

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»

Факультет радіоелектроніки, комп'ютерних систем та інфокомунікацій

Кафедра радіоелектронних та біомедичних комп'ютеризованих засобів
і технологій

Пояснювальна записка до кваліфікаційної роботи

магістра

(освітньо-кваліфікаційний рівень)

на тему «Синтез автоматизованого комплексу з визначення патологій легень
за акустичними шумами»

ХАІ.502.560М/2.23о.172.10766637 ПЗ

Виконав: студент 2 курсу групи №560М/2

Галузь знань 17 Електроніка та
телекомунікації

Спеціальність 172 «Телекомунікації та
радіотехніка»

Освітня програма «Радіоелектронні
комп'ютеризовані засоби»

(код і найменування напрямку підготовки)

Цуканов-Капітульський А.Г.

(прізвище й ініціали студента)

Керівник: к.т.н., проф. Бабаков М.Ф.

(прізвище й ініціали)

Рецензент: д.т.н., проф. Невлюдов І.Ш.

(прізвище й ініціали)

Міністерство освіти і науки України
Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет радіоелектроніки, комп'ютерних систем та інфокомунікацій

(повне найменування)

Кафедра радіоелектронних та біомедичних комп'ютеризованих засобів і технологій

(повне найменування)

Рівень вищої освіти другий (магістерський)

Галузь знань 17 Електроніка та телекомунікації

Спеціальність 172 Телекомунікації та радіотехніка

(код та найменування)

Освітня програма Радіоелектронні комп'ютеризовані засоби

(найменування)

ЗАТВЕРДЖУЮ
Завідувачка кафедри

О.В. Висоцька

(підпис)

(ініціали та прізвище)

«10» 10 2023 р.

ЗАВДАННЯ
ДО КВАЛІФІКАЦІЙНОЇ РОБОТИ МАГІСТРА

Цуканов-Капітульський Андрій Григорович

(прізвище, ім'я та по батькові)

Тема роботи Синтез автоматизованого комплексу з визначення патологій легень за акустичними шумами

керівник роботи к.т.н., проф. Бабаков М.Ф.

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом Університету № 1978 від «7» 11 2023 р.

2. Термін подання здобувачем вищої освіти кваліфікаційної роботи:
18.01.2024 р.

3. Вихідні дані до роботи: аудіозаписи акустичних шумів легень різної патології, технічна документація до мікроконтролера Arduino Nano 33 BLE Sense, програмне забезпечення EdgeImpulse.

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

4.1 Аналітичний огляд засобів для діагностики стану легень за акустичними шумами.

4.2 Огляд методу машинного навчання для вирішення задачі класифікації акустичних шумів легень.

4.3 Реалізація моделі машинного навчання для класифікації акустичних шумів легень.

4.4 Апаратна реалізація пристрою з визначення патологій легень за акустичними шумами.

5. Перелік графічного матеріалу:

5.1 Функціональна схема пристрою класифікації звуків легень за акустичними шумами (плакат, арк. А4).

5.2 Результати класифікації шумів легень при використанні навчальних і тестових даних (плакат, арк. А4).

5.3 Апаратна реалізація пристрою для класифікації шумів легень (плакат, арк. А4).

5.4 Алгоритм роботи програми класифікації звуків легень (плакат, арк. А4).

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв
Усі розділи	к.т.н., проф. Бабаков М.Ф.	07.11.2023	

Нормоконтроль В.М. Олійник «18» 01 2024 р.
(підпис) (ініціали та прізвище)

7. Дата видачі завдання «7» 11 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів випускної роботи	Строк виконання етапів роботи	Примітка
1	Аналіз завдання	07.09.13 – 15.11.23	
2	Аналітичний огляд засобів для діагностики стану легень за акустичними шумами	16.11.23 – 20.11.23	
3	Огляд методу машинного навчання для вирішення задачі класифікації акустичних шумів легень	21.11.23 – 29.11.23	
4	Створення наборів даних аудіозаписів шумів різних патологій легень	30.11.23 – 7.12.23	
5	Проведення експерименту з розпізнавання шумів легень за допомогою машинного навчання	8.12.23 – 17.12.23	
6	Апаратна реалізація пристрою з визначення патологій легень за акустичними шумами	18.12.23 – 28.12.23	
7	Оформлення пояснювальної записки	29.12.23 – 18.01.24	
8	Передзахист випускної роботи	23.01.24	
9	Захист випускної роботи	26.01.24	

Здобувач вищої освіти

(підпис)

Цуканов-Капітульський А.Г.

(прізвище та ініціали)

Керівник кваліфікаційної роботи

(підпис)

Бабаков М.Ф.

(прізвище та ініціали)

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи магістра: 147 с., 67 рис., 9 табл., 5 дод., 45 джер.

АУСКУЛЬТАЦІЯ ЛЕГЕНЬ, ЕЛЕКТРОННИЙ СТЕТОСКОП, МАШИННЕ НАВЧАННЯ, ШТУЧНІ НЕЙРОННІ МЕРЕЖІ, МІКРОКОНТРОЛЛЕР.

Об'єктом дослідження є процес визначення патології легень за акустичними шумами.

Предметом дослідження є автоматизований пристрій для визначення патологій легень за акустичними шумами, розроблений на базі апаратної обчислювальної платформи Arduino.

Метою роботи є підвищення якості діагностики органів дихання внаслідок застосування автоматизованого пристрою для електронної аускультатії легень.

Методи дослідження – метод машинного навчання, штучні нейронні мережі, аналіз, синтез, моделювання, проведення експерименту.

У даній роботі проводиться аналітичний огляд науково-технічної літератури щодо існуючих автоматизованих засобів діагностики стану легень за акустичними шумами, розглядаються існуючі цифрові методи аналізу звуків легень.

Створений набір даних акустичних шумів різних патологій легень для подальшого навчання моделі класифікації шумів легень.

Розроблена модель машинного навчання з класифікації патологічних шумів легень. Проведена експериментальна перевірка моделі на навчальних і тестових даних.

Удосконалено засіб для апаратної аускультатії легень для розпізнавання патологічних шумів легень, що дозволяє більш об'єктивно використовувати метод аускультатії у клінічній практиці.

Розроблений прототип пристрою для електронної аускультатії легень на базі мікроконтролера Arduino Nano 33 BLE Sense, який є більш дешевим аналогом іноземних пристроїв. Пристрій може використовуватись лікарями як допоміжний засіб при діагностиці захворювань органів дихання.

ABSTRACT

Explained a note to the master's qualification robot: 147 p., 67 pict., 9 tables, 5 attach., 45 ref.

LUNG AUSCULTATION, ELECTRONIC STETHOSCOPE, MACHINE LEARNING, ARTIFICIAL NEURAL NETWORKS, MICROCONTROLLER.

The object of study is the process of determining lung pathology by acoustic noise.

The subject of the study is an automated device for detecting lung pathologies by acoustic noise, developed on the basis of the Arduino hardware computing platform.

The aim of the study is to improve the quality of respiratory diagnostics by using an automated device for electronic lung auscultation.

Research methods: machine learning, artificial neural networks, analysis, synthesis, modeling, experimentation.

This paper conducts an analytical review of scientific, technical and patent literature on existing automated means of diagnosing the condition of the lungs by acoustic noise, and considers existing digital methods of analyzing lung sounds.

A dataset of acoustic sounds of various lung pathologies was created for further training of the lung sound classification model.

A machine learning model for classifying pathological lung sounds is developed. The model was experimentally tested on training and test data.

A tool for hardware lung auscultation for the recognition of pathological lung sounds was improved, which allows for a more objective use of the auscultation method in clinical practice.

A prototype of a device for electronic lung auscultation based on the Arduino Nano 33 BLE Sense microcontroller, which is a cheaper analog of foreign devices, has been developed. The device can be used by doctors as an aid in the diagnosis of respiratory diseases.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ.....	8
ВСТУП.....	9
1 АНАЛІТИЧНИЙ ОГЛЯД ЗАСОБІВ ДЛЯ ДІАГНОСТИКИ СТАНУ ЛЕГЕНЬ ЗА АКУСТИЧНИМИ ШУМАМИ.....	12
1.1 Сучасні проблеми аускультатії легень.....	12
1.2 Електронні засоби аускультатії та реєстрації звуків легень.....	19
1.3 Існуючі цифрові методи аналізу звуків легень.....	28
1.4 Висновки до розділу 1 та постановка задачі дослідження.....	36
2 ОГЛЯД МЕТОДУ МАШИННОГО НАВЧАННЯ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ КЛАСИФІКАЦІЇ АКУСТИЧНИХ ШУМІВ ЛЕГЕНЬ.....	38
2.1 Застосування машинного навчання і штучного інтелекту як наукового методу	38
2.2 Машинне навчання на вбудованих пристроях	41
2.3 Використання програмного забезпечення Edge Impulse для реалізації машинного навчання на вбудованих пристроях	45
2.4 Отримання ознак аудіоданих	50
2.5 Згорткові нейронні мережі	56
2.6 Висновки до розділу 2	66
3 РЕАЛІЗАЦІЯ МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ КЛАСИФІКАЦІЇ АКУСТИЧНИХ ШУМІВ ЛЕГЕНЬ	68
3.1 Формування наборів даних для навчання моделі	68
3.2 Визначення оптимальних параметрів для навчання моделі в ПЗ Edge Impulse	79
3.3 Реалізація моделі на вбудованому пристрої	87
3.4 Висновки до розділу 3	91
4 АПАРАТНА РЕАЛІЗАЦІЯ ПРИСТРОЮ З ВИЗНАЧЕННЯ ПАТОЛОГІЙ ЛЕГЕНЬ ЗА АКУСТИЧНИМИ ШУМАМИ	93
4.1 Вибір складових компонентів пристрою	93

4.2 Огляд технічних можливостей MCU Arduino Nano 33 BLE Sense..	94
4.3 Характеристики сенсора мікрофону MP34DT05-A	99
4.4 Характеристики дисплею OLED 128×32 I ² C	103
4.5 Апаратна реалізація пристрою для класифікації акустичних шумів легень	107
4.6 Висновки до розділу 4.....	113
ВИСНОВКИ.....	114
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ.....	117
Додаток А «Сертифікати участі в наукових конференціях»	122
Додаток Б «Лістинг аудіозаписів шумів легень за трьома діагнозами і чотирма місцями запису звуків легень».....	125
Додаток В«Схема мікроконтролера Arduino Nano 33 BLE Sense з выводами і сенсорами»	133
Додаток Г«Схема електрична принципова мікроконтролера Arduino Nano 33 BLE Sense»	135
Додаток Д «Програмний код пристрою для класифікації акустичних шумів легень (файл nano_ble33_sense_microphone_continuous.ino)».....	137

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ

AI – Artificial Intelligence (штучний інтелект)

FFT – Fast Fourier Transform (швидке перетворення Фур'є)

MCU – Microcontroller Unit (мікрокомп'ютер, виконаний на окремій материнській платі, що має свої власні процесор, пам'ять та програмовані входи та виходи)

MFCC – Mel-frequency cepstral coefficients (мелчастотні кепстральні коефіцієнти)

ML – Machine Learning (машинне навчання)

PDM – Pulse-density modulation (модуляція щільності імпульсу)

ВСТУП

Актуальність теми. Хвороби органів дихання є глобальною проблемою охорони здоров'я насамперед внаслідок їх значної поширеності серед працездатного населення, постійного прогресування, частого поєднання різних патології легень та обтяжливого впливу на супутні захворювання.

Одним із найбільш поширених неінфекційних захворювань є хронічна обструктивна хвороба легень (ХОЗЛ), яка займає третє місце серед причин смерті у світі [1]. В 2019 році ця хвороба унесла життя понад 3 мільйони людей у працездатному віці. За оцінками Всесвітньої організації охорони здоров'я (ВООЗ), близько 600 мільйонів людей страждають на цю недугу – це близько 11% всього населення планети. За останні 15-20 років відзначено значне зростання захворюваності на ХОЗЛ у різних країнах світу.

Туберкульоз залишається одним із найбільш смертоносних інфекційних захворювань. За даними Глобальної доповіді ВООЗ про боротьбу з туберкульозом 2022 р., у 2021 році на туберкульоз захворіло 10,6 млн осіб, що на 4,5% більше, ніж у 2020 році, а чисельність померлих від нього склала 1,6 млн осіб (включаючи 187 тис. ВІЛ позитивних осіб) [2].

Пневмонія є основною окремою інфекційною причиною смерті серед дітей у всьому світі. У 2019 р. від пневмонії померло 740 180 дітей віком до п'яти років, що становить 14% усіх випадків смерті дітей віком до п'яти років, а серед дітей від одного до п'яти років на пневмонію припадає 22% всіх випадків смерті [3].

Астма є одним із основних неінфекційних захворювань, яким страждають як діти, так і дорослі. Нині на неї страждає понад 339 мільйонів людей. Згідно з оцінками, у 2019 р. на астму страждали 262 мільйони людей і 461 тис. померли від астми [4].

Згідно даних Державної служби статистики України, в Україні щорічно зростає кількість померлих за причиною «хвороби органів дихання». У 2021 році захворювання органів дихання призвели до смерті 26 428 людей (з них

21 176 – грип і пневмонія), в попередньому 2020 році цей показник був у півтора рази менший – 16 479 (з них 11 133 – грип і пневмонія) [5].

Значній частині передчасних смертей можна було б запобігти належним лікуванням, яке, своєю чергою, вимагає своєчасної та надійної діагностики цієї групи захворювань.

Одним з ефективних методів діагностики пневмонії на ранніх стадіях є аускультация легень – вислуховання акустичних шумів з грудної клітини пацієнта за допомогою стетоскопа. Але аускультация вважається суб'єктивним методом дослідження. На сприйманий лікарем звук суттєво впливає стетоскоп, відмінності при цьому спостерігаються не лише між різними типами, але навіть між приладами з однієї партії. Для того, щоб зробити метод аускультации більш об'єктивним, збільшити його точність і довіру до нього, перш за все необхідна точна реєстрація і аналіз дихальних шумів апаратними засобами.

Апаратна аускультация легень на сьогоднішній день є перспективним високоінформативним методом діагностики органів дихання і наразі створюються передумови для поширення апаратної аускультации у клінічній практиці. Тому тема роботи, яка спрямована на розробку автоматизованого пристрою для визначення патологій легень за акустичними шумами є актуальною.

Мета і завдання дослідження. Метою роботи є підвищення якості діагностики органів дихання внаслідок застосування автоматизованого пристрою для електронної аускультации легень.

Для досягнення поставленої мети необхідно вирішення наступних завдань:

- здійснити аналітичний огляд науково-технічної літератури щодо існуючих автоматизованих засобів діагностики стану легень за акустичними шумами;

- провести аналіз технічних можливостей мікроконтролера Arduino Nano 33 BLE Sense для вирішення задачі розпізнавання шумів легень;

- запропонувати прототип пристрою для електронної аускультатії легень;
- створити програмне забезпечення для розпізнавання типів патологій;
- створити набори даних акустичних шумів різних патологій легень для навчання і тестування моделі класифікації шумів легень;
- провести модельний експеримент з розпізнавання типів патологій;
- здійснити аналіз отриманих результатів.

Об'єктом дослідження є процес визначення патології легень за акустичними шумами.

Предметом дослідження є автоматизований пристрій для визначення патологій легень за акустичними шумами, розроблений на базі апаратної обчислювальної платформи Arduino.

Методи дослідження – метод машинного навчання, штучні нейронні мережі, аналіз, синтез, моделювання, проведення експерименту.

Наукова новизна отриманих результатів полягає в тому, що удосконалено засіб для апаратної аускультатії легень для розпізнавання патологічних шумів легень, що дозволяє більш об'єктивно використовувати метод аускультатії у клінічній практиці.

Практична значущість отриманих результатів полягає в тому, що проведено експериментальну перевірку автоматизованого пристрою визначення патологій легень за акустичними шумами. Спроектований пристрій є більш дешевим аналогом іноземних апаратних пристроїв для електронної аускультатії легень без погіршення функції розпізнавання патологічних шумів. Пристрій має компактні розміри і може використовуватись лікарями як допоміжний засіб при діагностиці захворювань органів дихання.

Публікації. За матеріалами магістерської роботи опубліковано тези доповіді на двох наукових конференціях [6, 7]. Підтверджуючі сертифікати щодо участі в конференціях представлені у додатку А.

1 АНАЛІТИЧНИЙ ОГЛЯД ЗАСОБІВ ДЛЯ ДІАГНОСТИКИ СТАНУ ЛЕГЕНЬ ЗА АКУСТИЧНИМИ ШУМАМИ

1.1 Сучасні проблеми аускультатії легень

З найдавніших часів відомі такі звукові явища при аускультатії легень, як патологічні шуми та хрипи, які вислуховуються на відстані. Ще У V ст. до н. е. Гіппократ почав використовувати метод прямої аускультатії – прослуховування шляхом прикладання вуха до грудей пацієнта, щоб почути звуки, що передаються через грудну стінку [8].

Вперше детально і науково основи аускультатії розробив французький лікар Рене Лаеннек, який у 1816 р. винайшов стетоскоп. Він запропонував модель стетоскопа у вигляді трубки зі згорнутого папера. Після багатьох експериментів з різними матеріалами він зупинився на порожній трубці з дерева, діаметром 1,5 і довжиною 12 дюймів (рис. 1.1).

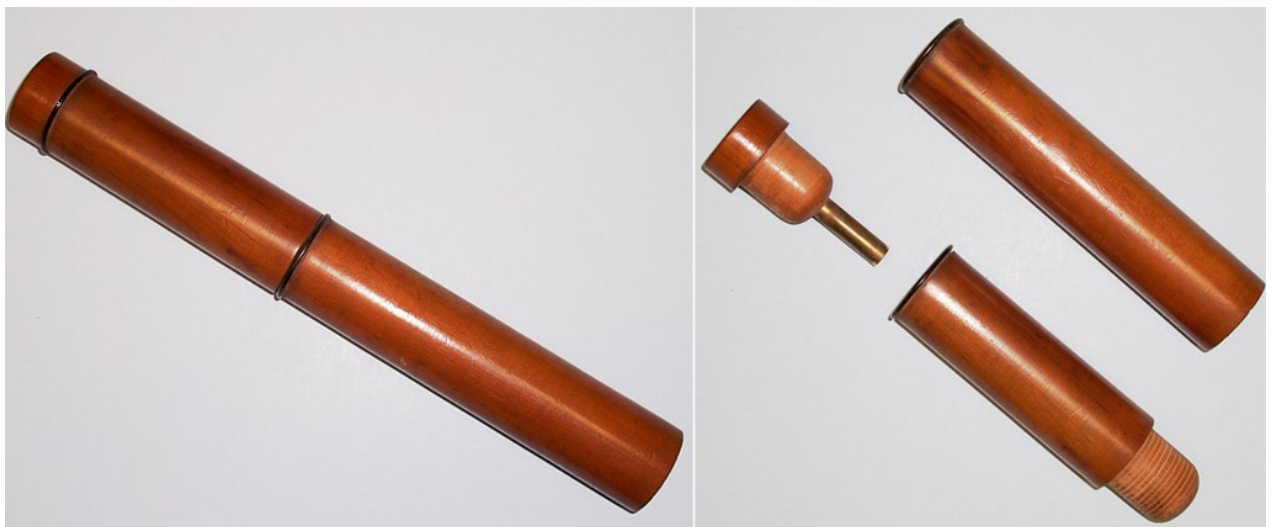


Рисунок 1.1 – Стетоскоп Лаеннека

Історія аускультатії, як науково обгрунтованого клінічного методу дослідження людини, веде свій початок з 1818 р., коли з'явилася робота Лаеннека: «Traite de l'auscultation mediate et des poumons et du coeur» («Про опосередковану аускультатію, або трактат про розпізнавання захворювань легень та серця»).

У роботі Лаеннека були блискуче описані деякі захворювання органів грудної клітки, багато з них – уперше. Серед них – бронхіт, бронхоектази, плеврит, пневмонія, гідроторакс, емфізема, пневмоторакс, набряк легень, гангрена та інфаркт легені, мітральний стеноз, езофагіт, перитоніт та, туберкульоз. Він звернув свою увагу на вивчення «спонтанного» дихання з відкритим ротом, щоб далі розвинути свою техніку «аускультатії дихальних шумів за допомогою циліндра». Здорові суб'єкти створювали везикулярні легеневі звуки, «шум дихання», у той час як пацієнти з патологічними станами виробляли додаткові легеневі звуки «...достатньо чіткі, щоб дозволити ідентифікувати більшість органічних розладів грудної клітки».

За свідченням сучасників, Р. Лаеннек зробив революцію у практичній медицині: «...Стетоскоп вводить лікаря у окремий світ, де симптоми приходять щодо нього відразу з тіла пацієнта» [9]. Відбулась переорієнтація медиків на акустичний метод діагностики. Аускультатія швидко поширилась і зайняла міцне становище у світовій медицині, оскільки була швидким, дешевим, легко здійсненним «польовим» методом дослідження, що застосовується при будь-якому стані та віці пацієнта [10].

У наступні роки стетоскопи модернізувалися у бінауральні («двовушні»). Винайшов їх доктор Меш (Marsh) із Цинциннаті (США) у 1861 р. У його моделі вперше була використана мембрана, що прикладається до грудної клітки (фонендоскоп). Через століття, в 1954 р. у Нью-Йорку Джордж Камман (G. Camman) суттєво покращив конструкцію свого співвітчизника та довів її до промислового зразка. Але навіть сучасний дорогий бінауральний стетоскоп залишається насамперед простим провідником звуку від грудної клітки хворого до вухалікаря. Замість точної передачі всіх звуків більшість сучасних стетоскопів селективно посилюють шуми з частотою нижче 112 Гц (що зручно для кардіологів, яким потрібно почути низькочастотні III і IV тони серця), але одночасно послаблюють високочастотні шуми (що небажано для пульмонологів, яким необхідно вислуховувати високочастотні дихальні шуми) [11].

З введенням у широку практику рентгенологічних та ендоскопічних методів дослідження значення аускультативної легень було понижене [12]. Таке ставлення мало під собою вагомі підстави, пов'язані з суб'єктивізмом методу. Невипадково аускультативна розцінюється як медичне мистецтво [13]. Так, при оцінці наявності додаткових дихальних шумів частота збігів їх оцінки між двома лікарями сягає трохи більше 47%. Особливу важкість викликають такі характеристики, як якість, час виникнення та тривалість хрипів [14]. Крім того, на сприйманий звук суттєво впливає стетоскоп, відмінності при цьому спостерігаються не лише між різними типами, а й навіть між приладами з однієї партії.

Основним способом об'єктивізації сприйняття дихальних шумів, що дозволяє проводити їх точну реєстрацію та аналіз, стало використання систем апаратної аускультативної [12]. Перші спроби створення подібних пристроїв були здійснені вже в першій чверті минулого століття Р. Martini та Н. Mueller. Методи роботи таких апаратів були засновані на автоматичному записі дихальних шумів за допомогою мікрофона. Далі сигнали проходили через фільтр на екран осцилоскопа або виводилися графічним записуючим пристроєм [10].

У 50-х роках ХХ століття швидко розвивалася звукозаписна техніка, що дозволило покращити реєстрацію дихальних шумів, дало можливість проводити їх акустичний аналіз та розширити уявлення про причини виникнення.

Використання з початку 70-х років комп'ютерної техніки для запису дихальних шумів дозволило подолати такі труднощі апаратної аускультативної легень, як взаємодія звукових хвиль та великий обсяг інформації [13]. Вперше дана методика була використана R. Murphy та R. Sorensen у хворих на азбестоз.

У наступні роки більшість робіт, присвячених аналізу дихальних шумів, виконували з використанням швидкого перетворювання Фур'є [15].

Нині методика аналізу дихальних шумів передбачає аналогічні етапи: визначення, запис, графічне представлення сигналу [12]. Записувальні системи складаються з датчика, який уловлює звукові коливання, підсилювача, фільтрів низьких частот та магнітного накопичувача (касета або комп'ютерний диск). Особливе значення надається якості датчиків. Найбільш поширені датчики – конденсаторні, електретні та п'єзоелектричні (акселерометри) мікрофони [15].

У спеціальній літературі йде широка дискусія, у ході якої розглядаються переваги та недоліки різних типів мікрофонів та пропонуються нові моделі, проте відомо, що всі типи цих пристроїв тією чи іншою мірою спотворюють сигнал [13]. Для зменшення завад при реєстрації сенсорами акустичних шумів процесу дихання потрібно використовувати в перехідних муфтах між фонендоскопом (стетоскопом) та мікрофоном акустичних демпферів у вигляді отворів певного діаметру, або гумового патрубку, що штатно використовують в фонендоскопах [16]. Для фіксації мікрофонів використовують різні методи: рукою, ременем, приклеюванням, присосками, загубником. До теперішнього часу ідеального рішення не знайдено, тому що під час дихального циклу постійно змінюються напруга і щільність підлягаючої тканини, що призводить до появи додаткових перешкод.

Перед записом проводиться підсилення та фільтрація дихальних шумів. При підсиленні відбувається збільшення реального сигналу, вольтаж якого на виході малий, до рівня, достатнього для запису. Фільтрація необхідна для відсіву артефактів, що виникають внаслідок роботи м'язів серця, пульсації судин [12]. Надалі електричні імпульси надходять у блок збору даних комп'ютера, потім за допомогою швидкого перетворення Фур'є будуються просторові графіки залежності амплітуди від частоти та часу [10, 17, 18]. В даний час створені рентабельні системи, у тому числі працюючі в реальному часі, які дозволяють проводити підрахунок та аналіз коротких та тривалих додаткових дихальних шумів [19].

Існують різні підходи до вибору місця реєстрації шумів при апаратній аускультатії, що пов'язане з метою дослідження [19]. Зазвичай вибирають або комплекс точок на поверхні грудної клітки, розподілених над проекцією легень, або область трахеї. Перший спосіб має певну перевагу: близьке положення датчика по відношенню до осередку теоретично дозволяє проводити топічну діагностику. Насправді, щоб точно встановити глибину джерела шуму, потрібна його одночасна реєстрація (пеленгування) з кількох точок, що створює додаткові технічні проблеми [10]. Крім того, реєстрація дихальних шумів за даної методики зазвичай займає багато часу.

У 1989 р. G.L. Shechter та ін. впровадили метод трахеофонографії, в якому датчик розташовується на шії – тобто трахеальну аускультатію легень. У 1992 р. Ю.В. Кулаковим та ін. запатентовано виконання трахеофонографії на форсованому видиху [18].

В даний час виконано багато досліджень, присвячених порівнянню трахеальної та легеневої аускультатії, де показано подібність частотної постійної та огинаючої основних дихальних шумів (дані відмінності пояснюють фільтраційною здатністю легень) [20]. Відзначено, що легенева паренхіма і грудна стінка діють як фільтр, послаблюючи звуки із частотою понад 200 Гц [17], і ще сильніше – із частотою понад 1000 Гц [12]. Це підтверджено і внутрішньостравохідною аускультатією.

Деякі дослідники позитивно оцінюють трахеальну аускультатію при оцінці додаткових дихальних шумів, підкреслюючи краще проведення звукових хвиль по стінках легневих шляхів, ніж по повітрю, що міститься в них [10, 12]. Недоліком трахеальної аускультатії була сумація дихальних шумів у короткому проміжку часу та накладення шумів, що виникають при роботі серця та пульсації великих судин. Сучасні методи обробки сигналів дозволяють подолати ці проблеми [9, 10].

Іншим недоліком називають неможливість визначення локалізації джерела звуку, проте це завдання не вирішується і при зніманні сигналу з поверхні грудної стінки через ефекти інтерференції, а також відмінності у

проведенні звуку над різними ділянками грудної клітини [17]. Низка авторів показала можливість визначення при трахеальній аускультатії рівня бронхів, в яких знаходиться джерело звукових коливань, насамперед при ураженні дистальних бронхів 8-9 порядку [18].

Можливість застосування апаратної аускультатії для диференціальної діагностики захворювань легень, при поєднанні різних дихальних шумів, була вперше переконливо показана у 1966 р. при акустичній диференціації екзогенного алергічного альвеоліту, бронхіальної астми та хронічного обструктивного бронхіту у фазі загострення.

Тривалі шуми частіше пов'язують із обструкцією дихальних шляхів [12]. Однак роботи, що аналізують зв'язок їх акустичних характеристик (тривалість та амплітуду) з величинами, що характеризують вентиляторну функцію легень, показали, що результати спірографії є більш точними в оцінці ступеня бронхіальної обструкції [9, 14, 17, 18].

Реєстрація нетривалих додаткових дихальних шумів (потріскувань) свідчить про локалізацію патологічного процесу в ділянці дрібних бронхіол. У ранній стадії патологічного процесу вони реєструються в базальних відділах, при прогресуванні починають вислуховуватись над відділами легень, що лежать вище, і їх кількість збільшується. Відзначено пряму залежність між тяжкістю враження та кількістю потріскувань.

У табл. 1.1 наводиться порівняльна характеристика шумів цього типу при різних захворюваннях залежно від часу виникнення, фази дихального циклу, величини початкового відхилення хвиль та тривалості двох повних комплексів [19].

Використання об'єктивних комп'ютерних методів аналізу здійснило справжній прорив у галузі аускультатії легень. Апаратна аускультатія легень є перспективним високоінформативним методом, що забезпечує об'єктивізацію аускультатії Р. Лаеннека. Зі збільшенням комп'ютеризації лікувально-діагностичних установ є передумови для поширення апаратної аускультатії в клінічній практиці.

Таблиця 1.1 – Порівняльна характеристика дихальних шумів при різних захворюваннях легень

Захворювання чи патологічний процес	Частота дихального шуму	Два повні комплекси та характеристика тембру	Відношення дихального шуму до фаз дихання
Фіброзуєчий альвеоліт	Низькі частоти (до 200 Гц)	<8 мс (середні)	У середині чи наприкінці вдиху
Азбестоз	Високі частоти (понад 500 Гц)	<6 мс (ніжні)	У середині чи наприкінці вдиху
Бронхоектази	Високі частоти (понад 400 Гц)	>9 мс (грубі)	На початку вдиху, видиху
Хронічний обструктивний бронхіт	Низькі частоти	>10 мс (грубі)	У середині вдиху
Серцева недостатність	Низькі частоти	>10 мс (грубі)	Наприкінці вдиху (при вираженій серцевій недостатності протягом дихального циклу)
Пневмонія (стадія приливу)	Високі частоти	6 мс (ніжні)	На початку вдиху
Пневмонія (стадія гепатизації)	Високі частоти	9-10 мс (грубі)	Наприкінці вдиху
Мікоплазмова пневмонія	Середні частоти (до 500 Гц)	4 та 8 мс (грубі)	Наприкінці вдиху

1.2 Електронні засоби аускультатії та реєстрації звуків легень

Аускультатія є одним з основних засобів діагностування захворювань дихання і суть її полягає в вислуховуванні легневих звуків. Для цього використовують стетоскопи – медичні діагностичні прилади для аускультатії звуків. Засобом сприймання звуків внутрішніх органів у стетоскопі є спеціальна головка, яка прикладається до поверхні тіла пацієнтів. Гнучким звукопроводом звук направляєтся в слуховий канал дослідника. Вдосконаленим стетоскопом є фонендоскоп, який має мембрану для того, щоб посилювати звук (рис. 1.2).



Рисунок 1.2 – Сучасний фонендоскоп

Електронний стетоскоп – прилад для аускультатії з посиленням звуку, що складається з мікрофона, підсилювача і динаміка або телефону. Мікрофон перетворює акустичний звук в електронні сигнали, які обробляються і знову перетворюються в звук за допомогою динаміків, вбудованих у вушні раковини (відбувається подвійне перетворення звуку).

Існує велика кількість електронних стетоскопів, що можуть передавати реєстровані звуки безпосередньо на комп'ютер або телефон. Таким чином, електронна аускультация корисна не тільки для діагностики, а й для дистанційного моніторингу пацієнтів та телемедицини.

Електронний стетоскоп 3М Littmann 3200 (рис.1.3) виглядає як звичайний фонендоскоп, але з вбудованим цифровим пристроєм, що дає можливість записувати і зберігати звуки з використанням Bluetooth для передачі даних на комп'ютер. Програмне забезпечення стетоскопу дає можливість візуалізації звуків серця і легень у вигляді фонокардіограми.



Рисунок 1.3 – Стетоскоп 3М Littmann 3200

Вбудовані фільтри стетоскопу призначені для скорочування фонових шумів навколишнього середовища та організму людини на 85%, однак смуга пропускання в частотному діапазоні має досить велику нелінійність [21]. Як заявлено в інструкції до експлуатації стетоскопу, звук підсилюється у 24 рази, що є важливим в умовах навколишнього шуму або конструктивних особливостей пацієнта, наприклад, ожиріння. Серед додаткових опцій пристрою – кристалічний дисплей, що дає можливість відобразити

інформацію про заряд батареї та робочий стан стетоскопу. Є можливість обирати режим роботи стетоскопу на високих частотах, на низьких або в широкій смузі. Крім цього, стетоскоп має можливість відображати серцевий ритм [22].

Наразі стетоскоп 3M Littmann 3200 виведений з продажу, але свого часу ця модель була знаковою для розвитку комерційних електронних стетоскопів.

Розвиток технологічної думки інженерів 3M Littmann еволюціонував у сучасну модель цифрового стетоскопу 3M Littmann CORE Digital Stethoscope. Ця модель складається із звичайного стетоскопа 3M Littmann Cardiology IV з подвійною мембраною, до якої під'єднано цифрову насадку Eko CORE Digital Attachment. Стетоскоп продається у зібраному стані, також цифрову насадку CORE можна придбати окремо (рис. 1.4).



Рисунок 1.4 – Стетоскоп 3M Littmann CORE Digital Stethoscope

У порівнянні із попередньою моделлю 3М Littmann 3200, характеристики нового стетоскопу було значно покращено. Цифровий стетоскоп Littmann CORE може перемикається між аналоговим і посиленим режимами прослуховування, забезпечує підключення через Bluetooth до мобільного пристрою, має двосторонню насадку з двома діафрагмами, що налаштовуються. Електронний стетоскоп Littmann моделі 3200 працює лише у посиленому режимі, він не сумісний із мобільними пристроями та має лише односторонню електронну діафрагму [23].

Насадка Eko CORE Digital Attachment (рис. 1.5) може перетворювати будь-який класичний стетоскоп в цифровий та надає можливість записувати звук. Це невеликий за розміром пристрій, з кріпленням для звукоприймаючої головки і звукопередавальної трубки, з можливістю швидкого перемикання між аналоговим і цифровим режимами. За допомогою Bluetooth сигнали передаються на комп'ютер або смартфон. Підсилення сигналу складає до 40 разів [24].



Рисунок 1.5 – Насадка Eko CORE Digital Attachment

Цифровий стетоскоп Thinklabs One (рис. 1.6) має класичний вигляд фонендоскопа з запатентованою електромагнітною мембраною і надає можливість підключення будь-яких навушників. Посилення сигналу можливе до ста разів. Він також має можливість обрати режим фільтрації для зняття низьких або високих частот та ще декілька видів цифрової фільтрації. Зареєстровані сигнали можна записати на комп'ютер [25].



Рисунок 1.6 – Стетоскоп Thinklabs One

Цей стетоскоп позиціонується виробником як найменший існуючий цифровий стетоскоп і має чималу вартість.

Розглянемо також розробку індійських інженерів, модель AyuSynk 2Pro (рис. 1.7). Цей стетоскоп за устроєм схожий на модель 3M Littmann CORE Digital Stethoscope і складається з голівки фонендоскопа та під'єданого до нього модулю обробки та посилення звукового сигналу до 40 разів. Стетоскоп також здійснює придушення небажаних шумів. За допомогою Bluetooth є можливість передавати аудіосигнал на комп'ютер або смартфон, здійснювати запис звуку і передавати його лікарю для використання у телемедицині [26].



Рисунок 1.7 – Стетоскоп AyuSynk 2Pro

Наразі існуючі іноземні зразки електронних стетоскопів є досить коштовними, тому не кожна клініка може собі дозволити їх використання. В таблиці 1.2 здійснено порівняння основних характеристик розглянутих електронних стетоскопів із зазначенням їхньої вартості. В середньому сучасний електронний стетоскоп буде коштувати від \$230 і дорожче.

Таблиця 1.2 – Порівняльна характеристика електронних стетоскопів

Назва	Основні функції	Вартість
3M Littmann 3200	<ul style="list-style-type: none"> • Запис і зберігання даних. Записи можна прослуховувати на самому фонендоскопі (доріжка на 6 хвилин), або передавати за допомогою Bluetooth на комп'ютер. Є можливість програвання записів на повільній швидкості для детального аналізу. • Аудіозаписи можна обробляти в спеціальному програмному забезпеченні Zargis StethAssist, яке може формувати осциляційну фонокардіограму. 	\$530 (на момент продажу, наразі ця модель знята з продажу)

Продовження таблиці 1.2

Назва	Основні функції	Вартість
3M Littmann 3200	<ul style="list-style-type: none"> • Зниження небажаних фонових шумів за допомогою цифрових фільтрів в цілому на 85%, не приглушаючи важливі звуки органів людини. • 24-х кратне посилення звуку. Регулюється клавішами на інтерфейсі «плюс» і «мінус». • Програма Zargis® Cardioscan™ для ідентифікації хрипів серця (продається окремо). • Автономне живлення (батарея типу AA). 	\$530 (на момент продажу, наразі ця модель знята з продажу)
3M Littmann COREDigital Stethoscope	<ul style="list-style-type: none"> • Складається із стетоскопа 3M Littmann Cardiology IV з подвійною мембраною, до якої під'єднано цифрову насадку Eko CORE Digital Attachment 	\$330
Еко COREDigital Attachment	<ul style="list-style-type: none"> • Цифрова насадка Eko CORE перетворює звичайний стетоскоп на електронний. Сумісна із більшістю популярних моделей аналогових стетоскопів. • Посилення та фільтрація звуку. Посилення до 40 разів. • Активне придушення шумів. • Можливість бездротового прослуховування за допомогою пристроїв Bluetooth. • Можливість запису аудіосигналів на смартфон. • Аналоговий та посилений режими прослуховування. Швидке переключення між режимами. • Виявлення захворювань за допомогою штучного інтелекту за допомогою мобільного додатку Eko (завантаження додатку – безкоштовно). • Автономне живлення (Li-ion акумулятор). Можливість підзаряджання. 	\$230
Thinklabs One	<ul style="list-style-type: none"> • Найменший електронний стетоскоп. • Підсилює аудіосигнал більш ніж у 100 разів. Висока вихідна потужність аудіосигналу. • В комплект входять аудіонавушки Thinklabs Studio Performance. Можна використовувати з будь-якими високоякісними навушками. 	\$500

Закінчення таблиці 1.2

Назва	Основні функції	Вартість
Thinklabs One	<ul style="list-style-type: none"> Програмне забезпечення Thinklink (входить у вартість стетоскопа) забезпечує з'єднання з мобільними пристроями або ноутбуками для захоплення та відправлення звуків через iPhone, iPad, пристрої Android та Mac/ПК. Легке перемикання між прослуховуванням пацієнтів та відтворенням записів. Автономне живлення (Li-ion акумулятор). Можливість підзаряджання. 	\$500
AyuSynk 2Pro	<ul style="list-style-type: none"> Посилення звуку до 40 разів. Передача у реальному часі за допомогою Bluetooth. Шумопридушення у додатку та на пристрої. Фільтр серця та легень у додатку та на пристрої. Генерація звіту в режимі реального часу, звукова візуалізація. Має вбудовану систему штучного інтелекту, яка може надавати візуальні зображення s1, s2, систоли та діастоли, що допомагає об'єктивізувати аускультацию. Автономне живлення (Li-ion акумулятор). Можливість підзаряджання. 	\$265

Проаналізувавши існуючі аналоги, можна зробити висновки щодо загальних вимог до сучасного електронного стетоскопу.

Як правило, електронні стетоскопи використовують головку від звичайного стетоскопу з мембраною як засіб для вислуховування звуків легень. За мембраною йде мікрофон, як правило електретний, який виступає в якості датчика звуку легень. Оскільки встановлено, що шуми легень мають частоти переважно від 100 до 1500 Гц, доцільно використовувати аналогові або цифрові фільтри, які будуть обмежувати аудіосигнал перед його обробкою до цього діапазону частот. В якості засобу для обробки аудіосигналу доцільно використовувати алгоритми вбудованого машинного навчання. Також пристрій має мати дисплей для виведення на нього результатів роботи електронного стетоскопу.

Загальна функціональна схема такого пристрою може бути представлена на рис. 1.8.

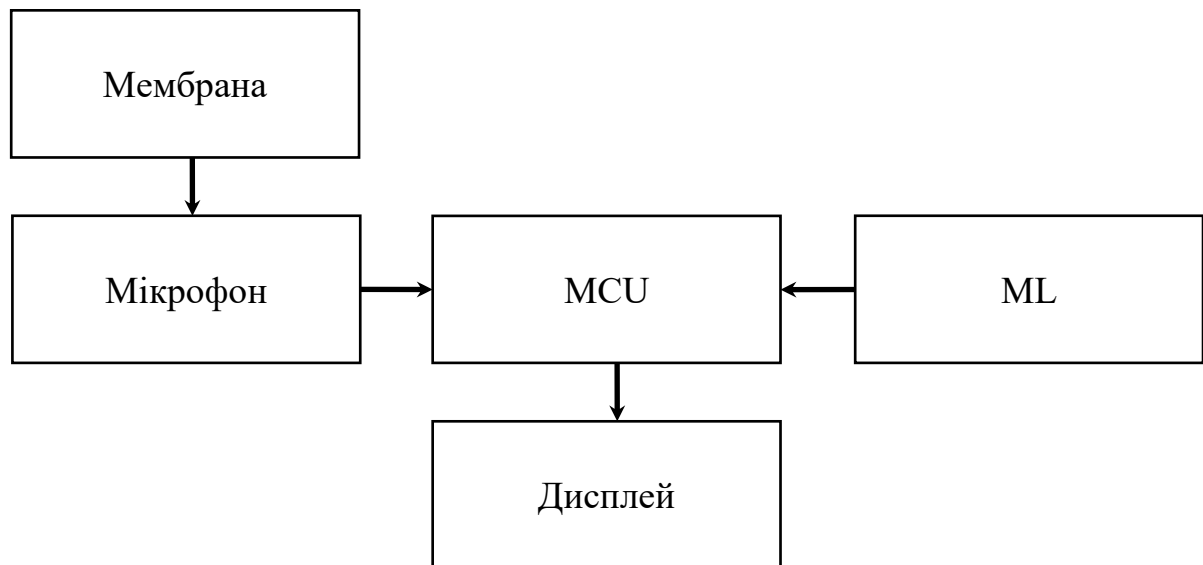


Рисунок 1.8 – Функціональна схема пристрою з класифікації звуків легень за акустичними шумами

Додатковими вимогами до пристрою має бути ергономічність використання: пристрій повинен бути компактних розмірів, мати можливість працювати автономно без підключення до ПК, мати можливість працювати від павербанку. Також пристрій бажано повинен мати нижчу вартість, ніж аналогічні пристрої, при збереженні основного функціоналу.

Одним з найважливіших завдань, яке необхідно розв'язувати під час електронної аускультативної діагностики звуків життєдіяльності людини, є об'єктивізація аускультативних ознак. Використання електронних стетоскопів, цифрової обробки і аналізу сигналів в аспекті діагностики та телемедицини, істотно підвищує ефективність об'єктивізації стану організму людини і скорочує час встановлення діагнозу [27].

1.3 Існуючі цифрові методи аналізу звуків легень

В останні роки стрімко зріс інтерес саме до цифрової обробки і аналізу звуків легень. Вчені з усього світу проводять безліч досліджень з метою виявлення як нових діагностичних ознак, так і задля отримання автоматизованих систем реєстрації, ідентифікації та класифікації звуків дихання.

Дихальні шуми – це шуми, що виникають у процесі дихальних рухів легень при переміщенні у них повітря. Ці шуми, головним чином, є аускультативними феноменами і реєструються при вислуховуванні грудної клітини. До них не відносять чутні на відстані від хворого так зване гучне дихання, стридор, кашель.

Серед дихальних шумів розрізняють основні (нормальні), що мають місце при везикулярному і бронхіальному диханні та вислуховуються у нормальному стані здорових людей, а також додаткові (паталогічні). До останніх належать крепітація, хрипи, шум тертя плеври. Зміна основних дихальних шумів за їх інтенсивністю (наприклад, ослаблене дихання), місця вислуховування, тембру (жорстке чи амфоричне дихання), безперервності (саккадоване дихання) свідчать про відхилення від норми і мають діагностичне значення, так само як і поява додаткових шумів [28].

Везикулярне дихання — шум розтягування легеневих альвеол при заповненні їх повітрям. Він вислуховується у здорових дорослих людей практично над усіма ділянками легень, але найкраще в підключичних і підлопаткових областях. Це тихий, однорідний шум руху повітря (нагадує звук «ф»), добре чутний протягом усього вдиху, але зникає або різко слабшає на самому початку видиху. Якщо шум на вдиху і видиху однаковий по гучності і тривалості, то його визначають як жорстке дихання, яке частіше вислуховується при бронхіті.

Бронхіальне дихання відрізняється від везикулярного і жорсткого дихання тим, що шум у фазі видиху триваліший, ніж у фазі вдиху, гучніший і

має специфічний тембр: нагадує шум видиху через рот при положенні мови для вимовляння звуку «хі». У нормі вислуховується над проекцією трахеї і великих бронхів. Над периферійними ділянками легких бронхіальна подих вислуховується тільки при патології: у випадках значного ущільнення легеневої тканини.

Ослаблене дихання характеризується значним зниженням гучності нормальних дихальних шумів. Може визначатися при потовщеній грудній стінці у огрядних людей, а також при поверхневому диханні [28].

До патологічних дихальних шумів належать хрипи — додаткові шуми, які виникають у трахеї та бронхах, що можуть бути вологими та сухими відповідно до механізмів утворення, що відбивається на звуковому сприйнятті. За розташуванням над однією обмеженою ділянкою, над кількома такими ділянками або більш широко над великими ділянками грудної клітини, хрипи можуть класифікуватись як локальні, розсіяні чи поширені [29]. Говорять також про поодинокі, множинні та рясні хрипи.

Вологі хрипи є звуками, що утворюються при русі повітря через бронхи чи поєднані із ними порожнини унаслідок розриву бульбашок повітря, що проходять через рідке мокротиння. Практика свідчить, що такі хрипи вислуховуються головним чином на вдиху, ніж на видиху. Як було відзначено, вологі хрипи поділяються на дрібно-, середньо- і крупно пухирцеві. Дослідження показують, що поширені крупно-пухирцеві хрипи відповідають пізній фазі розвитку набряку легень на тлі рясних середньо- і дрібно пухирцевих хрипів. За звуковим ефектом вологі хрипи розмежовують на звучні та незвучні. Перші пов'язуються із ущільненням легеневої тканини при пневмонії. Другі свідчать про наявність мокротиння у бронхах, однак вони не засвідчують ущільнення навколишньої легеневої тканини, а характерні для бронхіту [29].

Сухі хрипи у бронхах являють собою звуки з різним тембром, схожі на дзижчання або свистячі. Утворення перших пов'язують із звучанням повітря при його проходженні через «ниткоподібні перемички» з мокротиння при

запаленні бронхів. Другі виникають унаслідок спазму дрібних бронхів і набряку слизової, вважаються характерними для нападів бронхіальної астми.

Крепітація («crepitare» — скрипіти, хрустіти) — побічний дихальний шум, який утворюється при розправленні зволжених більше, ніж зазвичай стінок альвеол, що втратили еластичність. Цей шум є характерним для максимуму вдиху. Він лунає як короткий звуковий ефект «спалаху», «вибуху» і схожий на той, що може мати місце при розминці пальцями пучка волосся біля вуха. Крепітацію інколи важко відрізнити від дрібнопухирцевих вологих хрипів, однак її можна почути лише наприкінці вдиху і після кашлю вона не змінюється. Крепітацію часто вважають проявом крупозної пневмонії, що супроводжує фази появи і розсмоктування ексудату. Інколи цей дихальний шум виникає на початку набряку легень.

Шум тертя плеври має місце при захворюванні на сухий плеврит. Він з'являється при проходженні повітря, якщо поверхня плеври стає нерівною і шорсткою у результаті покриття фібрином. У цьому випадку при диханні може виникати звук, що схожий на скрип шкіри або снігу. Інколи цей шум може бути схожим на крепітацію або хрипи. Стосовно цього, у публікаціях фахівців вказується, що шум тертя плеври вислуховується у двох фазах дихання, а при натисканні на грудну клітину стетоскопом спостерігається його посилення. Вони також зазначають, що шум тертя плеври має місце й при імітації дихання із закритими носом і ротом [28].

Очевидно, що характер дихальних шумів складний і різноманітний. Тому не дивно, що аускультативна семіотика сповнена такими епітетами, як жорстке дихання, хрипи гудячі, тріскучі, свистячі, крупно пухирцеві, хрипи і т.п. За такої класифікації важко безпосередньо поставити у відповідність кожному з наведених описів будь-які об'єктивні інформаційно значущі кількісні характеристики [30].

Розмах частот шумів настільки великий, що в один діапазон потрапляють звуки з різних категорій, що робить частотний аналіз неоднозначним.

В табл. 1.3 показані частотні діапазони, де найбільш себе проявляють шуми дихання.

Таблиця 1.3 – Частотні діапазони шумів дихання

Тип дихання	Частота, Гц													
	160	250	350	400	450	500	550	600	650	700	750	800	900	1500
Бронхіальне	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Везикулярне	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Крепітація	□	□	■	■	■	■	■	■	■	■	■	■	■	■
Тертя плеври	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Сухі хрипи	□	■	■	■	■	■	■	■	■	■	■	■	■	■
Вологі хрипи	■	■	■	■	■	■	■	■	■	■	■	■	■	□

Як видно з таблиці, бронхіальне дихання та сухі хрипи мають широкі частотні діапазони, які значно перекриваються. Всі інші дихальні шуми так чи інакше знаходяться в межах цих широких діапазонів, тому можна зробити висновок, що ідентифікація дихальних шумів лише в частотній області не є ефективною. Доцільні подальші дослідження у цій області.

Основними методами аналізу звуків легень є:

- аналіз сигналів у частотній області;
- аналіз сигналів у частотно-часовій області;
- інші види аналізу сигналів (кепстральний аналіз та інші).

Перетворення Фур'є є одним з найпоширеніших методів аналізу сигналів у частотній області. При аналізі звуків легень перетворення Фур'є є корисним у виявленні присутності деяких хрипів [31].

Перетворення Фур'є функції $f(t)$ математично визначається як комплекснозначна функція $F(\omega)$, яка задається інтегралом:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt. \quad (1.1)$$

На рис. 1.9 наведено сигнали звуків нормального дихання та дихання з сухими хрипами та їх відповідні спектри.

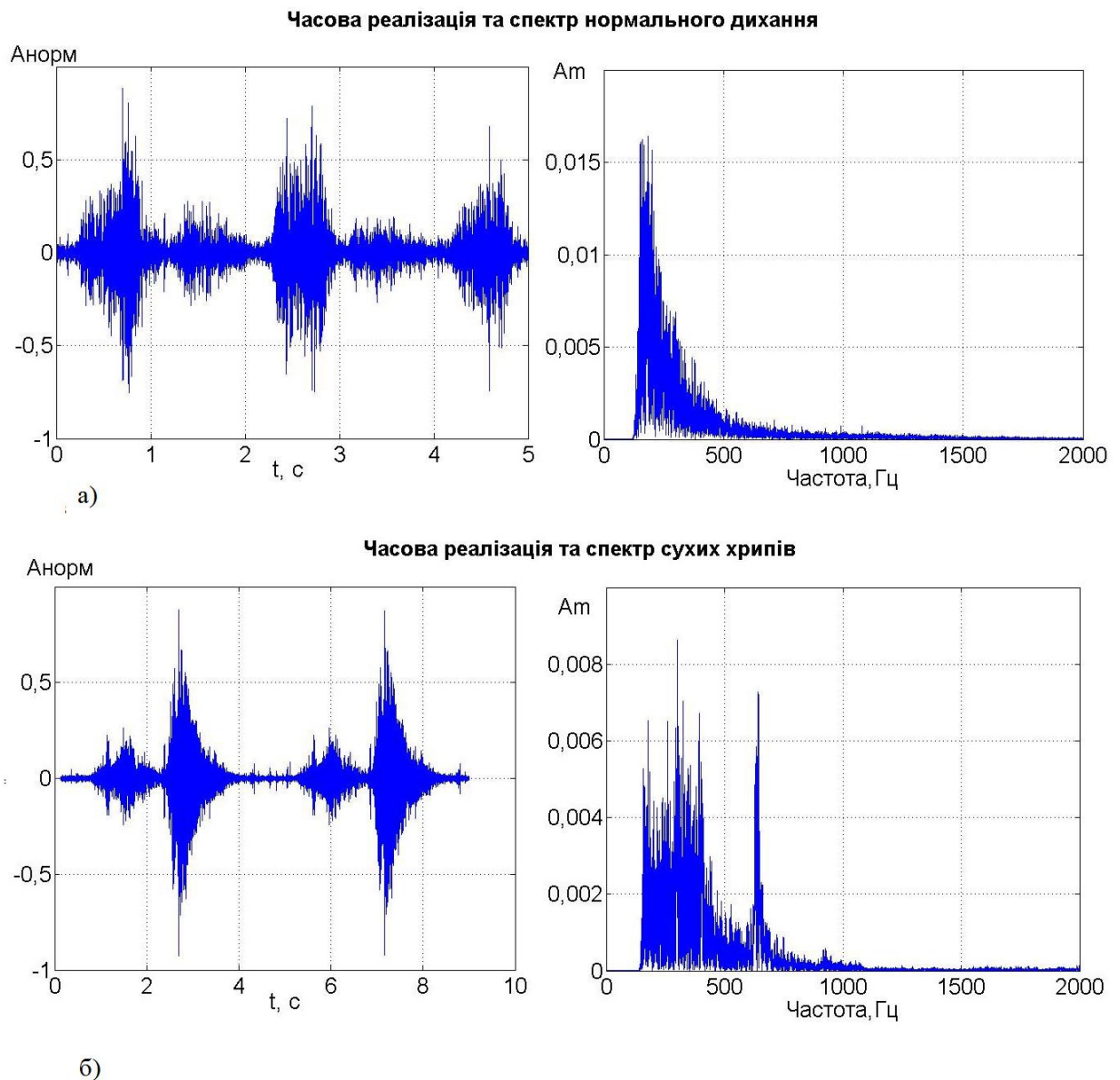


Рисунок. 1.9—Сигнали звуків дихання з хрипами (а) та їх амплітудні спектри (б)

Багато методів, особливо в більш ранніх дослідженнях, засновані тим чи іншим чином на спектральному аналізі [29].

Спектральний аналіз має суттєву перевагу через простоту розрахунків. Але, як вже зазначалося, нормальні шуми дихання (бронхіальне та везикулярне) займають широкий діапазон частот. Це викликає серйозні

труднощі в частотному аналізі звуків дихання. Як правило, це призводить до труднощів у розрізненні додаткових шумів на фоні основних через перекриття частотних діапазонів та малої різниці в амплітудах цих видів шумів. Досить проблематично розрізняти звуки шунтування дихання на тлі основного дихання через перекриття частотних діапазонів і невелику різницю в амплітудах.

Тобто, серед недоліків такого аналізу – недостатня інформативність аналізу нестационарних сигналів та їх особливостей, а також відсутність інформації про локальні властивості сигналу в умовах швидкої зміни в часі його спектрального складу.

З огляду на нестационарний характер звуку легень, все більш поширеним вибором для аналізу легневих звуків стає аналіз звуків дихання в частотно-часовій області. Одним з найширше використовуваних методів є короткочасне перетворення Фур'є (Short-Time Fourier Transform, STFT).

STFT – це Фур'є-перетворення, яке виконується на одному сегменті даних і формулюється відповідно до:

$$X(t, f) = \int_{-\infty}^{+\infty} x(t)w(t - \tau)e^{-j2\pi m/\tau} d\tau, \quad (1.2)$$

де $w(t - \tau)$ – функція вікна.

На основі спектрально-часового аналізу отримують різні числові параметри [32].

Переваги методу STFT полягають у тому, що він обчислювально простий і легкий у стеженні за частотою сигналу в певному часовому вікні. До недоліків цього методу відносяться досить низька роздільна здатність і невизначеність часу, до якого відноситься частота, тому що частоти обчислюються з певними інтервалами.

Інший підхід полягає в аналізі та обробці зображення, а саме спектрограми, що останнім часом використовується досить широко [33].

За зовнішнім виглядом спектрограми у деяких дослідженнях пропонується визначати наявність тих чи інших хрипів у звуках легень (рис. 1.10).

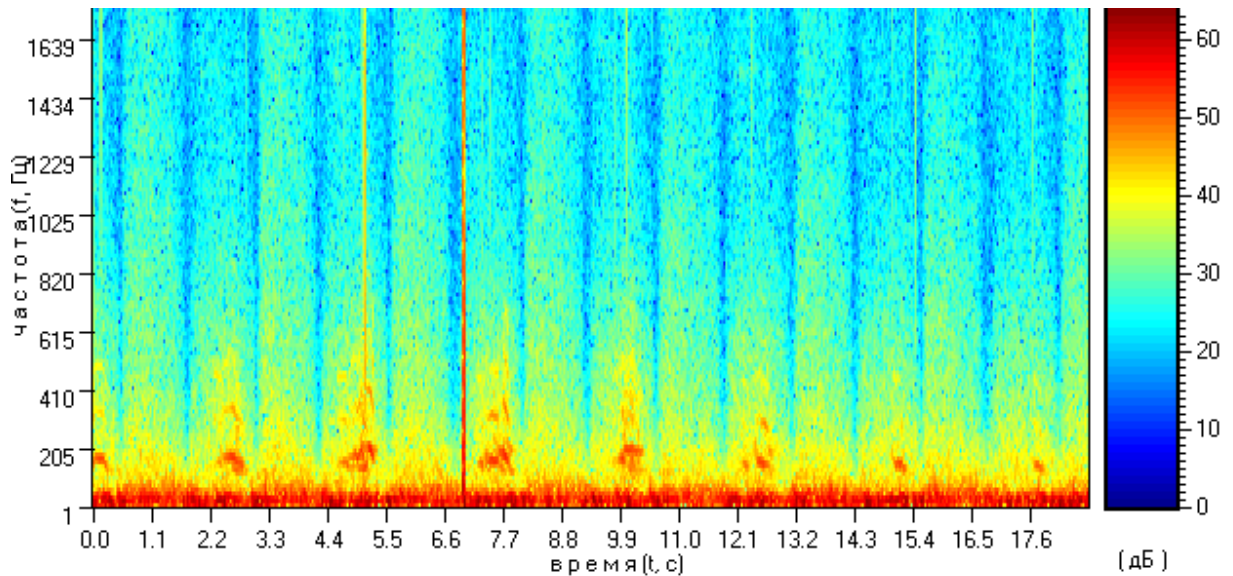


Рисунок. 1.10 – Спектрограма (респіросонограма) дихання хворого на ХОЗЛ

Спектрограми дозволяють вивчати динаміку зміни спектрального складу шумів в часі. При цьому наявність кріпотури або свистів проявляється в спектрограмах. Якщо тривалість різних фаз дихання для здорових легень становить одиниці секунд, то при патологіях тривалість може зменшуватися до часток секунди. Тому для аналізу динаміки зміни поточного спектра акустичних шумів дихання оцінки поточних спектрів потрібно здійснювати за сегментами тривалістю близько 0,1 с. Візуальні образи спектрограм здорової легені і з патологією суттєво відрізняються, що може бути покладено в основу діагностики [34].

Окрім вищезазначених методів в останній час використовується дослідження на основі кепстрального аналізу. Під кепстром розуміють спектр логарифма спектру початкового сигналу $S(\omega)$:

$$Ks(q) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \ln|S(\omega)|^2 e^{j\omega q} d\omega. \quad (1.3)$$

Між частотними складовими в традиційному спектрі та частотними складовими в кепстрі є принципові відмінності, тому складність полягає в правильній інтерпретації результатів. Необхідно мати певний досвід та підготовку для того, щоб зробити правильні висновки за результатами кепстрального аналізу. Але проте роботи в цьому напрямку ведуться і, напевно, в майбутньому будуть отримані корисні інформативні характеристики сигналів звуків дихання [35].

Дихальні звуки містять багато інформації, пов'язаної з будовою дихальних шляхів. Наприклад, звуження дихальних шляхів, порожнини в легенях, набряки та наявність сторонніх тіл можуть змінити структуру дихальних шляхів або еластичність середовища передачі звуку, що, у свою чергу, створює ненормальні дихальні звуки або змінює характеристики передачі звуку дихальна система.

Традиційні стетоскопи мають низьку роздільну здатність і для якісного аналізу покладаються лише на досвід лікаря. Відповідно до контуру рівної гучності, чутливість людського вуха до звуку є сукупним ефектом інтенсивності звуку та частоти, і деякі низькочастотні звуки низької інтенсивності не можуть бути сприйняті людським вухом. Штучно розрізнити зміни інтенсивності та частоти дихальних звуків, спричинені хворобою, важко, тому деякі патологічні ознаки важко вловити та підтвердити. У поєднанні з сучасною цифровою технологією обробки сигналів кількісне порівняння дихальних шумів у часовій області та частотній області може точно проаналізувати механізм генерації дихальних шумів, вивчити характеристики нормальних/ненормальних дихальних шумів і надалі розробити методи автоматичної класифікації дихальних шумів. хвороби. Це закладає теоретичну основу для портативної та зручної системи сімейного лікаря [36].

1.4 Висновки до розділу 1 та постановка задачі дослідження

1. Захворювання легень є одними з найпоширеніших захворювань у світі, що можуть призводити до серйозних ускладнень та до смерті. Своєчасне діагностування хвороби є однією з важливих умов вибору лікарем адекватної стратегії лікування. Аускультация, в тому числі електронна, є одним з головних неінвазивних методів дослідження пацієнта.

2. Звуки легень поділяються на нормальні та додаткові (патологічні). Важливою задачею є ідентифікація та диференціація дихальних шумів за певними категоріями, оскільки за слуховим сприйняттям ці звуки можуть бути дуже подібними, а за своєю природою свідчать про різні захворювання. Актуальною задачею є об'єктивізація дихальних ознак на основі цифрової обробки та аналізу сигналів.

3. Проведено огляд та аналіз сучасних засобів реєстрації звуків дихання. Одноканальні стетоскопи імпортного виробництва з ліцензійним програмним забезпеченням є доволі коштовними. В середньому вартість сучасного електронного стетоскоп складає від \$230 і дорожче. Актуальною залишається задача створення доступного за коштами портативного пристрою. Також було узагальнено функціональну схему пристрою з класифікації звуків легень за акустичними шумами.

4. Встановлено, що більшість сучасних методів цифрового аналізу звуків легень суттєво залежить від якості реєструючих засобів та умов реєстрації, тобто не є універсальними. Патологічні дихальні звуки часто мають схожий спектральний склад. Актуальним завданням є розробка методу, який, по-перше, мало залежить від умов реєстрації сигналу, а по-друге, є універсальним з точки зору широкого спектру патологій легень.

Проаналізувавши сучасний стан проблеми та розглянувши існуючі засоби для аускультации легень був зроблений висновок щодо актуальності обраної теми роботи і доцільності розробки автоматизованого пристрою з визначення патологій легень за акустичними шумами.

Відповідно до поставленої мети в роботі необхідно вирішити наступні завдання:

- провести аналіз технічних можливостей мікроконтролера Arduino Nano 33 BLE Sense для вирішення задачі розпізнавання аудіосигналів;
- запропонувати прототип пристрою для електронної аускультатії легень;
- створити програмне забезпечення для розпізнавання типів патологій;
- створити базу даних акустичних шумів різних патологій легень для подальшого використання в якості еталонів;
- провести модельний експеримент з розпізнавання типів патологій;
- здійснити аналіз отриманих результатів.

2 ОГЛЯД МЕТОДУ МАШИННОГО НАВЧАННЯ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ КЛАСИФІКАЦІЇ АКУСТИЧНИХ ШУМІВ ЛЕГЕНЬ

2.1 Особливості застосування машинного навчання і штучного інтелекту як наукового методу

Штучний інтелект (AI – Artificial Intelligence) – це здатність машин, систем, моделей, комп'ютерів мати можливість імітувати та покращувати аспекти інтелекту людини чи інтелекту в цілому. AI сьогодні дійсно добре вирішує конкретні проблеми в чітко визначених контекстах.

Розглянемо ключові терміни, які стосуються простору штучного інтелекту (рис. 2.1).

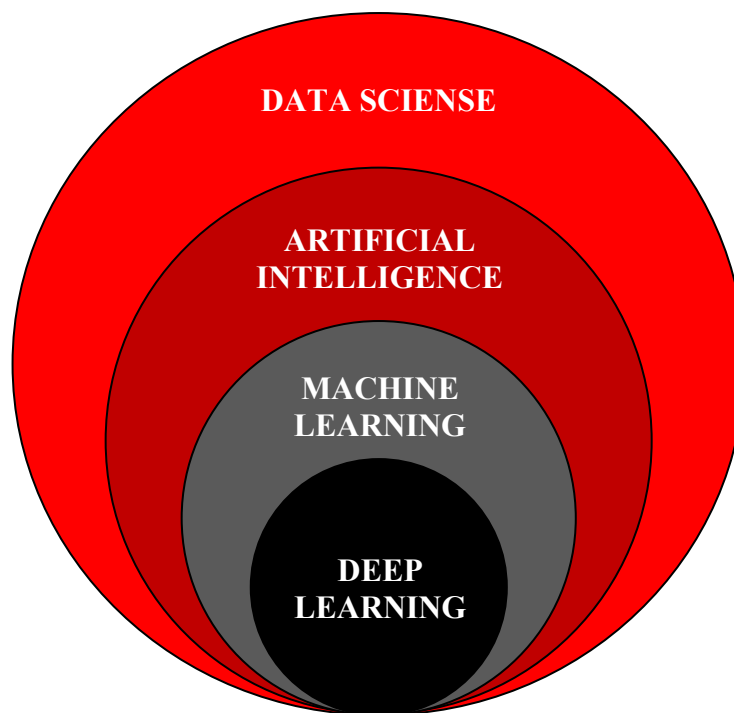


Рисунок 2.1 – Ключові терміни AI

Найширший термін – це наука про дані (Data Science). Науковець з даних – це кожен, хто використовує наукові методи та алгоритми для того, щоб приймати кращі рішення чи прогнози.

Підмножиною науки про дані є штучний інтелект (Artificial Intelligence). Визначення штучному інтелекту ми дали раніше.

Підмножиною штучного інтелекту є машинне навчання (ML – Machine Learning). Алгоритми машинного навчання – це сукупність математичних та статистичних моделей, які вивчають уявлення з базових навчальних даних. Алгоритми машинного навчання отримують шаблони з вхідних даних, щоб визначити правила і параметри цих моделей, щоб вони могли приймати розумні прогнози та рішення.

Підмножиною машинного навчання є глибоке навчання (Deep Learning). Моделі глибокого навчання – це в основному дуже складні моделі машинного навчання. Вони вивчають уявлення послідовно шарами. Глибоке навчання – це також сфера штучного інтелекту, яка досягла значних успіхів за останні роки з початку 2010 року. Наразі ми перебуваємо в розпалі хвилі ажіотажу зі штучним інтелектом, яку називають революцією глибокого навчання.

На рис. 2.2 зображено основні категорії алгоритмів і моделей машинного навчання.

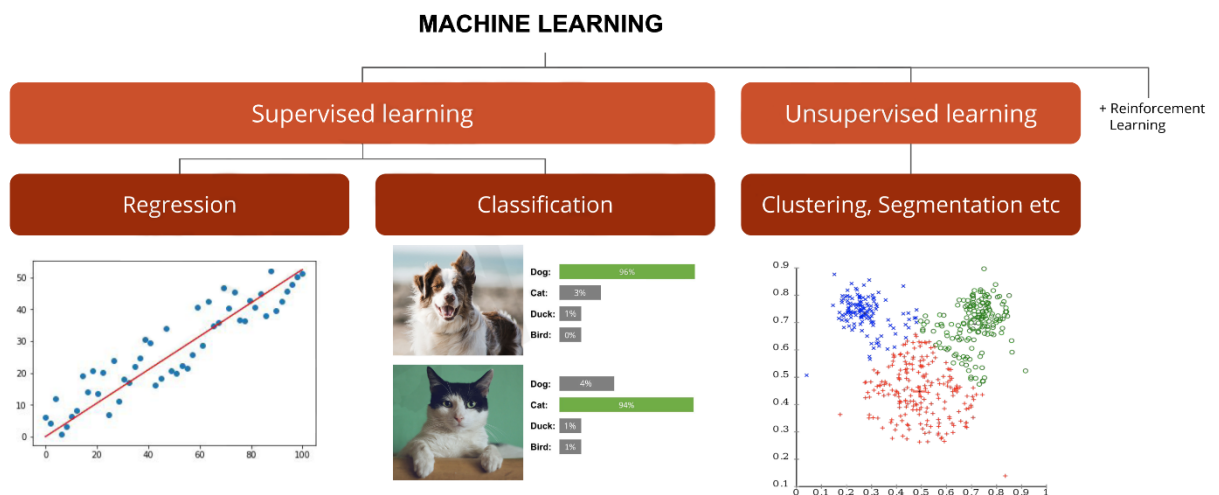


Рисунок 2.2 – Категорії алгоритмів і моделей машинного навчання

Контрольоване машинне навчання – це спроба знайти функцію, яка може зіставляти деякі вхідні дані з деякими вихідними даними, які можуть

бути передбаченням або класифікацією. У контрольованому машинному навчанні потрібно надати дані під час навчання цих алгоритмів машинного навчання, які є правильними парами введення та виводу.

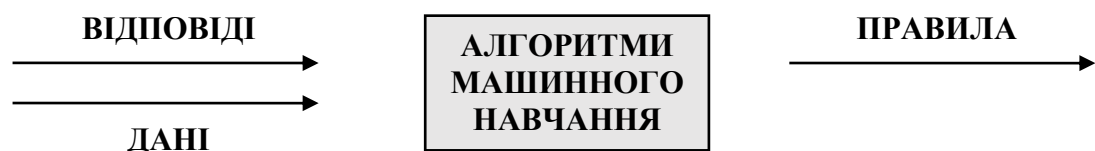
Контрольовані алгоритми машинного навчання можуть бути регресійними моделями. Вони передбачають безперервну змінну результату, або це можуть бути алгоритми та моделі класифікації, що передбачають або класифікують певний набір категорій або міток.

Також існує неконтрольоване машинне навчання. Для моделей машинного навчання без контролю не надаються правильні мітки або вихідні дані у навчальних даних. Натомість намагаються отримати та проаналізувати шаблони, щоб наприклад, могли здійснити кластеризацію, зменшення розмірності, виявлення викидів, сегментацію тощо.

Підводячи підсумок, відзначимо відмінність традиційних алгоритмів від алгоритмів машинного навчання (рис. 2.3).



а) Традиційні алгоритми: спираються на правила, визначені людьми



б) Алгоритми машинного навчання:

- на першому етапі отримують правила з даних (етап навчання);
- на другому етапі використовують отримані правила для здійснення прогнозів (етап висновків)

Рисунок 2.3 – Традиційні алгоритми і алгоритми машинного навчання

Для традиційних алгоритмів, наприклад, експертної системи, параметри та правила алгоритму визначаються людиною. Потім дані обробляються відповідно до цих правил, щоб отримати результати, відповіді та висновки з цих моделей.

При машинному навчанні значення параметрів та правила алгоритмів машинного навчання визначаються в процесі навчання. Для того, щоб навчити контрольований алгоритм машинного навчання, потрібно надати йому відповіді разом із вхідними даними, а на виході з алгоритму ми отримуємо правила (етап навчання). Далі на другому етапі ми можемо використати ці правила та цю навчену модель машинного навчання для того, щоб отримати прогнози (етап висновків).

2.2 Машинне навчання на вбудованих пристроях

Останні кілька років було проведено багато досліджень, спрямованих на підвищення ефективності алгоритмів машинного навчання. Стало можливим запускати їх на невеликих і більш енергоефективних пристроях.

Вбудована система – це будь-яка комп'ютерна система, вбудована в іншу більшу механічну або електричну систему. Одноплатні комп'ютери, такі як Raspberry Pi, можна вважати вбудованими системами, якщо їх вживити в будь-який інший пристрій, наприклад робота. Так само мікроконтролери (MCU), або плати розробника, зазвичай вважаються вбудованими системами і часто використовуються для управління фізичними або електричними пристроями практично без участі людини (рис. 2.4).

Одноплатні комп'ютери зазвичай оснащені більш потужним мікропроцесором з окремою пам'яттю і можуть працювати під управлінням повноцінної операційної системи, наприклад Linux, можуть надавати повноцінний інтерфейс користувача, будь то командний рядок або графічний інтерфейс. Частота процесора зазвичай становить сотні мегагерц чи гігагерц. Обсяг оперативної пам'яті обчислюється сотнями мегабайт чи гігабайт.

Мікроконтролер зазвичай набагато дешевший, але набагато менш потужний, ніж одноплатний комп'ютер. Він програмується за допомогою суперциклу (superloop) або, при необхідності, можливе використання легкої операційної системи, наприклад, ОС реального часу. Більшість мікроконтролерів не мають інтерфейсу користувача або, якщо він є, він складається з декількох кнопок і простого РК-екрана. Крім того, для їхньої роботи зазвичай потрібно набагато менше енергії, ніж для одноплатних комп'ютерів.

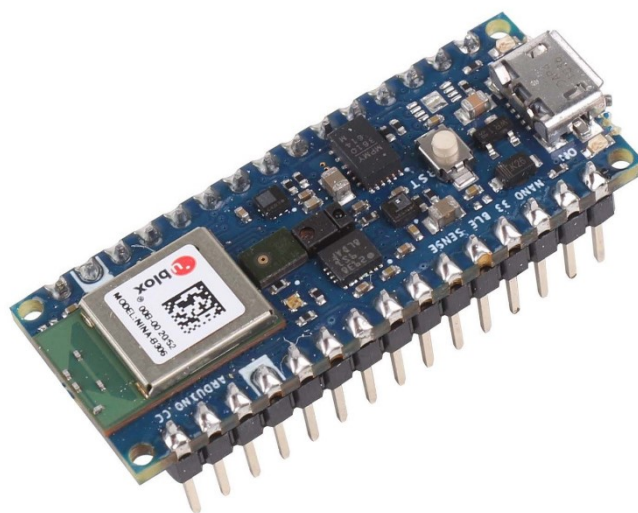
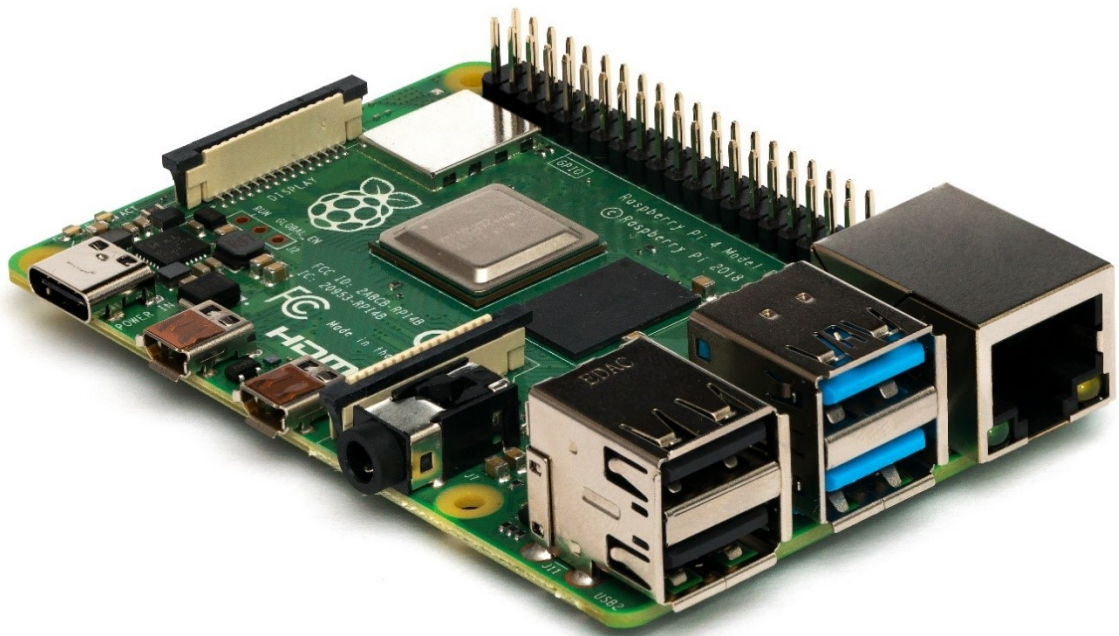


Рисунок 2.4 – Зовнішній вигляд одноплатного комп'ютеру Raspberry Pi та плати розробки Arduino Nano 33 BLE Sense

Більшість сучасних фреймворків машинного навчання були розроблені для настільних комп'ютерів і серверів, що працюють з використанням мов програмування високого рівня, таких як Python, тому часто їх простіше запускати на одноплатному комп'ютері, ніж на мікроконтролері, за умови, що одноплатний комп'ютер має достатню потужність для своєчасного виконання поставленого завдання. Можна запустити на Raspberry Pi такі програми, як scikit-learn, TensorFlow, PyTorch та Caffe. Однак донедавна було зроблено дуже мало для запуску цих фреймворків на мікроконтролерах.

Кілька років тому Google TensorFlow випустила урізану версію TensorFlow Lite, спеціально розроблену для мікроконтролерів. Це дозволяє запускати прості нейронні мережі на мікроконтролерах на високому рівні без необхідності програмувати вручну операції з матрицями. Крім того, виробники мікроконтролерів провели безліч оптимізацій коду та мікроконтролерів, щоб ці фреймворки та інші алгоритми машинного навчання працювали на цих крихітних контролерах набагато швидше.

Дамо визначення tinyML і tinyML-додатків, а також розглянемо, яке апаратне забезпечення є доступним і придатним для таких робочих навантажень. Найпоширеніше визначення, яке використовується для tinyML—це додатки машинного навчання на пристроях потужністю в один міліват і нижче. Зазвичай це мікроконтролерні системи, що живляться від невеликих батарейок і, можливо, використовують технології збору енергії [37].

Найпоширеніші додатки tinyML можна умовно розділити на три основні області у порядку зростання обчислень: це вібрація, голос і зір. До вібрації відносяться системи, які керують двигуном і виконують прогнозне обслуговування, локалізоване для конкретного двигуна. До голосових відносяться системи з одним або декількома мікрофонами, які використовуються для голосового визначення ключових слів і розпізнавання мови. Зір включає в себе промислові системи розпізнавання об'єктів для сортування товарів, виявлення дефектів, або системи розпізнавання облич для розблокування пристроїв.

Всі ці випадки використання мають різні вимоги до продуктивності робочого навантаження і вимагають масштабованих рішень. Це вимагає обробки сигналів, щоб зробити вхідні дані придатними для використання, і обробки машинного навчання, щоб забезпечити більш корисний інтелект і дії, роблячи обчислення на периферії реальністю. Сьогодні компанія Arm надає можливість обробки сигналів у ядрах процесорів Cortex-M4, доступну через бібліотеки CMSIS з відкритим кодом.

Існують практичні причини, чому варто використовувати ML на мікроконтролерах, зокрема:

- функціональність – «розумний пристрій» забезпечує високу швидкість виконання операцій на локальному рівні (незалежно від Інтернету);

- вартість – забезпечується простотою використання та не високою вартістю обладнання;

- конфіденційність – забезпечує захист даних і не дозволяє загального доступу ззовні до даних;

- ефективність – проявляється у меншому форм-факторі пристрою, економії енергії та довшому терміну служби акумулятора.

Машинне навчання за своєю природою є імовірнісним. Воно включає в себе розробку математичних моделей, що загалом відображають динаміку даних, що використовуються для навчання. Навіть 99-відсотковий показник ефективності буде недостатнім для того, щоб медичний пристрій з розпізнавання патологічних шумів легень був сертифікований для використання в закладах охорони здоров'я. Але вбудоване машинне навчання може доповнювати інші детерміновані функції або запропонувати унікальні рішення проблем.

Щоб створити проект машинного навчання ми будемо використовувати інструмент під назвою Edge Impulse, оскільки він значно полегшує процес навчання моделі. Він також дає графічне уявлення про те, як відбувається машинне навчання.

2.3 Використання програмного забезпечення Edge Impulse для реалізації машинного навчання на вбудованих пристроях

Edge Impulse була заснована Заком Шелбі та Яном Джонгбумом у 2019 році. Ідея полягала в тому, щоб створити онлайн-інструмент, який допоможе полегшити машинне навчання для застосуванні на вбудованих пристроях. Замість того, щоб вивчати один з фреймворків програмування для машинного навчання, можна використати цей інструмент, щоб виконати всю важку роботу щодо навчання моделі і зосередити увагу на проблемі, яку необхідно вирішити. Edge Impulse має власну наукову статтю, яка була написана у співпраці з деякими дослідниками з Harvard SEAS [38]. У ній детально описано багато проблем у вбудованому машинному навчанні і те, як Edge Impulse вирішує ці проблеми.

Для початку роботи з Edge Impulse необхідно створити обліковий запис на сайті <https://edgeimpulse.com>. Далі створити новий проект і відкрити його. Ліва панель показує кроки, необхідні для збору даних, створення конвеєра машинного навчання, який називається імпульсом, навчання моделі та розгортання її на вбудованому пристрої (рис. 2.5).

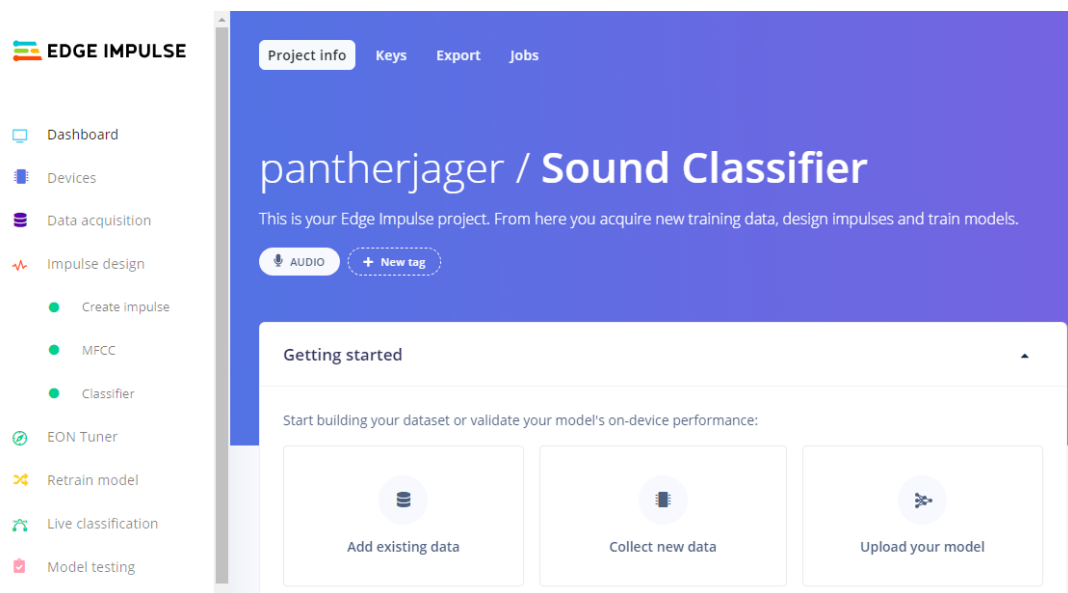


Рисунок 2.5 – Створення проекту з розпізнавання аудіосигналів в інструменті Edge Impulse

Існує кілька способів відправки даних в сервіс Edge Impulse. Найпростіший метод – завантажувати файли безпосередньо.

Edge Impulse приймає аудіофайли, файли зображень та необроблені дані, що зберігаються у форматі CBOR або JSON (рис. 2.6).

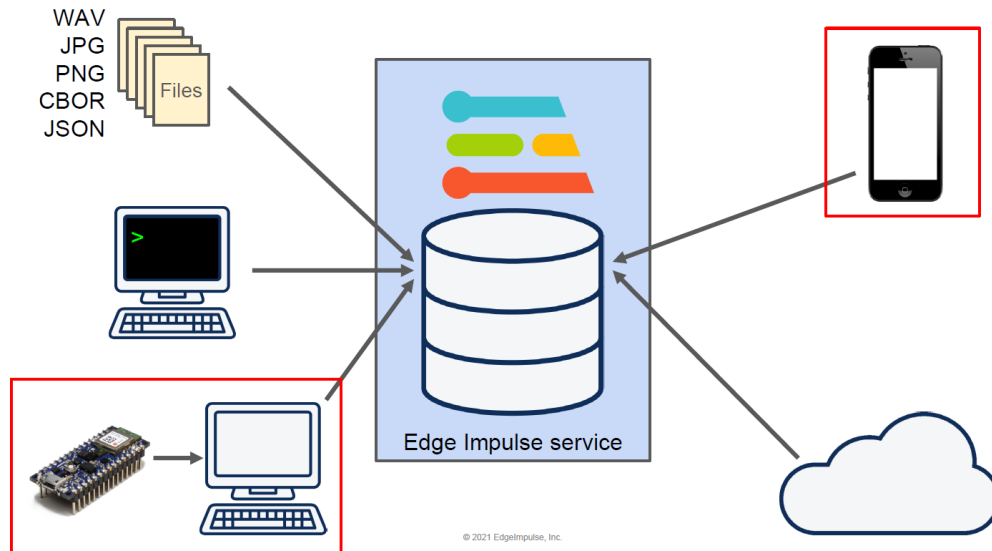


Рисунок 2.6 – Формати файлів, з якими працює Edge Impulse

Edge Impulse підтримує плату розробки Arduino Nano 33 BLE Sense, тому можна запустити спеціальну прошивку на цій платі та інструмент командного рядка на своєму комп'ютері для надсилання даних з датчиків мікроконтролера до сервісу Edge Impulse. Смартфон також може виступати як джерело даних, інформацію з його датчиків можна передавати до Edge Impulse. Також Edge Impulse підтримує прямі підключення до інших хмарних сервісів за допомогою корпоративного облікового запису.

Існує багато способів збору даних. Загальна рекомендація від розробників інструменту Edge Impulse – використовувати той самий тип датчика для збору даних, що і для розгортання моделі на вбудованому пристрої. Наприклад, якщо планується розгорнути модель на Arduino Nano 33 BLE Sense, то рекомендується спочатку зібрати дані за допомогою цього MCU. Це допоможе пояснити будь-яку невідповідність різних моделей датчиків.

Якщо зібрати дані з іншого датчика, наприклад, смартфона, а модель машинного навчання розгорнути на MCU, то можна зіткнутися з ситуацією, коли модель не працює так само добре з новим типом датчика. Один із способів виправити це – витягти з даних ознаки, які не залежать від виконавчої природи цього датчика. Для звуку це може бути перетворення Фур'є, оскільки воно дає частотні компоненти вхідного сигналу. В цьому випадку можна зібрати дані з одних джерел (наприклад, смартфона), потім навчити модель і розгорнути її на іншому обладнанні (рис 2.7)

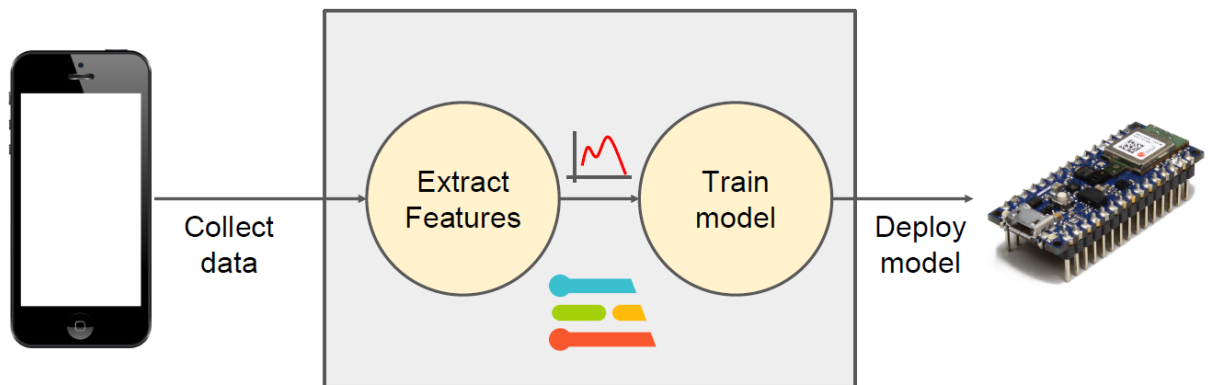


Рисунок 2.7 – Збір даних для машинного навчання на вбудованих пристроях

Для того, щоб оцінити якість роботи класифікатора, будемо використовувати найпростіший метод – метод Holdout [39].

Зібраний набір даних поділяється на два набори, які називають навчальним набором та тестовим набором. Основний спосіб зробити це – випадковим чином змішати дані, потім розділити дані на навчальний набір (Training Set), набір перевірки (Validation Set) та тестовий набір (Test Set). Часто від 10 до 20% даних відкладається у тестовий набір, ще від 10 до 20% відводиться у набір для перевірки, а решта використовується для навчання (рис 2.8).

Holdout Method

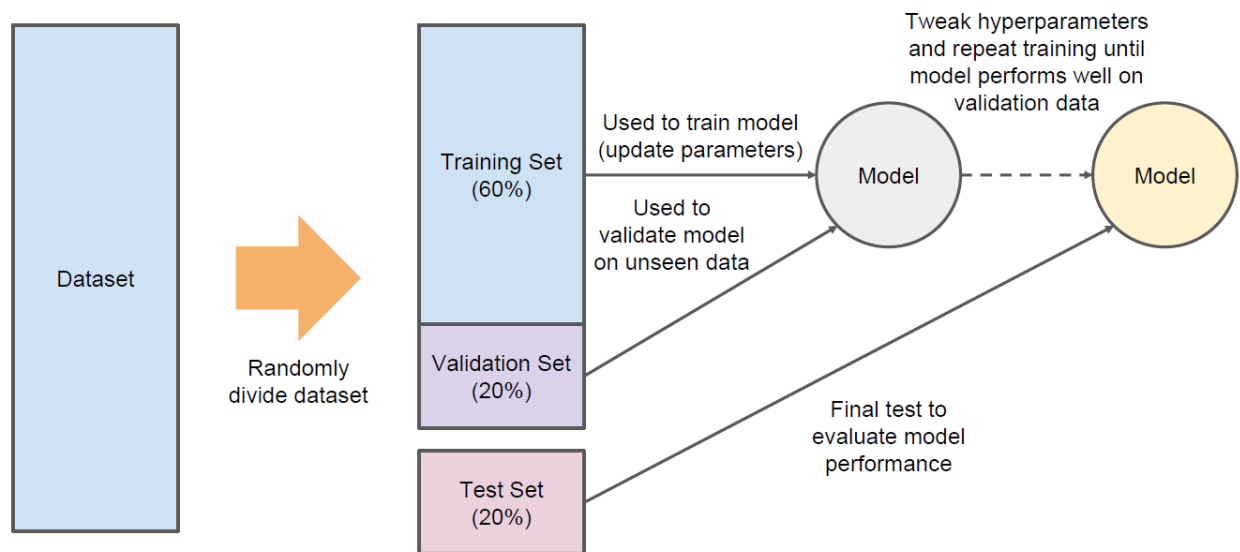


Рисунок 2.8 – Метод Holdout

Алгоритм машинного навчання використовує значення, знайдені в навчальному наборі, та пов'язані з ними мітки для оновлення параметрів у моделі машинного навчання. Потім використовується набір перевірки для перевірки продуктивності моделі, оскільки модель раніше не бачила даних у цьому наборі.

Якщо модель погано працює з цим набором перевірки, можна налаштувати гіперпараметри, такі як розмір і форма моделі, кількість навчальних кроків та швидкість навчання. Далі знову використовується навчальний набір, щоб перекваліфікувати модель з цими новими гіперпараметрами та знову перевірити її продуктивність за допомогою набору валідації.

Лише після того, як буде досягнута задовільна продуктивність моделі, відбувається тестування моделі на тестовому наборі даних. Якщо набір перевірки працює добре, але тестовий набір працює погано, це може бути ознакою того, що модель була перепристосована до вхідних даних, які включають набір перевірки.

Для створення проекту машинного навчання, який повинен повідомляти, чи є в шумах легень патологічні шуми, необхідно зібрати репрезентативний набір аудіоданих без патологічних шумів легень, а також репрезентативний набір аудіоданих із патологічними шумами. Також набір даних для навчання моделі повинен бути збалансованим.

Припустимо, що ми подаємо алгоритму машинного навчання набір даних, який у 99% буде містити звуки легень без патології і лише 1% даних з патологічними шумами. Далі цей набір даних використовується для навчання класифікатора. Тоді при переході на етап тестування моделі виявиться, що незалежно від того, який аудіосигнал ми подаємо класифікатору, він завжди буде класифікувати його як «без патологічних шумів». Такий класифікатор носить назву «наївний класифікатор» (рис. 2.9).

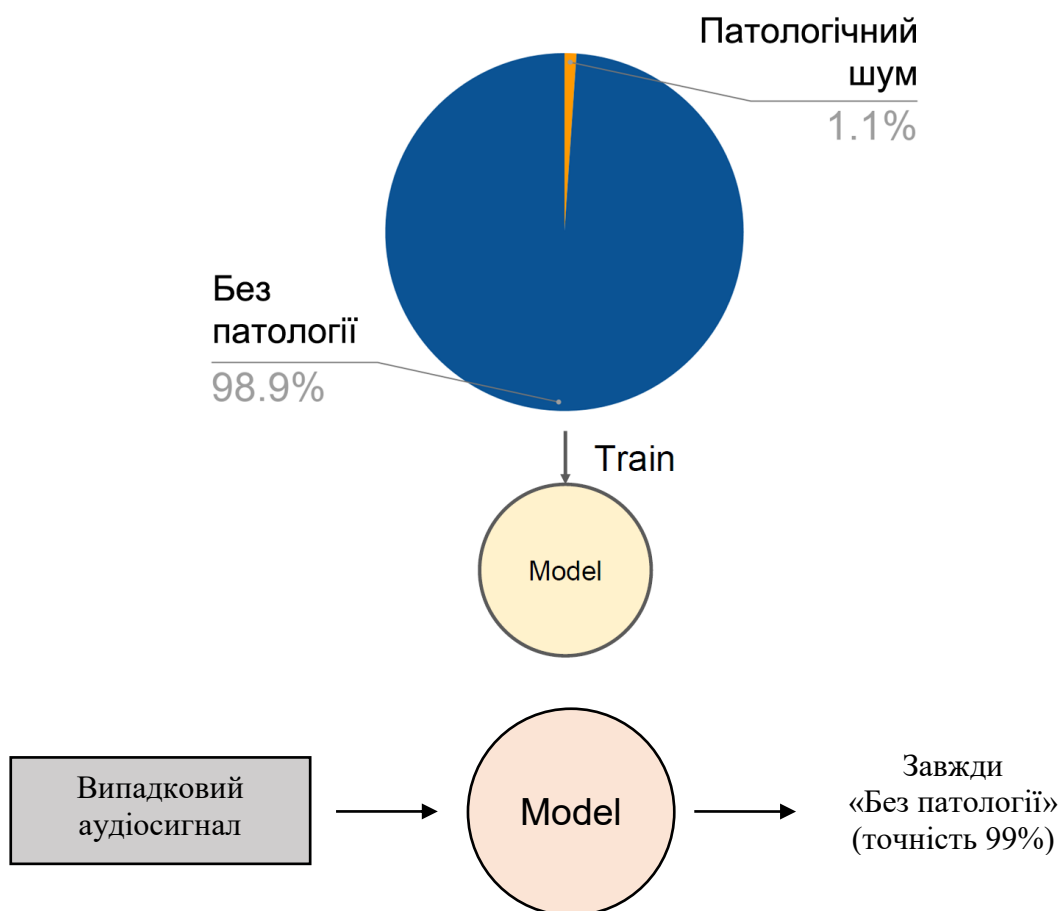


Рисунок 2.9 – Наївний класифікатор

Щоб запобігти цій ситуації, кількість вибірок у кожному класі повинна бути приблизно рівною. Для проблеми двох класів (патологічний шум, без патологічних шумів) це може означати поділ 50% на 50%. Для чотирьох класів прагнуть отримати близько 25% зразків для кожного класу і т.д.

2.4 Отримання ознак аудіоданих

При наявності достатньої кількості аудіоданих (записів шумів легень), можна створити модель з розпізнавання звуків легень (рис. 2.10).

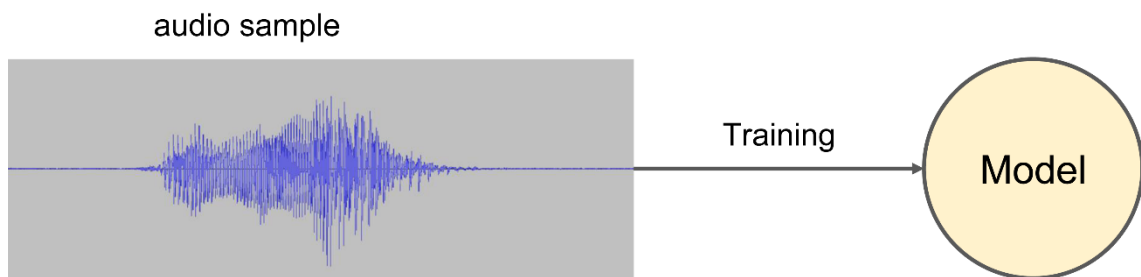


Рисунок 2.10 – Схема навчання моделі

Необроблені аудіодані не будуть мати найкращих ознак для навчання. Модель може наприклад, спробувати вивчити такі характеристики, як гучність, довжина файлу і положення в кадрі, що призведе до погіршення якості моделі. Щоб виправити це, необхідно виділити деякі ознаки самостійно. Фактично, ми виділимо ознаки, які імітують те, як люди сприймають звуки.

У 1800-х роках французький математик Жозеф Фур'є відкрив, що будь-який повторюваний сигнал можна представити у вигляді нескінченної суми синусоїдальних і косинусоїдальних хвиль на різних частотах. Це дало початок перетворенню Фур'є, яке дозволяє розбити будь-який сигнал на його частотні складові. Якби ми застосували перетворення Фур'є до всієї вибірки, ми могли б побачити деякі загальні тенденції в частотах, але ми б втратили будь-яку інформацію про те, як частоти змінюються з часом. Тому ми

візьмемо невеликий проміжок часу і обчислимо швидко перетворення Фур'є або ШПФ.

Оскільки аудіозапис складається з серії рівновіддалених відліків, а не з безперервної лінії, нам потрібно використовувати дискретне перетворення Фур'є. А ШПФ – це просто більш обчислювально ефективний спосіб його обчислення. У цьому невеликому проміжку часу ми можемо побачити, які частоти присутні. Звук шумів легень, як правило, знаходиться в діапазоні від 100 до 1500 Гц. Низькі тони ближче до 100 Гц, більш високі – до 1500 Гц.

Тепер можна трохи пересунути вікно, можливо, навіть перекриваючи попереднє вікно. Оскільки тепер це інша ділянка звуку, частоти можуть дещо змінитися, а отже, зміниться і результат нашого ШПФ. Якщо зробити це для всієї вибірки, ми можемо записати кожне ШПФ у стовпчик, щоб побудувати тривимірний графік, відомий як спектрограма. У спектрограмі вісь y – це частота, вісь x – час, вісь z – амплітуда кожної частотної смуги або спектральної складової з ШПФ. Замість повністю тривимірного графіка часто представляють вісь z у вигляді кольорів. У цьому випадку білі та червоні кольори вказують на високу амплітуду певної частоти, а сині – на низьку амплітуду. Спектрограма для аудіосигналу до 8 кГц зображена на рис. 2.11.

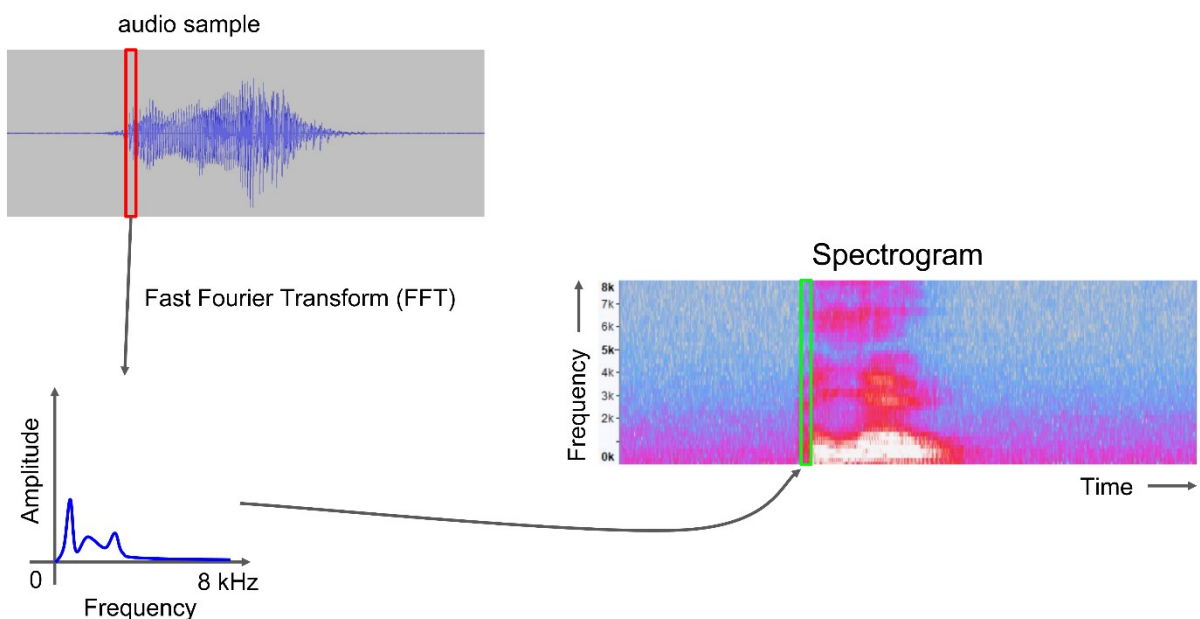


Рисунок 2.11 – Зображення результату ШПФ у вигляді спектрограми

Спектрограми – це, по суті, зображення, а звуки легень створюють унікальні спектрограми. Тому можна використовувати ті ж методи машинного навчання, які розрізняють зображення, для ідентифікації звуків. Тобто спектрограма звуку є одним зі шляхів отримання набіру ознак для нейронних мереж для ідентифікації аудіосигналів.

Дискретне перетворення Фур'є та швидке перетворення Фур'є мають важливе обмеження, про яке слід пам'ятати при розробці систем, що їх використовують. Згідно з теоремою дискретизації Найквіста-Шеннона, для того, щоб повністю захопити всі деталі сигналу, тобто відновити вихідний безперервний сигнал, необхідно робити вибірку на частоті, що вдвічі перевищує найвищу частотну складову цього сигналу.

Наприклад, є фрагмент звуку, який містить деякий аудіосигнал на частоті 10 кГц. Через те, як працює математика дискретного перетворення Фур'є і, аналогічно, ШПФ, ми можемо бачити лише частоти до половини нашої частоти дискретизації. У цьому випадку наша частота дискретизації (f_s) становить 16 кГц. Половина цієї частоти дорівнює 8 кГц. Ця половина частоти відома як частота Найквіста завдяки математичному перетворенню Фур'є. Будь-який компонент частоти, що перевищує цю частоту, дзеркально відображається і додається назад у вихідне ШПФ. Це називається псевдонімізацією.

Кінцевий результат ШПФ виглядатиме приблизно так, як зображено на рис. 2.12 – з високочастотною складовою, яка тепер присутня на частоті 6 кГц. Оскільки нас цікавив сигнал в діапазоні 8 кГц, то ми отримали «зайвий» сигнал.

Поширеним способом запобігання згладжування є використання фільтра низьких частот. Ми встановлюємо частоту зрізу на рівні або поблизу частоти Найквіста. Базовий ідеальний фільтр низьких частот помножить всі частоти нижче зрізу на 1, а всі частоти вище – на 0. Це ефективно усуне будь-яку складову вище частоти Найквіста і забезпечить відсутність аліасингу в нашому остаточному ШПФ (рис. 2.13).

Nyquist-Shannon Sampling Theorem

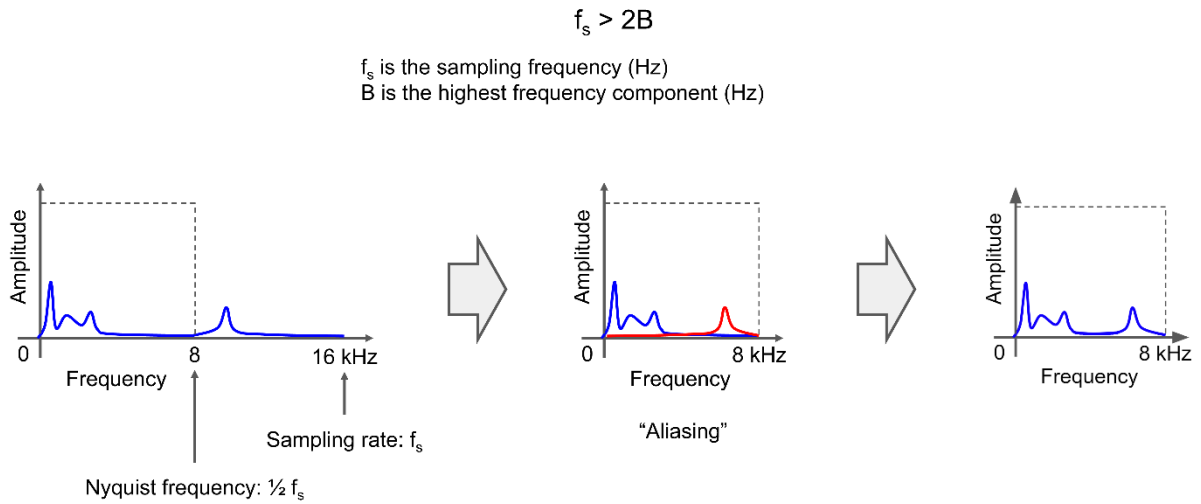


Рисунок 2.12 – Результат роботи ШПФ – зайвий сигнал на частоті 6 кГц

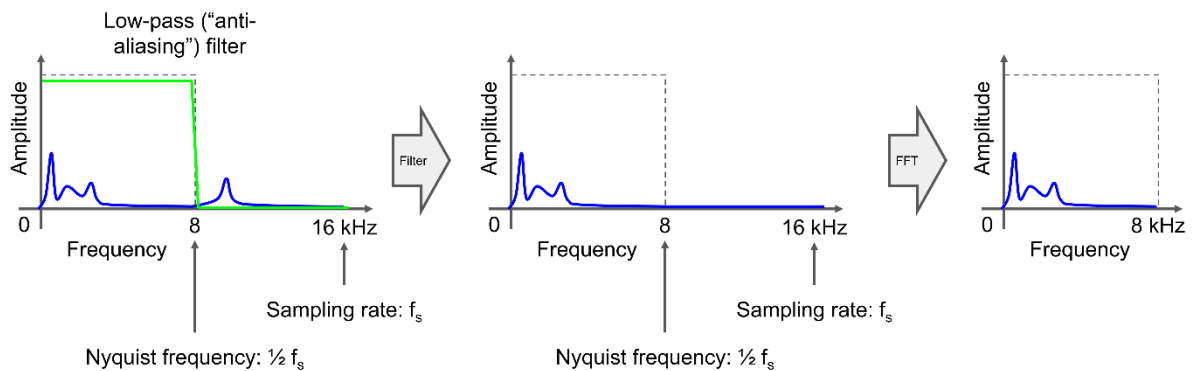


Рисунок 2.13 – Результат застосування фільтру низьких частот

Отже, звук легень, повинен мати частоту, меншу за половину частоти дискретизації (семпсування). Оскільки ми будемо використовувати звуки легень, записані за допомогою електронного стетоскопу 3М Littmann 3200, які мають частоту дискретизації 4 000 Гц, то ці аудіофайли підходять для аналізу звуків легень. Звуки легень, що нас цікавлять, вислуховуються на частотах 100-1 500 Гц, а подвійна максимальна їх частота 3 000 Гц менша за частоту дискретизації 4 000 Гц.

При виділенні ознак з аудіоданих звуків роботи легень ми будемо використовувати також обчислення частотних цепстральні коефіцієнтів Мела (MFCC) – підход для перетворення звуку в набір чисел, який наближено відповідає тому, як людське вухо інтерпретує звук (рис. 2.14).

Mel Frequency Cepstral Coefficients (MFCCs)

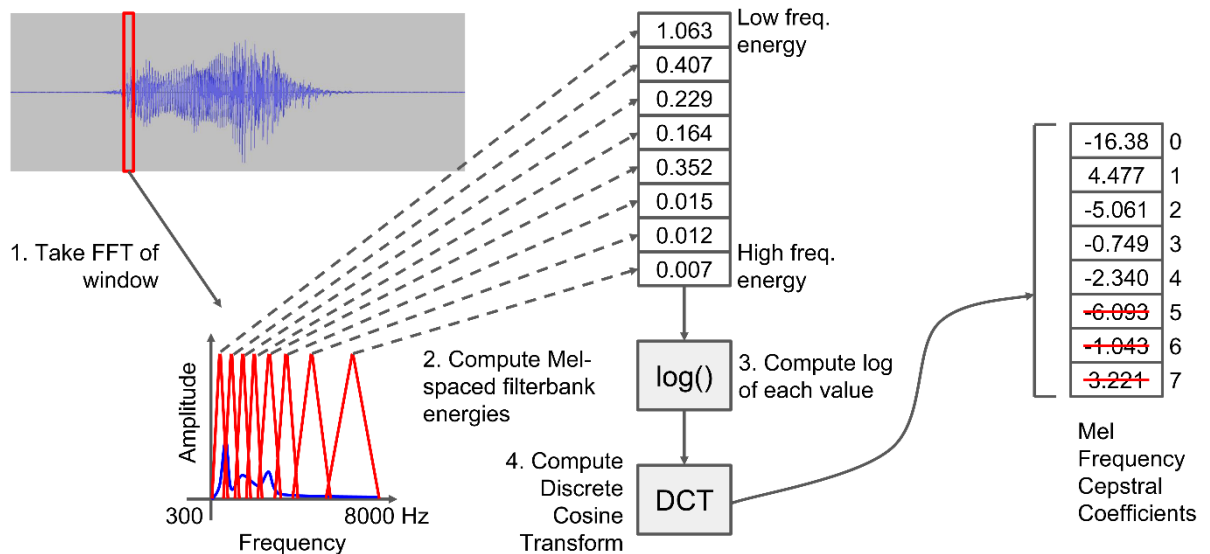


Рисунок 2.14 – Обчислення MFCC

Спочатку береться ШПФ віконного зрізу нашої форми сигналу, як і раніше. Потім створюється набір трикутних фільтрів з інтервалом Мела для цього ШПФ. Для прикладу візьмемо 8 фільтрів. Фільтри для частот менше одного кілогерца розташовані приблизно рівномірно один від одного з однаковою шириною. Понад один кілогерц вони стають ширшими і логарифмічно віддаленими один від одного. Це імітує те, як вухо сприймає звук, оскільки вухо розділяє частоти нелінійно.

Далі обчислюється енергія кожного фільтра. Вона знаходиться шляхом підсумовування площі під кривою після множення ШПФ на фільтр, ці значення енергії зберігаються в масиві. На рисунку цей масив показано з низькочастотними енергіями вгорі і високочастотними внизу.

Потім обчислюється логарифм кожного значення енергії. Знову ж таки, це імітує те, як людина чує звук. Ми сприймаємо гучність не за лінійною шкалою, а за логарифмічною.

Далі береться дискретне косинусне перетворення (ДКП) цієї групи значень енергії. Оскільки фільтри накладаються один на одного, ДКП дозволяє нам корелювати значення енергії і стискає інформацію. Результатом цієї операції є частотні кепстральні коефіцієнти Мела або MFCC.

Кількість вихідних значень з ДКП дорівнює кількості вхідних значень. Отже, якщо ми почали з 32 фільтрів, то на виході ми отримаємо 32 MFCC. Функція стиснення ДКП дозволяє нам відкинути вищі MFCC. Нижчі коефіцієнти описують загальну форму вхідного сигналу, тоді як вищі коефіцієнти описують швидкі зміни форми. Коли мова йде про розпізнавання звуків легень, нас цікавить лише загальна форма спектрограми, і ДКП допомагає нам досягти цього.

Щоб створити повний набір ознак, потрібно обчислити MFCC кілька разів (рис. 2.15).

Mel Frequency Cepstral Coefficients (MFCCs)

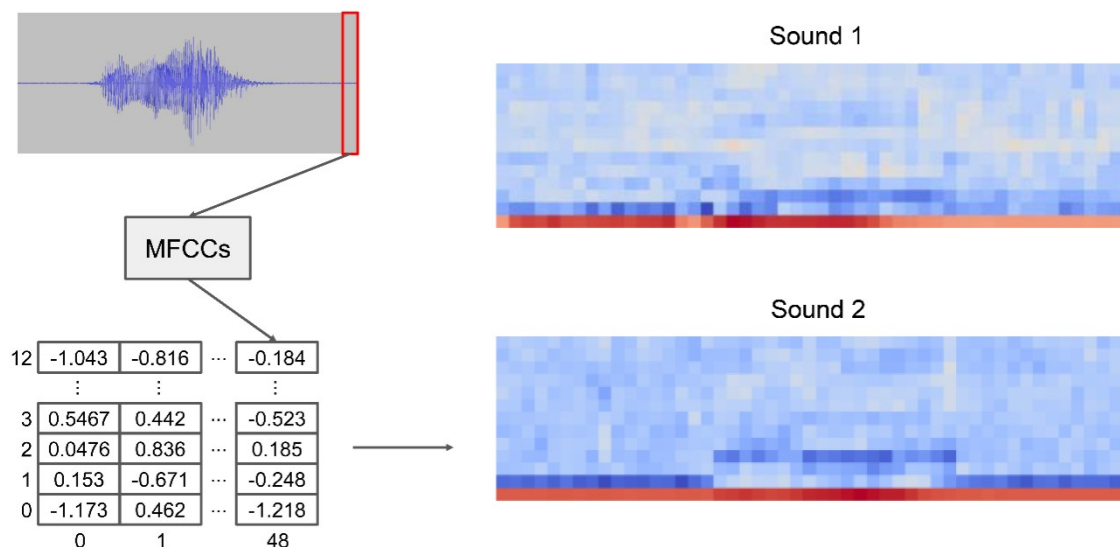


Рисунок 2.15 – Результат обчислення MFCC у вигляді зображень

Для цього прикладу ми починаємо розрахунок з вікна над першою частиною звуку і обчислимо 32 MFCC, далі відкинемо останні 19, оскільки вони нам не потрібні. Це дасть 13 MFCC для першого вікна. Потім зсунемо вікно, трохи перекриваючи його з першим вікном. Обчислюємо MFCC з цього вікна і зберігаємо їх поруч з першим набором. Повторюємо цей процес, поки не отримаємо 49 наборів по 13 значень MFCC.

Як і у випадку зі спектрограмою, ми отримуємо результат у вигляді зображень. Далі з метою класифікації звуків легень ми можемо використовувати нейронні мережі, призначені для класифікації зображень. Людське око може побачити невеликі відмінності в цих зображеннях, але комп'ютеру набагато легше помітити відмінності.

2.5 Згорткові нейронні мережі

Нейронні мережі бувають різних форм і розмірів і мають безліч застосувань. Розглянемо популярну мережу, яка використовується для класифікації зображень – згорткову нейронну мережу. Згорткова нейронна мережа, також відома як CNN або ConvNet, складається з серії операцій, які автоматично знаходять і виділяють особливості або шаблони у вхідних даних. Саме це робить її такою корисною для аналізу та класифікації зображень. Хоча CNN може мати й інші шари або вузли, які допомагають у фільтрації та класифікації, унікальними її роблять саме шари згортки [40].

Шари згортки містять фільтри, параметри яких оновлюються так само, як і параметри у звичайних вузлах нейронної мережі. Це означає, що їх можна навчати за допомогою зворотного поширення. Ці фільтри застосовуються до зображення і автоматично виділяють різні особливості. До них відносяться такі речі, як широкі форми, наприклад, очевидні кола, і повторювані візерунки, такі як хутро, трава або текстура деревини. Вони також можуть виокремлювати краї, наприклад, межу між двома об'єктами (рис. 2.16).

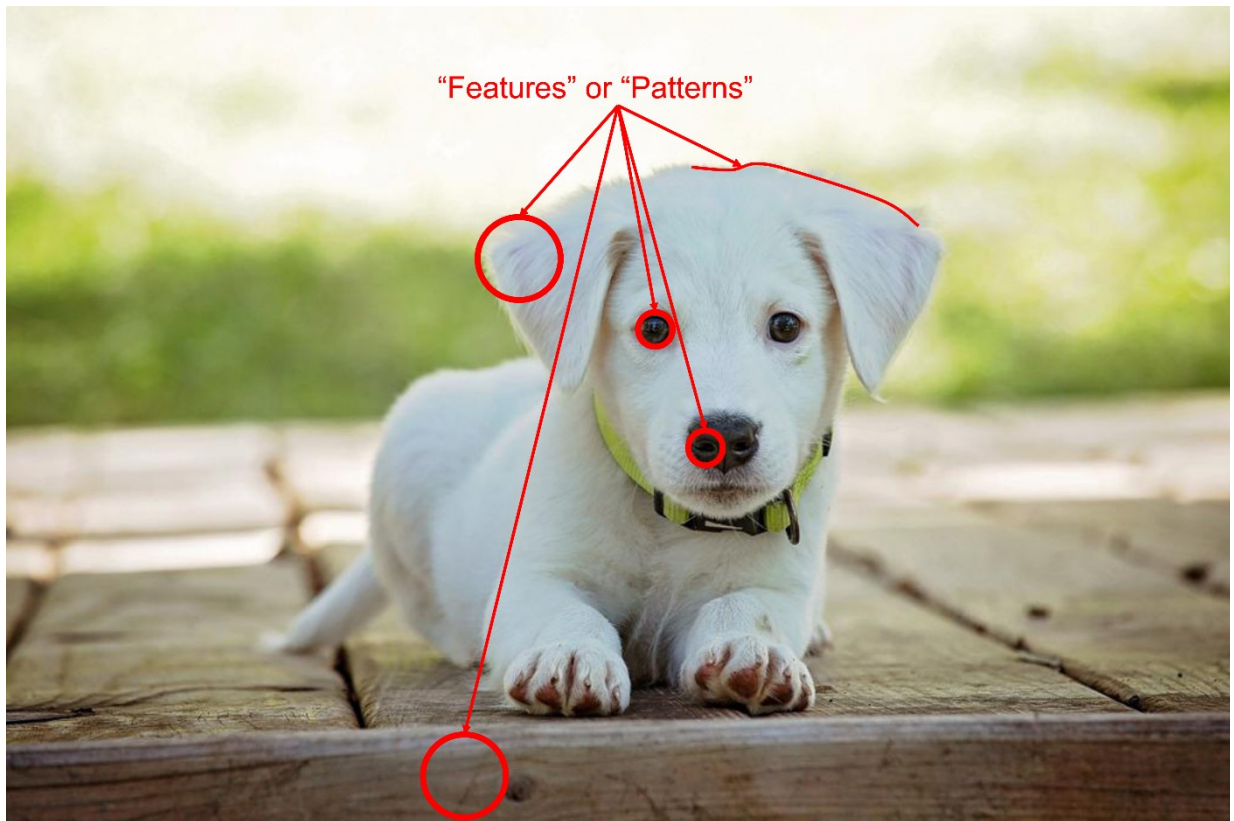


Рисунок 2.16 – Застосування фільтрів або шаблонів до зображення

Зазвичай CNN застосовується з більш ніж одним згортковим шаром, і кожен наступний шар виділяє складніші об'єкти, ніж попередній. Ці наступні шари починають виділяти вуха, очі, носи та інші складні форми. За наявності достатньої кількості шарів програма може почати розпізнавати цілі обличчя, людей, тварин, машини, дерева тощо.

Кожен шар згортки складається з декількох фільтрів. За замовчуванням в Edge Impulse для першого шару використовується вісім фільтрів. Зображення або матриця Mel-частотних кепстральних коефіцієнтів у цьому випадку копіюється на кожен з цих фільтрів. Після фільтрації та максимального об'єднання, вихід кожного вузла копіюється на вхід кожного іншого вузла в наступному шарі. Якщо наступний шар складається з 16 фільтрів, ці вісім вихідних зображень копіюються на кожен фільтр. Тепер ми маємо зберігати 128 зображень. Процес фільтрації та максимальний крок об'єднання зменшує розмір кожного відфільтрованого зображення, що допомагає зменшити використання пам'яті та процесора (рис. 2.17).

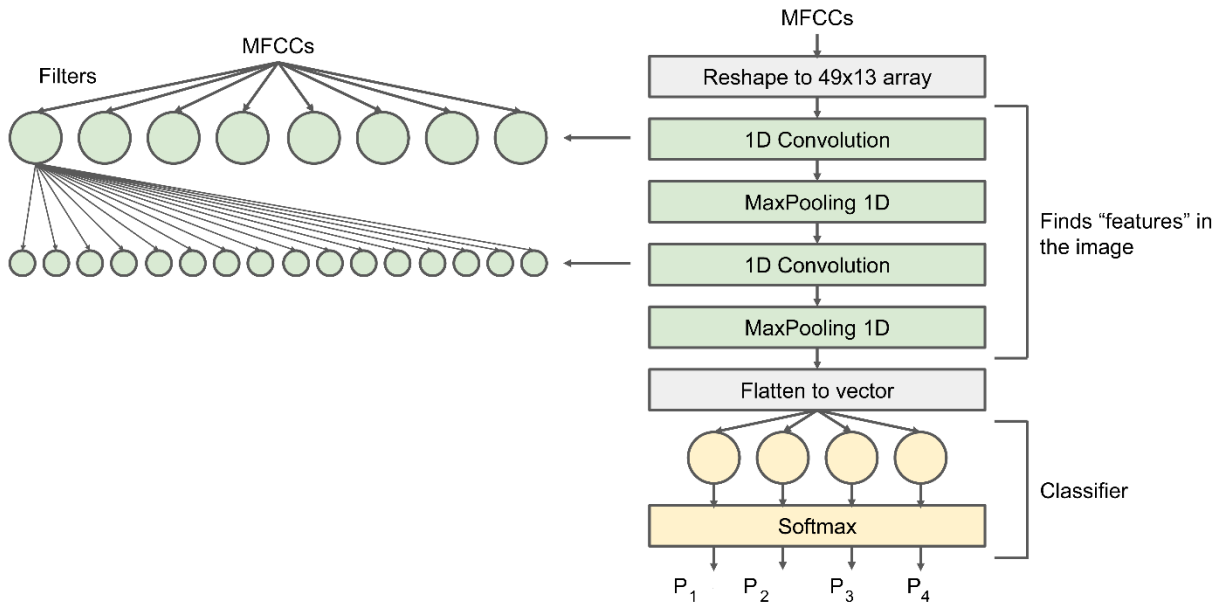


Рисунок 2.17 – Процес фільтрації зображення за допомогою CNN

Розглянемо зображення у відтінках сірого. Кожен піксель такого зображення – це просто число. У великих налаштуваннях машинного навчання вони можуть бути перетворені на значення з плаваючою комою і нормалізовані до значень між нулем і одиницею (рис. 2.18).

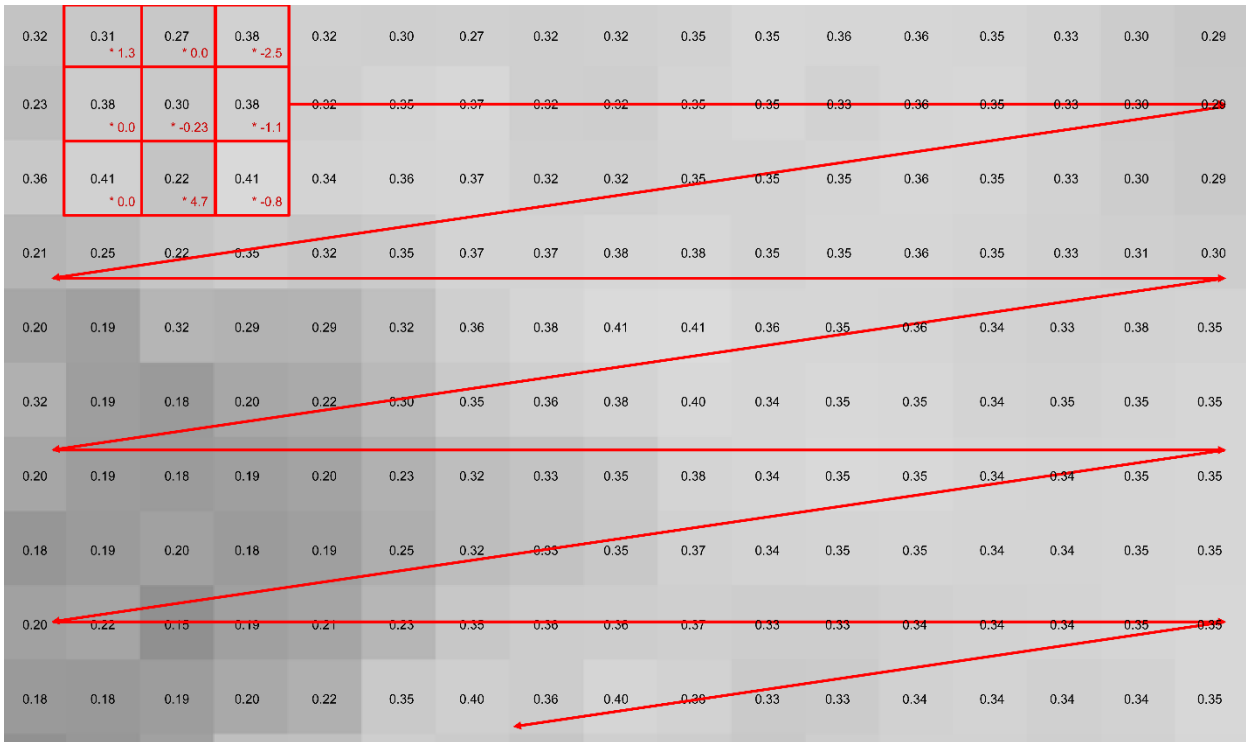


Рисунок 2.18 – Фільтр згортки

Фільтр, також відомий як ядро, – це матриця з параметрами, які перемножують значення пікселів. Ці добутки потім підсумовуються разом, щоб сформувати нове значення, яке стає одним пікселем відфільтрованого зображення (точковий добуток). Фільтр з тими самими параметрами переходить до наступного набору пікселів і виконує ту саму операцію, обчислюючи нове значення пікселя для цих дев'яти вихідних пікселів. Фільтр продовжує робити це для всього зображення, створюючи нове відфільтроване зображення. Процес фільтрації з використанням множника і деякого ядра називається згортка.

Ядро має розмір три на три для двовимірного фільтра згортки. Зміна розміру ядра – це гіперпараметр, який можна змінювати, щоб побачити, чи покращить це роботу класифікатора зображень. Параметрами є значення множників у ядрі. Вони автоматично оновлюються в процесі навчання, щоб допомогти всій моделі дізнатися, які ознаки найкраще підходять для класифікації.

У згортковій нейронній мережі кожен піксель вихідного зображення часто пропускається через деяку нелінійну функцію, відому як активація. Часто використовується випрямлена лінійна одиниця, яка просто відкидає всі від'ємні значення результуючого зображення. Це допомагає зробити будь-яку особливість ще більш помітною (рис. 2.19).

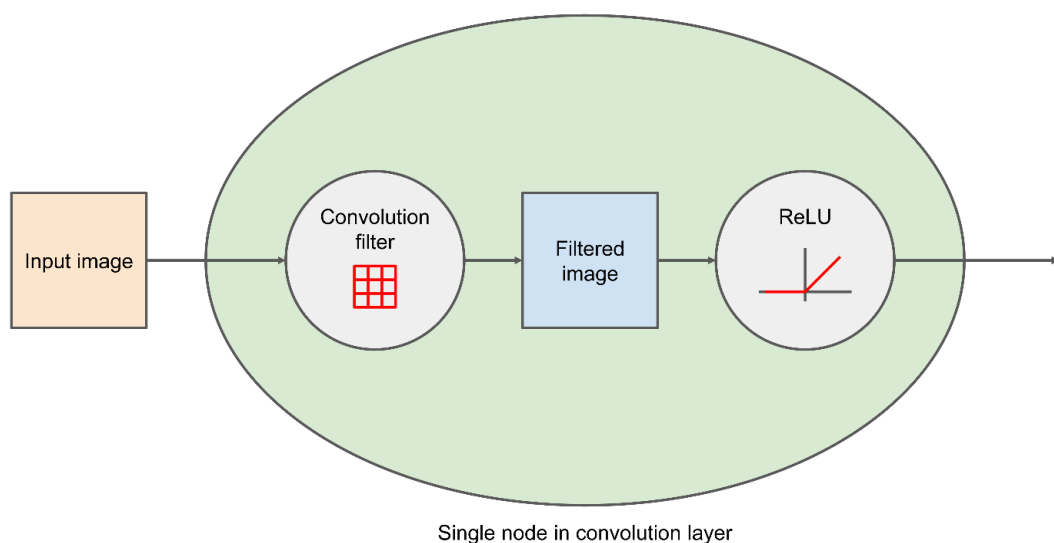


Рисунок 2.19 – Активація

2-D згортка – дуже трудомісткий процес, оскільки зазвичай використовується невелике ядро для проходження всього зображення. Іноді це не потрібно, і можна використовувати ядро того ж розміру, що й зображення в одному вимірі (одновимірна згортка).

Зображення MFCC виглядає як зображення у відтінках сірого, оскільки це просто двовимірна матриця числових значень. Кольори на зображенні додаються як ефект за допомогою програмного забезпечення Edge Impulse, оскільки вони можуть допомогти помітити більшу різницю між значеннями. Червоний колір тепліший і, відповідно, більший за величиною, ніж синій. Фільтр або ядро 1-D згортки на прикладі зображення MFCC 13x49 пікселів має вигляд, представлений на рис. 2.20.

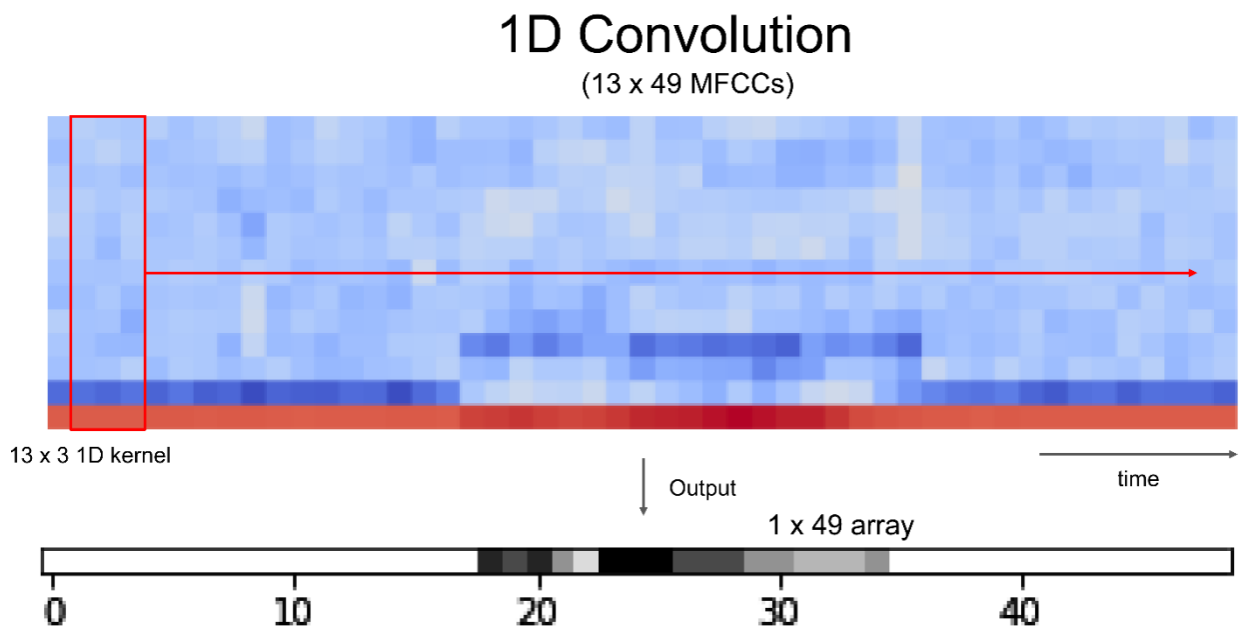


Рисунок 2.20 – Ядро одновимірної згортки на прикладі MFCC

Рображення має ширину 13 пікселів і довжину 49 пікселів. Є два варіанти фільтра: можна створити фільтр, що йде зверху вниз, або фільтр, що йде збоку в бік. Для MFCC нам потрібен фільтр, що йде від краю до краю, щоб зберегти часову інформацію, закодовану в матриці, оскільки частотні діапазони змінюються з часом. Таким чином, ми отримуємо ядро шириною 13 пікселів. Можна змінювати довжину ядра, але за замовчуванням довжина у Edge Impulse становить три пікселі.

Як і раніше, фільтр бере точковий добуток своїх параметрів і значень пікселів під ними. Це значення зберігається, і фільтр зсувається, щоб знову обчислити точковий добуток. Так продовжується для всього набору MFCCs. Результатом кожного кроку фільтрації в цьому першому шарі є масив один на 49 елементів, де кожен піксель позначає ознаку на кожному часовому зрізі. Тепер, коли ми маємо вісім різних відфільтрованих зображень, кожне з яких позначає набір ознак вихідного зображення, потрібно виконати крок максимального об'єднання (рис. 2.21).

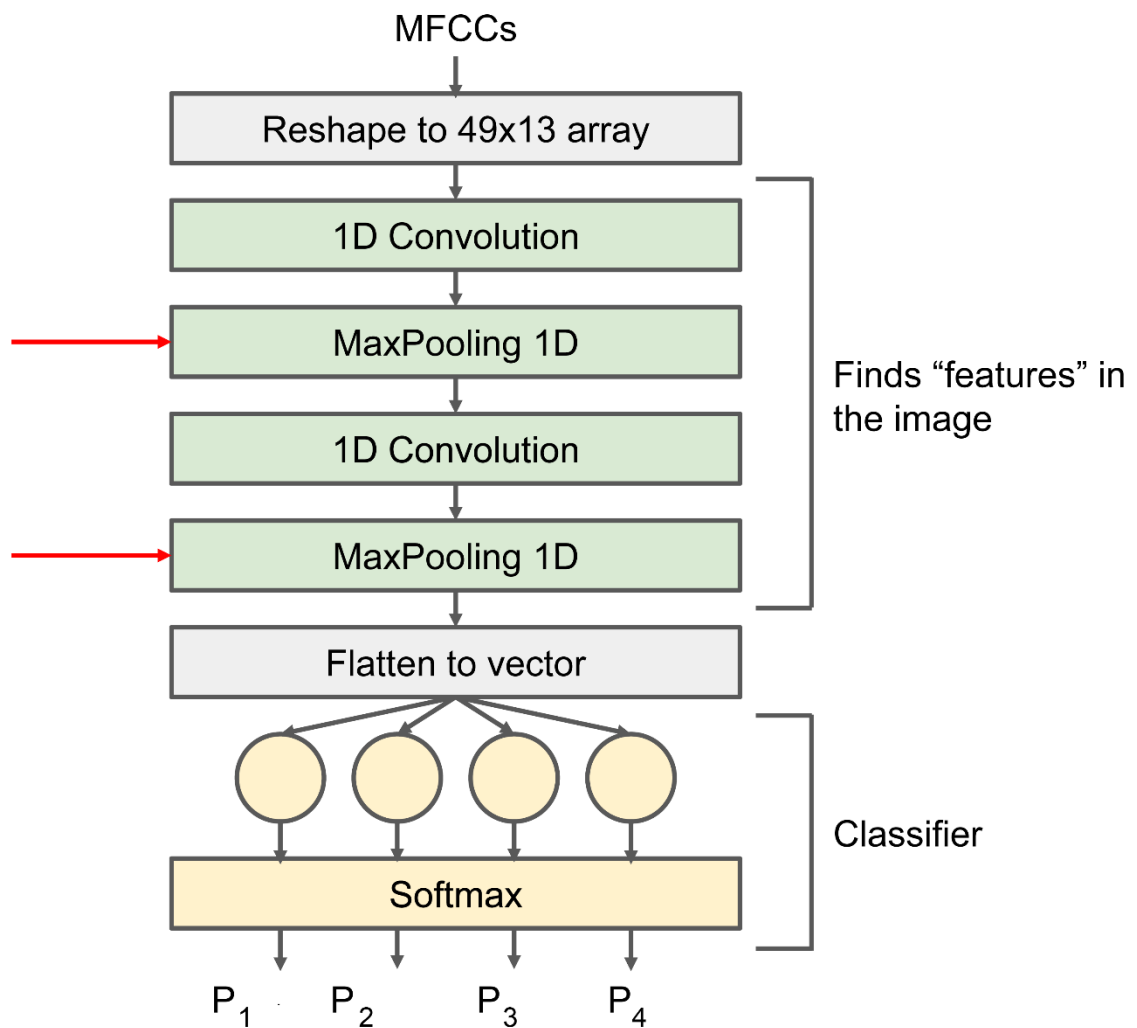


Рисунок 2.21 – Кроки максимального об'єднання

У CNN часто використовують шари максимального або середнього об'єднання, щоб зменшити розмір відфільтрованих зображень, що допомагає знизити обчислювальні витрати. Максимальне об'єднання шарів починається

так само, як і згортка. Застосовується вікно до нашого зображення замість того, щоб обчислювати точковий добуток, однак, ми просто обираємо найбільше число в цьому пулі. У цьому випадку це буде 0,3. Якби ми використовували середнє об'єднання, ми б усереднили всі значення пікселів у вікні разом, щоб отримати одне число. Ми застосовуємо його до виходу шару згортки, тому це буде відфільтроване зображення (рис. 2.22)

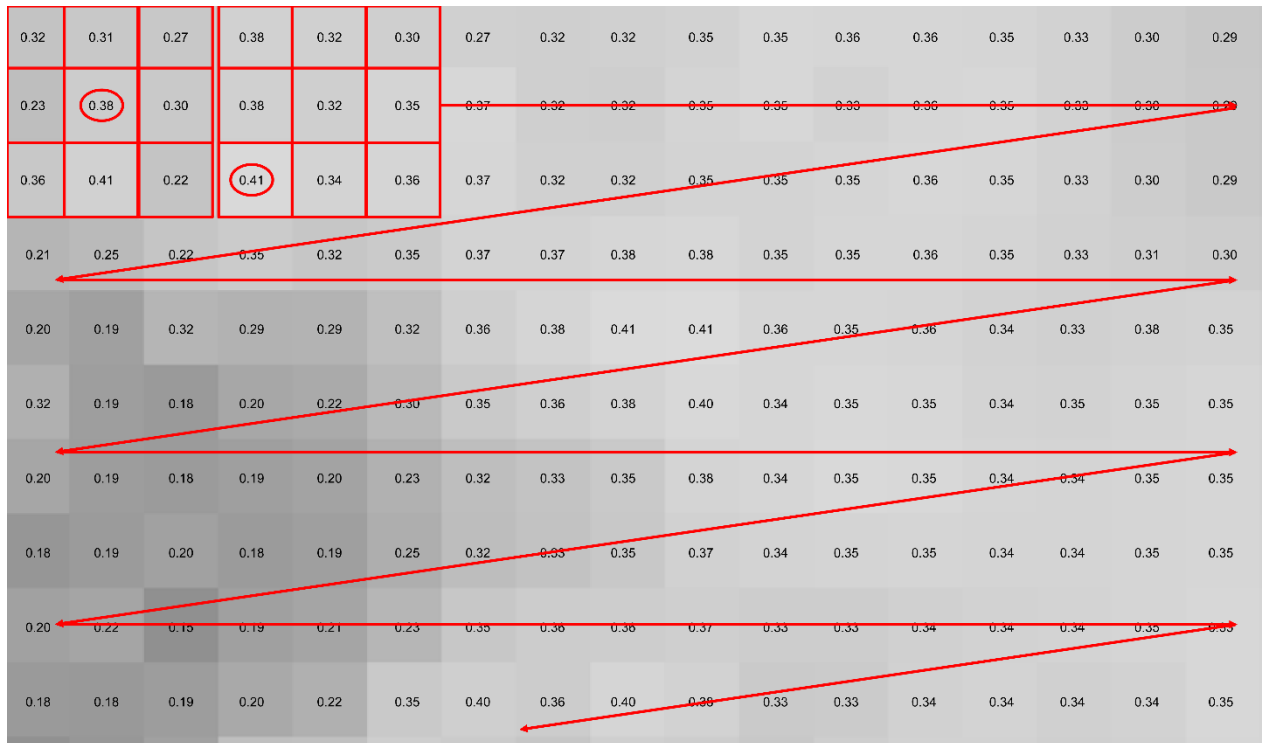


Рисунок 2.22 – Шар максимального об'єднання

Потім вікно пересувається і процес повторюється, знаходячи найвище значення в цьому пулі. Можна вибрати перекриття або перескакувати фільтр стільки разів, скільки потрібно. Чим менше перекриття між вікнами, тим більше стиснеться зображення, хоча в подальшому можна заощадити на обчисленнях, але при цьому втратити деякі функції. Тому часто потрібно трохи поекспериментувати, щоб знайти розмір вікна і довжину кроку, які найкраще підходять для програми. Цей процес триває для всього зображення і називається двовимірним процесом максимального об'єднання.

Оскільки ми працюємо з одновимірними масивами на виході кроку згортки, можна легко виконати одновимірний крок максимального об'єднання. Тут вікно накладається на підмножину пікселів у масиві, максимальне значення цих пікселів зберігається у вихідному масиві. Пул ковзає або перескакує, і процес повторюється для всього масиву.

В Edge Impulse розмір вікна за замовчуванням – один на два пікселі, і воно перескакує на два пікселі, створюючи масив вдвічі менший за вихідний. Якщо в масиві буде непарна кількість пікселів, Edge Impulse додасть додатковий нуль спереду або ззаду оригінального масиву, щоб зробити його парним. Це робиться для того, щоб запобігти втраті деяких значень, якщо стрибки вікна не будуть ідеально вирівняні. Таким чином можна отримати на виході 25 пікселів, навіть якщо на вході 49 (рис. 2.23).

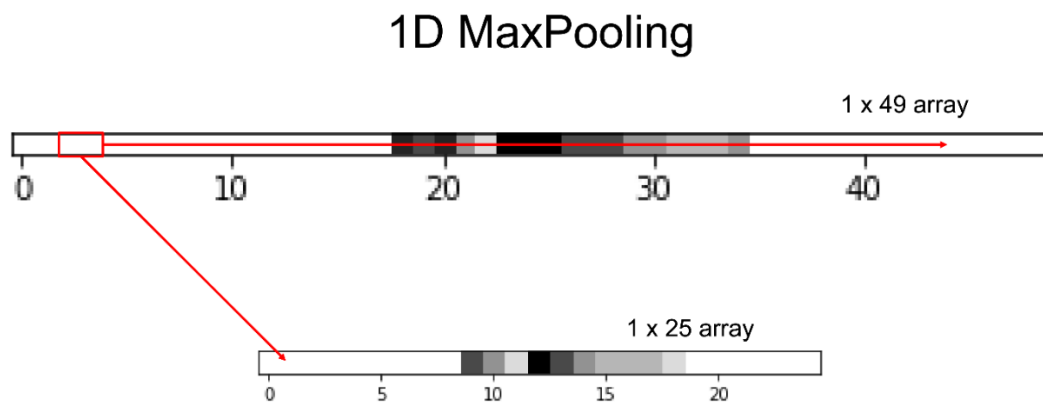


Рисунок 2.23 – Одновимірне максимальне об'єднання

Зазвичай використовується шар з максимальним об'єднанням після кожного кроку згортки, щоб зменшити розмір результуючого відфільтрованого зображення. Згортка з максимальним кроком об'єднання повторюється стільки разів, скільки вважається за потрібне. У прикладі з Edge Impulse ми бачимо це двічі. Другий шар згортки має 16 вузлів замість восьми, як у першому. Другий шар максимального об'єднання ефективно зменшує розмір масиву вдвічі, тому на виході кожного вузла з 16 вузлами ми отримуємо матрицю розміром один на 13 елементів, що дає 16 матриць розміром один на 13 елементів.

Заключною частиною CNN майже завжди є класифікатор, який зазвичай є нейронною мережею з будь-якою кількістю шарів. Перша частина вибирає відповідні ознаки, а друга частина класифікує ці ознаки, оскільки класифікатор очікує одновимірний вхід, ми повинні спочатку звести наші 16 масивів в один довгий масив. Ми просто складаємо їх кінець в кінець, перш ніж надсилати класифікатору. Якби масиви були двовимірними або тривимірними матрицями, їх також потрібно було б перетворити у одновимірний вхід (рис. 2.24).

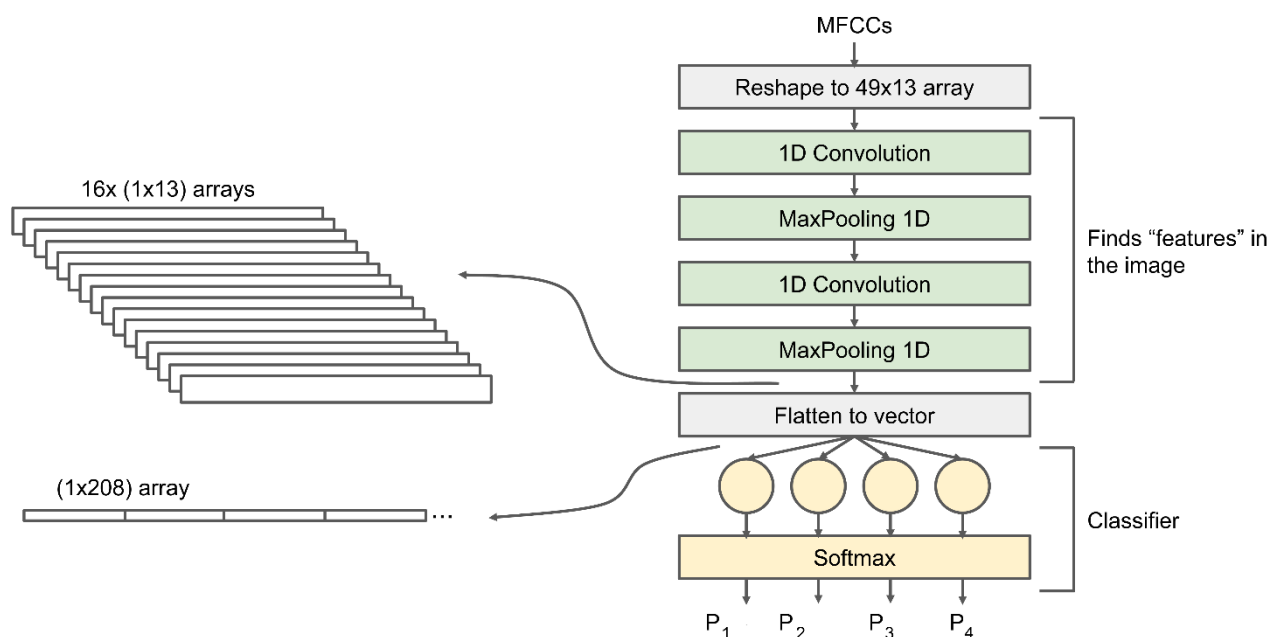


Рисунок 2.24 – Заклучна частина CNN

Частина класифікатора за замовчуванням в Edge Impulse складає чотири вузли, по одному вузлу для кожного з чотирьох класів.

Функція softmax перетворює результати роботи CNN в набір ймовірностей. Кожна ймовірність повинна бути між нулем і одиницею, і всі ймовірності в сумі повинні дорівнювати одиниці. Звідси випливає, що нам необхідно шукати вихід з найбільшою ймовірністю. Наприклад, якщо P_3 є найвищим, то ми можемо сказати, що модель прогнозує, що вхідні ознаки належать до третього класу.

В процесі навчання оновлюються параметри фільтра згортки, а також ваги та умови зміщення в повністю пов'язаному шарі класифікатора. Коли це буде зроблено, ми зможемо подати частотні цефстральні коефіцієнти Мела до цієї моделі і змусити її передбачити клас, до якого вона належить.

Розглянемо також шари відсікання (рис. 2.25).

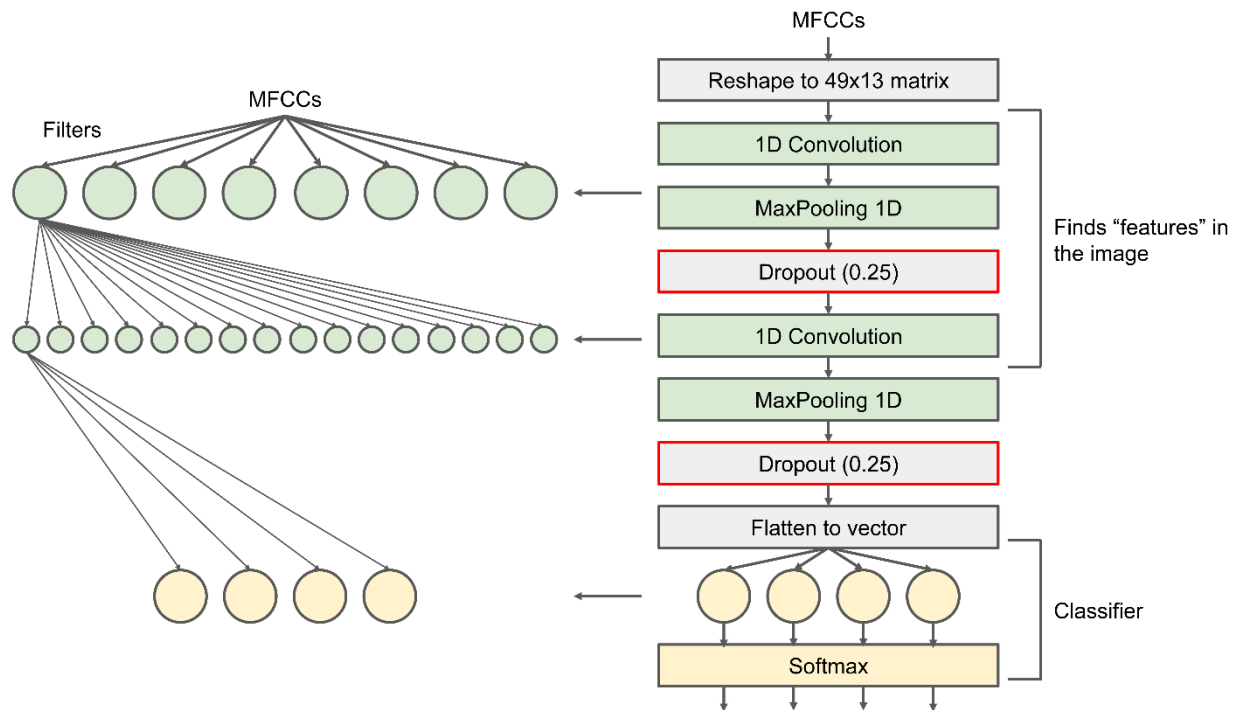


Рисунок 2.25 – Шари відсікання

Ці шари можна додавати між іншими шарами, щоб запобігти надмірному підлаштуванню. Вони призводять до того, що певний відсоток результатів попереднього шару повністю ігнорується під час навчання. У цьому випадку ігнорується 25 відсотків виходів з об'єднаних шарів. Вузлам наступного шару доводиться брати на себе більшу відповідальність за вхідні дані, що залишилися, які зазвичай відкидаються або випадково вибираються перед тим, як кожна вибірка зробить прямий прохід і мережа під час навчання. Це робить процес навчання шумним і часто може допомогти з надмірним підлаштуванням. Шари відсікання застосовуються лише під час навчання і не діють під час розгортання моделі.

2.6 Висновки до розділу 2

Для вирішення задачі класифікації аудіосигналів може бути застосоване машинне навчання. Для того, щоб навчити контрольований алгоритм машинного навчання розрізняти нормальні і патологічні шуми легень, необхідно надати йому на етапі навчання збалансований набір вхідних даних. Це важливо, щоб уникнути ситуації «наївного класифікатора».

В процесі навчання алгоритм виділяє ознаки серед отриманих даних, в нашому випадку в якості ознак ми будемо використовувати Mel-частотні кепстральні коефіцієнти, отримані при обробці аудіофайлів з шумами легень. Далі ми можемо використати цю навчену модель машинного навчання для того, щоб здійснювати класифікацію звуків легень на патологічні і нормальні з точністю моделі, яку вона виявить в процесі тестування.

Для того, щоб оцінити якість роботи класифікатора, можна використовувати метод Holdout, при якому 20% даних відкладається у тестовий набір і не використовується на етапі навчання. Як правило, точність класифікації на етапі тестування моделі нижча, ніж на етапі навчання, тому що алгоритм класифікації може бути перепристосованим до навчальних даних, але тестові дані він побачить вперше тільки на етапі тестування.

Завдяки останнім розробкам процесорів стало можливим застосування алгоритмів машинного навчання на вбудованих системах. Це було досягнуто завдяки появі процесорів, що надають можливість обробки сигналів у ядрах – наприклад, процесорів Arm Cortex-M4 у мікроконтролерах ArduinoNano 33 BLESense. Став можливим запуск алгоритмів машинного навчання на більш компактних обчислювальних пристроях.

Розглянута робота програмного забезпечення EdgeImpulse, що надає можливість створювати моделі машинного навчання і завантажувати їх до мікроконтролерів. Таким чином, класифікація аудіосигналів легень може

відбуватись на автономному пристрої (мікроконтролері) без підключення до комп'ютера.

Розглянутий метод цифрової обробки аудіосигналів – швидке перетворення Фур'є. Зазначено, що звуки легень, записані за допомогою електронного стетоскопу 3M Littmann 3200, які мають частоту дискретизації 4 000 Гц, підходять для аналізу звуків легень, оскільки шуми легень вислуховуються на частотах 100-1 500 Гц і подвійна максимальна їх частота 3 000 Гц (частота Найквіста) менша за частоту дискретизації аудіо 4 000 Гц.

Розглянуті згорткові нейронні мережі (CNN), призначення їх шарів та побудова конвеєру машинного навчання у ПЗ EdgeImpulse.

Узагальнюючи, можна виділити такі етапи побудування моделі машинного навчання для класифікації шумів легень у ПЗ EdgeImpulse:

1. Отримання репрезентативного і збалансованого набору аудіоданих для навчання і тестування моделі. Частота дискретизації аудіофайлів повинна бути не менше ніж 3 000 Гц (частота Найквіста).
2. Виконується швидке перетворення Фур'є.
3. Знаходяться Mel-частотні кепстральні коефіцієнти MFCC. В результаті отримуємо зображення.
4. Зображення MFCC обробляється алгоритмом машинного навчання з використанням згорткової нейронної мережі.
5. На виході з алгоритму машинного навчання ми отримуємо результат класифікації у відсотках.
6. Здійснюється тестування моделі на тестовому наборі даних, який не приймав участі на етапі навчання.
7. Якщо результат класифікації на тестовому наборі даних є задовільним, здійснюється експорт коду моделі і її завантаження до мікроконтроллера для подальшої автономної роботи.

3 РЕАЛІЗАЦІЯ МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ КЛАСИФІКАЦІЇ АКУСТИЧНИХ ШУМІВ ЛЕГЕНЬ

3.1 Формування наборів даних для навчання моделі

Для отримання аудіозаписів різних шумів легень скористаємось готовими наборами даних легеневих звуків, записаних із грудної стінки за допомогою електронного стетоскопа 3M Littmann 3200.

Перший набір даних [41] включає звуки семи захворювань: астми, серцевої недостатності, пневмонії, бронхіту, плеврального випоту, фіброзу легень і хронічного обструктивного захворювання легень (ХОЗЛ), а також звуки нормального дихання. Розташування стетоскопа на обстежуваному визначав лікар-спеціаліст, який проводив діагностику. Записи передавався на комп'ютер за допомогою адаптера Bluetooth.

Набір даних включає дихальні шуми від ста дванадцяти суб'єктів (35 здорових і 77 хворих). Середній вік досліджуваних від 21 до 90 років, з них 43 жінки та 69 чоловіків. Детальна демографічна інформація та кількість суб'єктів із відповідним станом здоров'я описані в таблиці 3.1.

Таблиця 3.1 –Стан здоров'я досліджуваних, включених у набір даних, і їх демографічна інформація

Стан здоров'я	Кількість	Вік	Стать
Нормальний	35	18–81	11 жінок, 24 чоловіки
Астма	32	12–72	17 жінок, 15 чоловіків
Пневмонія	5	36–70	2 жінки, 3 чоловіки
ХОЗЛ	9	42–76	1 жінка, 8 чоловіків
Бронхіт	3	20–68	1 жінка, 2 чоловіки
Серцева недостатність	21	20–83	9 жінок, 12 чоловіків
Фіброз легень	5	44–90	2 жінки, 3 чоловіки
Плевральний випіт	2	70–81	0 жінок, 2 чоловіка

Набір даних містить один запис на суб'єкта. Тривалість кожного запису коливається від 5 до 30 секунд, чого достатньо, щоб охопити хоча б один дихальний цикл. Максимальна тривалість обмежена можливостями запису електронного стетоскопа. Кожен запис був відтворений три рази відповідно до різних частотних фільтрів, які підкреслюють певні тілесні звуки.

Ім'я кожного файлу даних починається з типу фільтра, закодованою літерою В, D або E. За нею йде літера Р, унікальний порядковий номер пацієнта, починаючи з 1, і підкреслення. Після цього ім'я файлу містить діагноз, тип звуку, місце вимірювання на грудях, вік і стать суб'єкта.

До даних було включено три типи фільтрів. Буква В використовується з фільтрацією в режимі Белла, яка посилює звуки в діапазоні частот [20-1000] Гц, але підкреслює низькочастотні звуки в діапазоні [20-200] Гц. Буква D використовується з фільтрацією в режимі діафрагми, яка підсилює звуки в діапазоні частот [20-2000] Гц, але підкреслює звуки частоти в діапазоні [100-500] Гц. Буква E використовується з фільтрацією розширеного режиму, яка підсилює звуки в діапазоні частот [20-1000] Гц, але підкреслює частоту звуків в діапазоні [50-500] Гц.

Зона грудей (рис. 3.1) кодується трьома впорядкованими буквами з наборів {A, P}, {L, R} і {L, M, U} відповідно.

Букви мають такі значення:

- Передня: A (Anterior);
- Задня: P (Posterior);
- Ліворуч: L (Left);
- Праворуч: R (Right);
- Нижня: L (Lower);
- Верхня: U (Upper);
- Середня: M (Middle).

Крім того, тип звуку кодується як: вдих: I (Inspiratory), видих: E (Expiratory), свист: W (Wheezes), потріскування: C (Crackles) або крепітація: Crp (Crepitations), N (Normal): нормальний звук легень.

Стать закодована літерою F (Female) для жіночої або M (Male) для чоловічої статі.

Наприклад, файл із назвою «BP60_heart failure,Сrep,PLL,83,F» – це відфільтрований в режимі Белла звук крепітації, отриманий із задньої лівої нижньої зони грудної клітки 83-річної пацієнтки із серцевою недостатністю.

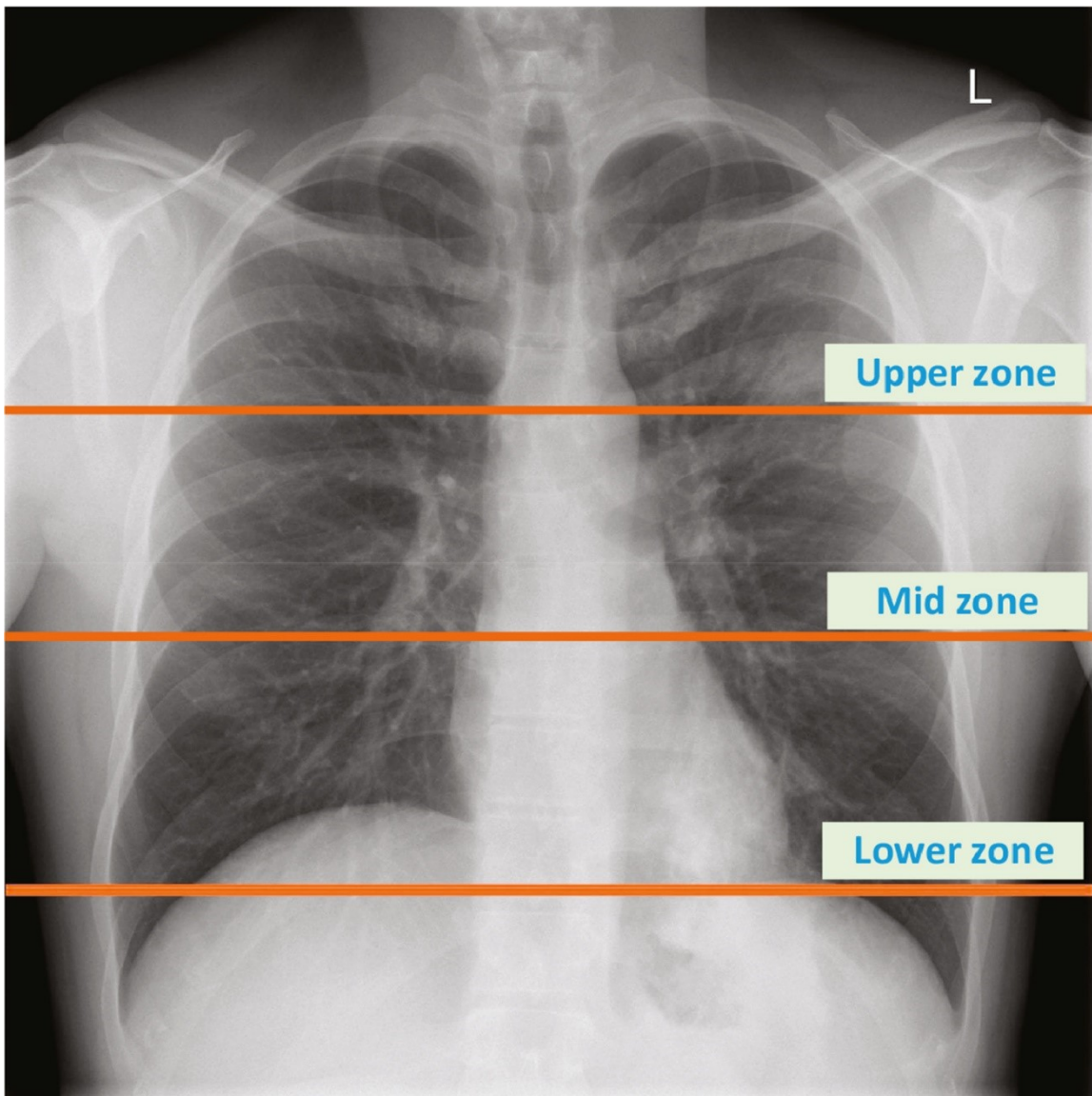


Рисунок 3.1 – Розташування зон грудної клітини, які використовувалися для запису легеневих звуків

Цей набір даних можна використовувати для розробки автоматизованих методів, які виявляють легеневі захворювання за легневими звуками або ідентифікують правильний тип легневих звуків.

Для навчання моделі класифікації звуків легень візьмемо з цієї бази аудіофайли за двома діагнозами: нормальний звук легень і астма – тому що цих аудіозаписів представлено найбільше у базі.

Для класифікації звуків легень візьмемо аудіозаписи, що записані з одного місця грудної клітини, тобто наприклад, будемо порівнювати записи астми, записаної з місця PLM (Posterior left middle – задня ліва зона легень посередині) з аудіозаписами нормального звуку легень з цього ж місця PLM.

Будемо використовувати 4 вибірки даних з цієї бази, кожна вибірка буде відповідати місцям запису PLM, PRM, PLU, PRU – тому що аудіозаписів астми і нормальних звуків легень з цих місць грудної клітини зроблено найбільше. Тобто будемо використовувати максимально можливу кількість аудіоданих виходячі з діагнозу (астма, нормальні звуки легень) і місця запису (PLM, PRM, PLU, PRU). Перелік аудіофайлів з цього набору даних, що був використаний для моделі машинного навчання, наведений у лістингу 3.1.

Лістинг 3.1 – Перелік файлів з діагнозом астма і нормальними звуками легень за місцем запису звуків легень

PLM

asthma (12 файлів)

BP43_asthma,E W,P L M,15,M.wav
 BP47_asthma,E W,P L M,60,M.wav
 BP58_asthma,E W,P L M,40,M.wav
 BP101_Asthma,E W,P L M,12,F.wav
 DP43_asthma,E W,P L M,15,M.wav
 DP47_asthma,E W,P L M,60,M.wav
 DP58_asthma,E W,P L M,40,M.wav
 DP101_Asthma,E W,P L M,12,F.wav
 EP43_asthma,E W,P L M,15,M.wav

EP47_asthma,E W,P L M,60,M.wav
 EP58_asthma,E W,P L M,40,M.wav
 EP101_Asthma,E W,P L M,12,F.wav

normal (18 файлів)

BP31_N,N,P L M,27,F.wav
 BP51_N,N,P L M,30,M.wav
 BP89_N,N,P L M,70,M.wav
 BP95_N,N,P L M,18,M.wav
 BP109_N,N,P L M,26,M.wav
 BP112_N,N,P L M,30,M.wav
 DP31_N,N,P L M,27,F.wav
 DP51_N,N,P L M,30,M.wav
 DP89_N,N,P L M,70,M.wav
 DP95_N,N,P L M,18,M.wav
 DP109_N,N,P L M,26,M.wav
 DP112_N,N,P L M,30,M.wav
 EP31_N,N,P L M,27,F.wav
 EP51_N,N,P L M,30,M.wav
 EP89_N,N,P L M,70,M.wav
 EP95_N,N,P L M,18,M.wav
 EP109_N,N,P L M,26,M.wav
 EP112_N,N,P L M,30,M.wav

PRM

asthma (18 файлів)

BP27_asthma,E W,P R M,53,F.wav
 BP33_Asthma,E W,P R M,43,F.wav
 BP38_Asthma,E W,P R M,38,M.wav
 BP53_asthma,E W,P R M,72,F.wav
 BP56_asthma,E W,P R M,56,F.wav
 BP78_Asthma,E W,P R M,20,M.wav

DP27_asthma,E W,P R M,53,F.wav
DP33_Asthma,E W,P R M,43,F.wav
DP38_Asthma,E W,P R M,38,M.wav
DP53_asthma,E W,P R M,72,F.wav
DP56_asthma,E W,P R M,56,F.wav
DP78_Asthma,E W,P R M,20,M.wav
EP27_asthma,E W,P R M,53,F.wav
EP33_Asthma,E W,P R M,43,F.wav
EP38_Asthma,E W,P R M,38,M.wav
EP53_asthma,E W,P R M,72,F.wav
EP56_asthma,E W,P R M,56,F.wav
EP78_Asthma,E W,P R M,20,M.wav

normal (21 файл)

BP30_N,N,P R M,18,F.wav
BP49_N,N,P R M,21,M.wav
BP74_N,N,P R M,74,M.wav
BP87_N,N,P R M,72,M.wav
BP93_N,N,P R M,75,M.wav
BP99_N,N,P R M,50,M.wav
BP100_N,N,P R M,70,F.wav
DP30_N,N,P R M,18,F.wav
DP49_N,N,P R M,21,M.wav
DP74_N,N,P R M,74,M.wav
DP87_N,N,P R M,72,M.wav
DP93_N,N,P R M,75,M.wav
DP99_N,N,P R M,50,M.wav
DP100_N,N,P R M,70,F.wav
EP30_N,N,P R M,18,F.wav
EP49_N,N,P R M,21,M.wav
EP74_N,N,P R M,74,M.wav

EP87_N,N,P R M,72,M.wav

EP93_N,N,P R M,75,M.wav

EP99_N,N,P R M,50,M.wav

EP100_N,N,P R M,70,F.wav

PLU

asthma (15 файлів)

BP46_asthma,E W,P L U,41,F.wav

BP64_asthma,E W,P L U,60,M.wav

BP104_Asthma,E W,P L U,45,F.wav

BP106_Asthma,E W,P L U,45,F.wav

BP107_Asthma,E W,P L U,59,F.wav

DP46_asthma,E W,P L U,41,F.wav

DP64_asthma,E W,P L U,60,M.wav

DP104_Asthma,E W,P L U,45,F.wav

DP106_Asthma,E W,P L U,45,F.wav

DP107_Asthma,E W,P L U,59,F.wav

EP46_asthma,E W,P L U,41,F.wav

EP64_asthma,E W,P L U,60,M.wav

EP104_Asthma,E W,P L U,45,F.wav

EP106_Asthma,E W,P L U,45,F.wav

EP107_Asthma,E W,P L U,59,F.wav

normal (15 файлів)

BP48_N,N,P L U,31,M.wav

BP70_N,N,P L U,52,F.wav

BP75_N,N,P L U,73,F.wav

BP81_N,N,P L U,33,M.wav

BP82_N,N,P L U,38,M.wav

DP48_N,N,P L U,31,M.wav

DP70_N,N,P L U,52,F.wav

DP75_N,N,P L U,73,F.wav

DP81_N,N,P L U,33,M.wav

DP82_N,N,P L U,38,M.wav

EP48_N,N,P L U,31,M.wav

EP70_N,N,P L U,52,F.wav

EP75_N,N,P L U,73,F.wav

EP81_N,N,P L U,33,M.wav

EP82_N,N,P L U,38,M.wav

PRU

asthma (18 файлів)

BP10_Asthma,E W,P R U,59,M.wav

BP15_Asthma,E W,P R U,49,F.wav

BP44_asthma,E W,P R U,40,M.wav

BP55_asthma,E W,P R U,72,F.wav

BP91_asthma,E W,P R U,43,F.wav

BP97_Asthma,E W,P R U,24,M.wav

DP10_Asthma,E W,P R U,59,M.wav

DP15_Asthma,E W,P R U,49,F.wav

DP44_asthma,E W,P R U,40,M.wav

DP55_asthma,E W,P R U,72,F.wav

DP91_asthma,E W,P R U,43,F.wav

DP97_Asthma,E W,P R U,24,M.wav

EP10_Asthma,E W,P R U,59,M.wav

EP15_Asthma,E W,P R U,49,F.wav

EP44_asthma,E W,P R U,40,M.wav

EP55_asthma,E W,P R U,72,F.wav

EP91_asthma,E W,P R U,43,F.wav

EP97_Asthma,E W,P R U,24,M.wav

normal (12 файлів)

BP71_N,N,P R U,36,M.wav

BP72_N,N,P R U,24,F.wav

BP86_N,N,P R U,68,F.wav
BP103_N,N,P R U,81,F.wav
DP71_N,N,P R U,36,M.wav
DP72_N,N,P R U,24,F.wav
DP86_N,N,P R U,68,F.wav
DP103_N,N,P R U,81,F.wav
EP71_N,N,P R U,36,M.wav
EP72_N,N,P R U,24,F.wav
EP86_N,N,P R U,68,F.wav
EP103_N,N,P R U,81,F.wav

Розширимо цей перелік, додавши аудіозаписи пацієнтів з ХОЗЛ, отримані з іншого набору даних [42]. Це база даних респіраторних звуків, що містить зразки звуку, зібрані незалежно двома дослідницькими групами в двох різних країнах протягом кількох років. Більшість бази даних складається із зразків аудіо, записаних дослідницькою групою Школи наук про здоров'я Університету Авейру (ESSUA) у Лабораторії респіраторних досліджень та реабілітації (Lab3R) ESSUA та в лікарні Infante D. Pedro, Авейру, Португалія. Друга дослідницька група з Університету Аристотеля в Салоніках (AUTH) і Університету Коїмбри (UC) отримала звуки дихання в Загальній лікарні Папаніколау в Салоніках і в Загальній лікарні Іматії (відділення охорони здоров'я Наузи), Греція.

База даних складається із записів загальною тривалістю 5,5 годин, що містять 6898 дихальних циклів, з яких 1864 містять потріскування, 886 містять свисти, 506 містять як потріскування, так і свисти, у 920 анотованих звукових зразках від 126 суб'єктів.

Записи збиралися за допомогою різноманітного обладнання, їх тривалість коливалася від 10 до 90 секунд. Також надано місця розташування грудної клітки, звідки було отримано записи. Рівень шуму в деяких циклах дихання високий, що імітує реальні умови життя.

Кожне ім'я файлу в цьому наборі даних поділено на 5 елементів, розділених символом підкреслення (_).

1. Номер пацієнта (101,102,...,226).

2. Індекс запису.

3. Місце запису:

- Трахея (Tc);
- Передня ліва частина (Al – Anterior left);
- Передня права частина (Ar – Anterior right);
- Задня ліва частина (Pl – Posterior left);
- Задня права частина (Pr – Posterior right);
- Бічна ліва частина (Ll);
- Бічна права частина (Lr).

4. Режим отримання:

- послідовний/одноканальний (sc);
- одночасний/багатоканальний (mc).

5. Записуюча апаратура:

- Мікрофон AKG C417L (AKGC417L);
- Стетоскоп 3M Littmann Classic II SE (LittC2SE);
- Електронний стетоскоп 3M Littmann 3200 (Litt3200);
- Електронний стетоскоп WelchAllyn Meditron Master Elite (Meditron).

Аудіофайли надані за такими діагнозами:

- ХОЗЛ: хронічне обструктивне захворювання легень;
- LRTI: інфекція нижніх дихальних шляхів;
- URTI: інфекція верхніх дихальних шляхів.

З цього набору даних були відібрані аудіозаписи звуків легень за діагнозом ХОЗЛ, записані за допомогою електронного стетоскопу 3M Littmann 3200 із зон грудної клітини Pl – Posterior left і Pr – Posterior right (табл. 3.2).

Таблиця 3.2 – Вибірка даних за досліджуваними з діагнозом ХОЗЛ

Patient number	Recording index	Chest location	Acquisition mode	Recording equipment	Diagnosis
104	1b1	Pl	sc	Litt3200	COPD
109	1b1	Pl	sc	Litt3200	COPD
112	1p1	Pl	sc	Litt3200	COPD
113	1b1	Pl	sc	Litt3200	COPD
118	1b1	Pl	sc	Litt3200	COPD
124	1b1	Pl	sc	Litt3200	COPD
139	1b1	Pl	sc	Litt3200	COPD
175	1b1	Pl	sc	Litt3200	COPD
185	1b1	Pl	sc	Litt3200	COPD
195	1b1	Pl	sc	Litt3200	COPD
218	1p1	Pl	sc	Litt3200	COPD
104	1b1	Pr	sc	Litt3200	COPD
109	1b1	Pr	sc	Litt3200	COPD
112	1p1	Pr	sc	Litt3200	COPD
113	1b1	Pr	sc	Litt3200	COPD
118	1b1	Pr	sc	Litt3200	COPD
124	1b1	Pr	sc	Litt3200	COPD
139	1b1	Pr	sc	Litt3200	COPD
175	1b1	Pr	sc	Litt3200	COPD
185	1b1	Pr	sc	Litt3200	COPD
195	1b1	Pr	sc	Litt3200	COPD
218	1p1	Pr	sc	Litt3200	COPD

Після додавання аудіозаписів звуків легень пацієнтів з ХОЗЛ до даних пацієнтів з астмою і нормальним звуком легень було сформовано чотири вибірки даних: PLM, PRM, PLU, PRU. Кожна з вибірок має аудіозаписи трьох діагнозів: астма (asthma), ХОЗЛ (copd), нормальні звуки легень (normal). Всі аудіозаписи зроблені на однаковому обладненні – за допомогою електронного стетоскопу 3М Littmann 3200 з однакових зон легень, мають однакову частоту дискретизації 4 000 Гц. Кінцевий перелік файлів бази для навчання моделі вказаний у додатку Б.

3.2 Визначення оптимальних параметрів для навчання моделі в ПЗ Edge Impulse

Обчислимо наші ознаки для чотирьох вибірок аудіоданих в ПЗ Edge Impulse. Для цього на сторінці створення проектів створимо 4 проекти з назвами: Sound Classifier PLU, Sound Classifier PLM, Sound Classifier PRM, Sound Classifier PRU (рис. 3.2). В кожному проекті будемо використовувати відповідну вибірку аудіоданих звуків легень (PLU, PLM, PRM, PRU).

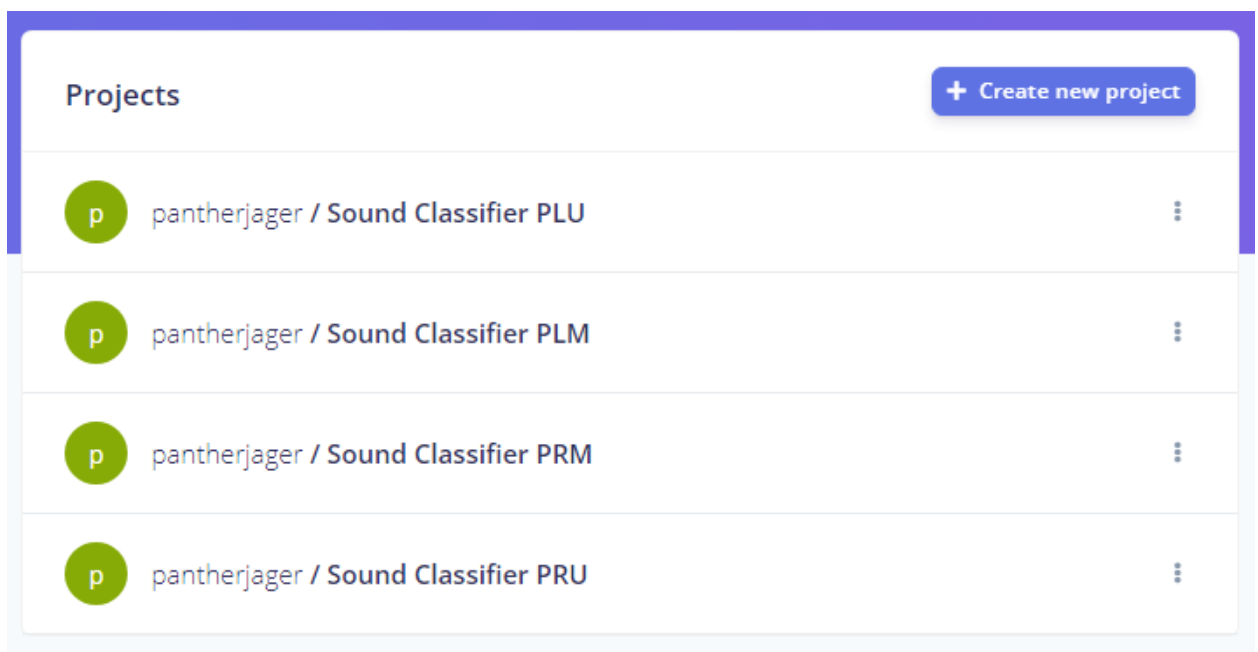


Рисунок 3.2 – Сторінка створення проектів у ПЗ Edge Impulse

Далі на прикладі проекту Sound Classifier PRU здійснено вибірку ознак з аудіофайлів і подальше навчання нейронної мережі.

На першому етапі здійснимо завантаження аудіофайлів з вибірки PRU до навчального набору даних за допомогою веб-інтерфейсу Edge Impulse на сторінці Data acquisition. При завантаженні присвоємо аудіофайлам відповідні 3 класи: asthma, cord, normal. З-за того, що файли були відібрані з двох різних баз, вони розрізняються за своєю тривалістю в секундах.

На другому етапі за допомогою вбудованого інструменту Split sample (розділити зразок) випадковим чином розіб'ємо завантажені аудіофайли на більш менші зразки, тривалістю 2 секунди (рис. 3.3). Розмір 2 секунди був обраний тому, що за цей проміжок часу людиною може бути здійснений вдих або видих.



Рисунок 3.3 – Розбиття аудіофайлів зі зразками звуків легень на більш менші зразки

В результаті застосування операції Split sample до всіх аудіофайлів отримуємо однакові зразки аудіоданих тривалістю 2 секунди, розбиті на 3 категорії: asthma (50 зразків), copd (76 зразків), normal (47 зразків). Вікно веб-інструменту Edge Impulse зі зразками аудіофайлів зображено на рис. 3.4.

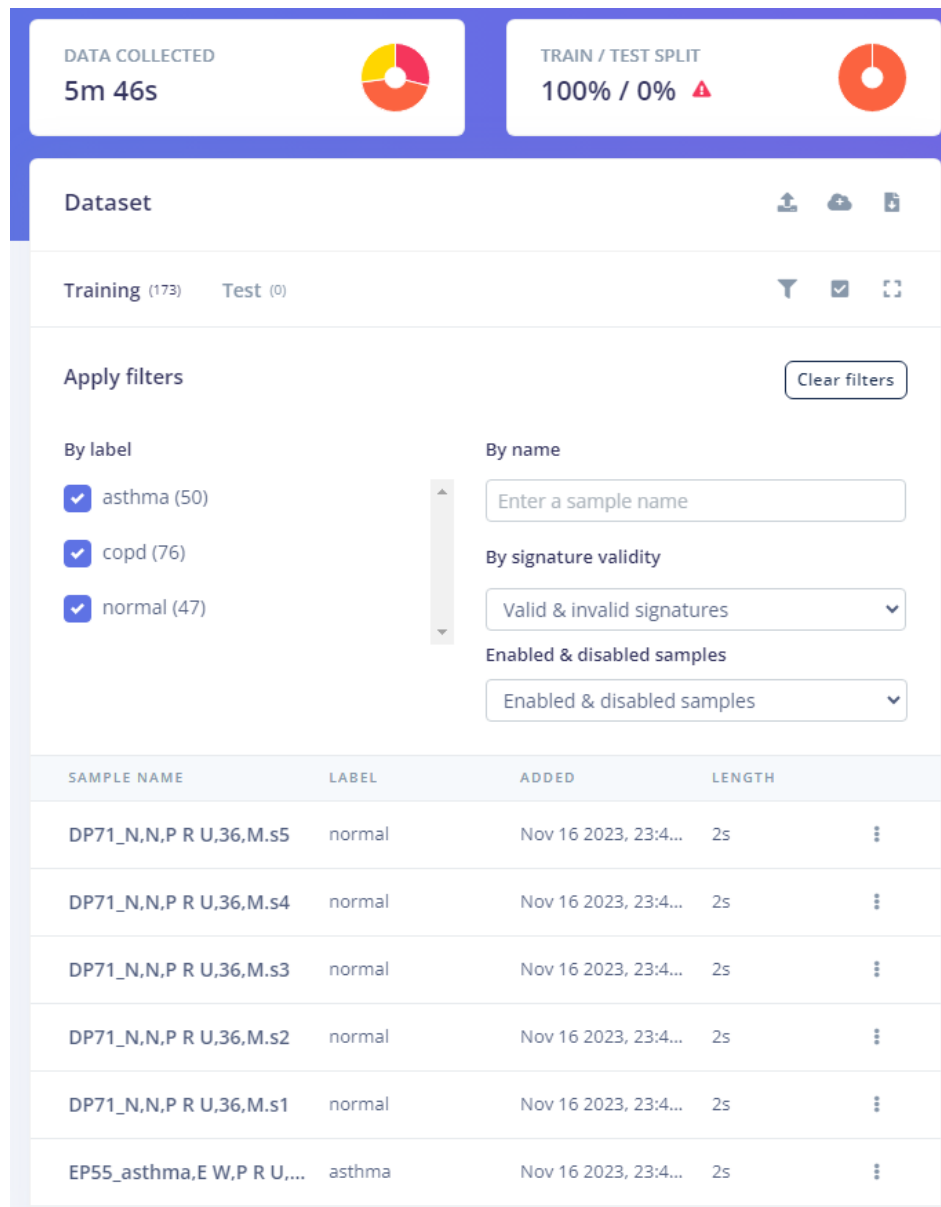


Рисунок 3.4 – Сформований набір навчальних даних за вибіркою PRU

Оскільки дані в подальшому будуть використовуватися для попередньої оцінки навчання моделі, розбиття даних на навчальний та тестовий набори не здійснювалось, всі наявні аудіодані будемо вважати навчальними (Training).

Після завантаження та попередньої підготовки навчальних даних, переходимо на сторінку проектування імпульсу (Impulse design). Встановимо розмір вікна на рівні 1000 мілісекунд, розмір збільшення вікна встановимо 100 мілісекунд (рис. 3.5).

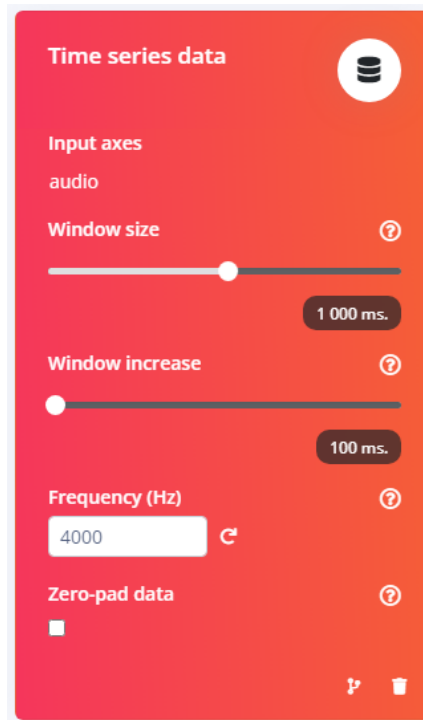


Рисунок 3.5 – Зображення вікна з налаштуванням розміру вікна даних

Це означає, що з кожного аудіофайла тривалістю 2 с програма буде брати для аналізу ознак вікно тривалістю 1 с, потім буде здійснювати зсув на 100 мс і брати інше вікно даних і так далі до кінця аудіофайлу.

На наступному етапі додаємо блок обробки даних MFCC і нейронну мережу як блок навчання (рис. 3.6).

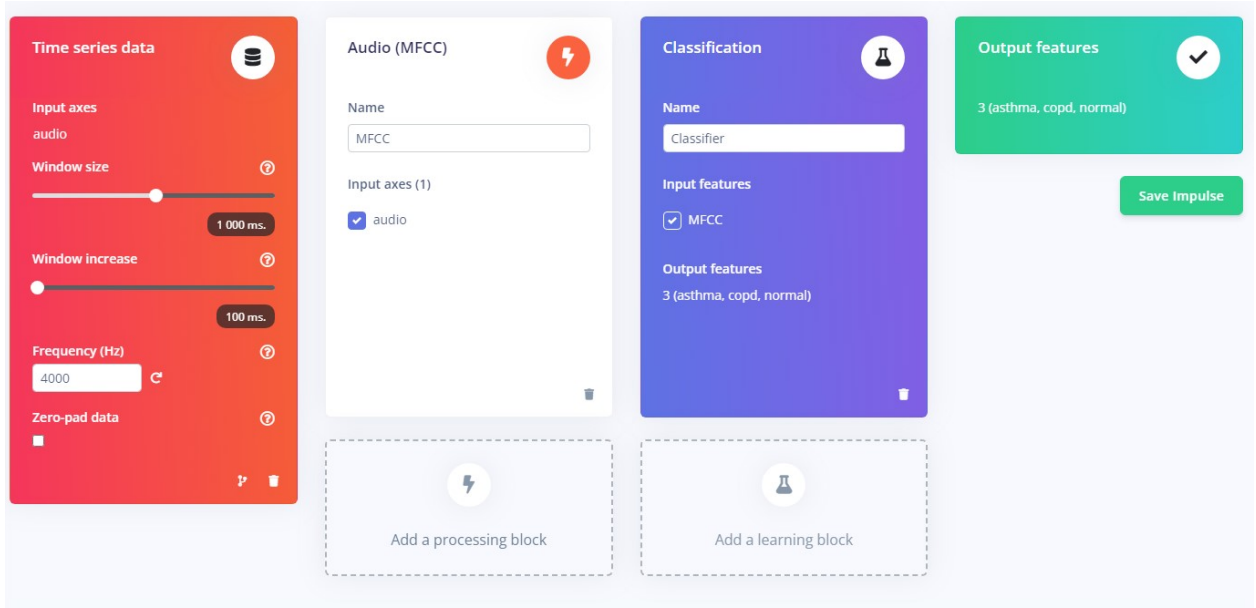


Рисунок 3.6 – Зображення вікна проектування імпульсу

Далі натискаємо «Зберегти імпульс» і переходимо на сторінку налаштувань MFCC. Розглянемо значення кожного з налаштувань (рис 3.7).

Mel Frequency Cepstral Coefficients

Number of coefficients ?	<input type="text" value="13"/>
Frame length ?	<input type="text" value="0.02"/>
Frame stride ?	<input type="text" value="0.02"/>
Filter number ?	<input type="text" value="32"/>
FFT length ?	<input type="text" value="256"/>
Normalization window size ?	<input type="text" value="0"/>
Low frequency ?	<input type="text" value="100"/>
High frequency ?	<input type="text" value="1500"/>

Pre-emphasis

Coefficient ?	<input type="text" value="0.98"/>
---------------	-----------------------------------

[Save parameters](#)

Рисунок 3.7 – Вікно налаштувань MFCC

Number of coefficients – налаштування, що ми використовуємо 13 частотних цепстральних коефіцієнтів Mel.

Frame length – розмір кожного вікна для розрахування MFCC в секундах. Становить 20 мілісекунд.

Frame stride – зсув вікна для розрахування MFCC в секундах. Становить 20 мілісекунд, тобто вікна не перекриваються.

Filter number – кількість фільтрів. Налаштування 32 означає, що ми будемо використовувати 32 фільтри, які дадуть розрахунок 32 MFCC. Далі ми відкинемо останні 19 значень коефіцієнтів Mel, оскільки вони нам не потрібні. І це дасть в результаті 13 MFCC для вікна.

FFT length – розмір ШПФ. Це визначає кількість діапазонів частот ШПФ, а також роздільну здатність частотних піків, які можна розділити. Менше число означає, що більше сигналів буде усереднено разом у тому самому діапазоні FFT, але також зменшує кількість ознак і розмір моделі. Більше число розділить більше сигналів в окремі контейнери, але генерує більшу модель. Встановимо це налаштування на 256.

Normalization window size – розмір ковзаючого вікна для нормалізації локального кепстрального середнього. Встановимо це значення на 0, що означає вимкнути локальну нормалізацію.

Low frequency, High frequency – можливість обмежити частоту обробки сигналу засобом MFCC. Може використовуватись для усунення будь-якої складової вище частоти Найквіста. Так як ми знаємо, що звуки легень вислуховуються на частотах 100-1500 Гц, ми можемо вказати ці значення як межі низької і високої частоти.

Pre-emphasis – фільтр попереднього підсилення, використовується для підсилення високих частот у сигналі перед обчисленням MFCC. Наразі ми залишимо це налаштування за замовчуванням.

Після встановлення і збереження налаштувань MFCC переходимо на вкладку Generate Features (розрахувати ознаки) і натискаємо Generate features (рис. 3.8).

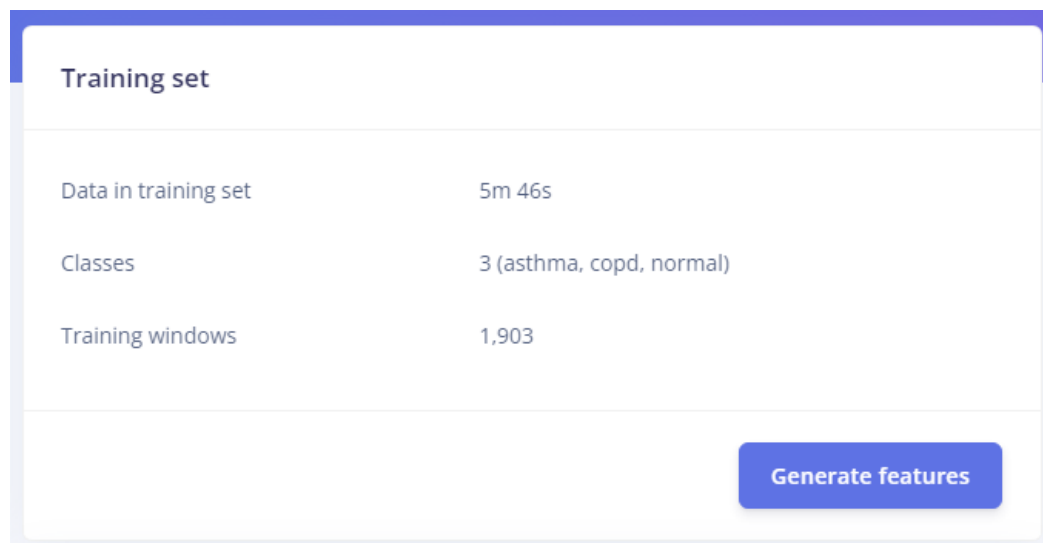


Рисунок 3.8 – Вікно розрахування ознак

Зачекаємо кілька хвилин, поки всі MFCC будуть обчислені. Коли це буде зроблено, можна побачити широке представлення даних у вікні «дослідницьких ознак». Оскільки візуалізувати 1 903 вимірів майже неможливо, Edge Impulse об'єднує MFCC кожного зразка, щоб створити 3D-графік. Для цього ПЗ виконує алгоритм UMAP для кожного набору MFCC, щоб зменшити розмірність до трьох. Це схоже на алгоритм кластеризації, який бере всі пікселі з зображення MFCC і намагається згрупувати їх разом на основі подібності. Спектрограми MFCC, які виглядають схожими, з'являються на цьому графіку у вигляді точок, розташованих близько одна до одної. А ті, що не виглядають схожими, стають точками, розташованими далі. Якщо візуально можна помітити деяке розмежування між групами, то це означає, що алгоритм машинного навчання також зможе їх відокремити (рис. 3.9).

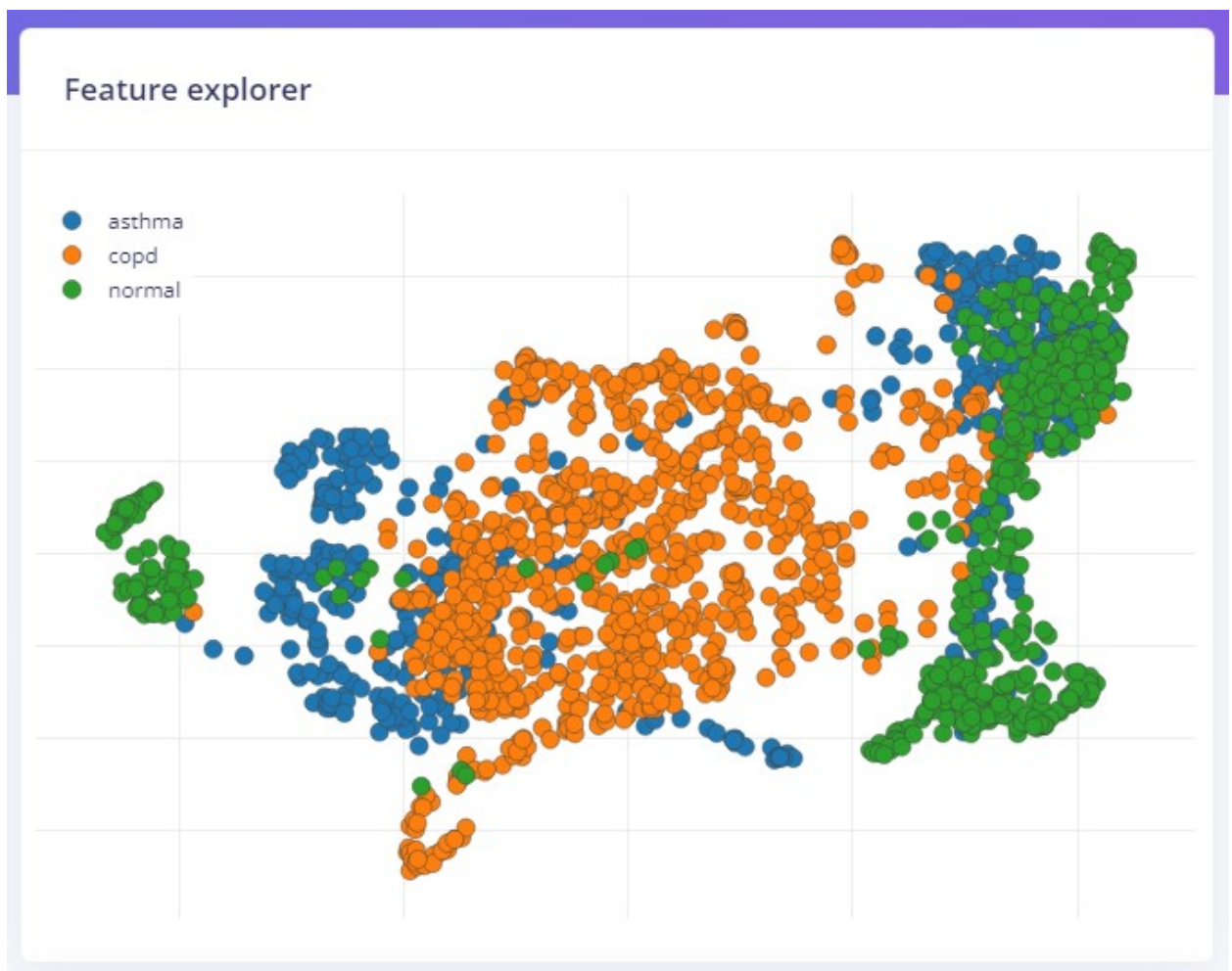


Рисунок 3.9 – Зображення вікна «дослідницьких ознак»

Спектрограми MFCC можна переглянути у блоці DSP result на сторінці з налаштуваннями MFCC (рис. 3.10).

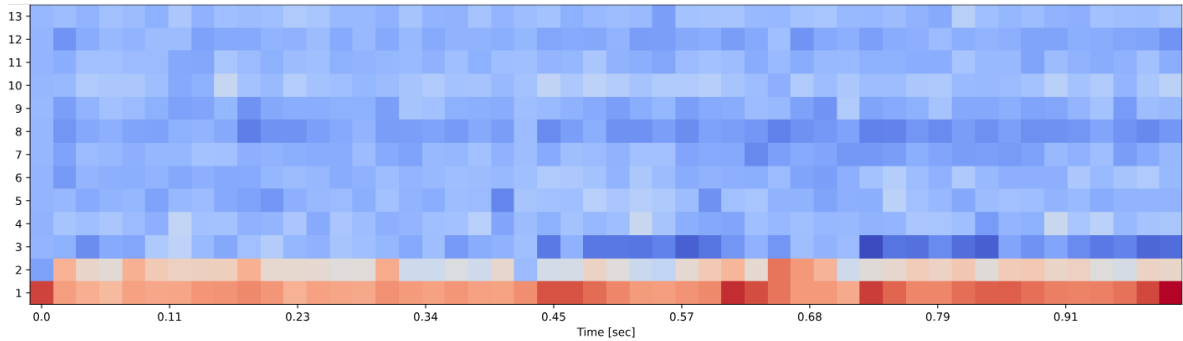


Рисунок 3.10 – Сформована спектрограма MFCC для одного з вікон даних

Герейдемо до навчання моделі. На сторінці налаштувань нейронної мережі залишимо налаштування за замовчуванням і розпочнемо навчання. Результати навчання представлені на рис. 3.11.

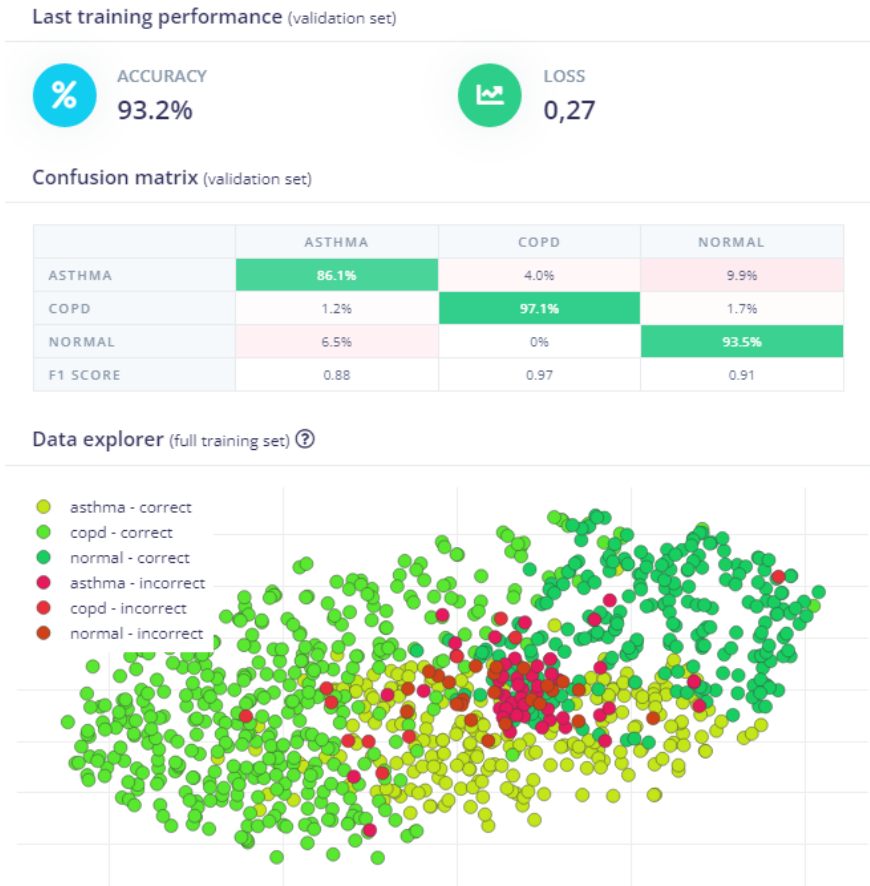


Рисунок 3.11 – Результати навчання моделі з класифікації звуків легень з використанням вибірки даних PRU

Бачимо, що модель демонструє здатність до класифікації шумів легень.

Для знаходження оптимальних налаштувань моделі, проведемо той самий експеримент, обравши в налаштуваннях зсуву вікна даних замість 100 мс величини 50 мс та 200 мс. Також отримаємо результати класифікації і для інших вибірок: PRU, PLM, PRM. Результати з точності розпізнавання зведено у таблицю 3.3.

Таблиця 3.3 – Результати роботи моделей машинного навчання з розпізнавання акустичних шумів легень

Вибірка даних	Кількість даних	Зсув вікна 50 ms	Зсув вікна 100 ms	Зсув вікна 200 ms
PRU	5m 46s	89,3%	93,2%	91,8%
PLU	8m 6s	86,3%	86,5%	80,5%
PRM	8m 58s	88,4%	86,5%	87,3%
PLM	7m 38s	91,6%	92,7%	83,6%

Бачимо, що найбільш ефективно нейрона мережа класифікує звуки легень при застосуванні зсуву вікна розміром 100 мс. Найбільшого відсотку успішних класифікацій було отримано у вибірках PRU та PLM. Загалом можна зробити висновок, що модель здатна класифікувати звуки легень на етапі навчання за всіма чотирма вибірками аудіоданих.

3.3 Реалізація моделі на вбудованому пристрої

Після того, як ми визначили оптимальний розмір налаштування зсуву вікна даних (100 мс), можемо приступити до реалізації моделі на вбудованому пристрої. Для цього будемо використовувати одну вибірку даних PRM, тому що після розділення файлів на частини тривалістю по 2 с вона мала найбільшу кількість зразків аудіоданих трьох класів (asthma – 75 файлів, cold – 76 файлів, normal – 118 файлів).

Для побудування моделі необхідно, щоб кількість даних у трьох класах була однаковою. Це викликано тим, що модель при навчанні може надавати

перевагу при класифікації тим даним, яких при навчанні було найбільше (може вважати їх більш важливими). Тому для усунення цього фактору необхідна однакова кількість навчальних даних у різних класах.

Також за методом Holdout 20% всіх даних не буде приймати участі у навчанні моделі а буде використовуватись лише при тестуванні моделі.

Створимо в Edge Impulse новий проект з назвою «Sound Classifier PRM (with test data)» (рис. 3.12).

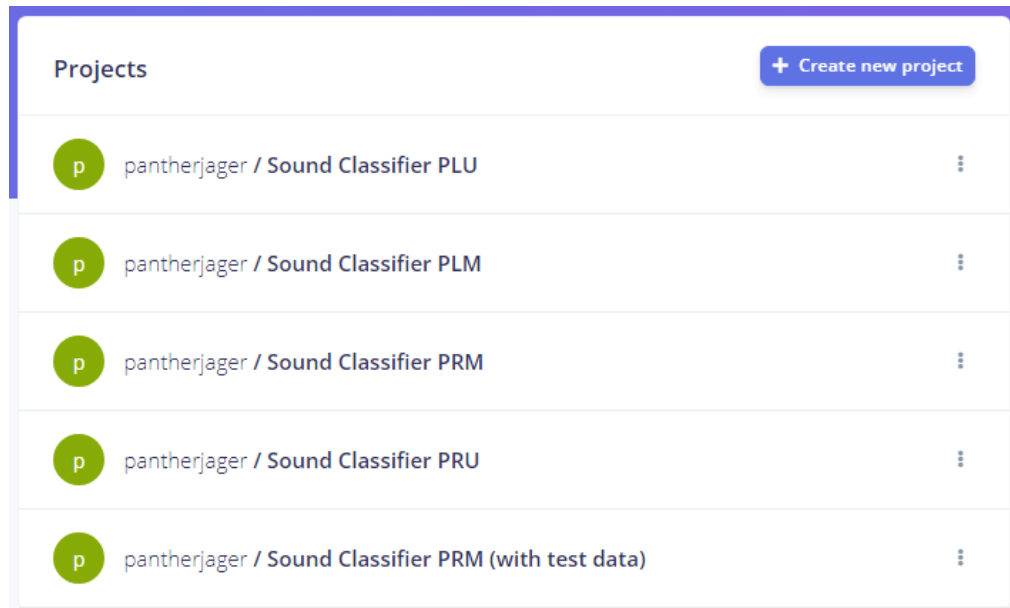


Рисунок 3.12 – Створення нового проекту «Sound Classifier PRM (with test data)» в Edge Impulse

Далі зробимо експорт аудіоданих з проекту «Sound Classifier PRM» і імпортуємо їх в новий створений проект. Оскільки нам необхідна однакова кількість даних у всіх категоріях класифікатора, ми випадковим чином видалимо надлишок аудіоданих, в результаті отримаємо більш гармонійну вибірку даних:

- asthma – 75 файлів;
- cord – 75 файлів;
- normal – 75 файлів.

Далі випадковим чином перенесемо 20% даних з кожної категорії до тестового набору даних, 80% даних залишимо у навчальному наборі даних. Таким чином, отримаємо:

- навчальний набір:
 - asthma – 60 файлів;
 - copd – 60 файлів;
 - normal – 60 файлів;
- тестовий набір:
 - asthma – 15 файлів;
 - copd – 15 файлів;
 - normal – 15 файлів.

Проведемо експеримент з навчання моделі на навчальному наборі даних і далі протестуємо модель на тестовому наборі даних, який не буде приймати участь у навчанні моделі.

За результатом експерименту, на етапі навчання точність класифікації моделі склала 90,7% (астма – 77,8%; ХОЗЛ – 95,5%; без патології – 97,8%). При застосуванні тестового набору даних модель показала гірший результат класифікації 67,47% (астма – 73,9%; ХОЗЛ – 73,9%; без патології – 54,5%).

Отже, з вірогідністю близько 0,7 модель може вірно класифікувати захворювання астма і ХОЗЛ. В подальшому можлива оптимізація параметрів моделі для досягнення більшої точності у класифікації.

Результати експерименту з класифікації аудіосигналів шумів легень на навчальному наборі даних представлено на рис. 3.13, на тестовому наборі даних – на рис. 3.14.

Далі отримана модель класифікації шумів легень може бути експортована із програмного забезпечення Edge Impulse у вигляді бібліотеки Arduino і завантажена на вбудований пристрій (в нашому випадку – MCU Arduino Nano 33 BLE Sense) для автономної роботи – без використання комп'ютеру.

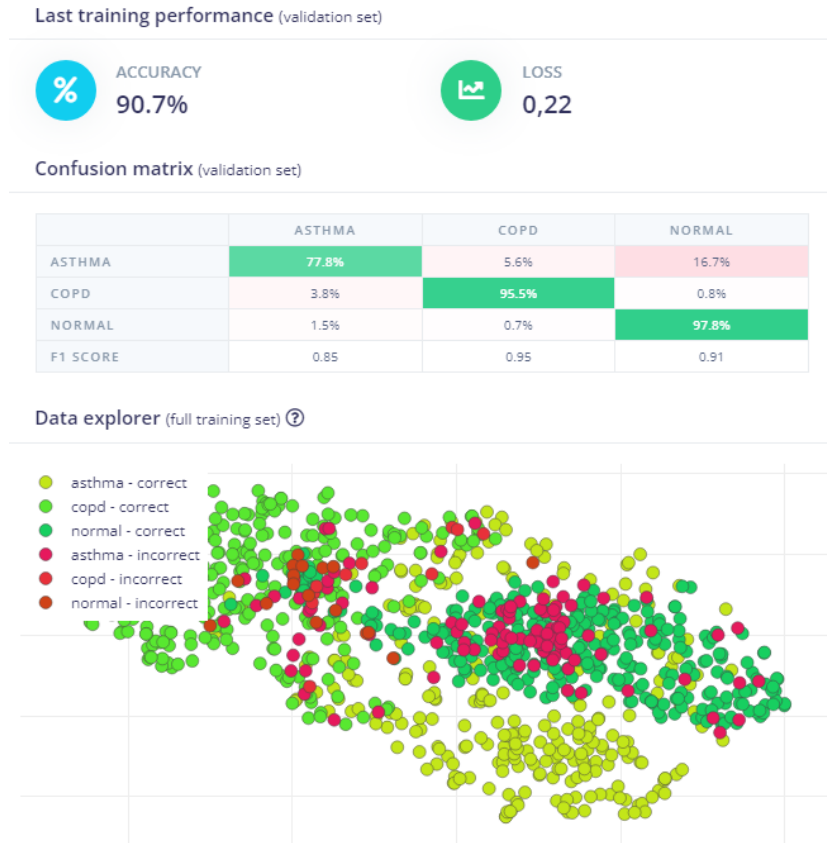


Рисунок 3.13 – Результати класифікації на навчальних даних

Model testing results

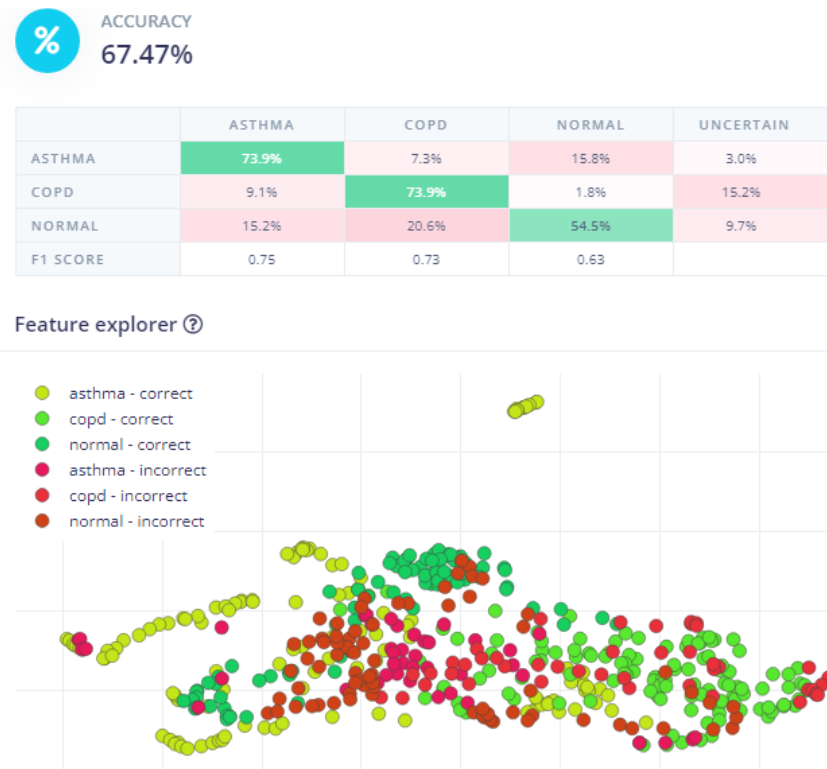


Рисунок 3.14 – Результати класифікації на тестових даних

3.4 Висновки до розділу 3

З двох джерел було сформовано чотири вибірки даних: PLM, PRM, PLU, PRU. Кожна з вибірок має аудіозаписи трьох діагнозів: астма (asthma), ХОЗЛ (copd), нормальні звуки легень (normal). Всі аудіозаписи зроблені на однаковому обладненні – за допомогою електронного стетоскопу 3M Littmann 3200 з однакових зон легень, мають однакову частоту дискретизації 4 000 Гц:

- PLM:asthma (12 файлів), normal (18 файлів), copd (11 файлів);
- PRM:asthma (18 файлів), normal (21 файл), copd (11 файлів);
- PLU:asthma (15 файлів), normal (15 файлів), copd (11 файлів);
- PRU:asthma (18 файлів), normal (12 файлів), copd (11 файлів).

Далі був проведений експеримент щодо виявлення розміру оптимального зсуву вікна даних. Файли чотирьох вибірок були засобами EdgeImpulse розбиті на частини тривалістю 2 секунди. В налаштуваннях розміру зсуву вікна даних обирались значення 50 мс, 100 мс, 200 мс і здійснювалось навчання моделі. За результатом експерименту виявлено, що оптимальний розмір зсуву вікна даних складає 100 мс, при цьому значенні був отриманий найбільший відсоток успішних класифікацій в чотирьох вибірках на етапі навчання (близько 90%).

В подальшому на базі однієї вибірки PRM було здійснено навчання і тестування моделі машинного навчання. Після розділення файлів на частини тривалістю по 2 с ця вибірка мала найбільшу кількість зразків аудіоданих трьох класів (asthma – 75 файлів, copd – 76 файлів, normal – 118 файлів). Ці дані спочатку були збалансовані – приведені до однієї кількості файлів у кожному класі (по 75 файлів). Далі за методом Holdout були сформовані навчальний набір даних (asthma – 60 файлів, copd – 60 файлів, normal – 60 файлів) і тестовий набір даних, до якого віднесено 20% усіх даних (asthma – 15 файлів, copd – 15 файлів, normal – 15 файлів).

За результатом експерименту, на етапі навчання точність класифікації моделі склала 90,7% (астма – 77,8%; ХОЗЛ – 95,5%; без патології – 97,8%).

При застосуванні тестового набору даних модель показала гірший, але прийнятний результат класифікації 67,47% (астма – 73,9%; ХОЗЛ – 73,9%; без патології – 54,5%). Можна зробити висновок, що отримана модель машинного навчання на вибірці PRM з вірогідністю близько 0,7 може вірно класифікувати захворювання астма і ХОЗЛ. Подальша оптимізація параметрів моделі для досягнення більшої точності у класифікації не здійснювалась.

Отриману модель можна завантажити на вбудований пристрій (Arduino Nano 33 BLE Sense) для побудування автономного пристрою класифікації шумів легень з вірогідністю 0,7.

4 АПАРАТНА РЕАЛІЗАЦІЯ ПРИСТРОЮ З ВИЗНАЧЕННЯ ПАТОЛОГІЙ ЛЕГЕНЬ ЗА АКУСТИЧНИМИ ШУМАМИ

4.1 Вибір складових компонентів пристрою

Для спрощення розробки в якості бази була обрана плата розробки Arduino Nano 33 BLE Sense. Ця плата вже має вбудований мікрофон та містить процесор ARM Cortex-M4, який підтримує виконання алгоритмів машинного навчання.

Для зняття звуків легень з грудної клітини знадобиться голівка і звукопровідна трубка від стетоскопу для прослуховування легень, наприклад, стетоскопу «GIMA» (рис. 4.1). Кінець звукопровідної трубки стетоскопу буде під'єднаний до мікрофону мікроконтролера.



Рисунок 4.1 – Стетоскоп для прослуховування легень «GIMA» (Італія)

Для забезпечення автономної роботи MCU Arduino Nano 33 BLE Sense (без підключення до ПК) необхідний засіб для виводу інформації – дисплей. Для цієї мети пропонується використовувати дисплей OLED 128×32 І²С. Характеристики дисплею буде вказано нижче.

При проектуванні системи розпізнавання і класифікації звукових сигналів програмне забезпечення для машинного навчання буде створено за допомогою інструменту Edge Impulse ідоповнено за допомогою середовища для програмування Arduino IDE. Після завантаження кінцевої прошивки у пристрій він буде працювати автономно, без підключення до ПК і виводити результати своєї роботи на дисплей.

4.2. Огляд технічних можливостей MCU Arduino Nano 33 BLE Sense

Для реалізації системи розпізнавання та класифікації звукових сигналів на мікроконтролері Arduino Nano 33 BLE Sense необхідні наступні компоненти:

- плата Arduino Nano 33 BLE Sense;
- кабель Micro USB для підключення плати Arduino до настільного пристрою;
- веб-редактор Arduino або Arduino IDE для програмування плати.

Ця плата поєднує в собі мікроконтролер з вбудованими сенсорами, тому її можна застосовувати у багатьох випадках без додаткового обладнання або зовнішніх пристроїв. Плата Arduino Nano 33 BLE Sense також досить невелика за габаритами і її можна використовувати як кінцевий мобільний пристрій (рис. 4.2).

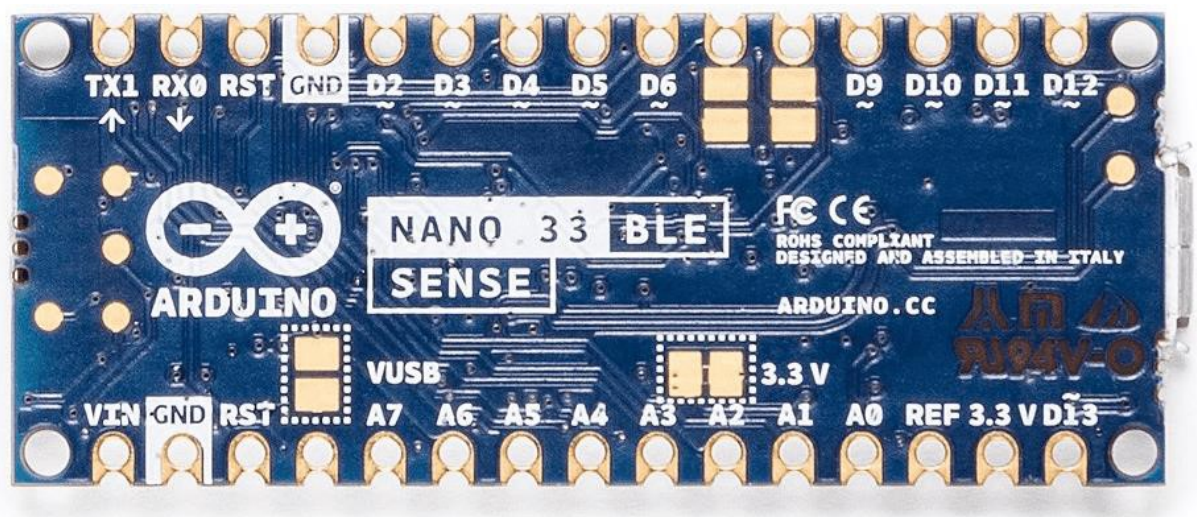


Рисунок 4.2 – Плата Arduino Nano 33 BLE Sense

Arduino Nano 33 BLE Sense – це плата Arduino з підтримкою моделей штучного інтелекту, яка живиться джерелом постійної напруги 3,3 В і доступна в найменшому форм-факторі 45*18 мм. Плата поставляється з серією вбудованих сенсорів:

- 9-ти осьовий інерційний сенсор – робить цю плату ідеальною для реалізації носимих пристроїв;
- сенсор вологості і температури – використовується для високоточних вимірювань умов навколишнього середовища;
- барометричний сенсор – забезпечує можливість створення простої метеостанції
- мікрофон – використовується для захоплення та аналізу звуку в режимі реального часу;
- сенсор жестів, наближення, кольору та інтенсивності світла – дозволяє оцінювати освітленість приміщення, а також визначати чи наближається хтось до місця, де встановлена плата.

Сему розташування сенсорів представлено на рис. 4.3.

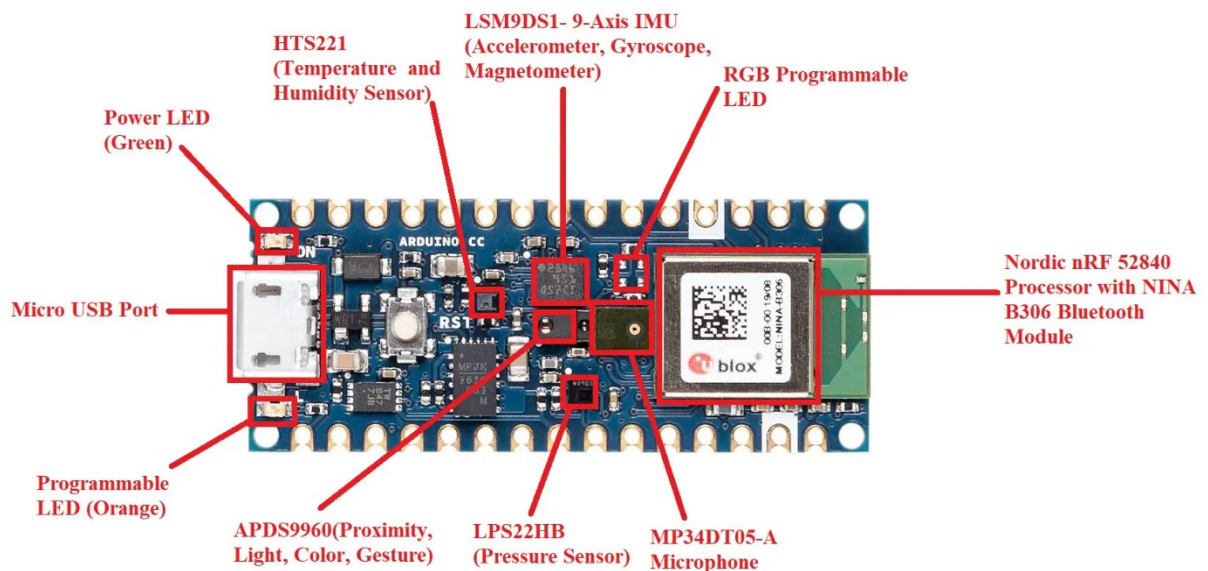


Рисунок 4.3 – Схема розташування сенсорів Arduino Nano 33 BLE Sense

Основною особливістю цієї плати є можливість запускати на ній програми Edge Computing (AI) за допомогою TinyML. Розробники можуть

створювати свої моделі машинного навчання за допомогою TensorFlow Lite і завантажувати їх на плату за допомогою Arduino IDE.

Arduino Nano 33 BLE Sense має потужний процесор nRF52840 від Nordic Semiconductors, 32-розрядний процесор ARM® Cortex®-M4, що працює на частоті 64 МГц. Це дозволяє створювати більші програми, ніж з Arduino Uno (він має 1 МБ програмної пам'яті, в 32 рази більше), і з набагато більшою кількістю пам'яті (ОЗП в 128 разів більше).

Основні технічні параметри [43] цього виду мікроконтролерів Arduino показано у таблиці 3.1.

Таблиця 4.1 – Технічні параметри Arduino Nano 33 BLE Sense

Компонент	Характеристика
Мікроконтролер	nRF52840
Робоча живлення	3.3В
Вхідна напруга (обмеження)	21В
Постійний струм на виводах вводу/виводу	15 мА
Тактова частота	64МГц
Об'єм флеш-пам'яті	1МБ (nRF52840)
Оперативна пам'ять	256КБ (nRF52840)
Постійна пам'ять	відсутня
Цифрові входи / виходи	14
PWM виходи	Всі цифрові виходи
UART	1
SPI	1
I2C	1
Аналогові входи	8 (ADC 12 bit 200 ks)
Аналогові виходи	Тільки через PWM
Зовнішні переривання	Всі цифрові виходи
LED BUILTIN	13
USB	Власний nRF52840
IMU	LSM9DS1
Мікрофон	MP34DT05
Жест, світло, близькість	APDS9960
Барометричний тиск	LPS22HB
Температура, вологість	HTS221

Організація пам'яті та регістрів представлено на рисунку 4.4.

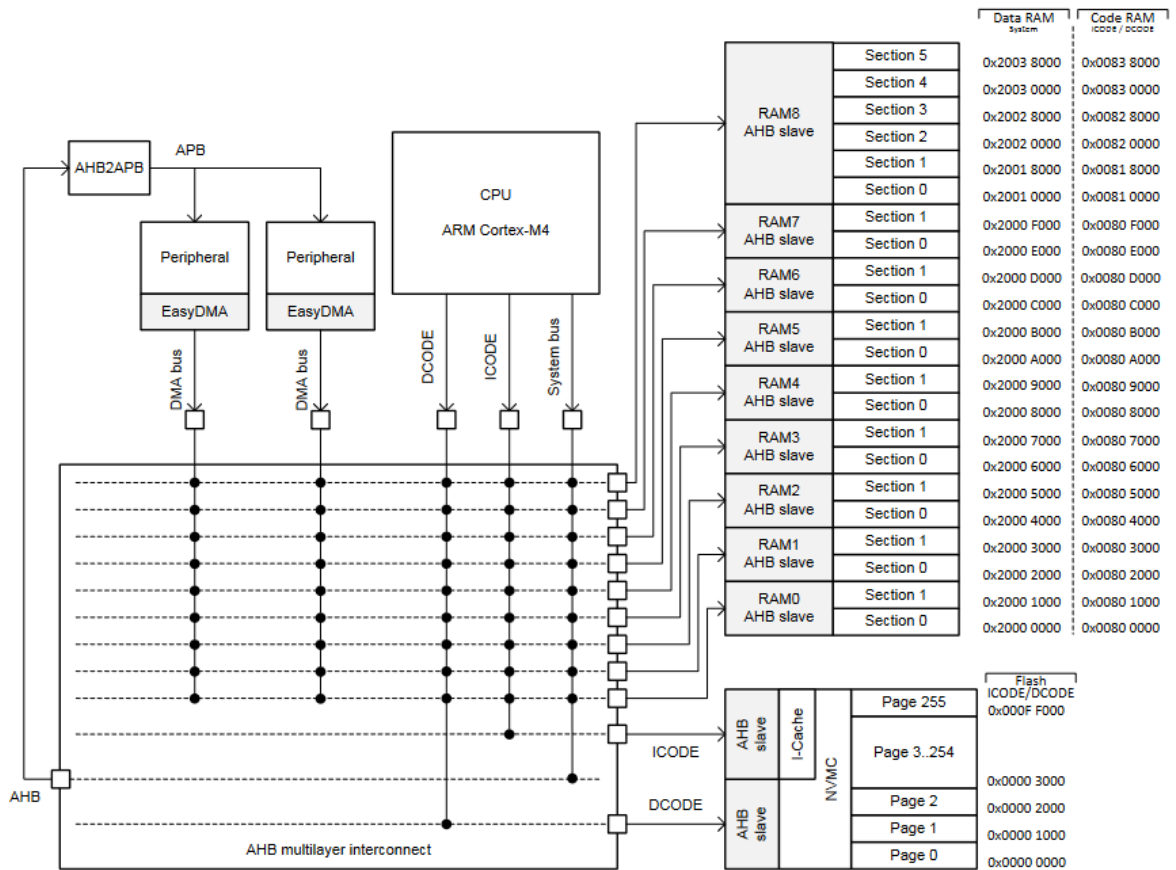


Рисунок 4.4 – Організація пам'яті процесора nRF52840

Процесор nRF52840 містить 1 МБ флеш-пам'яті та 256 КБ оперативної пам'яті, які можна використовувати для зберігання коду та даних.

Процесор і периферійні пристрої з EasyDMA можуть отримати доступ до пам'яті через багаторівневе з'єднання АНВ. ЦП також може отримати доступ до периферійних пристроїв через багаторівневе з'єднання АНВ.

Джерело живлення складається з ряду LDO і DC/DC стабілізаторів, які використовуються для максимізації енергоефективності системи.

Цей пристрій має такі характеристики джерела живлення:

- вбудовані регулятори LDO та DC/DC;
- режими ввімкнення/вимкнення глобальної системи;
- індивідуальне управління живленням секції RAM для всіх режимів системи;
- пробудження аналогового або цифрового виводу у вимкненій системі;

- обладнання диспетчера для керування скиданням увімкнення живлення, відключенням та збоєм живлення;
- автоматично керовані режими оновлення для регуляторів LDO та DC/DC для максимальної ефективності;
- зовнішнє живлення схем;
- окреме джерело USB.

Загальна структура та організація виводів Arduino Nano 33 BLE Sense показана на рисунку 4.5.

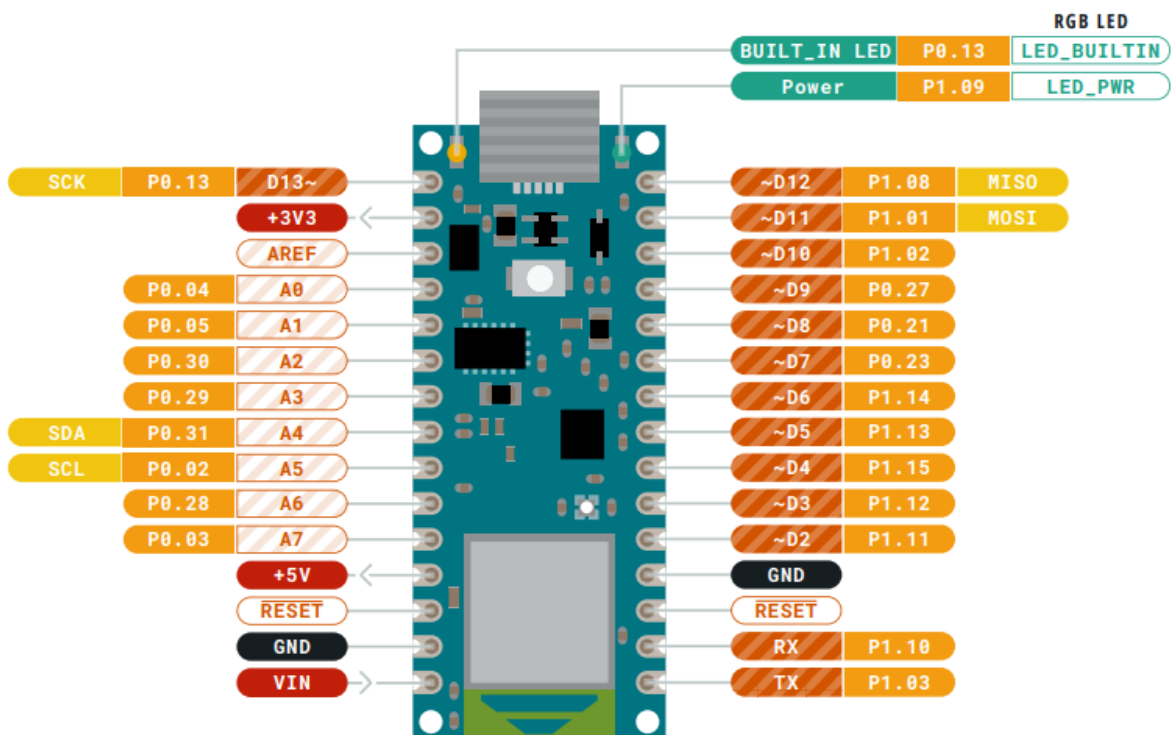


Рисунок 4.5 – Загальна структура та організація виводів Arduino Nano 33 BLE Sense

Більш детальна організація Arduino Nano 33 BLE Sense з виводами і сенсорами представлена у додатку В. Організація схеми електричної принципової представлена у додатку Г.

Для підключення Arduino Nano 33 BLE Sense к Edge Impulse необхідно почати з підключення його до комп'ютера, далі натиснути кнопку Reset двічі

поспіть, щоб перевести плату в режим завантажувача (почне імпульсувати жовтий світлодіод). Далі необхідно повернутись до Edge Impulse, перейти до «Пристрої» (Devices) та натиснути «Підключити новий пристрій» (Connect a new device).

Після виконання інструкцій і підключення плати до Edge Impulse можна використовувати пристрій Arduino Nano 33 BLE Sense для завантаження на нього для нашого проекту машинного навчання з розпізнавання аудіосигналів. Перевірити підключення можна на сторінці пристроїв у своєму проєкті Edge Impulse (рис 4.6).

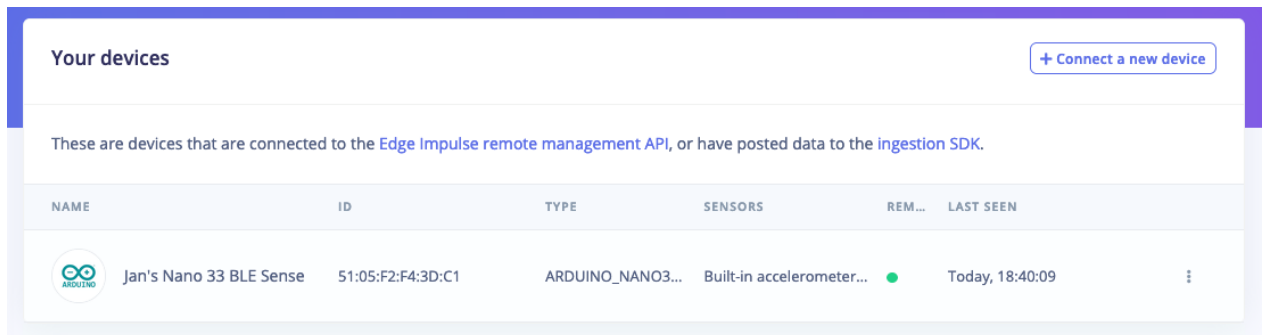


Рисунок 4.6 – Вікно Пристрої (Devices) після підключення плати розробника Arduino Nano 33 BLE Sense

4.3 Характеристики сенсора мікрофону MP34DT05-A

При організації системи розпізнавання та класифікації звукових сигналів важливу роль відіграє мікрофон. Тому доцільно розглянути його технічні характеристики та особливості застосування для того, щоб мати уявлення про обмеження та якість прийому звукового сигналу.

Arduino Nano 33BLE Sense має вбудований сенсор мікрофону MP34DT05-A, що представляє собою ультракомпактний, малопотужний, всеспрямований цифровий MEMS-мікрофон із ємнісним чутливим елементом та інтерфейсом I²C.

Чутливий елемент, здатний виявляти акустичні хвилі і виготовляється за допомогою спеціального процесу мікрообробки кремнію, призначеного для виробництва аудіо сенсорів. Інтерфейс I²C виготовляється з використанням CMOS, що дозволяє розробити спеціальну схему, здатну надавати цифровий сигнал ззовні у форматі PDM (модуляція щільності імпульсу).

Мікрофон синхронізується з тактовою частотою 1 МГц – 3 МГц і на лінії даних видає однобітний потік у вигляді імпульсів різної щільності. Лінія даних на виході може бути 0 або 1, а прямокутна хвиля створює щільність, яка при усередненні дає на виході аналогове значення (рис. 4.7).

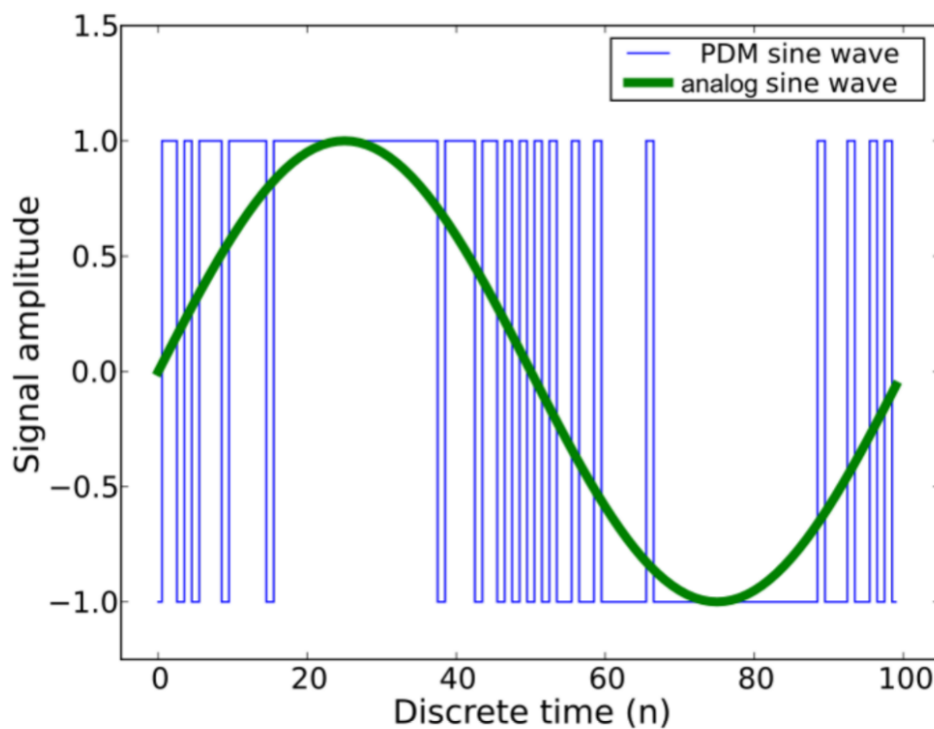


Рисунок 4.7 – Принцип дії MEMS-мікрофонів

Часта зміна логічних рівнів з однаковою шириною відповідає нульовій амплітуді сигналу. Якщо ширина імпульсів з логічною одиницею, що йдуть один за одним, більша за ширину імпульсу з логічним нулем, то можна говорити про позитивне значення амплітуди сигналу. Що більша ширина імпульсу з логічною одиницею, то більшим є значення позитивної амплітуди,

і навпаки, за від'ємного значення амплітуди переважає ширина імпульсу з логічним нулем.

MP34DT05-A – цифровий мікрофон з низьким рівнем спотворень із співвідношенням сигнал/шум 64 дБ і чутливістю $-26 \text{ DBFS} \pm 3 \text{ дБ}$. MP34DT05-A доступний у сумісному з SMD, захищеному від електромагнітних шумів корпусі та гарантовано працюватиме в розширеному діапазоні температур від $-40 \text{ }^\circ\text{C}$ до $+85 \text{ }^\circ\text{C}$ [44].

На рис. 4.8 показано зовнішній вигляд MP34DT05-A.

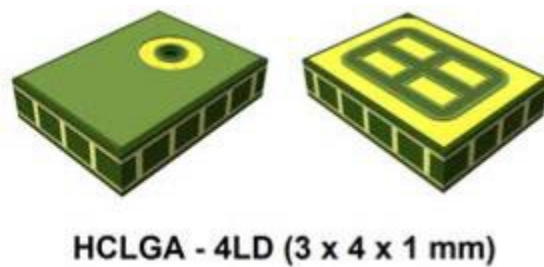


Рисунок 4.8 – Зовнішній вигляд MP34DT05-A

Чутливий елемент MP34DT05-A представляє собою акустичний датчик, що складається з провідної рухомої пластини та нерухомої пластини, розміщених у крихтному кремнієвому чіпі. Цей сенсор перетворює звуковий тиск у зміни пов'язаної ємності між цими двома пластинами.

Позначення та розташування виводів мікросхеми мікрофона показано на рисунку 4.9, а призначення кожного виводу – у таблиці 4.3.

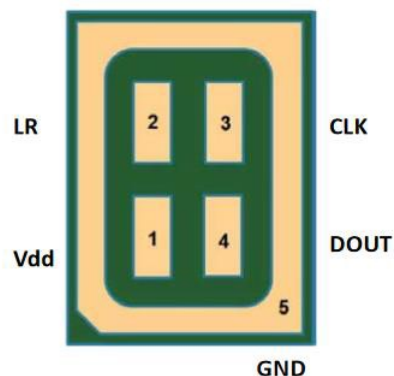


Рисунок 4.9 – Виводи MP34DT05-A

Таблиця 4.2 – Призначення і мітки виводів

№ виводу	Мітка виводу	Призначення
1	Vdd	Живлення
2	LR	Вибір лівого або правого каналу
3	CLK	Синхронізація входу
4	DOUT	Лівий/Правий канал виходу даних PDM
5	GND	Заземлення

Геометричні розміри та розташування основних компонентів MP34DT05-A представлено на рис. 4.10.

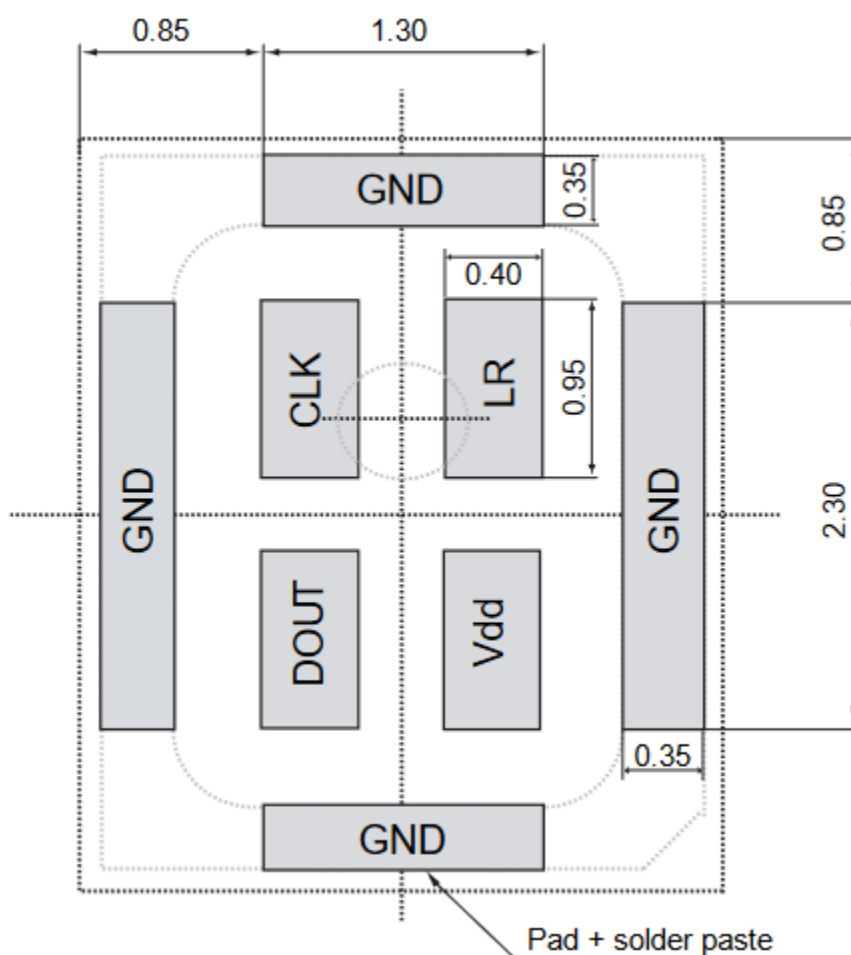


Рисунок 4.10 – Геометричні розміри (мм) та розташування компонентів мікрофона

АЧХ мікрофону представлена на рис. 4.11.

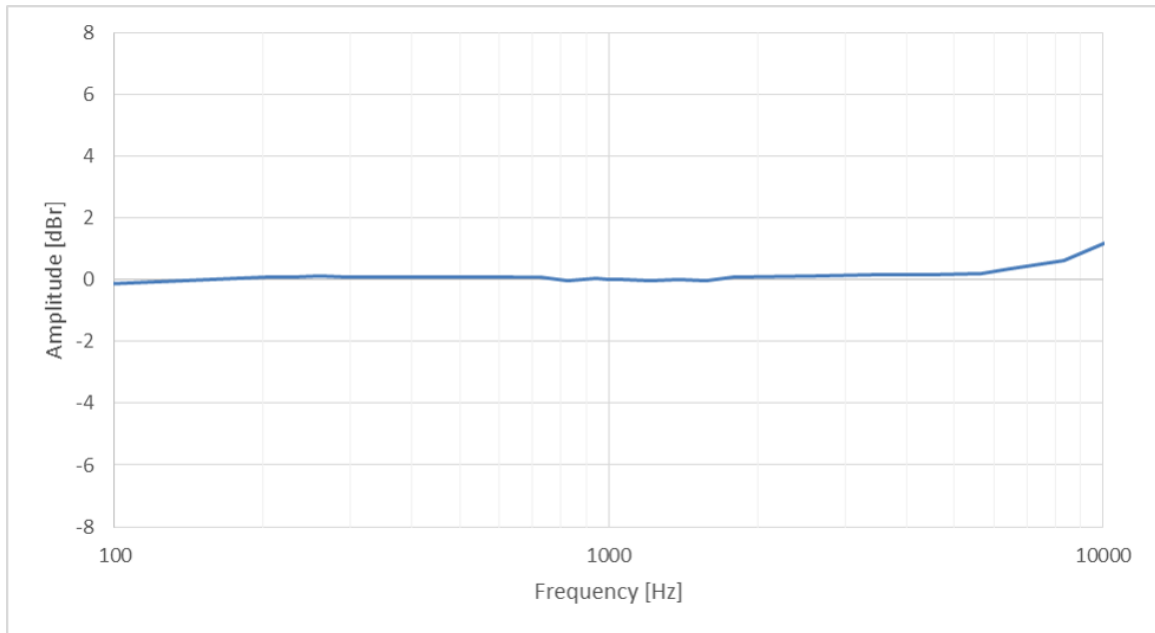


Рисунок 4.11 – АЧХ MEMS-мікрофону MP34DT05-A

Бачимо, що мікрофон має лінійну амплітудно-частотну характеристику на частотах 100 Гц – 5 000 Гц і може використовуватись для вимірів звуку у цьому діапазоні. Оскільки шуми легень фіксуються на частотах переважно від 100 до 1 500 Гц, цей мікрофон може застосовуватись для визначення шумів легень.

Загальні переваги MEMS-мікрофонів перед електретними мікрофонами:

- висока повторюваність параметрів під час виробництва;
- слабка залежність параметрів від температури;
- велике відношення сигнал/шум;
- дуже низька чутливість до ВЧ- та електромагнітних перешкод;
- менші розміри та хороша інтегрованість [45].

4.4 Характеристики дисплею OLED 128×32 I²C

Щоб бачити результати роботи мікроконтролера без підключення до ПК, необхідне додавання невеликого дисплея для відображення інформації.

На ринку є багато дуже маленьких дисплеїв на основі технології OLED (органічний світловипромінюючий діод). Це енергоефективні пристрої, вони дешеві, і порівняно з РК-дисплеями мають велику перевагу в тому, що їм не потрібне фонове освітлення. Також вони швидше оновлюють інформацію. OLED-дисплеї можуть працювати від різних драйверів. Найпопулярнішими є SH1106 і SSD1306.

Розглянемо підключення популярного монохромного графічного OLED-дисплею 128×32 пікселів. Дисплей використовує чіп SSD1306. Зв'язок здійснюється через інтерфейс I²C, тому крім живлення та GND потрібні лише два контакти. Також дисплей має компактні розміри і підходить для встановлення на макетній платі.

OLED-пристрій розміром 128×32 можна розглядати як матрицю з 4096 окремих пікселів, кожен з яких окремо може бути увімкнений або вимкнений контролером за допомогою програмних інструкцій. Оскільки фон дисплею чорний, увімкнені пікселі виділяються досить гарно, а читабельність ідеальна. Однією з приємних особливостей конфігурації масиву пікселів є те, що можна виводити не тільки текст, а й відображати графічні растрові зображення.

Зовнішній вигляд дисплею представлений на рис. 4.12, розміри друкованої плати – на рис. 4.13.

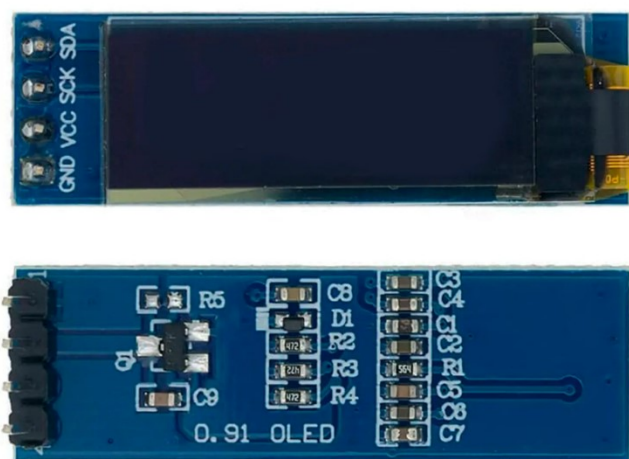


Рисунок 4.12 – Зовнішній вигляд OLED дисплею 128×32 I²C

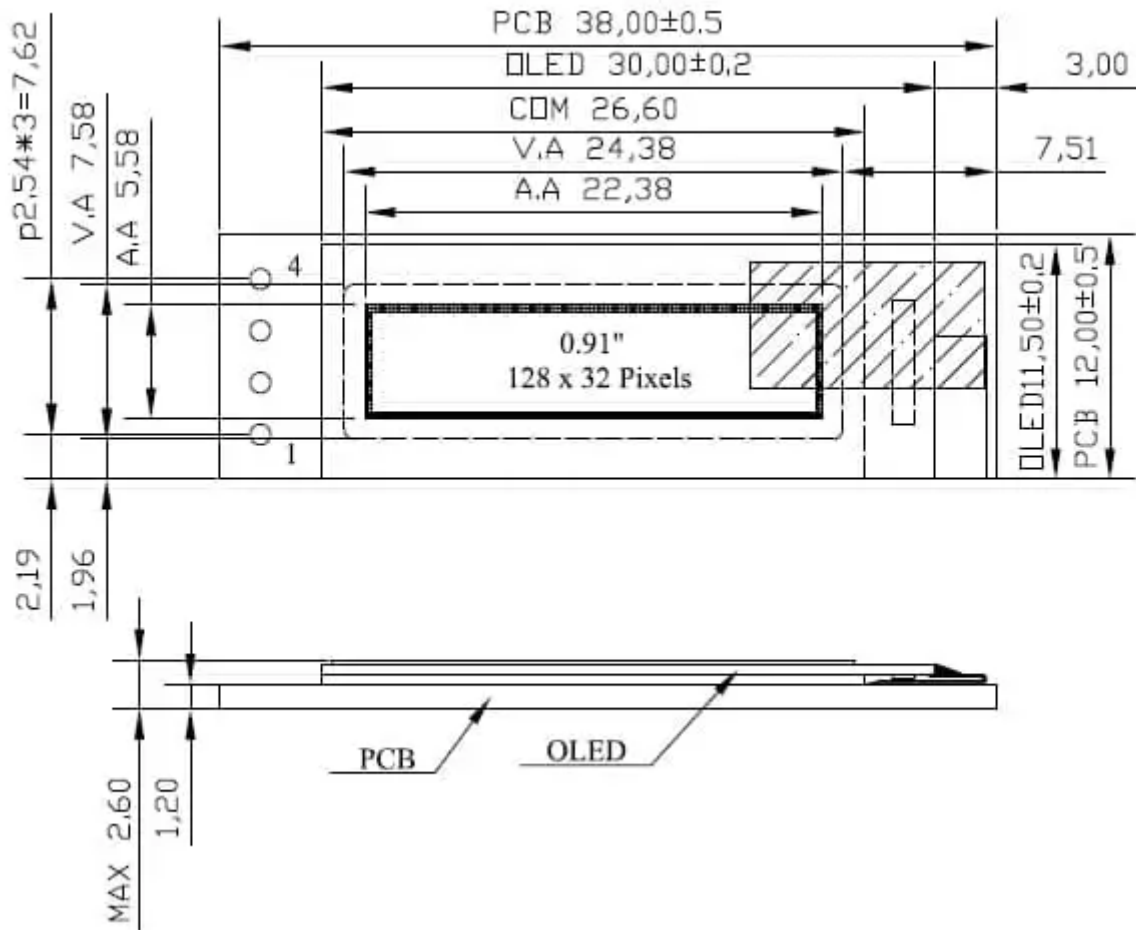


Рисунок 4.13 – Розміри друкованої плати OLED-дисплею 128×32 I²C

OLED-дисплей може працювати від напруги 5 В чи 3,3 В, вхідні сигнали перетворюються на 3,3 В за допомогою вбудованого перетворювача напруги. Це означає, що OLED-дисплей можна підключити до плати мікроконтролера Arduino Nano 33 BLE Sense без необхідності додаткових резисторів.

Для керування OLED-дисплеєм потрібні бібліотеки `adafruit_SSD1306.h` і `adafruit_GFX.h`, актуальну версію яких можна зручно завантажити через програмне забезпечення Arduino IDE.

Дисплей підключається до контактів мікроконтролера через послідовну шину даних I²C: SDA (A4), SCL (A5), GND (GND) і VCC (3v3) відповідно. Функціональна схема підключення на макетній платі наведена на рис. 4.14, схема електрична принципова наведена на рис. 4.15.

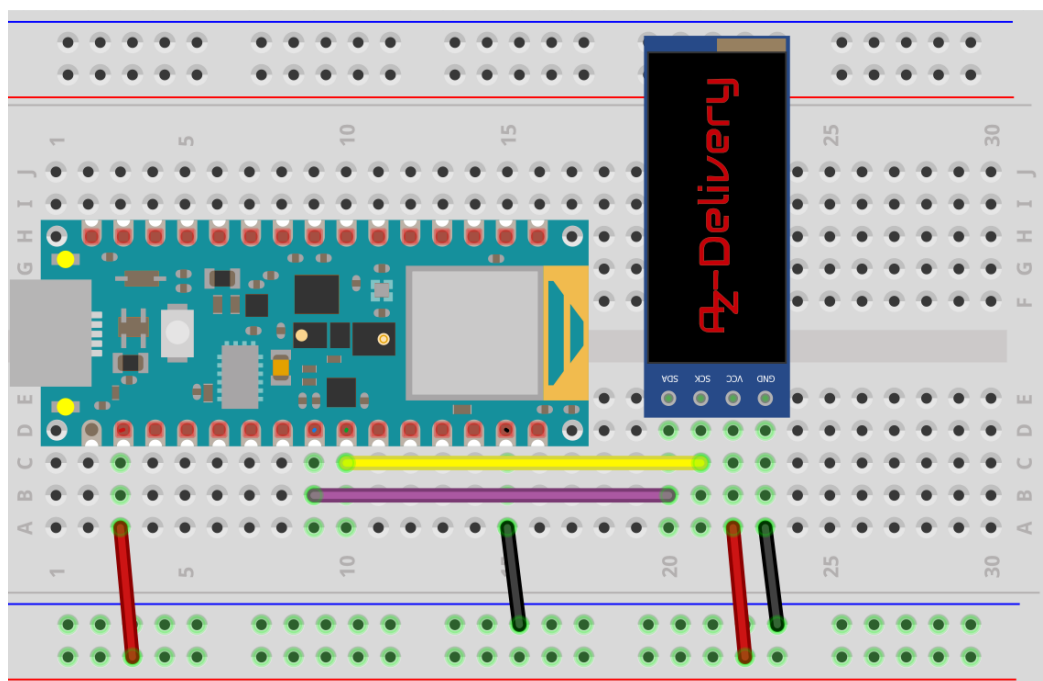


Рисунок 4.14 – Функціональна схема підключення OLED дисплею 128×32 I²C до мікроконтролера Arduino Nano 33 BLE Sense

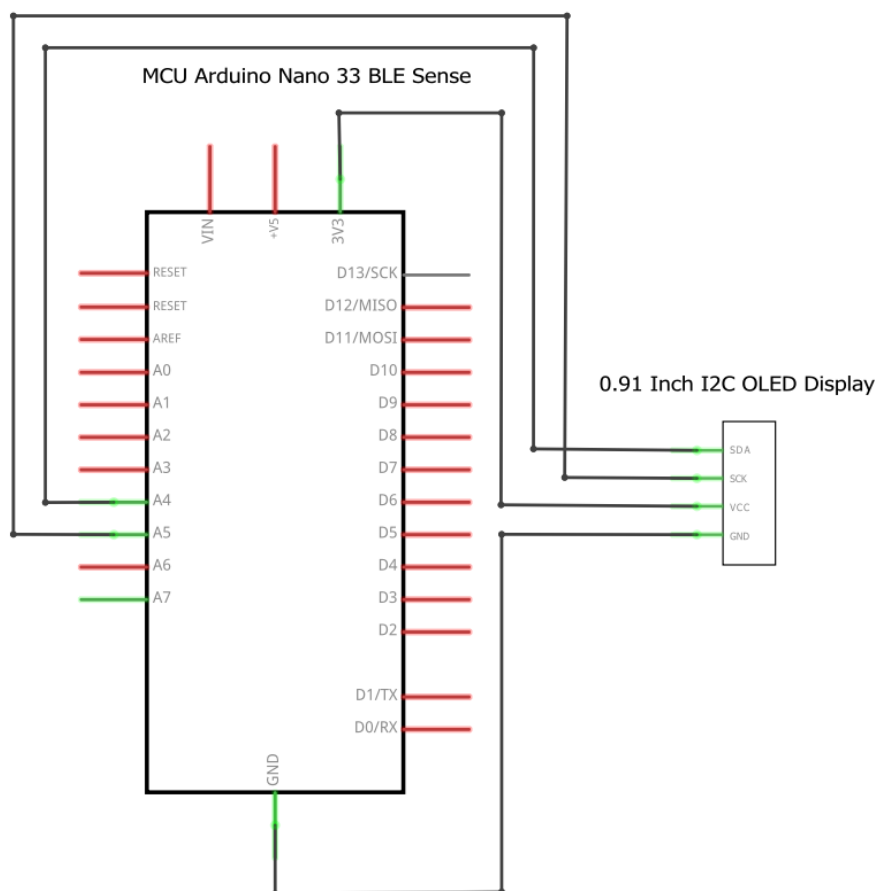


Рисунок 4.15 – Схема електрична принципова підключення OLED дисплею 128×32 I²C до мікроконтролера Arduino Nano 33 BLE Sense

4.5 Апаратна реалізація пристрою для класифікації акустичних шумів легень

Після розгляду технічних характеристик компонентів пристрою і обґрунтування їх вибору, з'єднаємо компоненти разом згідно рис. 4.16.

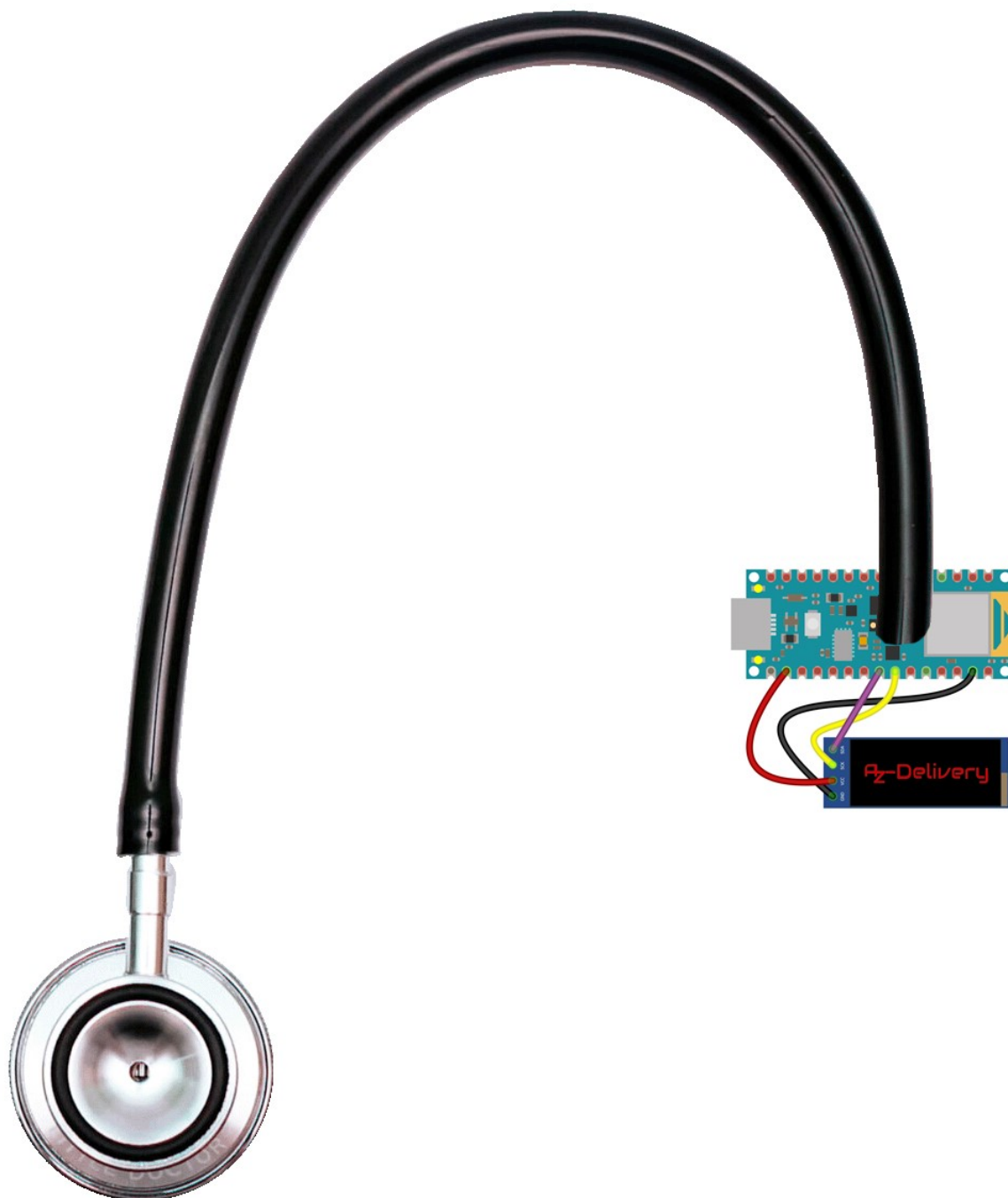
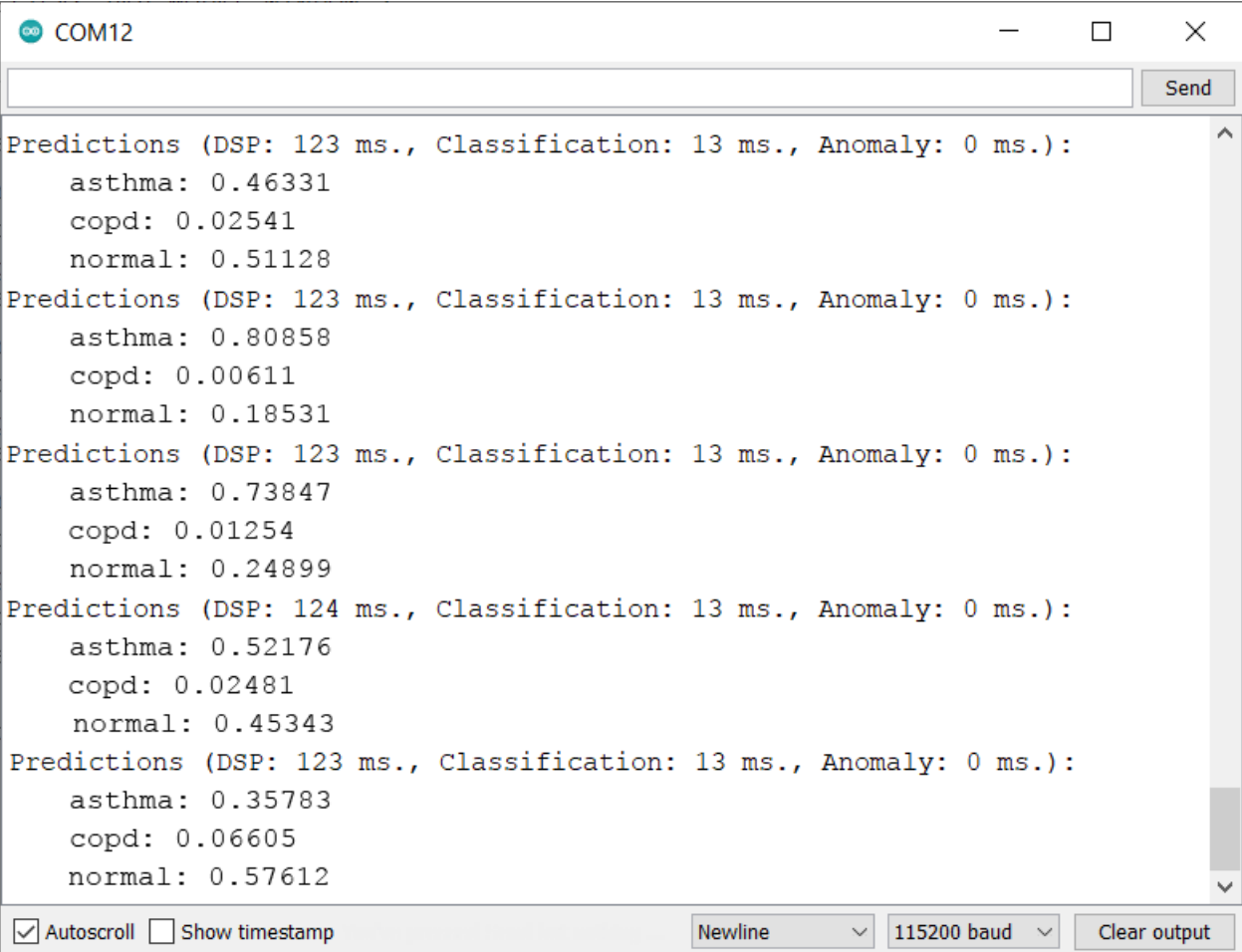


Рисунок 4.16 – Апаратна реалізація пристрою для класифікації шумів легень

Під'єднаємо головку стетоскопу із звукопровідною трубкою до MEMS-мікрофону MP34DT05-A, розташованому на платі Arduino Nano 33 BLE Sense. Далі під'єднаємо OLED-дисплей 128×32 I²C до мікроконтролера. Після цього апаратну реалізацію пристрою для отримання і розпізнавання шумів легень можна вважати завершеною.

Після включення мікроконтролера, він у «суперциклі» loop буде чекати аудіоданих з мікрофону. Далі завдяки вбудованій моделі машинного навчання відбуватиметься класифікація аудіосигналу на 3 категорії (астма, ХОЗЛ, без патології) і вивід результатів класифікації у серійний порт у вигляді вірогідностей віднесення аудіосигналу до кожного з трьох класів. На цьому етапі можна побачити результати класифікації на комп'ютері, якщо підключити до нього мікроконтролер (рис. 4.17). Алгоритм роботи програми класифікації шумів легень представлений на рис. 4.18.



```

COM12
Send
Predictions (DSP: 123 ms., Classification: 13 ms., Anomaly: 0 ms.):
  asthma: 0.46331
  copd: 0.02541
  normal: 0.51128
Predictions (DSP: 123 ms., Classification: 13 ms., Anomaly: 0 ms.):
  asthma: 0.80858
  copd: 0.00611
  normal: 0.18531
Predictions (DSP: 123 ms., Classification: 13 ms., Anomaly: 0 ms.):
  asthma: 0.73847
  copd: 0.01254
  normal: 0.24899
Predictions (DSP: 124 ms., Classification: 13 ms., Anomaly: 0 ms.):
  asthma: 0.52176
  copd: 0.02481
  normal: 0.45343
Predictions (DSP: 123 ms., Classification: 13 ms., Anomaly: 0 ms.):
  asthma: 0.35783
  copd: 0.06605
  normal: 0.57612
 Autoscroll  Show timestamp Newline 115200 baud Clear output

```

Рисунок 4.17 – Вивід результатів класифікації у серійний порт

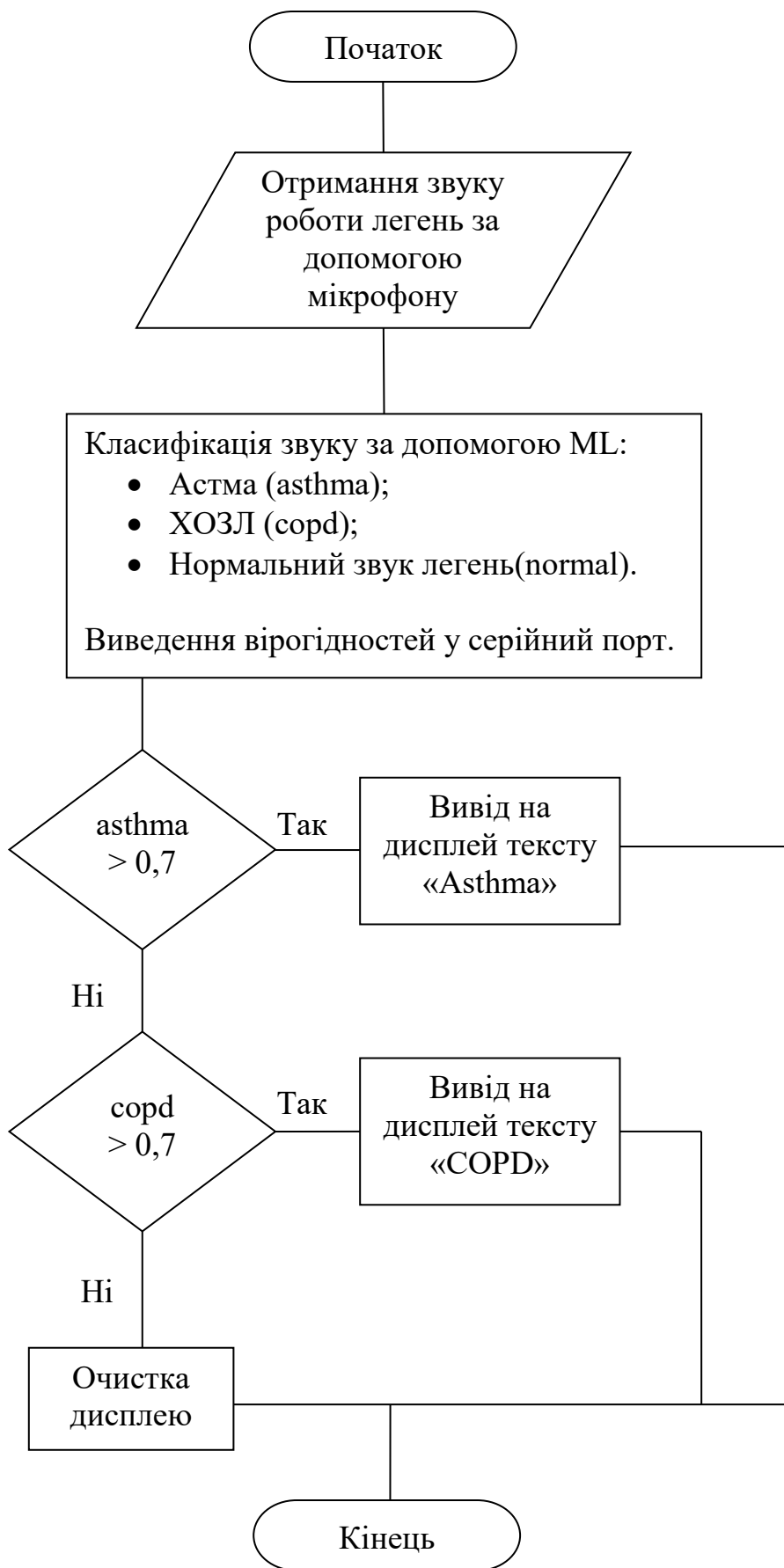


Рисунок 4.18 – Алгоритм роботи програми класифікації звуків легень

Після отримання результатів класифікації здійснюється їх вивід на OLED-дисплей 128×32. Якщо вірогідність класифікації астма або ХОЗЛ більше 0,7 – на екран виводиться повідомлення «Asthma» або «COPD» відповідно. В інших випадках на екран нічого не виводиться. Таким чином, якщо пристрій виявить патологічний звук легень із заданою мінімальною вірогідністю 0,7 – він повідомить про це користувача.

Розглянемо докладніше процес інтеграції програмного коду на мікроконтролер.

В Edge Impulse перейдемо на сторінку розгортання та натиснимо опцію бібліотеки Arduino. Проект буде створено та завантажено на комп'ютер у вигляді бібліотеки Arduino з прикладами її використання. Далі відкриємо програмне забезпечення Arduino IDE і додамо завантажену з Edge Impulse бібліотеку.

Відкриємо приклад «nano_ble33_sense_microphone_continuous.ino», який постачається разом із бібліотекою. Даний код містить приклад для безперервного отримання аудіоданих через вбудований мікрофон мікроконтролера Arduino Nano 33 BLE Sense і далі їх обробки моделлю машинного навчання. Розглянемо, як працює цей приклад коду.

Функція `run_classifier_init()` – функція ініціалізації Edge Impulse, яка встановлює всі необхідні функції MFCC і нейроної мережі.

Далі йде цикл, в якому програма чекає, поки заповниться буфер мікрофону – функція `microphone_inference_record()`. Плата використовує апаратну модуляцію щільності імпульсів або контролер PDM, щоб отримати аудіодані, цей функціонал застосовується викликом бібліотеки Arduino PDM. Створюється подвійний буфер таким чином, що коли один буфер заповнюється, інший використовується для виводу даних.

Далі код викликає функцію `run_classifier_continuous()`, яка обчислює MFCC з захопленого аудіо та запускає вивід за допомогою нейроної мережі.

`result.classification[].label` – вивід з масиву назви категорії класифікації.

`result.classification[].value` – вивід з масиву значення класифікації.

Назви категорій і значень зберігаються у масиві в такому порядку:

- [0] – asthma (Астма);
- [1] – copd (ХОЗЛ);
- [2] – normal (без патології);

Знаючи це, ми можемо вдосконалити код прикладу, додавши до нього вивід на дисплей результатів класифікації.

Підключимо кілька бібліотек, необхідних для роботи дисплею:

```
#include <SPI.h>
#include <Wire.h>
#include <Adafruit_GFX.h>
#include <Adafruit_SSD1306.h>
```

Далі додаємо команди для підключення дисплею:

```
#define SCREEN_WIDTH 128 // OLED display width, in pixels
#define SCREEN_HEIGHT 32 // OLED display height, in pixels
#define OLED_RESET -1 // Reset pin
#define SCREEN_ADDRESS 0x3C /// Address for 128x32
Adafruit_SSD1306 display(SCREEN_WIDTH, SCREEN_HEIGHT,
&Wire, OLED_RESET);
```

Для зручності створюємо 3 змінні для збереження вірогідностей класифікації:

```
float asthma, normal, noise;
```

Для зручності виводу на дисплей додаємо нову функцію, результатом якої буде вивід тексту на дисплей:

```
void displayText(String text, int x, int y, int size, boolean d) {
    display.setTextSize(size); // розмір шрифту
    display.setTextColor(WHITE); // колір шрифту
    display.setCursor(x,y); // початкові координати тексту
    display.println(text); // вивід тексту
    if(d){
        display.display(); // вивід інформації на дисплей
    }
```

```

    }
    // delay(100);
}

```

До функції налаштувань мікроконтролера `setup()` додаємо код ініціалізації дисплею (з прикладу, що надається разом із бібліотекою «Adafruit_SSD1306»):

```

if(!display.begin(SSD1306_SWITCHCAPVCC, SCREEN_ADDRESS)) {
    Serial.println(F("SSD1306 allocation failed"));
    for(;;);
}

```

До функції `loop()` мікроконтролера додаємо код для виводу результатів класифікації на OLED-дисплей:

```

asthma = result.classification[0].value;
copd = result.classification[1].value;
normal = result.classification[2].value;
display.clearDisplay(); // очистка дисплею
if (asthma > 0.70) {
    displayText("Asthma", 0, 0, 2, true);
    delay(1000);
}
if (copd > 0.70) {
    displayText("COPD", 0, 0, 2, true);
    delay(1000);
}

```

Далі зберігаємо файл «`nano_ble33_sense_microphone_continuous.ino`» і здійснюємо його компіляцію і завантаження до мікроконтролеру за допомогою Arduino IDE. Лістинг коду файлу представлений у додатку Д.

4.6 Висновки до розділу 4

Розглянуті складові компоненти пристрою для автоматичної класифікації шумів легень на базі мікроконтролера Arduino Nano 33 BLE Sense. Спроектовано пристрій, що являє собою приклад застосування машинного навчання на вбудованій системі.

Пристрій з вірогідністю 0,7 здійснює виявлення патологічних звуків легень – ХОЗЛ або астми, та здійснює вивід на ЖК дисплей повідомлення у разі такого виявлення. Пристрій працює автономно без підключення до ПК, модель машинного навчання працює як вбудована у пристрій. Живлення пристрою – від зовнішнього павербанку.

Вартість компонентів розробленого пристрою узагальнена в табл. 4.3:

Таблиця 4.3 – Вартість компонентів пристрою з розпізнавання шумів легень на базі мікроконтролера Arduino Nano 33 BLE Sense

Назва компоненту	Вартість, дол. США
Мікроконтролер Arduino Nano 33 BLE Sense	\$40,5
Стетоскоп для прослуховування легень «GIMA»	\$22,5
Дисплей OLED 128×32 I ² C	\$3,5

Загальна вартість пристрою складає близько \$50. Зважаючи на те, що вартість сучасного електронного стетоскопа складає в середньому від \$230 і дорожче, застосування MCU Arduino Nano 33 BLE Sense а також додаткових компонентів для побудови пристрою з розпізнавання шумів легень є як технічно, так і економічно обґрунтованим.

Модель машинного навчання, що здійснює класифікацію шумів легень на нормальні і патологічні, була створена в інструменті Edge Impulse. Далі за допомогою Arduino IDE були зроблені зміни в коді для виводу результатів класифікації на зовнішній дисплей мікроконтролера і модель була інтегрована у мікроконтролер для автономності роботи пристрою.

ВИСНОВКИ

Для досягнення мети – підвищення якості діагностики органів дихання, основним завданням було отримання працюючого зразка пристрою для визначення патологій легень за акустичними шумами, який за технічними характеристиками був би наближений до іноземних комерційних зразків, але при цьому мав нижчу вартість.

В першому розділі роботи визначено, що метод аускультатії із застосуванням звичайного стетоскопу має суб'єктивні ознаки, оскільки лікар вислуховує шуми легень покладаючись на свій власний слух. Для більшої об'єктивізації методу аускультатії легень перспективним є напрямок апаратної аускультатії за допомогою електронних стетоскопів і подібних пристроїв. Розробка пристрою для електронної аускультатії легень залишається актуальним технічним завданням.

Здійснено аналіз науково-технічної літератури щодо існуючих автоматизованих засобів діагностики стану легень за акустичними шумами. В результаті аналізу була узагальнена функціональна схема роботи такого пристрою. Було з'ясовано, що в середньому вартість сучасного електронного стетоскопа складає від \$230 і дорожче, тому доцільно шукати шляхи зменшення вартості подібних пристроїв.

Також було з'ясовано, що існуючі цифрові методи аналізу звуків легень, такі як аналіз сигналів у частотній і частотно-часовій області, не є універсальними і часто залежать від якості реєструючих засобів та умов реєстрації. Тому актуальним завданням є розробка методу, який мало залежить від умов реєстрації сигналу. Зроблено висновок, що таким методом може виступати машинне навчання на вбудованих пристроях.

У другому розділі роботи розглянутий метод машинного навчання його застосування для вирішення задачі класифікації аудіосигналів. Розглянуто програмне забезпечення EdgeImpulse, яке дозволяє розробляти проекти з використанням машинного навчання на вбудованих системах.

Докладно розглянуті питання отримання ознак з аудіоданих у вигляді MFCCі застосування згорткових нейронних мереж у навчанні моделі класифікації аудіосигналів. З'ясована характеристика аудіофайлів із записом звуків легень, які можуть бути використані для навчання моделі. За умови, що максимальна частота звуку легень, що досліджується, складає 1 500 Гц, аудіофайли повинні бути записані із частота дискретизації не менше 3 000 Гц.

В якості методу оцінки ефективності класифікатора запропоновано використовувати метод Holdout, в якому 20% всіх даних відкладаються у тестовий набір даних і не використовуються на етапі навчання моделі.

Третій розділ роботи присвячений формуванню бази шумів легень для проведення експерименту з розпізнавання патологічних шумів. Зроблено акцент на тому, що дані повинні бути репрезентативними та збалансованими. Для покращення репрезентативності запропоновано використовувати аудіозаписи шумів легень, отримані з однієї області грудної клітини. Збалансованість даних забезпечується їх однаковою кількістю у групах класифікації.

Був проведений експеримент щодо виявлення оптимальних вихідних параметрів моделі. Виявлено, що оптимальний розмір зсуву вікна даних складає 100 мс, при цьому значенні був отриманий найбільший відсоток успішних класифікацій в чотирьох вибірках на етапі навчання (близько 90%).

В подальшому на базі однієї вибірки даних було здійснено навчання і тестування моделі машинного навчання. За результатом експерименту, на етапі навчання точність класифікації моделі склала 90,7% (астма – 77,8%; ХОЗЛ – 95,5%; без патології – 97,8%). На етапі тестування модель показала гірший результат класифікації 67,47% (астма – 73,9%; ХОЗЛ – 73,9%; без патології – 54,5%). Відзначено, що отримана модель машинного навчання з вірогідністю близько 0,7 може вірно класифікувати захворювання астма і ХОЗЛ і може бути застосована для побудування автономного пристрою класифікації шумів легень.

В четвертому розділі роботи розглянуто апаратну реалізацію пристрою з визначення патологій легень за акустичними шумами на базі мікроконтролера Arduino Nano 33 BLE Sense. Розглянуті складові компоненти пристрою. Загальна вартість розробленого пристрою складає близько \$50, що в чотири рази дешевше, ніж комерційні західні зразки електронних сетоскопів.

В якості програмного забезпечення розробленого прототипу виступає модель машинного навчання з класифікації звуків легень, розроблена за допомогою ПЗ EdgeImpulse, в яку додано алгоритм виводу результатів класифікації на зовнішній дисплей.

Спроектований пристрій має компактні розміри і може використовуватись лікарями як допоміжний засіб при діагностиці захворювань органів дихання.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. World Health Organization / The top 10 causes of death. URL: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> (дата звернення: 10.05.2023)
2. World Health Organization / Global Tuberculosis Report 2022. URL: <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022> (дата звернення: 10.05.2023)
3. World Health Organization / Pneumonia in children. URL: <https://www.who.int/news-room/fact-sheets/detail/pneumonia> (дата звернення: 10.05.2023)
4. World Health Organization / Asthma URL: <https://www.who.int/news-room/fact-sheets/detail/asthma> (дата звернення: 10.05.2023)
5. Державна служба статистики України / Кількість померлих за окремими причинами смерті. URL: https://ukrstat.gov.ua/operativ/operativ2022/ds/krops/arh_krops2022_u.html (дата звернення: 09.03.2023).
6. «Світ наукових досліджень. Випуск 17»: матеріали Міжнародної мультидисциплінарної наукової інтернет-конференції (м. Тернопіль, Україна – м. Переворськ, Польща, 16-17 березня 2023 р.) / [редкол.: О. Патряк та ін.]; ГО «Наукова спільнота»; WSSG w Przeworsku. – Тернопіль: ФО-П Шпак В.Б. – 318 с.
7. Сучасний стан та перспективи біомедичної інженерії: матеріали Міжнародної науково-практичної конференції, присвяченої 125-річному ювілею Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» (13-14.12.2023, м. Київ): ел.збірник / Упоряд.: О.І. Голембіовська – Київ: КПІ ім. Ігоря Сікорського, 2023. – 239 с.
8. Гиппократ. Избранные книги; Пер. с греч. – М.: Сварог, 1994.
9. Wilkins R.L. // Respir. Care. – 2004. – Vol. 49, No. 12. – P. 1488-1489.

10. Ademovis E., Pesquet J.C., Charbonneau G. // *Technol. Health. Care.* – 1998. – Vol. 6, No. 1. – P. 53-56.
11. Рене Лаэннек: от скепсиса до признания // *Клінічна іммунологія. Алергологія. Інфектологія.* – 2011, №9-10 (48-49). – С. 55-58. URL: <https://kiai.com.ua/ua/archive/2011/9-10%2848-49%29/pages-55-58/rene-laennek-ot-skepsisa-do-priznaniya> (дата звернення: 09.03.2023).
12. Postiaux G., Lens E. // *Rev. Mal. Respir.* – 1999. – Vol. 16. – P. 1075-1090.
13. Kraman S.S., Hamm G. // *Ann. Intern. Med.* – 1999. – Vol. 21, No. 12. – P. 963-967.
14. Gavriely N., Nissan M., Cugell D.W., Rubin A.H. // *Eur. Respir. J.* – 1994. – Vol. 7. – P. 35-42.
15. Kraman S.S., Wodicka G.R., Kiyokawa H., Pasterkamp H. // *Biomed. Instrum. Technol.* – 2002. – Vol. 36, No. 3. – P. 177–182.
16. Особливості побудови сенсорів реєстрації акустичних шумів процесу дихання / М.Ф. Бабаков, Є.О. Кривенко, О.В. Кривенко, В.І. Луценко, В.А. Клименко, О.В. Піонтковська, Н.М. Дробова // IV Міжнародна науково-практична конференція «Інформаційні системи та технології в медицині» (ІСМ–2021) [Текст]: зб. наук. пр. – Харків: Нац. аерокосм. ун-т ім. М.Є. Жуковського «Харків. авіац. ін-т», 2021. – 260 с., С. 206-208.
17. Кулаков Ю.В., Малышенко И.Ю., Коренбаум В.И. // *Пульмонология.* – 2002. – № 5. – С. 29–32.
18. Почекутова И.А., Коренбаум В.И., Кулаков Ю.В., Авдеева Е.В., Тагильцев А.А. // *Физиология человека.* – 2001. – Т.27, № 4. – С. 61–65.
19. Sun X., Cheetham B.M., Earis J.E. // *Technol. Health. Care.* – 1998. – Vol. 6. – P. 3-10.
20. Chardonneau G., Meslier N., Racineux J.L. // *Eur. Respir. J.* – 1995. – Vol. 8, No. 11. – P. 1942–1948.
21. V. Oliynik, «On potential effectiveness of integration of 3M Littmann 3200 electronic stethoscopes into the third-party diagnostic systems with

auscultation signal processing», 2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO), Kiev, 2015, pp. 417-421. – doi: 10.1109/ELNANO.2015.7146923

22. Littmann3200: <https://medilife.com.ua/ua/stetoskop-1.html> (дата звернення: 25.09.2023)

23. 3M Littmann CORE Digital Stethoscope: https://www.littmann.com/3M/en_US/littmann-stethoscopes/advantages/core-digital-stethoscope/ (дата звернення: 25.09.2023)

24. <https://www.ekohealth.com/products/core-digital-attachment> (дата звернення: 25.09.2023)

25. Thinklabs One: <https://www.thinklabs.com/> (дата звернення: 25.09.2023)

26. AyuSynk 2Pro: <https://www.ayusynk.ai/product/ayusynk-2pro/> (дата звернення: 25.09.2023)

27. A. Rao, E. Huynh, T. J. Royston, A. Kornblith, and S. Roy, "Acoustic methods for pulmonary diagnosis", IEEE Rev. Biomed. Eng., vol. 12, pp. 221-239, 2019. doi: 10.1109/RBME.2018.2874353

28. R. Palaniappan, K. Sundaraj, N. U. Ahamed, A. Arjunan, and S. Sundaraj, "Computer-based respiratory sound analysis: a systematic review, IETE Techn. Rev., vol. 30, no 3, pp. 248-256, 2013.

29. А.А. Гусейнов, З.Р. Айсанов, А.Г. Чучалин, "Акустический анализ дыхательных звуков: состояние вопроса," Пульмонология, № 6, с.105-112, 2005.

30. А.А. Макаренкова, "Акустические характеристики звуков дыхания у больных пневмокониозом" Акустичний вісник, т. 11, №1, с. 51-59, 2008.

31. U. Patel, Computerized Respiratory Sound Analysis: An Exploration of Methods, B.S. Thesis, Case Western Reserve Univ., 2011. http://www.phys.cwru.edu/undergrad/Senior%20Projects/PreviousProjects/papers/papers2012/Patel_Cechner.pdf

32. D.S. Morillo, S.A. Morenoemail, M.Á.F. Graneroemail, and A.L. Jiménezemail, "Computerized analysis of respiratory sounds during COPD exacerbations," *Comput. Biol. Med.*, vol. 43, no 7, pp. 914-921, 2013. - doi: 10.1016/j.combiomed.2013.03.011

33. V. Oliynik, "Time-domain fragment-similarity processing of the auscultation records for detection of their periodic components," *Proc. Int. Conf. Electron. Nanotechnol. (ELNANO-2018)*, Kyiv, 2018, pp. 340-345. - doi: 10.1109/ELNANO.2018.8477549

34. Технології виділення сигнатур акустичних шумів при патології легень / М.Ф. Бабаков, В.І. Луценко, І.В. Луценко, Ло Иян // II Міжнародна науково-практична конференція «Інформаційні системи та технології в медицині» (ICM–2019) [Текст]: зб. наук. пр. – Харків: Нац. аерокосм. ун-т ім. М.Є. Жуковського «Харків. авіац. ін-т», 2019. – 220 с., С. 30-31.

35. H.M.H. Johari, N.A. Malik, and K.A. Sidek, "Distinctive features for classification of respiratory sounds between normal and crackles using cepstral coefficients," *Proc. Int. Conf. Comput. Communic. Eng. (ICCCE-2018)*, Kuala Lumpur, pp. 476-479. - doi 10.1109/ICCCE.2018.8539305

36. Simulation Model of Respiratory Sound and Technology for Separating Characteristics of Pulmonary / Luo, Y., Lutsenko, V.I., Shulgar, S.M., Anh, N.X. // Disease. In: Yang, X.S., Sherratt, S., Dey, N., Joshi, A. (eds) *Proceedings of Seventh International Congress on Information and Communication Technology. Lecture Notes in Networks and Systems*, vol 448. Springer, Singapore. https://doi.org/10.1007/978-981-19-1610-6_13 (2023) – P. 143-151.

37. TinyML: Making Smart Devices Tinier than Ever: <https://www.plugandplaytechcenter.com/resources/tinyml-making-smart-devices-tinier-ever/> (дата звернення: 25.09.2023)

38. Edge Impulse: An MLOps Platform for Tiny Machine Learning: <https://arxiv.org/abs/2212.03332>(дата звернення: 25.09.2023)

39. Holdout Method: <https://www.geeksforgeeks.org/introduction-of-holdout-method/> (дата звернення: 25.09.2023)

40. 1D Convolutional Neural Networks and Applications – A Survey:
<https://arxiv.org/ftp/arxiv/papers/1905/1905.03554.pdf>

41. A dataset of lung sounds recorded from the chest wall using an electronic stethoscope: <https://doi.org/10.1016/j.dib.2021.106913> (дата звернення: 22.10.2023)

42. ICBHI 2017 Challenge. Respiratory Sound Database:
https://bhichallenge.med.auth.gr/ICBHI_2017_Challenge (дата звернення: 22.10.2023)

43. Arduino Nano 33 BLE Sense Datasheet:
<https://docs.arduino.cc/resources/datasheets/ABX00031-datasheet.pdf> (дата звернення: 22.10.2023)

44. MEMS audio sensor omnidirectional digital microphone MP34DT05-A Datasheet: <https://www.st.com/resource/en/datasheet/mp34dt05-a.pdf> (дата звернення: 22.10.2023)

45. MEMS-микрофоны компании Analog Devices / В. Макаренко // Журнал «Электронные компоненты и системы» № 6, июнь 2012, – С. 34-40:
http://www.ekis.kiev.ua/UserFiles/Image/pdfArticles/V.Makarenko_MEMS_microphones_EKIS_6_2012-3.pdf (дата звернення: 22.10.2023)

ДОДАТОК А
«Сертифікати участі в наукових конференціях»

www.economy-confer.com.ua

СЕРТИФІКАТ

цей Сертифікат засвідчує, що

Цуканов-Капітульський Андрій Григорович

взяв(ла) участь у роботі Міжнародної наукової інтернет-конференції

«Світ наукових досліджень. Випуск 17» з публікацією на тему:

**«РОЗРОБКА АВТОМАТИЗОВАНОГО СМАРТ КОМПЛЕКСУ ВИЗНАЧЕННЯ
ЗАХВОРЮВАНЬ ОРГАНІВ ДИХАННЯ ЗА АКУСТИЧНИМИ ШУМАМИ»**

Форма участі: дистанційна, тривалість конференції 18 годин 0,6 ECTS credits

Конференція проведена за
сприянням та участю Громадської
організації "Наукова спільнота" та
Wyższej Szkoły Społeczno-
Gospodarcza w Przeworsku (Польща)



Голова Оргкомітету

м.Тернопіль (Україна) -
м.Переворськ (Польща),
16-17 березня 2023 року



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
ГРОМАДСЬКА ОРГАНІЗАЦІЯ «ВСЕУКРАЇНСЬКА АСОЦІАЦІЯ БІОМЕДИЧНИХ
ІНЖЕНЕРІВ І ТЕХНОЛОГІВ»



СЕРТИФІКАТ

Цим підтверджуємо, що

Цуканов-Капітульський Андрій

приймав(-ла) участь в

Міжнародній науково-практичній конференції
«СУЧАСНИЙ СТАН ТА ПЕРСПЕКТИВИ БІОМЕДИЧНОЇ ІНЖЕНЕРІЇ»,
присвяченій
125-річчю Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»,
13-14 грудня 2023

Відповідно до критеріїв нарахування балів безперервного професійного розвитку нараховано 10 годин / 0,15 кредитів ECTS

Віталій МАКСИМЕНКО
Професор, д.м.н

Kyiv, Ukraine

Олександр ГАЛКІН
Професор, д.м.н.
Факультет
біомедичної
інженерії



Додаток Б

«Лістинг аудіозаписів шумів легень

за трьома діагнозами і чотирма місцями запису звуків легень»

PLM

asthma (12 файлів)

BP43_asthma,E W,P L M,15,M.wav
BP47_asthma,E W,P L M,60,M.wav
BP58_asthma,E W,P L M,40,M.wav
BP101_Asthma,E W,P L M,12,F.wav
DP43_asthma,E W,P L M,15,M.wav
DP47_asthma,E W,P L M,60,M.wav
DP58_asthma,E W,P L M,40,M.wav
DP101_Asthma,E W,P L M,12,F.wav
EP43_asthma,E W,P L M,15,M.wav
EP47_asthma,E W,P L M,60,M.wav
EP58_asthma,E W,P L M,40,M.wav
EP101_Asthma,E W,P L M,12,F.wav

normal (18 файлів)

BP31_N,N,P L M,27,F.wav
BP51_N,N,P L M,30,M.wav
BP89_N,N,P L M,70,M.wav
BP95_N,N,P L M,18,M.wav
BP109_N,N,P L M,26,M.wav
BP112_N,N,P L M,30,M.wav
DP31_N,N,P L M,27,F.wav
DP51_N,N,P L M,30,M.wav
DP89_N,N,P L M,70,M.wav
DP95_N,N,P L M,18,M.wav
DP109_N,N,P L M,26,M.wav
DP112_N,N,P L M,30,M.wav
EP31_N,N,P L M,27,F.wav
EP51_N,N,P L M,30,M.wav
EP89_N,N,P L M,70,M.wav

EP95_N,N,P L M,18,M.wav

EP109_N,N,P L M,26,M.wav

EP112_N,N,P L M,30,M.wav

copd (11 файлів)

104_1b1_Pl_sc_Litt3200.wav

109_1b1_Pl_sc_Litt3200.wav

112_1p1_Pl_sc_Litt3200.wav

113_1b1_Pl_sc_Litt3200.wav

118_1b1_Pl_sc_Litt3200.wav

124_1b1_Pl_sc_Litt3200.wav

139_1b1_Pl_sc_Litt3200.wav

175_1b1_Pl_sc_Litt3200.wav

185_1b1_Pl_sc_Litt3200.wav

195_1b1_Pl_sc_Litt3200.wav

218_1p1_Pl_sc_Litt3200.wav

PRM

asthma (18 файлів)

BP27_asthma,E W,P R M,53,F.wav

BP33_Asthma,E W,P R M,43,F.wav

BP38_Asthma,E W,P R M,38,M.wav

BP53_asthma,E W,P R M,72,F.wav

BP56_asthma,E W,P R M,56,F.wav

BP78_Asthma,E W,P R M,20,M.wav

DP27_asthma,E W,P R M,53,F.wav

DP33_Asthma,E W,P R M,43,F.wav

DP38_Asthma,E W,P R M,38,M.wav

DP53_asthma,E W,P R M,72,F.wav

DP56_asthma,E W,P R M,56,F.wav

DP78_Asthma,E W,P R M,20,M.wav

EP27_asthma,E W,P R M,53,F.wav

EP33_Asthma,E W,P R M,43,F.wav

EP38_Asthma,E W,P R M,38,M.wav

EP53_asthma,E W,P R M,72,F.wav

EP56_asthma,E W,P R M,56,F.wav

EP78_Asthma,E W,P R M,20,M.wav

normal (21 файл)

BP30_N,N,P R M,18,F.wav

BP49_N,N,P R M,21,M.wav

BP74_N,N,P R M,74,M.wav

BP87_N,N,P R M,72,M.wav

BP93_N,N,P R M,75,M.wav

BP99_N,N,P R M,50,M.wav

BP100_N,N,P R M,70,F.wav

DP30_N,N,P R M,18,F.wav

DP49_N,N,P R M,21,M.wav

DP74_N,N,P R M,74,M.wav

DP87_N,N,P R M,72,M.wav

DP93_N,N,P R M,75,M.wav

DP99_N,N,P R M,50,M.wav

DP100_N,N,P R M,70,F.wav

EP30_N,N,P R M,18,F.wav

EP49_N,N,P R M,21,M.wav

EP74_N,N,P R M,74,M.wav

EP87_N,N,P R M,72,M.wav

EP93_N,N,P R M,75,M.wav

EP99_N,N,P R M,50,M.wav

EP100_N,N,P R M,70,F.wav

copd (11 файлів)

104_1b1_Pr_sc_Litt3200.wav

109_1b1_Pr_sc_Litt3200.wav

112_1p1_Pr_sc_Litt3200.wav
113_1b1_Pr_sc_Litt3200.wav
118_1b1_Pr_sc_Litt3200.wav
124_1b1_Pr_sc_Litt3200.wav
139_1b1_Pr_sc_Litt3200.wav
175_1b1_Pr_sc_Litt3200.wav
185_1b1_Pr_sc_Litt3200.wav
195_1b1_Pr_sc_Litt3200.wav
218_1p1_Pr_sc_Litt3200.wav

PLU

asthma (15 файлів)

BP46_asthma,E W,P L U,41,F.wav
BP64_asthma,E W,P L U,60,M.wav
BP104_Asthma,E W,P L U,45,F.wav
BP106_Asthma,E W,P L U,45,F.wav
BP107_Asthma,E W,P L U,59,F.wav
DP46_asthma,E W,P L U,41,F.wav
DP64_asthma,E W,P L U,60,M.wav
DP104_Asthma,E W,P L U,45,F.wav
DP106_Asthma,E W,P L U,45,F.wav
DP107_Asthma,E W,P L U,59,F.wav
EP46_asthma,E W,P L U,41,F.wav
EP64_asthma,E W,P L U,60,M.wav
EP104_Asthma,E W,P L U,45,F.wav
EP106_Asthma,E W,P L U,45,F.wav
EP107_Asthma,E W,P L U,59,F.wav

normal (15 файлів)

BP48_N,N,P L U,31,M.wav
BP70_N,N,P L U,52,F.wav
BP75_N,N,P L U,73,F.wav

BP81_N,N,P L U,33,M.wav
BP82_N,N,P L U,38,M.wav
DP48_N,N,P L U,31,M.wav
DP70_N,N,P L U,52,F.wav
DP75_N,N,P L U,73,F.wav
DP81_N,N,P L U,33,M.wav
DP82_N,N,P L U,38,M.wav
EP48_N,N,P L U,31,M.wav
EP70_N,N,P L U,52,F.wav
EP75_N,N,P L U,73,F.wav
EP81_N,N,P L U,33,M.wav
EP82_N,N,P L U,38,M.wav

copd (11 файлів)

104_1b1_Pl_sc_Litt3200.wav
109_1b1_Pl_sc_Litt3200.wav
112_1p1_Pl_sc_Litt3200.wav
113_1b1_Pl_sc_Litt3200.wav
118_1b1_Pl_sc_Litt3200.wav
124_1b1_Pl_sc_Litt3200.wav
139_1b1_Pl_sc_Litt3200.wav
175_1b1_Pl_sc_Litt3200.wav
185_1b1_Pl_sc_Litt3200.wav
195_1b1_Pl_sc_Litt3200.wav
218_1p1_Pl_sc_Litt3200.wav

PRU

asthma (18 файлів)

BP10_Asthma,E W,P R U,59,M.wav
BP15_Asthma,E W,P R U,49,F.wav
BP44_asthma,E W,P R U,40,M.wav
BP55_asthma,E W,P R U,72,F.wav

BP91_asthma,E W,P R U,43,F.wav
BP97_Asthma,E W,P R U,24,M.wav
DP10_Asthma,E W,P R U,59,M.wav
DP15_Asthma,E W,P R U,49,F.wav
DP44_asthma,E W,P R U,40,M.wav
DP55_asthma,E W,P R U,72,F.wav
DP91_asthma,E W,P R U,43,F.wav
DP97_Asthma,E W,P R U,24,M.wav
EP10_Asthma,E W,P R U,59,M.wav
EP15_Asthma,E W,P R U,49,F.wav
EP44_asthma,E W,P R U,40,M.wav
EP55_asthma,E W,P R U,72,F.wav
EP91_asthma,E W,P R U,43,F.wav
EP97_Asthma,E W,P R U,24,M.wav

normal (12 файлів)

BP71_N,N,P R U,36,M.wav
BP72_N,N,P R U,24,F.wav
BP86_N,N,P R U,68,F.wav
BP103_N,N,P R U,81,F.wav
DP71_N,N,P R U,36,M.wav
DP72_N,N,P R U,24,F.wav
DP86_N,N,P R U,68,F.wav
DP103_N,N,P R U,81,F.wav
EP71_N,N,P R U,36,M.wav
EP72_N,N,P R U,24,F.wav
EP86_N,N,P R U,68,F.wav
EP103_N,N,P R U,81,F.wav

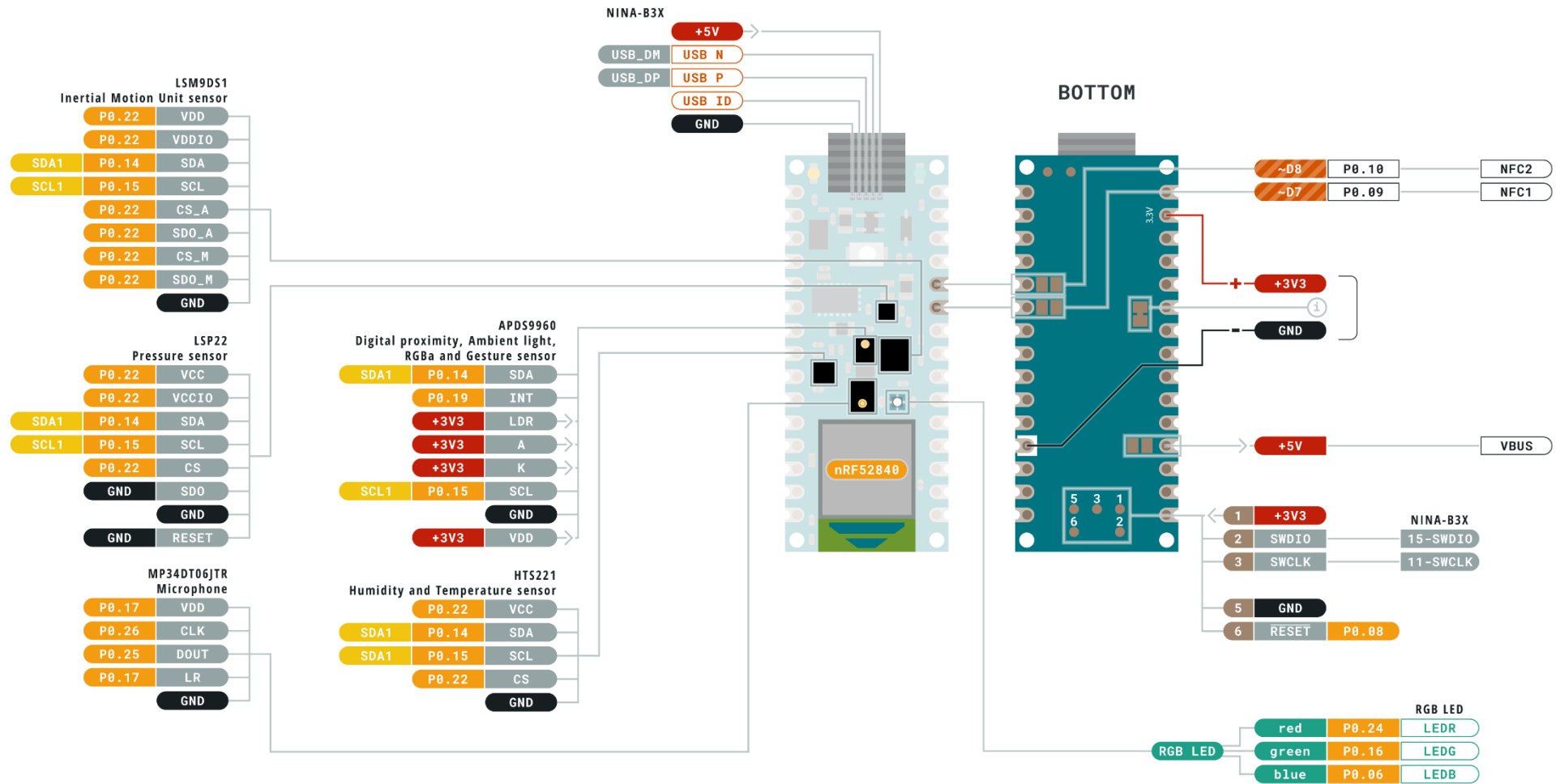
copd (11 файлів)

104_1b1_Pr_sc_Litt3200.wav
109_1b1_Pr_sc_Litt3200.wav

112_1p1_Pr_sc_Litt3200.wav
113_1b1_Pr_sc_Litt3200.wav
118_1b1_Pr_sc_Litt3200.wav
124_1b1_Pr_sc_Litt3200.wav
139_1b1_Pr_sc_Litt3200.wav
175_1b1_Pr_sc_Litt3200.wav
185_1b1_Pr_sc_Litt3200.wav
195_1b1_Pr_sc_Litt3200.wav
218_1p1_Pr_sc_Litt3200.wav

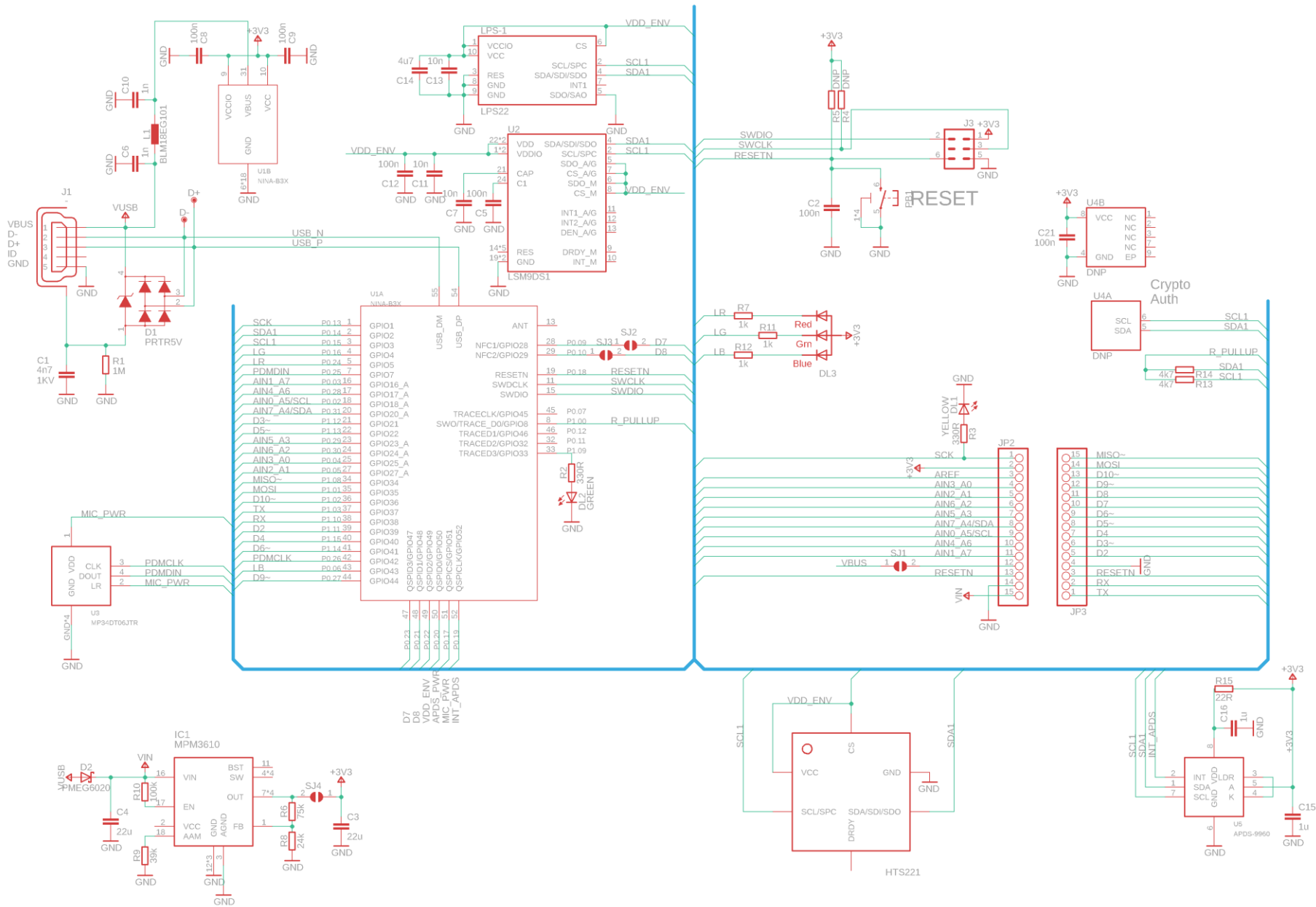
ДОДАТОК В

«Схема мікроконтролера Arduino Nano 33 BLE Sense з виводами і сенсорами»



ДОДАТОК Г

«Схема електрична принципова
мікроконтролера Arduino Nano 33 BLE Sense»



Додаток Д

«Програмний код пристрою для класифікації акустичних шумів легень
(файл nano_ble33_sense_microphone_continuous.ino)»

```

/* Edge Impulse ingestion SDK
 * Copyright (c) 2022 EdgeImpulse Inc.
 *
 * Licensed under the Apache License, Version 2.0 (the "License");
 * you may not use this file except in compliance with the License.
 * You may obtain a copy of the License at
 * http://www.apache.org/licenses/LICENSE-2.0
 *
 * Unless required by applicable law or agreed to in writing, software
 * distributed under the License is distributed on an "AS IS" BASIS,
 * WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
 * See the License for the specific language governing permissions and
 * limitations under the License.
 *
 */

// If your target is limited in memory remove this macro to save 10K RAM
#define EIDSP_QUANTIZE_FILTERBANK 0

/**
 * Define the number of slices per model window. E.g. a model window of 1000 ms
 * with slices per model window set to 4. Results in a slice size of 250 ms.
 * For more info: https://docs.edgeimpulse.com/docs/continuous-audio-sampling
 */
#define EI_CLASSIFIER_SLICES_PER_MODEL_WINDOW 4

/*
 ** NOTE: If you run into TFLite arena allocation issue.
 **
 ** This may be due to may dynamic memory fragmentation.
 ** Try defining "-DEI_CLASSIFIER_ALLOCATION_STATIC" in boards.local.txt (create
 ** if it doesn't exist) and copy this file to
 **
`<ARDUINO_CORE_INSTALL_PATH>/arduino/hardware/<mbed_core>/<core_version>/'
 **

```

```

** See
** (https://support.arduino.cc/hc/en-us/articles/360012076960-Where-are-the-installed-cores-located-)
** to find where Arduino installs cores on your machine.
**
** If the problem persists then there's not enough memory for this model and application.
*/

/* Includes ----- */
#include <PDM.h>
#include <Sound_Classifier_PRM_with_test_data__inferencing.h>

// Підключення бібліотек
#include <SPI.h>
#include <Wire.h>
#include <Adafruit_GFX.h>
#include <Adafruit_SSD1306.h>

/** Audio buffers, pointers and selectors */
typedef struct {
    signed short *buffers[2];
    unsigned char buf_select;
    unsigned char buf_ready;
    unsigned int buf_count;
    unsigned int n_samples;
} inference_t;

static inference_t inference;
static bool record_ready = false;
static signed short *sampleBuffer;
static bool debug_nn = false; // Set this to true to see e.g. features generated from the raw signal
static int print_results = -(EI_CLASSIFIER_SLICES_PER_MODEL_WINDOW);

float asthma, normal, noise; // змінні для збереження вірогідностей класифікації

```

```

/** Display */

#define SCREEN_WIDTH 128 // OLED display width, in pixels
#define SCREEN_HEIGHT 32 // OLED display height, in pixels

// Declaration for an SSD1306 display connected to I2C (SDA, SCL pins) - створюємо об'єкт
дисплею SSD1306, підключеного до I2C (контакти SDA, SCL)
// The pins for I2C are defined by the Wire-library.
// On an arduino UNO:   A4(SDA), A5(SCL)
// On an arduino MEGA 2560: 20(SDA), 21(SCL)
// On an arduino LEONARDO:  2(SDA),  3(SCL), ...
#define OLED_RESET   -1 // Reset pin # (or -1 if sharing Arduino reset pin) - встановлюємо
пін reset
#define SCREEN_ADDRESS 0x3C ///< See datasheet for Address; 0x3D for 128x64, 0x3C for
128x32 - адреса 0x3C для дисплею 128x32
Adafruit_SSD1306 display(SCREEN_WIDTH, SCREEN_HEIGHT, &Wire, OLED_RESET);

/**
 * @brief   Arduino setup function
 */
void setup()
{
    // put your setup code here, to run once:
    Serial.begin(115200);
    // comment out the below line to cancel the wait for USB connection (needed for native USB)
    while (!Serial);
    Serial.println("Edge Impulse Inferencing Demo");

    // summary of inferencing settings (from model_metadata.h)
    ei_printf("Inferencing settings:\n");
    ei_printf("\tInterval: %.2f ms.\n", (float)EI_CLASSIFIER_INTERVAL_MS);
    ei_printf("\tFrame size: %d\n", EI_CLASSIFIER_DSP_INPUT_FRAME_SIZE);
    ei_printf("\tSample length: %d ms.\n", EI_CLASSIFIER_RAW_SAMPLE_COUNT / 16);
    ei_printf("\tNo. of classes: %d\n", sizeof(ei_classifier_inferencing_categories) /
        sizeof(ei_classifier_inferencing_categories[0]));
}

```

```

run_classifier_init();
if (microphone_inference_start(EI_CLASSIFIER_SLICE_SIZE) == false) {
    ei_printf("ERR: Could not allocate audio buffer (size %d), this could be due to the window
length of your model\r\n", EI_CLASSIFIER_RAW_SAMPLE_COUNT);
    return;
}

// SSD1306_SWITCHCAPVCC = generate display voltage from 3.3V internally - генерувати
внутрішню напругу на дисплеї від піну 3v3
if(!display.begin(SSD1306_SWITCHCAPVCC, SCREEN_ADDRESS)) { // у випадку
невдалої ініціалізації дисплею
    Serial.println(F("SSD1306 allocation failed")); // вивести повідомлення про помилку
for(;;); // Don't proceed, loop forever - далі нічого не робити - нескінченний цикл
}
}

/**
 * @brief   Arduino main function. Runs the inferencing loop.
 */
void loop()
{
    bool m = microphone_inference_record();
    if (!m) {
        ei_printf("ERR: Failed to record audio...\n");
        return;
    }

    signal_t signal;
    signal.total_length = EI_CLASSIFIER_SLICE_SIZE;
    signal.get_data = &microphone_audio_signal_get_data;
    ei_impulse_result_t result = {0};

    EI_IMPULSE_ERROR r = run_classifier_continuous(&signal, &result, debug_nn);
    if (r != EI_IMPULSE_OK) {

```

```

    ei_printf("ERR: Failed to run classifier (%d)\n", r);
    return;
}

// Вивід результатів класифікації на дисплей

asthma = result.classification[0].value;
copd = result.classification[1].value;
normal = result.classification[2].value;

display.clearDisplay(); // очистка дисплею
if (asthma > 0.70) {
    displayText("Asthma", 0, 0, 2, true);
    delay(1000);
}
if (copd > 0.70) {
    displayText("COPD", 0, 0, 2, true);
    delay(1000);
}

if (++print_results >= (EI_CLASSIFIER_SLICES_PER_MODEL_WINDOW)) {
    // print the predictions
    ei_printf("Predictions ");
    ei_printf("(DSP: %d ms., Classification: %d ms., Anomaly: %d ms.)",
        result.timing.dsp, result.timing.classification, result.timing.anomaly);
    ei_printf(": \n");
    for (size_t ix = 0; ix < EI_CLASSIFIER_LABEL_COUNT; ix++) {
        ei_printf("  %s: %.5f\n", result.classification[ix].label,
            result.classification[ix].value);
    }
}
#if EI_CLASSIFIER_HAS_ANOMALY == 1
    ei_printf("  anomaly score: %.3f\n", result.anomaly);
#endif

print_results = 0;

```

```

    }
}

/**
 * @brief   PDM buffer full callback
 *          Get data and call audio thread callback
 */
static void pdm_data_ready_inference_callback(void)
{
    int bytesAvailable = PDM.available();

    // read into the sample buffer
    int bytesRead = PDM.read((char *)&sampleBuffer[0], bytesAvailable);

    if (record_ready == true) {
        for (int i = 0; i < bytesRead >> 1; i++) {
            inference.buffer[inference.buf_select][inference.buf_count++] = sampleBuffer[i];

            if (inference.buf_count >= inference.n_samples) {
                inference.buf_select ^= 1;
                inference.buf_count = 0;
                inference.buf_ready = 1;
            }
        }
    }
}

/**
 * @brief   Init inferencing struct and setup/start PDM
 *
 * @param[in] n_samples The n samples
 *
 * @return   { description_of_the_return_value }
 */
static bool microphone_inference_start(uint32_t n_samples)

```

```
{
inference.buffers[0] = (signed short *)malloc(n_samples * sizeof(signed short));

if (inference.buffers[0] == NULL) {
    return false;
}

inference.buffers[1] = (signed short *)malloc(n_samples * sizeof(signed short));

if (inference.buffers[1] == NULL) {
    free(inference.buffers[0]);
    return false;
}

sampleBuffer = (signed short *)malloc((n_samples >> 1) * sizeof(signed short));

if (sampleBuffer == NULL) {
    free(inference.buffers[0]);
    free(inference.buffers[1]);
    return false;
}

inference.buf_select = 0;
inference.buf_count = 0;
inference.n_samples = n_samples;
inference.buf_ready = 0;

// configure the data receive callback
PDM.onReceive(&pdm_data_ready_inference_callback);

PDM.setBufferSize((n_samples >> 1) * sizeof(int16_t));

// initialize PDM with:
// - one channel (mono mode)
// - a 16 kHz sample rate
```



```

if (!PDM.begin(1, EI_CLASSIFIER_FREQUENCY)) {
    ei_printf("Failed to start PDM!");
}

// set the gain, defaults to 20
PDM.setGain(127);

record_ready = true;

return true;
}

/**
 * @brief Wait on new data
 *
 * @return True when finished
 */
static bool microphone_inference_record(void)
{
    bool ret = true;

    if (inference.buf_ready == 1) {
        ei_printf(
            "Error sample buffer overrun. Decrease the number of slices per model window "
            "(EI_CLASSIFIER_SLICES_PER_MODEL_WINDOW)\n");
        ret = false;
    }

    while (inference.buf_ready == 0) {
        delay(1);
    }

    inference.buf_ready = 0;

    return ret;
}

```

```

}

/**
 * Get raw audio signal data
 */
static int microphone_audio_signal_get_data(size_t offset, size_t length, float *out_ptr)
{
    numpy::int16_to_float(&inference.buffers[inference.buf_select ^ 1][offset], out_ptr, length);

    return 0;
}

/**
 * @brief Stop PDM and release buffers
 */
static void microphone_inference_end(void)
{
    PDM.end();
    free(inference.buffers[0]);
    free(inference.buffers[1]);
    free(sampleBuffer);
}

#if !defined(EI_CLASSIFIER_SENSOR) || EI_CLASSIFIER_SENSOR !=
EI_CLASSIFIER_SENSOR_MICROPHONE
#error "Invalid model for current sensor."
#endif

/**
 * displayText(String text, int x, int y,int size, boolean d) - функція для виводу тексту на
дисплей
 * text is the text string to be printed
 * x is the integer x position of text
 * y is the integer y position of text
 * size is the text size, 1, 2, 3 etc

```

```
* d is either "true" or "false". Not sure, use true
*/
void displayText(String text, int x, int y, int size, boolean d) {
    display.setTextSize(size); // розмір шрифту
    display.setTextColor(WHITE); // колір шрифту
    display.setCursor(x,y); // початкові координати тексту
    display.println(text); // вивід тексту
    if(d){
        display.display(); // вивід інформації на дисплей
    }
    //delay(100);
}
```