

Програма автентифікації користувача по клавіатурному почерку.

Програма працює в двох режимах:

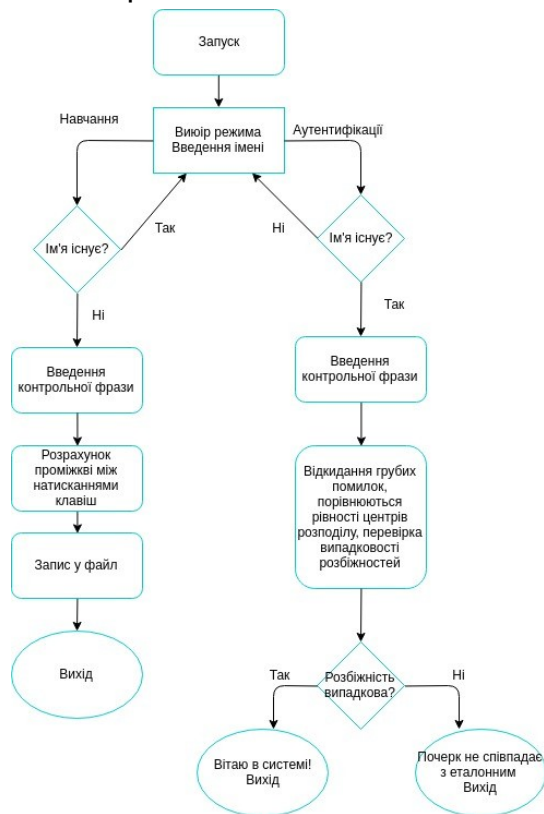
Навчання (створення біометричного ідентифікатора).

Ідентифікація (порівняння з біометричним ідентифікатором).

Тимчасові інтервали між натисканням сусідніх букв при наборі ключової фрази, як правило, підпорядковуються нормальному закону розподілення.

Автентифікація користувача здійснюється шляхом вирішення задачі виключення грубих помилок у спостереженнях і рішення задачі перевірки гіпотези про рівність центрів розподілу двох нормальних генеральних сукупностей.

Схема роботи:



Алгоритм рішення задачі виключення грубих помилок у спостереженнях:

Нехай $y = \{y_1, y_2, \dots, y_n\}$ - множина тимчасових інтервалів між натисканням клавіш, p - число тимчасових інтервалів.

Необхідно вирішити задачу перевірки значимості елементів y_i множини y ($i = 1, \dots, n$). Для цього сформуємо множину $y' = y \setminus \{y_i\}$, тобто всі математичні параметри розраховуються без урахування елемента y_i в вихідній множині y .

n' - число елементів у множині y' , $n' = n - 1$.
математичне очікування M_i

$$M_i = \frac{\sum_{k=1}^{n'} y'_k}{n'};$$

Розраховуємо дисперсію

$$S_i^2 = \frac{\sum_{k=1}^{n'} (y_k - M_i)^2}{n' - 1};$$

Розраховуємо середньоквадратичне відхилення

$$S_i = \sqrt{S_i^2};$$

Коефіцієнт Стюдента t_p

$$t_p = \left| \frac{y_i - M_i}{S_i} \right|;$$

Для числа ступенів свободи $n' - 1$ і рівня значимості $\alpha = 0.05$ визначаємо табличний коефіцієнт Стюдента t_T . Якщо $t_p > t_T$ то елемент y_i - відкидається і розрахунок починається знову до тих пір, поки $i \neq n$. В іншому випадку елемент y_i оголошується значимим.

Алгоритм рішення задачі перевірки гіпотези про рівність центрів розподілу двох нормальних генеральних сукупностей.

Вирішуємо задачу перевірки однорідності двох вибірових дисперсій. Нехай n - число елементів у множині тимчасових інтервалів.

s_1^2 - Вибіркова дисперсія для еталонних параметрів.

s_2 - Вибіркова дисперсія для безлічі тимчасових інтервалів, отримана в режимі автентифікації, тоді

$$S_{\max}^2 = \max(S_1^2, S_2^2); \quad S_{\min}^2 = \min(S_1^2, S_2^2);$$

Розраховуємо коефіцієнт Фішера - розрахунковий **F_p**

$$F_p = \frac{S_{\max}^2}{S_{\min}^2};$$

Для числа ступенів свободи **n-1** і рівня значимості **α = 0.05** визначаємо табличний коефіцієнт Фішера **FT**. Якщо **F_p > FT**, то дисперсії неоднорідні, в іншому випадку дисперсії рівні.

Вирішуємо задачу перевірки гіпотези про рівність центрів розподілу двох нормальних генеральних сукупностей.

Нехай **y = {y1, y2, ..., yn}** - множина тимчасових інтервалів між натисканням клавіш отриманих в режимі автентифікації, **n** - число тимчасових інтервалів. Нехай **ke** - число еталонних спроб набору ключової фрази,

xα = {x1, x2, ..., xn} - множина еталонних часових інтервалів (**λ = 1, k**), тоді

$$M_{x_{\lambda}} = \frac{\sum_{i=1}^n x_{\lambda_i}}{n}, \quad M_y = \frac{\sum_{i=1}^n y_i}{n};$$

Тепер необхідно перевірити, чи можна з певною ймовірністю **p** вважати, що розбіжність між **Mxy** і **My** викликані випадковими причинами. Для цього по таблиці розподілу Стюдента (рівень значимості **α = 1 - p**, число ступенів свободи **n-1**) визначимо табличне значення **tT**, після чого розрахуємо величину **tp**. Якщо **tp > tT**, то розбіжність не випадкова.

$$S_{x_{\lambda}}^2 = \frac{\sum_{i=1}^n (x_{\lambda_i} - M_{x_{\lambda}})^2}{n-1}; \quad S_y^2 = \frac{\sum_{i=1}^n (y_i - M_y)^2}{n-1};$$

$$S = \sqrt{\frac{((S_{x_{\lambda}}^2)^2 + (S_y^2)^2) \cdot (n-1)}{2n-1}};$$

$$t_p = \frac{|M_{x_{\lambda}} - M_y|}{S \cdot \sqrt{\frac{2}{n}}};$$