CS5284 : Graph Machine Learning

Administrative (Week 4)

Semester 1 2025/26

Xavier Bresson

https://x.com/xbresson

Department of Computer Science
National University of Singapore (NUS)

QR Attendance

The International Mathematical Olympiad "IMO" is the world's most prestigious competition for young mathematicians since 1959.
Each country taking part is represented by six elite, pre-university mathematicians who compete to solve six exceptionally difficult problems in algebra, combinatorics, geometry, and number theory.
Medals are awarded to the top half of contestants, with approximately 8% receiving a prestigious gold medal.
July 2025

https://deepmind.google/discover/blog/advanced-version-of-gemini-with-deep-think-officially-achieves-gold-medal-standard-at-the-international-mathematical-olympiad/

After vibe coding, now comes vibe math !

Please, scan the new below QR image for attendance.
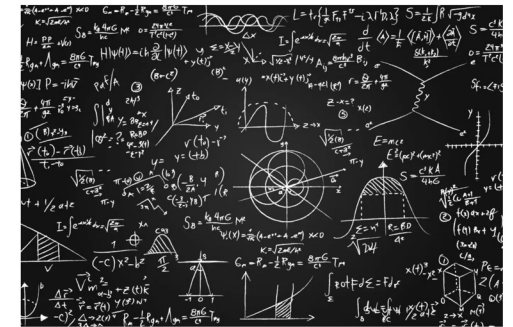
Mathematics

DeepMind and OpenAI claim gold in International Mathematical Olympiad

Two AI models have achieved gold medal standard for the first time in a prestigious competition for young mathematicians – and their developers claim these AIs could soon crack tough scientific problems

By Alex Wilkins

📅 22 July 2025

Breakthrough Performance at IMO 2025 with Gemini Deep Think

An advanced version of Gemini Deep Think solved five out of the six IMO problems perfectly, earning 35 total points, and achieving gold-medal level performance. The solutions can be found online here.

" "We can confirm that Google DeepMind has reached the much-desired milestone, earning 35 out of a possible 42 points — a gold medal score. Their solutions were astonishing in many respects. IMO graders found them to be clear, precise and most of them easy to follow."

In-lecture questions

In-lecture question [Answer]

- Consider an ε-NN graph, i.e. a graph with the following weights :

$$A_{ij}^{\epsilon} = \begin{cases} e^{-\frac{\text{dist}(x_i, x_j)^2}{2\epsilon^2}} & \text{if } \text{dist}(x_i, x_j) \leq \epsilon \\ 0 & \text{otherwise} \end{cases}$$
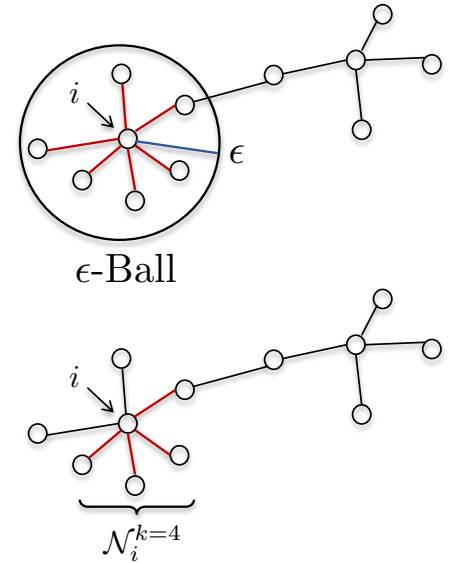
- The k-NN graph's adjacency matrix is defined as :

$$A_{ij}^{\text{kNN}} = \begin{cases} e^{-\frac{\text{dist}(x_i, x_j)^2}{\sigma^2}} & \text{if } j \in \mathcal{N}_i^k \\ 0 & \text{otherwise} \end{cases}$$

- Which graph is the most used in practice? Justify.

- Answer :

  - ε-NN graphs are not used in practice because they are usually either very sparse or fully connected.

  - k-NN graphs are sparse graphs by construction with controlled sparsity degree.
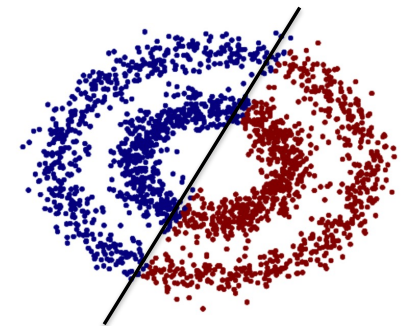
In-lecture question [Answer]

- Does k-means effectively cluster these two datasets? What assumptions does k-means make about data distribution to be successful? Justify.

- Answer : Standard k-means assume that data points are linearly separable, s.a. dataset follows a GMM, i.e. clusters are linearly separable and spherical.

- Standard k-means is not effective for non-linearly separable data. However, kernel k-means, a non-linear variant, can possibly handle non-linear datasets.
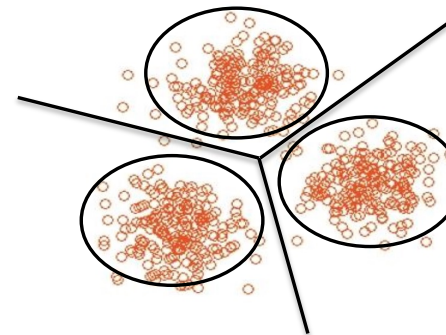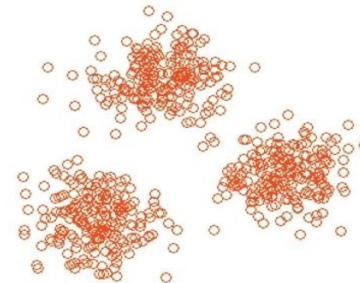
Linearly separable data points

Non-linear data points

In-lecture question [Answer]

- Is kernel k-means guaranteed to converge with all kernels? What is the computational complexity of this algorithm?

- In Slack #lectures

  - Identify the question and Reply in thread with a short response

- Answer : Like standard k-means, the kernel k-means algorithm is monotonic and guaranteed to converge as long as the kernel is positive semi-definite, $x^T A x \geq 0 \; \forall x$. And all machine learning kernels are PSD by construction.

- The kernel matrix K is full, leading to a memory complexity of $O(n^2)$ and a speed complexity of $O(n^2 dk n_i)$, where n is the number of data points, d is the data dimension, k is the number of clusters and n_i is the number of iterations to convergence.

Questions?