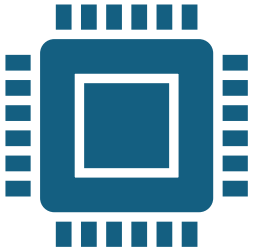
Quantum Ensemble Learning

Why ensembles for quantum ML?



Hardware bottlenecks: circuit count, depth, measurements

Ensembles trade one large model for many small/shallow learners

We compare classical vs quantum ensembling + overhead

Briefings in Bioinformatics, 2026, 27, bbag280

https://doi.org/10.1093/bib/bbag280
Published: 4 June 2026

Problem Solving Protocol

OXFORD

Quantum ensembling methods for healthcare and life science

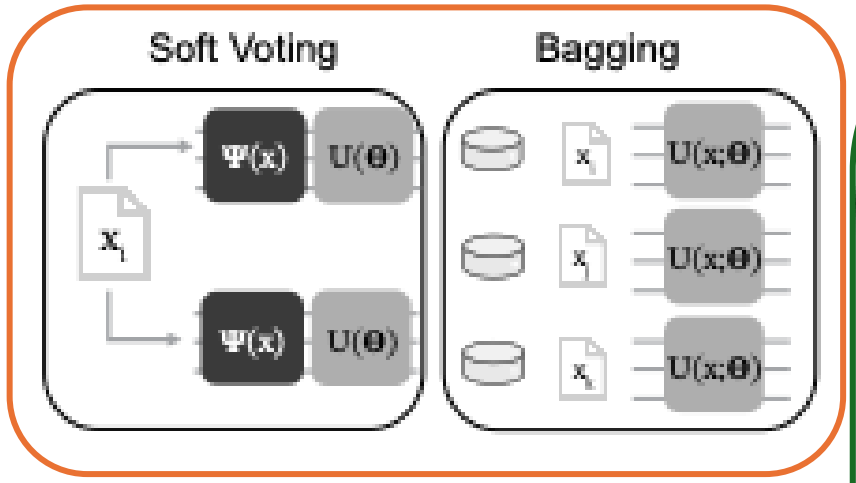
Kahn Rhrissorrakrai [1,*], Kathleen E. Hamilton [2], Prerana Bangalore Parthasarathy [3], Aldo Guzmán-Sáenz [1], Shreya Gupta [3], Tyler Alban [3], Filippo Utro [1], Laxmi Parida [1]

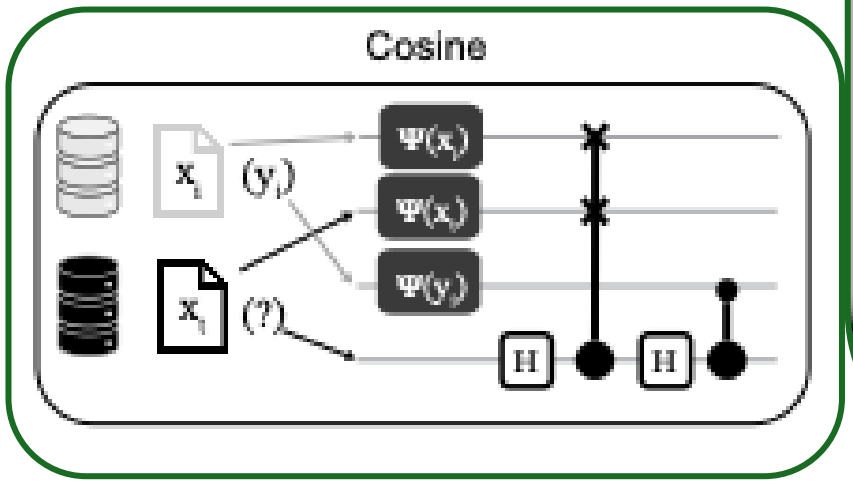
[1]IBM Research, Yorktown Heights, NY 10598, United States
[2]Computational Science and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830, United States
[3]Lerner Research Institute, Cleveland Clinic, Cleveland, OH44195, United States

Briefings in Bioinformatics

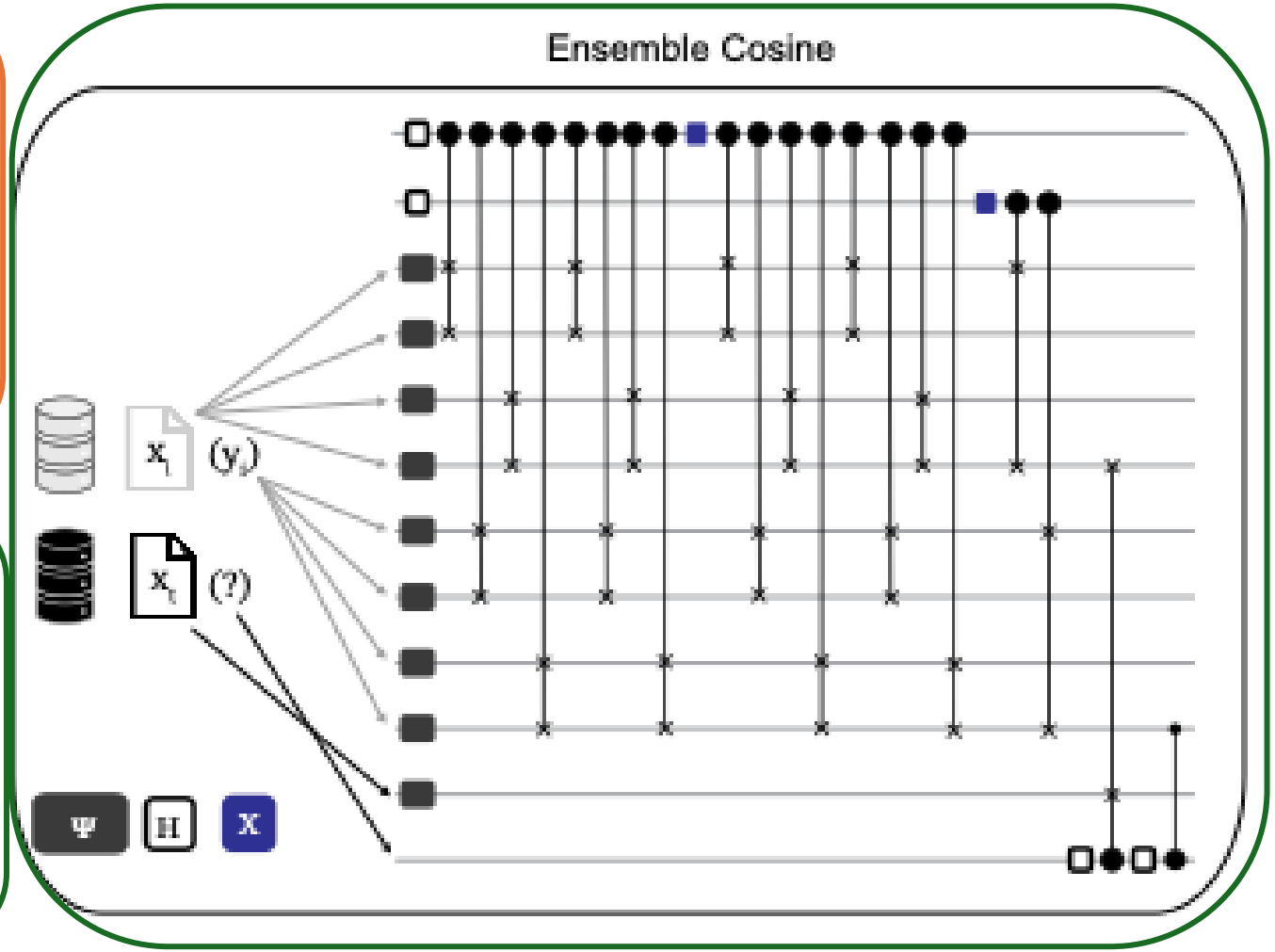
Tested ensemble classifiers

Variational classifiers





Cosine classifiers

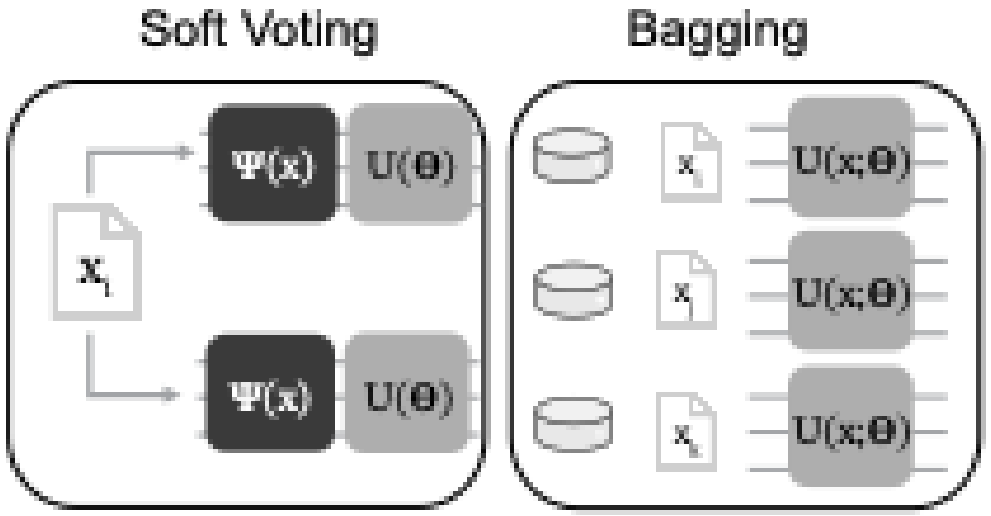


Variational classifiers

Each learner predicts the label of one encoded feature vector x_i at a time using a parameterized unitary

Soft Voting Ensemble

- An ensemble of parameterized circuits are trained using batch gradient descent (cross entropy loss)
- One sample is embedded and each wire (learner) makes an independent prediction

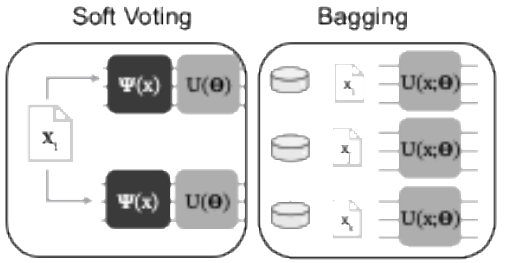


Bagging Ensemble

- An ensemble is composed from multiple learners (parameterized circuits) each are trained on a different subset of the data
- One learner is trained at a time then during inference each learner makes a prediction on an unknown sample

- With (n) qubits per learner, (L) learners per ensemble and (p) layers in each parameterized circuit the soft voting ensemble has a total of $(3(L+1)n\ell)$ trainable parameters
- With (n) qubits per learner, and (p) layers in each circuit, the bagging ensemble has a total of $(3(L+1)n\ell)$ trainable parameters but only $3(L+1)$ are trained at a time

Variational classifiers

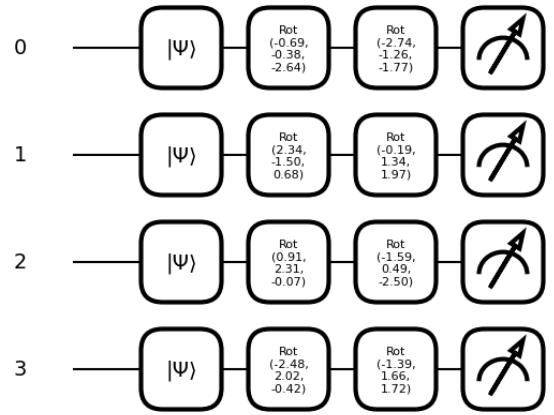


Boosted classifier

- We tested a 3rd variational classifier that used AdaBoost
- AdaBoost iteratively updates the weight of samples in the training set based on the error of the weak learner

Configurations

- Amplitude encoding of features
- A parameterized single qubit rotation decomposed as a RZ-RY-RZ gate sequence is applied to each qubit
- If learner has 2+ qubits, a CNOT is applied between qubits
- Another parameterize rotation is applied to each qubit
- Trained using mini batch with Adam

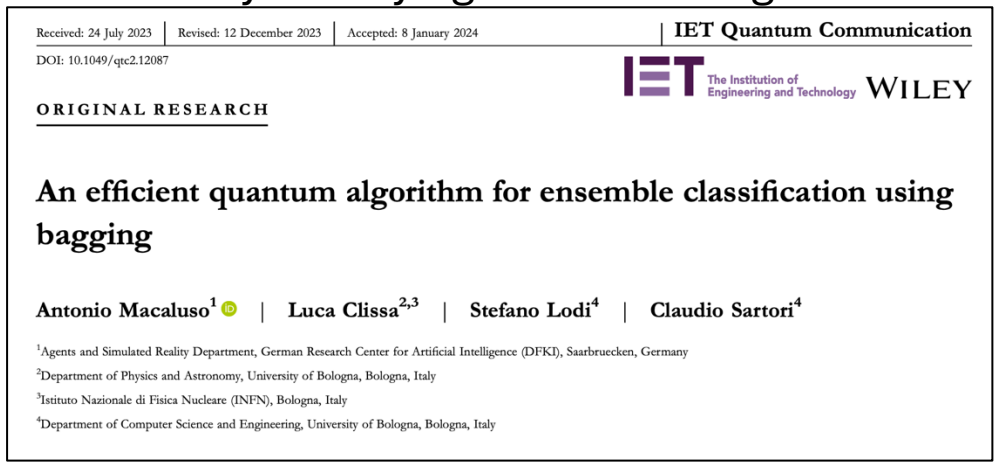


90 configurations tested over the following parameters:

- Adam learning rates $\alpha \in [1 \times 10{-}3, 1 \times 10{-}2, 1 \times 10{-}1]$
- Batch size $b \in [1, 2, 4, 8, 16]$
- Ensemble sizes $n_\ell \in [1, 2, 3, 4, 5, 6, 7]$

Cosine classifiers

- Predicts using multiple encoded feature vectors x_i, x_t and known label information y_i
- Quantum cosine classifier (QCC) is a weak learner that calculates the cosine distance between sample vectors in quantum state via interference using a single training sample
- Quantum ensemble cosine classifier (QEC) captures independent trajectories via sampling in superposition from 2^d transformations of x_i where d is ancilla control register qubits → exponential increase in ensemble size with linear increase in circuit depth
- Implemented by modifying and extending Macaluso et al (2024)



Received: 24 July 2023 | Revised: 12 December 2023 | Accepted: 8 January 2024

IET Quantum Communication

DOI: 10.1049/qtc2.12087

ORIGINAL RESEARCH

An efficient quantum algorithm for ensemble classification using bagging

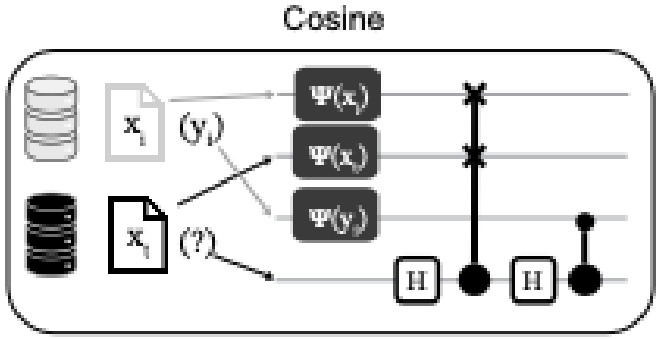
Antonio Macaluso[1] | Luca Clissa[2,3] | Stefano Lodi[4] | Claudio Sartori[4]

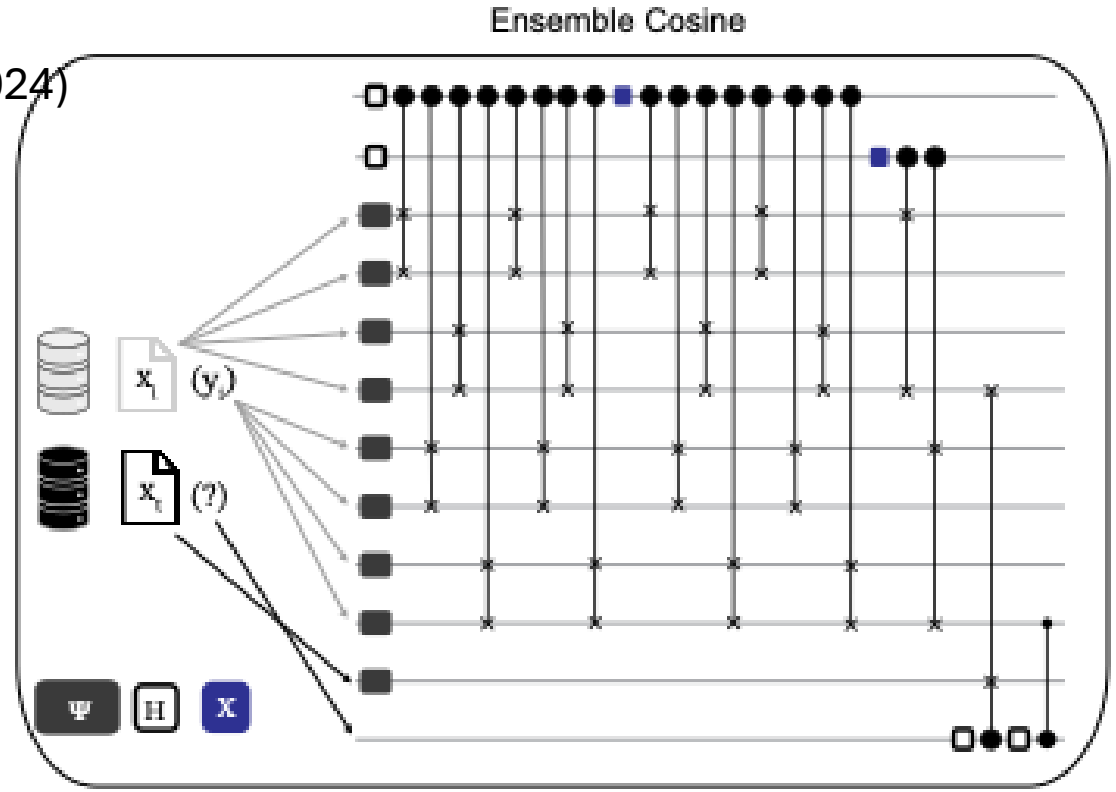
[1]Agents and Simulated Reality Department, German Research Center for Artificial Intelligence (DFKI), Saarbruecken, Germany

[2]Department of Physics and Astronomy, University of Bologna, Bologna, Italy

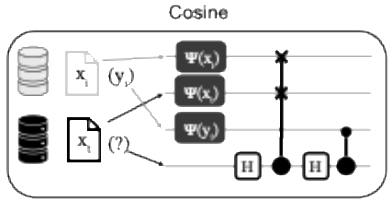
[3]Istituto Nazionale di Fisica Nucleare (INFN), Bologna, Italy

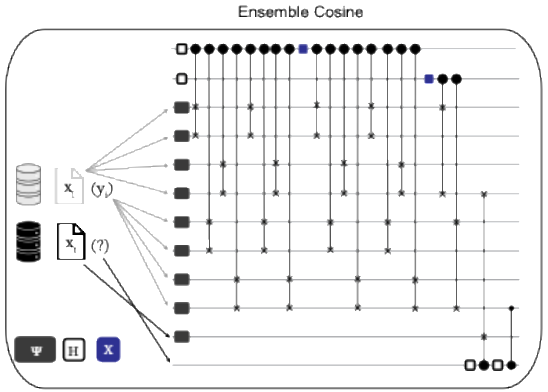
[4]Department of Computer Science and Engineering, University of Bologna, Bologna, Italy





Cosine classifiers





QEC with Random Unitaries (QECRU)

- QEC's structure yields unitaries that are generally uncorrelated
- We tested a 3rd quantum cosine classifier that considers other forms random sampling on *U(n)* to make all distribution on *U(n)* as a possible choice in the sampling, in this study used a uniform distribution on *U(n)*

Simulation configurations tested over the following parameters:

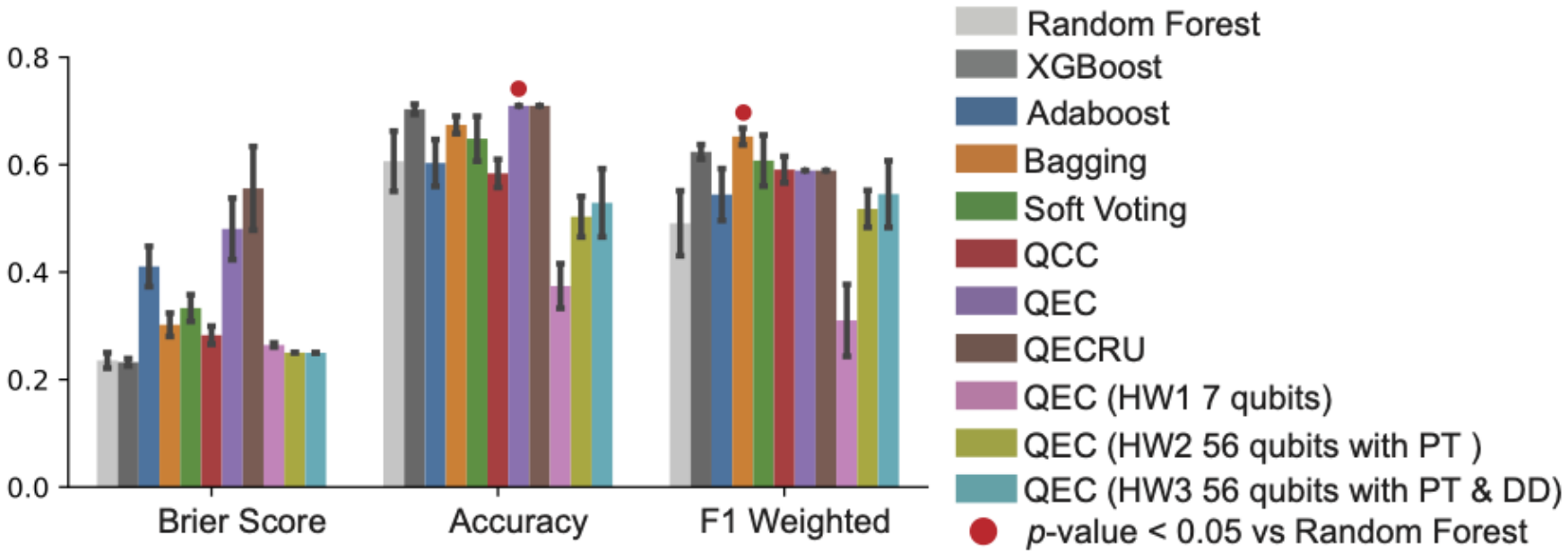
- Control registers $d \in [1,2,3]$
- Training size $n\_train \in [2,4]$
- Number of random swaps $n\_swap \in [1,2,4]$
- Number of features $n\_feature \in [2,4,8]$

Results on renal cell carcinoma + QPU experiments

Quantum ensembles comparable to strong classical baselines in simulation

Early hardware runs show trends with mitigation (PT + DD)

Key message: small ensembles + few training samples can be useful



Other findings

Variational methods could learner with fewer samples if the number of learners is increased

Bagged and boosted ensembles can be simulated quite fast as a learners are trained one at a time on disjoint subsets

Boosted ensembles can’t be easily distributed because of the need to adaptive weighting of samples

Soft voting is the slowest with the highest memory requirements as the gradient update step is expensive

- There can be opportunities to better optimize the gradient update

Share Your Feedback

Scan the QR code to let us know your thoughts on this tutorial.

ISMB 2026
WASHINGTON, DC
JULY 12–16

Everyone who completes the survey will be entered into a drawing to receive a discount code for up to $250 off an ISCB conference of their choice.

Thank you!