Implementation of Deep Learning Based Traffic Classification Model Using Genetic Artificial Intelligence Algorithm

Dr. Gangolu Yedukondalu[1], Prof. M. Srinivasa Rao[2], Prof. J. Anand Chandulal[3]
[1]CSE Dept, Vignan Institute of Technology and Science, Hyderabad.
[2] Retd. Professor, School of Information Technology, JNTU Hyderabad.
[3]SASI Institute of Technology and Engineering, Andhra Pradesh.

Abstract: Traffic classification is broadly utilized in different organization works, for example, programming designed systems administration and organization interruption location frameworks. Numerous traffic classification strategies have been proposed for grouping scrambled traffic by using a profound learning model without examining the bundle payload. Notwithstanding, they have a significant test in that the component of profound learning is mysterious. A glitch of the profound learning model might happen if the preparation dataset incorporates pernicious or mistaken information. Explainable artificial knowledge (XAI) can give some understanding for further developing the profound learning model by making sense of the reason for the glitch. In this paper, we propose a technique for making sense of the working component of profound learning-based traffic classification as a strategy for XAI in view of a hereditary calculation. We portray the instrument of the profound learning-based traffic classifier by measuring the significance of each element. Furthermore, we influence the hereditary calculation to create an element choice cover that chooses significant elements in the whole list of capabilities. To exhibit the proposed clarification technique, we executed a profound learning-based traffic classifier with an exactness of roughly 98.15%. In expansion, we present the significance of each element got from the proposed clarification strategy by defining the predominance rate.

Key words: Deep learning, AI, genetic algorithm, classification, network traffic.

I. INTRODUCTION

With the multiplication of Internet-associated gadgets and different Internet providers, it is critical to control the enormous traffic volume in an effective way. Traffic classification can be utilized to control different sorts of traffic in software designed organizing (SDN) or to identify pernicious traffic in network interruption recognition framework (NIDS) [1]. On account of SDN, QoS the board is vital to alleviate the weight of the whole organization and to fulfill the necessities of each
kind of administration [2]. As the Internet administrations are more different, It is critical to give every Internet administration the differential QoS. Dynamic QoS can give the differential QoS by partitioning the QoS class to help a more intricate QoS. Also, in light of the fact that various gadgets are associated with the Internet, the significance of advances for distinguishing and shielding against different assaults that might happen on the organization has been underscored. NIDS fills in as a center capability in network security by recognizing assaults like the disavowal of administration (DoS) assault in light of traffic characterization.

Customary traffic classification (TCs) are typically founded on a payload-review, which is known as a payload-based TC. A payload-based TC straightforwardly investigates the payload of bundles and matches the pre-characterized designs. Deep learning model a payload-based TC shows an elite presentation, there are two basic issues. One issue is that payload-based TC can't review the encoded payload. Since secure communication plans, for example, SSH and TLS encode the payload, the payload-based approaches can't review the payload mixed by the encryption conspire. Another issue is that investigating the payload of parcels requires huge computational assets.

A stream conduct-based TC has been proposed to resolve basic issues of a customary TC. A stream conduct put together TC is based with respect to AI innovations that can recognize designs without reviewing the

payload. In a stream conduct based TC, the AI model learns various factual highlights showing up in the organization, for example, the between appearance time or parcel size. The measurable highlights show contrasts in every application in light of the fact that the organization applications utilize various conventions and the examples of behavior change for every application client. Thusly, a stream conduct based TC enjoys benefits that can work inside the scrambled traffic and fulfill the prerequisites of crucial applications requiring low idleness.
In any case, there is an extreme test to a stream conduct put together TC that happens based with respect to the idea of AI. The discovery issue is gotten from the trouble of making sense of the aftereffects of the AI model. With an absence of dependability of the outcomes, the black-box issue has turned into a significant issue in AI [3].

We propose a prevailing aspect determination strategy to make sense of how the proposed profound learning-based traffic classifier operates. We characterize a fitting score as a component significance quantification and make an element choice cover that finds the ideal compromise between the high characterization exactness and a decrease of the superfluous highlights in view of a hereditary calculation. A hereditary calculation is a transformative calculation that can tackle the different NP-difficult issues like the mobile sales rep issue (TSP) or the plan of exceptionally huge scope reconciliation (VLSI). At long last, we portray the profound learning-based traffic classifier by characterizing a strength rate demonstrating the degree to which every profound learning model alludes to each element. Taking everything into account, the proposed strategy has two specialized commitments.

•We propose a predominant aspect determination strategy utilizing a hereditary calculation to make sense of how the deep learning-based traffic classifier works. Specifically, the supportive of presented technique can figure out what part of the whole component the classifier centers around by measuring the significance of each element.

•We carry out the stream conduct based traffic classifier as the assessment strategy that arranges the traffic and produces the exactness to process the fitting score. Deep learning model the proposed strategy likewise functions admirably in any granularity of the sorts of traffic, we execute a help explicit traffic grouping model to sort out the qualities of internet providers.

II. RELATED WORKS

The most pivotal issue of late examinations in rush hour gridlock classification is to group the scrambled traffic. Straightforwardly reviewing the payload was a boundary to scrambled traffic order. Conduct insights turned into a piece of information for characterizing encoded traffic in light of the fact that the measurements can be separated without examine a mixed payload. Stream conduct based approaches empower scrambled traffic to be grouped by utilizing the conduct insights. The creators of [6] presented three representative encryption systems of traffic and separated the measurements from the scrambled traffic. Besides, they evaluated the execution of a few AI calculations, for example, a help vector machine, irregular woods, innocent Bayes, strategic relapse, and brain organizations. They introduced the common sense of stream conduct based approaches by assessing different AI advances.

With the critical advances in profound learning, many studies on traffic order have taken on its assets. The center benefit of profound learning over customary AI innovations is to empower the classifier to automatically extricate highlights from the crude information. The creators of [9] took on a convolutional brain organization (CNN) for traffic characterization. Portrayal learning is a technique used to consequently extricate highlights from crude information and the CNN is a normal strategy for portrayal learning in profound learning. The convolution layer empowers a CNN to extricate the neighborhood highlights from the crude information. The creators incorporate element extraction and preparing by utilizing the upsides of the CNN. In [10], the creators assessed the two unique kinds of run of the mill profound learning models, a CNN and a repetitive brain organization (RNN). A RNN is intended to deal with consecutive information like time-series information. A few kinds of measurements can introduce the time-related nature, and the creators manage the time-related insights utilizing a RNN.

The creators of [1] favorable to represented a profound learning-based traffic order plot for versatile encoded traffic. The creators proposed that traffic order plans utilizing a physically removed include set for versatile traffic produced by a moving objective are impractical. Likewise, they address the restrictions of conventional traffic order plans by utilizing the benefits of profound realizing, which can consequently extricate the list of capabilities. The creators of [1] proposed a traffic order plot utilizing a hierarchical grouping. Stream conduct based approaches have a drawback in that they can't characterize obscure traffic classes in light of the idea of AI. More-finished, expanding the granularity of the traffic class grouping execution. The creators make the sub-classifier progressively founded on the granularity of the traffic class. In [1], the creators resolved an issue in which an obscure traffic class can't ordered use meta-learning. Profound advancing necessities adequate dataset for the calibrating when an obscure traffic class shows up. In any case, it is challenging to gather an adequate dataset of an obscure traffic class. Barely any shot learning, in particular meta-learning, empowers profound figuring out how to prepare the relationship of every information.

III.EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI)

Logical explainable AI (XAI) procedures have been considered to exhibit the system of the AI model. In [6], the creators proposed the idea of XAI, and contrived an original logical model that permits the AI model to determine such order results in view of the component subset of the info information. In [16], the creators imagined the important highlights that are utilized for grouping specific information and made sense of why the profound learning model can perceive such information. To make sense of this, the creators proposed a responsiveness examination (SA) that makes sense of the effect of the change of every pixel. Additionally, the creators likewise proposed layer-wise significance engendering (LRP), which makes sense of the significance of every pixel. In [1], the creators proposed EXPLAIN-IT, a system that makes sense of how for bunch an unlabeled YouTube traffic dataset gained from the organization utilizing an unsupervised learning strategy. Make sense of IT makes sense of the clustering strategy utilizing LIME, which chooses the element generally pertinent to a particular choice from the information. Consequently, the key element determination can make sense of why the profound mastering model orders the information. In [8], the creators portray the connection between the information and result by embedding fake irritations in specific elements. The info yield relationship can give some translation rules to discovery indicators like profound learning. Brain picture subtitle age with a visual consideration plot is proposed in [9]. The creators separated the vital highlights in the picture utilizing convolutional highlight extraction. The removed highlights are utilized to prepare the RNN for picture inscribing. During this method, the consideration instrument carried out through a convolutional include extraction can feature a significant piece of the picture.

Many investigations on XAI mean to make sense of the AI model for picture order. Notwithstanding, the traffic arrangement issue has various qualities from the picture characterization issue. In the picture order issue, all components of the information have the equivalent semantic, for example, RGB variety esteem. The consideration system proposed in [19] chooses an element subset by distinguishing an item from information composed of pixels having a similar importance. In the rush hour gridlock characterization issue, the element of the information is more modest than that of the picture information. Furthermore, since every component of the information has various implications, an element determination strategy that can mirror these qualities is required. We planned a predominant aspect determination strategy that is reasonable for the low-layered social insights in light of a hereditary calculation.

IV. THE PROPOSED METHOD

The outline of the proposed predominant aspect determination strategy is delineated in Figure 1. The proposed technique con-sists of two sections: (1) the development of a traffic classifier, and (2) prevailing aspect determination. The traffic classifier is planned by a leftover organization (ResNet), which is known as a cutting-edge profound learning method. The traffic characterization applies information pre-handling and preparing step. The information pre-handling step gathers parcels from traffic streams and concentrates the measurable elements of each stream. After the pre-handling step, in which the traffic dataset is made, the traffic classifier is prepared utilizing the dataset made out of factual elements.

After the classifier is prepared, the proposed predominant feature determination strategy creates an element choice cover in light of a hereditary calculation. The predominant aspect choice directs a veil determination and a posterity cover age. The veil choice assesses the covers by counting the zero-components and computing the exactness utilizing a concealed info dataset and a pre-prepared classifier. After the assessment, with the veil determination picks a couple of covers are picked utilizing a roulette wheel choice strategy for the veil production of the future. With the roulette wheel choice strategy, the likelihood of choosing the veils with a higher fitting score is higher than the others. The posterity cover age makes a veil pool utilizing the chose covers and gives assortment to the veil pool through a hybrid and transformation. The veil pool created by the posterity cover age is acquired by the future. After the cycle of two stages, the element choice covers for each help are made and the veils are utilized to pick the highlights important to arrange each assistance from every measurable component. At long last, we investigate the mechanism of the traffic classifier by registering the significance of each element utilizing the component choice covers.

The development of a traffic classifier comprises of three stages: parcel gathering, information pre-handling, and classifier preparing. The parcel gathering step gathers bundles and gatherings them by the traffic to build the preparation dataset. Since most parcels are scrambled, a bundle itself can't be utilized as a train-ing dataset, deep learning model the gathered parcel dataset that has a similar start to finish network address, for example, the IP address or TCP port number is required. The parcels in a gathered dataset may serve a similar application administration since they have a similar application source, and such bundles structure an organization stream. The parcels in an organization stream have a comparative way of behaving, which is addressed by the measurable highlights, for example, the packet size and between appearance time. Subsequent to get-together bundles from an organization stream, the information pre-handling step figures the measurable highlights from the gathering of assembled parcels. At long last, the classifier preparing step prepares the profound learning-based classifier utilizing the dataset.

V.DATA PREPROCESSING

The information pre-handling step processes the measurable highlights of the bidirectional stream set displayed as

The ways of behaving of the bundles in the organization are represented as measurable highlights, which are fundamentally uncovered by the between appearance time, parcel size, number of bundles, and number of bytes. Deep learning model the bundles are encoded, the parcels serving a similar application layer convention have special ways of behaving, and the conventions that serve a comparable sort of administration show comparable ways of behaving. For instance, moment messaging administrations can cause burst traffic, which can be displayed in the factual highlights, for example, a somewhat short between appearance time and parcel size. In this manner, the profound learning-based traffic. Information preprocessing includes highlight extraction and administration marking. The previous part expects to remove highlights from the stream F p1, p2, p3, . . ., pn, where F has n bundles and pi is the I-th parcel. Since we manage a bidirectional stream, the highlights of the converse course stream F¯ are moreover required. Hence, 10 sorts of factual highlights in one course can be extricated, and there are 20 kinds of elements in one bidirectional stream made by F and F¯. In stream F, we process measurable highlights, for example, the between appearance time and bundle size as follows:

VI. DOMINANT FEATURE SELECTION

We proposed a predominant aspect choice strategy to make sense of how the profound learning model groups traffic. In classification issues, there are key components inside the information that are the reason for order. For instance, in regular language handling (NLP), the subject and action word are the key components, and the others are qualifiers used to make sense of them in the word tokens. Using information with such a large number of or pointless parts for preparing may cause a higher intricacy of the model. As a matter of fact, information with a large number might prompt a higher exactness. As such, the characterization accuracy shocking additionally diminishes in light of the fact that low-layered information have less data for the choice. In this way, there is a compromise between the characterization exactness and aspects of the information, and the classifier needs an aspect decrease technique that boosts the precision.

We propose a predominant aspect choice strategy in view of a hereditary calculation as an aspect decrease method. The point of the proposed technique is to find the ideal component choice veils, limiting the quantity of chosen includes and amplifying the arrangement precision. Subsequently, we formulated the goal capability as a direct blend of two variables, to be specific, the quantity of covered components and the characterization exactness. Here, ρ1 is the quantity of dropped highlights and ρ2 is the order exactness. In addition, we amplify ρ1 on the grounds that boosting the quantity of dropped highlights is equivalent to limiting the quantity of chosen features.

The component determination strategy is led as a few iterations comprising of two stages, in particular, best veil choice and posterity cover age. The best cover determination step evaluates the fitting score of the parent veils and chooses a couple of the best veils through a roulette-wheel choice. After veil determination, the offspring are made through a hybrid and change. The ideal covers are made through adequate emphases of the above strides toward a boost of the goal capability addressed by the fitting score. Note that the fitting score is a marker that addresses the optimality.

The proposed Traffic Classification

Require: Training packet trace *P* collected with short time
duration, the traffic flow *F* composed of packets p_i, the
function Ω (*F*) returning the 5-tuple of the _ow *F*.
Ensure: Pre-trained traffic classifier.
1: *D* --π
2: Ω D f Ω (F_1); Ω (F_2); : : : ; Ω(F_N)g
3: Perform clustering packets in P by 5-tuple set Ω to form
a bidirectional _ow set fF_1; F_2; : : : ; F_N g
4: for *i* D 1 ! *N* do
5: t ← S_{n-1}
$_{j}$D1 f_ (p_{j}C1) _ (p_j)g
6: s D fs_k js_k is packet size of packet p_k ; 1 _ *k* _ *n*g
7: Compute total bytes *b* in the _ow
8: Compute feature vector by using traffic flow statistical features
D [m_t M_t _t _t m_s M_s _s _s *n* *b*]
9: Compute reverse directional feature vector N
10: x_i D [; N]
11: Detect the application layer l_i by the packet gathering step
12: *D* *D* [f(x_i; l_i)g
13: end for
14: Normalize dataset *D*
15: for *i* D 0 ! *N* do
16: Pick (x_i; l_i) 2 *D*
17: for *j* D 0 !number of ResNet layers do
18: *e* VD x_i
19: x_i VD batch_normalization (x_i)
20: x_i VD ReLU (x_i)
21: x_i VD convolution (x_i)
22: x_i VD *e* C x_i
23: end for
24: Calculate the loss between the result of ResNet and
l_i, and backpropagate the gradient of the loss to the model.
25: end for

The proposed technique has two hyper-boundaries that have a likelihood of working a hybrid and change, for example, the hybrid and transformation rates. It works a hybrid and transformation with a specific likelihood, and in this way some chromosomes have been changed, deep learning model some others have been saved. Thusly, the proposed technique gives optimal highlight choice veils in a stochastic way. Since posterity cover age strategy depend on the hereditary calculation, they are of a probabilistic sort. All in all, with the hereditary calculation, a hybrid and change randomly happen. Subsequently, tracking down better chromosomes and keeping up with the best people creates the ideal component choice veil.

VII. PERFORMANCE EVALUATION AND EMPIRICAL ANALYSIS

In this part, we portray the exhibition of the profound learning-based traffic classifier used to assess the exactness, learning cost. To assess the presentation of the proposed strategy, we completed various investigations utilizing true information. For fair assessments, the public pcap datasets are utilized to construct the preparation dataset. We took on open pcap datasets from ISCX VPN-non VPN,MACCDC, and WRCCDC, which have likewise been often utilized in different examinations in rush hour gridlock grouping and incorporate both scrambled and non-encoded parcels. Deep learning model the public pcap dataset has numerous parcels that work different conventions, the quantity of streams gathered by bundles that share a similar 5-tuple is inadequate to prepare profound learning-based traffic order model. To supply additional preparation information, we accumulate the extra pcap information utilizing the server which produces bundles of different conventions from a ground's organization. Thus, the whole dataset is made out of 49 applications, as displayed. The number of segments addresses the quantity of streams. We set the quantity of information by each help like keep away from one-sided preparing.

Besides, for commonsense use, parcels of one stream are accumulated for 900 seconds disregarding a TCP meeting break. To make a preparation dataset, we executed the typed and non-scrambled bundles. Deep learning model the public pcap dataset has numerous bundles that work different conventions, the quantity of streams gathered by parcels that share a similar 5-tuple is lacking to prepare profound learning-based traffic grouping model. To supply additional preparation information, we assemble the extra pcap information utilizing the server which produces bundles of different conventions from a grounds organization. Thus, the whole dataset is made out of 49 applications. The number of section addresses the quantity of streams. We set the quantity of information by each assistance like stay away from one-sided preparing. Besides, for viable use, bundles of one stream are accumulated for 900 seconds disregarding a TCP meeting break. To make a preparation dataset, we executed the information pre-handling program utilizing the nDPI library, which can identify the application.

The nDPI is an open-source DPI library that gives data on both the payload and header of the bundle. In view of the nDPI, we accumulate information about the application layer convention, IP address, and TCP port number. Deep learning model the public pcap dataset applies the preassigned marks, we really want to allot names to the extra dataset gathered by our ground's organizations. We influence the nDPI library to allot the extra dataset gathered by our ground's organizations. In the wake of social occasion, the data, we implemented the remainder of the pre-handling program that bunches the bundles into the stream and concentrates the stream measurements. The profound learning-based traffic classifier is carried out utilizing TensorFlow. The whole tests are directed by a server with an Intel i9-8980XE CPU, 64GB of RAM, and a NVIDIA GeForce GTX 2080.

VIII.PERFORMANCE EVALUATIONS OF TRAFFIC CLASSIFICATION

For preparing the profound learning model, we isolated the 70% of the dataset into preparing dataset and 30% into test dataset, and all assessments depend on the test dataset. The boundaries are instated aimlessly, and a clump standardization layer is utilized to relieve the work expected to regularize the parameters by framing a comparable circulation in each layer. There are some hyper-boundaries to be tuned before the preparation, for example, bunch size and number of ages. We found the two hyper-boundaries above through a satisfactory number of tests with a bunch size of 300 and 5,000 ages. More-finished, we directed tests by changing other hyper-boundaries, for example, the quantity of channels in the convolution layer and the quantity of layers in the lingering block. Note that one leftover block comprises of a few convolution layers and cluster standardization layers, and the whole model is developed by stacking a few lingering blocks. Figures 2 show the test cost and test precision as per the cycles. ResNet enjoys a benefit that effectively controls the intricacy of the model by tuning the quantity of remaining blocks and convolution channels. Subsequently, the model ought to have adequate intricacy to enough depict the dataset by expanding the quantity of layers and channels.

Figures 1 show the test cost and test exactness as indicated by hyper-boundaries like the layers and channels. As displayed in the figures, the model with 128 channels shows the greatest precision and least expense. It tends to be seen that when the profound learning model is prepared by 128 channels, the classifier can accomplish an adequate exhibition. Subsequently, we utilize an adequately prepared model whose quantities of layers and channels are 16 and 128, separately.

The disarray grid shows the amount of the information are accurately grouped and assists with working out a few measurements like genuine positive, genuine negative, bogus positive, and misleading negative. Figure 2 shows the disarray network as per the classes. In the disarray network, it tends to be seen that the general arrangement exactness of the model is roughly 97.24%. Additionally, the accuracy, review, and F1-score of each help can likewise be determined in view of the disarray framework, as displayed in Figure 2. Accuracy is the proportion of how much genuine information 'A' from the aggregate sum of information anticipated as 'A', and review is the proportion of how much information anticipated as 'A' from the aggregate sum of genuine information 'A'. Nonetheless, assessing the exhibition in light of accuracy and review might be confounded on the grounds that an imbalanced dataset may cause various patterns in the accuracy and review. The F1-score is the symphonious mean of accuracy and review, and in this manner the F1-score can make sense of the presentation of the model, which has various patterns in the accuracy and review.

IX. EMPIRICAL ANALYSIS OF SERVICE-SPECIFIC TRAFFIC CLASSIFICATION

In this segment, we dissect how profound learning-based traffic classifier orders traffic into administrations. We led tests by changing the two hyper-boundaries, λ1 and λ2, of the proposed hereditary calculation based logical technique and assessed the presentation as per the hyper-boundaries. Note that we set the amount of λ1 and λ2 as 1 to reasonably quantify the impacts of the two factors ρ1 and ρ2. In light of the component determination covers created by each examination, we characterize the predominance rate, which addresses the significance of elements and dissects the factual highlights of each help.

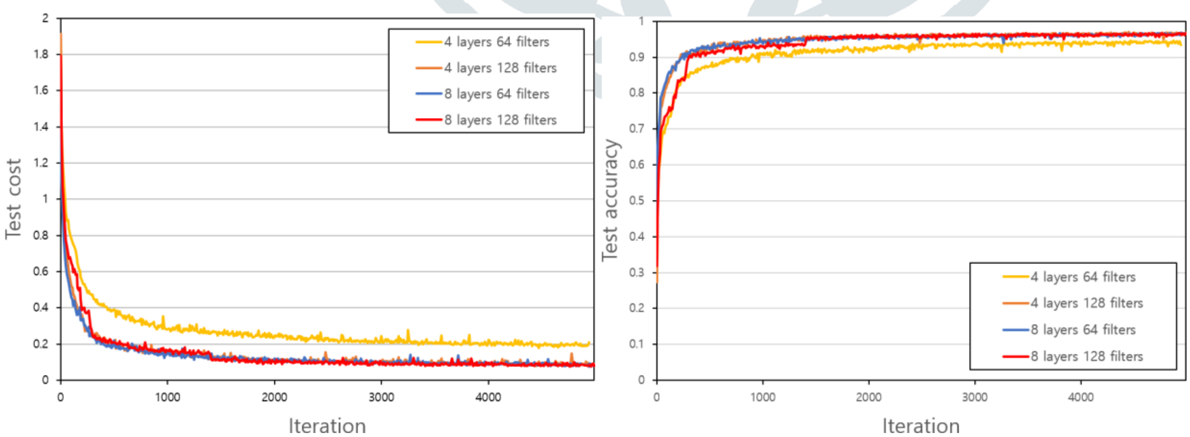

Fig 1. (a) Test cost according to iterations and (b) test accuracy according to number of iterations.

Figure 3 shows how the fitting score, exactness, and number of dropped highlights change all through the ages relying upon the loads λ1 and λ2. From each help, it tends to be seen that, as λ1 builds, the quantity of dropped highlights increments. Simultaneously, as λ2 diminishes, the typical exactness of the veil pool diminishes. The primary column is the outcome wherein λ1 is set to 0.1 and λ2 is set to 0.9, where the proposed strategy expects to look for covers with a higher characterization precision. Since it is for the most part better to have more information for a profound learning model concerning characterizing the administrations, a higher exactness and more strength in the grouping are shown. The subsequent segment shows the outcomes in which both λ1 and λ2 are set to 0.5, implying that the methodology expects to track down veils with both a higher precision and a bigger number of dropped highlights. It very well may be seen that the proposed strategy will in general show a fair and corresponding way of behaving among precision and the quantity of dropped highlights. That's what the third segment shows, when λ1 is set to 0.9 and λ2 is set to 0.1, it demonstrates that the proposed technique intends to track down veils with a bigger number of dropped highlights, as opposed to accomplishing a higher precision. Hence, the outcome shows that the quantity of dropped highlights is a lot higher than that of the other two outcomes with various loads.

In any case, it shows the most reduced precision bring about broad. What's more, it very well may be seen that the consequence of a "web surfing" administration shows a generally higher precision in any event, when the weight λ1 is higher and λ2 is lower. That is on the grounds that a web riding administration is a nonexclusive class and consequently has more broad qualities than other explicit administrations, it permits the profound learning model to require many less elements to characterize the "web surfing" than different administrations.

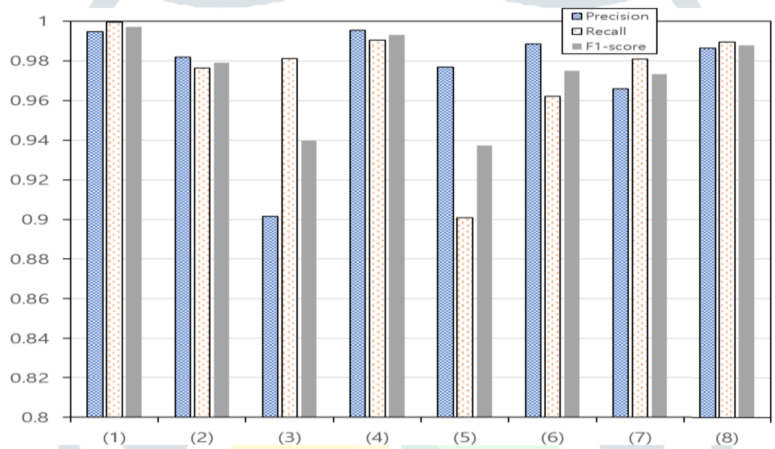

Fig 2. Precision, recall, and F1-score according to each service. Services

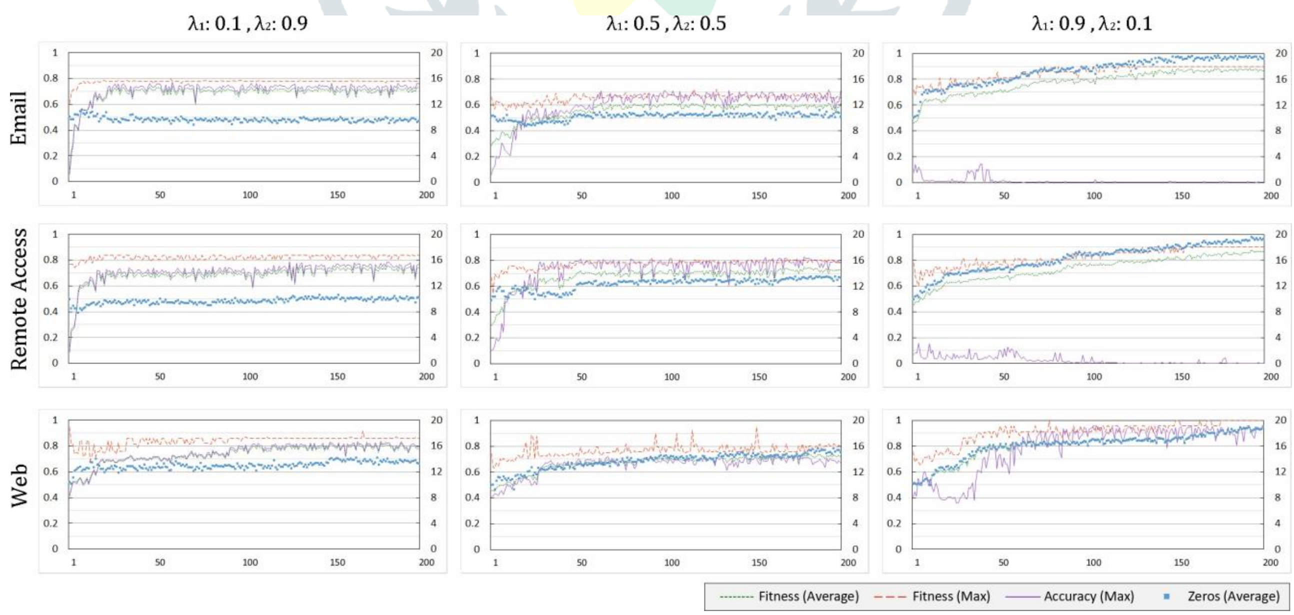

Fig 4. Fitting score, accuracy, and number of zeros per generation for 3 services

Figure 4 shows the fitting score, precision, and number of dropped highlights in the last age of the proposed technique. From each assistance, it very well may be seen that the affirm age exactness and number of dropped highlights are subject to the weight values λ1 and λ2. The fitting scores will generally be higher when the weight λ1 is higher on the grounds that it requires less investigation while tracking down veils with additional dropped highlights. Conversely, when λ2 is higher, moderately lower fitting scores are shown in light of the fact that a significantly more careful investigation is expected to track down the legitimate covers with higher exactness.

As a general rule, the CNN model works the grouping supportive of access by gathering a few nearby elements. It is vital to find key elements that are utilized as standards utilized for the profound learning model to arrange. The XAI is an informative technique that depicts how the profound learning model can classify the qualities of a specific item by separating the key highlights like eyes, nose, and ear.

It very well may be seen that the profound learning-based traffic classifier doesn't group traffic into a help utilizing all elements, however orders it utilizing just unambiguous highlights. We directed nine analyses by changing the λ1 and λ2 and arrived at the midpoint of the consequences of the examinations. At the point when the proposed technique finishes an adequate number of emphases, it creates 200 component determination covers of each help for one examination and picks the best 10 veils, which accomplish the most noteworthy. Figure 4 shows the strength rate for each factual element that influences the precision. In the event that the predominance rate is high, it can assist with expanding the characterization exactness or fitting score, in particular, it tends to be a possibility for the key component. In any case, these elements have less impact on the traffic arrangement, and consequently they can be competitors of pointless highlights. A predominance pace of 100 percent infers that the classifier generally utilizes the elements to group the traffic into the help, in particular, the component is utilized as the center component of the assistance. Since the 0% predominance rate suggests that the element doesn't influence the arrangement, it shows that the component is insignificant for grouping.

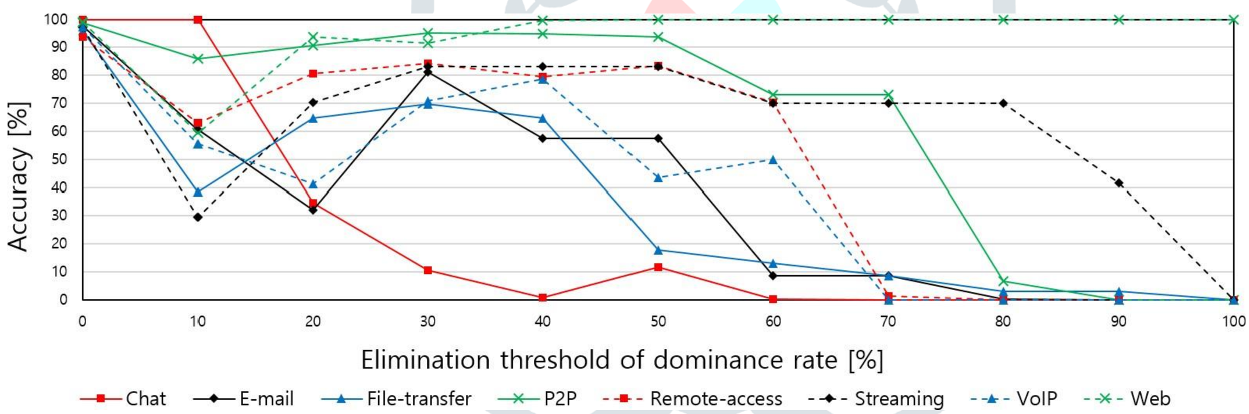

Fig 3. Test accuracy according to the elimination threshold of the dominance rate.

To gauge how much each component adds to the grouping precision, we characterize the disposal edge and measure the exactness of every end limit. In the event that the strength pace of an element is lower than the elimination edge, the component is taken out. Figure 3 shows the precision as per the disposal limit. As the elimination limit expands, the quantity of elements eliminated increments, and in this manner the general characterization exactness will in general diminish. Since highlights with a low strength rate are just taken out through the end edge, the correlations between highlights are not thought of. Thus, while eliminating an element that is connected with different highlights, a vacillation that briefly diminishes the precision might happen. On account of the "texting" class, the relationship between the quantity of taken out highlights and accuse shocking has a dull lessening, and that intends that there is little connection for each element. On account of the "web surfing" class, there might be slight changes in precision, deep learning model the general exactness doesn't diminish fundamentally as the quantity of eliminated highlights increments on the grounds that the majority of the highlights have a low strength rate.

A large portion of the administrations have a few key highlights, deep learning model the "web surfing" class has not many up-and-comers. As such, the profound learning model of our strategy will in general order traffic that it can't characterize into the "web surfing" class. This can be made sense of by the way that the traffic of "web surfing" administrations will in general show a widespread property that addresses the overall ways of behaving of Internet administrations. For instance, most Internet administrations depend on the hyper-text move convention (HTTP), and subsequently the traffic of a "web surfing" administration in light of HTTP can show different ways of behaving in the organization. Consequently, all of the other traffic that doesn't show the particular ways of behaving of different administrations can be characterized into the "web surfing" class, which is viewed as a nonexclusive class. Paradoxically, there are numerous relationships among elements, and subsequently a change in the exactness can oftentimes happen.

X.CONCLUSION

In this study, we proposed a explanatory method of the deep learning-based traffic classifier based on a genetic algorithm. Further, we implemented the deep-learning-based traffic classifier based on the ResNet model for demonstrating the proposed explanatory method. We designed the dominant feature selection method as a explanatory method based on a genetic algorithm to generate an optimal feature selection mask. The proposed explanatory method generates the optimal feature selection masks by grafting the deep-learning based traffic classifier's result onto the evaluation of the chromosome in a genetic algorithm. The feature selection masks are used to extract the key feature subset from the entire feature set by considering the trade-off between the classifier's accuracy and the number of unnecessary features. We conducted several experiments for reflecting the stochastic property of a genetic algorithm and computed the importance rate through the feature selection masks. Through the importance rate, we explained the mechanism of the deep-learning-based traffic classifier by investigating the key features of each Internet service. In the future, we plan to design a key feature selection algorithm for finger-grained application-specific traffic classifiers. In addition, we will improve the convergence speed of the genetic algorithm to enable real-time key feature selection.

REFERENCES

[1] D. Gunning, ``Explainable articial intelligence (XAI)," in Defense Advanced Research Projects Agency (DARPA), nd Web. Arlington, VA, USA: Defense Advanced Research Projects Agency (DARPA), 2017.

[2] K. Zhou, W. Wang, C. Wu, and T. Hu, ``Practical evaluation of encrypted traffic classification based on a combined method of entropy estimation and neural networks," ETRI J., vol. 42, no. 3, pp. 311-323, Jun. 2020.

[3] JEEHYEONG KIM,"Explaining Deep Learning-Based Traffic Classification Using a Genetic Algorithm", IEEE Access.

[4] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescape, ``Mobile encrypted trafc classication using deep learning: Experimental evaluation, lessons learned, and challenges," IEEE Trans. Netw. Service Manage., vol. 16, no. 2, pp. 445458, Jun. 2019.

[5] A. Montieri, D. Ciuonzo, G. Bovenzi, V. Persico, and A. Pescape, ``A dive into the dark Web: Hierarchical trafc classication of anonymity tools," IEEE Trans. Netw. Sci. Eng., vol. 7, no. 3, pp. 1043-1054, Jul. 2020.

[6] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, ``Show, attend and tell: Neural image caption generation with visual attention," in Proc. Int. Conf. Mach. Learn., 2015, pp. 20482057.

[7] K. He, X. Zhang, S. Ren, and J. Sun, ``Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770-778.

[8] H. Kim and N. Feamster, ``Improving network management with software dened networking," IEEE Commun. Mag., vol. 51, no. 2, pp. 114-119, Feb. 2013.

[9] M. Karakus and A. Durresi, ``Quality of service (QoS) in software defined networking (SDN): A survey," J. Netw. Comput. Appl., vol. 80, pp. 200218, Feb. 2017.