International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 14, No.1 (Aug-2023)

http://dx.doi.org/10.12785/ijcds/140142

Online QoE Assessment Model Based on Incremental Stacked Multiclass Classifier

Radhia Elwerghemmi[1], Maher Heni[1], Riadh Ksantini[2] and Ridha Bouallegue[1]

[1]*Innov'COM Laboratory, Higher School of Communication (Sup'com), Ariana, Tunisia*
[2]*Department of Computer Science, College of IT, University of Bahrain, Kingdom of Bahrain*

Received 13 Feb. 2023, Revised 08 Jul. 2023, Accepted 13 Jul. 2023, Published 01 Aug. 2023

Abstract: The enormous growth of streaming services in the last decade leads to the emergence of the Quality of Experience (QoE) metric, which aims to improve and optimize the delivery of video streaming service, thus strengthening the loyalty of end-users to the provided services. Yet, predicting QoE of a multimedia stream is a challenging task because it is dependent on several different influencing factors. Moreover, it should handle dynamic environments with large-scale data. Machine learning methods offer a method for quantifying the intricate connections between various influencing factors and QoE. Thus, in this paper, a new online QoE prediction method is proposed, namely, Incremental Stacked Support Vector Machine (ISSVM). The proposed approach uses a developed stacked generalization technique to increase the global accuracy and minimize the execution time, by combining predictions of several parallel Multi-class Incremental SVM (ISVM) learners trained with different types of sub-features. Then another ISVM model is used as a meta-classifier instead of a simple linear regression model in order to build a robust fully incremental model. In fact, using the ISVM model as weak classifiers aims to handle non-stationary and very huge volumes of data in real-time contexts. The findings show that the suggested model is more effective over the rest of the state-of-the-art methods.

Keywords: Quality of experience, Ensemble Learning, Online Learning, Incremental Support Vector Machine, Video Streaming service.

1. Introduction

Recently, there has been tremendous growth in streaming video services. Cisco Visual Networking Index announced that 79% of the total mobile data traffic will be video streaming by 2022 [1]. The quality of video transfer is seriously affected by the unstable bandwidth availability of recent wireless networks. For that reason, network service providers, such as Netflix and YouTube, should continuously control the quality of the transferred videos to offer high-performance services. Hence the emergence of QoE concept, defined by The International Telecommunication Union (ITU-T, 2008) as "the overall acceptability of an application or service, as perceived subjectively by the end-user" [2].

In literature, there have been various techniques used to quantify the QoE level of video streaming service. There are three types of approaches for assessing the QoE: subjective methods (the QoE prediction is directly performed by service users), objective methods (the prediction of QoE is automatically applied via technical factors), and hybrid methods (combine subjective and objective models) [3].

Subjective QoE assessment is a realistic assessment, performed by a population of users for estimating the overall QoE of a given service. Mean Opinion Score (MOS) is the most commonly employed measure in subjective quality of experience estimation tests [4]. This score is a numeric value from 1 (excellent) to 5 (bad). However, subjective tests are costly and time-consuming which makes the real-time assessment of QoE very difficult.

Objective models use mathematical formulas or algorithms, for instance, Peak Signal-to-Noise Ratio ($PSNR$), mean squared error (MSE), Video Quality Metric (VQM) [5], [6], [7]. Yet, an extra control channel is necessary, since those measurements require a reference to the original video source which causes the overhead of the network bandwidth. Moreover, authors in [8] demonstrate that objective models do not always match the perception of a human.

The last category is called hybrid models, which are a mixture of the previously mentioned technologies (subjective and objective models). This approach is based on Machine learning (ML) algorithms [9]. It maps network parameters to subjective feedback values, which can solve real-time QoE assessment. Moreover, ML techniques quantify the complex relationship between user-perceived quality and various Influencing Factors (IFs) affecting QoE. As a result, no explicit and unique link between IFs and QoE is required.

Many state-of-the-art techniques have been developed using

E-mail address: radhia.werghemmi@supcom.tn, maher.heni@supcom.tn, rksantini@uob.edu.bh, ridha.bouallegue@supcom.tn

ML models to estimate the QoE. For example, in [10] an application of various ML techniques for QoE estimation was debated. Other research works have employed a neural network model (NN) to assess perceptual video quality [11]. Moreover, the Decision Tree (DT) model was used in [12] to build a No reference objective model to estimate End-user QoE of video stream services. Recently, Support Vector Machines (SVMs) have become widely and successfully employed for QoE assessment [13]. The main advantages of SVM are: providing a unique solution because the optimum problem is convex. Additionally, because the solution is based only on support vectors, it reduces computing complexity.

The study of [14] proposes a transfer learning model for the video perception assessment, which stacks the predictions of a generic pre-trained model with a specific trained model, to enhance the global accuracy. Although this model gives a marginal improvement, fine-tuning the proposed method requires big volumes of QoE data to add new weights or data points, as it requires working with layers in the pre-trained algorithm to get to where it provides value for developing the new model. Moreover, authors have limited feature set size, related only to the content type and they do not consider other context features, related for example to the user, the device, and the application.

Recently, deep learning models have been widely used for the estimation of the QoE. For Example, authors in [15] propose a DeepQoE framework based on deep models, which generates QoE score in an end-to-end manner. Moreover, the work in [16] presents a hybrid deep learning model for medical video QoE prediction, based on the integration of the LSTM (Long Short Term Memory) model with the boosting model. Also, the study of [17] proposes a deep learning method that employs an integrated framework made up of Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) networks.

All the techniques used in the prediction of user's perception stated above are called batch learning Models. Thus, the training procedure needs the whole data to be provided in advance. For this reason, they tend to fail in a real-time context, since data are added sequentially. Moreover, when a new model is trained, if there is new data to learn, the whole training procedure should be repeated. This is a common problem for video streaming services, since equipment must continually adjust to changing conditions.

Therefore, incremental methods are a good candidate for QoE assessment for many reasons. In effect, they are frequently significantly quicker, especially when the data set is redundant; they can also be used when the data set is not fixed. Also, they are more suitable for tracking non-stationary environments.

Only a few learning models have adopted incremental learning in the context of QoE. Among these we mention two.

The work in [18] presents an online quality of experience prediction method based on Hoeffding Trees, an incremental decision tree induction technique that can learn from huge data streams. The authors use four variations of this algorithm, namely, Standard Hoeffding Trees (HT), HT with Nave Bayes (NB), HT with adaptive NB and Hoeffding Option Trees with NB and adaptive approach (HOTNBAdaptive). This approach shows high accuracy and strong flexibility to concept drift in the database, however, this model uses only Quality of Service (QoS) available data as input to predict the QoE.

The authors in [19] presented an incremental multi-class support vector machine method for predicting the QoE of streaming video. Experimental results prove its superiority over other ML techniques based on batch learning, in terms of QoE assessment accuracy and computational complexity. Despite their high performance, all the above-cited techniques rely on single learners that have little knowledge concerning the QoE dataset. Therefore, building an ensemble learning model will assist in reducing the classification difficulty task by dividing it into a number of sub-problems. Besides, it increases global accuracy by mixing the outputs of the different learners. The most known techniques are bagging, boosting, and Stacking.

Thus, in the present paper, we propose a novel incremental QoE assessment model for video streaming service, namely, Incremental Stacked Support Vector Machine (ISSVM), which predicts the user perception via various preprocessed factors. So we have combined two principal concepts in one high-performance model, which are the Stacked Generalization approach and the online learning concept more specifically, the multi-class ISVM.

This model is based on the divide-and-conquer theory. Firstly, it decomposes the classification problem into several sub-problems processed in parallel using a pool of multi-class ISVMs, which decrease the required execution time and computational power. After that, it combines the outputs of the different classifiers to provide greater predictive accuracy. In fact, the ISVM algorithm is a good tool to handle non-stationary and huge volumes of data in a real-time context. Furthermore, because dynamic modification of the model is necessary, it reduces the complexity of the training process. Also, it helps in building expert knowledge about the problem thanks to the use of kernel tricks and ISVM convex objective function. Incremental SVM is a binary classification model, yet we extended it for solving multi-classification problems using the One-against-all method [20].

This work is an extension of our prior conference version [21]. The following are the major enhancements: 1) We run additional experiments on the new dataset and adapt our model to the new scenario. 2) We give further performance assessments to demonstrate the efficacy and robustness of our model. 3) We replace the classic meta-classifier, which is the logistic regression model (Batch model), that decrease the performance of our model with the growth of databases (Big Data) with an ISVM model to build a powerful fully incremental model (the base-classifiers and the meta-classifier are incremental models).

The remainder of this paper is arranged as follows. The background section presents the video streaming service and explains the incremental learning mechanism. In section 3 the proposed ISSVM algorithm is presented in detail. Section 4 evaluates the performance of the prediction algorithm. Finally, section 5 provides some concluding remarks and perspectives.

2. Background

A. Video Streaming service

A video streaming service allows the end user to watch the video content which is being delivered continuously from a source. Typically, the video begins playing after downloading the initial part, and the remaining part is then downloaded as it goes into the buffer memory so that the playback continues smoothly, as explained in Figure 1.
For online video distribution, video streaming protocols such as RTSP (Real Time Streaming Protocol) [22] and RTP (Real Time Protocol) [23] used to be employed. Those protocols are built on top of the UDP protocol to ensure the transmission of videos with a minimum delay. Nevertheless, using UDP, visual impairments may be present during video playing, i.e., certain video frames may be lost or corrupted as a result of packet losses [24]. Recent video streaming platforms employ the HTTP protocol to deliver videos [25]. The usage of this protocol for video streaming is simple since the majority of firewalls accept HTTP/HTTPS traffic; therefore no additional network setups to handle video traffic are required. For this reason, the widespread use of HTTP-based video streaming has become very successful. HTTP uses protocols like TCP or QUIC, so the video content will be reliably delivered to the end-user. Also, there will not be visual distortions (frame drops) in video playback.

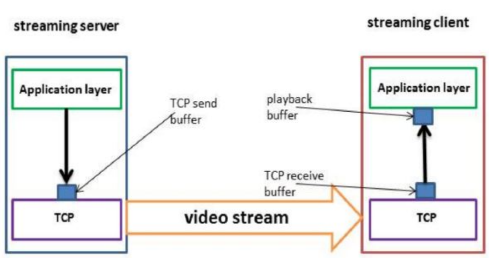

Figure 1. Video streaming service architecture

B. Incremental learning

In a classification task, the objective is to generate a model $y = f(x)$ from n training samples of the form (x, y), where y represents a class label and x a vector of d attributes. This model should be well able to predict the class y from any future sample of x.
For solving this problem, traditional batch learning models should load the whole training dataset into memory in the initial stage. Yet, with the tremendous and incessant growth of databases, a high computational effort is required. Incremental learning methods are efficient techniques to overcome those problems, which corresponds to a system capable of receiving and integrating new examples without having to carry out complete learning. A learning algorithm is incremental if, for any examples $x_1, ..., x_n$ it can produce models $f_1, ..., f_n$ such that f_{i+1} depends only on f_i and the current example x_i. This learning model is used either when the dataset is too large to be used at one time, or when the training dataset is not available in its entirety and the training samples arrive incrementally.

3. Incremental Stacked Multi-class SVMs for QoE assessment

In this section, first, the extracting features approach and the Incremental SVM algorithm are presented. Then, the multi-class incremental SVM built using the one-against-all technique is described. Following that, it introduces the ensemble learning model used in our research, the Stacked Generalization model. Finally, a flowchart is used to describe the suggested online QoE assessment approach.

A. Feature extraction

QoE is an index of users' subjective feeling, which is influenced by various end-to-end factors. To define our categories of QoE Influence Factors (IFs) we will rely on the source of parameters (user, network, application and devices) [26], as shown in Figure 2:

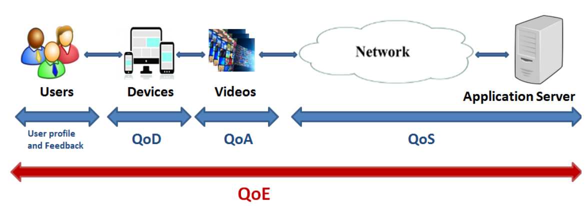

Figure 2. QoE Influence Factors.

Hence, we have five QoE IFs categories:

- Quality of Application (QoA): Specify the application variables.
- Quality of Service (QoS): Represent the infrastructure and the QoS parameters.
- Quality of Device (QoD): Related to the device characteristics.
- User Profile (UP): Represent Psychological features or human features.
- User FeedBack (UF): Linked to information gathered from the experimentation entity.

For building our ISSVM-based QoE prediction model, we run a subjective test, to generate learning databases for connecting objective metrics(User Profile, User FeedBack, QoA, QoD and QoS parameters), with the subjective quality of experience received in terms of single rating score.

B. Incremental Support Vector Machine (ISVM)

In this part, we start by presenting the classical batch SVM technique, since it is the basis of the proposed ISVM model. Then, we describe in detail the ISVM algorithm that will be used as a base classifier and a meta-classifier in our ISSVM model.

1) Batch Support Vector Machine (BSVM)

Giving a training set $X = \{(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)\}$ where $x_i = \{QoE\_IFs\} \in R^k (k \geq 1), i = 1, ..., N$ are the input QoE IFs, $y_i = QoEscore \in R \in \{1, 2, 4, 5\}$ and k is the input feature vectors dimension. BSVM tries to find an optimal hyperplane that can separate two classes while maximizing the distance between these two classes (maximizing the margin) [27]. To build such a hyperplane, the following quadratic problem must be solved:

$$0 < \underset{min}{\alpha_i} < C : W = \frac{1}{2} \sum_{i,j} \alpha_i Q_{ij} \alpha_j - \sum_i \alpha_i + b \sum_i y_i \alpha_i. \quad (1)$$

With α_i are the Lagrange multipliers, b is the offset, $Q_{ij} = y_i y_j K(x_i, x_j), K(x_i, x_j)$ is the kernel function and C is a parameter that present the misclassification cost.
After the resolution of this problem, the following equation can be used to define the separation hyperplane:

$$f(x) = \sum_{i=1}^{N} y_i \alpha_i K(x_i, x) + b \quad (2)$$

2) The Incremental version of the SVM

In batch learning, the overall dataset is available at the beginning of the learning procedure, however, in incremental learning samples can arrive at any time [28]. If a new sample is inserted and cannot be classified by the current solution, the Lagrange multipliers should be updated, while retaining the Karush-Kuhn-Tucker (KKT) conditions on all previously acquired samples, without retraining old data again.
1) KKT condition:
The saddle point of the problem expressed by eq.1 is defined using the KKT conditions:

$$g_i = \frac{\partial W}{\partial \alpha_i} = \sum_j Q_{ij} \alpha_j + y_i b - 1. \quad (3)$$

$$\frac{\partial W}{\partial b} = \sum_j y_j \alpha_j = 0. \quad (4)$$

The KKT conditions split the targeted training dataset (D) into three groups:

- The subset S denote support vectors ($g_i = 0, 0 < \alpha_i < C$).
- The subset E denote error vectors ($g_i < 0, \alpha_i = C$).
- The subset R denote non-support vectors ($g_i > 0, \alpha_i = 0$).

2) Adiabatic increments:
Maintaining the KKT conditions' equilibrium simultaneously for all previously seen training samples, we express them in the following equations:

$$\Delta g_i = Q_{ic} \Delta \alpha_c + \sum_{j \in S} Q_{ij} \Delta \alpha_j + y_i \Delta b \qquad \forall i \in D \cup \{c\}. \quad (5)$$

$$0 = y_c \Delta \alpha_c + \sum_{j \in S} y_j \Delta \alpha_j. \quad (6)$$

Where α_c is the coefficient that should be incremented. These equations are as follows:

$$Q. \begin{bmatrix} \Delta b \\ \Delta \alpha_S \end{bmatrix} = - \begin{bmatrix} y_c \\ Q_{S,c} \end{bmatrix} \Delta \alpha_c. \quad (7)$$

$$With \quad Q = \begin{bmatrix} 0 & y_S^T \\ y_S & Q_S \end{bmatrix}.$$

Where $\Delta \alpha_S$ is a vector containing the matching $\Delta \alpha_i : i \in S(\alpha)$, Q_S is a kernel matrix containing S_s and $Q_{S,c}$ is a kernels vector between S_s and x_c.
Thus, in equilibrium

$$\Delta_b = \beta \Delta \alpha_c. \quad (8)$$

$$\Delta \alpha_j = \beta_j \Delta \alpha_c \qquad \forall j \in D. \quad (9)$$

The coefficients β are calculated as follow:

$$\begin{bmatrix} \beta \\ \beta_S \end{bmatrix} = -R. \begin{bmatrix} y_c \\ Q_{S,c} \end{bmatrix}. \quad (10)$$

Where $R = Q^{-1} \quad and \quad \beta_j \equiv 0 \ \ \forall j \notin S$.

$$\Delta g_i = \gamma_i \Delta \alpha_c \ \ \forall i \in D \cup \{c\}. \quad (11)$$

Where
$\gamma_i = Q_{ic} + \sum_{j \in S} Q_{ij} \beta_j + y_i \beta, \gamma_i = 0 \ \ , \forall i \in S$.

3) The resulting updates:
For the ISSVM, if a new candidate example x_c is present, it should be added to the support vector subset, the error vector subset, or the remaining vector subset depending on g_c values and α_c. For instance, when x_c is added as a support vector, we should update the subset S. Furthermore, we can see from Eq.(10) that only the R matrix has to be computed to obtain all updated values. For adding a new point c, the R matrix should be extended as follow:

$$R \leftarrow \begin{bmatrix} R & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\gamma_c} \begin{bmatrix} \beta \\ \beta_S \\ 1 \end{bmatrix} . \begin{bmatrix} \beta \, \beta_S \ 1 \end{bmatrix}. \quad (12)$$

For removing a support vector x_k from S, the R matrix will be contracted as:

$$R_{ij} \leftarrow R_{ij} - R_{kk}^{-1} R_{ik} R_{kj} \qquad \forall i, j \in S \cup \{O\}; i, j \neq k. \quad (13)$$

In which the index O denotes the b term.
The use of the two previous formulas will decrease the computational complexity of the ISVM model from $O(n^3)$ to $O(ns^2)$, where n represents the global number of examples

in the dataset and ns is the number of support vectors.

C. Stacking model

By combining several models, ensemble learning is applied to increase machine learning performance [29]. This technique offers better predictions compared to the individual classification model.

The stacking model is an ensemble learning model, where various outputs of base-learners are combined using a meta-learner, in order to predict the QoE class of any instance x as follow :

$$\hat{y}(x) = \sum_{j=1}^{m} \beta_j h_j(x). \tag{14}$$

The level-0 predictors h_j are trained using the whole dataset, and then the level-1 algorithm (meta-learner) is trained using the predicted class labels of the level-0 classifiers (base-learners) as inputs. That means, the meta-learner is employed for learning the weights β_j of the level-0 predictors.

A stacked generalization model trains all classifiers on the entire dataset, whereas in our work each base classifier will be learned using one different type of feature subset, which minimizes the global computational complexity. Moreover, the linear logistic regression method is frequently employed as a meta-learner for multi-class problems. Yet, in this work, we will employ an ISVM model to maximize the prediction accuracy, and essential for building a fully incremental model.

D. Stacked multi-class ISVM model algorithm

This section describes the proposed QoE model based on an incremental stacked multi-class support machine algorithm using a flowchart presented in Figure 3.

For the training of the ISSVM model, we use n training samples as input: $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, where $x_i = \{QoE\_IFs\}$ and $y_i = QoEscore$.

The database features are spliced into various QoE_IFs categories presented earlier (QoS, QoA, QoD, UF, UP). Then, the multi-class ISVM models are trained in parallel using the corresponding categories as input. Since we represent the user's QoE prediction using 5 classes (rating scores), each multi-class ISVM is composed of 5 binary ISVM models. The outputs of each base-classifiers is a decision function modeling the predicted QoE score expressed as:

$$ISVM_k(x) = h_k(x) = \sum_{j=1}^{m} \alpha_{opt} k(x, x_j) + b_{opt}. \tag{15}$$

Where k is the ISVM algorithm's index. The combination of different methods into a global system is made by the meta-classifier, trained utilizing a new predictions database $D = \{x_i', y_i\}$, where $x_i' = \{h_1(x_i), h_2(x_i), h_3(x_i)\}$ whose output is a general decision function which improve the QoE score predictions.

The pseudo-code of the proposed algorithm is described in Algorithm 1. We continue moving parameters sequentially until reaching the equilibrium. The objective is to find the greatest possible increase α_c while keeping the set's decomposition intact. We must take into account the movements of some elements from one category to another during the procedure of updating. This is how adiabatic increments work [28].

Algorithm 1 Incremental Stacked multi-class SVM model algorithm: high-level summary.

1: **Input:** An example x_c, y_c
2: **Output:** A global decision function for the prediction of QoE score
3: **Initialization:**
4: Divide the dataset features into T subsets
5: Read example x_c, y_c
6: **for** t=1 to T **do**
7: Calculate R, and employ it to find β and γ using Eqs. (8)-(11)
8: Set α_c and $\Delta\alpha_c = 0$
9: Compute g_c using Eq. (3)
10: **while** while $g_c < 0$ and $\alpha_c < C$ do **do**
11: **if** $g_c = 0$ **then**
12: Add x_c to S and equilibrium is reached
13: Set $\alpha_c = \Delta\alpha_c$
14: Update $(\alpha_i)_{i=1...n}$
15: Update R according to (12)
16: **end if**
17: **if** $g_c < 0$ **then**
18: Add x_c to E and equilibrium has been attained
19: Set $\alpha_c = c$
20: **end if**
21: Update the subsets S, E, and R
22: Update R recursively according to Eqs. (12)-(13).
23: **end while**
24: Compute $h_k(x)$ according to Eqs. (15)
25: **end for**
26: construct new example $\{x_i^{'}, y_i\}$ where $x_i^{'} = \{h_1(x_i), ..., h_T(x_i)\}$
27: Update ISVM solution state of the Meta-classifier based on $\{x_i^{'}, y_i\}$ (Step 7 to Step 21)
28: Predict the final QoE score

4. EXPERIMENTAL RESULTS

To investigate the effectiveness of the proposed method for the QoE assessment in a real-time context, we have realized extensive experiments. Our dataset is described first; after that, we detailed the used experimental protocol and the classifiers. Finally, a discussion and interpretation of the experimental results are provided. All experiments are conducted on windows 10 OS with an Intel i5 CPU (single processor) and 8 GB of RAM. As a simulator, we have used MATLAB software.

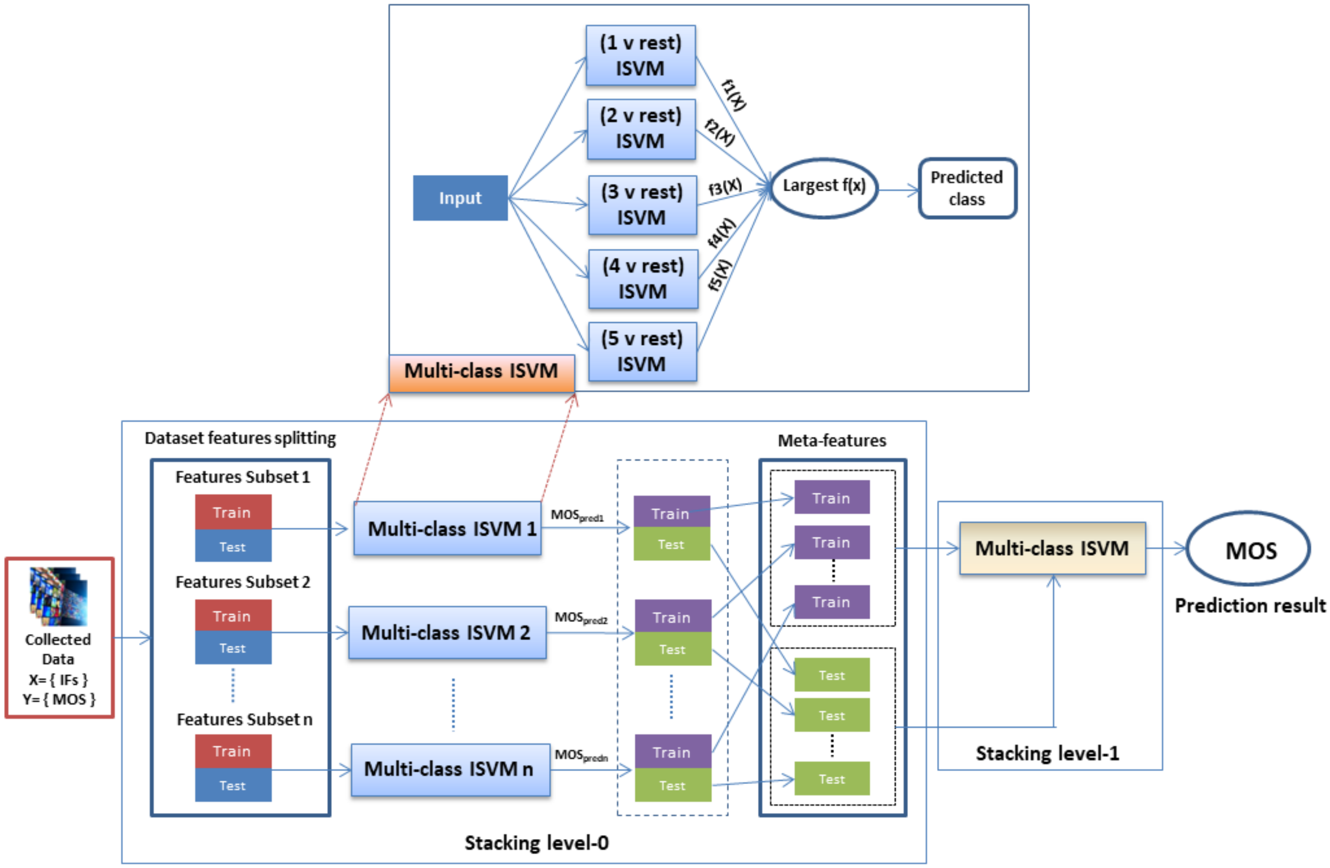

Figure 3. Flowchart of the proposed Incremental Stacked multi-class SVM model algorithm

A. Dataset used

The algorithm presented in this paper was evaluated using two different datasets, built in the LiSSi laboratory to collect a lot of QoE IFs using a VLC media player, as illustrated in Table I [30].

- Dataset 1 contains 300 samples covering 23 QoE IFs, which is constructed on the basis of a controlled laboratory testbed. The employed videos have different types and complexities, which are divided into 3 groups: the QoS , the QoA, and the QoD.
- The second dataset comprises 1543 examples that cover 20 QoE Impact Factors (QoE IFs). The utilized videos have various types and complexities, which are organized into five categories: QoS, QoA, QoD, UF, and UP.

B. Performance metrics

The evaluation of the performance of the suggested quality of experience assessment approach is organized into two steps:

First, to demonstrate the advantage of mixing various parallel multi-class Incremental SVMs, we perform a comparison between our ISSVM model, the ISVMs base classifiers, and a single ISVM.

Second, we compare the ISSVM approach to other pertinent models to illustrate the superiority of incremental learning over batch learning, and the ensemble learning models over single models. The same features databases and experimental settings are used to evaluate these classifiers. An RBF kernel, using the same kernel ”width” and regularization parameter C, was used for batch SVM, ISVM, SSVM, and ISSVM.

To evaluate the performance of the proposed ISSVM model we use the measure Prediction Accuracy, which is defined as the percentage of correct results that a classifier has achieved out of the total number of observations in the dataset.

For investigating the prediction accuracy, we employed a cross-validation method [31]. We split randomly the original database into 10 subsets. This way, a single subset is retained for testing, and the rest are employed as training data. Lastly, the accuracies are averaged to get the global

TABLE I. Details of the real QoE datasets.

Dataset	Number of features	Number of instances	Number of features subsets
Dataset 1	23	300	3
Dataset 2	20	1543	5

accuracy.

C. Classifiers

The Incremental Stacked multiclass SVM-based QoE model was compared against the following single and ensemble learning models:

- Stacked Support Vector Machine (SSVM): This classifier is the batch version of our proposed ISSVM model, in which a pool of batch SVMs are combined as base classifiers using another batch SVM as a meta classifier, in order to make a final prediction.
- Random Forest (RF): This model is an ensemble learning classifier developed by Breiman. RF model performs parallel learning on multiple decision trees randomly constructed. These single models are trained on distinct subsets of data using the bagging concept, with a random selection of features using the "random projections" approach. It predicts by taking the average or mean of the output from several trees [32].
- Adaboost Decision Tree (ADT): The ADT algorithm generates a set of weak learners and combines them to build a very efficient classifier. The weak classifiers in AdaBoost model are decision trees with a single split. Each weak learner is trained taking into account the previous learner's classification errors, by increasing the weight of instances that are not correctly classified [33].
- Support Vector Machine (SVM): SVM is a binary classification model invented by Vapnik. This model looks for the optimal hyperplane that separates two classes, with a maximum margin between the closest points of these two classes [27].
- Artificial Neural Networks (ANN): ANNs are densely linked networks of basic processors that run in parallel. Based on the input it obtains, each elementary processor computes a single output. ANN consists of three layers, input layer, hidden layer, and output layer. This model learns by adjusting its weight iteratively until the outputs are consistent with the inputs [34].

D. Results and discussion

In this section, we will show the findings of the experimental comparisons and evaluations that we conducted in order to highlight the superiority of our incremental ensemble learning model over other well-known singles and ensemble learning methods. First, we evaluate the ISSVM model on different levels. For that reason, we compare it to the used ISVM base-classifiers, each one of which uses one type of the IFs subsets, and to a single ISVM model that uses all the IFs parameters for the user's perception estimation.

Table II and Table III demonstrate the advantage of our suggested method in terms of accuracy. In fact, the estimation of QoE using only one IFs subset can give better results than the obtained ones using all IFs parameters, which confirms the assumption made by [35] that the interaction between parameters may lead to worse results than expected. As a result, it's essential to weight QoE IFs subsets. Our proposed model can take into consideration this weighting thanks to the meta-classifier, which learns the weights of the base classifiers taking into consideration the performance of each model in the QoE prediction. Furthermore, the execution time of our implemented method is lower than the execution time of one ISVM, trained using the whole dataset features. The ISSVM divides the dataset into various features subsets that will be trained in parallel by various ISVMs, which decreases the classification complexity task by splitting it into an ensemble of sub-problems. To assess the efficacy of

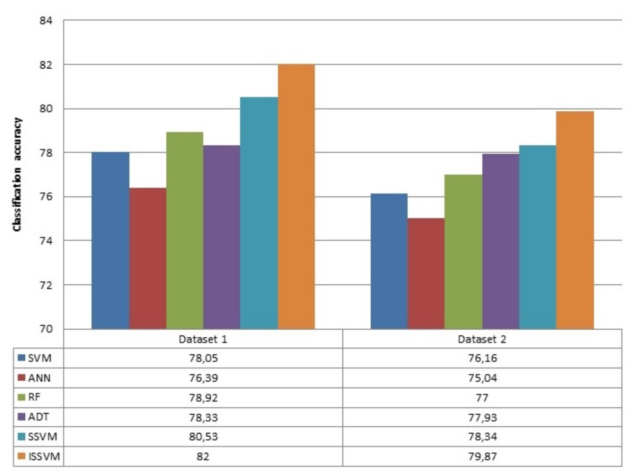

Figure 4. Average accuracy in % of ISSVM vs other batch models for two real-world datasets.

the proposed approach, we compare the ISSVM classifier to six different learning models. As we can see in Figure 4, our proposed model produces superior results in terms of accuracy measures. That was expected given that the ISSVM is built on an ensemble of multi-class ISVMs rather than a single learner (ANN and SVM). Besides, contrary to the RF and ADT ensemble learning methods, that use batch decision tree classifier as base classifiers, our ISSVM

TABLE II. QoE estimation performance of the ISSVM over the First Dataset Boldface indicates the best result.

Machine learning model	Accuracy (%)	Execution time (s)
ISVM (QoS subset)	79.67	2.43
ISVM (QoA subset)	71.54	**1.89**
ISVM (QoD subset)	65	2.13
ISVM (all IFs)	78.82	3.47
ISSVM	**82**	2.67

TABLE III. QoE estimation performance of the ISSVM over the second Dataset Boldface indicates the best result.

Machine learning model	Accuracy (%)	Execution time (s)
ISVM (QoS subset)	74.33	42.50
ISVM (QoA subset)	56.5	47.89
ISVM (QoD subset)	62	46.41
ISVM (UP subset)	69	**40.78**
ISVM (UF subset)	78	42.28
ISVM (all IFs)	*76.89*	*57.07*
ISSVM	**79.87**	49.12

TABLE IV. Average execution time in second of ISSVM vs other models for two databases (best method in bold).

Machine learning model	Dataset 1	Dataset 2
SVM	14.98	159.32
ANN	12.31	124.01
RF	8.77	95.63
ADT	7.04	87.12
SSVM	12.12	147.84
ISSVM	**2.67**	**49.12**

technique includes an ensemble of ISVMs. These classifiers are described via convex optimization problems which can be easily solved applying a numerical procedure.

Moreover, as we can see, the ISVM-based classifiers, i.e. ISSVM outperforms the SVM-based classifiers, i.e. SVM and SSVM. This is because when executing using classical SVM models, complete data are provided in advance during the training period, and the test of the classifier is apart from the latter. The objective of the training procedure is to minimize the cost function as well as to find maximum margins between classes. However, this method is carried out in a single step on the whole dataset, which may result in classification errors due to the incorrect classification of examples.

When using ISVM, the margins are adjusted, and improperly classified examples may be included in the support vector set. As a result, each choice made for new data will be employed incrementally to update and enhance the ISVM classifier's preceding result. Furthermore, because the support vectors are assessed incrementally, multi-class ISVM gives a cleaner solution.

As it can be observed from Table IV, the ISSVM model has a faster training speed relative to other Batch classifiers and SVM-base classifiers (SVM and SSVM). That is because firstly, unlike SVM models where the complexity is equal to $O(n^3)$, where n represents the number of samples used for training, the incremental SVM complexity is $O(ns^2)$, where ns denotes the number of support vectors and $ns \leq n$. This can be explaining by the employing of the Woodbery formula to recalculate the gradient, β and γ, requiring matrix-vector multiplication and recursively updating of the matrix R (section 3.2), which has a dimension equal to the number of support vectors ns. As a result, the execution time required for updating R is quadratic in the number of support vectors. Secondly, the base-classifiers of our model are various ISVM models trained in parallel with different input subsets features (QoS, QoA, QoD, UP, UF). So, the trained time will be lower than training all the dataset with one ISVM.

5. Conclusion

We have presented a novel incremental stacked multi-class support vector machine model for the online assessment of the quality of experience video streaming service. The proposed approach is based on the combination of a pool of ISVMs, trained in parallel on different regions of the feature space, and another ISVM to construct a powerful fully incremental model.

We implemented rigorous experiments on several datasets, and we compared our proposed method against incremental base classifiers, ISVM, Batch single classifiers, and Batch ensemble learning classifiers. In terms of precision and execution speed.

Evaluation results demonstrate the superiority of the ISSVM approach. In fact, the proposed approach inherits the benefits of employing a multi-class SVM classifier, incremental learning process, and stacked generalization method.

In future work, we will add a convolutional neural network model to extract deep features which have discriminative power and lead to performance improvement.

6. ACKNOWLEDGMENT

We thank Michael Hutton of the öbex project for language editing support.

7. Conflicts of interest

Not applicable

References

[1] Cisco, "Cisco annual internet report (2018–2023)," Cisco Systems Inc., Tech. Rep. C11-741490-01, 2022.

[2] I.-T. R. P.10, "Vocabulary for performance and quality of service, amendment 2: New definitions for inclusion in recommendation itu-t p.10/g.100," 2008.

[3] Y. Wang, W. Zhou, and P. Zhang, *QoE Management in Wireless Networks*. Springer International Publishing, 2017.

[4] R. ITU-T, "Itu-t rec. p.800.1: Mean opin- ion score (mos) terminology," 2003.

[5] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electronics Letters*, vol. 44, no. 13, p. 800, 2008.

[6] A. Tanchenko, "Visual-PSNR measure of image quality," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 874–878, Jul. 2014.

[7] R. Serral-Gracià, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro, and X. Masip-Bruin, "An overview of quality of experience measurement challenges for video applications in IP networks," in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2010, pp. 252–263.

[8] S. Winkler, *Digital video quality*. Chichester, England: John Wiley & Sons, Jan. 2005.

[9] V. Menkovski, A. Oredope, A. Liotta, and A. C. Sánchez, "Predicting quality of experience in multimedia streaming," in *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia - MoMM 09*. ACM Press, 2009.

[10] "Survey on machine learning-based QoE-QoS correlation models," in *2014 International Conference on Computing, Management and Telecommunications (ComManTel)*. IEEE, Apr. 2014.

[11] W. Kaiyu, W. Yumei, and Z. Lin, "A new three-layer QoE modeling method for HTTP video streaming over wireless networks," in *2014 4th IEEE International Conference on Network Infrastructure and Digital Content*. IEEE, Sep. 2014.

[12] A. Hameed, R. Dai, and B. Balas, "A Decision-Tree-Based perceptual video quality prediction model and its application in FEC for wireless multimedia communications," *IEEE Transactions on Multimedia*, vol. 18, no. 4, pp. 764–774, Apr. 2016.

[13] L. Qian, H. Chen, and L. Xie, "SVM-based QoE estimation model for video streaming service over wireless networks," in *2015 International Conference on Wireless Communications & Signal Processing (WCSP)*. IEEE, Oct. 2015.

[14] S. Ickin, M. Fiedler, and K. Vandikas, "Customized video QoE estimation with algorithm-agnostic transfer learning," Mar. 2020.

[15] H. Zhang, L. Dong, G. Gao, H. Hu, Y. Wen, and K. Guan, "DeepQoE: A multimodal learning framework for video quality of experience (QoE) prediction," *IEEE Transactions on Multimedia*, vol. 22, no. 12, pp. 3210–3223, Dec. 2020.

[16] K. Lavanya, K. V. Devi, and M. Subramaniam, "Quality of experience (QOE) content aware hybrid lean predictive models for medical video transmission over internet of things (IOT) networks," *International Journal of Communication Systems*, Jun. 2021.

[17] M. Ghosh, D. C. Singhal, and R. Wayal, "DeSVQ: Deep learning based streaming video QoE estimation," in *23rd International Conference on Distributed Computing and Networking*, ser. ICDCN 2022. New York, NY, USA: Association for Computing Machinery, 2022, pp. 19–25.

[18] V. Menkovski, G. Exarchakos, and A. Liotta, "Online QoE prediction," in *2010 Second International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, Jun. 2010.

[19] Y. B. Youssef, M. Afif, R. Ksantini, and S. Tabbane, "A novel online QoE prediction model based on multiclass incremental support vector machine," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*. IEEE, May 2018.

[20] T. Hastie and R. Tibshirani, "Classification by pairwise coupling," *The Annals of Statistics*, vol. 26, no. 2, Apr. 1998.

[21] R. Elwerghemmi, M. Heni, R. Ksantini, and R. Bouallegue, "Online QoE prediction model based on stacked multiclass incremental support vector machine," in *2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO)*. IEEE, Apr. 2019.

[22] H. Schulzrinne, A. Rao, and R. Lanphier, "Real time streaming protocol (rtsp)," 1998, https://www.ietf.org/rfc/rfc2326.txt.

[23] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "Rtp: A transport protocol for real-time applications," 1996, https://www.rfc-editor.org/rfc/pdfrfc/rfc1889.txt.pdf.

[24] C. G. Bampis and A. C. Bovik, "Learning to predict streaming video QoE: Distortions, rebuffering and memory," *arXiv [cs.MM]*, Mar. 2017.

[25] Y. Sani, A. Mauthe, and C. Edwards, "Adaptive bitrate selection: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2985–3014, 2017.

[26] K. Brunnström, K. De Moor, A. Dooms, S. Egger-Lampl, M.-N. Garcia, T. Hossfeld, S. Jumisko-Pyykkö, C. Keimel, C. Larabi, B. Lawlor, P. Le Callet, S. Möller, F. Pereira, M. Pereira, A. Perkis, A. Pinheiro, U. Reiter, P. Reichl, R. Schatz, and A. Zgank, *Qualinet White Paper on Definitions of Quality of Experience*, 2013.

[27] V. N. Vapnik, *Statistical learning theory*, ser. Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Commun ications and Control. Nashville, TN: John Wiley & Sons, Sep. 1998.

[28] P. Laskov, C. Gehl, S. Krueger, and K.-R. Müller, "Incremental support vector learning: Analysis, implementation and applications," *Journal of Machine Learning Research (JMLR)*, vol. 7, 2006.

[29] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992.

[30] L. Amour, S. Sami, S. Hoceini, and A. Mellouk, "An open source platform for perceived video quality evaluation," in *Proceedings of the 11th ACM Symposium on QoS and Security for Wireless and Mobile Networks*. ACM, Nov. 2015.

[31] C. Alippi and M. Roveri, "Virtual k-fold cross validation: An effective method for accuracy assessment," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*. IEEE, Jul. 2010.

[32] L. Breiman, *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[33] Y. Freund and R. E. Schapire, "A Decision-Theoretic generalization of On-Line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, Aug. 1997.

[34] T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *IEEE Communications Surveys and Tutorials*, vol. 10, pp. 56–76, 2008.

[35] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang, "Developing a predictive model of quality of experience for internet video," in *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*. ACM, Aug. 2013.

Radhia Elwerghemmi received the M.Sc. degree in computer science and multimedia from the Higher Institute of Computing and Multimedia of Gabes, Gabes University, Tunisia, in 2015. She is currently pursuing the Ph.D. degree with the research laboratory Innov'COM, Sup'Com, where she has worked on QoE prediction for video streaming services.

Maher Heni received the national engineering diploma in telecommunication from the National Engineering School of Tunis (ENIT), the master diploma in Telecommunications from ENIT with collaboration of IEF institute in parissud university, and the Ph.D. degree in Telecommunications in 2013 from the Higher School of Communication (Sup'COM), Tunisia. His research interests are ad hoc networks, mobile communication systems, and Big Data.

Riadh Ksantini received the M.Sc. and Ph.D. degrees in Computer Science from the Université de Sherbrooke, Sherbrooke, QC, Canada, in 2003 and 2007, respectively. From 2001 to 2007, he was a graduate research associate with Bell Canada Laboratories and the research center MOIVRE (MOdElisationen Imagerie, Vision et REseaux de neurones). Presently, he is Associate Professor at the Department of Computer Science, College of IT, University of Bahrain, Adjunct Associate Professor at the School of Computer Science, within the Faculty of Science of the University of Windsor, Windsor, Ontario, Canada, and Adjunct Professor at the Department of Computer Science, Université du Québec à Montréal (UQAM). Prior to that, he was Postdoctoral Research Associate in the Ecole de Technologie Superieure, University du Quebec, Montreal, Canada. He has also served as Visiting Fellow Research Scientist at the Canadian Space agency, and Postdoctoral Research Associate in the School of Computer Science, within the Faculty of Science of the University of Windsor. In 2008, he was awarded a fellowship (of excellence) for postdoctoral research from the granting agency ¨Fonds quebecois de la recherche sur la nature et les technologies ¨ (FQRNT). His PhD was evaluated and ranked third Ph.D. in Quebec for 2007 by the committee of Information Technology and Communications. His research interests include Artificial Intelligence, Machine/Deep Learning, Pattern Recognition and Computer Vision. His research work on Artificial Intelligence and Data Science has always involved collaboration between academia and industry. More than 70 Articles stemming from his research work have been published in several prestigious journals and conferences.

Ridha Bouallegue (Member, IEEE) received the M.S., Ph.D., and H.D.R. degrees in telecommunications from the National Engineering School of Tunis (ENIT), Tunisia, in 1990, 1994, and 2003, respectively. He is currently a Professor with ENIT and the Director of the research laboratory Innov'COM/Sup'Com. His current research interests include mobile and satellite communications, access techniques, intelligent signal processing, code division multiple access (CDMA), multi-in multi-out (MIMO), OFDM, and UWB systems..