

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

Факультет Комп'ютерних наук
Кафедра Програмної інженерії

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

другий (магістерський)
(рівень вищої освіти)

Дослідження методів виявлення помилок проектування сайтів

Виконав:

Студент 2 курсу групи ПЗЗдм-20-1

Лозиченко А.В.

(прізвище, ініціали)

Спеціальність 121 – Інженерія програмного
забезпечення

Тип програми Освітньо-наукова

Керівник доц. Мельнікова Р.В.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. Кафедри

проф. Дудар З.В.

2022 р.

Харківській національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____
Кафедра _____ Програмної інженерії _____
Рівень вищої освіти _____ другий (магістерський) _____
Спеціальність _____ 121 – Інженерія програмного забезпечення _____
Тип програми _____ Освітньо-наукова програма _____
Освітня програма _____ Інженерія програмного забезпечення _____

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«__» _____ 202_ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Студентки _____ Лозиченко Андрія Володимировича _____
(прізвище, ім'я, по батькові)

1. Тема роботи: Дослідження методів виявлення помилок проектування сайтів.
затверджена наказом університету від «24» __03__ 2022 р. № 31 Стз
2. Термін подання студентом роботи до екзаменаційної комісії «22» __05__ 2022 р.
3. Вихідні дані до роботи: Технічне завдання, календарний план, методичні вказівки, алгоритми категоризації
4. Перелік питань, що потрібно опрацювати в роботі: аналіз предметної галузі, аналіз існуючих алгоритмів, опис розробки системи, опис досліджень.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної галузі та постановка задачі	31.03.2022	виконано
2	Створення алгоритму	07.04.2022	виконано
3	Проектування системи	14.04.2022	виконано
4	Реалізація системи	26.04.2022	виконано
5	Підготовка пояснювальної записки, презентації та доповіді	02.05.2022	виконано
6	Перевірка на плагіат, нормоконтроль	16.05.2022	виконано
7	Рецензування, занесення матеріалів у електронний архів	20.05.2022	виконано
8	Попередній захист	22.05.2022	виконано
9	Допуск до захисту у зав. кафедри	24.05.2022	виконано

Дата видачі завдання 17.01.2022 р.

Керівник доц. _____ (Мельнікова Р.В.)

Завдання прийняв до виконання _____ (Лозиченко А.В.)

РЕФЕРАТ / ABSTRACT

Кваліфікаційна робота магістра містить: 62 с., 30 рис., 4 додат., 11 джер.

БІЗНЕС, СИСТЕМА, ВЕБ-ДИЗАЙН, ОНОВЛЕННЯ ДИЗАЙНУ, СПОСТЕРЕЖЕННЯ, АНАЛІЗ ПОВЕДІНКИ, АНАЛІЗ СТАТИСТИЧНИХ ДАНИХ.

Об'єктом роботи є дослідження для впровадження системи виявлення невдалих дизайнерських рішень.

Метою роботи є проведення дослідження та проектування моделі для системи виявлення невдалих дизайнерських рішень за допомогою збору статистичних даних та аналізу поведінки користувача. Результатом дослідження має бути систематичний підхід до визначення невдалих рішень у дизайні веб-сторінки, який складається з методу збору статистичних даних та методу їх структурування.

Під час дослідження було розглянуто існуючі можливі автоматизовані та виконуванні спеціалістами рішення, які частково або повністю вирішували проблеми та мали різні рівні доступності та складності використання. Також проаналізовано сильні та слабкі сторони кожного з них з метою проектування найточнішого та найдоступнішого методу виявлення слабких місць у дизайні та оцінка можливості його автоматизації.

Основною метою є винайдення алгоритму та систематичного підходу, який дозволить пов'язати між собою статистичні дані спостережень, нетипову або неочікувану поведінку користувача та елементи дизайну веб-сторінки.

BUSINESS, SYSTEM, WEB DESIGN, DESIGN UPDATES, OBSERVATIONS, BEHAVIOR ANALYSIS, STATISTICAL ANALYSIS.

The object of the work is research to implement a system for detecting failed design decisions. The aim of the work is to conduct research and design a model for the system of detection of failed design decisions through the collection of statistics and analysis of user behavior. The result of the research should be a systematic approach to identifying

bad decisions in web design, which consists of a method of collecting statistics and the method of structuring them.

The study examined existing possible automated and specialist solutions that partially or completely solved problems and had different levels of accessibility and complexity of use. The strengths and weaknesses of each of them are also analyzed in order to design the most accurate and accessible method of identifying weaknesses in the design and assess the possibility of its automation.

The main goal is to invent an algorithm and a systematic approach that will link the statistics of observations, atypical or unexpected user behavior and web design elements.

Умови публікації пояснювальної записки

Я,

_____ Лозиченко Андрій Володимирович _____
(прізвище, ім'я, по батькові)

студент(ка) групи _____ здобувач вищої освіти на другому (магістерському) рівні

кафедра _____ програмної інженерії _____,
(повна назва кафедри)

заявляю: моя кваліфікаційна робота на тему

_____ Дослідження методів виявлення помилок проектування сайтів _____,
(назва роботи)

що буде представлена до ЕК для публічного захисту, виконана самостійно, в ній не містяться елементи плагіату і вона може бути опублікована в електронному архіві відкритого доступу EIArKhNURE. Всі запозичення з друкованих та електронних джерел мають відповідні посилання.

Я ознайомлений (а) з діючим положенням «Про протидію академічному плагіату в ХНУРЕ», згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування дисциплінарних заходів.

ЗМІСТ

ВСТУП	7
1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ	9
1.1 Загальний аналіз галузі	9
1.2 Виявлення проблем	13
1.3 Постановка задачі	15
2 АНАЛІЗ ІСНУЮЧИХ АЛГОРИТМІВ	17
2.1 Опис існуючих методів	17
2.2 Вибір алгоритму	26
3 ОПИС РОЗРОБЛЕНОГО ПРОТОТИПУ ПРОГРАМНОЇ СИСТЕМИ	30
3.1 Налаштування сервісу спостереження	30
3.2 Реалізація та налаштування карти Кохонена	31
3.3 Реалізація алгоритму виявлення помилок	35
3.4 Архітектура системи	38
4 ОПИС ДОСЛІДЖЕНЬ	41
ВИСНОВКИ	49
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	51
ДОДАТОК А	52
ДОДАТОК Б	53
ДОДАТОК В	54
ДОДАТОК Г	55

ВСТУП

В даний час Інтернет використовують більше 2,5 млрд осіб і ця кількість постійно зростає. До всесвітньої мережі вже підключено більшість розвинених держав, а можливості супутникового зв'язку показують, що у найближчому майбутньому інтернет буде проведений навіть у важкодоступні регіони.

У світі, де глобалізація виходить на перші позиції в економічних, соціальних і політичних відносинах, неможливо уявити собі життя без цього універсального засобу зв'язку.

З плином часу все більше закладів з надання послуг переходять до роботи у режимі онлайн, особливо поширеною стала ця тенденція підчас введення карантинних обмежень через CovId-19. Також зі стрімким переходом сервісів до Інтернет, пришвидшився й розвиток веб-технологій, що відображається у функціонування сучасних веб-сайтів, не відстають і технології дизайну, які з роками стають все важливішими для користувача.

Заходячи на сайт, користувач відразу ж звертає увагу на його графічне оформлення і зручність навігації. У більшості випадків, від того враження, яке справляє зовнішній вигляд вашого сайту, залежить рішення відвідувача: залишитися на сторінці або продовжити пошуки на інших ресурсах.

Дизайн сайту - це не просто його зовнішнє оформлення, він також вирішує такі важливі завдання, як функціональність, зручність у використанні, те що на мові веб-майстрів називається юзабіліті. Відвідувач сайту повинен легко орієнтуватися на його сторінках, знаходити потрібну інформацію, мати можливість швидко повернутися на головну сторінку, а головне – не задаватися зайвими запитаннями.

Тенденції веб-дизайну змінюються досить часто. Якщо ваш сайт не оновлювався протягом довгого часу, зараз він, ймовірно, виглядає застарілим та зовсім непрофесійним. Це прирівнює ваші шанси залучити нових клієнтів до нуля. Ніхто не довірятиме вам, якщо на вашому сайті незручна навігація або важко знайти будь-яку важливу інформацію, або ж сайт просто жахливо виглядає візуально. Навпаки, він має створювати довіру до вашої компанії та представляти ваш бренд найбільш професійним способом.

Є чимало причин, по яким користувач може передумати користуватись веб-сайтом, більшість з яких пов'язані з дизайном.

Для вирішення цієї проблеми, а саме – виявлення проблемних місць дизайну веб-сторінки, й проводилось дослідження. Під час дослідження було розглянуто існуючі можливі автоматизовані та виконуванні спеціалістами рішення, які частково або повністю вирішували проблеми та мали різні рівні доступності та складності використання. Також проаналізовано сильні та слабкі сторони кожного з них з метою проектування найточнішого та найдоступнішого методу виявлення слабких місць у дизайні та оцінка можливості його автоматизації.

Проект з розробленим алгоритмом дозволить, з одного боку, поліпшити економічний стан підприємців та власників сайтів, з іншого – покращить досвід використання, дозволить не відмовлятися від сервісів та зробить взаємодію з сайтом приємнішою для користувачів Інтернет. Спроектвана система дозволить дизайнерам зосередитись на вирішенні проблеми та не витратити забагато часу та коштів на збори статистичних даних, проведення спостережень та аналізу поведінки користувачів та пов'язування її з причинами.

Під час проектування, основною метою є винайдення алгоритму та систематичного підходу, який дозволить пов'язати між собою статистичні дані спостережень, нетипову або неочікувану поведінку користувача та елементи дизайну веб-сторінки.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

1.1 Загальний аналіз галузі

На сучасному рівні розвитку інформаційних технологій використання комп'ютера для збереження будь-яких видів інформації стає єдиним засобом, що надає широкі можливості керування інформацією. Важливу роль у процесі отримання інформації відіграє мережа Інтернет.

Internet сьогодні це найбільш розвинена у світі інформаційна система, за допомогою якої здійснюється комунікація між мільйонами користувачами. За допомогою мережі Internet забезпечується доступ до провайдерів послуг багатьох спектрів, наприклад:

- робота з документами на державному рівні;
- замовлення доставки побутових приладів та їжі;
- організація зустрічей різного формату;
- навчання різних рівнів та напрямів;
- комунікація та розваги;
- ведення бізнесу та операції з грошима;
- онлайн-шопінг та ін.

Всього цього стає забагато, а значить у власників бізнесу у Інтернет збільшується кількість конкурентів, а у користувачів, з свою чергу, стає більший вибір. І тому у підприємців першостепенної важливості у боротьбі з конкурентами виходить те, що здібне захопити користувача з першої ж митті – дизайн веб-сторінки.

Абсолютно природньо, що людську увагу спершу приваблює візуальна складова, а потім вони переходять (або не переходять) до читання контенту. Результати багатьох досліджень користувачів показують, що відвідувачі формують думку про ваш бізнес на сайті менш ніж за дві десятих секунди. Допомагає їм у цьому візуальний контекст, який формує гарне враження і затримує користувачів на сайті довше. Картинка впливає на користувачів сильніше текстового контенту, так як з її допомогою можна більш ефективно і швидше передати повідомлення.

На словах важко пояснити читачеві, що письменник хоче передати. Те ж саме і у веб-дизайні. Сайту необхідний візуальний контекст, який зв'яже життєвий досвід і історії організації.

Візуальний контекст пов'язаний з поданням контенту таким чином, щоб краще донести основну ідею сайту.

Наприклад, на сайті архітектора може бути портфоліо з успішними проектами (див. рис. 1.1). Користувач сам вирішить, чи потрібен йому цей розділ чи ні. Візуальний контекст з самого початку (зображення, кольору, відступи) дасть користувачеві розуміння про знаннях архітектора в дизайні, досвід в будівництві та адміністративної експертизі.

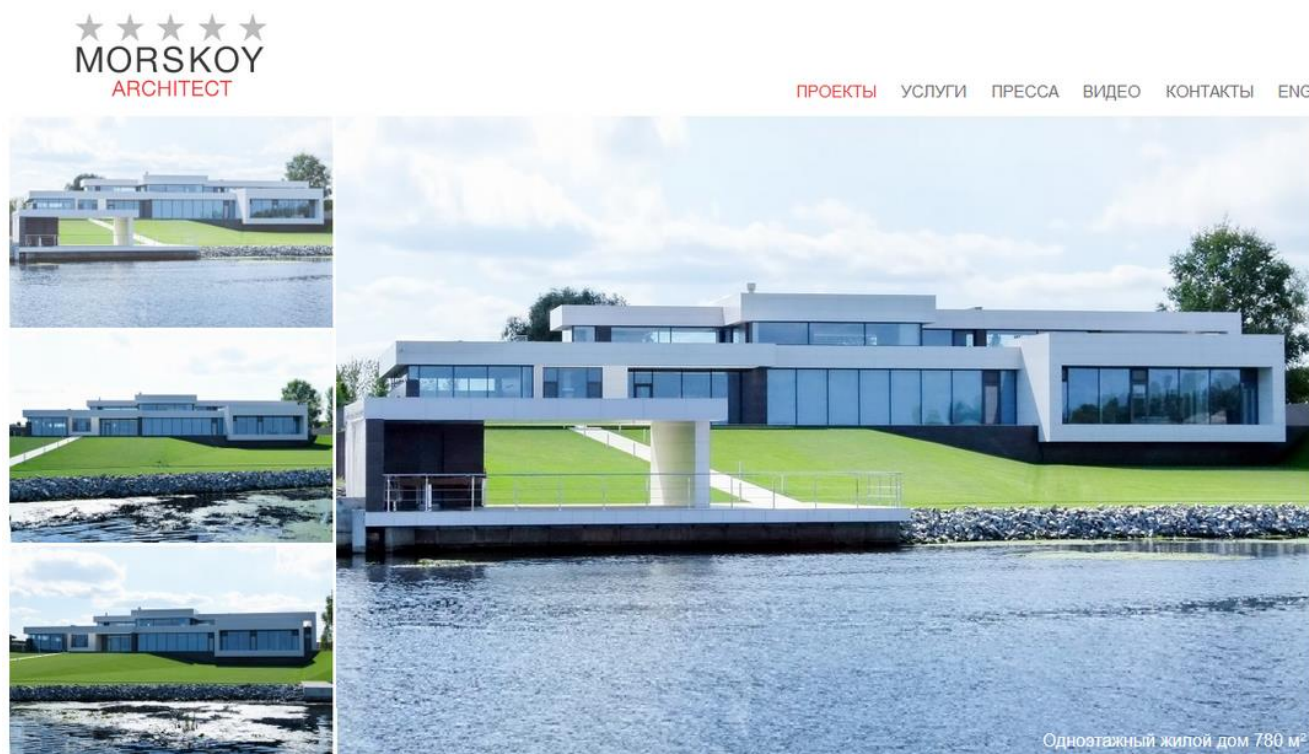


Рисунок 1.1 – Портфоліо архітектора

Розглянемо ще один приклад з компанією з виробництва аерокосмічної техніки (див. рис. 1.2). Смілива картинка літака і техніки відразу дає зрозуміти про стан цієї компанії в авіаційно-космічній промисловості. Користувачеві навіть не потрібно читати текст на банері.

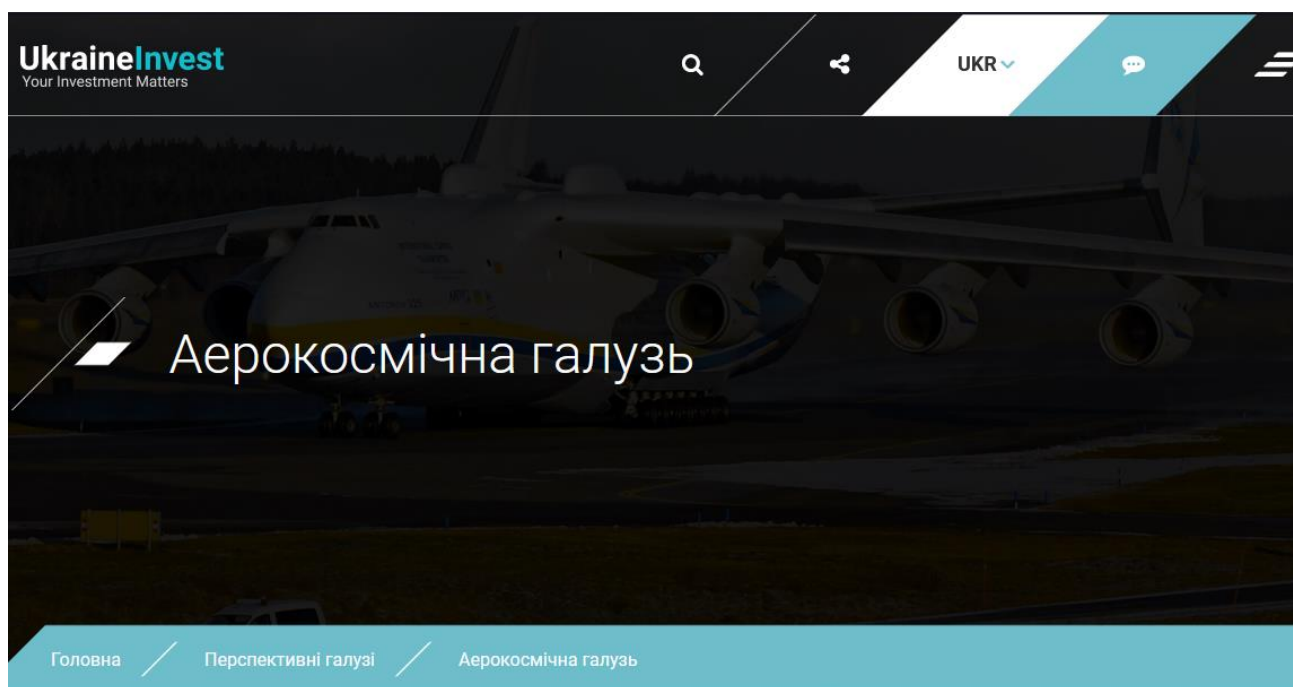


Рисунок 1.2 – Сайт аерокосмічної техніки

Створення візуального контексту залежить від розуміння ваших користувачів, перспективи зміни користувачів з віком, статтю та етнічною приналежністю. Наприклад, зображення з білим кольором можуть викликати щастя у одних людей і смуток у інших. Тому для створення ідеального візуального контексту на вашому бізнес сайті необхідно оцінити розуміння вашої аудиторії, того, з чим вам може допомогти користувач.

Розглянемо декілька причин значущості візуального контексту:

- перше враження - вище вже сказали, що користувачам для формування думки про вашому сайті потрібні частки секунди. Більш повна думка складається за наступні кілька секунд. Бренд, зображення, іконки повинні створити хороше перше враження за короткий проміжок часу. Щоб користувачі залишилися на сайті довше, тому важливо переконатися, що візуальний контекст дає інформативну картинку того, про що ваш сайт і для чого;

- допомагає з донесенням інформації - візуальна складова сайту «чуттєво» привабливіший, ніж текстовий контент. Користувач, швидше за все, зверне увагу на ще декілька зображень, що підтверджують загальний контекст, а не

на фрази та параграфи тексту. Неважливо, наскільки добре автор передає історію, перший емоційний відгук від користувача сильно залежить від зображень, що допомагає в передачі інформації;

- користувачі набагато краще розуміють візуальний контекст - візуальний контекст оживляє історію, він допомагає користувачеві зрозуміти ваш бізнес і ваші операції. Він або вона зрозуміють базову ідею вашого бізнесу з контексту і почнуть шукати опції, товари і послуги.

Візуальний контекст допомагає користувачам, особливо тим, хто на вашому сайті перший раз. Він інтерпретує характеристики і значення з сайту. Як дизайнер, ви можете використовувати все вищеописане для створення важливого контексту (зображення, кольори, візуальні дані, фото, відео, анімація, іконки, напрями та інші візуальні інструменти).

Розробка дизайну веб-сторінок є досить складним процесом – необхідно проаналізувати потенційних користувачів, мету та призначення створюваного сайту, дослідити найсучасніші технології та тенденції створення дизайну, а також існує велика сукупність правил, яку треба брати до уваги. Все це потребує не аби яких досвіду та знань, якими володіють далеко не всі, тому на даний момент існує дуже багато незручних або просто непривабливих сайтів.

Для вирішення цієї проблеми пропонується декілька варіантів – з самого початку створення сайту найняти професіоналів та доручити цю справу їм, або коли сайт вже створено, знову ж таки, найняти професіоналів справи для проведення аналізу та робіт з поліпшення дизайну [1].

Всі перелічені вище варіанти коштують досить дорого, а з безкоштовних рішень – скористатися застосунками для збору статистичних даних, по яких можна лише дізнатися дії користувача та їй потребує знань та досвіду для самостійного діагностування причини незадовільної статистики взаємодії з сайтом. Також можна скористатися зворотнім зв'язком від користувачів сайту, але далеко не всі користувачі мають бажання витратити на це час та іноді їх відгуки можуть бути занадто суб'єктивними або спровоковані іншими, не пов'язаними з дизайном,

факторами, наприклад, негативний відгук про не зручний дизайн був написаний лише тому, що користувач не отримав очікуваного результату.

1.2 Виявлення проблем

Під час створення веб-сторінки для переносу свого бізнесу до Інтернет, можна допустити чимало помилок дизайну веб-сторінки, починаючи ще з аналізу цільового користувача та етапі проектування та планування архітектури, закінчуючи вибором технологій та шаблонів сайту.

Розглянемо детальніше існуючі проблеми:

- ігнорування потреб цільової аудиторії - в ідеалі, потреби, бачення та попередній досвід цільової аудиторії варто враховувати ще на етапі проектування та розробки веб-ресурсу. Необхідно добре розуміти чого саме хоче користувач, з якою метою відвідує сайт та чому може його завчасно покинути. Саме тому професійні дизайн-агентства часто створюють «портрет» потенційного відвідувача. Хорошою ідеєю є перевірка готового концепту на певній «фокус-групі» ще до етапу розробки. Все це необхідно передбачати оскільки бачення одних і тих же речей може сильно відрізнятись у різних людей. Розробка веб-продукту «на власний смак» збільшує ризики того, що користувач залишиться незадоволений;

- автозапуск відео/аудіо – багато підприємців, отримавши пораду від маркетологів, бажають вразити користувача цікавим відео або оригінальним аудіо-привітанням. В більшості випадків краще, аби користувач сам захотів відтворити ті чи інші файли;

- надлишковість - надлишковість є прямою протилежністю одного з основних принципів правильного дизайну – простоти. Ідея розмістити на одній сторінці максимум корисної інформації, сама по собі, не несе в собі нічого поганого. З іншого боку є два нюанси на які варто звертати увагу. Перш за все, це має бути справді потрібна інформація (до якої явно не належать яскраві, миготливі, анімовані банери або гіфки). Другий нюанс полягає в тому, що реалізація ідеї, сама по собі, є доволі складною та вимагає хороших навичок дизайнера. В

протилежному випадку є загроза отримати результат, наведений на скріншоті (див. рис. 1.3);



Рисунок 1.3 – Надлишковий дизайн

- неправильні форми - веб-форми використовуються дуже часто. Зворотній зв'язок, реєстрація, авторизація – це далеко не повний перелік їх можливого застосування. Проте навіть в такому простому елементі як форми можна допустити цілу низку помилок. Найбільш поширеними проблемами є надто велика кількість запитань, надто велика кількість обов'язкових полів та некоректно налаштована валідація;

- проблеми авторизації або виходу - ще одна поширена проблема дизайну, яку допускають власники веб-сайтів – непродумана реєстрація/авторизація, а також механізм виходу з сайту. Найчастіше це: відсутність входу через соцмережі, занадто велика кількість даних для реєстрації та складна процедура виходу з особистого кабінету;

- мобільні пристрої – користувачів, які користуються Інтернет з телефонів майже стільки ж як і тих, хто віддає перевагу ПК. Проте, як показує практика, частина власників сайтів досі не подбала про своїх мобільних користувачів;

- відсутність інтуїтивності - під це поняття підпадає простота, зручність та зрозумілість для більшості людей, які користуватимуться ресурсом. Спроба зробити крутий дизайн часто може перетворитись на пошук “золотої середини” між чимось простим та зрозумілим і вишуканим та креативним;

- колірна гама - вплив кольору на людину давно відомий та вдало використовується маркетологами останніх декілька десятиліть. Відповідно до досліджень, зроблених в США, при купівлі якогось з продуктів, 85% користувачів зупиняють свій вибір на товарі завдяки кольору;

- технічна оптимізація – використання занадто «важких» для обробки елементів дизайну може значно погіршити час обробки даних та взаємодії сайту. Та навпаки – занадто велике навантаження операціями може привести до спотворення елементів дизайну.

1.3 Постановка задачі

Як можна побачити, після проведення аналізу предметної області та виявлення в ній проблем, їх було виявлено чимало. Та допомогти вирішити їх всі лише одним підходом не можливо, бо всі вони пов’язані з різними сферами, такими як психологія, бізнес аналітика, проектування архітектури, маркетинг, аналіз перформансу та технологій.

Але наслідком багатьох з них є невдало створені окремі елементи дизайну, дослідження методів відстеження яких і стає задачею даної роботи. Необхідно розглянути існуючі автоматизовані та виконувані спеціалістами методи спостереження за поведінкою користувача на сайті, збір статистичних даних щодо виконаних взаємодій з елементами дизайну. Для цього треба ознайомитись з їх переліком, обрати різні за ступенем складності використання, доступності та точності і інформативності даних, що надається, навантаження на сайт та систему

у цілому через постійний збір даних. Також важливо визначити необхідність втручання спеціаліста у процес збору даних чи він потрібен лише для початково налаштування, або не потрібен взагалі. Оцінити достатність статистичних даних для виявлення «проблемних місць».

Майже по такими ж критеріям необхідно відібрати й оцінити методи аналізу отриманих статистичних даних, бо лише отримати результати спостережень за користувачами багатьом власникам бізнесу у Інтернет нічого не дасть і їм доведеться долучати спеціалістів або з самого початку наймати команду аналітиків, яка виконуватиме весь процес самостійно.

А отже результатом дослідження має бути систематичний підхід до визначення невдалих рішень у дизайні веб-сторінки, який складається з методу збору статистичних даних та методу їх структурування та аналізу і повинен надавати достатньо інформації для обмеження кола пошуку елементів дизайну та повторювані події, які з цим елементом виконувались [2].

2 АНАЛІЗ ІСНУЮЧИХ АЛГОРИТМІВ

2.1 Опис існуючих методів

Методи спостереження за діями користувачів на сайті та збір статистичних даних не обмежується лише програмним забезпеченням і не завжди потребує чималих технічних ресурсів.

Почнемо огляд з одного з таких методів – зворотній зв'язок.

Одним з найпростіших і найменш витратних способів дослідження є проведення онлайн-опитування, з використанням простих, безкоштовних програм, які можуть надати результати вже за кілька годин після початку опитування. Онлайн-опитування – це чудовий спосіб зібрати інформацію від користувачів послуг.

В залежності від обставин та засобів отримання відповіді може розділятися на підвиди:

- організація User Acceptance Testing (UAT). Тестування схвалювання користувачем визначає, чи «прийнятний» веб-продукт кінцевими користувачами або замовником, і чи працює він згідно з їхніми очікуваннями. В процесі UAT цільова аудиторія взаємодіє з сайтом, вказує на недоліки і розповідає, що варто виправити. Не йдеться про пошук багів, йдеться про перевірку того, чи виконує сайт свою місію так, як передбачалося;

- запит відгуків після кожної взаємодії з сервісом, який пропонує сайт. Полягає у надсиланні користувачам, за умови, що ми маємо їх контактні данні, листів до електронної пошти або повідомлення у профіль на сайті, з використанням внутрішніх засобів сервісу, запиту на отримання зворотнього зв'язку після кожного сеансу взаємодії з сайтом з певною затримкою між запитами;

- залишити форму для добровільного зворотнього зв'язку.

Розглянемо конкретні приклади сервісів, що дозволяють проводити опитування клієнтів.

Testograf – сервіс, призначений для створення професійних онлайн-опитування та тестів для всіх сфер бізнесу. Познайомитися з конструктором і

зробити опитування онлайн, включаючи налаштування логіки та дизайну, можна безкоштовно. Зібрати відповіді та переглянути результати можна лише після оплати, але на запит надається демо-доступ на кілька днів.

Кількість опитувань та відповідей не обмежена. Є готові шаблони і 31 тип питань: від різних матриць, рейтингів, текстових полів до списків, що випадають, і тестових типів питань (див. рис. 2.1.).

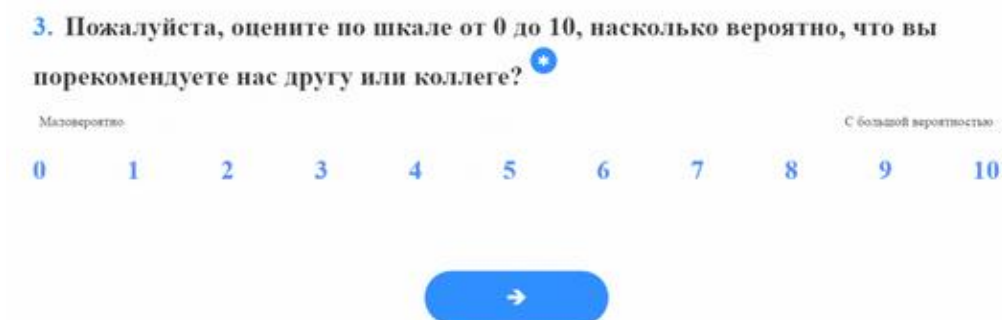


Рисунок 2.1 - Testograf

Також для порівняння розглянемо ще Google форми. Це абсолютно безкоштовний сервіс онлайн-опитування. Щоб почати користуватися ним, потрібен лише обліковий запис Google.

У Google Формах можна створювати необмежену кількість опитувань будь-якої довжини (див. рис. 2.2). Під час створення опитування у вашому розпорядженні 11 типів питань. Серед них: текстові поля, перемикачі та списки, що розкриваються. Є можливість додавати до питань фотографії чи відео з YouTube.

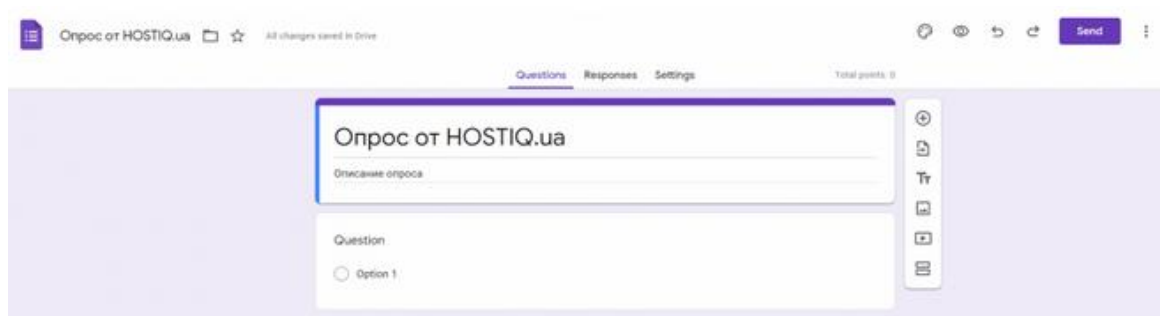


Рисунок 2.2. - Google форми

Інтернет (онлайн) опитування не може вважатися репрезентативним, його результати не є такими, на які можна покладатися, як на дані, що свідчать про реальний стан ситуації. Дані, отримані шляхом такого опитування не можуть слугувати кількісними показниками, що використовуватимуться як доказові. Отримана інформація є такою, що демонструє думки передусім користувачів Інтернету, це думки активних користувачів, які самі вирішили пройти це опитування, так званий самовідбір. Це відсутність гарантій, що відповіді надавалися без впливу інших осіб які могли бути поряд, тож відповіді, по суті, матимуть умовно «колективний характер» .

Тому: дані, які отримано шляхом онлайн опитування можна використовувати як допоміжні, корисні для аналізу поточної ситуації, орієнтирні в суперечливих моментах, підказові в непередбачуваних, невідомих невидимих ефектах чи діях.

Отримавши результати такого опитування, можна побачити умовно кількісну перевагу тих чи інших дій або вподобань, з можливим зв'язком цього з конкретними групами за статтю, проживанням, інтересами, віком чи іншими параметрами , - що може постати підказкою в портфель пропозицій подальших змін дизайну.

Наступний метод полягає у зборі статистичних даних і аналізу їх самотужки. Тому у якості джерела даних розглянемо наступні додатки [3].

Google Analytics - зручний і багатофункціональний сервіс від компанії Google для аналізу інтернет-сайтів та мобільних додатків. Дозволяє веб-майстрам перевірити стан індексування, створити детальну статистику аудиторії сайту та оптимізувати видимість своїх веб-сторінок.

Google Analytics використовується для відстеження активності користувачів на веб-сайті, такої як тривалість сеансу, сторінки за сеанс, показник відмов тощо, а також інформації про джерело трафіку (див. рис. 2.3).

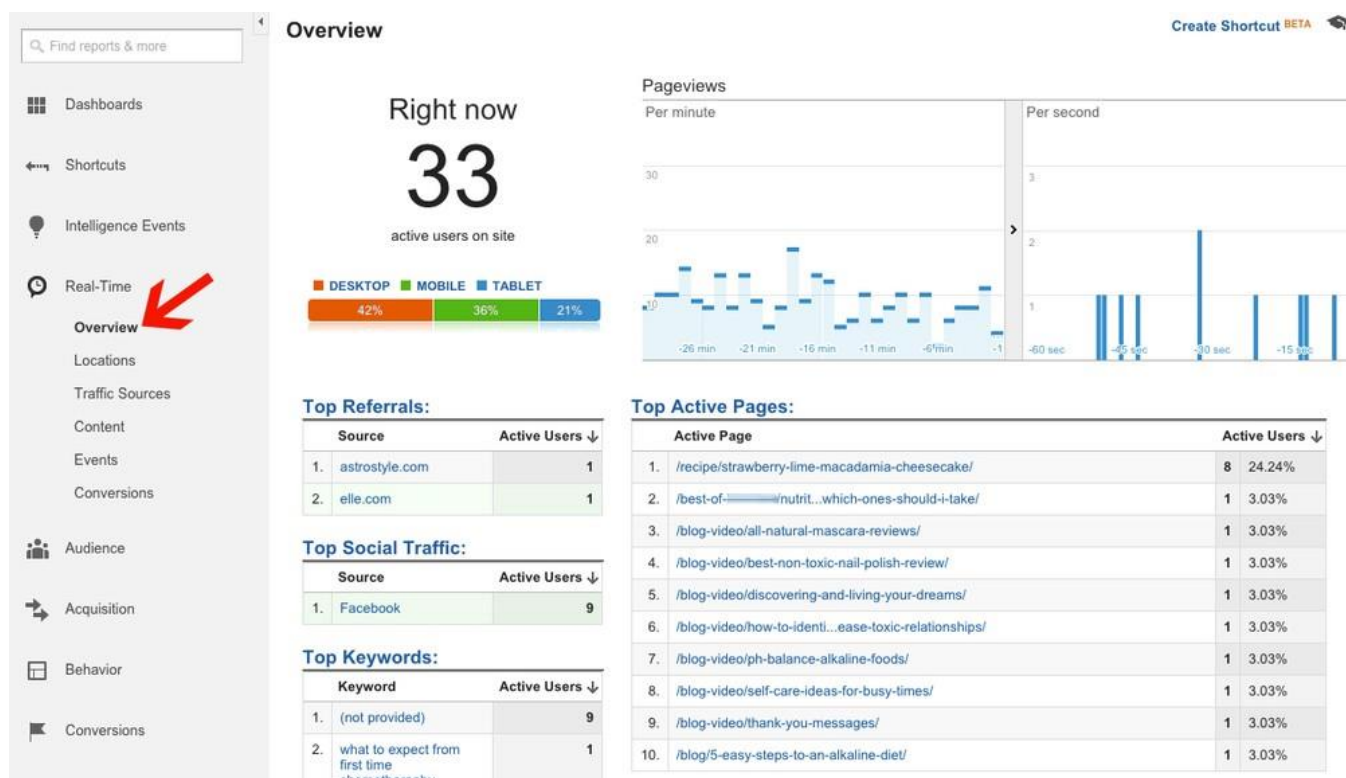


Рисунок 2.3 – Google Analytics

Доступна інтеграція з Google Ads, яка дозволяє веб-майстрам, відстежувати якість цільової сторінки та конверсією. Аналіз Google Analytics дозволяє виявити неефективні сторінки, враховуючи кількість сеансів, тривалість перебування користувача на веб-сторінці, вірогідність конверсії. У сервісі також доступна функція власної сегментації відвідувачів.

Clicky — це онлайн-сервіс, який дозволяє виконувати контроль, розширений аналіз трафіку та реагувати на його зміни, пропонуючи повноцінну статистику.

Базові функції Clicky:

- аналіз контенту, закладок, джерел та ін.;
- відстежувати індивідуальних користувачів та їх дії;
- дані у режимі реального часу;
- налаштуванні дашборди;
- аналіз версій сайтів, адаптованих під iPhone та Android;
- пропонує розширену сегментацію;
- відстеження зон активності;
- відео-аналітика;

- формує звіти;
- відстеження цілей користувачів;
- підтримка роботи із довгостроковими показниками;
- аналіз конверсії. (див. рис. 2.4)

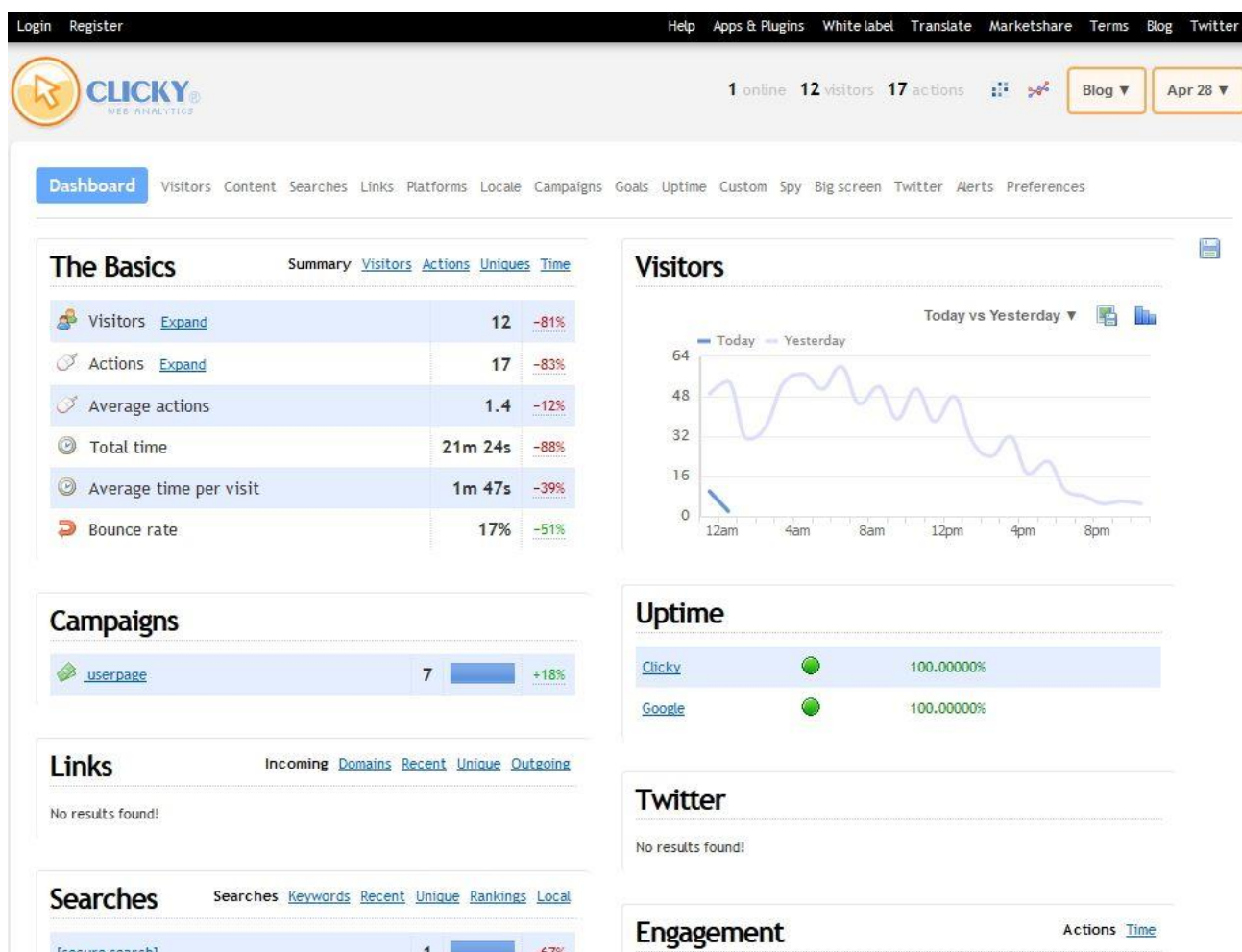


Рисунок 2.4 – Clicky

SimilarWeb — це веб-сайт, який надає послуги веб-аналітики для бізнесу. Компанія пропонує своїм клієнтам інформацію та статистику про обсяги трафіку своїх клієнтів та конкурентів; джерела трафіку, включаючи аналіз ключових слів, час перебування на сайті, перегляд сторінок, показник відмов та ін (див. рис. 2.5).

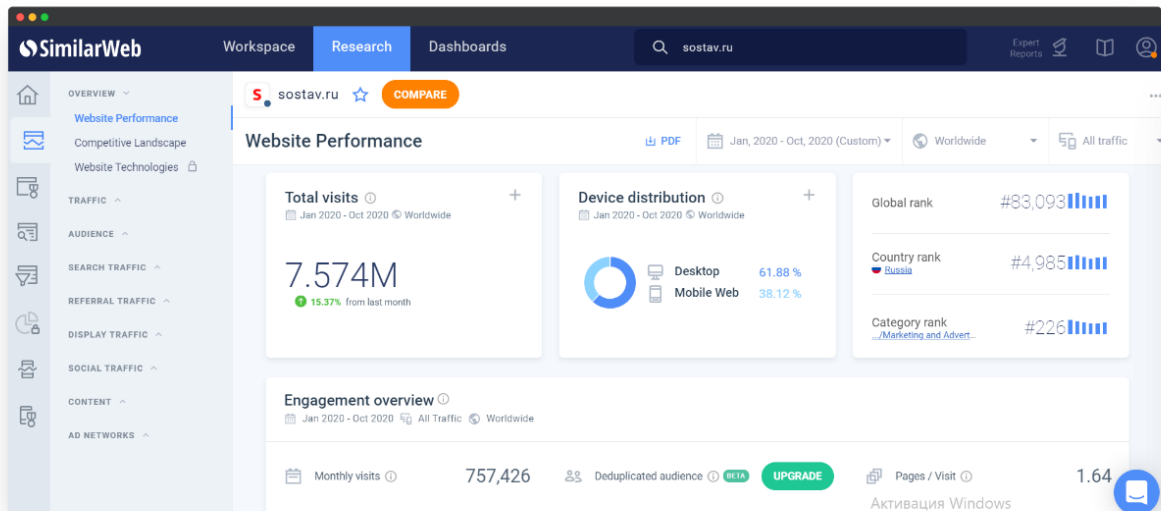


Рисунок 2.5 - SimilarWeb

Як і Alexa Internet, він екстраполює дані з панелі веб-користувачів, які дозволяють відстежувати їхню діяльність в Інтернеті, поєднуючись із прямими спостереженнями щодо підмножини Інтернет-ресурсів, таких як власна статистика відвідувань веб-сайтів.

Hotjar це потужний інструмент, який онлайн-користувачі можуть використовувати для оптимізації конверсії та зручності використання своїх веб-сайтів. Його найбільш затребувані функції включають конверсійні вирви, теплові карти і записи відвідувачів (див. рис. 2.6).

Він також об'єднує інструменти зворотного зв'язку та аналізу, щоб допомогти користувачам дізнатися, де відвідувачі клікають, скільки часу вони витрачають на кожну сторінку та їх шаблони прокручування. Все це допоможе вам визначити області, які потребують покращення.

З великою кількістю людей, які не наважуються ділитися особистими даними, Hotjar надає простий спосіб по-справжньому зрозуміти ваші мобільні та веб-сайти, не просячи їх заповнити форму.

Hotjar отримав свою популярність тому, що став першим інструментом аналізу, що пропонував записи відвідувачів. Він записує екран сеансів користувача,

щоб власник сайту міг бачити, як відвідувачі веб-сайту взаємодіють з різними елементами сторінки, наприклад, як вони прокручуються.

Власник онлайн-бізнесу отримує за лаштунками погляд на те, як відвідувачі веб-сайту споживають його контент. Це допоможе йому краще зрозуміти, які проблеми виникають на шляху покупки надаваних послуг.

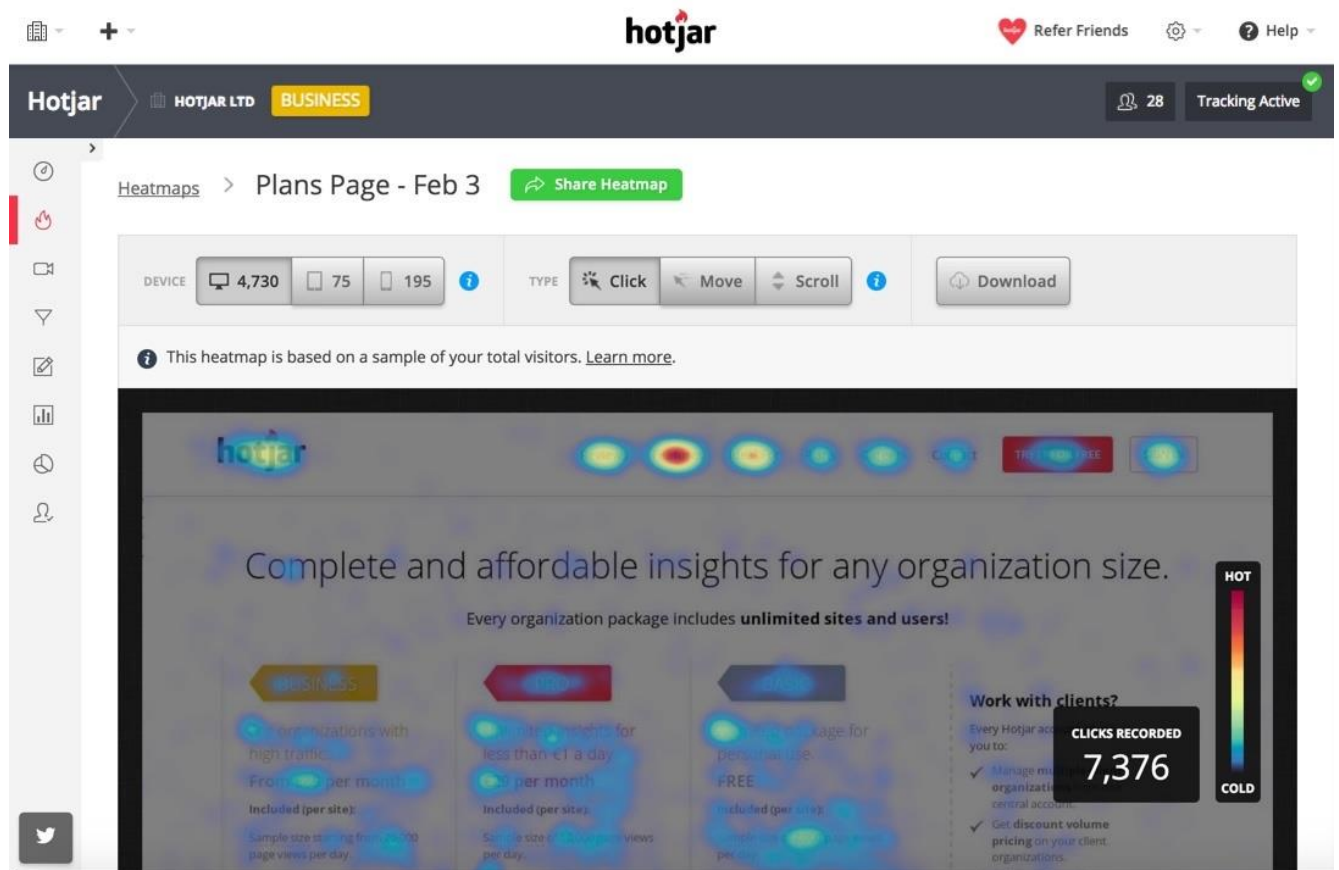


Рисунок 2.6 - Hotjar

Як можна побачити з описань, перелічені сервіси надають дуже багато інформації, але далеко не весь її обсяг може бути корисним для рішення досліджуваної проблеми. Тим паче, вони всі лише надають статистичні дані і результати спостережень, отримані дані необхідно ще певним чином обробити, проаналізувати та зробити висновки.

Тому далі розглянемо рішення, що зможуть вказати на конкретну причину відхилення від типової поведінки та, саме, надати інформацію у чому різниця була найдена – алгоритми аналізу даних [4].

Розглянемо представника нейронних мереж, що навчаються самі - самоорганізована карта Кохонена. Це змагальна нейронна мережа з навчанням без вчителя, виконує завдання візуалізації та кластеризації. Є методом проєкціювання багатовимірного простору в простір з більш низькою розмірністю (найчастіше, двовимірне), застосовується також для вирішення завдань моделювання, прогнозування та ін. Є однією з версій нейронних мереж Кохонена [5].

Самоорганізована карта складається з компонентів, які називаються вузлами або нейронами. Їх кількість задається аналітиком. Кожен з вузлів описується двома векторами. Перший – так званий вектор ваги m , що має таку ж розмірність, що і вхідні дані. Другий - вектор r , що представляє собою координати вузла на карті. Зазвичай вузли розташовують у вершинах регулярної решітки з квадратними або шестикутними осередками (див. рис. 2.7).

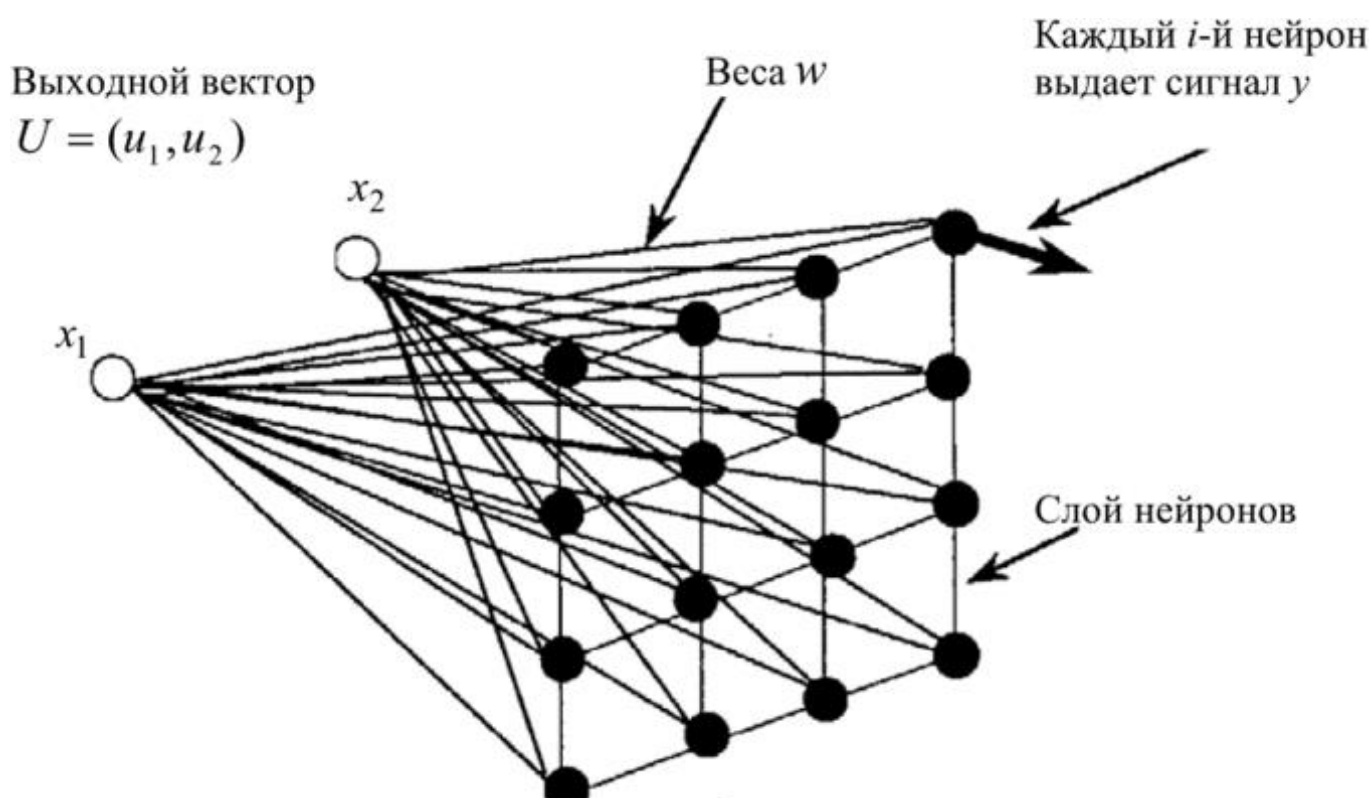


Рисунок 2.7 – Карта Кохонена

Спочатку відома розмірність вхідних даних, по ній деяким чином будується початковий варіант карти. У процесі навчання вектори ваги вузлів наближаються

до вхідних даних. Для кожного спостереження (семпла) вибирається найбільш схожий по вектору ваги вузол, і значення його вектора ваги наближається до спостереження. Також до спостереження наближаються вектори ваги декількох вузлів, розташованих поруч, таким чином якщо в множині вхідних даних два спостереження були схожі, на карті їм будуть відповідати близькі вузли. Циклічний процес навчання, перебираючий вхідні дані, закінчується по досягненні картою допустимої (заздалегідь заданою аналітиком) похибки, або по скоєнні заданої кількості ітерацій.

На противагу розглянутій нейронній мережі, що навчається сама, розглянемо ту, що потребує вчителя – Перцептрон. Це математична або комп'ютерна модель сприйняття інформації мозком (кібернетична модель мозку [6]). Незважаючи на свою простоту, перцептрон здатен навчатися і розв'язувати досить складні завдання. Основна математична задача, з якою він здатний впоратися — це лінійне розділення довільних нелінійних множин, так зване забезпечення лінійної сепарабельності (див. рис. 2.8).

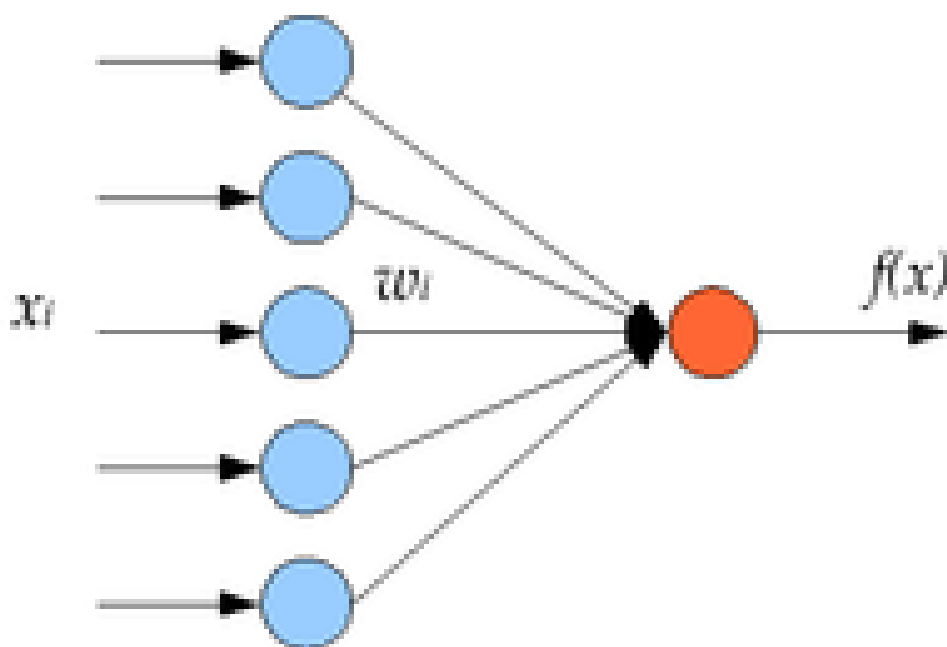


Рисунок 2.8 - Перцептрон

Перцептрон складається з трьох типів елементів, а саме: сигнали, що надходять від давачів, передаються до асоціативних елементів, а відтак до

реагуючих. Таким чином, перцептрони дозволяють створити набір «асоціацій» між вхідними стимулами та необхідною реакцією на виході. В біологічному плані це відповідає перетворенню, наприклад, зорової інформації у фізіологічну відповідь рухових нейронів [7].

У алгоритмі перцептрон приймає набір входів і повертає набір результатів. Вони часто візуально представлені на графіках для користувачів. У багатьох мовах комп'ютерного програмування алгоритм перцептрона може мати форму циклу "для" або "поки", де кожен вхід обробляється для отримання результату. Результати показують, як ці вдосконалені типи алгоритмів вчать з даних - одна з визначальних характеристик перцептрона полягає в тому, що це не просто ітераційний набір процесів, а еволюціонуючий процес, коли машина вчиться з часу прийому даних.

Також, для аналізу отриманих даних можна залучити до процесу розробки та підтримки веб-додатку професіоналів, які займаються дизайном. Це рішення можна вважати достатнім для вирішення проблеми, але воно потребує досить багато затрат часу професіонала для ознайомлення з проектом, збору даних, аналізу статистики і т.д., та великих затрат коштів від власника.

2.2 Вибір алгоритму

Проаналізувавши існуючі методи виявлення невдалих дизайнерських рішень, можна зробити висновок, що єдиного і достатнього, та у той же час фінансово доступного, нема. Тому було вирішено спроектувати та реалізувати комплексний підхід вирішення проблеми – об'єднати метод збору статистичних даних та алгоритм аналізу отриманої інформації.

Для збору статистичних даних було обрано метод Heatmap (теплокарта), бо він надає достатньо багато даних для визначення місць, де може бути помилка (див. рис. 2.9). Також великим плюсом є те, що на теплокатах «нагріваються» місця, де користувачі бувають найчастіше, що дозволить виключити зі статистики одиничні випадки, коли у користувача виникли проблеми з використанням сайту або погане враження через чинники, не пов'язані з дизайном сайту.



Рисунок 2.9 - Теплокарта

Є не менш важливою інформація, яку сервіси такого типу здатні надавати - виконувані дії з зафіксованим часом у продовж сеансу – це дозволить будувати ланцюги послідовностей, над якими буде проводити аналіз наступний алгоритм.

Для аналізу отримуваних послідовностей дій, записаних з сайту, буде використовуватись самоорганізована карта Кохонена.

Перевагу було надано карті Кохонена через її простоту організації та відсутності необхідності її навчання додатковими засобами, що значно спрощує, прискорює та робить дешевшим в плані ресурсів процес розгортання (див. рис. 2.10). Тим паче це значно зменшує обсяг підготовчої роботи власника сайту, що захоче підключити систему до свого сайту.

Алгоритм буде будувати карту для кожної сторінки з нейронів, кожний з яких складатиметься з послідовності дій у часі. Даний алгоритм дозволить розпізнавати кожну послідовність та відносити її до певного класу, якщо події не відхиляються від очікуваних, та виявляти нові «міцніючі» кластери, які створилися у результаті нетипової поведінки, та привертати до них увагу [8].

Codes plot

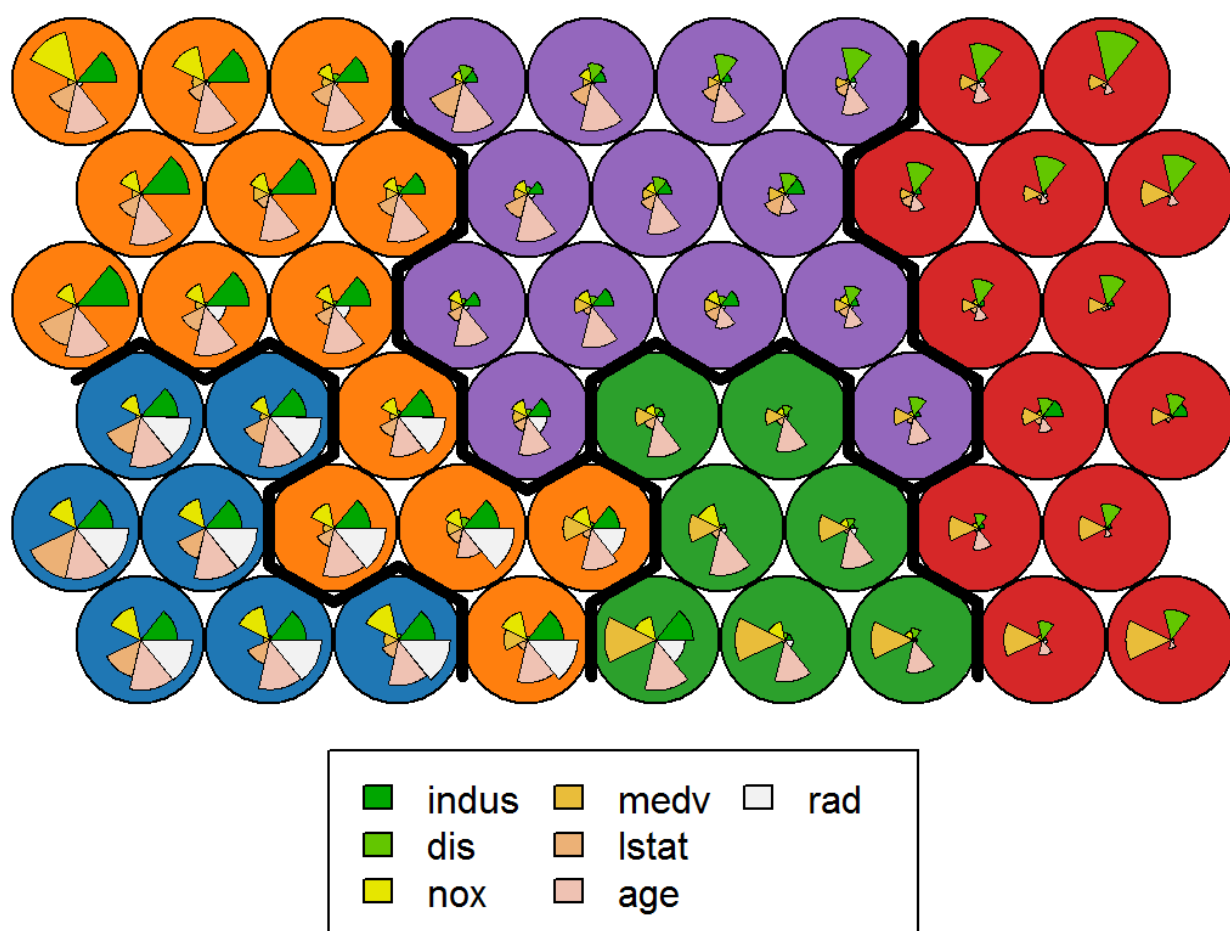


Рисунок 2.10 – Приклад карти Кохонена

Отримавши таким чином нетипову послідовність дій, пов'язати її з ділянкою на веб-сторінці не проблема.

Для алгоритму є два важливих показника, які експериментально будуть підібрані під час калібрування – тип сусідства та її розміри (див. рис. 2.11).

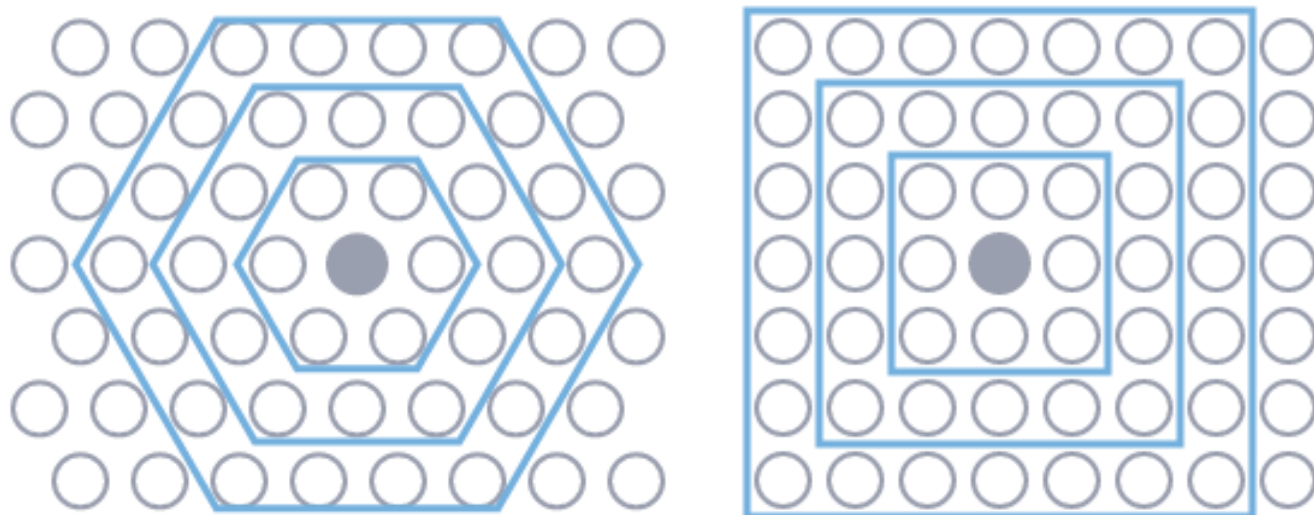


Рисунок 2.11 – Типи сусідства

Тип сусідства, - квадратний або шестикутний, впливатиме на кількість сусідів, а отже на швидкість поширення «заряду» нейрону, а розміри – на кількість кластерів, що можуть утворитися, та їх площу.

3 ОПИС РОЗРОБЛЕНОГО ПРОТОТИПУ ПРОГРАМНОЇ СИСТЕМИ

3.1 Налаштування сервісу спостереження

Для реалізації прототипу програмної системи виявлення помилок дизайну веб-сторінок було використано сервіс для спостереження Plerdy. Він надає зручний доступ до інформації, має безкоштовні тарифи та досить інформативні звіти з можливістю фільтрувати відображення дій користувачів (див. рис. 3.1).

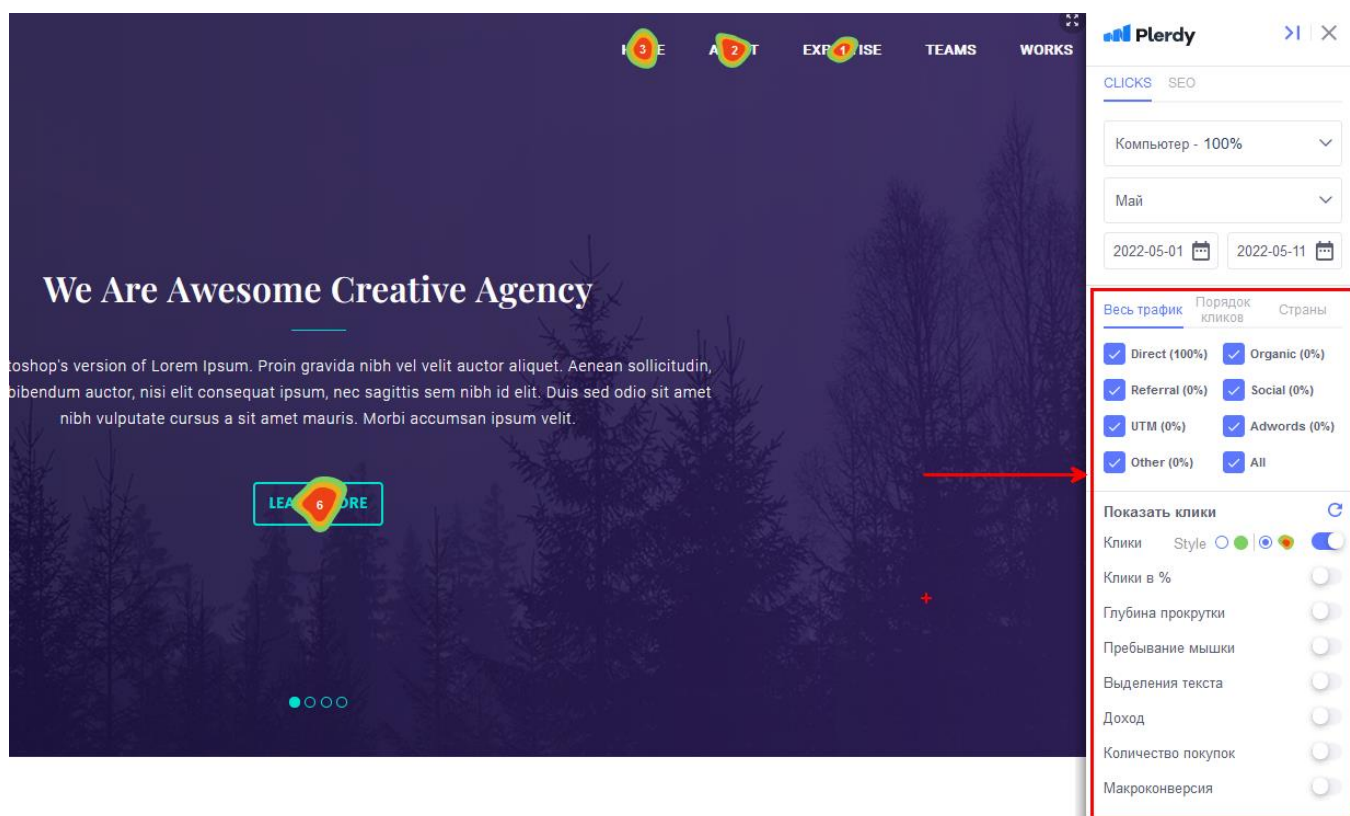


Рисунок 3.1 – Панель керування Plerdy

Під час виконання експериментів у сервісі було зареєстровано піддослідний сайт за допомогою надання до нього посилання та «вживлення» фрагменту коду до розмітки сторінки, який надає сервіс для передачі даних з сайту до бази даних.

Нажаль, жоден з безкоштовних тарифів різних сервісів не надає можливості експорту даних спостережень у файлах csv (Comma-Separated Values) формату, тому в рамках проведення експериментів дані переносились в ручному режимі.

В подальшому, при розробці повноцінно функціонуючої системи, можна розгорнути сервер та з'єднати його з модулем спостереження та отримувати дані за допомогою його API (див. рис. 3.2).

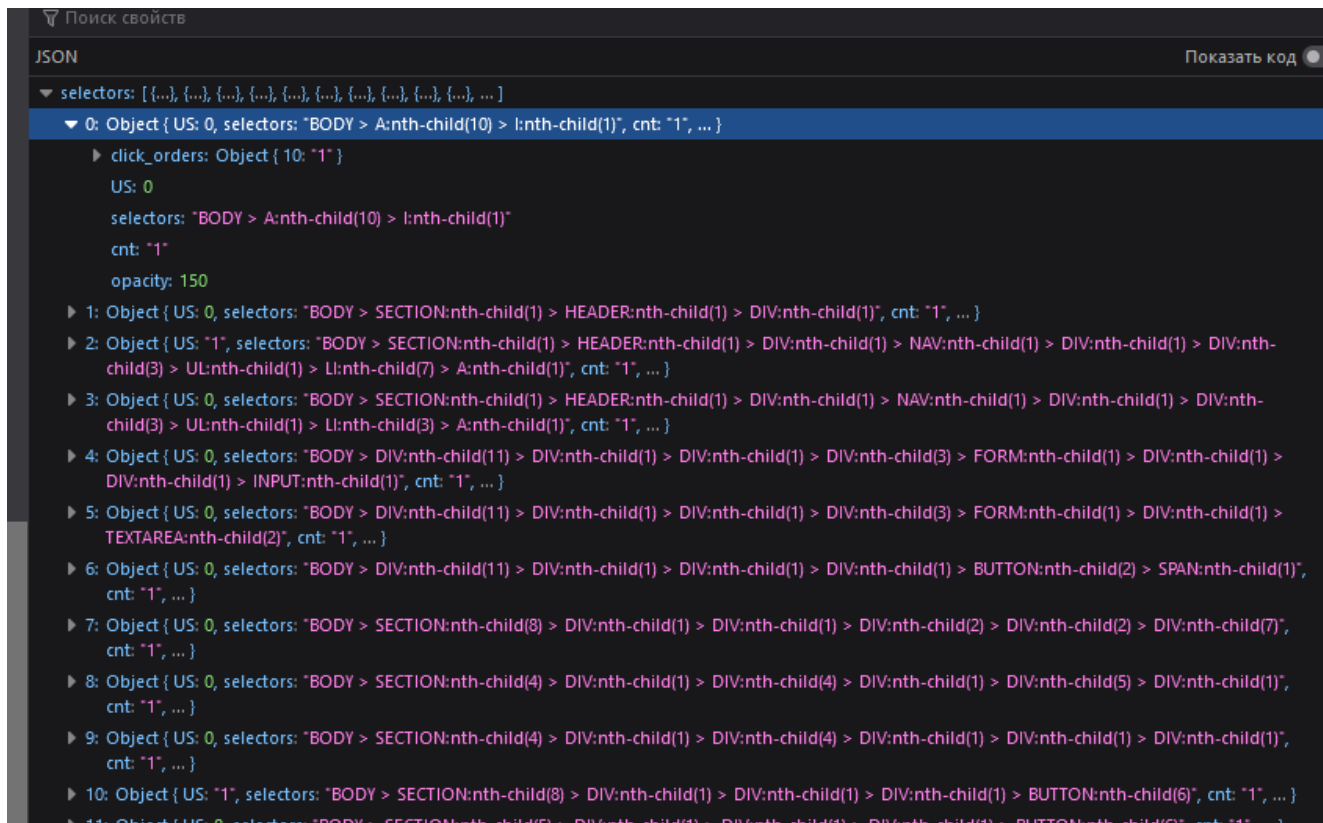


Рисунок 3.2 – Результат запиту API

У пакеті даних передається список, з якими користувач може взаємодіяти з їх місцем знаходження та кількістю використання.

3.2 Реалізація та налаштування карти Кохонена

Карта Кохонена працює за рахунок виконання наступних етапів:

- формування початкової матриці з нейронів, згенерованих випадковим чином;
- циклічне «навчання» нейронної мережі, кожна ітерація якого полягає у зміцненні нейрона та розповсюдженні його «знань» серед сусідів. В результаті такої обробки карти, будуються кластери, в середині якого знаходиться нейрон найбільш подібний до вхідних даних, а навколо нього – множина сусідів з подібними даними;

- виконання пошуку відповідного до вхідних даних кластеру або створення нового у разі неможливості знайти [9].

Для реалізації було обрано квадратний тип сусідства, тобто при візуалізації нейрони карти малюються квадратними, а для зберігання і обробки використовується звичайна матриця (див. рис. 3.3).

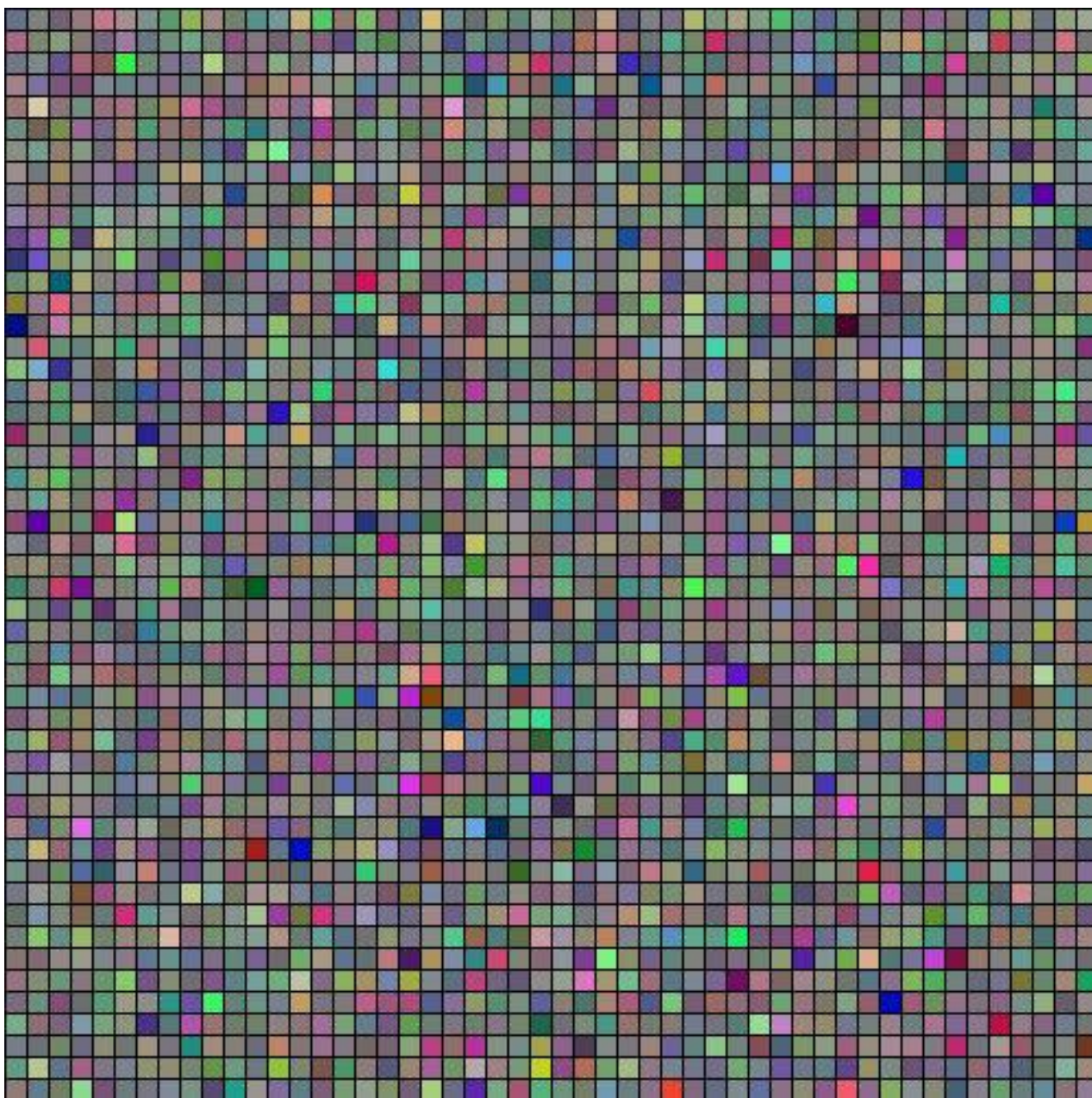


Рисунок 3.3 – Приклад візуалізації карти Кохонена з випадковими значеннями

Саме поняття нейрону є універсальним для кожної окремої карти Кохонена і містить наступні зміни:

- координати x та y, яке визначають місце нейрона на карті;
- вага нейрона – числова характеристика, яка за рахунок нормалізації завжди зводиться до однакових величин, а саме числа від 0 до 1 або від -1 до 1, незалежно від сенсу й змісту параметрів, які потрібні для порівняння [10].

Але є унікальним процес нормалізації, тобто перетворення комплексного і різноманітного пакету даних у одне досить обмежене число.

Вхідні дані для карти є множина теплових карт, які збирались з різних окремих сеансів відвідувань користувачів. Кожна теплова карта являє собою послідовність об'єктів, з якими виконувались дії, з кількістю натискань, координатами об'єкта на сторінці та радіусом зони входження курсора (див. рис. 3.4).

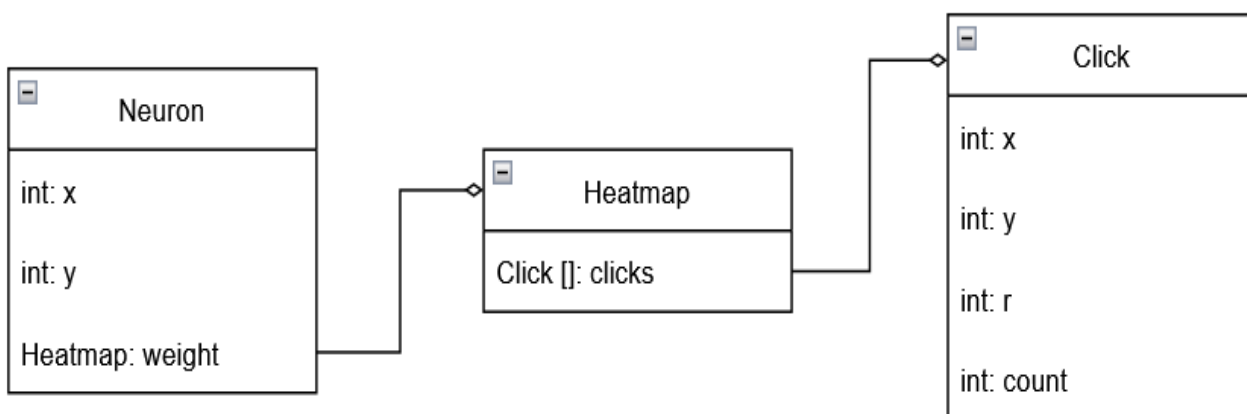


Рисунок 3.4 – Структура даних

Нормалізація вагових коефіцієнтів нейронів карти відбувається у наступній послідовності:

- перебір всіх можливих значень параметрів, що будуть порівнюватись у наступних етапах роботи з метою пошуку найбільших та найменших серед них;
- повторний перебір, але вже з заміною параметрів на нормалізовані значення у заданому діапазоні.

Нормалізація кожного значення виконується за наступною формулою

$$V_n = \frac{(x-x_{min})*(d_{max}-d_{min})}{x_{max}-x_{min}} + d_{min} ,$$

де x – значення параметру, яке потребує нормалізації;

x_{max} та x_{min} – максимальне та мінімальне значення параметру;

d_{max} та d_{min} – максимальне та мінімальне значення діапазону нормалізації відповідно.

Після виконання нормалізації відбувається «навчання» нейронної мережі з так званим «пошуком переможців».

«Пошук переможців» полягає у переборі вхідних даних та випадково згенерованих нейронів карти та у пошуку таким чином пари нейрона та даних, що мають найменшу різницю значень ніж у інших подібних парах.

Після знаходження такої пари виконується формування кластеру. Під час цього процесу нейрон-переможець та всі його сусідні нейрони проходять обробку, яка модифікує значення вагових коефіцієнтів, таким чином щоб вони стали трохи подібнішим до нейрона-переможця.

При візуалізації навченої карти чітко виділяються сформовані кластери, де колір нейрона-переможця поступово охоплює сусідів (див. рис. 3.5).

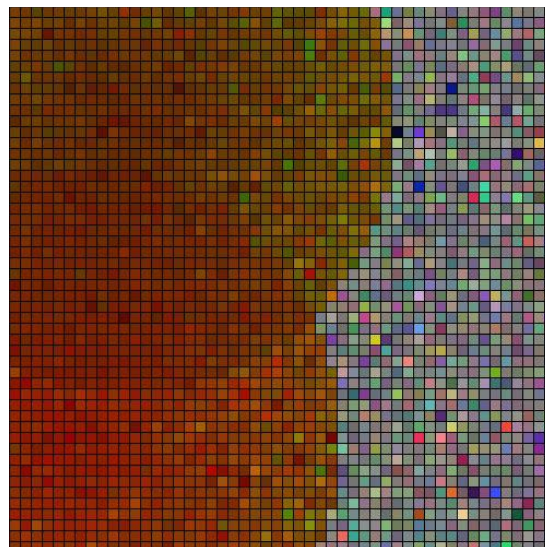


Рисунок 3.5 – Приклад «навченої» карти Кохонена

Модифікація значень параметрів виконується по наступній формулі

$$V_a = V_w + (\theta * \alpha_i * (V_o - V_w)) ,$$

де V_a – модифіковане значення;

V_w – значення параметру нейрона-переможця;

θ – коефіцієнт впливу нейрона-переможця на сусідні нейрони;

α_i – швидкість розповсюдження значень нейрона-переможця серед сусідніх нейронів;

V_o – значення, що треба модифікувати.

Швидкість для формули буває абсолютною та ітеративною, яка розраховується для кожної ітерації розповсюдження окремо. Абсолютне значення задається у конфігураціях системи та потребує калібрування.

3.3 Реалізація алгоритму виявлення помилок

Алгоритм виявлення помилок дизайну веб-сторінки базується на принципі виявлення відхилень у спостереженнях від очікуваної або складеної стереотипної поведінки.

В залежності від стадії існування, на якому сайт було під'єднано до системи, можливі наступні сценарії роботи:

- у випадку, коли систему використовують на етапі розробки сайту, коли дизайн сайту ще доробляється і не було реального досвіду використання користувачами – є необхідність згенерувати еталонні теплові карти для кожної категорії користувачів та сторінки, у яких дизайнер має відобразити очікувану ним поведінку користувачів на сайті. Тобто, базові «знання» нейронної мережі створюються штучно;

- якщо ж систему використовують вже на повноцінно працюючих сайтах, тоді потрібно лише ввімкнути її. Нейронна мережа буде самостійно «навчатись», але тоді є необхідність участі спеціаліста, бо система буде сповіщати про кожний новий створений кластер, тобто кожну нову модель поведінки, і її

треба вказувати що з цього є нормальною поведінкою, а що ні. Звісно, таке можливо лише при умові, що є впевненість у правильності поточного стану дизайну сайту.

Для створення еталонної теплової карти достатньо лише відтворити очікувану поведінку користувача у окремому сеансі на сайті, до якого підключено сервіс спостереження. Більшість сервісів надають можливість отримувати дані з окремих обраних сеансів та сторінок (див. рис. 3.6).

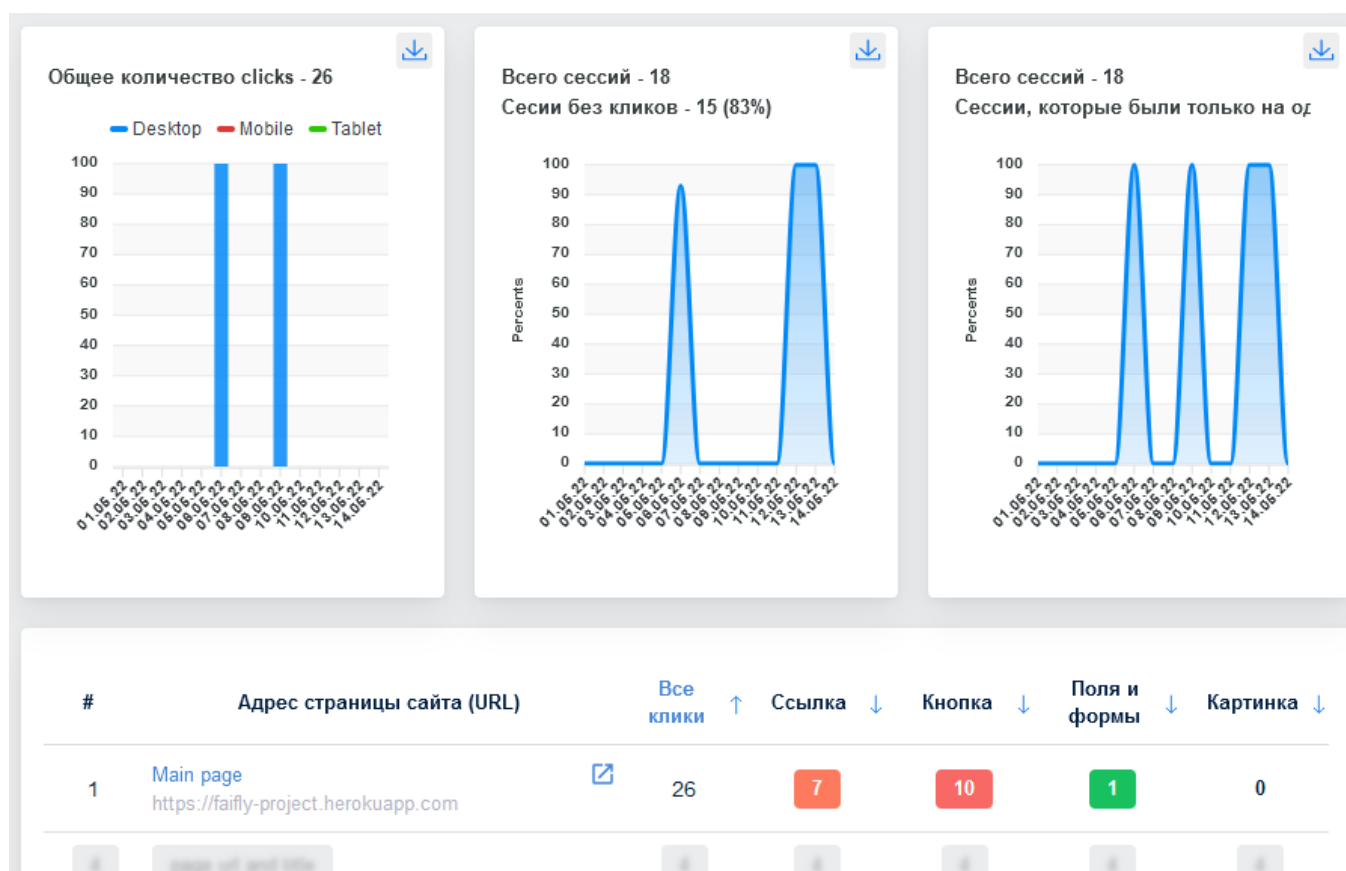


Рисунок 3.6 – Панель сеансів сервісу

Результатом роботи алгоритму виявлення помилок є звіт, у якому сповіщається про виявлення відхилення від типової поведінки користувача з вказанням саме того «місця», де користувач почав працювати з сайтом неочікуваним чином.

Сам алгоритм працює в наступній послідовності:

- фіксування послідовності дій у вигляді теплових карт;

- спроба віднести отриману теплову карту до одного з існуючих кластерів у карті Кохонена;
- у разі якщо вдалося знайти подібний до вхідних даних кластер, то виконується зміцнення кластеру, тобто корегування його центрального нейрону та сусідів;
- у випадку невдачі – відбувається формування нового кластеру, про появу якого складається звіт;
- для звіту виконується порівняння теплової карти, яка прив’язана до кластеру, з тепловою картою, наданою з вхідними даними, виявляються ті елементи послідовностей взаємодій, які створюють різницю [11].

Отримані елементи дизайну додаються до звіту з відповідним маркуванням (див. рис. 3.7).

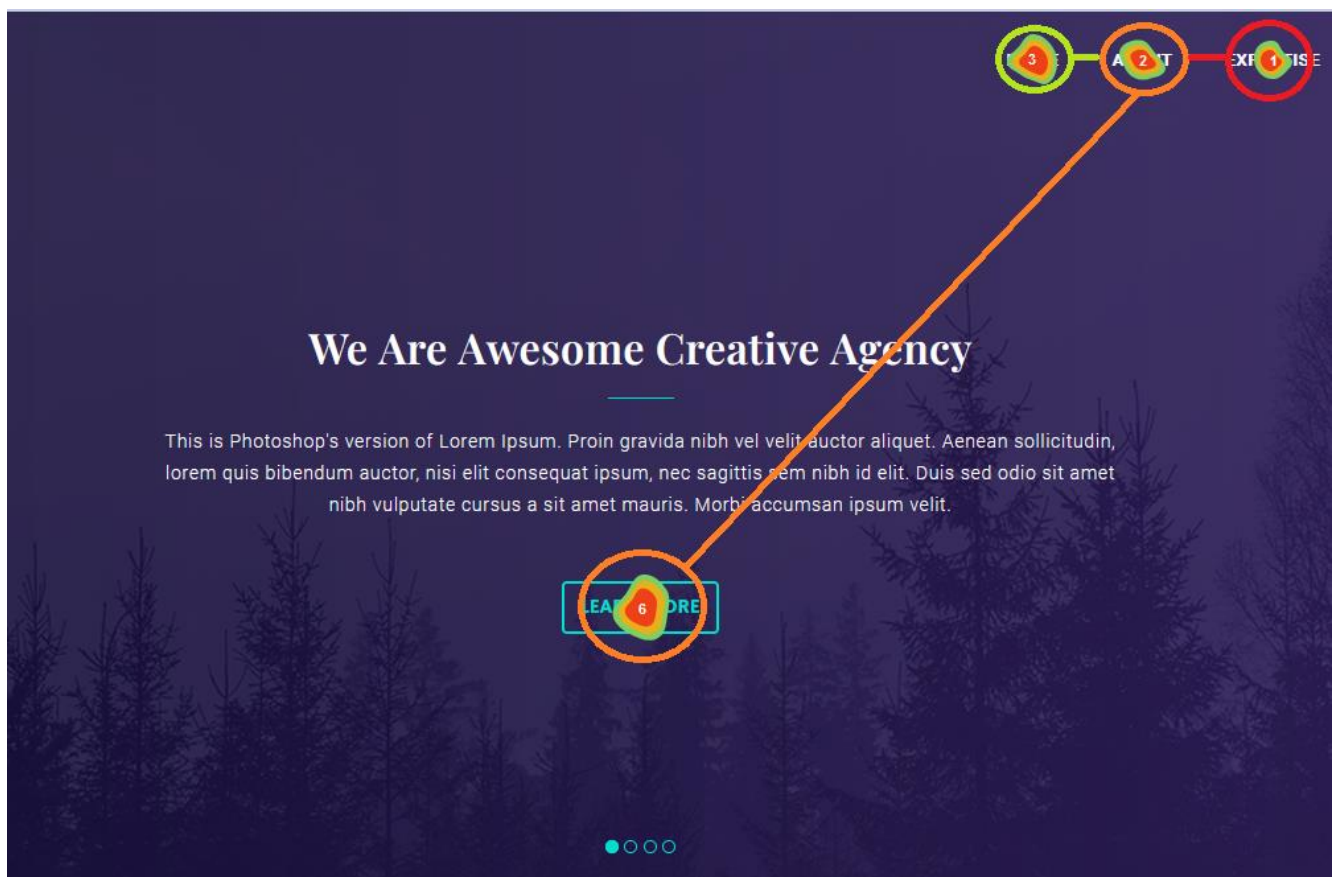


Рисунок 3.7 – Приклад звіту

У звіті для відображення порівнюваних теплових карт використовуються

наступні маркування:

- круги – для виділення об’єктів, з якими виконувались взаємодії;
- лінії, що з’єднують круги у ланцюг – відображають послідовність, у якій виконувались дії.

Також ці фігури мають забарвлення відповідно до послідовності, якою вони вважаються:

- зелений колір – колір еталонної послідовності або послідовності, правильність якої підтвердили спеціалісти;
- помаранчевий колір – колір, яким позначаються дії та послідовності, що співпадають у очікуваної та піддослідної поведінки;
- червоний колір – колір, який використовується для виділення відхилень піддослідної послідовності від тієї, що вважається правильною.

У результаті аналізу, помилки дизайну імовірно можуть бути в об’єктах з зеленим або червоним маркуванням. Тобто, в місцях де думки користувача та дизайнера розходяться.

В такому випадку власнику сайту слід передивитися дизайн сторінки та або змінити об’єкт з зеленим маркуванням, щоб він був більш бажаним для користувача, або модифікувати елемент з червоним маркуванням таким чином, щоб користувачу він здавався менш важливим.

3.4 Архітектура системи

Для подальшої повноцінної програмної реалізація системи виявлення помилок дизайну веб-сторінки розроблено схему, яка складається з наступних елементів:

- централізована база даних для розміщення найпоширеніших поведінкових шаблонів;
- база даних в межах одного бізнес-середовища для проведення аналізу дизайну цільового веб-додатку;
- залучена зовнішня система збору статистичних даних дій користувача на веб-сторінці з використанням технологій теплових карт;

- модуль, що відповідає за аналіз отриманих даних.

Централізована база даних буде використовуватись для надання нейронній мережі можливостей по нарощуванню знань, які у глобальних масштабах роботи системи дозволять значно підвищити точність та швидкість виявлення помилок, оптимізувати процес навчання нейронної мережі та пришвидшити розгортання системи до стану, коли вона працює виключно на пошук та аналіз помилок, а не отримання початкових знань (див. рис. 3.8).

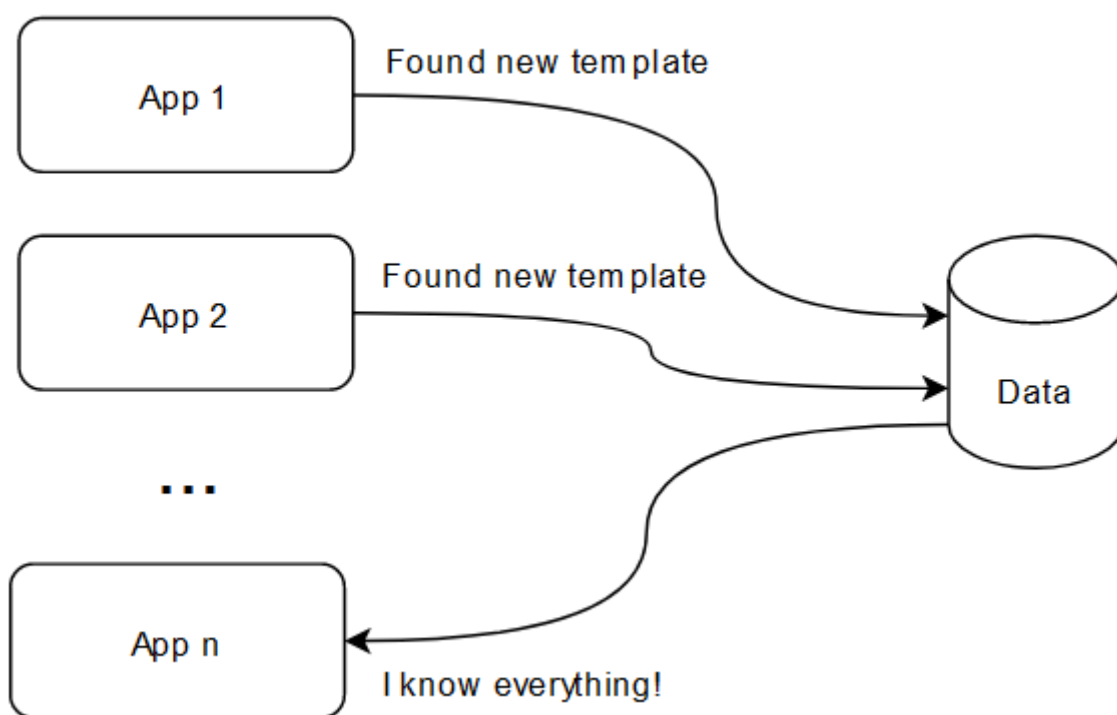


Рисунок 3.8 – Схема комунікації

Спілкування окремого екземпляру додатку з централізованою базою даних виконується при накопиченні відповідної «порції» досліджень, що зробила локальна система.

Розроблена таким чином схема дозволить виділити спільні для всіх сайтів помилки, які для кожного наступного сайту будуть виявлятися все швидше і швидше.

Локальна, в межах одного бізнес-контексту, база даних виконує функцію сховища, у якому зберігаються дані, підходящі лише для певного середовища, та виконує функцію буфера, у якому і буде накопичуватись «порція» даних для переміщення у централізоване сховище (див. рис. 3.9).

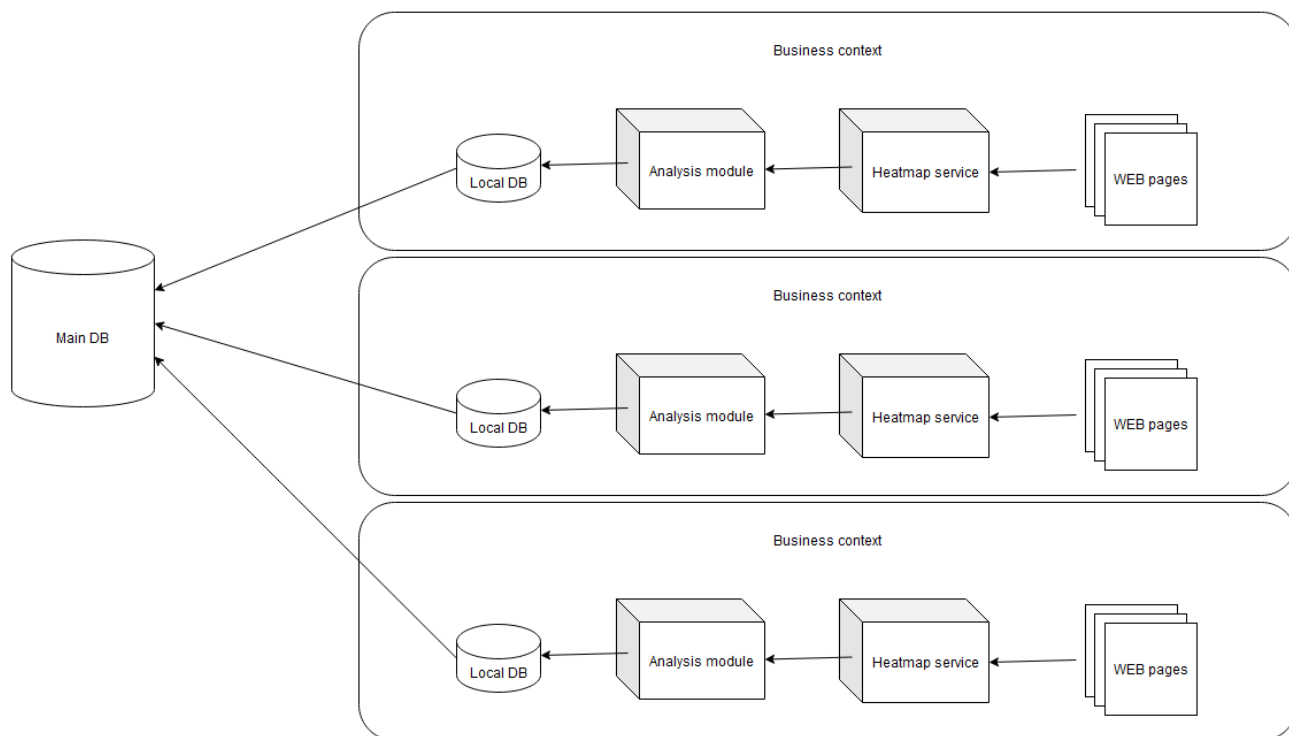


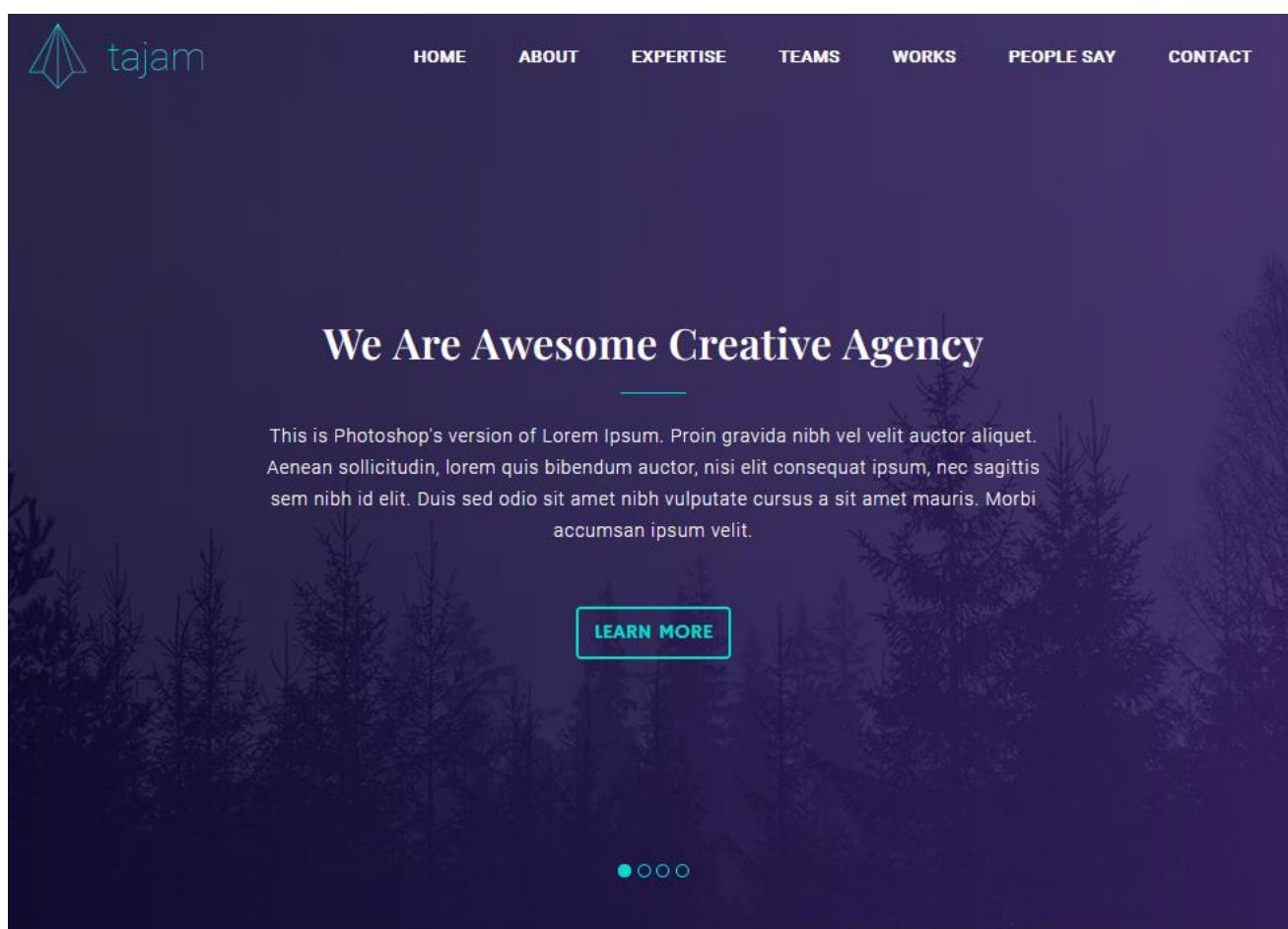
Рисунок 3.9 – Схема взаємодії у системі

Для проведення експериментів під час дослідження реалізована частина схеми, що спрямована лише на використання у локальному просторі.

Провести експеримент з накопиченням шаблонів та передачі їх через централізоване сховище можливо лише в умовах коли працездатні локальні системи напрацювали чималу базу знань на реальних об'єктах, а це потребує багато часу на доведення системи до задовільного стану, розповсюдження її та налаштування та ін.

4 ОПИС ДОСЛІДЖЕНЬ

Дослідження проводились на автономному екземплярі програми без повноцінного розгортання системи через необхідність використання багатьох ресурсів для хостингу та організації з'єднання та взаємодії по мережі. Також неможливим було і проведення дослідження на реальних працюючих сайтах з реальними їх відвідувачами – замість цього було використано попередньо створений та розміщений у Інтернеті сайт, на якому симулювалась поведінка користувачів (див. рис. 4.1).



OUR STORY

This is Photoshop's version of Lorem Ipsum. Proin gravida nibh vel velit auctor aliquet. Aenean sollicitudin, lorem quis bibendum auctor, nisi elit consequat



Рисунок 4.1 – Піддослідний сайт

В наведених умовах є можливим проведення дослідження тільки із сценарієм, коли кластери будуються на основі еталонних теплових карт, вказаних спеціалістом. Проведення досліду за другим сценарієм потребує значно більше часу і має зовсім велику невпевненість у строках через недостатню відомість піддослідного сайту.

Для проведення досліду було створено 3 еталонних теплових карти, на яких зафіксовано 3 різних можливих послідовностей дій відповідно.

Перша послідовність створена для того випадку, коли користувач ознайомлюється з сайтом (див. рис. 4.2).

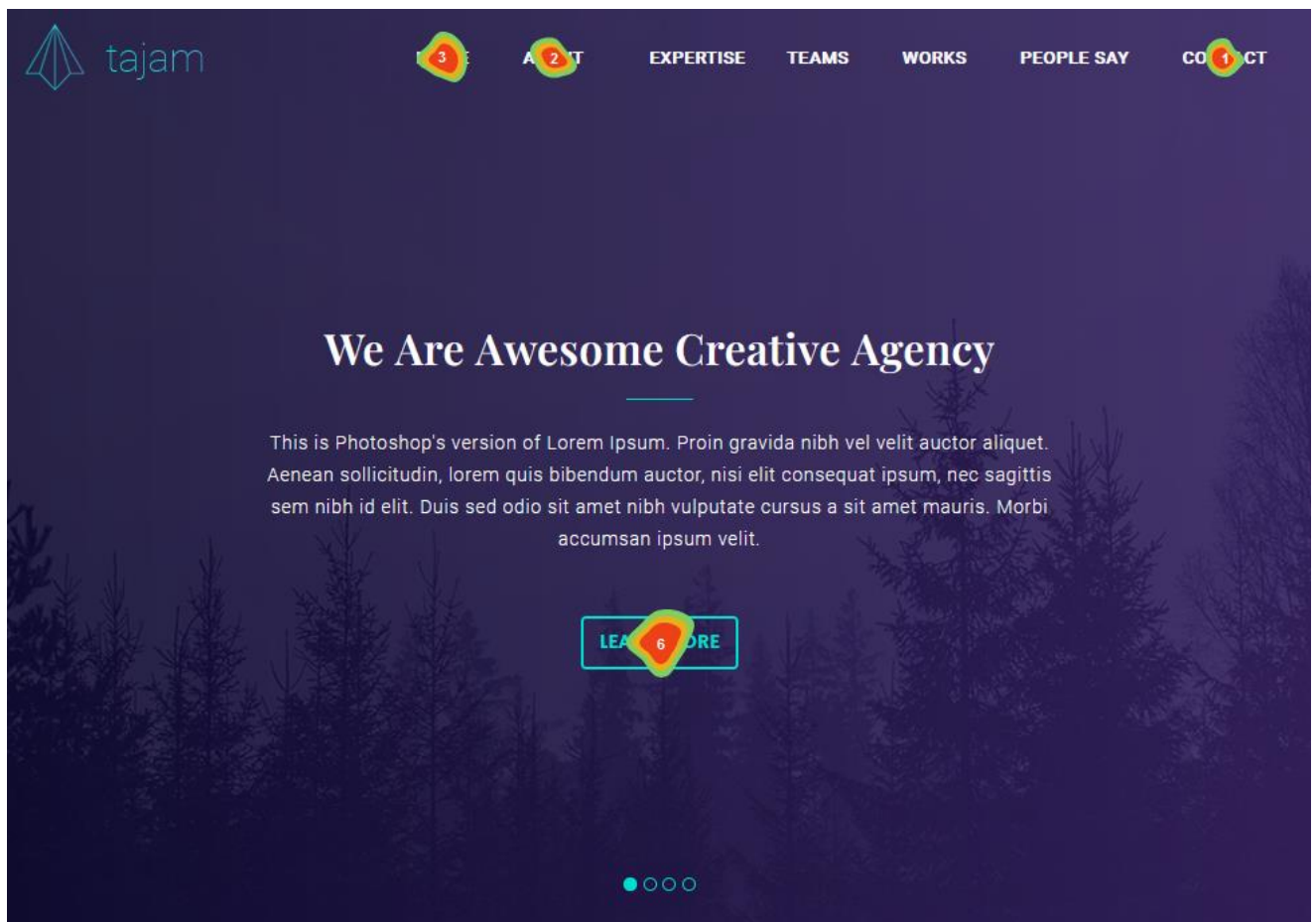


Рисунок 4.2 – Теплова карта першої послідовності

У цій послідовності користувач натискає на кнопку «Home», далі на головній сторінці натискає «Learn more», повертається назад на сторінку і натискає «About», а після ознайомлення переходить до контактів для отримання консультацій.

Для другої послідовності було обрано сценарій пошуку та дослідження послуг на сайті (див. рис. 4.3).

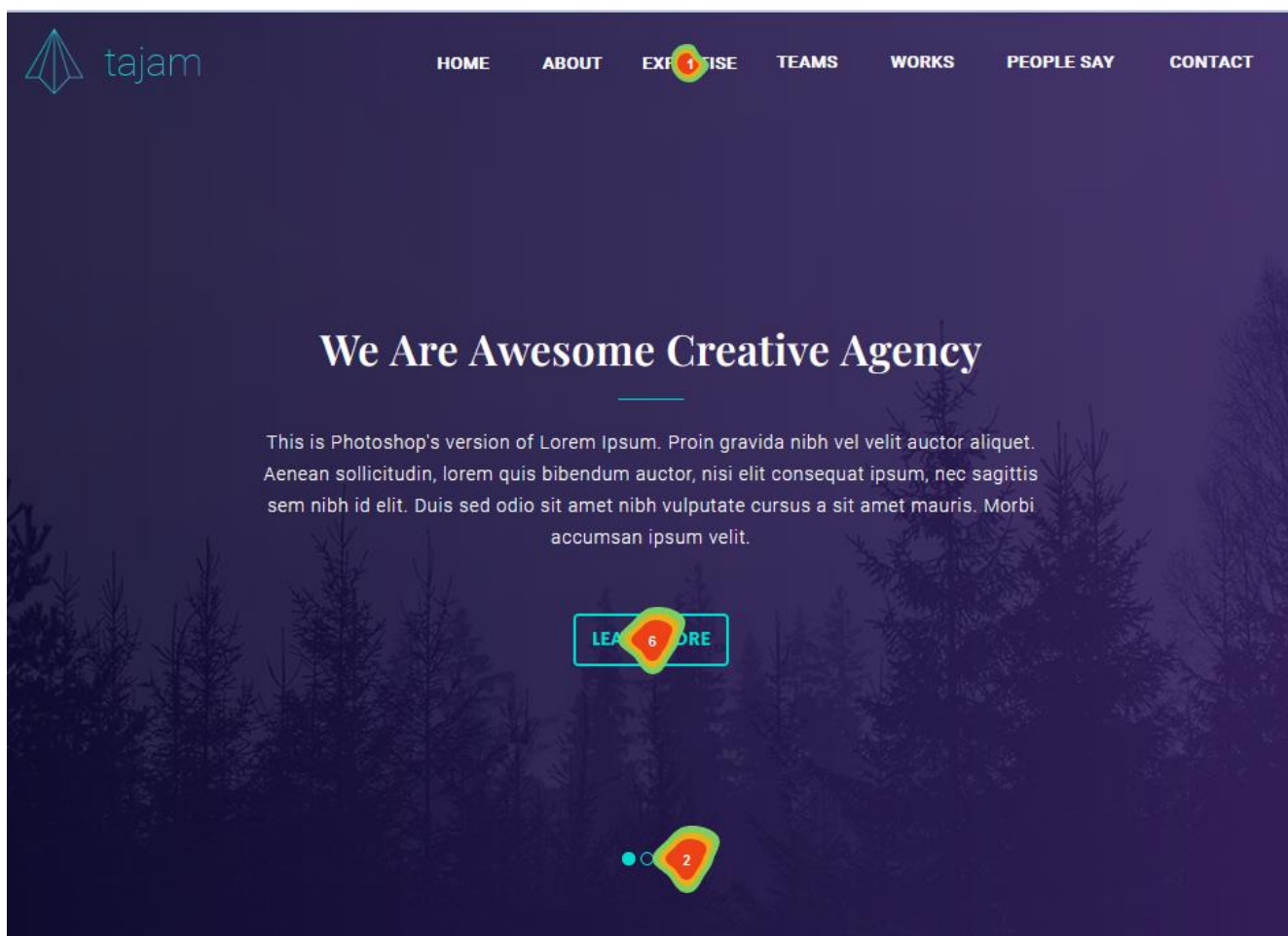


Рисунок 4.3 – Теплова карта другої послідовності

Така послідовність припускає взаємодію з кнопкою «Expertise», перебір промо зображень у «каруселі» та ознайомлення з сайтом у розділі «Learn more».

Наступна еталона послідовність складається з переліку дій, які імітують поведінку користувача, що хоче найняти спеціаліста.

Для цього користувач спершу натискає на кнопку «Teams» та досліджує біографії членів команди, потім викликає вікно чату за допомогою «плаваючої» кнопки. Взаємодії користувача, що відбувались на сторінках з профілями спеціалістів під час експериментів не враховувались, бо відносились до іншої сторінки (див. рис. 4.4).

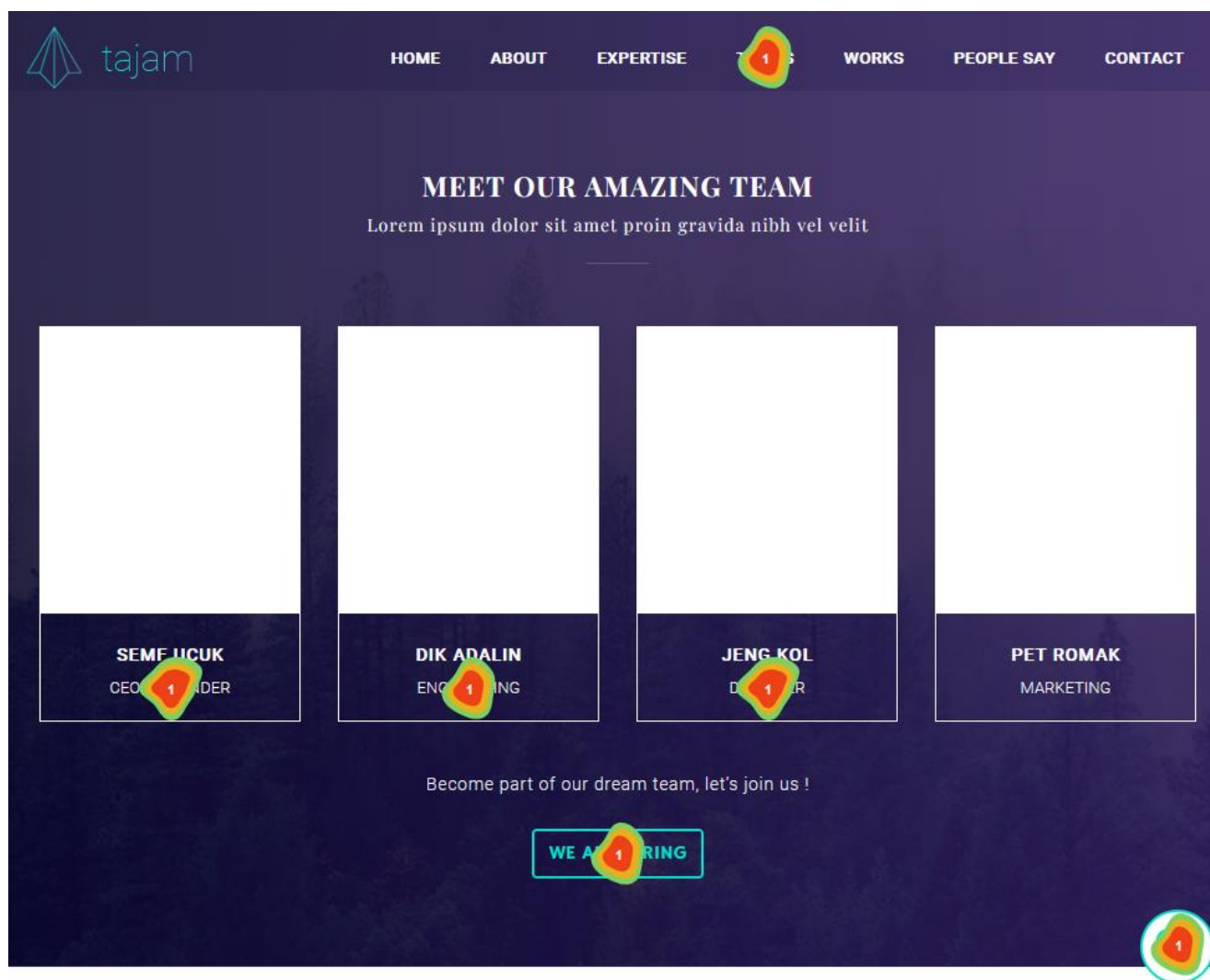


Рисунок 4.4 – Теплова карта третьої послідовності

Отримані дані теплових карт було занесено до системи та виконано навчання карти Кохонена. Навчання карти проводилося з різними показниками коефіцієнту впливу та кількості ітерацій навчання, тобто були використані комбінації двох.

Значення коефіцієнту впливу були від 0.1 до 0.7, а значення кількості ітерацій – 10, 100 та 1000.

Отримані результати навчання карти Кохонена було відображено на окремих малюнках для кожної комбінації. Після дослідження всіх зображень карти, було обрано для подальшої роботи наступні значення:

- коефіцієнт швидкості розповсюдження – 0.1;
- кількість ітерацій – 100.

При таких значеннях нейрони-сусіди, які увійшли до кластера піддаються меншій модифікації, що зберігає достатньо велику різноманітність значень, що увійшли до кластеру (див. рис. 4.5).

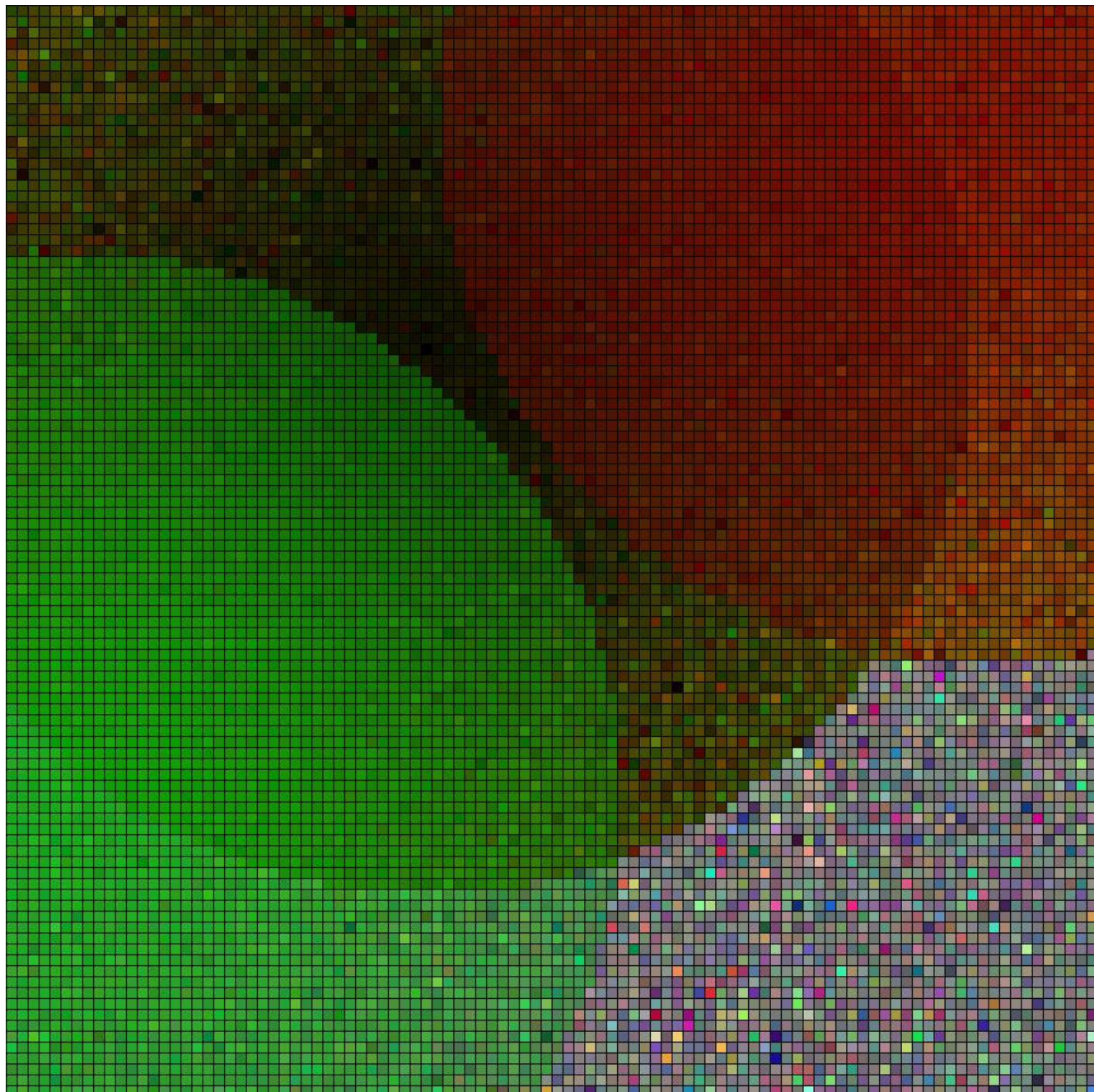


Рисунок 4.5 – Візуалізація карти Кохонена при значеннях швидкості розповсюдження 0.1 та кількості ітерацій 100

На зображенні чітко видно сформовану відповідну вхідним даним кількість кластерів та достатньо різноманітні значення сусідів.

Далі наступний етап дослідження – перевірка на скільки система здатна розпізнавати у поведінці користувача існуючі моделі поведінки.

Для цього у систему було додано дані теплових карт, на яких виконувались дії, що відповідають хоча б одній існуючій моделі. Очікуваним результатом досліджу є зміцнення існуючих кластерів та не створення жодного нового.

Після процесу перенавчання з новими даними, спостерігається корегування значень існуючих кластерів на користь уточнення та урізноманітнення даних (див. рис. 4.6).

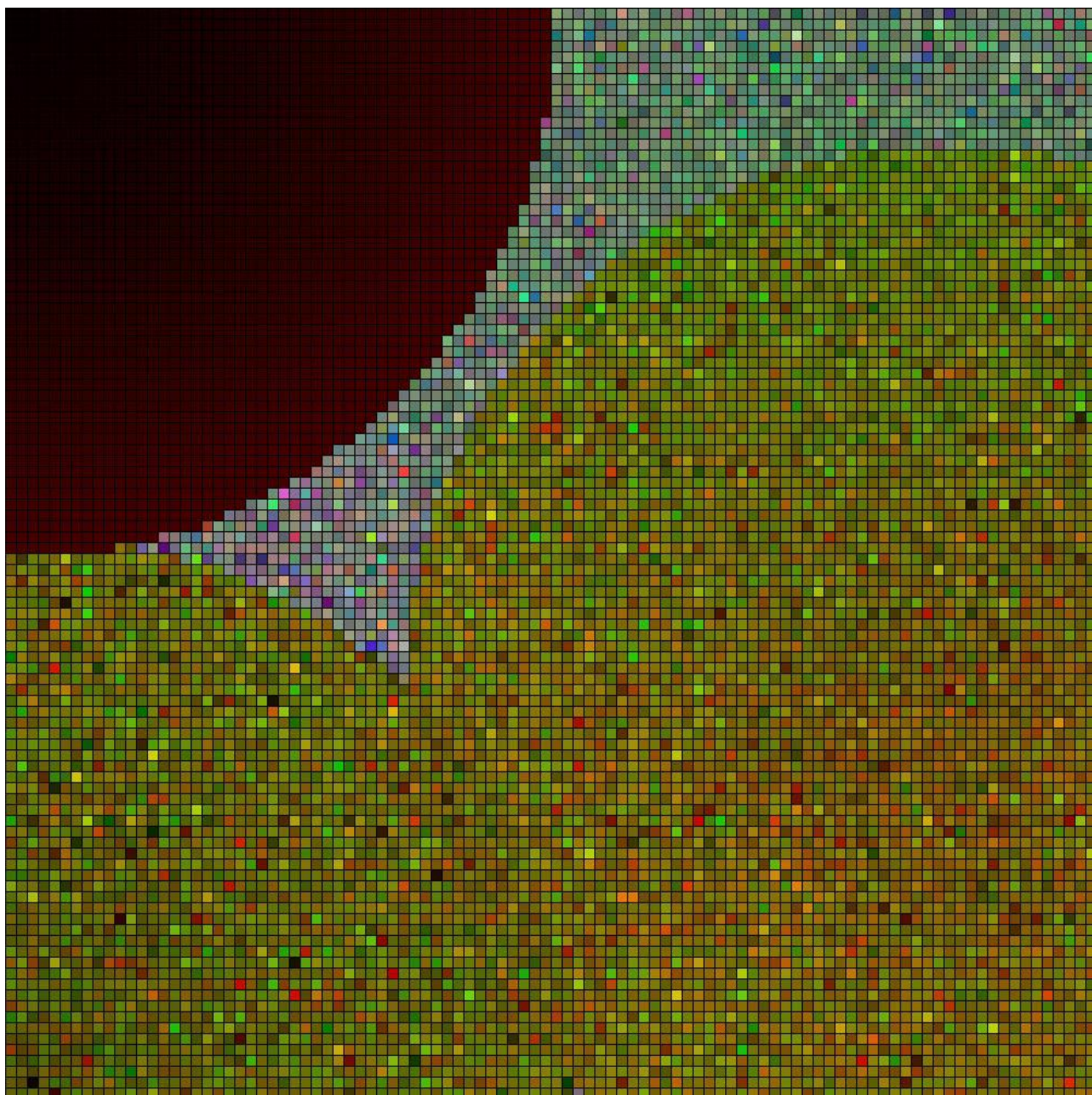


Рисунок 4.6 – Візуалізація карти Кохонена після перенавчання

Також кластерів спостерігається все ще 3 та стає більш наочна подібність двох кластерів, що відповідають дійсно трохи схожим першій та другій послідовності, в той час як третя послідовність зовсім інша та відповідно формує кластер зовсім іншого кольору.

Наступним і останнім етапом дослідження є перевірка спроможності системи виявляти відмінність поведінки користувача від існуючих моделей. Для цього до даних додали ще одну теплову карту з послідовністю, яка повторює одну з існуючих послідовностей, але має хоча б одну дію значно відмінну від раніше зафіксованих.

Після чергового перенавчання нейронної мережі було помічено утворення нового кластеру, що означає фіксацію картою нової послідовності (див. рис. 4.7).

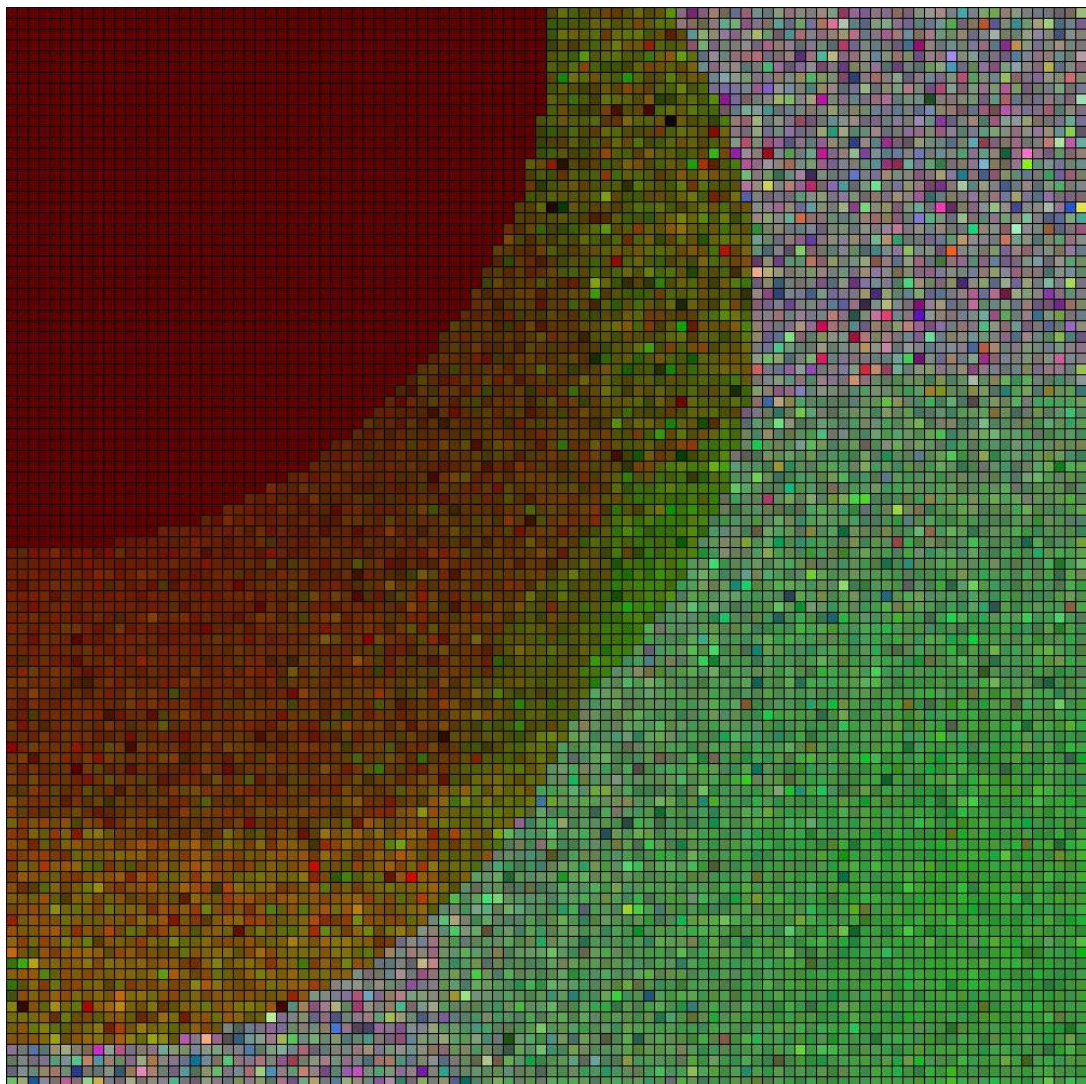


Рисунок 4.7 – Візуалізація карти Кохонена на останньому експерименті

Після виявлення появи нового кластеру система починає порівняння теплових карт, що зберігаються у нейронах, які являються найближчим сусідом.

Знайдені таким чином відмінності у послідовностях можна вважати тими самими «проблемними зонами» у дизайні – як ті елементи, що були у еталонних теплових картах, так і елементи, з якими взаємодівав користувач. Але варто також зазначити, що виявлені таким чином послідовності не всі є помилкою, а можуть бути ще однією гілкою взаємодій із сайтом, тому потребується участь спеціаліста у процесі вдосконалення сайту – для визначення чи є нова послідовність помилкою. Необхідність такого втручання може бути спровокована й недостатньо ретельного аналізу потенційних користувачів та проробки варіантів використання, та як слід невдалу теплову карту.

Але ця проблема – надлишкове залучення спеціалістів – є лише наслідком відсутності досить великої бази знань і має вирішитись сама собою. Необхідність підтвердження чи спростування правильності кожної послідовності буде зменшуватись з часом по експоненційно спадаючому графіку, поки не залишаться виключні випадки. А з реалізацією у майбутньому схеми з централізованою бази «знань» нейронної мережі також зменшиться необхідність у спеціалісті й на начальних етапах.

Отриманий звіт з такими даними дозволить досить швидко зорієнтуватись дизайнеру на будь-якому етапі розробки сайту.

ВИСНОВКИ

В результаті роботи було проведено аналіз існуючих комплексних заходів по виправленню помилок дизайну, починаючи з етапу розгортання бізнесу у Інтернеті та закінчуючи супроводом сайту після його виконання, та окремих методів спостереження та збору статистичних даних та методів і алгоритмів аналізу даних.

Спроектowana система потребує вказання місцеположення сторінок сайту у Інтернеті у якості вхідних даних та для коректної роботи необхідно наростити базу знань, тобто «навчити» нейронну мережу. Для цього можна підвергнути сайт стрес-тестам або додатково перевіряти висновки, що робить система перший час.

Додатково до сайту необхідно підключити сервіс, який зчитуватиме дії користувачів з сайту і передаватиме до аналізатора.

У наступних релізах планується розробити самостійний модуль теплокарт та вживити його до системи для зменшення початкових конфігурацій.

Також планується централізація даних, що дозволить повністю витіснити процес «навчання» нейронної мережі під час кожного наступного розгортання та значно підвищить точність виявлення помилок.

Не менш важливою є модифікація - підключення системи до хостів, на яких розміщені сайти, для відстеження того, як саме було виправлено певні помилки та буде підказувати конкретні дії щодо поліпшення стану дизайну у разі виявлення помилки.

Було проведено експерименти з використанням реалізованого прототипу системи з достатніми для дослідів можливостями, у яких досліджувались можливості методу спостереження та алгоритму виявлення помилок формувати кластери з різними моделями поведінки, відносити отримані спостереження до зафіксованих раніше послідовностей дій та виявляти появу нового кластеру. Під час проведення експериментів було використано заздалегідь створений набір даних, розміщений на хостингу сайт з «вживленим» модулем сервісу спостереження.

У порівнянні з існуючими методами, зокрема з тими, де майже всю роботу виконує людина, вдалося досягнути значного пришвидшення, полегшення та

збільшення точності у процесі пошуку та виявлення факту утворення відхилень від очікуваної поведінки користувачів.

Також вдалося значно зменшити присутність спеціалістів під час роботи над вдосконаленням сайту. Але, нажаль, не повністю виключити, бо модулю для проведення аналізу потрібна допомога у формуванні уявлення про «правильне» та «не правильне».

В поточному стані спроектованої системи роль спеціаліста стає більше обслуговуючою, а ніж ключовою, коли людина самостійно виконує всю роботу – дослідження зібраної статистики, аналіз результатів спостережень багатьох сервісів, відстеження закономірностей у діях користувачів, формування тих же самих поділів моделей поведінки на різні категорії користувачів та цілі таких послідовностей.

За рівнем новизни результати дослідження можна оцінити як «вперше одержано», бо подібних поєднань методів збору даних та аналізу з певними модифікаціями не траплялось, тим паче у розглянутій сфері.

ПЕРЕЛІК ДжЕРЕЛ ПОСИЛАННЯ

1. Лаптон Е., Філліпс Дж.-К.: Характеристики Графічний дизайн. Нові основи – Україна: ArtHuss, 2020.
2. Мельнікова Р.В., Лозиченко А.В. Дослідження методів виявлення помилок проектування сайтів. Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління. Тези доповідей дванадцятої міжнародної науково-технічної конференції. Том 2: секція 5. 27-28 квітня 2022, м. Харків, Україна, с. 180.
3. Хьорф С.: Усе про книжку Как создать продукт, который полюбят. Опыт успешных менеджеров и дизайнеров. – Україна: Манн, Іванов і Фарбер, 2019
4. Breiman, L.: Bagging predictors. Machine Learning, Vol. 24, 1996, 123-140 p.
5. Кохонен, Т. Ассоциативные запоминающие устройства. М.: Мир, 1982. – 384 с.
6. Минский М., Пейперт С. Перцептроны. М.: Мир, 1971 - 261 с.
7. Терехов В.А., Ефимов Д.В., Тюкин И.Ю. Нейросетевые системы СПб.: Издательство С.-Петербургского университета, 1999. – 265 с.
8. Хайкин С. Нейронные сети: полный курс, 2-е изд., испр. : Пер. с англ. - М.: ООО "И.Д. Вильямс", 2006. – 1104 с. управления: Учеб. Пособие для вузов - М.: Высш. шк. 2002. – 183 с.
9. Дударь З.В., Егоров С.В. Оптимизация алгоритмов кластерного анализа в задачах распознавания образов. – г. Херсон: Вестник ХНТУ №2(41) 2011. – 172-173 с.
10. Dudar Z.V., Bondarenko M.F., Chetverikov G.G. Multiple-valued structures of Intellectual Systems: int. book ser. «Information Science and Computing». 2008 – 125-130 p.
11. Дударь З.В., Шуклин Д.Е. Реализация нейронов в семантических нейронных сетях. – г. Харьков: Радиоэлектроника и информатика, №4. 2000 – 24-29 с.