Multispectral Satellite Data Analysis Using Support Vector Machines With Quantum Kernels

Artur Miroszewski

European Space Agency project:
‘Quantum machine learning for analyzing multi- and hyperspectral satellite images‘

Quantum Cosmos Lab, Jagiellonian University

JAGIELLONIAN UNIVERSITY IN KRAKÓW

Quantum Cosmos Lab

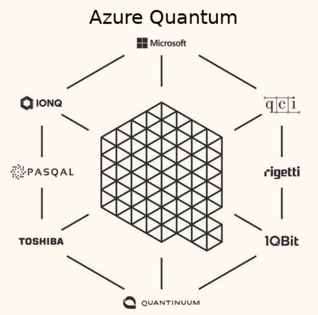
Outline

- "Introduction" to quantum computers
- Support Vector Machines
  - Linear classifiers
  - Kernel trick
- Quantum Kernel Methods
  - Hybrid SVM design
  - Data embedding
  - Quantum Kernel Estimation
  - Target Kernel Alignment
- Experimental setup
  - Satellite data
  - Results

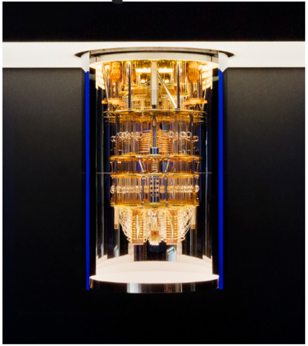
“Introduction” to quantum computers

aws

Amazon Braket



IBM Quantum



Quantum AI

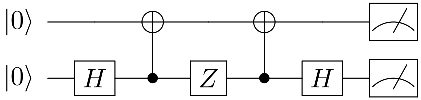
Noisy Intermediate-Scale Quantum (NISQ) devices era

Main types of QC

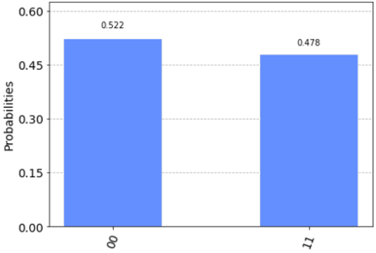
- **Gate based**
- Adiabatic

Gate based quantum computers

- **superconducting qubits**
- photonic
- ion traps
- topological
- quantum dots
- ...



- The measurement is customarily done in a Z-basis
- Probabilities are obtained by repeating the circuit runs



- Classification supervised learning algorithm
- Maximization of margins

$$\min_{w,b} \frac{1}{2}|w|^2,$$

$$such\ that:\ y^{(i)}(w \cdot w^{(i)} + b) \geq 1, i = 1, \ldots, m$$

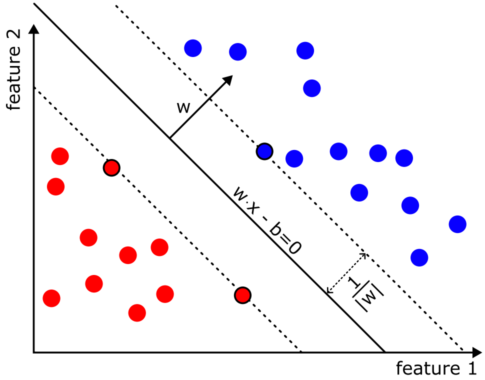
- Dual formulation

$$\max_{\alpha} \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} y^{(i)} y^{(j)} \alpha_i \alpha_j \langle x^{(i)}, x^{(j)} \rangle$$

$$such\ that:\ \alpha_i \geq 0, \sum_{i=1}^{m} \alpha_i y^{(i)} = 0$$

- Test phase

$$decision:\ \text{sign}\left(\sum_{i=1}^{m} y^{(i)} \alpha_i \langle x, x^{(i)} \rangle + b\right)$$

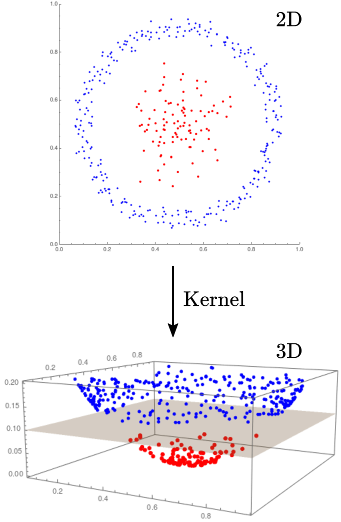


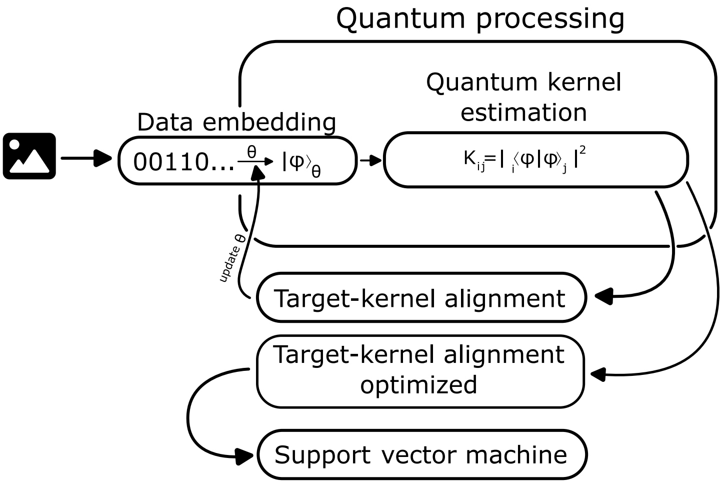
SVM kernels

- Both training and test phase depends on $\langle x^{(i)}, x^{(j)} \rangle$
- Kernel trick: exchange inner product for some 'arbitrary' similarity measure

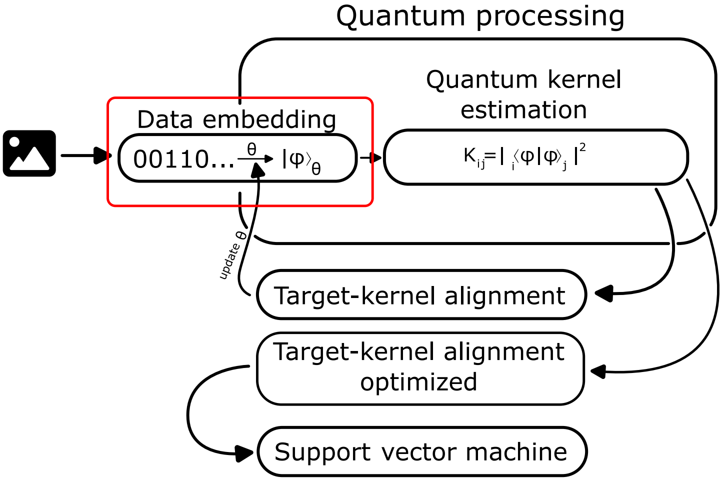
$$\langle x^{(i)}, x^{(j)} \rangle \mapsto K_{ij} = \langle \phi\left(x^{(i)}\right), \phi\left(x^{(j)}\right) \rangle$$

- The transformation ϕ:
  - leads to higher dimensional space - improved separability
  - allows for nonlinear class boundaries

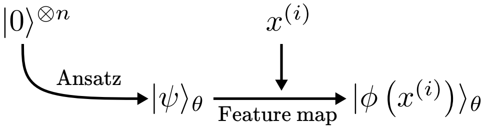



Quantum processing
Quantum kernel estimation
Data embedding
00110... θ |φ⟩θ
Kij=| i⟨φ|φ⟩j |2
update θ
Target-kernel alignment
Target-kernel alignment optimized
Support vector machine

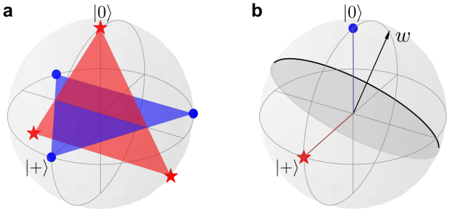
Hybrid SVM design



Data embedding

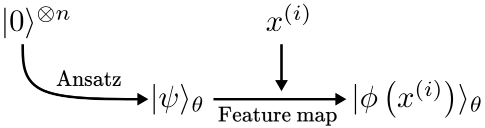


- Two stages of data embedding
  - Preparation of parameterized initial state via Ansatz
  - Application of the feature map to parameterized initial state
- The target state depends both on the parameters θ of the initial state and the data point $x^{(i)}$



Glick, Jennifer R., et al. "Covariant quantum kernels for data with group structure." arXiv:2105.03406 (2021).

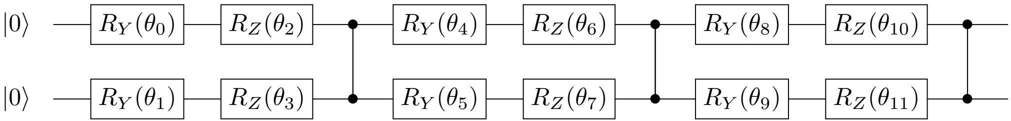
Data embedding



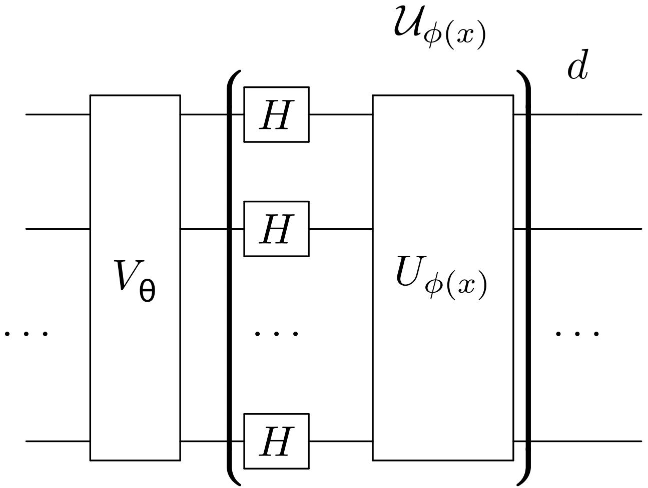
- Two stages of data embedding
  - Preparation of parameterized initial state via Ansatz
  - Application of the feature map to parameterized initial state
- The target state depends both on the parameters θ of the initial state and the data point $x^{(i)}$

<u>Ansatz</u>

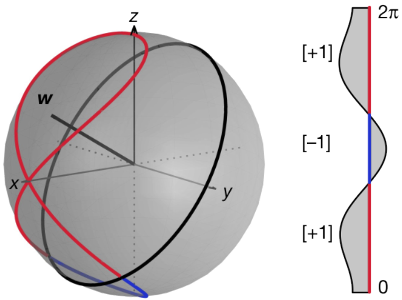
For two qubits, the ansatz consists of 3 layers of 1-qubit rotations around Y and Z axes and $ctrl - Z$ 2-qubit gates.



ZZ feature map



ZZ feature map



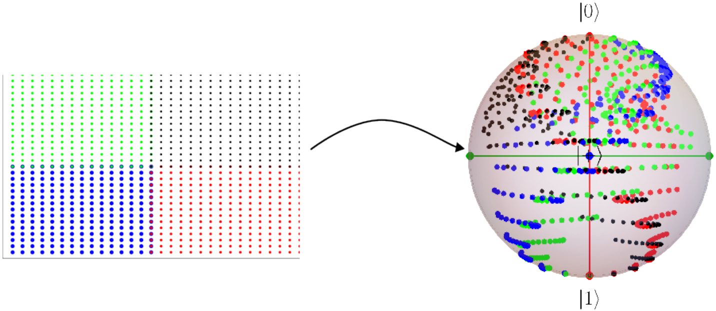
$$U_{\phi(x)} = \exp\left(i \sum_{S \subseteq [n]} \phi_S(x) \prod_{i \in S} Z_i\right)$$

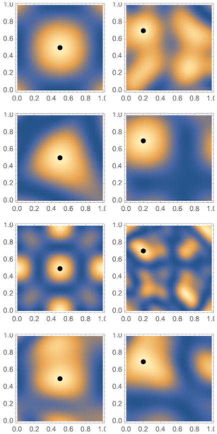
- $U_{\phi(x)}$ is consists of feature-parameterized [with $\phi_S(x)$] Z rotations
- S describes the connectivities between different qubits: $S \in \{\binom{n}{k}\ \mathit{combinations}, k \in \{1, \ldots, n\}\}$
- For a normalized data $x \in [0,1]^n$ we choose

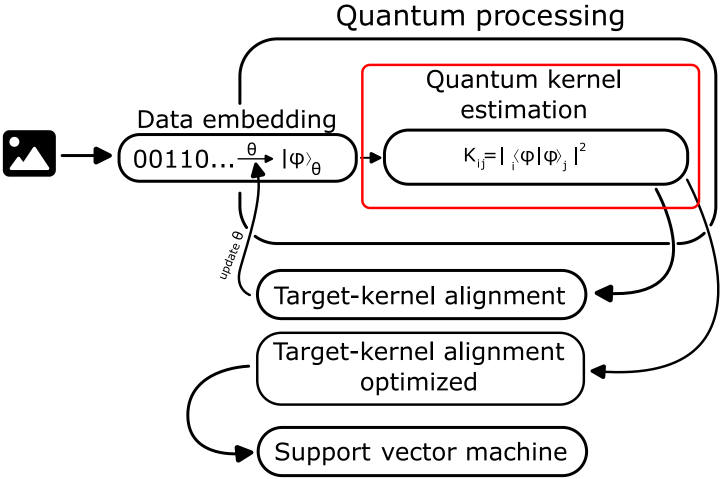
$$\phi_j(x) = \pi x_j,\ \phi_{i\ldots j} = \pi(1 - x_i)\ldots(1 - x_j)$$

Havlíček, Vojtěch, et al. "Supervised learning with quantum-enhanced feature spaces." Nature 567.7747 (2019): 209-212.

Visualizations






Quantum processing
Quantum kernel estimation
Data embedding
00110... $\xrightarrow{\theta}$ $|\varphi\rangle_\theta$
$K_{ij} = |{}_i\langle\varphi|\varphi\rangle_j|^2$
update θ
Target-kernel alignment
Target-kernel alignment optimized
Support vector machine

Quantum Kernel Estimation

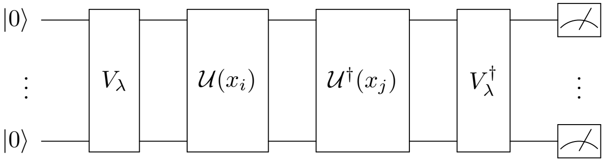
Kernel matrix elements given by an embedding state's overlap

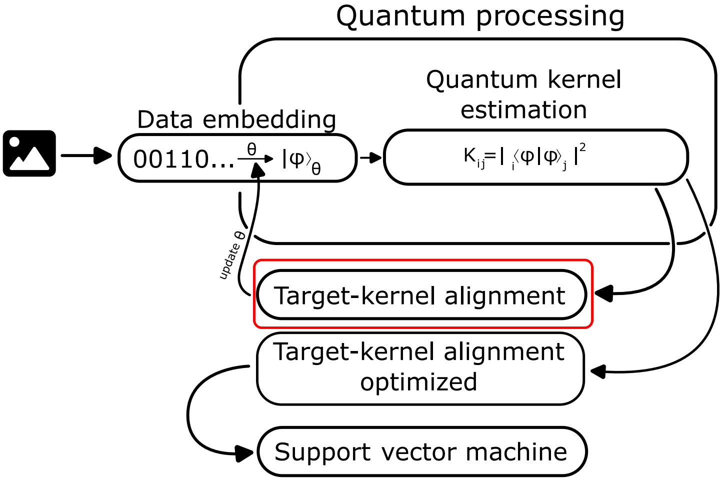
$$\mathcal{K}_{ij} = |\langle \phi(x_j) | \phi(x_i) \rangle|^2$$

Calculated by:

- Swap test
- Hadamard test
- Concatenating Hermitian conjugate

$$\mathcal{K}_{ij} = \langle 0|^{\otimes n} \; V_\lambda^\dagger \; \mathcal{U}_{\phi(x_j)}^\dagger \mathcal{U}_{\phi(x_i)} \; V_\lambda |0\rangle^{\otimes n}$$




Quantum processing
Quantum kernel estimation
Data embedding
00110... θ |φ⟩θ
Kij=| i⟨φ|φ⟩j |2
update θ
Target-kernel alignment
Target-kernel alignment optimized
Support vector machine

Target kernel alignment

Ideal kernel

$$\bar{\mathcal{K}}_{ij} = \begin{cases} +1 & \text{if } x_i \text{ and } x_j \text{ are in the same class} \\ -1 & \text{if } x_i \text{ and } x_j \text{ are in different classes.} \end{cases}$$

For supervised learning problems

$$\bar{\mathcal{K}}_{ij} = y_i y_j$$

Target kernel alignment

$$\mathcal{T}(\mathcal{K}) = \frac{\langle \mathcal{K}, \bar{\mathcal{K}} \rangle_F}{\sqrt{\langle \mathcal{K}, \mathcal{K} \rangle_F \langle \bar{\mathcal{K}}, \bar{\mathcal{K}} \rangle_F}},$$

$$\langle \mathcal{A}, \mathcal{B} \rangle_F = Tr\{\mathcal{A}^T \mathcal{B}\}$$

- Similarity measure between kernel $\mathcal{K}$ and the ideal kernel $\bar{\mathcal{K}}$
- Related to "angle" between matrix vectors

$$cos(\alpha) \text{``}=\text{''} \frac{\mathcal{K} \cdot \bar{\mathcal{K}}}{||\mathcal{K}|| \, ||\bar{\mathcal{K}}||}$$

- For general kernels

$$\mathcal{T}(\mathcal{K}) \leq 1,$$

for QKE

$$\mathcal{T}(\mathcal{K}) \leq \frac{1}{\sqrt{2}}.$$

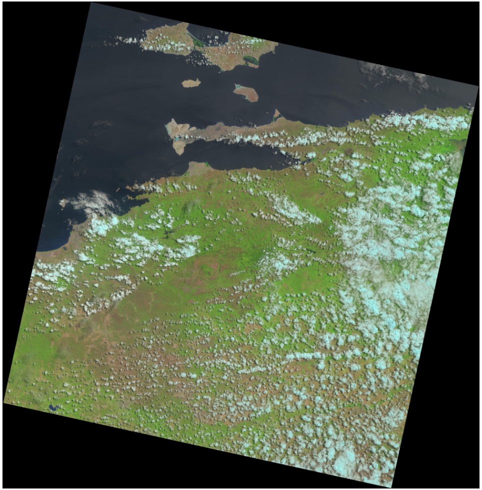
Data: 38-Cloud data set

<u>About:</u>

- 38 Landsat-8 scene images cropped into 17601 $384px \times 384px$ patches
- Four spectral bands + ground-truth cloud mask:
  - red: 630-680 nm
  - green: 520-600 nm
  - blue: 450-515 nm
  - NIR: 845-885 nm

<u>Task:</u> Cloud classification

- Data reduction through extraction of useful information.
  Cloudy regions can be removed for further processing.
- Cloud cover important for meteorological and climate research.



S. Mohajerani et al. "A Cloud Detection Algorithm for Remote Sensing Images Using Fully Convolutional Neural Networks," 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), Vancouver, BC, 2018

Experimental results

Setup:

- Comparison between classical RBF Gaussian (SVM) and quantum kernels (hSVM)
- Quantum kernels simulated with Qiskit Aer
- 20 runs with randomly selected pixels: 800 training, 200 test
- 2 features for each pixel extracted with PCA

Results:

- Two methods achieve comparable accuracy
- Wilcoxon matched-pairs signed rank test: no statistically significant difference between hSVMs and classic SVM (at $p < 0.05$).

Table 1: The results of the circuit simulations on 20 different training-test 38-Cloud splits: hSVM and SVM indicate classification accuracy for hybrid SVM and classic SVM methods. $\mathcal{T}_i$ and $\mathcal{T}_f$ show kernel target alignment before and after optimization.

	$\mathcal{T}_i$	$\mathcal{T}_f$	**hSVM**	**SVM**
Average	0.049	0.081	0.778	0.788
Standard deviation	0.018	0.024	0.038	0.029

Thank you for your attention!