



О. С. Васютук, П. В. Сердюк

Національний університет "Львівська політехніка", м. Львів, Україна

ІНТЕЛЕКТУАЛЬНА ГЕОІНФОРМАЦІЙНА СИСТЕМА ДЛЯ ПРОГНОЗУВАННЯ МІСЬКИХ ПРОБЛЕМ ІЗ ЗАСТОСУВАННЯМ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ ТА ГЕЙМІФІКАЦІЇ

Розроблено інтелектуальну геоінформаційну систему для управління міськими інцидентами на підставі клієнт-серверної архітектури (ASP.NET Core, React), яка інтегрує геопросторовий сервіс, аналітичну панель і модуль прогнозування. Основною інновацією є поєднання двох типів прогностичних моделей (LightGBM (англ. *Light Gradient Boosted Machine*) для тривалості виконання заявки та SSA (англ. *Singular Spectrum Analysis*) для кількості звернень) з подієво-орієнтованим модулем гейміфікації, який не тільки підвищує залученість громадян до збирання даних про міські проблеми, але й створює механізм зворотного зв'язку для цілеспрямованого покращення якості вхідних даних для навчання і тестування самих моделей. Проаналізовано відкритий набір даних "Звернення на Гарячу лінію міста Львова" (2022-2024 рр.) та на його основі впроваджено моделі машинного навчання засобами ML.NET. Проведено багатоетапну підготовку даних для регресійної задачі: початковий набір даних із 362 тис. записів очищено від дублікатів, відфільтровано за тривалістю виконання (60-17500 хв) та категоріями (>5000 прикладів), сформувавши фінальний навчальний набір даних із 118 708 записів. Виконано детальну інженерію ознак (англ. *Feature Engineering*) шляхом генерування циклічних часових (sin/cos) та геопросторових (широта, довгота) ознак, а також доповнення погодними характеристиками. Для забезпечення високої продуктивності та потокобезпечного доступу до моделей у режимі реального часу обґрунтовано та реалізовано інженерне рішення з використанням механізму PredictionEnginePool. Проведено детальну валідацію розроблених моделей. Встановлено, що регресійна модель LightGBM пояснює 64 % варіативності тривалості виконання заявки ($R^2 = 0,640$, MAE = 1707,9 хв). Виявлено через аналіз важливості ознак домінування операційних факторів ("Виконавець", "Категорія") над зовнішніми ознаками (погодні умови). З'ясовано, що багатомодельна архітектура SSA для прогнозування кількості звернень, попри помірну середню відносну точність (середня абсолютна помилка MAE = 32,6 %), правильно відтворює ключові часові патерни, зокрема – тижневу сезонність, що є придатним для операційного планування. Продемонстровано візуалізацію результатів в інтерфейсі, враховуючи інтерактивну карту з кластеризацією та спеціалізовані панелі для відображення прогнозів. Визначено, що запропонований гібридний підхід відкриває перспективи для побудови проактивних систем управління містом, де якість прогнозування та якість вхідних даних посилюють одне одного.

Ключові слова: інженерія ознак; геопросторовий аналіз; модель LightGBM; аналіз сингулярного спектра; аналіз часових рядів; подієво-орієнтована архітектура.

Вступ / Introduction

Проблему ефективного управління міською інфраструктурою в межах концепції "розумних міст" (англ. *Smart Cities*) активно досліджують у світовій науковій літературі. Значна частина наявних рішень, таких як муніципальні гарячі лінії чи портали звернень, зосереджена на розробленні систем для збору та оброблення скарг громадян. Використання інтерактивних карт, форм звітності та систем відстеження статусу заявок дає змогу підвищити прозорість роботи міських служб. Однак, більшість наявних рішень працюють за реактивним принципом, не використовуючи нагромаджені дані для

прогнозування майбутніх інцидентів. Окрім цього, вони часто характеризуються низькою залученістю громадян до вирішення міських проблем, що призводить до неповних і нерегулярних даних про їхню активність.

Актуальність дослідження полягає в необхідності розроблення такої інтелектуальної системи, яка давала б змогу не тільки збирати та візуалізувати дані про міські проблеми, але й використовувати методи машинного навчання моделей для прогнозування їх динаміки та інтегрувати елементи гейміфікації для підвищення активності громадян, цим самим економлячи бюджетні кошти та вдосконалюючи роботу комунальних служб.

Інформація про авторів:

Васютук Остап Степанович, магістр, кафедра програмного забезпечення. Email: ostap.vasiutyk.mpzip.2024@lpnu.ua;

<https://orcid.org/0009-0006-7688-1168>

Сердюк Павло Віталійович, канд. техн. наук, доцент, кафедра програмного забезпечення. Email: pavlo.v.serdyuk@lpnu.ua;

<https://orcid.org/0000-0002-2677-3170>

Цитування за ДСТУ: Васютук О. С., Сердюк П. В. Інтелектуальна геоінформаційна система для прогнозування міських проблем із застосуванням моделей машинного навчання та гейміфікації. Науковий вісник НЛТУ України. 2025, т. 35, № 5. С. 87–96.

Citation APA: Vasiutyk, O. S., & Serdyuk, P. V. (2025). Intelligent geographic information system for urban problem prediction using machine learning and gamification. *Scientific Bulletin of UNFU*, 35(5), 87–96. <https://doi.org/10.36930/40350510>

Основним напрямом дослідження є розроблення комплексної системи моніторингу та прогнозування тривалості виконання заявок і кількості звернень громадян із застосуванням машинного навчання моделей та елементів гейміфікації для підвищення їх активності.

Об'єкт дослідження – управління міськими інцидентами на підставі інтелектуального аналізу даних.

Предмет дослідження – методи машинного навчання моделей для прогнозування тривалості виконання заявки та кількості звернень громадян, що дасть змогу побудувати геоінформаційну систему та елементи гейміфікації для підвищення громадської активності у зборі даних.

Мета роботи – розробити інтелектуальну геоінформаційну систему для прогнозування міських проблем із застосуванням методів машинного навчання моделей та досягнення якості вхідних даних засобами гейміфікації, що дасть змогу підвищити ефективність управління міською інфраструктурою через прогнозування тривалості вирішення інцидентів і зменшення їхньої кількості.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- проаналізувати офіційні дані міських звернень та розробити на їхній основі прогностичні моделі, що дасть змогу кількісно оцінювати майбутнє навантаження на комунальні служби та тривалість виконання заявок;
- розробити архітектуру інформаційної системи, яка забезпечить інтеграцію розроблених моделей із геоінформаційним сервісом для візуалізації проблемних зон міста з аналітичною панеллю для моніторингу ключових показників;
- інтегрувати в інформаційну систему механізми гейміфікації (систему балів, рейтингів, досягнень), що сприятиме підвищенню громадської активності та якості збору первинних даних про міські інциденти;
- провести тестування інформаційної системи на історичних даних для валідації прогностичних моделей, що дасть змогу визначити їхню точність та практичну придатність для ефективного управління.

Аналіз останніх досліджень та публікацій. Сучасні підходи до управління міською інфраструктурою дедалі частіше спираються на аналіз великих даних і методи машинного навчання моделей для ефективного управління інфраструктурою. Значна увага приділяється розробленню прогностичних систем, які, використовуючи методи машинного навчання моделей, аналізують історичні дані для прогнозування міських процесів [3, 8]. Паралельно розвиваються платформи для залучення громадян за моделлю їхнього звітування [2]. Проте, ключовою проблемою, яку підтверджують аналітичні роботи [2, 10], є те, що ці два напрями існують переважно ізольовано, оскільки прогностичні системи залежать від офіційних даних, ігноруючи їхню якість, а платформи залучення громадян обмежуються збором даних, не використовуючи їхнього аналітичного потенціалу.

У роботі [14] продемонстровано застосування сучасних методів машинного навчання моделей для вирішення міських проблем. У ній детально описано, як ансамблеві методи, зокрема – метод LightGBM (англ. *Light Gradient Boosted Machine*), можна використати для класифікації рівня забруднення повітря. Однією з ключових переваг, зазначених у роботі, є проведення порівняльного аналізу кількох потужних моделей (XGBoost, LightGBM, CatBoost) та використання методу SHAP (англ. *Shapley Additive Explanations* – метод поясню-

вального ШП) для інтерпретації результатів, що забезпечує високу прозорість моделі. Однак, автори вказують на істотні обмеження, оскільки фінальна модель показала значний розрив у продуктивності для різних класів. Окрім цього, основна увага роботи зосереджена на завданні класифікації, а не на складніших для операційного планування завдань регресії та аналізу часових рядів.

Потрібну аналітичну спрямованість має і робота [18]. У ній детально описано процес розроблення гібридної моделі для прогнозування обсягів пасажиропотоку в метро. Зокрема, автори зазначають, що модель LightGBM забезпечує високу швидкість навчання та точність завдяки своїй стратегії росту дерев (leaf-wise), яка дає змогу швидше досягати оптимального результату на великих масивах даних. Особливо варто відзначити, що автори наголошують на критичній важливості інженерії ознак, оскільки виокремлення специфічних часових характеристик дає змогу моделі краще вловлювати приховані патерни. Важливість цього підходу підтверджується і в роботі [16], де автори стверджують, що саме погодні та часові характеристики (день тижня, тиждень року) є ключовими для точності прогнозу. Проте, основна увага обох робіт є суто аналітичним і не розглядає інженерну реалізацію у вигляді комплексної програмної системи. Отже, бракує розуміння, як інтегрувати подібні прогностичні моделі з інструментами активного збору даних.

У роботі [3] автори розробляють комплексну хмарну систему для прогнозування завантаженості парковок. Їхня модель складається з мобільного застосунку та прогностичного сервісу на підставі нейронної мережі LSTM (англ. *Long Short-Term Memory* – довга короткочасна пам'ять), яку автори навчали на наборі даних із 4,5 млн записів. Дослідження довело, що запропонований метод є ефективним для моделювання часових залежностей у даних і значно перевершує класичні підходи, такі як ARIMA (англ. *Autoregressive Integrated Moving Average* – авторегресійне інтегроване ковзне середнє), досягаючи високої точності прогнозу (MAE = 0,92). Водночас, недоліком їхнього підходу є повна залежність від офіційних даних, оскільки архітектура системи побудована на пасивному, однобічному потоці даних.

У роботі [8] автори вирішують подібне завдання спрогнозувати кількість викликів швидкої допомоги. Їхній підхід полягає у розробленні стекінг-ансамблю, який поєднує чотири базові моделі та дає змогу використовувати переваги різних алгоритмів, що підвищує стійкість і точність прогнозу. Однією з ключових переваг їхнього підходу є глибока інженерія ознак, яка містить просторове агрегування викликів за допомогою методу геосітки, що дає змогу моделі виявляти географічні патерни, які є невидимими на рівні окремих інцидентів. Отримані ними результати показали, що їхній метод забезпечує значно покращену продуктивність порівняно з іншими моделями ($R^2 = 0,9954$). Однак, незважаючи на високу точність, їхня модель повністю залежить від офіційних історичних даних. Окрім цього, автори вирішують тільки одне з двох ключових завдань операційного планування – прогнозування кількості звернень, не прогнозуючи час їхнього виконання.

Робота [2] є фундаментальним оглядом, де автори аналізують сучасні програмні додатки для управління містом. Ключовим у цьому дослідженні є класифікація наявних систем, де найпоширенішою категорією вияви-

лася така, як "звітування за участі громадян". Проте, незважаючи на велику кількість розроблених прототипів, самі автори роблять висновок про необхідність поглиблювати механізми для гарантування якості даних. Це вказує на невирішене інженерне завдання щодо того, як саме перевіряти та підвищувати достовірність даних, отриманих від громадян.

У дослідженні [10] аналізують цю саму проблему даних, але з більш технічного боку. У ньому подано фундаментальний аналіз, що систематизує стан та проблеми міських даних. Особливо варто відзначити, що автори прямо вказують на існування великого розриву між збором та використанням даних, що підтверджує актуальність розроблення систем, які його долали б. Однак, у цій роботі немає детального розгляду конкретних інженерних рішень, де гейміфікація могла б виступити як ефективний інструмент для підвищення повноти та достовірності даних.

Окремо варто відзначити роботу [13], яка фокусується на систематичному огляді застосування гейміфікації для підвищення громадської активності. Ключовим висновком авторів є те, що ефективна гейміфікація виходить за рамки стандартних механік (балів, рейтингів) і вимагає розроблення продуманого досвіду, орієнтованого на весь життєвий цикл користувача. Отже, гейміфікація може виступати не просто як спосіб заохочення, а як складний інженерний інструмент для цілеспрямованого збирання та покращення якості даних.

Отже, внаслідок проведеного аналізу останніх досліджень та публікацій можна зробити висновок, що залишається актуальною проблема розроблення програмного забезпечення, яке забезпечувало б комплексне прогнозування міських проблем та одночасно вирішувало б проблему якості даних через інтегрований модуль гейміфікації.

Матеріали та методи дослідження. Досягнення мети дослідження потребувало застосування специфічного інструментарію та методів. Для розроблення серверної частини використано платформу ASP.NET Core, яка забезпечує побудову логіки у вигляді REST API та інтеграцію розроблених моделей машинного навчання. Для розроблення інтерактивного користувацького інтерфейсу використано бібліотеку React з мовою TypeScript. Для побудови прогностичних моделей тривалості виконання заявки та кількості звернень застосовано ансамблевий метод LightGBM та метод аналізу часових рядів SSA відповідно, реалізовані за допомогою бібліотеки ML.NET [9]. Окрім цього, застосовано метрики коефіцієнта детермінації (R^2) та середньої абсолютної помилки (MAE) [15], щоб оцінити ефективність розроблених моделей на реальних історичних даних з офіційного набору даних міста Львова "Звернення на гарячу лінію міста Львова" [6]. Для забезпечення інтерактивної візуалізації геоданих використано бібліотеку Leaflet.

Результати дослідження та їх обговорення / Research results and their discussion

Аналіз останніх досліджень та публікацій виявляє три основні, але переважно ізольовані, напрями розвитку міських інтелектуальних систем. По-перше, суто аналітичні дослідження демонструють потужність окремих алгоритмів, але не є комплексними інженерними рішеннями. По-друге, наявні прогностичні системи покладаються винятково на офіційні дані, ігноруючи про-

блему їхньої якості. По-третє, платформи залучення не використовують свого аналітичного потенціалу для прогнозування. Тому ключовою невирішеною проблемою залишається розроблення єдиної інтегрованої геоінформаційної системи, яка ефективно поєднувала б ці три підходи.

Інтелектуальна геоінформаційна система (ГІС) – це система, яка поєднує функції звичайної геоінформаційної системи (збір, зберігання, аналіз та візуалізація просторових даних) з елементами штучного інтелекту, що дає змогу їй приймати більш складні рішення, аналізувати великі обсяги даних і надавати рекомендації. Така система, наприклад, може використовувати алгоритми машинного навчання для прогнозування, автоматично виявлення закономірностей та оптимізації процесів.

Розроблення та валідація прогностичних моделей. Ключовим компонентом розробленої інтелектуальної геоінформаційної системи є модуль прогнозування на підставі моделей машинного навчання, реалізований за допомогою фреймворку ML.NET. Цей модуль вирішує два основні завдання – прогнозування тривалості виконання заявки та прогнозування кількості майбутніх звернень громадян, використовуючи відкритий набір даних "Звернення на Гарячу лінію міста Львова" за 2022-2024 роки [6].

Прогнозування тривалості виконання заявки. Для кількісного оцінювання очікуваної тривалості виконання заявки розроблено регресійну модель на підставі ансамблевого методу LightGBM, що є високоефективною реалізацією градієнтного бустингу.

Підготовку даних для цієї моделі проведено у кілька етапів. Початковий набір даних у 362 869 записів очищено від дублікатів (до 355 867 записів). Далі, специфічно для цієї регресійної задачі, вибірку відфільтровано за тривалістю виконання (діапазон 60-17500 хв), що зменшило її до 284 983 записів. Фінальний навчальний набір даних сформовано шляхом відбору тільки тих категорій, що мали понад 5000 прикладів, і він становив 118 708 записів.

Важливим етапом є інженерія ознак (англ. *Feature Engineering*) – процес застосування знань предметної галузі для створення ознак, які забезпечують роботу алгоритмів машинного навчання, де згенеровано та підготовлено різноманітні характеристики. Додано циклічні часові ознаки (синус/косинус перетворення для дня тижня, години доби, місяця року) для кращого врахування їхньої періодичності. Враховано геопросторові ознаки, такі як широта, довгота та відстань до центру міста. Інтегровано погодні умови на момент подання заявки, зокрема – температуру, опади, глибину снігу, швидкість вітру та атмосферний тиск. Разом з цими згенерованими ознаками використано також початкові характеристики, такі як рік, тип звернення, категорія, група категорії заявки, виконавець, джерело, статус, район та сезон.

Усі категоріальні поля перетворено у числовий формат методом One-Hot Encoding. Числові ознаки нормалізовано методом Min-Max для приведення до спільного діапазону [0, 1]. Підготовлені ознаки об'єднано у єдиний вектор ознак Features, а цільовою змінною обрано тривалість виконання заявки у хвилинах. Модель навчали на 80 % підготовлених даних, використовуючи 20 % для тестування. Результати оцінювання ефективності на тестовій вибірці продемонстровано у табл. 1.

Табл. 1. Результати валідації моделі прогнозування тривалості виконання заявки (LightGBM) / Validation results of the model for predicting the application execution time (LightGBM)

Метрика	Значення
R-Squared	0,640
MAE, хв	1707,9
RMSE, хв	2676,7

Отримане значення R-Squared (0,640) свідчить, що розроблена модель пояснює 64 % варіативності реальної тривалості виконання заявок у тестовій вибірці. Середня абсолютна помилка (MAE), що становить 1707,9 хв (приблизно 28,5 год), вказує на задовільну точність для первинного оцінювання очікуваних термінів, враховуючи значний природний розкид у реальному часі виконання заявок (від годин до тижнів). Цей результат демонструє потенціал моделі для надання орієнтовної інформації користувачам та підтримки операційного планування.

Для глибшого розуміння логіки моделі та визначення факторів, що найбільше впливають на прогноз, проведено аналіз важливості ознак PFI (англ. *Permutation Feature Importance*). Метод PFI вимірює, наскільки зростає середня абсолютна помилка моделі (MAE), якщо випадково перемішати одну конкретну ознаку. Результати PFI агреговано за ознаками для оцінювання їх сумарного впливу (табл. 2).

Табл. 2. Агрегована важливість ознак (PFI) за типом / Aggregate feature importance (PFI) by type

Ознака	MAE, хв
Виконавець	1244,17
Категорія	511,99
Максимальна температура	94,84
День тижня (sin)	93,63
Середня температура	72,29

Унаслідок аналізу сумарного впливу ознак на точність прогнозу виявлено чітку перевагу операційних факторів над зовнішніми. Якщо сумарний внесок "Назва виконавця" та "Категорія" становить 1244,17 хв та 511,99 хв до MAE відповідно, то релевантні інженерні ознаки, такі як "Максимальна температура" (94,84 хв) та "День тижня (sin)" (93,63 хв), хоч і є статистично значущими, але демонструють менший вплив на остаточний прогноз.

Прогнозування кількості заявок. Для прогнозування щоденної кількості нових звернень громадян обрано архітектуру з окремими, незалежними моделями для кожного району Львова, що дає змогу краще враховувати унікальні локальні патерни динаміки звернень. Як алгоритм прогнозування часових рядів використано SSA (англ. *Singular Spectrum Analysis* – метод аналізу сингулярного спектра), що ефективно розкладає часовий ряд на компоненти тренду, сезонності та шуму без строгих припущень щодо стаціонарності даних.

Навчання моделей SSA проведено на даних після базового очищення (видалення неповних записів, некоректних дат, дублікатів тощо). Унаслідок цього залишилось 194 445 унікальних чистих записів, до яких не застосовувалася специфічна фільтрація, використана для методу LightGBM. Ключовим етапом підготовки стало формування неперервного часового ряду для кожного району міста: спочатку дані агреговано для підрахунку щоденної кількості звернень, а потім розроблено повний діапазон дат, де дні без звернень заповнено нульовим значенням. Цей крок є вимогою для коректної роботи алгоритму SSA.

Моделі SSA навчали для кожного району міста індивідуально, прогнозуючи кількість заявок на найближчі 14 днів. Основні параметри SSA встановлено як `windowSize = 60` (розмір вікна для аналізу структури ряду) та `seriesLength = min(365, кількість днів у ряді)` (довжина ряду для побудови траєкторної матриці). Результати валідації точності отриманих моделей на тестових даних для 14 днів подано у табл. 3.

Табл. 3. Результати валідації моделей прогнозування кількості заявок (SSA) за районами міста / Validation results of request volume forecasting models (SSA) by city districts

Заявок у день Район	MAE	RMSE	Середня кількість	Відносна MAE, %
Галицький	6,123	7,748	20,74	29,5
Залізничний	12,554	14,286	29,21	43,0
Личаківський	9,810	13,010	25,78	38,1
Сихівський	9,009	11,332	30,14	29,9
Франківський	11,371	14,012	33,87	33,6
Шевченківський	8,996	11,776	37,55	24,0
Середнє	9,644	12,027	29,55	32,6

Примітка: кількість днів у ряді – 1097, Галицький – 1096.

Аналіз результатів валідації моделей SSA показує різну точність прогнозування для окремих районів міста. Середня абсолютна помилка (MAE) варіюється в діапазоні від 6,1 (Галицький район) до 12,6 (Залізничний район) заявок на день. У середньому по місту цей показник становить 9,6 заявок на день. Відносна помилка (MAE відносно середньої історичної кількості заявок) становить близько 32,6 % у середньому, що свідчить про помірну точність для прогнозування точного навантаження на конкретний день, особливо для районів міста з вищою похибкою, ніж Залізничний (~43 %).

Водночас, візуальний аналіз прогнозів, згенерованих моделями SSA, демонструє їхню здатність відтворювати ключові часові патерни у даних. На прикладі Сихівського району міста (рис. 1) видно, що модель правильно вловлює тижневу сезонність – прогнозує характерне зниження кількості звернень громадян у вихідні дні, що відповідає реальній динаміці.

Прогноз на найближчі 14 днів для Сихівський район:

01.01.2025: 38 заявок
02.01.2025: 36 заявок
03.01.2025: 34 заявок
04.01.2025: 23 заявок
05.01.2025: 16 заявок
06.01.2025: 25 заявок
07.01.2025: 32 заявок
08.01.2025: 28 заявок
09.01.2025: 26 заявок
10.01.2025: 26 заявок
11.01.2025: 17 заявок
12.01.2025: 15 заявок
13.01.2025: 30 заявок
14.01.2025: 40 заявок

--- Оцінка якості моделі ---

Середня абсолютна помилка (MAE): 9,009

Корінь із середньоквадратичної помилки (RMSE): 11,332

Рис. 1. Прогноз кількості звернень для Сихівського району міста / Forecast of the number of appeals for Sykhiv district

Отже, розроблені моделі SSA, хоча й мають обмеження щодо точності денних прогнозів, забезпечують цінну інформацію про основні тенденції та сезонні коливання потоку звернень громадян. Ця здатність робить їх придатними для використання як інструмент підтримки прийняття рішень у завданнях середньотерміно-

вого планування роботи комунальних служб, таких як тижневий розподіл ресурсів чи ідентифікація очікуваних періодів пікового навантаження.

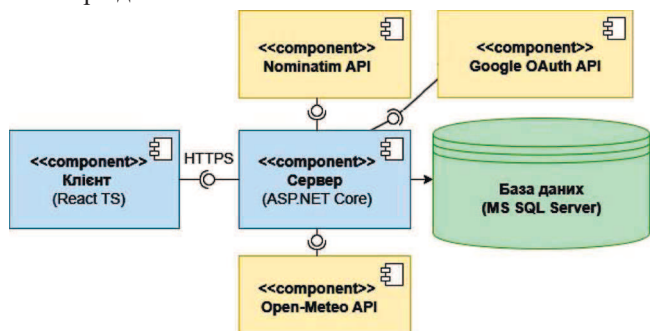


Рис. 2. Загальна архітектура розробленої інтелектуальної геоінформаційної системи / General architecture of the developed intelligent geoinformation system

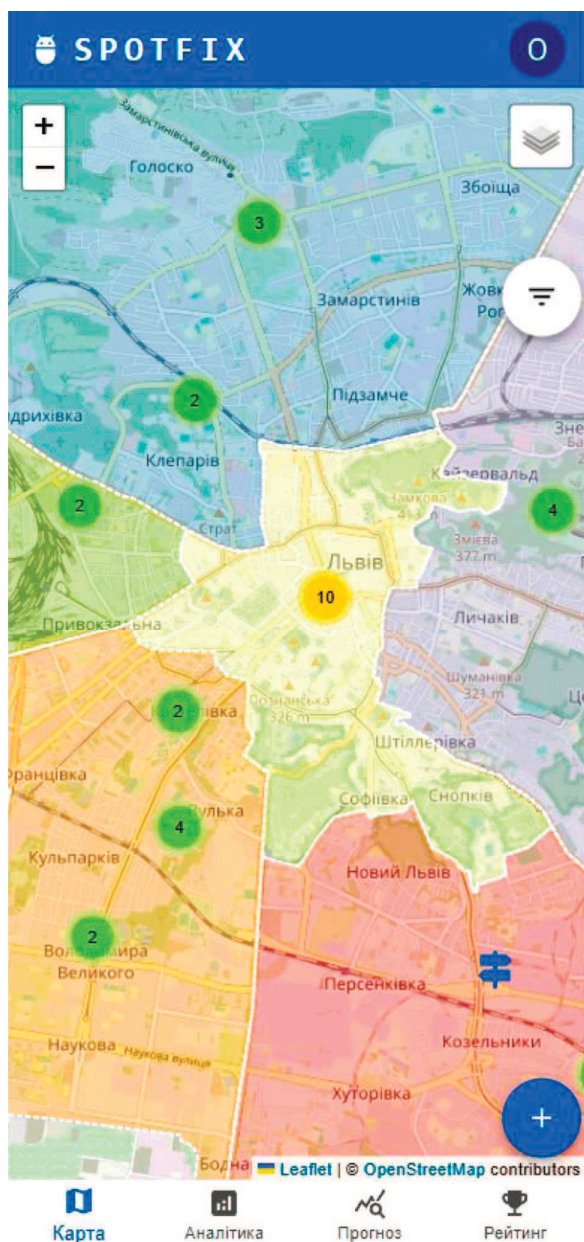


Рис. 3. Головний інтерфейс геоінформаційної системи з візуалізацією міських проблем на карті Львова / Main interface of the geoinformation system with visualization of urban problems on the map of Lviv

Архітектура інтегрованої системи та візуалізація результатів. Результатом дослідження є інтелектуальна геоінформаційна система, що функціонує за кліє-

нт-серверною архітектурою, яку подано на рис. 2. В її основу покладено серверний компонент (ASP.NET Core), клієнтський веб-інтерфейс (React) та базу даних (MS SQL Server). Серверна частина відповідає за оброблення даних, інтеграцію прогностичних моделей (ML.NET), а також використовує зовнішні сервіси: "Google OAuth API" для автентифікації, "Nominatim API" для геокодингу та "Open-Meteo API" для збагачення моделей погодними даними. Ця архітектура забезпечує надійну основу для реалізації всього функціоналу системи.

Клієнтський застосунок надає користувачам інтерактивну карту для візуалізації міських інцидентів (рис. 3). На карті відображено географічні межі районів міста та маркери зареєстрованих проблем, які автоматично групуються у кластери за зміни масштабу. Система дає змогу фільтрувати інциденти за категоріями та переглядати детальну інформацію (категорія, автор, дата, адреса) під час вибору конкретного маркера. Така візуалізація та інтерактивність забезпечують користувачам наочне просторове уявлення про актуальні проблеми міста.

Окрім цього, для моніторингу ключових показників реалізовано аналітичну панель (рис. 4). Вона надає користувачам агреговану статистику у вигляді ключових показників та набору графіків, що ілюструють динаміку звернень громадян та їх розподіл за категоріями і районами, даючи змогу відстежувати поточну ситуацію у місті.

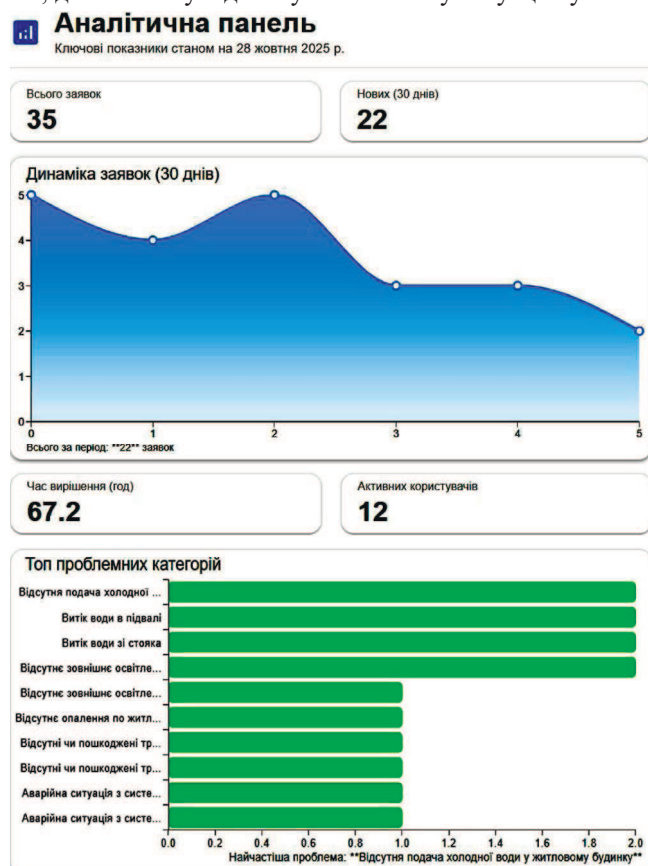


Рис. 4. Інтерфейс аналітичної панелі для моніторингу / Interface of the analytical dashboard for monitoring

Ключовим результатом є ефективна інтеграція прогностичних моделей. Процес використання моделі для отримання прогнозу тривалості виконання заявки детально ілюструє рис. 5. Ініційований користувачем запит з вебзастосунку надходить до PredictionController (API), який делегує його ExecutionTimeForecastService. Цей сервіс виконує критичний етап збагачення даних

необхідними ознаками, після чого передає підготовлений запит до механізму PredictionEnginePool. Його використання забезпечує високопродуктивний та потокобезпечний доступ до завантаженої моделі LightGBM за одночасних запитів від великої кількості користувачів. Отримавши розрахований прогноз від пулу, сервіс повертає його через контролер до вебзастосунку, який ві-

дображає результат користувачу, реалізуючи завдання надання орієнтовного оцінювання термінів.

Результатом роботи цього інтегрованого механізму є прогноз очікуваної тривалості виконання заявки у хвилинах, який відображається користувачу безпосередньо в інтерфейсі (рис. 6) після вибору категорії заявки та району міста.

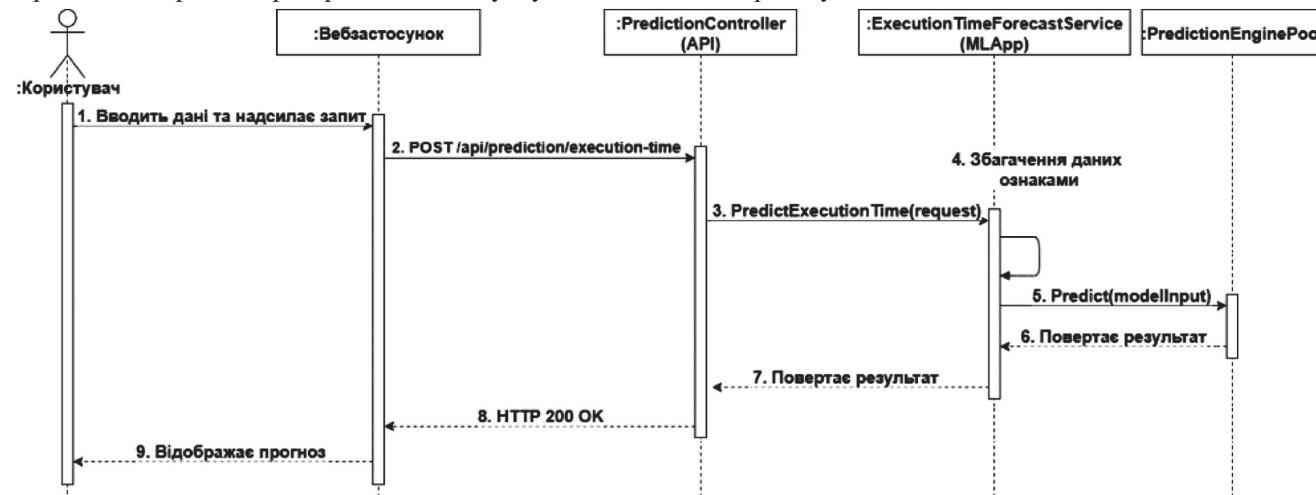


Рис. 5. Діаграма послідовності процесу отримання прогнозу тривалості виконання заявки / Sequence diagram of the process for obtaining the request execution time forecast

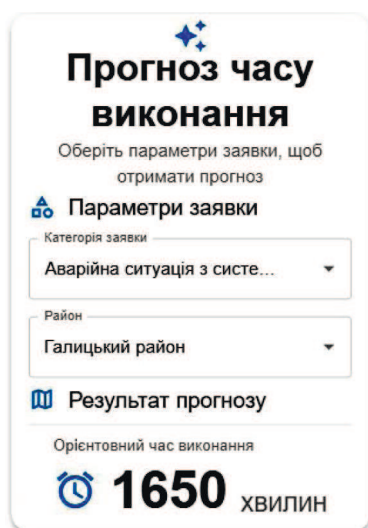


Рис. 6. Відображення прогнозованої тривалості виконання заявки / Display of the forecasted request execution time

Для візуалізації прогнозу кількості майбутніх звернень громадян розроблено спеціалізований інтерфейс (рис. 7). Він дає змогу динамічно відображати графік прогнозу SSA для обраних районів міста та часового діапазону, надаючи інструмент для аналізу очікуваного та виявлення сезонних трендів пікового навантаження.

Для реалізації прогнозу кількості майбутніх звернень громадян застосовано багатомодельний підхід, де кожен район міста має власну SSA-модель. Для ефективного керування цими окремими моделями розроблено спеціфічну архітектуру їх завантаження через механізм PredictionEnginePool (рис. 8). Це дає змогу динамічно отримувати потрібну модель для конкретного району міста, забезпечуючи швидку відповідь інформаційної системи.

Інтеграція модуля гейміфікації. Для підвищення громадської активності та якості збору первинних даних, що є поширеною проблемою муніципальних систем, розроблено та інтегровано в геоінформаційну сис-

тему спеціалізований програмний модуль гейміфікації. Цей модуль використовує ігрові елементи у неігровому контексті та дає змогу мотивувати користувачів через систему винагород та визнання їхнього внеску.

Основу модуля становлять такі взаємопов'язані ігрові механіки [8]:

- система балів досвіду та рівнів, яка забезпечує базову винагороду за активність. Користувачі отримують певну кількість балів за виконання цільових дій, зокрема – 10 балів за розроблення заявки. Нагромадження балів призводить до підвищення рівня, що слугує візуальним індикатором внеску користувача;
- система досягнень, яка доповнює систему рівнів, надаючи автоматично віртуальні нагороди за виконання критеріїв, таких як "Новачок" за першу заявку або "Картограф" за звіти з усіх районів міста. Цей механізм слугує інструментом визнання внеску користувача;
- система завдань, що вводить елементи цілеспрямованості, пропонуючи коротко- або довготермінові завдання, наприклад, повідомити про певну кількість проблем конкретного типу. Успішне виконання завдань винагороджується додатковими балами, що дає змогу стимулювати активність звітування громадян у пріоритетних напрямках для муніципальних служб;
- рейтинг лідерів, який відображає перелік користувачів за кількістю балів, впроваджуючи елемент змагання та соціального визнання.

Для забезпечення надійності та захисту від маніпуляцій ігровими показниками, логіку оброблення ігрових подій (нарахування балів, перевірка умов досягнень та завдань) реалізовано на серверній частині застосунку. Робота модуля ґрунтується на подієво-орієнтованому принципі. Це означає, що система реагує на конкретні події, які виникають при діях користувача, наприклад, розроблення нової заявки, що ініціюються іншими сервісами. Відповідний серверний компонент відстежує ці події та запускає логіку нарахування балів чи перевірки досягнень, не маючи прямої залежності від сервісу, що згенерував подію. Такий підхід сприяє гнучкості та масштабованості архітектури. Історію ігрових дій (транзак-

ції балів, прогрес завдань, досягнення) зберігають у базі даних для аудиту та відображення.

Візуальне подання елементів гейміфікації реалізовано у клієнтському застосунку у вигляді розділу "Ігровий Центр" (рис. 9). Цей інтерфейс дає змогу користувачам відстежувати прогрес, переглядати досягнення, завдання та рейтинг.

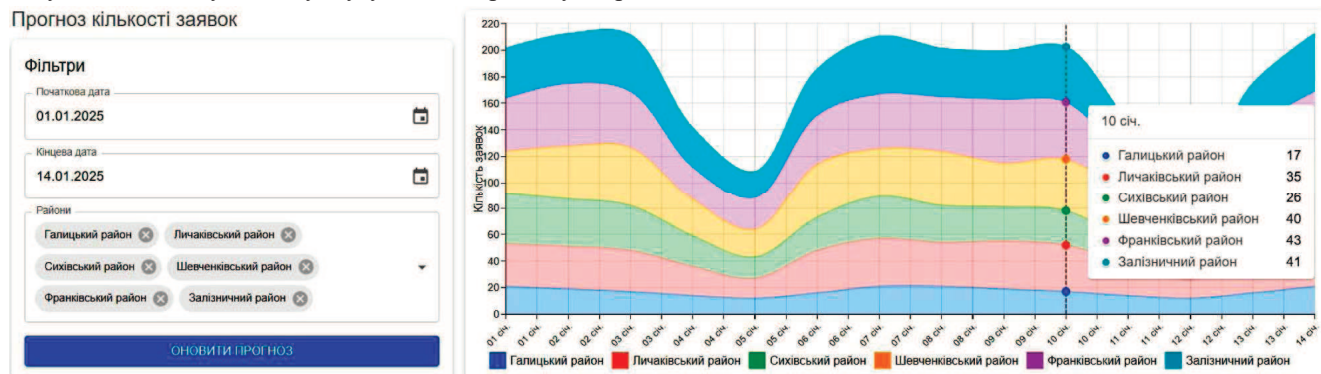


Рис. 7. Інтерфейс візуалізації прогнозу кількості звернень громадян / Interface for visualizing the citizen request volume forecast

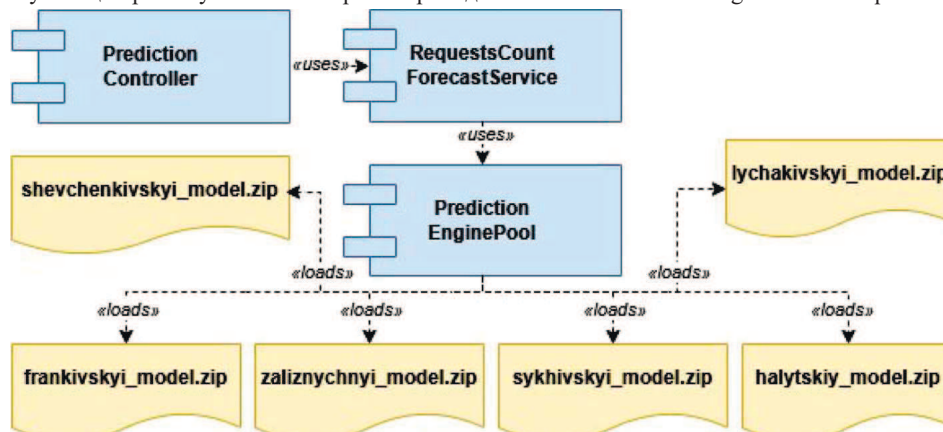


Рис. 8. Архітектура модуля прогнозування кількості заявок / Architecture of the request volume forecasting module

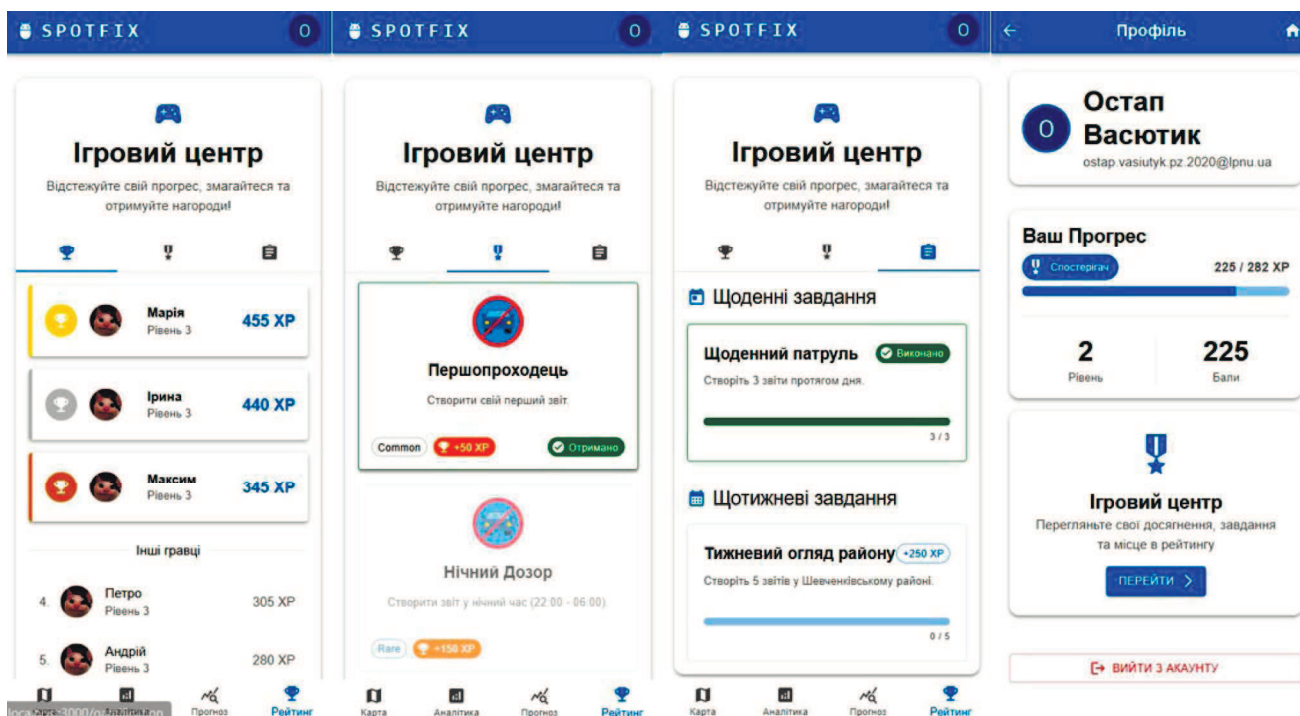


Рис. 9. Візуальне подання елементів гейміфікації у клієнтському застосунку / Visual representation of gamification elements in the client application

Упровадження описаних таких ігрових механізмів, як гейміфікації, спрямоване на підвищення мотивації громадян до використання інформаційної системи. Це, водночас, сприятиме збільшенню частоти подання звітів про міські проблеми та покращенню якості наданої інформації. Зокрема, механізм гейміфікації завдань дає

змогу стимулювати збір інформації про конкретні типи проблем або у певних географічних зонах, де недостатньо даних для навчання прогностичних моделей, шляхом розроблення відповідних цільових завдань, що сприятиме формуванню збалансованого та репрезентативного набору даних. Отримання більш повних і якіс-

них даних є критично важливим, оскільки вони безпосередньо вплинуть на точність роботи розроблених прогностичних моделей та, як наслідок, на ефективність реагування міських служб.

Обговорення результатів дослідження. У роботі [15] автори досліджують проблему прогнозування тривалості вирішення інцидентів, порівнюючи три алгоритми машинного навчання моделей: Decision Tree, Random Forest та XGBoost. Результати експерименту показали, що найвищу точність прогнозу забезпечує модель випадкового лісу Random Forest (Accuracy = 70,04 %, MAE = 24,17, RMSE = 41,45). Перевагами цього підходу автори також називають здатність працювати з категоріальними та числовими змінними без складного попереднього їх оброблення. Водночас, серед недоліків методу виділяють залежність від історичних даних, обмежену гнучкість та, що є ключовим для нашого дослідження, низьку ефективність за наявності рідкісних категорій. У нашому дослідженні, натомість, застосовуємо модель LightGBM для подолання саме цих обмежень. Його технічні переваги, такі як градієнтний бустинг для роботи зі складними даними, histogram-based підхід для швидкого оброблення та кращої роботи з рідкісними категоріями, а також вища гнучкість, дають змогу отримати точніший та стабільніший прогноз.

У роботі [1] автори розглядають проблему прогнозування тривалості виконання муніципальних заявок, представляючи MuST²-Learn-архітектуру глибокого навчання на підставі підходу багатовидового навчання (multi-view learning). Перевагою їхнього підходу є одночасне моделювання чотирьох типів кореляцій у даних: просторових, часових, міжтипів і внутрішньотипових. Для врахування унікального контексту кожної заявки вони впровадили intra-type variation encoder, що аналізує текстовий опис за допомогою LLM (LLaMA-3 8B). Такий підхід продемонстрував значне зниження помилки прогнозування (до 84,6 % за метрикою MAPE). Хоча ця комплексна архітектура доводить високу ефективність, її недоліками є висока обчислювальна складність та необхідність великої кількості даних для тренування. Наше дослідження, на відміну від цього суто алгоритмічного підходу, пропонує системне рішення. Ключовою відмінністю є вирішення проблеми якості вхідних даних через гейміфікацію та збагачення їх зовнішніми факторами, наприклад погодними умовами. Такий підхід створює активний зворотний зв'язок, де залученість громадян безпосередньо підвищує точність прогнозів.

У роботі [11] автори описують процес розроблення мобільного застосунку для збирання даних про міські проблеми в Чіангмай, Таїланд. Система поєднує геоінформаційну складову із комплексною гейміфікацією, використовуючи ігрові механіки із гри Pokemon GO (інтерактивний "збір проблем", нарахування балів та система рівнів). Такий підхід довів свою ефективність, залучивши понад 500 користувачів за два місяці. Водночас, ключовим недоліком системи є її винятково реактивний характер, оскільки вона тільки збирає та візуалізує наявні проблеми, не маючи функціоналу прогнозування. Саме тому наше дослідження доповнює цей підхід, впроваджуючи прогностичні моделі та аналітичну панель для переходу до проактивного управління міськими проблемами.

У роботі [4] автори наведено результати розроблення цифрового двійника міста Патри – комплексної сис-

теми, яка поєднує конвеєр для оброблення даних, модель машинного навчання та інтерактивну візуалізацію у вебзастосунку. Для аналізу та прогнозування вони використали Random Forest, досягнувши точності 79,04 % у завданні багатокласової класифікації інфраструктурних інцидентів. Хоча перевагою роботи є її комплексна реалізація, ключовим недоліком є основна увага на відносно простому завданні класифікації, а не на прогнозуванні динамічних показників. Наше дослідження, на відміну від цього, розв'язує складніші завдання регресійного прогнозування тривалості виконання заявки за допомогою моделі LightGBM та прогнозування кількості нових заявок за допомогою часових рядів. Це дає змогу перейти від простої категоризації інцидентів до проактивного планування ресурсів.

У роботі [12] автори експериментально доводять ефективність гейміфікації для підвищення залученості громадян, показавши, що гейміфікована версія застосунку забезпечила на 45 % більше виконаних місій та на 68 % більше завантажених фотографій. Водночас, їхнє дослідження виявляє й істотні недоліки цього підходу. Зокрема, їхні власні дані показали, що гейміфікований додаток створює вище когнітивне навантаження на користувача, а короткотривалість експерименту в один тиждень не дає змоги оцінити довготермінову мотивацію та вилучити ефект новизни. Окрім цього, автори згадують і про етичні ризики використання гейміфікації як інструменту державного примусу.

У роботі [5] автори проводять комплексний аналіз найсучасніших методів прогнозування в урбаністиці – STGNN (англ. *Spatio-Temporal Graph Neural Networks* – просторово-часові графові нейронні мережі). Автори стверджують, що для складних просторово-часових даних, таких як міські мережі, класичні моделі машинного навчання є недостатньо ефективними. Водночас, як впливає з огляду, недоліками STGNN-підходів є їхня висока обчислювальна складність, потреба у великих наборах даних і складність впровадження й інтерпретації. Саме тому в нашому дослідженні зроблено свідомий інженерний вибір на користь моделі LightGBM. Хоча метод STGNN є більш потужнішими, модель LightGBM забезпечує оптимальний баланс між точністю, швидкістю навчання та інтерпретованістю, що є критичним для розроблення практичної та масштабованої муніципальної системи.

У роботі [7] автори аналізують дані викликів швидкої допомоги у США за 2010–2017 рр., розробляючи методи прогнозування їхньої кількості на погодинному рівні. Вони порівняли класичні статистичні моделі з нейронною мережею MLP, провівши навчання на записах за 2010–2016 рр. та тестування на даних 2017 року. Встановлено, що найкращий результат серед класичних методів забезпечило ковзне середнє (MAPE = 27,14 %), тоді як MLP продемонструвала точність (MAPE = 27,31 %) з вищим потенціалом масштабування. Водночас, недоліком їхнього підходу є використання універсального алгоритму MLP, який є менш ефективним на розріджених даних, що ускладнює аналіз складних часових рядів. Наше дослідження розв'язує аналогічну задачу прогнозування кількості, але, на відміну від універсального MLP, застосовує більш спеціалізований інструмент – алгоритм SSA. Він призначений саме для декомпозиції часового ряду та виявлення складних періодичностей, що дає змогу точніше моделювати сезонні та трендові компоненти в даних.

У роботі [19] автори розв'язують задачу прогнозування енергоспоживання електричних автобусів, використовуючи модель LightGBM з гіперпараметрами, налаштованими через оптимізацію методом Байєса. Перевагою їхнього підходу є детальна інженерія ознак, що містить розроблення фізичної моделі тяги, динамічну оцінку пасажиропотоку та аналіз 13 параметрів стилю водіння, які згодом спрощують за допомогою PCA (англ. *Principal Component Analysis* – метод головних компонент). Завдяки цьому вони досягли високої точності прогнозу (помилка MAPE = 3,92 %), а для аналізу впливу ознак використали SHAP (англ. *Shapley Additive Explanations* – метод пояснювального ШП). Водночас, недоліком такого підходу, що базується на глибокому моделюванні фізичних процесів, є його вузька спеціалізація, яка обмежує застосовність до інших типів міських задач. Наше дослідження, на відміну від цього фізично-орієнтованого методу, застосовує більш універсальний, керований даними (data-driven) підхід, що поєднує декомпозицію наявних даних і збагачення їх зовнішніми факторами, забезпечуючи вищу адаптивність системи.

У роботі [17] автори розв'язують задачу прогнозування попиту на швидку допомогу, пропонуючи гібридну модель SSA-ARIMA. Їхній метод використовує SSA (англ. *Singular Spectrum Analysis* – аналіз сингулярного спектра) як крок попереднього оброблення для декомпозиції часового ряду, після чого кожна з компонент (тренд, періодичність, залишки) прогнозується окремою моделлю ARIMA (англ. *Autoregressive Integrated Moving Average* – авторегресійне інтегроване ковзне середнє). Хоча робота доводить високу ефективність SSA як інструменту декомпозиції, недоліком такого гібридного підходу є його складність, оскільки він вимагає побудови та налаштування окремих моделей для кожної компоненти. Наше дослідження, на відміну від цього, використовує SSA як самостійний, інтегрований інструмент прогнозування для завдання визначення кількості заявок. Це дає змогу застосувати більш прямий підхід, уникаючи складності, пов'язаної з побудовою та налаштуванням окремих моделей для кожної з компонент часового ряду.

Отже, проведене обговорення результатів дослідження у контексті аналізованих вище робіт підтверджує доцільність проведення саме цього дослідження, оскільки воно не виявило жодного комплексного рішення, яке одночасно вирішувало б два ключові прогностичні завдання, а саме – оцінювання тривалості та прогнозування кількості заявок, і поєднувало це з механізмами покращення якості вхідних даних. Як було показано, вони фокусуються або на виборі конкретного алгоритму, ігноруючи якість даних [1, 5, 7, 15, 17, 19], або ж розглядають гейміфікацію винятково як реактивний інструмент збору даних, без інтеграції з прогнозуванням [11, 12], тоді як інші [4] обмежуються простими завданнями класифікації. Це підтверджує відсутність інженерного підходу, де залучення громадян є не просто остаточною метою, а засобом для створення активного зворотного зв'язку, що цілеспрямовано підвищує точність прогностичних моделей. Запропоновані в цій роботі результати дослідження спрямовані саме на заповнення цієї прогалини.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – розроблено методіку комплексного прогнозування міських проблем, яка забезпечує узгоджений їхній аналіз завдяки поєднанню регресійної моделі для оцінювання тривалості їхнього вирішення, моделі часових рядів для прогнозування кількості заявок у визначеному часовому проміжку та інтегрованому механізмі гейміфікації, який підвищує якість навчальних даних.

Практична значущість результатів дослідження – результати дослідження реалізовано у вигляді програмного забезпечення, яке можна застосувати для підвищення ефективності управління міською інфраструктурою, що дасть змогу муніципальним службам перейти до проактивного планування та оптимізації ресурсів на підставі прогностичних моделей, якість яких забезпечується через зручний та гейміфікований інструмент для підвищення залученості громадян, що стане корисним у сферах міського планування, управління комунальними послугами та розвитку інструментів громадської участі.

Висновки / Conclusions

Розроблено інтелектуальну геоінформаційну систему для прогнозування міських проблем із застосуванням методів машинного навчання моделей та досягнення якості вхідних даних засобами гейміфікації, що дає змогу підвищити ефективність управління міською інфраструктурою через прогнозування тривалості вирішення інцидентів і зменшення їхньої кількості. За результатами проведеного дослідження можна зробити такі основні висновки.

1. Розроблено та валідовано прогностичні моделі (LightGBM, SSA) на підставі офіційних даних міських звернень громадян Львова, що забезпечило можливість прогнозувати кількість міських проблем та термін їх вирішення. Зокрема, модель LightGBM ($R^2 = 0,640$, MAE = 1707,9 хв) дає змогу прогнозувати тривалість виконання звернень, демонструючи ключовий вплив операційних факторів ("Категорія", "Виконавець"), а архітектура з SSA-моделями (середня відносна MAE = 32,6 %) надає інструмент для планування ресурсів, успішно відтворюючи сезонні шаблони.
2. Розроблено клієнт-серверну архітектуру (ASP.NET Core, React), яка забезпечила успішну інтеграцію прогностичних моделей (через механізм PredictionEnginePool) з геоінформаційним сервісом (Leaflet) та аналітичним інтерфейсом, що дало змогу візуалізувати проблемні зони міста та відображати прогнози для моніторингу в реальному часі.
3. Розроблено та інтегровано подієво-орієнтований модуль гейміфікації, що надало механізм для підвищення громадської активності та цілеспрямованого покращення якості первинних даних, створюючи потенційний зворотний зв'язок для підвищення точності прогностичних моделей.
4. Проведене тестування геоінформаційної системи на історичних даних підтвердило працездатність інтегрованих компонент та практичну придатність запропонованого підходу для підтримки прийняття рішень у сфері управління міськими інцидентами.

References

1. Asif, N., Hong, Z., Ren, S., Zhang, X., Shang, X., & Yuan, Y. (2025). MuST²-Learn: Multi-view spatial-temporal-type learning for heterogeneous municipal service time estimation. *arXiv*. <https://doi.org/10.48550/arXiv.2508.16503>
2. Bastos, D., Fernández-Caballero, A., Pereira, A., & Rocha, N. P. (2022). Smart City Applications to Promote Citizen Participation

- in City Management and Governance: A Systematic Review. *Informatics*, 9(4), article ID 89. <https://doi.org/10.3390/informatics9040089>
3. Canli, H., & Toklu, S. (2021). Deep learning-based mobile application design for smart parking. *IEEE Access*, 9, 61171–61183. <https://doi.org/10.1109/ACCESS.2021.3074887>
 4. Gkontzis, A. F., Kotsiantis, S., Feretzakis, G., & Verykios, V. S. (2024). Enhancing Urban Resilience: Smart City Data Analyses, Forecasts, and Digital Twin Techniques at the Neighborhood Level. *Future Internet*, 16(2), article ID 47. <https://doi.org/10.3390/fi16020047>
 5. Jin, G., Liang, Y., Fang, Y., Huang, J., Zhang, J., & Zheng, Y. (2023). Spatio-temporal graph neural networks for predictive learning in urban computing: A survey. *arXiv*. <https://doi.org/10.48550/arXiv.2303.14483>
 6. Lviv City Council. (2025). *Dataset – Open Data Portal of Lviv*. Lviv Municipal Enterprise "City Institute of Information Technologies". [In Ukrainian]. URL: <https://opendata.city-adm.lviv.ua/dataset>
 7. Martin, R. J., Mousavi, R., & Saydam, C. (2021). Predicting emergency medical service call demand: A modern spatiotemporal machine learning approach. *Operations Research for Health Care*, 28, article ID 100285. <https://doi.org/10.1016/j.orhc.2021.100285>
 8. Megouo, T. G. P., & Pierre, S. (2024). A stacking ensemble machine learning model for emergency call forecasting. *IEEE Access*, 12, 115820–115837. <https://doi.org/10.1109/ACCESS.2024.3445591>
 9. Microsoft. (2025). *ML.NET documentation – Tutorials, API reference*. Microsoft Learn. URL: <https://docs.microsoft.com/uk-ua/dotnet/machine-learning/>
 10. Okojie, J., Ike, P., Idu, J., Nnabueze, S. B., Filani, O., & Ihwughwawve, S. (2023). Predictive analytics models for monitoring smart city emissions and infrastructure risk in urban ESG planning. *International Journal of Multidisciplinary Futuristic Development*, 4(1), 45–57. <https://doi.org/10.54660/IJMF.D.2023.4.1.45-57>
 11. Puritat, K. (2019). A gamified mobile-based approach with web monitoring for a crowdsourcing framework designed for urban problems related smart government: A case study of Chiang Mai, Thailand. *International Journal of Interactive Mobile Technologies* (iJIM), 13(12), 55–66. <https://doi.org/10.3991/ijim.v13i12.10989>
 12. Romano, M., Díaz, P., & Aedo, I. (2022). Gamification-less: May gamification really foster civic participation? A controlled field experiment. *Journal of Ambient Intelligence and Humanized Computing*, 13(9), 4451–4465. <https://doi.org/10.1007/s12652-021-03322-6>
 13. Santos, A., Aquino, F., Furtado, L., Silva, T., & Cruz, L. (2021). Gamification and civic engagement in digital government applications: A review. *Proceedings of the 9th Workshop on Applied Computing in E-Government*, (pp. 167–178). SBC. <https://doi.org/10.5753/wcge.2021.15986>
 14. Toharudin, T., et al. (2023). Boosting algorithm to handle unbalanced classification of PM2.5 concentration levels by observing meteorological parameters in Jakarta-Indonesia using AdaBoost, XGBoost, CatBoost, and LightGBM. *IEEE Access*, 11, 35680–35696. <https://doi.org/10.1109/ACCESS.2023.3265019>
 15. Tong, T.-E., Haw, S.-C., Ng, K.-W., Al-Tarawneh, M., & Tong, G.-K. (2024). Performance evaluation on resolution time prediction using machine learning techniques. *JOIV: International Journal on Informatics Visualization*, 8(2), 583–591. <https://doi.org/10.62527/joiv.8.2.2305>
 16. Wang, B., Wu, P., Chen, Q., & Ni, S. (2021). Prediction and analysis of train passenger load factor of high-speed railway based on LightGBM algorithm. *Journal of Advanced Transportation*, 2021, article ID 9963394. <https://doi.org/10.1155/2021/9963394>
 17. Wang, J., Peng, X., Wu, J., Ding, Y., Ali, B., Luo, Y., Hu, Y., & Zhang, K. (2023). Singular spectrum analysis (SSA) based hybrid models for emergency ambulance demand (EAD) time series forecasting. *IMA Journal of Management Mathematics*, 35(1), 45–64. <https://doi.org/10.1093/imaman/dpad019>
 18. Zhang, Y., Zhu, C., & Wang, Q. (2020). LightGBM-based model for metro passenger volume forecasting. *IET Intelligent Transport Systems*, 14(11), 1815–1823. <https://doi.org/10.1049/iet-its.2020.0396>
 19. Zhao, J., He, J., Wang, J., & Liu, K. (2025). Energy Consumption Prediction for Electric Buses Based on Traction Modeling and LightGBM. *World Electric Vehicle Journal*, 16(3), article ID 159. <https://doi.org/10.3390/wevj16030159>

O. S. Vasiutyk, P. V. Serdyuk

Lviv Polytechnic National University, Lviv, Ukraine

INTELLIGENT GEOGRAPHIC INFORMATION SYSTEM FOR URBAN PROBLEM PREDICTION USING MACHINE LEARNING AND GAMIFICATION

Modern urban management increasingly requires intelligent systems that combine predictive analytics with citizen engagement to enhance municipal service delivery. This study developed and validated an intelligent geographic information system using 362,000 municipal service requests from Lviv, Ukraine (2022–2024), implementing dual predictive models within the ML.NET framework: LightGBM regression for issue resolution time forecasting and district-specific Singular Spectrum Analysis (SSA) for daily request volume prediction. The regression pipeline filtered the dataset to 118,708 records (resolution times: 60–17,500 minutes, categories: >5,000 instances) and incorporated cyclical temporal encodings, geospatial coordinates, and meteorological variables. The LightGBM model achieved $R^2 = 0.640$, explaining 64 % of resolution time variance with MAE = 1,707.9 minutes. Moreover, Permutation Feature Importance analysis identified executor assignment and problem categorization as the dominant contributors to MAE (1,244.17 and 511.99 minutes, respectively), substantially exceeding environmental factors such as maximum temperature (94.84 minutes). The SSA ensemble achieved an average daily MAE of 9.6 requests, corresponding to a 32.6 % relative error; consequently, the models effectively reproduced weekly seasonality patterns, validating their utility for operational resource planning. The system architecture employed ASP.NET Core with PredictionEnginePool for thread-safe, high-throughput model inference, React for the user interface, and Leaflet for interactive mapping and clustering. Critically, the event-driven gamification module – featuring experience points, achievements, task mechanics, and leaderboards – established a feedback mechanism whereby targeted tasks incentivize data collection in underrepresented areas and categories, as a result systematically improving the quality of the training corpus and subsequent model refinement. The benefits of this study include a deployable framework that demonstrates synergy between participatory engagement and predictive accuracy, thereby opening avenues for further exploration of citizen-driven data enhancement in smart city implementations.

Keywords: feature engineering; geospatial analysis; LightGBM model; singular spectrum analysis; time series analysis; event-driven architecture.