Malware Technical Exchange Meeting
26-28 July 2022

Can Feature Engineering Help Quantum Machine Learning for Malware Detection?

Ran Liu, Maksim Eren, and Charles Nicholas
University of Maryland Baltimore County

UMBC
DREAM LAB

Introduction

The Android App Store receives hundreds if not thousands of new applications every day. Most of them are benign, but some of them are malware, which leads us to ask "can we use machine learning to detect malware specimens and classify them into the correct families?" To answer this question, two problems need to be solved. The first one is known as the imbalanced dataset problem.The second problem is to detect malware efficiently, which can be reduced to the problem of model training, followed by a relatively fast classification process. In this research, we explored the feature engineering technique usage based on classical ML methods to help train classifiers on quantum gate model machines for malware detection.

Qubit and Their Proprieties

'Qubit' or 'Quantum bit' is the fundamental analogous concept in Quantum computation, which is defined in the Hilbert Space. Using Dirac notation, given an orthonormal basis in Hilbert Space, qubits can be measured to a basis state with probability. We can express a quantum system using state vector language. For example, we can generate superpositions by form linear combinations of states with computational basis $|0\rangle$ and $|1\rangle$:

$$|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad (1)$$

where $\alpha, \beta \in \mathbb{C}$, and $|\alpha|^2 + |\beta|^2 = 1$. Unlike classical bits which can only be in state 0 or 1, qubit can be in the continuum of basis until it has been observed. For example, 1 means qubit can be in state the $|0\rangle$ with probability $|\alpha|^2$ or in the state $|1\rangle$ with probability $|\beta|^2$.

Variational Quantum Classifier

In our research, we use the Variational Quantum Classifier proposed by Havlíček, V., Córcoles, A.D., Temme, K. et al. [1]. A classical data point in the interval $(0, 2\pi]$ is first mapped onto the Bloch sphere using the non-linear circuits constructed as:

$$\mathcal{U}_{\Phi}(\vec{x}) = U_{\Phi(\vec{x})} H^{\otimes n} U_{\Phi(\vec{x})} H^{\otimes n}$$

H is the Hadamard gate and

$$U_{\Phi(\vec{x})} = \exp\left(i \sum_{S \subseteq [n]} \phi_S(\vec{x}) \prod_{i \in S} Z_i \right)$$

Then variational circuits are constructed for the optimization purpose, which is called as the l-layer circuits. The l means repeat constructed variational circuits l times. Next the Z-basis measurements $f : \{0,1\}^n \longrightarrow \{+1, -1\}$are applied to the output bit strings. In the end, measurements are repeated n times to give the probability distribution. And a label as assigned to the input vectors.

Experiment Results using Feature Selection

In our experiments, we use the Drebin dataset consisting of +15k benign and malware samples with 215 dynamic and static analysis based features [2]. We first implemented feature selection strategies to reduce the data size and malware classifier training time. We utilize XGBoost and Decision Tree (DT) to identify the top 20 most important features. Using these features, we next conducted our experiments on the IBM Qiskit Simulator with Variational Quantum Classifier (VQC). The preliminary results show that VQC with XGBoost selected features can get a 78.91% test accuracy score on 10,000 samples with a 50% test set split. Differently, VQC with DT feature selection got 62.41% test accuracy. We also conducted tests using IBM 5 qubits machines. Our experiments in this round were limited by the number of qubits in IBM Quantum Machine. Therefore we ran our experiment ten times with 20 randomly selected samples on each run with a 50% test set split. To show that our results are statistically significant, we report our final accuracy score with a 95% Confidence Interval (CI), using QSVM. The final average accuracy for the model trained using the features selected with XGBoost was 74% (+- 11.35%).

Experiment Results using Sampling Strategy

We next explored oversampling and under-sampling strategies based on classical ML methods to help train classifiers on quantum gate model machines for malware detection. Our preliminary experimental results show:

- In the case of using oversampling techniques, the accuracy increased up to 80.15% from 74% with SMOTE, and up to 78.06% with Adaptive Synthetic (ADASYN).
- In the case of using under-sampling techniques, the accuracy increased up to 78.62% from 74% by removing samples using K-means, and up to 77.26% by removing samples which do not agree "enough" with their neighborhood.
- To overcome noisy samples introduced by SMOTE, we combined the use of oversampling and under-sampling, the accuracy increased up to 83.78%.
- We conducted our experiments on the IBM 5 qubits Quantum Machine. Due to the limited number of qubits available on the Quantum Machine, we use 20 samples with a 50% test set split. The Quantum SVM got 56% accuracy and VQC achieved 80% success for malware detection.

References

[1] Vojtech Havlicek, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, Jay M. Gambetta.
Supervised learning with quantum enhanced feature spaces, 2018.

[2] Suleiman Y. Yerima and Sakir Sezer.
Droidfusion: A novel multilevel classifier fusion approach for android malware detection.
IEEE Transactions on Cybernetics, 49(2):453–466, 2019.

Presented at the *Malware Technical Exchange Meeting* (MTEM), MIT Lincoln Laboratory, July 26-28, 2022.