doi: 10.32620/oikit.2025.105.19

УДК 004.89

М. О. Кушнарьов, Ю. С. Аверін,
І. В. Шостак

Методи машинного навчання в інтелектуальних системах «розумний будинок»: класифікація, проблеми та перспективи

Національний аерокосмічний університет «Харківський авіаційний інститут»

У статті представлено систематизований огляд методів машинного навчання, що застосовуються в інтелектуальних системах «Розумний будинок». Наведено класифікацію типових задач, пов'язаних із розпізнаванням активності користувачів, прогнозуванням споживання ресурсів, виявленням аномалій у процесі функціонування, персоналізацією побутових сценаріїв, управлінням комфортом та моніторингом стану здоров'я мешканців. Детально проаналізовано ключові класи алгоритмів: класичні методи (k-NN, SVM, дерева рішень, Random Forest), ансамблеві підходи (Gradient Boosting, XGBoost), методи глибокого навчання (CNN, RNN, LSTM), а також легковагові моделі, оптимізовані для периферійних пристроїв (TinyML, TensorFlow Lite). Проведено порівняльний аналіз їхніх характеристик, зокрема точності, ресурсних вимог, здатності до адаптації, інтерпретованості та можливостей інтеграції у реальному часі. Окрему увагу приділено питанням організації даних, сенсорних мереж та архітектурних стратегій створення «розумних будинків». Визначено ключові виклики: забезпечення приватності та кібербезпеки, нестача якісних датасетів, апаратні обмеження пристроїв, проблеми стандартизації та інтероперабельності. Проаналізовано сучасні підходи до їх подолання: методи аугментації й трансферного навчання, технології підвищення конфіденційності диференційна приватність, гомоморфне шифрування, безпечні багатосторонні обчислення), федеративне навчання та edge-архітектури. Значну увагу приділено етичним аспектам застосування ШІ, пов'язаним із упередженістю моделей, справедливістю, прозорістю, підзвітністю та збереженням автономії користувача. Окремо розглянуто проблеми безпеки моделей машинного навчання: атаки змагального типу, отруєння даних, крадіжка моделей, а також відповідні механізми захисту — очищення даних, виявлення аномалій, навчання змагального типу, обфускація та водяні знаки. Висвітлено роль людиноцентричного підходу та Explainable AI у підвищенні довіри користувачів. Окреслено перспективи розвитку галузі, включаючи поширення Edge AI і TinyML, удосконалення алгоритмів федеративного навчання, впровадження мультимодальних систем, розвиток методів навчання з підкріпленням та створення децентралізованих архітектур з локальним самонавчанням. Запропонований огляд формує цілісне уявлення про сучасний стан і тенденції використання МН у «розумному будинку» та визначає практичні шляхи подолання існуючих проблем.
***Ключові слова*:** розумний будинок; машинне навчання; інтелектуальні системи; персоналізація; енергоефективність; виявлення аномалій; федеративне навчання; Edge AI; етика штучного інтелекту; безпека моделей.

Вступ

Мотивація даного дослідження зумовлена необхідністю подолання низки актуальних викликів у сфері створення, розгортання та підтримки функціонування інтелектуальних систем керування «розумний будинок». Серед них можна виділити: недостатню адаптивність традиційних сценаріїв управління до індивідуальних звичок користувача; обмеженість класичних алгоритмів у контексті роботи з великими обсягами гетерогенних даних, зібраних у реальному часі; потребу в забезпеченні енергоефективності та зниженні навантаження на обчислювальні ресурси системи; необхідність дотримання високого рівня конфіденційності персональних даних мешканців. Технології «розумного

будинку» на даний час інтенсивно розвиваються, активно інтегруючи при цьому елементи Інтернету речей (IoT), хмарних сервісів і систем штучного інтелекту (ШІ). Основною метою розробки таких систем є автоматизація побутових процесів, підвищення комфорту, забезпечення належного рівню енергоефективності та безпеки як у житлових, так і у промислових чи спеціалізованих приміщеннях, зокрема в медичних закладах.

Проте, аналіз існуючих рішень, що наведено у роботах [1, 4] показує, що більшість із них залишаються обмеженими в адаптивності та персоналізації поведінки користувача. Застосування методів машинного навчання (МН) дозволяє значно розширити функціональні можливості «розумних будинків». МН здатне опрацьовувати великі масиви даних, прогнозувати майбутні стани систем, адаптуватися до нових сценаріїв і навіть виявляти небажані або аномальні ситуації. Це, в свою чергу, дає змогу не лише автоматизувати, але й «інтелектуалізувати» житловий простір, створюючи середовище, здатне самостійно приймати рішення в інтересах користувача.

Дане дослідження спрямоване на пошук раціональних шляхів вирішення зазначених вище проблем, та надає відповіді на такі питання: якою мірою сучасні методи МН здатні подолати іманентні практичні проблеми середовищ «розумний будинок», включаючи проблеми забезпечення конфіденційності, безпеки, етики та інтероперабельності; які найбільш критичні компроміси мають місце при розгортанні різних класів алгоритмів МН в «розумних будинках» з огляду на ефективність, приватність, пояснюваність та ресурсні обмеження.

Метою даної роботи є систематизація існуючих підходів до застосування методів машинного навчання у «розумних будинках», зокрема шляхом класифікації типових задач, аналізу найбільш ефективних алгоритмів, оцінки їх практичного застосування в умовах обмежених ресурсів, високих вимог до конфіденційності та автономності, а також визначення актуальних викликів (включаючи розширені етичні аспекти, загрози безпеці моделей МН) і перспектив розвитку локальних, самонавчальних інтелектуальних систем, здатних до децентралізованого прийняття рішень.

Новизна даного огляду полягає у застосуванні комплексного підходу, що поєднує аналіз методів машинного навчання з поглибленим аналізом ключових практичних викликів їх впровадження в інтелектуальних системах «розумний будинок» та пропонує збалансований огляд перспектив їх подолання. На відміну від робіт, що часто концентруються на окремих алгоритмах або обмеженому колі проблем, ця стаття представляє цілісну картину. Вона охоплює класифікацію задач, порівняльний аналіз алгоритмів, специфіку даних та архітектур, а також детально аналізує такі критичні обмеження, як забезпечення приватності, нестача якісних даних, апаратні ліміти, та розглядає новітні проблеми, пов'язані з етикою ШІ (упередженість, справедливість, підзвітність) [6, 12, 24, 25, 36] та безпекою моделей МН (атаки змагального типу, отруєння даних, викрадення моделей) [2, 4, 9, 21]. Огляд пропонує сучасні рішення, такі як edge-архітектури, федеративне навчання, методи аугментації даних, технології підвищення приватності (PETs) та пояснюваний ШІ (XAI), що робить його актуальним та релевантним для сучасного стану досліджень.

Попри існуючи досягнення, залишається відкритим питання створення інтелектуальних систем, які були б водночас автономними, приватними та здатними до локального самонавчання. Більшість наявних рішень базуються на використанні хмарних сервісів або статичних правил, що обмежує їхню

адаптивність та створює ризики для конфіденційності даних користувача. Відсутність можливості навчання моделей у межах самого житлового середовища знижує персоналізацію та чутливість до динамічного контексту. Таким чином, виникає потреба у впровадженні локальних інтелектуальних агентів, які виконують повний цикл обробки даних — від збору до адаптивного прийняття рішень — без передачі інформації за межі домогосподарства. Саме це відкриває шлях до побудови наступного покоління «розумних будинків» — інтерпретованих, безпечних, етичних та по-справжньому персоналізованих.

Стаття структурується у такий спосіб. У розділі першому розглянуто основні типи задач, які виникають під час впровадження МН у системи «розумного будинку». Розділ другий присвячено огляду ключових методів машинного навчання. У розділі третьому об'єднано опис характерних типів даних, сенсорних мереж і архітектурних підходів. Розділ четвертий містить аналіз основних проблем, обмеження та перспективи. Розділ п'ятий присвячений новим критичним темам: розширеним етичним аспектам, безпеці моделей МН, людиноцентричному ШІ та стандартизації. У шостому розділі наведені висновки і опис напрямів подальших досліджень.

1. Типи задач машинного навчання в системах «розумний будинок»

Впровадження МН у системи «розумного будинку» спрямоване на вирішення низки специфічних задач, серед яких:

- Розпізнавання активності користувачів: Ідентифікація дій мешканців (сон, приготування їжі, перегляд телевізора) для адаптації освітлення, клімату, роботи приладів.
- Прогнозування споживання ресурсів: Передбачення витрат електроенергії, води, тепла для оптимізації та економії.
- Виявлення аномалій та безпека: Детектування нетипової поведінки, потенційних загроз (вторгнення, витік газу, пожежа) або несправностей обладнання.
- Персоналізація та адаптивне управління: Налаштування параметрів системи (освітлення, температура, мультимедіа) відповідно до індивідуальних звичок та вподобань користувачів.
- Управління комфортом: Автоматичне підтримання оптимальних умов мікроклімату, освітлення та акустичного середовища.
- Моніторинг здоров'я: Відстеження показників життєдіяльності та поведінкових патернів для раннього виявлення проблем зі здоров'ям, особливо для літніх людей або осіб з обмеженими можливостями.

2. Методи машинного навчання у системах «Розумний будинок»

Ефективне вирішення задач, окреслених у попередньому розділі, вимагає ретельного вибору методів машинного навчання, які відповідають як типу задачі, так і технічним обмеженням конкретної системи. У цьому розділі розглянуто основні класи ML-методів, що використовуються у «розумних будинках», з акцентом на їх характеристики, переваги, недоліки та приклади застосування. Варто також згадати навчання з підкріпленням (Reinforcement Learning, RL), яке набуває все більшого значення для задач оптимізації та адаптивного керування в «розумних будинках», зокрема в управлінні енергоспоживанням [5, 20, 39].

2.1. Класичні методи машинного навчання

Як зазначено в [1, 4], класичні підходи залишаються релевантними для простих задач класифікації та детекції у середовищах з обмеженими ресурсами.

До класичних методів машинного навчання, які застосовуються у системах розумного будинку, належать алгоритми, що демонструють хорошу інтерпретованість і помірну обчислювальну складність.

Зокрема, алгоритм k-найближчих сусідів (k-NN) використовується для класифікації активностей користувачів на основі схожості з попередніми спостереженнями (наприклад, k-NN класифікує активності шляхом відображення вектора сенсорних даних X реального часу до класу k його найближчих сусідів в історичному, маркованому наборі даних активностей Y).

Метод підтримувальних векторів (SVM) ефективний для задач бінарної класифікації, зокрема у виявленні аномалій.

Дерева рішень та Випадковий ліс (Random Forest) дозволяють інтерпретувати логіку прийняття рішень і добре працюють із невеликими датасетами.

Незважаючи на обмежену здатність до узагальнення складних патернів, класичні методи залишаються актуальними для систем із обмеженими ресурсами.

2.2. Глибоке навчання

Згідно з [2], глибокі нейронні мережі демонструють високу ефективність у задачах розпізнавання зображень та прогнозування поведінки, хоча і потребують значної кількості даних та ресурсів.

Глибокі нейронні мережі, зокрема згорткові нейронні мережі (CNN) та рекурентні нейронні мережі (RNN, LSTM), дозволяють працювати з великими масивами даних, включаючи зображення, аудіо та часові ряди.

У системах розумного будинку CNN застосовуються для розпізнавання облич або аналізу відео з камер, а RNN/LSTM — для передбачення поведінкових шаблонів мешканців або споживання ресурсів.

Глибоке навчання демонструє високу точність, проте потребує значних обчислювальних потужностей та великих навчальних вибірок. Окрім цього, варто згадати проблеми недостатньої інтерпретовності, що є центральною темою досліджень в області Explainable AI (XAI) [10, 23, 26, 32], та вразливості до атак змагального типу [1, 4, 9, 19].

2.3. Ансамблеві методи

Як показано в [3, 4], ансамблеві підходи дозволяють зменшити ймовірність переобучення та добре підходять для задач з високими вимогами до точності.

Методи ансамблю, такі як Random Forest, Gradient Boosting та XGBoost, поєднують результати декількох слабких моделей для підвищення точності прогнозу.

Їх часто використовують у задачах прогнозування енергоспоживання або виявлення аномалій.

Ансамблеві підходи дозволяють зменшити ймовірність переобучення та забезпечують високу стабільність результатів [4].

Проте, для розгортання в реальному часі такі моделі потребують попередньої оптимізації.

2.4. Легковагові моделі та МН на периферійних пристроях (Edge ML)

Останні дослідження [21, 22] доводять ефективність TinyML та подібних фреймворків у задачах реального часу без підключення до хмари, що особливо важливо для приватності та енергоефективності.

Зважаючи на обмежені обчислювальні ресурси типових розумних пристроїв, зростає популярність легковагових МН-моделей, які можуть працювати локально.

Платформи TensorFlow Lite та TinyML дозволяють реалізовувати нейронні мережі на мікроконтролерах, що відкриває можливості для автономної роботи без потреби у хмарній обробці [21, 22]. Такі моделі використовуються для локальної класифікації подій, виявлення руху або зміни стану середовища.

Хоча вони мають обмежену складність, їх ефективність доведена у практичних застосуваннях. Часто для розгортання складних моделей на периферійних пристроях потрібні методи стиснення та оптимізації моделей, такі як pruning, quantization та knowledge distillation [21, 22].

Таким чином, застосування легковагових ML-моделей є не лише відповіддю на апаратні обмеження, а й ключовим етапом у побудові автономних, локально навчальних інтелектуальних агентів у межах «розумного будинку».

Нижче наведено таблицю, що узагальнює ключові характеристики методів машинного навчання, розглянутих у цьому розділі:

Таблиця 1

Порівняльна характеристика методів машинного навчання для систем «розумного будинку»

Метод	Точ-ність	Ресурси	Адаптивність (до нових даних / поведінки)	Складність реалізації	Edge-суміс-ність	Інтерпретовність
k-NN	Середня	Низькі	Низька	Низька	Так	Висока
SVM	Висока	Середні	Середня	Середня	Частково	Середня
Дерева рішень	Середня	Низькі	Середня	Низька	Так	Висока
Random Forest	Висока	Середні	Середня	Середня	Так, з оптимізацією	Середня
XGBoost	Висока	Високі	Висока	Висока	Так, з оптимізацією	Низька
CNN	Дуже висока	Високі	Висока	Висока	Так (через TinyML/TF Lite)	Низька
RNN/LSTM	Висока	Високі	Дуже висока	Висока	Так (через TinyML/TF Lite)	Низька
TensorFlow Lite/TinyML	Середня	Дуже низькі	Середня	Середня	Так	Залежить від моделі

Примітка: Edge-сумісність для CNN, RNN/LSTM та XGBoost часто вимагає методів стиснення та оптимізації моделей. Для порівняння алгоритмів було використано умовну шкалу, засновану на результатах аналізу літератури та практичних кейсів, наведених у джерелах [1, 2, 4, 21, 22]. Оцінювання є узагальненим та носить орієнтовний характер.

3. Дані, сенсорні мережі та архітектура систем «розумний будинок»

Ефективність МН-компонентів у системах розумного будинку значною мірою залежить від типів використовуваних даних, структури сенсорної мережі та архітектури обробки. Без якісних вхідних даних і правильно організованої інфраструктури навіть найсучасніші алгоритми машинного навчання не здатні забезпечити очікуваний рівень автономності та ефективності. Важливими є не лише технічні характеристики сенсорів, а й способи їх інтеграції, методи попередньої обробки сигналів та формат зберігання. Також необхідно враховувати реальні обмеження: обмеження енергоспоживання, мережеву нестабільність, проблеми з оновленням моделей у польових умовах. У цьому розділі розглядаються ключові джерела даних, типові проблеми їх збору та обробки, а також архітектурні стратегії реалізації МН у реальному середовищі розумного будинку.

3.1. Типи даних і джерела

Зібрані дані можуть мати різну частоту оновлення, обсяг, формат та ступінь точності. Важливим є попереднє опрацювання даних — фільтрація шумів, нормалізація, обробка пропущених значень. Це значно впливає на точність моделей. Проблеми якості даних з недорогих IoT-сенсорів (шум, пропущені значення, проблеми калібрування) є поширеними та вимагають надійних методів попередньої обробки [21, 22]. Для навчання моделей часто формуються спеціальні датасети з анотаціями дій або подій, які збираються за участі користувачів або симуляційно [1, 2]. У розумному будинку збираються гетерогенні дані, які класифікуються за такими основними групами: Дані від сенсорів руху, температури, вологості, світла — використовуються для контролю клімату, освітлення, енергоспоживання. Аудіо- та відеодані — обробляються для розпізнавання голосу, ідентифікації осіб, безпеки [1, 2, 4]. Дані від побутових приладів (IoT) — генерують інформацію про активність, яку можна аналізувати для побудови поведінкових моделей [21, 22]. Наприклад, профілі енергоспоживання зі смарт-лічильників, інформація про завершення циклів роботи пральних машин чи холодильників, дані про стан фільтрів у системах вентиляції є цінними для МН. Дані календарів, графіків, мобільних додатків — застосовуються для адаптації сценаріїв до звичок користувача.

3.1.1. Методи аугментації даних для систем розумного будинку

З огляду на проблему дефіциту даних, особливо для персоналізованих середовищ «розумного будинку», методи аугментації даних відіграють важливу роль. Для часових рядів сенсорних даних застосовуються такі техніки, як jittering (додавання невеликого випадкового шуму), scaling (масштабування амплітуди), time warping (викривлення часової шкали), а також більш складні генеративні моделі, такі як варіаційні автоенкодери (VAE) та генеративно-змагальні мережі (GAN) [13, 15]. Ці методи допомагають збільшити обсяг та різноманітність

навчальних даних, підвищуючи надійність та здатність до узагальнення моделей МН.

3.2. Сенсорна інфраструктура

Ключову роль відіграє топологія мережі — зіркова, деревоподібна або mesh-структура. Вибір визначається площею покриття, рівнем навантаження на вузли, кількістю повторювачів сигналу. Для стабільності передачі важливим є також протокол синхронізації та живлення пристроїв. Бездротові сенсори можуть працювати від батарей, що обмежує обсяг переданих даних і частоту оновлення. У сучасних рішеннях також використовуються багатофункціональні датчики з інтегрованою локальною обробкою [21, 22]. Сенсорна мережа може включати десятки пристроїв, які взаємодіють між собою та з центральним вузлом — локальним сервером або хмарною платформою. Важливим фактором є вибір комунікаційного протоколу (Zigbee, Z-Wave, Wi-Fi), що визначає стабільність і масштабованість системи [27]. Спостерігається тенденція до фьюжн мультимодальних сенсорних даних, тобто поєднання даних з різних типів сенсорів (наприклад, камера + мікрофон + температура) для отримання більш насиченого контексту та надійніших висновків [36]. Також важливу роль відіграє стандартизація у комунікації сенсорів та форматах даних, наприклад, через ініціативи типу Matter [29, 33], що може спростити збір даних для МН.

3.3. Архітектура системи

При побудові архітектури важливо передбачити механізми оновлення моделей ML, логування подій, розмежування прав доступу та захист від атак типу MITM або spoofing. Часто використовується контейнеризація (наприклад, Docker), що дозволяє ізолювати обробку ML-даних від загальної логіки системи. У випадку гібридних архітектур усе більше застосовується підхід федеративного навчання (Federated Learning, FL), коли локальні моделі тренуються на пристроях і передають тільки оновлені ваги на центральний сервер для агрегації [2, 4, 11, 21, 40]. FL може мати різні конфігурації, наприклад, edge-to-cloud (пристрої надсилають оновлення до хмарного сервера) або device-to-device (пристрої обмінюються оновленнями безпосередньо), кожна з яких має свої компроміси щодо приватності, комунікаційних витрат та обчислювальних ресурсів. Важливим архітектурним аспектом є також забезпечення оновлення та підтримки моделей МН у реальному часі в розгорнутих системах «розумний будинок», включаючи механізми перенавчання або доналаштування моделей на нових даних без порушення користувацького досвіду. Такий поділ дозволяє враховувати обмеження апаратної бази, потреби в конфіденційності та вимоги до швидкодії.

У довгостроковій перспективі саме локальні ML-модулі із можливістю самонавчання без передачі сирих даних формують базу для децентралізованих і безпечних інтелектуальних середовищ, здатних самостійно адаптуватися до користувача.

4. Проблеми, обмеження, перспективи та нові критично важливі аспекти впровадження МН у розумних будинках

Незважаючи на значний потенціал використання машинного навчання у розумних будинках, його практичне впровадження супроводжується низкою комплексних проблем.

4.1. Існуючі проблеми та обмеження

Хоча машинне навчання відкриває значні можливості для розвитку розумних будинків, його впровадження супроводжується низкою викликів, які необхідно враховувати на етапі проєктування та реалізації систем. Приватність: По-перше, системи розумного будинку працюють з великим обсягом персональної інформації, включаючи поведінкові шаблони, розклад, переміщення, голосові команди тощо. Зберігання та обробка цих даних у хмарі створює потенційні ризики витоків або несанкціонованого доступу [1, 40]. Для їх мінімізації впроваджуються підходи до локальної обробки, шифрування, а також федеративного навчання [2, 4, 11, 21, 40]. Федеративне навчання зберігає приватність, дозволяючи обробляти дані децентралізовано, не передаючи їх на центральний сервер [2]. Інші технології підвищення приватності (Privacy-Enhancing Technologies, PETs), такі як диференційна приватність [11], гомоморфне шифрування або безпечні багатосторонні обчислення (Secure Multi-Party Computation, SMC) [11], також можуть бути застосовані, хоча й мають свої компроміси щодо обчислювальної складності та точності. Фреймворк PRASH [31] може слугувати прикладом підходу до аналізу ризиків приватності. Впровадження периферійних обчислень (Edge AI) також позитивно впливає на приватність даних, оскільки дані обробляються локально [21, 22].

Дефіцит даних: По-друге, багато МН-моделей потребують великих розмічених датасетів для ефективного навчання. Особливою складністю є той факт, що кожне середовище є унікальним — як за структурою, так і за звичками мешканців. Це вимагає адаптивних або онлайнових методів навчання, а також генерації синтетичних даних або застосування transfer learning [1, 2]. Як було зазначено в розділі 3.1.1, методи аугментації даних [13, 15] є важливим інструментом для подолання цієї проблеми. Трансферне навчання та навчання на малій кількості прикладів (few-shot learning) є перспективними рішеннями, особливо для персоналізації моделей у нових середовищах «розумного будинку» з обмеженими даними конкретного користувача. Наприклад, існують фреймворки, що зберігають приватність, поєднуючи динамічну кластеризацію та трансферне навчання для прогнозування енергоспоживання будівель [2].

Апаратні обмеження: По-третє, технічні обмеження пристроїв розумного будинку не завжди дозволяють використовувати складні моделі. Нестача ресурсів — як-от процесора, оперативної пам'яті чи енергоживлення — вимагає застосування оптимізованих рішень або переносу обробки на периферійні сервери [21]. Досягнення в TinyML [21, 22] пропонують рішення, забезпечуючи наднизьке енергоспоживання та можливість виконання задач МН на мікроконтролерах. Периферійні обчислення (Edge AI) є ключовою відповіддю на ці обмеження [21, 22].

Труднощі інтеграції: Окрім того, викликом є інтеграція МН до вже наявної інфраструктури. Потрібна підтримка сумісності, стандартизації протоколів передачі даних та гнучкі механізми оновлення [4, 27]. Ініціативи стандартизації, такі як Matter [29, 33], мають на меті покращити інтероперабельність пристроїв та даних, що може спростити інтеграцію МН.

4.2. Перспективи впровадження МЛ

Насамкінець, перспективи розвитку включають ширше застосування Edge AI. Це передбачає не лише ширше застосування, але й досягнення в платформах

Edge AI, апаратних прискорювачах (наприклад, NPU, Google Edge TPU [21, 22]) та можливість запуску складних моделей на периферійних пристроях [21, 22]. Однак, виникають проблеми, пов'язані з компромісами стиснення моделей та оновленням периферійних моделей.

Федеративне навчання (ФН) є ще одним перспективним напрямком [2, 4, 11, 21, 40]. Останні досягнення включають алгоритми, такі як FedAvg, FedProx, FedNova [2]. Практичні проблеми включають системну гетерогенність, комунікаційні накладні витрати, неоднорідні (non-IID) дані з різних будинків та вибуття клієнтів (пристроїв) [2, 4, 21, 40].

Поєднання різних джерел даних за допомогою мультимодального МН [36] відкриває нові можливості для глибшого розуміння контексту в «розумних будинках» (наприклад, поєднання аудіо, відео та сенсорних даних для розпізнавання активності або безпеки).

Використання методів навчання з підкріпленням (RL) [5, 20, 39] для підвищення автономності систем, особливо для систем управління енергоспоживанням будинку (HEMS), оптимізації комфорту та адаптивного керування. Глибоке навчання з підкріпленням (DRL) та багатоагентне навчання з підкріпленням (MARL) [5, 20] є перспективними для складного, розподіленого прийняття рішень. Проблеми включають ефективність вибірки в динамічних домашніх середовищах та безпеку автономного керування, а також інтеграцію з глобальними платформами «розумного міста» [11, 40].

На відміну від традиційного хмарного підходу, локальне самонавчання забезпечує не лише швидку реакцію та приватність, а й дозволяє системі «розумного будинку» еволюціонувати на основі унікальних даних користувача без стороннього впливу.

4.3. Розширені етичні аспекти ШІ в «розумному будинку» (поза приватністю)

Окрім приватності, існує ширший спектр етичних проблем, пов'язаних із застосуванням ШІ в «розумних будинках» [12, 24, 25, 36, 38]. Ці питання є критично важливими для надійного ШІ, оскільки системи МН приймають рішення, що впливають на повсякденне життя, комфорт, безпеку та споживання ресурсів.

Упередженість та справедливість (Bias and Fairness): Упередження в даних (наприклад, демографічні, поведінкові) можуть призвести до несправедливих або дискримінаційних результатів (наприклад, голосові асистенти гірше розпізнають певні акценти) [6, 12, 25, 36]. Важливо використовувати різноманітні дані та інструменти для виявлення та пом'якшення упередженості [6, 12, 25].

Підзвітність та відповідальність (Accountability and Responsibility): Визначення відповідальності за помилки системи МН, що спричинили шкоду, є складним питанням [24, 25, 36]. Потрібні чіткі регуляторні рамки.

Прозорість та пояснюваність (Transparency and Explainability): Користувачі мають право розуміти, як ШІ приймає рішення. Технології XAI (Explainable AI) спрямовані на вирішення цієї проблеми [10, 23, 24, 25, 26, 32, 36].

Безпека та надійність (Safety and Reliability): Забезпечення надійності систем МН, особливо тих, що контролюють критичні функції (безпека, моніторинг здоров'я), є етичним імперативом [5, 20, 24, 25, 39].

Автономія користувача (User Autonomy): Автоматизація на основі ШІ не повинна зменшувати автономію користувача або створювати надмірну залежність [25].

Таблиця 2

Розширені етичні аспекти МН в «розумних будинках»

Етичний принцип	Прояв в «розумних будинках»	Ключові питання / проблеми	Підходи до пом'якшення / керівні принципи проектування
Справедливість/Недискримінація	Упереджене розпізнавання активності, розбіжності у роботі голосових асистентів.	Як забезпечити різноманітність даних? Як проводити аудит моделей на упередженість?	Різноманітні джерела даних, інструменти виявлення упередженості, алгоритми, що враховують справедливість [6, 12, 25].
Підзвітність	Визначення відповідальності за помилки системи МН, що спричинили збитки/втрати.	Як встановити чіткі лінії відповідальності?	Чіткі регуляторні рамки, механізми зворотного зв'язку від користувачів [25].
Прозорість/ Пояснюваність	Розуміння користувачем рішень, прийнятих ШІ (наприклад, чому система змінила налаштування).	Як зробити «чорні скриньки» ШІ зрозумілими для неекспертів?	Методи XAI, зрозумілі користувацькі інтерфейси, що пояснюють рішення [10, 25, 26, 32].
Безпека та надійність	Забезпечення стабільної та безпечної роботи систем, що контролюють опалення, безпеку, здоров'я.	Як гарантувати відмовостійкість та захист від збоїв у критичних системах?	Ретельне тестування, механізми відмовостійкості, регулярні оновлення безпеки [24, 25, 39].
Автономія користувача	Ризик надмірної залежності від автоматизації, зменшення контролю користувача над власним середовищем.	Як збалансувати зручність автоматизації та збереження контролю користувача?	Надання користувачам можливості легко переозначати автоматичні дії, прозоре інформування про функції пристрою [25].

4.4. Розширені аспекти безпеки моделей МН в «розумних будинках»

Безпека моделей МН виходить за межі загальної кібербезпеки, включаючи специфічні вразливості, такі як отруєння даних та атаки змагального типу [1, 2, 4, 9, 19]. Скомпрометована модель МН може призвести до порушення приватності, збоїв у фізичній безпеці або фінансових втрат.

Атаки змагального типу (Evasion Attacks): Невеликі, непомітні для людини зміни у вхідних даних можуть змусити моделі МН помилково їх класифікувати (наприклад, обман систем розпізнавання облич) [7, 8, 9, 11, 14].

Отруєння даних (Data Poisoning): Шкідливі дані, введені під час навчання, можуть пошкодити модель МН або створити «чорні ходи» [1, 2, 4, 7, 8, 9, 11, 14, 19, 34, 35, 37]. Типи включають невибіркові, цільові атаки та атаки через «чорний хід» [1, 4].

Крадіжка/екстракція моделей (Model Theft/Extraction): Ризик викрадення пропрієтарних моделей МН або виведення чутливих навчальних даних [7, 8, 9, 11, 14].

Захисні механізми: проти отруєння даних - очищення даних (data sanitization) [7, 14, 16, 34], надійні алгоритми навчання, виявлення аномалій [11], захист на основі блокчейну [7, 34], машинне забування (machine unlearning) [35, 37]; проти атак змагального типу - навчання змагального типу, захисна дистиляція; проти крадіжки моделей: водяні знаки, обфускація моделей, диференційна приватність [8, 9].

Таблиця 3

Загрози безпеці та захисні механізми для моделей МН в «розумних будинках»

Тип загрози	Опис та приклад в «розумному будинку»	Можливий вплив	Стратегії захисту
Отруєння даних (Label Flipping, Clean-Label, Backdoor)	Зловмисне змінення навчальних даних для розпізнавання активності, що призводить до неправильної класифікації критичних подій.	Несправність системи, порушення приватності, несанкціонований доступ, ризик для безпеки.	Очищення даних [1, 16], надійне агрегування в FL, виявлення аномалій у внесках [11], машинне забування [35, 37].
Атаки змагального типу (Evasion)	Створення аудіокоманди, яка звучить нешкідливо для людини, але інтерпретується голосовим асистентом як «відчинити двері».	Несанкціонований доступ, маніпулювання системою.	Навчання змагального типу, валідація/трансформація вхідних даних.
Крадіжка/екстракція моделей	Викрадення моделі розпізнавання облич для створення клону або виведення даних про мешканців.	Порушення інтелектуальної власності, порушення приватності, створення умов для цільових атак.	Диференційна приватність, водяні знаки в моделі, моніторинг запитів [8, 9].

4.5. Людино центричний ШІ: сприйняття користувачами, довіра та взаємодія

Успіх МН в «розумних будинках» критично залежить від сприйняття та довіри користувачів [3, 18, 30, 38]. Це вимагає людиноцентричних підходів, таких як Human-in-the-Loop (HITL) та Explainable AI (XAI).

Фактори, що впливають на сприйняття та довіру: Психологічні фактори, сприйнята корисність, простота використання, задоволення, вартість, ризики безпеки/приватності, сумісність [3, 18, 30, 38].

Машинне навчання з людиною в циклі (Human-in-the-Loop, HITL): Залучення користувачів до процесу навчання/доопрацювання МН може покращити персоналізацію та побудувати довіру (наприклад, навчання преференцій) [18, 30].

Пояснюваний ШІ (Explainable AI, XAI) для довіри користувачів: Методи XAI можуть зробити рішення МН більш прозорими та зрозумілими, сприяючи довірі [10, 23, 26, 32]. Використання великих мовних моделей (LLM) для генерації пояснень є перспективним напрямком [23].

4.6. Стандартизація, інтероперабельність та управління даними для МН

Практичне розгортання МН залежить від стандартів для пристроїв та даних.

Комунікаційні протоколи для IoT/Edge ML: Роль специфічних протоколів (Zigbee, Z-Wave, Wi-Fi, BLE, Thread) важлива для збору даних та підтримки Edge ML [27].

Стандарти форматів даних: Відсутність стандартів для форматів сенсорних даних, що полегшують МН, є викликом [17, 28].

Інтероперабельність моделей МН: Проблема інтероперабельності між моделями МН від різних виробників. Зусилля щодо стандартизації форматів обміну моделями (наприклад, ONNX) можуть бути корисними.

Вплив стандартів типу Matter: Стандарти, такі як Matter [29, 33], можуть полегшити доступ до даних з різноманітних пристроїв, але вони не завжди безпосередньо стосуються інтероперабельності моделей МН [29].

Стандартизуючі організації: Ключові організації включають CSA (для Matter), IEEE, IETF, ISO/IEC JTC 1/SC 42 (для стандартів ШІ, таких як ISO/IEC 5338, 42001) [10, 17, 27, 28, 29, 33].

5. Висновки та напрями подальших досліджень

У цій статті було здійснено огляд основних напрямів застосування методів машинного навчання в інтелектуальних системах розумного будинку.

Проаналізовано ключові задачі, такі як розпізнавання активності, прогнозування споживання ресурсів, виявлення аномалій та адаптивне управління побутовими сценаріями.

Розглянуто сучасні алгоритми МН, включаючи класичні, ансамблеві, глибокі та edge-орієнтовані моделі, а також специфіку їх застосування в умовах обмежених ресурсів та високих вимог до приватності. Особливу увагу було приділено типам даних, сенсорній інфраструктурі та архітектурним рішенням, що визначають ефективність впровадження МН у розумні середовища.

Крім того, окреслено основні проблеми — нестача якісних даних, обмеження обчислювальних потужностей, складність інтеграції, необхідність

забезпечення кібербезпеки, а також новітні етичні та безпекові проблеми. Загалом, використання МН в «розумних будинках» демонструє значний прогрес, але фундаментальні проблеми, пов'язані з приватністю, безпекою, етикою та інтероперабельністю, все ще потребують більш ефективних варіантів вирішення для широкого впровадження МН у практику створення даних систем.

Системи, здатні до локального самонавчання, поступово стають новим стандартом у галузі інтелектуального житла. Поєднання Edge AI, TinyML та людино центричного дизайну створює умови для виникнення справді автономних та адаптивних середовищ.

Подальші дослідження можуть бути зосереджені на:

Розробці гібридних моделей, що поєднують фізично-обґрунтоване МН з підходами, керованими даними, для надійного прогнозування енергоспоживання в будинках з розрідженими сенсорними даними.

Оптимізації edge-рішень, включаючи розробку апаратних прискорювачів та методів стиснення моделей для складних задач на пристроях з обмеженими ресурсами.

Побудові масштабованих децентралізованих архітектур, зокрема, дослідження ефективності та безпеки різних конфігурацій федеративного навчання для non-IID даних.

Вивченні впливу таких систем на якість життя користувачів, включаючи лонгітюдні дослідження довіри та сприйняття XAI-посилених «розумних будинків» [7, 18, 30, 39].

Поєднанні МН із технологіями семантичного моделювання, мультиагентних систем та когнітивних підходів до адаптації середовища.

Етичний ШІ: Розробці культурно-адаптованих та контекстно-чутливих етичних рамок для ШІ в «розумному будинку»; методи безперервного аудиту моделей МН на предмет упередженості та справедливості [6, 12, 25].

Безпека МН: Розробці більш надійних та обчислювально ефективних захисних механізмів проти нових атак змагального типу та отруєння даних у розподілених середовищах (наприклад, для FL або Edge ML). Дослідження верифікованого машинного забування [1, 4, 7, 8, 9, 19, 34, 35, 37].

Людино центричний ШІ: Розробці інтуїтивно зрозумілих інтерфейсів HITL для неекспертних користувачів для персоналізації їхнього ШІ в «розумному будинку» [18, 30].

Стандартизація: Дослідженні доменно-специфічних онтологій даних та стандартизованих форматів моделей/даних МН для застосувань у «розумному будинку» з метою покращення інтероперабельності [17, 28, 29, 33].

References

1. Saleem A. A. Smart Homes Powered by Machine Learning: A Review / A. A. Saleem, M. M. Hassan, I. A. Ali // 2022 International Conference on Computer Science and Software Engineering (CSASE) : proc. of the conf. – 2022. – P. 355–361.

2. Machine Learning in Smart Buildings: A Review of Methods, Challenges, and Future Trends / F. El Husseini, H. N. Noura, O. Salman, K. Chahine // Applied Sciences. – 2025. – Vol. 15, № 14. – Art. no. 7682.

3. da Silva D. C. Evaluating Multi-Label Machine Learning Models for Smart Home Environments / D. C. da Silva, D. R. D. Boaventura, M. dos Santos Oliveira // Software: Practice and Experience. – 2025.

4. Smart Home System: A Comprehensive Review / A. Chakraborty [et al.] //

Journal of Electrical and Computer Engineering. – 2023. – Vol. 2023. – P. 1–30.

5. Latoń D. Applications of Deep Reinforcement Learning for Home Energy Management Systems: A Review / D. Latoń, J. Grela, A. Ożadowicz // Energies. – 2024. – Vol. 17, № 24. – P. 6420.

6. A Survey on Bias and Fairness in Machine Learning [Electronic resource] / N. Mehrabi [et al.] // arXiv preprint arXiv:1908.09635. – 2022. – Available at: https://arxiv.org/abs/1908.09635.

7. Detecting and Preventing Data Poisoning Attacks on AI Models / H. Kure, P. Sarkar, A. B. Ndanusa, A. O. Nwajana // 2025 PhotonIcs & Electromagnetics Research Symposium (PIERS) : proc. of the symp., Abu Dhabi, 2025. – 2025.

8. Dunn C. Robustness Evaluations of Sustainable Machine Learning Models against Data Poisoning Attacks in the Internet of Things / C. Dunn, N. Moustafa, B. Turnbull // Sustainability. – 2020. – Vol. 12, № 16. – Art. no. 6434.

9. Iot device security risks: a comprehensive overview and mitigation strategies / U. Iwuanyanwu [et al.] // Journal of Technology & Innovation. – 2023. – Vol. 3, № 1. – P. 38–43.

10. Explainable Artificial Intelligence: A Survey of Needs, Techniques, Applications, and Future Direction [Electronic resource] / M. Mershaa [et al.] // arXiv preprint arXiv:2409.00265. – 2024. – Available at: https://arxiv.org/abs/2409.00265.

11. A Detailed Survey on Federated Learning Attacks and Defenses / H. S. Sikandar [et al.] // Electronics. – 2023. – Vol. 12, № 2. – Art. no. 260.

12. Bias in data-driven artificial intelligence systems—An introductory survey / E. Ntoutsi [et al.] // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. – 2020. – Vol. 10, № 3.

13. Iwana B. K. An empirical survey of data augmentation for time series classification with neural networks / B. K. Iwana, S. Uchida // PLOS One. – 2021. – Vol. 16, № 7. – P. e0254841.

14. Biggio B. Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning / B. Biggio, F. Roli // Pattern Recognition. – 2018. – Vol. 84. – P. 317–331.

15. Time Series Data Augmentation for Deep Learning: A Survey [Electronic resource] / Q. Wen [et al.] // arXiv preprint arXiv:2002.12478. – 2022. – Available at: https://arxiv.org/abs/2002.12478.

16. Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations / NIST. – Gaithersburg, MD : National Institute of Standards and Technology, 2023. – (NIST AI 100-2).

17. The FAIR Guiding Principles for scientific data management and stewardship / M. D. Wilkinson [et al.] // Scientific Data. – 2016. – Vol. 3. – Art. no. 160018.

18. The Human in Interactive Machine Learning: Analysis and Perspectives for Ambient Intelligence / K. Delcourt, S. Trouilhet, J.-P. Arcangeli, F. Adreit // Journal of Artificial Intelligence Research. – 2024. – Vol. 81.

19. A Taxonomy of Attacks in Federated Learning / N. Pitropakis [et al.] // Applied Sciences. – 2021. – Vol. 11, № 21. – Art. no. 10300.

20.Wang Y. Multi-Agent Reinforcement Learning for Smart Community Energy Management / Y. Wang, T. Huang, K. Li // Energies. – 2024. – Vol. 17, № 20. – P. 5211.

21. Thakur P. Edge AI Enabled IoT Framework for Secure Smart Home Infrastructure / P. Thakur, S. Goel, E. Puthooran // Procedia Computer Science. – 2024. – Vol. 235. – P. 3369–3378.

22. Reis M. J. C. S. Edge AI for Real-Time Anomaly Detection in Smart Homes

/ M. J. C. S. Reis, C. Serôdio // Future Internet. – 2025. – Vol. 17, № 4. – Art. no. 179.

23. Leveraging Large Language Models for Explainable Activity Recognition in Smart Homes: A Critical Evaluation [Electronic resource] / M. Fiori, G. Civitarese, P. Choudhary, C. Bettini // arXiv preprint arXiv:2503.16622. – 2025. – Available at: https://arxiv.org/abs/2503.16622.

24. Ethical Dilemmas and Privacy Issues in Emerging Technologies: A Review / L. L. Dhirani, N. Mukhtiar, B. S. Chowdhry, T. Newe // Sensors. – 2023. – Vol. 23, № 3. – P. 1151.

25. Chung J. Ethical Considerations Regarding the Use of Smart Home Technologies for Older Adults: An Integrative Review / J. Chung, G. Demiris, H. J. Thompson // Annual Review of Nursing Research. – 2016. – Vol. 34, № 1. – P. 155–181.

26. Vilone G. Explainable AI for Smart Buildings: A Survey on Current Trends, Applications, and Challenges / G. Vilone, L. Longo // Buildings. – 2023. – Vol. 13, № 1. – Art. no. 211.

27. Utilization of 5G Technologies in IoT Applications: A Systematic Literature Review / I. Makhdoom [et al.] // Sensors. – 2023. – Vol. 23, № 8. – P. 3995.

28. Stojkoska B. L. R. A review of Internet of Things for smart home: Challenges and solutions / B. L. R. Stojkoska, K. V. Trivodaliev // Journal of Cleaner Production. – 2017. – Vol. 140. – P. 1454–1464.

29. Matter Specification [Electronic resource] / Connectivity Standards Alliance. – 2022. – Available at: https://csa-iot.org/all-solutions/matter/#spec.

30. A Survey of Human-in-the-loop for Machine Learning [Electronic resource] / Y. Wu [et al.] // arXiv preprint arXiv:2108.00941. – 2022. – Available at: https://arxiv.org/abs/2108.00941.

31. Xu X. Privacy Risk Assessment of Smart Home System Based on a STPA–FMEA Method / X. Xu, Z. Li, C. Zhang // Sensors. – 2023. – Vol. 23, № 10. – Art. no. 4664.

32. Angelov P. Human-Centered XAI for Smart Home Applications [Electronic resource] / P. Angelov, E. Soares // arXiv preprint arXiv:2404.16074. – 2024. – Available at: https://arxiv.org/abs/2404.16074.

33. Matter: The Foundation for Connected Things [Electronic resource] / Connectivity Standards Alliance. – 2022. – Available at: https://csa-iot.org/all-solutions/matter/#main.

34. Kim T. A Novel Data Sanitization Method Based on Dynamic Dataset Partition and Inspection Against Data Poisoning Attacks / T. Kim, K. Chae // Sensors. – 2024. – Vol. 24, № 2. – Art. no. 374.

35. Machine Unlearning: A Comprehensive Survey [Electronic resource] / J. Zhang [et al.] // arXiv preprint arXiv:2405.07406. – 2024. – Available at: https://arxiv.org/abs/2405.07406.

36. Novak P. Multimodal ML in Ambient Intelligence / P. Novak // Computers in Human Behavior. – 2024.

37. A Survey of Machine Unlearning Techniques / A. Pawlicka, J. Pawlicki, A. Kozik, M. Choraś // Applied Sciences. – 2023. – Vol. 13, № 24. – Art. no. 13180.

38. Marikyan D. A systematic review of the literature on the smart home and its users: The pillars of adoption and continued use / D. Marikyan, S. Papagiannidis, E. Alamanos // Technological Forecasting and Social Change. – 2019. – Vol. 146. – P. 660–681.

39. Challenges and Approaches of Reinforcement Learning for Power System

Decision and Control: A Review / J. Zhang [et al.] // Renewable and Sustainable Energy Reviews. – 2023. – Vol. 182.

40. Recent Advancements in Federated Learning: State of the Art, Fundamentals, Principles, IoT Applications and Future Trends / L. U. Khan [et al.] // IEEE Access. – 2021. – Vol. 9. – P. 130095–130129.

Надійшла до редакції 26.08.2025, розглянута на редколегії 26.08.2025

Machine Learning Methods in Intelligent «Smart Home» Systems: Classification, Problems, and Prospects

The article presents a systematic review of machine learning methods applied in intelligent «Smart Home» systems. A classification of typical tasks is provided, including user activity recognition, resource consumption forecasting, anomaly detection during operation, personalization of household scenarios, comfort management, and health monitoring of residents. Key classes of algorithms are analyzed in detail: classical methods (k-NN, SVM, decision trees, Random Forest), ensemble approaches (Gradient Boosting, XGBoost), deep learning methods (CNN, RNN, LSTM), as well as lightweight models optimized for edge devices (TinyML, TensorFlow Lite). A comparative analysis of their characteristics is conducted, in particular accuracy, resource requirements, adaptability, interpretability, and real-time integration capabilities.

Special attention is given to data organization, sensor networks, and architectural strategies for building «Smart Homes». Key challenges are identified: ensuring privacy and cybersecurity, the lack of high-quality datasets, hardware limitations of devices, as well as issues of standardization and interoperability. Current approaches to overcoming these challenges are analyzed, including data augmentation and transfer learning methods, privacy-enhancing technologies (differential privacy, homomorphic encryption, secure multiparty computation), federated learning, and edge architectures.

Significant attention is devoted to the ethical aspects of AI application, related to model bias, fairness, transparency, accountability, and preservation of user autonomy. Security issues of machine learning models are examined separately: adversarial attacks, data poisoning, and model stealing, as well as respective defense mechanisms — data sanitization, anomaly detection, adversarial training, obfuscation, and watermarking. The role of human-centric approaches and Explainable AI in increasing user trust is highlighted.

Future prospects of the field are outlined, including the spread of Edge AI and TinyML, advancement of federated learning algorithms, implementation of multimodal systems, development of reinforcement learning methods, and the creation of decentralized architectures with local self-learning. The proposed review forms a holistic understanding of the current state and trends of ML use in «Smart Home» systems and defines practical ways to overcome existing problems.

Key words: Smart Home; machine learning; intelligent systems; personalization; energy efficiency; anomaly detection; federated learning; Edge AI; AI ethics; model security.

Відомості про авторів:

Кушнарьов Максим Олександрович – аспірант, Національний аерокосмічний університет «ХАІ», м. Харків, Україна, maksimkushnarov@gmail.com, ORCID: 0009-0003-6322-1740

Аверін Юрій Сергійович – аспірант, Національний аерокосмічний університет «ХАІ», м. Харків, Україна, averin.yury@gmail.com, ORCID: 0009-0007-5533-1789

Шостак Ігор Володимирович – професор, доктор технічних наук, професор кафедри 603 «Інженерії програмного забезпечення», Національний аерокосмічний університет «ХАІ», м. Харків, Україна, i.shostak@khai.edu, ORCID: 0000-0002-3051-0488

About the authors:

Kushnarov Maksym– PhD Student, National Aerospace University «KhAI», Kharkiv, Ukraine, maksimkushnarov@gmail.com, ORCID: 0009-0003-6322-1740

Averin Yurii– PhD Student, National Aerospace University «KhAI», Kharkiv, Ukraine, averin.yury@gmail.com, ORCID: 0009-0007-5533-1789

Shostak Ihor – Professor, Doctor of Technical Sciences, Professor of the Department of Software Engineering, National Aerospace University "KhAI", Kharkiv, Ukraine, i.shostak@khai.eduiv, ORCID: 0000-0002-3051-0488