YIDDISH Vol 13 Issue 04,Dec 2023
ISSN NO: 0364-4308

SURVEY ON MACHINE LEARNING IN 5G

1.T. SWATHI ,2. S. VARSHITHA REDDY,3. S. GAYATHRI,4. S. SUPRAJA,5. S. PRAVALIKA

1.ASSISTANT PROFESSOR,2,3,4&5 UG SCHOLAR

DEPARTMENT OF ECE, MALLA REDDY ENGINEERING COLLEGE FOR WOMEN, HYDERABAD

ABSTRACT: The core of next generation 5G wireless network is heterogeneous network. The upcoming 5G heterogeneous network cannot be fulfilled until Artificial Intelligence is deployed in the network. The existing traditional 4G technology approaches are centrally managed and reactive conception-based network which needs additional hardware for every update and when there is a demand for the resources in the network. 5G helps in giving solution to the problem of 4G network using prediction and traffic learning to increase performance and bandwidth. Heterogeneous network provides more desirable Quality of Service (QOS) and explores the resources of the network explicitly. The assortment of heterogeneous network brings difficulty in traffic control of the network. The problem in heterogeneous network is network traffic which cannot be controlled and managed due to different protocols and data transfer rate. To solve the problem in heterogeneous network advanced techniques like Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) are employed in 5G Network which are self pro-active, predictive and adaptive. In this paper we discuss about above mentioned advanced techniques that are deployed in 5G to reduce traffic in a network which increases efficiency of the network.

Keywords:5GNetwork, Artificial Intelligence, Machine Learning, Deep Learning.

INTRODUCTION In recent years with the prosperous development of the Internet, networking has allured a lot of recognition in both industry and academia. The improvement of mobile communication has increased the data transfer rate significantly, for large amount of data with multimedia communication service. The mobile communication is now stepping into 5G. To satisfy the data traffic demand the network technology are moving towards heterogeneous network which provides ubiquitous internet access and enhanced public services [1]. The next generation network is service-driven where a single infrastructure should efficiently provide different service such as low latency communication, enhanced mobile broadband and immense machine type communication for heterogeneous network. The heterogeneous network different layers of cells like femto, macro, micro, pico, relays,

diverse user devices and application. In this paper we will discuss about how AI's potential is used in next generation wireless network using basic learning algorithms like ML, DL, etc.,

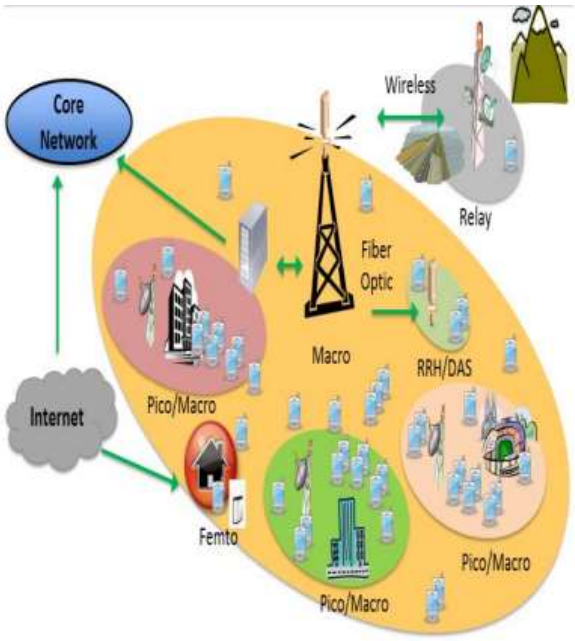

Fig: Components of Heterogeneous Network

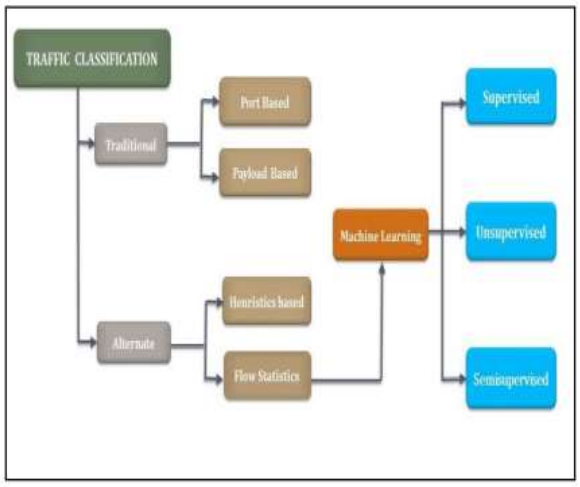

Fig: Traffic Classification approaches.

A. Port based IP traffic classification TCP and UDP give multiplexing of different streams between IP endpoints with the assistance of port numbers. Generally numerous applications use an 'outstanding' port to which different hosts may start correspondently. The application is deduced by looking into the TCP SYN parcel's objective port number in the Internet Assigned Numbers Authority (IANA's) rundown of enlisted ports. In any case, this methodology has constraints. Right off the bat, a few applications might not have their ports enrolled with

IANA (for instance, distributed applications, for example, Napster and Kazaa). An application may utilize ports other than its outstanding ports to rescue from working framework get to control confinements. Additionally, at times server ports are powerfully assigned as required. Though port-based traffic grouping is the quickest and straightforward strategy, a few examinations have demonstrated that it performs ineffectively, e.g., under 70% precision in characterizing streams [11]12].

B. Payload based IP traffic classification This methodology reviews the packet header to decide the applications. Packet payloads are analyzed a little bit at a time to find the bit streams that contain signature. In the event such piece of streams is discovered, at that point bundles can be precisely named. This methodology is regularly utilized for P2P traffic discovery and system interruption identification. Real impediments of this methodology is that the protection laws may not enable directors to assess the payload; it additionally forces huge multifaceted nature and preparing load on traffic ID gadget; requires significant computationally power and capacity limit since it examinations the full payload [2]

C. Protocol Behavior or Heuristics Based Classification In this method the classification of networks is based on connection level patterns and network protocol behavior. This method is based on identifying and observing patterns of host behavior at the transport layer. The advantage of this classification is that packet pay load access is not needed [10][2].

D. Classification based on flow statistics traffic properties: The preceding techniques are restricted by their dependence on the inferred linguistics of the information gathered through deep review of packet content (payload and port numbers). Newer approaches depend on traffic's statistical characteristics to identify the applying [4][7][6][9]. associate degree assumption underlying such ways in which is that traffic at the network layer has mathematical properties that are distinctive definitely classes of applications and modify wholly totally different offer applications to be distinguished from each other. It uses network or transport layer that has applied mathematics properties like distribution of flow length, flow idle time, packet interarrival time, packet lengths etc. These are distinctive sure categories of applications and thence facilitate to {differentiate|to tell apart} different applications from one another. This methodology is possible to see application sort however not usually the particular consumer type. as an example, it can't verify if flow belongs to Skype or MSN traveller voice traffic specifically. The advantage of this approach is that there's no packet payload scrutiny concerned.

MACHINE LEARNING In every possible field machine learning has been used to leverage its astonishing power. In variety of application such as speech recognition, bio informatics and computer vision, ML techniques have been used efficiently. Machine learning is mainly used for prediction and classification and also in networking it is mainly used for performance prediction and intrusion detection. To make decision directly Machine learning constructs models that can learn themselves from data without being explicitly programmed or without following some set of rules. Machine learning enables the model to get into selflearning mode without being explicitly programmed. The model can be trained by providing data sets to them, when exposed to new data, models are enabled to learn, predict and develop by themselves. Machine learning algorithm can be classified into three categories. They are supervised learning, unsupervised learning, reinforcement learning [3]. In Supervised learning the model is trained on a labeled data set which then learns on its own and when new testing data is given it compares with the training data set and predicts the output. Supervised learning is mainly used for regression and classification problems. In unsupervised learning the training data set is unlabelled, and it finds pattern and relationship among data. It is mainly used in clustering and association problems. In reinforcement learning the model learns on its own without any training data.

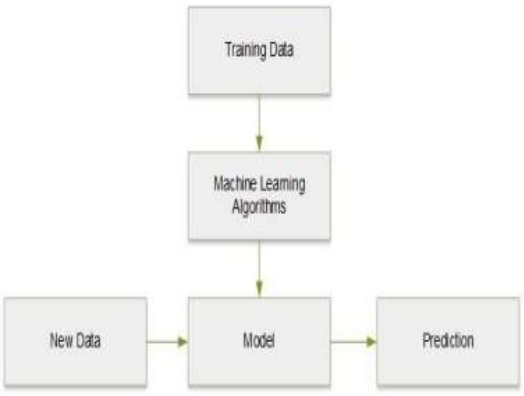

Fig: Machine Learning Work Flow

WORK FLOW OF MACHINE LEARNING IN NETWORK In next generation network machine learning plays an important role. The steps that are involved in networking are:

• Step1: Problem formulation

• Step2: Data collection

• Step3: Data analysis

• Step4: Model construction

• Step5: Model validation

• Step6: Deployment and interference

A. Problem formulation In machine learning the training process is time consuming so it is mandatory that the problem should be formulated correctly at the beginning of the process. There should be a strong relation between the problem and the data that has been collected. The machine learning model is classified as clustering, classification and decision making and the problem statement should also fall under this category [8][15][16]. This help in identifying the learning model and also for collecting data. When the problem formulation is not done properly it leads to unsuitable learning model and un satisfactory performance.

B. Data collection. There are two types of data collection. They are offline data collection and online data collection. In online data collection the real time data are collected and they can be used as feedback for the model and it can also be used as a re-training data for the model. Offline data can be collected from repositories [17][18][19]. For the purpose of classification of network traffic, we are utilizing the datasets that are made from this present reality traffic flow named as 'wide'. The wide dataset comprises of traffic streams which are haphazardly chosen from the wide follow and cautiously perceived by the manual examination. It comprises of 3416 occasions with 7 classes, for example, (bt, dns, ftp, http, smtp, yahoomsg, ssh) and 22 traits. The features that are extracted from the process [13] is recorded in Table 1 By using monitoring and measurement tool online and offline data can be collected effectively which provides security in various data collection aspects. It can also be stored for model adaption. After data collection the process is categorized as training or learning phase, validation and testing phase.

C. Data analysis Data analysis consists of two phases. They are:

• pre-processing

• feature extraction.

Pre-processing is done to remove noise from the data that has been collected. Then the features of the data are extracted which is a prior step for learning and training[10]. The types of features that can be extracted from the network are:

• Packet level features.

• Flow level features

In packet level features the extracted features are packet size, mean, root and variance. In flow level mean flow duration and mean number of packet flow features are extracted

Types Of Features	Feature Description	Number
Packets	Number of packets transferred	2
Bytes	Volume of bytes transferred	2
Packet Size	Min, Max, Median and Standard. Deviation of packet size	8
Inter Packet Time	Min, Max, Median and Standard. Deviation of inter packet time	8
	Total	20

TABLE1: STATISTICAL FEATURES

D. Model construction In this process model selection, training and tuning are involved. According to the size of the data set a suitable learning model and algorithm needs to be selected. Training involves training of the model along with the data set that is bee collected at the beginning of the stage. The tuning process helps in making the model to learn themselves by comparing them with the trained data.

E. Model validation It involves cross validation of the testing process to test the accuracy of the model. This helps in optimizing the model and maintains the overall performance of the system.

F. Deployment and interference In deployment and interference stage all the trade off and stability of the model is maintained to check the accuracy and finds the best way in which steps has been followed

CONCLUSION AI is used as a tool to improve 5G technologies in recent technologies. The reason for not using AI algorithm in networking is due to the lack of learning process that has been left in past few years. Heterogeneous network is the basic for next generation network

where traffic in a network plays a major role in disturbing the performance of the network. In this paper we discussed about machine learning techniques and its implementation in 5G heterogeneous network to increase its performance by reducing traffic in a network.

REFERENCES

[1] L. C. I, S. Han, Z. Xu, et al, New paradigm of 5G wireless internet [J], IEEE Journal on Selected Areas in Communications,Vol.34, No.3,474-482,2016.

[2] R. Li, Z. Zhao, X. Zhou, et al, Intelligent 5G: when cellular networks meet artificial intelligence[J], IEEE Wireless Communications, PP(99) ,2-10,2017

[3] S. Dzulkifly, L. Giupponi, F. Sai, et al, Decentralized Q learning for uplink power control[C], IEEE, International Workshop on Computer Aided Modelling and Design of Communication Links and Networks, IEEE, 54-58, 2015.

[4] M. Shafi et al.,"5G: A Tutorial Overview of Standards, Trials,Challenges, Deployment, and Practice," IEEE JSAC, vol. 35,no. 6, , pp. 1201–21. June 2017

[5] D. Soldani, A. Manzalini, Horizon 2020 and beyond: on the5G operating system for a true digital society [J], IEEE Vehicular Technology Magazine, Vol.10, No.1, 32-42,2015.

[6] Y. Zhang et al., "Home M2M Networks: Architectures, Standards, and QoS Improvement," IEEE Commun. Mag.,vol. 49, no. 4, Apr. 2011, pp. 44–52.

[7] Chen Z, Wen J, Geng Y. Predicting future traffic using hidden markov models. In: Proceddings of 24th IEEE International Conference on Network Protocols (ICNP). IEEE; 2016. p. 1–6.

[8] I. Portugal, P. Alencar, and D. Cowan, "The Use of MachineLearning Algorithms in Recommender Systems: A Systematic Review," Expert Systems with Applications, vol. 97, May 2018, pp. 205–27.

[9] S. Maharjan et al., "Dependable Demand Response Management in the Smart Grid: A Stackelberg Game Approach," IEEE Trans. Smart Grid, vol. 4, no. 1, 2013, pp. 120–32.

[10] G. Alnwaimi, S. Vahid, and K. Moessner, "Dynamic heterogeneous learning games for opportunistic access in LTE-based macro/femtocell deployments," IEEE Transactions on Wireless Communications, vol. 14, no. 4, pp. 2294–2308, 2015.

[11] D. D. Nguyen, H. X. Nguyen, and L. B. White, “Reinforcement Learning with Network-Assisted Feedback for Heterogeneous RAT Selection,” IEEE Transactions on Wireless Communications, vol. 16, no. 9, pp. 6062–6076, 2017.

[12] U. Challita, L. Dong, and W. Saad, “Deep learning for proactive resource allocation in LTE-U networks,” in European Wireless 2017- 23rd European Wireless Conference, 2017.

[13] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training recurrent neural networks,” Tech. Rep., 2013.

[14] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo, “Machine Learning Paradigms for Next-Generation Wireless Networks,” IEEE Wireless Communications, 2017.

[15] T. E. Bogale, X. Wang, and L. B. Le, “Machine Intelligence Techniques for Next-Generation Context-Aware Wireless Networks,” ITU Special Issue: The impact of Artificial Intelligence (AI) on communication networks and services., vol. 1, 2018