

Хмельницький національний університет
Факультет інформаційних технологій
Кафедра комп'ютерної інженерії та інформаційних систем

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

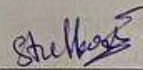
Галузь знань 12 – Інформаційні технології

Спеціальність 123 – Комп'ютерна інженерія

на тему «Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини»

КвРКІ. 016018.17.01.01 ПЗ

Виконала: студентка 2 курсу, група КІ2м-20-1

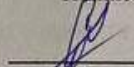


Підпис

Стрелкова О.Г.

Ініціали, прізвище

Керівник кандидат техн. наук, доцент
Науковий ступінь, вчене звання



Підпис

Нічепорук А.О.

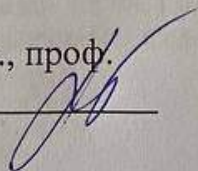
Ініціали, прізвище

До захисту допускаю:

Зав. кафедри КІС, д.т.н., проф.

Т.О. Говорущенко

04 05 2022 р.



Хмельницький, 2022

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОРМАЦІЙНИХ СИСТЕМ

Освітній рівень МАГІСТР

Галузь знань 12 ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Спеціальність 123 КОМП'ЮТЕРНА ІНЖЕНЕРІЯ

Освітня програма ОСВІТНЬО-НАУКОВА ПРОГРАМА «КОМП'ЮТЕРНА ІНЖЕНЕРІЯ ТА ПРОГРАМУВАННЯ»

ЗАТВЕРДЖУЮ

Зав. кафедри _____

Т.О.Говорущенко

“ 01 ” 09 2021 р.

ЗАВДАННЯ

НА ДИПЛОМНИЙ ПРОЕКТ (РОБОТУ)

Стрелковій Олександрі Георгіївні

Прізвище, ім'я, по батькові студента

1. Тема проекту (роботи) Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Керівник проекту (роботи) Нічепорук А.О., к.т.н., доцент

Прізвище, ім'я, по батькові, науковий ступінь, вчене звання

Затверджена наказом ректора університету від 06.01.2022 р. № 1

2. Строк подання студентом проекту (роботи) на кафедру 03.05.2022 р.

3. Вихідні дані до проекту (роботи) Завдання на дипломне проектування

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)

Огляд існуючих систем інтелектуалізації процесів у розумних будинках





Підготовка даних та модель процесу класифікації активностей людини у розумному будинку

Метод розпізнавання активностей людини в розумному будинку

Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень)

6. Консультанти розділів дипломного проекту (роботи)

| Розділ | Прізвище, ініціали та посада консультанта | Підпис, дата | |
|---------------|---|---|---|
| | | завдання видав | завдання прийняв |
| Нормоконтроль | Лисенко С.М., професор кафедри КІС |  |  |
| Антиплагіат | Нічепорук А.О., доцент кафедри КІС |  |  |

7. Дата видачі завдання « 06 » 09 2021р.

КАЛЕНДАРНИЙ ПЛАН

| №з/п | Назва етапів (розділів) дипломного проекту (роботи) | Термін виконання етапів проекту (роботи) | Примітка |
|------|---|--|----------|
| 1 | Вибір напрямку дослідження та узгодження тематики ДРМ з керівником | 05.09.2021 | виконано |
| 2 | Ознайомлення з предметною областю; формулювання мети та задач дослідження; визначення об'єкта та предмета дослідження | 05.10.2021 | виконано |
| 3 | Робота над розділом 1 – аналіз відомих моделей, методів за темою; постановка задачі | 05.11.2021 | виконано |
| 4 | Робота над розділом 2 – розробка моделей для вирішення поставленої задачі | 05.12.2021 | виконано |
| 5 | Робота над тезою доповіді | 05.01.2022 | виконано |
| 6 | Робота над розділом 3 – розробка методів для вирішення поставленої задачі | 15.02.2022 | виконано |
| 7 | Робота над розділом 4 – проектування та розробка ПЗ для вирішення поставленої задачі, експериментальна частина | 05.04.2022 | виконано |
| 8 | Оформлення пояснювальної записки згідно вимог | 15.04.2022 | виконано |
| 9 | Попередній захист ДРМ | 18.04.2022 | виконано |
| 10 | Захист ДРМ на засіданні ЕК | До 10.05.2022 | |

Студентка



Підпис

О.Г. Стрелкова

Ініціали, прізвище

Керівник проекту (роботи)



Підпис

А.О. Нічепорук

Ініціали, прізвище

РЕФЕРАТ

Тема дипломної роботи: Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини.

Автор роботи: Стрелкова Олександра Георгіївна.

Керівник роботи: Нічепорук Андрій Олександрович.

Пояснювальна записка: 97 с., 25 рис., 5 табл., 1 дод., 50 джерел.

Інтернет речей, алгоритм класифікації, штучний інтелект, розпізнавання діяльності, розумний будинок, автоматизація.

Об'єктом дослідження є процес класифікації видів діяльності зафіксованих у розумних будинках.

Предметом дослідження є інтелектуальна система класифікації видів діяльності, зафіксованих у розумних будинках.

Метою дипломної роботи є підвищення ступеню автоматизації та реагування на потреби мешканців розумного будинку шляхом проектування інтелектуальної системи класифікації видів діяльності, зафіксованих у розумних будинках.

Для розв'язання поставлених задач використовувалися аналітичні та математичні методи дослідження, засоби комп'ютерних мереж, теорія графів та множин, методи оцінки ефективності.

Наукова новизна отриманих результатів:

– набула подальшого розвитку модель процесу розпізнавання людської активності на основі наївного байєсового класифікатора, яка дозволяє визначити найбільш ймовірну активність, виконувану мешканцем розумного будинку за набором ознак, що дозволило формалізувати основні етапи розпізнавання людської діяльності в середовищі розумного будинку;

– удосконалено інтелектуальну систему прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини, яка на відмінну від відомих залучає методи машинного

навчання, що дозволило підвищити ступінь автоматизації процесів у розумних будинках на основі видів людської діяльності та налаштовувати параметри розумних пристроїв під потреби користувачів та/або оптимізацію енергоспоживання.

На основі проведених досліджень запропонована інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини.

Практична значимість отриманих результатів полягає у тому, що запропонована інтелектуальна система дозволить підвищити ступінь автоматизації та реагування на потреби мешканців розумного будинку шляхом проектування інтелектуальної системи класифікації видів діяльності, зафіксованих у розумних будинках.

ЗМІСТ

| | |
|--|----|
| Скорочення та умовні позначки | 4 |
| Вступ..... | 5 |
| 1 Огляд існуючих систем інтелектуалізації процесів у розумних будинках | 8 |
| 1.1 Концепція Інтернету речей..... | 8 |
| 1.2 Архітектура IoT системи | 9 |
| 1.3 Методи аналізу даних розпізнавання людської діяльності | 12 |
| 1.3.1 Штучні нейронні мережі | 12 |
| 1.3.2 K-найближчі сусіди | 14 |
| 1.3.3 Support Vector Machine (SVM)..... | 15 |
| 1.3.4 Random Forests (RF) | 15 |
| 1.3.5 CART (classification and regression tree) | 16 |
| 1.3.6 Hidden Markov Model (HMM) | 17 |
| 1.4 Огляд відомих методів та засобів розпізнавання людської діяльності штучним інтелектом в контексті IoT..... | 17 |
| 1.5 Постановка задачі дослідження | 24 |
| 2 Підготовка даних та модель процесу класифікації активностей людини у розумному будинку..... | 25 |
| 2.1 Опис предметної області | 25 |
| 2.2 Функціональні та нефункціональні вимоги | 30 |
| 2.3 Попередня обробка даних | 32 |
| 2.4 Модель процесу класифікації життєдіяльності людини | 35 |
| 2.5 Висновки | 41 |
| 3 Метод розпізнавання активностей людини в розумному будинку..... | 42 |
| 3.1 Метод розпізнавання активностей людини в розумному будинку на основі використання Naïve Bayes Classifier..... | 42 |
| 3.2 Обробка даних багатовимірних часових рядів | 42 |
| 3.3 Класифікація часових рядів на основі Bayesian Network Classifier | 46 |
| 3.4 Класифікатор на основі використання Naïve Bayes Classifier | 51 |

| | |
|--|----|
| 3.5 Висновки | 57 |
| 4 Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини | 58 |
| 4.1 Структура інтелектуальної системи прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини | 58 |
| 4.2 Алгоритми колекціонування та обробки даних датчиків розумного будинку | 64 |
| 4.3 Алгоритм визначення активності та оцінка ефективності системи..... | 74 |
| 4.4 Висновки | 78 |
| Висновки | 79 |
| Перелік посилань..... | 81 |

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

IoT – Internet of Things

ITU – International Telecommunication Union

GPS – Global Positioning System

AR – Augmented Reality

ML – Machine Learning

AI – Artificial Intelligence

CRISP-DM – Cross-Industry Standard Process for Data Mining

CNN – Convolutional neural network

LSTM – Long short-term memory

BLE – Bluetooth Low Energy

CPU – Central Processing Unit

GPU – Graphics Processing Unit

RSSI – Received Signal Strength Indication

MSE – Mean Squared Error

TSDB – Time series database

API – Application Programming Interface

BSON – Binary JavaScript Object Notation

JSON – JavaScript Object Notation

CHAP – Challenge Handshake Authentication Protocol

REST – Representational State Transfer

MQTT – Message Queue Telemetry Transport

UUID – Universally Unique Identifier

NFC – Near Field Communication

ВСТУП

На сьогоднішній день людство постійно переосмислює застосування існуючих речей, шукаючи тим самим способи вирішення нових задач, які перед ним ставляться. Схожим чином була запропонована та реалізована концепція Інтернету речей, в якій, шляхом інтеграції інтелектуальної складової у фізичні об'єкти та створення комунікаційного середовища між ними, забезпечувалась автоматизація повсякденних процесів у виробництві, житті, здоров'ї та інших галузях людського суспільства.

Одне із застосувань Інтернету речей, це розпізнавання людської діяльності в середовищах, що містять пристрої з додатковим технологічним обладнанням, що підключає їх до Інтернету, що тим самим робить їх «розумними». Прикладом для такого середовища є розумний дім, де датчики, наприклад, виявляють, коли людина лягає спати до спальні. Як тільки система розумного будинку підхоплює діяльність за допомогою датчиків, вона класифікує подію як «сплячу» за допомогою алгоритму класифікації. Система допомагає людині та вимикає систему опалення приміщення без будь-якого втручання людини. Таким чином, моніторинговий розумний будинок збирає дані датчиків, виявляє події, реагує та допомагає мешканцю, коли це потрібно. Проте, існуючі системи домашньої автоматизації не дозволяють в повному обсязі автоматизувати цей процес, покладаючись виключно на формування правил виду «якщо – то».

Метою роботи є підвищення ступеню автоматизації та реагування на потреби мешканців розумного будинку шляхом проектування інтелектуальної системи класифікації видів діяльності, зафіксованих у розумних будинках.

Об'єктом дослідження є процес класифікації видів діяльності, зафіксованих у розумних будинках.

Предметом дослідження є інтелектуальна система класифікації видів діяльності, зафіксованих у розумних будинках.

Методи дослідження. У роботі було застосовано наступні теорії та засоби:

1. Аналітичні та математичні методи дослідження.
2. Засоби комп'ютерних мереж.
3. Теорія графів та множин.
4. Методи оцінки ефективності.
5. Сучасні програмні засоби проектування та дослідження.
6. Персональний комп'ютер.

Наукова новизна роботи:

1. Набула подальшого розвитку модель процесу розпізнавання людської активності на основі наївного байєсового класифікатора, яка дозволяє визначити найбільш ймовірну активність, виконувану мешканцем розумного будинку за набором ознак, що дозволило формалізувати основні етапи розпізнавання людської діяльності в середовищі розумного будинку.

2. Удосконалено інтелектуальну систему прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини, яка на відмінну від відомих залучає методи машинного навчання, що дозволило підвищити ступінь автоматизації процесів у розумних будинках на основі видів людської діяльності та налаштовувати параметри розумних пристроїв під потреби користувачів та/або оптимізацію енергоспоживання.

Практична цінність роботи полягає в тому, що запропонована інтелектуальна система дозволить підвищити ступінь автоматизації та реагування на потреби мешканців розумного будинку шляхом проектування інтелектуальної системи класифікації видів діяльності, зафіксованих у розумних будинках.

Важливість роботи і висновки. Робота має важливе значення для розвитку та впровадження технологій домашньої автоматизації.

За результатами роботи зроблені наступні висновки:

1. Проведено аналіз існуючих методів, моделей та засобів класифікації людської діяльності в системах розумних будинків.
2. Розроблено модель процесу класифікації активностей людини у розумному будинку.
3. Розроблено метод розпізнавання активностей людини в розумному будинку.
4. Запропоновано інтелектуальну систему прийняття рішень щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини.
5. Досліджено ефективність класифікації діяльності людини на основі запропонованого рішення.

Публікації. Дослідження проведені у даній роботі були представлені на Всеукраїнській науково-практичній конференції «Актуальні проблеми комп'ютерних наук АПКН-2021», за результатами якої опубліковано у збірнику конференції тезу доповіді «Інтелектуальна система класифікації видів діяльності, зафіксованих у розумних будинках» [1].

Структура та об'єм дипломної роботи. Дипломна складається з вступу, чотирьох розділів, висновку та додатків, її повний зміст 97 сторінок, основний зміст викладено на 76 сторінках, 1 додаток на 11 сторінках, містить 25 рисунків, 5 таблиць, включає 50 найменувань вітчизняної та зарубіжної літератури.

1 ОГЛЯД ІСНУЮЧИХ СИСТЕМ ІНТЕЛЕКТУАЛІЗАЦІЇ ПРОЦЕСІВ У РОЗУМНИХ БУДИНКАХ

1.1 Концепція Інтернету речей

Термін «Інтернет речей» (IoT) вперше ввів Кевін Ештон у 1999 році. Немає офіційного і загальноприйнятого визначення IoT. Автор терміну зазначив: «Завдяки Інтернету речей комп'ютери можуть відчувати речі самі». Міжнародний союз телекомунікацій (ITU) визначив IoT як «Глобальну інфраструктуру для інформаційного суспільства, що забезпечує передові послуги шляхом взаємозв'язку (фізичних і віртуальних) речей на основі існуючих та розроблюваних інформаційно-комунікаційних технологій». Це визначення можна розглядати як основу технологій IoT.

Системи Інтернету речей підключають фізичний світ до Інтернету. IoT працює шляхом приєднання фізичних речей до Інтернету, таких як сенсори, які надають дані і реагують на середовище навколо них. В результаті IoT системи забезпечують технології та засоби для вимірювання, оцінки та реагування на навколишній світ.

Підключення датчиків додає дані фізичного світу в Інтернет. Ця трансформаційна зміна долає розрив між фізичним і віртуальним світом, який зберігався з моменту початку комп'ютеризації. IoT доповнює Інтернет з усіма його функціями і можливостями, додаючи до нього вимір реального світу. З впровадженням IoT Інтернет стає мережею людей, інформації, послуг і речей – Інтернетом всього.

В IoT Інтернет поширюється на фізичні пристрої та предмети побуту. З вбудованою електронікою, підключеним Інтернетом та іншими видами апаратного забезпечення (наприклад, датчиками), ці пристрої можуть спілкуватися та взаємодіяти з іншими через Інтернет, і вони можуть контролюватись віддалено. IoT додає можливість IoT-пристроєм взаємодіяти з існуючою інфраструктурою Інтернет.

Аналітики прогнозують, що встановлена база підключених пристроїв налічуватиме десятки мільйонів у найближчі кілька років, а кількість датчиків, за прогнозами, зросте до сотні мільярдів у найближчому майбутньому. У той же час прогнозується, що вплив Інтернету речей на бізнес становитиме трильйони доларів США, що можна порівняти з ВВП найбільших економік світу. Незалежно від прогнозів, це приголомшливі цифри, які вказують на справді трансформаційні зміни.

З'єднання фізичного та віртуального світів за допомогою мережі Інтернет і протоколів дає змогу створювати багато нових типів гібридних взаємодій і, таким чином, потенціал для створення безлічі нових застосувань, додатків та бізнес-моделей. Розвиток цієї сфери впливає не тільки на галузь, але й на багато аспектів повсякденного життя та добробуту людей.

1.2 Архітектура IoT системи

Архітектура IoT складається з чотирьох головних функціональних рівнів: рівня додатків, рівня датчиків, рівня шлюзів і мереж, серверного рівня, що утворюють взаємопов'язуюче середовище комунікації між фізичним світом та кіберсередовищем (рис. 1.1). Зокрема нижній рівень формують додатки IoT, такі як офіс, автомобіль, транспорт, виробництво, сільське господарство, будинок, офісна будівля, розумне місто. Рівень вище ілюструє деякі репрезентативні категорії IoT датчиків та речей.

Деякі датчики та речі призначені для безпосереднього підключення до Інтернету та спілкування з додатками та сервісами, що знаходяться в хмарному середовищі. Ними є різноманітні пристрої, такі як камери безпеки, пожежні сенсори, термостати, прилади та лічильники електроенергії.

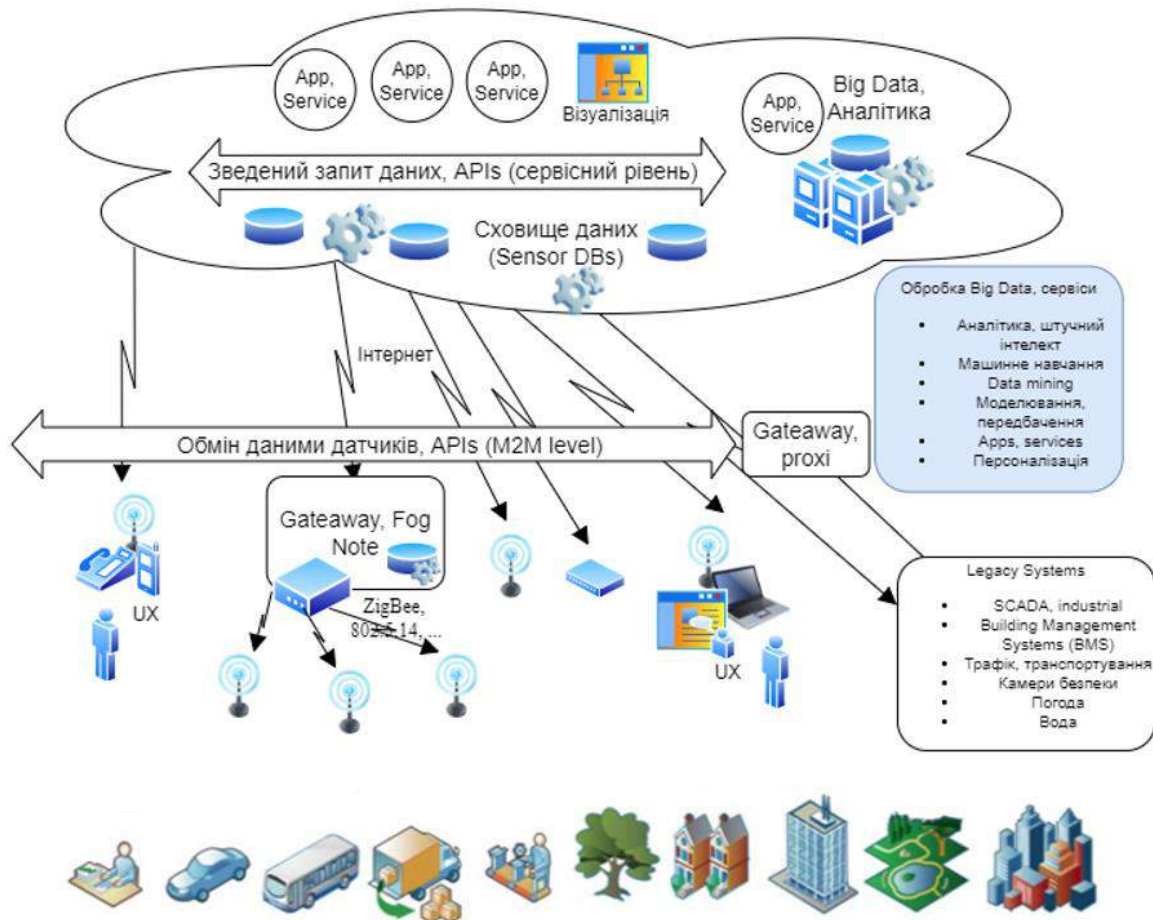


Рисунок 1.1 – IoT системи

Інші датчики підключаються до решти IoT системи та Інтернету за допомогою посередників, наприклад, шлюзів. Шлюзи і туманні вузли це зазвичай більш потужні пристрої, підключені до наборів менш потужних датчиків, приєднаних до них через локальні мережі, часто бездротові, такі як ZigBee, варіанти 802.14 мереж, Bluetooth і Wi-Fi з низькою потужністю.

Шлюзи забезпечують широку зону підключення та обробку для підключених датчиків, які можуть бути у формі перетворення протоколів, збереження та фільтрування даних, обробки подій та аналітики.

Смартфони відіграють важливу роль в IoT системі. Вони можуть бути задіяні як розумні підключені датчики, шлюзи і пристрої взаємодії з користувачем.

Телефони містять різноманітні вбудовані датчики: зчитувач відбитків пальців, датчик тиску, датчик світла, датчик Холла, барометр, термометр, акселерометр, гіроскоп, глобальну систему позиціонування (GPS) – яку можна використовувати для повідомлення про місцезнаходження пристрою. Ці датчики є програмно доступними для додатків і тому підключені до Інтернету або мають можливість підключення до нього.

Багато інших додатків, які зчитують локацію або додатків з доповненою реальністю (AR) можуть працювати шляхом поєднання їх з даними інших датчиків телефону, таких як камери та мікрофон.

«Back-end» обробка, яка часто виконується віддаленими серверами в системах IoT, зображена схематично загальною хмарою. Хмара – верхній шар в ієрархії системи Інтернету речей, де дані з різних джерел агрегуються та обробляються для оптимізації та виявлення глобальних тенденцій та відносин. Залежно від їх характеру та поточних вимог, вхідні дані датчиків і події можуть оброблятися в режимі реального часу у вигляді потоків, які зберігаються для наступної обробки або архівування.

Крім підтримки та запуску програм, IoT хмара також може містити деякі поширені сервіси: велике сховище, системи аналітичної обробки, візуалізацію даних і створення графіків, а також додаткові функції, такі як безпека.

Зазвичай алгоритми машинного навчання (ML) і штучного інтелекту (AI) працюють у хмарі з великими об'ємами даних. Такі алгоритми працюють краще на великих наборах даних, що покращує їх навчання і здатність до передбачення.

З точки зору типових потоків даних IoT, датчики контролюють стан фізичного світу та забезпечують відповідне цифрове зчитування цих даних. Програми, які обробляють дані датчиків і реагують на них, можуть перебувати у хмарі або поширювати свої функції через декілька компонентів системи.

Таким чином архітектура IoT дає змогу безперервно відчувати та вимірювати стан і поведінку фізичного світу, реагувати на них, та покращувати аспекти, які можна контролювати. Вона є функціональною, масштабованою,

гнучкою, економічно ефективною та може запропонувати різноманітні рішення для різних галузей.

1.3 Методи аналізу даних розпізнавання людської діяльності

Оскільки IoT є основним джерелом нових даних, методи отримання даних і машинного навчання зробили великий внесок у створення більш інтелектуальних додатків IoT. Ці методи складаються з різноманітних алгоритмів, відповідних для різних областей.

1.3.1 Штучні нейронні мережі

Штучні нейронні мережі вперше були представлені в сучасній формі Розенблатом, 1958 як концепція, що має назву «одношаровий перцептрон», відома як найпростіша форма нейронної мережі. Основна ідея полягає у відтворенні того, як працює мозок людини. Біологи виявили, що клітини мозку складаються з мережі нейронів, які з'єднані один з одним за допомогою синапсів: синапси передають інформацію через хімічну передачу, що призводить до реакції тільки в разі комбінованого імпульсу. Перцептрон моделюється подібно до цього відкриття. На рисунку 1.2 показано схематичне представлення нейронної мережі з прямим зв'язком.

Один нейрон отримує сигнали від кількох інших нейронів і, на основі зваженої суми сигналів, нейрон може відправити сигнал від себе.

Перцептрон використовує цю ідею для формування своїх елементарних структурних блоків. Кілька вхідних даних зважуються, а потім підсумовуються. Результат порівнюється з граничною межею. Згенерований прогнозний вихід з цього порогу двійкова перевірка: 1, якщо результат перевищує межу, 0, якщо він нижче.

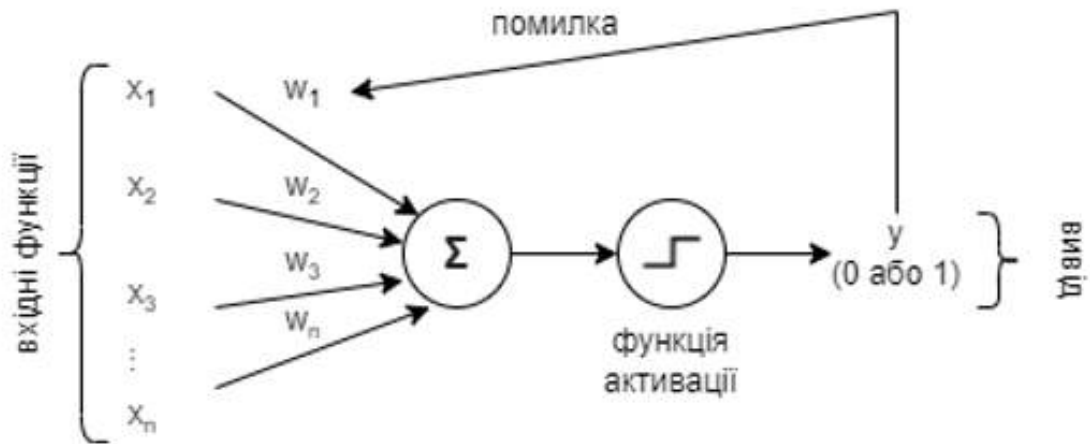


Рисунок 1.2 – Схематичне представлення нейронної мережі з прямим зв'язком

Основними компонентами сучасної нейронної мережі є один вхідний шар, один або кілька прихованих шарів і один вихідний шар (рис. 1.3).

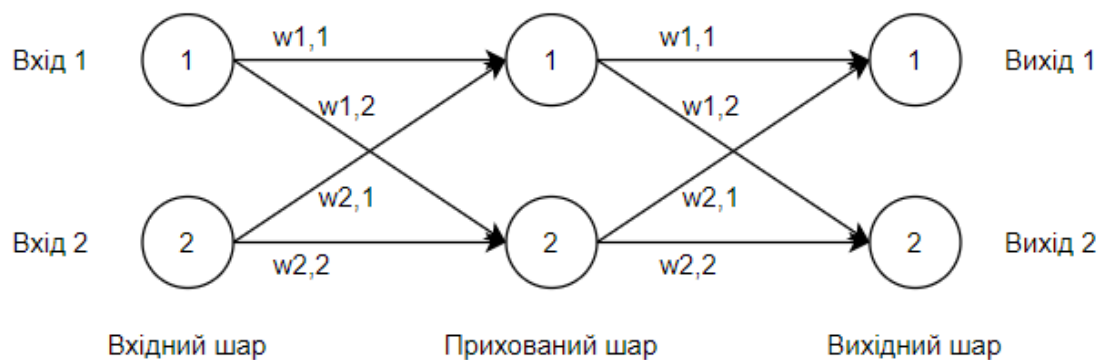


Рисунок 1.3 – Побудова нейронної мережі з одним прихованим шаром

Взаємозв'язані нейронні одиниці вивчають і виводять правила з даних спостережень. Шари з'єднані вагами, які пом'якшують вхідні імпульси.

Вхідний рівень не виконує жодних операцій над даними, він просто подає дані до першого прихованого шару.

Прихований шар приймає вхідні дані, множить ваги на вхідні дані, створює суму і використовує функцію активації для генерування вихідних даних нейрона.

Вихідний рівень перетворює дані з останнього прихованого шару в прогнозований вихід для класифікації. Нейронна мережа лише з одним прихованим шаром є неглибокою нейронною мережею.

Особливою підсферою машинного навчання є глибинне навчання. Глибинне навчання нейронних мереж створює конвеєр різноманітних простих перетворень, використовуючи для навчання більше ніж один прихований шар між входом і виходом. Більш глибокий рівень абстракції ознак встановлюється на етапі навчання.

Нейронні мережі глибинного навчання визначаються як багат шарові перцептрони. Популярність глибинного навчання зростає, оскільки проблеми з великими обсягами даних з'являються все частіше в сфері IoT, а такий тип машинного навчання показує високу ефективність при роботі з великою кількістю вхідних даних.

1.3.2 K-найближчі сусіди

Метод K-найближчого сусіда – це простий, але ефективний алгоритм машинного навчання для класифікації та розпізнавання образів, який широко використовується.

Навчальні кортежі описуються N атрибутами. Кожен кортеж представлений як крапка в N-вимірному просторі атрибутів. Алгоритм вчиться класифікувати за аналогією. Подібні кортежі групуються разом формуючи простір. Алгоритм розміщує невідомий кортеж, найближчий до навчальних кортежів K-найближчих сусідів, у шаблонний простір.

Алгоритм застосовує метричну функцію відстані, наприклад, евклідову відстань, щоб знайти K-найближчого сусіда і класифікувати невідомий кортеж у найбільш загальний клас серед його K-найближчих сусідів.

Кортеж класифікується як його єдиний найближчий навчальний кортеж у шаблонному просторі, коли $K=1$.

Значення K збільшується зі збільшенням розміру набору навчальних кортежів.

1.3.3 Support Vector Machine (SVM)

Support Vector Machine – це класифікуючий метод, який створює розділову гіперплощину для поділу нанесених точок даних. Алгоритм застосовний для обох лінійних та нелінійних наборів даних.

Відповідно до цього підходу, навчальні дані перетворюються за допомогою нелінійного відображення у вищий вимір, у якому машина виділяє лінійну оптимальну гіперплощину, що пронизує дані двох класів. У процесі пошуку використовуються гіперплощина, опорні вектори та поля.

Опорні векторні машини відомі своєю підвищеною точністю класифікації, оскільки вони здатні моделювати нелінійні межі рішення, але ця точність пов'язана з повільною тривалістю навчання.

У порівнянні з іншими методами класифікації, цей менш сприйнятливий до переобладнання. Лінійний алгоритм опорної векторної машини виділяє пряма лінія, щоб класифікувати точки даних, якщо вони є лінійно роздільні.

Нелінійно роздільні точки даних можна класифікувати шляхом розширення лінійних для пошуку нелінійних гіперпросторів.

1.3.4 Random Forests (RF)

Алгоритм Random Forests заснований на методі ансамблевого машинного навчання. Ліс складається з дерев передбачень.

Вузли кожного дерева розбиваються за допомогою значень випадкового вектора, відібраного спеціально для цього дерева, тоді як випадковий вектор слідує розподілу з усіх лісових дерев (рис. 1.4).

Додатковими ключовими факторами, що впливають на похибку узагальнення алгоритму, є міцність окремих дерев та їх взаємовідносини.

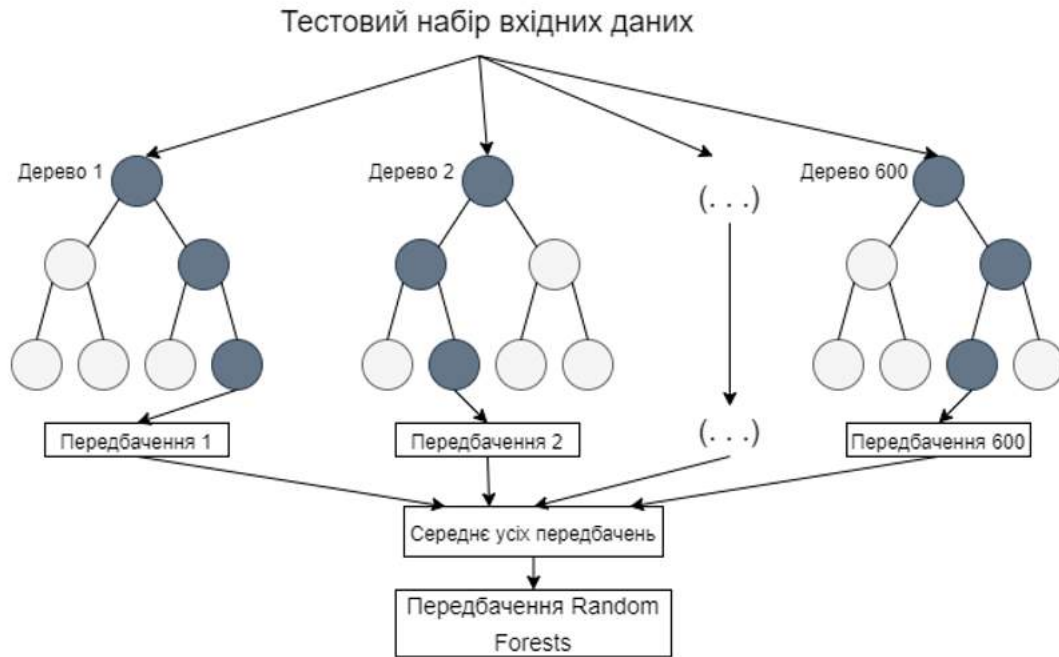


Рисунок 1.4 – Формування передбачення методом Random Forests

Розбиття кожного вузла на випадково вибрані набори функцій створює частоту помилок, яка може збільшити стійкість до шуму.

Оцінюється точність моделі через оцінки, які відстежують помилки, силу та кореляцію.

Ці помилки відображають реакцію на збільшення кількості розділювальних функцій. Також внутрішні оцінки вказують на змінну важливість. Вище згадані поняття також застосовні до регресії.

1.3.5 CART (classification and regression tree)

CART – це алгоритм дерева рішень, який використовує стратегію скорочення мінімальних витрат і складності для створення моделей класифікації та прогнозування.

У дереві рішень вузли розбиваються на підвузли на основі порогового значення атрибута. Алгоритм CART робить це шляхом пошуку найкращої однорідності для підвузлів за допомогою критерію Gini Index.

Кореневий вузол береться як навчальний набір і розбивається на два, враховуючи найкращий атрибут і порогове значення. Підмножини також розбиваються з використанням цієї логіки.

Це триває до тих пір, поки на дереві не буде знайдено останню чисту підмножину або максимальну кількість листків у цьому зростаючому дереві. Цей процес відомий під назвою «обрізка дерев».

1.3.6 Hidden Markov Model (HMM)

HMM – це ймовірнісна модель, яка використовується для пояснення чи виведення ймовірнісної характеристики будь-якого випадкового процесу.

В основі лежить твердження, що спостережувана подія не буде відповідати її покроковому статусу, а буде пов'язана з набором розподілів ймовірностей. Система, яка моделюється, вважається ланцюгом Маркова, і в цьому процесі є деякі приховані стани, які залежать від неї.

Основна мета HMM – дізнатися про ланцюг Маркова, спостерігаючи за його прихованими станами. Розглядаючи процес X з прихованими станами Y , HMM затверджує, що для кожної позначки часу розподіл ймовірності Y не повинен залежати від історії X відповідно до цього часу.

HMM використовується для опису еволюції спостережуваних подій, які залежать від внутрішніх факторів, що не спостерігаються безпосередньо.

Це клас ймовірнісних графічних моделей, які дозволяють передбачити послідовність невідомих змінних з набору спостережуваних змінних.

1.4 Огляд відомих методів та засобів розпізнавання людської діяльності штучним інтелектом в контексті IoT

Огляд літератури показав, що проблемі розпізнавання людської діяльності системами IoT присвячується значна увага серед науковців. Основними методами вирішення є Finite State Machine (FSM), Weighted Finite

Automata (WFA), приховані ланцюги Маркова, наївний баєсів класифікатор, Support Vector Machine, Random Forests, бутстрепова агрегація (табл. 1.1). Розглянемо детальніше дані методи.

У роботі [2] представлено предметне рішення проблеми, засноване на знаннях. В даному рішенні використовувались лише вхідні дані від двійкових датчиків з чотирьох джерел даних, включених до набору даних.

В документах стверджується, що звичайні алгоритми машинного навчання, такі як Support Vector Machine, не враховують тимчасові зв'язки між видами діяльності. Представлене рішення представляє собою експертну систему, схожу на скінченний автомат з точністю 81,3%.

Скінченний автомат моделює послідовність станів і переходи між цими станами. Стани представляють діяльності. Цей тип системи має перевагу в тому, що виведені правила експертної системи легко зрозуміти людям. Особливо в секторі електронної охорони здоров'я користувачі системи повинні мати можливість зрозуміти та відтворити як була проведена класифікація. Проблема такого типу рішення полягає в тому, що коли мешканець розумного будинку змінює щоденні звички, в системи з'являються труднощі з правильною класифікацією видів діяльності.

Ще одна проблема з послідовностями і даним набором даних полягає в тому, що дані були зібрані в контрольованому середовищі. Особа, яка тестувалась, могла поводитись інакше, тому що їй було відомо, що її діяльність фіксується. Крім того, люди можуть змінювати свої звички або їх дії можуть бути перервані. Це означає, що послідовності, змодельовані у скінченному автоматі могли би вийти з ладу і не змогли би більше бути класифікованими.

Іншим фактором, який згадали автори, є те, що набір даних досить малий, щоб забезпечити надійний ввід для навчання моделі.

Підхід, запропонований у роботі [3], описує комбінацію ймовірнісної та описової моделі.

Таблиця 1.1 – Таблиця публікацій, присвячених розпізнаванню людської діяльності системами IoT, упорядкована за появою в тексті

| Автор | Рік | Алгоритм | Особливості | Точність тестування |
|-------------------------|------|--|---|---------------------|
| Карвонен і Клейко | 2018 | Експертна система, схожа на скінченний автомат | Бінарні датчики | 81,3% |
| Лаго та Іноуе | 2018 | Гібридна модель (приховані ланцюги Маркова і логічна модель) | Бінарні датчики та датчики наближення | 45,0% |
| Саломон і Тірнеуке | 2018 | Weighted Finite Automata | Бінарні датчики, датчики наближення та датчики підлоги | 90,65% |
| Хіменес і Секо | 2018 | Багатоподієвий наївний баєсів класифікатор | Бінарні датчики, датчики наближення, підлоги та прискорення | 60,5% |
| Раззак і ін. | 2018 | FilteredClassifier (Weka tool) | Бінарні датчики, датчики наближення, підлоги та прискорення | 47,0% |
| Церон, Лопес і Ескоф'єр | 2018 | J48, Ib1, Support Vector Machine, Random Forests, AdaBoostM1 (найкращий результат) і бутстрепова агрегація | Бінарні датчики, датчики наближення, підлоги та прискорення | 62,77% |

Зазначений підхід застосовував дані двійкового датчика і датчика наближення як вхідні дані для процесу розпізнавання активності. Імовірнісна

модель заснована на прихованому ланцюгу Маркова і представляє кожен стан діяльності як стан.

Переходи прихованого ланцюга Маркова моделюють послідовність дій. Оскільки особа, що тестується в лабораторії розумного будинку, перемикається з однієї діяльності на іншу, результатом є послідовність спостережень, що підходить для застосування прихованих ланцюгів Маркова.

Ймовірності викидів для ланцюга Маркова були розраховані за допомогою нейронної мережі, враховуючи середню тривалість кожної з 24 діяльностей. Таким чином можна було розрахувати ймовірність перебування у тій самій діяльності або зміну на іншу діяльність.

Описова частина моделі спиралася на словесний опис кожної діяльності, який був заданий набором даних. Логіка була реалізована за допомогою класів Java, але через складність перетворення описів діяльності в логіку, були пропущені діяльності «мий руки» та «мий посуд».

На основі опису кожного виду діяльності, можна було визначити події та використати їх, щоб позначити кінець і початок кожної діяльності. Загальна точність склала 45,0%.

Гібридна модель погано виконувала дії з короткою тривалістю, наприклад, «покласти прання в пральну машину» або «відвідати SmartLab». Класифікатор не міг розпізнати жодну з цих діяльностей.

Перевага моделі опису полягає в тому, що вона може класифікувати дії, яких модель ніколи не бачила раніше, оскільки модель спирається на правильний опис класу.

Проте, якщо датчик, який позначає завершення дії, не спрацює, закінчення діяльності не розпізнається, і зміна стану повністю пропускається. Тоді модель прихованого ланцюга Маркова не може оновити стан і змінити на найбільш ймовірну подальшу діяльність.

Авторами в роботі [4] представлено класифікацію часових рядів із заданим набором даних за допомогою напівкерovanого алгоритму машинного

навчання з Weighted Finite Automata (схожими на скінченний автомат з роботи [2]) та регулярними виразами.

У підході запропоновано додатково до бінарних датчиків використати дані датчика підлоги, щоб визначити положення досліджуваного в кімнаті. Один автомат був навчений для кожного сегменту дня, щоб визначати потік діяльності в сегментах. Вага між станами скінченного автомата розраховувалась за тим, як часто конкретний шлях обирався.

При напівконтрольованому процесі опис кожної діяльності було отримано шляхом спочатку навчання автомата кожної дії і перетворення в регулярний вираз, а потім ручного налаштування результату. Регулярний вираз позначав, які датчики зазвичай активні і в якому порядку для кожної окремої діяльності.

Комбінуючи ймовірності переходу автомата для кожного сегмента та регулярні вирази, які представляють кожен вид діяльності, ці види діяльності можуть бути класифіковані. Результат запропонованої системи має точність 90.65%. Помилки в моделі передбачення зазначеного підходу в основному були викликані неправильними прогнозами початкового та кінцевого часу діяльності. Найбільш ймовірно, що це сталося через помилки, зроблені людьми, які обробляли результат та додали шум в набір даних.

У роботі [5] запропоновано наївний баєсів класифікатор для оцінки того, які діяльності були виконані. Наївний баєсів класифікатор – це імовірнісний метод, заснований на теоремі Баєса. Цей підхід збирає вхідні дані зі всіх чотирьох джерел даних. В результаті виконання класифікації з тестовим набором даних досягнуто точності 60.5%.

У дослідженні [6] досягнуто 94.0% точності для навчальних даних та точності 47.0% для набору тестових даних, використовуючи FilteredClassifier від Weka Tool Weka 3: Data Mining Software in Java 2019.

Велика різниця між навчанням і тестуванням є ознакою переобладнання класифікатора. Ця система використовувала всі доступні дані датчиків для навчання моделі. FilteredClassifier було реалізовано з стандартним фільтром

StringToWordVector і з Random Forests як базовим класифікатором. Фільтр Weka взяв на себе попередню обробку – очищення набору даних та зміни його для класифікатора.

Авторами роботи [7] представлено підхід з використанням Міжгалузевого стандартного процесу для інтелекту даних (CRISP-DM). Є шість етапів цієї методології за [8]:

1. Розуміння бізнесу.
2. Розуміння даних.
3. Підготовка даних.
4. Моделювання.
5. Оцінка.
6. Розгортання.

Система, представлена у [7], бере вхідні дані з потоків подій двійкових датчиків, дані близькості, дані прискорення зі смарт-годинника і дані про місцезнаходження зі смарт-підлоги. Використовувалися моделі J48, I_b1, Support Vector Machine, Random Forests, AdaBoostM1 і бутстрепова агрегація. З цього списку AdaBoostM1 дала найкращі результати.

Проблема підходу в тому, що набір даних має класовий дисбаланс: не вистачає достатньо зразків діяльностей «мити посуд» та «грати у відеоігри», щоб надійно передбачити ці діяльності. Наприклад, були підсумовані «вечеря», «обід» і «сніданок» як діяльність «вживати їжу». Цей крок також було виконано до діяльностей, пов'язаних з приготуванням страв для цих прийомів.

Тілесні рухи цих видів діяльності не розрізняються, різний лише час доби. Злиття цих діяльностей підвищило точність класифікації на 13%.

Незважаючи на те, що 10-кратна перехресна перевірка на даному навчальному етапі призвела до точності класифікації 92.1%, точність класифікації тестового набору даних становила лише 60.1%, 62.77% без нульового класу (клас, що використовується для зразків, які не пов'язані з жодним із 24 визначених класів).

В результаті досліджень [6] основна проблема зі стандартним шаблоном методів розпізнавання активності на основі датчиків (Support Vector Machine, приховані моделі Маркова, наївний баєсів класифікатор) полягає в тому, що ознаки завжди добуваються за допомогою ручних підходів, заснованих на досвіді і знаннях предметної області науковця, який програмує модель.

Вилучені ознаки в кінцевому підсумку виявляються досить неглибокими, наприклад, статистичні особливості, як середнє значення або дисперсія. Цього обмеження може бути недостатньо для вивчення складних видів діяльності.

Ще одна проблема з розпізнаванням діяльності полягає в тому, що зазвичай маркований набір даних невеликий через те, що його хтось повинен був позначити, і потрібна велика кількість даних, щоб навчитися діяльності, яка не завжди може давати однакові дані датчика однаковим чином.

Крім того, більшість стандартних алгоритмів розпізнавання образів навчаються на статистичних даних, але дані датчиків для розпізнавання активності надходять у послідовності спостережень з датчиків. Кожен зразок має різні події датчика, які мають сенс лише тоді, коли їх розглядають послідовно.

Беручи до уваги природу набору даних цієї роботи, моделі глибокого навчання можуть виконувати вилучення абстрактних ознак і побудову моделі одночасно. Один із варіантів моделі глибокого навчання, що стосується проблемної області розпізнавання діяльності, є CNN. Використовуючи дані часових рядів, CNN можуть використовувати можливості виявлення локальних шаблонів [6].

Підхід [9] дозволив дослідити кілька алгоритмів глибокого навчання для розпізнавання людської діяльності за допомогою розумного одягу. Для порівняння нейронних мереж із прямим зв'язком, CNN та рекурентних нейронних мереж на основі LSTM були використані різні набори даних. Це дало змогу виявити, що двонаправлені LSTM працюють значно краще, ніж інші сучасні алгоритми класифікації.

1.5 Постановка задачі дослідження

Огляд існуючих методів та засобів розпізнавання людської активності в системах розумних будинків показав, що відомі рішення не дозволяють ефективно автоматизувати процес розпізнавання та мають низькі показники точності. Таким чином постає завдання розробки інтелектуальної системи та нового методу розпізнавання активностей людини в розумному будинку, що дозволили б підвищити ступінь автоматизації процесів та реагування на потреби мешканців. Для вирішення поставленого завдання необхідне виконання наступних етапів:

1. Проаналізувати існуючі методи, моделі та засоби класифікації людської діяльності в системах розумних будинків, виокремити недоліки.
2. Розробити модель процесу класифікації активностей людини у розумному будинку.
3. Розробити метод розпізнавання активностей людини в розумному будинку.
4. Розробити інтелектуальну систему прийняття рішень щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини.
5. На основі розробленої системи оцінити ефективність класифікації діяльності людини.

2 ПІДГОТОВКА ДАНИХ ТА МОДЕЛЬ ПРОЦЕСУ КЛАСИФІКАЦІЇ АКТИВНОСТЕЙ ЛЮДИНИ У РОЗУМНОМУ БУДИНКУ

2.1 Опис предметної області

Дослідження даної роботи проводиться над набором даних, зібраними в розумній лабораторії на Ubiquitous Computing and Ambient Intelligence Cup 2018 [10].

Використаємо цей набір даних для навчання моделі розпізнавання людської активності і розробимо систему, методи та інструменти якої дозволять створити рішення для інтеграції штучного інтелекту в розумних будинках.

Предметною областю є розумна лабораторія з датчиками, розміщеними всередині неї, для фіксації дій людини (рис. 2.1).

Розумна лабораторія відтворювала середовище реальної квартири та була розділена на п'ять різних зон: передпокій, кухню, вітальню і спальню з вбудованою ванною кімнатою.

В наборі даних, що розглядається, визначені 24 різні види повсякденної діяльності.

Датчики, які збирають дані пов'язані з цими видами, були розміщені в безпосередній близькості до спостережуваних об'єктів або прикріплені до них (табл. 2.1).

Для аналізу та наступної обробки доступні дані зібрані з чотирьох джерел даних:

1. 30 бінарних датчиків, які передають двійкове значення для магнітного контакту, руху або тиску.
2. Дані про наближення між розумним наручним годинником і набором з 15 маяків Bluetooth Low Energy (BLE), розташованих в лабораторії (рис. 2.2).
3. Дані прискорення з 3 осей, згенеровані розумним годинником.

4. Інформація про місцезнаходження, яка надається розумною підлогою, поділеною на модулі.

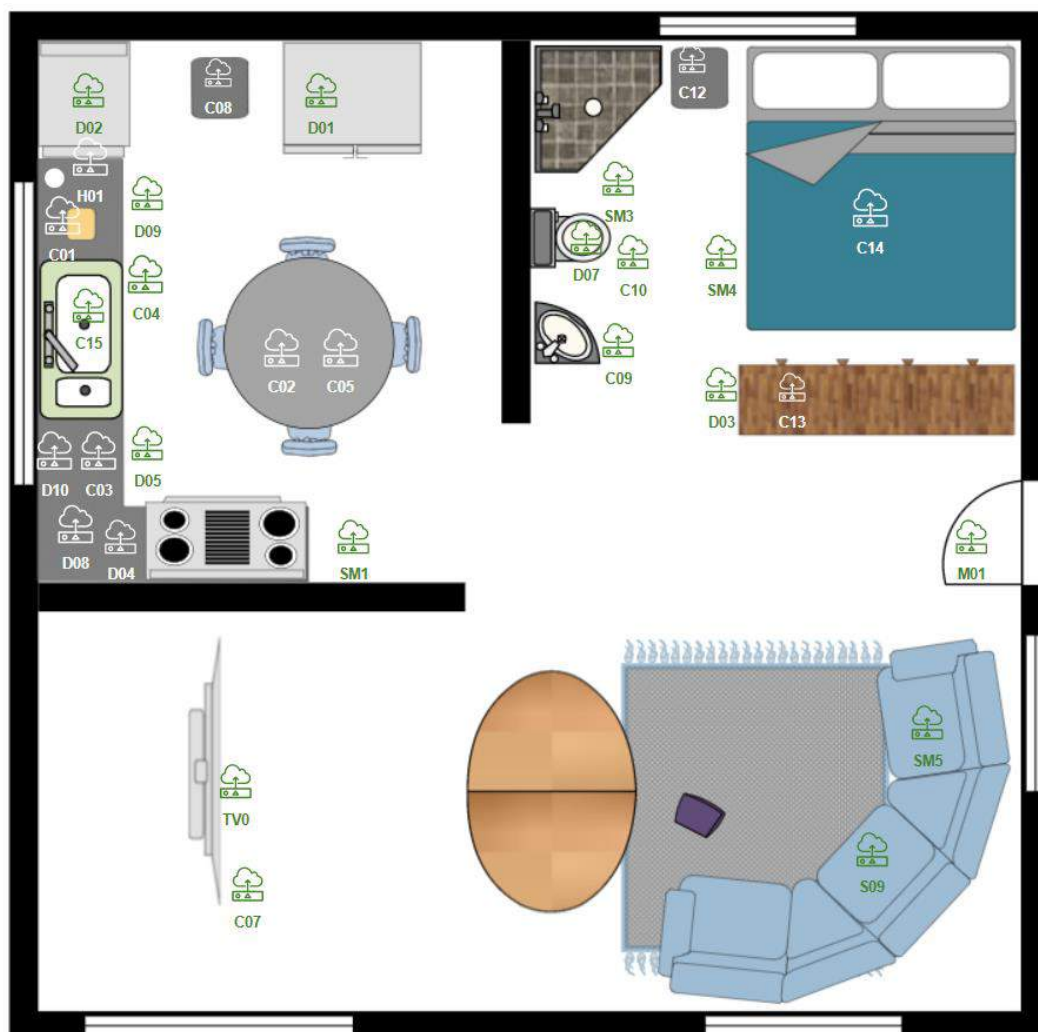


Рисунок 2.1 – Схематичне представлення розміщення бінарних датчиків в розумній лабораторії

Таблиця 2.1 – Позначення датчиків лабораторії розумного будинку

| Позначення | Спостережуваний об'єкт |
|------------|----------------------------|
| M01 | Двері |
| TV0 | Телевізор |
| SM1 | Датчик руху кухні |
| SM3 | Датчик руху ванної кімнати |

Кінець таблиці 2.1 – Позначення датчиків лабораторії розумного будинку

| Позначення | Спостережуваний об'єкт |
|------------|-------------------------|
| SM4 | Датчик руху спальні |
| SM5 | Датчик руху дивану |
| D01 | Холодильник |
| D02 | Мікрохвильовка |
| D03 | Шафа |
| D04 | Сервіз |
| D05 | Машина для миття посуду |
| D07 | Туалет |
| D08 | Кухонна шафа |
| D09 | Пральна машина |
| D10 | Буфетна |
| H01 | Чайник |
| C01 | Коробка з ліками |
| C02 | Блюдо з фруктами |
| C03 | Столове приладдя |
| C04 | Каструлі |
| C05 | Пляшка з водою |
| C07 | XBOX |
| C08 | Відро для сміття |
| C09 | Кран |
| C10 | Резервуар |
| C12 | Корзина для прання |
| C13 | Ящик для піжами |
| C14 | Ліжко |
| C15 | Кухонний кран |
| S09 | Диван |

В якості бінарних датчиків були використані бездротові датчики, які працюють за принципами протоколу Z-Wave. Наприклад, якщо пляшка води використовувалась, то її стан був визначений як «Відкрита», і іншому разі - «Закрита». Переміщення мешканців контролювалось за допомогою PIR-сенсорів, які працюють з основним протоколом ZigBee. Переміщення представлене двійковим значенням: «Рух» або «Без руху». М'які меблі оснащені текстильним шаром для спостереження за тиском, який передавав двійкове значення «Тиск присутній» або «Тиску немає» за допомогою протоколу Z-Wave.

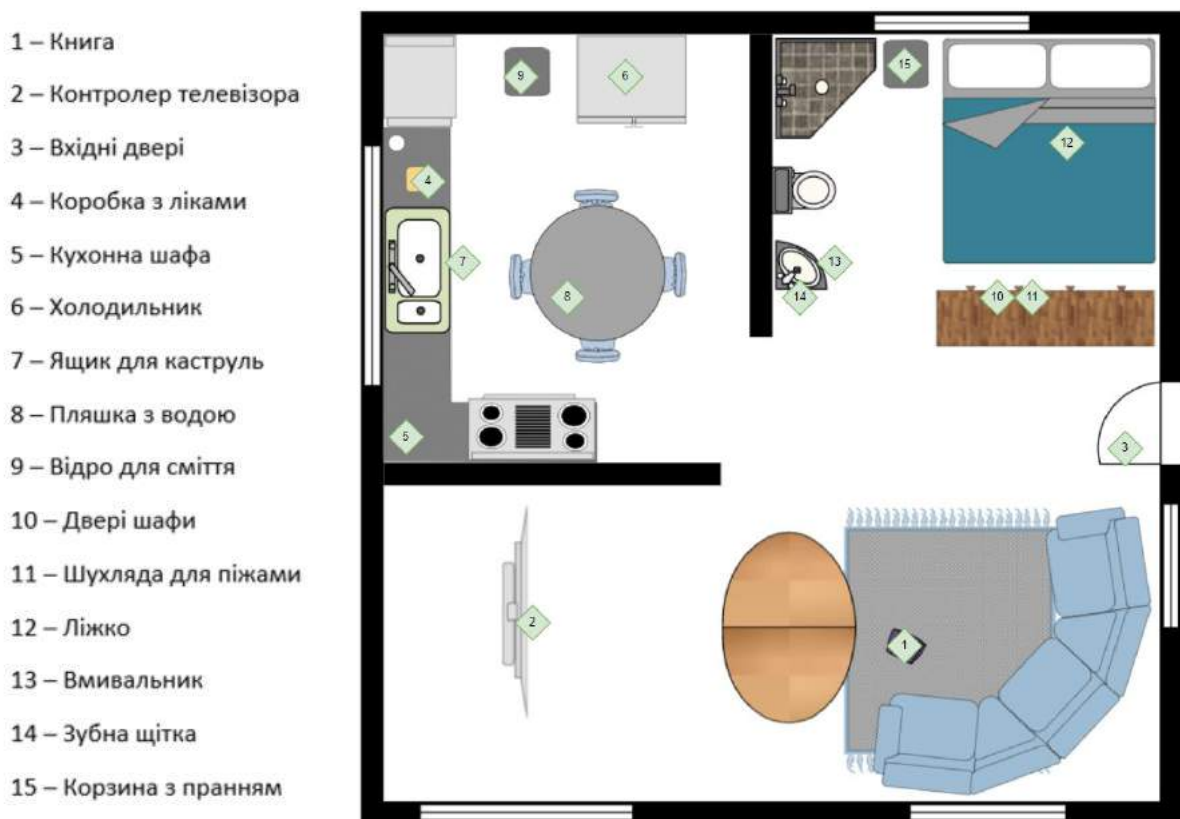


Рисунок 2.2 – Схематичне представлення розташування BLE маяків в розумному будинку

Дані підлоги зібрані за допомогою технології SensFloor, яка може бути встановлена майже під будь-яким типом покриття підлоги і відслідковує

переміщення людини. Набір даних SensFloor містить опис восьми сенсорних полів, ідентифікованих індивідуальним ідентифікатором (рис. 2.3).

Технологія працює, посилаючи невеликий електричний заряд через провідну тканину, що містить вбудовані сенсорні пластини (32 пластини на квадратний метр) і радіомодулі. Коли людина проходить через датчик, невелика зміна ємності заряду запускає систему.

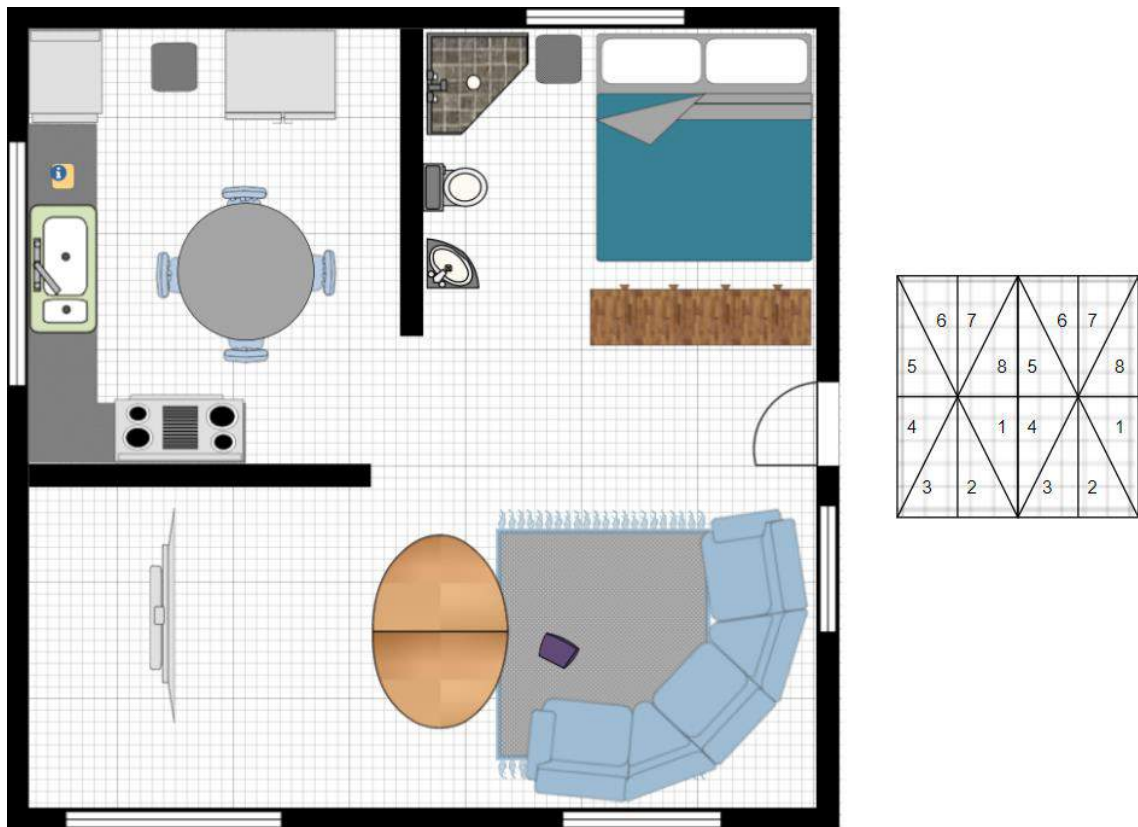


Рисунок 2.3 – Схематичне представлення макету модулів розумної підлоги у лабораторії розумного будинку

Дані про наближення були зібрані з 15 маяків BLE, які надсилали інформацію до розумного годинника. Для взаємодії з маяками на годиннику було встановлене спеціальне програмне забезпечення. Маяки BLE також можуть бути використані, щоб виконувати дії на основі місцезнаходження, коли людина з годинником входить у певну область.

Проводилось достатньо досліджень з таким же набором даних, однак кожне рішення стикається з проблемами при складних та більш реалістичних завданнях класифікації діяльності. В даній роботі розробимо алгоритм для підвищення ефективності класифікації та оптимізації розпізнавання людської діяльності IoT.

2.2 Функціональні та нефункціональні вимоги

Функції, які повинна забезпечувати система мають бути індивідуально адаптованими до мешканців розумного будинку, щоб підвищити якість життя за допомогою автоматизації. Розглянемо потенційний сценарій використання з якими може зіткнутися розумний дім.

Одним із можливих сценаріїв є сценарій відслідковування та нагадування системою розумного будинку про прийом ліків. Мешканець може приймати гормональні препарати для щитовидної залози або препарати для підтримання рівня цукру на постійній основі. Регулярний прийом таких ліків може бути критично важливим для здоров'я людини. Мешканець може забути про цей прийом, і в цьому випадку має спрацювати система сповіщень.

Система відслідкуватиме прийом ліків за допомогою даних розміщення розумної підлоги, магнітного контакту і датчика наближення коробки з ліками, та даних прискорення розумного годинника. При завершенні сегменту дня (ранок, обід або вечір), в який повинен був відбутись прийом ліків має існувати перевірка чи датчики коробки з ліками спрацювали, якщо ні – діяльність була пропущена.

У випадку пропуску система має нагадати про обов'язковий прийом ліків за допомогою сповіщення, надіслане на пристрій мешканця з встановленим програмним забезпеченням для цього.

Платформа IoT повинна відповідати також великій кількості нефункціональних вимог. Перш за все, повинен бути надійним алгоритм класифікації, який використовується для розпізнавання людської діяльності.

Надійність процесу прийняття рішень у системі особливо важлива, коли штучний інтелект повинен приймати рішення в ситуаціях, які мають значний вплив на життя людини, таких як, сфера охорони здоров'я. Тому рішення системи мають бути точними, прозорими та зрозумілими, чого важко досягти в існуючих алгоритмах штучного інтелекту.

Система повинна мати можливість обробляти дані в режимі реального часу та забезпечувати взаємну сумісність різних протоколів та стандартів. Додатки IoT створюються як взаємопов'язана мережа, щоб мати можливість реалізувати зв'язок між пристроями, службою зберігання даних і фактичною службою додатків.

У розумній лабораторії, дані якої використані в цій роботі, датчики виявляли дії та надсилали дані до місця призначення у вбудованій системі. Завдяки численним можливостям датчиків, які можна використовувати в контексті розпізнавання активності, система повинна мати можливість керувати зв'язком з низкою різних пристроїв.

Оскільки була підключена велика кількість розподілених пристроїв, потрібно впровадити виявлення несправностей, щоб підготуватись до того, що датчики можуть надіслати неправильні дані або мають низький заряд акумулятора.

У роботі [11] визначені такі основні вимоги безпеки для системи IoT:

1. Доступність сервісу. Розумний дім має бути постійно доступним і, отже захищений від атак відмови в обслуговуванні.
2. Аутентифікація, авторизація, облік: лише люди, які володіють привілеєм на вхід до системи, повинні мати можливість нею користуватися.
3. Конфіденційність та цілісність даних: оскільки розумні будинки працюють з розумними даними, дуже важливо, щоб ці дані були захищені належними механізмами шифрування.
4. Енергоефективність: при роботі з пристроями, призначеними для тривалого використання, важливе використання енергоефективних продуктів.

Ще одним важливим атрибутом є масштабованість. У IoT масштабованість є значною проблемою через зростаючу складність використання великої кількості різних пристроїв у мережі та забезпечення потоку даних. Система IoT генерує величезну кількість інформації, тому для зберігання та обробки даних необхідно налаштувати відповідну архітектуру.

Через велику кількість обчислень, які пов'язані з глибоким навчанням, для прискорення навчання алгоритму необхідне відповідне обладнання. Наприклад, замість центральних процесорів (CPU) можна використовувати графічні процесори (GPU), щоб зробити навчання швидшим.

Зручність використання панелей керування або інших типів входів для керування системою розумного дому має бути перевірена, наприклад, шляхом проведення тестів зручності з реальними користувачами.

2.3 Попередня обробка даних

Застосуємо попередню обробку даних, щоб підготувати їх до використання для моделі машинного навчання. Для цього відформатуємо та впорядкуємо вхідні дані.

Представимо процес попередньої обробки даних у вигляді послідовності кроків (рис. 2.4):

1. Синхронізація даних датчика у вибірки.
2. Перевірка даних на відповідність категорійним значенням.
3. Нормалізація ознак.
4. Очищення даних.
5. Розбиття набору даних.

Розглянемо детальніше кроки процесу попередньої обробки даних.

Крок 1. Синхронізація даних датчиків в вибірки, що представляють проміжок попередньо визначеного часу.

Цей крок необхідний, оскільки датчики мають різні частоти, на яких передаються дані і виникає проблема синхронізації бездротових мереж при

об'єднані даних датчика, отриманих з різних джерел. Для цього оберемо оптимальний проміжок короткого часу, на якому об'єднаємо кілька значень датчика до одного для кожного джерела.



Рисунок 2.4 – Алгоритм попередньої обробки даних перед подачею в алгоритм машинного навчання

Крок 2. Перевірка даних на відповідність категорійним значенням.

Дані бінарних датчиків представлені у двійковому форматі, що дозволяє одразу використовувати їх, та більшість даних не можуть бути використані моделлю в їх початковому вигляді. Приведемо ці дані до категорійних значень.

Дані про підлогу перетворимо в одну функцію для кожного модуля плитки підлоги і встановимо в 1, якщо на часовому проміжку надходили дані від модуля, в 0 – якщо ні.

Оскільки дії, пов'язані з різними прийомами їжі протягом дня, сильно залежать від сегмента (ранок, день, вечір), сегмент додамо як власну функцію, яка буде повертати значення 0, 1 або 2. Таким чином модель зможе розрізнити сніданок, обід і вечерю.

Дані про наближення виміряні сигналом Received Signal Strength Indicator (RSSI) таким чином, що будь-яке вимірювання, яке перевищувало -97, означало близькість. Чим вищий показник, тим ближча відстань до об'єкта розміщеного всередині квартири. Зведемо це значення в масштаб від 0 до 1, де 1 – означає найбільшу близькість, на наступному кроці.

Крок 3. Нормалізація.

Виконаємо нормалізацію на небінарних даних, таких як дані прискорення, за допомогою MinMaxScaler для підвищення ефективності моделі машинного навчання, а також зменшення часу необхідного для вивчення даних.

Небінарні дані мають різну шкалу, і це може призвести до меншої ефективності атрибута з меншою шкалою, через те, що інші атрибути мають значення у більшому масштабі. Використаємо формулу нормалізації для кожного значення:

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)}, \quad (2.1)$$

де x – значення з даних.

Це перетворить всі значення в діапазон $[0,1]$, що означає, що мінімальне та максимальне значення будуть 0 та 1 відповідно.

Крок 4. Очищення даних.

Деякі вибірки в початкових даних позначені як «без активності», хоча і містять дані датчиків. Вилучимо зразки з цією міткою, оскільки вони можуть погіршити якість навчання моделі.

Крок 5. Розбиття набору даних.

Розіб'ємо набір даних для моделі машинного навчання на два окремих набори – навчальний набір і тестовий набір методом k -кратного перехресного затвердження:

1. Довільно розділимо набір даних на k груп приблизно однакового розміру.
2. $k-1$ частини використаємо для навчання моделі, а одну k частину використаємо для випробування і оцінки ефективності.
3. Повторимо процедуру k разів, щоб отримати k кількість оцінок продуктивності для кожної ітерації.
4. Результируючу ефективність обрахуємо як середнє значення k кількості оцінок.

В результаті виконання кроків, описаних вище, покращили якість даних, забрали суперечливості, що могли порушити загальне навчання моделі і призвести до помилкових прогнозів. Отримали дані, які готові для побудови моделі.

2.4 Модель процесу класифікації життєдіяльності людини

Оскільки об'єктом дослідження є процес класифікації видів діяльності, зафіксованих у розумних будинках, то важливою задачею є розроблення його моделі, що дозволить формалізувати основні його складові. Для побудови моделі процесу класифікації видів діяльності, зафіксованих у розумних будинках використаємо теоретико-множинний підхід.

Представимо модель процесу розпізнавання людської діяльності на основі використання Naive Bayes Classifier у вигляді кортежу:

$$M = \langle K, C, T, S, \gamma \rangle, \quad (2.2)$$

де K – модель класифікатора;

C – множина видів діяльності;

T – багатовимірний часовий ряд;

S – множина сенсорів розумного будинку;

γ – функція визначення відповідності до певного виду діяльності.

Розглянемо детальніше складові даної моделі процесу розпізнавання людської діяльності.

Визначення виду діяльності відбувається у відповідності до одного з 24 типів видів діяльності людини з вхідних даних. Представимо множину видів повсякденної діяльності:

$$C = \{C_i\}_{i=1}^{24}. \quad (2.3)$$

Представимо множину сенсорів розумного будинку наступним чином:

$$S = \{s_i\}_{i=1}^4 = \langle s_f, s_o, s_a, s_m \rangle, \quad (2.4)$$

де s_f – датчики розумної підлоги;

s_o – датчики, що розміщені на об'єктах всередині розумного будинку;

s_a – дані прискорення, згенеровані розумним годинником на руці людини;

s_m – бінарні датчики, які відслідковують наявність магнітного контакту.

Задамо модель класифікатора на основі Naive Bayes Classifier у вигляді наступного кортежу:

$$K = \langle \gamma, X, Av \rangle, \quad (2.5)$$

де γ – функція визначення людської діяльності на основі поданих ознак;

X – вхідний вектор ознак;

Av – наявність людей у розумному будинку (0 або 1), у разі нульового значення жодна діяльність не може бути визначена.

Функція визначення активності на основі вхідного вектора ознак побудована на основі теореми Баєса:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)}, \quad (2.6)$$

де $P(C|X)$ – апостеріорна ймовірність даного виду діяльності при даному наборі ознак X ;

$P(C)$ – апріорна ймовірність цього виду діяльності;

$P(X|C)$ – правдоподібність;

$P(X)$ – апріорна ймовірність даного значення ознаки.

Вхідний вектор ознак визначимо на основі вхідних даних від різних датчиків:

$$X = \{x_i\}_{i=1}^5 = \langle x_z, x_o, x_D, x_{mv}, x_{mg} \rangle, \quad (2.7)$$

де x_z – дані про зону розміщення людини в розумному будинку;

x_o – дані про об'єкт, на якому спрацював датчик (чайник, телевізор, ліжка тощо);

x_D – дані про період доби, в який була виконана діяльність (ранок, обід або вечір);

x_{mv} – наявність руху під час виконання діяльності (активний рух або знаходження на одному місці);

x_{mg} – бінарні дані про наявність або відсутність магнітного контакту.

Представимо дані про розміщення людини в розумному будинку у вигляді кортежу:

$$x_z = \langle F, \delta \rangle, \quad (2.8)$$

де F – дані про переміщення датчиків розумної підлоги;

δ - функція визначення належності пройдених модулів підлоги до одного з п'яти можливих варіантів приміщень: передпокою, кухні, вітальні, спальні або ванної кімнати.

Дані про переміщення збираються з 38 модулів розумної підлоги з датчиками і визначаються як:

$$F = \{f_i\}_{i=1}^{38}, \quad (2.9)$$

де f – представлення одного модуля розумної підлоги.

Дані про період доби розраховуються на основі часу виконання діяльності:

$$x_D = \nu(t), \quad (2.10)$$

де t – час виконання діяльності;

ν – функція визначення відповідності часу виконання діяльності певному періоду доби (ранковому часу - 0, обідньому - 1 або вечірньому - 2).

Використаємо вектор та ланцюгове правило теорії ймовірності і отримаємо теорему Баеса в розширеному вигляді:

$$P(c|x_1, \dots, x_n) = \frac{P(x_1|c)P(x_2|c)\dots P(x_n|c)P(c)}{P(x_1)P(x_2)\dots P(x_n)}. \quad (2.11)$$

Для всіх варіантів поєднань ознак діяльності знаменник не зміниться, відкинемо його і введемо в рівняння пропорційність:

$$P(c | x_1, \dots, x_n) \propto P(c) \prod_{i=1}^n P(x_i | c). \quad (2.12)$$

Отримаємо функцію визначення людської діяльності на основі використання Naïve Bayes Classifier.

$$\gamma = \arg \max_c P(c) \prod_{i=1}^n P(x_i | c). \quad (2.13)$$

Результатом функції буде вид діяльності, який за вхідним набором ознак підходить з найбільшою ймовірністю.

Правдоподібність належності ознаки до певного виду діяльності розрахуємо як:

$$P(x_i, c) = \frac{N_i}{N}, \quad (2.14)$$

де N_i – частота, з якою ознака зустрічається з видом c у вибірці;

N – частота, з якою зустрічається вид c у вибірці.

Оскільки неможливо навчити систему всім можливим поєднанням ознак та видів, в початковому наборі даних може виникнути проблема нульової частоти, яка є поширеним недоліком даного алгоритму. Коли ознака не відповідає виду діяльності оцінка ймовірності дорівнює 0. В результаті добутку всіх ймовірностей отримаємо 0. Це призведе до неможливості класифікації через нульову ймовірність для кожного виду діяльності.

Для вирішення цієї проблеми використаємо згладжування Лапласа та штучно додамо по одиниці до кожної частоти:

$$P(x_i, c) = \frac{\alpha + N_i}{\alpha + N}, \quad (2.15)$$

де α – параметр для згладжування, дорівнює 1.

Подемо алгоритм навчання класифікатора у вигляді кортежу:

$$L = \langle D, \tau, \mathcal{G} \rangle, \quad (2.16)$$

де D – набір початкових даних, в якому для певного набору ознак визначений певний вид діяльності (X, C) , X – вектор ознак, C – вид діяльності, $C = \{C_1, C_2, \dots, C_{24}\}$;

τ – функція перетворення початкових даних для подальшої обробки;

\mathcal{G} – функція побудови моделі та навчання класифікатора.

У функції перетворення даних використаємо k -кратне перехресне затвердження:

$$\tau = \frac{D}{k} = \langle A, B \rangle, \quad (2.17)$$

де k – кількість рівних частин, на які розбиваються початкові дані;

A – тренувальний набір, $k-1$ частина даних;

B – випробувальний набір, одна k -підвибірка.

Представимо функцію навчання класифікатора як:

$$\mathcal{G} = \langle \phi, \mu \rangle, \quad (2.18)$$

де ϕ – процес перехресного затвердження, що повторюється k разів. Щоразу використовується кожна з k підвбірок як затверджувальна;

μ – функція оцінки ефективності моделі на наборі даних.

Результуюча ефективність обраховується як середнє з k результатів. Для цього вимірювання використаємо середньоквадратичну похибку:

$$\mu = (1/k) * \sum MSE_i, \quad (2.19)$$

де k – кількість частин;

MSE_i – середньоквадратична похибка i -тої ітерації.

Щоб оцінити ефективність моделі потрібно виміряти наскільки добре передбачення моделі відповідають спостережуваним даним. Для цього розрахуємо похибку i -тої ітерації:

$$MSE = (1/n) * \sum (y_i - f(x_i))^2, \quad (2.20)$$

де n – загальна кількість спостережень;

y_i – отримане значення i -того спостереження;

$f(x_i)$ – передбачуване значення i -того спостереження.

2.5 Висновки

Проведено аналіз початкових даних, зібраних в розумній лабораторії з встановленими датчиками та маяками для відслідковування діяльності людини, описано предметну область та визначено основні вимоги для оцінки якості роботи системи. Виконано попередню обробку даних, створено навчальну та тестові вибірки.

На основі оброблених даних датчиків розроблено модель процесу розпізнавання людської активності з застосуванням Naive Bayes Classifier, яка дозволяє визначити найбільш ймовірну активність, виконувану мешканцем розумного будинку за набором ознак.

3 МЕТОД РОЗПІЗНАВАННЯ АКТИВНОСТЕЙ ЛЮДИНИ В РОЗУМНОМУ БУДИНКУ

3.1 Метод розпізнавання активностей людини в розумному будинку на основі використання Naïve Bayes Classifier

Для розпізнавання активностей людини в розумному будинку запропоновано метод, в результаті якого буде отримано вид діяльності на основі поданих на вхід ознак. Узагальнену схему методу наведено на рис. 3.1, його кроки будуть розглянуті в наступних підрозділах цього розділу.

3.2 Обробка даних багатовимірних часових рядів

В цій роботі набір даних складається з часових рядів, зведених в один багатовимірний часовий ряд, який розглядає декілька змінних. Часовий ряд спостерігає одну або кілька випадкових величин протягом певного часу. Час дискретний і складається з серії рівних часових кроків. Часовим кроком обрано проміжок тривалістю 5 секунд. Набір даних часового ряду містить одне значення x_t однієї функції датчиків на етапі часу t .

Багатовимірний часовий ряд T – це серія впорядкованих спостережень, зібраних послідовно з часом і визначається таким чином:

$$T = \chi_i(t); [i = 1, \dots, n; t = 1, \dots, m] \quad (3.1)$$

де i – індекс вимірювань, зібраних в кожній часовій точці t ;

n – кількість змінних, спостережуваних на протязі часу, для багатовимірного ряду $n > 1$;

m – кількість спостережень.

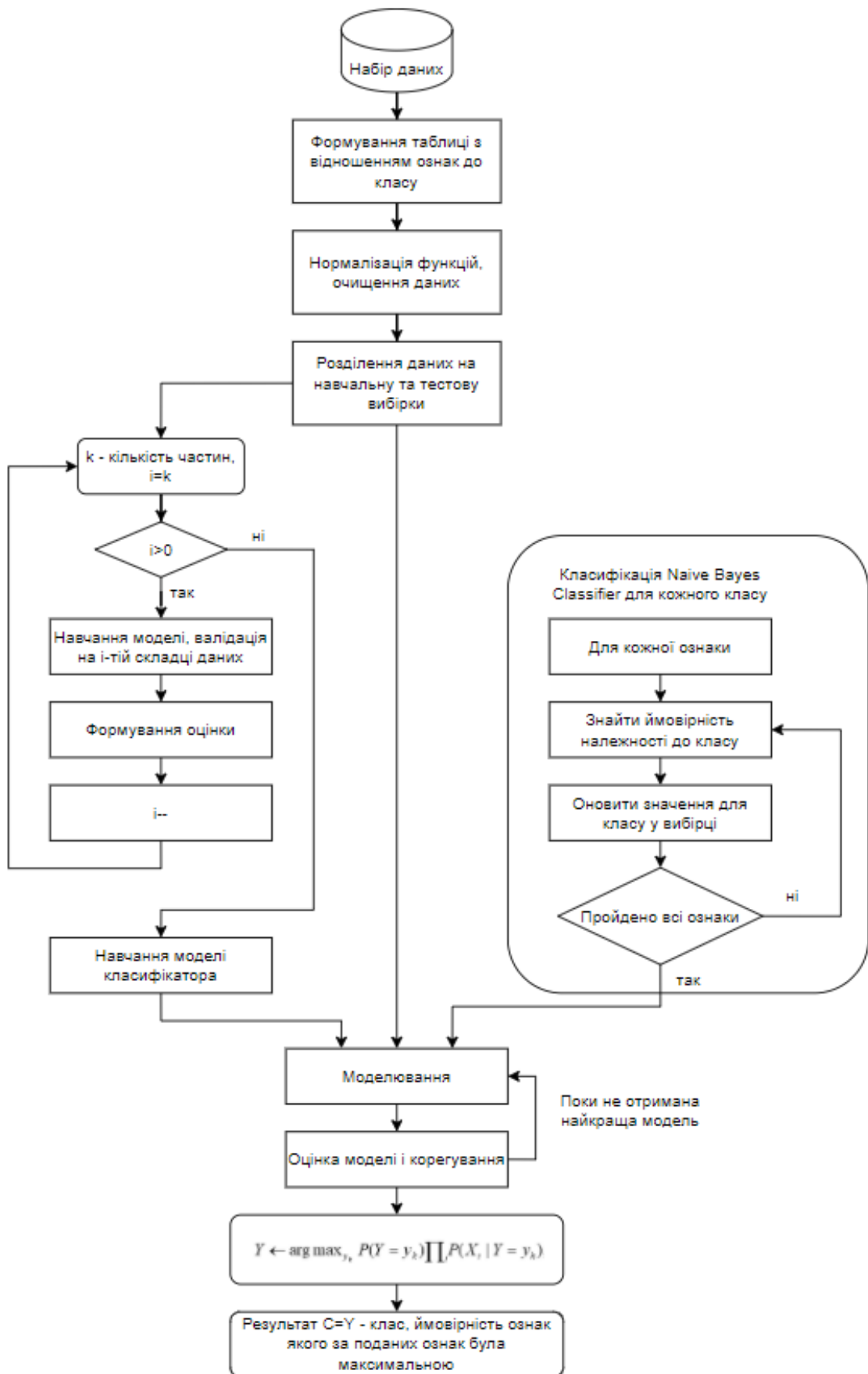


Рисунок 3.1 – Узагальнена схема методу розпізнавання активностей людини в розумному будинку

Часовий ряд для бінарних датчиків показує двійкове значення: 1 сигналізує про те, що є рух, тиск або магнітний контакт, отримання 0 означає, що більше немає руху, тиску або відсутній магнітний контакт.

Якщо на протязі одного часового інтервалу, тобто протягом п'яти секунд, є кілька різних показників одного і того ж об'єкта, то значення підсумовується до 1, коли хоча б одне значення з вибірки 1, в іншому випадку – 0 (рис. 3.2).

Часовий ряд бінарного датчика $S(t)$ складається з послідовності значень (S_1, S_2, \dots) , відібраних через регулярні інтервали часу t секунд.

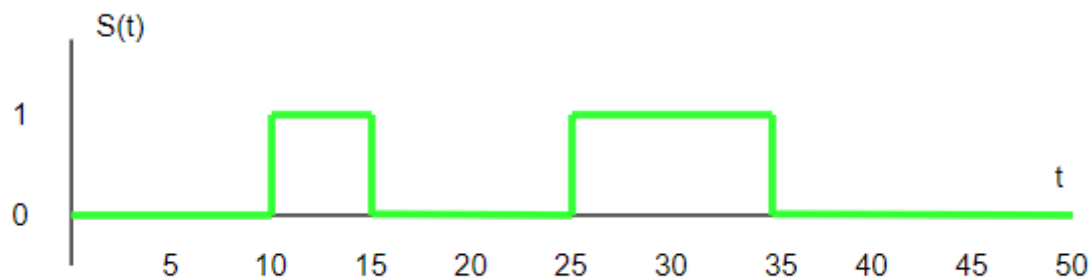


Рисунок 3.2 – Часовий ряд бінарного датчика $S(t)$

Часовий ряд для небінарних даних зобразимо на прикладі датчика про наближення, дані якого отримані за допомогою RSSI. Ці дані містять числові від'ємні значення $S(t)$, що змінюються з часом в залежності від пересування мешканця (рис. 3.3).

Щоб підвищити ефективність моделі, проведемо нормалізацію даних. Це призведе до зміни чисел в діапазон дійсних чисел від 0 до 1 (рис. 3.4).

Маючи однакову шкалу виміру для всіх часових рядів, часові ряди з різних джерел зведемо в один багатовимірний ряд. Перед початком класифікації проведемо розбиття на навчальну і тестову вибірку.

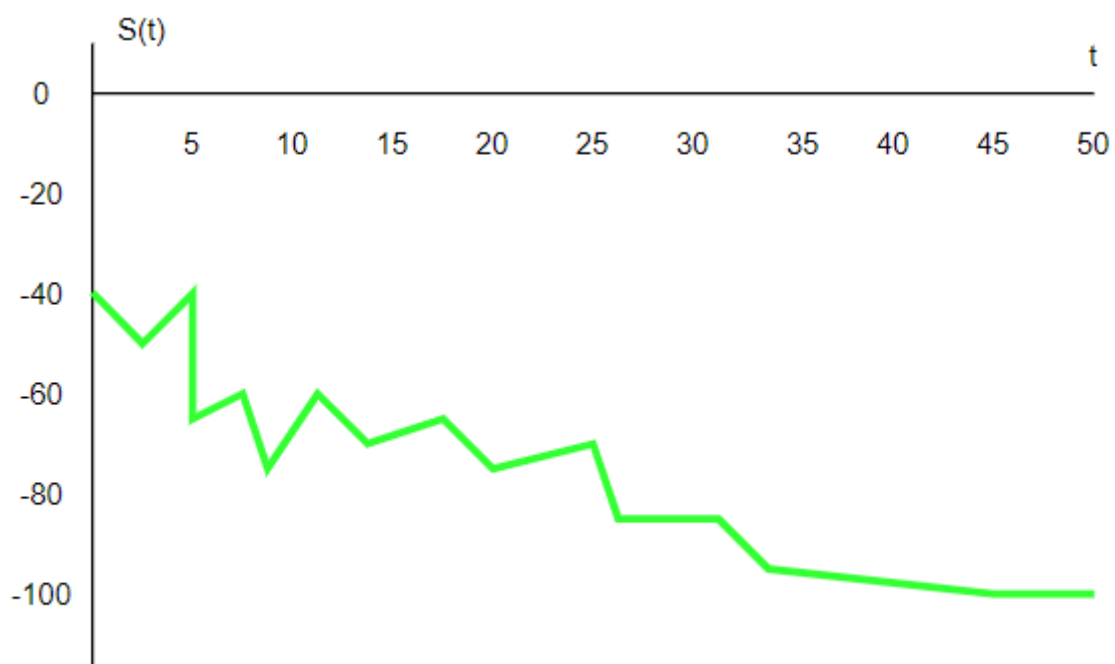


Рисунок 3.3 – Часовий ряд для небінарного датчика $S(t)$ на прикладі показників RSSI

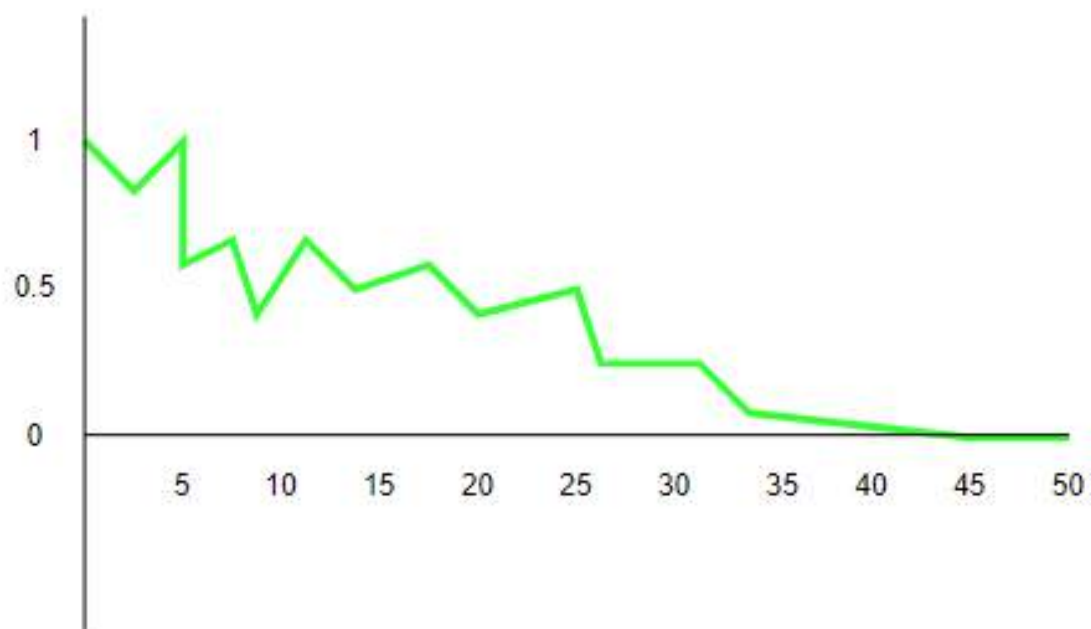


Рисунок 3.4 – Часовий ряд для небінарного датчика $S(t)$ на прикладі показників RSSI після нормалізації

Мета машинного навчання – розробити більш узагальнену модель даних, яка може добре працювати для точок даних, які ніколи не використовувалися під час навчального процесу. Для ефективного використання даних і побудови більш узагальненої моделі використаємо як стратегію розподілу даних k -кратне перехресне затвердження.

Перед початком процесу навчання встановлюємо значення параметрів моделі при її створенні і отримуємо простір гіперпараметрів. Під час виконання алгоритм приймає одну комбінацію параметрів за раз із вказаного нами простору гіперпараметрів.

Використовуючи ці значення параметрів, модель тренується і оцінюється методом k -кратного перехресного затвердження. Оцінка продуктивності (MSE) зберігається.

Потім алгоритм бере іншу комбінацію значень параметрів і робить те ж саме. Після отримання всіх комбінацій алгоритм зберігає всі оцінки продуктивності. З цих оцінок обирається найкраща.

В результаті отримуємо найкращу модель, яка включає значення параметрів, що дали найкращу оцінку продуктивності (рис. 3.5).

Підбір найкращої моделі з використанням стратегії k -кратного перехресного затвердження дозволяє отримати більш точну і надійну модель.

3.3 Класифікація часових рядів на основі Bayesian Network Classifier

Для аналізу часових рядів використаємо Байєсівську мережу безперервного часу. Це графічна модель, вузли якої пов'язані з випадковими величинами і стан якої безперервно розвивається з часом. Як наслідок, еволюція кожної змінної залежить від стану її батьківських елементів на графіку. Мережа складається з двох основних компонентів:

1. Початковий розподіл ймовірностей
2. Динаміка, яка керує еволюцією в часі розподілу ймовірностей, пов'язаних з мережею.

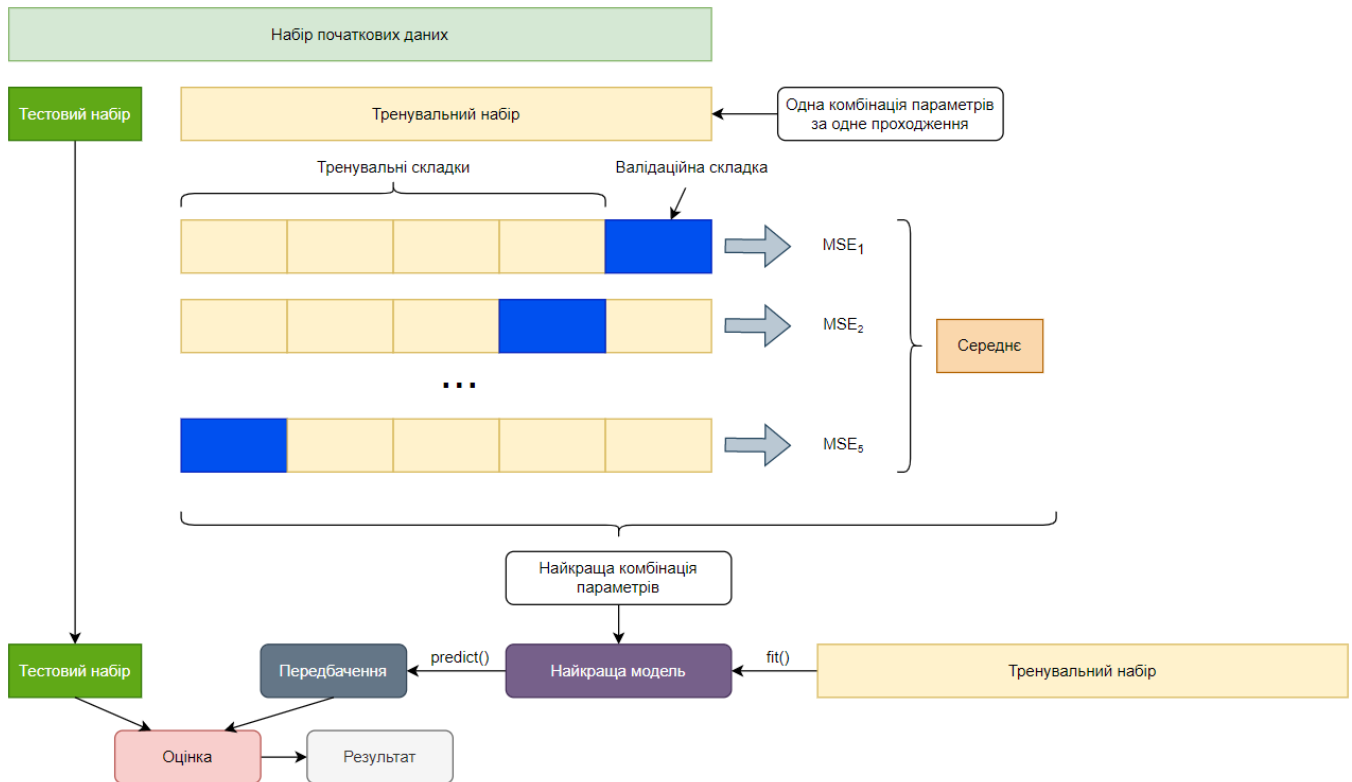


Рисунок 3.5 – k-кратне перехресне затвердження при k=5

Опишемо Байєсівську мережу безперервного часу. Нехай X – набір випадкових значень X_1, X_2, \dots, X_N . Кожне X_n має скінченну область значень $Val(X_n) = \{x_1, x_2, \dots, x_n\}$. Байєсівська мережа безперервного часу η містить 2 компоненти, перший з яких це початковий розподіл P_x^0 , заданий як Байєсівська мережа β над X . Другий – це модель безперервного переходу, яка визначається як:

1. Орієнтований (можливо циклічний) граф δ , вузлами якого $\in X_1, X_2, \dots, X_N$; $pa(X_n)$ позначає батьків X_n в графі δ .
2. Матриця умовної інтенсивності $Q_{X_n}^{pa(X_n)}$, для кожної змінної $X_n \in X$.

Враховуючи випадкову змінну X_n , матриця умовної інтенсивності $Q_{X_n}^{pa(X_n)}$ складається з набору матриць інтенсивності, одна матриця інтенсивності для кожного екземпляру $pa(x_n)$ батьків $pa(X_n)$ вузла X_n , де $q_{x_i}^{pa(x_n)} = \sum_{x_j \neq x_i} q_{x_i x_j}^{pa(x_n)}$ інтерпретується як миттєва ймовірність залишити x_i для конкретного

екземпляра $pa(x_n)$ з $pa(X_n)$ тоді як $q_{x_i x_j}^{pa(x_n)}$ інтерпретується як миттєва ймовірність переходу від x_i до x_j для конкретного екземпляра $pa(x_n)$ з $pa(X_n)$:

$$Q_{X_n}^{pa(x_n)} = \begin{bmatrix} -q_{x_1}^{pa(x_n)} & q_{x_1 x_2}^{pa(x_n)} & \cdot & q_{x_1 x_1}^{pa(x_n)} \\ q_{x_2 x_1}^{pa(x_n)} & -q_{x_2}^{pa(x_n)} & \cdot & q_{x_2 x_1}^{pa(x_n)} \\ \cdot & \cdot & \cdot & \cdot \\ q_{x_1 x_1}^{pa(x_n)} & q_{x_1 x_2}^{pa(x_n)} & \cdot & -q_{x_1}^{pa(x_n)} \end{bmatrix}. \quad (3.2)$$

Байєсівська мережа дозволяє точкові докази та безперервні докази. Точковий доказ – це спостереження значення x зі змінної X_n у певний момент часу, тобто $X_n(t)=x$, тоді як безперервні докази надають значення змінної протягом усього інтервалу, який приймається як напівзакритий інтервал $[t_1, t_2)$. Мережа створюється за допомогою послідовності доказів і це має назву потік доказів.

Для того, щоб описати потік доказів, потрібен часовий потік. J -часовий потік за закритим інтервалом часу $[0, T)$ є розбиттям на J закритих ліворуч інтервалів $[0, t_1); [t_1, t_2); \dots; [t_{j-1}, T)$.

Опишемо J -потік доказів. Маємо Байєсівську мережу безперервного часу \mathcal{N} , що складається з N вузлів та J -часовий потік $[0, t_1); [t_1, t_2); \dots; [t_{j-1}, T)$. J -потік доказів це набір спільних екземплярів $X=x$ для будь-якої підмножини випадкових величин X_n , $n=1, 2, \dots, N$ пов'язаний з кожним із J відрізків часу. J -потік доказів позначимо як $(X^1=x^1, X^2=x^2, \dots, X^J=x^J)$ або скорочено як (x^1, x^2, \dots, x^J) . J -потік доказів (x^1, x^2, \dots, x^J) є повністю спостережуваним у випадку, коли стан усіх змінних X_n відомий на всьому інтервалі часу $[0, T)$. В іншому випадку – є частково спостережуваним.

Для класифікації людської діяльності на основі багатовимірних даних часових рядів набору даних X_n , $n=1, 2, \dots, N$ використаємо модель Bayesian Network Classifier. Щоб визначити новий клас моделі класифікації необхідний додатковий вузол Y , пов'язаний з класом.

Bayesian Network Classifier безперервного часу представимо у вигляді:

$$C = \{\eta, P(Y)\}, \quad (3.3)$$

де η – модель Байєсівської мережі з вузлами атрибутів X_1, X_2, \dots, X_N ;

$P(Y)$ – гранична ймовірність у станах $Val(Y) = \{y_1, y_2, \dots, y_k\}$.

Представимо Bayes Network Classifier у вигляді графа G з дотриманням наступних умов:

1. G – зв'язаний граф.
2. $pa(Y) = \emptyset$, змінна класу $Y \in$ кореневим вузлом.
3. Y повністю визначається $P(Y)$ і не залежить від часу.

Модель (рис. 3.6) містить два цикли: перший включає атрибути X_3, X_2, X_4 і X_5 , а другий включає атрибут X_1 , який викликає атрибут X_3 і потім знову X_1 .

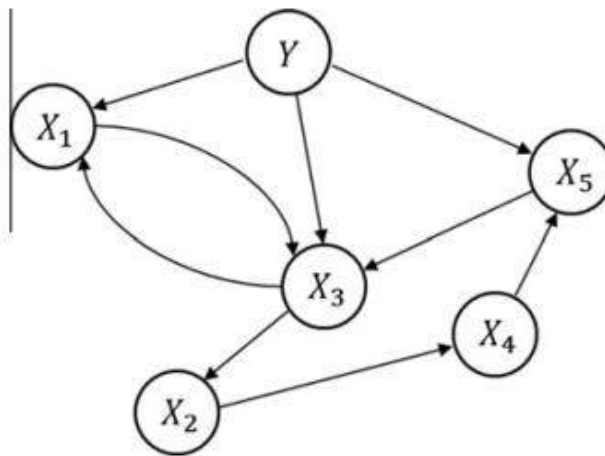


Рисунок 3.6 – Bayes Network Classifier у вигляді графа

Маючи набір даних D , що складається з повністю спостережуваних потоків доказів, класифікатор можна навчити максимізуючи функцію оцінки:

$$S_\eta(G : D) = \ln P(D | G) + \ln P(G), \quad (3.4)$$

де D – навчальні дані;

G – граф класифікатора.

Для точного навчання потрібно заздалегідь встановити максимальну кількість батьків L для вузлів X_1, X_2, \dots, X_N . У випадку, коли L набуває великого значення, потрібні значні обчислювальні зусилля, щоб знайти оптимальну структуру графа G , тобто таку, яка максимізує оцінку виразу (3.4).

До класу Bayes Network Classifier безперервного часу належать такі типи класифікаторів: Naive Bayes Classifier безперервного часу і класифікатор дерева безперервного часу, доповнене Naive Bayes Classifier.

Naive Bayes Classifier безперервного часу – це такий класифікатор Байєсівської мережі виразу (3.3), при якому $pa(X_n)=Y, n=1, 2, \dots, N$. Класифікатор дерева безперервного часу, доповнене Naive Bayes Classifier – це Bayes Network Classifier з дотриманням наступних умов:

1. $Y \in pa(X_n), n=1, 2, \dots, N$
2. Значення вузлів $X_n, n=1, 2, \dots, N$, формують дерево, $\exists j: |pa(X_j)|=1$ при цьому $i \neq j, |pa(X_i)|=2$.

Цей класифікатор (рис. 3.7) є таким, що змінна класу Y є кореневим вузлом, при цьому кожен атрибут $X_n, n=1, 2, \dots, N$ може мати за батьківський елемент змінну класу Y і один з інших атрибутів.

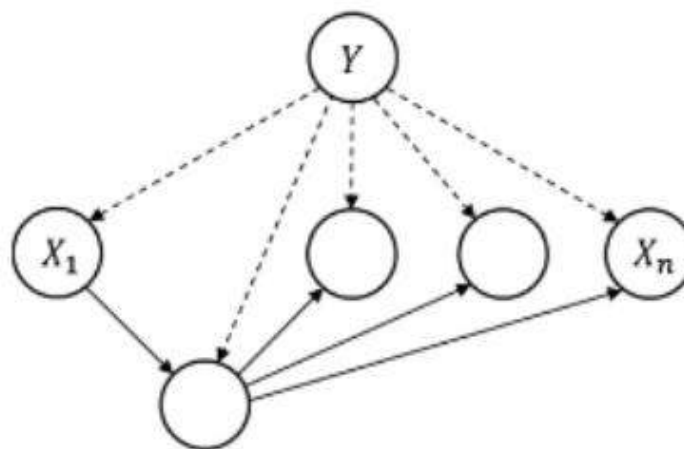


Рисунок 3.7 – Класифікатор дерева безперервного часу, доповнене Naive Bayes Classifier

Коли даний класифікатор представлений повністю спостережуваним J -потокм доказів (x^1, x^2, \dots, x^J) , він класифікується на основі стандартного Naive Bayes Classifier.

3.4 Класифікатор на основі використання Naive Bayes Classifier

В якості класифікатора, який використовується в методі обрано Naive Bayes Classifier, оскільки він працює дуже швидко, може легко передбачити клас тестового набору даних та, в порівнянні з іншими методами, працює краще при невеликому наборі даних з багатьма параметрами.

Крім того, Naive Bayes Classifier дає гарні результати, коли модель містить сценарії відсутні в навчальних даних, спрощує оновлення при надходженні нових даних, має високу масштабованість, здатний робити імовірнісні прогнози, не потребує великої кількості навчальних даних, є нечутливим до невідповідних функцій і здатний використовувати їх для прогнозування в режимі реального часу. Він має також високу точність і швидкість на великих наборах даних.

Це точний і надійний алгоритм.

Ці переваги зумовили вибір Naive Bayes Classifier для класифікації.

Naive Bayes Classifier – це імовірнісний класифікатор, який застосовує теорему Байєса із припущенням, що всі предиктори незалежні один від одного. Він передбачає, що конкретна ознака в класі не пов'язана з наявністю інших ознак.

Завдання даного класифікатора полягає в тому, щоб знайти клас спостереження за значеннями ознак, тобто обчислити ймовірність класу з урахуванням набору значень ознак $P(Y/X)$:

$$f : X \rightarrow Y, \quad (3.5)$$

де Y – змінна, що представляє клас;

X – вектор, який містить n ознак (рис. 3.8).

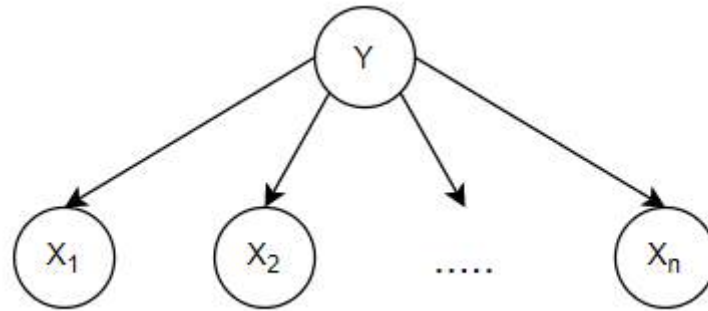


Рисунок 3.8 – Структура Naive Bayes Classifier

В цьому дослідженні класом виступає одна з двадцяти чотирьох визначених заздалегідь активностей людини, а вектором ознак X дані від різних датчиків розумного будинку:

$$X = \langle X_1, X_2, \dots, X_n \rangle, \quad (3.6)$$

де X_i – значення i -тої ознаки вектора X .

Застосовуючи теорему Байєса, $P(Y=y_i/X)$ представимо як:

$$P(Y = y_i | X = x_k) = \frac{P(X = x_k | Y = y_i)P(Y = y_i)}{\sum_j P(X = x_k | Y = y_j)P(Y = y_j)}, \quad (3.7)$$

де y_i – i -те можливе значення Y ;

y_j – j -те можливе значення Y ;

x_k – k -те можливе значення вектора X .

Щоб визначити $P(Y/X)$ використаємо навчальні дані для оцінки $P(X/Y)$ і $P(Y)$. Алгоритм повинен зберігати розподіли ймовірностей ознак для кожного класу окремо, щоб потім використати отримані оцінки для визначення $P(Y/X=x_k)$ для будь-яких нових екземплярів x_k .

Для такого обчислення потрібні надзвичайно великі набори даних, які дадуть оцінити розподіл ймовірностей для всіх можливих комбінацій значень ознак. Naive Bayes Classifier долає цю проблему, передбачаючи, що всі функції незалежні одна від одної, тобто кожна ознака X_i умовно незалежна від кожної іншої X_k , а також незалежна від кожної підмножини іншої X_k , заданого Y .

Значення цього передбачення значно спрощує представлення $P(X/Y)$ і проблему його оцінювання на основі навчальних даних. Враховуючи умовне припущення незалежності для X , що містить n ознак отримаємо функцію:

$$P(X_1 \dots X_n | Y) = \prod_{i=1}^n P(X_i | Y). \quad (3.8)$$

Виведемо алгоритм Naive Bayes з Y , що представляє змінну з дискретним значенням, і атрибутами $X_1 \dots X_n$ з дискретними ознаками. Метою є навчити класифікатор, який виводить розподіл ймовірності за можливими значеннями Y для кожного нового екземпляра X , який подається для класифікації. Вираз ймовірності того, що Y набуде k -го можливого значення, згідно з правилом Байєса:

$$P(Y = y_k | X_1 \dots X_n) = \frac{P(Y = y_k) P(X_1 \dots X_n | Y = y_k)}{\sum_j P(Y = y_j) P(X_1 \dots X_n | Y = y_j)}, \quad (3.9)$$

де $\sum_j P(Y = y_j) P(X_1 \dots X_n | Y = y_j)$ береться за всіма можливими значеннями y_j з Y .

Припускаючи, що X_i умовно незалежне від Y і використовуючи вираз 3.8 перепишемо вираз ймовірності 3.9 як:

$$P(Y = y_k | X_1 \dots X_n) = \frac{P(Y = y_k) \prod_i P(X_i | Y = y_k)}{\sum_j P(Y = y_j) \prod_i P(X_i | Y = y_j)}. \quad (3.10)$$

Отриманий вираз є фундаментальним рівнянням нашого класифікатора. При поданні на вхід класифікатора нового вектора ознак X , цей вираз показує як обрахувати ймовірність, що Y прийме будь-яке з заданих значень, враховуючи спостережувані значення ознак X та розподіли, оцінені з навчальних даних.

Оскільки результатом виконання методу буде найбільш ймовірне значення Y , опишемо виведення Y :

$$Y \leftarrow \arg \max_{y_k} \frac{P(Y = y_k) \prod_i P(X_i | Y = y_k)}{\sum_j P(Y = y_j) \prod_i P(X_i | Y = y_j)}. \quad (3.11)$$

Оскільки знаменник не залежить від y_k спростимо вираз до наступного:

$$Y \leftarrow \arg \max_{y_k} P(Y = y_k) \prod_i P(X_i | Y = y_k). \quad (3.12)$$

Таким чином алгоритм класифікації на основі аналізу даних визначає вихід нового екземпляра даних (рис. 3.9).

Результатом виконання методу буде вид діяльності $C=Y$ для якої ймовірність за поданих вхідних ознак була максимальною.

Оскільки набір ознак містить неперервні змінні, потрібно обрати оптимальний тип Naive Bayes Classifier.

Найкраще з таким набором ознак працює Gaussian Naive Bayes Classifier. Цей підхід полягає в припущенні, що для кожного можливого дискретного значення y_k з Y розподіл кожного безперервного X_i є розподілом Гауса і визначається математичним сподіванням і середнім квадратичним відхиленням характерним для X_i і y_k .

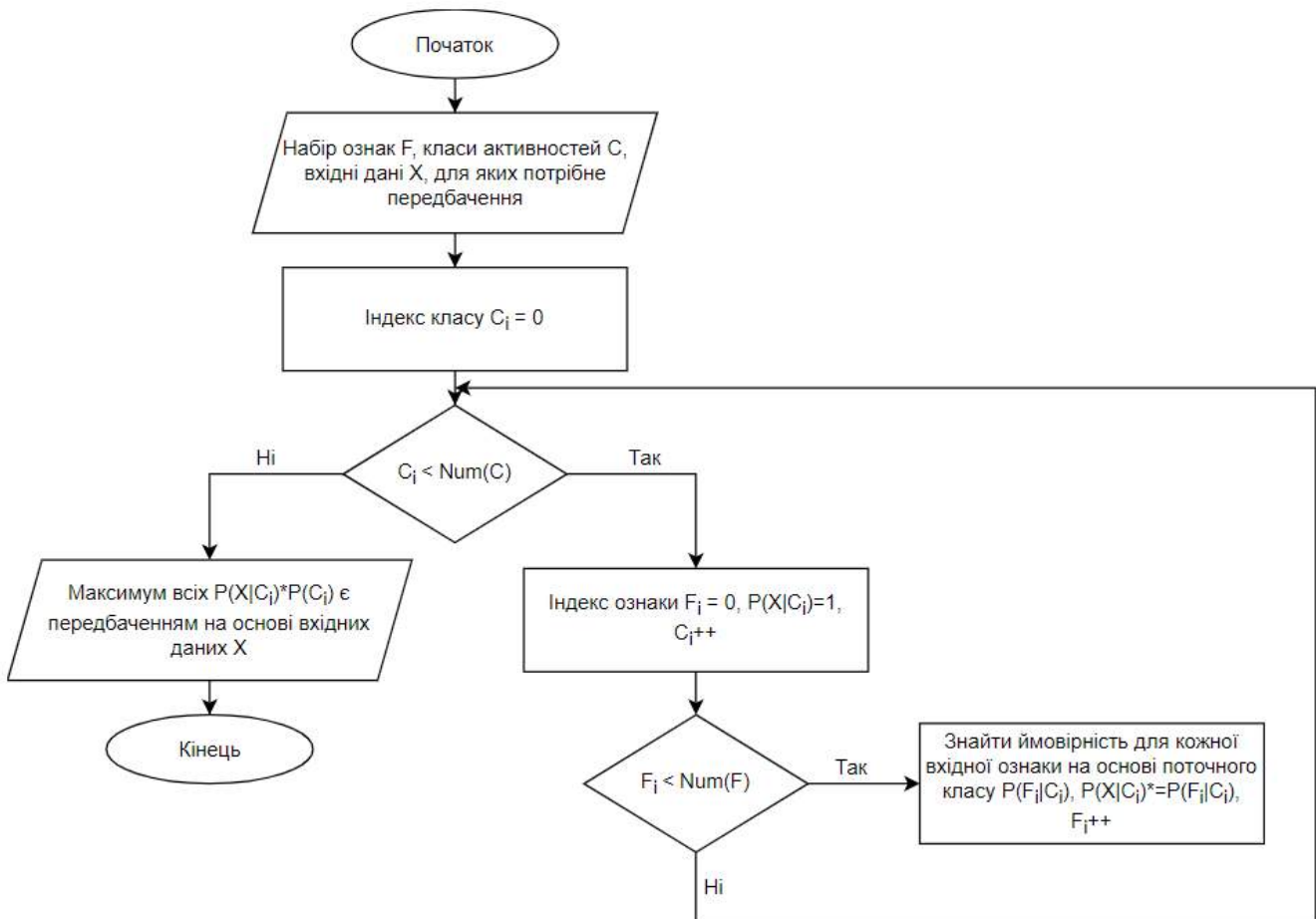


Рисунок 3.9 – Блок-схема роботи алгоритму класифікації на основі Naive Bayes Classifier

При нанесенні на графік розподілу Гауса зображається як крива, симетрична відносно середніх значень характеристик (рис. 3.10).

Щоб навчити такий класифікатор, ми повинні оцінити математичне сподівання, яке є максимумом графіка і щодо якого графік завжди симетричний та середнє квадратичне відхилення.

Математичне сподівання знайдемо за формулою:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i, \quad (3.13)$$

де n – кількість ознак;

x_i – i -те значення ознаки x .

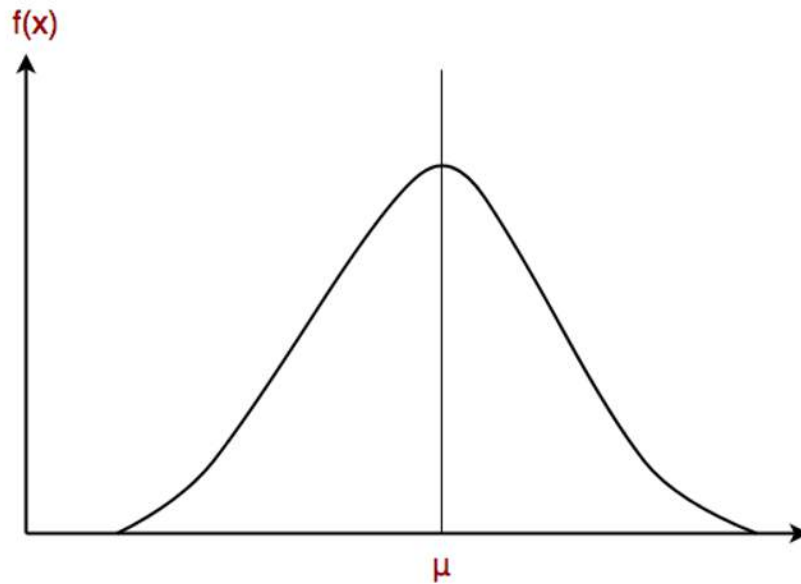


Рисунок 3.10 – Розподіл Гауса на графіку

Середнє квадратичне відхилення залежить від математичного сподівання та визначається як:

$$\sigma = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 \right]^{0.5}, \quad (3.14)$$

де μ – математичне сподівання.

Щільність розподілення ймовірностей для розподілу Гауса визначимо формулою:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (3.15)$$

де σ – середнє квадратичне відхилення;

μ – математичне сподівання.

Подана формула щільності розподілення ймовірностей буде застосована в основі класифікатора.

3.5 Висновки

Розроблено метод розпізнавання людської активності в розумному будинку на основі даних багатовимірного часового ряду та Bayesian Network Classifier. Запропонований Bayesian Network Classifier є класифікатором ймовірнісних графічних моделей, яка моделює час. Він дозволяє відновити розподіл ймовірностей у часі, коли відбуваються конкретні події.

Алгоритм висновку класифікації застосовує для прийняття рішення Naïve Bayes Classifier, точність і надійність якого дозволяє працювати добре з будь-яким об'ємом навчальних даних.

Даний метод здатний вирішити складність проблеми класифікації при великій кількості класів, оскільки визначає одну з багатьох активностей і досягає достатню велику точність при цьому.

4 ІНТЕЛЕКТУАЛЬНА СИСТЕМА ПРИЙНЯТТЯ РІШЕННЯ ЩОДО ФУНКЦІОНУВАННЯ ЕЛЕМЕНТІВ РОЗУМНОГО БУДИНКУ НА ОСНОВІ КЛАСИФІКАЦІЇ ВИДІВ ДІЯЛЬНОСТІ ЛЮДИНИ

4.1 Структура інтелектуальної системи прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

На основі розроблених методу та моделі розпізнавання людської активності в розумному будинку з застосуванням Naïve Bayes Classifier розроблено інтелектуальну систему прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини. Узагальнену структуру системи наведено на рис. 4.1.

Розроблена система підтримує інтеграцію з різноманітними датчиками, виробленими різними виробниками. Система підтримує модульний веб-інтерфейс, який дозволяє візуалізацію даних датчиків та додаток для керування даними. Ці дані використовуються для аналізу і тренування моделі машинного навчання, яка визначає вид людської активності.

Систему розроблено, щоб запропонувати масштабовану високу продуктивність, що включає перевірені технології з відкритим кодом. Розглянемо детальніше основні складові представленої системи.

Загальні записи датчиків і записи метаданих зберігаються за допомогою високопродуктивного, масштабованого механізму зберігання даних з низькою затримкою, заснованого на перевірених технологіях з відкритим вихідним кодом.

Цей механізм зберігання містить дві перевірені системи баз даних – MongoDB і InfluxDB.

InfluxDB – це база даних часових рядів (TSDB), оптимізована для зберігання та пошуку даних, яка використовує часові позначки як індекс. Індекс часової позначки є унікальним, і для підтримки великого обсягу даних має високу роздільну здатність.

Індексована мітка часу в InfluxDB має точність до наносекунд. Змішування записів з індексом такої високої роздільної здатності неможливо. Цей клас баз даних призначений для обробки високої швидкості послідовних операцій читання та запису у великій кількості.

InfluxDB дозволяє зберігати довільні дані для кожного збереженого запису.

Кластеризована операція InfluxDB дає можливість масштабованої операції.

Ці характеристики роблять базу даних придатною до зберігання великих даних датчиків.

Описана база даних має високу послідовність і не оптимізована для підтримки випадкової вставки, видалення або модифікації записів, що зберігаються.

Враховуючи це обмеження, необхідна незалежна база даних для зберігання та керування даними, які не є датчиками.

Ці дані включають метадані датчиків, профілі користувачів, ключі API та експериментальні метадані.

Для підтримки цих типів записів обрано базу даних MongoDB.

Ця база даних використовує парадигму документів, де записи моделюються як документи, які можна створювати, оновлювати, читати та видаляти.

У MongoDB документи зберігаються у форматі BSON і представлені JSON. MongoDB обрано завдяки її доведеній масштабованості та вбудованій моделі зберігання, яка підходить для цієї системи.

InfluxDB і MongoDB мають відкритий вихідний код і не потребують ліцензійних зборів для використання. Це зменшує ризик, пов'язаний із залежністю від постачальника, який може припинити підтримку бази даних або змінити умови обслуговування. Безкоштовність цих баз даних забезпечує масштабованість без будь-яких фінансових міркувань.

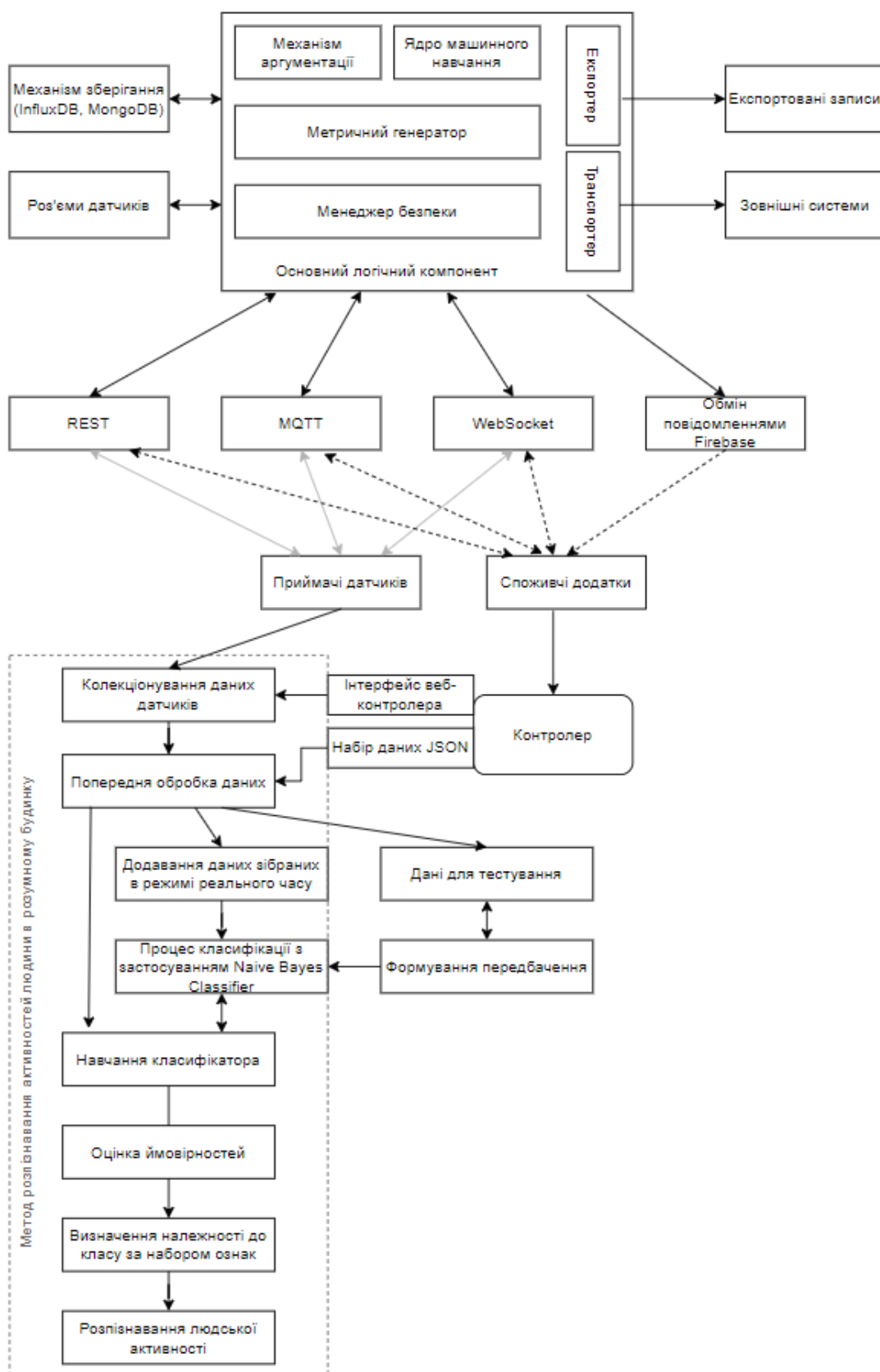


Рисунок 4.1 – Узагальнена схема системи прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Цей механізм зберігання інтегрований в архітектуру системи.

Механізм зберігання підключений до основного логічного компонента. Цей компонент містить ряд елементів: менеджер безпеки, підпрограми генерації показників, механізм аргументації на основі правил, ядро машинного навчання та експортер і транспортер записів.

Менеджер безпеки використовується для аутентифікації, верифікації та пропонування криптографічних послуг. Послуги аутентифікації включають:

1. Аутентифікація виклик-відповідь та CHAP аутентифікація.
2. Надання та перевірка ключа API (512 біт).
3. Керування профілем користувача.
4. Верифікація API та перевірка токена входу.

Криптографічні послуги обробляються за допомогою бібліотеки `bouncy castle`. Це забезпечує систему високоякісним генератором псевдовипадкових чисел і набором криптографічних функцій.

Логіка генерації вимірювання на стороні сервера дозволяє споживчим програмам (програмам пристроїв) полегшити їх обробку, зменшуючи мережевий трафік і скорочуючи час обробки завдяки перевагам локальності даних.

Для класифікації вікон показників використовується `rule-based engine`. Правила підтримують ряд логічних операторів, які включають: більше, менше і дорівнює. Також підтримуються заперечення цих операторів. Правила вказуються за допомогою веб-інтерфейсу або через Java API.

Послуги машинного навчання реалізуються за допомогою пакета Python для машинного навчання `Scikit Learn`. Він надає прості та ефективні інструменти для прогнозного аналізу даних, побудований на `NumPy`, `SciPy` та `matplotlib` та має відкритий вихідний код. Це дозволяє створювати прототипи сумісних і передбачуваних рішень.

Основна логіка підготовки даних містить функції пересилання та експорту датчиків. Логіка пересилання датчиків дозволяє пересилати записи в реальному часі для груп записів даних датчиків до визначених зовнішніх

систем. Пересилання цих даних відбувається через виклики REST та WebSockets.

Функція експорту дозволяє експортувати записи експериментів і пов'язані з ними дані датчиків в окремі документи і набори даних. Система створює один документ JSON, що містить усі пов'язані дані датчиків та експериментальні дані.

Основна логіка створена як бібліотека Java. Ця бібліотека дозволяє зовнішнім рішенням інтегруватись з системою та обробляти і використовувати дані всередині неї.

Ця логіка поєднується з трьома кінцевими точками, які забезпечують інтеграцію з іншими компонентами: датчиками, веб-інтерфейсами та мобільними додатками. Ці кінцеві точки засновані на REST, MQTT і WebSocket.

Ендпоінт на основі REST є основним ендпоінтом веб-служби, який дає змогу програмам взаємодіяти з системою. Ці програми включають веб-програми, приймачі датчиків, мобільні додатки та інше споживче програмне забезпечення. Цей ендпоінт заснований на технології Java EE без збереження стану і тому пропонує масштабовану операцію.

Представлені ендпоінти також для MQTT і WebSockets. MQTT – це інтерфейс на базі pub-sub для споживчих додатків та інтеграції з сенсорними пристроями. WebSockets пропонують додатковий інтерфейс для споживчих додатків і датчиків.

Основна логіка містить бібліотеки та код для інтеграції служб обміну повідомленнями для мобільних і десктопних додатків через кросплатформену платформу обміну повідомленнями Firebase Cloud.

Приймачі датчиків зчитують дані датчиків низького рівня з сенсорних пристроїв, перетворюють дані у формат JSON, а потім передають дані в цьому форматі на вихід.

Подібними до приймачів датчиків є роз'єми даних датчиків. Роз'єми даних датчиків полегшують інтеграцію з іншими сенсорними платформами і

мережами. Один з цих роз'ємів об'єднує всі дані датчиків із серверів RaZberry, а інший об'єднує підключену мережу.

Роз'єм RaZberry забезпечує безперебійну інтеграцію з датчиками на основі Z-Wave. Цей роз'єм передає всі дані від датчиків, які зареєстровані в екземплярі RaZberry (рис. 4.2).

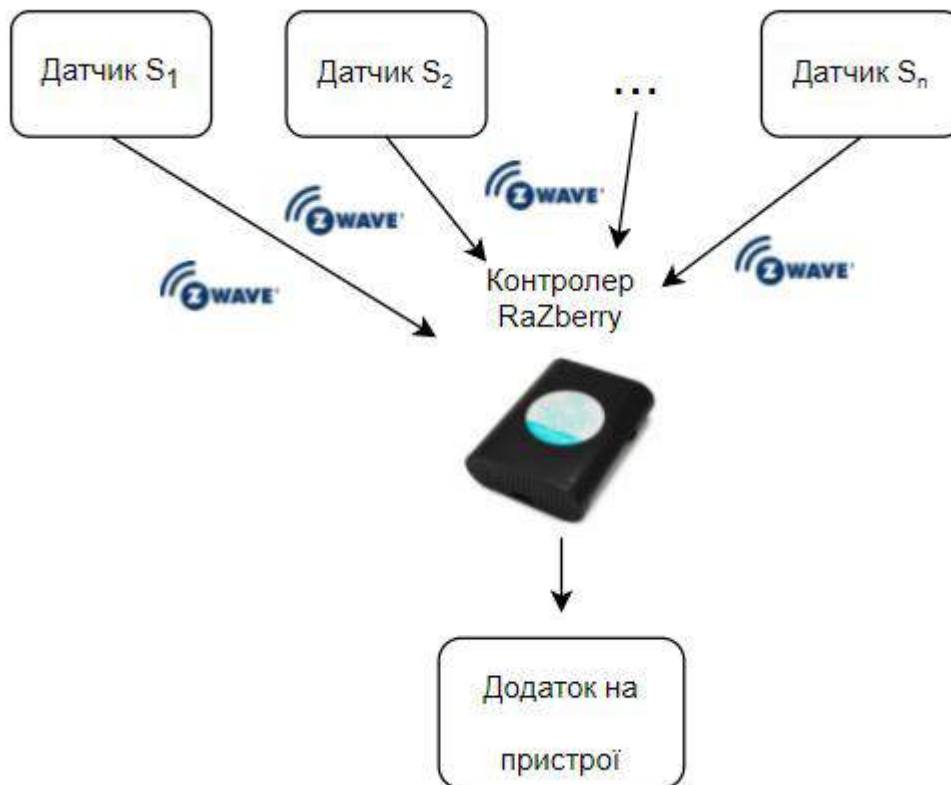


Рисунок 4.2 – Збір даних контролером RaZberry з датчиків розумного будинку

Для керування даними використовується модульний веб-додаток, який включає: керування датчиками, візуалізацію даних датчиків, керування доступом користувачів, керування ключами API та керування експериментальними налаштуваннями.

Цей веб-додаток має бути розроблено з використанням сучасних технологій, а саме фреймворку Angular та фреймворку інтерфейсів bootstrap.

Веб-додаток має бути розширюваним і модульним, що дозволить дослідникам і розробникам створювати модулі, які можна буде інтегрувати в

загальну платформу. Це дасть можливість адаптувати існуючу веб-програму в менший модуль, що дозволить швидко створити спеціальний інтерфейс з меншою складністю.

Такий інтерфейс дозволить повторне використання програмного забезпечення для використання сторонніми веб-додатками.

Цей модульний інтерфейс може бути доставлений із серверів, відмінних від REST ендпоінт, якщо послабити обмеження Cross-Origin Resource Sharing. Це обмеження вмикається для підвищення безпеки.

4.2 Алгоритми колекціонування та обробки даних датчиків розумного будинку

Датчики підтримують зв'язок через протоколи, які включають Bluetooth, радіочастоту RF, LoRaWAN, Wi-Fi, Ethernet, IEEE 802.15.4 і Z-wave (табл. 4.1). Список відображає датчики, які мають інтегруватися в систему і може бути розширеним.

Інші датчики, які можуть підтримуватись системою, це: монітори артеріального тиску, пульсоксиметри, детектори витoku води, розумні годинники та ваги.

Два загальні роз'єми для датчиків розроблені для безперешкодної підтримки інтеграції всіх датчиків, розгорнутих на двох взаємодоповнюючих платформах: підключеній мережі і сервері z-wave RaZberry. Сервер RaZberry Z-wave – це прослуховувач для датчиків, які локально комунікують за допомогою протоколу Z-wave.

Підтримка датчиків запропонованої системи забезпечується завдяки:

1. Розробці стратегії призначення глобально унікальних ідентифікаторів датчиків.
2. Застосуванні загального формату запису датчика, що включає схему на принципах зчитування.

Таблиця 4.1 – Датчики та протоколи зв'язку

| Тип датчика | Виробники | Протокол зв'язку |
|---------------------------------|---|---|
| Акселерометр | Bosch Microchip Sun Microsystems Texas Instruments | I ² C з Wi-Fi LoRaWAN IEEE 802.15.4 Bluetooth |
| Датчик якості повітря | Elsys | LoRaWAN |
| Аналогова напруга | Adeunis | LoRaWAN |
| Bluetooth маяк | Various | LoRaWAN |
| Контактний датчик | Everspring Nexa Tynetec | Z-Wave RF (433MHz) RF (169MHz) |
| Розташування GPS | Adeunis RF GlobalSat Ulster University | LoRaWAN Wi-Fi/4G (через додаток) |
| Датчик вологості повітря | Adeunis RF Microchip Texas Instruments Ulster University | LoRaWAN Bluetooth Wi-Fi |
| Інерційна одиниця вимірювання | Slever Technologies | Bluetooth/USB |
| Вимірювачі інтенсивності світла | Sun Microsystems Texas Instruments | IEEE 802.15.4 Bluetooth |
| Магнітометри | Bosch Texas Instruments | Wi-Fi/Bluetooth Bluetooth |
| Теги NFS | Various | Wi-Fi/4G/Ethernet |

Кінець таблиці 4.1 – Датчики та протоколи зв'язку

| Тип датчика | Виробники | Протокол зв'язку |
|---|----------------------|------------------|
| Пасивні інфрачервоні датчики руху | Belkin | Wi-Fi |
| | Elsys | LoRaWAN |
| | Nexa | RF (433MHz) |
| Монітор споживання електроенергії | Belkin | Wi-Fi |
| | NKE Wattecco | LoRaWAN |
| Натискання кнопки | FLIC | Bluetooth |
| Розумна підлога | Future-Shape GMBH | RF (868MHz) |
| Звуковий тиск | Ulster University | Wi-Fi/Bluetooth |
| Температура | Adeunis RF | LoRaWAN |
| | Microchip | |
| | Sun Microsystems | IEEE 802.15.4 |
| | Texas Instruments | Bluetooth |
| | Ulster University | Wi-Fi |
| Теплове бачення | Heimann GMBH | Ethernet |
| | IOTTech | USB |
| | Ulster University | Wi-Fi&Bluetooth |

Стратегія отримання глобально унікальних ідентифікаторів датчиків (UUID) гарантує, що пари даних датчиків не містять помилкового значення з несподіваних джерел, зокрема інших датчиків. Зазвичай такі UUID генеруються за допомогою програмного забезпечення для приймачів датчиків.

Ці приймачі реєструють датчики в системі, створюючи запис метаданих датчиків.

Записи метаданих передаються і обробляються у вигляді JSON. Після реєстрації приймачі переходять до передачі даних датчиків.

Дані з експериментальних екземплярів експортуються в один файл JSON, що сприяє спрощеному обміну даними датчиків.

Набір даних доступний через REST запит.

Спочатку впроваджується JSON парсер для читання вхідного файлу JSON, отриманого в результаті процесу обробки даних датчиків.

В наборі даних включено 24 різних види людської активності (табл. 4.2).

Таблиця 4.2 – Види людської активності, записані в наборі даних

| ID активності | Назва активності | Сценарій |
|----------------------------------|---|---|
| C ₁ | Прийом ліків | Піти на кухню – взяти пляшку з водою – дістати ліки з коробки – проковтнути ліки. |
| C ₂ | Приготування сніданку | Піти на кухню – взяти продукти для сніданку – розмістити їжу на столі. Може включати: заварювання чаю з застосуванням електричного чайника або приготування гарячого шоколаду з молоком, використовуючи мікрохвильову піч. |
| C ₃ C ₄ | Приготування обіду Приготування вечері | Піти на кухню – дістати деякі продукти з холодильника і комори – розмістити їжу на столі. Може включати: приготування гарячої їжі на плиті або нагрівання їжі швидкого приготування в мікрохвильовій печі. |
| C ₅ | Сніданок | Піти на кухню зранку – сісти за стіл для прийому їжі – поїсти – прибрати посуд в раковину або посудомийну машину. |

Продовження таблиці 4.2 – Види людської активності, записані в наборі даних

| ID активності | Назва активності | Сценарій |
|-----------------|---------------------------|---|
| C ₆ | Обід | Піти на кухню в обід – сісти за стіл для прийому їжі – поїсти – прибрати посуд в раковину або посудомийну машину. |
| C ₇ | Вечеря | Піти на кухню ввечері – сісти за стіл для прийому їжі – поїсти – прибрати посуд в раковину або посудомийну машину. |
| C ₈ | Перекус | Піти на кухню – взяти фрукт або снєк – з'їсти на кухні або у вітальні. Може включати: прибирання посуду в раковину або посудомийну машину. |
| C ₉ | Перегляд телевізора | Піти у вітальню – взяти пульт – сісти на диван – подивитись телевізор – залишити пульт біля телевізора. |
| C ₁₀ | Вхід до розумного будинку | Зайти до будинку, використовуючи вхідні двері – залишити ключ в невеликій корзині. |
| C ₁₁ | Гра у відеоігри | Піти у вітальню – взяти пульт від телевізора та джойстик від ігрової приставки – сісти на дивані – пограти – встати з дивану – покласти пульт і джойстик біля телевізора. |
| C ₁₂ | Відпочинок на дивані | Піти у вітальню – сісти на дивані – посидіти декілька хвилин – встати з дивану. |

Продовження таблиці 4.2 – Види людської активності, записані в наборі даних

| ID активності | Назва активності | Сценарій |
|-----------------|-------------------------------|---|
| C ₁₃ | Вихід з розумного будинку | Підійти до вхідних дверей – відкрити вхідні двері – вийти з будинку – закрити двері. |
| C ₁₄ | Відвідування мешканця будинку | Підійти до вхідних дверей – відкрити двері – поговорити з кимось біля дверей – закрити двері. |
| C ₁₅ | Викидання сміття | Піти на кухню – взяти сміття з відра для сміття – взяти ключі з маленької корзини при вході – вийти з будинку – повернутися – залишити ключі в маленькій корзині. |
| C ₁₆ | Миття рук | Піти у ванну кімнату – відкрити кран – намити руки – промити руки – закрити кран – висушити руки. |
| C ₁₇ | Чистка зубів | Піти у ванну кімнату – відкрити кран – почистити зуби – закрити кран. |
| C ₁₈ | Використання туалету | Піти у ванну кімнату – підняти кришку унітазу – використати туалет – опустити кришку унітазу – натиснути на зливний бачок. |
| C ₁₉ | Миття посуду | Піти на кухню – покласти брудний посуд в посудомийну машину – запустити – покласти посуд на місце. |

Кінець таблиці 4.2 – Види людської активності, записані в наборі даних

| ID активності | Назва активності | Сценарій |
|-----------------|------------------|--|
| C ₂₀ | Прання | Піти у спальню – взяти корзину для білизни – піти на кухню – покласти одяг в пральну машину – зачекати кінця прання – витягнути одяг з пральної машини – покласти випраний одяг до шафи. |
| C ₂₁ | Робота за столом | Сісти на робочому місці – зробити роботу – встати після завершення. |
| C ₂₂ | Переодягання | Піти до спальні – покласти брудний одяг у корзину для білизни – відкрити шафу – одягнути чистий одяг – зачинити шафу. |
| C ₂₃ | Відхід до сну | Піти у спальню – лягти у ліжко – заснути. |
| C ₂₄ | Пробудження | Прокинутись – встати з ліжка. |

В результаті роботи системи може бути визначена лише одна з визначених активностей.

Діяльність, що виконується, анується за допомогою NFC тегів і смартфона. Цей процес використовується для позначення початку та кінця кожної дії.

NFC – це популярна бездротова технологія, яка дозволяє передавати дані між двома пристроями, які знаходяться в безпосередній близькості один від одного за допомогою електромагнітної індукції.

Теги NFC допомагають передавати невеликі обсяги даних.

Щоразу, коли зчитувач з живленням NFC підноситься до тегу, тег отримує живлення і передає будь-які збережені дані в мікročіпі на пристрій.

Зчитувач NFC живить теги і зчитує дані з пасивних тегів. Щоб робити це зчитувач потребує джерело живлення і пропускає електричний струм через власну котушку. Це створює змінне магнітне поле в безпосередній близькості від зчитувача.

Введення тегу в діапазон цього поля призводить до індуктивного зв'язку між двома котушками завдяки закону індукції Фарадея (рис. 4.3).

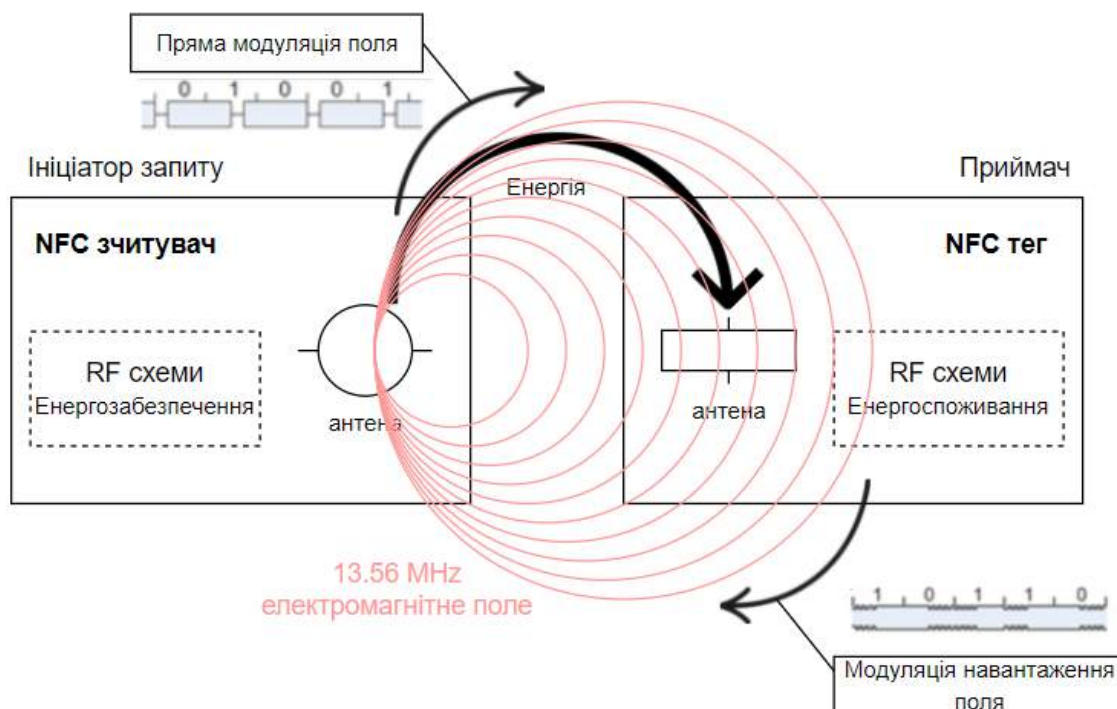


Рисунок 4.3 – Процес передачі даних за допомогою NFC

Після того, як тег увімкнено, зчитувач може виявити, як NFC тег модулює електромагнітне поле.

Виконується сегментація даних, при якій всю послідовність подій датчиків ділимо на менші послідовні однакові за розміром часові інтервали.

Результат сегментації дає послідовність часових рядів, що дозволить застосувати методи класифікації в режимі реального часу.

Аналізуючи тривалість та набори ознак активностей в початковому наборі, відбуватиметься накопичення підказок, які підвищуватимуть

ймовірність виконання певної діяльності за вхідним набором ознак і тривалості. Далі відбуватиметься динамічний процес класифікації, використовуючи ймовірність переходу від однієї діяльності до іншої та ймовірність отримання події для кожного виду діяльності.

Якщо виконана діяльність не підходить під жодну з визначених 24 видів, в додатку для NFC створюється 25-тий вид під назвою «Без діяльності». При попередній обробці даних проводиться очищення і такі зразки відкидаються.

Класифікатор даної системи визначає вид діяльності на основі поданих ознак. Всі ознаки зводимо до одного формалізованого цифрового вигляду. Бінарні дані позначаються нулем або одиницею, сегмент дня цифрою від 0 до 2, інші мають своє цифрове значення.

Дані прискорення генеруються розумним годинником. Годинник має вбудований акселерометр, датчик якого розраховує силу прискорення, яка впливає на пристрій по трьох осях (X, Y і Z). Вимірюються зміни швидкості вздовж кожної осі, якщо годинник лежить на поверхні циферблатом вгору (рис. 4.4):

1. +X – горизонтально вбік розташованих кнопок управління, -X в протилежний бік від +X.
2. +Y – горизонтально до верхнього краю годинника, -Y в протилежний бік від +Y.
3. +Z – вгору вертикально від циферблата, -Z – вниз вертикально від циферблата.

Дані наближення вимірюються датчиками RSSI. RSSI є вимірюванням сили отриманого сигналу. Це значення завжди негативне, оскільки показує скільки сили сигналу було втрачено.

Чим вище значення RSSI, тим сильніший сигнал, тобто -50 – досить хороший сигнал, а -100 – взагалі відсутній.

Змінні, які вимірюються в різних масштабах, не можуть внести однаковий внесок у функції налаштування і навчання моделі і це зробить модель менш ефективною.

Для того, щоб впоратись з цією потенційною проблемою, над такими даними як прискорення та наближення виконується нормалізація за допомогою MinMaxScaler.

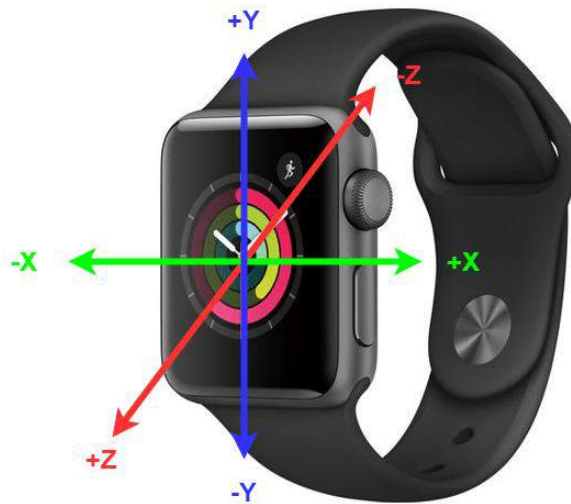


Рисунок 4.4 – Розташування осей розумного годинника

Атрибути нормалізуються так, щоб вони лежали у фіксованому діапазоні – від нуля до одиниці – шляхом віднімання мінімального значення та ділення на діапазон між максимальним і мінімальним значенням. При цьому зв'язок між вихідними значеннями даних зберігається.

Набір оброблених даних містить записи ознак і класи, які визначені за набором цих ознак.

Наприклад, маємо такий набір ознак: місце розташування мешканця визначено як кухня за даними датчиків розумної підлоги, спрацював датчик магнітного сигналу електричного чайника, пора доби визначена як ранок, акселерометром розумного годинника зафіксовано горизонтальний рух. Такий набір позначений класом C_2 «Приготування сніданку».

Модель класифікації навчатиметься на основі таких записів. За методом k-кратного перехресного затвердження ділимо набір даних на навчальну і тестову вибірки.

4.3 Алгоритм визначення активності та оцінка ефективності системи

Модель запропонованої системи використовує контрольоване навчання. Воно включає зворотний зв'язок, щоб вказати правильний прогноз чи неправильний (рис. 4.5).

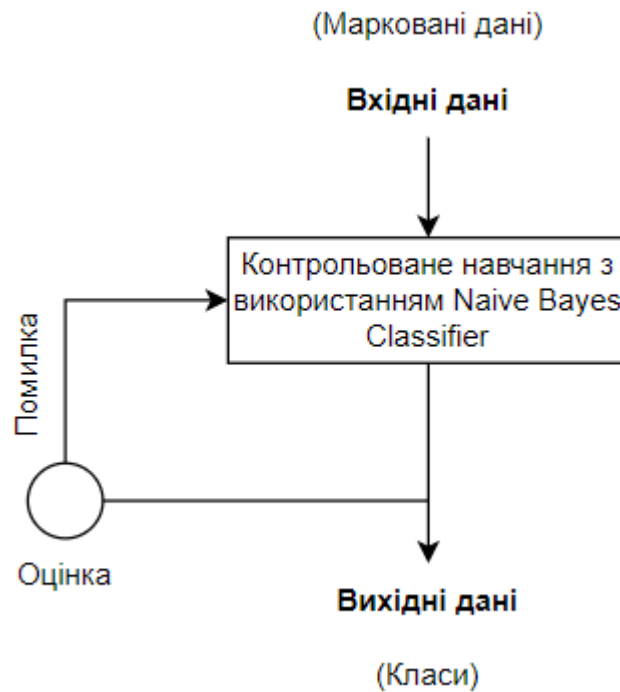


Рисунок 4.5 – Модель контрольованого навчання

Набір даних містить мітки, щоб функція могла обчислити помилку для заданого передбачення.

Навчання моделі відбувається шляхом створення функції, яку можна навчити за допомогою навчального набору даних, а потім застосувати до невидимих даних для досягнення певної прогнозованої продуктивності.

Мета полягає в тому, щоб побудувати функцію так, аби вона добре узагальнювала дані, яких ніколи не бачила.

Після виконання попереднього етапу маємо набір даних розділений на навчальні і тестові дані. Модель навчається на наборі навчальних даних поки не

досягне певного рівня продуктивності (показник того, наскільки точно функція відображає навчальні дані у відповідний бажаний результат).

Це відбувається з кожним навчальним зразком, де використовується помилка (відмінність фактичного результату від бажаного) для зміни навчальної функції. Далі навчена модель тестується на тестовому наборі даних.

Тестові дані не використовуються під час навчання і забезпечують хороший показник того, наскільки добре модель визначає результат за невидимими даними (рис. 4.6).

Навчена модель класифікатора здатна розпізнавати активність за поданим на вхід набором ознак. Результати класифікатора зберігаються в матриці невідповідностей. Це дозволяє візуалізувати ефективність алгоритму навчання (табл. 4.3).

Таблиця 4.3 – Позначення матриці невідповідності

| | | Справжній клас | |
|--------------------|----------------|-------------------------|-------------------------|
| | | Позитивний (P) | Негативний (N) |
| Прогнозований клас | Позитивний (P) | Істинно Позитивний (TP) | Хибно Позитивний (FP) |
| | Негативний (N) | Хибно Негативний (FN) | Істинно Негативний (TN) |

Істинно Позитивний результат означає кількість позитивних випадків, які були класифіковані як позитивні, Істинно Негативний – кількість негативних випадків, які були класифіковані як негативні, Хибно Позитивний – кількість негативних випадків, які були класифіковані як позитивні, Хибно Негативний – кількість позитивних випадків, які були класифіковані як негативні.

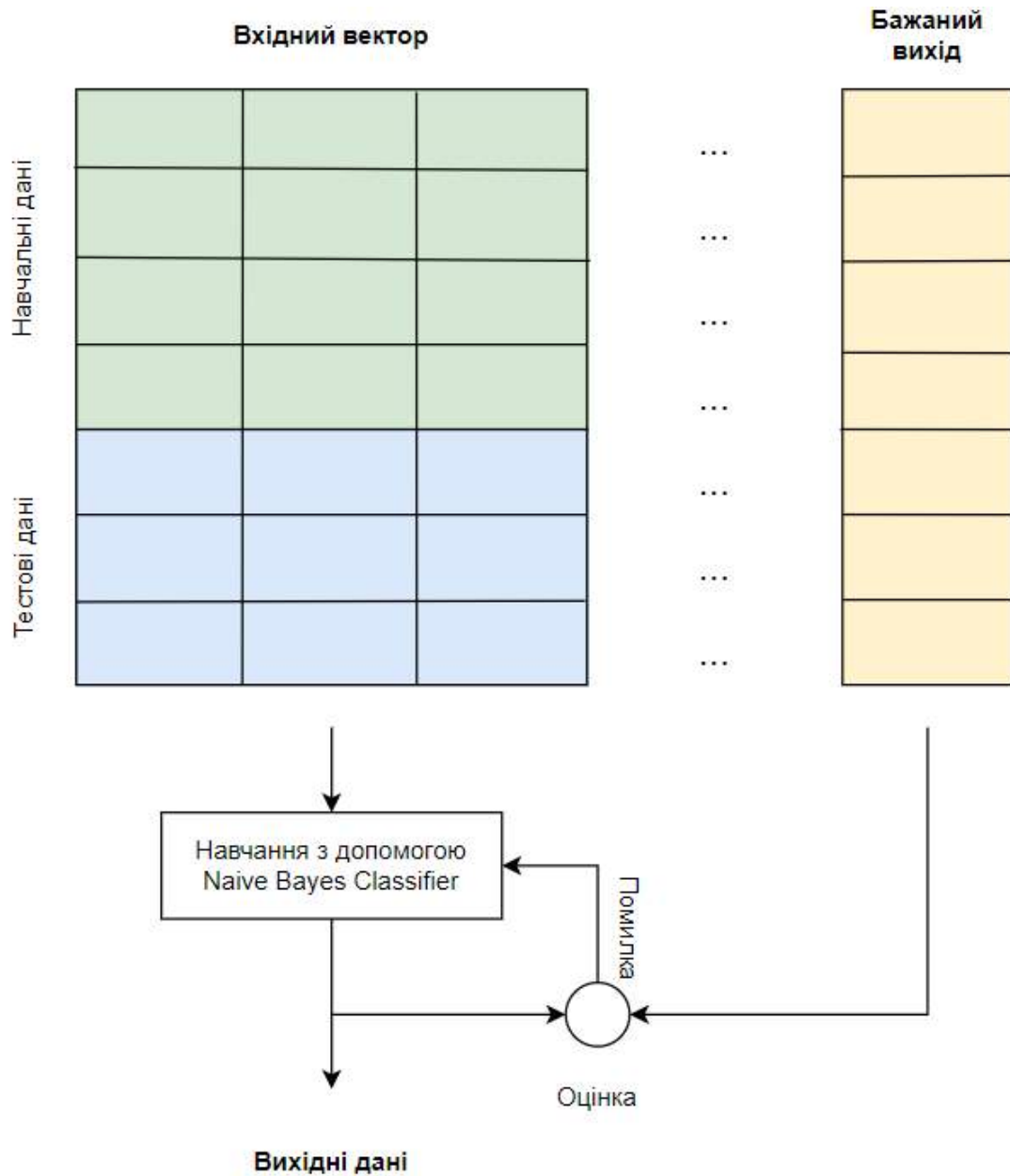


Рисунок 4.6 – Навчання та налаштування моделі на основі набору даних, розділеного на навчальну і тестову вибірки

Маючи ці показники, можемо обчислити точність системи, яка є найбільш широко використовуваним показником ефективності в задачах класифікації. Обчислимо її за наступною формулою:

$$acc = \frac{TN + TP}{(TP + TN + FP + FN)} \quad (4.1)$$

Також обчислимо позитивне прогнозоване значення – відношення кількості правильно класифікованих екземплярів до всього набору екземплярів, класифікованих як позитивні:

$$pr = \frac{TP}{(TP + FP)}. \quad (4.2)$$

Виміряємо справжній позитивний коефіцієнт – відношення кількості правильно класифікованих позитивних випадків до всього набору позитивних випадків:

$$rec = \frac{TP}{(TP + FN)}. \quad (4.3)$$

Цей параметр вказує здатність моделі виявляти позитивні зразки.

Також визначимо оцінку F1 – загальний показник точності моделі, який визначається на основі позитивного прогнозованого значення та справжнього позитивного коефіцієнту:

$$F1 = 2 * \frac{pr * rec}{(pr + rec)}. \quad (4.4)$$

Загальні результати розпізнавання активності з використанням Naïve Bayes Classifier зображені на рис. 4.7. Числа по горизонталі представляють передбачувані види активності, числа по вертикалі – активності, визначені за допомогою алгоритму, діагональна лінія представляє правильне виявлення людської активності. Отримано досить високу точність для більшості видів діяльності.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | Сума | TP | FP | FN | TN | pr | rec | F1 | acc | |
|----|-----|-----|-----|-----|-----|----|-----|-----|----|-----|-----|-----|----|-----|----|----|-----|-----|-----|----|----|----|----|----|------|-----|-----|----|----|--------|-------|--------|--------|-------|
| 1 | 152 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 152 | 152 | 0 | 0 | 0 | 100,00 | 100 | 100,00 | 100,00 | |
| 2 | 0 | 145 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 148 | 145 | 3 | 0 | 0 | 97,97 | 100 | 98,98 | 97,97 | |
| 3 | 0 | 0 | 153 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 161 | 153 | 8 | 0 | 0 | 95,03 | 100 | 97,45 | 95,03 | |
| 4 | 0 | 0 | 0 | 139 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 177 | 139 | 38 | 0 | 0 | 78,53 | 100 | 87,97 | 78,53 | |
| 5 | 0 | 18 | 0 | 0 | 143 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 184 | 143 | 41 | 0 | 0 | 77,72 | 100 | 87,46 | 77,72 | |
| 6 | 5 | 0 | 22 | 0 | 0 | 99 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 0 | 0 | 0 | 0 | 144 | 99 | 45 | 0 | 0 | 68,75 | 100 | 81,48 | 68,75 | |
| 7 | 42 | 0 | 0 | 14 | 0 | 0 | 130 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 190 | 130 | 60 | 0 | 0 | 68,42 | 100 | 81,25 | 68,42 | |
| 8 | 0 | 0 | 0 | 0 | 33 | 0 | 0 | 153 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 188 | 153 | 35 | 0 | 0 | 81,38 | 100 | 89,74 | 81,38 | |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 58 | 0 | 35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 124 | 58 | 66 | 0 | 0 | 46,77 | 100 | 63,74 | 46,77 | |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 123 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 127 | 123 | 4 | 0 | 0 | 96,85 | 100 | 98,40 | 96,85 | |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 152 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 15 | 0 | 0 | 0 | 0 | 170 | 152 | 18 | 0 | 0 | 89,41 | 100 | 94,41 | 89,41 | |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 144 | 120 | 24 | 0 | 0 | 83,33 | 100 | 90,91 | 83,33 | |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 0 | 0 | 139 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 160 | 139 | 21 | 0 | 0 | 86,88 | 100 | 92,98 | 86,88 | |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 115 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 125 | 115 | 10 | 0 | 0 | 92,00 | 100 | 95,83 | 92,00 | |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 123 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 125 | 123 | 2 | 0 | 0 | 98,40 | 100 | 99,19 | 98,40 | |
| 16 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 26 | 1 | 0 | 0 | 96,30 | 100 | 98,11 | 96,30 | |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 127 | 120 | 7 | 0 | 0 | 94,49 | 100 | 97,17 | 94,49 | |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 123 | 120 | 3 | 0 | 0 | 97,56 | 100 | 98,77 | 97,56 | |
| 19 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 153 | 0 | 0 | 0 | 0 | 0 | 176 | 153 | 23 | 0 | 0 | 86,93 | 100 | 93,01 | 86,93 | |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 89 | 0 | 27 | 0 | 0 | 144 | 89 | 55 | 0 | 0 | 61,81 | 100 | 76,39 | 61,81 | |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 60 | 0 | 0 | 0 | 0 | 61 | 60 | 1 | 0 | 0 | 98,36 | 100 | 99,17 | 98,36 | |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 156 | 151 | 5 | 0 | 0 | 96,79 | 100 | 98,37 | 96,79 | |
| 23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 168 | 133 | 35 | 0 | 0 | 79,17 | 100 | 88,37 | 79,17 | |
| 24 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 140 | 169 | 140 | 29 | 0 | 0 | 82,84 | 100 | 90,61 | 82,84 |

Рисунок 4.7 – Матриця невідповідності з результатами досліджень

Найнижчі показники в діяльностей під номером 9 і 20 – «Перегляд телевізора» і «Прання». Діяльність «Перегляд телевізора» багато разів було визначено як «Гра у відеоігри», оскільки в обох випадках спрацьовували датчики дивану, телевізора, датчики підлоги у вітальні. Діяльність «Прання» визначалась як «Миття посуду» через розташування пральної машини на кухні або «Переодягання» через показники датчика корзини для білизни. Найвищі показники у проміжку 98%-100% отримані у визначенні діяльностей 1, 15, та 21 – «Прийом ліків», «Викидання сміття» і «Робота за столом». Це можна пояснити тим, що набір ознак для цих діяльностей мало перетинається з ознаками інших діяльностей.

4.4 Висновки

Таким чином в результаті дослідження було удосконалено інтелектуальну систему прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини, яка на відмінну від відомих залучає методи машинного навчання, що дозволило підвищити ступінь автоматизації процесів у розумних будинках на основі видів людської діяльності та налаштовувати параметри розумних пристроїв під потреби користувачів та/або оптимізацію енергоспоживання.

ВИСНОВКИ

За результатами проведеного дослідження запропоновано інтелектуальну систему прийняття рішення щодо функціонування елементів розумного будинку на основі методу розпізнавання активностей людини в розумному будинку. Запропонований метод забезпечує ефективне визначення людської діяльності та демонструє високі показники точності.

В першому розділі розглянуто концепцію Інтернету речей, наведено загальну архітектуру IoT системи. Проведено огляд існуючих методів аналізу даних розпізнавання людської діяльності. Проаналізовано переваги та недоліки існуючих рішень розпізнавання людської діяльності штучним інтелектом в контексті Інтернету речей.

В другому розділі представлений опис предметної області та визначені основні функціональні і нефункціональні вимоги системи. Запропоновано модель процесу класифікації життєдіяльності людини в розумному будинку, яка заснована на Naïve Bayes Classifier та дозволяє визначити найбільш ймовірну активність, виконувану мешканцем будинку за набором ознак.

В третьому розділі запропоновано метод розпізнавання людської активності в розумному будинку, що заснований на використанні байєсівських мереж. Набір початкових даних розділено на два набори, виведено алгоритм класифікації, в якому класифікатор навчається на тренувальному наборі та тестується на тестовому. В результаті виконання методу буде отримано вид діяльності людини. Перевагою методу є його надійність, яка забезпечує достатньо точний результат при роботі як з малим, так і з великим набором даних.

У четвертому розділі на основі моделі та методу розпізнавання людської активності в розумному будинку запропоновано удосконалення інтелектуальної системи прийняття рішення щодо функціонування елементів розумного будинку, яка застосовує методи машинного навчання, що дозволило підвищити ступінь автоматизації процесів у розумних будинках на основі видів людської

діяльності та налаштовувати параметри розумних пристроїв під потреби користувачів та/або оптимізацію енергоспоживання.

За темою дипломної роботи опубліковано тезу у збірнику Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2021».

ПЕРЕЛІК ДЖНЕРЕЛ ПОСИЛАНЬ

1. Стрелкова О.Г., Нічепорук А.А., Казанцев А.Д., Нічепорук А.О. Інтелектуальна система класифікації видів діяльності, зафіксованих у розумних будинках. *Збірник наукових праць за матеріалами XIII Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2021»*. Хмельницький. – 2021. – С. 238-241.
2. Karvonen, Niklas and Denis Kleyko. A Domain KnowledgeBased Solution for Human Activity Recognition: The UJA Dataset Analysis. *The 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018)*. Vol. 2. 19 1261. Punta Cana, Dominican Republic: MDPI. URL: <https://doi.org/10.3390/proceedings2191261>
3. Lago, Paula and Sozu Inoue. A Hybrid Model Using Hidden Markov Chain and Logic Model for Daily Living Activity Recognition. *The 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018)*. Vol. 2. 19 1266. Punta Cana, Dominican Republic: MDPI. URL: <https://doi.org/10.3390/proceedings2191266>
4. Salomon, Sergio and Cristina Tirnauce. Human Activity Recognition through Weighted Finite Automata. *The 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018)*. Vol. 2. 19 1263. Punta Cana, Dominican Republic: MDPI. URL: <https://doi.org/10.3390/proceedings2191263>
5. Jimenez, Antonio R. and Fernando Seco. Multi-Event Naive Bayes Classifier for Activity Recognition in the UCAmI Cup. *The 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018)*. Vol. 2. 19 1264. Punta Cana, Dominican Republic: MDPI. URL: <https://doi.org/10.3390/proceedings2191264>
6. Razzaq, Muhammad Asif et al. Multimodal Sensor Data Fusion for Activity Recognition Using Filtered Classifier. *The 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018)*. Vol. 2. 19 1262.

Punta Cana, Dominican Republic: MDPI. URL: <https://doi.org/10.3390/proceedings2191262>

7. Ceron, Jesus, Diego Lopez, and Bjoern Eskofier. Human Activity Recognition Using Binary Sensors, BLE Beacons, an Intelligent Floor and Acceleration Data: A Machine Learning Approach. *The 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018)*. Vol. 2. 19 1265. Punta Cana, Dominican Republic: MDPI. URL: <https://doi.org/10.3390/proceedings2191265>

8. Wirth, Rudiger and Jochen Hipp. CRISP-DM: Towards a Standard Process Model for DataMining. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, pp. 29–39.

9. Hammerla, Nils Y. and Shane Halloran Thomas Plotz. Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables. *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)*, pp. 1533–1540, 2016. URL: <https://www.ijcai.org/Proceedings/16/Papers/220.pdf>

10. Espinilla, Macarena, Javier Medina, and Chris Nugent. UCAmI Cup. Analyzing the UJA Human Activity Recognition Dataset of Activities of Daily Living. *The 12th International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI 2018)*. Vol. 2. 19 1267. Punta Cana, Dominican, Dec. 2018. URL: <https://www.mdpi.com/2504-3900/2/19/1267>

11. Priyadarshini, Sushree Bibhuprada B., Amiya Bhusan Bagjadab, and Brojo Kishore Mishra. The Role of IoT and Big Data in Modern Technological Arena: A Comprehensive Study. *Intelligent Systems Reference Library. Springer International Publishing*. Chap. 2, vol. 154, pp. 13–25, 2019. isbn: 9783030042035. URL: https://link.springer.com/chapter/10.1007/978-3-030-04203-5_2

12. Pantazis, N. A., Stefanos, A. N., Dimitrios, D. V. Energy-efficient routing protocols in wireless sensor networks: A survey. *IEEE Commun. Surveys Tutor*. 2013, 15, 551–591.

13. Liu, X. A survey on clustering routing protocols in wireless sensor networks. *Sensors* 2012, 12, 11113–11153.
14. De Silva, L.C.; Morikawa, C.; Petra, I.M. State of the art of smart homes. *Eng. Appl. Artif. Intell.* 2012, 25, 1313–1321.
15. Zhang, D.; Shah, N.; Papageorgiou, L.G. Efficient energy consumption and operation management in a smart building with microgrid. *Energy Convers. Manag.* 2013, 74, 209–222.
16. Pedrasa, M.; Spooner, T.; MacGill, I. Coordinated Scheduling of Residential Distributed Energy Resources to Optimize Smart Home Energy Services. *IEEE Trans. Smart Grid* 2010, 1, 134–143.
17. United States Department of Commerce, National Telecommunications and Information Administration, United States Department of Commerce. URL: <http://www.its.bldrdoc.gov>
18. Gu, H.; Diao, Y.; Liu, W.; Zhang, X. The design of smart home platform based on Cloud Computing. *Proceeding of the International Conference on the Design of Smart Home Platform Based on Cloud Computing*, Harbin, China, 12–14 August 2011.
19. Płaczek, B.; Bernaś, M. Uncertainty-based information extraction in wireless sensor networks for control applications. *Ad Hoc Netw.* 2014, 14, 106–117.
20. Saponara, S.; Bacchillone, T. Network Architecture, Security Issues, and Hardware Implementation of a Home Area Network for Smart Grid. *J. Comput. Netw. Commun.* 2012
21. Gonzalez-Tablas, A.; Ferreres, A.I.; Ramos Alvarez, B.; Garnacho, A. Guaranteeing the Authenticity of Location Information. *IEEE Pervasive Comput.* 2008, 7, 72–80.
22. Xing, F.; Wang, W. Analyzing Resilience to Node Misbehaviors in Wireless Multi-Hop Networks. *Proceedings of the IEEE Wireless Communications and Networking Conference, WCNC 2007*, Kowloon, Hong Kong, China, 11–15 March 2007.

23. Liu, X. A survey on clustering routing protocols in wireless sensor networks. *Sensors* 2012, 12, 11113–11153.
24. Wang, P.; Yao, C.; Zheng, Z.; Sun, G.; Song, L. Joint Task Assignment, Transmission, and Computing Resource Allocation in Multilayer Mobile Edge Computing Systems. *IEEE IoT J.* 2018, 6, 2872–2884.
25. Ren, J.; He, Y.; Huang, G.; Yu, G.; Cai, Y.; Zhang, Z. An edge-computing based architecture for mobile augmented reality. *IEEE Netw.* 2019, 33, 162–169.
26. Heinzelman, W.R.; Chandrakasan, A.; Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, Maui, HI, USA, 4–7 January 2000.
27. Younis, O.; Sonia, F. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans. Mobile Comput.* 2004, 3, 366–379.
28. Cherchi, R.; Colistra, G.; Pilloni, V.; Atzori, L. Energy consumption management in Smart Homes: An M-Bus communication system. *Proceedings of the 2014 International Conference on Telecommunications and Multimedia (TEMU)*, Heraklion, Greece, 28–30 July 2014
29. Hu, Q.; Li, F. Hardware Design of Smart Home Energy Management System with Dynamic Price Response. *IEEE Trans. Smart Grid* 2013, 4, 1878–1887.
30. Jo, H.; Kim, S.; Joo, S. Smart heating and air conditioning scheduling method incorporating customer convenience for home energy management system. *IEEE Trans. Consum. Electron.* 2013, 59, 316–322.
31. Peruzzini, M.; Germani, M.; Papetti, A.; Capitanelli, A. Smart Home Information Management System for Energy-Efficient Networks. *Collaborative Systems for Reindustrialization*; Springer: Berlin, Germany; Heidelberg, Germany; Dresden, Germany, 2013; Volume 408, pp. 393–401.
32. Han, D.M.; Lim, J.H. Smart home energy management system using IEEE 802.15.4 and ZigBee. *IEEE Trans. Consum. Electron.* 2010, 56, 1403–1410.

33. Han, D.M.; Lim, J.H. Design and implementation of smart home energy management systems based on ZigBee. *IEEE Trans. Consum. Electron.* 2010, 56, 1417–1425.
34. R. Want. An introduction to RFID technology. *IEEE Pervasive Computing*, vol. 5, no. 1, pp. 25–33, Jan 2006.
35. L. Atzori, A. Iera, and G. Morabito. The Internet of Things: A Survey. *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
36. M. Friedli, L. Kaufmann, F. Paganini, and R. Kyburz. Energy Efficiency of the Internet of Things - Technology and Energy Assessment Report. *International Energy Agency*, vol. 1, 2016.
37. M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, A. P. Shath. Machine learning for Internet of Things data analysis: A survey. *Digital Community Network*, pp. 163-167, 2017.
38. E. M. Tapia, S. S. Intille, K. Larson. Activity recognition in the home using simple and ubiquitous sensors. *Pervasive*, vol. 4, pp. 158-175, 2004.
39. D. Pavithra, R. Balakrishnan. IoT based monitoring and control system for home automation. *Global Conference on Communication Technologies-GCCT*, pp. 169-173, Apr 2015.
40. M. Pandurov, B. Petelj, R. Pavlovic, N. Teslic. Platform for extending home automation gateway's functionality with plugin mechanism. *IEEE 5th International Conference on Consumer Electronics*, pp. 354-357, 2015.
41. J. Wei. How Wearables Inersect with the Cloud and the Internet of Things: Considerations for the developers of wearables. *IEEE Consumer Electronics Magazine*, 3(3), 53-56, 2014.
42. J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems*, 29(7), pp. 1645-1660, 2013.
43. N. Li, N. Zhang, S. Das, B. Thuraisingham. Privacy preservation in wireless sensor networks: A state-of-the-art survey. *Ad Hoc Networks*, vol. 7, no. 8, pp. 1501-1514, 2009.

44. P. Kinney. Zigbee technology: Wireless control that simply works. *Communications design conference*, vol. 2, pp. 1-7, October, 2003.
45. D. Minoli, K. Sohraby, B. Occhiogrosso. IoT Considerations, Requirements, and Architectures for Smart Buildings – Energy Optimization and Next-Generation Building Management Systems. *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 269-283, Feb. 2017.
46. J. Wan et al. Software-Defined Industrial Internet of Things in the Context of Industry 4.0. *IEEE Sensors Journal*, vol. 16, no. 20, pp. 7373-7380, Oct. 15, 2016.
47. C. C. Aggarwal, N. Ashish, A. Sheth. The Internet of Things: A Survey from the Data-Centric Perspective. *Managing and Mining Sensor Data*, Springer US, pp. 383-428, 2013.
48. D. Miorandi, S. Sicari, F. De Pellegrini, I. Chlamtac. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497-1516, 2012.
49. M. Palattela, N. Accettura, X. Vilajosana, T. Watteyne, L. Grieco, G. Boggia, M. Dohler. Standardized Protocol Stack for the Internet of (Important) Things. *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1389-1406, 2013.
50. A. Jara, A. Olivieri, Y. Bocchi, M. Jung, W. Kastner, A. Skarmeta. Semantic Web of Things: an analysis of the application semantics for the IoT moving towards the IoT convergence. *International Journal of Web and Grid Services*, vol. 10, no. 23, p. 244, 2014.

ДОДАТОК А
(обов'язковий)
Презентація доповіді

Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Виконав: студент групи KI2м-20-1 Стрелкова О. Г.

Науковий керівник: к.т.н., Нічепорук А. О.

Хмельницький, 2022




Мета, об'єкт та предмет дослідження

Об'єкт дослідження – процес класифікації видів діяльності, зафіксованих у розумних будинках.

Предмет дослідження – інтелектуальна система класифікації видів діяльності, зафіксованих у розумних будинках.

Мета роботи – підвищення ступеню автоматизації та реагування на потреби мешканців розумного будинку шляхом проектування інтелектуальної системи класифікації видів діяльності, зафіксованих у розумних будинках.



Наукова новизна

Удосконалено інтелектуальну систему прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини, яка на відмінну від відомих залучає методи машинного навчання, що дозволило підвищити ступінь автоматизації процесів у розумних будинках на основі видів людської діяльності та налаштувати параметри розумних пристроїв під потреби користувачів та/або оптимізацію енергоспоживання.

Набула подальшого розвитку модель процесу розпізнавання людської активності на основі наївного байєсового класифікатора, яка дозволяє визначити найбільш ймовірну активність, виконувану мешканцем розумного будинку за набором ознак, що дозволило формалізувати основні етапи розпізнавання людської діяльності в середовищі розумного будинку.

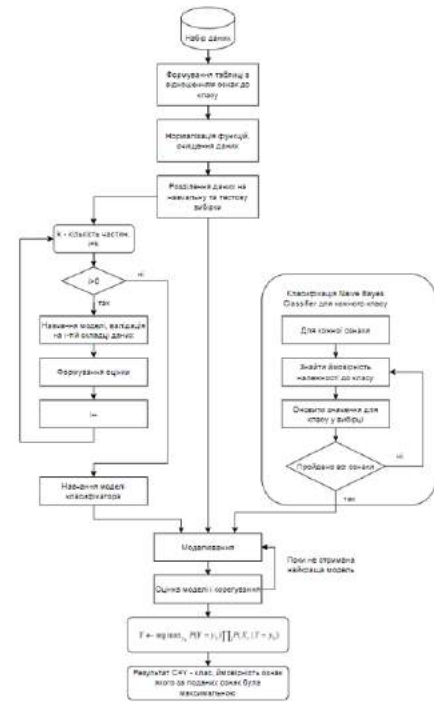
Потенційний сценарій

Відслідковування та нагадування системою розумного будинку про прийом ліків.

Мешканець може приймати гормональні препарати для щитовидної залози або препарати для підтримання рівня цукру на постійній основі. Регулярний прийом таких ліків може бути критично важливим для здоров'я людини. Мешканець може забути про цей прийом, і в цьому випадку має спрацювати система сповіщень.

Система відслідкуватиме прийом ліків за допомогою даних розміщення розумної підлоги, магнітного контакту і датчика наближення коробки з ліками, та даних прискорення розумного годинника. При завершенні сегменту дня (ранок, обід або вечір), в який повинен був відбутись прийом ліків має існувати перевірка чи датчики коробки з ліками спрацювали, якщо ні – діяльність була пропущена. У випадку пропуску система має нагадати про обов'язковий прийом ліків за допомогою сповіщення, надіслане на пристрій мешканця з встановленим програмним забезпеченням для цього.

- Формування таблиці з відношенням ознак до класу
- Нормалізація функцій, очищення даних
- Розділення даних на навчальну та тестову вибірки
- Класифікація
- Моделювання
- Оцінка моделі і корегування
- Визначення класу на основі поданих ознак



Метод розпізнавання активностей людини в розумному будинку на основі використання Naïve Bayes Classifier

Метод розпізнавання активностей людини в розумному будинку на основі використання Naïve Bayes Classifier

Виконаємо нормалізацію на небінарних даних, таких як дані прискорення, за допомогою MinMaxScaler для підвищення ефективності моделі машинного навчання, а також зменшення часу необхідного для вивчення даних. Небінарні дані мають різну шкалу, і це може призвести до меншої ефективності атрибута з меншою шкалою, через те, що інші атрибути мають значення у більшому масштабі. Використаємо формулу нормалізації для кожного значення:

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

де x – значення з даних.

Це перетворить всі значення в діапазон $[0,1]$, що означає, що мінімальне та максимальне значення будуть 0 та 1 відповідно.

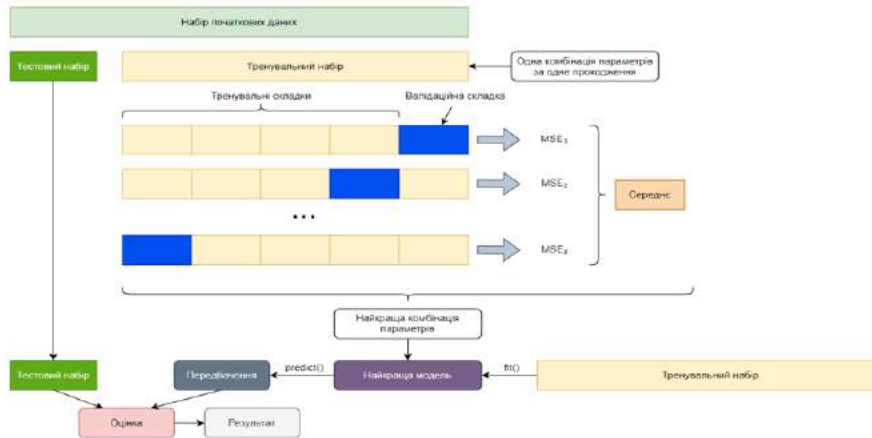
Метод розпізнавання активностей людини в розумному будинку на основі використання Naïve Bayes Classifier

Деякі вибірки в початкових даних позначені як «без активності», хоча і містять дані датчиків. Вилучимо зразки з цією міткою, оскільки вони можуть погіршити якість навчання моделі.

Розіб'ємо набір даних для моделі машинного навчання на два окремих набори – навчальний набір і тестовий набір методом k-кратного перехресного затвердження:

- Довільно розділимо набір даних на k груп приблизно однакового розміру
- k-1 частини використаємо для навчання моделі, а одну k частину використаємо для випробування і оцінки ефективності
- Повторимо процедуру k разів, щоб отримати k кількість оцінок продуктивності для кожної ітерації
- Результуючу ефективність обрахуємо як середнє значення k кількості оцінок.

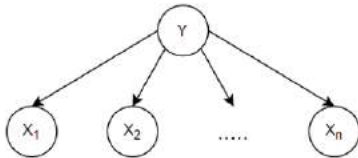
Розділення даних на вибірки



Класифікація з використанням Naive Bayes Classifier

Завдання даного класифікатора полягає в тому, щоб знайти клас спостереження за значеннями ознак, тобто обчислити ймовірність класу з урахуванням набору значень ознак $P(Y|X)$:

$$f: X \rightarrow Y, \quad \begin{matrix} \text{де } Y - \text{змінна, що представляє клас,} \\ X - \text{вектор, який містить } n \text{ ознак.} \end{matrix} \quad (1)$$



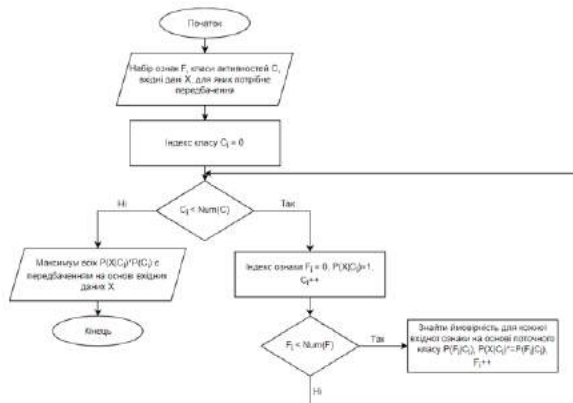
Структура Naive Bayes Classifier

В цьому дослідженні класом виступає одна з двадцяти чотирьох визначених заздалегідь активностей людини, а вектором ознак X дані від різних датчиків розумного будинку:

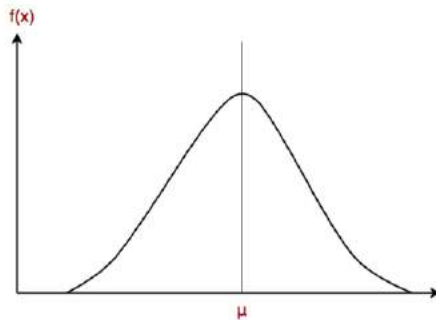
$$X = \langle X_1, X_2, \dots, X_n \rangle, \quad \begin{matrix} \text{де } X_i - \text{значення } i\text{-тої ознаки вектора } X. \end{matrix} \quad (2)$$

Визначення діяльності людини за поданими ознаками

$$Y \leftarrow \arg \max_y P(Y = y_i) \prod_j P(X_j | Y = y_i), \quad (3)$$



Знаходження ймовірності для кожної ознаки



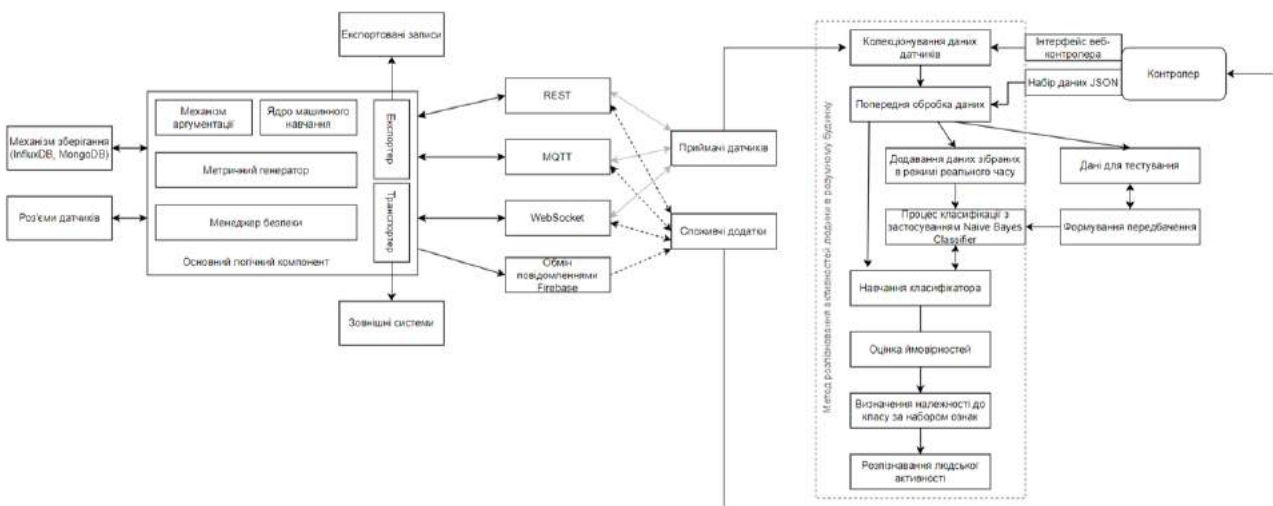
Розподіл Гауса на графіку

Використаємо Gaussian Naive Bayes Classifier. Цей підхід полягає в припущенні, що для кожного можливого дискретного значення y_k з Y розподіл кожного безперервного X_i є розподілом Гауса і визначається математичним сподіванням і середнім квадратичним відхиленням характерним для X_i і y_k . При нанесенні на графік розподілу Гауса зображається як крива, симетрична відносно середніх значень характеристик

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

де σ – середнє квадратичне відхилення,
 μ – математичне сподівання.

(4)



Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Загальні записи датчиків і записи метаданих зберігаються за допомогою високопродуктивного, масштабованого механізму зберігання даних з низькою затримкою, заснованого на перевірених технологіях з відкритим вихідним кодом. Цей механізм зберігання містить бази даних MongoDB і InfluxDB.



InfluxDB – це база даних часових рядів, оптимізована для зберігання та пошуку даних, яка використовує часові позначки як індекс. В ній зберігаються дані датчиків.



MongoDB використовується для зберігання та керування метаданими датчиків, профілями користувачів, ключами API тощо.

Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Механізм зберігання підключений до основного логічного компонента. Цей компонент містить ряд елементів: менеджер безпеки, підпрограми генерації показників, механізм аргументації на основі правил, ядро машинного навчання та експортер і транспортер записів.

Менеджер безпеки використовується для аутентифікації, верифікації та надання криптографічних послуг.

Підпрограми генерації на стороні сервера дозволяють споживчим програмам полегшити обробку, зменшуючи мережевий трафік і скорочуючи час обробки.

Механізм аргументації використовується для класифікації показників, правила для цього вказуються за допомогою веб-інтерфейсу або через Java API.

Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Машинне навчання реалізовується за допомогою Python пакета [Scikit Learn](#).

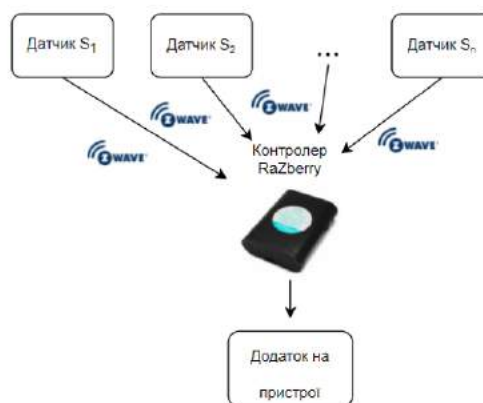
Пересилання записів даних в реальному часі відбувається за допомогою викликів REST та [WebSockets](#).

Система створює один документ JSON, що містить усі пов'язані дані датчиків та експериментальні дані, який експортується за допомогою функції експорту.

Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Приймачі датчиків зчитують дані датчиків низького рівня з сенсорних пристроїв.

Роз'єми даних датчиків полегшують інтеграцію з іншими сенсорними платформами і мережами. Роз'єм RaZberry забезпечує безперебійну інтеграцію з датчиками на основі Z-Wave. Цей роз'єм передає всі дані від датчиків, які зареєстровані в екземплярі RaZberry.

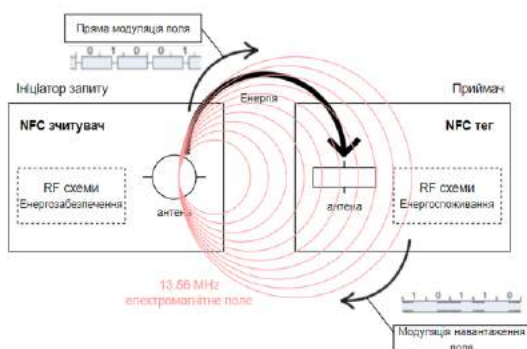


Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Для керування даними використовується модульний веб-додаток, розроблений з використанням сучасних технологій, а саме фреймворку Angular та фреймворку інтерфейсів bootstrap.

Веб-додаток має бути розширюваним і модульним, що дозволить дослідникам і розробникам створювати модулі, які можна буде інтегрувати в загальну платформу. Це дасть можливість адаптувати існуючу веб-програму в менший модуль, що дозволить швидко створити спеціальний інтерфейс з меншою складністю. Такий інтерфейс дозволить повторне використання програмного забезпечення для використання сторонніми веб-додатками.

Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини



Діяльність, що виконується, анується за допомогою NFC тегів і смартфона. Цей процес використовується для позначення початку та кінця кожної дії.

Щоразу, коли зчитувач з живленням NFC підноситься до тегу, тег отримує живлення і передає будь-які збережені дані в мікрочіпі на пристрій. Зчитувач NFC живить теги і зчитує дані з пасивних тегів. Щоб робити це зчитувач потребує джерело живлення і пропускає електричний струм через власну котушку. Це створює змінне магнітне поле в безпосередній близькості від зчитувача. Введення тегу в діапазон цього поля призводить до індуктивного зв'язку між двома котушками завдяки закону індукції Фарадея.

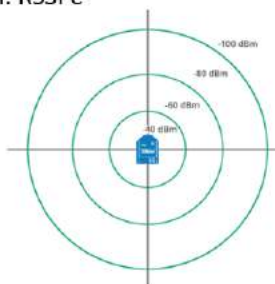
Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Дані прискорення генеруються розумним годинником. Годинник має вбудований акселерометр, датчик якого розраховує силу прискорення, яка впливає на пристрій по трьох осях (X, Y і Z).



Дані наближення вимірюються датчиками RSSI. RSSI є вимірюванням сили отриманого сигналу.

Це значення завжди негативне, оскільки показує скільки сили сигналу було втрачено. Чим вище значення RSSI, тим сильніший сигнал, тобто -50 – досить хороший сигнал, а -100 – взагалі відсутній.



| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | Сума | TP | FP | FN | TN | pr | rec | F1 | acc | |
|----|-----|-----|-----|-----|-----|----|-----|-----|----|-----|-----|-----|-----|-----|-----|----|----|-----|----|-----|----|----|-----|----|------|-----|-----|----|----|--------|-------|--------|--------|-------|
| 1 | 152 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 152 | 152 | 0 | 0 | 0 | 100,00 | 100 | 100,00 | 100,00 | |
| 2 | 0 | 145 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 148 | 145 | 3 | 0 | 0 | 97,97 | 100 | 98,98 | 97,97 |
| 3 | 0 | 0 | 153 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 161 | 153 | 8 | 0 | 0 | 95,03 | 100 | 97,45 | 95,03 |
| 4 | 0 | 0 | 0 | 139 | 0 | 0 | 0 | 38 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 177 | 139 | 38 | 0 | 0 | 78,53 | 100 | 87,97 | 78,53 |
| 5 | 0 | 18 | 0 | 0 | 143 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 184 | 143 | 41 | 0 | 0 | 77,72 | 100 | 87,46 | 77,72 |
| 6 | 5 | 0 | 22 | 0 | 0 | 99 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 144 | 99 | 45 | 0 | 0 | 68,75 | 100 | 81,48 | 68,75 |
| 7 | 42 | 0 | 0 | 14 | 0 | 0 | 130 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 190 | 130 | 60 | 0 | 0 | 68,42 | 100 | 81,25 | 68,42 |
| 8 | 0 | 0 | 0 | 0 | 33 | 0 | 0 | 153 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 188 | 153 | 35 | 0 | 0 | 81,38 | 100 | 89,74 | 81,38 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 58 | 0 | 35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 31 | 0 | 0 | 0 | 124 | 58 | 66 | 0 | 0 | 46,77 | 100 | 63,74 | 46,77 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 123 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 127 | 123 | 4 | 0 | 0 | 96,85 | 100 | 98,40 | 96,85 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 152 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 15 | 0 | 0 | 0 | 0 | 0 | 170 | 152 | 18 | 0 | 0 | 89,41 | 100 | 94,41 | 89,41 |
| 12 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 144 | 120 | 24 | 0 | 0 | 83,33 | 100 | 90,91 | 83,33 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 0 | 0 | 139 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 160 | 139 | 21 | 0 | 0 | 86,88 | 100 | 92,98 | 86,88 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 115 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 125 | 115 | 10 | 0 | 0 | 92,00 | 100 | 95,83 | 92,00 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 123 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 125 | 123 | 2 | 0 | 0 | 98,40 | 100 | 99,19 | 98,40 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 26 | 1 | 0 | 0 | 96,30 | 100 | 98,11 | 96,30 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 127 | 120 | 7 | 0 | 0 | 94,49 | 100 | 97,17 | 94,49 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 123 | 120 | 3 | 0 | 0 | 97,56 | 100 | 98,77 | 97,56 |
| 19 | 3 | 0 | 0 | 0 | 0 | 0 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 153 | 0 | 0 | 0 | 0 | 0 | 176 | 153 | 23 | 0 | 0 | 86,93 | 100 | 93,01 | 86,93 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 89 | 0 | 27 | 0 | 0 | 0 | 0 | 0 | 144 | 89 | 55 | 0 | 0 | 61,81 | 100 | 76,39 | 61,81 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 60 | 0 | 0 | 0 | 0 | 0 | 61 | 60 | 1 | 0 | 0 | 98,36 | 100 | 99,17 | 98,36 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 156 | 151 | 5 | 0 | 0 | 96,79 | 100 | 98,37 | 96,79 |
| 23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 35 | 133 | 0 | 0 | 168 | 133 | 35 | 0 | 0 | 79,17 | 100 | 88,37 | 79,17 |
| 24 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 169 | 140 | 29 | 0 | 0 | 82,84 | 100 | 90,61 | 82,84 |


Результати досліджень

Числа по горизонталі представляють передбачувані види активності, числа по вертикалі – активності, визначені за допомогою алгоритму, діагональна лінія представляє правильне виявлення людської активності. Отримано досить високу точність для більшості видів діяльності.


Оцінка ефективності системи

Найнижчі показники в діяльностей під номером 9 і 20 – «Перегляд телевізора» і «Прання». Діяльність «Перегляд телевізора» багато разів було визначено як «Гра у відеоігри», оскільки в обох випадках спрацьовували датчики дивану, телевізора, датчики підлоги у вітальні. Діяльність «Прання» визначалась як «Миття посуду» через розташування пральної машини на кухні або «Переодягання» через показники датчика корзини для білизни.

Найвижчі показники у проміжку 98%-100% отримані у визначенні діяльностей 1, 15, та 21 – «Прийом ліків», «Викидання сміття» і «Робота за столом». Це можна пояснити тим, що набір ознак для цих діяльностей мало перетинається з ознаками інших діяльностей.



Дякую за увагу!



Ім'я користувача:
Кафедра КІ

ID перевірки:
1010957762

Дата перевірки:
26.04.2022 17:47:15 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
26.04.2022 17:47:36 EEST

ID користувача:
100005591

Назва документа: Інтелектуальна система прийняття рішення щодо функціонування елементів розумного бу...

Кількість сторінок: 72 Кількість слів: 12189 Кількість символів: 93735 Розмір файлу: 2.07 MB ID файлу: 1010865618

1.57% Схожість

Найбільша схожість: 0.56% з джерелом з Бібліотеки (ID файлу: 1007705660)

0.53% Джерела з Інтернету

34

Сторінка 74

1.12% Джерела з Бібліотеки

93

Сторінка 74

0.72% Цитат

Цитати

3

Сторінка 75

Не знайдено жодних посилань

0% Вилучень

Немає вилучених джерел

Модифікації

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Замінені символи

5

Anti-Plagiarism v-15.257

Максимальное совпадение с одним документом 1.0%

Словари проверки: en_US, ru_RU, ua_UA. **Ошибок в документах: 8%**

| | | | | |
|---|----------|---------|-------------------------------------|---------|
| ID: 103084 Название: Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини Добавлено в БД: 2022-04-26 Автори: Стрелкова О.Г. Руководители: Нічепорук А.О. Консультанти: Опоненти: | Документ | | Суммарное совпадение по Базе Данных | |
| | Символы | Лексемы | Символы | Лексемы |
| | 80312 | 725 | 960 (1%) | 21 (3%) |

Источник плагиата

| ID | Описание | Наличие плагиата в документе | |
|----|----------|------------------------------|---------|
| | | Символы | Лексемы |

РЕЦЕНЗІЯ НА ДИПЛОМНУ РОБОТУ

Дипломник: Стрелкова Олександра Георгіївна

Тема: Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Спеціальність: 123 «Комп'ютерна інженерія»

Обсяг дипломної роботи:

Кількість листів креслень —; кількість сторінок записки 97

1. Короткий зміст роботи та прийнятих рішень У роботі запропоновано структуру інтелектуальної системи прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини.

2. Висновок про відповідність роботи дипломному завданню Дипломна робота відповідає виданому завданню

3. Характеристика виконання кожного розділу, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У першому розділі проведено огляд концепції Інтернету речей та наведено загальну архітектуру IoT системи. Розглянуто існуючі методи аналізу даних розпізнавання людської діяльності. У другому розділі запропоновано модель процесу класифікації життєдіяльності людини в розумному будинку. У третьому розділі запропоновано метод розпізнавання людської активності в розумному будинку, що заснований на використанні байєсівських мереж. У четвертому розділі запропоновано структуру інтелектуальної системи прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини.

4. Позитивні сторони роботи: Запропоновані модель процесу, метод та інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини в комплексі

дозволяють підвищити ступінь автоматизації процесів у розумних будинках на основі видів людської діяльності.

5. Негативні сторони роботи: В дослідженні було б бажано також провести експерименти на власному наборі даних від розумних пристроїв, з описом процесу збору, обробки та зберігання цих даних. Не наведено практичні рекомендації по реалізації запропонованої інтелектуальної системи підтримки прийняття рішення.

6. Оцінка графічного оформлення та пояснювальної записки роботи: =

7. Відгук про роботу в цілому: В загальному робота виконана на достатньому рівні.

8. Інші зауваження: =

9. Оцінка дипломної роботи:

Розглянувши позитивні та негативні сторони представленої дипломної роботи вважаю, що робота заслуговує оцінки «добре» 4,5 (В)

Рецензент (прізвище, ім'я, по батькові, посада, місце роботи)

Мартиш Валерій Володимирович, д.т.н., проф.
Зав. каф. АКТ ХНУ



" 4 " травня 2022р.

Завідувачу кафедри КПС
д-р.техн.наук, проф. Говорушенко Т. О.

Стрельцоваї Александри Гергіївни
ІІБ здобувача вищої освіти

ФІТ, 2 курсу, групи КІ2М-20-1

ЗАЯВА

З правилами чинного Положення «Про дотримання академічної доброчесності в Хмельницькому національному університеті» від 26.09.2020 (зі змінами від 26.11.2020), згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування заходів дисциплінарної та академічної відповідальності, ознайомлений (а). Про використання програмно-технічних засобів для перевірки кваліфікаційних робіт здобувачів вищої освіти на плагіатоповіщений (а) та надаю свою згоду на обробку та збереження університетом моєї роботи в інституційному репозитарії університету.

Також надаю університету право на передачу моєї роботи для обробки та збереження в базах даних програмно-технічних засобів (Unicheck та Anti-Plagiarism) та використання роботи для виявлення плагіату в інших роботах, які перевіряються програмно-технічними засобами та користувачами, що мають доступ до цих програмно-технічних засобів, виключно в обмежених цілях для виявлення плагіату в текстах робіт.

Робота для перевірки університетом надається в друкованому та електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

04.05.2022

дата

Strelkova

підпис

РІШЕННЯ ЕКСПЕРНОЇ КОМІСІЇ
КАФЕДРИ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ ТА ІНФОМАЦІЙНИХ СИСТЕМ
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Підтверджуємо ознайомлення з результатом звіту подібності щодо роботи, генерованого системою виявлення текстових збігів/ідентичності/схожості:

Назва: Інтелектуальна система прийняття рішення щодо функціонування елементів розумного будинку на основі класифікації видів діяльності людини

Автор: Стрелкова Олександра Георгіївна

Спеціальність: 123 – Комп'ютерна інженерія

Освітня програма: освітньо-наукова

Науковий керівник: Нічепорук Андрій Олександрович, к.т.н, доцент

Після аналізу звіту подібності зроблено такий висновок:

| № | Висновок | Позначка про відповідність |
|---|---|----------------------------|
| 1 | Запозичення, виявлені в роботі, є законними і не є плагіатом. Робота приймається до захисту. | відповідає |
| 2 | Виявлені запозичення не є плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота приймається до захисту, але має бути відкоригована. Відкоригований варіант має бути поданий на кафедру за 2 дні до захисту, разом із заявою щодо самостійності виконання письмової роботи та ідентичності друкованої та електронної версії роботи | |
| 3 | Виявлені запозичення не є плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту (наступного року) після того як буде відкоригована та допрацьована і успішно пройде повторну перевірку на академічний плагіат. | |
| 4 | Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту. | |

Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:

- 1) запозичення розміщені в розділах аналізу існуючих аналогів та прототипів, які не описують безпосередньо авторське дослідження і не стосуються результатів роботи;
- 2) всі зафіксовані системою ознаки модифікації тексту відносяться до комбінування латинських символів зі україномовними скороченнями індексів в формулах, що не є модифікацією тексту.

Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ідентичності/схожості, складає 1.57% і адресується до 127 першоджерела, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

Керівник роботи

Гарант ОП

Завідувач кафедри КПС

А. О. Нічепорук

О. С. Савенко

Т. О. Говорущенко