International Journal of Technology 13(5) 1035-1044 (2022)
Received May 2022 / Revised August 2022 / Accepted August 2022

International Journal of Technology

http://ijtech.eng.ui.ac.id

Modelling and Predicting Quality-of-Experience of Online Gaming Users in 5G Networks

Keat Han Tan[1], Heng Siong Lim[1*], Kah Seng Diong[2]

[1]*Faculty of Engineering and Technology, Multimedia University, 75450 Melaka, Malaysia*
[2]*Technical Service, ZTE (Malaysia) Corporation Sdn. Bhd., 55100 Kuala Lumpur, Malaysia*

Abstract. 5G technology will greatly improve quality of human life by enabling new use cases that will fully leverage on the improved throughput, connections, and latency of the 5G networks. Enhanced Mobile Broadband (eMBB), which supports ultra-high throughput, is one of the most important features in 5G networks. This service is expected to improve users' quality of experience (QoE) when using resource-intensive and far more interactive applications such as playing online games. It is widely known that 5G networks can be used for gathering network monitoring data and application metrics; however, the correlation between the data and the users' QoE is not well understood. Since large amount of data can be collected, machine learning approach is well suited for predicting users' QoE when playing online games in 5G networks. In this paper, an artificial neural network (ANN) model is proposed to predict the users' QoE based on the network monitoring data of a 5G network during an online gaming session and the model's performance is evaluated. The ANN model consists of four layers which include one input layer, two hidden layers, and one output layer. The Unified Management Expert (UME) system is used to collect the network monitoring data from a 5G NSA indoor private campus network. The proposed ANN model achieves prediction accuracy of close to 80% using 30 most relevant features derived from the radio access network monitoring data.

Keywords: 5G; Artificial Neural Network (ANN); Online Gaming; Quality of Experience (QoE)

1. Introduction

Human life will be transformed by 5G technology. This is because 5G technology has significant advantages over its predecessor, which not only offers a higher data rate, but it also provides a technological platform to support machine to machine communications. For example, connecting millions of Internet of Things (IoT) sensors together to a 5G network, and advanced services requiring very low latency communication, such as remote surgery and autonomous driving. 5G services can be classified into three scenarios - enhanced mobile broadband (eMBB), ultra-reliable and low latency communications (URLLC), and massive machine-type communications (mMTC). To meet the technical requirements of 5G networks, a better quality-of-experience (QoE) model is required than what is currently used for 2G, 3G, and 4G services.

eMBB, which provides users with higher throughputs and data rates, is one of the most important use cases in 5G networks. It is able to support applications with very high video

*Corresponding author's email: hslim@mmu.edu.my, Tel.: +6062523822
doi: 10.14716/ijtech.v13i5.5866

traffic such as VR/AR, live streaming, UHD video, online gaming, cloud gaming, etc. Online gaming and cloud gaming contribute a significant amount of data traffic to 5G networks. Traditional video games are designed to depict a virtual 2D environment through a monitor. The first gaming methodology, known as "arcade" allow users to interact with the game. Then came multiplayer games, which allow users to play offline games with a few other players simultaneously. Online gaming and cloud gaming are becoming more popular. They are the mainstream of current and future gaming approaches, as communication networks improve to provide higher data rates and lower latency. 5G networks is expected to enhance online gaming and cloud gaming experience by connecting users around the world to play games simultaneously with high connection quality and ultra-low latency.

Traditionally, performance monitoring of the 2G/3G/4G networks are based on the existing standards which focus on technical Key Performance Indicators (KPIs) and Quality of Service (QoS). However, these KPIs and QoS hide the real experience of users and do not directly measure the target user experience. This is because QoS is defined as the technical point of view on service quality of the service providers. In order to provide better services to 5G users, a more holistic approach based on Quality of Experience (QoE) is needed which directly measures the target user experience. The definition of QoE according to ITU is the “degree of delight or annoyance of the users of an application or service” (ITU, 2017). However, there is no standard way on how to measure the user’s QoE for 5G networks. Therefore, this project aims to develop a machine learning model for modelling and predicting QoE based on the network monitoring data collected from an indoor 5G non-standalone (NSA) private campus network.

The remainder of this paper is structured as follows: Related work is presented in Section 2. Section 3 describes the methodology of the project. The result and discussion are described in Section 4. Section 5 concludes the paper.

2. Related Works

According to (Banović-Ćurguz & Ilišević, 2019), the relationship between QoS and QoE may be service-specific, non-parametric, non-linear and not straightforward. This is because QoE is not directly reflected in measured network data and indeterminism of human behaviour. The authors suggested a 5G KPIs-QoS-QoE mapping framework to determine the end-to-end (E2E) user QoE based on both subjective and objective assessments. Due to the highly non-linear relationship between the KPIs and the QoE, data-driven modelling will be able to provide the required prediction accuracy.

In (Liotou et al., 2016), a roadmap for QoE metrics and frameworks in 5G networks is proposed. The link between QoS and QoE mapping, QoE influencing variables, QoE assessment and estimate approaches, and QoE estimation models are discussed in this work. Perception-centric QoS-QoE mapping and stimulus-centric QoS-QoE mapping are the two methodologies offered for the QoS and QoE mapping process. The importance, changes, and influence on QoE as we move toward 5G are discussed in (Malik, 2020). It provides an overview of the QoE and the reason why does not address all aspects of network performance, including user experience.

The authors in (Laselva et al., 2018) propose a customizable QoE service model and assessment approach based on key quality indicators (KQIs) for objective user QoE measurement. To begin, data of network KPIs is gathered from a variety of sources, KPI normalisation is the following phase, which is used to compare and process the specified KPIs. The normalised KPIs are then used to compute the KQIs. The weighted-mean technique is applied, in which each low-level KPI is given a weight. This is done to see how important the indicator is when it's mapped to the KQI. The QoE score can be calculated based on the weighted combination of KQI, which captures the objective user’s QoE.

In (Pierucci, 2015), a QoE prediction framework based on neural network (NN) (or ANN) approach is suggested for 5G networks. The concept is to utilise a NN to link the QoS parameter to the QoE values. Network KPIs (e.g. channel quality indicator, user throughput, data volume, modulation order, and coding) can be used as input to the NN. The NN produces one of the QoE values as the output. A test on real KPIs collected from the HSPA network of TIM using a basic multi-layer perception (MLP) network with two hidden layers delivers satisfactory results. The author argues that the use of the NN approach enables simple adaptability of the categorization to changes in KPI features, which is the main benefit of the MLP NN. However, the proposed framework has not been tested on a real 5G network.

In (Schwarzmann et al., 2019), estimation of video streaming QoE in the 5G architecture using machine learning (ML) approach is presented. Network data analytics function (NWDAF) is introduced to collect data from application functions (AFs), control plane, management plane, and user plane. The network-level monitoring data statistics allow the application of ML techniques to predict user's QoE. The proposed framework can be separated into three phases. In the first phase, network performance data are collected by NWDAF through the third-party AF. In the second phase, the network features derived from the collected data are ranked based on their significance in influencing QoE. Then, a subset of the features and ground truth QoE are selected to train the ML models for QoE prediction. After completing each test, the performance based on the selected feature sets and ML-based models are evaluated. If the performance is not achieving requirement, the process can be repeated with different feature sets and ML-based models. The third phase is the deployment phase, where the feature sets and ML models selected are applied to estimate the QoE in real time. The results show QoE score can be reliably estimated using support vector regression based solely on network monitoring data. However, traces generated within an OMNeT++ simulation are used; the framework still needs to be validated within a real 5G deployment.

Online gaming is gaining increasing attention and popularity. QoE evaluation of online gaming applications is very important for both the game providers and the network service providers. However, due to the interactive nature of the online game applications among human and machines, it is very difficult to accurately predict online gaming's QoE.

Many researchers have explored the network characteristics for gaming QoE evaluation, yet there is no common QoE model for online mobile gaming (OMG). A gaming QoE paradigm for OMG is explored in (Moller et al., 2018). It also discusses a few research questions, such as whether the cloud gaming (CG) QoE assessment model proposed in (Yang et al., 2019) can be used to measure OMG QoE, how important is jitter when exploring the impact of packet loss on gaming QoE, which features influenced the most to the overall gaming QoE during playing online mobile game, and how can gaming QoE be predicted using only network monitoring data. Some ML-based QoE predictive models for video gaming and multimedia applications are highlighted in (Huang et al., 2018; Anwar et al., 2020; Kougioumtzidis et al., 2022) but most of them are relying on application-level metrics. Moreover, they are not targeting 5G networks. In this work, we study the feasibility of an ANN approach for QoE prediction based on real-life physical-layer traffic patterns and statistics and evaluates their performance for an online gaming use-case in 5G network.

3. Methodology

Figure 1 shows the proposed ML-based QoE prediction framework. The 5G NSA network service used in this study is an indoor private network available in the MMU-ZTE Training Centre. In the data collection phase, the Unified Management Expert (UME) software is used to collect different types of raw radio access network monitoring data, such as UE uplink average MCS, UE downlink average MCS, Channel Quality Indicator (CQI), Precoding Matrix Indicator (PMI), SRS SINR, Downlink Physical Throughput, Uplink Physical Throughput, PUCCH Average SINR, PUSCH Frequency Offset, and PUSHC Average SINR. The true QoE is the user's experience during their gaming session. It is collected from the users in the form of mean opinion score (MOS) after every gaming session using a survey form. The features derived from the collected raw network monitoring data will act as the inputs for training the ML model, while the true QoE data are the expected outputs. All the collected data, including network monitoring data from UME and users' true QoE, are processed and labelled accordingly. In our data collection campaign, a UE which consists of a laptop and a 5G customer premise equipment (CPE) is used to connect to the 5G network. The user will then play the online game, Apex Legends, at different locations within the coverage of the 5G service. All the network monitoring data mentioned above are collected using the UME software with a time resolution of 1s. Figure 2 shows the 5G NSA network employed in this work and the data collection locations.

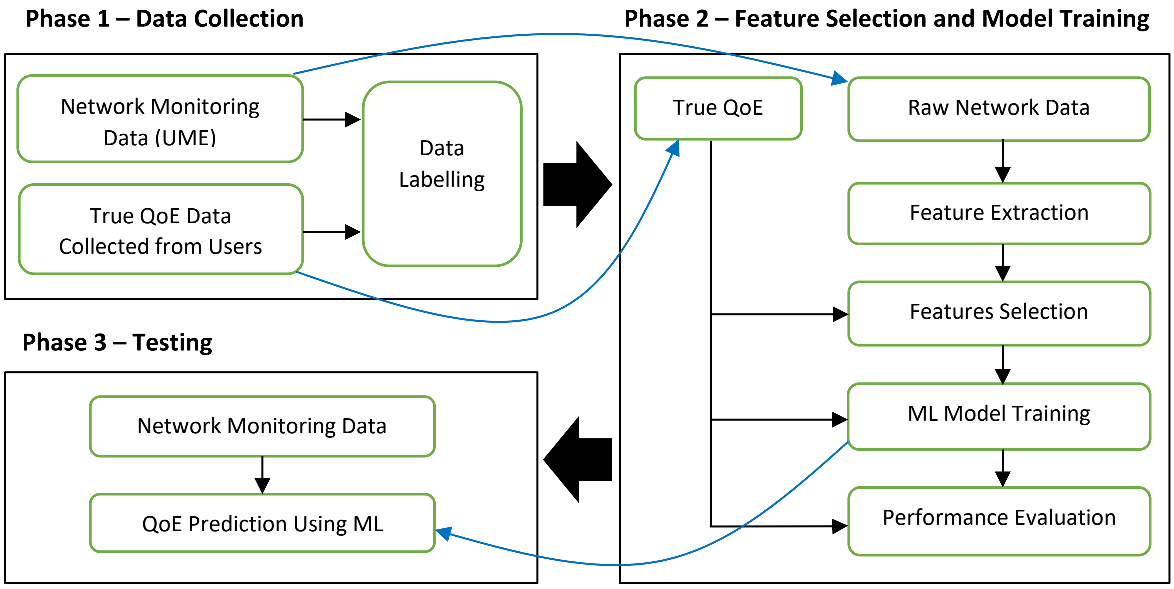
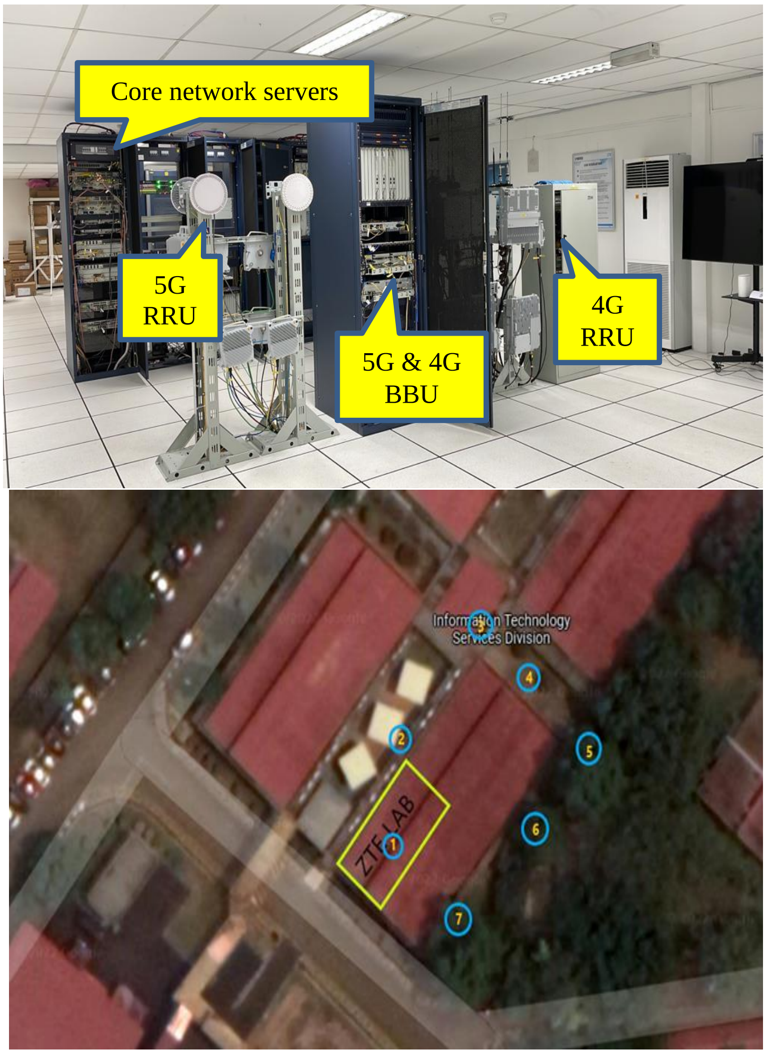


Figure 1 Proposed ML-based QoE modelling and prediction framework

The second phase of the framework consists of feature selection and machine learning model training. Features extraction is first performed for every set of data. The statistics (or features) for each network monitoring data are calculated, including the mean, standard error, median, mode, standard deviation, sample variance, kurtosis, skewness, range, maximum, and minimum. In this work, 10 different types of network monitoring data are collected for each online gaming session. Therefore, every instance of data in our dataset has 110 features (or statistics). F-Test is used for feature selection and ranking in order to select the most relevant features for the ML model. Those irrelevant features which negatively affect the model's performance will be removed. The selected features are used to train a ML model for QoE prediction. In this work, an artificial neural network (ANN) model with four layers (one input layer, two hidden layers, and one output layer) is proposed. The number of nodes in the hidden layer 1 and 2 are set to 250 and 200 respectively. Rectified Linear Unit (ReLU) activation function is used for the hidden nodes. For the output layer, 5 nodes are used to predict the 5 different levels of user's experience (from very poor to excellent). Softmax activation function is used for the output nodes. The performance evaluation is performed after each epoch of model training. If the performance evaluation does not satisfy the requirement for accuracy, the features are reselected and the ML model is trained again. This process is repeated until the performance converges or meets requirement.



(a)

(b)

Figure 2 (a) Industrial-grade 5G NSA network equipment, (b) Measurement locations

In phase 3, which is the testing phase, after selecting the features that achieve the required accuracy for the ML model, the UME is used to collect the required network monitoring data only. Finally, the ML model uses only the features that are selected to predict the user's QoE in 5G networks when playing online games. The novelty of this work is mainly in the proposed feature extraction and selection as well as the ML modeling to accurately predict user's QoE.

4. Results and Discussion

4.1. Feature Selection

Table 1 shows the result of feature selection using ANOVA F-Test. It ranks the features from the most influential feature that affects the output to the least influential feature. From this table, the most relevant features can be selected for training the ANN model. The result shows that the variance, standard deviation and standard error derived from the uplink (UL) data and control channels (i.e., PUSCH and PUCCH) and the sounding reference signal (SRS) data have critical influence on the user's QoE. These features measure the amount of variability of the UL channels' conditions during an online gaming session. The radio access network data which are most useful for determining the user's QoE are SINR, MCS index, and physical throughput.

Table 1 Features ranking based on ANOVA F-Test

Rank	Feature	Rank	Feature
1	VARIANCE_PUSCHAVGSINR	16	MODE_PUCCHSINR
2	VARIANCE_PUCCHSINR	17	RANGE_PUCCHSINR
3	VARIANCE_SRSSINR	18	MEDIAN_PUCCHSINR
4	VARIANCE_ULAvgMCS	19	RANGE_PUSCHAVGSINR
5	SERROR_PUCCHSINR	20	MEDIAN_SRSSINR
6	SDEVIATION_PUCCHSINR	21	MIN_PUCCHSINR
7	SERROR_SRSSINR	22	MODE_PUSCHAVGSINR
8	SDEVIATION_SRSSINR	23	VARIANCE_DLAvgMCS
9	SDEVIATION_PUSCHAVGSINR	24	RANGE_SRSSINR
10	SERROR_PUSCHAVGSINR	25	MEDIAN_ULPHYTHROUGHPUT
11	SERROR_ULAvgMCS	26	MODE_ULPHYTHROUGHPUT
12	SDEVIATION_ULAvgMCS	27	MIN_SRSSINR
13	MODE_SRSSINR	28	SDEVIATION_DLAvgMCS
14	MEAN_PUCCHSINR	29	SERROR_DLAvgMCS
15	MEAN_SRSSINR	30	MIN_PUSCHAVGSINR

4.2. Number of Features

Table 2 shows the QoE prediction performance of the ANN model employing different number of features. The number of epochs is set to 100 and the train-test splitting is set to 20%. This means that 80% of the dataset is used for training, while the remaining 20% is for testing. In our case, the dataset has 270 instances of data. The ANN model used in this test has four layers - one input layer, two hidden layers, and one output layer.

Table 2 Performance of the ANN model employing different number of features (100 training epochs)

No. of Features	Train Accuracy	Train Precision	Train Recall	Train Loss	Test Accuracy	Test Precision	Test Recall
5	81.85	53.30	44.91	36.88	44.44	44.44	44.44
10	85.37	68.35	50.00	30.65	50.00	50.00	50.00
15	90.37	79.79	69.44	22.00	59.26	59.26	59.26
20	92.04	81.86	77.31	18.78	70.37	70.37	70.37
25	94.44	86.79	85.19	13.89	66.66	66.66	66.66
30	95.93	90.95	88.43	11.73	**72.22**	**72.22**	**72.22**
35	97.69	94.42	93.98	8.00	64.81	64.81	64.81
40	98.89	97.22	97.22	4.91	66.66	66.66	66.66
50	99.26	98.15	98.15	2.72	55.55	55.55	55.55

It can be observed that by increasing the number of features from 5 to 50 features, the training accuracy improves from 81.85% to 99.26%, the precision improves from 53.3% to 98.15%, the recall rate improves from 44.91% to 98.15%, and the training loss is reduced from 36.88% to 2.72%. But for the testing, the accuracy improves from 44.44% to 72.22% when the number of features increases from 5 to 30 features. The performance of the ANN model drops when the number of features used is more than 30. This means that the ANN model is overfitting. Overfitting refers to the situation where the model performs well during the training stage, but performs poorly in the testing stage. It happens when a model learns the details and noise in the training data to the extent that it negatively impacts the prediction performance of the model.

Table 3 shows the performance of the ANN model considering a higher number of training epochs of 500. The purpose of this test is to make comparison between the models with 100 and 500 training epochs, as well as the impact of the different number of features used. The other parameters remain the same as in the previous test.

Table 3 Performance of the ANN model trained with 500 epochs

No. of Features	Train Accuracy	Train Precision	Train Recall	Train Loss	Test Accuracy	Test Precision	Test Recall
5	82.96	57.92	54.17	33.98	48.15	48.15	48.15
10	89.35	77.30	66.20	23.14	59.26	59.26	59.26
15	97.87	94.88	94.44	7.33	61.11	61.11	61.11
20	98.33	95.83	95.83	5.46	59.26	59.26	59.26
25	99.26	98.15	98.15	1.64	**66.66**	**66.66**	**66.66**
30	100.00	100.00	100.00	0.58	62.96	62.96	62.96
35	99.63	99.07	99.07	0.77	64.82	64.82	64.82
40	100.00	100.00	100.00	0.04	59.26	59.26	59.26
50	100.00	100.00	100.00	0.01	55.55	55.55	55.55

It can be observed that by increasing the number of epochs to 500, the training accuracy of the ANN model is improved. However, due to the overfitting problem, the testing accuracy is reduced when 500 training epochs is considered. For example, for 30 features, the testing accuracy is reduced from 72.22% to 62.96% when the number of epochs is increased from 100 to 500.

4.3. Number of Training Epochs

Table 4 shows the QoE prediction performance of the ANN model using different number of training epochs. The ANN model has four layers - one input layer, two hidden layers, and one output layer. The train-test splitting is set to 20%. The number of features is fixed to the top 20 features from the feature selection result (see Table 1).

Table 4 Performance of the ANN model employing different number of training epochs

Epochs	Train Accuracy	Train Precision	Train Recall	Train Loss	Test Accuracy	Test Precision	Test Recall
25	84.54	65.61	47.69	30.00	55.56	55.56	55.56
50	88.89	76.97	63.43	24.23	62.96	62.96	62.96
100	92.04	81.86	77.31	18.78	**70.37**	**70.37**	**70.37**
150	94.72	87.68	85.65	14.74	64.81	64.81	64.81
200	95.00	88.94	85.65	12.66	64.81	64.81	64.81
250	96.39	91.16	90.74	10.31	64.81	64.81	64.81
300	97.50	94.37	93.06	8.55	64.81	64.81	64.81
400	97.31	93.49	93.06	6.13	64.81	64.81	64.81
500	98.33	95.83	95.83	5.46	59.26	59.26	59.26
600	98.89	97.22	97.22	3.11	61.11	61.11	61.11
700	99.07	97.69	97.69	2.77	61.11	61.11	61.11
800	99.63	99.07	99.07	1.79	64.81	64.81	64.81
900	99.26	98.15	98.15	1.83	64.81	64.81	64.81
1000	99.81	99.54	99.54	1.08	64.81	64.81	64.81

Based on the result in Table 4, it can be observed that when the number of training epochs increases from 25 to 1000, the training accuracy, precision, recall rate and loss are all improved. However, the testing outcome shows the best performance is achieved for 100 epochs. After 100 epochs of training, the testing performance reduces, which means that the ANN model is overfitting.

Table 5 shows the performance of the ANN model versus the number of epochs when the number of features is increased to 25.

Table 5 Performance of the ANN model employing 25 features

Epochs	Train Accuracy	Train Precision	Train Recall	Train Loss	Test Accuracy	Test Precision	Test Recall
25	87.13	73.05	56.48	27.17	57.41	57.41	57.41
50	91.39	81.54	73.61	20.16	68.52	68.52	68.52
55	91.02	79.60	74.07	20.62	**70.37**	**70.37**	**70.37**
100	94.44	86.79	85.19	13.89	66.66	66.66	66.66
150	96.94	92.96	91.67	9.19	64.81	64.81	64.81
200	97.04	92.59	92.59	7.40	64.81	64.81	64.81
250	98.52	96.73	95.83	5.26	66.66	66.66	66.66
300	98.98	97.67	97.22	3.41	62.96	62.96	62.96
400	99.44	98.61	98.61	1.63	64.81	64.81	64.81
500	99.26	98.15	98.15	1.64	66.66	66.66	66.66
600	99.81	99.54	99.54	0.90	64.81	64.81	64.81
700	99.81	99.54	99.54	0.61	66.66	66.66	66.66
800	100.00	100.00	100.00	0.39	66.66	66.66	66.66
900	100.00	100.00	100.00	0.19	68.52	68.52	68.52
1000	100.00	100.00	100.00	0.10	66.66	66.66	66.66

By increasing the number of features from 20 to 25 features, the training accuracy, training precision, and training recall rate are also improved. But this improvement is not reflected in the testing performance. The testing performance reaches its maximum at 55 epochs, which is 70.37% for testing accuracy, and then drops down and remains roughly the same from 100 epochs onward toward the end of the testing using 25 features.

Table 6 shows the QoE prediction performance of the ANN model employing 30 features versus the number of training epochs. Higher number of features is considered in this test to understand the effect on the prediction performance.

Table 6 Performance of the ANN model employing 30 features

Epochs	Train Accuracy	Train Precision	Train Recall	Train Loss	Test Accuracy	Test Precision	Test Recall
25	86.67	70.93	56.48	27.26	61.11	61.11	61.11
50	90.74	79.29	72.69	19.70	68.52	68.52	68.52
100	95.93	90.95	88.43	11.73	**72.22**	**72.22**	**72.22**
150	97.96	95.33	94.44	7.16	68.52	68.52	68.52
200	97.96	94.91	94.91	6.31	66.67	66.67	66.67
250	99.44	98.61	98.61	2.76	66.67	66.67	66.67
300	99.63	99.07	99.07	1.70	66.67	66.67	66.67
400	100.00	100.00	100.00	0.79	68.52	68.52	68.52
500	100.00	100.00	100.00	0.58	62.96	62.96	62.96
600	100.00	100.00	100.00	0.34	62.96	62.96	62.96
700	100.00	100.00	100.00	0.15	59.26	59.26	59.26
800	100.00	100.00	100.00	0.10	62.96	62.96	62.96
900	100.00	100.00	100.00	0.08	62.96	62.96	62.96
1000	100.00	100.00	100.00	0.05	62.96	62.96	62.96

By increasing the number of features to 30, it can be observed that the training performance has improved as well. At 400 epochs, the training accuracy reaches 100%. The testing accuracy improves to 72.22% at 100 epochs, However, after 100 epochs, the testing performance starts to degrade and remain roughly the same until the end of the testing. This is because the ANN model is overfitting after 100 epochs of training. By making a comparison between the results of these three tests, we can observe that the best performance is achieved with 30 features at 100 training epochs, which gives 72.22% of testing accuracy.

4.4. Number of Hidden Layer Nodes

Table 7 shows the QoE prediction performance of the ANN model versus different combinations of number of hidden layer nodes. In this test, the top 30 features are considered, and the settings of the other parameters are given in Table 7.

Table 7 Performance of the ANN model versus different combinations of number of hidden layer nodes (100 training epochs)

Hidden Layer Nodes	Train Accuracy	Train Precision	Train Recall	Training Loss	Test Accuracy	Test Precision	Test Recall
50, 50	90.37	81.11	67.59	21.52	72.22	72.22	72.22
80, 50	91.39	83.98	70.37	20.41	74.07	74.07	74.07
80, 80	91.85	82.00	75.93	18.78	72.22	72.22	72.22
100, 80	93.89	87.50	81.02	16.91	**77.78**	**77.78**	**77.78**
100, 100	94.07	88.00	81.48	16.19	70.37	70.37	70.37
150, 100	93.89	87.50	81.02	14.67	70.37	70.37	70.37
150, 150	94.54	87.56	84.72	13.13	68.52	68.52	68.52
200, 150	94.35	86.73	84.72	14.26	74.07	74.07	74.07
200, 200	95.28	89.10	87.04	12.52	70.37	70.37	70.37
250, 200	95.93	90.95	88.43	11.73	72.22	72.22	72.22
250, 250	95.83	90.14	88.89	10.96	68.52	68.52	68.52
300, 300	96.57	92.02	90.74	10.23	70.37	70.37	70.37
400, 300	96.94	92.56	92.13	10.28	68.52	68.52	68.52
400, 400	97.22	93.46	92.59	8.46	68.52	68.52	68.52
500, 400	96.85	92.52	91.67	8.40	68.52	68.52	68.52
500, 500	97.13	93.43	92.13	8.05	68.52	68.52	68.52

It can be observed that different combination of the number of hidden layer nodes gives different performance. When the number of nodes increases, the training performance of the model improves as well. However, the testing performance didn't show similar improvement. The best performance that can be achieved is 77.78% when the number of hidden layer nodes is set to 100 nodes for the first hidden layer and 80 nodes for the second hidden layer.

5. Conclusions

In the past, networks KPIs and QoS were used to measure the user experience. However, they do not directly target the actual user experience. The current trend in measuring user experience is towards QoE evaluation. But there is still a lack of study regarding data-driven 5G user's QoE prediction using real-life network monitoring data. Since QoE is also dependent on specific use case, online gaming is selected in this study which is one of the important use cases of 5G. A machine learning model, namely an artificial neural network (ANN) model, is proposed to predict the users' QoE in a 5G network based on features extracted from radio access network monitoring data. Ten types of physical-layer data are collected using an industrial-grade 5G NSA indoor campus network for an online gaming use-case. Out of the total 110 statistical features, the top 30 features are selected based on ANOVA F-test for training the ANN model. The proposed ANN model with four layers achieves prediction accuracy close to 80% and very fast convergence in training. The performance of the ANN model may be affected by the small dataset and the inconsistency of the users when rating their experience using the survey form after each gaming session. At present there are not sufficient data instances in the constructed dataset due to the difficulty in field measurement. Moreover, different users may have different expectations, and the inconsistency of the rating will lead to reduced performance of the

ANN model. It is also worth mentioning the status of the game server is also a factor influencing the user's experience. Our current work focuses on expanding the dataset, developing more advanced machine learning model such as deep learning and including external factors such as game server status to improve the prediction accuracy.

Acknowledgements

We thank the supports of MMU-ZTE Training Centre (i.e., ZTE Lab) for providing the 5G NSA network facilities to carry out the experiments and data collection.

References

Anwar, M.S., Wang, J., Khan, W., Ullah, A., Ahmad, S., Fei, Z., 2020. Subjective QoE of 360-Degree Virtual Reality Videos and Machine Learning Predictions. *IEEE Access*, Volume 8, pp. 148084–148099

Banović-Ćurguz, N., Ilišević, D., 2019. Mapping of QoS/QoE in 5G Networks. *In*: 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 404–408

Huang, R., Wei, X., Gao, Y., Lv, C., Mao, J., Bao, Q., 2018. Data-driven QoE Prediction for IPTV Service. *Computer Communications,* Volume 118, pp. 195–204

International Telecommunication Union (ITU), 2017. Vocabulary for Performance, Quality of Service and Quality of Experience, ITU-T P.10/G.100, Geneva, Switzerland

Kougioumtzidis, G., Poulkov, V., Zaharis, Z.D., Lazaridis, P.I., 2022. A Survey on Multimedia Services QoE Assessment and Machine Learning-Based Prediction, *IEEE Access,* Volume 10, pp. 19507–19538

Laselva, D., Mattina, M., Kolding, T.E., Hui, J., Liu, L., Weber, A., 2018. Advancements of QoE Assessment and Optimization in Mobile Networks in The Machine Era. *In*: 2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), pp. 101–106

Liotou, E., Tsolkas, D., Passas, N., 2016. A Roadmap on QoE Metrics and Models. *In:* 23rd International Conference on Telecommunications (ICT), pp. 1–5

Malik, S.U.R., 2020. Moving Toward 5G: Significance, Differences, and Impact on Quality of Experience. *IEEE Consumer Electronics Magazine,* Volume 9(6), pp. 9–14

Moller, S., Schmidt, S., Zadtootaghaj, S., 2018. New ITU-T Standards for Gaming QoE Evaluation and Management. *In*: Tenth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6

Pierucci, L., 2015. The Quality of Experience Perspective Toward 5G Technology. *IEEE Wireless Communications*, Volume 22(4), pp. 10–16

Schwarzmann, S., Marquezan, C.C., Bosk, M., Liu, H., Trivisonno, R., Zinner, T., 2019. Estimating Video Streaming QoE in the 5G Architecture Using Machine Learning. *In*: Proceedings of the 4th Internet-QoE Workshop on QoE-based Analysis and Management of Data Communication Networks, pp. 7–12

Yang, K., Zhang, X., Zhao, Y., Fan, Q., Gao, Q., Lyu, Y., Yin, H., Ma, Z. 2019. Looking into Online Gaming from Measurement Perspective. *In*: 2019 IEEE International Conference on Service-Oriented System Engineering (SOSE), pp. 203–2035