

Аналіз жестів руки: k-NN та k-means

Шпагін Павло

10 лютого 2026 р.

1. Постановка задачі

Об'єкт дослідження: гіпотетичне зображення руки.

Вектор ознак: (w, h) – ширина та висота долоні (см).

Класи: Кулак, Долоня (розкрита).

2. Генерація даних

Дані згенеровано з нормального розподілу для двох базових жестів: **Кулак** ($w \sim N(8, 0.8)$, $h \sim N(4, 0.9)$; 30 зразків) та **Долоня** ($w \sim N(16, 1.5)$, $h \sim N(16, 1.8)$; 30 зразків).

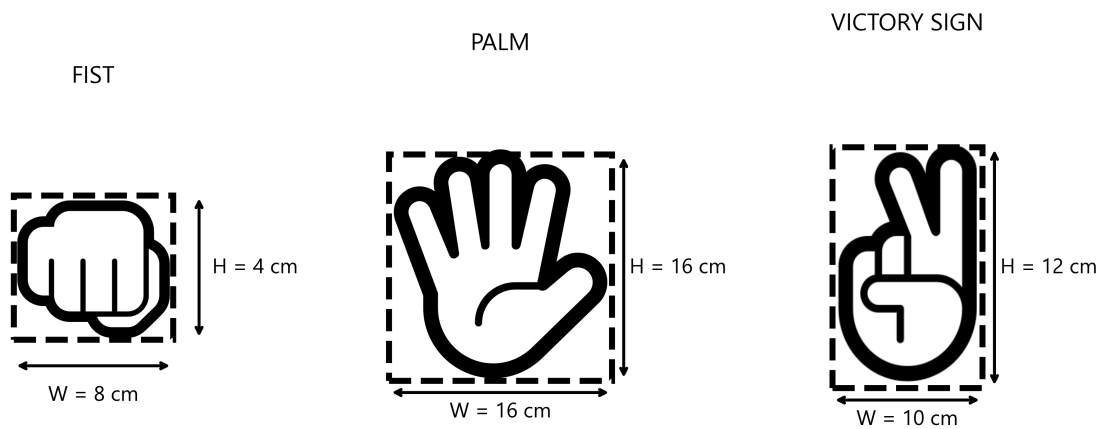


Рис. 1: Приклади обмежувальних рамок для жестів руки

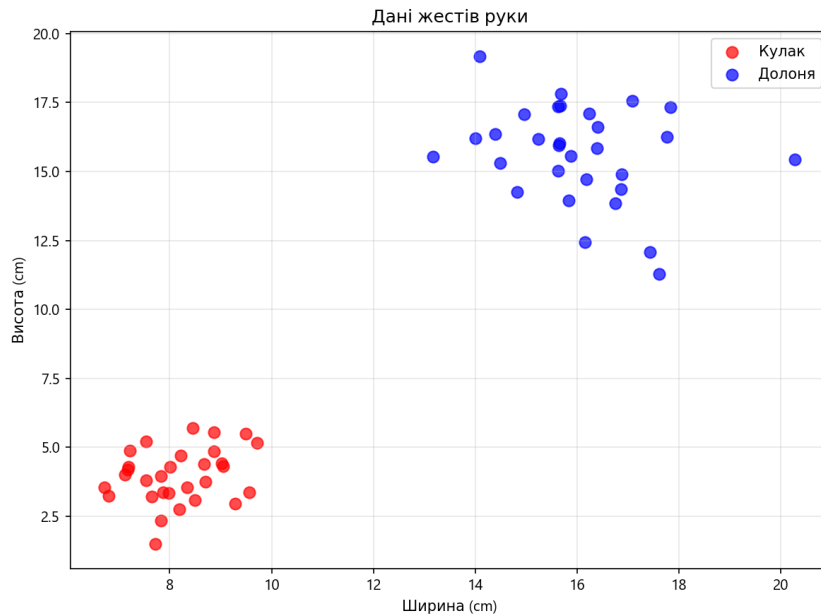


Рис. 2: Розподіл даних у просторі ознак

3. k-NN класифікація

Реалізовано алгоритм k-NN без використання бібліотек. Відстань – евклідова:

$$d(p_1, p_2) = \sqrt{(w_1 - w_2)^2 + (h_1 - h_2)^2} \quad (1)$$

Експериментальна перевірка

Додатково проведено перевірку точності на **окремо згенерованих тестових даних** (навчальна та тестова вибірки генерувалися незалежно). Порівнювалися два сценарії: (і) класифікація двох класів (Кулак/Долоня) та (іі) класифікація трьох класів після додавання “Знаку перемоги” до тренувальної і тестової вибірок.

Сценарій	k	Точність
2 класи (Кулак/Долоня)	1	100.0%
2 класи (Кулак/Долоня)	3	100.0%
2 класи (Кулак/Долоня)	5	100.0%
2 класи (Кулак/Долоня)	7	100.0%
3 класи (+ Знак перемоги)	1	98.9%
3 класи (+ Знак перемоги)	3	98.9%
3 класи (+ Знак перемоги)	5	98.9%
3 класи (+ Знак перемоги)	7	98.9%

Табл. 1: Точність k-NN на незалежних тестових вибірках для 2 та 3 класів

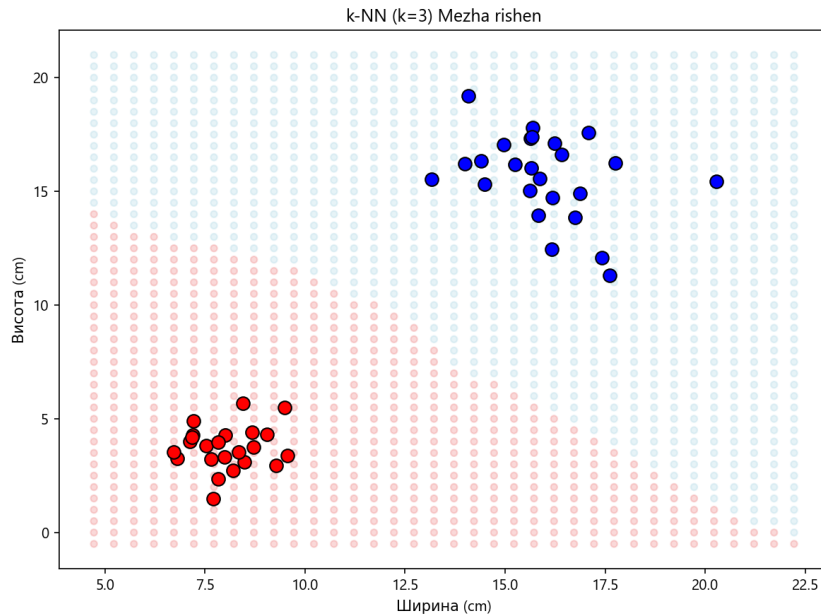


Рис. 3: Межа рішень k-NN (k=3)

4. Проблема класу “Знак перемоги”

Якщо збільшити кількість класів у задачі k-NN з 2 до 3 та додати клас “Знак перемоги” ($w \sim N(10, 1.5)$, $h \sim N(12, 2.0)$), то класифікація в просторі ознак (w, h) ускладнюється: у деяких випадках області класів можуть перекриватися, а “найближчі сусіди” для нового жесту виявляються змішаними.

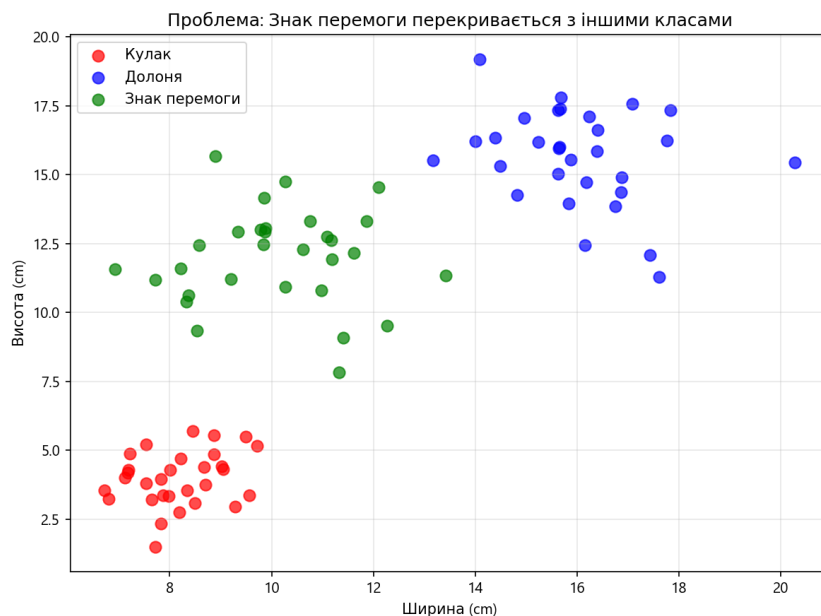


Рис. 4: Перекриття класів при додаванні знаку перемоги

Практичне вирішення для k-NN полягає в тому, щоб додати приклади жесту “Знак перемоги” до тренувальної вибірки як окремий клас. Алгоритм k-NN є багатокласовим за своєю природою (через голосування сусідів), тому він може працювати і з 3+ класами за умови, що тренувальні дані містять приклади кожного класу. З експериментальної

перевірки видно, що після додавання третього класу якість залежить від вибору k і від перекриття розподілів у просторі ознак.

Водночас, використання лише 2D-ознак (w, h) є обмеженням: вони описують розмір, але не геометрію жесту. Для підвищення якості розпізнавання доцільно розширювати вектор ознак (наприклад, кількість пальців, кути між пальцями, співвідношення сторін, морфологічні характеристики контурів).

5. k-means кластеризація

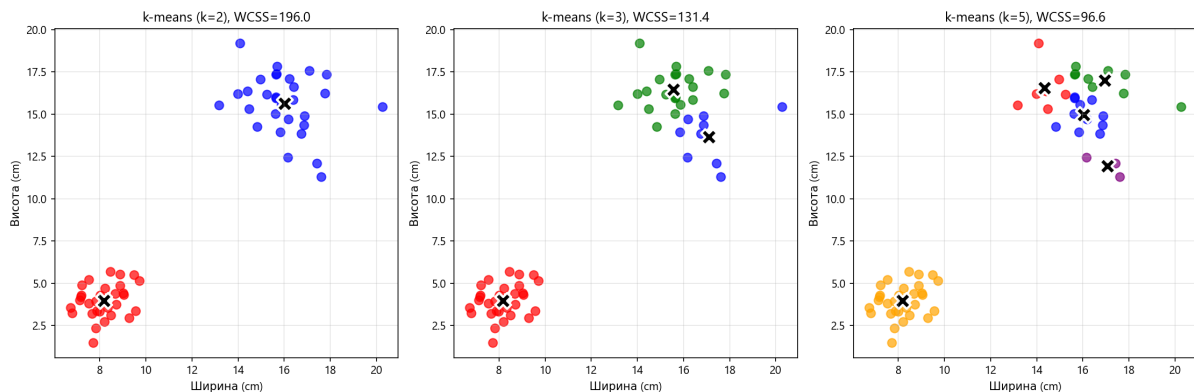


Рис. 5: Результати k-means для $k=2,3,5$

k	WCSS
2	196.0
3	131.4
5	96.6

Табл. 2: Within-Cluster Sum of Squares

6. Висновки

Отримані результати показують, що k-NN ефективно розділяє два базові класи у 2D просторі ознак. Після додавання третього класу точність залишається високою, але стає більш чутливою до вибору параметра k та до перекриття класів. Метод k-means підтверджує природну двокластерну структуру базових даних ($k=2$). Загалом, для класифікації більшої кількості жестів доцільно використовувати розширений набір ознак, який описує конфігурацію пальців, а не лише габаритні розміри.