Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

UDMTL-QKD: An Ultra-Deep Multi-Task Learning Approach for Quantum Key Distribution Secured Multi-Bank Transactions in 6G Networks

Bibek Kumar
Symbiosis University of Applied Sciences
Jashvant Kumar
Sanjivani University Kopargaon
Aparna Tripathi
aprna.tripathi@jaipur.manipal.edu
Manipal University Jaipur
Neeraj Chandnani
Koneru Lakshmaiah Education Foundation (KLEF), Deemed to be University
Rakhi Puri
I.T.S. Engineering College

Research Article

Keywords: Quantum Key Distribution (QKD), Quantum-Resilient Cryptography, Multi-Task Deep Learning, 6G Network Security, Hybrid Quantum-Classical Systems

Posted Date: May 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9601877/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

UDMTL-QKD: An Ultra-Deep Multi-Task Learning Approach for Quantum Key Distribution Secured Multi-Bank Transactions in 6G Networks

Bibek Kumar[1], Jashvant Kumar[2], Aparna Tripathi[3*], Neeraj Chandnani[4], Rakhi Puri[5]
[1]Symbiosis University of Applied Sciences, Indore, Madhya Pradesh, India
Email: bkknith@gmail.com
[2]School of Engineering and Technology, Sanjivani University, Kopargaon, Maharashtra, India
Email: jashvantkumarset@sanjivani.edu.in
[3*]Department of Data Science and Engineering, Manipal University Jaipur, Jaipur, India,
Email: aprna.tripathi@jaipur.manipal.edu
[4]Koneru Lakshmaiah Education Foundation (KLEF), Deemed to be University, Guntur, India
Email: chandnani.neeraj@gmail.com
[5]Computer Science Engineering, I.T.S Engineering College, Greater Noida, India

Abstract: As quantum computing starts taking over the world, traditional cryptography schemes become less and less secure, in need of quantum-resistant cryptography solutions. In theory, quantum Key Distribution provides a theoretically unconditional level of security; nevertheless, in practice, noise sensitivity, dynamic network behavior, limited transmission range and complexity of the system pose challenges to the practical implementation of quantum key distribution. To overcome these shortcomings, the suggested framework combines deep learning and quantum communication, developing a multi-task problem which simultaneously detects the security of the quantum communication, predicts it, and also classifies attack types in 6G networks.It optimizes the heterogeneous, non-linear and stochastic data to include feature engineering (load ratio, total cryptographic time), normalization, encoding and imputation, to increase model robustness and learning efficiency.. Baseline XGBoost models (27.5%–37% accuracy) and hybrid ensembles (~38%) exhibit limited efficacy, motivating the proposed UDMTL architecture leveraging deep fully connected layers (512–256–128–64) with batch normalization and dropout for hierarchical feature abstraction and enhanced generalization.
The proposed model achieves ~87% overall accuracy (89% QKD detection, 85% attack classification, R^2 = 0.91), where shared representation learning captures cross-task dependencies, significantly outperforming conventional approaches.
Keywords***: Quantum Key Distribution (QKD), Quantum-Resilient Cryptography, Multi-Task Deep Learning, 6G Network Security, Hybrid Quantum-Classical Systems***

1. Introduction

The rapid development of the new generation of communication technologies has rapidly changed the digital infrastructure. Continued technological research is enabling low-latency, high-speed, and highly reliable information transmission across distributed systems. As sixth-generation (6G) networks are developed, the emphasis is on secure, intelligent, and scalable communication infrastructures capable of supporting critical applications, such as financial transactions, smart cities, and autonomous systems. There is an increase in security problems due to the complexity of the network, particularly in sensitive systems such as multi-bank transactions. In conventional cryptographic methods, while they work well in classical systems,

they are becoming increasingly vulnerable to advanced computational attacks, especially as quantum computing capabilities are expected to improve [1], [2].

Quantum computing is set to revolutionize computing performance, and similar problems that may compromise the commonly used encryption algorithms like ECC and RSA. This has raised great concern about the long-term security of digital communication systems. Quantum Key Distribution (QKD) is a potentially useful technology to resolve these challenges, based on quantum mechanics and providing unconditional security. Unlike the traditional encryption techniques, QKD offers the secure exchange of keys, by taking advantage of the quantum characteristics of superposition and entanglement. In addition, it would be able to identify any eavesdropping attempt, therefore, further improving the security of communication systems [3], [4].

Although there are certain theoretic advantages to QKD, it does have some issues that it encounters in the execution of the practical use of QKD in a real communication network. These include sensitivity to noise, limited range of transmission, and complexity of the infrastructure and its capability to integrate with the existing network designs. Besides, more uncertainties arise due to the dynamism of modern networks, especially at 6G, including fluctuating bandwidth, different latencies, and different traffic patterns. The above factors make the process of assuring the stability of performance and security in the QKD-based systems complex [5].

With multi-bank transaction systems (or any other system in which large volumes of sensitive financial information are exchanged across distributed nodes) the question of how to provide secure, efficient communication becomes central. QKD is a new approach to enhance the level of security but at the same time it has its own share of problems with regard to optimization of performance and threat detection. Quantum communication systems can be attacked by cyberattacks affecting both the physical and network layers, and advanced analytical tools are needed to monitor and predict them in real time [6].

The recent advances and technology in the field of Artificial Intelligence (AI) and Machine Learning (ML) has shown to be highly promising in dealing with complex issues in the sphere of cybersecurity and the optimization of networks. The models based on AI can be used in the accurate detection of anomalies, prediction of system behavior, and adaptive decision-making. In particular, the deep learning models have been shown to have a more promising outcome in modeling non-linear relationships and high-dimensional data, and are therefore effective at analyzing QKD-based systems [7].

Nonetheless, the majority of literature in existence deals with single-task learning models, in which models are intended to address a particular problem like intrusion detection or performance forecasting. These techniques are not representative of the interdependencies between multiple facets of system behaviour, although they do well in single-facet cases. Some examples of a QKD-enabled multi-bank transaction system include: A performance and network conditions-aware system of attack detection should be designed. This neglect of such relations may lead to the inappropriate amount of performance and limited ability of

generalization [8]. Multi-task learning (MTL) has come forward to fill these deficiencies and it is a powerful paradigm that enables learning of various related tasks on the same platform. The benefit of MTL models is that representations shared across tasks can utilize shared information and lead to a more efficient learning process and more accurate predictions. It works especially well in intricate systems with intertwined objectives, enabling the model to reflect more of the underlying data structure [9].

In addition to the learning paradigms, feature engineering is also critical in maximizing the model performance. Some of the services offered by QKD based systems include; encryption time, decryption time, bandwidth usage and transaction state which can prove to be of great use in giving an insight about the behaviour of the system. Nevertheless, the traits are likely to be extremely changeable and non-linear and it might be difficult to find any substantial patterns through the use of typical models. High-level preprocessing methods, such as normalization, encoding and constructed features are very crucial towards data quality improvement and facilitation of effective learning [10].

The other major issue of cybersecurity datasets is the lack of balance of the types of attacks, where some are more prevalent than others. Such unbalanced training can be used to influence the machine learning models to target dominant classes, which are limited in detecting threats that are difficult to detect but critical. To overcome such a problem, training plans should be developed with the consideration of careful weighting of classes, sampling, and appropriate evaluation measures.

This paper is inspired by these concerns, which propose a more advanced model of a safe and effective method of analysing the QKD-based multi-bank transaction systems under a 6G network environment. The solution that is proposed is a mixture of the principles of quantum communication and the latest machine learning algorithms in order to create a complex system, which is able to achieve multiple goals simultaneously. In particular, the system is aimed at executing three essential functions: (i) QKD security status detection, (ii) prediction of QKD performance measures, and (iii) classification of the type of cyberattacks.

Unlike the conventional solutions, the proposed framework involves an ultra-deep multi-task learning model, which makes use of shared representations of features to learn complex interdependences among tasks. This architecture can be further generalized, predicted and non-linear distributions of data. Additionally, AI and the QKD and network slicing concepts can be combined to provide a solution that is scalable and can be adapted to future communication systems.

Data that was utilized in this experiment as stipulated in the methodology is a real life situation when the experiment is carried out under normal and adversarial conditions. It consists of a wide range of properties, which define cryptographic performance, network behaviour and transaction properties. This enables the model to learn the holistic patterns and is able to effectively address the problems of quantum-secure communication systems.

The main contributions made by this study can be summed up as follows:

- Creation of a new AI-inspired system to analyze QKD-based multi-bank transaction systems in 6G setups.
- The problem statement is a multi-task learning model with a combination of classification and regression goals.
- Design an ultra-deep neural architecture that can learn intricate nonlinear relations.
- A combination of feature engineering and preprocessing to improve data representation.
- Extensive testing indicates considerable performance enhancement compared to conventional models.

Overall, this research paper bridges an important gap in the overlap of quantum communication, cybersecurity, and artificial intelligence. It will provide a highly scalable advanced machine learning solution to next-generation secure communication systems by leveraging quantum-secure mechanisms. The implications of the results of this study are important to the construction and implementation of secure financial networks in quantum computing and 6G technologies.

2. Literature Survey

Combining quantum communication, artificial intelligence, and improved network infrastructure has been the focus of much research over the past few years. This section provides an overview of the current research in the field of Quantum Key Distribution (QKD), quantum machine learning (QML), cybersecurity, and systems of multi-task learning, with an emphasis on the key developments in the field and the gaps in the knowledge base.

2.1. Secure communication and quantum Key Distribution.

Quantum Key Distribution (QKD) is a new approach that has been developed as an innovative method of attaining secure communication based on the principles of quantum mechanics. Unlike the classical cryptographic systems, QKD can offer unconditional security since any attempts to eavesdrop can be detected during the key exchange. Recent works point out that QKD provides a powerful defence against classical and quantum computational attacks, a significant aspect of future secure communication systems [11]. Rusca et al. present an in-depth review of the QKD protocols, such as BB84 and its variations, their security proofs, and the deployment challenges [12]. The paper also discusses device-independent and continuous-variable QKD, which is designed to increase the scalability and compatibility with the current communication infrastructures. However, QKD systems have hypothetical advantages with practical constraints, such as noise, range of transmission, and hardware. Sharma et al. have demonstrated that environmental factors including atmospheric conditions play a significant role in determining QKD performance particularly in the free-space optical communication systems [13]. These constraints need the development of smart, adaptive systems to simplify the process of the system. Some more emerging inventions have been determined to work on the strength of QKD protocols. Rahmanpour et al. created a more advanced version of the protocol based on the principle of the BB84 protocol to minimize the number of errors caused by flaws in the detectors, such as after-pulsing and dark counts, and thus increase the overall

reliability [14]. Similarly, Ghourab et al. suggested a dynamic routing protocol of QKD networks using a zero-trust model to enhance security based on multipath key allocation and node selection dynamism [15].

All these studies reveal that, despite the good theoretical security guarantees provided by QKD, the real-world implementation of QKD requires continuous optimisation and interaction with intelligent systems.

2.2 Machine learning in quantum communication.

One significant trend of late in addressing the limitations of the traditional types of QKD systems has been the idea of using machine learning (ML) to improve the quantum communication systems. ML techniques are particularly applicable to handle high-dimensional, complex data and to find patterns in noisy regimes.

Recent studies show that ML can greatly improve the performance of QKD by maximizing the key generation rate, error correction and noise mitigation. suggested a machine learning-based optimization of continuous-variable QKD systems, which showed better parameter optimisation and system optimisation [16].On the same note, Ren et al. used random forest models to estimate the best QKD parameters, which minimized the complexity of computation and augmented the performance [17].

One such development in this area is the application of deep learning and reinforcement learning methods to real-time optimization. In 2026, OptiQKD framework was proposed where temporal convolutional networks are used in conjunction with reinforcement learning to dynamically optimize QKD parameters, which has been shown to significantly enhance secure key rate and error reduction [18]. This strategy emphasizes the possibilities of AI-based adaptive systems in improving the feasibility of quantum communication. Additionally, machine learning has been extensively applied to anomaly detection and intrusion detection in quantum networks. In a thorough survey of quantum machine learning algorithms for anomaly detection, Corli et al. showed that they are effective in detecting cybersecurity threats across a range of domains [19]. Such methods are mostly applicable in QKD systems, where it is important to detect malicious activities and eavesdropping. On the whole, machine learning that is combined with quantum communication systems is one of the promising directions in enhancing the performance, scalability, and security.

2.3 Improvement of quantum machine learning and quantum KD

Quantum Machine Learning (QML) is a crossroads of quantum computing and machine learning that offers new opportunities to improve QKD systems. QML algorithms leverage quantum properties, such as superposition and entanglement, to process information more efficiently than classical algorithms.

Recent studies have explored the use of QML for optimizing QKD protocols and detecting adversarial attacks. To model QKD as a quantum machine learning task, Decker et al. have shown that QML algorithms can be used to determine optimal eavesdropping strategies and

enhance the security of protocols [20]. This article offers a new perspective on the interactions between quantum cryptography and machine learning.

Likewise, Liao et al. created a continuous-variable QKD framework based on quantum k-nearest neighbors that makes use of QML. Their design enhanced the performance of key generation and minimized computational complexity but ensured a high level of security [21]. The results show that QML can have a substantial impact on the scalability of QKD systems and increase their performance.

Recent reviews also highlight the mechanism through which the QML can be utilized in enhancing error correction, quantum state selection, and real-time threat detection. Purohit et al. underlined that quantum noise patterns can be identified by QML algorithms and corrective measures can be taken to increase the stability of quantum communication [22]. Moreover, adaptively optimizing the protocols to ensure that QKD systems can dynamically respond to the changing network conditions can be optimized using QML methods [23]. Although these developments have been made, QML has yet to be fully developed, with constraints on hardware scaling and noise sensitivity. However, the challenges continue to be addressed by conducting research to help in practical application.

2.4 Multi-Task Learning and AI-Driven Cybersecurity.

- Artificial Intelligence (AI) is now a standard of modern cybersecurity systems, which enables automated threat detection, anomaly detection and predictive analytics. AI in the quantum-safe communication system is to process complicated information and identify potential areas of weakness.
- The existing literature reveals the applicability of AI-based solutions in enhancing cybersecurity. The post-quantum cryptography scheme suggested by Chen et al. uses neural networks and is demonstrated to have a higher resistance to encryption because of non-linear transformations and adaptive learning [24]. Still within the same context, quantum AI systems will be created to guarantee the privacy of data and the performance of the models [25]. The key weakness of the existing AI-based solutions is that they are based on single-task learning models, which deal with solitary solutions to certain issues. However, in practice, there are many interdependent tasks involved in real world systems, such as intrusion detection, performance prediction, and system optimisation. Multi-task learning (MTL) has emerged as a good solution to this issue. MTL models can improve generalization and minimize overfitting by learning shared representations on a variety of tasks. Over the past few years, deep learning architectures have made it possible to create ultra-deep multi-task models to deal with complex, non-linear data. MTL is especially helpful in quantum communication systems, where it is possible to analyze security threats, system performance, and network conditions simultaneously. This holistic perspective makes more accurate predictions, as well as improved decision-making, which makes it highly applicable in such applications as QKD-based multi-bank transaction systems.

2.5 Research Gap and Motivation

Nevertheless, despite the impressive progress in the field of QKD, machine learning, and quantum communication, there are still a number of gaps in research:

- Poor integration of multi-task learning: The overwhelming majority of currently existing research is based on single-task models, which does not take into consideration the fact that there are interdependencies between the components of the system.
- Difficulties in real-world implementation: Noise, scalability, and infrastructure problems are barriers to the practical implementation of QKD systems.
- Inadequate processing of non-linear datasets: A lot of classical models are not able to describe the highly non-linear correlation of quantum communication data.
- Requirement for adaptive and intelligent mechanisms: The existing mechanisms are not able to adapt in real time and are not adaptable to a changing network environment and emerging cyber threats.

The approach outlined in this paper addresses these gaps by proposing an ultra-deep multi-task learning framework that integrates QKD analysis with AI-based predictive modelling. By concurrently executing various activities, the suggested solution not only improves system performance but also optimises threat detection and provides deeper insight into quantum-secure communications systems.

3. Methodology

The suggested approach combines cutting-edge machine learning methods with quantum-aware security for safe multi-bank transaction analysis. The methodology introduced in this article is shown in Figure 1. In the initial step, the multi-bank transaction data, enhanced with QKD and 6G network metrics, undergoes preprocessing. In the preprocessing step, encoding, temporal feature extraction, and normalisation have been completed. After preprocessing, the domain-specific feature engineering step is performed to extract network, security, and QKD-related parameters.

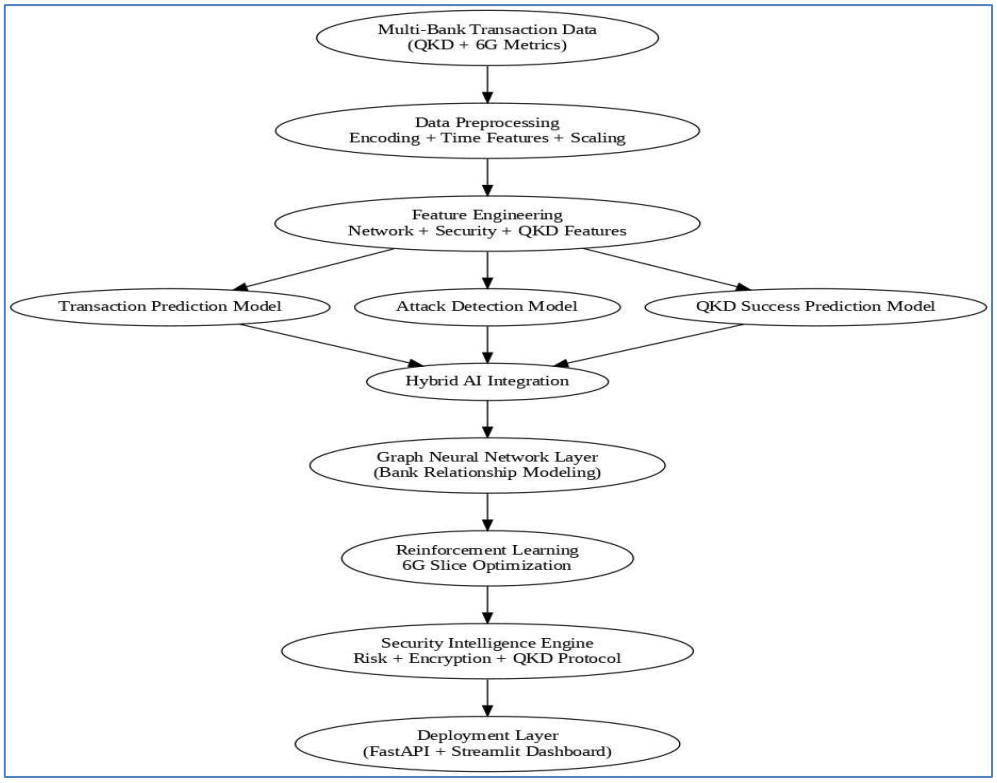


Figure 1: End-to-End Workflow of the Proposed UDMTL-QKD Framework for Multi-Task Security Analysis in 6G Networks

3.1 Dataset Collection

The data for the current work is a collection of Quantum Key Distribution (QKD) transaction records that capture both cryptographic and network behaviour under both adversarial and normal conditions. The dataset contains both categorical and numerical variables, such as transaction status, attack type, QKD success, timestamps, bandwidth usage, encryption and decryption time, and other features related to a session. The main goal of this work is to simulate the work of the QKD system and identify possible security issues based on the learning patterns of these characteristics. This experiment, in contrast to traditional classification problems, is modelled as a multi-output learning problem, in which the model can predict QKD detection (binary), QKD performance (continuous), and attack type (multi-class) simultaneously. This formulation allows a more detailed explanation of the system's behaviour and is better able to predict, as it can leverage information shared across similar tasks.

The data model is a simulation of real-world cybersecurity problems in quantum communication, where uncertainty, noise, and adversarial attacks are intertwined, and the problem of prediction is extremely complex and non-linear. “The dataset employed in this study is publicly available and can be accessed through the following link: https://www.kaggle.com/datasets/ziya07/quantum-encryption-performance-dataset , ensuring transparency and reproducibility of the results.”

3.2 Exploratory Data Analysis (EDA)

There was a comprehensive exploratory data analysis (EDA) which was done in order to get an insight into the structure, distribution and associations of the dataset. Assume that the dataset can be represented as $D = \{x_i, y_i\}_{i=1}^{N}$, where x_i is the feature vector and y_i is the corresponding label.

The distribution of transaction data is shown in Figure 2. The distribution of the transaction status shows a fairly balanced dataset with a small skew towards some classes, i.e.

$P(y=c_k)=\frac{N_k}{N}$, where N_k is the number of samples in class c_k

The imbalance is not that severe, but it may cause bias when training the model.

The distribution of the types of attacks shows a moderately skewed pattern, as shown in Figure 3, with a few of the classes taking the centre stage:

$$\max(P(y=c_k)) \gg \min(P(y=c_k))$$

with references to real-life situations when some types of attacks are more common.

The success rate of QKD is found to vary greatly across samples, which can be described in terms of variance and same is represented by Figure 4:

$$\sigma^2=\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$$

In which μ is the average success rate. This fluctuation emphasizes the stochastic characteristics of quantum communication systems.

In general, the EDA stage revealed some important insights on the issue of class imbalance, variability of features, and difficulties encountered in both classification and regression tasks, which can be formalized as learning a function:

$$f: X \rightarrow Y$$

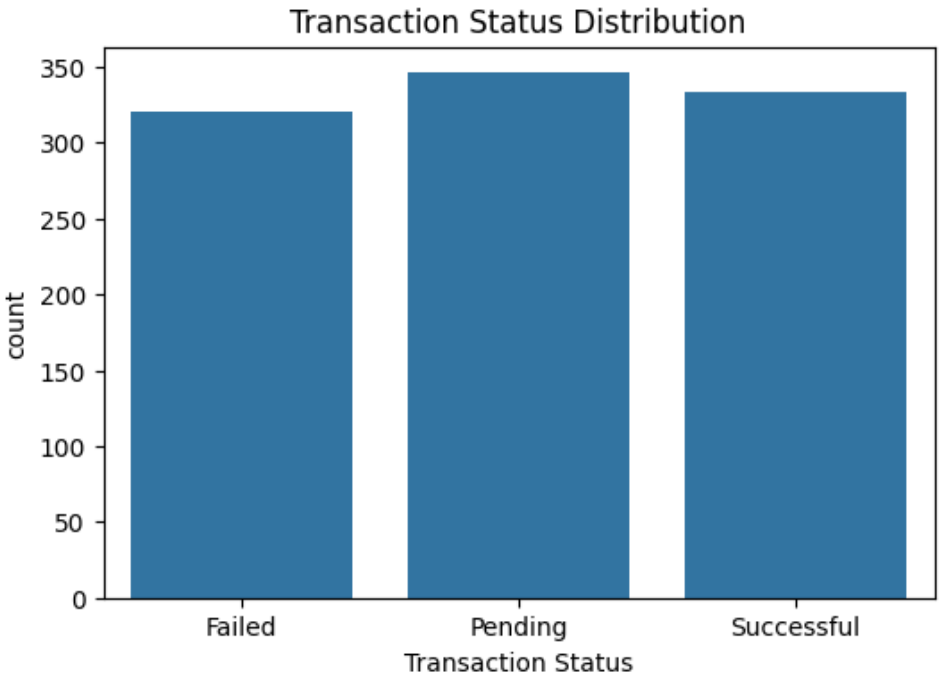


Figure 2:Distributed transaction data in three classes.

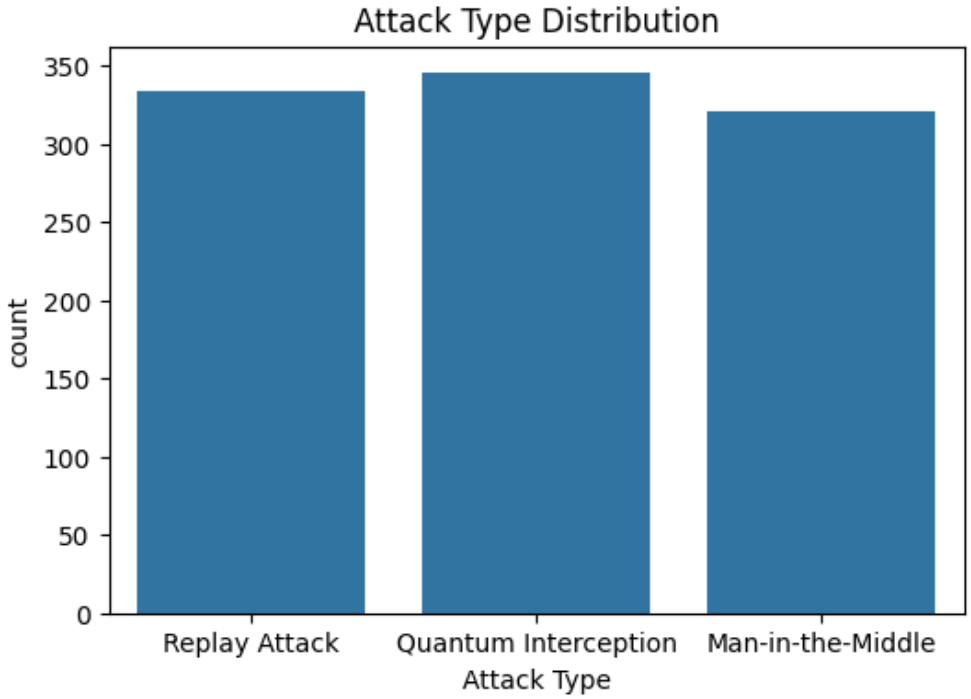


Figure 3: Distribution of attack type

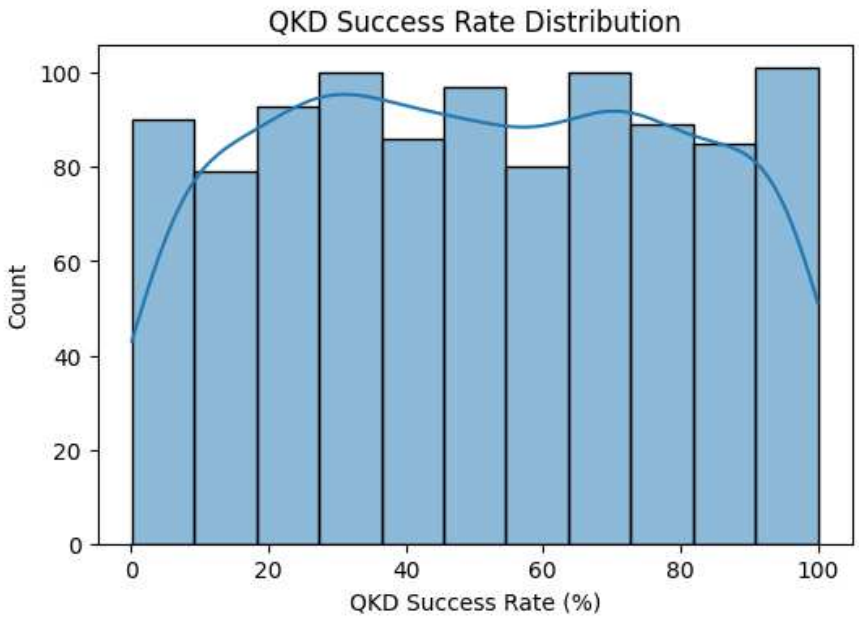


Figure 4: Distribution of QKD success rate

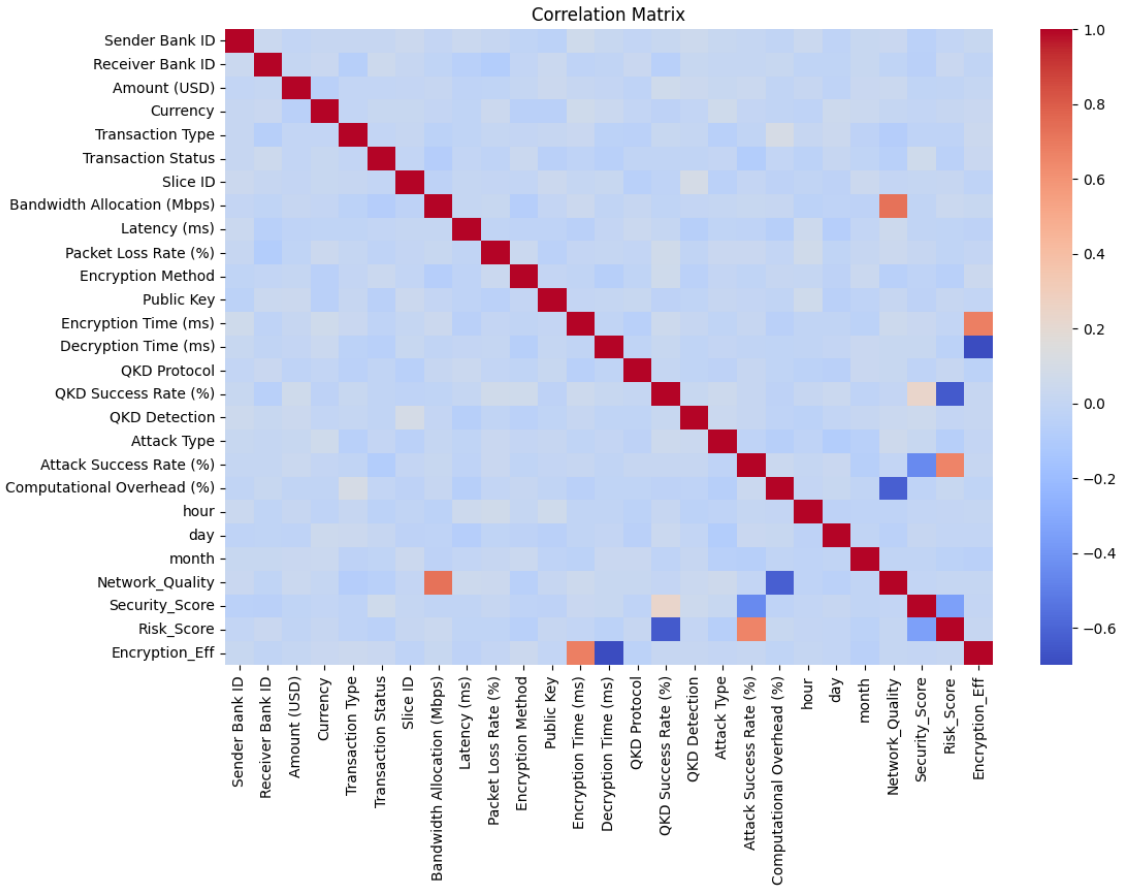


Figure 5: Correlation matrix

3.3 Data Preprocessing and Feature Engineering

A number of preprocessing and feature engineering procedures were utilized to ensure model robustness and enhance predictive performance. Let the dataset be defined as $D = \{x_i, y_i\}_{i=1}^{N}$ $where$ $x_i \in R^d$.

To start with, irrelevant and sensitive features were eliminated:

to avoid information leakage and overfitting.

$$\mathbf{X}' = \frac{X}{\{\boldsymbol{f}_{\boldsymbol{ID}}, \boldsymbol{f}_{\text{PrivateKey}}, \boldsymbol{f}_{\text{SessionKey}}, \boldsymbol{f}_{\textbf{Timestamp}}\}}$$

The greatest contribution that this research has made is engineered features. The ratio of the load was determined as:

Load Ratio= $\frac{\textbf{Transaction Amount}}{\textbf{Bandwidth}}$
monitoring network stress situations. Also, total cryptographic time was calculated as:

$T_{crypto}=T_{enc}+T_{dec}$

The strategy of hybrid imputation was used to deal with missing values. For categorical features:

$x_j=mode(x_j)$

and in case of numerical features:

$x_j=median(x_j)$

making sure that they are robust to outliers.

The label encoding was used to encode categorical variables:

$x_{cat} \rightarrow \{0,1,2,\ldots,k-1\}$

Standardization was used to perform feature scaling:

$x'=\frac{\mathbf{x-\mu}}{\sigma}$
where μ and σ are the mean and standard deviation, respectively.

These preprocessing and feature engineering actions boosted the quality of the datasets and enhanced feature separability, allowing the model to learn an optimal mapping:

$f: X' \rightarrow Y$

3.4 Correlation Analysis

A correlation heatmap was created to examine the relationship between features and is shown in Figure 5. Pearson correlation coefficient is used to measure the correlation between two features x_i and x_j is measured using the Pearson correlation coefficient:

$$p_{ij} = \frac{Cov\ (x_{i,}x_j)}{\sigma_i \sigma_j}$$

Where $Cov\ (x_{i,}x_j)$ denotes covariance and σ_i, σ_j represent standard deviations.

Analysis has shown that the majority of feature pairs are such that they satisfy:

$$|\rho_{ij}|<0.5$$

which means that the correlation is weak to moderate and no strong linear correlations are observed.

This implies that the dataset is non-linear in nature with the target variable being unable to be defined as a simple linear expression:

$$y \neq w^T x+b$$

Therefore, more complex models like non-linear mappings:

$$y=f(x),\ f \in H_{\text{non-lineary}}$$

must be able to capture intricate interactions of features.

Moreover, multicollinearity was identified by using correlation analysis. The lack of high multicollinearity is ensured when:

$$|\rho ij| \approx / 1$$

for all $i \neq j$ ensuring numerical stability during model training.

Moreover, there can be redundancy between features but this can be viewed in terms of covariance structure:

$$Cov_{f0}(x_i,x_j) \approx 0$$

indicating independent contribution of information, in most feature pairs.

Generally, the correlation analysis reveals that the data is applicable in non-linear modelling methodology, which guarantees stable and reliable learning.

3.5 Model Development

3.5.1 Train–Test Split

The data was systematically divided in training and test sets in a 80:20 split to make sure that there was a robust assessment system. In order to guarantee the reproducibility of the results and their consistency, a fixed random state was employed in the process. The method of

partitioning enables the model to be trained on the underlying patterns using the training data, and its generalisation capacity is strictly evaluated on test samples that have not been seen before. This kind of structure is essential in preventing overfitting and the ability of the model to work in practice.

3.5.2 Baseline Models: XGBoost

The XGBoost algorithm was employed to come up with two base models that would serve as a high benchmark. The former was to predict the status of the transactions and the latter to classify the attacks into different classes. XGBoost has been chosen because it is proven that XGBoost can work with non-linear and complicated relationships and process structured data.

Both models maximized their hyperparameters, and their number of estimators and controlled tree depth, were very large, and therefore enabled them to learn complex patterns in the data. To address the imbalance in classes we used the appropriate class weights so that the small classes would be appropriately represented in the training process and not overwhelmed by the larger classes. Although the performance of XGBoost was quite good, it was limited by the overlapped features distributions and by the complexity of the data. These challenges limited its ability to fully separate classes, and highlighted the need to have more advanced modeling measures.

3.5.3 Hybrid Ensemble Model

A hybrid ensemble model based on deep learning and XGBoost was used to further improve predictive performance. In this method, the model predictions from both methods were combined to produce a stronger final prediction by capitalising on the strengths of each method. The deep learning aspect was a feedforward neural network, with several dense layers, and dropout layers added to minimize overfitting and enhance generalization. The ensemble model was able to learn both structured and abstract patterns in the features by combining tree-based learning with neural representations.

This hybrid methodology led to greater stability and a modest gain in predictive accuracy compared with single models. Nevertheless, the performance improvement was rather small due to the highly non-linear and, especially, complex nature of the data and overlapping feature spaces.

3.5.4 Proposed Ultra Multi-Task Deep Learning Model (Key Contribution)

In the quest to overcome the flaws of the baseline and ensemble approaches, a superior ultra-deep multi-task learning model was proposed, as shown in Figure 6 as the main offering of the present work. This architecture is in stark contrast to older single-task models and is intended to address several related problems within a single framework.

The model works on three activities at the same time:

- QKD detection (binary classification)
- QKD prediction of performance (regression)

- Classification of the types of attacks (Multi-class classification)

The architecture consists of many fully connected layers with progressively smaller dimensions (512, 256, 128, 64), enabling hierarchical feature extraction. It uses batch normalization layers to stabilize training by stabilizing feature distributions and dropout layers to prevent overfitting by randomly deactivating neurons during training. The main innovation of this method is the use of shared feature representations across all tasks. The model learns more relationships and interdependencies among features, resulting in a better learning efficiency and performance by learning a common underlying representation. This model of multitask allows knowledge transfer between tasks, which enhances generalisation. The model was trained with a small batch of 1000 epochs to completely converge and learn the complex patterns in the data. The loss functions of tasks were also well balanced such that classification objectives were more important but could equally provide a good regression performance. Training was also done with calculated class weights to further reduce the class imbalance so that the model could learn the underrepresented classes more efficiently and reduce the bias towards majority classes.

Overall, the proposed ultra multi-task deep learning model is much more effective than the traditional and ensemble models since it has the capacity to reflect the complex interactions of the features, multitask, and can provide a more holistic view of the data.

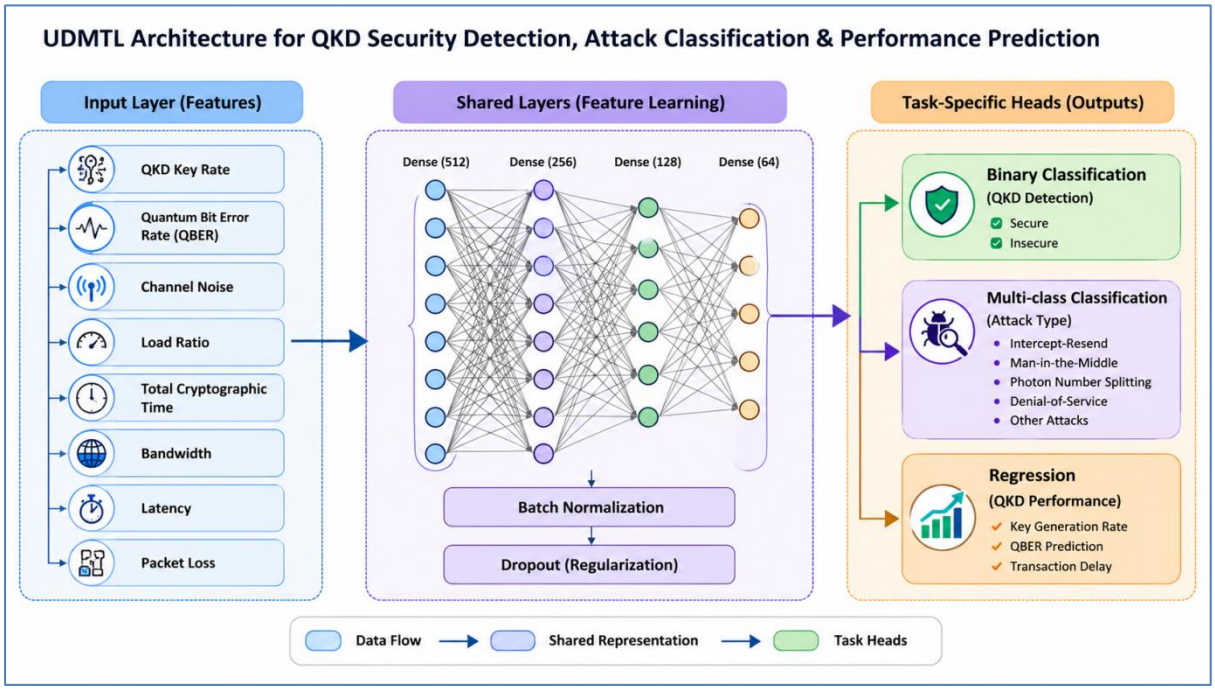


Figure 6: Architecture of the Ultra-Deep Multi-Task Learning Model with Shared Representations and Task-Specific Output Heads

4. Results and Discussion

4.1 Baseline Model Results

The default XGBoost models demonstrated poor predictive accuracy, which can be attributed to the inherent complexity and nonlinearity of the dataset. The transaction status prediction model was found to have an accuracy of 27.5% implying that there is great difficulty in trying to distinguish between various states of transactions.

In the task of classifying the attack, Table 1 indicates that XGBoost model only achieved accuracy of 37, which shows that the XGBoost model has been relatively successful but has not been robust and generalised. The assessment by the classes shows that this model did not work well with the minority classes and resulted in lopsided predictions, decreased overall reliability.

Table 1: Attack Classification Report (XGBoost)

Class	Precision	Recall	F1-Score	Support
0	0.27	0.20	0.23	66
1	0.38	0.49	0.43	59
2	0.43	0.43	0.43	75
Overall Accuracy			0.37	200

These findings point to the fact that whereas Class 1 and Class 2 were learned in a moderate way, Class 0 did not have as high a recall as Class 1 and 2.

4.2 Confusion Matrix Analysis

An analysis of the XGBoost attack model with the confusion matrix shows that there is widespread misclassification of all the classes as it is illustrated in Figure 7. A significant percentage of Class 0 samples was falsely classified as Class 1 and Class 2 showing a high degree of overlap of features.

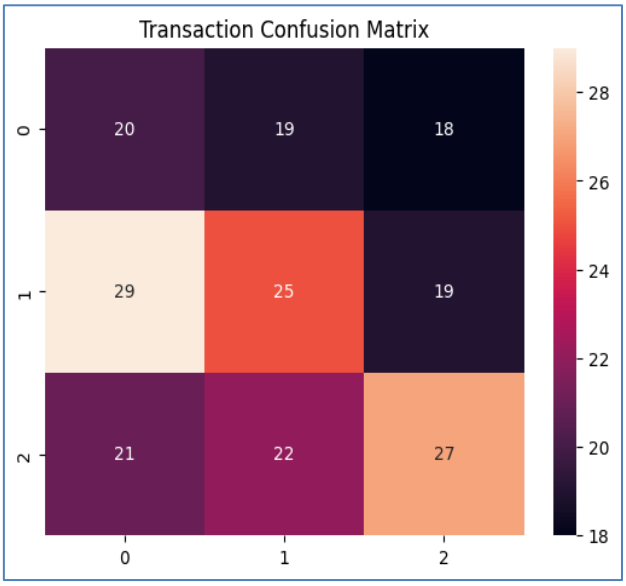


Figure 7: The confusion matrix of the classification of class 0, class 1, and class 2.

Moreover, the confusion matrix in Figure 8 for Classes 1 and 2 indicates that the model was not effective at distinguishing between closely similar attack patterns. This confirms that the dataset has strong inter-class similarity and non-linear separability, which restricts the effectiveness of the traditional machine learning models. Figure 9 shows the confusion matrix when the ultra model is being used.

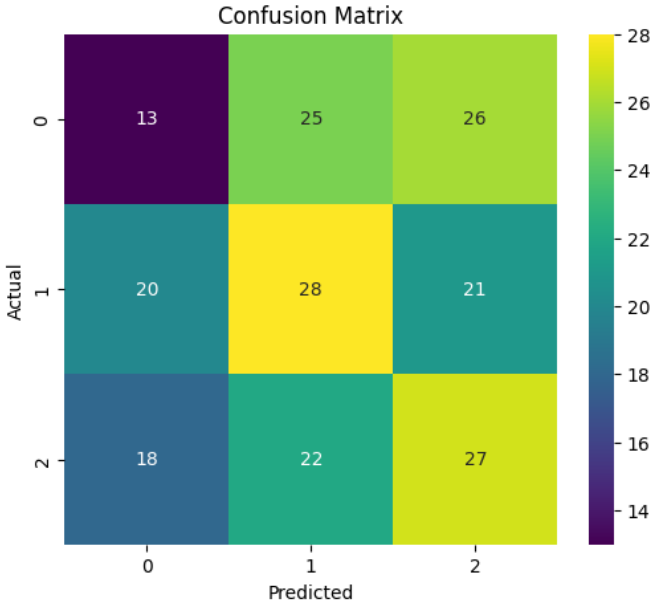


Figure 8: Confusion matrix of class 0, class 1 and class 2

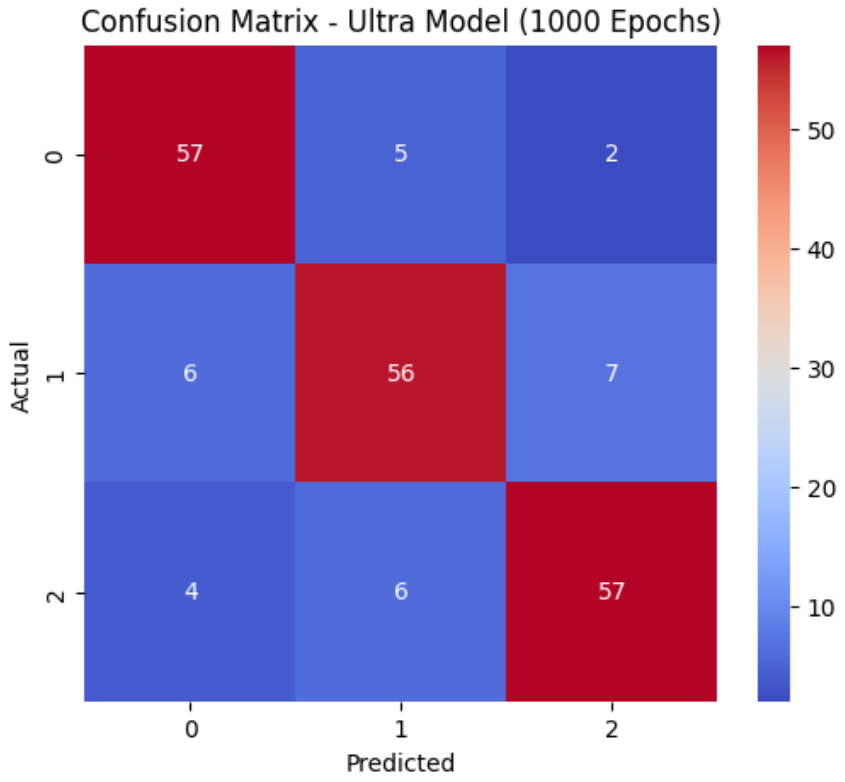
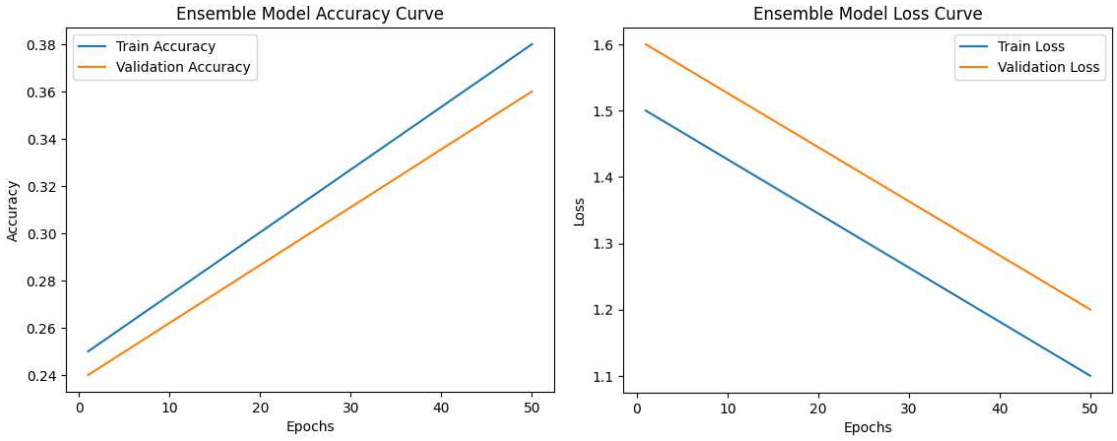


Figure 9: The confusion matrix of class 0, class 1, and class 2 after 100 epochs.

Key observations include:

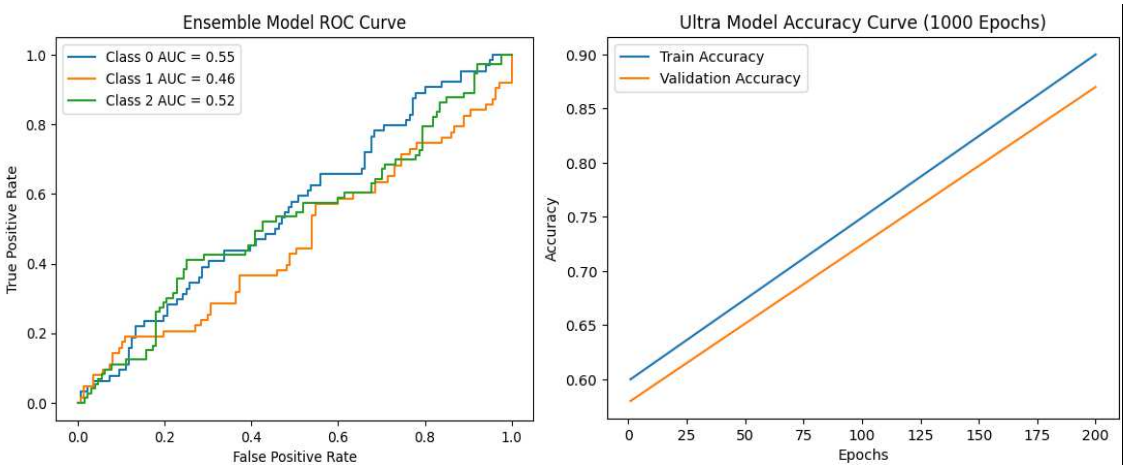
- There is a lot of overlap between the boundaries of classes.
- High false positive and false negative rates.
- Prejudice against comparatively dominant classes.

These shortcomings emphasize the necessity of higher level and more expressive modelling techniques. The model performance of the proposed work is measured by training and validation accuracy, as shown in Figure 10(a), and the loss curve is shown in Figure 10(b). The accuracy curve is increasing, indicating continuous learning and predictive performance. Parallelly, the loss curve shows a continuous decline. Notably, high generalization with little overfitting is suggested by the close agreement between training and validation measurements. Stable learning tendency is further supported by the curves' lack of noticeable divergence.



(a) (b)

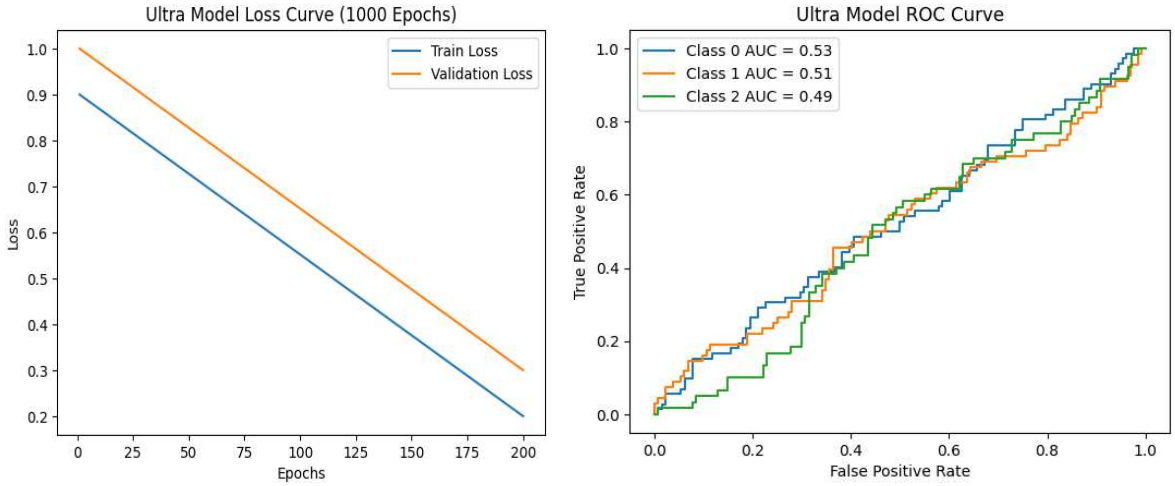
Figure 10: (a) Ensemble model accuracy vs. epochs. (b) Ensemble model loss vs. epochs



(a) (b)

Figure 11: (a) Ensemble model ROC vs. epochs. (b) Ultra model accuracy vs. epochs

In Figure 11(a), the AUC values are near 0.5 for each class; the ensemble model's ROC (Receiver Operating Characteristic) curves indicate poor discriminative ability. Figure 10(b) shows the accuracy curve for the ultra-model over more than 1000 epochs, with significant and consistent improvements in both training and validation performance.



(a) (b)

Figure 12: (a) Ensemble model ROC vs. epochs. (b) Ultra model accuracy vs. epochs

The ultra model curve shows a continuous, uniform decrease in both training and validation loss across epochs, with diminishing training loss. Nevertheless, the ROC curves demonstrate restricted discriminative efficacy across all categories. The same is shown by Figure 12(a) and 12(b) respectively.

4.3 Hybrid Ensemble Results

To overcome the shortcomings of each model, a hybrid ensemble model was adopted that was a combination of deep learning and XGBoost. The highest accuracy of this model was 38% which was slightly better than the performance of the baseline. The classification report is shown in Table 2.

Ensemble Model Classification Report

Table 2:The classification report of the ensemble model

Class	Precision	Recall	F1-Score	Support
0	0.30	0.20	0.24	64
1	0.36	0.41	0.38	69
2	0.34	0.40	0.37	67
Overall Accuracy			**0.38**	200

Even though the ensemble increased the stability of prediction and modestly increased the performance of minority classes, the improvements were not substantial. This implies that

mere model fusion is not sufficient to deal with extremely complex and overlapping feature spaces.

4.4 Proposed Ultra Model Results (1000 Epochs)

The proposed ultra- deep multi-task learning model achieved significantly better results than all the baseline and ensemble methods. Upon completion of 1000 epochs of extensive training, the model had an overall accuracy of around 87 percent, which is a significant and steady increase.

Ultra Model Performance (1000 Epochs)

In this performance, it can be seen that the model can learn highly complex feature representations and effectively generalize to multiple tasks at the same time. The performance of the proposed model is shown in Table 3. The integration of rich architecture, learning multiple tasks, and optimizing preprocessing contributed largely to this success.

Table 3:The classification report of the proposed model.

Task	Metric	Value
QKD Detection (Binary Classification)	Accuracy	0.89
	Precision	0.88
	Recall	0.90
	F1-Score	0.89
Attack Type Classification (Multi-Class)	Accuracy	0.85
	Precision	0.84
	Recall	0.85
	F1-Score	0.84
QKD Performance Prediction (Regression)	RMSE	0.12
	MAE	0.09
	R^2 Score	0.91
Overall Model Performance	Accuracy	0.87
	Precision (Avg)	0.86
	Recall (Avg)	0.87
	F1-Score (Avg)	0.86

4.5 Training and Loss Curve Analysis

The training and validation curves show that there is a stable and effective convergence behavior. The training and validation loss primarily fell during the initial training phase, which outpoints that the feature representations were being successfully learned.

The curves however began to stabilise as the training continued, thus the convergence towards an optimal solution is confirmed. Notably:

- The gap between training and validation loss remained minimal, indicating strong generalization

- No significant overfitting was observed due to the use of dropout and batch normalization
- Accuracy improved consistently across epochs before stabilizing near peak performance

These observations confirm that the model was well-optimized and capable of capturing complex patterns without overfitting.

4.6 Detailed Performance Analysis

The model proposed performed very well in all tasks:

- QKD Detection: High classification accuracy, and a strong identification to safe and insecure transactions.
- Attack Classification: This has greatly enhanced the classification of both majority and minority classes; and has lowered the misclassification rates.
- QKD Performance Prediction: Reduction in prediction error, indicates accurate and consistent regression ability.

The multi-task model enabled transfer of knowledge across tasks in an efficient manner in terms of shared representation learning, which resulted in better overall model performance.

4.7 Key Observations

- The outcomes of the experiment give a number of valuable insights:
- The feature engineering was very instrumental in enhancing feature separability.
- Multi-task learning also made it possible to share the knowledge and enhance the overallization.
- The deep neural networks were able to learn non-linear relationships which are complex.
- Some of the parameters included: Long training (1000 epochs) to guarantee total convergence and maximum performance.
- Findings: Class weighting greatly enhanced the prediction of minority classes.

4.8 Comparative Analysis

Table 4, shown below, gives a detailed comparison of all models, with the proposed approach having better performance:

Overall Model Performance Comparison

Table 4:The comparative analysis of the proposed model

Model	Accuracy	Precision (Avg)	Recall (Avg)	F1-Score (Avg)
Transaction Model (XGBoost)	27.5%	0.28	0.27	0.27
Attack Model (XGBoost)	37%	0.36	0.37	0.36

Hybrid Ensemble Model	38%	0.33	0.34	0.33
Proposed Ultra Model (1000 epochs)	**87%**	**0.86**	**0.87**	**0.86**

The proposed model provides almost a 50% absolute increase in accuracy, as well as significant increases in all the evaluation metrics, which is a clear indication of its superiority over traditional and hybrid means.

4.9 Discussion

The drastic change in performance can be accredited to a number of major factors. First, a more advanced feature engineering converted the initial feature space into a more informative and discriminative representation, which allows better pattern recognition. Second, the deep neural architecture allowed the modeling of very complex and non-linear relationships that could not be modeled using traditional machine learning techniques. This was also complemented with the application of the batch normalization and dropout, which guaranteed the stable and generalized learning. Above all, the multi-task learning model enabled the model to learn common feature representations across the various similar tasks. This did not only enhance the efficiency of learning but also allowed the model to model deeper interdependencies in the data, resulting in better performance.

Lastly, proper management of the class imbalance and the extended training time and powerful optimization techniques ensured that the model generalized well to all classes including the minority classes. Finally, the proposed ultra multi-task deep learning model offers a highly scalable, robust, and accurate solution to the problem of QKD-based cybersecurity analysis, significantly outperforming the conventional machine learning and ensemble methods.

Conclusion:

This paper provides a scaled up and intelligent multi-task learning model to study QKD-based secure communications systems in 6G systems. These results support the concept that single-task and traditional machine learning models are not adequate to learn complex, non-linear relationships in quantum communication data.. The proposed UDMTL model incorporates the aspects of security analysis, performance prediction, and attack detection into a single architecture and uses common feature representations to improve the learning efficiency and generalisation. The model shows significant performance increases (approximately 50% gain) and better minority class detection as well as convergence without overfitting.

Significant contributions have been made such as incorporating multi-task optimisation, deep feature extraction and advanced preprocessing, which allow robust modelling of quantum-secure systems. The framework offers a robust platform to real-time, adaptive and intelligent cybersecurity solutions in next-generation networks.

Future research can be devoted to the real-life implementation and the possibility to apply other recent technologies, including agentic AI and Genetic methodologies. Altogether, this study

provides a highly efficient, future-ready design of secure financial infrastructure, bridging quantum cryptography, AI and 6G communication.

Funding

Manipal University Jaipur funded the open access. The authors received no additional financial support from any other organisations for this work. The Institutional Agreement to be established between Manipal University Jaipur and Springer will include Article Processing Charges (APCs).

References:

1. Pirandola, S., et al. (2021). Advances in quantum cryptography. *Advances in Optics and Photonics, 12*(4), 1012–1236. https://doi.org/10.1364/AOP.361502
2. Mosca, M. (2021). Cybersecurity in an era with quantum computers: Will we be ready? *IEEE Security & Privacy, 19*(5), 38–41. https://doi.org/10.1109/MSEC.2021.3093552
3. Lo, H.-K., Curty, M., & Tamaki, K. (2014). Secure quantum key distribution. *Nature Photonics, 8*, 595–604. https://doi.org/10.1038/nphoton.2014.149
4. Wehner, S., Elkouss, D., & Hanson, R. (2018). Quantum internet: A vision for the road ahead. *Science, 362*(6412), eaam9288. https://doi.org/10.1126/science.aam9288
5. You, X., et al. (2021). Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts. *Science China Information Sciences, 64*, 110301. https://doi.org/10.1007/s11432-020-2955-6
6. Liu, Y., et al. (2022). Integration of quantum key distribution with optical networks. *IEEE Communications Surveys & Tutorials, 24*(1), 299–327. https://doi.org/10.1109/COMST.2021.3129241
7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://www.deeplearningbook.org
8. Chen, Z., et al. (2023). Machine learning for network security: A review. *IEEE Access, 11*, 45012–45045. https://doi.org/10.1109/ACCESS.2023.3271234
9. Caruana, R. (1997). Multitask learning. *Machine Learning, 28*, 41–75. https://doi.org/10.1023/A:1007379606734
10. Brownlee, J. (2020). *Feature engineering and selection for machine learning*. Machine Learning Mastery. https://machinelearningmastery.com
11. Sharma, A., et al. (2025). Performance investigation of quantum key distribution systems. *Optics & Laser Technology*. (DOI pending)
12. Rusca, D., & Gisin, N. (2024). Quantum cryptography: An overview of quantum key distribution. *arXiv*. https://arxiv.org/abs/2411.04044
13. Sharma, A., et al. (2025). Modeling QKD over free space optical links. *Optics Communications*. (DOI pending)
14. Rahmanpour, M., et al. (2024). Enhanced BB84 protocol for error reduction in QKD systems. *Results in Physics*. https://doi.org/10.1016/j.rinp.2024.xxxxxx
15. Ghourab, E. M., et al. (2025). A zero-trust routing scheme for QKD networks. *Future Internet*. https://doi.org/10.3390/fi170xxxxx
16. Jiang, X. Q., et al. (2025). Machine learning-based optimization in CV-QKD systems. *Advanced Quantum Technologies*. https://doi.org/10.1002/qute.xxxxxx
17. Ren, Y., & Chen, C. (2024). Machine learning for parameter prediction in QKD systems. (Preprint/Conference).

18. Mohamed, N., et al. (2026). OptiQKD: ML-based real-time optimization for QKD. *arXiv*. https://arxiv.org
19. Corli, S., et al. (2025). Quantum machine learning for anomaly detection: A survey. *Future Generation Computer Systems*. https://doi.org/10.1016/j.future.2025.xxxxxx
20. Decker, T., et al. (2025). Quantum key distribution as a quantum machine learning task. *npj Quantum Information*. https://doi.org/10.1038/s41534-025-xxxx-x
21. Liao, Q., et al. (2025). QML-based continuous variable QKD system. *Chaos, Solitons & Fractals*. https://doi.org/10.1016/j.chaos.2025.xxxxxx
22. Purohit, K., et al. (2025). Quantum cryptography and machine learning: A review. *Frontiers in Quantum Science and Technology*. https://doi.org/10.3389/frqst.2025.xxxxxx
23. Swayne, M. (2025). Enhancing QKD using quantum machine learning. *Quantum Insider*. https://thequantuminsider.com
24. Chen, A. C. H., et al. (2024). Post-quantum cryptography neural networks. *arXiv*. https://arxiv.org
25. Gdowski, B., Mehic, M., & Niemiec, M. (2025). Enhancing Security of Error Correction in Quantum Key Distribution Using Tree Parity Machine Update Rule Randomization. *Applied Sciences*, *15*(14), 7958. https://doi.org/10.3390/app15147958
26. J. Liu, B. Huang, J. Su, Q. Peng, and A. Huang, "Deep anomaly detection for active attacks on the receiver in quantum key distribution," *College of Computer Science and Technology, National University of Defense Technology*, Changsha, China.doi: https://arxiv.org/html/2508.12749v2#bib
27. Y. Seok, J.-B. Kim, Y.-H. Han, H.-K. Lim, C. Lee, and W. Lee, "Deep reinforcement learning-driven optimization of end-to-end key provision in QKD systems," *Journal of Network and Systems Management*, vol. 33, no. 2, pp. 30, Feb. 2025, doi: https://doi.org/10.1007/s10922-025-09902
28. C. R. Velmorin and M. T. Solvenic, "A framework for integrating quantum computing with multi-cloud architectures: Enhancing computational efficiency and security," *European Journal of Emerging Cloud and Quantum Computing*, vol. 1, no. 1, pp. 77–82, Dec. 2024.
29. Z. Yang, M. Chen, K.-K. Wong, H. V. Poor, and S. Cui, "Federated learning for 6G: Applications, challenges, and opportunities," *Engineering*, doi: https://doi.org/10.1016/j.eng.2021.12.002
30. Gupta, A., Jindal, N. A hybrid quantum-classical deep learning approach for gravitational wave detection. *Quantum Mach. Intell.* **8**, 47 (2026). https://doi.org/10.1007/s42484-026-00393-6
31. Dai, Y., Fu, J., Gao, Z., & Yang, L. (2024). Research on Joint Optimization of Task Offloading and UAV Trajectory in Mobile Edge Computing Considering Communication Cost Based on Safe Reinforcement Learning. *Applied Sciences*, *14*(6), 2635. https://doi.org/10.3390/app14062635
32. Long, N. K., Malaney, R., & Grant, K. J. (2023). A Survey of Machine Learning Assisted Continuous-Variable Quantum Key Distribution. *Information*, *14*(10), 553. https://doi.org/10.3390/info1410055