Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Authentication using Dynamics Keystrokes and Quantum Machine Learning

Namisha Bhasin [1*], Sanjay Kumar Sharma[2], Rajesh Mishra[3]
[*1, 2, 3] *Gautam Buddha University, Greater Noida,201312, Uttar Pradesh, India, namishabhasin@gmail.com*

ARTICLE INFO

Received: 20 Oct 2024

Revised: 22 Nov 2024

Accepted: 10 Dec 2024

ABSTRACT

Authenticating a user based on his/her typing pattern is known as the keystroke dynamics. Here, Authentication is based on user typing data and user typing data cannot be copy by anyone including machine. Authenticating a user at the time of login is called static keystroke dynamics, whereas authenticating a user after login is called free text authentication. To date, free-text/dynamics/continuous authentication statistical and classical machine learning algorithms have been used. However, in this study, we solve the problem of authentication using classical, and quantum algorithms. The given dataset contained two types of information 1) text and 2) typing rhythm. We use typing rhythm data to solve the authentication problem. Out of classical, and quantum machine learning algorithms, the best performance was achieved by the QSVM algorithms. QSVM is able to solve with 100% accuracy and 0% EER. Among classical fusion of MLP and RF is able to solve the problem with 99.3% accuracy and EER of 0.007%.

Keywords: QSVM, QNN, RF, fusion of MLP and RF, VQC

1 INTRODUCTION

There are situations such as writing an email or writing about a given topic, where users initially make a blueprint about what they are going to write and then think and type about the given topic. Typing mistakes might be committed and can be corrected using the backspace or delete key. It is important to confirm whether the entered data are by a genuine user. Here, authenticating a user based on the email received/online typed text response in their own words is known as free-text/continuous/dynamics authentication. Based on the rhythm/pattern of typing a decision, the typed text is taken by the user[1].

For an online exam, a user must type about a given topic and the teacher must verify whether the answer submitted by the student is entered by him/her. Free-text authentication was used to overcome this challenge. It is known that the human brain can easily solve complex problems, such as recognizing a person using their walking style. This is also known as gait recognition and is a simple task for the human brain. This was observed during World War 2 (from 1939-1945). The army could confidently identify whether the message they received, irrespective of the content, was sent by an authorized person. It is easy for the human brain to recognize a person based on typing rhythm, but it is not easy for a machine. Hence, a system is required in which a user can be authenticated according to his/her typing style[2].

Authenticating a user by typing rhythm is an inexpensive method, as no extra device is required, as in gait recognition cameras, microphones, or acoustic sensors, as keyboards are the only hardware required. Because imposter and genuine users type about the same topic, their content and typing rhythm will be different. The authentication of a user problem is solved by how typed- based on typing rhythm and ignoring the content and not what is typed- considering the time for typing and content[3].

This problem of authentication has been solved by classical, quantum, and hybrid algorithms. QC considers qubits instead of bits which are actually vectors[4]. Classical data, in the form of bits, must be converted into qubits for processing using quantum computers. Qubits are entities that convert data into a Hilbert state(like projecting classical data to higher dimension), as shown in Figure 1. Qubit data are converted back to classical data using the principal component analysis(PCA) method, as shown in Figure 2.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Hilbert state is a complete vector space equipped with an inner product that can be helpful in calculating the lengths and angles of the parameters. Ket (|ψ⟩) also known as state vector in QC describes the state of a quantum system. For a single qubit, the state vector can be written as: |ψ⟩=α|0⟩+β|1⟩ where α and β are complex numbers, and |0⟩ and |1⟩ are the basis vectors of the qubit's Hilbert space. The coefficients α and β must satisfy the normalization condition: $|\alpha|^2+|\beta|^2=1$. Superposition, which makes a quantum system in multiple states simultaneously, and Entanglement, where the state of one particle cannot be described independently of the state of the other(s) while the distance between them does not matter. When applied to the |0> state, the Hadamard gate converts into $H|0>=\frac{|0>+|1>}{\sqrt{2}}$ and when applied to |1>, it converts into $H|1>=\frac{|0>-|1>}{\sqrt{2}}$. The entangled state can be described by the Bell state[5].

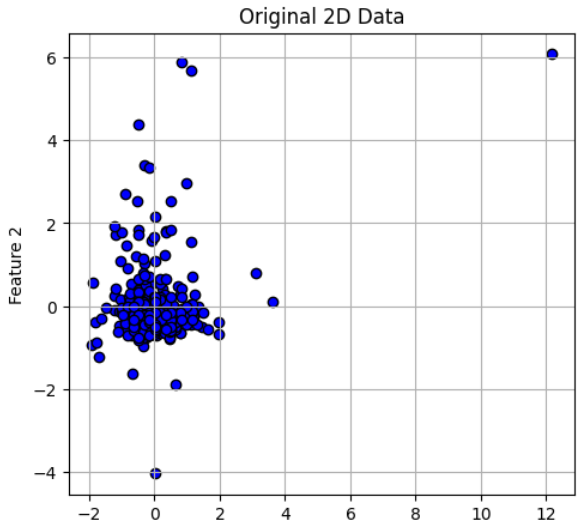

Figure 1: representation of classical data

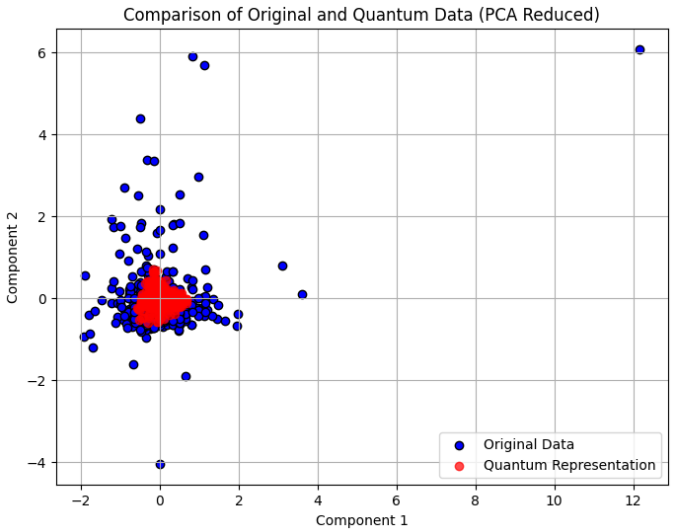

Figure 1: conversion of qubit to classical data with PCA reduced method

Quantum gates are used by in quantum computers to convert data into qubits(where data is projected to higher dimension), whereas for projection of data to higher dimension kernels are used by machine-learning algorithms(known as kernel trick). As kernels transform data into a higher-dimensional space, they make classification tasks easier, but computationally expensive. Quantum kernels can map classical data into quantum Hilbert space by exploiting quantum properties, such as superposition and entanglement, to improve the algorithm

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

performance. Kernel matrix K for a dataset{x_i} is defined as $kij =< \phi(xi) \mid \phi(xj) >$ [7]. In a hybrid system combining qubits and kernels, algorithm's performance can be enhanced.

In QC, various gates exist:

1) the Hadamard gate(represented as H), which helps to create superposition,

2) the Pauli-X gate(X), which can flip the state of a qubit, and

3) the CNOT gate(CX), which can entangle two qubits[6].

This paper contains section 2 - background, section 3 explains the dataset, Section 4 describes data processing techniques, Section 5 discusses architecture, Section 6 discusses the results and discussion, Section 7 is used for the conclusion, and Section 8 represents future work.

2 BACKGROUND

In their study, the authors used the flight time calculated by dividing the keyboard's key into left, right, and four-line zones and classifying the data by forming a vector based on the matching string of two users. Based on the Euclidean distance, they achieved an FRR of 4.0 and FAR of 2.0[8].

The authors used the Pierce criteria for outlier removal. Then, from the remaining data, they calculated 1-graph and 2-graph and combined them for processing. The model was trained with the neural network, and the authors also used an NN to obtain missing values of 2-graph parameters. They were able to solve the problem with an FAR of 0.0152% and an FRR of 4.82% for 53 users. When this concept was applied to the data of 17 fresh users, the authors achieved an FAR of 0% and an FRR of 5.01%[9].

The authors used the concept of Instance-based Tail Area Density (ITAD), where the "tail areas" of each user were compared, and they were able to predict the correct user in less than 500 strokes. They used two features for their research paper: 1) monographs and 2) digaphs. They applied a random forest classifier to assess how effectively each feature could identify who was typing, using two public datasets for validation. A fused matching score is created by combining the results from each feature, which improves the overall accuracy of user authentication. The study achieved significantly lower Equal Error Rate (EER) of 9.7% and 7.8% for tests using 100 and 200 digraphs from the Clarkson II dataset, respectively[10].

The authors considered that every user has a hand movement pattern during typing, which can act as a way of authenticating a user. The authors recorded the hand movements of 63 users when they were typing on the keyboard for both types of data, that is, static and dynamic, using a webcam and recorded the hand movements of the user while typing. For these test data, the authors considered both cross-features(where different types of hand features are considered) and cross-temporal(where movement over time is checked) patterns. The authors used static datasets as training data and dynamic/continuous/ free text as the testing datasets. A window size of 5-20 seconds was used. The authors were able to achieve the best performance with an area under the curve(AUC) of 99.96% by using a bag of multi-dimensional phases for free-text (letter writing)[11].

Typing using a keyboard and touchscreen are different in some ways. Keyboard keystroke features such as hold time, latency, and rhythm are considered, whereas in a touchscreen system, tap features such as tap duration, location, force, and rhythm are considered. Tap duration explains how much time a user takes to press a particular spot on the screen, tap location explains where the user taps, force explains how much pressure is put on a particular spot, rhythm explains speed and timing taken by a user for multiple taps, and all these features describe how users interact with the touchscreen devices. The authors used hand movement, orientation, and how a user holds, moves, and interacts with the device as features. The authors collected data from 100 users and used them to authenticate them based on their typing, during which they could be in a sitting, standing, or walking position. They collected 1) resistance features that notice small shifts, and 2) stable features that determine how fast the device regains its balance. The authors found that features collected using an accelerometer and gyroscope were better than those obtained using a magnetometer, as the former were able to capture micromovements. The authors found that the EER value was lower in the walking than in the sitting position because of unique body movements. The authors were able to have a reliable system with tap features(an EER of 15.1%) rather than with a combination of tap features,

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

keyhold, and swipe(an EER of 25.7%-34.2%). The authors collected data 16 times per second, which is sufficient to capture fine details of movement with a slight increase in power consumption; this increase is only 7.9%[12].

The operational benefits of quantum computing can be used in various fields of healthcare, such as drug discovery(quantum computers combined with AI help to identify genetic factors and target drugs), personalized medicine(an efficient DNA analysis can lead to personalized healthcare solutions), DNA sequencing(), medical imaging, and operational optimization[13].

The authors used photonic quantum devices known as Gaussian Boson Samplers(GBS), which process information using the quantum state of light as a unique property of photons, to solve complex mathematical problems that are difficult to handle using classical computers. The authors used GBS to predict the molecular docking configurations, which is a crucial task in drug design. This concept plays an important role, as it allows us to understand how a drug interacts with its targets. The authors solved this problem using the graph method, where they converted ligand binding to proteins into a clique and then found the maximum weighted clique. The authors applied this concept to ligand binding of thiol-containing aryl sulphonamide with the tumor necrosis factor-α converting enzyme (TACE). The authors found that when they applied the coarse coarse-grained modelling (CGM) approach to classical computers, they were not able to obtain full detail information, and for quantum computers, this CGM was not required. They extracted features using RDKit software; for node selection, they used pharmacophore pairs within 4 Å, followed by graph construction. to increase the sample efficiency, random search, greedy shrinking, and greedy shrinking with local search are augmented with GBS. The authors found that the performance of GBS-based hybrid methods was better than that of classical approaches[14].

The authors analyzed the genetic sequences using quantum computing. Nucleotides (A, T, C, and G) that store genetic information can be extremely long, and genetic diversity can be checked by knowing the difference between these sequences, which is crucial for fields such as medicine. However, identifying these differences is difficult because of the size and complexity of the data, especially when changes occur at the level of individual nucleotides. The authors solved this problem with human image processing-like capability and for this, they had a Flexible Representation of Quantum Images (FRQI) where genetic sequences are converted into “quantum images” where nucleotide acts as pixels in an image. The sequences were processed to generate quantum circuits representing each sequence, and the rotation differences between the sequences were compared. The authors found accurate similarity assessments by measuring the phase angle differences between nucleotides, recorded as expected values. Using these values allows for an efficient comparison process that captures subtle sequence differences, while requiring fewer computational resources than traditional methods[15].

The authors used a Quantum Support Vector Classifier (QSVC), Variational Quantum Classifier (VQC), Estimator Quantum Neural Network (QNN), and Sampler QNN for fraud detection. The authors achieved an F1 score of 0.98, whereas other algorithms also seem promising with certain limitations. The authors found that improvements in quantum algorithms and handling of larger datasets are necessary to further optimize these models. The authors employed Principal Component Analysis (PCA) for feature selection so that the model focuses on factors most likely to impact fraud detection. The authors found that transaction amounts showed a weak negative correlation with fraud. Thus, as the transaction amount increased, the likelihood of the transaction being fraudulent also increased slightly, even if the correlation was not particularly strong. This association suggests that higher transaction amounts may be a red flag for potential fraud, allowing the model to prioritize transactions with higher amounts for closer scrutiny. The authors used a correlation heatmap to determine the influence of features on fraud detection. The authors found that ZFeatureMap is more suitable for optimizing the VQC model, leading to better accuracy, and that Sampler QNNs may not be the best fit for applications that demand complex pattern recognition across large datasets[16].

3 DATASET

The authors collected data[17] for free text where they used crowdsourcing through Amazon Mechanical Turk to collect truthful and deceptive texts for research. For the data, each participant was assigned one of three topics: restaurant reviews, gay marriage, or gun control, and they had to type two texts: one truthful and one deceptive. For restaurant reviews, participants wrote a real review of the place they liked and a deceptive review of the place they

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

had not visited or did not like. The copy/paste was disabled to ensure that participants typed their responses. Participants were randomly assigned to one of two writing orders: they either wrote the truthful text first, followed by the deceptive one, or vice versa, and were not allowed to switch between these task flows. The dataset for free-text has two modules, where in one module text data that is greater than or equal to 100 words is there; in the second module, the stored data has 1) a time stamp, 2) a key code, and 3) the key status - up and down. Feature selection aims to obtain the most relevant features in a dataset that can increase a model's performance. In machine learning, the feature selection process involves various subsets of features, and chooses the optimal subset based on factors such as accuracy, information gain, or mutual information. The increase in the number of features in datasets increases the number of possible feature subsets, making exhaustive searches impractical. Hence, we can solve this problem by using a feature selection algorithm.

3.1 Feature Engineering

Feature engineering problem for keystroke dynamics authentication is solved as follows

1.In this approach-the number of elements considered is known as the window size. This window size can be 1,2 or n, where 1 means only one element is considered at one time, 2 means two elements are considered at one time, and n means n elements are considered at one time. For this study, a window size of 2 is considered. The manner in which one element is influenced by others is referred to as the number of constructs. The number of constructs can be 1-graph, 2-graph or n-graph. For the 1- graph only one letter's hold time(time taken by a key to press and release) is considered. For 2-graph 2, the hold time and flight time(time taken to release one key and pressing of the second key) are considered, while for n-graph n, the hold time and flight time are considered. In this research paper 5- graph construct is considered. Here 5-graph is considered because the average word length of the content written by the author, whose data we are testing against an imposter, was 5. Hence, this parameter depends on the given data and can vary depending on the given data. In addition, ASCII(American standard code for information interchange) value of, 1) key press and 2) key up are used. For feature selection random forest method is used.

3.2 Data preprocessing

For this approach, four steps are followed as represented below:

1. Outlier Removal

A box plot is used for outlier removal. As shown in Figure 3, with the help of a box plot, hold time values above 200 and below 0 were considered outliers.

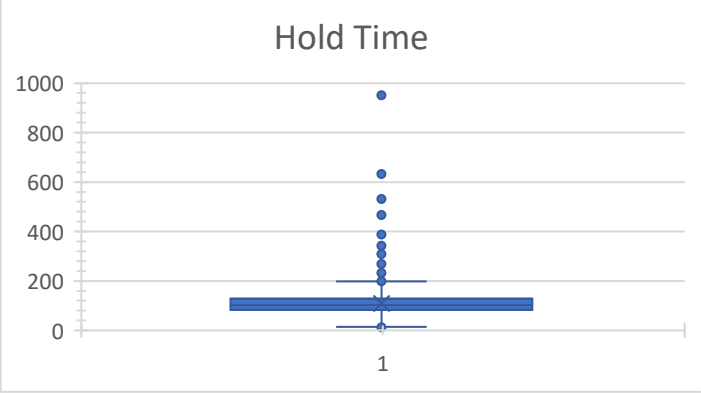

Figure 3: box plot for outlier detection

2. Imbalanced data

A conditional tabular generative adversarial network (CT-GAN) was used for data augmentation. It is a deep-learning-based model used to generate synthetic tabular data. It is based on a mode-specific normalization method in which there is a minimum impact of outliers. The CT-GAN can generate synthetic data and can be used to overcome the problem of data scarcity. CT-GAN consists of three parts: 1) generator used to generate the synthetic data, 2) discriminators that are used to challenge the generated data and force the generator for further

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

improvement, and 3) conditional sampling CTGAN[18] learns from the distribution of columns, which helps to generate realistic data based on specific conditions.

3. Normalization of data

These data were normalized with the help of the Standard and Minmax scaler functions. The standard scaler transforms the data such that its mean becomes zero, and the standard deviation becomes 1. The Minmax scaler transforms data in the range of 0 to 1 depending on the minimum and maximum values.

4. Feature selection and data used for training and testing in ratio 80:20.

4 DATA PROCESSING TECHNIQUES

Various techniques have been used to solve authentication problems using free- text preprocessing data. They are as follows:

4.1 Support vector machine(SVM) [19]

The SVM machine learning algorithm also known as the maximum margin classifier works on the idea of projecting data points in a way so that there is a visible partition between two classes. For classification, it uses a hyperplane that separates the closest data points(also known as support vectors) of two classes. A larger margin between the data points and the hyperplane reduces the risk of misclassification. The SVM also uses kernel tricks if the data points are not linearly separable. Kernel tricks help project data points to different dimensions, making it easy to separate the two classes. If the data are linearly separable, then linear kernels can be used, but if not, polynomial or radial basis function kernels must be used. A sigmoid kernel that is similar to the neural network activation function is also present. But sometimes, it is robust to overfitting when used with the kernel trick. A hyperplane for an n-feature set can be described as $w^T x + b = 0$ where w is a weight vector normal to the hyperplane, x is a feature of a data point, and b is a bias. For testing of a data value, $w^T x + b > 0$ function is used which represents one class if value is positive otherwise it represents another class. Depending upon the type of data different kernel trick functions can be used which are as follows:

1) If data is linearly separable then linear kernel function $K(x_i, x_j) = x_i^T x_j$ is used which is computationally inexpensive. Otherwise,
2) polynomial kernel which is represented as $K(x_i, x_j)=(x_i^T x_j+1)^d$ but still here in this case dataset is well structured, or
3) the radial basis function(RBF) is used when data is complex too and can be represented as $K(x_i, x_j)=\exp(-\gamma||x_i - x_j||^2)$, or
4) sigmoid function can be used to convert SVM output into probability when for classification probability estimation is required and is represented as $K(x_i, x_j)=\tanh(\alpha x_i^T x_j+r)$ can be used.

But there are also cases when perfect separation between nonlinearly separable classes was not possible, then soft margin is used in the SVM. For this slack variable (ξi) is used which represents the number of data points that are misclassified but if $\xi_i = 0$, the point is correctly classified; and if $0< \xi_i <=1$, then the point is inside the margin and correctly classified; otherwise, it is misclassified. Classification based on SVM for KD authentication dataset is presented in Figure 4.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

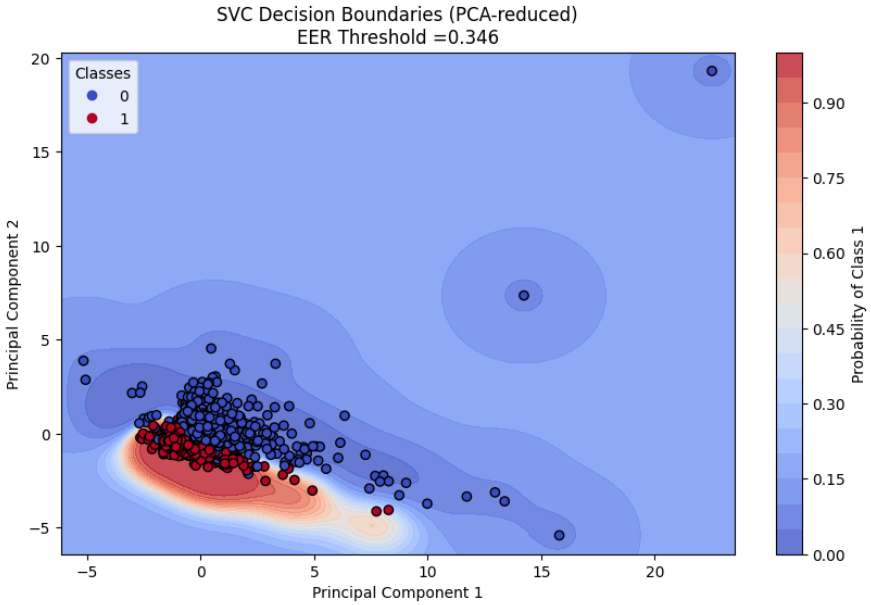

Figure 4: presentation of classification by SVM

4.2.2 One Class-Support Vector Classifier(OC-SVC)[21]

OC-SVC is an unsupervised approach and can be used to identify malicious typing patterns. In this approach, malicious typing patterns is considered as outliers. For this approach, one class data is considered as normal data, and any deviation from this is considered as an anomaly data and for this a decision boundary is created which separates legitimate user data from malicious typing data. OC-SVC can also use kernel tricks(depending upon the type of data) to map data into a higher-dimensional space; then, it will be easy to separate anomalies from legitimate user datasets. OC-SVC is sensitive to the scaling of features; hence, normalization is essential before training, which ensures that all features should contribute equally. The classification based on keystroke dataset using top rank features and OC-SVC with kernel RBF is used.

4.2.3 The Random Forest algorithm[22]

The Random Forest algorithm is a machine learning algorithm and can be used for classification. For this approach, multiple decision trees are created and then they are combined to improve accuracy and reduce overfitting. A decision tree in a random forest may have limited predictive power; however, when many trees are combined, they form a robust model. Each tree has nodes in which data is split based on feature values which can be classified into target classes. They used a technique called bagging, in which they built a tree on a unique subset of data. They drew random samples and features with replacement from the dataset. These samples are then used for training a decision tree, and during this process, they capture different patterns of the data. They considered $\sqrt{f}$ features for training once out of f features. Each tree predicted a class label for the test sample. Considering the result from each decision tree, a class decision is made based on the majority of the trees' decisions. They reduced overfitting by averaging the predictions of multiple uncorrelated trees. Each tree is trained on a different subset of data and features; hence, it prevents the model from fitting any single training sample, noise, or outlier. The higher the number of trees, the higher the accuracy, which also stabilized the output.

4.2.4 Multilayer Perceptron [23]

A multilayer Perceptron(MLP) whose architecture can be used for classification and is a type of neural network that can act as a deep learning model by having at least one hidden layer between the input and output layers. A neural network consists of an input layer, an output layer, an activation function(for different layers relu and for final layer sigmoid activation function is used), weights and biases, a loss function, backpropagation, and gradient descent with at least one hidden layer. The input layer consists of neurons, where one neuron corresponds to a

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

feature. The hidden layer consisted of one or more neurons. In this layer, the calculation consists of the weighted sum of its input. After applying the activation function, the output of the processing is passed to the output layer. More hidden layers mean more neurons; hence, more complex patterns can be identified. Optimization is performed using ADAM function. For binary classification, it uses one neuron with a sigmoid function, whereas for multi-class classification, it uses one neuron for each class, and softmax is used as an activation function. Sigmoid is a nonlinear activation function whose range lies between 0 and 1, whereas for the softmax range is same to sigmoid activation function, but the outputs of all classes sum to 1.

4.2.5. A fusion of Multi-Layer Perceptron (MLP) and Random Forest (RF)

A soft fusion of MLP and Rf provides benefit from MLP's deep learning power, and leverage RF's ensemble stability and interpretability. As MLP captures non-linear patterns from data through deep neural networks. It learns complex interactions through backpropagation. Whereas, RF which is an ensemble of decision trees is robust to noise, overfitting, and can handles tabular data well. It aggregates decisions through majority vote or average probability.

Let P_{MLP} be the predicted probability from MLP, P_{RFP} be the predicted probability from RF and $w \in [0,1]$ be the fusion weight

Then the fused probability:

$$P_{Hybrid} = w \cdot P_{MLP} + (1-w) \cdot P_{RFP} \quad (1)$$

This is a convex combination, satisfying:

$$0 \le w \le 1 \text{ and } w+(1-w)=1 \quad (2)$$

Then a threshold (like 0.5 or the EER threshold) is applied to make the final decision:

$$Y' = \begin{cases} 1, & PHybrid > threshold \\ 0, & otherwise \end{cases}$$

4.3 Quantum algorithms the following part describe the quantum algorithms performance for free-text authentication.

4.3.1 Quantum support vector machine(QSVM) [20]

SVM can be computationally expensive when applied to data with high-dimensional feature sets, and this can be overcome using quantum computers that can perform computations more efficiently with the help of default properties such as entanglement, superposition, and parallelism. Classical data used to map quantum algorithms must be converted into qubits, which can map the data by default into higher dimensions. Hence, quantum computers by default provide this facility, for which the SVM applies after the kernel trick. For QSVM implementation, classical data in the form of a . The csv file is converted into qubits with the help of a ZZ Feature map(which applies controlled z-gates between a pair of qubits to create entanglement between two qubits). After applying the ZZ feature map, Rz creates an interaction between qubit1 and qubit2. In the next step, the Real Amplitude is applied to enhance expressivity. The combined state of ZZ Feature Map and the real amplitude can be used as a quantum feature map in the quantum kernel for QSVM. In this circuit, a Fidelity Quantum Kernel is used. Fidelity measures the similarity between quantum states. If they are orthogonal or opposite, then fidelity is 0; otherwise, it is 1. It captures feature correlations that are difficult to model classically. After applying the kernel function, a quantum support vector classifier was applied.

4.3.2 Variational Quantum Circuits[24]

Variational Quantum Circuits (VQC) are quantum circuits with tunable parameters that can be trained in a manner similar to neural networks. They are typically used in hybrid classical quantum algorithms. Here, quantum computer computations such as data encoding and feature extraction are performed, and the optimization of classical computer functions is performed. There are 3 encoding techniques are available for converting classical data into Qubits. 1.) amplitude encoding, which maps data onto the amplitude of the quantum states; 2) angle encoding maps the data into qubit rotation angles, and 3.) basis, which encodes map data directly onto basis states. Amplitude encoding

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

performed better on large datasets, whereas the other two were more suitable for small datasets. In this study, we considered amplitude encoding. After encoding, an ansatz is used, which helps to find the best solution to a quantum problem. The free-text authentication problem is solved using two ansatzes, as follows:

a) Real amplitude ansatz, where real values can be used as angles for gate rotation. This circuit can be implemented by using RX-like gates. The circuit is composed of parameterized RX gates interleaved with entangling layers(the CNOT gate is used for entanglement). The parameters of the rotation gates are updated during training. The real amplitudes work well for certain machine learning tasks. The real ansatz in the VQC was used for the binary classification. Real Amplitude circuits use only rotation gates with real parameters, such as RY, which apply rotations around the y-axis to simplify computations while still exploring a vast solution space.

b) An Efficient SU2(special unitary group of degree 2) ansatz can construct a highly expressive circuit using minimal layers to approximate any rotation within the SU2 group for each qubit. This circuit is efficient, as only a few parameters are sufficient for functioning compared to an unstructured ansatz. It can take both real and complex values as rotations and combine multiple single-qubit rotations, which helps maximize the expressivity of the quantum circuit. Each layer has three types of rotations per qubit(RX, RY, and RZ) to cover the full range of SU2 transformations. The parameters used for the rotations were optimized to minimize loss. These layers are combined with entangling gates, such as CNOT, to create entanglement between the layers. Entangling enables the representation of the nontrivial higher-dimensional functions required for classification. If the circuit has more layers, more complex patterns can be captured by the circuit.

4.3.3 Quantum neural network [25] use unique properties of quantum computing, such as superposition, entanglement, and interference. Because superposition qubits can be in three states at the same time, entanglement qubits are correlated in such a way that the state of one qubit can influence another even at a distance. Each layer of the QNN is a circuit where instead of layers of neurons, quantum gates act as building blocks and serve as layers. Quantum gates such as the Pauli gate, Hadamard gate, and controlled gates are replaced with classical neuron layers to manipulate the qubit states. The core layers can be 1) parameterized quantum circuits and contain adjustable parameters that are similar to weights and are used to minimize the loss function, as in a classical NN. Real numbers are angles in rotation gates such as Ry/Rz, which can manipulate the qubits; and 2) real amplitude, as shown in Figure 5, quantum circuits where parameters are real numbers and uses only RY gates for real-value rotations. This makes them computationally efficient, while keeping the model simple and less sensitive to noise. For training, the QNN parameters of each gate were initialized randomly, which is similar to initializing the weights in the classical NN. Forward pass is used to pass data through quantum gates, where each gate modifies the qubit states based on the initialized parameters to have qubits in a state where the desired output can be achieved. After processing the data through a quantum circuit, the qubits are measured, which leads to the collapse of their states to classical values. The results are compared with the target output and a loss function with the help of the mean squared error(MSE). During backpropagation, the gradients of the loss functions for each parameter are calculated from the quantum gradient descent algorithm. The QNN uses the parameter-shift rule to calculate gradients(helps to optimize the kernel parameters) in quantum systems, which leads quantum systems to compute gradients without needing to access all qubit states explicitly. Quantum circuits can represent and compute complex functions more efficiently than classical networks can.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 5: flow of data from classical data to qubits by ZZ feature Map which process by RealAmplitudes(for 9 features)

4.3.4 Sampler QNN [26] Instead of estimating the expected value of an observable it generates probabilistic output. Sampler QNN aims to match the target probability distributions. Similar to the estimator QNN, it converts classical data to qubits and then uses variational circuits, which are used to find the value of weights for the optimized output. For classification, the cross-entropy loss function(loss=-$[y\,log(\hat{y}) + (1-y)\,log(1-\hat{y})]$) was used. But the output of this provides a probabilistic output rather than expectation value which can be written as $P(x|\theta) = |\langle x|\psi(\theta)\rangle|^2$.

5 ARCHITECTURE

Here, architecture is described for various algorithms.

5.1Quantum support vector classifier(QSVC): Here, input data are encoded with the help of a feature map, and for this research paper, ZZFeatureMap(a second-order Pauli z circuit) is used, as shown in Figure 6. After encoding the data in the next step, a quantum kernel was applied to measure the similarities between the different quantum states, as shown in Figure 7. Here, a quantum kernel named Fidelity Quantum Kernel(can transform features to high-dimensional space) is used for training. During testing, the parity of the quantum measurement was used to determine the datapoint class.

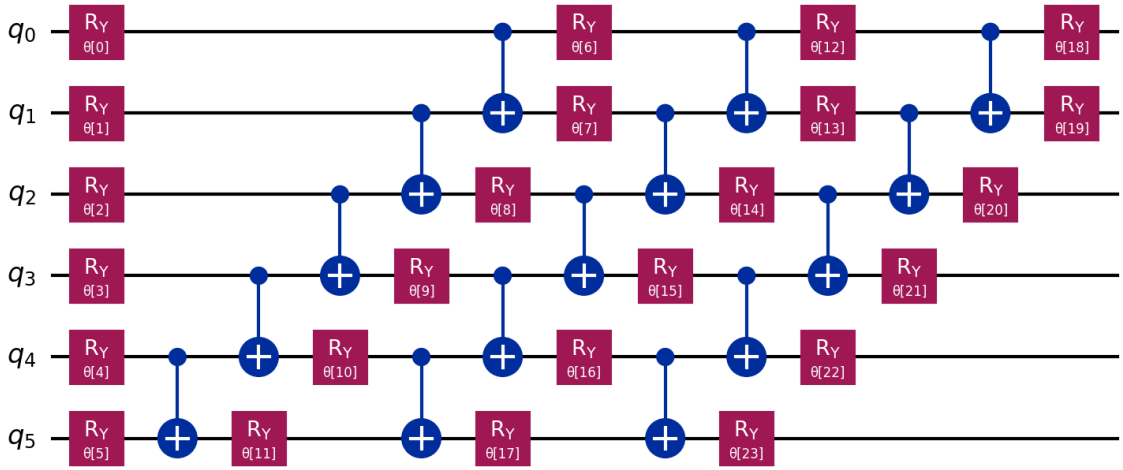

Figure 6: presentation of ZZFeatureMap with Ry and CNOT gate(6 features)

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

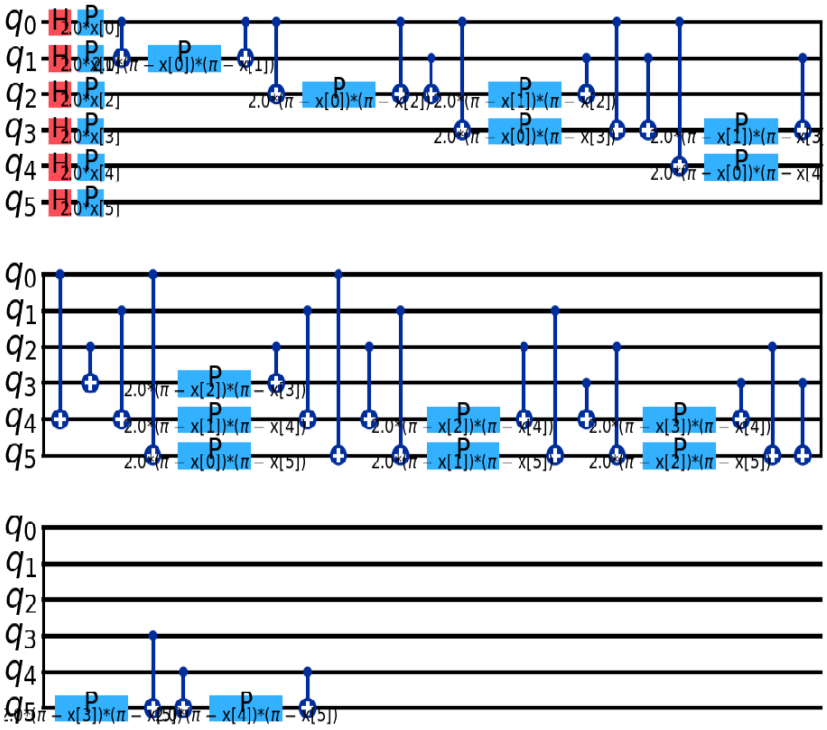

Figure 7: background circuit for QSVC using H gate(for superposition in red), phase gate(labelled as P in blue), CNOT(used for entanglement) and values for parameterized rotation gates

5.2 The variational quantum circuit(VQC) in this algorithm for feature maps used in VQC for data encoding is shown in Figure 8. Ansatz, which is shown in Figure 9, is a parameterized QC that converts the qubits received from the feature map to a quantum state and introduces entanglement and correlations. Two ansatzs can be used in VQC: 1) EfficientSU2 ansatz is considered hardware efficient, where SU2 efficiently represents general single qubit transformation, as shown in Figure 10;and 2) Real Amplitudes, as shown in Figure 11 ansatz, which uses only real parameters and avoids complex numbers. By avoiding complex numbers, the computational cost can be reduced so they are avoided in this experiment. Figure 12 shows how ZZFeatureMap when connected to the Real Amplitudes.

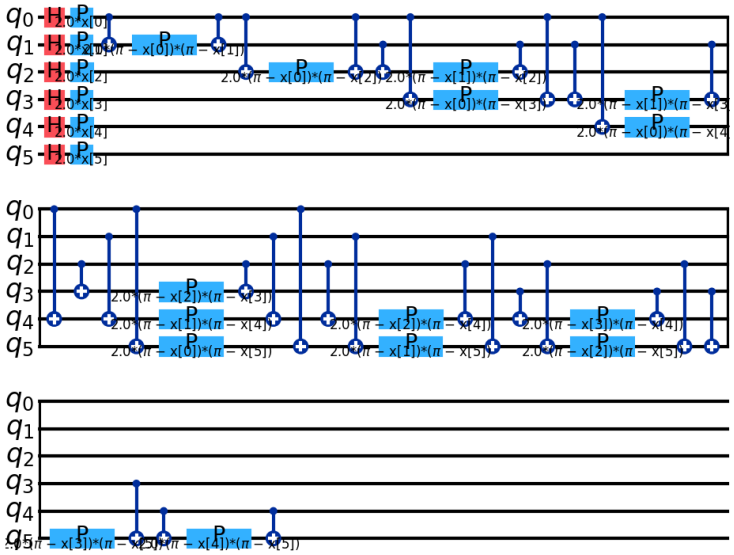

Figure 8: feature map for VQC having H (Hadamard) gate, CNOT gate, phase gate, and values for parameterized rotation gates

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

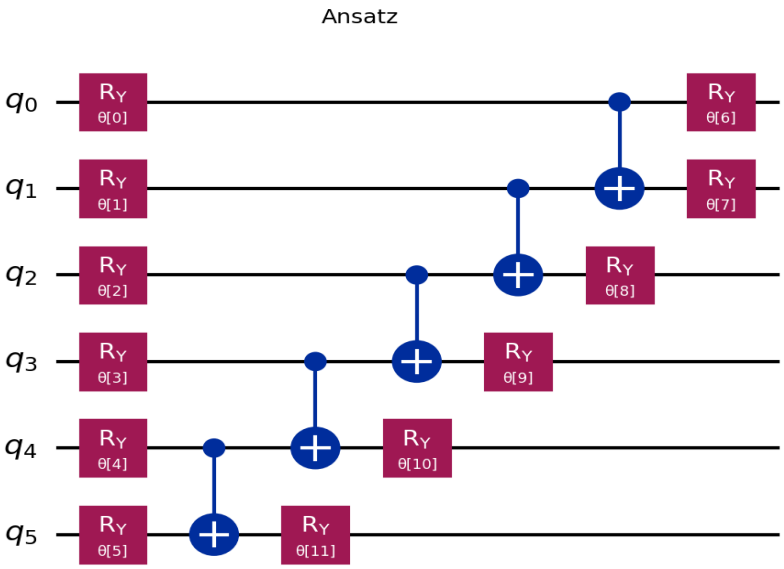

Figure 9: ansatz used in VQC having Ry gate, CNOT gate

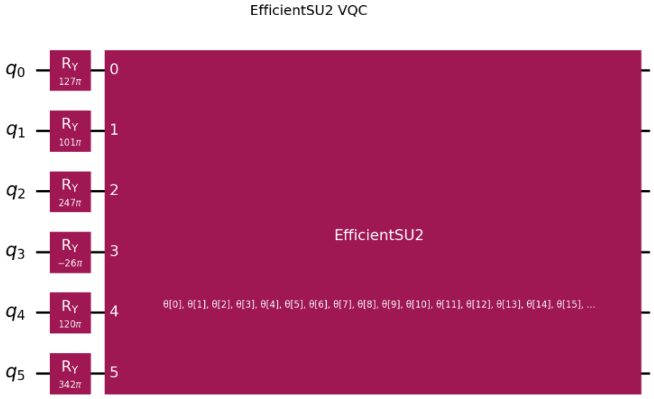

Figure 10: EfficientSU2 as ansatz in VQC

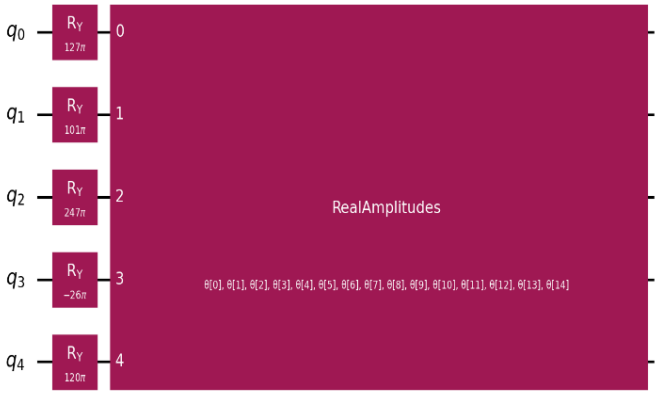

Figure 11: Real Amplitudes as ansatz in VQC

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

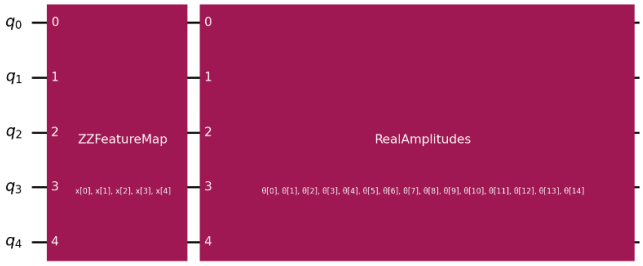

Figure 12: combination of real amplitude with ZZ feature map

6 Results and Discussion

- The result has been checked based on accuracy and EER. As accuracy describes the proportion of correct predictions (both true positives and true negatives) over all predictions. It's also a basic measure of how often the model is right. It is best suited when classes are balanced as is with this dataset. Precision-Recall is where precision predict how many positives are actually correct whereas recall tells how many actual positives are captured. Describe more correctly positive class performance.

The performance of different algorithms based on accuracy is represented in Table 1.

Table 1: accuracy, precision, recall and f1-score for different algorithms

SVC	**Precision**	**recall**	**f1-score**
0.0	0.95	0.98	0.96
1.0	0.98	0.95	0.96
accuracy		0.96	
macro avg	0.96	0.96	0.96
weighted avg	0.96	0.96	0.96
OCSVC			
0.0	0.9414	0.8438	0.8899
1.0	0.8584	0.9475	0.9008
accuracy		0.8956	
macro avg	0.8999	0.8956	0.8953
weighted avg	0.8999	0.8956	0.8953
Random Forest			
0.0	0.9798	0.9700	0.9749
1.0	0.9703	0.9800	0.9751
accuracy		0.9750	
macro avg	0.9750	0.9750	0.9750
weighted avg	0.9750	0.9750	0.9750
MLP			
0.0	0.9801	0.9850	0.9825

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

1.0	0.9849	0.9800	0.9825
accuracy		0.9825	
macro avg	0.9825	0.9825	0.9825
weighted avg	0.9825	0.9825	0.9825
Fusion of MLP+RF			
0.0	0.9929	0.9929	0.9929
1.0	0.9929	0.9929	0.9929
accuracy		0.9929	
macro avg	0.9929	0.9929	0.9929
weighted avg	0.9929	0.9929	0.9929
QSVC			
0.0	1.0000	1.0000	1.0000
1.0	1.0000	1.0000	1.0000
accuracy		1.0000	
macro avg	1.0000	1.0000	1.0000
weighted avg	1.0000	1.0000	1.0000
VQC(Real Amplitude)			
0.0	0.5882	1.0000	0.7407
1.0	1.0000	0.3000	0.4615
accuracy		0.6500	
macro avg	0.7941	0.6500	0.6011
weighted avg	0.7941	0.6500	0.6011
VQC(EfficientSU2)			
0.0	0.8333	1.0000	0.9091
1.0	1.0000	0.8000	0.8889
accuracy		0.9000	
macro avg	0.9167	0.9000	0.8990
weighted avg	0.9167	0.9000	0.8990
QNN			
0.0	0.52	0.58	.49
1.0	0.52	0.52	0.55
accuracy		0.52	
macro avg	0.52	0.52	0.52
weighted avg	0.52	0.52	0.52
Sampler QNN			
0.0	**0.5962**	0.58	0.61
1.0	0.61	0.60	0.59
accuracy		0.60	
macro avg	0.6	0.6	0.6
weighted avg	0.6	0.6	0.6

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

From the table, it is clear that the QSVC accuracy performance is the best among all the algorithms implemented above. The performance of other quantum algorithms is lower than expected because they are implemented in Noisy Intermediate-Scale Quantum devices(NISQ). Quantum computing will be increasingly used in the coming days and needs to be explored further. The result is further analysis on basis of Equal Error Rate(EER). EER is a point on ROC where FAR(False acceptance rate)=FRR(False Rejection Rate). Low value of EER means better performance of the algorithm. It is clear from the EER value too performance of QSVC is better than other algorithms.

Receiver Operating Characteristic - Area Under Curve -The ROC AUC score reflects the model's ability to distinguish between positive and negative classes. A higher ROC AUC value indicates better overall model performance.

1.Support vector machine able to solve KD problem with an Equal Error Rate (EER) of 0.0393 at threshold 0.3462 presented in Figure 13.

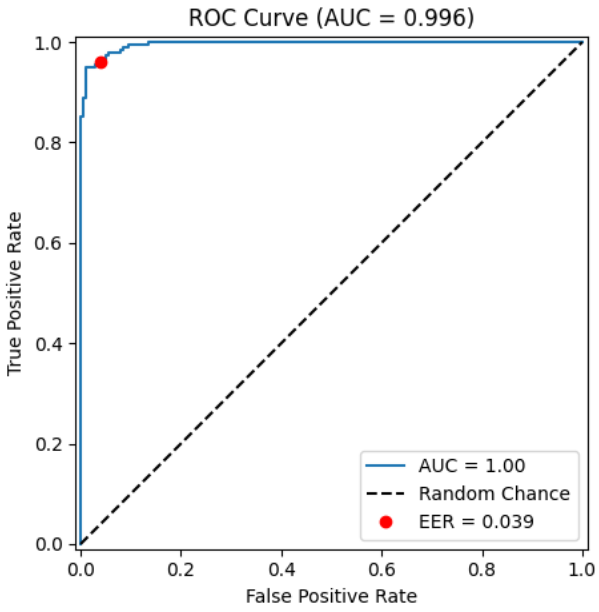

Figure 13: EER presentation for SVM

2. One class-svc able to solve KD problem with an EER of 0.0900 at threshold of 0.1094 presented in Figure 14 .

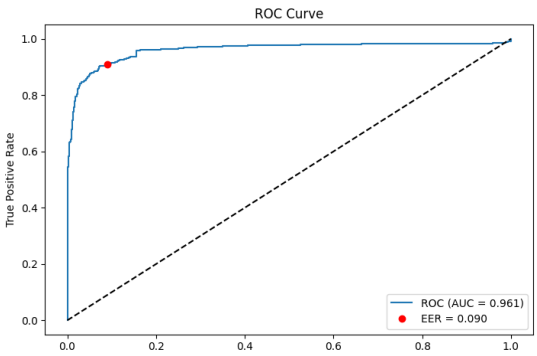

Figure 14: EER presentation for OC-SVM

3. Random-Forest is able to solve KD problem with an EER of 0.0300 at threshold ≈ 0.5333 presented in Figure 15.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

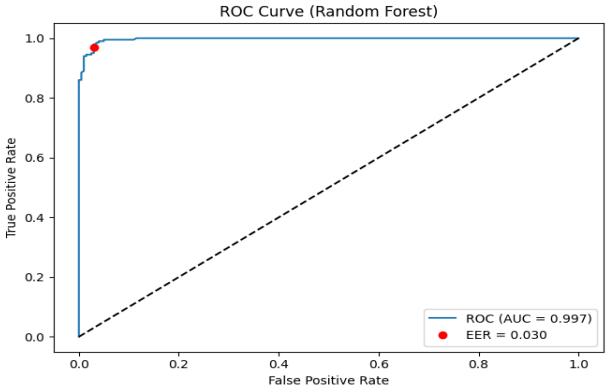

Figure 15: EER presentation for RF

4. EER presentation for Multilayer Perceptron in Figure 16 with EER of 0.0200 at threshold ≈ 0.4897 is presented.

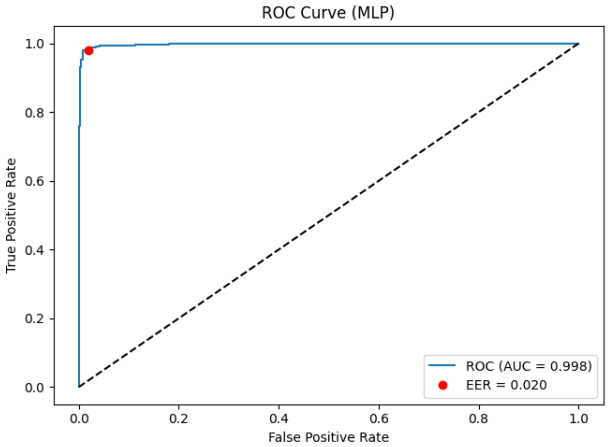

Figure 16: EER presentation for MLP

5. Fusion of MLP and RF is able to solve the KD authentication problem with an EER of 0.0071 at threshold ≈ 0.5973 is presented in Figure 17.

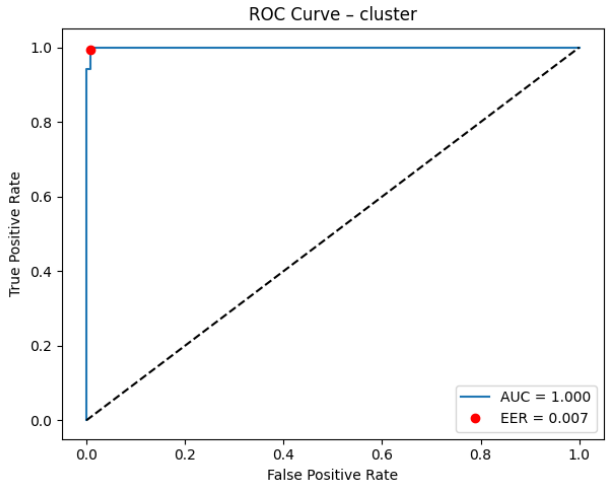

Figure 17: EER presentation for fusion of MLP and RF

6. Quantum Support Vector Machine able to solve with an EER of 0.0000 as shown in Figure 18.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

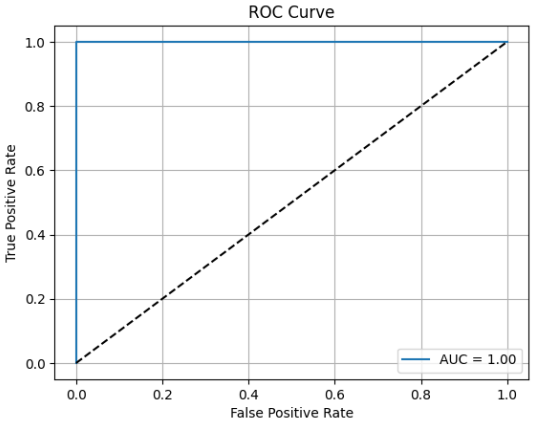

Figure 18: EER presentation for QSVC

7. a. VQC with Real Amplitude ansatz of EER of 35% with an AUC is presented in Figure 19.

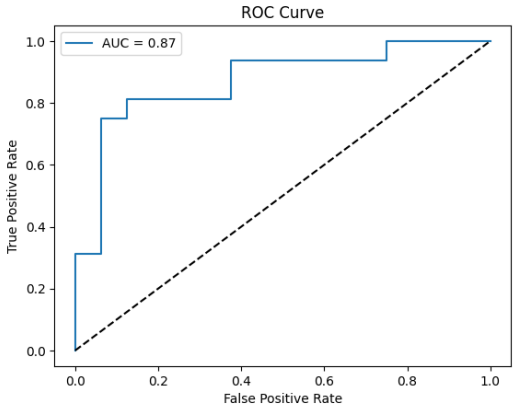

Figure 19: EER presentation for VQC(real amplitude)

7 .b. VQC with EfficientSU2 ansatz of EER of 0.00% with an AUC is presented in Figure 20.

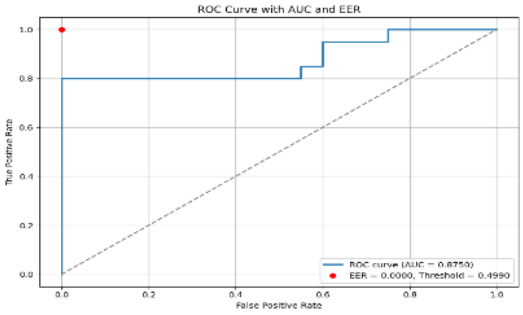

Figure 20: EER presentation for VQC(EfficientSU2)

8. Quantum Neural Network(QNN)

QNN's EER of 30% with an AUC is presented in Figure 21.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

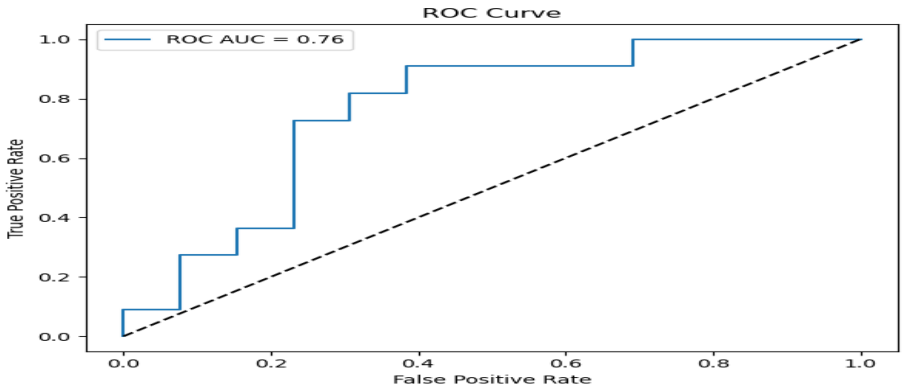

Figure 21: EER presentation for QNN

9. Sampler QNN

Sampler QNN's EER of 31% with an AUC is presented in Figure 22.

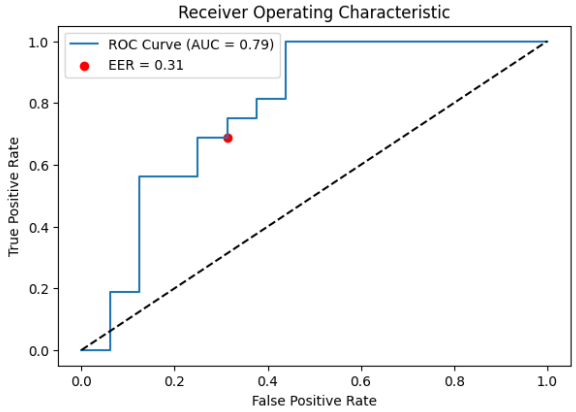

Figure 22: EER presentation for Sampler QNN

7 CONCLUSION

Various quantum and, classical algorithms have been applied. Although machine learning algorithms are efficient in finding various patterns, some special properties of quantum computing, such as entanglement, superposition, and interference, render them capable of finding patterns that classical algorithms cannot find. The overall performance of QSVM among classical and quantum algorithms is better than any algorithm as can be checked with above results. The performance given by fusion on MLP and Rf is best among the applied classical algorithms.

8 FUTURE WORK

Quantum computing is in the future, and there is a need for further exploration of quantum algorithms for keystroke dynamics. The following issues can result in the low accuracy of quantum algorithms.

1.Quantum circuits being sensitive to environment can lead to loss of quantum information and introduce errors known as quantum decoherence.

2. Current quantum computers have limited qubit counts and are prone to errors.

3. During measurement quantum states collapse to classical state which limits the ability to perform backpropagation functions.

4. Quantum circuits are highly sensitive to parameter changes which can lead to challenges in optimisation like barren plateaus, where gradients vanish.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

These limitations must be minimized to improve quantum algorithms. A quantum computer with photons can solve complex mathematical problems more efficiently, so we will try to implement Keyboard Dynamics with them in the future.

"Statements and Declarations"

Conflict of Interest

The authors declare that they have no conflict of interest related to the content of this manuscript.

REFERENCES

1. Sharma A, Jureček M, Stamp M. Keystroke Dynamics for User Identification. arXiv preprint arXiv:2307.05529. 2023 Jul 7
2. Tsimperidis I, Asvesta OD, Vrochidou E, Papakostas GA. IKDD: A Keystroke Dynamics Dataset for User Classification. Information. 2024 Aug 23;15(9):511.
3. Shadman R, Wahab AA, Manno M, Lukaszewski M, Hou D, Hussain F. Keystroke dynamics: Concepts, techniques, and applications. arXiv preprint arXiv:2303.04605. 2023 Mar 8.
4. Yang Z, Zolanvari M, Jain R. A survey of important issues in quantum computing and communications. IEEE Communications Surveys & Tutorials. 2023 Mar 8;25(2):1059-94.
5. Di Meglio A, Jansen K, Tavernelli I, Alexandrou C, Arunachalam S, Bauer CW, Borras K, Carrazza S, Crippa A, Croft V, De Putter R. Quantum computing for high-energy physics: state of the art and challenges. PRX Quantum. 2024 Aug 1;5(3):037001.
6. Kim Y, Eddins A, Anand S, Wei KX, Van Den Berg E, Rosenblatt S, Nayfeh H, Wu Y, Zaletel M, Temme K, Kandala A. Evidence for the utility of quantum computing before fault tolerance. Nature. 2023 Jun 15;618(7965):500-5.
7. Nofer M, Bauer K, Hinz O, van der Aalst W, Weinhardt C. Quantum Computing. Business & Information Systems Engineering. 2023 Aug;65(4):361-7.
8. Singh S, Arya KV. Key classification: a new approach in free text keystroke authentication system. In2011 Third Pacific-Asia Conference on Circuits, Communications and System (PACCS) 2011 Jul 17 (pp. 1-5). IEEE.
9. Ahmed AA, Traore I. Biometric recognition based on free-text keystroke dynamics. IEEE transactions on cybernetics. 2013 May 13;44(4):458-72.
10. Ayotte B, Banavar M, Hou D, Schuckers S. Fast free-text authentication via instance-based keystroke dynamics. IEEE Transactions on Biometrics, Behavior, and Identity Science. 2020 Jun 24;2(4):377-87.
11. Roth J, Liu X, Metaxas D. On continuous user authentication via typing behavior. IEEE Transactions on Image Processing. 2014 Aug 15;23(10):4611-24.
12. Sitová Z, Šeděnka J, Yang Q, Peng G, Zhou G, Gasti P, Balagani KS. HMOG: New behavioral biometric features for continuous authentication of smartphone users. IEEE Transactions on Information Forensics and Security. 2015 Dec 8;11(5):877-92.
13. Ur Rasool R, Ahmad HF, Rafique W, Qayyum A, Qadir J, Anwar Z. Quantum computing for healthcare: A review. Future Internet. 2023 Feb 27;15(3):94.
14. Banchi L, Fingerhuth M, Babej T, Ing C, Arrazola JM. Molecular docking with Gaussian boson sampling. Science advances. 2020 Jun 5;6(23):eaax1950.
15. Kösoglu-Kind B, Loredo R, Grossi M, Bernecker C, Burks JM, Buchkremer R. A biological sequence comparison algorithm using quantum computers. Scientific Reports. 2023 Sep 4;13(1):14552.
16. Innan N, Khan MA, Bennai M. Financial fraud detection: a comparative study of quantum machine learning models. International Journal of Quantum Information. 2024 Mar 16;22(02):2350044.
17. Banerjee R, Feng S, Kang JS, Choi Y. Keystroke patterns as prosody in digital writings: A case study with deceptive reviews and essays. InProceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) 2014 Oct (pp. 1469-1473).
18. Xu L, Skoularidou M, Cuesta-Infante A, Veeramachaneni K. Modeling tabular data using conditional gan. Advances in neural information processing systems. 2019;32.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

19. Vishwanathan SV, Murty MN. SSVM: a simple SVM algorithm. InProceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290) 2002 May 12 (Vol. 3, pp. 2393-2398). IEEE.
20. Park JE, Quanz B, Wood S, Higgins H, Harishankar R. Practical application improvement to Quantum SVM: theory to practice. arXiv preprint arXiv:2012.07725. 2020 Dec 14.
21. Li KL, Huang HK, Tian SF, Xu W. Improving one-class SVM for anomaly detection. InProceedings of the 2003 international conference on machine learning and cybernetics (IEEE Cat. No. 03EX693) 2003 Nov 5 (Vol. 5, pp. 3077-3081). IEEE.
22. Rigatti SJ. Random forest. Journal of Insurance Medicine. 2017 Jan 1;47(1):31-9.
23. Maheshwari D, Sierra-Sosa D, Garcia-Zapirain B. Variational quantum classifier for binary classification: Real vs synthetic dataset. IEEE access. 2021 Dec 30;10:3705-15.
24. Taud H, Mas JF. Multilayer perceptron (MLP). Geomatic approaches for modeling land change scenarios. 2018:451-5.
25. Tripathi SM, Upadhyay H, Soni J. Quantum Neural Network Classification-Based Cyber Threat Detection in Virtual Environment. In2023 International Conference on Computational Science and Computational Intelligence (CSCI) 2023 Dec 13 (pp. 391-396). IEEE.
26. Roth TM. Achieving Quantum Speed and Scale in Full Stack Quantum Computing Designs Using Quantum Machine Learning (Doctoral dissertation, National University).