QIntern 2025

Q.2 Upload a detailed research plan (approx. 1-2 pages), describing a 6 week plan (including background study plan for the interns), how you would guide the process and plans for future work, if any.
This will be used for evaluating the viability of the project.

Mentorship Approach

The internship will follow a clear, step-by-step structure involving weekly deliverables, foundational reading materials, and frequent mentoring sessions (at least twice per week). Interns will gain both theoretical insights and practical experience using tools such as Qiskit, PyTorch, Tensorly, Pennylane, and RDKit.

Weekly Breakdown

Week 1: Foundational Learning

- **Objective:** Familiarize interns with the basics of chemical language processing.
- Content:
    - Introduction to quantum computing principles (states, gates, measurements)
    - Overview of quantum machine learning
    - Basics of drug discovery workflows and molecular representations (SMILES, SELFIES, Morgan Fingerprinting, etc)
- **Output:** Literature summary, completed tutorials, initial molecular representation data handling scripts

Week 2 – 3 : Molecular Encoding and Data Processing

- **Objective:** Prepare molecular data for quantum model input.
- Content:
    - Parsing and cleaning molecular data
    - Feature engineering with molecular fingerprints and descriptors
    - Encoding schemes for quantum circuits after converting the SMILES to the quantum states
    - Comparison of descriptor framework and its respective compressions using PCA, t-SNE, UMAP visualizations.
- **Output:** Encoded dataset and preprocessing pipeline

Week 4 – 6 : Benchmarking

- **Objective:** Benchmark our encoding and molecular representation strategies into QML models for molecular property predictions.
- Content:

    - Designing quantum circuit embedded into the QML
    - Parameter initialization and sampling logic
    - Understanding model training objectives

- **Output:** Pre – Trained QML Models capable of molecular property prediction with high accuracy.