Tornado-Q: The Basic vs. The Excellent Submission

A **basic, complete submission** will successfully execute the core tasks:

- Correctly preprocess the data to create a binary classification problem.
- Train at least one classical model to serve as a baseline.
- Build a functional quantum-hybrid model that runs without errors.
- Produce a final set of metrics comparing the two approaches.
- Show clearly the training, validation, and test set results

---

An **excellent, winning submission** will do all of the above, but will differentiate itself through **depth, justification, and insight**:

- **Justification:** The team won't just *do* things; they will explain *why* they did them. Why did they choose a specific method to handle class imbalance? Why did they design their quantum circuit that way? Their code comments and final summary will tell a story.
- **Rigor:** The team will treat the problem like a scientific experiment. They will thoughtfully select features, possibly tune their models, and use a robust evaluation framework.
- **Analysis:** The team won't just show a table of scores. They will interpret the results, use confusion matrices, ROC and AUC diagrams to discuss the types of errors the model makes and what that means for real-world applications, providing relevant clear, insightful and meaningful visualizations.

Hackathon Scoring Rubric (100 Points Total)

1. Code Quality & Reproducibility (15 Points)

- **(5 pts) - Foundational:** The code runs, but may be disorganized or require manual steps to get from A to B.
- **(10 pts) - Proficient:** The code is well-organized, commented, and can be run from top to bottom to reproduce their results easily.
- **(15 pts) - Exceptional:** The code is clean, modular (uses functions effectively), and follows best practices. It's not just reproducible; it's a pleasure to read.

---

2. Data Preprocessing & Feature Strategy (20 Points)

- **(7 pts) - Foundational:** Correctly sets up the binary (weak vs. strong) classification problem, resolves any data issues, and scales the data.
- **(14 pts) - Proficient:** Effectively identifies and addresses the class imbalance using a standard technique (e.g., SMOTE, class weights, etc.).
- **(20 pts) - Exceptional:** Thoughtfully selects the most impactful features to use in the quantum model and provides a clear justification for their choice. May have experimented with different feature sets or balancing techniques.

---

3. Classical Baseline Model (15 Points)

- **(5 pts) - Foundational:** Trains a single classical model and reports basic metrics.
- **(10 pts) - Proficient:** Implements at least one strong classical baseline (e.g., Random Forest, Gradient Boosting, etc.), feature generation. Establish clear performance benchmark.

- **(15 pts) - Exceptional:** Trains multiple strong baselines and provides a clear, concise summary of which classical approach works best and why.

---

4. Quantum-Hybrid Model Implementation (30 Points)

- **(10 pts) - Foundational:** Implements a standard, functional quantum classifier circuit that integrates with the classical data.
- **(20 pts) - Proficient:** Demonstrates clear decision-making in their quantum model design (e.g., choice of feature map, circuit structure/ansatz, number of qubits), explains the choices.
- **(30 pts) - Exceptional:** Shows a deep understanding of the quantum components. They may have experimented with different circuit architectures, optimizers, or encoding strategies and feature generation to improve performance. Their hybrid model is seamlessly and efficiently integrated.

---

5. Results, Analysis & Interpretation (20 + 10 Points)

- **(7 pts) - Foundational:** Produces a final comparison of metrics showing the classical vs. quantum model performance.
- **(14 pts) - Proficient:** Meets or exceeds the ROC and AUC chart, showing a clear performance lift from the random guessing and provided quantum model. Uses confusion matrices and metrics like Precision and Recall to analyze the results for the "strong tornado" class.
- **(20 pts) - Exceptional:** Provides a compelling narrative around their results. They not only show *what* happened but offer insightful hypotheses for *why* it happened. Their conclusions are clear, well-supported, and consider the real-world implications of their findings. Able to explain why (or why not) the tool could be incorporated into an emergency response system, including what else should be considered and how it can be tested and accepted.
- **(30 pts) Gold Standard:** If all above are accomplished at the Exceptional level, extend the code to tackle the F-strength classification problem by individual F-scale.