

Ідентифікація диктора по голосу

Автор: студент 4-го курсу
Групи КА-23

Деундяк Олександр

Науковий керівник – к.т.н. доц. Дідковська М.В.

Об'єкт та предмет дослідження

- Об'єкт дослідження – методи розпізнавання голосу, методи Data Mining
- Предмет дослідження – методи ідентифікації дикторів по голосу

Актуальність роботи

- Радіо-розвідка, антитерористичний моніторинг
- Системи безпеки
- Розблокування електронних пристроїв голосом (мобільні телефони, ноутбуки)

Існуючі методи ідентифікації диктора

- Марківські моделі (складність використання в алгоритмічних композиціях)
- Нейронні мережі (навчання займає багато часу)
- Метод найближчого сусіда (не створює ніяких моделей чи правил, узагальнюючих попередній досвід)

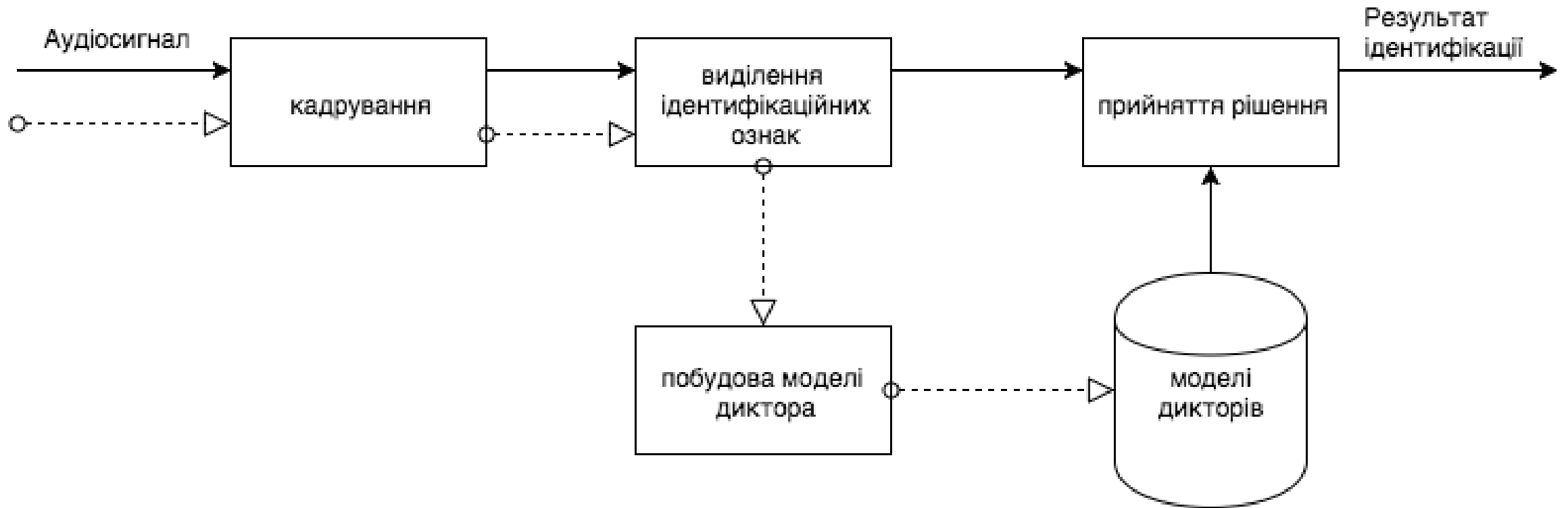
Постановка задачі

- Провести аналіз існуючих систем ідентифікації диктора, дослідити їх основні етапи роботи
- Розробити архітектуру системи ідентифікації диктора
- Розробити систему прийняття рішень на основі композиції двох класифікаторів
- Реалізувати цю систему у вигляді програмного продукту

Вимоги до метода ідентифікації

- При навчанні має враховувати попередній досвід
- Повинен мати інтуїтивно-зрозумілу структуру, що дасть можливість створити його композицію з іншими алгоритмами

Функціональна схема роботи системи



Етапи роботи алгоритму

Навчання:

- Підготовка даних
- Виділення ознак
- Навчання класифікаторів

Тестування:

- Підготовка даних
- Виділення ознак
- Класифікація

Схема роботи алгоритму виділення ознак

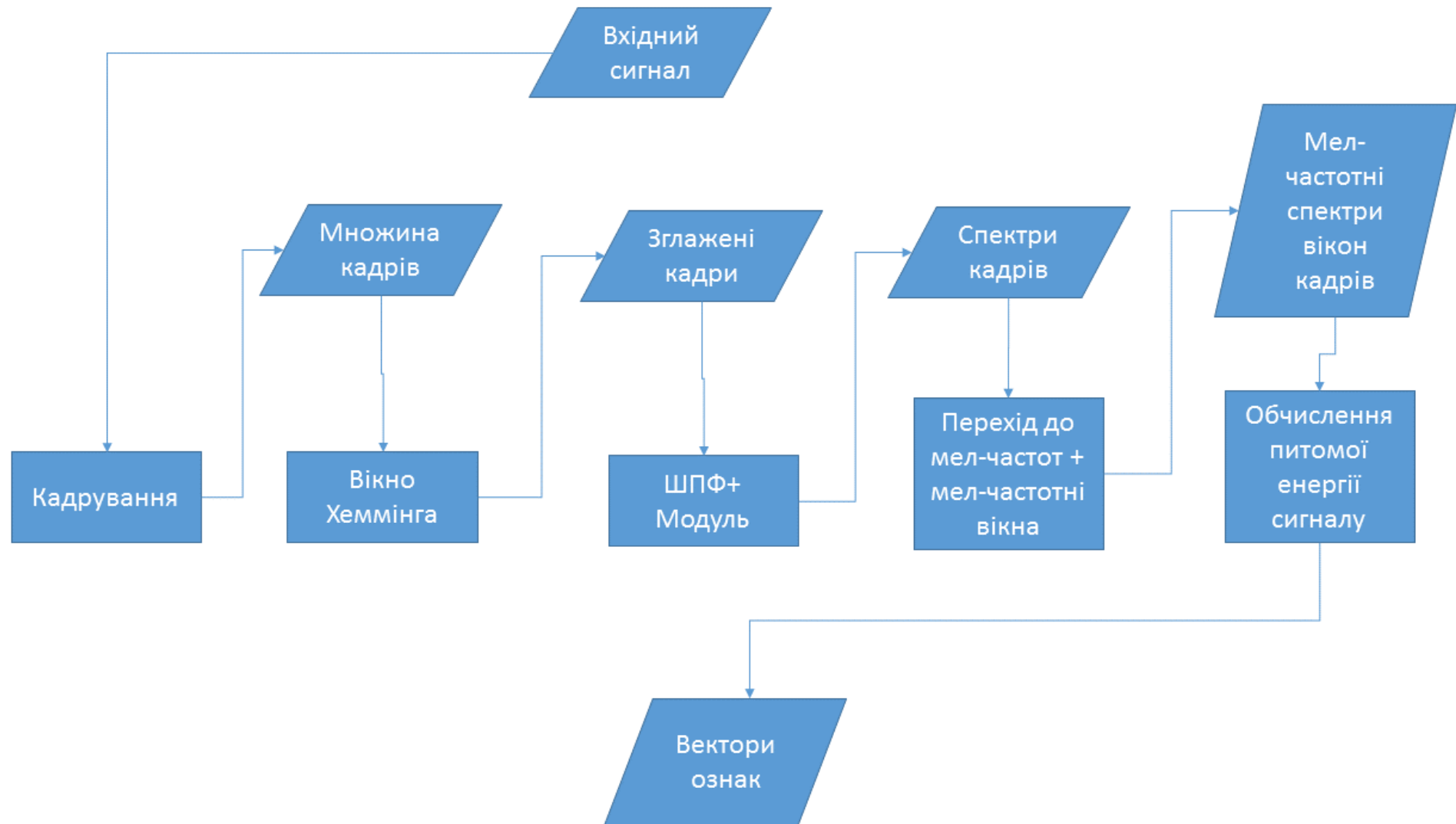


Схема роботи алгоритму навчання

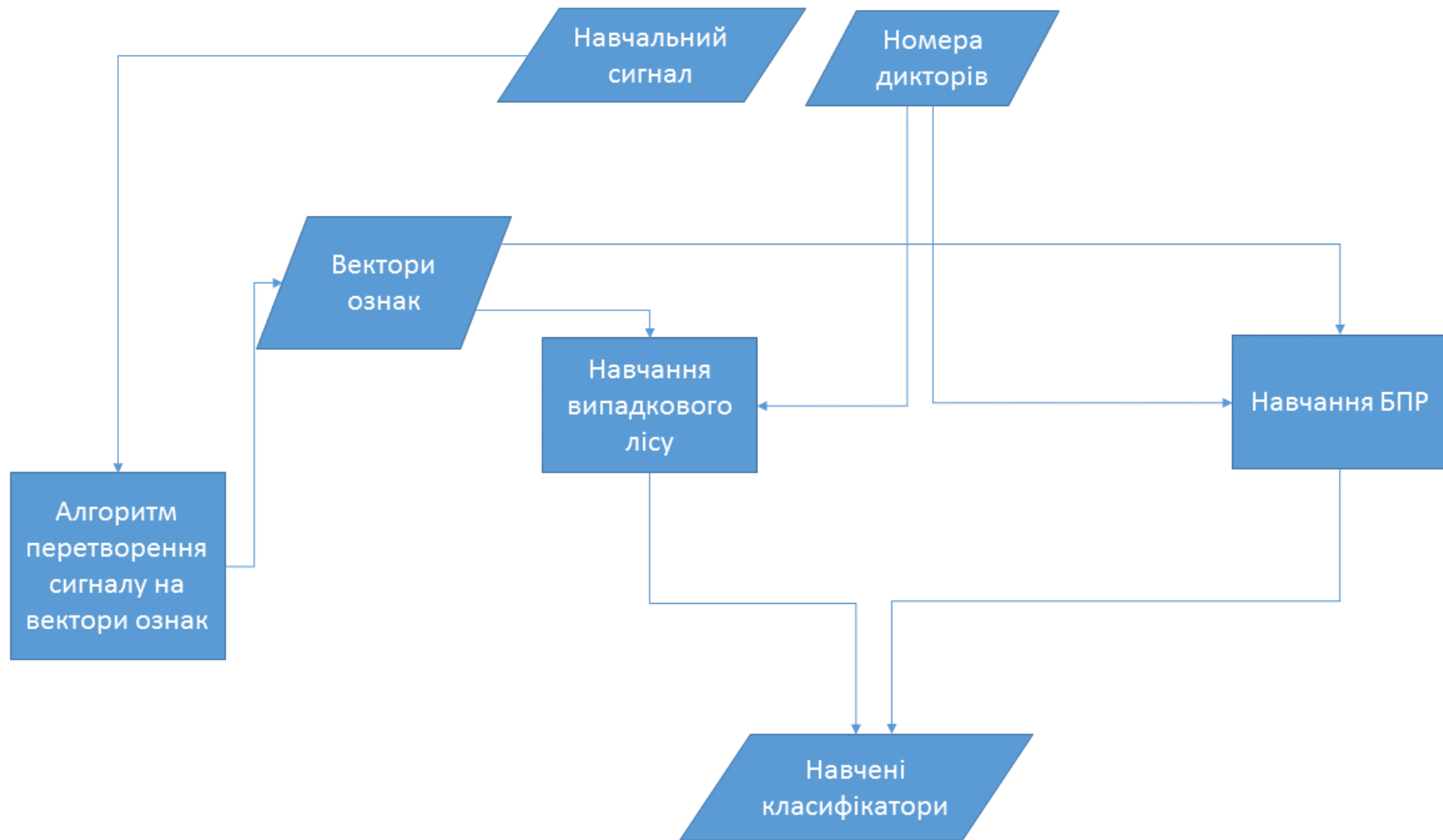
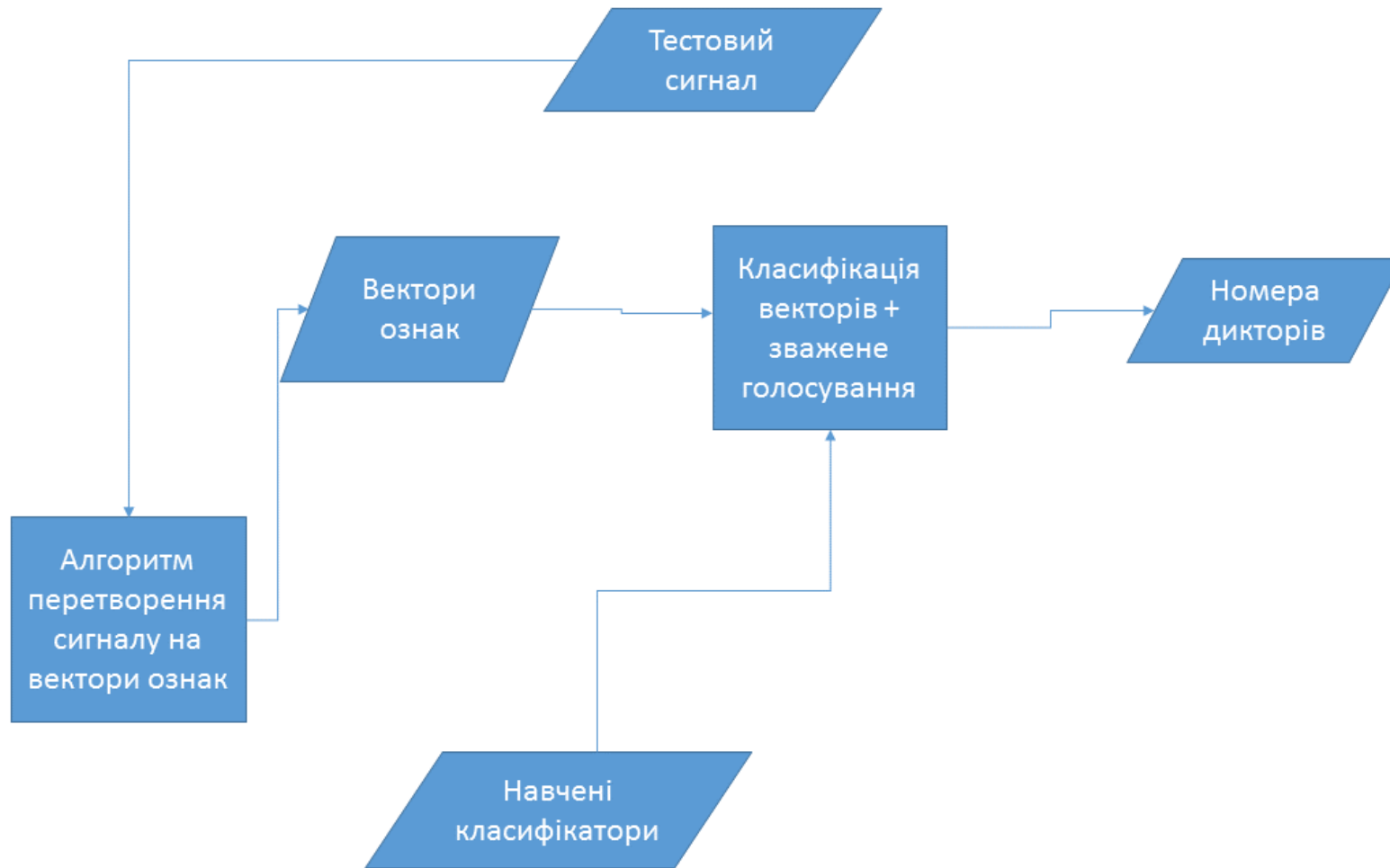
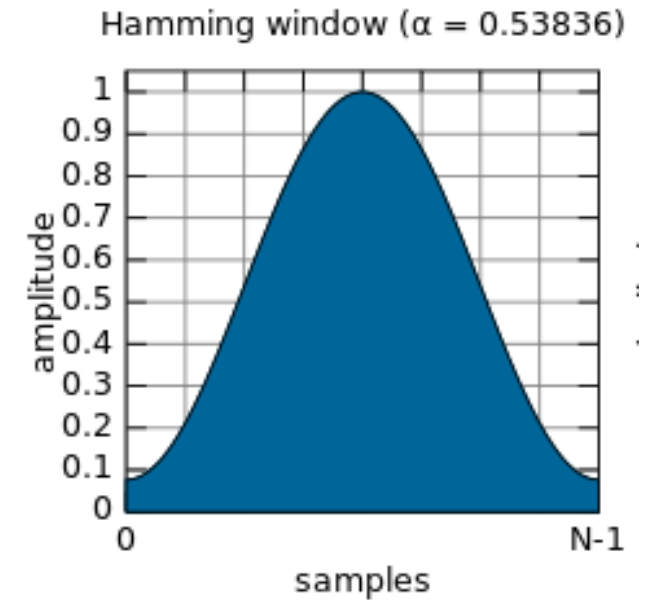
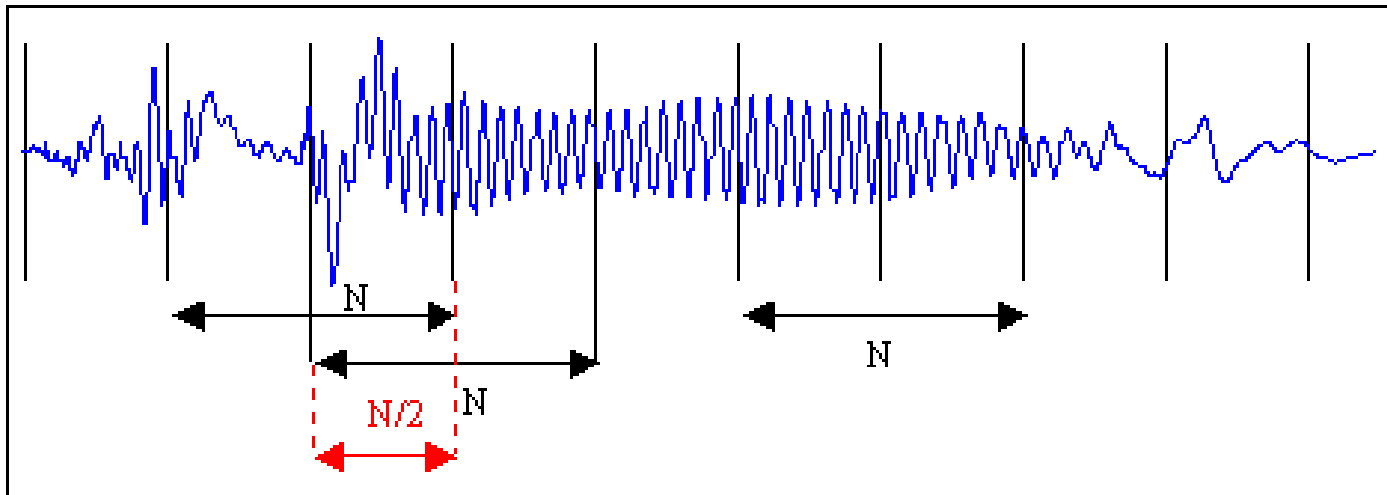


Схема роботи алгоритму тестування



Етап 1. Підготовка даних (Кадрування)



Етап 2. Виділення ознак

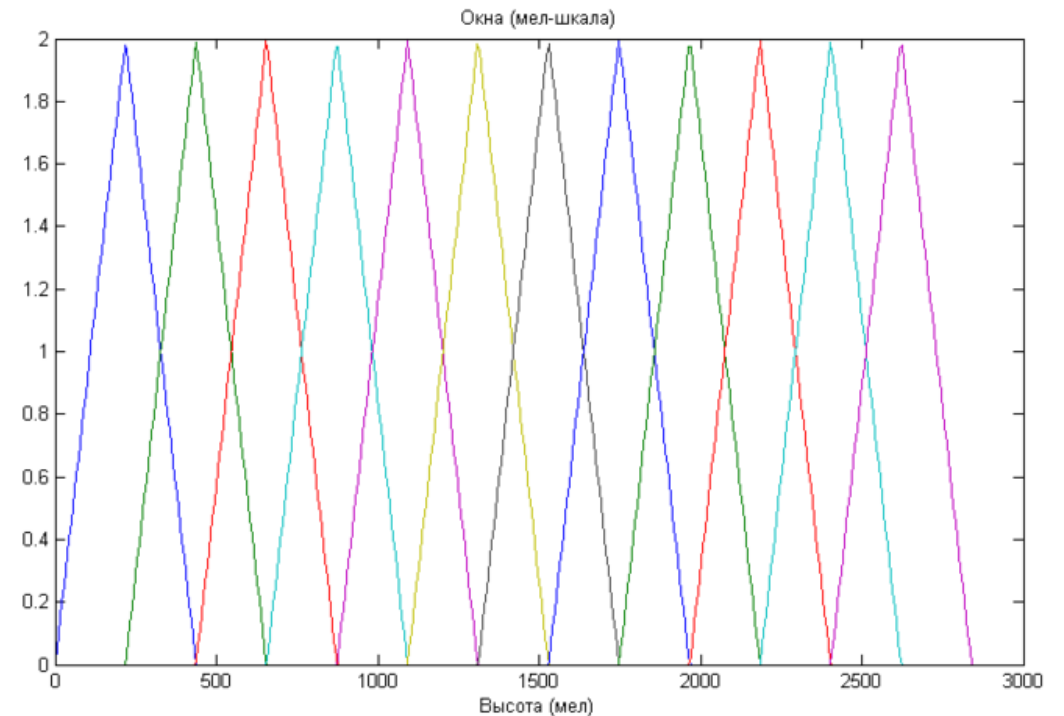
1) Швидке перетворення Фур'є

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k \frac{n}{N}} \quad k = 0, \dots, N-1$$

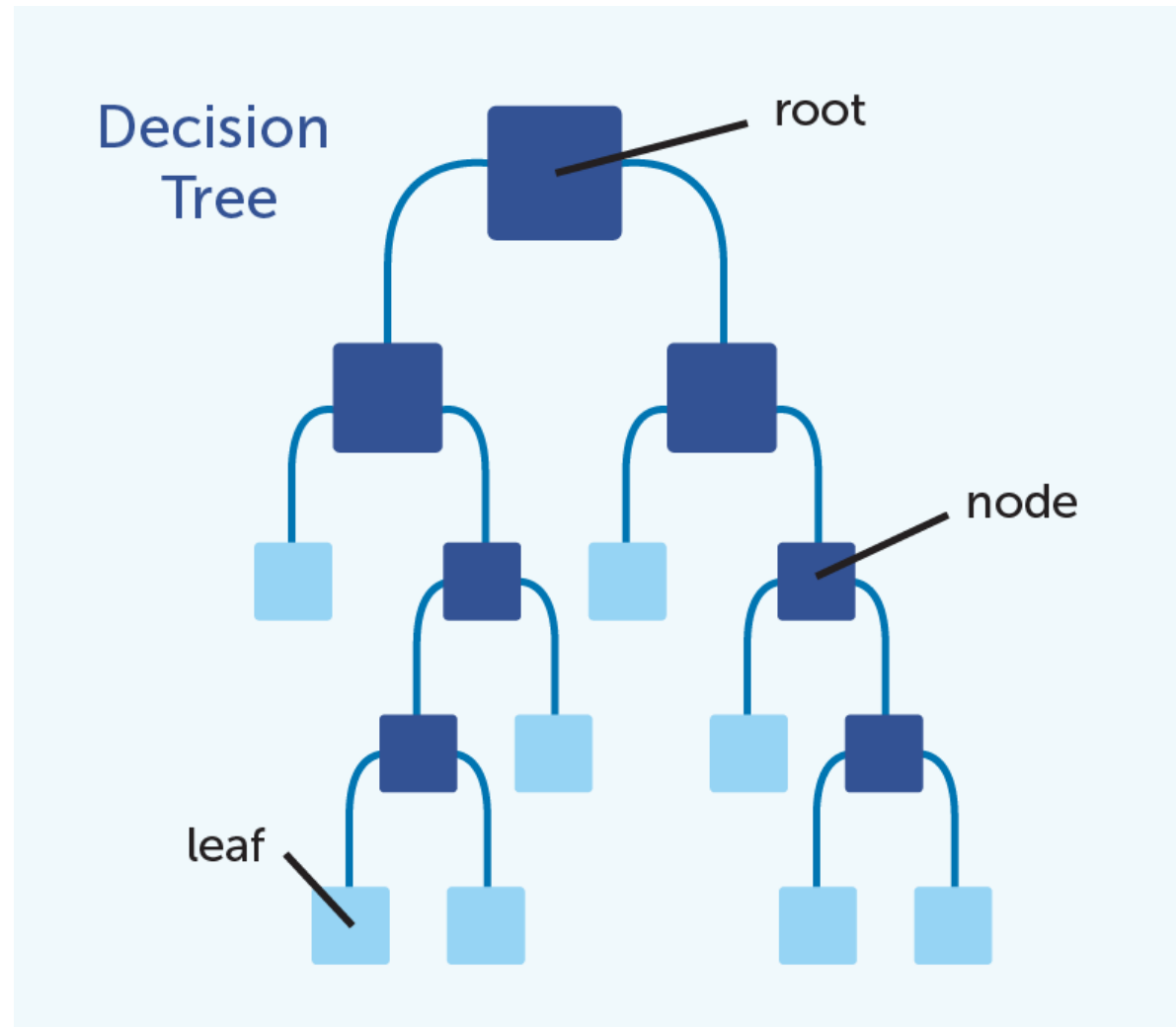
2) Мел-частотна гребінка

3) Отримання питомої енергії

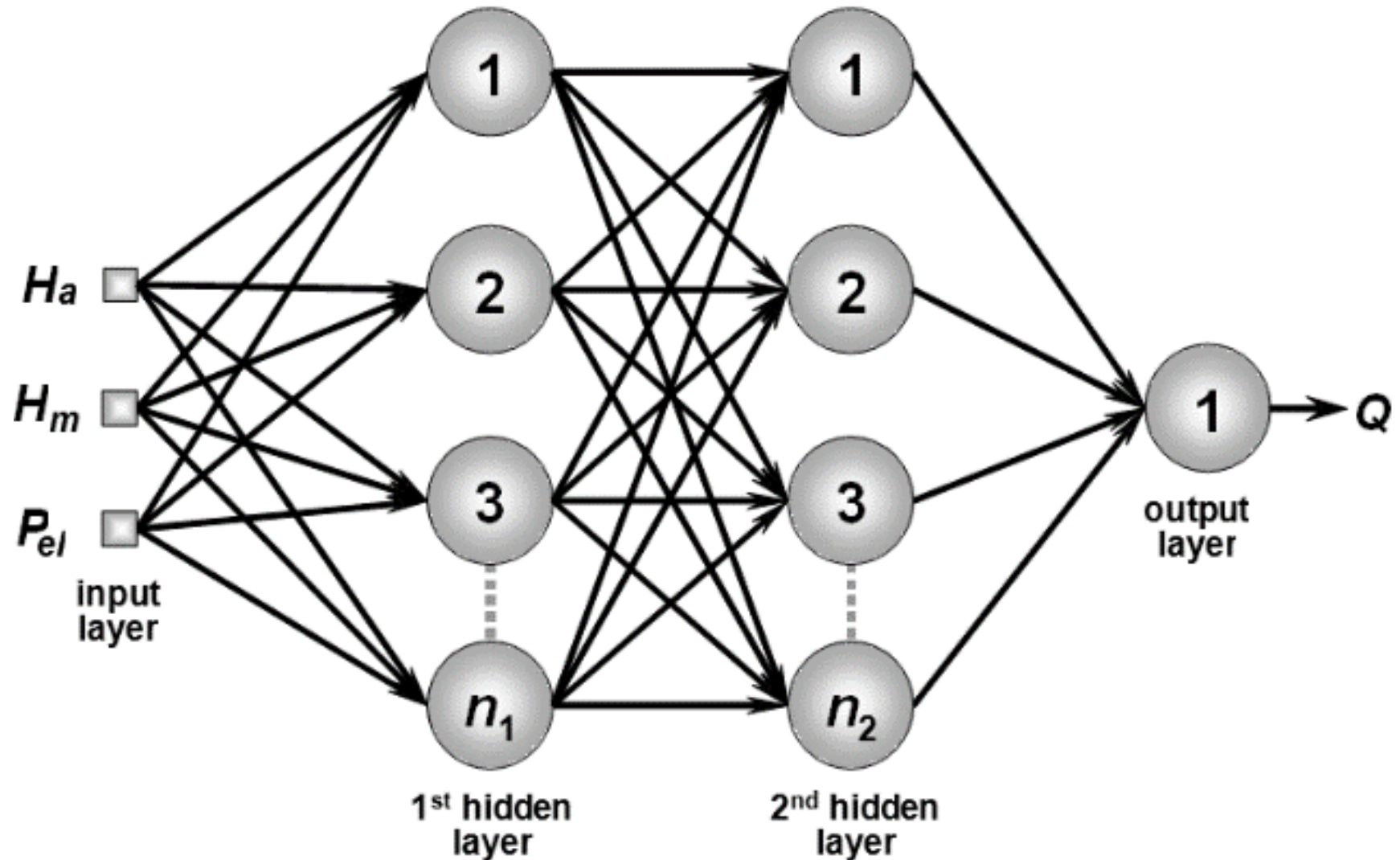
$$E_D = \sum_{i=0}^{N-1} s^2(x_i)$$



Етап 3. Класифікація (Random Forest)



Етап 3. Багатошаровий перцептрон



Етап 4. Зважене голосування

Критерії оцінки моделі

1. Процент правильно ідентифікованих дикторів.
2. Коефіцієнт достовірності ідентифікації диктора.

Аналіз результатів (власний датасет з 10 дикторів)

Назва методу	Критерій 1	Критерій 2
Випадковий ліс	100%	59.4%
Багатошаровий перцептрон Румельхарта	100%	86.6%
Розроблена композиція двох методів	100%	96.0%
Розроблена композиція двох методів з власною модифікацією	100%	97.4%

Аналіз результатів (сторонній датасет з 8 дикторів)

	Критерій 1	Критерій 2
Розроблена композиція двох методів з власною модифікацією	75%	80%

ВИСНОВКИ

- Запропоновано та реалізовано власний метод виділення ознак, як модифікацію методу мел-кепстральних коефіцієнтів, що підвищило точність та швидкодію результуючої моделі.
- Запропоновано власний метод бустингу між випадковим лісом та багатошаровим перцептроном Румельхарта, як спосіб підвищення точності моделі. Проаналізовано отримані результати. Порівняно результати роботи композиції класифікаторів з результатами кожної моделі окремо. Зроблено висновок про доцільність використання бустингу. Порівняно результати роботи звичайної композиції класифікаторів та власної модифікації композиції класифікаторів. Зроблено висновок про доцільність власної модифікації та про її позитивний вплив на точність моделі.
- Розроблено програмний продукт, що дає змогу розпізнавати дикторів за їхнім голосом та працює з точністю більше 95% на вибірках з записам, довшими за 1 секунду, та 75% на коротших записах.