Design of Quantum Classifier for Hate Video Classification in Online Media

Mentor: Dr. Kumar Prateek, Dr. Simranjit Singh and Dr. Vijay Kumar

Overview: The project aims to harness the potential of quantum machine learning; specifically, quantum classification algorithms - to develop an efficient and accurate quantum classifier for identifying and categorizing hate videos on online media platforms. Hate videos, which propagate harmful and discriminatory content, present significant challenges for classical machine learning approaches due to the high dimensionality of video data, computational inefficiency, and nuanced contextual analysis. We aim to utilize quantum computing's inherent parallelism and superior optimization capabilities and seek to design a quantum-powered solution that outperforms classical methods in detection accuracy and scalability. Also, we aim to try for hybrid classical-quantum methods in detection accuracy and compare with state-of-art techniques. The research will focus on **five key objectives**: a) Conducting a comprehensive survey of QML algorithms (e.g., quantum neural networks, quantum support vector machines) and their applications in video classification, with emphasis on hate speech detection. b) Investigating bottlenecks in traditional approaches, including data imbalance, high computational costs, and interpretability challenges. c) To design and implement QML models tailored for hate video classification, optimizing quantum circuits for feature extraction and decision boundaries. d) To compare the quantum classifier's accuracy, speed, and robustness against state-of-the-art classical models (e.g., CNNs, Transformers) and quantum-classical hybrid models. e) To publish findings on QML's advantages, limitations, and future directions for moderating online hate content. The project scopes include: a) To explore hybrid quantum-classical architectures (e.g., QSVM, Quantum Boltzmann Machines) for scalable video analysis. b) To address quantum encoding techniques for high-dimensional video data (e.g., quantum feature maps). c) To use quantum simulators (IBM Qiskit, Google Cirq) and real-world hate video datasets (e.g., Hatebase) for testing.

We expect Interns to dedicate 2–3 days per week (8-hour workdays) as follows:
1) Weeks 1 and 2: Interns will perform literature review & dataset curation.
2) Weeks 3 and 4: Interns will design novel quantum circuits and perform simulation.
3) Weeks 5 and 6: Performance analysis of developed quantum model/circuit for the hate dataset and to perform benchmarking with state-of-art models available in literature.

Guidance Approach: As the research lead, we will: a) **Curate reading materials** for interns on both purely quantum and hybrid classical-quantum model bibliographies of essential papers in addition to performing following activities. b) **Lead writing workshops** (weekly) focusing on: i) Scientific writing best practices and Implementation documentation ii) **Hold weekly standup meetings** (15-20 minutes with each intern) to: Address technical challenges and ensure alignment with research goals; Facilitate knowledge sharing between team members c) **Provide structured feedback** on all written materials within 24 hours.