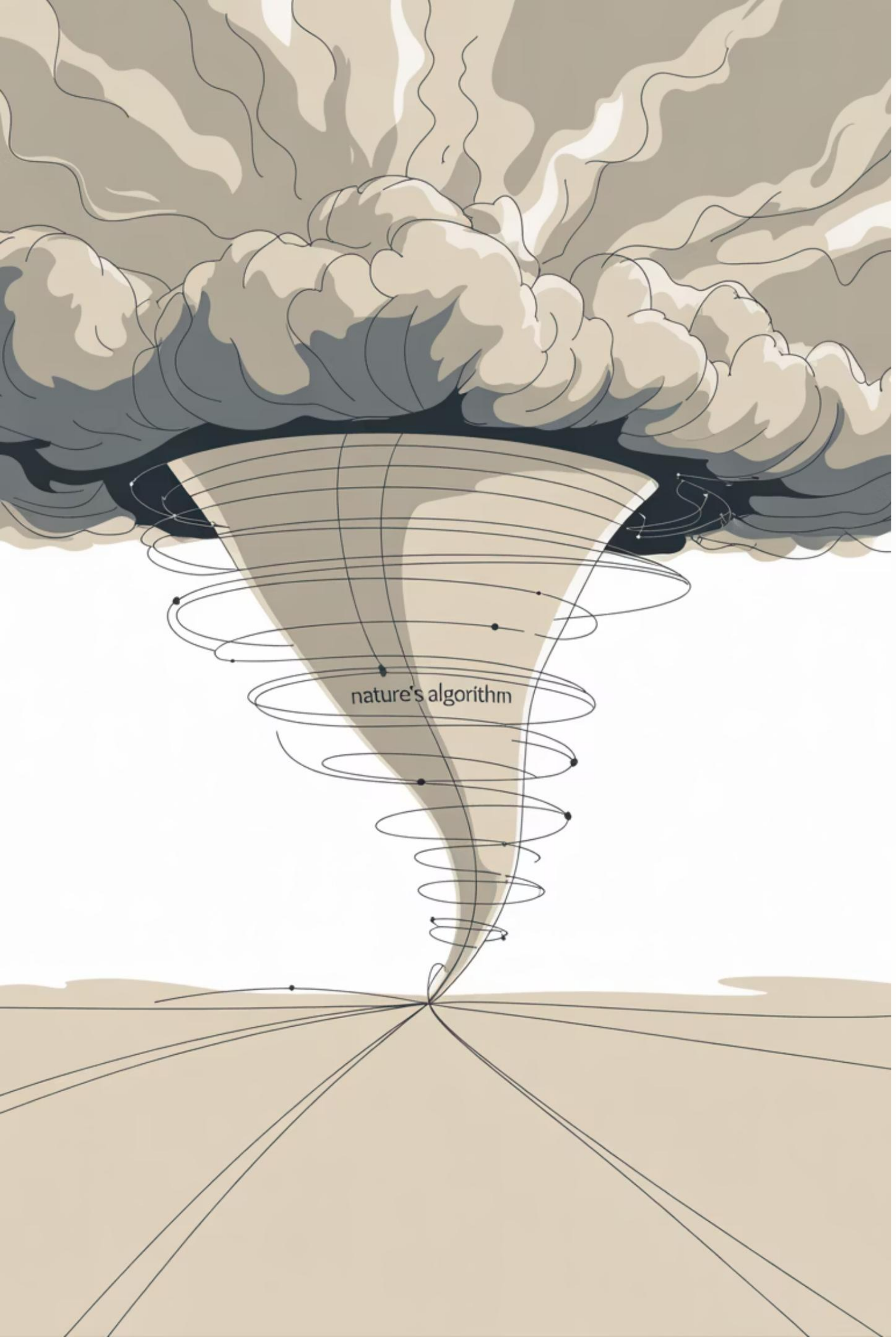


Project Twister-Q: A Quantum Hackathon Challenge

Quantum Computing for Severe Weather Prediction

Presenter: Deacon Larry M. Deschaine, PhD

Leader: Indigenous Peoples Ministry - Diocese of Charleston

Team Lead: Data Sciences - Savannah River National Laboratory

IQC25 Championship: QuantathonV2

October 9, 2025 - October 12, 2025

University of South Carolina - Darla Moore School of Business - Columbia, SC 29208

Contact: Larry.M.Deschaine@alum.mit.edu
Larry.Deschaine@srnl.doe.gov

Overview

- **Introduction & Motivation:** We'll start by framing the problem and highlighting the "quantum for good" mission.
- **Data Deep Dive:** Next, we'll guide you through exploring the dataset, nudging you toward key insights like class imbalance and how to define the target variable.
- **Methodology:** Then, we'll suggest a path forward, from building classical baselines to integrating quantum components. We'll introduce concepts like hybrid models.
- **Evaluation & The Goal:** Finally, we'll show them what success looks like by presenting the ROC and AUC chart and discussing the right metrics to use, including presentation of the scoring metrics

The Mission

Using Quantum for Humanity's Good 🌍

Our challenge is clear: Can we leverage the power of quantum computing to make a measurable improvement in predicting tornado intensity?

Even a small predictive edge could translate into earlier, more accurate warnings, giving communities more time to prepare and potentially saving lives.

This is about using quantum technology with a purpose.

Agenda

Your Hackathon Journey

01

The Meteorological Challenge: Predicting Tornado Strength

02

Your Primary Information: The Tornado Dataset

03

Building a Foundation: The Classical Approach

04

The Quantum Leap: A Hybrid Strategy

05

Measuring What Matters: Evaluation & Metrics

06

Your Target: The Goal to Beat

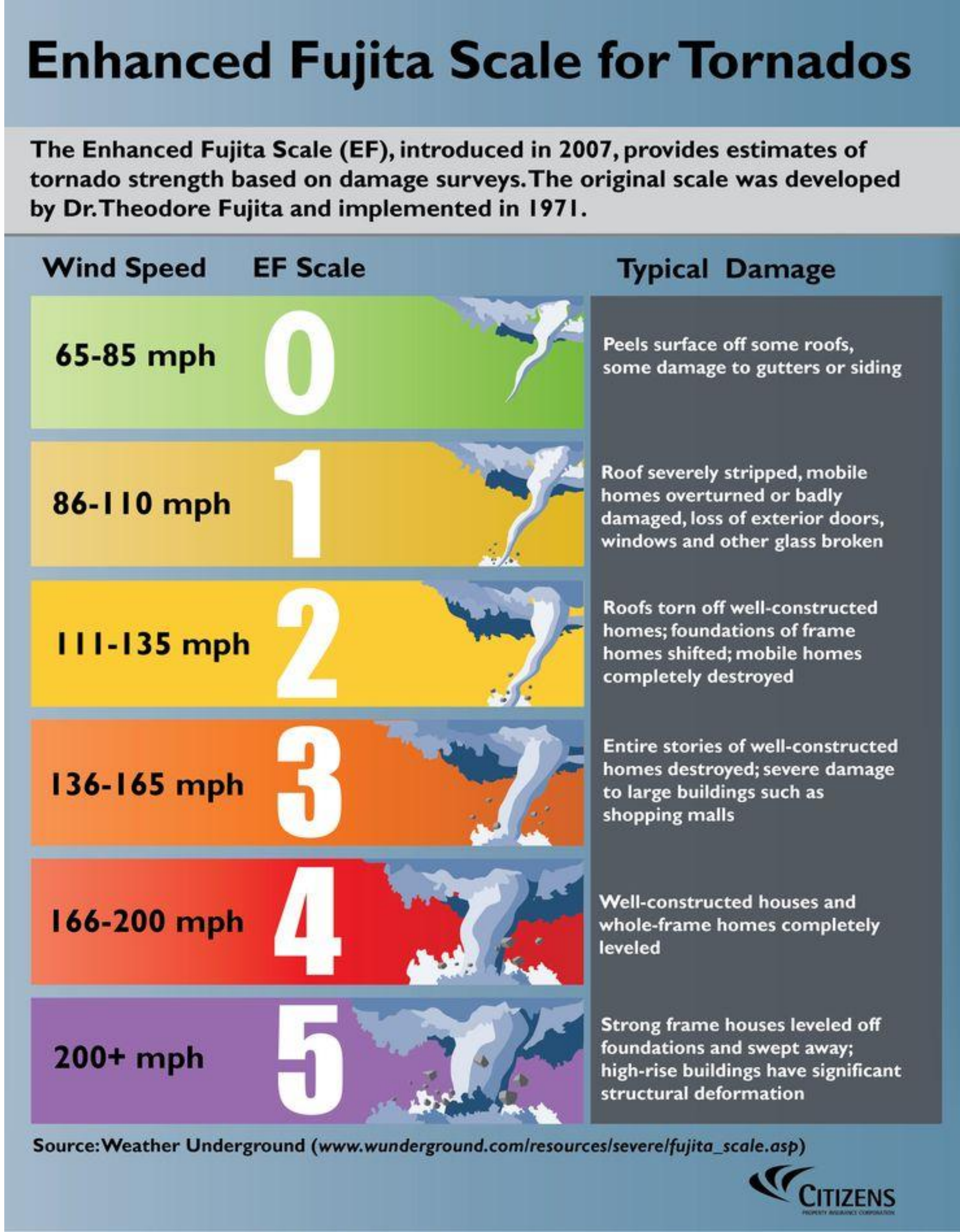
The Challenge: Predicting Tornado Strength

The Core Problem

Tornadoes are rated on the Enhanced Fujita (EF) scale from 0 (weakest) to 5 (most violent).

While we can often predict the conditions that form tornadoes, predicting their strength is incredibly difficult.

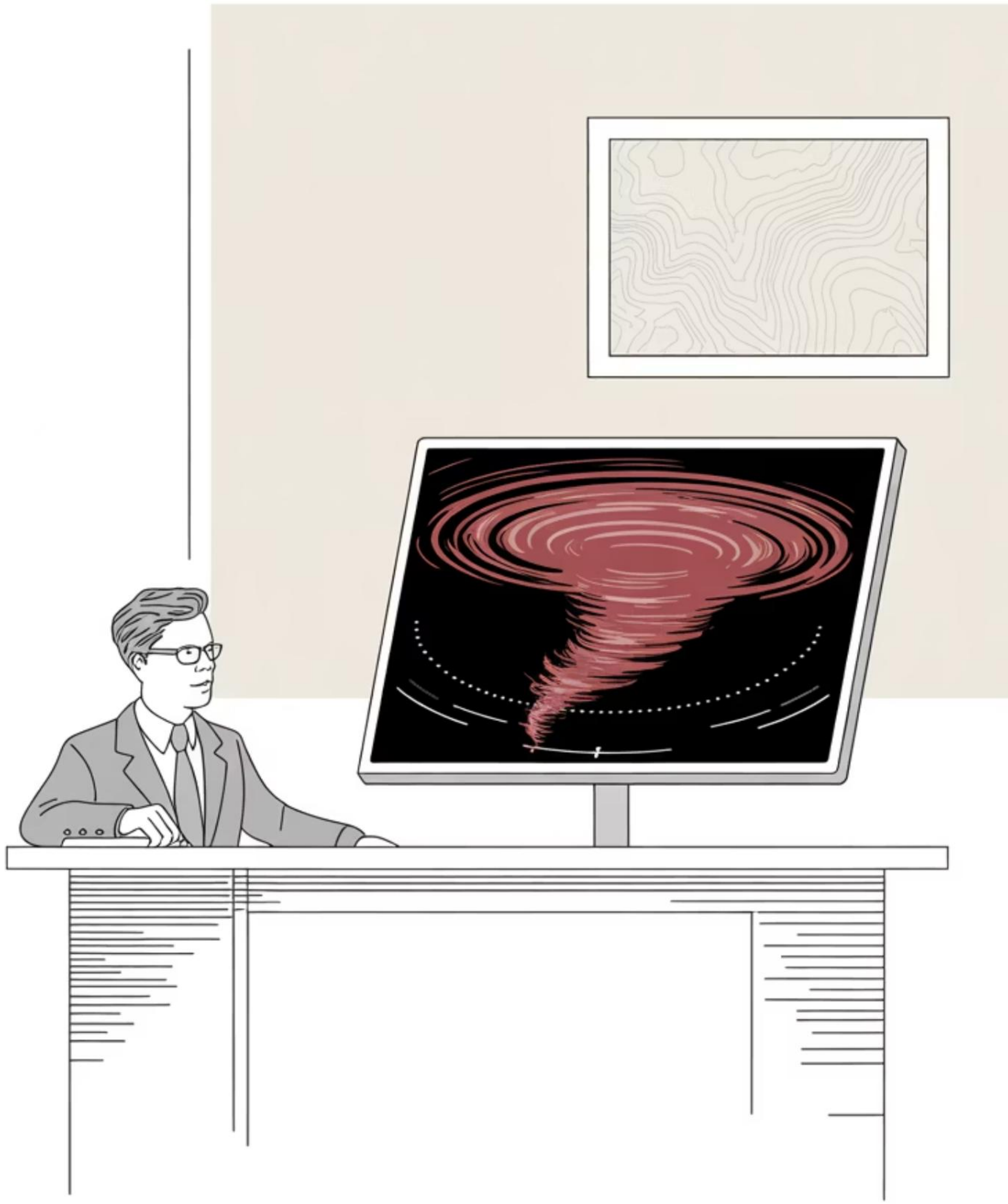
Can we find a hidden signal in the atmospheric data that points to a tornado becoming "significant" (i.e., strong and destructive)?

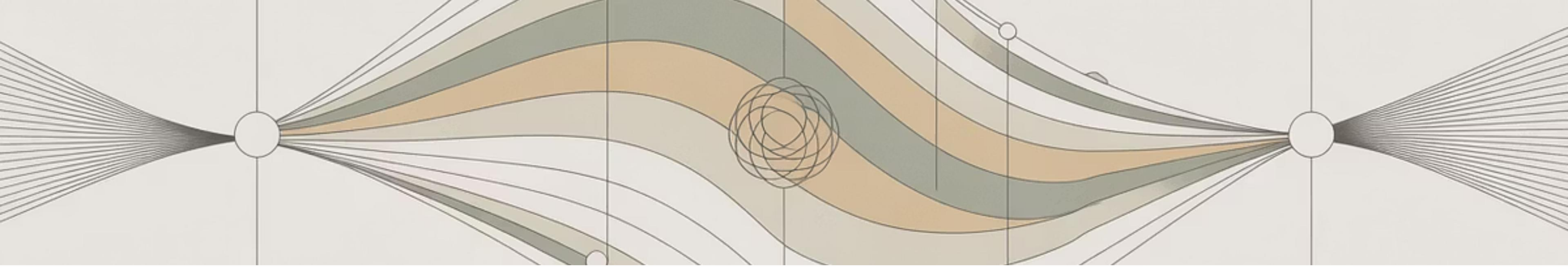


Why It Matters: The Human Impact

> "Strong tornadoes (EF2+) account for a disproportionate number of fatalities."

"The goal of a warning is not just to say, 'a tornado is coming,' but to accurately convey the threat level. This is where a breakthrough is needed."





The Core Question

Can Quantum Give Us an Edge?

Classical models have limitations in navigating complex, high-dimensional probability spaces.

Quantum computing offers a new paradigm for pattern recognition.

Hypothesis: Can a quantum or quantum-enhanced machine learning model identify subtle correlations in atmospheric data that classical models miss?

The Tornado Dataset

Exploring the Data of 1,000 Tornado Events

DATASET SUMMARY:

Total samples: 1,000
Training: 640 | Validation: 160 | Test: 200

Base problem*:

- Predict weak versus strong (ef_binary)

Advanced Problem (extra credit):

- Predict class (ef_class)

- Solve the binary problem first.
- You will be provided with all three data sets
- Provide the model results for all three data sets.

cape	cin	dewpoint	temp_2m	tcwv	surface_p	shear_0_1	shear_0_3	ef_class	ef_binary
4207.768	192.0946	294.767	299.4913	33.25276	99651.69	10.14138	22.5396	1	0
87.97427	317.0112	290.7365	304.0492	20.08835	92164.19	0.008923	6.829328	0	0
0		287.0265	297.8693	29.98212	100696.1	4.367172	9.160995	1	0
20.59766		286.7823	294.1244	18.8224	101153.3	8.598077	14.70581	0	0
0		283.4138	290.9977	20.81953	98071.88	12.41787	6.327886	1	0
729.1318	333.1519	288.4998	290.0753	27.7108	98545.55	13.56244	2.748345	1	0
0.726563		285.7498	289.0021	29.54581	99164.69	10.15079	7.062698	1	0
1036.965	248.3568	293.3568	301.4514	26.25888	97715.34	2.788389	5.070676	0	0
457.5419	84.23776	293.4752	298.8862	45.59843	98458.39	4.275392	1.975444	1	0
6.979492		286.4398	289.4167	16.87723	97343.25	6.347144	12.30718	1	0
31.95291		289.0826	291.9341	20.04114	95666	2.802124	3.251769	0	0
814.4762		282.3411	301.0153	24.32449	84731.06	0.414269	26.57933	0	0
401.0068		290.4292	295.5862	32.24719	97482.22	2.673092	4.656228	0	0
23.8082		282.0573	292.9504	15.38755	99817.97	3.538581	0.63982	2	1
43.81656		291.5903	293.4585	31.23024	97899.63	3.640186	8.863858	0	0
0		287.7668	300.5901	31.39842	96989.94	1.029864	13.17346	1	0
2578.624	212.7451	291.355	293.4812	29.90898	92827.13	1.454837	9.231501	2	1
1663.813	102.1952	293.0188	300.3083	31.23633	95520.5	5.703114	6.01422	0	0
319.6089	20.26369	291.9138	296.3585	42.76965	97048.41	1.543523	16.634	1	0
531.56	947.406	291.0023	303.1697	31.02458	95654.56	2.829055	13.84124	2	1
59.83009		285.5024	287.5751	33.57976	99548.81	10.63117	3.004904	3	1

The Tornado Dataset

cape - Convective Available Potential Energy: measure of atmospheric instability and thunderstorm potential (J/kg)
cin - Convective Inhibition: atmospheric energy barrier that suppresses convective development (J/kg)
dewpoint_2m - 2-meter dewpoint temperature: moisture content indicator at surface level (°C)
temp_2m - 2-meter air temperature: surface-level atmospheric temperature measurement (°C)
tcwv - Total Column Water Vapor: integrated atmospheric moisture content from surface to top of atmosphere (kg/m²)
surface_pressure - Surface atmospheric pressure: weight of air column at ground level (hPa or mb)
shear_0_1km - Wind shear 0-1km: difference in wind speed/direction between surface and 1km altitude (m/s)
shear_0_3km - Wind shear 0-3km: difference in wind speed/direction between surface and 3km altitude (m/s)
ef_class - Enhanced Fujita Scale class: tornado intensity rating from EF0 (weakest) to EF3+ (strongest)
ef_binary - Binary tornado classification: simplified weak (EF0-1) versus strong (EF2+) tornado categories

cape	cin	dewpoint_	temp_2m	tcwv	surface_p	shear_0_1	shear_0_3	ef_class	ef_binary
4207.768	192.0946	294.767	299.4913	33.25276	99651.69	10.14138	22.5396	1	0
87.97427	317.0112	290.7365	304.0492	20.08835	92164.19	0.008923	6.829328	0	0
0		287.0265	297.8693	29.98212	100696.1	4.367172	9.160995	1	0
20.59766		286.7823	294.1244	18.8224	101153.3	8.598077	14.70581	0	0
0		283.4138	290.9977	20.81953	98071.88	12.41787	6.327886	1	0
729.1318	333.1519	288.4998	290.0753	27.7108	98545.55	13.56244	2.748345	1	0
0.726563		285.7498	289.0021	29.54581	99164.69	10.15079	7.062698	1	0
1036.965	248.3568	293.3568	301.4514	26.25888	97715.34	2.788389	5.070676	0	0
457.5419	84.23776	293.4752	298.8862	45.59843	98458.39	4.275392	1.975444	1	0
6.979492		286.4398	289.4167	16.87723	97343.25	6.347144	12.30718	1	0
31.95291		289.0826	291.9341	20.04114	95666	2.802124	3.251769	0	0
814.4762		282.3411	301.0153	24.32449	84731.06	0.414269	26.57933	0	0
401.0068		290.4292	295.5862	32.24719	97482.22	2.673092	4.656228	0	0
23.8082		282.0573	292.9504	15.38755	99817.97	3.538581	0.63982	2	1
43.81656		291.5903	293.4585	31.23024	97899.63	3.640186	8.863858	0	0
0		287.7668	300.5901	31.39842	96989.94	1.029864	13.17346	1	0
2578.624	212.7451	291.355	293.4812	29.90898	92827.13	1.454837	9.231501	2	1
1663.813	102.1952	293.0188	300.3083	31.23633	95520.5	5.703114	6.01422	0	0
319.6089	20.26369	291.9138	296.3585	42.76965	97048.41	1.543523	16.634	1	0
531.56	947.406	291.0023	303.1697	31.02458	95654.56	2.829055	13.84124	2	1
59.83009		285.5024	287.5751	33.57976	99548.81	10.63117	3.004904	3	1

Defining the Target

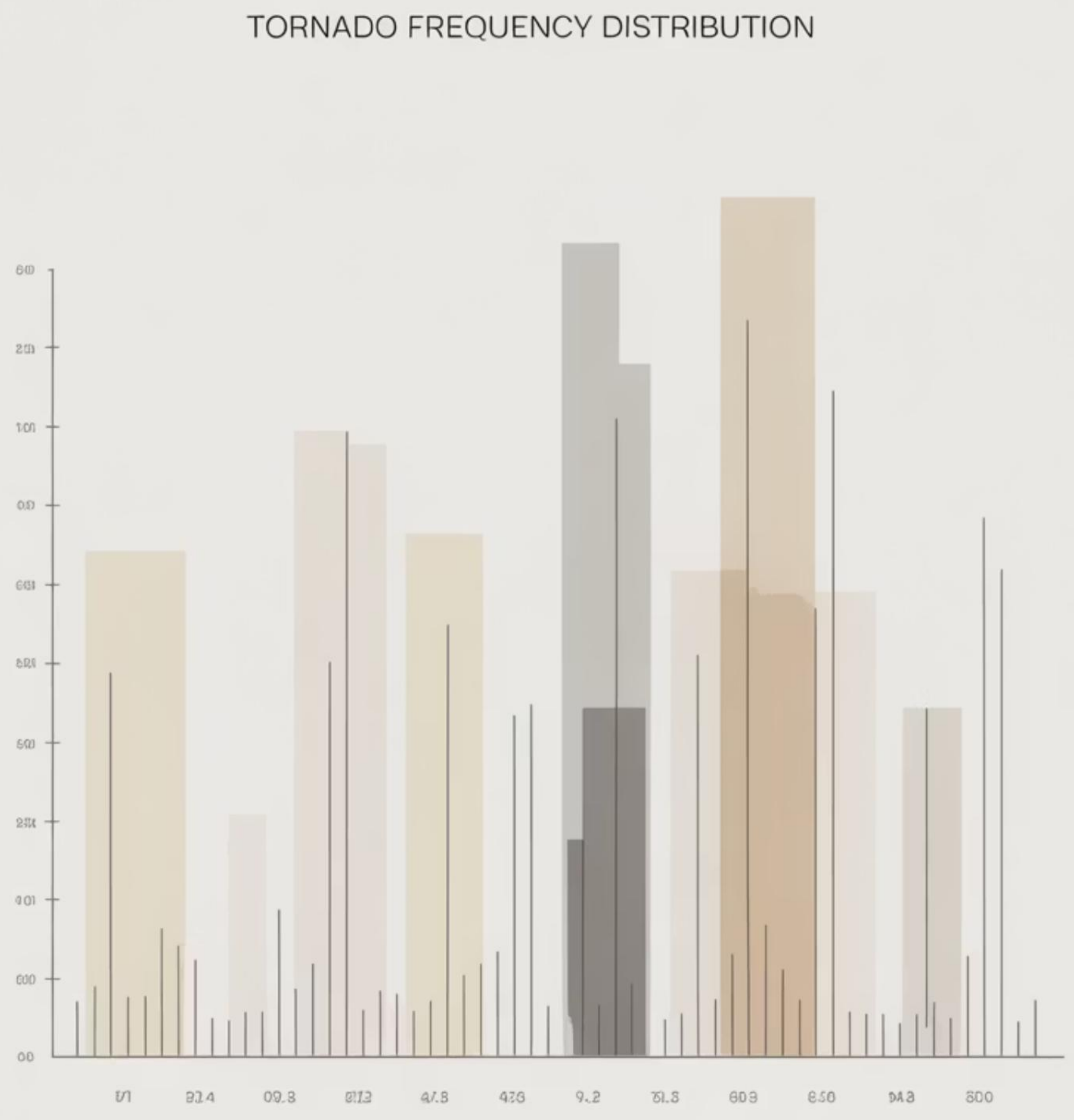
A Critical First Step

The ef_scale is categorical (0-5). You could treat this as a multi-class classification problem.

Guiding Question: However, is it easier and more impactful to start with a simpler question?

- For instance, what is the difference between a weak tornado (EF0-EF1) and a strong tornado (EF2+)?
- How might you re-frame the problem?

Hint: Think Binary Classification. It's often a more robust starting point.



A Key Challenge: Data Imbalance

The Rarity of a Storm

Most tornadoes are weak. Your dataset will reflect this reality—it is imbalanced.

Critical Question: If you train a naive model, it might achieve high accuracy by simply always predicting "weak tornado."

- How will you ensure your model learns to detect the rare but critical strong tornado cases?
- What techniques can address data imbalance?

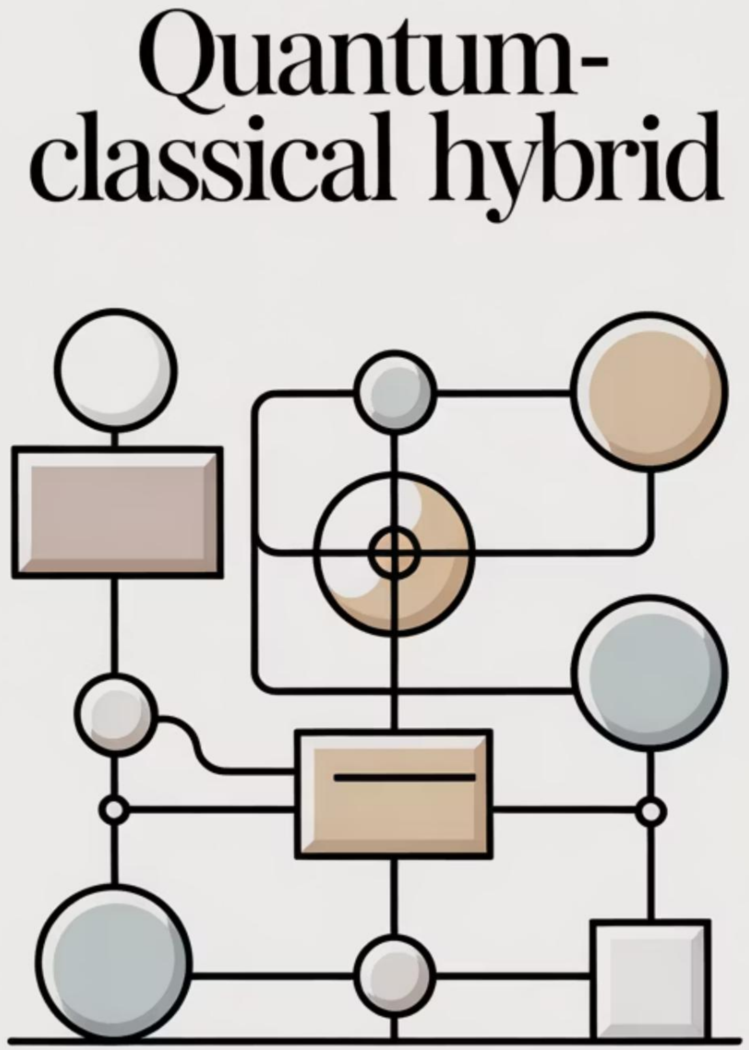
Step 1: Establish a Classical Baseline

Know Your Starting Point

Before diving into quantum, build the best possible classical model. This is your benchmark. If your quantum model can't beat it, you know you need to rethink your approach.

Suggested Models to Try t Get off to a Good Start

Logistic Regression	Random Forest	Gradient Boosting Machines (like LightGBM or XGBoost)



Thinking Quantum: The Hybrid Approach

Combining the Best of Both Worlds

Concept: A pure quantum classifier is impractical for today's Challenge Problem. A hybrid model is the way to go for this exercise.

Classical Part:

Handles data preprocessing, feature engineering, and even some parts of the classification.

Quantum Part:

A specialized approach for feature engineering and quantum circuit design that processes data to find patterns the classical computer struggles with or misses.

Encoding the Storm: From Classical to Quantum

The Data-to-Qubit Bridge

Your first quantum challenge:

- How do you represent classical feature vectors (like CAPE, shear, temp) in a quantum state?
- This is called feature encoding.

Questions to Consider

- Will you encode data into the rotation angles of your qubits?
- Will you use amplitude encoding?
- How many classical features should you select to encode onto your available qubits?
- How many and what type of quantum features will you develop?
- Which quantum framework(s) to use?
- Which algorithms from which frameworks should be used?

Feature Selection: Finding the Signal

Not All Data is Created Equal

Guiding Questions:

- How can you determine which features are the most important?
- How could you use the feature importance scores from your classical baseline model to decide which variables to encode in your quantum circuit?
- How else can you best develop the features?

Suggested Tech Stack

Some Tools for the Challenge used in the example solution:

Programming Language

Python

Classical ML

Scikit-learn, Pandas, NumPy

Quantum SDKs

We encourage you to explore cutting-edge quantum libraries. Consider frameworks like stand-alone or combined:

- Qiskit
- Pennylane
- Cirq

(Feel free to use others you are familiar with!)

Evaluation Metrics: Beyond Accuracy

Measuring Success Correctly

With an imbalanced dataset, accuracy is misleading. Focus on metrics that tell you how well you're predicting the rare "strong tornado" class.

Precision

How many of the tornadoes you predicted were strong?

Recall

Of all the strong tornadoes that occurred, how many did you correctly identify?

F1-Score

The harmonic mean of Precision and Recall.

ROC AUC Score

A great overall measure of a classifier's performance.

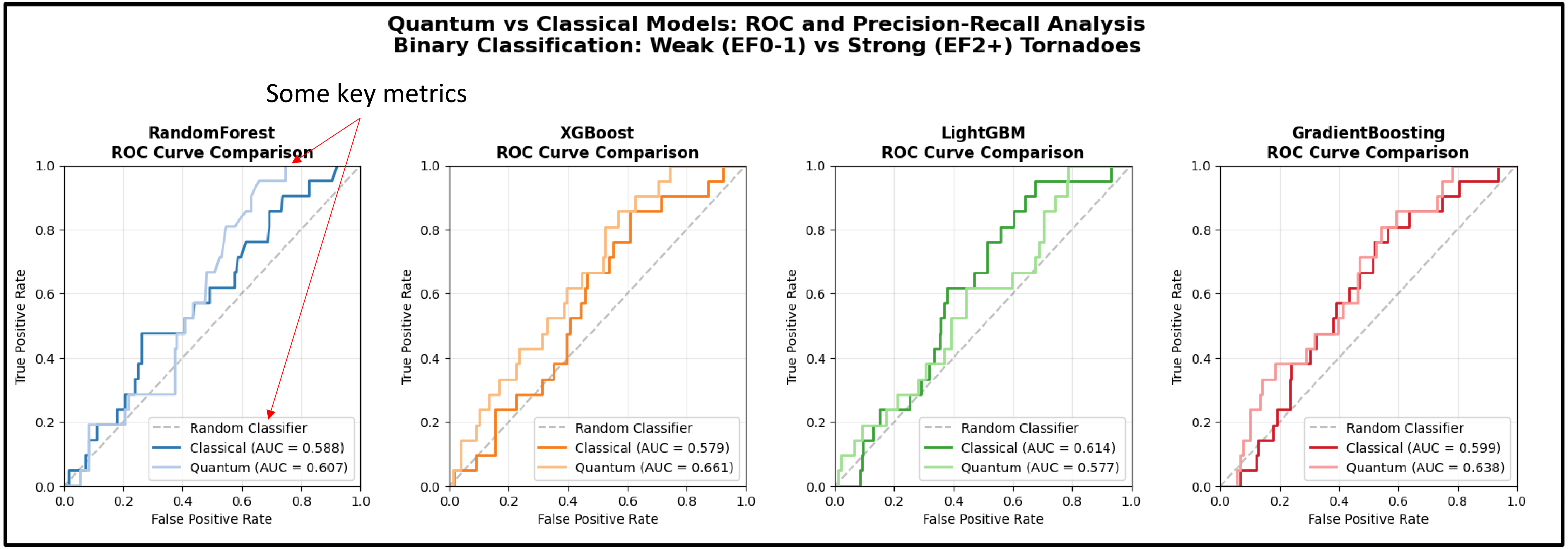
The Goal: The ROC and AUC Curves

This is Your Baseline Target 🎯

The following charts show the
- Receiver Operating Characteristic (ROC) curve and
- Area Under the Curve (AUC) for several models.

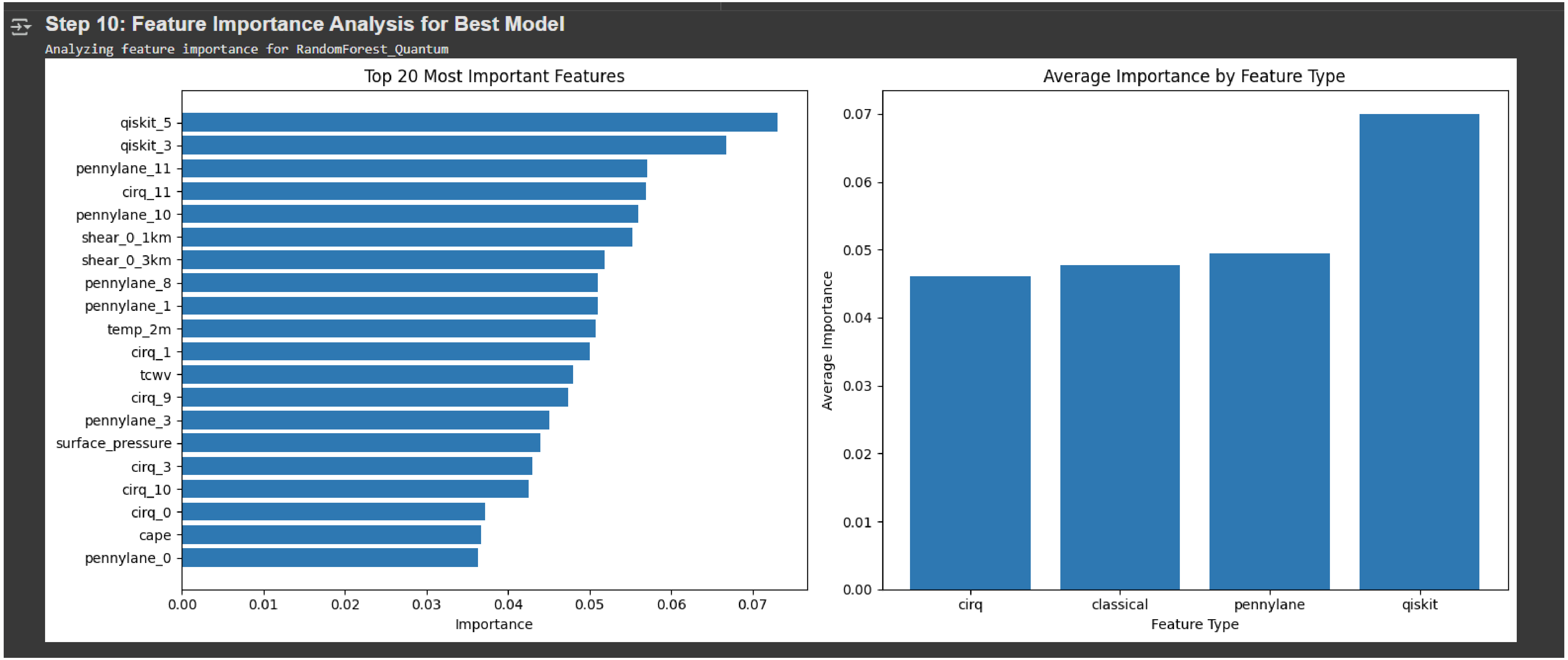
- **Your Goal:** Recreate and exceed the following results! Can you develop a quantum-hybrid model that provides a clear performance lift (pushing the curve up and to the left) over its classical-only version?

Benchmarked Evaluation Metrics: Achievable Solutions



Remember to remove one of the target variables from the calculations (ef_binary, or ef_class).
Else, you will get a 100% accurate model, but it won't have any value for emergency response teams

Benchmarked Evaluation (Examples, not Challenge Targets)
What Drives the Better Solutions?

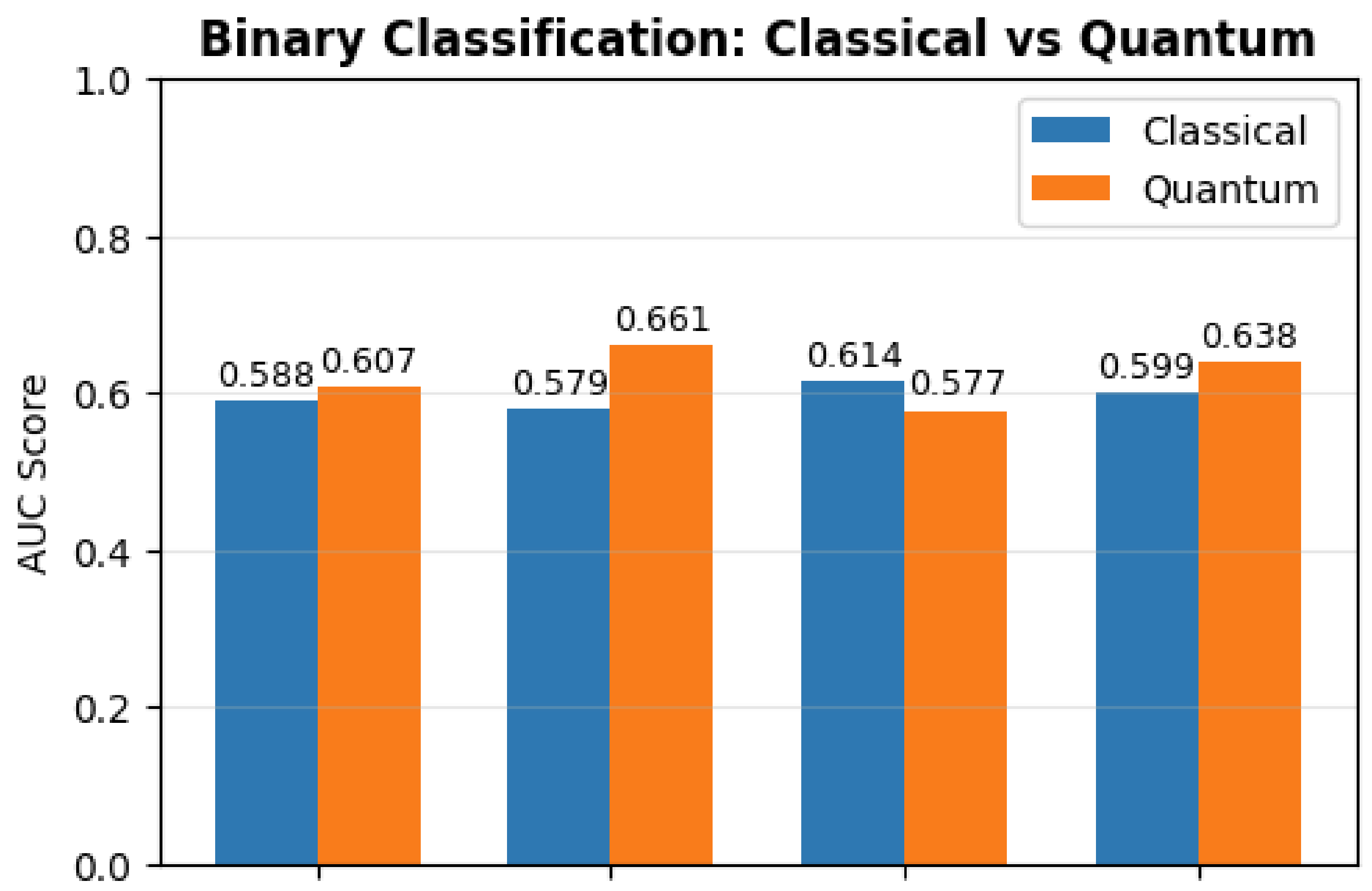


Understand Your Model's Accuracy and Mistakes

A confusion matrix is essential. It will show you exactly where your model is succeeding and failing.

Key Quadrant: Pay special attention to False Negatives:

- The strong tornadoes that your model incorrectly classified as weak.
- Minimizing these is critical for a useful warning system.



The Confusion Matrix

Understand Your Model's Mistakes

A confusion matrix is essential. It will show you exactly where your model is succeeding and failing.

Key Quadrant: Pay special attention to False Negatives:

- The strong tornadoes that your model incorrectly classified as weak.
- Minimizing these is critical for a useful warning system.
- What does your model show???

Beyond the Data: What's Missing?

Pushing the Boundaries:

This dataset is a great starting point, but real-world forecasting uses a much richer tapestry of data.
A model is only as good as the data it's trained on. What other information could paint a more complete picture of a storm's potential?

Missing Data Types

- **Real-Time Radar Data:** Where is the storm's rotation (the mesocyclone)? Are there signatures like a hook echo or a debris ball? This is the primary tool for real-time warnings.
- **3D Atmospheric Profiles:** We have surface and other data; storms are three-dimensional. What are the temperature, wind, and moisture conditions at different altitudes?
- **Satellite Imagery:** Cloud-top temperatures and storm structure seen from space provide crucial macro-level context about the storm system's organization and strength.
- **High-Resolution Topography:** Does the local landscape (hills, valleys, bodies of water) influence the storm's behavior?

The Next Frontier: Fusing New Data

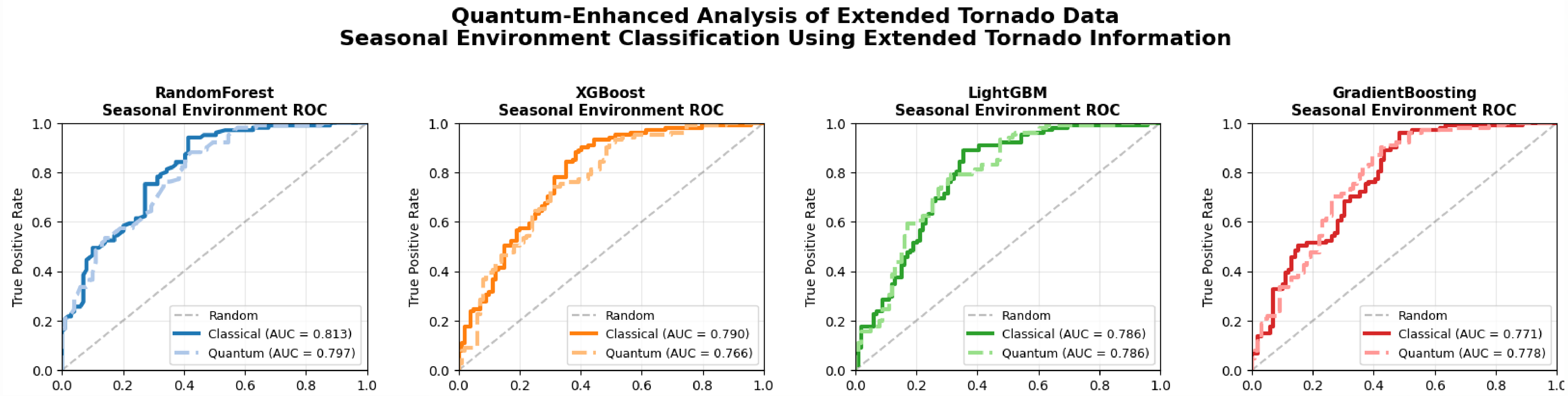
The Next Frontier: Quantum Approaches for Complex Data ⚛️

If you had access to this new data, how could you use it? Fusing complex data types is a massive challenge for classical ML and an exciting frontier for quantum computing.

Conceptual Ideas to Inspire Them

- **For Radar/Satellite Images:** Could a Quantum Convolutional Neural Network (QCNN) be trained to find subtle patterns in storm imagery that are indicative of intensification?
- **For 3D Atmospheric Data:** Could you model the atmosphere as a complex graph and use Quantum Graph Algorithms to analyze the relationships between different atmospheric layers?
- **The Ultimate Challenge:** The holy grail is Sensor Fusion—combining all these data types into one cohesive and predictive model. This is a high-dimensional problem space where quantum's unique capabilities might one day provide a significant breakthrough.

Work in Progress: Using Extended Information (For Information Only: Not part of Challenge)



Searching for a Quantum Advantage: Starting with nearly equivalent predictions….where can you take this?

Tips for Success

Hackathon Wisdom

Start Simple

Get a basic end-to-end pipeline working first, then add complexity.

Iterate Quickly

Don't spend hours perfecting one component. Build, test, learn, and repeat.

Handle the Imbalance

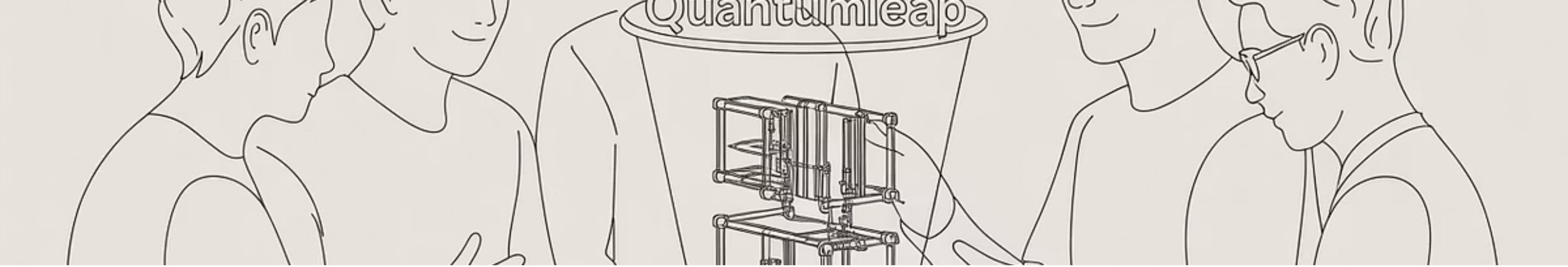
Don't forget to address the data issues (imbalance, missing, scale, etc.) early!

Collaborate

Share ideas with your team. Use any AI system(s) you'd like. Remember that different perspectives can unlock a new approach.

Manage Resources

Develop a plan and use resources effectively and efficiently.



Good Luck!

Go Build the Future

You have the data, the tools, the skills, and a meaningful challenge. We are incredibly excited to see the innovative quantum solutions you create.

Remember that the successes are not only in getting good results, but also in being able to explain and defend them, both the good and the bad.

Good luck, and let's build something that can make a real difference!