних відгуків на ІТ-компанії як інформаційна система	11
1.2 Огляд і аналіз веб-ресурсів для оцінювання ІТ-компаній з використанням відгуків	11
1.3 Існуючі дослідження з оцінювання ІТ-компаній	14
1.4 Використання системи рекомендацій для платформи анонімних відгуків	15
1.5 Існуючі підходи до побудови систем рекомендацій в предметній області	16
1.6 Постановка задачі дослідження	17
2 РОЗРОБКА АЛГОРИТМІЧНОГО ТА МАТЕМАТИЧНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ПОБУДОВИ ПЛАТФОРМИ АНОНІМНИХ ВІДГУКІВ ДЛЯ ОЦІНЮВАННЯ ІТ-КОМПАНІЙ	19
2.1 Опис методів, які використовуються в системах рекомендацій, та їх порівняння	19
2.2 Розробка математичного методу для системи рекомендацій ІТ-відгуків та компаній	20
3 РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ПОБУДОВИ ПЛАТФОРМИ АНОНІМНИХ ВІДГУКІВ ДЛЯ ОЦІНЮВАННЯ ІТ-КОМПАНІЙ	25
3.1 Розробка вимог до програмного забезпечення	25
3.1.1 Функціональні вимоги	25
3.1.2 Нефункціональні вимоги	26
3.2 Розробка структури бази даних	27
3.3 Розробка програмного забезпечення	29
4 ЗАСТОСУВАННЯ ПЛАТФОРМИ АНОНІМНИХ ВІДГУКІВ ДЛЯ ОЦІНЮВАННЯ ІТ-КОМПАНІЙ	32
4.1 Робота користувача з програмним забезпеченням	32
4.2 Платформа в дії	38
5 ЕКОНОМІЧНЕ ОБҐРУНТУВАННЯ ПРОЕКТУ ЩОДО РОЗРОБКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ	40
5.1 Обґрунтування доцільності розробки ПЗ	40
5.2 Оцінка конкурентоспроможності ПЗ у порівнянні з аналогом	40
5.3 Планування комплексу робіт з розробки ПЗ і оцінка трудомісткості робіт	41
5.4 Розрахунок проектних витрат на розробку ПЗ	43

5.5 Розрахунок витрат на впровадження ПЗ	46
ВИСНОВКИ	48
СПИСОК ДЖЕРЕЛ ІНФОРМАЦІЇ	49

ПЕРЕЛІК ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

ПЗ – програмне забезпечення
ІС – інформаційна система
СР – система рекомендацій
СУБД – система управління базами даних
HTML – Hypertext Markup Language
HTTP – Hypertext Transfer Protocol
IT – information technology
IDF – inverse document frequency
JSON – JavaScript Object Notation
JWT – JSON Web Token
NoSQL – not only SQL
REST – Representational State Transfer
SSR – Server Side Rendering
SQL – Structured Query Language
TF – term frequency
UML – Unified Modeling Language
UTM – Urchin Tracking Module.

ВСТУП

Сьогодні ІТ-ринок України розвивається дуже швидко. Відкривається велика кількість продуктових та аутсорсингових компаній, залучаються клієнти з-за кордону. Незважаючи на кількість пропозицій, вибір працедавця в ІТ-індустрії є нелегким процесом, який потребує багато часу.

З розумом обране місце роботи є запорукою успіху і подальшого розвитку ІТ-спеціаліста. Більше того, правильна людина на правильному місці здатна продуктивно працювати та позитивно впливати на економіку країни, рухаючи її вперед.

При виборі працедавця в ІТ-індустрії основним джерелом інформації людини є сайти компаній. Проте, дуже часто сайти компаній не містять достатньої кількості інформації. В цьому випадку людина може звернутися до ресурсів з оцінками ІТ-компаній. Найпоширенішими методами оцінювання є відгуки на роботодавців, результати опитувань та експертні оцінки.

Важливість відгуків та їх вплив на стан справ в ІТ-компанії доводиться в дослідженні Janis Cloos [1]. Компанії з гарною репутацією приваблюють більшу кількість потенційних співробітників, ніж компанії з поганою репутацією. Гарні оцінки компанії приваблюють більш досвідчених спеціалістів. Репутація також впливає на продуктивність та зацікавленість в роботі діючих працівників.

Варто зауважити, що існуючі платформи з відгуками на ІТ-компанії мають ряд проблем:

- неінформативні відгуки;
- велика кількість упереджених та куплених відгуків;
- відгуки з персональними даними працівників чи іншими чутливими даними;
- відсутність української локалізації.

Звідси виникає потреба в інформаційній системі, яка б не мала перелічених вище недоліків.

В даній роботі об'єктом дослідження є процес підтримки створення анонімних відгуків на ІТ-компанії з використанням інтелектуальних систем. Предметом дослідження є математичне, алгоритмічне та програмне забезпечення для створення анонімних відгуків на ІТ-компанії.

Метою даної роботи є розробка методів оцінювання ІТ-компаній з використанням анонімних відгуків з метою пришвидшення процесу пошуку підходящої компанії.

1 ОСНОВНІ ПРОБЛЕМИ, ЩО ПОВ'ЯЗАНІ З ОЦІНЮВАННЯМ ІТ-КОМПАНІЙ

1.1 Платформа анонімних відгуків на ІТ-компанії як інформаційна система

За визначенням інформаційна система (ІС) – це сукупність організаційних і технічних засобів для збереження та обробки інформації з метою забезпечення інформаційних потреб користувачів. В даному випадку такою ІС може бути платформа анонімних відгуків на ІТ-компанії, яка:

- є веб-застосунком, що складається з декількох компонент;
- дозволяє зберігати та обробляти дані щодо ІТ-компаній та відгуків;
- забезпечує користувачів інформацією щодо оцінок компаній;
- вирішує проблему пошуку працедавця.

Така платформа повинна мати механізм для створення анонімних та інформативних відгуків, ручну перевірку рецензій адміністраторами та інтелектуальну частину в якості вбудованої системи рекомендацій відгуків та компаній.

Доцільність анонімності доводиться в дослідженні [1]. В цій роботі сказано, що високий ступінь анонімності сприяє чесній оцінці роботодавця.

Описана ІС за сферою діяльності є галузевою так як повинна застосовуватися в ІТ сфері. За рівнем автоматизації процесів управління вона є інформаційно довідковою, через те що орієнтована на розв'язування завдань пошуку інформації. За ступенем централізації обробки інформації - це централізована ІС, бо обробка інформації та обчислення могли б виконуватися на єдиному сервері.

1.2 Огляд і аналіз веб-ресурсів для оцінювання ІТ-компаній з використанням відгуків

Вебсайт DOU [2], який позиціонує себе як спільноту розробників, надає змогу користувачам залишати відгуки на ІТ-компанії (рис. 1.1).

По-перше, на DOU можна створити лише текстовий відгук. Тобто, відсутні будь які інші інтерактивні елементи для оцінювання компаній. По-друге, відгуки на DOU не є анонімними і для когось це може бути перешкодою через побоювання

написати правду. Проте, кожен може взаємодіяти з відгуками інших людей і підтримувати їх. Представники компаній мають змогу відповідати на відгуки від імені компанії. Відгук може залишити лише зареєстрований користувач. Вхід на сайт виконується за допомогою пошти, LinkedIn, GitHub або Facebook.

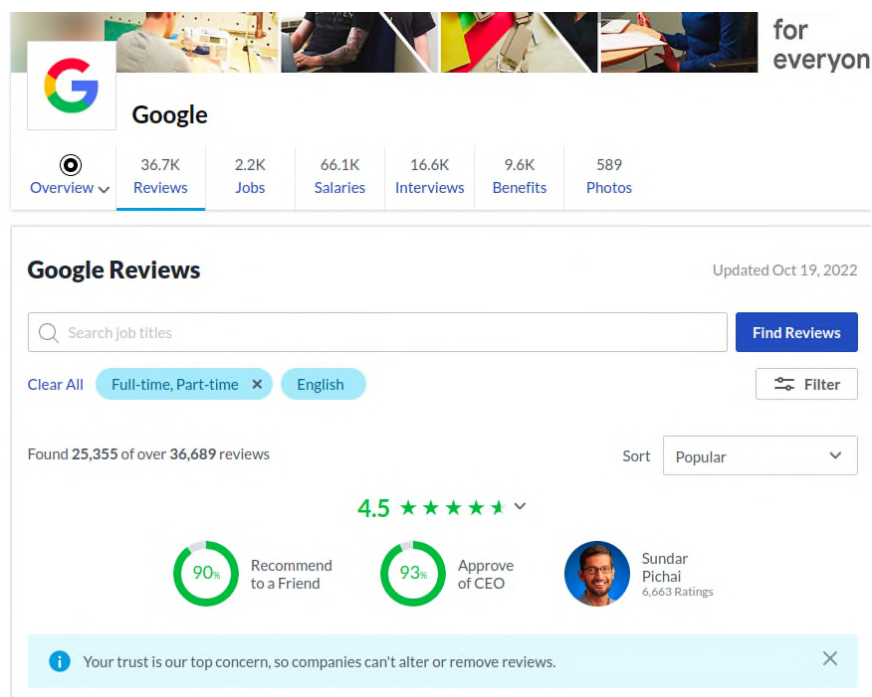


Рисунок 1.1 – Відгук на DOU

Іншим ресурсом, який вартий уваги є Glassdoor [3] (рис. 1.2). На відміну від DOU відгуки на Glassdoor є анонімними. Можуть використовуватися інтерактивні елементи, наприклад, щоб оцінити компанію за п'ятизначною шкалою. Також Glassdoor збирає від співробітників анонімну статистику заробітних плат і запитання та відповіді, які задають під час співбесіди в компанії.

Незважаючи на популярність Glassdoor у світі, він не є популярним в Україні і дуже багато українських компаній на сайті відсутні. Для входу може використовуватись пошта, LinkedIn або Google.

Також розглянемо Kununu [4] (рис. 1.3). Використовуючі п'ятизначну шкалу можна оцінити різні аспекти компанії. Наприклад, зарплату, пільги, корпоративну культуру, цікавість завдань та робоче середовище.



Google

Overview 36.7K Reviews 2.2K Jobs 66.1K Salaries 16.6K Interviews 9.6K Benefits 589 Photos

Google Reviews Updated Oct 19, 2022

Search job titles Find Reviews

Clear All Full-time, Part-time x English Filter

Found 25,355 of over 36,689 reviews Sort Popular

4.5 ★★★★★

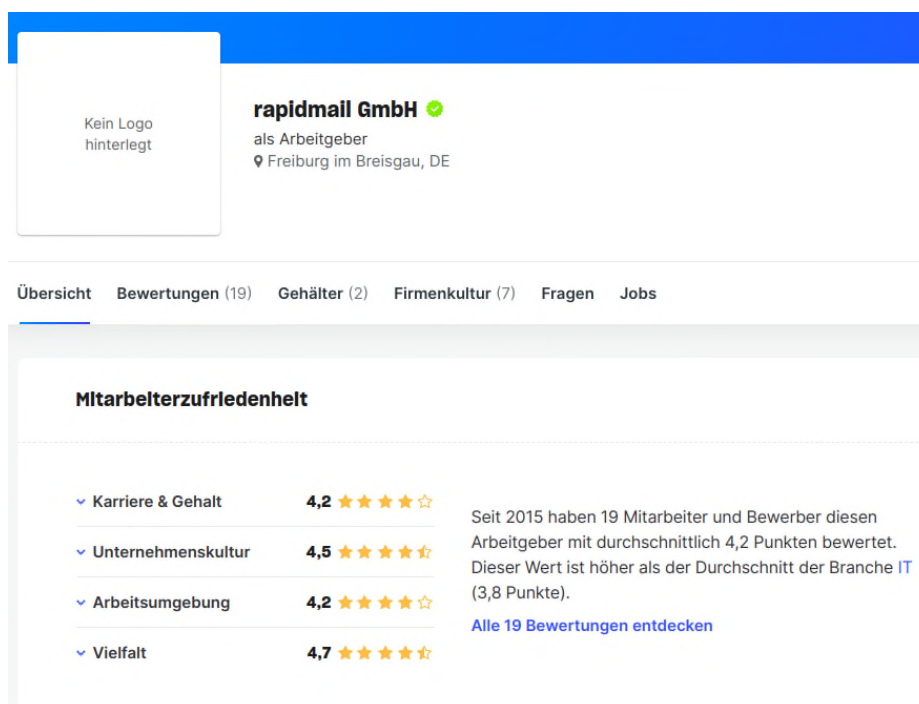
90% Recommend to a Friend 93% Approve of CEO

Sundar Pichai 6,663 Ratings

Your trust is our top concern, so companies can't alter or remove reviews.

Рисунок 1.2 – Сторінка компанії на Glassdoor

Недоліком Kijiji є цілковита орієнтованість на ІТ-ринок Німеччини. Серед доступних мов інтерфейсу лише німецька. Увійти на сайт можна лише з використанням пошти.



Kein Logo hinterlegt

rapidmail GmbH als Arbeitgeber
Freiburg im Breisgau, DE

Übersicht Bewertungen (19) Gehälter (2) Firmenkultur (7) Fragen Jobs

Mitarbeiterzufriedenheit

Karriere & Gehalt 4,2 ★★★★★

Unternehmenskultur 4,5 ★★★★★

Arbeitsumgebung 4,2 ★★★★★

Vielfalt 4,7 ★★★★★

Seit 2015 haben 19 Mitarbeiter und Bewerber diesen Arbeitgeber mit durchschnittlich 4,2 Punkten bewertet. Dieser Wert ist höher als der Durchschnitt der Branche IT (3,8 Punkte).
[Alle 19 Bewertungen entdecken](#)

Рисунок 1.3 – Сторінка компанії на Kijiji

В усіх рішеннях, які були розглянуті, спостерігаються проблеми описані у підрозділі 1.1. У випадку з DOU це малоінформативні, текстові відгуки, які доволі часто є упередженими, або містять чутливі дані. Glassdoor та Kununu не мають української локалізації та вбудованих систем рекомендацій для поліпшення досвіду користувача.

1.3 Існуючі дослідження з оцінювання ІТ-компаній

В аналітичному звіті Development of Ukrainian IT Industry [5] за 2018 рік представлено дуже багато інформації щодо ІТ-ринку України. Згідно цього дослідження в Україні на 2018 рік було більше 12.6 тисяч ІТ-компаній. Для порівняння, в 2015 було 8.2 тисячі. Більшість компаній є аутсорсинговими, і приблизно половина зареєстрована в Києві. Також багато компаній є в Харкові, Львові, Дніпрі та Одесі. Відзначається, що у 2017 році комп'ютерні послуги були третьою за величиною статтею експорту українських послуг.

Також цікавим поглядом на українську індустрію ІТ є дослідження Lviv IT Cluster 2018 року [6]. Наприклад, там відзначається, що у 2018 році у Львові стало на 28% більше ІТ-компаній і на 25% більше ІТ-спеціалістів у порівнянні з 2017 роком. Також на 8% зросла середня заробітна плата. Тое можна зробити висновок, що ця сфера в Україні дуже стрімко розвивається.

Автори [7] провели дослідження щоб дослідити фактори, які мотивують працівників в сфері інформаційних технологій. Для цього вони проаналізували 15000 відгуків на ІТ-компанії зі Сполучених Штатів з сайту Glassdoor. В результаті було виявлено, що на задоволення від роботи впливає комфортне робоче середовище та гарні стосунки між колегами. Найпопулярнішим фактором, який викликає незадоволення є втома.

Miruna Florina Lungu в своїй роботі [8] висвітлює як різні фактори впливають на продуктивність ІТ-компаній. Були опитані працівники індустрії інформаційних технологій і після використання регресійного аналізу було встановлено, що в якості головного рушія ефективності компанії відзначається стратегічна гнучкість.

1.4 Використання системи рекомендацій для платформи анонімних відгуків

Системи рекомендацій (CR) – це програмні інструменти та техніки, які пропонують певні елементи користувачу. Пропозиції, які надаються користувачу, стосуються різних процесів прийняття рішень, наприклад, які товари купувати, яку музику слухати, чи які новини в інтернеті читати [9].

CR є системами обробки інформації, які активно збирають різні види даних. Дані в першу чергу стосуються елементів, що пропонуються, та користувачів, які отримують ці рекомендації. Оскільки джерела даних, що є в розпорядженні цих систем, можуть бути дуже різноманітними, можливість їх використання залежить від конкретної техніки рекомендацій.

Елемент – загальний термін, що використовується для позначення того, що система рекомендує користувачам. CR зазвичай фокусується на конкретному типі елементів. Наприклад, на фільмах чи музиці.

Користувачі CR можуть мати дуже різноманітні цілі та характеристики. Для того, щоб персоналізувати рекомендації, CR використовує широкий спектр інформації про користувачів, яка може бути структурована різними способами. Наприклад, користувачів можна описати патернами їх поведінки: історією переглядів веб-сайтів, історією пошуку, історією використання тих чи інших інструментів тощо.

Дві основні стратегії створення систем рекомендацій – «фільтрація на основі вмісту» (Content-based filtering) та «колаборативна фільтрація» (Collaborative filtering). При фільтрації на основі вмісту система вчиться рекомендувати елементи, які схожі на ті, що сподобалися користувачеві в минулому. Подібність елементів обчислюється на основі їх ознак. Наприклад, якщо користувач позитивно оцінив фільм, що належить до комедійного жанру, то система може навчитися рекомендувати інші фільми цього жанру. При колаборативній фільтрації система рекомендує активному користувачу елементи, які сподобались іншим користувачам зі схожими смаками в минулому. Схожість вподобань двох користувачів обчислюється на основі подібності їх історій рейтингів [9].

В платформі анонімних відгуків, яка описується, система рекомендацій може використовуватися для надання користувачам рекомендацій щодо найбільш підходящих ІТ-компаній (потенційних працедавців) або вакансій.

1.5 Існуючі підходи до побудови систем рекомендацій в предметній області

В своїй роботі Roshan G. Belsare та Dr. V. M. Deshmukh [10] описують різні можливості створення системи рекомендацій для надання користувачам пропозицій щодо вакансій. Розглядаються і порівнюються три підходи:

- matching based recommendation design;
- content based recommendation design;
- collaborative filtering based recommendation design.

В першому підході користувачам рекомендуються вакансії, які їм підходять по параметрам їх профіля. Наприклад, місто, штат, досвід роботи. В другому підході, користувачу пропонуються вакансії, які максимально схожі на ті, що привернули багато його уваги в минулому. В третьому підході вибудовується схожість користувачів на основі їх профілів. Подібним користувачам рекомендуються подібні вакансії.

На думку авторів така система дозволить людям витратити менше часу на пошук вакансій в різноманітних каталогах і великих базах даних. В своїй іншій роботі [11] дослідники описують складнощі під час розробки таких систем. В якості найскладнішого випробування описується обробка вхідних даних через велику кількість різноманітних параметрів користувачів і вакансій.

В роботі Jeevankrishna [12] описується створення веб-застосунку з системою рекомендацій вакансій всередині. Вхідними даними є штучно створені профілі користувачів на основі результатів щорічного опитування StackOverflow та великий набір даних вакансій. В кінці демонструється робота застосунку, де через графічний інтерфейс користувача задається ID користувача для подальшого відображення вакансій у вигляді таблиці.

В роботі доводиться доцільність використання стратегії фільтрації на основі змісту (content-based filtering). Для визначення подібності використовується косинусна міра подібності (cosine similarity).

1.6 Постановка задачі дослідження

Кожна ІТ-компанія з точки зору працівника характеризується множиною характеристик привабливості. Наприклад, гнучкий графік, дружній колектив, гарний офіс, висока зарплата, кваліфікована команда тощо. Такі характеристики, як позитивні так і негативні, можуть бути складовими відгука на компанію. У попередніх розділах відзначено, що відгуки є джерелом важливої інформації для існуючих співробітників, потенційних співробітників та потенційних клієнтів (замовників).

Головною задачею даної роботи є розробка методів оцінювання ІТ-компаній з використанням анонімних відгуків з метою пришвидшення процесу пошуку підходящої компанії.

Об'єктом дослідження є процес підтримки створення анонімних відгуків на ІТ-компанії з використанням інтелектуальних систем. Предметом дослідження є математичне, алгоритмічне та програмне забезпечення для створення анонімних відгуків на ІТ-компанії.

Попередньо було виконано аналіз предметної області. В тому числі було здійснено дослідження процесу та існуючих методів оцінки ІТ-компаній; виявлення особливостей розробки інформаційних систем з системою рекомендацій, огляд та аналіз існуючих продуктів на ринку, їх переваг та недоліків. Далі для досягнення мети роботи повинна бути виконана низка задач.

1 Математичний опис задачі та розробка технології оцінки ІТ-компаній з використанням анонімних відгуків. На цьому етапі описується запропонований метод оцінки, розробляється алгоритмічне та математичне забезпечення для системи рекомендацій відгуків.

2 Розробка програмного забезпечення для побудови платформи анонімних відгуків, яка базується на розробленому методі оцінювання ІТ-компаній. Даний етап

розпочинається з розробки функціональних та нефункціональних вимог, і створення основних програмних специфікацій. Далі необхідно, на основі набутих під час виконання попередніх задач знань, розробити модель даних та інші моделі для опису системи. Заключною задачею є безпосередня розробка ПЗ.

3 Опис застосування програмного забезпечення. На останньому етапі потрібно описати роботу користувача з платформою; навести приклади оцінювання ІТ-компаній з використанням анонімних відгуків і роботи системи рекомендацій; провести аналіз результатів.

2 РОЗРОБКА АЛГОРИТМІЧНОГО ТА МАТЕМАТИЧНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ПОБУДОВИ ПЛАТФОРМИ АНОНІМНИХ ВІДГУКІВ ДЛЯ ОЦІНЮВАННЯ ІТ-КОМПАНІЙ

2.1 Опис методів, які використовуються в системах рекомендацій, та їх порівняння

Більшість систем рекомендацій, що базуються на підході «фільтрація на основі вмісту», використовують такі моделі пошуку, як відповідність ключових слів (keyword matching) або векторну модель (Vector Space Model, VSM) з простим TF-IDF зважуванням та косинусною мірою подібності.

Векторна модель – це алгебраїчне представлення колекції документів векторами, які належать до одного, спільного для всієї колекції векторного простору. У цій моделі кожен документ представлений вектором у мірному просторі, де кожен вимір відповідає терміну із загального словника даної колекції документів.

TF-IDF – статистичний показник, що використовується для оцінки важливості термінів (слів) в контексті документа, що є частиною колекції документів (корпусу). Вага (значимість) терміну пропорційна кількості вживань цього терміну в документі, і обернено пропорційна частоті вживання терміну у інших документах колекції.

Якщо вихідні дані представлені у вигляді векторів бінарних значень то замість TF-IDF зважування та косинусної міри подібності можна, в якості альтернативи, використовувати бінарні міри подібності.

Коефіцієнт Жаккара є бінарною мірою подібності та використовується в інформатиці для пошуку подібних документів, плагіату тощо. Коефіцієнт Жаккара розраховується наступним чином [9]:

$$sim(d_i, d_j) = \frac{c}{a + b - c}, \quad i \neq j, \quad i = \overline{1, m}, \quad j = \overline{1, m}.$$

де a – це число термінів t_k , $k = \overline{1, n}$, які входять в документ d_i ;

b – число термінів t_k , $k = \overline{1, n}$, які входять в документ d_j ;

c – число термінів t_k , $k = \overline{1, n}$, які входять в і документ d_i , і в документ d_j .

Іншою бінарною мірою подібності є коефіцієнт Соренсена [13]:

$$\text{sim}(d_i, d_j) = \frac{2c}{a+b}, \quad i \neq j, \quad i = \overline{1, m}, \quad j = \overline{1, m}.$$

Ще однією бінарною мірою є коефіцієнт Сімпсона [14]:

$$\text{sim}(d_i, d_j) = \frac{c}{\min(a, b)}, \quad i \neq j, \quad i = \overline{1, m}, \quad j = \overline{1, m}.$$

Варто відзначити, що при значно більшій кількості параметрів, підхід з TF-IDF зважуванням може бути точнішим, завдяки акцентуванню уваги на більш важливі терміни під час порівняння.

2.2 Розробка математичного методу для системи рекомендацій IT-відгуків та компаній

Запропонована система базується на принципах які були описані в роботі [21]. Система рекомендацій використовує підхід «Content-based filtering». Користувач переглядає відгуки на IT-компанії та відзначає ті, що йому сподобались. Далі користувачу рекомендуються інші відгуки, які схожі на ті, що йому сподобались. У більшості випадків це будуть відгуки на інші IT-компанії, а, отже, це сприятиме виявленню нових компаній, які за своїми ознаками та оцінками підходять користувачу. Три основні кроки процедури формування рекомендацій:

- 1) визначення трьох останніх відгуків, які сподобались користувачу;
- 2) визначення по два найбільш схожих відгуків на ті, що сподобались користувачу;

3) формування списку з шести рекомендованих відгуків.

Пропонується TF-IDF підхід, в якому в якості документу використовується відгук на IT-компанію. Кожен відгук складається текстового опису та ознак, якими рецензент описав компанію. Наприклад, гарний офіс, гнучкий графік, висока зарплата тощо. Колекцією документів є загальний набір відгуків.

Нехай $D = \{d_1, d_2, \dots, d_m\}$ – це набір документів (корпус), де $d_j, j = \overline{1, m}$ – це відгук, представлений вектором довжини n , а m – кількість документів в колекції. $T = \{t_1, t_2, \dots, t_m\}$ – це загальний словник колекції документів, який містить n термінів. Кожен термін словника це ознака за якою можна оцінити IT-компанію, наприклад, $t_1 = \text{" гнучкий графік"}$.

Кожен документ колекції $d_j, j = \overline{1, m}$ може бути представлений вектором $d_j = \{w_{1j}, w_{2j}, \dots, w_{nj}\}$, де w_{kj} – це вага терміну $t_k, k = \overline{1, n}$ в документі $d_j, j = \overline{1, m}$.

Представлення документів з використанням векторної моделі дозволяє приділити увагу двом питанням: зважуванню термінів та визначенню подібності векторів. Методика зважування TF-IDF базується на наступних емпіричних спостереженнях щодо тексту:

- рідкісні терміни не менш релевантні, ніж ті, що зустрічаються частіше (так зване IDF припущення);
- багаторазове входження терміну в документ є менш релевантним, ніж одиничне (так зване TF припущення);
- великі за об'ємом документи не мають переваги перед малими (так зване припущення нормалізації).

Іншими словами, терміни, які часто зустрічаються в одному документі, але рідко в інших документах, частіше мають відношення до теми документа. Крім того, нормалізація результуючих векторів ваг дозволяє позбутися переваги великих документів.

Описані вище припущення добре ілюструє функція TF-IDF [9]:

$$\text{TF-IDF}(t_k, d_j) = \frac{n_j}{\underbrace{\sum_{l=1}^n n_l}_{\text{TF}}} \cdot \log \frac{m}{\underbrace{n_k}_{\text{IDF}}},$$

де m – кількість документів (SQL-кодів запитів) в колекції;

n_k – кількість документів в колекції, в яких термін t_k , $k = \overline{1, n}$ зустрічається хоча б один раз;

n_j – число входжень терміну t_k , $k = \overline{1, n}$ в документ d_j , $j = \overline{1, m}$;

n_l – число входжень кожного терміну t_l , $l = \overline{1, n}$ в документ d_j , $j = \overline{1, m}$.

Для того, щоб значення ваги знаходилося в інтервалі $[0, 1]$ і документи були представлені векторами однакової довжини, ваги повинні бути нормалізовані:

$$w_{kj} = \frac{\text{TF-IDF}(t_k, d_j)}{\sqrt{\sum_{s=1}^{|T|} \text{TF-IDF}(t_s, d_j)^2}}, \quad k = \overline{1, n}, j = \overline{1, m}.$$

Для визначення подібності двох векторів необхідний показник схожості, наприклад, косинусна міра подібності [9]:

$$\text{sim}(d_i, d_j) = \frac{\sum_{k=1}^n (w_{ki} \cdot w_{kj})}{\sqrt{\sum_{k=1}^n w_{ki}^2} \cdot \sqrt{\sum_{k=1}^n w_{kj}^2}}, \quad i \neq j, i = \overline{1, m}, j = \overline{1, m}.$$

З використанням формули подібності, ознаки трьох останніх відгуків, які сподобались користувачу, порівнюються з іншими відгуками колекції D . Для кожного відгука визначаються по два найбільш схожих.

Для прикладу розглянемо спрощену ситуацію з п'ятьма відгуками, кожен з яких має дві ознаки. Всього ознак – сім. Дані представлені в таблиці 2.1.

Таблиця 2.1 – Ознаки п'яти відгуків

d_j	Ознака 1	Ознака 2	t_1	t_2	t_3	t_4	t_5	t_6	t_7
d_1	Гарний офіс	Курси	1	0	0	0	1	0	0
d_2	Гарний офіс	Цікаві завдання	1	0	0	0	0	1	0
d_3	Гнучкий графік	Дружній колектив	0	1	0	0	0	0	1
d_4	Висока зарплата	Дружній колектив	0	0	1	0	0	0	1
d_5	Міжнародна	Курси	0	0	0	1	1	0	0

Застосуємо TF-IDF зважування та, з використанням косинусної міри подібності, знайдемо відгуки, які найбільш схожі на перший. Результати представлені в таблиці 2.2.

Таблиця 2.2 – Результати TF-IDF зважування та порівняння

d_j	w_{1j}	w_{2j}	w_{3j}	w_{4j}	w_{5j}	w_{6j}	w_{7j}	$sim(d_1, d_j)$
d_1	0.707	0	0	0	0.707	0	0	1
d_2	0.572	0	0	0	0	0.819	0	0.405
d_3	0	0.819	0	0	0	0	0.573	0
d_4	0	0	0.819	0	0	0	0.573	0
d_5	0	0	0	0.819	0.573	0	0	0.405

В результаті, на перший відгук в однаковій мірі схожі відгуки номер два та номер п'ять. Через те, що в прикладі використовувалась мала кількість відгуків з

малою кількістю параметрів, можна легко перевірити достовірність результатів. Відгук номер два схожий на відгук номер один значенням параметру «Ознака 1», а запит номер п'ять схожий на запит номер один значенням параметру «Ознака 2».

3 РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ПОБУДОВИ ПЛАТФОРМИ АНОНІМНИХ ВІДГУКІВ ДЛЯ ОЦІНЮВАННЯ ІТ-КОМПАНІЙ

3.1 Розробка вимог до програмного забезпечення

3.1.1 Функціональні вимоги

Взаємодію з платформою анонімних відгуків здійснюють два типи користувачів: клієнти та адміністратори. У переліку нижче представлені функціональні вимоги, які ставляться до програмного забезпечення.

1 Клієнт повинен мати можливість зареєструватися на платформі самостійно, використовуючи LinkedIn або GitHub.

2 Клієнт повинен мати можливість переглядати ІТ-компанії та анонімні відгуки на них.

3 Клієнт повинен мати можливість відмічати відгуки інших користувачів через механізм “сподобалось”.

4 Клієнт повинен мати можливість залишити анонімний відгук на бажану ІТ-компанію. Повинна бути передбаченою можливість залишити корпоративну пошту для підтвердження відношення до компанії.

5 Клієнт повинен мати можливість переглядати рекомендовані відгуки.

6 Клієнт повинен мати можливість додати нову ІТ-компанію.

7 Клієнт не може додавати відгуки на одну й ту ж компанію частіше ніж раз на півроку.

8 Клієнт повинен отримувати листи з оновленнями про статус доданого відгука/компанії на електронну пошту.

9 Адміністратор повинен мати можливість переглядати список відгуків зі статусом “очікують підтвердження”. З метою перевірки відгука адміністратор повинен мати можливість переглядати ім'я користувача, що залишив відгук, домен його корпоративної пошти та сторінку LinkedIn, якщо вона є.

10 Адміністратор повинен мати можливість підтвердити або відхилити відгук клієнта.

11. Адміністратор повинен мати можливість переглядати список ІТ-компаній зі статусом “очікують підтвердження”.

12. Адміністратор повинен мати можливість підтвердити або відхилити додавання нової ІТ-компанії клієнтом.

Таким чином були описані функціональні вимоги до програмного забезпечення.

3.1.2 Нефункціональні вимоги

До програмного забезпечення ставляться вимоги щодо ефективності, зручності використання, безпеки та надійності. У переліку нижче представлені ці нефункціональні вимоги.

1 Повинна бути забезпечена швидка реакція системи з метою інформування користувача про стан задачі, яка виконується.

2 Всі дії в системі повинні виконуватись через максимально зручний графічний інтерфейс користувача.

3 Система має забезпечувати належний рівень безпеки. Відгуки повинні бути повністю анонімними і не містити персональних даних. Тільки адміністратор повинен мати змогу підтверджувати додані відгуки та компанії. Після додавання корпоративної електронної пошти клієнт повинен підтвердити її через отримання відповідного листа.

4 Система повинна бути стійкою до збоїв, а помилкові дії користувача не повинні пошкоджувати дані в базі даних.

Можливості клієнта та адміністратора при взаємодії з платформою анонімних відгуків на ІТ-компанії наведені на рисунку 3.1.

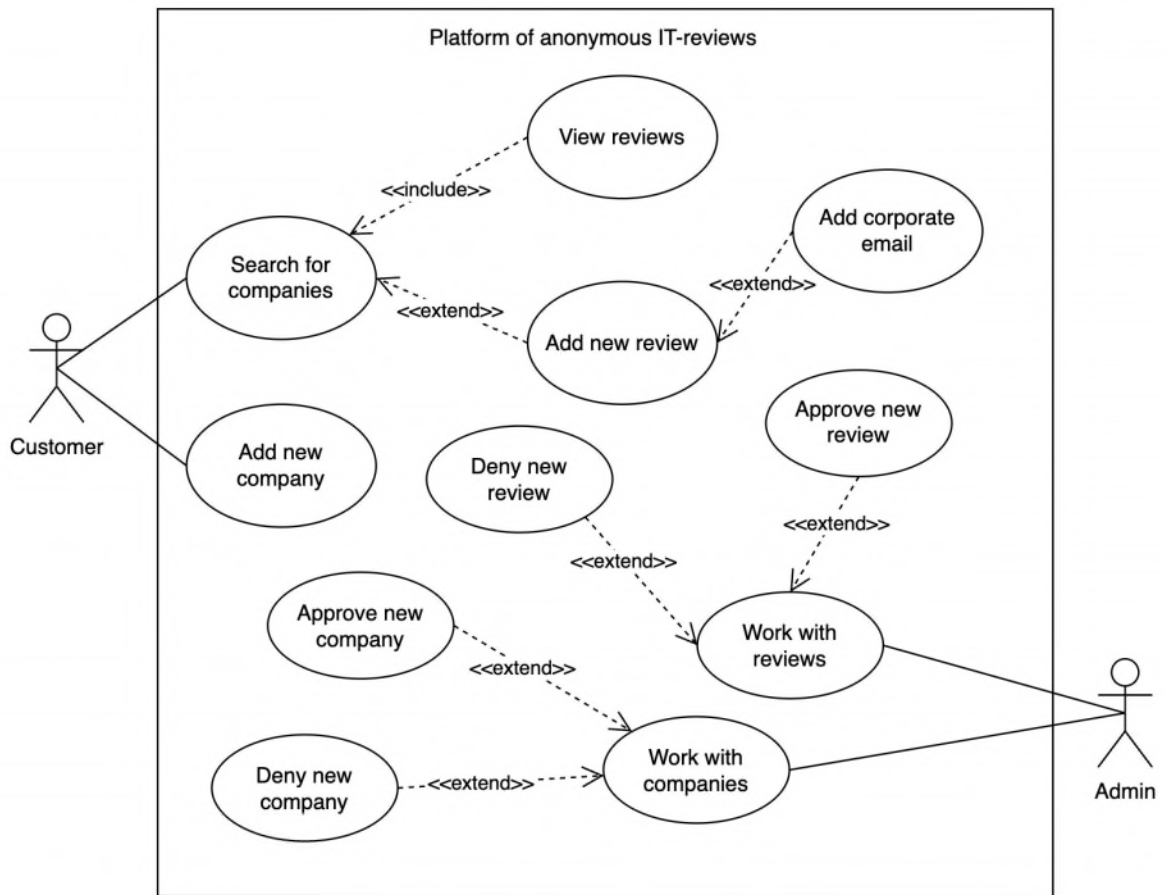


Рисунок 3.1 – Діаграма варіантів використання

Таким чином були описані нефункціональні вимоги до програмного забезпечення.

3.2 Розробка структури бази даних

Система повинна мати власну базу даних для зберігання інформації про користувачів, IT-компанії та відгуки. Логічна модель даних в нотації IDEF1X представлена на рисунку 3.2.

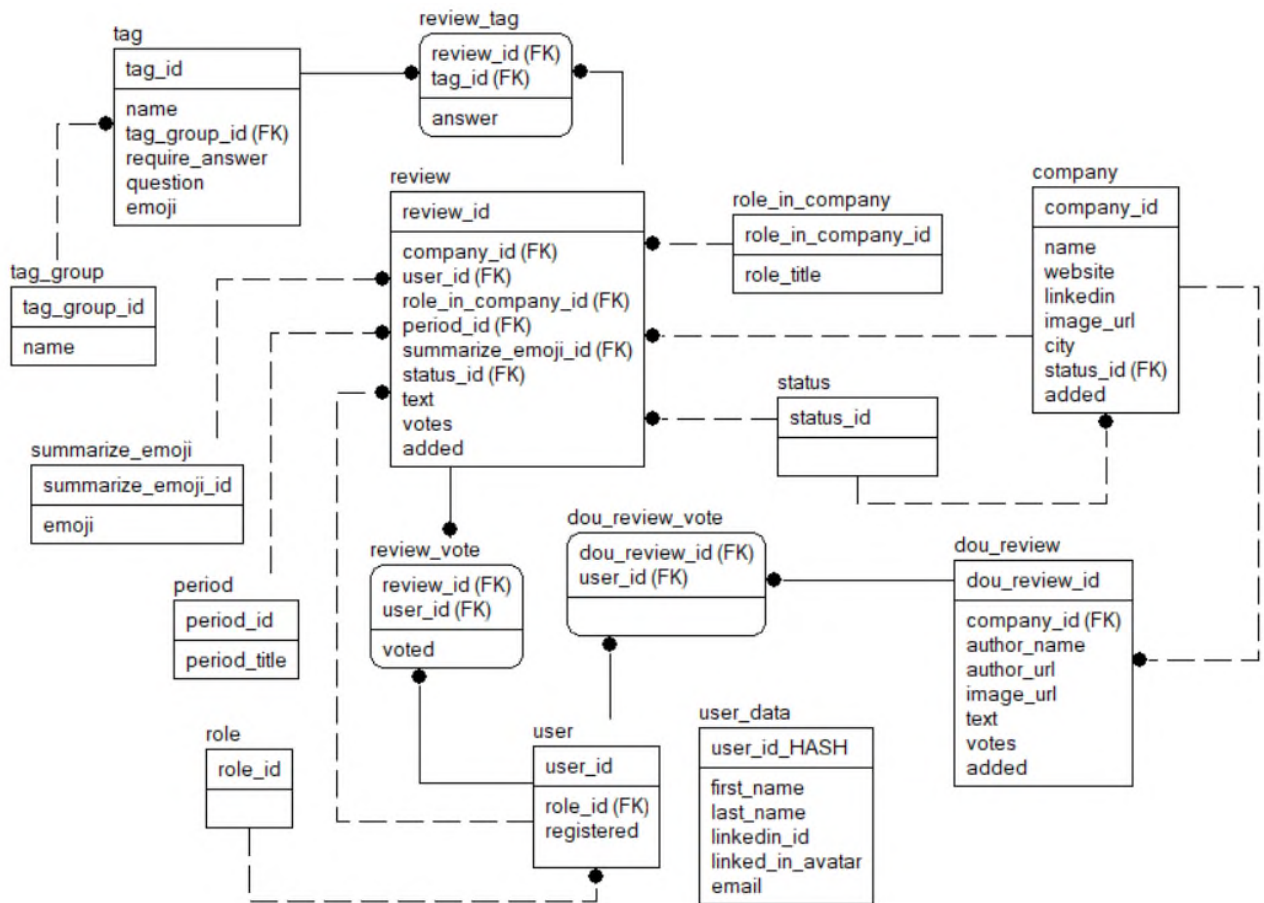


Рисунок 3.2 – Логічна модель даних

Ключові сутності представленої моделі даних описані в таблиці 3.1.

Таблиця 3.1 – Призначення ключових сутностей логічної моделі даних

№	Сутність	Призначення
1	user	Користувач системи (клієнт або адміністратор)
2	user_data	Чутливі дані користувача для їх зберігання в зашифрованому вигляді
3	company	ІТ-компанія
4	review	Відгук на ІТ-компанію
5	status	Статус відгука або компанії
6	tag	Ознака компанії для відгука

Таким чином була розглянута запропонована модель даних.

3.3 Розробка програмного забезпечення

Програмне забезпечення пропонується будувати на клієнт-серверній архітектурі, компонентами якої є сервер баз даних, сервер backend-застосунку, сервер frontend-застосунку та браузер в якості клієнта. Діаграма розгортання представлена на рисунку 3.3.

Серверна частина має REST архітектуру та працює на JavaScript-фреймворку Nest.js для Node.js. Застосунок реалізує бізнес-логіку, взаємодіє з сервером frontend-застосунку і клієнтом (браузером) через HTTPS протокол, а з сервером баз даних через відповідний драйвер. Клієнтська частина реалізована з використанням JavaScript-бібліотеки React та фреймворку Next.js. В якості бази даних використовується безкоштовна та потужна PostgreSQL.

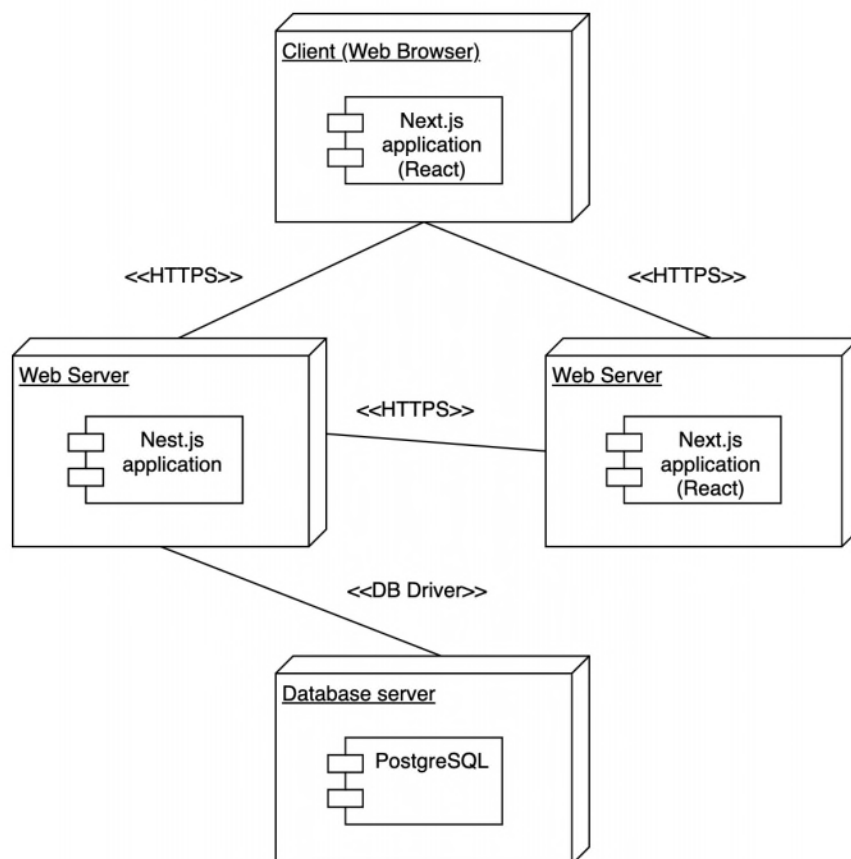


Рисунок 3.3 – Архітектура програмного забезпечення

Фреймворк Next.js дозволяє використовувати SSR (Server Side Rendering) та здійснювати перше будівництво веб-сторінки ще на сервері frontend-застосунку. Завдяки цьому Google Search Engine зможе отримувати більше інформації про сторінку та відображати її в пошуку, якщо вона релевантна.

Для авторизації пропонується використовувати стандарт OAuth 2, для входу до системи через LinkedIn або GitHub, та JSON Web Token (JWT). Вибір LinkedIn та GitHub в якості провайдерів зумовлений тим, що цільовою аудиторією платформи є IT-спеціалісти і ці ресурси є найбільш популярними серед них. Отже, користувач входить до системи з використанням LinkedIn або GitHub, сервер відправляє клієнту сгенерований JWT токен і той його зберігає в cookies. При наступних зверненнях до серверу клієнт використовує цей токен в заголовку HTTP запити. Сервер перевіряє токен та у випадку його валідності надає клієнту доступ до захищених ресурсів. Ця процедура представлена на діаграмі послідовності, яка представлена на рисунку 3.4.

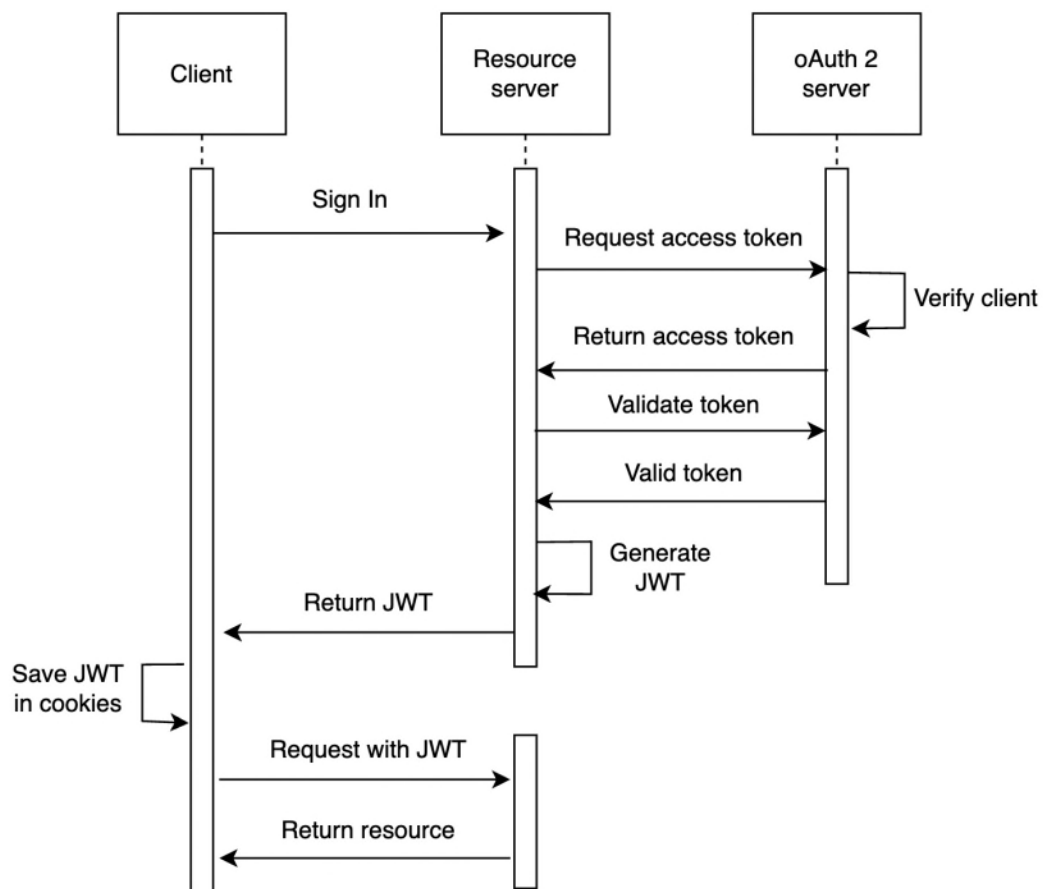


Рисунок 3.4 – Авторизація з використанням JWT токену

В якості інтегрованого середовища розробки пропонується використовувати WebStorm від JetBrains, який є безкоштовним для студентів вищих навчальних закладів. В якості системи контролю версій – GitHub.

4 ЗАСТОСУВАННЯ ПЛАТФОРМИ АНОНІМНИХ ВІДГУКІВ ДЛЯ ОЦІНЮВАННЯ ІТ-КОМПАНІЙ

4.1 Робота користувача з програмним забезпеченням

Головна сторінка сайту представлена на рисунку 4.1. На верхній панелі розташована кнопка для додавання нового відгука. З лівої частини – поле для пошуку ІТ-компаній за назвою. З правої частини – п'ять найсвіжіших відгуків.

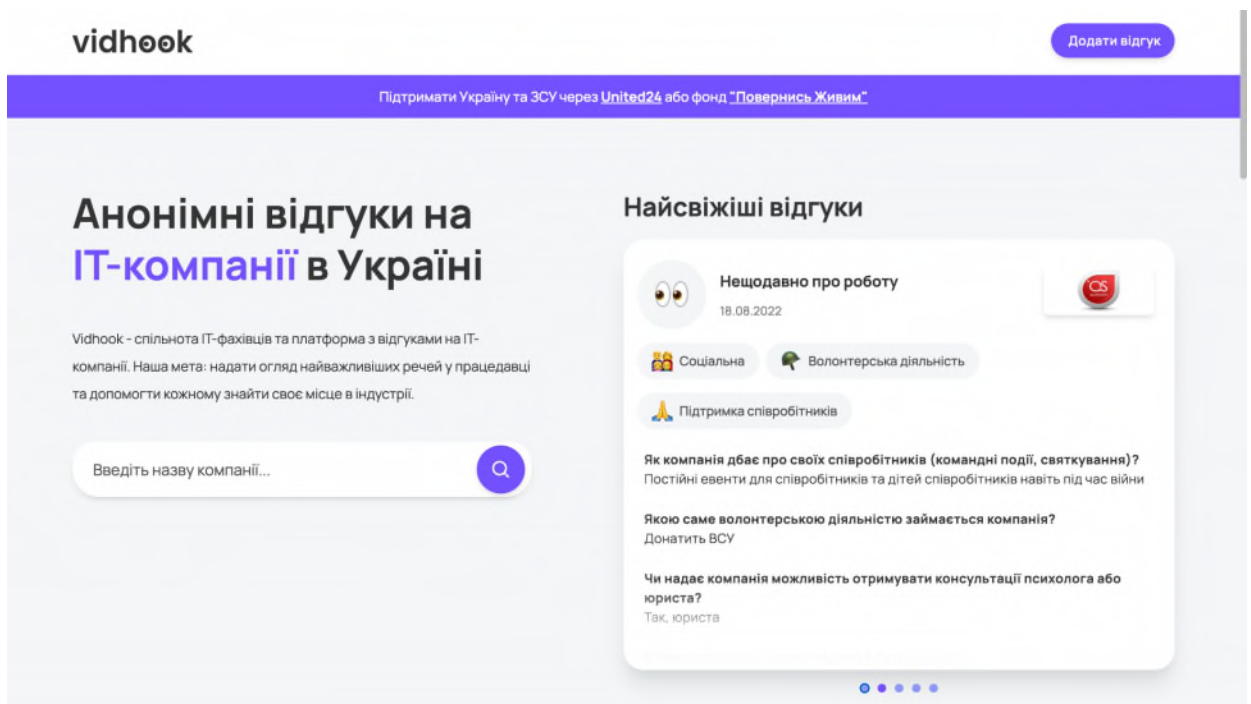


Рисунок 4.1 – Головна сторінка сайту

Щоб переглянути відгуки на бажану ІТ-компанію, клієнту необхідно скористатися пошуком компаній за назвою. Для пошуку доступні понад 8 тисяч компаній. Модальне вікно з інтерфейсом для пошуку зображене на рисунку 4.2.

На сторінці компанії можна побачити її назву, логотип та та сумарну кількість ознак, якими користувачі охарактеризували компанію. Кількість ознак розраховується на основі залишених відгуків. Так наприклад, якщо в чотирьох відгуках відзначено гарний офіс компанії, то в сумарній кількості буде чотири ознаки “Гарний офіс”. Сторінка компанії представлена на рисунку 4.3.

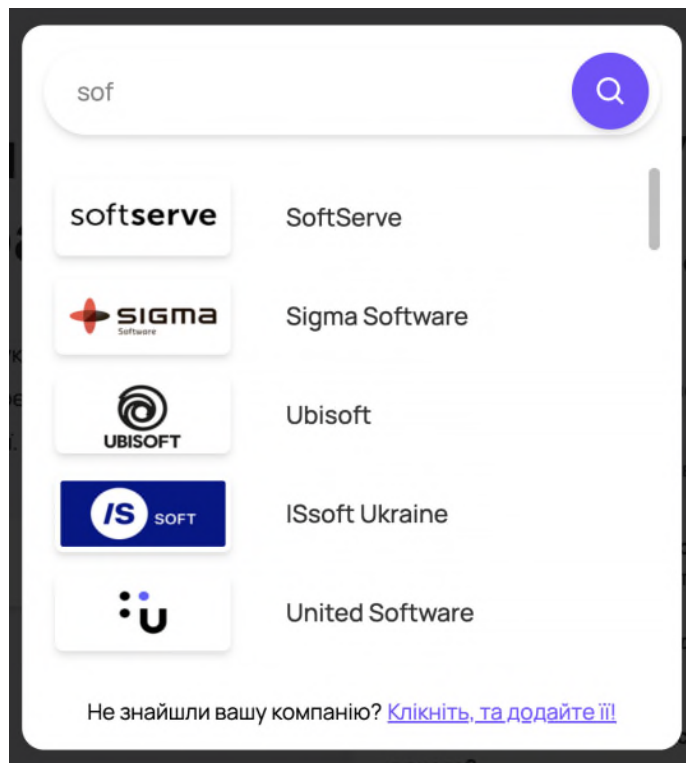


Рисунок 4.2 – Модальне вікно для пошуку ІТ-компаній

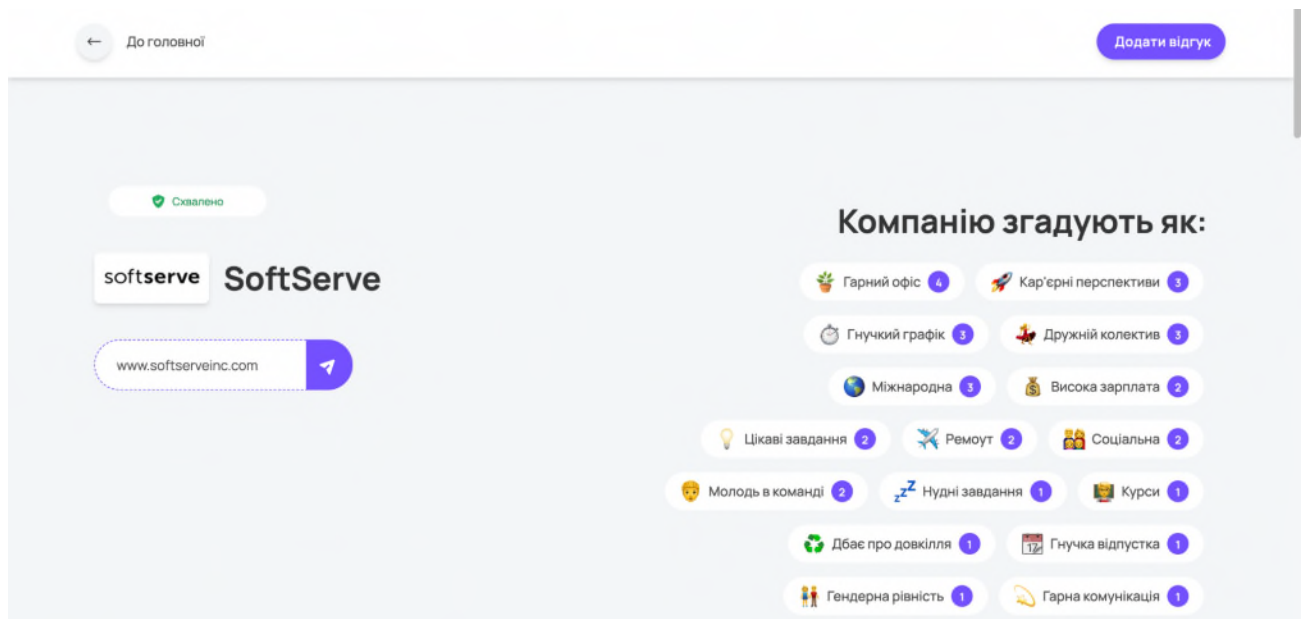


Рисунок 4.3 – Сторінка компанії

Нижче на сторінці компанії представлені відгуки на неї. Відгук складається з ознак, якими користувач охарактеризував компанію, відповідей на питання та текстового повідомлення. Питання підбираються в залежності від обраних ознак.

Так наприклад якщо користувач обрав ознаку “Дбає про довкілля”, буде задане питання “Що робить компанія для довкілля?”. Відгук наведено на рисунку 4.4.

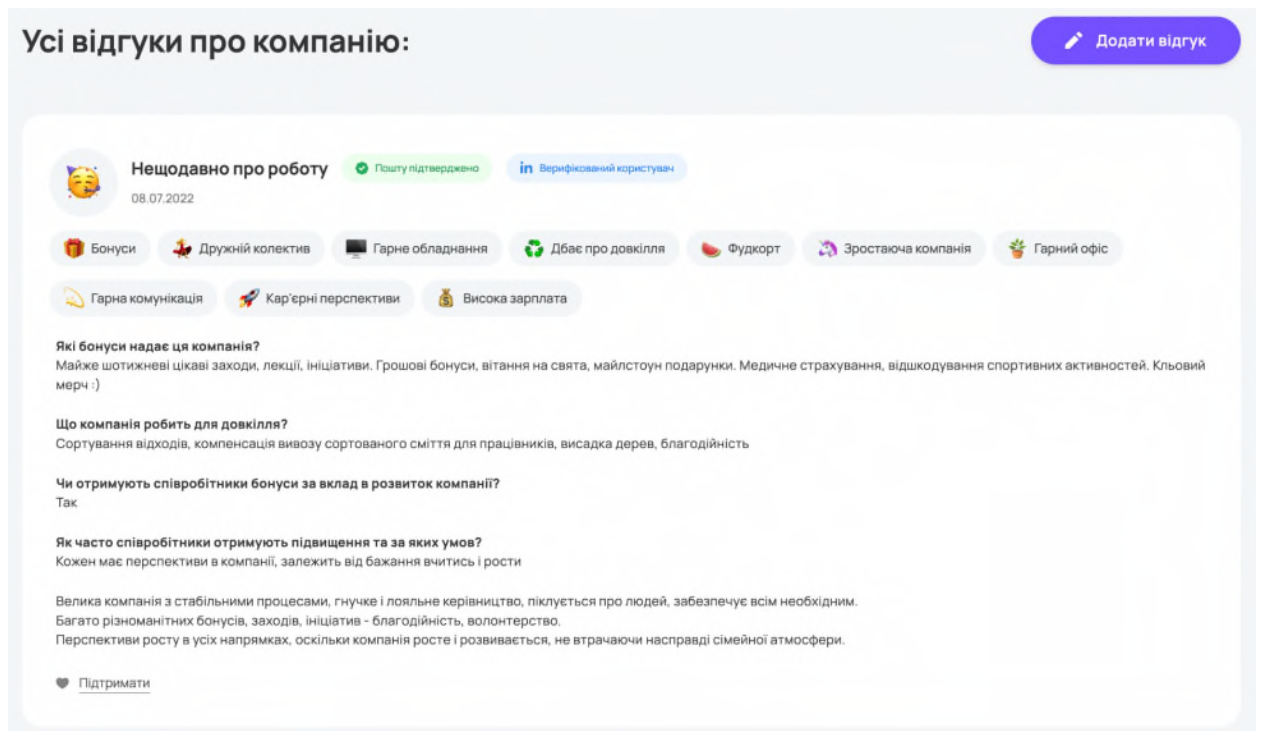


Рисунок 4.4 – Анонімний відгук

Щоб додати відгук потрібно на сторінці бажаної компанії натиснути на кнопку “Додати відгук”. У модальному вікні, на першому кроці, користувачу потрібно буде вказати свою роль та час коли він працював в компанії. На другому кроці компанію потрібно охарактеризувати ознаками та відповісти на поставлені запитання в текстовій формі. На третьому кроці необхідно додати своє повідомлення та вказати загальне враження від роботи обравши одне з п’яти запропонованих емої. На заключному етапі є можливість вказати свою корпоративну пошту, щоб додатково підтвердити своє відношення до компанії. Важливим моментом є створення облікового запису або вхід до існуючого через LinkedIn або Github. Це пропонується зробити в самому кінці, перед відправкою відгука на розгляд. Другий крок створення відгука представлено на рисунку 4.5, третій – на 4.6. Модальне вікно для входу до облікового запису наведено на рисунку 4.7.

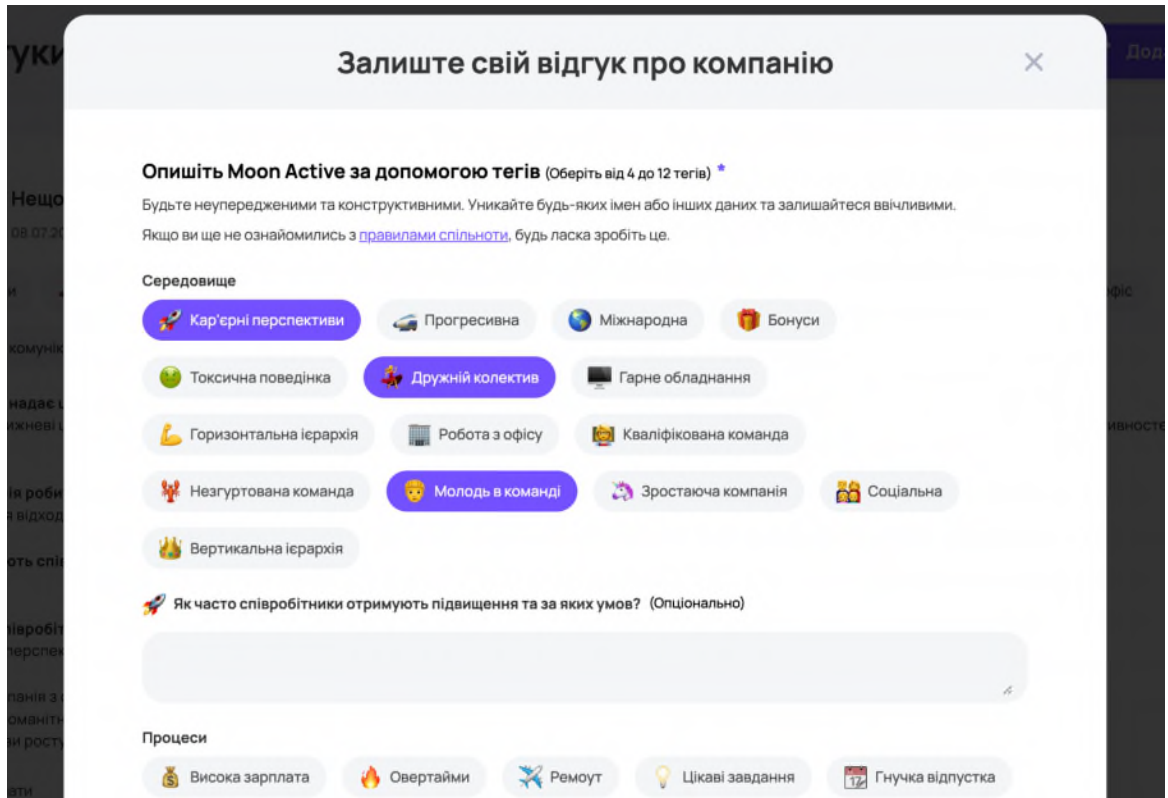


Рисунок 4.5 – Процес опису компанії запропонованими ознаками

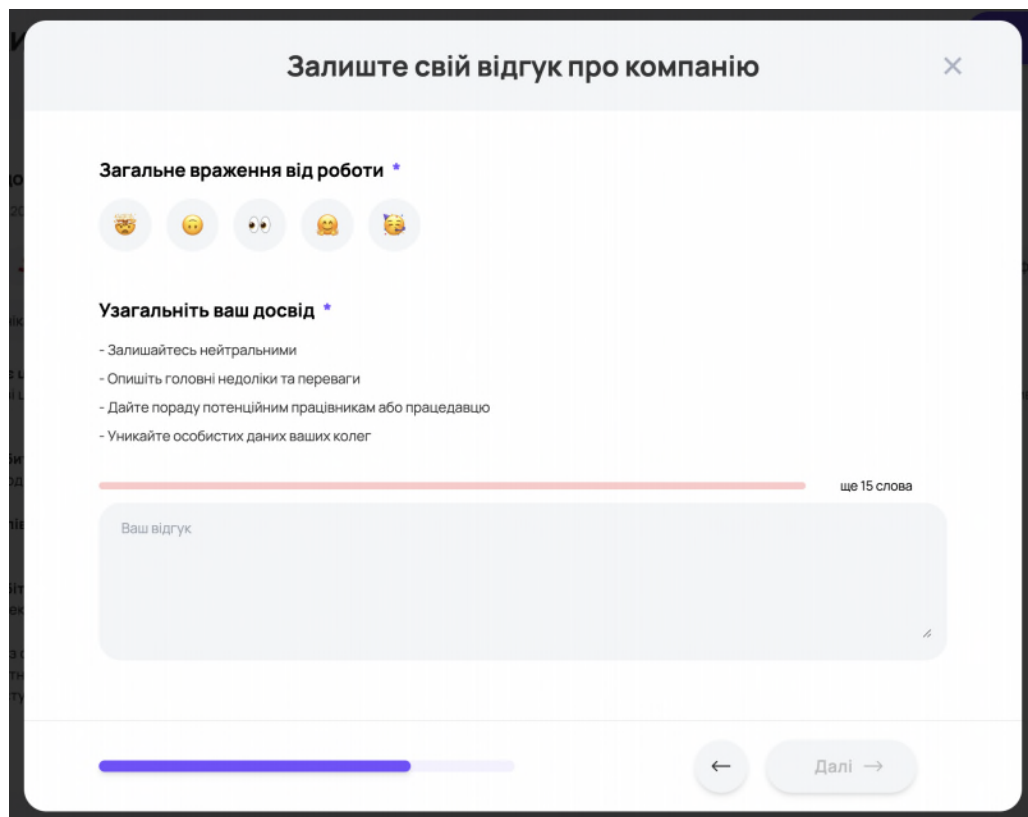


Рисунок 4.6 – Додавання текстового повідомлення до відгука

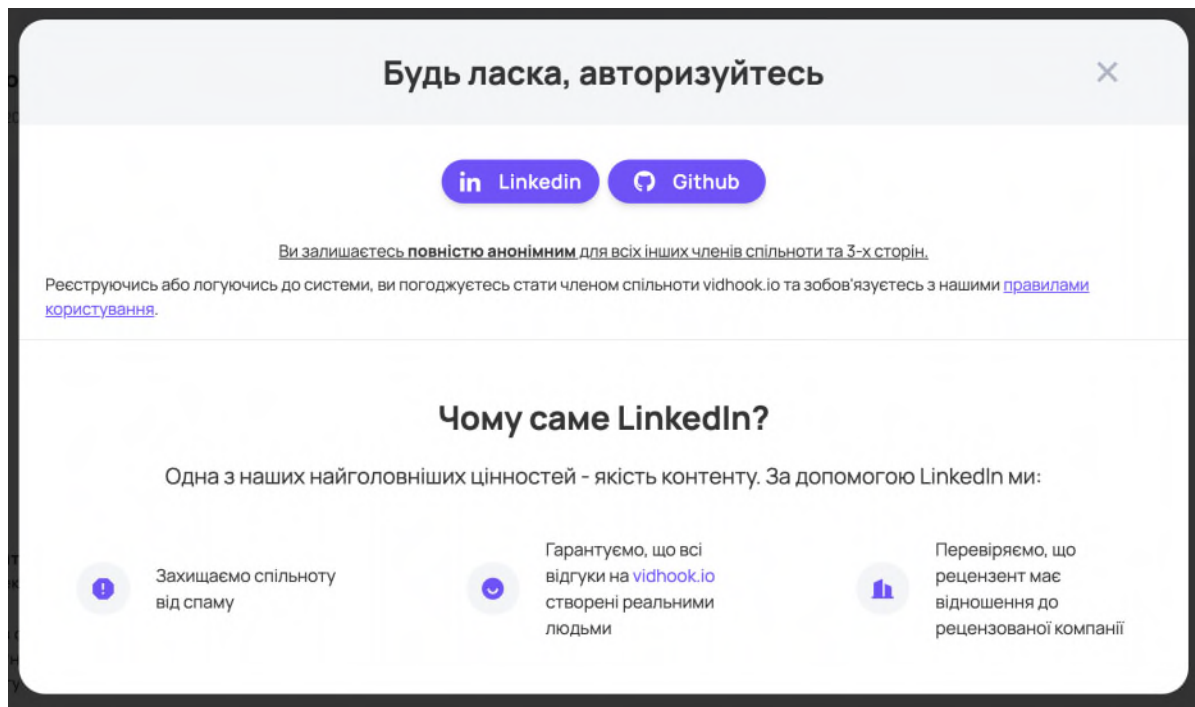


Рисунок 4.7 – Модальне вікно для входу до облікового запису

Оразу після того як новий відгук був створений клієнтом, адміністратор має змогу його побачити в списку відгуків, які очікують підтвердження. Адміністратор зобов'язаний перевірити наступні речі:

- відгук не містить персональних даних;
- відгук не є упередженим;
- відсутня ненормативна лексика;
- користувач має відношення до компанії.

Те, що користувач має відношення до компанії перевіряється насамперед через його сторінку в LinkedIn. Після того як адміністратор підтверджує відгук, він стає загальнодоступним на сторінці відповідної IT-компанії.

Якщо відгук було схвалено, користувач, який його залишив, отримує сповіщення на електронну пошту. Такий лист представлено на рисунку 4.8. У випадку, коли відгук було відхилено, користувач отримує листа, що наведений на рисунку 4.9.

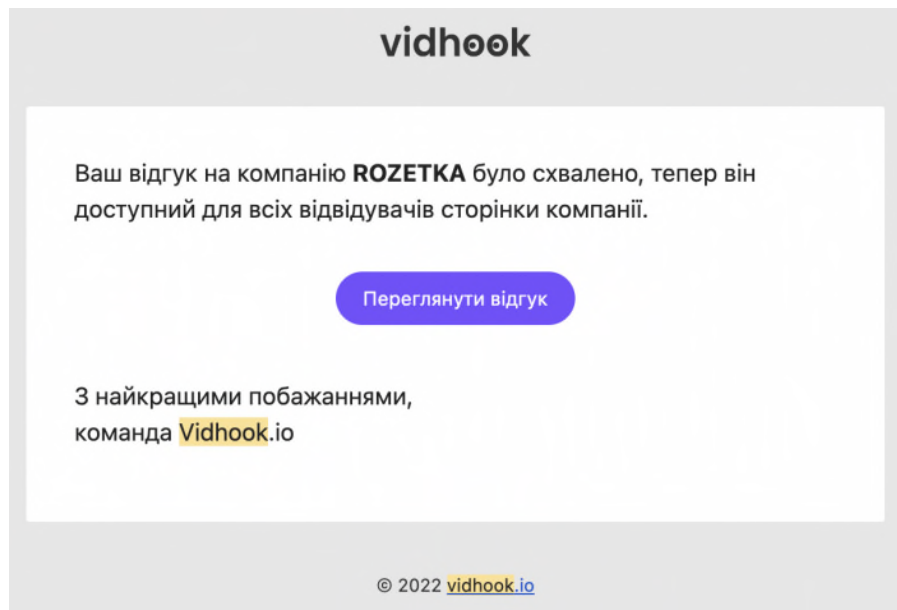


Рисунок 4.8 – Лист про схвалення відгука

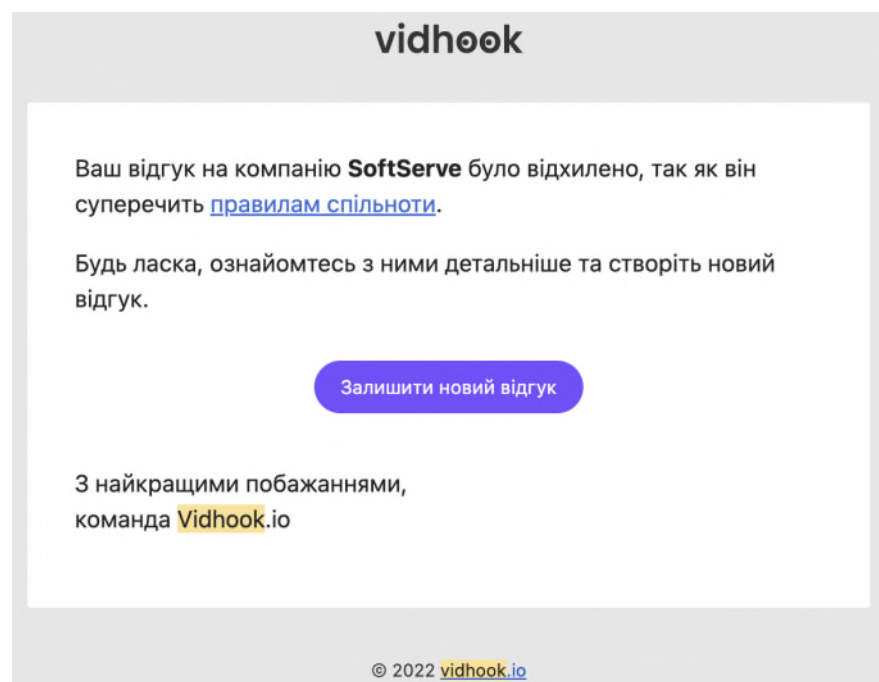


Рисунок 4.9 – Лист про відхилення відгука

Таким чином була описана робота користувача, а саме клієнта та адміністратора, з розробленим програмним забезпеченням.

4.2 Платформа в дії

На теперішній час описана в цій роботі платформа анонімних відгуків на ІТ-компанії має назву Vidhook. Вона вже працює і є доступною для українців [24]. Станом на сьогодні база даних платформи налічує понад 8 тисяч ІТ-компаній, 127 користувачів та 63 відгука.

Був здійснений експеримент з метою тестування способу залучення аудиторії через рекламу в Telegram. Так в одному з українських Telegram-каналів з новинами була куплена реклама за 2000 грн. В результаті, за добу на сайт перейшов 3761 користувач, зареєструвався 31 і було залишено 12 нових відгуків. Тобто конверсія склала трохи більше ніж 3%, що є доволі низьким показником. Виходить, що користувач коштував 64.5 грн, а відгук – 166.7 грн. Трафік відстежувався через UTM-мітки і платформу Google Analytics [25].

Також для залучення аудиторії використовувалась зовнішня реклама – у центрі міста Львів був придбаний сітілайт, який представлено на рисунку 4.10. Трафік з сітілайта не відстежувався.



Рисунок 4.8 – Рекламний сітілайт в центрі Львова

В майбутньому планується активний розвиток платформи з подальшими експериментами з маркетинг-стратегіями і реалізацією описаної в даній роботі системи рекомендацій відгуків та ІТ-компаній. Наведена математична модель буде ядром цієї системи і вкупі з запланованим розділом вакансій дозволить монетизувати ресурс.

5 ЕКОНОМІЧНЕ ОБҐРУНТУВАННЯ ПРОЕКТУ ЩОДО РОЗРОБКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

5.1 Обґрунтування доцільності розробки ПЗ

В першій частині роботи було проаналізовано існуючі продукти з метою виявлення їх переваг і недоліків. Головними недоліками є:

- неінформативні відгуки;
- велика кількість упереджених та куплених відгуків;
- відгуки з персональними даними працівників чи іншими чутливими даними;
- відсутність української локалізації.

Було доведено доцільність створення платформи анонімних відгуків для оцінювання ІТ-компаній, яка б не мала перелічених вище недоліків. Така платформа вирішує проблему оцінювання компаній і пришвидшує процес пошуку підходящої компанії.

5.2 Оцінка конкурентоспроможності ПЗ у порівнянні з аналогом

Для того щоб визначити чи є розробка ПЗ доцільною необхідно оцінити його конкурентоспроможність у порівнянні з аналогом. В якості такого аналогу був обраний портал DOU.

Для здійснення аналізу конкурентоспроможності ПЗ використовується оцінка його експлуатаційно-технічного рівня (ЕТУ). В свою чергу, ЕТУ визначається застосовуючи індекс ($J_{\text{ЕТР}}$) експлуатаційно-технічного рівня, який розраховується як сума приватних індексів, куди входять показники якості програмного продукту.

Визначення значущості окремих параметрів відбувається з використанням бально-індексного методу. Результати розрахунку $J_{\text{ЕТР}}$ бально-індексним методом за п'ятибальною шкалою наведені в таблиці 5.1, де B_j – коефіцієнт вагомості j -го показника в долях одиниці, який призначається відповідно до потреб організації

замовника програмного продукту, а X_j – експертна оцінка j -го показника якості за обраною шкалою оцінювання.

Таблиця 5.1- Розрахунок показника якості $J_{\text{ЕТР}}$

Показники якості програмного продукту	Коефіцієнт вагомості, B_j	Проект		Аналог	
		X_j	$X_j \cdot B_j$	X_j	$X_j \cdot B_j$
1. Зрозумілість інтерфейсу користувача	0.1	4	0.4	3	0.3
2. Простота використання	0.2	5	1	2	0.4
3. Швидкість навчання персоналу	0.1	4	0.4	3	0.3
4. Часова ефективність	0.15	4	0.6	3	0.45
5. Стабільність роботи	0.1	4	0.4	4	0.4
6. Надійність (захист даних)	0.2	4	0.8	5	1
7. Гнучкість щодо налаштування	0.05	2	0.1	4	0.2
8. Здатність до впровадження на підприємстві	0.1	5	0.5	3	0.3
Узагальнений показник якості $J_{\text{ЕТР}}$		4.2		3.35	

При порівнянні ПЗ, яке розробляється, з аналогом коефіцієнт технічного рівня буде дорівнювати:

$$A_k = 4.2 / 3.35 = 1.25.$$

Оскільки отриманий коефіцієнт більше 1, то розробка ПЗ з технічної точки зору є доцільною.

5.3 Планування комплексу робіт з розробки ПЗ і оцінка трудомісткості робіт

Розробкою ПЗ займаються дві людини: керівник та виконавець (інженер-програміст). Керівник поводить регулярні консультації та керує розробкою. Виконавець же відповідає за проектування архітектури ПЗ, розробку алгоритмів та бази даних, написання коду, тестування та багато інших дій, відповідно до плану проекту. Комплекс робіт представлений у таблиці 5.2.

Таблиця 5.2- Комплекс робіт проекту по розробці ПЗ

Зміст робіт	Виконавець	Тривалість, день	Навантаження викон.		
			Трив., день	%	Людино- дні
1. Аналіз предметної області	Керівник, Програміст	7	3 7	25 100	0.75 7
2. Постановка задачі дослідження	Керівник, Програміст	5	2 5	50 100	1 5
3. Розробка алгоритмічного та математичного забезпечення	Керівник, Програміст	7	3 7	25 100	0.75 7
4. Розробка вимог до ПЗ	Керівник, Програміст	5	3 5	50 100	1.5 5
5. Розробка моделей для опису ПЗ	Керівник, Програміст	7	3 7	40 100	1.2 7
6. Розробка ПЗ	Керівник, Програміст	35	5 35	25 100	1.25 35
7. Опис застосування ПЗ	Керівник, Програміст	5	2 5	30 100	0.6 5
8. Проведення розрахунків безпеки життєдіяльності	Програміст	3	3	100	3
9. Проведення економічних розрахунків	Програміст	3	3	100	3
10. Оформлення пояснювальної записки	Керівник, Програміст	20	8 20	30 100	2.4 20
Разом за проектом	Керівник, Програміст	97	35 97		9.45 97

Використовуючи дані таблиці 5.2, було розроблено календарний графік виконання робіт за проектом, який представлений в таблиці 5.3.

Таблиця 5.3- Календарний графік виконання робіт проекту

№	Зміст роботи	Виконавець	Тривалість, день	Графік роботи	
				Початок	Кінець
1	Аналіз предметної області	Керівник	3	17.09.2022	19.09.2022
		Програміст	7	17.09.2022	23.09.2022
2	Постановка задачі дослідження	Керівник	2	24.09.2022	25.09.2022
		Програміст	5	24.09.2022	28.09.2022

Продовження таблиці 5.3

№	Зміст роботи	Виконавець	Тривалість, день	Графік роботи	
				Початок	Кінець
3	Розробка алгоритмічного та математичного забезпечення	Керівник	3	01.10.2022	02.10.2022
		Програміст	7	01.10.2022	06.10.2022
4	Розробка вимог до ПЗ	Керівник	3	08.10.2022	10.10.2022
		Програміст	5	08.10.2022	12.10.2022
5	Розробка моделей для опису ПЗ	Керівник	3	13.10.2022	15.10.2022
		Програміст	7	13.10.2022	19.10.2022
6	Розробка ПЗ	Керівник	5	20.10.2022	24.10.2022
		Програміст	35	20.10.2022	23.11.2022
7	Опис застосування ПЗ	Керівник	2	24.11.2022	25.11.2022
		Програміст	5	24.11.2022	28.11.2022
8	Проведення розрахунків безпеки життєдіяльності	Програміст	3	29.11.2022	01.12.2022
9	Проведення економічних розрахунків	Програміст	3	02.12.2022	04.12.2022
10	Оформлення пояснювальної записки	Керівник	8	05.12.2022	12.12.2022
		Програміст	20	05.12.2022	24.12.2022

Таким чином були виконані планування комплексу робіт з розробки ПЗ та оцінка трудомісткості витрат.

5.4 Розрахунок проектних витрат на розробку ПЗ

Виконаємо розрахунки витрат проекту щодо розробки ПЗ.

1 Основна заробітна плата виконавців проекту представлена в таблиці 5.4. Визначення витрат часу на розробку ПЗ кожним виконавцем відбувається виходячи з його завантаження за календарним графіком (табл. 5.2). Наведена основна заробітна плата виконавців проекту представлена з урахуванням того, що в місяці 21 робочий день.

Таблиця 5.4- Основна заробітна плата виконавців проекту

Посада	Посадовий оклад, грн	Середня денна ставка C_{Ti} , грн	Витрати часу на розробку t_i , людино-дні	Витрати на основну заробітну плату C_{Hi} , грн
Керівник	9000	9000/21=428.6	9.45	4050
Програміст	2000	2000/21=95.2	97	9234
Усього, C_H				13284

2 Витрати на використання машинного часу M_T . Виходячи з того, що ПЗ розробляється на комп'ютерах, до сумарних витрат на розробку мають бути додані витрати на використання машинного часу:

$$M_T = t_{M_T} \cdot S_{M_T} \cdot K_M = 776 \cdot 5 \cdot 1 = 3880 \text{ грн.},$$

де t_{M_T} – машинний час комп'ютера, необхідний для розробки програмного

продукту; $t_{M_T} = 97 \cdot 8 = 776$ годин;

S_{M_T} – вартість 1 год машинного часу; $S_{M_T} = 5$ грн. / год.;

K_M – коефіцієнт мультипрограмності, який показує частку машинного часу, яка надається для реалізації проекту; у випадку виконання ДР $K_M = 1$.

3 Витрати на матеріали C_M . В даному випадку у процесі виконання проекту був придбаний зошит загальний, компакт-диск CD-RW, тонер для принтера та офісний папір А4. Всі придбані матеріали та їх вартість приведені в таблиці 5.5.

Таблиця 5.5- Витрати на матеріали

Матеріал	Од. вимірювання	Необхідна кількість	Ціна за од., грн.	Сума, грн.
Зошит загальний	шт.	1	60	60
Компакт-диск CD-RW	шт.	2	12	24
Тонер для принтера	шт.	1	115	115

Продовження таблиці 5.5

Матеріал	Од. вимірювання	Необхідна кількість	Ціна за од., грн.	Сума, грн.
Офісний папір А4	Пачка	1	110	110
Усього, C_M				309

Отже, капітальні вкладення проекту дорівнюють:

$$K_p = ((1 + W_d)(1 + W_c) + W_H) \sum_{i=1}^m C_{Hi} + C_M + M_T =$$

$$= ((1 + 0.5)(1 + 0.22) + 0.6) \cdot 13284 + 309 + 3880 = 36469.12 \text{ грн.}$$

де m – кількість виконавців проекту, які беруть участь у розробці ПЗ;

C_{Hi} – витрати на основну заробітну плату виконавці i -ї категорії, грн;

W_d – коефіцієнт, який враховує додаткову заробітну плату в частках до основної заробітної плати (встановлюється кожною організацією окремо, наприклад, це може бути кваліфікаційні надбавки за стаж, наукове звання, науковий ступінь тощо);

W_c – коефіцієнт, який враховує єдиний соціальний внесок (податок з фізичної особи, та інші податки та збори), в частках до суми основної та додаткової заробітної плати виконавця проекту ($W_c = 22\%$);

W_H – коефіцієнт, який враховує накладні витрати організації, у частках до основної заробітної плати виконавця проекту (приймається за фактичними даними, $W_H = 0.6$);

C_M – витрати на матеріали;

M_T – витрати на використання машинного часу.

Кошторис – перелік, в якому відображаються статті витрат проекту на розробку програмного забезпечення, наведено у таблиці 5.6.

Таблиця 5.6 – Кошторис на розробку ПЗ

Статті витрат	Сума, грн.
Основна заробітна плата	13284
Додаткова заробітна плата	6642
Відрахування	4383.72
Витрати на матеріали	309
Витрати на машинний час	3880
Накладні витрати організації	7970.4
Всього	36469.12

В результаті, на розробку ПЗ витрати K_P склали 36469.12 грн.

5.5 Розрахунок витрат на впровадження ПЗ

В даному випадку для розрахунку капітальних вкладень проекту на впровадження ПЗ потрібно врахувати витрати на основне і допоміжне обладнання (K_{AE}), а також витрати на придбання типових розробок (K_{PP}).

Так як розроблене ПЗ є веб-застосунком, то у підприємства немає необхідності в закупівлі додаткових ПК і програмного забезпечення, отже, ці витрати будуть дорівнювати нулю. Потрібно тільки здійснити оплату віртуальної машини на віддаленому сервері від DigitalOcean вартістю 3313 грн. (12 міс.) [21] і доменного імені на GoDaddy вартістю 300 грн [22]. Отже, необхідні кошти на:

Для даного проекту витрати на його реалізацію складуть:

$$K_R = K_{AE} + K_{PP} = 0 + (3313 + 300) = 3613 \text{ грн.}$$

Отже, сумарні витрати на реалізацію проекту щодо розробки ПЗ складуть:

$$K_1 = K_P + K_R = 36469 + 3613 = 40082 \text{ грн.}$$

Таким чином був виконаний розрахунок витрат на впровадження ПЗ. В результаті знайдено усі дані, які доводять технічну доцільність і економічну ефективність розробленого ПЗ.

ВИСНОВКИ

Головною задачею даної роботи була розробка методів оцінювання ІТ-компаній з використанням анонімних відгуків з метою пришвидшення процесу пошуку підходящої компанії.

Спочатку був проведений аналіз предметної області. Під час аналізу було:

- встановлено актуальність проблеми;
- розглянуто існуючі дослідження в предметній області;
- розглянуто існуючі продукти на ринку;
- здійснено ознайомлення з принципами систем рекомендацій;
- доведено доцільність створення платформи анонімних відгуків на ІТ-компанії.

На основі отриманих під час аналізу предметної області знань було розроблено метод оцінювання ІТ-компаній з використанням ознак для їх характеристики. Був розроблений математичний алгоритм для формування списку рекомендованих відгуків. У рамках розробки платформи анонімних відгуків були розроблені:

- функціональні та нефункціональні вимоги до ПЗ;
- структура бази даних;
- архітектура системи;
- база даних, серверна та клієнтська частина ПЗ.

В результаті виконання роботи було проведено дослідження процесу підтримки створення анонімних відгуків на ІТ-компанії з використанням з використанням інтелектуальних систем, а також розроблений метод оцінювання ІТ-компаній з використанням анонімних відгуків і відповідна web-платформа.

СПИСОК ДЖЕРЕЛ ІНФОРМАЦІЇ

- 1 Employer Review Platforms – Do the Rating Environment and Platform Design affect the Informativeness of Reviews? Theory, Evidence, and Suggestions / Janis Cloos, 2021, 152 – 181
- 2 Спільнота програмістів – DOU. URL <https://dou.ua/> (дата звернення 20.10.2022)
- 3 Glasdoor. URL <https://www.glassdoor.com/> (дата звернення 20.10.2022)
- 4 Kununu. URL <https://www.kununu.com/> (дата звернення 20.10.2022)
- 5 IT Ukraine Association, better Regulation Delivery Office / Development of Ukrainian IT Industry - Analytical Report. Kyiv, October, 2018
- 6 IT Research 3.0. Lviv IT Cluster. URL https://itcluster.lviv.ua/wp-content/uploads/2018/12/IT_RESEARCH_final_English.pdf (дата звернення 20.10.2022)
- 7 What drives job satisfaction in IT companies? / R. F. Ramos., S. Moro, P. Rita. International Journal of Productivity and Performance Management, March 2020.
- 8 Factors determining company performance in the IT industry / Miruna Florina Lungu. Management & Marketing, 2020 - p. 59-77.
- 9 Recommender Systems Handbook / [Ricci F., Rokach L., Shapira B., Paul B. Kantor]. Springer New York Dordrecht Heidelberg London, 2011 – 842 p.
- 10 Employment Recommendation System using Matching, Collaborative Filtering and Content Based Recommendation / Roshan G. Belsare, Dr. V. M. Deshmukh, 2018, p 215-220
- 11 Employment Recommendation System: A Review / Roshan G. Belsare, Dr. V. M. Deshmukh, 2018, p 64-67
- 12 Job Recommendation System Using Machine Learning And Natural Language Processing / Jeevankrishna, 2020
- 13 Sørensen T. A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and its Application to Analyses of the Vegetation on Danish Commons / T. Sørensen. Kobenhavn, 1948. –41 p.

- 14 Vijaymeena M.K. A Survey on Similarity Measures in Text Mining / M. K. Vijaymeena, K. Kavitha // Machine Learning and Applications: An International Journal (MLAIJ). – 2016. vol. 3, no. 1. – P. 19-28.
- 15 On the Distortion of Locality Sensitive Hashing / F. Chierichetti, R. Kumar, A. Panconesi, E. Terolli // SIAM Journal on Computing. – 2019. vol. 48, no. 2. – P. 350-372.
- 16 Unhelkar B. Software Engineering with UML / Bhuvan Unhelkar. CRC Press, Taylor & Francis Group, 2018 – 389 p.
- 17 JSON Web Tokens. URL <https://jwt.io/> (дата звернення 18.11.2022)
- 18 Flanagan D. JavaScript: The Definitive Guide: Master the World's Most-Used Programming Language / D. Flanagan. – 7th edition. O'Reilly Media. 2020. – 706 p.
- 19 Padolsey J. Clean Code in JavaScript / J. Padolsey. Packt Publishing Ltd, 2020. –527 p.
- 20 Casciaro M. Node.js Design Patterns: Design and implement production-grade Node.js applications using proven patterns and techniques / M. Casciaro, L. Mammino. – 3rd Edition. Packt Publishing Ltd. Birmingham, UK, 2020. – 630 p.
- 21 Dmytro Orlovskiy, Andrii Kopp, Ivan Bilous "Development of Adaptive Parametric Interface for Analytical Queries: Electrical Equipment Management System Case Study." Annual Scientific Conference «Information Technology and Automation — 2020» Conference proceeding (2020): 232-235.
- 22 Digital Ocean - Pricing. URL <https://www.digitalocean.com/pricing> (дата звернення 18.11.2022)
- 23 Пошук доменного імені GoDaddy. URL <https://ua.godaddy.com/domains/domain-name-search> (дата звернення 18.11.2022)
- 24 Анонімні відгуки на українські ІТ-компанії. URL <https://vidhook.io> (дата звернення 18.12.2022)
- 25 Google аналітика. URL <https://analytics.google.com> (дата звернення 18.12.2022)