Automated Quantum Support Vector Machines

Darsh Khandelwal[1]

[1] *Department of Physics, Indian Institue of Science Education and Reserach Bhopal (IISER Bhopal), India*

(Dated: January 26, 2025)

This report deals with explaining the working of Support Vector Machines (SVMs), and their quantum counterpart, Quantum Support Vector Machines (QSVMs) using mathematical approach. Then, we perform the Exploratory data analysis (EDA) of the Statlog German Credit Data and discuss the key results. Thereafter, we implement Automated Quantum Support Vector classifier showed in [1] on German Credit dataset and discuss the results.

I. INTRODUCTION

Support Vector Machines (SVMs) are supervised machine learning models widely used for classification and regression tasks. Their quantum counterpart, Quantum Support Vector Machines (QSVMs), employ quantum feature maps which encode the translation of classical information to quantum space. These feature maps transforms data to higher dimensional hilbert space leading to better solutions for classification of non linearly seperable datasets. These work combined with concepts of genetic algorithms can be used to automate the selection of feature maps for design of QSVM.

II. SUPPORT VECTOR MACHINES (SVMS)

Support Vector Machines (SVMs) is a robust ML model for classification purposes. They are designed to find the optimal hyperplane by maximizing the distance between the marginal hyperplanes which are consturcted using parallely transporting the candidate hyperplane. The equation of marginal hyperplanes can be described as below:

$$w \cdot x + b = +1 \quad \text{(Positive Hyperplane)} \tag{1}$$

$$w \cdot x + b = -1 \quad \text{(Negative Hyperplane)}.d = \frac{2}{\|w\|}. \tag{2}$$

where, d is the distance between these hyperplanes. The optimization problem is defined as:

$$\min \frac{1}{2}\|w\|^2 \quad \text{subject to} \quad y_i(w \cdot x_i + b) \geq 1, \forall i. \tag{3}$$

The decision function depends solely on the support vectors:

$$f(x) = \text{sign}\left(\sum_i \alpha_i y_i (x_i \cdot x) + b\right), \tag{4}$$

where, x_i are Support vectors, α_i are Lagrange multipliers and y_i are class labels ($+1$ or -1). This optimization is known as hard margin SVM because of the strict constraint present in the above equation.

In real-world data, perfect distinction is often not possible. The soft margin SVM introduces slack variables ξ_i to allow for some misclassification:

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0. \tag{5}$$

The optimization problem becomes:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \xi_i, \tag{6}$$

where C is a regularization parameter controlling the trade-off between maximizing the margin and minimizing classification errors. The kernel trick avoids explicit computation of $\phi(\mathbf{x})$ by using a kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$:

1. Linear: $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$,
2. Polynomial: $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + c)^d$,
3. RBF: $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\big(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\big)$.

III. QUANTUM SUPPORT VECTOR MACHINES (QSVM)

QSVMs uses quantum circuits which acts as feature maps to encode classical data into quantum states, enabling the computation of quantum kernels:

$$K(\mathbf{x}_i, \mathbf{x}_j) = |\langle \psi(\mathbf{x}_i) | \psi(\mathbf{x}_j) \rangle|^2, \tag{7}$$

where $\psi(\mathbf{x})$ is a parameterized quantum feature map. This kernel is then used in a classical SVM to classify data.

IV. EXPLORATORY DATA ANALYSIS OF DATASET

The Statlog German Credit Data [2] includes 1000 data points labeled with 1 as Good or 2 as Bad entry. Dataset was divided into 80% train and 20% test data points. Each entry contained 20 features which gave information about Credits, Balance in Savings and Checking account, Duration of loans, Age, Sex, etc.

We perform the exploratory data analysis to gain insights about the dataset. It provides the information about measure of the importance of each feature and confirms their inclusivity in the dataset.

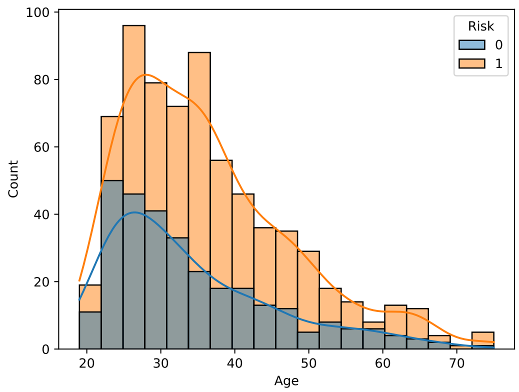

FIG. 1. Distribution of Age with risk parameters

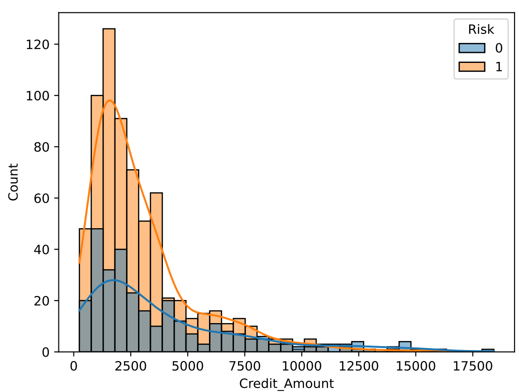

FIG. 2. Credit Amount distribution by risk.

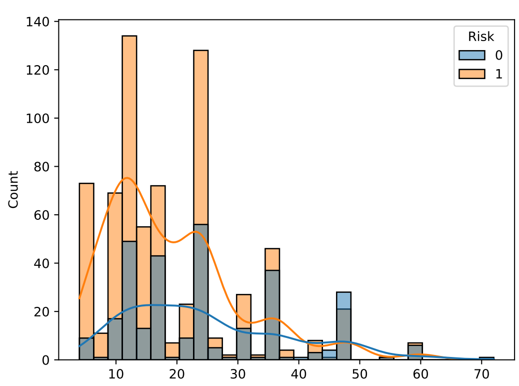

FIG. 3. Duration Distribution by risk.

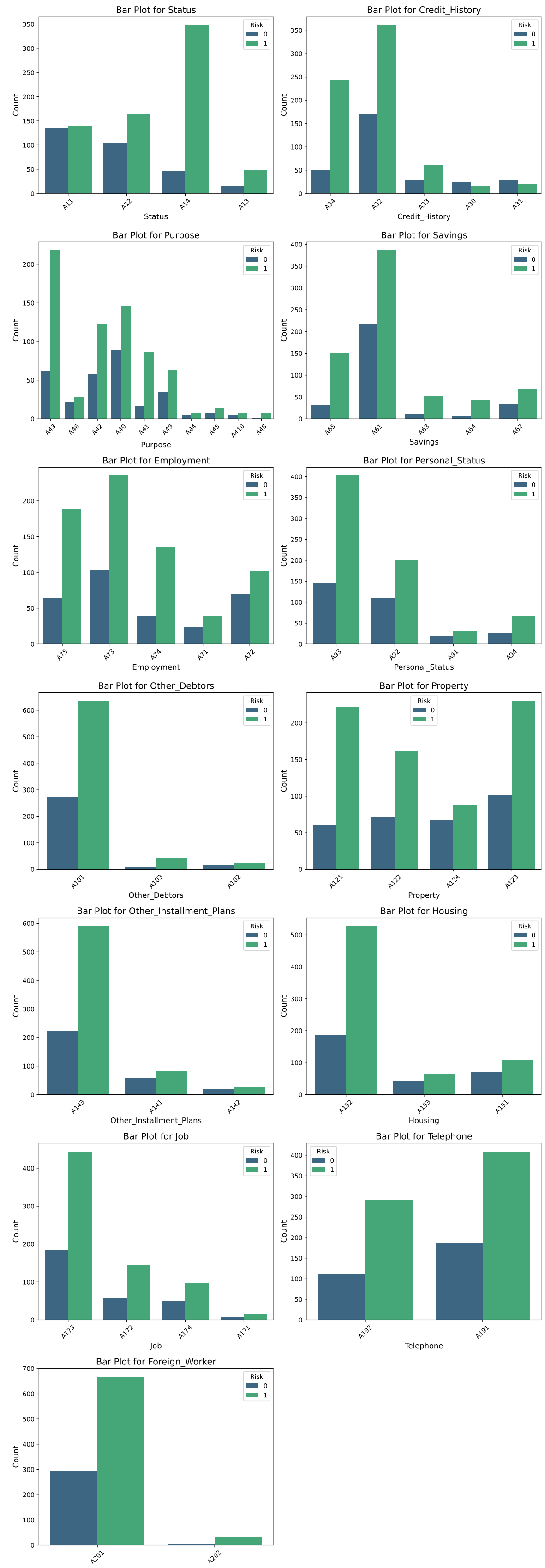

FIG. 4. Bar plot of categorical variables by risk.

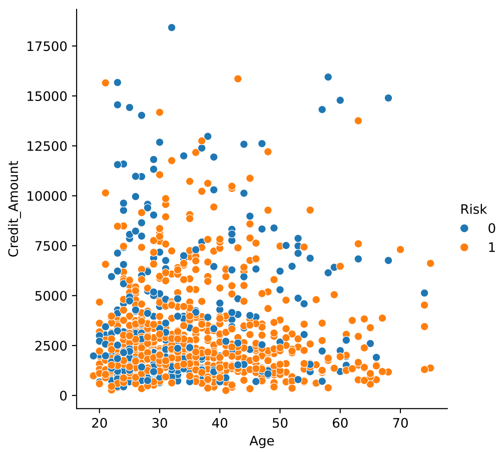

FIG. 5. Relationship between Age and Credit Amount

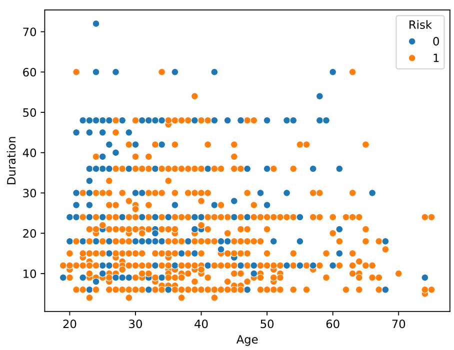

FIG. 6. Relationship between Age and duration

V. AUTOMATED QSVC (AQSVC)

Automated QSVC optimizes the feature maps in order to maximize the accuracy of the model and minimize the number of gates employed in the model. We implement the AQSVC shown in [1] for German credit data

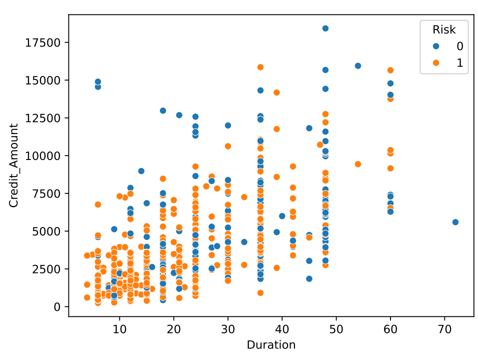

FIG. 7. Relationship between duration and credit

set. They employ multi-objective genetic algorithms to determine the best feasible feature map having the least expressibility.
Genetic algorithms are adaptive heuristic search algorithms which belong to larger part of evolutionary algorithms. Working of these algorithms is similar to the working of reinforcement learning model of machine learning.

A. Working of Genetic Algorithms

1. Initial population of the solution is taken randomly as string of binary digits. The number of elements are obtained as $n_{qubits} \times n_{depth} \times n_{gate}$.

2. Fitness function is evaluated based on the objective of the optimization problem.

3. Depending on the output of the fitness function, the genes in a particular generation are classified as accurate or not.

4. Accurate genes are used for generating new offsprings for the next generation and non useful genes are substituted by the new offsprings. The generation of new offspring are done by crossover and mutation.

5. When the number of iterations reach the number of generation. It stops evaluating and stores the best optimized output in pareto front.

VI. RESULTS AND DISCUSSION

We run the algorithm for achieving three different optimization tasks for the feature maps. The number of generations for which the algorithm ran was 1000 ($n_{gen} = 1000$), the model was trained using all the parameters, and the feature map was allowed to use maximum of 10 parameters ($n_{param} = 10$) into the feature map.

1. Maximizing accuracy and minimizing expressiblity.

2. Maximizing recall and minimizing expressiblity.

3. Maximizing the recall.

The feature map for condition 1 is maximum accuracy (acc = 0.76) and minimum number of gates (gates = 6) is given in Fig.8. For condition 2, maximum recall (recall = 1.0) and minimum gates (gates = 3) is given in Fig.9. For condition 3, maximum recall (recall = 1.0) is given in Fig.10

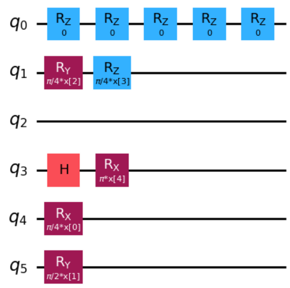

FIG. 8. Feature map 1

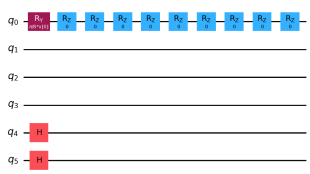

FIG. 9. Feature map 2

VII. OTHER APPROACHES TO HANDLING AQSVC

We can employ several other model to approach the problem. Firstly, I would like to propose the use of reinforced learning instead of genetic algorithms. Secondly, We can also employ neural networks to modify the feature map of the circuit. Another approach could be by emplying Quantum Machine Learning technique by training the parameters of the variational circuit rather than just obtaining the feature map.

VIII. INTEGRATION OF QUANTUM COMPUTING AND ML, AND SHIFT TO QML

Machine Learning has already developed its significance and trust in today's market by accurately providing the optimal solutions to the problems such as predictive maintenance, fraud detection, risk analysis, etc. This trust is not only developed through its predictive capabilities but also through its explainability.

While, QML is still in its infancy period and its development would also require the transparency about its working. This can be developed through focusing our research more on Quantum Explainable AI (Q-XAI) [3]. Frameworks like SHAP and LIME are already widely used in ML. SHAP is used to describe the contribution of each input feature to predict the output. I believe the use of quantum explianable AI would be drastically helpful for smooth transition and integration of both the fields. I propose to apply Explainable AI and evalaute the SHAP scores of models to explain the working and influence of the model on different parameters.

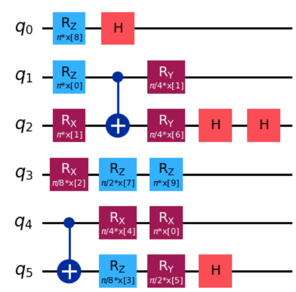

FIG. 10. Feature map 3

[1] S. Altares-López, A. Ribeiro, and J. J. García-Ripoll, Quantum Science and Technology **6**, 045015 (2021).

[2] H. Hofmann, Statlog (German Credit Data), UCI Machine Learning Repository (1994), DOI: https://doi.org/10.24432/C5NC77.

[3] R. Heese, T. Gerlach, S. Mücke, S. Müller, M. Jakobs, and N. Piatkowski, Explaining quantum circuits with shapley values: Towards explainable quantum machine learning (2023), arXiv:2301.09138 [quant-ph].