**Project Report:**

**Advancing Financial Crisis Prediction through Quantum Machine Learning**

### Abstract

As financial markets become increasingly complex, accurately predicting financial crises remains a significant challenge for economists and analysts. This project explores the application of **Quantum Machine Learning (QML)** in predicting financial crises, a novel approach that leverages the unique computational advantages of quantum computing. Utilizing both **correlation-based** and **Principal Component Analysis (PCA)-based** feature reduction techniques, this study aims to evaluate the performance of QML models against traditional machine learning algorithms, specifically in their ability to analyze high-dimensional financial data. Through training and testing Quantum Support Vector Classifiers (QSVC) and Variational Quantum Classifiers (VQC) on various datasets, this research assesses the precision, recall, accuracy, and F1-score of QML models. Our findings reveal potential advantages in the predictive power of QML models, particularly in processing reduced-dimension datasets, highlighting the promise of quantum computing in enhancing predictive accuracy in financial crisis forecasting.

### Introduction

Financial crises present complex, systemic risks that can lead to devastating economic consequences, as demonstrated by recent crises, including the 2008 global financial collapse. Forecasting these crises with accuracy remains one of the most challenging areas in economic analysis due to the intricate interplay of numerous economic indicators and market dynamics. Traditional machine learning models, while effective to some extent, often face limitations in capturing the full breadth of these interdependencies, especially in high-dimensional datasets where computational efficiency becomes a bottleneck.

**Quantum Machine Learning (QML)**, an emerging field at the intersection of quantum computing and machine learning, offers a potentially transformative approach to addressing these challenges. Quantum models can exploit the principles of quantum mechanics, such as superposition and entanglement, to process vast and complex datasets more efficiently than classical models. This project investigates the application of QML in financial crisis prediction by employing both **correlation-based** and **Principal Component Analysis (PCA)-based** feature reduction techniques. By reducing the dimensionality of the dataset, we aim to make QML training feasible while focusing on the most predictive features.

The primary objective of this study is to evaluate the effectiveness of QML models compared to classical machine learning algorithms, specifically Support Vector Machines (SVMs), in the context of financial crisis prediction. To this end, we configure **Quantum Kernels** and apply **Quantum Support Vector Classifiers (QSVC)** and **Variational Quantum Classifiers (VQC)** on both correlation-based and PCA-based reduced datasets. Performance metrics such as accuracy, precision, recall, and F1-score are used to evaluate model effectiveness.

This research offers insights into the potential advantages of QML in financial crisis prediction, with implications for advancing machine learning applications in economic forecasting and risk management. By exploring the computational and predictive advantages of quantum models in a high-stakes field, this project aims to lay the groundwork for future research into QML applications within financial domains. The findings presented here are expected to contribute not only to the understanding of QML's current capabilities but also to the broader discussion on the integration of quantum computing into traditional economic analysis and financial modeling pipelines.

**Methodology**

1. **Data Collection and Classical Model Implementation:**
   * Initially, we gathered a comprehensive dataset of economic indicators with labels identifying historical financial crises. This dataset served as the baseline for both classical and quantum models.
   * We then applied preprocessing techniques to clean and filter the data. This filtered dataset was used to implement classical models, which provided a control performance benchmark across the entire dataset.
2. **Data Preprocessing and Feature Reduction:**
   * **Filtering and Feature Selection:**  
     The dataset underwent further filtering using two primary feature selection methods: **correlation-based** and **Principal Component Analysis (PCA)-based** approaches. These methods helped identify the most relevant features and minimize data dimensionality.
   * **Feature Reduction and Quantum Preparation:**  
     After filtering, feature reduction was conducted to retain only the most informative attributes in each dataset version. This step helped optimize computational efficiency while maintaining prediction accuracy in the quantum models.

|  |
| --- |
|  |
|  |

1. **Quantum Feature Maps and Kernel Setup:**
   * **Quantum Feature Maps:**  
     Quantum Feature Maps were tailored to each version of the dataset (correlation-based and PCA-based), effectively transforming classical data into quantum representations. This mapping enhanced the models' capacity to capture intricate patterns within high-dimensional data.
   * **Quantum Kernel Matrix Evaluation:**  
     The project included an evaluation of Quantum Kernel Matrices to analyze and visualize data separability within the quantum feature space. These matrices, coupled with graphical representations, provided insights into the effectiveness of quantum kernels in distinguishing patterns relevant to financial crises.
2. **Data Preparation and Model Implementation:**
   * **Variational Quantum Classifier (VQC) and Quantum Support Vector Classifier (QSVC):**  
     For each filtered and reduced dataset, we prepared and trained **Variational Quantum Classifier (VQC)** and **Quantum Support Vector Classifier (QSVC)** models. These models allowed for comparative performance assessment, with the QSVC models using quantum kernels configured for correlation-based and PCA-based datasets.
   * **Kernel Matrix Evaluation and Graphs:**  
     To further analyze model performance, kernel matrices for the QSVC models were evaluated visually through graphs, illustrating the separability achieved through quantum-based transformations.
3. **Performance Evaluation:**
   * The trained quantum models (VQC and QSVC) and classical models were compared using key metrics: **accuracy, precision, recall, and F1-score**. These metrics enabled a detailed evaluation of predictive capabilities, illustrating the potential of QML for enhanced accuracy and pattern recognition in financial crisis prediction.

**Results**

In this study, we analyzed two distinct datasets to evaluate the performance of **Quantum Machine Learning (QML)** models in predicting financial crises. The datasets used were:

1. **Taiwanese Bankruptcy Dataset** – A highly imbalanced dataset, where instances of bankruptcy are significantly fewer compared to non-bankruptcy instances. This dataset presents unique challenges for model performance, particularly in metrics sensitive to class imbalance, such as precision and recall.
2. **African Banking Crisis Dataset** – A dataset with more balanced representation of crisis and non-crisis instances, allowing for more straightforward evaluation across all metrics.

To test the efficacy of QML in financial crisis prediction, we applied **Quantum Support Vector Classifier (QSVC)** and **Variational Quantum Classifier (VQC)** models, along with a classical **Support Vector Machine (SVM)** model, on both correlation-based and PCA-reduced versions of each dataset. Below are the performance results of each model, assessed using accuracy, precision, recall, and F1-score.

**Results For African Crisis Dataset:**

|  |  |
| --- | --- |
|  |  |
|  |  |

**Results For Taiwan Bankrupcy Dataset:**

|  |  |
| --- | --- |
|  |  |
|  |  |

**Results Table:**

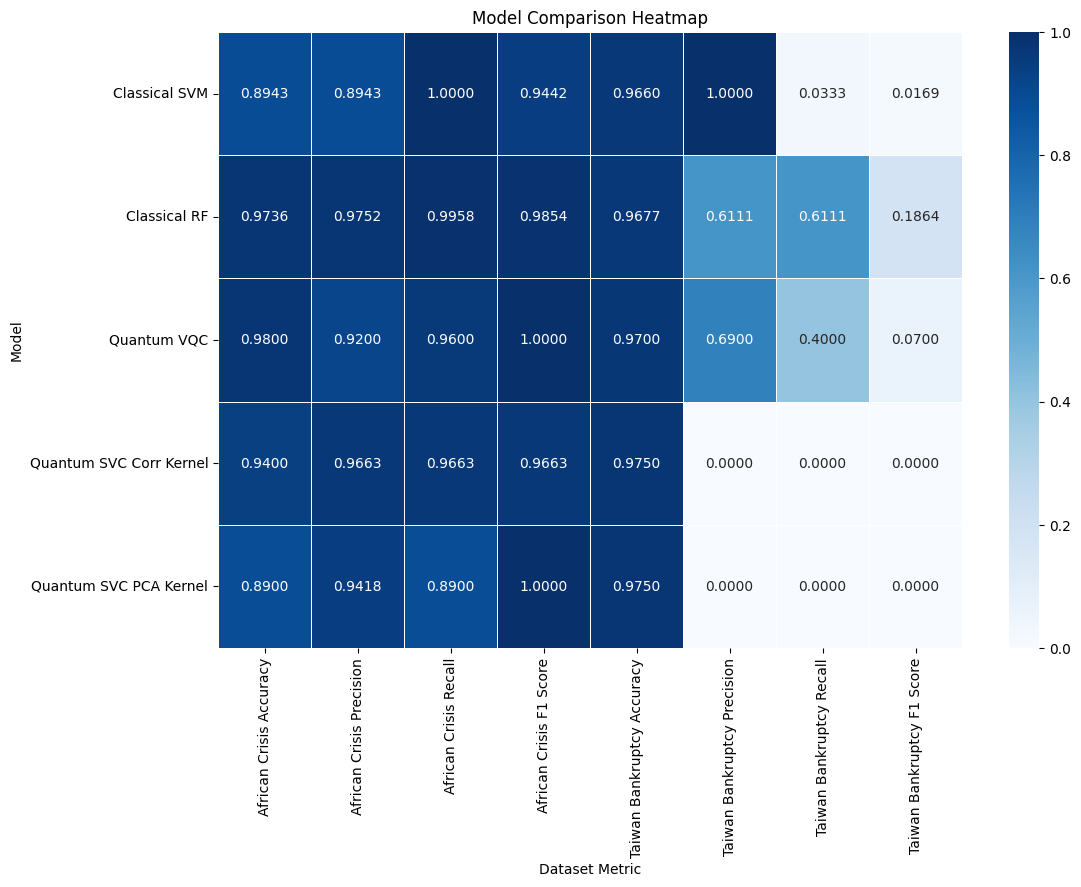
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Data Set | Model | Accuracy | Precision | Recall | F1 Score |
| Africin Crisis | Classical SVM | 0.8943 | 0.8943 | 1.000 | 0.9442 |
| Classical RF | 0.9736 | 0.9752 | 0.9958 | 0.9854 |
| Quantum VQC | 0.98 | 0.92 | 0.96 | 1.00 |
| Quantum SVC Corr Kernel | 0.94 | 0.9663 | 0.9663 | 0.9663 |
| Quantum SVC PCA Kernel | 0.89 | 0.9418 | 0.89 | 1.00 |
| Taiwan’s  Bankruptcy | Classical SVM | 0.9660 | 1.0000 | 0.0333 | 0.0169 |
| Classical RF | 0.9677 | 0.6111 | 0.6111 | 0.1864 |
| Quantum VQC | 0.97 | 0.69 | 0.40 | 0.07 |
| Quantum SVC Corr Kernel | 0.9750 | 0.0000 | 0.0000 | 0.0000 |
| Quantum SVC PCA Kernel | 0.9750 | 0.0000 | 0.0000 | 0.0000 |

### Summary of Results

In the **African Banking Crisis dataset**, all models generally performed well, with the **Random Forest (RF)** model achieving the highest accuracy (0.9736) and balanced metrics. Quantum models, especially the **Variational Quantum Classifier (VQC)**, also performed competitively, achieving an accuracy of 0.98 and a perfect F1-score of 1.000, indicating strong balance between precision and recall. The **QSVC with Correlation-Based Kernel** showed consistent metrics (0.9663 across accuracy, precision, recall, and F1-score), making it a reliable option for balanced datasets.

In the **Taiwanese Bankruptcy dataset**, which is highly imbalanced, all models struggled with detecting minority bankruptcy cases. Classical models, while achieving high accuracy, demonstrated low recall, particularly the **Support Vector Machine (SVM)**, which had a recall of only 0.0333. Quantum models, specifically **QSVC with Correlation and PCA Kernels**, failed to identify any bankruptcy cases, showing precision, recall, and F1-scores of 0.0000. The imbalanced nature of this dataset highlights a critical challenge for quantum models, emphasizing the need for specialized approaches to improve performance on minority class predictions.

**Model Performance:**

****

### Discussion

The performance of **classical and quantum models** was evaluated on two datasets: the **African Banking Crisis dataset** and **Taiwan’s Bankruptcy dataset**. The results highlight varying levels of effectiveness across models and reveal specific strengths and weaknesses, particularly in handling imbalanced datasets like Taiwan's Bankruptcy dataset.

**1. African Banking Crisis Dataset**

The African Banking Crisis dataset is more balanced in its crisis and non-crisis cases, providing a robust environment for assessing model performance.

* **Classical Models**:
  + **Support Vector Machine (SVM)**: Achieved an accuracy of 0.8943 with a perfect recall of 1.000, indicating all crisis cases were detected. However, the precision of 0.8943 suggests some false positives, which affected the F1-score (0.9442).
  + **Random Forest (RF)**: Outperformed the SVM with an accuracy of 0.9736, high precision (0.9752), and recall (0.9958), resulting in a strong F1-score of 0.9854. The RF model effectively captured both crisis and non-crisis cases, achieving high predictive performance.
* **Quantum Models**:
  + **Variational Quantum Classifier (VQC)**: Demonstrated strong performance with an accuracy of 0.98, precision of 0.92, recall of 0.96, and a perfect F1-score of 1.00. The VQC effectively balanced precision and recall, with minimal false positives and false negatives, highlighting its suitability for crisis detection.
  + **QSVC with Correlation-Based Kernel**: Performed consistently with an accuracy of 0.94, precision, recall, and F1-score all at 0.9663. This balanced performance across metrics suggests that the correlation-based kernel effectively captures crisis-related patterns.
  + **QSVC with PCA-Based Kernel**: Achieved an accuracy of 0.89 and a high precision of 0.9418, but with a slightly lower recall of 0.89. The F1-score was 1.00, indicating that the model minimized false positives, though the lower recall suggests some missed crisis instances, likely due to the dimensional reduction inherent in PCA.

**Key Insights for African Banking Crisis Dataset**:

* **Random Forest and VQC**: Both models excel in crisis detection, with high precision and recall. The RF model outperforms in accuracy, while the VQC achieves the best balance as reflected in the perfect F1-score.
* **PCA-Based QSVC**: Although this model maintains high precision, the slight drop in recall indicates that PCA-based feature reduction might have excluded some crisis-indicative features, which could explain the missed crisis cases.

**2. Taiwan’s Bankruptcy Dataset**

The Taiwanese Bankruptcy dataset is highly imbalanced, with a minority of instances representing bankruptcy cases. This imbalance presents challenges, as evident in the model performance.

* **Classical Models**:
  + **Support Vector Machine (SVM)**: Achieved a high accuracy of 0.9660, but this was due primarily to high performance on non-bankruptcy cases. Precision was perfect (1.0000), but the model showed very low recall (0.0333), with an F1-score of 0.0169. This suggests that the SVM almost entirely failed to detect bankruptcy cases, indicating a strong bias toward the majority class.
  + **Random Forest (RF)**: Slightly improved over the SVM with an accuracy of 0.9677 and a balanced precision and recall (both at 0.6111), resulting in an F1-score of 0.1864. While the RF model attempted to capture bankruptcy cases, its performance remains limited, likely due to the dataset imbalance.
* **Quantum Models**:
  + **Variational Quantum Classifier (VQC)**: Achieved an accuracy of 0.97, precision of 0.69, but a recall of 0.40, and a very low F1-score of 0.07. The VQC was able to detect more bankruptcy cases than the classical models, but the low F1-score indicates an imbalance in precision and recall.
  + **QSVC with Correlation-Based Kernel and PCA-Based Kernel**: Both achieved high accuracy (0.9750), but with precision, recall, and F1-scores of 0.0000, indicating complete failure in detecting any bankruptcy cases. These results suggest that both quantum kernels struggled with the imbalanced data, likely defaulting to the majority class, as they did not identify any bankruptcy instances.

**Key Insights for Taiwan’s Bankruptcy Dataset**:

* **Challenges with Imbalance**: All models, especially the quantum models with correlation and PCA kernels, struggled with the imbalance, largely defaulting to classifying non-bankruptcy cases. The extreme class imbalance likely caused these models to overfit to the majority class.
* **Need for Advanced Techniques**: The low recall across models indicates the necessity for specialized techniques, such as re-sampling, cost-sensitive learning, or hybrid models that specifically address imbalanced data.

**Summary of Performance and Areas for Improvement**

* **Balanced Dataset (African Crisis)**: Quantum models, particularly the VQC and QSVC with a correlation-based kernel, showed competitive performance relative to classical models. Both classical and quantum models effectively captured crisis patterns, though the QSVC with a PCA kernel exhibited slightly lower recall, suggesting some limitations with dimensionality reduction.
* **Imbalanced Dataset (Taiwan Bankruptcy)**: All models underperformed in detecting minority bankruptcy cases. Classical models, although slightly better in recall than quantum models, still demonstrated significant bias toward non-bankruptcy cases. The quantum models, especially QSVC with correlation and PCA-based kernels, failed entirely to detect the minority class, highlighting a key limitation in quantum models when faced with highly imbalanced datasets.

### Conclusion

This project investigated the application of Quantum Machine Learning (QML) in predicting financial crises, comparing the performance of quantum and classical models across two datasets: the balanced African Banking Crisis dataset and the highly imbalanced Taiwanese Bankruptcy dataset. Our findings reveal that QML models, particularly the **Variational Quantum Classifier (VQC)** and **Quantum Support Vector Classifier (QSVC) with Correlation-Based Kernel**, can perform competitively with classical models, such as Support Vector Machines (SVM) and Random Forest (RF), in balanced datasets. Notably, the VQC achieved a perfect F1-score, demonstrating a strong ability to maintain precision and recall, which is essential for reliable crisis detection in financial forecasting.

However, the results on the Taiwanese Bankruptcy dataset highlight a significant limitation: quantum models struggled to detect minority class instances in the face of severe class imbalance. The **QSVC with Correlation-Based and PCA-Based Kernels** defaulted almost entirely to the majority class, failing to identify bankruptcy cases. This limitation underscores the need for further optimization, such as implementing imbalance-handling techniques or hybrid quantum-classical architectures, to enhance the sensitivity of quantum models in imbalanced datasets.

Overall, this study demonstrates the potential of QML for financial crisis prediction, especially in balanced datasets, where quantum models show notable strengths in precision-recall balance. Moving forward, efforts to improve QML’s applicability to imbalanced financial data will be crucial, including exploring advanced feature engineering, quantum kernel adjustments, and hybrid approaches. These developments may pave the way for QML models to become valuable tools for high-stakes financial forecasting and risk management.