Technical Report

Remote Data Mirroring
Logistic Regression Results

July, 2022

Version: 0.9

1 Remote Data Mirroring

The self-adaptive Remote Data Mirroring (RDM) network [1] is based on the operational model presented in [2]. The RDM network under consideration consists of 25 RDM mirrors (servers), to hold multiple copies of data, with 300 physical links in total that can used to transfer data between the mirrors. Each network link has an associated operational cost[1] and a measurable throughput, latency and loss rate used to determine the reliability, performance and cost of the RDM system [3, 4]. The goal here is to satisfy the non-functional requirements (NFRs) of Minimization of Costs (MC), Maximization of Performance (MP)[2] and Maximization of Reliability (MR) under environmental uncertainty of link failures and varying ranges of bandwidth consumption [1]. For this purpose, the network is required to continuously take adaptive actions of switching between the topological configurations of Minimum Spanning Tree (MST) and Redundant Topology (RT) to maintain better levels of satisfaction of NFRs. Both the configurations offer a different impact on the NFRs' satisfaction. The topological configuration of RT provides a higher level reliability than MST topology but it has a negative impact on the satisfaction of the MC and MP as the cost of maintaining non-stop RT topology will be high and due to data redundancy, the performance can be reduced. On the other hand, MST topology supports the satisfaction of MC and MP by maintaining a minimum spanning tree for the network. The simulation tool of *RDMSim* [5] has been used to simulate the RDM network for experiments.

1.1 Experiment Scenarios

Experiments were performed by introducing different scenarios of the environmental conditions under which the RDM network works as follows:

Stable Scenario S_0 representing the RDM network performing under normal environmental conditions

Detrimental Scenario representing the RDM network performing under the dynamic environment when different disturbance levels (increased packet loss or network link failures) in the network environment are introduced. There are 6 scenarios of detrimental situations defined by the simulated environment of *RDMSim*:

Scenario S_1 **Unexpected Packet Loss during MST:** An unexpected data packet loss during the execution of MST generates a decrease in the reliability of the system. A MST topology connects all remote sites in the RDM network by the identification of a minimum spanning tree on the network of links among each remote site. Data packet loss may represent network link failures in the RDM system, which may be caused due to problems with the equipment such as failure of a switch or a router.

Scenario S_2 **Unexpected Packet Loss during RT:** Unexpected data packet loss during the execution of the RT Topology, are generating an unusual rate of data forwarding, which would increase cost in terms of bandwidth consumption and would decrease the system's performance. In case of RDM, the cost for inter-site links communication refers to the data sent over them. Therefore, during RT topology, that involves a bigger number of inter-site network links than MST, is more expensive. Costs increase as the number of network links increases and a decrease in the system's performance could also be expected.

Scenario S_3 **Simultaneous occurrence of Scenario** S_1 **and** S_2**:** It represents the situation where simultaneous occurrence of Scenarios S_1 and S_2 is observed.

Scenario S_4 **MST topology execution failures:** An unexpected data packet loss during the execution of MST generates a decrease in the reliability of the system and effects the operational cost and performance of the system.

[1] In RDM system, Operational Cost is measured in terms of inter-site network traffic. [3]

[2] In case of RDM network, performance is measured in terms of total time to write data i.e. the sum of the time to write each copy of data on each remote site. [3]

Scenario S_5 RT topology execution failures: Unexpected data packet loss during the execution of the RT Topology, are generating an unusual rate of data forwarding, which would increase cost in terms of bandwidth consumption and would decrease the system's performance along with affecting reliability of the network.

Scenario S_6 Significant Site Failure: It represents the situation where simultaneous occurrence of Scenarios S_4 and S_5 is observed.

The remaining report is organized as follows. In Section 2, an overview of the Logistic Regression technique followed by the experiments for evaluation of Multi-Reward Partially Observable Markov Decision Process Plus Plus (MR-POMDP++) model are provided.

2 Logistic Regression

Logistic Regression is a binary classification technique [6]. It is a generalization of linear regression model to the binary classification setting. Both linear regression and logistic regression fall in the category of supervised learning models as they both make use of labeled data to make predictions. As opposed to linear regression, which is used to predict continuous values, logistic regression is used to predict discrete values. The model takes as an input a number of independent variables and the output is a discrete valued variable. For example, in an Internet of Things (IoT) network, given the traffic load and interference on the network links, we want to predict whether it satisfies the NFR of minimization of energy consumption (MEC) for the network. Here, the traffic load and link interference are the independent variables and the MEC satisfaction is a discrete valued dependent variable having a binary value. The MEC Satisfaction can have a value of True when it meets a required satisfaction threshold and False otherwise.

Hence, Logistic regression is considered as an extension of linear regression where the dependent variable can only have k number of classes. When k=2, it is known as binary logistic regression and for k>2, it is known as multiple or multinomial logistic regression. For a formal definition, we can use the following notation:

Notation:

Let $x^{(i)}$ denotes the "input independent variables" (input features) and $y^{(i)}$ denotes the "output dependent variable" that we want to predict. We denote a pair $(x^{(i)}, y^{(i)})$ to represent a training example and the list of m training examples is called a training set denoted as $(x^{(i)}, y^{(i)}); i = 1, ..., m$. Hence, X represents the input space and Y represents the output space.

The formal definition for the typical regression problem is as follows:

***Formal Definition**: Given a training set, the goal is to learn a hypothesis function h which is a good predictor of the value of y given the corresponding input x. Formally defined as:*

$$h : X \to Y$$

The hypothesis function h is typically represented as a linear function of x to approximate y as follows:

$$h(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + ... + \theta_n x_n \tag{1}$$

where $x_1 - x_n$ represent the input features and the θ_i's are the parameters (also called weights) parameterizing the space of linear function's mapping from X to Y. Futhremore, to simplify the notation, the convention of letting $x_0 = 1$ (the *intercept* term) is introduced, so that:

$$h(x) = \sum_{i=0}^{n} \theta_i x_i = \theta^T x \tag{2}$$

The hypothesis function presented in equation 2 gives us a continuous real value whereas the output variable y in logistic regression is discrete valued. Hence, for this purpose, the functions like Logistic Sigmoid function or Softmax function are used to modify the existing hypothesis function as follows:

Logistic Function:

The logistic function [7] is a model of the sigmoid function representing an S shaped curve that can be mathematically represented as follows:

$$P(y) = \frac{1}{1 + e^{-z}} \tag{3}$$

The goal is to compute the probability that an observed set of input features x belongs to class y=1, using the Sigmoid function. In this case, our z parameter, is the linear hypothesis function *h(x)*. We will denote this new hypothesis function as g(z):

$$g(z) = P(y = 1) = \frac{1}{1 + e^{-(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + ... + \theta_n x_n)}} \tag{4}$$

The parameters θ values, are selected to maximize the likelihood of predicting the probability of belonging to a particular class. If probability for observations is greater that 0.5, the observation will belong to class 1 and vice versa.

Softmax Function:

The Softmax function [7], as opposed to Sigmoid Function, gives a vector of values for each output class y. Given an observation of input features x, it gives a vector of values to represent the probability of belonging to each output class. The sum of all the probability values in the vector is equal to 1. Mathematically, the Softmax function is computed as follows:

$$P(y^{(i)} = k | x^{(i)}; \theta) = \frac{exp(\theta^T x)}{\sum_{j=1}^{k} exp(\theta^T x)} \tag{5}$$

As Softmax function provides us with the membership probability for each class, Softmax function has been used as part of the logistic regression classification presented in this report.

Learning parameters θs:

Given a training set, the learning of θs can be done by making the result of hypothesis function close to y. Formally, a cost function known as the *least square cost function* is defined to measure how close $g(z^{(i)})$ is to the corresponding $y^{(i)}$ as follows:

$$J(\theta) = \frac{1}{2} \sum_{i=1}^{m} \left(g_\theta(z^{(i)}) - y^{(i)}\right)^2 \tag{6}$$

The θ parameters are chosen so that the cost function $J(\theta)$ is minimized. For the purpose of optimization of θ, the Stochastic Gradient Descent (SGD) algorithm [8] is used.

The experimental evaluations for all the scenarios are presented in the following subsections:

2.1 Stable Scenario S_0

Parameter Tuning

For the purpose of tuning of the parameters of the model, the Stochastic Gradient Descent (SGD) algorithm is used. The SGD algorithm is executed for 20000 iterations with different learning rates for the results computed for each NFR separately as shown in Figs. 1, 2 and 3 respectively. On the basis of loss and the accuracy scores presented in Table 1, the learning rate of 0.05 has been selected for the model to evaluate the test set.

Table 1: Learning Rate Accuracy Score under Scenario 0 (Cross Validation Set)

Learning Rate	Accuracy Score MC	Accuracy Score MR	Accuracy Score MP
0.0001	0.9323308270676691	0.9924812030075187	0.9147869674185464
0.0003	0.9699248120300752	0.9899749373433584	0.9573934837092731
0.0005	0.9774436090225563	0.9899749373433584	0.9674185463659147
0.0007	0.9799498746867168	0.9899749373433584	0.9674185463659147
0.0009	0.9849624060150376	0.9899749373433584	0.9699248120300752
0.001	0.9849624060150376	0.9899749373433584	0.9699248120300752
0.003	0.9924812030075187	0.9899749373433584	0.9924812030075187
0.005	0.9924812030075187	0.9899749373433584	0.9974937343358395
0.007	0.9924812030075187	0.9899749373433584	0.9974937343358395
0.009	0.9949874686716792	0.9899749373433584	0.9974937343358395
0.01	0.9949874686716792	0.9899749373433584	0.9974937343358395
0.03	1.0	1.0	1.0
0.05	1.0	1.0	1.0
0.07	1.0	1.0	1.0
0.09	1.0	1.0	1.0

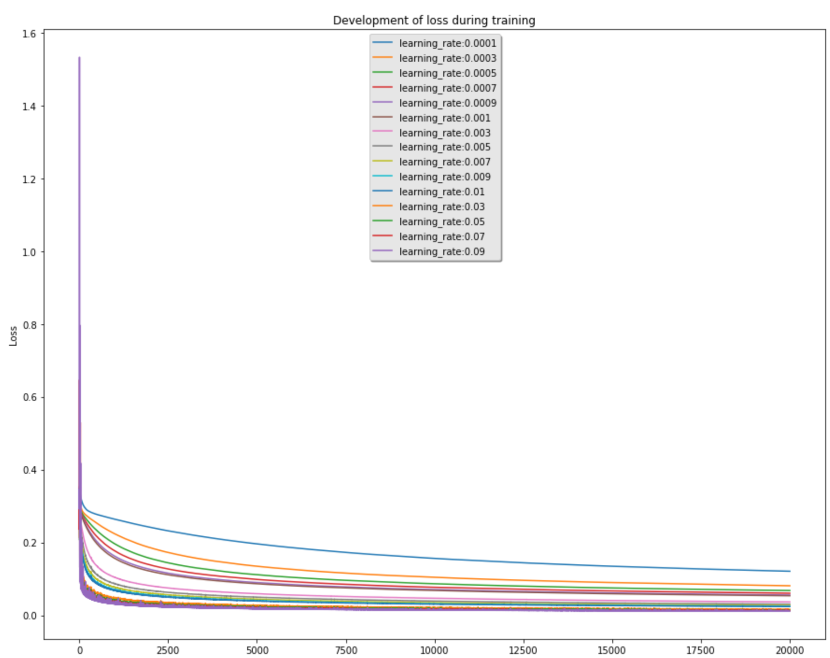


Figure 1: Learning Rate Optimization MC under stable scenario

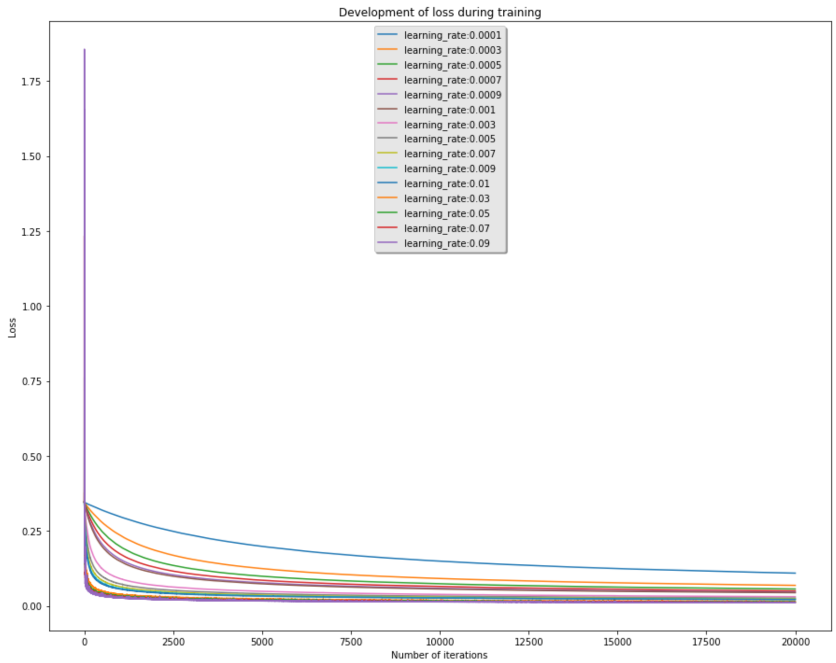


Figure 2: Learning Rate Optimization MR under stable scenario

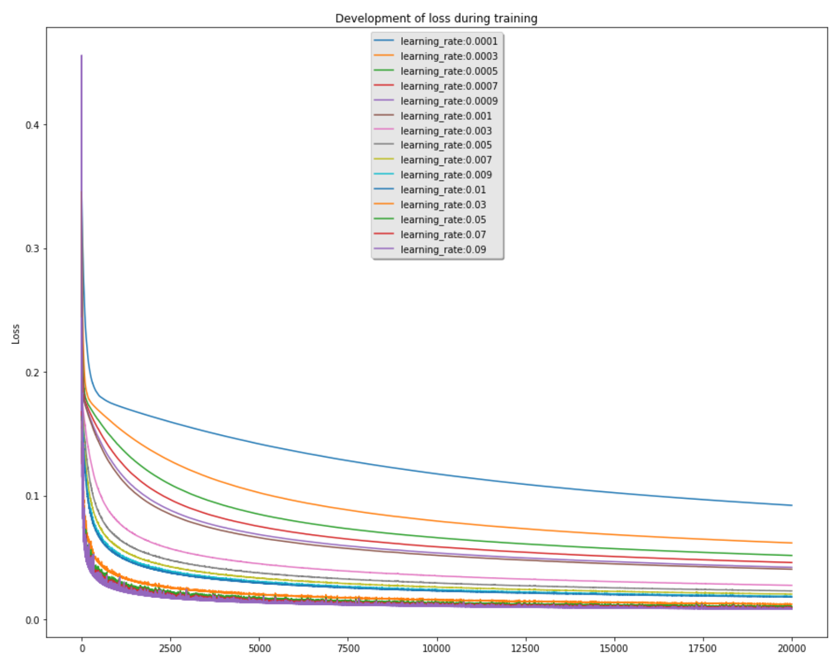


Figure 3: Learning Rate Optimization MP under stable scenario

Classification Results

The classification results for the NFRs using logistic regression under stable scenario are presented in Tables 2, 3 and 4. The results show an accuracy score of 0.994 for MC, 0.998 for MR and 0.966 for MP with a precision of 0.9921, 0.996 and 1.0 respectively. The F1 scores for MC, MR and MP are 0.996, 0.998 and 0.9808 respectively.

Table 2: Classification Results for Minimization of Operational Cost for time steps 487-493 under stable scenario

Time Step	**Bandwidth Consumed**	**Satisfaction Probability**	**Actual Satisfaction State**	**Predicted Satisfaction State**	**Prediction Probability for Satisfaction**
487	2142	0.878681003	True	True	[2.47829847e-09 9.99999998e-01]
488	4785	0.902529311	False	False	[9.99999709e-01 2.90762955e-07]
489	3260	0.816721721	True	True	[0.00659603 0.99340397]
490	3818	0.86974671	False	False	[0.90983172 0.09016828]
491	1728	0.875243403	True	True	[1.05236353e-11 1.00000000e+00]
492	567	0.903931281	True	True	[2.28132604e-18 1.00000000e+00]
493	1540	0.906253481	True	True	[8.60496423e-13 1.00000000e+00]

Table 3: Classification Results for Maximization of Reliability for time steps 487-493 under stable scenario

Time Step	**Active Links**	**Satisfaction Probability**	**Actual Satisfaction State**	**Predicted Satisfaction State**	**Prediction Probability for Satisfaction**
487	102	0.890476277	False	False	[0.71150418 0.28849582]
488	165	0.89421477	True	True	[3.79750627e-11 1.00000000e+00]
489	163	0.942541195	True	True	[9.94040136e-11 1.00000000e+00]
490	166	0.896645408	True	True	[2.57943671e-11 1.00000000e+00]
491	64	0.892796789	False	False	[9.9999988e-01 1.1996716e-07]
492	27	0.838290613	False	False	[1.00000000e+00 6.44833236e-14]
493	77	0.835466106	False	False	[9.99974896e-01 2.51039101e-05]

Table 4: Classification Results for Maximization of Performance for time steps 487-493 under stable scenario

Time Step	**Writing Time**	**Satisfaction Probability**	**Actual Satisfaction State**	**Predicted Satisfaction State**	**Prediction Probability for Satisfaction**
487	1224	0.831446787	True	True	[7.17404749e-11 1.00000000e+00]
488	1815	0.910477271	True	True	[1.44611620e-06 9.99998554e-01]
489	1956	0.847364808	True	True	[1.82596644e-05 9.99981740e-01]
490	1826	0.90915029	True	True	[1.74927234e-06 9.99998251e-01]
491	960	0.913417868	True	True	[6.70801965e-13 1.00000000e+00]
492	459	0.913903067	True	True	[1.30750763e-16 1.00000000e+00]
493	1155	0.91532629	True	True	[1.85678821e-11 1.00000000e+00]

Classification Results for MC

For the 500 simulation time steps, under stable scenario, the model correctly classifies the satisfaction state of MC with an exception of 3 simulation time steps as shown in Fig. 4. For these 3 simulation time steps, the actual satisfaction state for MC was False but it was predicted as True.

Moreover, based on the results presented in Table 2, we can deduce the extent to which MC can be considered as satisfied, when both the bandwidth consumption and satisfaction probability are considered. For example, at time step 488 the bandwidth consumption is 4785 GBps and the satisfaction probability is 0.902529311. The satisfaction state predicted by the model is False, similar to the actual state, with 9.99999709e-01 probability of being False. It means that given the input values, there is almost 99 percent chance of MC not being satisfied i.e. having satisfaction state as False.

Classification Results for MR

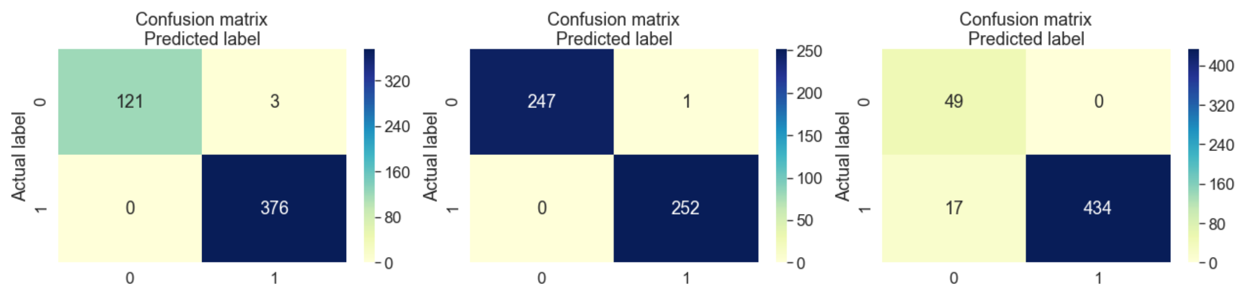
For the 500 simulation time steps, the model correctly classifies the satisfaction state of MR with an exception of only 1 simulation time step as shown in Fig. 4. For this 1 time step, the actual satisfaction state for MR was False but it was predicted as True.

Moreover, based on the results presented in Table 3, we can deduce the extent of the satisfaction state of MR, when both the active links and satisfaction probability are considered. For example, at time step 487, the active links are 102 and the satisfaction probability is 0.890476277. The output satisfaction state predicted by the model is False similar to the actual state with 0.71150418 probability of being False. It means that given the input values, there is around 71 percent chance of MR not being satisfied i.e. having satisfaction state as False.

Classification Results for MP

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MP with an exception of 17 simulation time steps as shown in Fig. 4. For these 17 time steps, the actual satisfaction state for MP was True but it was predicted as False.

Moreover, based on the results presented in Table 4, we can deduce the extent of the satisfaction state of MP, when both the writing time and satisfaction probability are considered. For example, at time step 487, the writing time is 1224 ms and the satisfaction probability is 0.831446787. The output satisfaction state predicted by the model is True similar to the actual state with the 1.00000000e+00 probability of being True. It means that given the input values, there is 100 percent chance of MP to be satisfied i.e. having satisfaction state as True.



(a) Minimization of Operational Cost (b) Maximization of Reliability (c) Maximization of Performance

Figure 4: Confusion Matrix for Classification of Satisfaction State of NFRs under stable scenario

2.2 Scenario S_1

Parameter Tuning

The SGD algorithm is executed for 20,000 iterations with different learning rates for the results computed for each NFR separately as shown in Figs. 5, 6 and 7 respectively. On the basis of loss and the accuracy scores presented in Table 5, the learning rate of 0.03 is selected for the model to evaluate the test set.

Table 5: Learning Rate Accuracy Score under Scenario S_1 (Cross Validation Set)

Learning Rate	Accuracy Score MC	Accuracy Score MR	Accuracy Score MP
0.0001	0.9624060150375939	0.9849624060150376	0.9398496240601504
0.0003	0.9799498746867168	0.9974937343358395	0.9674185463659147
0.0005	0.9849624060150376	0.9974937343358395	0.974937343358396
0.0007	0.9849624060150376	0.9974937343358395	0.9849624060150376
0.0009	0.9874686716791979	0.9974937343358395	0.9874686716791979
0.001	0.9874686716791979	0.9974937343358395	0.9874686716791979
0.003	0.9924812030075187	0.9974937343358395	0.9949874686716792
0.005	0.9974937343358395	0.9974937343358395	0.9949874686716792
0.007	0.9974937343358395	0.9974937343358395	0.9974937343358395
0.009	0.9974937343358395	0.9974937343358395	0.9974937343358395
0.01	0.9974937343358395	1.0	0.9974937343358395
0.03	1.0	1.0	1.0
0.05	0.9974937343358395	1.0	1.0
0.07	0.9974937343358395	1.0	1.0
0.09	0.9974937343358395	1.0	1.0

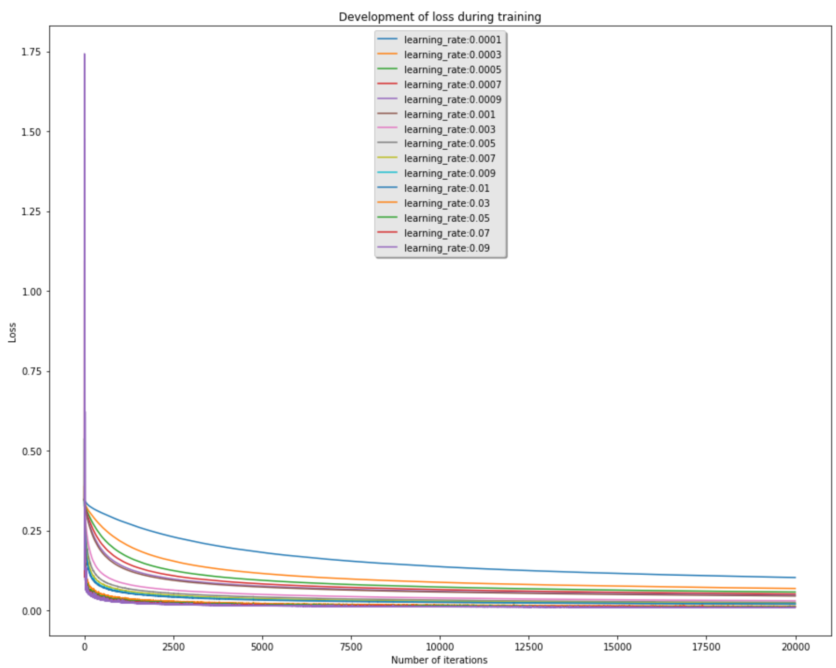


Figure 5: Learning Rate Optimization for MC under Scenario S_1

Classification Results

The classification results for the NFRs using logistic regression are presented in Tables 6, 7 and 8. The results show an accuracy score of 0.986 for MC, 0.998 for MR and 1.0 for MP with a precision of 1.0 for MC and MP, and 0.9967 for MR. The F1 scores for MC, MR and MP are 0.9889, 0.9983 and 1.0 respectively.

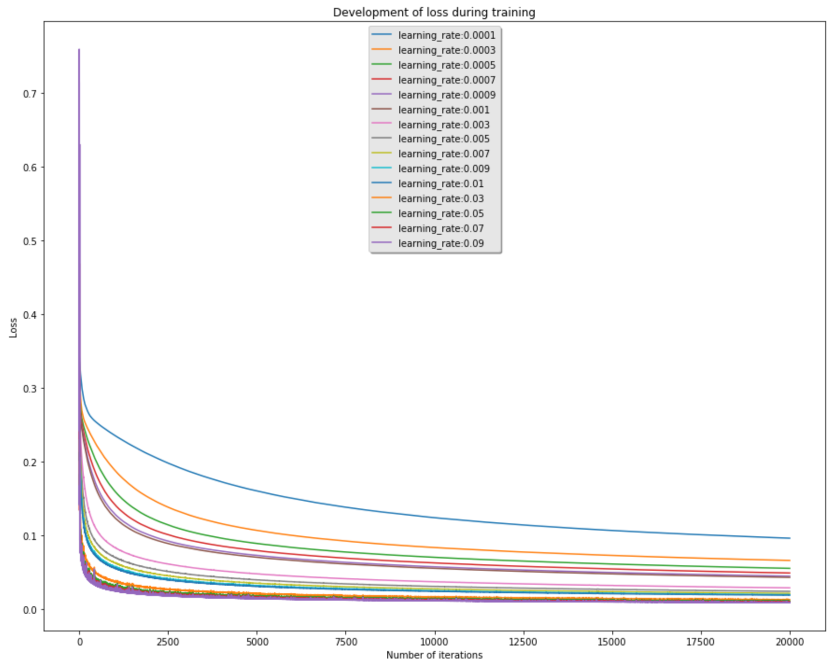


Figure 6: Learning Rate Optimization for MR under Scenario S_1

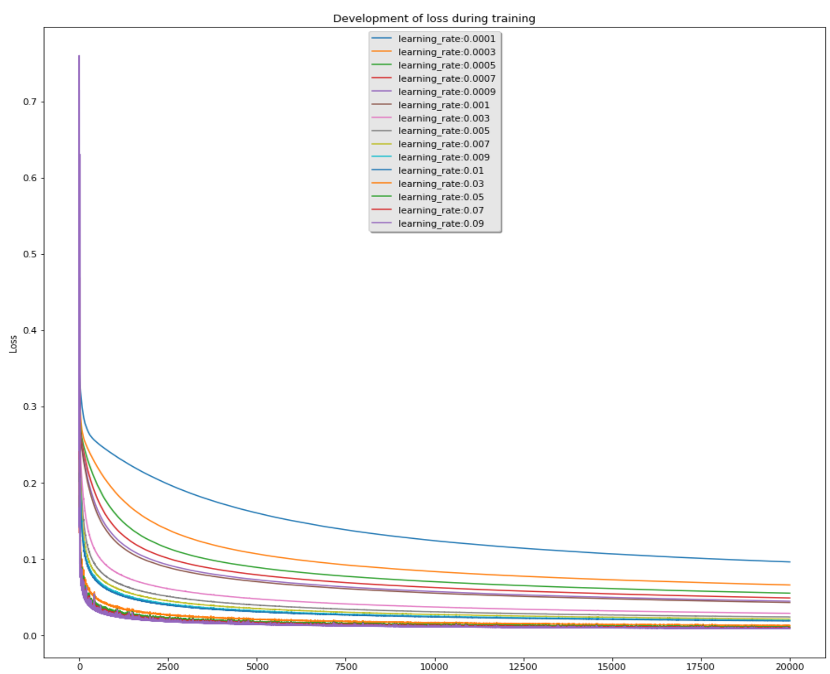


Figure 7: Learning Rate Optimization for MP under Scenario S_1

Classification Results for MC

For the 500 simulation time steps, under Scenario S_1 of *RDMSim*, the model correctly classifies the satisfaction state of MC with an exception of 7 simulation time steps as shown in Fig. 8. For these 7 simulation time steps, the actual satisfaction state for MC was True but it was predicted as False.

Moreover, based on the results presented in Table 6, we can deduce the extent to which MC can be considered as satisfied, when both the bandwidth consumption and satisfaction probability are considered. For example, at time step 482, the bandwidth consumption is 5550 GBps and the satisfaction probability is 0.725491276. The satisfaction state predicted by the model is also False with the 1.00000000e+00 probability of being False. It means that given the input values, there is 100 percent chance of MC not being satisfied i.e. having satisfaction state as False.

Classification Results for MR

For all of the 500 simulation time steps comprising the test set, the model correctly classifies the satisfaction state of MR with an exception of only 1 simulation time step as shown in Fig. 8. For this 1 simulation time step, the actual satisfaction state for MR was False but it was predicted as True.

Moreover, based on the results presented in Table 7, we can deduce the extent of the satisfaction state of MR, when both the active links and satisfaction probability are considered. For example, at time step 482, the active links are 222 and the satisfaction probability is 0.958693166. The output satisfaction state predicted by the model is True similar to the actual state with the 1.000e+00 probability of being True. It means that given the input values, there is 100 percent chance of MR being satisfied i.e. having satisfaction state as True.

Classification Results for MP

For all of the 500 simulation time steps in the test set, the model correctly classifies the satisfaction state of MP as shown in Fig. 8.

Moreover, based on the results presented in Table 8, we can deduce the extent of the satisfaction state of MP, when both the writing time and satisfaction probability are considered. For example, at time step 482, the writing time is 2220 ms and the satisfaction probability is 0.716539246. The output satisfaction state predicted by the model is True similar to the actual state with the 0.99878437 probability of being True. It means that given the input values, there is around 99 percent chance of MP to be satisfied i.e. having satisfaction state as True.

Table 6: Classification Results for MC for time steps 482-488 under Scenario S_1

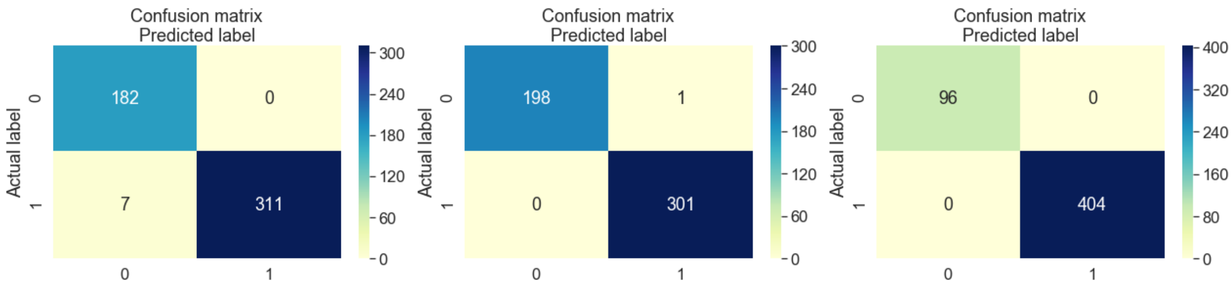
Step	Bandwidth Consumed	SatProb	Actual State	Predicted State	Probability per class
482	5550	0.725491276	False	False	[1.00000000e+00 4.28385292e-10]
483	4833	0.783454523	False	False	[9.99998958e-01 1.04154805e-06]
484	4725	0.791813864	False	False	[9.99996635e-01 3.36520499e-06]
485	2625	0.864214984	True	True	[5.38856619e-05 9.99946114e-01]
486	1944	0.902410103	True	True	[3.51550671e-08 9.99999965e-01]
487	5175	0.867215377	False	False	[9.9999996e-01 4.0101029e-08]
488	3525	0.873227415	True	True	[0.40579247 0.59420753]

Table 7: Classification Results for MR for time steps 482-488 under Scenario S_1

Step	Active Links	SatProb	Actual State	Predicted State	Probability per class
482	222	0.958693166	True	True	[2.3037888e-17 1.0000000e+00]
483	179	0.964496786	True	True	[2.57926305e-11 1.00000000e+00]
484	225	0.952357296	True	True	[8.70731399e-18 1.00000000e+00]
485	125	0.83784989	True	True	[9.89801317e-04 9.99010199e-01]
486	81	0.766044144	False	False	[9.99337825e-01 6.62174864e-04]
487	225	0.902908016	True	True	[8.61891872e-18 1.00000000e+00]
488	141	0.83331521	True	True	[5.55843368e-06 9.99994442e-01]

Table 8: Classification Results for MP for time steps 482-488 under Scenario S_1

Step	Writing Time	SatProb	Actual State	Predicted State	Probability per class
482	2220	0.716539246	True	True	[0.00121563 0.99878437]
483	2148	0.699744551	True	True	[4.48927419e-04 9.99551073e-01]
484	2475	0.699741136	True	True	[0.04798941 0.95201059]
485	1375	0.80810347	True	True	[4.86431874e-09 9.99999995e-01]
486	1215	0.910140268	True	True	[3.73099626e-10 1.00000000e+00]
487	2925	0.849803269	False	False	[0.95807766 0.04192234]
488	2820	0.823122431	False	False	[0.84301066 0.15698934]



(a) Minimization of Operational Cost (b) Maximization of Reliability (c) Maximization of Performance

Figure 8: Confusion Matrix for Classification of Satisfaction State of NFRs under Scenario S_1

2.3 Scenario S_2

Parameter Tuning

The SGD algorithm has been executed for 20,000 iterations with different learning rates for the results computed for each NFR separately as shown in Figs. 9, 10 and 11 respectively. On the basis of loss and the accuracy scores presented in Table 9, the learning rate of 0.05 has been selected for the model to evaluate the Test set. For the learning rate of 0.05, the accuracy score for MC is 0.9979959919839679, MR is 0.9959919839679359 and MP is 0.9979959919839679.

Table 9: Learning Rate Accuracy Score under Scenario S_2 (Cross Validation Set)

Learning Rate	Accuracy Score MC	Accuracy Score MR	Accuracy Score MP
0.0001	0.935871743486974	0.9899799599198397	0.9198396793587175
0.0003	0.9839679358717435	0.9919839679358717	0.9338677354709419
0.0005	0.9919839679358717	0.9939879759519038	0.9438877755511023
0.0007	0.9939879759519038	0.9959919839679359	0.9559118236472945
0.0009	0.9959919839679359	0.9939879759519038	0.9679358717434869
0.001	0.9959919839679359	0.9939879759519038	0.9679358717434869
0.003	0.9959919839679359	0.9979959919839679	0.9779559118236473
0.005	0.9959919839679359	0.9979959919839679	0.9819639278557114
0.007	0.9959919839679359	0.9979959919839679	0.9859719438877755
0.009	0.9959919839679359	0.9979959919839679	0.9859719438877755
0.01	0.9959919839679359	0.9979959919839679	0.9879759519038076
0.03	0.9979959919839679	0.9959919839679359	0.9919839679358717
0.05	0.9979959919839679	0.9959919839679359	0.9979959919839679
0.07	0.9979959919839679	0.9959919839679359	0.9979959919839679
0.09	0.9979959919839679	0.9959919839679359	0.9979959919839679

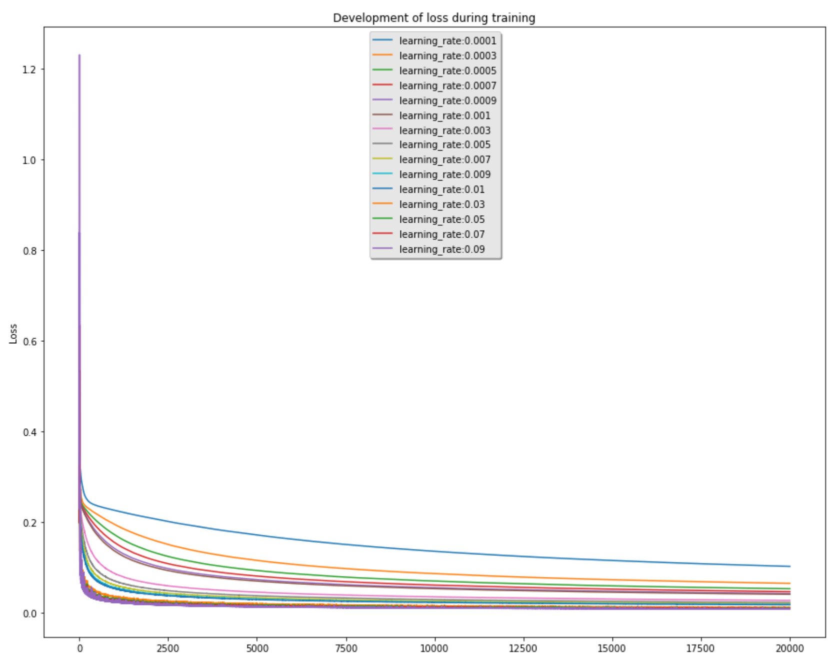


Figure 9: Learning Rate Optimization for MC under Scenario S_2

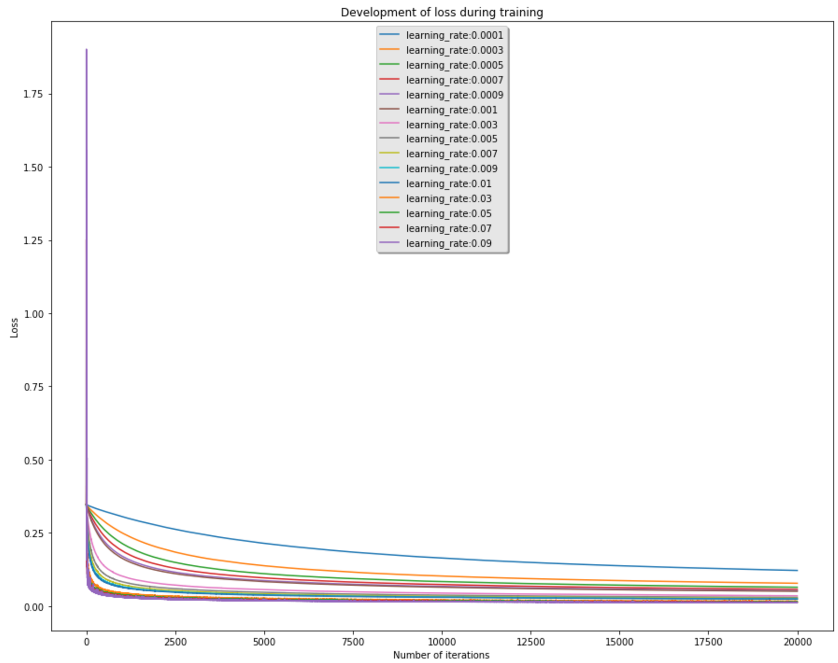


Figure 10: Learning Rate Optimization for MR under Scenario S_2

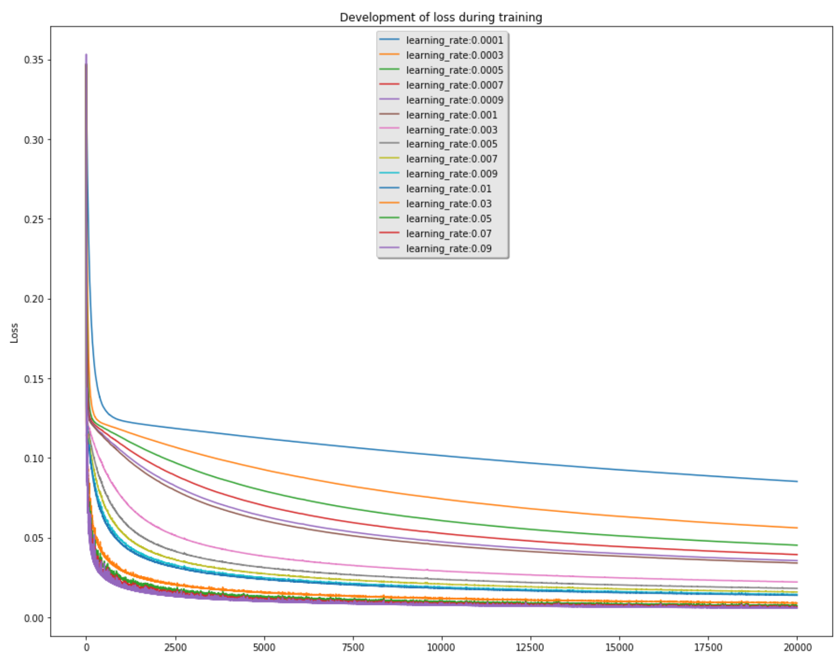


Figure 11: Learning Rate Optimization for MP under Scenario S_2

2.3.1 Classification Results

The classification results for the NFRs using logistic regression are presented in Tables 10, 11 and 12. The results show an accuracy score of 0.962 for MC, 0.972 for MR and 0.89 for MP with a precision of 0.9544 for MC and 1.0 for both MR and MP. The F1 scores for MC, MR and MP are 0.9767, 0.9708 and 0.9367 respectively.

Classification Results for MC

For the 500 simulation time steps, under Scenario S_2, the model correctly classifies the satisfaction state of MC with an exception of 19 simulation time steps as shown in Fig. 12. For these 19 simulation time steps, the actual satisfaction state for MC was False but it was predicted as True.

Moreover, based on the results presented in Table 10, we can deduce the extent to which MC can be considered as satisfied, when both the bandwidth consumption and satisfaction probability are considered. For example, at time step 491 the bandwidth consumption is 2592 GBps and the satisfaction probability is 0.904068481. The satisfaction state predicted by the model is True with 9.99999968e-01 probability of being True. It means that given the input values, there is almost 99 percent chance of MC to be satisfied i.e. having satisfaction state as True.

Classification Results for MR

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MR with an exception of 14 simulation time steps as shown in Fig. 12. For these 14 time steps, the actual satisfaction state for MR was True but it was predicted as False.

Moreover, based on the results presented in Table 11, we can deduce the extent of the satisfaction state of MR, when both the active links and satisfaction probability are considered. For example, at time step 488, the active links are 48 and the satisfaction probability is 0.835354651. The output satisfaction state predicted by the model is False similar to the actual state with the 1.00000000e+00 probability of being False. It means that given the input values, there is 100 percent chance of MR not being satisfied i.e. having satisfaction state as False.

Classification Results for MP

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MP with an exception of 55 simulation time steps as shown in Fig. 12. For these 55 time steps, the actual satisfaction state for MP was True but it was predicted as False.

Moreover, based on the results presented in Table 12, we can deduce the extent of the satisfaction state of MP, when both the writing time and satisfaction probability are considered. For example, at time step 489, the writing time is 1521 ms and the satisfaction probability is 0.915477939. The output satisfaction state predicted by the model is True similar to the actual state with 9.99998302e-01 probability of being True. It means that given the input values, there is around 99 percent chance of MP to be satisfied i.e. having satisfaction state as True.

Table 10: Classification Results for Minimization of Operational Cost for time steps 487-493 under Scenario S_2

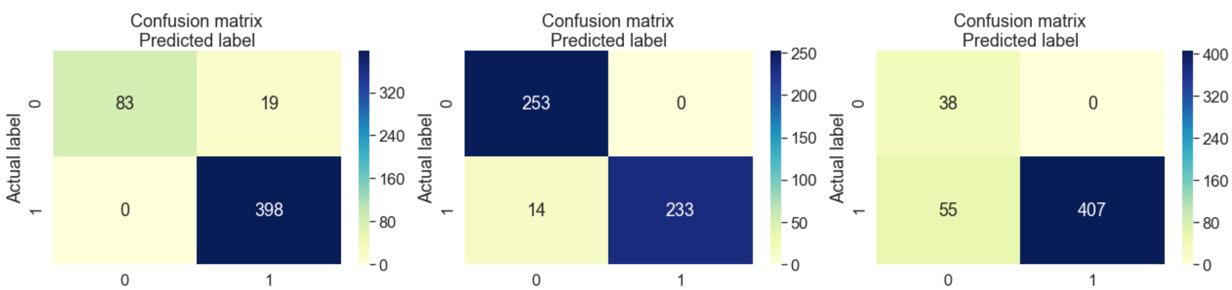
Time Step	Bandwidth Consumed	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
487	1265	0.905418947	True	True	[2.1350262e-16 1.0000000e+00]
488	1344	0.906354254	True	True	[6.3313197e-16 1.0000000e+00]
489	3393	0.906444334	True	True	[0.00252334 0.99747666]
490	2882	0.878908448	True	True	[4.64379415e-06 9.99995356e-01]
491	2592	0.904068481	True	True	[3.24650712e-08 9.99999968e-01]
492	4050	0.905254055	False	False	[0.96661554 0.03338446]
493	3608	0.878539014	False	True	[0.12075274 0.87924726]

Table 11: Classification Results for Maximization of Reliability for time steps 487-493 under Scenario S_2

Time Step	Active Links	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
487	55	0.837094713	False	False	[1.00000000e+00 1.70858955e-10]
488	48	0.835354651	False	False	[1.00000000e+00 8.94950487e-12]
489	117	0.83523595	True	True	[0.0245434 0.9754566]
490	131	0.890294742	True	True	[8.03355177e-05 9.99919664e-01]
491	96	0.892214353	False	False	[0.99526592 0.00473408]
492	135	0.837284883	True	True	[1.27068300e-05 9.99987293e-01]
493	164	0.890580591	True	True	[7.18767037e-11 1.00000000e+00]

Table 12: Classification Results for Maximization of Performance for time steps 487-493 under Scenario S_2

Time Step	Writing Time	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
487	990	0.914541562	True	True	[7.89528192e-11 1.00000000e+00]
488	528	0.915397479	True	True	[1.33639106e-14 1.00000000e+00]
489	1521	0.915477939	True	True	[1.69826685e-06 9.99998302e-01]
490	1310	0.91533598	True	True	[3.22201212e-08 9.99999968e-01]
491	1920	0.913995064	True	True	[0.00306633 0.99693367]
492	2430	0.914412491	True	False	[0.97809124 0.02190876]
493	3116	0.831266169	False	False	[9.99999954e-01 4.59557700e-08]



(a) Minimization of Operational Cost (b) Maximization of Reliability (c) Maximization of Performance

Figure 12: Confusion Matrix for Classification of Satisfaction State of NFRs under Scenario S_2

2.4 Scenario S_3

Parameter Tuning

The SGD algorithm has been executed for 20,000 iterations with different learning rates for the results computed for each NFR separately as shown in Figs. 13, 14 and 15 respectively. On the basis of loss and the accuracy scores presented in Table 13, the learning rate of 0.09 has been selected for the model to evaluate the test set.

Table 13: Learning Rate Accuracy Score under Scenario S_3 (Cross Validation Set)

Learning Rate	Accuracy Score MC	Accuracy Score MR	Accuracy Score MP
0.0001	0.9323308270676691	0.9649122807017544	0.9674185463659147
0.0003	0.9774436090225563	0.974937343358396	0.9674185463659147
0.0005	0.9899749373433584	0.9799498746867168	0.9699248120300752
0.0007	0.9899749373433584	0.9824561403508771	0.9699248120300752
0.0009	0.9899749373433584	0.9824561403508771	0.9774436090225563
0.001	0.9899749373433584	0.9849624060150376	0.9774436090225563
0.003	0.9924812030075187	0.9949874686716792	0.9849624060150376
0.005	0.9924812030075187	0.9949874686716792	0.9924812030075187
0.007	0.9924812030075187	0.9949874686716792	0.9924812030075187
0.009	0.9924812030075187	0.9949874686716792	0.9924812030075187
0.01	0.9924812030075187	0.9949874686716792	0.9924812030075187
0.03	0.9924812030075187	0.9924812030075187	0.9949874686716792
0.05	0.9924812030075187	0.9924812030075187	0.9949874686716792
0.07	0.9924812030075187	0.9924812030075187	0.9949874686716792
0.09	0.9949874686716792	0.9924812030075187	0.9949874686716792

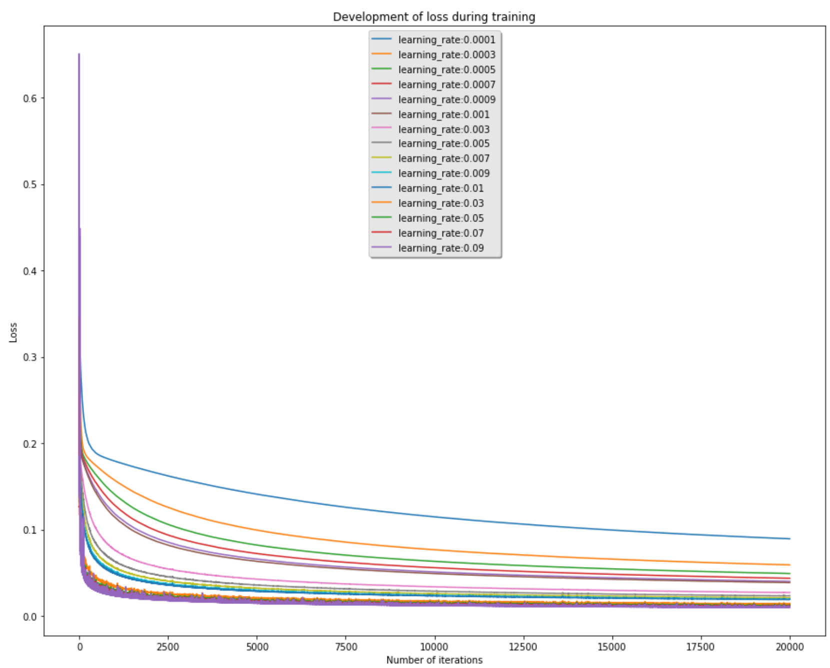


Figure 13: Learning Rate Optimization for MC under Scenario S_3

Classification Results

The classification results for the NFRs using logistic regression are presented in Tables 14, 15 and 16. The results show an accuracy score of 0.934 for MP and 1.0 for both MC and MR with a precision of 1.0 for all

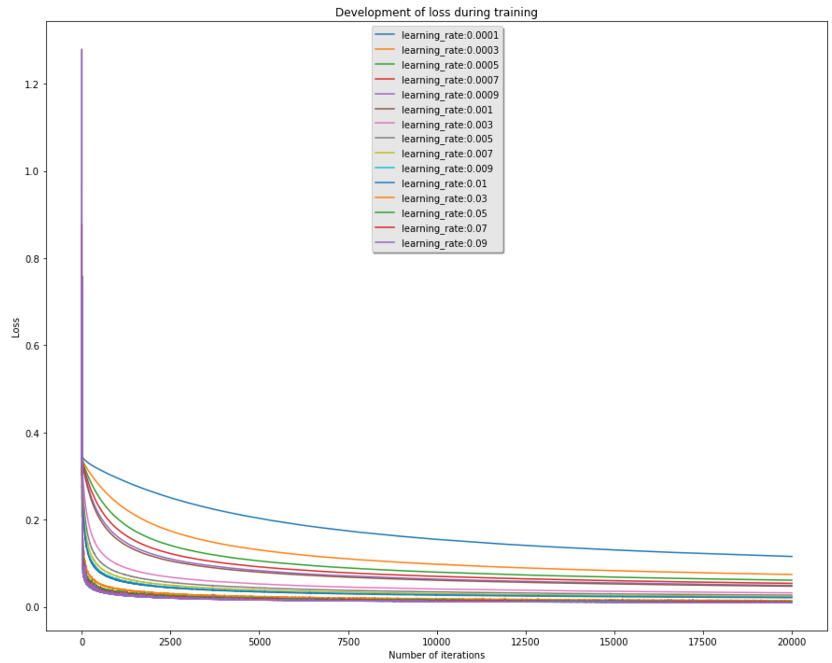


Figure 14: Learning Rate Optimization for MR under Scenario S_3

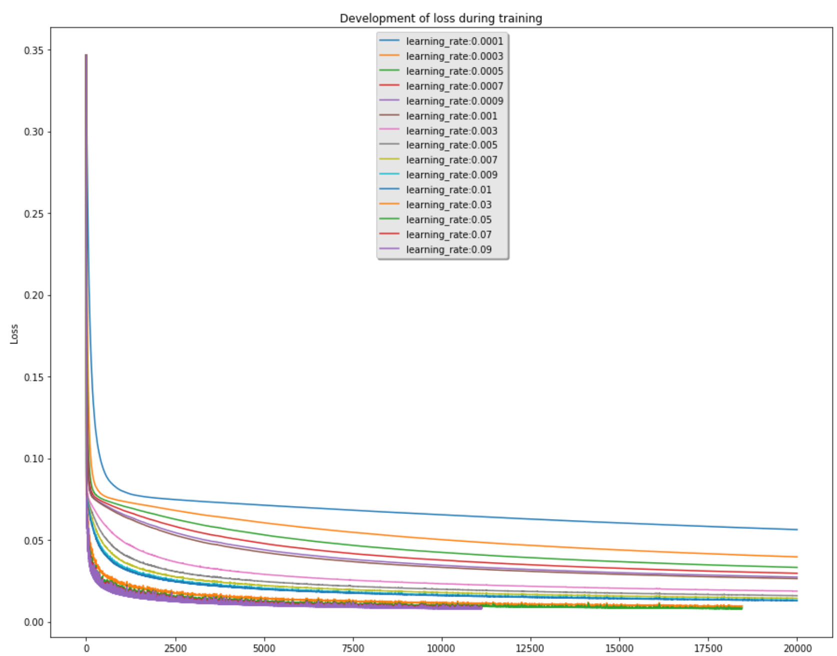


Figure 15: Learning Rate Optimization for MP under Scenario S_3

three NFRs. The F1 score for MC and MR is 1.0 and for MP is 0.9646.

Classification Results for MC

For the 500 simulation time steps, under Scenario S_3, the model correctly classifies the satisfaction state of MC as shown in Fig. 16.

Moreover, based on the results presented in Table 14, we can deduce the extent to which MC can be considered as satisfied, when both the bandwidth consumption and satisfaction probability are considered. For example, at time step 480, the bandwidth consumption is 3562 GBps and the satisfaction probability is 0.906331043. The satisfaction state predicted by the model is True, similar to the actual state, with 0.82924947 probability of being True. It means that given the input values, there is almost 82 percent chance of MC to be satisfied i.e. having satisfaction state as True.

Classification Results for MR

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MR as shown in Fig. 16.

Moreover, based on the results presented in Table 15, we can deduce the extent of the satisfaction state of MR, when both the active links and satisfaction probability are considered. For example, at time step 480, the active links are 137 and the satisfaction probability is 0.692851966. The output satisfaction state predicted by the model is True, similar to the actual state, with 9.99999996e-01 probability of being True. It means that given the input values, there is around 99 percent chance of MR not being satisfied i.e. having satisfaction state as True.

Classification Results for MP

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MP with an exception of 33 simulation time steps as shown in Fig. 16. For these 35 time steps, the actual satisfaction state for MP was True but it was predicted as False.

Moreover, based on the results presented in Table 16, we can deduce the extent of the satisfaction state of MP, when both the writing time and satisfaction probability are considered. For example, at time step 483, the writing time is 2540 ms and the satisfaction probability is 0.91293695. The output satisfaction state predicted by the model is False with 9.99999675e-01 probability of being False. It means that given the input values, there is around 99 percent chance of MP to be satisfied i.e. having satisfaction state as False. In contrast, the actual satisfaction state is True.

Table 14: Classification Results for Minimization of Operational Cost for time steps 480-486 under Scenario S_3

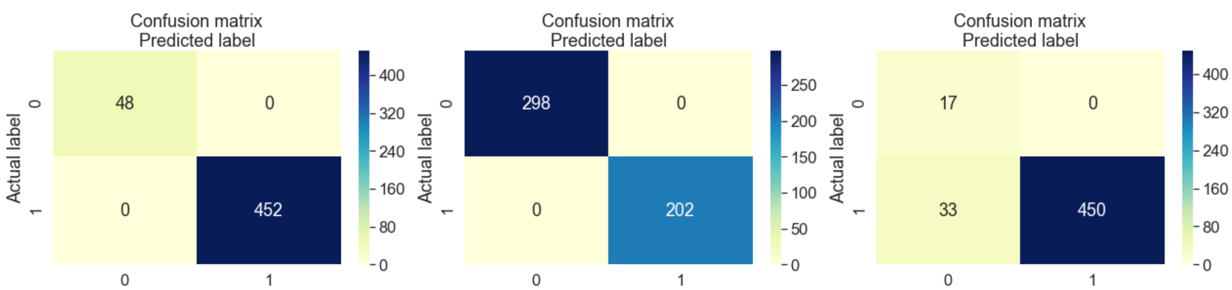
Time Step	Bandwidth Consumed	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
480	3562	0.906331043	True	True	[0.17075053 0.82924947]
481	2904	0.88220514	True	True	[4.88311634e-07 9.99999512e-01]
482	1848	0.90644988	True	True	[1.27956206e-16 1.00000000e+00]
483	3683	0.905914451	False	False	[0.71091073 0.28908927]
484	1050	0.882402351	True	True	[1.7353882e-23 1.0000000e+00]
485	3390	0.904734487	True	True	[0.00629665 0.99370335]
486	540	0.881979411	True	True	[5.23249897e-28 1.00000000e+00]

Table 15: Classification Results for Maximization of Reliability for time steps 480-486 under Scenario S_3

Time Step	Active Links	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
480	137	0.692851966	True	True	[3.93107583e-09 9.99999996e-01]
481	132	0.760228004	True	True	[6.72007666e-08 9.99999933e-01]
482	88	0.782844985	False	True	[9.99945392e-01 5.46076167e-05]
483	127	0.668750656	True	True	[1.67304856e-06 9.99998327e-01]
484	35	0.791035729	False	False	[1.0000000e+00 8.7760098e-19]
485	113	0.695257102	True	True	[0.00690514 0.99309486]
486	27	0.791429034	False	False	[1.00000000e+00 7.24970693e-21]

Table 16: Classification Results for Maximization of Performance for time steps 480-486 under Scenario S_3

Time Step	Writing Time	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
480	1507	0.916072125	True	True	[4.88541333e-09 9.99999995e-01]
481	2508	0.917745666	True	False	[0.9789962 0.0210038]
482	1232	0.833266282	True	True	[1.20363037e-11 1.00000000e+00]
483	2540	0.91293695	True	False	[0.98999413 0.01000587]
484	525	0.836301065	True	True	[1.05963072e-18 1.00000000e+00]
485	1582	0.912531818	True	True	[2.77046894e-08 9.99999972e-01]
486	324	0.917531265	True	True	[7.74977411e-21 1.00000000e+00]



(a) Minimization of Operational Cost (b) Maximization of Reliability (c) Maximization of Performance

Figure 16: Confusion Matrix for Classification of Satisfaction State of NFRs under Scenario S_3

2.5 Scenario S_4

Parameter Tuning

The SGD algorithm has been executed for 20,000 iterations with different learning rates for the results computed for each NFR separately as shown in Figs. 17, 18 and 19 respectively. On the basis of loss and the accuracy scores presented in Table 17, the learning rate of 0.05 has been selected for the model to evaluate the test set.

Table 17: Learning Rate Accuracy Score under Scenario S_4 (Cross Validation Set)

Learning Rate	Accuracy Score MC	Accuracy Score MR	Accuracy Score MP
0.0001	0.9724310776942355	0.9899749373433584	0.9423558897243107
0.0003	0.9799498746867168	0.9974937343358395	0.9774436090225563
0.0005	0.9899749373433584	0.9974937343358395	0.9849624060150376
0.0007	0.9899749373433584	0.9974937343358395	0.9849624060150376
0.0009	0.9899749373433584	0.9974937343358395	0.9874686716791979
0.001	0.9924812030075187	0.9974937343358395	0.9874686716791979
0.003	0.9949874686716792	0.9949874686716792	0.9899749373433584
0.005	0.9974937343358395	0.9949874686716792	0.9974937343358395
0.007	1.0	0.9974937343358395	0.9974937343358395
0.009	1.0	0.9974937343358395	0.9974937343358395
0.01	1.0	0.9974937343358395	0.9974937343358395
0.03	1.0	0.9974937343358395	0.9974937343358395
0.05	1.0	0.9974937343358395	0.9974937343358395
0.07	1.0	0.9974937343358395	0.9974937343358395
0.09	1.0	0.9974937343358395	0.9974937343358395

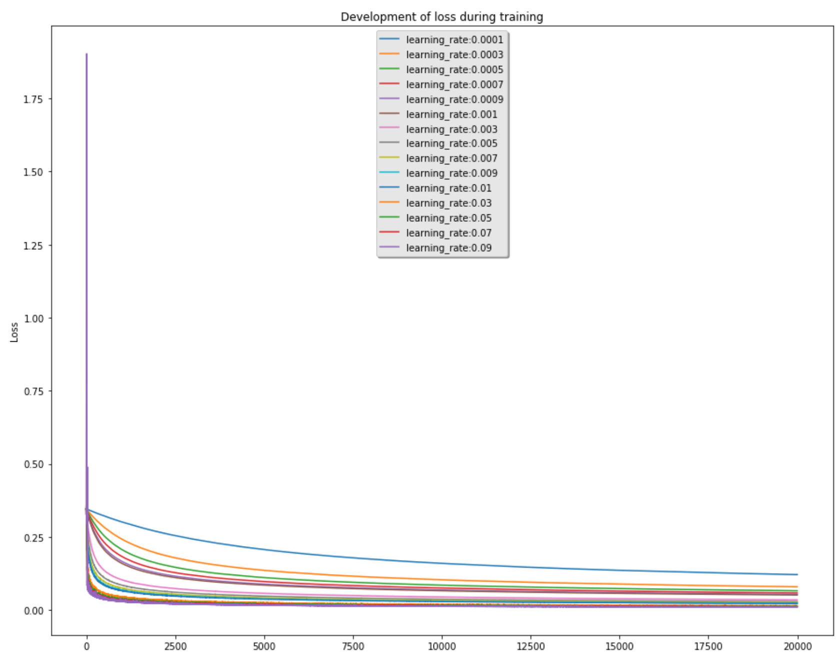


Figure 17: Learning Rate Optimization for MC under Scenario S_4

Classification Results

The classification results for the NFRs using logistic regression are presented in Tables 18, 19 and 20. The results show an accuracy score of 0.986 for MC, 0.998 for MR and 0.994 for MP with a precision of 1.0 for

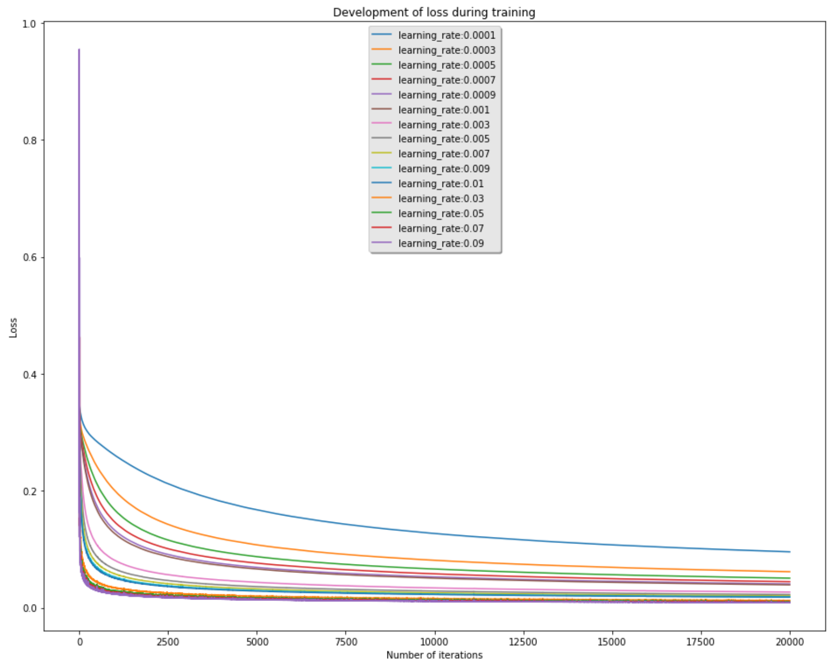


Figure 18: Learning Rate Optimization for MR under Scenario S_4

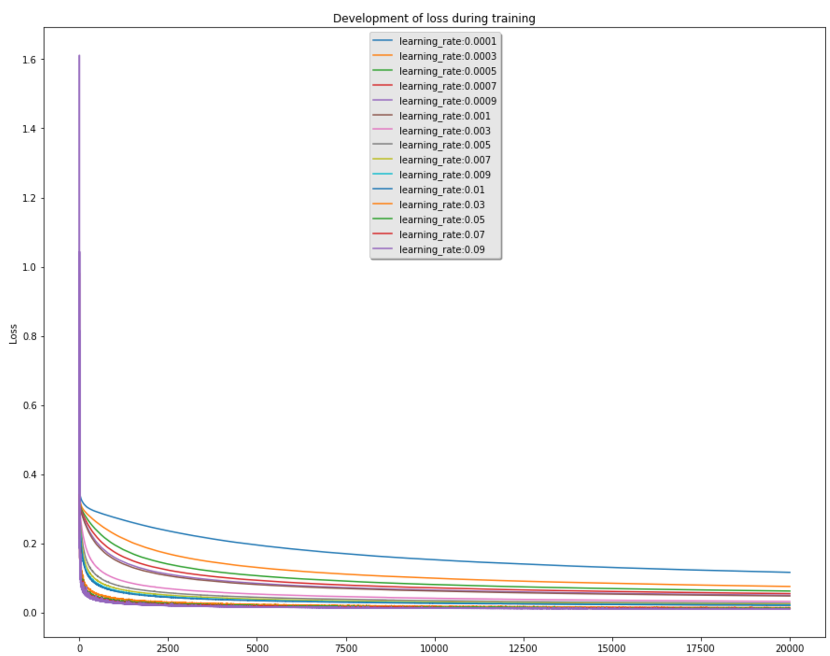


Figure 19: Learning Rate Optimization for MP under Scenario S_4

all three NFRs. The F1 scores for MC, MR and MP are 0.9861, 0.9984 and 0.9956 respectively.

Classification Results for MC

For the 500 simulation time steps, under Scenario S_4, the model correctly classifies the satisfaction state of MC with an exception of 7 simulation time steps as shown in Fig. 20. For these 7 simulation time steps, the actual satisfaction state for MC was True but it was predicted as False.

Moreover, based on the results presented in Table 18, we can deduce the extent to which MC can be considered as satisfied, when both the bandwidth consumption and satisfaction probability are considered. For example, at time step 482, the bandwidth consumption is 2317.8 GBps and the satisfaction probability is 0.829105937. The satisfaction state predicted by the model is True, similar to the actual state, with 9.99999969e-01 probability of being True. It means that given the input values, there is 99 percent chance of MC to be satisfied i.e. having satisfaction state as True.

Classification Results for MR

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MR with an exception of only 1 simulation time step as shown in Fig. 20. For this one time step, the actual satisfaction state for MR was True but it was predicted as False.

Moreover, based on the results presented in Table 19, we can deduce the extent of the satisfaction state of MR, when both the active links and satisfaction probability are considered. For example, at time step 482, the active links are 101 and the satisfaction probability is 0.707430978. The output satisfaction state predicted by the model is False, similar to the actual state, with 0.85522103 probability of being False. It means that given the input values, there is around 85 percent chance of MR not being satisfied i.e. having satisfaction state as False.

Classification Results for MP

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MP with an exception of 3 simulation time steps as shown in Fig. 20. For these 3 time steps, the actual satisfaction state for MP was True but it was predicted as False.

Moreover, based on the results presented in Table 20, we can deduce the extent of the satisfaction state of MP, when both the writing time and satisfaction probability are considered. For example, at time step 484, the writing time is 2295 ms and the satisfaction probability is 0.715493263. The output satisfaction state predicted by the model is True similar to the actual state with the 0.99883726 probability of being True. It means that given the input values, there is around 99 percent chance of MP to be satisfied i.e. having satisfaction state as True.

Table 18: Classification Results for Minimization of Operational Cost for time steps 482-488 under Scenario S_4

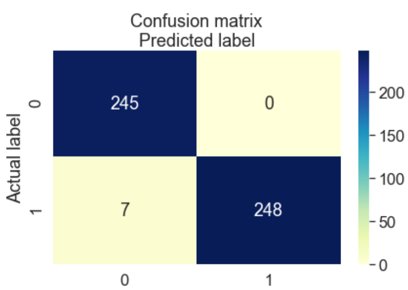
Time Step	Bandwidth Consumed	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
482	2317.812966	0.829105937	True	True	[3.05508203e-08 9.99999969e-01]
483	4420	0.858285042	False	False	[9.99992491e-01 7.50923106e-06]
484	3060	0.806895924	True	True	[8.82272771e-04 9.99117727e-01]
485	2391.931089	0.746101155	True	True	[8.45209605e-08 9.99999915e-01]
486	755.1363952	0.796469635	True	True	[1.22733952e-17 1.00000000e+00]
487	5771	0.854079338	False	False	[1.00000000e+00 5.67196746e-14]
488	4200	0.807092062	False	False	[9.99841367e-01 1.58632848e-04]

Table 19: Classification Results for Maximization of Reliability for time steps 482-488 under Scenario S_4

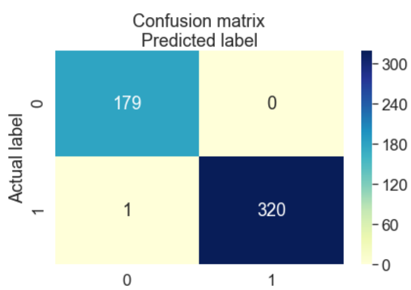
Time Step	Active Links	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
482	101	0.707430978	False	False	[0.85522103 0.14477897]
483	221	0.94123164	True	True	[2.90089793e-19 1.00000000e+00]
484	153	0.959312548	True	True	[2.44601351e-08 9.99999976e-01]
485	84	0.792773551	False	False	[9.99679771e-01 3.20229006e-04]
486	32	0.716732387	False	False	[1.00000000e+00 1.38381855e-12]
487	199	0.908432481	True	True	[1.00330684e-15 1.00000000e+00]
488	168	0.945876097	True	True	[9.53490399e-11 1.00000000e+00]

Table 20: Classification Results for Maximization of Performance for time steps 482-488 under Scenario S_4

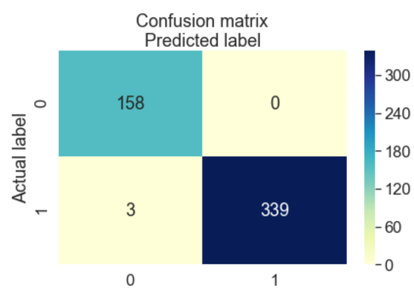
Time Step	Writing Time	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
482	1221.315878	0.844334395	True	True	[1.99731997e-12 1.00000000e+00]
483	4199	0.845519857	False	False	[1.00000000e+00 4.84985008e-13]
484	2295	0.715493263	True	True	[0.00116274 0.99883726]
485	916.1225406	0.662968754	True	True	[9.74539222e-15 1.00000000e+00]
486	422.4350357	0.811583637	True	True	[7.66597981e-19 1.00000000e+00]
487	3383	0.843091203	False	False	[9.99998160e-01 1.83951135e-06]
488	2352	0.716116614	True	True	[0.00333979 0.99666021]



(a) Minimization of Operational Cost



(b) Maximization of Reliability



(c) Maximization of Performance

Figure 20: Confusion Matrix for Classification of Satisfaction State of NFRs under Scenario S_4

2.6 Scenario S_5

Parameter Tuning

The SGD algorithm has been executed for 20,000 iterations with different learning rates for the results computed for each NFR separately as shown in Figs. 21, 22 and 23 respectively. On the basis of loss and the accuracy scores presented in Table 21, the learning rate of 0.3 has been selected for the model to evaluate the test set.

Table 21: Learning Rate Accuracy Score under Scenario S_5 (Cross Validation Set)

Learning Rate	Accuracy Score MC	Accuracy Score MR	Accuracy Score MP
0.0001	0.9273182957393483	0.9899749373433584	0.924812030075188
0.0003	0.9674185463659147	0.9899749373433584	0.9649122807017544
0.0005	0.974937343358396	0.9899749373433584	0.9849624060150376
0.0007	0.9824561403508771	0.9924812030075187	0.9874686716791979
0.0009	0.9824561403508771	0.9924812030075187	0.9899749373433584
0.001	0.9799498746867168	0.9924812030075187	0.9899749373433584
0.003	0.974937343358396	0.9924812030075187	0.9974937343358395
0.005	0.9774436090225563	0.9924812030075187	0.9974937343358395
0.007	0.9774436090225563	0.9974937343358395	1.0
0.009	0.9774436090225563	0.9974937343358395	1.0
0.01	0.9774436090225563	0.9949874686716792	1.0
0.03	0.9849624060150376	0.9949874686716792	0.9974937343358395
0.05	0.9849624060150376	0.9949874686716792	0.9974937343358395
0.07	0.9899749373433584	0.9924812030075187	0.9974937343358395
0.09	0.9949874686716792	0.9924812030075187	0.9974937343358395
0.1	0.9949874686716792	0.9924812030075187	0.9974937343358395
0.3	0.9974937343358395	0.9924812030075187	0.9974937343358395
0.5	0.9724310776942355	0.9949874686716792	0.9974937343358395
0.7	0.9849624060150376	0.9799498746867168	0.9974937343358395
0.9	0.9849624060150376	0.9974937343358395	0.9974937343358395

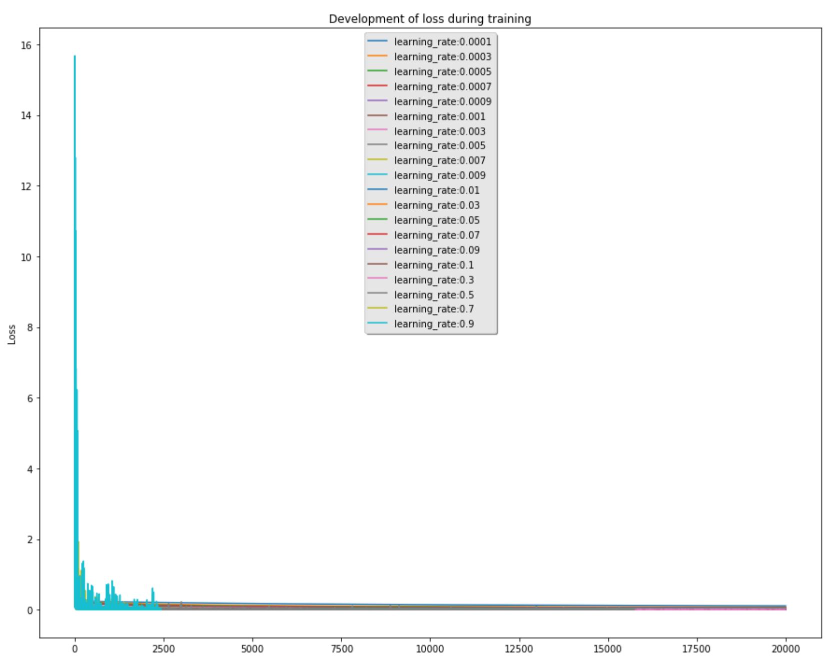


Figure 21: Learning Rate Optimization for MC under Scenario S_5

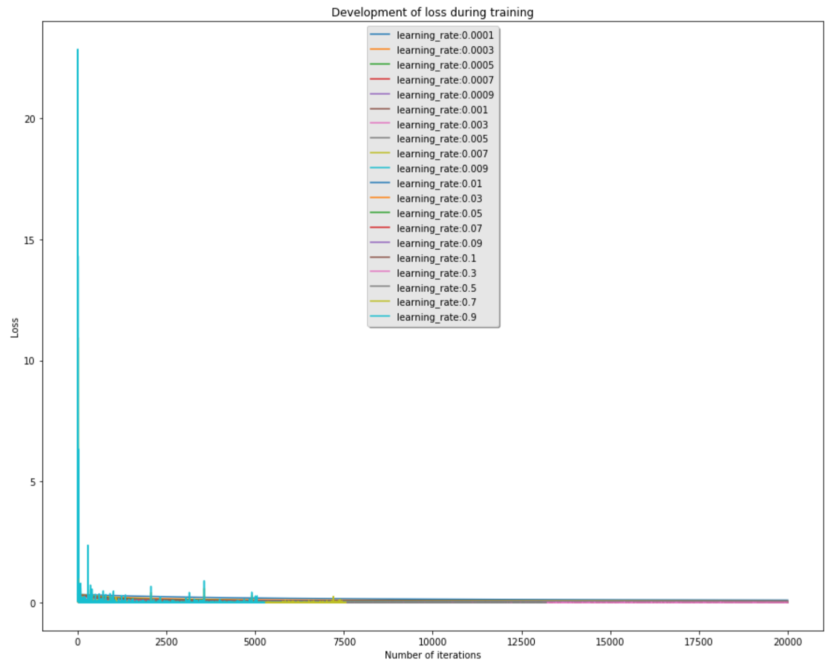


Figure 22: Learning Rate Optimization for MR under Scenario S_5

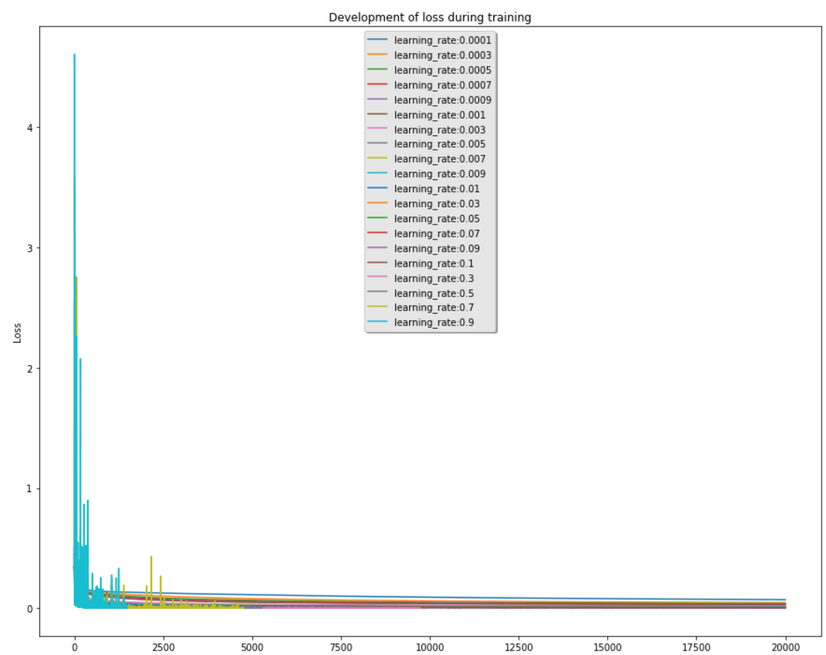


Figure 23: Learning Rate Optimization for MP under Scenario S_5

Classification Results

The classification results for the NFRs using logistic regression are presented in Tables 22, 23 and 24. The results show an accuracy score of 0.908 for MC, 0.902 for MR and 0.998 for MP with a precision of 1.0, 0.8172 and 0.9978 respectively. The F1 scores for MC, MR and MP are 0.9413, 0.8993 and 0.9989 respectively.

Classification Results for MC

For the 500 simulation time steps, under Scenario S_5, the model correctly classifies the satisfaction state of MC with an exception of 46 simulation time steps as shown in Fig. 24. For these 46 simulation time steps, the actual satisfaction state for MC was True but it was predicted as False.

Moreover, based on the results presented in Table 22, we can deduce the extent to which MC can be considered as satisfied, when both the bandwidth consumption and satisfaction probability are considered. For example, at time step 478 the bandwidth consumption is 3614 GBps and the satisfaction probability is 0.878204047. The satisfaction state predicted by the model is False, similar to the actual satisfaction state, with 9.99998964e-01 probability of being False. It means that given the input values, there is almost 99 percent chance of MC not being satisfied i.e. having satisfaction state as False.

Classification Results for MR

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MR with an exception of 49 simulation time steps as shown in Fig. 24. For these 49 time steps, the actual satisfaction state for MR was False but it was predicted as True.

Moreover, based on the results presented in Table 23, we can deduce the extent of the satisfaction state of MR, when both the active links and satisfaction probability are considered. For example, at time step 475, the active links are 152 and the satisfaction probability is 0.89023413. The output satisfaction state predicted by the model is True similar to the actual state with 1.00000000e+00 probability of being True. It means that given the input values, there is 100 percent chance of MR being satisfied i.e. having satisfaction state as True.

Classification Results for MP

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MP with an exception of only 1 simulation time step as shown in Fig 24. For this 1 time step, the actual satisfaction state for MP was False but it was predicted as True.

Moreover, based on the results presented in Table 24, we can deduce the extent of the satisfaction state of MP, when both the Writing Time and satisfaction probability are considered. For example, at time step 475, the writing time is 2280 ms and the satisfaction probability is 0.915390173. The output satisfaction state predicted by the model is True similar to the actual state with 9.99996228e-01 probability of being True. It means that given the input values, there is around 99 percent chance of MP to be satisfied i.e. having satisfaction state as True.

Table 22: Classification Results for Minimization of Operational Cost for time steps 475-481 under Scenario S_5

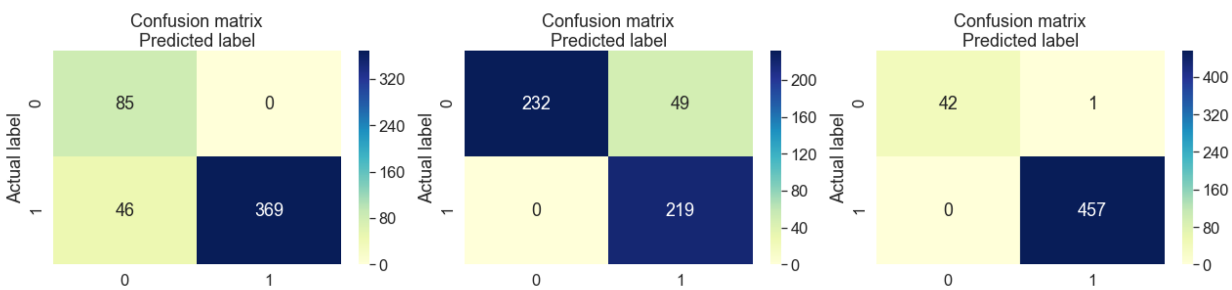
Time Step	**Bandwidth Consumed**	**Satisfaction Probability**	**Actual Satisfaction State**	**Predicted Satisfaction State**	**Prediction Probability for Satisfaction**
475	4408	0.906345948	False	False	[1.00000000e+00 1.06386093e-18]
476	648	0.877414365	True	True	[9.41670411e-40 1.00000000e+00]
477	3120	0.902542864	True	True	[0.02647624 0.97352376]
478	3614	0.878204047	False	False	[9.99998964e-01 1.03583654e-06]
479	3718	0.87377789	False	False	[9.99999973e-01 2.67485686e-08]
480	1365	0.873550311	True	True	[7.30562341e-29 1.00000000e+00]
481	1350	0.903755478	True	True	[3.71562161e-29 1.00000000e+00]

Table 23: Classification Results for Maximization of Reliability for time steps 475-481 under Scenario S_5

Time Step	Active Links	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
475	152	0.89023413	True	True	[1.94269255e-28 1.00000000e+00]
476	24	0.891779606	False	False	[1.00000000e+00 2.86201691e-31]
477	156	0.840988967	True	True	[2.63620486e-30 1.00000000e+00]
478	139	0.890831364	True	True	[1.60450105e-22 1.00000000e+00]
479	169	0.894566022	True	True	[3.59544574e-36 1.00000000e+00]
480	65	0.894716076	False	False	[1.00000000e+00 1.29770827e-12]
481	45	0.838531021	False	False	[1.00000000e+00 1.16065869e-21]

Table 24: Classification Results for Maximization of Performance for time steps 475-481 under Scenario S_5

Time Step	Writing Time	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
475	2280	0.915390173	True	True	[3.77221991e-06 9.99996228e-01]
476	360	0.829764862	True	True	[2.33771684e-31 1.00000000e+00]
477	2808	0.910490776	False	False	[0.96052477 0.03947523]
478	2780	0.830773311	False	False	[0.82733802 0.17266198]
479	1690	0.82250093	True	True	[3.49623337e-14 1.00000000e+00]
480	715	0.909877203	True	True	[2.02564887e-26 1.00000000e+00]
481	720	0.913696145	True	True	[2.44175868e-26 1.00000000e+00]



(a) Minimization of Operational Cost (b) Maximization of Reliability (c) Maximization of Performance

Figure 24: Confusion Matrix for Classification of Satisfaction State of NFRs under Scenario S_5

2.7 Scenario S_6

Parameter Tuning

The SGD algorithm has been executed for 20,000 iterations with different learning rates for the results computed for each NFR separately as shown in Figs. 25, 26 and 27 respectively. On the basis of loss and the accuracy scores presented in Table 25, the learning rate of 0.05 has been selected for the model to evaluate the test set.

Table 25: Learning Rate Accuracy Score under Scenario S_6 (Cross Validation Set)

Learning Rate	Accuracy Score MC	Accuracy Score MR	Accuracy Score MP
0.0001	0.9273182957393483	0.9824561403508771	0.9674185463659147
0.0003	0.9649122807017544	0.9924812030075187	0.974937343358396
0.0005	0.9649122807017544	0.9924812030075187	0.9799498746867168
0.0007	0.9674185463659147	0.9974937343358395	0.9849624060150376
0.0009	0.9699248120300752	0.9974937343358395	0.9874686716791979
0.001	0.9699248120300752	0.9974937343358395	0.9924812030075187
0.003	0.9799498746867168	1.0	0.9949874686716792
0.005	0.9824561403508771	1.0	0.9974937343358395
0.007	0.9849624060150376	1.0	0.9974937343358395
0.009	0.9899749373433584	1.0	1.0
0.01	0.9899749373433584	1.0	1.0
0.03	0.9924812030075187	1.0	1.0
0.05	0.9924812030075187	0.9974937343358395	1.0
0.07	0.9924812030075187	0.9974937343358395	0.9974937343358395
0.09	0.9924812030075187	0.9974937343358395	0.9949874686716792

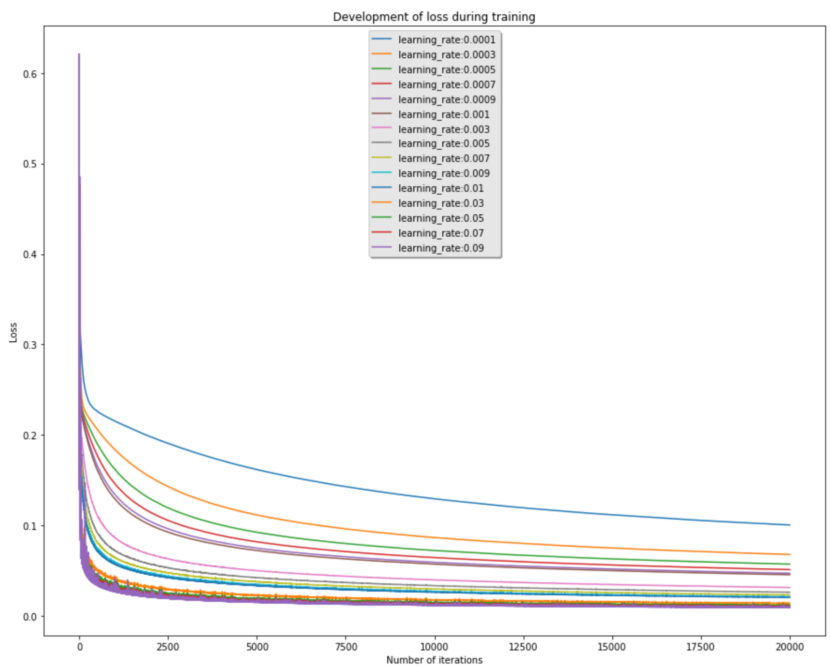


Figure 25: Learning Rate Optimization for MC under Scenario S_6

Classification Results

The classification results for the NFRs using logistic regression are presented in Tables 26, 27 and 28. The results show an accuracy score of 0.852 for MC, 0.976 for MR and 0.994 for MP with a precision of 0.848,

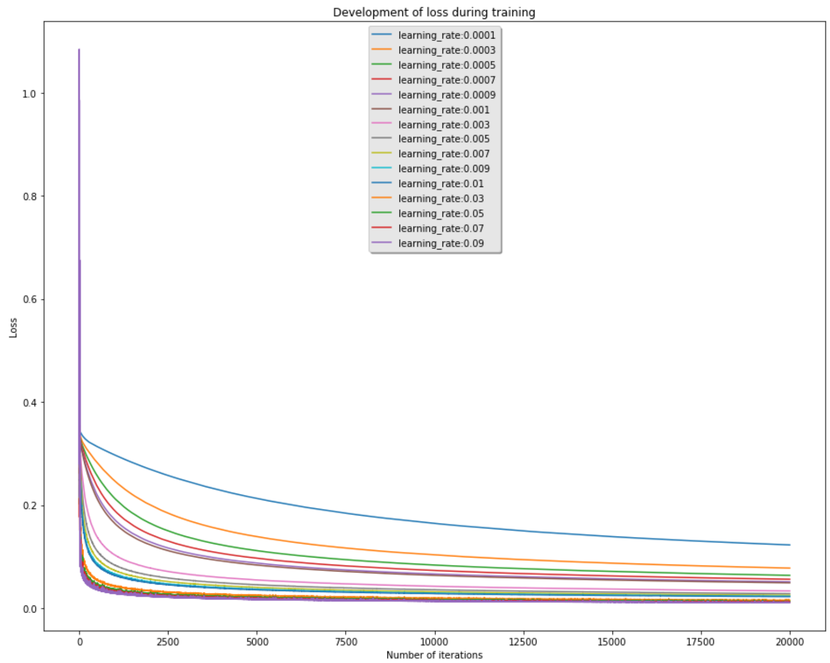


Figure 26: Learning Rate Optimization for MR under Scenario S_6

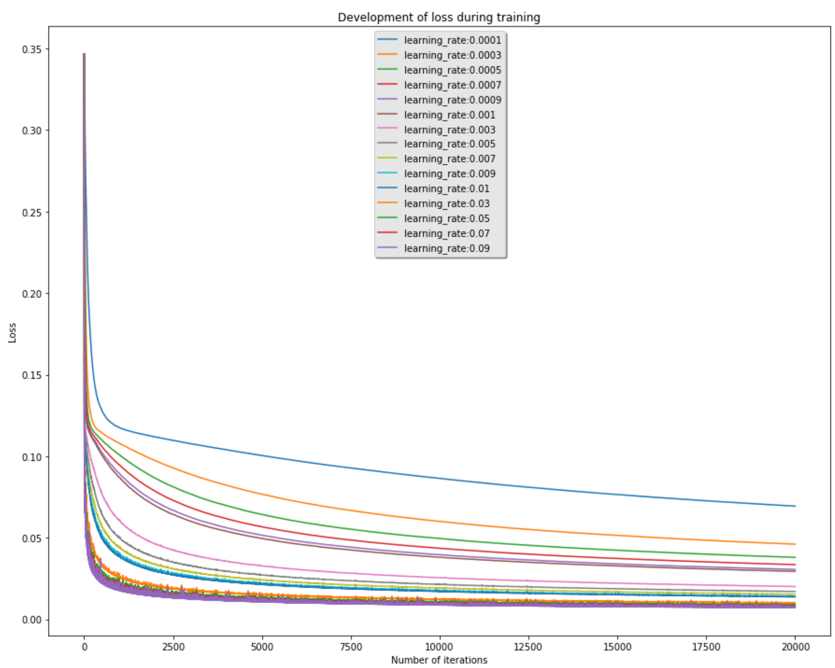


Figure 27: Learning Rate Optimization for MP under Scenario S_6

0.9409 and 0.9937 respectively. The F1 scores for MC, MR and MP are 0.9177, 0.9695 and 0.9968 respectively.

Classification Results for MC

For the 500 simulation time steps, under Scenario S_6, the model correctly classifies the satisfaction state of MC with an exception of 74 simulation time steps as shown in Fig. 28. For these 74 simulation time steps, the actual satisfaction state for MC was False but it was predicted as True.

Moreover, based on the results presented in Table 26, we can deduce the extent to which MC can be considered as satisfied, when both the bandwidth consumption and satisfaction probability are considered. For example, at time step 482, the bandwidth consumption is 2811.18 GBps and the satisfaction probability is 0.800333179. The satisfaction state predicted by the model is True, similar to the actual state, with 1.00000000e+00 probability of being True. It means that given the input values, there is 100 percent chance of MC being satisfied i.e. having satisfaction state as True.

Classification Results for MR

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MR with an exception of 12 simulation time steps as shown in Fig. 28. For these 12 time steps, the actual satisfaction state for MR was False but it was predicted as True.

Moreover, based on the results presented in Table.27, we can deduce the extent of the satisfaction state of MR, when both the active links and satisfaction probability are considered. For example, at time step 483, the active links are 131 and the satisfaction probability is 0.774904176. The output satisfaction state predicted by the model is True, similar to the actual state, with 1.00000000e+00 probability of being True. It means that given the input values, there is 100 percent chance of MR being satisfied i.e. having satisfaction state as True.

Classification Results for MP

For the 500 simulation time steps, the model correctly classifies the satisfaction state of MP with an exception of 3 simulation time steps as shown in Fig 28. For these 3 time steps, the actual satisfaction state for MP was False but it was predicted as True.

Moreover, based on the results presented in Table 28, we can deduce the extent of the satisfaction state of MP, when both the writing time and satisfaction probability are considered. For example, at time step 482, the writing time is 2047.139 ms and the satisfaction probability is 0.818595404. The output satisfaction state predicted by the model is True, similar to the actual state, with 9.99996634e-01 probability of being True. It means that given the input values, there is around 99 percent chance of MP to be satisfied i.e. having satisfaction state as True.

Table 26: Classification Results for Minimization of Operational Cost for time steps 482-488 under Scenario S_6

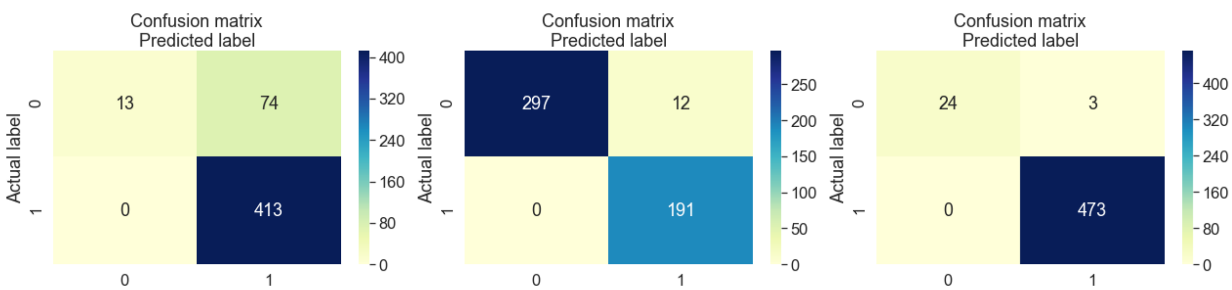
Time Step	Bandwidth Consumed	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
482	2811.180718	0.800333179	True	True	[1.38651595e-10 1.00000000e+00]
483	4088.490433	0.819174751	False	True	[5.65418723e-04 9.99434581e-01]
484	2056.129307	0.779324397	True	True	[1.82527215e-14 1.00000000e+00]
485	4550.743946	0.800471075	False	True	[0.14099679 0.85900321]
486	3098.344859	0.75329296	True	True	[5.87764568e-09 9.99999994e-01]
487	1921.293336	0.757097342	True	True	[4.16067614e-15 1.00000000e+00]
488	2695.519972	0.824877382	True	True	[2.95874514e-11 1.00000000e+00]

Table 27: Classification Results for Maximization of Reliability for time steps 482-488 under Scenario S_6

Time Step	Active Links	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
482	102	0.697876925	False	True	[0.4195637 0.5804363]
483	131	0.774904176	True	True	[5.13324505e-07 9.99999487e-01]
484	64	0.789328323	False	False	[9.99999986e-01 1.44312614e-08]
485	139	0.698271471	True	True	[1.12842716e-08 9.99999989e-01]
486	112	0.774522882	True	True	[0.00521161 0.99478839]
487	86	0.779436015	False	False	[9.99373649e-01 6.26351180e-04]
488	95	0.698377304	False	False	[0.95587702 0.04412298]

Table 28: Classification Results for Maximization of Performance for time steps 482-488 under Scenario S_6

Time Step	Writing Time	Satisfaction Probability	Actual Satisfaction State	Predicted Satisfaction State	Prediction Probability for Satisfaction
482	2047.138671	0.818595404	True	True	[3.36632577e-06 9.99996634e-01]
483	1728.466235	0.674782101	True	True	[1.74054149e-08 9.99999983e-01]
484	707.4143466	0.820974076	True	True	[1.68901772e-16 1.00000000e+00]
485	2978.726221	0.841578812	False	False	[0.97864493 0.02135507]
486	1727.069223	0.71163708	True	True	[1.54252846e-08 9.99999985e-01]
487	1324.932629	0.829141928	True	True	[9.21512659e-12 1.00000000e+00]
488	1763.638539	0.833946736	True	True	[2.14220808e-08 9.99999979e-01]



(a) Minimization of Operational Cost (b) Maximization of Reliability (c) Maximization of Performance

Figure 28: Confusion Matrix for Classification of Satisfaction State of NFRs under Scenario S_6

References

[1] M. Ji, A. C. Veitch, and J. Wilkes, "Seneca: remote mirroring done write," in *USENIX Annual Technical Conference, General Track*, 2003.

[2] K. Keeton, C. Santos, D. Beyer, J. Chase, and J. Wilkes, "Designing for disasters," 2004.

[3] L. Garcia-Paucar and N. Bencomo, "Knowledge base k models to support trade-offs for self-adaptation using markov processes," *13th IEEE Conference on Self-Adaptive and Self-Organizing Systems. Umea, Sweden*, 2019.

[4] N. Bencomo and L. Garcia-Paucar, "Ram: Causally-connected requirements-aware models using bayesian inference," *IEEE/ACM 22nd International Conference on Model Driven Engineering Languages and Systems (MODELS). Munich, Germany.*, 2019.

[5] H. Samin *et al.*, "RDMSim: An Exemplar for Evaluation and Comparison of Decision-Making Techniques for Self-Adaptation," in *SEAMS*, 2021.

[6] S. Menard, *Applied logistic regression analysis.* Sage, 2002, vol. 106.

[7] E. Bisong, "Logistic regression," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform.* Springer, 2019, pp. 243–250.

[8] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010.* Springer, 2010, pp. 177–186.