

# КЛАСИФІКАЦІЯ МНОЖИНИ 3D-ТОЧОК НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ

Вінницький національний технічний університет;

## Анотація

*У роботі розглянуто підходи до класифікації тривимірних точкових хмар із використанням алгоритмів машинного навчання. Описано особливості побудови моделей, що працюють з 3D-даними, а також проаналізовано переваги та недоліки методів KNN та Random Forest у задачах просторової класифікації.*

**Ключові слова:** класифікація, 3D точки, машинне навчання, KNN, Random Forest, хмара точок.

## Abstract

*The paper presents approaches to the classification of 3D point clouds using machine learning algorithms. It outlines the modeling challenges associated with spatial data and compares KNN and Random Forest methods in terms of their efficiency and accuracy.*

**Keywords:** classification, 3D points, machine learning, KNN, Random Forest, point cloud.

## Вступ

Класифікація множини 3D-точок є важливою задачею в комп'ютерному зорі, геоінформатиці, архітектурному скануванні, робототехніці та інших сферах, де обробка просторових даних має вирішальне значення. Основна мета такої класифікації полягає у присвоєнні кожній точці хмари певного класу, що відповідає її фізичному або функціональному призначенню (наприклад, «стіна», «земля», «рослинність» тощо).

Тривимірні точки, отримані шляхом лазерного сканування або фотограмметрії, не мають явно виражених ознак, тому їх класифікація є складною задачею. Сучасні методи машинного навчання, зокрема алгоритми K найближчих сусідів та Random Forest, дозволяють виконувати таку класифікацію шляхом аналізу просторових ознак: координат, нормалей, щільності тощо.

Для класифікації хмари точок спочатку здійснюється попередня обробка: видалення шуму, нормалізація координат, виділення локальних дескрипторів (напр., PCA-ознаки, розрахунок нормалей або щільності точки у локальному вікні). Далі формується навчальна вибірка з мітками для тренування моделей.

Метод, K найближчих сусідів, класифікує точку за більшістю класів серед K найближчих до неї сусідів. Він є інтуїтивно простим, але має обмеження при великих наборах даних через високу обчислювальну складність.

Random Forest – ансамблевий метод, що будує багато рішень дерев і використовує голосування для класифікації. Він менш чутливий до шуму та ефективніший при великій кількості ознак. У випадку 3D-точок часто використовується спільно з PCA-компонентами або іншими геометричними ознаками.

## Результати дослідження

В рамках експерименту було реалізовано класифікацію на тестовій хмарі точок, отриманій із відкритих датасетів. Було обрано два популярних алгоритми – Random Forest та KNN – для оцінки їх ефективності в умовах зашумлених даних. Кожен з алгоритмів було протестовано на одному й тому ж наборі точок з однаковими параметрами передобробки. Для кожного запуску обчислювалася середня точність класифікації. Діаграма порівняння точності алгоритмів зображена на рисунку 1.

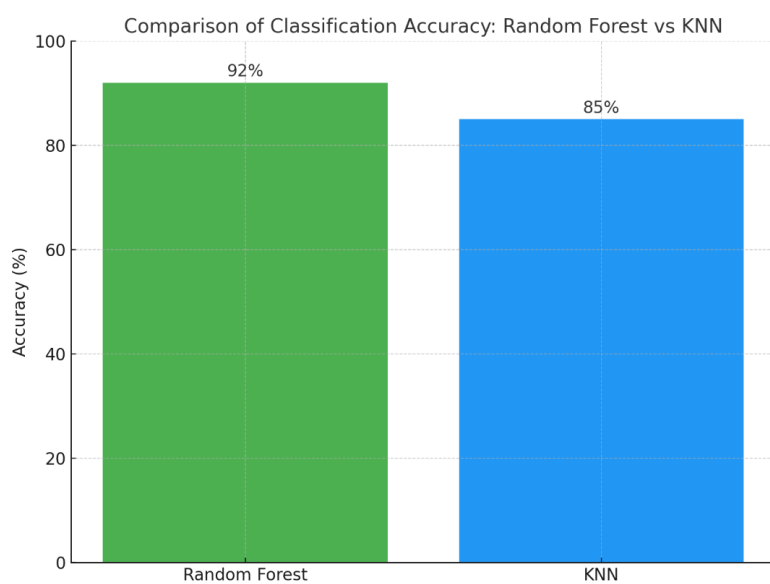


Рисунок 1 – Діаграма порівняння точності алгоритмів

Алгоритм Random Forest показав точність до 92%, порівняно з KNN, близько 85%, особливо у випадках із великою кількістю шуму або при складній геометричній структурі.

### Висновки

Класифікація множини 3D-точок – це актуальна задача в галузі комп'ютерного зору та аналізу просторової інформації. Вона дає змогу автоматично виявляти й розрізняти різні об'єкти або їх частини за формою та розташуванням у просторі. Для цього використовуються методи машинного навчання, які здатні працювати навіть з неструктурованими, зашумленими даними. Велике значення при цьому має етап попередньої обробки даних – зокрема, нормалізація координат і виділення корисних ознак.

Незважаючи на певні труднощі, пов'язані з обсягами даних і складністю розрахунків, ця тема продовжує активно розвиватися. У майбутньому очікується поява ще точніших та швидших моделей, здатних працювати з великомасштабними наборами 3D-точок. Це відкриває нові можливості для використання таких технологій у найрізноманітніших сферах – від картографії до автоматизованих систем управління.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Hackel T., Wegner J. D., Schindler K. "Fast Semantic Segmentation of 3D Point Clouds with Strongly Varying Density", ISPRS, 2016.
2. Qi C. R. et al. "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation", CVPR, 2017.
3. GeeksForGeeks. Random Forest Algorithm. URL: <https://www.geeksforgeeks.org/random-forest-classifier/>
4. GeeksForGeeks. K-Nearest Neighbors Algorithm. URL: <https://www.geeksforgeeks.org/k-nearest-neighbors/>

**Гуменюк Олександр Віталійович** — студент групи 4ПІ-216, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: nadij3122@gmail.com

Науковий керівник: **Рейда Олександр Миколайович** — к.т.н. доцент кафедри ПЗ, Вінницький національний технічний університет, м. Вінниця

**Humenyuk Oleksandr V.** — student of group 4PI-21b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: nadij3122@gmail.com

**Supervisor: Reyda Oleksandr M.** — Ph.D. associate professor of the Department of Computer Science, Vinnytsia National Technical University, Vinnytsia