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

Thilanka Munasinghe[1,2,3] (tmunasinghe@albany.edu / thilanka.munasinghe@nasa.gov), Jennifer. C. Wei[2] (jennifer.c.wei@nasa.gov), Phung Lai[1], (lai@albany.edu) James Hendler[3] (hendler@cs.rpi.edu)
University At Albany[1], NASA GSFC[2], Rensselaer Polytechnic Institute (RPI)[3]

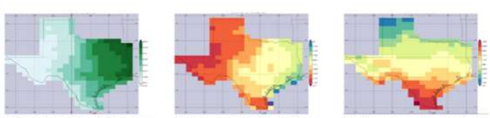
Introduction

Science Questions:

- **How effective are QML methods compared to classical approaches in predicting climate-induced phenomena like "crop-frosting" using NASA EO data?**
- **How can we leverage KG to organize information and integrate the ML results with DTs?**

This study explores the feasibility of leveraging quantum machine learning (QML) to analyze NASA Earth Observational (EO) data for climate change research, with a particular focus on the phenomenon of "*crop frosting*" which has become more prevalent due to climate change. We implemented and evaluated two QML models, the Variational Quantum Classifier (VQC) and Quantum Support Vector Classifier (QSVC), in both simulated and real quantum computing environments using a 127 qubit IBM quantum processor. Our study emphasizes the scientific rigor in comparing these quantum models with a classical Support Vector Machine (SVM) classifier, highlighting their performance in processing climate data. The results offer valuable insights into the potential scientific advantages, limitations, and scalability of QML for analyzing EO datasets, thus paving the way for more advanced climate modeling and predictive analytics using quantum computing. We showcased how Environmental Interaction Knowledge Graphs (EIKGs) and Digital Twins (DTs) can be integrated into this study. This research underscores the transformative potential of Classical and QML leveraging KGs and DT to address the multifaceted challenges posed by climate change.

Exploratory Data Analysis & Maps Generation



(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.

Fig 1. Time-averaged maps of the three variables over Texas from May 2016 to October 2021 were plotted using Giovanni's mapping

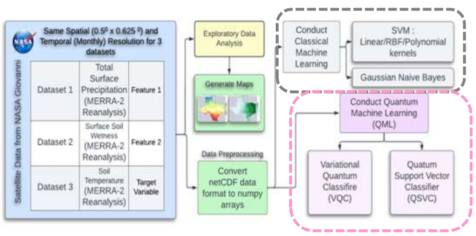


Fig 2. Classical & QML Workflow for Crop-Frosting Project

EIKG Integration with Classical & Quantum Machine Learning

Knowledge Graph Applications: Integration with Digital Twins

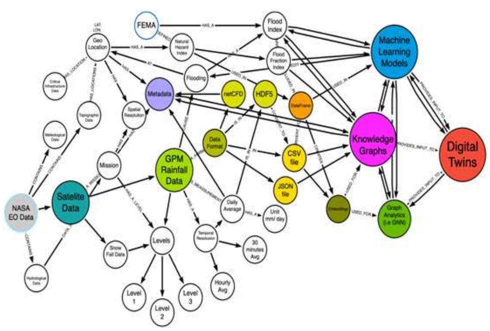


Fig 4. Schema: Integration for EIKGs, DTs and ML

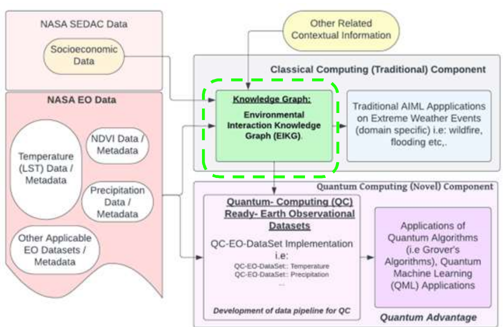


Fig 5. Big Picture: Future Work (content from NASA FINESST Proposal; FI: Thilanka Munasinghe; PI: Dr. Abdullah Canbaz; Co-I: Dr. Jennifer Wei)

Import
- Import Data from NASA Giovanni
- Conduct Exploratory Data Analysis
- Implementation of Classical Binary Classification Models

Encode
- Quantum Encoding: Encode classical data into the amplitudes of quantum states using IBM Qiskit
- Implement Quantum Support Vector Classifier (QSVC) and the Variational Quantum Classifier (VQC) models

Evaluate
- Run models on IBM Quantum Simulator
- Run models on IBM 127 qubits real Quantum Computer

Fig 6. A Common High-level Workflow for Quantum Machine Learning

Qiskit Machine Learning 0.7.2 Ecosystem:

- We used `FidelityQuantumKernel` class, which utilizes the `BaseStateFidelity` algorithm from Qiskit.
- This class simplifies the computation of kernel matrices for specific datasets and can be integrated with a Quantum Support Vector Classifier (QSVC).

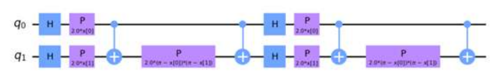


Fig 7. Visual representation of the quantum circuit generated using IBM Qiskit

Classical Machine Learning Results:

Classical Machine Learning algorithms such as Support Vector Machines (SVM: Linear, Polynomial Kernal, Radial Basis Function) and Gaussian Naïve Bayes classifier models were implemented to compare the classical ML results with QML results.

	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

Quantum Machine Learning Results:

	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: Results of QSVC on IBM Qiskit Simulator

	Precision	Recall	F1-score	Support
0	0.74	0.95	0.83	39
1	0.33	0.07	0.12	14
accuracy			**0.72**	53
macro avg	0.54	0.51	0.47	53
weighted avg	0.63	0.72	0.64	53

Table 6: Results of VQC on IBM Qiskit Simulator

	Precision	Recall	F1-score	Support
0	0.74	1.00	0.85	39
1	0.00	0.00	0.00	14
accuracy			**0.74**	53
macro avg	0.37	0.50	0.42	53
weighted avg	0.54	0.74	0.62	53

Table 7: Performance Metrics of VQC on the real IBM 127 qubits Quantum Computer (Real-Quantum-Computer)

Environmental Interaction Knowledge Graph (EIKG)

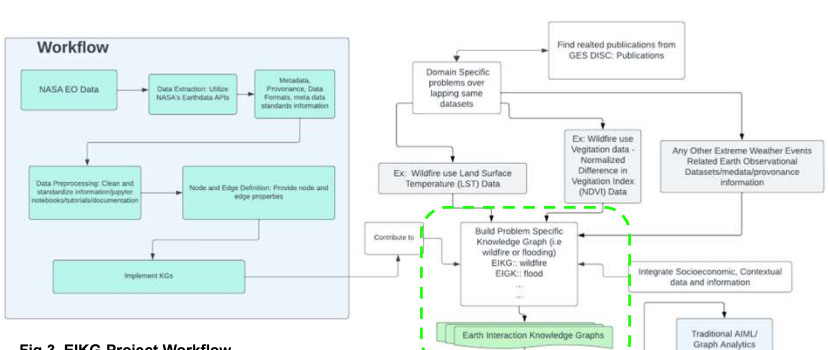


Fig 3. EIKG Project Workflow

Conclusion

This study introduced the Enhanced Integrated Knowledge Graph (EIKG) framework to model environmental factors impacting crop frosting. We assessed classical and quantum machine learning (QML) techniques using NASA Earth Observation data. Classical methods, such as Support Vector Machine (SVM) and Gaussian Naïve Bayes, were evaluated on a 16 GB RAM MacBook, while QML approaches, including Quantum Support Vector Classifier (QSVC) and Variational Quantum Classifier (VQC), were tested using the IBM quantum simulator and a 127-qubit IBM quantum computer. Our results show that the quantum classifiers performed comparably to the best classical model, with only a 6% difference in accuracy. Although modest, this performance gap suggests that QML techniques have the potential for climate-related predictions. The EIKG framework effectively integrates and organizes both heterogeneous and structured data, facilitating the fusion of classical and QML models with Earth Observation data. This integration highlights the promising role of QML in weather and climate forecasting and emphasizes the value of combining AIML insights with DTs to improve predictive accuracy and decision-making.

Glossary:

IV: Independent Variable
DV: Dependent Variable
netCDF: network Common Data Format
QML: Quantum Machine Learning
QSVC: Quantum Support Vector Classifier
VQC: Variational Quantum Classifiers

IV_1: Total Surface Precipitation
IV_2 : Surface Soil Wetness"
DV (Target variable). : Soil Temperature
We programmatically generated the labels as "Hot(Warm)" or "Cool" to conduct the binary classification task.
If "Soil Temperature" value > 295 K → labeled as "Hot(Warm) == 1"
If the "Soil Temperature" < 295 K, → labeled it as "Cool == 0".
295 Kelvin is 71.3 Fahrenheit (21.85 Celsius).

Resources:

IBM Qiskit Machine Learning 0.7.2 Ecosystem: https://qiskit-community.github.io/qiskit-machine-learning/
IBM Qiskit: https://www.ibm.com/quantum/qiskit
Scikit Learn Machine Learning Package: https://scikit-learn.org/stable/
NASA GES DISC: https://www.earthdata.nasa.gov/eosdis/daacs/gesdisc
NASA Giovanni: https://earth.gsfc.nasa.gov/ocean/data/giovanni

GitHub Code Repository: https://github.com/thilankam/Quantum4ClimateChange