

Магістерська кваліфікаційна робота на тему:

# ІДЕНТИФІКАЦІЯ КОРИСТУВАЧА ЗА ГОЛОСОМ НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ

Виконав студент гр. 1КН-15м

Семенюк О.В.

Науковий керівник: к.т.н., доц., Колесницький О.К.

## МЕТА ТА ЗАВДАННЯ ДОСЛІДЖЕННЯ

Метою дослідження магістерської кваліфікаційної роботи є підвищення достовірності ідентифікації користувача за голосом за рахунок використання нейронних мереж.

**Для досягнення поставленої мети необхідно розв'язати такі завдання:**

- провести аналіз проблеми розв'язання задачі ідентифікації користувача за голосом;
- розглянути існуючі методи вирішення задачі ідентифікації користувача за голосом та обрати й обґрунтувати вибір методу, який задовольняє мету даної роботи;
- розробити математичну модель обробки звукового сигналу;
- розробити структуру інформаційної системи ідентифікації користувача за голосом;
- сформулювати стадії інформаційної технології, розробити структуру та алгоритм роботи програмного засобу;
- виконати програмну реалізацію запропонованої інформаційної технології;
- провести тестування програмного продукту та виконати аналіз отриманих результатів.

## ОБ'ЄКТ, ПРЕДМЕТ ТА МЕТОДИ ДОСЛІДЖЕННЯ

Об'єкт дослідження – процес ідентифікації користувача за голосом з використанням інформаційних технологій.

Предмет дослідження – інформаційна технологія та програмні засоби ідентифікації користувача за голосом на основі нейронних мереж та достовірність їх роботи.

### Методи дослідження

У роботі використані наступні методи наукових досліджень:

- системного аналізу для аналізу структури інформаційної системи,
- теорії нейронних мереж для реалізації інформаційної технології ідентифікації користувача за голосом,
- методи математичної статистики для розробки процесу пошуку та обрахунків результатів експериментів із програмним засобом,
- об'єктно-орієнтованого програмування для програмної реалізації.

## НАУКОВА НОВИЗНА ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

полягає в наступному:

-- знайшла подальшого розвитку інформаційна технологія ідентифікації користувача за голосом за рахунок використання штучної нейронної мережі зустрічного поширення для розпізнавання ознак голосу користувача, що дозволило підвищити достовірність ідентифікації користувача;

-- удосконалено метод навчання нейронної мережі зустрічного поширення за рахунок введення правила «бажання працювати» та коефіцієнту  $\alpha_1$  - швидкість модифікації, що дозволило підвищити точність навчання нейронної мережі.

## ПРАКТИЧНЕ ЗНАЧЕННЯ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

1. Розроблено алгоритм ідентифікації користувача за голосом на основі нейронної мережі.
2. Розроблено алгоритм функціонування нейронної мережі для підвищення достовірності ідентифікації користувача за голосом.
3. Розроблено програмний засіб для ідентифікації користувача за голосом на основі нейронної мережі зустрічного поширення.

Розроблені алгоритми можуть бути впроваджені в начальний процес як лекція на тему «Нейромережевий метод ідентифікації користувача за голосом» дисципліни «Нейромережеві методи обчислювального інтелекту».

# Аналіз предметної області

Розпізнавання мови поділяється на три типи:

- розпізнавання окремих виділених слів;
- розпізнавання зливої мови;
- розпізнавання за зразком.



Механізм розпізнавання для перших двох типів



Механізм розпізнавання для третього типу

**Для вирішення задачі ідентифікації користувача було обрано змішаний тип розпізнавання, тобто розпізнавання за зразком та розпізнавання за окремими виділеними словами задля досягнення точнішого результату.**

Для зменшення шумів та полегшення ідентифікації користувача за голосом застосовують фільтри.

В результаті обробки отримуємо майже чітко виділений спектр голосу користувача з яким легко порівнювати наступні зразки.

# ОБҐРУНТУВАННЯ ВИБОРУ МЕТОДУ РОЗВ'ЯЗАННЯ ЗАДАЧІ

## Відомі методи ідентифікації користувача за голосом:

- ідентифікація за допомогою дерев рішень;
- ідентифікація методом опорних векторів;
- статистичні методи, зокрема , лінійна регресія;
- ідентифікація за допомогою методу "найближчого сусіда";
- ідентифікація за допомогою генетичних алгоритмів;
- **ідентифікація за допомогою штучних нейронних мереж .**

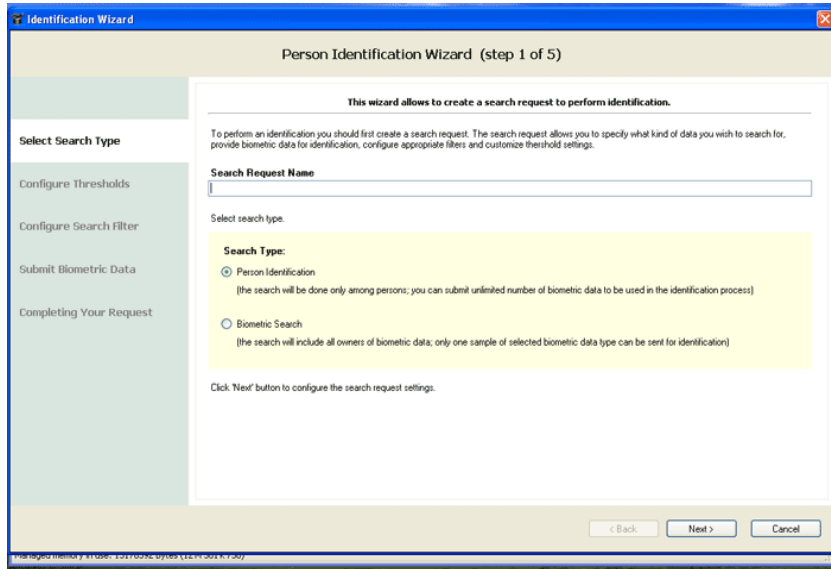
## Недоліки відомих методів

- 1) великий обсяг пам'яті, зростає пропорційно квадрату кількості прикладів для навчання
- 2) Велика обчислювальна трудомісткість (особливо при повному переборі навчальної вибірки при ідентифікації).

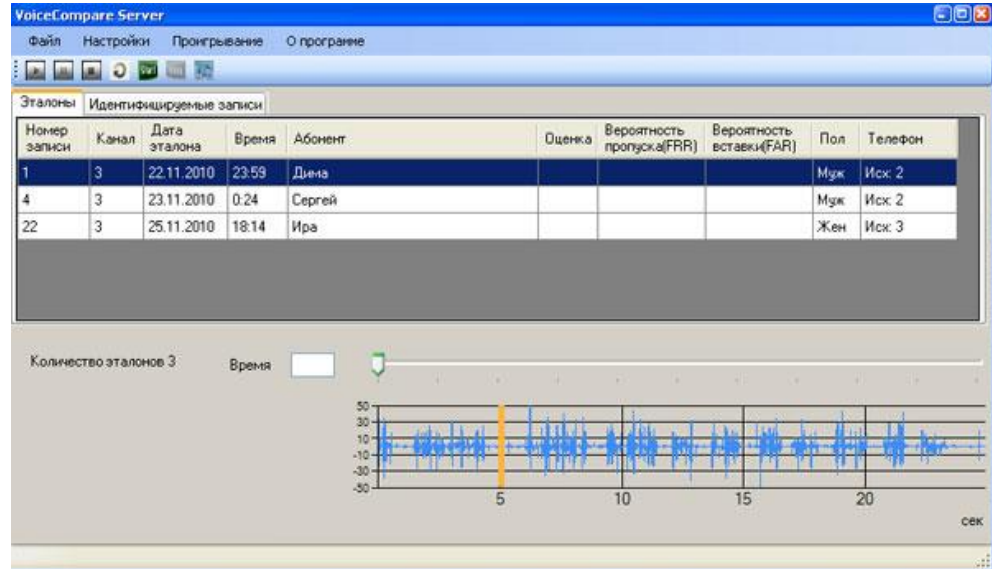
## Обрано метод нейронних мереж. Переваги методу:

- 1) Узагальнююча здатність нейронних мереж дозволяє ідентифікувати зашумлені сигнали.
- 2) Висока швидкість розпізнавання, але перед цим треба витратити час на навчання мережі

# АНАЛІЗ ІСНЮЮЧИХ ПРОГРАМНИХ ЗАСОБІВ ІДЕНТИФІКАЦІЇ КОРИСТУВАЧА ЗА ГОЛОСОМ

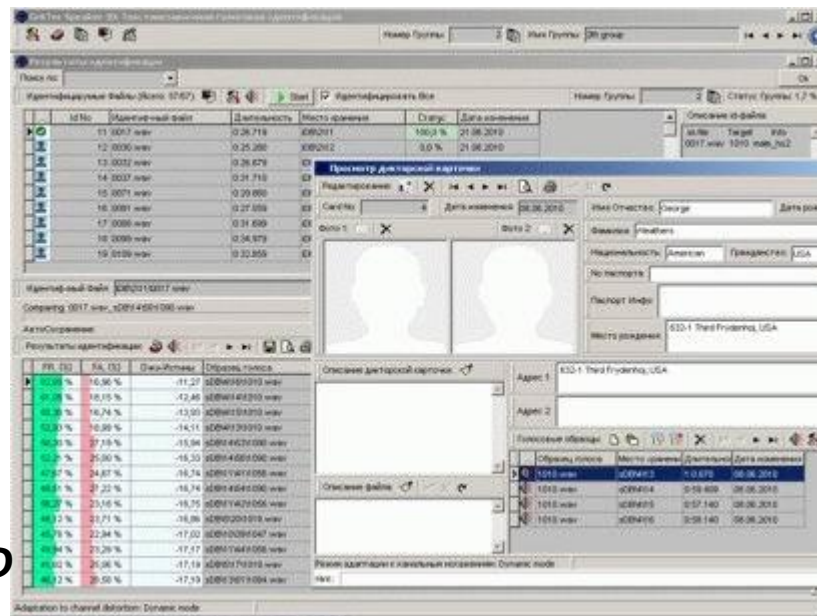


Програма  
BioLink AMIS

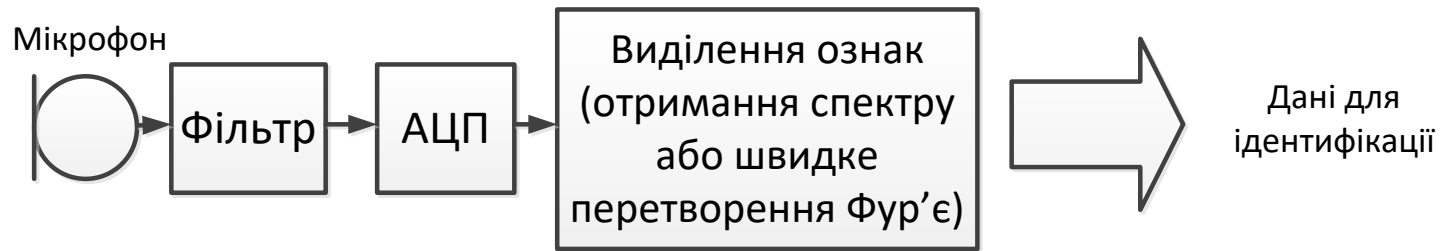


Програма  
VoiceCompare

Програма  
GritTec Speaker-ID



# Схема попередньої обробки звукового сигналу для ідентифікації користувача



## Фільтрація голосового сигналу

- За допомогою функції Кайзера :  
$$I_{01} = \frac{1}{I_0(b)} ;$$
$$X_i = X_i * \left[ b * \sqrt{1 - \left( \frac{2 * k}{n - 1} \right)^2} \right] * I_{01}$$

- За допомогою функції Ханнінга:  
$$0 \leq n \leq N_s ;$$
$$a_w = 0.54 ;$$
$$W(n) = \frac{\alpha_w - (1 - \alpha_w) * \cos\left(\frac{2\pi n}{N_s - 1}\right)}{\beta_w}$$



# Обробка звукового сигналу

- Для представлення звукового сигналу у вигляді спектру застосовують дискретне перетворення Фур'є.

$$A_f = \sum_{t=0}^{N/2-1} s_{2t} e^{-2\pi f t / (N/2)} + e^{-2\pi f / N} * \sum_{t=0}^{N/2-1} s_{2t+1} e^{-2\pi f t / (N/2)}$$

- Отримавши спектральне подання сигналу його потрібно відчистити від шумів.

# Математична модель визначення еталонного зразку голосу

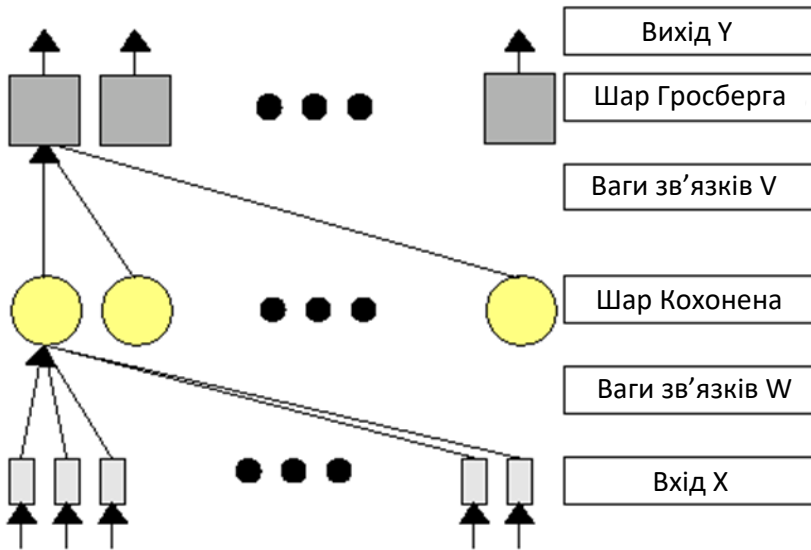
- Міра подібності двох звукових фрагментів визначається за формулою:

$$f_{xy} = \left| \frac{\sum_i (x_i - M_x)(y_i - M_y)}{\sqrt{\sum_i (x_i - M_x)^2} * \sqrt{\sum_i (y_i - M_y)^2}} \right|$$

де  $M_x$  і  $M_y$  – математичне очікування для масивів  $X[]$  і  $Y[]$  відповідно, обчислюються за такою формулою:

$$M_z = \frac{1}{N} \sum_1^N z_i$$

# Вдосконалення математичної моделі навчання нейронної мережі зустрічного поширення



## Архітектура нейронної мережі зустрічного поширення

В даній роботі застосовується правило «бажання працювати»: якщо який-небудь нейрон довго не знаходиться в активному стані, він підвищує ваги зв'язків до тих пір, поки не стане активним і не почне піддаватися навчанню. Цей метод дозволяє також вирішити проблему тонкої класифікації: якщо утвориться група вхідних сигналів, розташованих близько один до одного, з цією групою асоціюється і велике число нейронів Кохонена, які розбивають її на класи. Правило «бажання працювати» записується в наступній формі:

$$w_n = w_c + w_c \alpha_1 (1 - a), \quad (2.17)$$

де  $w_n$  - нове значення ваги,  $w_c$  - старе значення,  $\alpha_1$  - швидкість модифікації,  $a$  - активність нейрона.

Чим менше активність нейрона, тим більше збільшуються ваги зв'язків. Вибір коефіцієнта  $\alpha_1$  визначається такими міркуваннями: постійне зростання ваг нейронів за правилом (2.17) компенсується правилом (2.15) (активні нейрони прагнуть знову повернутися на гіперсферу одиничного радіуса), причому за одну ітерацію нейронмережі збільшать свою вагу практично всі нейрони, а зменшить тільки один активний нейрон або нейронний ансамбль. У зв'язку з цим коефіцієнт  $\alpha_1$  в (2.17) необхідно вибирати значно менше коефіцієнта  $\alpha$  в (2.15), враховуючи при цьому 1 число нейронів у шарі

Навчання нейронів шару Кохонена проводиться за правилом:

$$w_n = w_c + \alpha (x - w_c) a, \quad (2.15)$$

$w_n$  - нове значення ваги,

$w_c$  - старе значення,

$\alpha$  - швидкість навчання,  $\alpha < 1$

$x$  - нормований вхідний вектор,

$a$  - активність нейрона.

Навчання нейронів шару Гросберга проводиться за правилом:

$$w_{ijn} = w_{ijc} + \beta (y_j - w_{ijc}) x_i, \quad (2.16)$$

$w_{ijn}$ ,  $w_{ijc}$  - ваги зв'язків до і після модифікації

$\beta$  - швидкість навчання,  $\beta < 1$

$y_j$  - вихід нейрона

$x_i$  - вхід нейрона

# СТРУКТУРА ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ІДЕНТИФІКАЦІЇ КОРИСТУВАЧА ЗА ГОЛОСОМ



# ОБГРУНТУВАННЯ ВИБОРУ МОВИ ПРОГРАМУВАННЯ

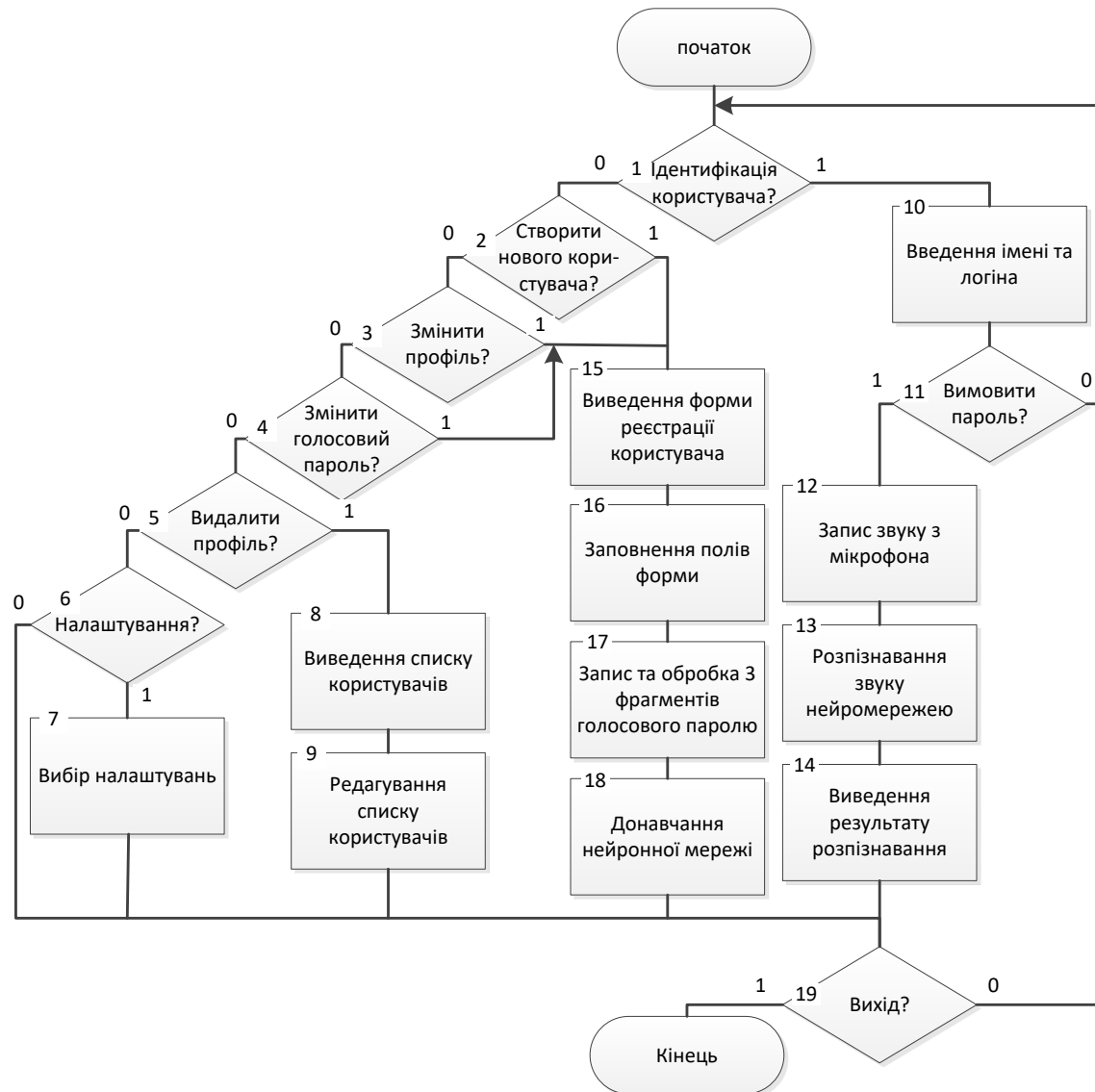
- Для написання програми обрано мову програмування. С#, реалізація програмного продукту велась в середовищі розробки Microsoft Visual Studio 2015 (надане кафедрою КН в межах програми DreamSpark (MSDN AA))

## ЗАГАЛЬНА СТРУКТУРА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

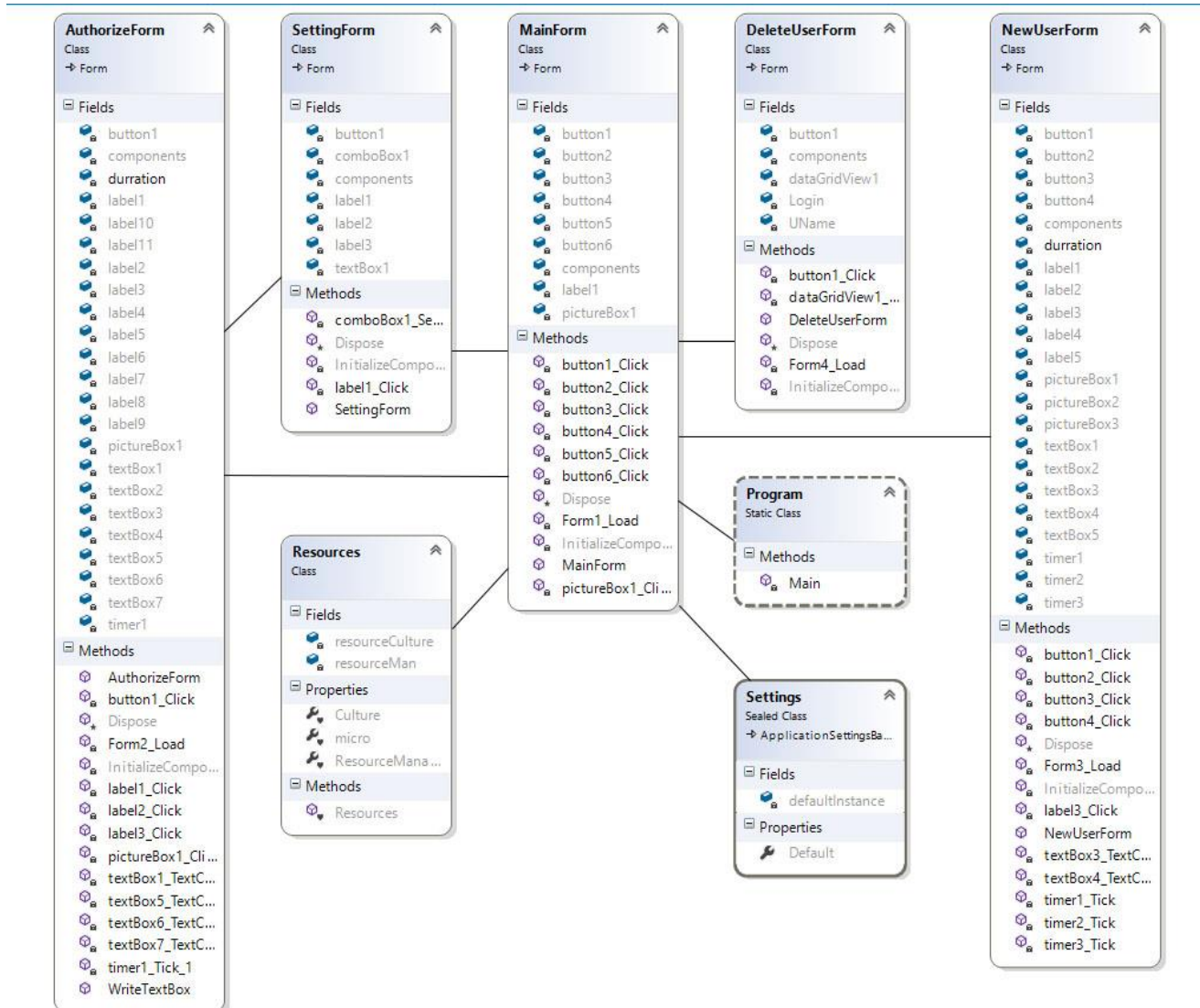
Програма складається з діалогів:

- додавання користувача;
- зміни параметрів зареєстрованого в системі користувача;
- налаштування програми;
- налаштування параметрів роботи адміністратора;
- перегляду статистики роботи програми;
- видалення користувача зареєстрованого в системі.

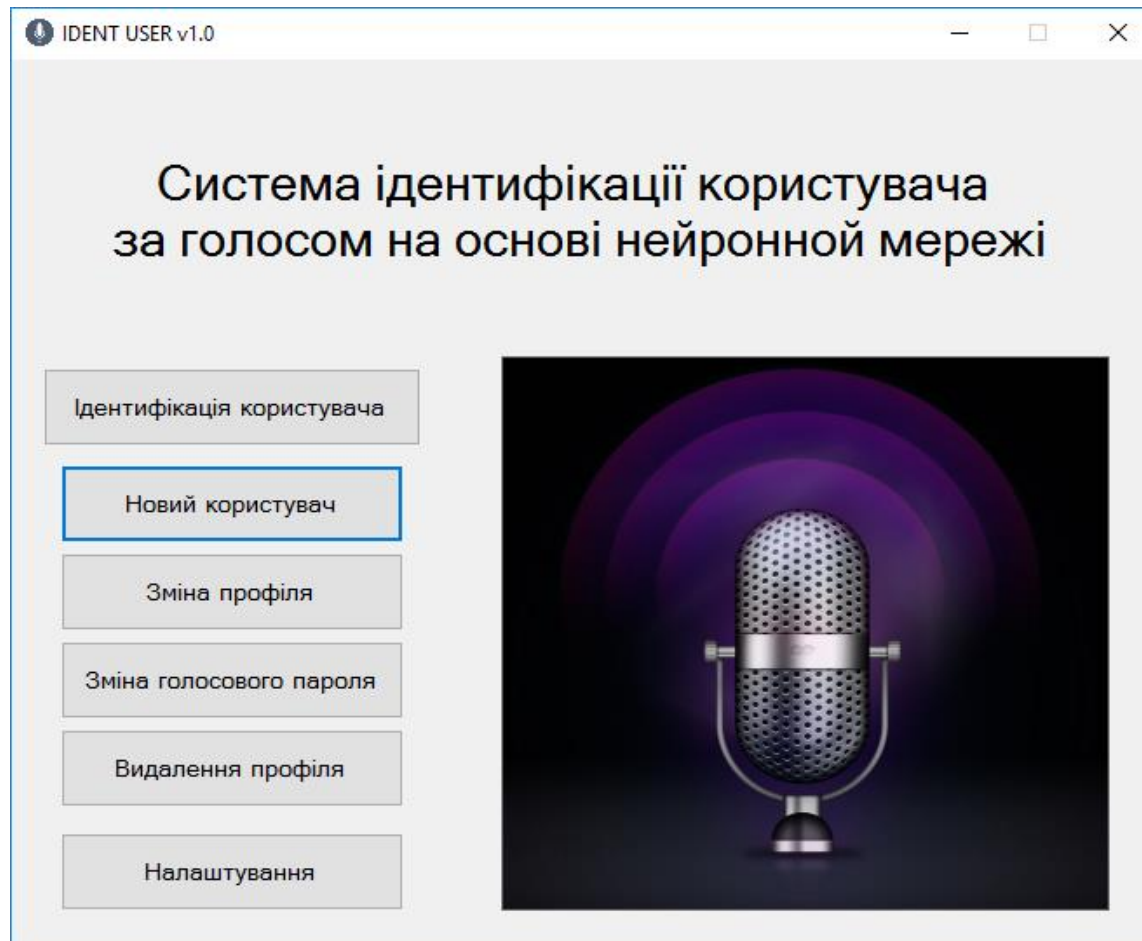
# СХЕМА АЛГОРИТМУ РОБОТИ ПРОГРАМИ ІДЕНТИФІКАЦІЇ КОРИСТУВАЧА ЗА ГОЛОСОМ



# ДІАГРАМА КЛАСІВ ПРОГРАМИ ІДЕНТИФІКАЦІЇ КОРИСТУВАЧА ЗА ГОЛОСОМ



# Головне вікно програми ідентифікації користувача за голосом





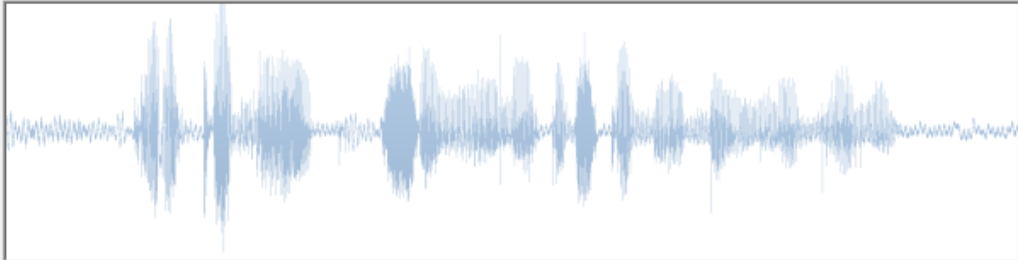
# Вікно авторизації zareєстрованого користувача

Ідентифікація Користувача

Введіть Ваше Ім'я:

Введіть Ваш Логін:

Скажіть пароль  сек.



Відповідність збереженим паролем:

1 -  %      2 -  %      3 -  %

# Робочі вікна програми ідентифікації користувача за голосом

Управління профілями

Введіть Своє Ім'я:

Введіть Свій Логін:

Записати Пароль 1

Записати Пароль 2

Записати Пароль 3

Вікно реєстрації нових користувачів

Видалення корист...

UName	Login
Chester	Cuceman
Yuzef	PROKARIOT
Dart	VoiceR
Fidel	RedminD
Sadam	Solid
Snake	Metal

Вікно «Видалення профілю»

Налаштування

Період запису голосового пароллю:

Мінімальна відповідність пароллю, необхідна для успішної авторизації:  %

Вікно «Налаштування»

# Порівняння результатів роботи розробленої програми та програми-аналога

Показники	Програма-аналог (GritTec Speaker-ID)	Розроблений модуль
тестованих користувачів	8	8
Кількість різних способів вимови однакової фрази кожним користувачами	15	15
Загальний обсяг тестового набору	120	120
Кількість вірно ідентифікованих фраз	98	109
Достовірність розпізнавання	81,7%	90,8%

Із таблиці видно, що розроблена програма має кращі показники достовірності ідентифікації користувача за голосом (90,8%), ніж аналог (81,7%). Таким чином, поставлена у роботі мета досягнута

## ЕКОНОМІЧНА ЧАСТИНА

Було виконано оцінювання комерційного потенціалу розробки інформаційної технології та програмного засобу ідентифікації користувача за голосом на основі штучної нейронної мережі.

- Загальні витрати на виконання науково-дослідної роботи - 78 058 грн.
- Приведена вартість всіх чистих прибутків підприємства від реалізації результатів наукової розробки - 563 113 грн.
- Абсолютна ефективність вкладених інвестицій - 485055 грн.
- Щорічна ефективність вкладених в наукову розробку інвестицій -83,8%.
- Термін окупності вкладених у реалізацію проекту інвестицій 1,19 років, що свідчить про доцільність фінансування нової розробки.

# АПРОБАЦІЯ РЕЗУЛЬТАТІВ РОБОТИ ТА ПУБЛІКАЦІЇ

## **Апробація результатів роботи.**

Результати роботи були апробовані на -- науково-практичній конференції «Сучасні тенденції розвитку системного програмування» (25-26 листопада 2016 р.), м. Київ [1],

**Публікації.** За результатами магістерської кваліфікаційної роботи опубліковано 1 тези доповіді на конференції [1].

# Висновки

В результаті виконання роботи було розроблено інформаційну технологію ідентифікації користувача за голосом. За допомогою розробленої інформаційної технології та програмних засобів користувач може захистити важливу інформацію від несанкціонованого доступу. Розроблена інформаційна технологія забезпечує високий рівень захисту, оскільки має достовірність ідентифікації 90,8%, а програма-аналог – 81,7%, що свідчить про досягнення мети, яка полягала у підвищенні достовірності ідентифікації користувача за голосом.

Дякую за увагу.