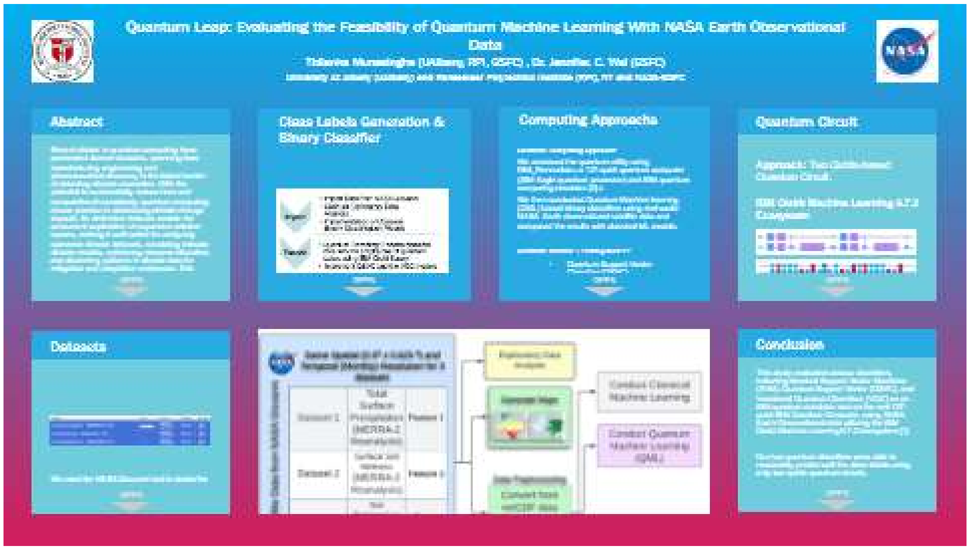
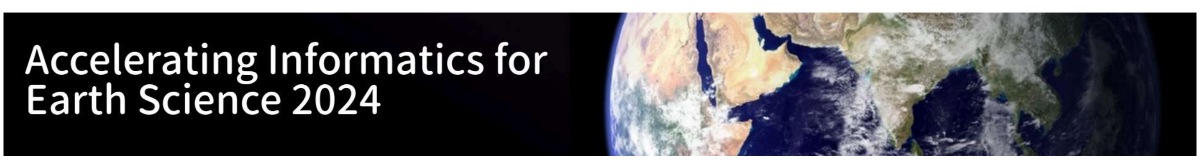
Quantum Leap: Evaluating the Feasibility of Quantum Machine Learning With NASA Earth Observational Data



Thilanka Munasinghe (UAlbany, RPI, GSFC) , Dr. Jennifer. C. Wei (GSFC)

University At Albany (UAlbany) and Rensselaer Polytechnic Institute (RPI), NY and NASA-GSFC

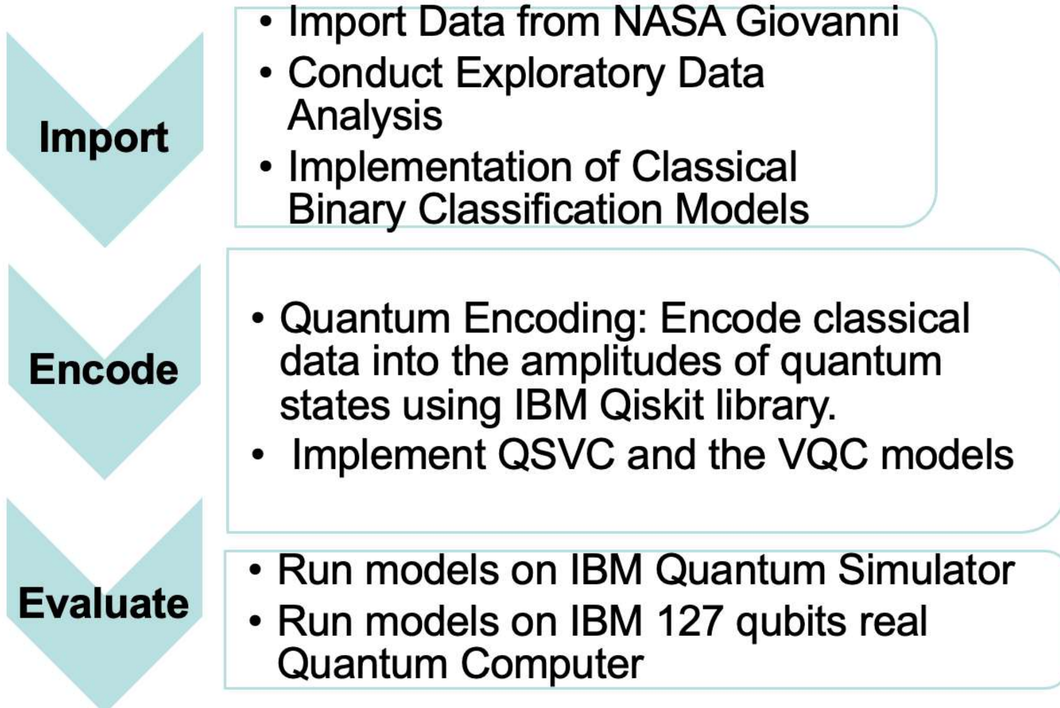
PRESENTED AT:



ABSTRACT

Recent strides in quantum computing have permeated diverse domains, spanning from manufacturing engineering and pharmaceutical discovery to the latest frontier of detecting climate anomalies. With the potential to substantially reduce time and computational complexity, quantum computing shows promise in addressing climate change impacts. Its distinctive features enable the concurrent exploration of expansive solution spaces, making it well-suited for analyzing extensive climate datasets, simulating intricate climate models, optimizing resource allocation, and discerning patterns in climate data for mitigation and adaptation endeavors. This study explores the potential of using Quantum machine learning (QML) techniques on climate and weather data obtained from NASA Giovanni*. We explored two QML algorithms, the Quantum Support Vector Classifier (QSVC) and the Variational Quantum Classifier (VQC) models, using the IBM Qiskit* Machine Learning 0.7.2 ecosystem*. The methodology and results sections describe the experiences gained from applying and evaluating quantum machine learning results on climate and weather data obtained from NASA satellites as a novel practical application of quantum computing.

CLASS LABELS GENERATION & BINARY CLASSIFIER



Downloaded data in netCDF data format and clean and preprocessed by removing the "fill-values" etc, and converted the cleaned data to NumPy arrays using the Python SciPy libraries [3,4].

Converted NumPy arrays were used for data embedding for QML applications.

<u>**Binary Classifier**</u>:
We programmatically generated the labels as “Hot(Warm)” or ”Cold(Cool)” to conduct the binary classification task.

If ”Soil Temperature” value > 295 K à labeled as “**Hot(Warm) == 1**”

If the ”Soil Temperature” < 295 K, à labeled it as “**Cool == 0**”.

295 Kelvin is 71.3 Fahrenheit (21.85 Celsius).

COMPUTING APPROACHS

Quantum Computing Approach:

We assessed the quantum utility using IBM_Rensselaer, a 127-qubit quantum computer (IBM Eagle quantum processor) and IBM quantum computing simulator [2].s

We then conducted Quantum Machine learning (QML) based binary classifiers using real-world NASA Earth observational satellite data and compared the results with classical ML models.

Quantum Machine Learning Models:

- Quantum Support Vector Classifier (QSVC)
- Variational Quantum Classifier (VQC)

Classical Machine Learning Models:

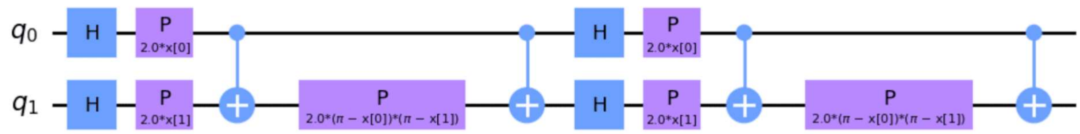
We conducted the classical machine learning using the Python-based Sckit-learn library [3] using the following models

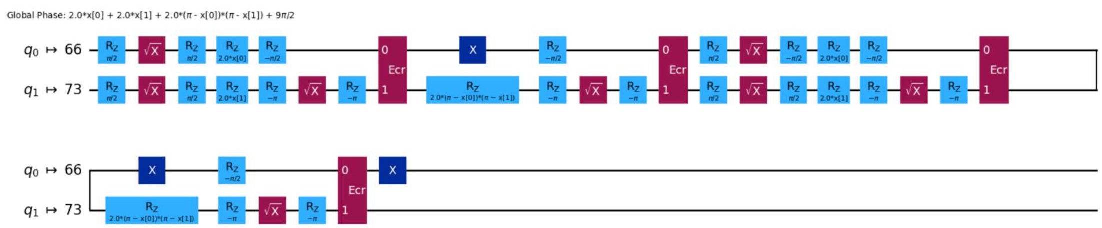
- SVM model - linear kernel
- SVM – Radial Basis Function
- SVM – Polynomial Kernel
- Gaussian Naïve Bayes model

QUANTUM CIRCUIT

Approach: Two Qubits-based Quantum Circuit.

IBM Qiskit Machine Learning 0.7.2 Ecosystem





- We Used the *ZZFeatureMap* function from qiskit.circuit.library and set the feature dimension to value two.
- The number of Independent Variables (IVs) are equal to the number of qubits (feature dimension) in our design.
- We used the *FidelityQuantumKernel* class, which utilizes the BaseStateFidelity algorithm from Qiskit.
- The *FidelityQuantumKernel*class simplifies the computation of kernel matrices for specific datasets and can be integrated with a Quantum Support Vector Classifier (QSVC).
- We Used the *ZZFeatureMap* function from qiskit.circuit.library, and set feature dimension to value two.
- We set the entanglement parameter to ”linear” and made the reps parameter to the value two. Along wih *Sampler()* function and using the FidelityQuantumKernel.

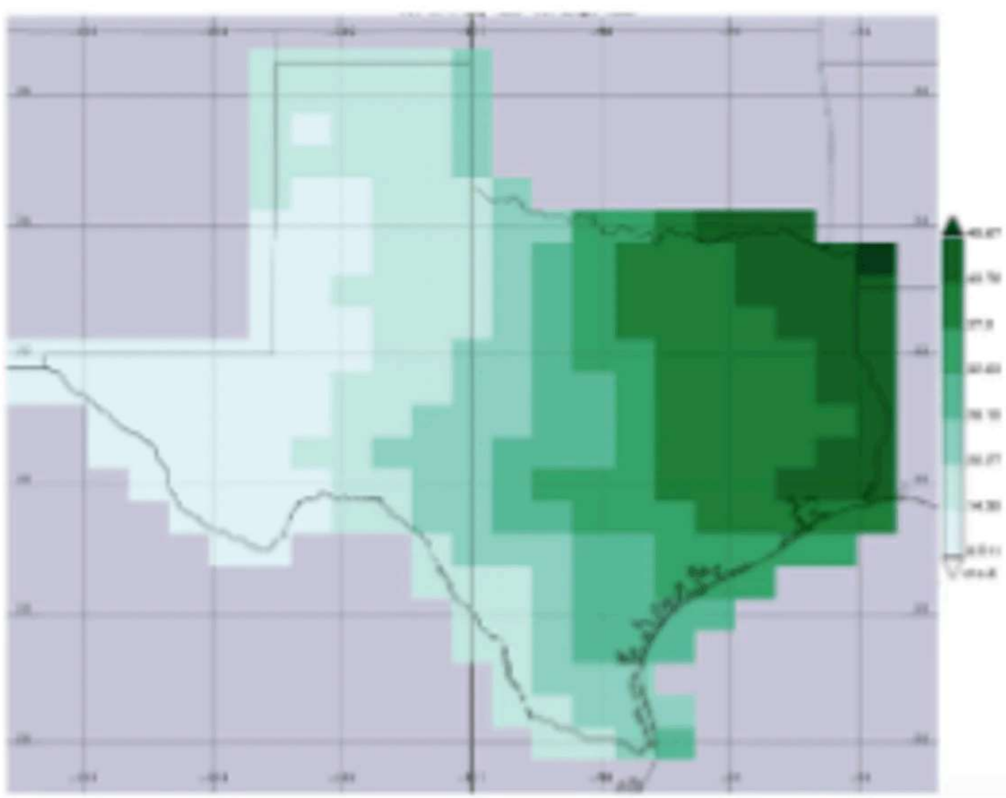
- We generated the feature map shown in the figure using the *draw()* function available in the Qiskit library.

DATASETS

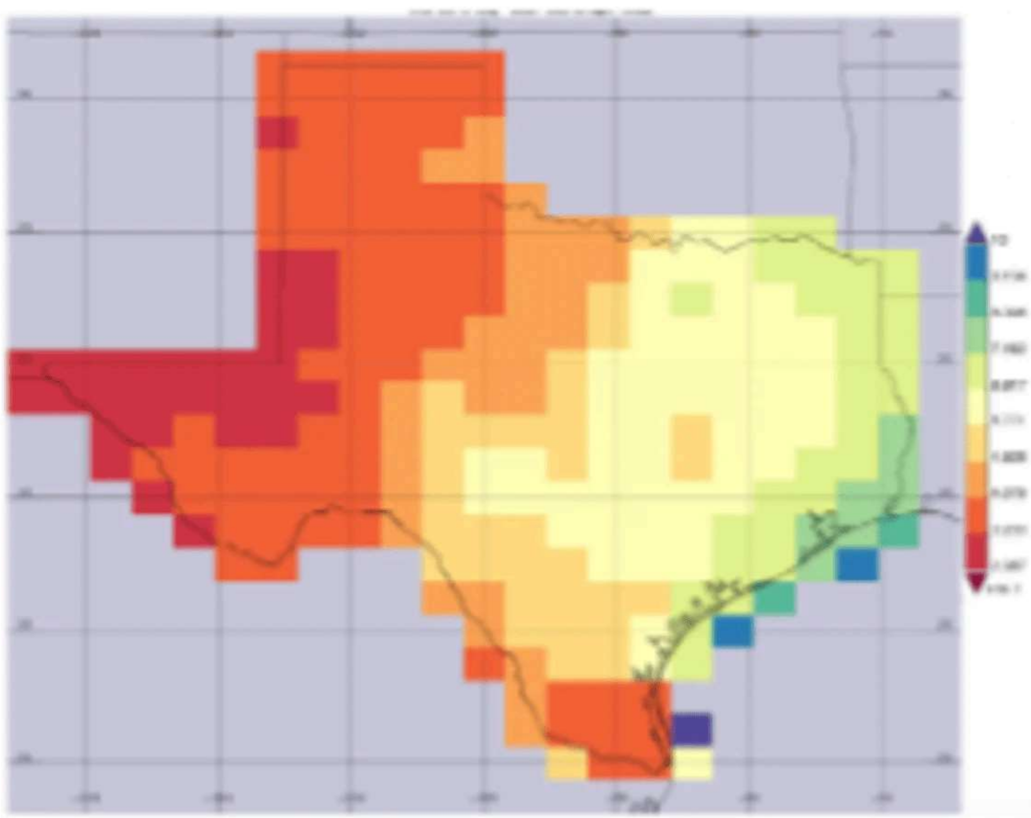
Variable	Units	Source	Temp.Re	Spat.Res
Total surface precipitation (M2TMNXFLX v5.12.4)	kg m-2 s-1	MERRA-2 Reanalysis	Monthly	0.5 x 0.625 °
Surface soil wetness (M2TMNXLND v5.12.4)	-	MERRA-2 Reanalysis	Monthly	0.5 x 0.625 °
Soil temperatures layer_1 (M2TMNXLND v5.12.4)	K	MERRA-2 Reanalysis	Monthly	0.5 x 0.625 °

We used the NASA Giovanni tool to obtain the data [6].

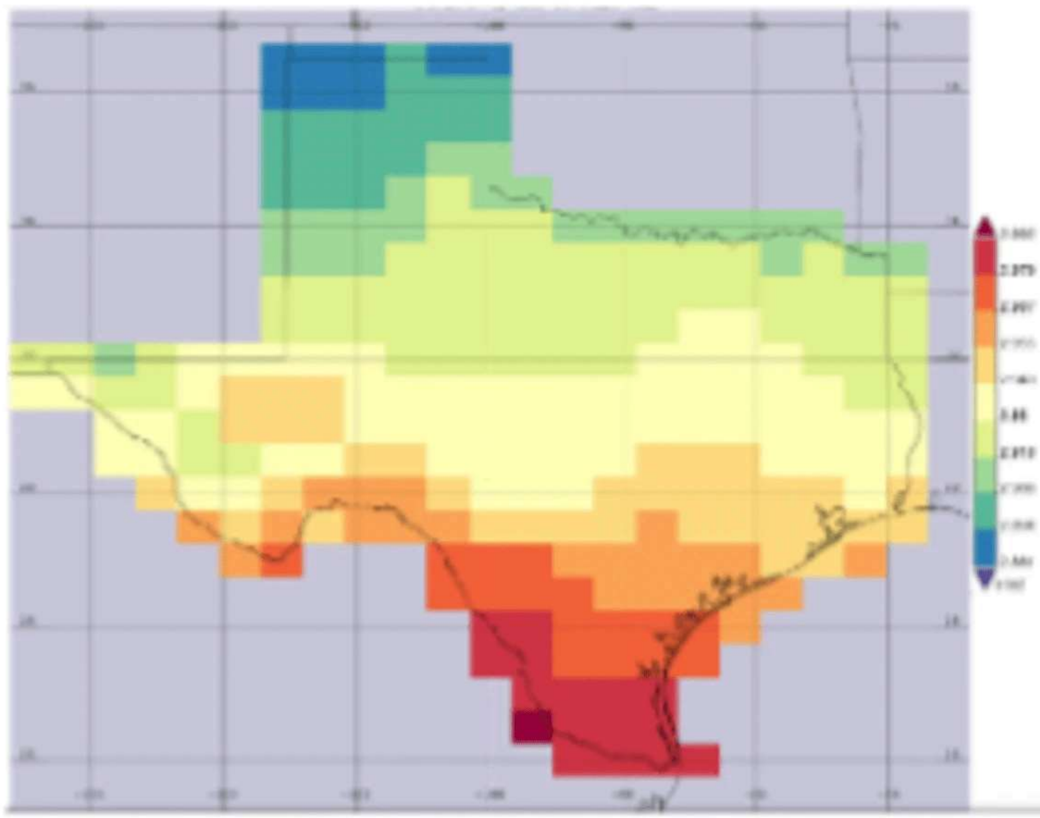
Time-averaged maps of the three variables over Texas from May 2016 to October 2021 were plotted using Giovanni's mapping tool.



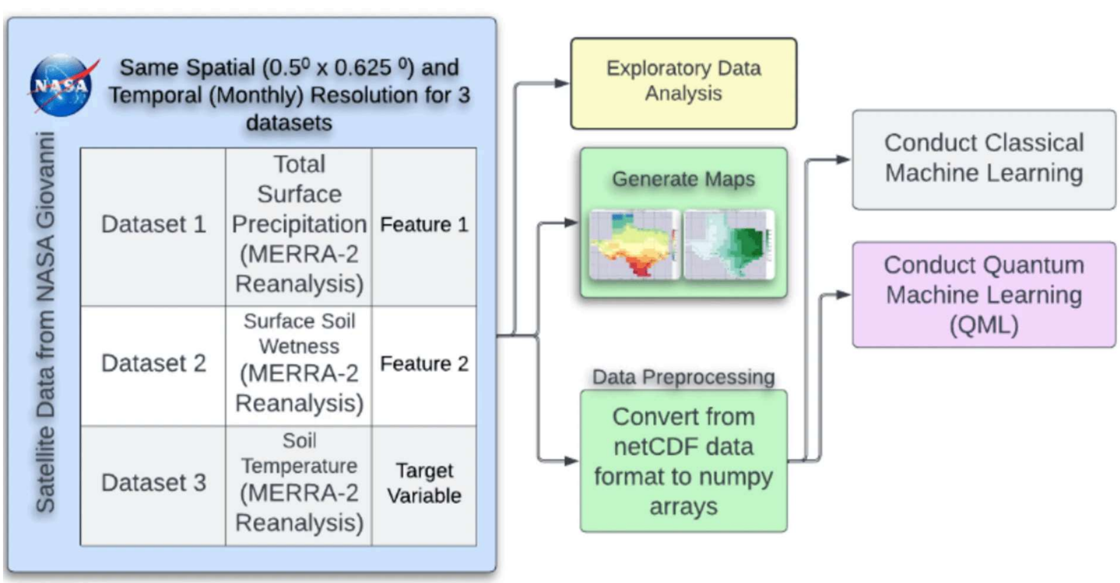
(a) Time Averaged Map of Total surface precipitation monthly 0.5 x 0.625 deg. [MERRA-2 Reanalysis M2TMNXFLX v5.12.4] kg m-2 s-1 from 2016-May to 2021-October, over Texas.



(b) Time Averaged Map of Surface soil wetness monthly 0.5 x 0.625 deg. [MERRA-2 Reanalysis M2TMNXLND v5.12.4] from 2016-May - 2021-Octobeer, over Texas.



(c) Time-averaged map of soil temperatures for Layer 1 (monthly, 0.5 x 0.625 deg) from the MERRA-2 Reanalysis (M2TMNXLND v5.12.4) in K, from May 2016 to October 2021, over Texas.



	Precision	Recall	F1-score	Support
0	0.82	0.95	0.88	39
1	0.75	0.43	0.55	14
accuracy			**0.81**	53
macro avg	0.79	0.69	0.71	53
weighted avg	0.80	0.81	0.79	53

Table 1: Performance Metrics for Classical ML with SVM model linear kernel

	Precision	Recall	F1-Score	Support
0	0.80	0.85	0.83	39
1	0.50	0.43	0.46	14
accuracy			**0.74**	53
macro avg	0.65	0.64	0.64	53
weighted avg	0.72	0.74	0.73	53

Table 2: Performance Metrics for Classical SVM – Radial Basis Function

	Precision	Recall	F1-Score	Support
0	0.80	1.00	0.89	39
1	1.00	0.29	0.44	14
accuracy			**0.81**	53
macro avg	0.90	0.64	0.67	53
weighted avg	0.85	0.81	0.77	53

Table 3: Performance Metrics for Classical SVM – Polynomial Kernel

	Precision	Recall	F1-Score	Support
0	0.74	1.00	0.85	39
1	1.00	0.00	0.00	14
accuracy			**0.74**	53
macro avg	0.87	0.50	0.42	53
weighted avg	0.81	0.74	0.62	53

Table 4: Performance Metrics for Classical Gaussian Naïve Bayes model

	Precision	Recall	F1-score	Support
0	0.75	1.00	0.86	39
1	1.00	0.07	0.13	14
accuracy			**0.75**	53
macro avg	0.88	0.54	0.50	53
weighted avg	0.82	0.75	0.67	53

Table 5: Performance Metrics for QSVC Classifier on IBM Qiskit Simulator

CONCLUSION

This study evaluated various classifiers, including classical Support Vector Machines (SVM), Quantum Support Vector (QSVC), and Variational Quantum Classifiers (VQC) on an IBM quantum simulator and on the real 127-qubit IBM Quantum Computer, using NASA Earth Observational data utilizing the IBM Qiskit Machine Learning 0.7.2 ecosystem [1].

Our two quantum classifiers were able to reasonably predict well the class labels using only two qubits quantum circuits.

The difference between the best-performing classical ML model and QML models is 0.6 (6%).

Overall, this analysis highlights the strengths and weaknesses of the two different QML models and the applicability of quantum computing technology using EO data.

Resources/Reference:

1. IBM Qiskit Machine Learning 0.7.2 Ecosystem: https://qiskit-community.github.io/qiskit-machine-learning/
2. IBM Qiskit: https://www.ibm.com/quantum/qiskit
3. Python pandas visualization: https://pandas.pydata.org/pandas-docs/stable/visualization.html#visualization-hist
4. Scikit Learn Machine Learning Package: https://scikit-learn.org/stable/
5. NASA GES DISC: https://www.earthdata.nasa.gov/eosdis/daacs/gesdisc
6. NASA Giovanni: https://earth.gsfc.nasa.gov/ocean/data/giovanni
7. GitHub Code Repository: https://github.com/thilankam/Quantum4ClimateChange

TRANSCRIPT