National Conference on Emerging Technologies for Sustainable Engineering & Management (NCETSEM'18)-2018

RESEARCH ARTICLE OPEN ACCESS

Advancement in Networking with Machine Learning Techniques

[1]Julie Joan N.S,[2]Dr.D.Rajinigirinath,[3]Dr.K.Poonkavithai
[1]Ph.D Scholar, [2]Professor and HOD
Department of Computer Science and Engineering
[3]KCG College of Technology,OMR,Chennai.
Sri Muthukumaran Institute of Technology
Mangadu, Chennai.
[1]juliejoan107@gmail.com, [2]dgirinathicici@gmail.com

Abstract:

Machine learning has been utilized in every possible field to leverage its astonishing puissance. The networking and distributed computing system is the key infrastructure which provides computational resources in efficient manner for machine learning. Networking itself can benefit from this promising technology. This article fixates on the application of MLN, which can not only avail solve the intractable old network questions but stimulate incipient network applications. In this article, we summarize the rudimentary workflow to explicate to apply machine learning technology in the network domain. This paper provides the latest representative advances with explications of their design principles and benefits. The advancement in machine learning is categorized into several network design objectives and the detailed information for the performance of machine learning work flow. This paper provides a wide research guideline on networking with machine learning to avail incentive researchers to develop innovative algorithms, standards and frameworks.

Keywords—**Machine Learning, Quality of Service, Bayesian classifier, Radial base function, Clustering, Redundancy.**

I. INTRODUCTION

ith the prosperous advancement of the Internet, organizing Wresearch has polarized a plenty of consideration in the previous quite a few years both in the scholarly community and industry. Specialists and system administrators can confront sundry sorts of systems (e.g., wired or remote) and applications (e.g., organize security and live gushing [1]). Each system application furthermore has its own particular highlights and execution requirements, which may change powerfully with time and space. In light of the assorted variety and involution of systems, downright calculations are regularly worked for various system situations predicated on the system attributes and utilizer requests. Creating effective calculations and frameworks to manage involute binds in various system situations is a testing errand. As of late, machine learning (ML) strategies have made leaps forward in an assortment of utilization zones, for example, bioinformatics, verbalization apperception and PC vision. Machine learning tries to develop calculations and models that machine learning calculations by and large fall into three classes: directed learning (SL), unsupervised learning (USL) and support learning (RL). All the more solidly, SL calculations figure out how to direct assignment or relapse undertakings from marked information, while USL calculations focus on consigning the example sets into various gatherings (i.e., bunches) with unlabeled information. In RL calculations, specialists figure out how to locate the best activity arrangement to expand the cumulated compensate (i.e., target work) by collaborating with nature. The most recent leaps forward, including profound learning (DL), exchange learning and generative ill-disposed systems (GAN), withal give potential research and application bearings in an inconceivable manner. Managing involute predicaments is a standout amongst the most noteworthy points of interest of machine learning. For a few assignments requiring

transfer, relapse and basic leadership, machine learning may perform proximate to or far and away superior to individuals.

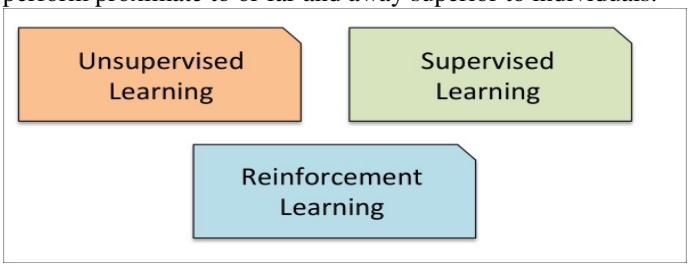

Fig 1: Types in Machine Learning

Since the system field frequently optically observes involute problems that request productive arrangements, it is promising to bring machine learning calculations into the system space to use the intense ML resources for higher system execution. The joining of machine learning into arrange outline and administration withal gives the likelihood of causing early system applications. Genuinely, ML strategies have been used in the system field for quite a while. The current advancement of foundations (e.g., computational creations like GPU and TPU, ML libraries like Tensor stream and Scikit-Learn) and appropriated information handling structures (e.g., Hadoop and Spark) gives a decent chance to release the enchantment energy of machine learning for seeking after the beginning potential in arrange frameworks. Completely, machine learning for systems administration (MLN) is consistent and proficient for the accompanying reasons. In the first place, as the best kenned capacities of ML, assignment and foretell assume simple however fundamental parts in arrange scrapes, for example, interruption identification and execution forecast [23]. In reconciliation, machine learning can withal benefit basic leadership, which will encourage organize planning and parameter adjustment as indicated by the present conditions of the earth. Second, numerous system binds need to cooperate with astounded framework situations. It isn't effortless to manufacture exact or explanatory models to speak to involute framework deportments, for example, stack transmuting examples of CDN and throughput attributes. Machine learning can give an expected model of these frameworks with worthy accuracy. Determinately, each system situation may have diverse trademark (e.g., activity examples and system states) and scientists frequently need to explain the scrape for every situation freely. Machine learning may give nascent conceivable outcomes to build the summed up demonstrate by means of a uniform preparing strategy. Among endeavors in MLN, profound learning has been explored and connected to give end-to-end arrangements.

The most recent work in conducts a thorough review on point of reference endeavors that apply profound learning innovation in arrange related territories. In this article, we explore how machine learning innovation can profit arrange outline and improvement. Completely, we abridge the regular work processes and essentials for applying machine learning methods in the system space, which could give a basic yet pragmatic rule for analysts to have a speedy begin in the zone of MLN. At that point we give a specific overview of the considerable systems administration propels with the stronghold of machine learning innovation, a large portion of which have been distributed over the most recent three years. We amass these advances into a few regular systems administration fields and explain how these earlier endeavors perform at each progression of the MLN work process. At that point the chances of this rising between teach zone are talked about.

II. MLN WORK FLOW

By applying machine learning in the network field, including quandary formulation, data collection, data analysis, model construction, model validation, deployment and inference. These stages are not independent but have inner relationships. This workflow is very kindred to the traditional workflow for machine learning, as network quandaries are still applications that machine learning can play a role.

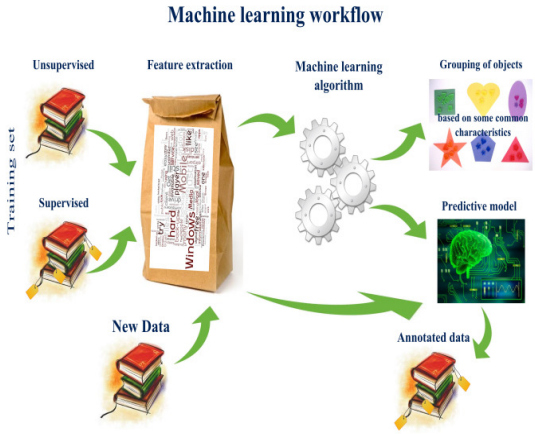

Fig 2: Machine learning workflow

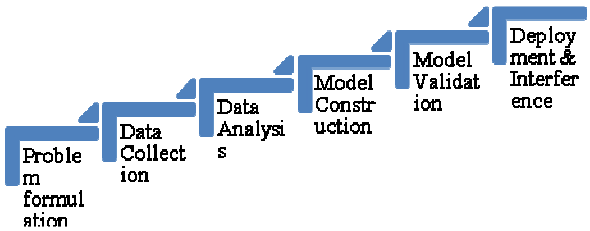

Fig 3: Stages in Machine learning

A. *Problem Formulation*

Since the training process of machine learning is often time consuming and involves high cost, it is important to correctly abstract and formulate the problem at the first step of MLN. A

target problem can be classified into one of the machine learning categories, such as classification, clustering and decision making. This helps decide what kind of and the amount of data to collect and the learning model to select. An improper problem abstraction may provide an unsuitable learning model, which can result in unsatisfactory learning performance. For example, it is better to cast the optimal quality of service (QoS) for live streaming into a real-time exploration-exploitation process rather than as a prediction-based problem [6] to well match the application characteristics.

B. Data Collection

The goal of this step is to collect a large amount of representative network data without bias. The network data (e.g., traffic traces and session logs with performance metrics) are recorded from different network layers according to the application needs. For example, the traffic classification problem often requires datasets containing packet-level traces labeled with corresponding application classes.

C. Data Analysis

Every network problem has its own characteristics and is impacted by many factors, but only several factors (i.e., feature) have the most effect on the target network performance metric. For instance, RTT and the inter-arrival time of ACK may be the critical features in choosing the best size of the TCP congestion window. In the learning paradigm, finding proper features is the key to fully unleashing the potential of data. This step attempts to extract the effective features of a network problem by analyzing the historical data samples, which can be regarded as a feature engineering process in the machine learning community. Before feature extraction, it is important to preprocess and clean raw data, through processes such as normalization, discretization, and missing value completion. Extracting features from cleaned data often needs domain-specific knowledge and insights of the target network problem [5], which is not only difficult but time-consuming. Thus in some cases deep learning can be a good choice to help automate feature extractions [24].

D. Model Construction

Model construction involves model selection, training and tuning. A suitable learning model or algorithm needs to be selected according to the size of the dataset, typical characteristics of a network scenario, the problem category, and so on. For example, accurate throughput prediction can improve the bitrate adaption of Internet video, and a Hidden-Markov model may be selected for prediction due to the dynamic patterns of stateful throughput. Then the historical data will be used to train a model with hyper-parameter tuning, which will take a long period of time in the offline phase. The parameter tuning process still lacks enough theoretical guidance, and often involves a search in a large space to find acceptable parameters or to tune by personal experiences.

E. Model Validation

Offline validation is an indispensable step in the MLN workflow to evaluate whether the learning algorithm works well enough. During this step, cross validation is usually used to test the overall accuracy of the model in order to show if the model is over fitting or under-fitting. This provides good guidance on how to optimize the model, e.g., increasing the data volume and reducing model complexity when there exists over fitting. Analyzing wrong samples helps find the reasons for errors to determine whether the model and the features are proper or the data are representative enough for a problem. The procedures in the previous steps may need to be re-taken based on the error sources.

F. Deployment and Inference

When implementing the learning model in an operational network environment, some practical issues should be considered. Since there are often limitations on computation or energy resources and requirements on the response time, the tradeoff between accuracy and the overhead is important for the performance of the practical network system. In addition, machine learning often works in a best-effort way and does not provide any performance guarantee, which requires system designers to consider fault tolerance. Finally, practical applications often require the learning system to take real-time input, and obtain the inference and output the corresponding policy online.

III. RECENT ADVANCEMENT

Recent breakthroughs of deep learning and other promising machine learning techniques have a non-ignorable influence on incipient endeavors of the network community. Subsisting efforts have led to several considerable advances in different subfields of networking. To illustrate the relationship between these au courant advances and the MLN workflow, show how they perform at each step of the MLN workflow. Without ML techniques, the typical solutions for these advances are involved with time-series analytics, statistical methods and rule-predicated heuristic algorithms, which are often more interpretable and more facile to implement. However, ML-predicated methods have a more vigorous faculty to provide a fine-grained strategy and can achieve higher presage precision by extracting information from historical data. As an astronomically immense challenge of ML-predicated solutions, the feasibility quandary is withal discussed in this section. Information Cognition Since data are the fundamental resource for MLN, information (data) cognition with high efficiency is critical to capture the network characteristics and monitor network performance. However, due to the involute nature of subsisting networks and the inhibitions of quantification implements and architectures, it is still not facile to access some types of data (e.g., trace route and traffic matrix) within acceptable granularity and cost. With its

capability for presage, machine learning can avail evaluate network reliability or the probability of a certain network state.As the first example, Internet route quantifications avail monitor network running states and troubleshoot performance quandaries. However, due to inadequate utilizable vantage points (VP) and a circumscribed probing budget, it is infeasible to execute each route query because the query may not match any antecedently quantified path or the path may have transmuted.

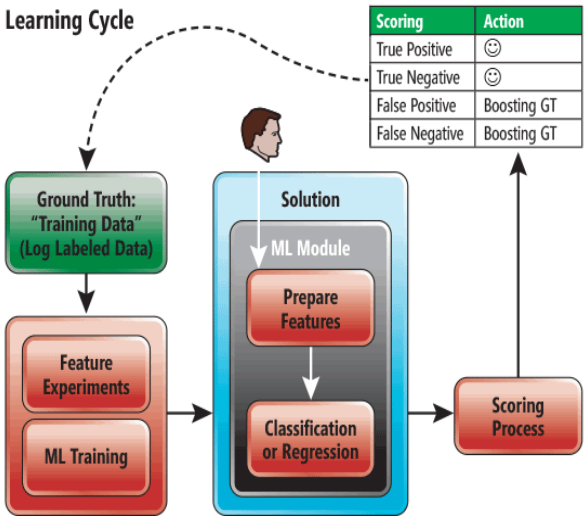

Fig 3: ML learning cycle

Sibyl [10] endeavors to prognosticate the unseen paths and assign confidence to them by utilizing a supervised machine learning technique called Rule Fit. The cognition relies on data acquisition, and MLN withal requires an incipient scheme of data cognition. In MLN, it often needs to maintain an au courant ecumenical network state and perform real-time replications to client demands, which needs to quantify and amass the information in the core network. In order to enable the network to perform diagnostics and make decisions by itself with the avail of machine learning or cognitive algorithms, a different network architecture, the Erudition Plane [12], was presented that can achieve automatic information cognition, which has inspired the following efforts that leverage ML or data-driven methods to enhance network performance. Traffic Prognostication and Relegation are two of the earliest machine learning applications in the networking field. Because of the well formulated question descriptions and authoritatively mandates from sundry subfields of networking, studies of the two topics always maintain a certain degree of popularity.

A. *Traffic Prognostication*

As a consequential research quandary, the precise estimation of traffic volume (e.g., the traffic matrix) is benign to congestion control, resource allocation, network routing, and even high-level live streaming applications. There are mainly two directions of research, time series analysis and network tomography, which can be simply distinguished depending on if it conducts traffic prognostication with direct observations or not. However, it is extravagant to directly measure traffic volume, especially in an immensely colossal-scale high speed network environment. Many subsisting studies fixate on reducing the quantification cost by utilizing indirect metrics rather than only endeavoring different ML algorithms. There are two methods to handle this quandary. One is to take more human effort to develop sophisticated algorithms by exploring domain-concrete cognizance and undiscovered data patterns. Another method is inspired by the terminus-to-end deep learning approach. It takes some facilely obtained information (e.g., bits of a header in the first few flow packets) as direct input and extract features automatically with the avail of the cognition model [9].

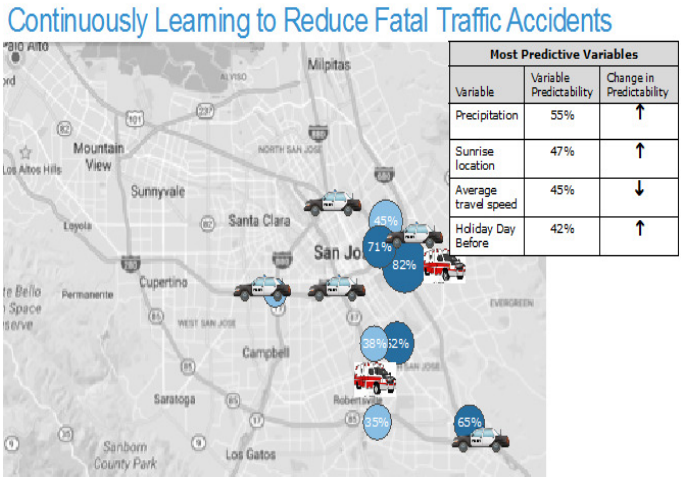

Fig 4: Leveraging Advanced Analytics to Power Digital Transformation

B. *Traffic Relegation*

As a principal work part in organize administration and security frameworks, activity assignment matches arrange applications and conventions with the relating movement streams. The customary activity assignment strategies incorporate the port-predicated approach and the payload-predicated approach. The port-predicated approach has been turned out to be insufficient because of unfixed or reused port assignments, while the payload-predicated approach experiences protection difficulties caused by profound parcel review, which can even bomb within the sight of encoded movement. Thus, machine learning approaches predicated on factual highlights have been broadly examined as of late, particularly in the system security domain. By and large, these investigations run from all-kenned assignment situations to a more credible circumstance with obscure movement (e.g., zero day application activity [7]). This exploration guide is extremely homogeneous to the machine taking in innovation that advances from regulated figuring out how to unsupervised and semi-directed realizing, which can be dealt with as a pioneer worldview to import machine learning

into systems administration fields.

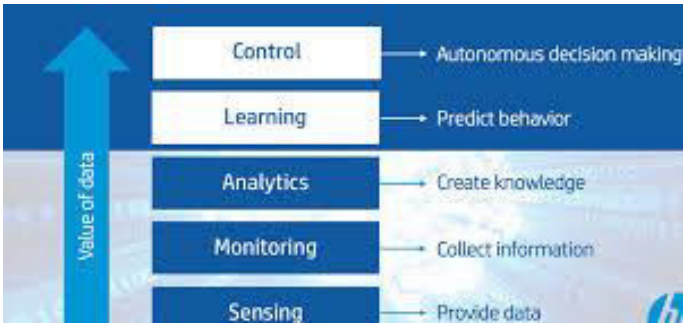

Fig 5: Payload prediction in Traffic Relegation

IV. Application of Machine Learning Approaches for Internet Traffic Classification

A. *Supervised (classification) Methods Supervised techniques as follows:*

1. Bayes Net Method

Bayes Net approach is generally known as Belief Network. It is a Probabilistic model which uses the graph model to represent the set of random variables and their conditional dependencies. Bayes Net uses the concept of directed acyclic graph (DAG) to represent the set, in which each node represent a variable and edges among the nodes represent the relative dependencies between random variables and these relative dependencies in the graph are calculated by well-known statistical and computational methods. There are two phases of Bayes net approach first phase is learning of network structure, in which uses various types of search algorithm like hill climbing, for identified a good network structure and second is estimate probabilistic table for each random variable. In 2012 JaspreetKaur et al. [5] uses five well known machine learning algorithms (Naïve Bayes, C4.5, RBF, MLP, Bayes Net) to classify the educational and non-educational websites. In this paper they use two types of data sets for classification, one is a full feature dataset and another one is reduced feature datasets with CFS (Correlation based feature selection) and CON (Consistency based feature selection) feature reduction algorithms. In case of the full feature dataset, the efficiency was decreases due to large number of features and that's why they use reduced feature dataset. In this Bayes Net gives 96.6% classification accuracy with full feature dataset but the number of samples in a dataset is low.

2. Feed forward Neural Network classifier:

The feed forward neural network was the first and simplest type of artificial neural network methods. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network. In 2011 Wengang Zhou et al. [13] proposed an approach based on a feed forward neural network for accurate traffic classification and combined it with FCBF (Fast Correlation Based Feature) feature selection algorithm. FCBF is used for eliminating the redundant features and chosen the valuable features and feed forward neural network work as classifier. In 2007 Tom Auld et al. [14] proposed a novel approach of machine learning based on a Bayesian neural network for Internet traffic classification. In this they use a Bayesian framework using a neural network model to classify traffic without accessing the port host umber information, data is collected from a set of flows which taken from two distinct days eight months apart from each other and they consist of ten sets of classified TCP (transport control protocol) traffic flow, author reported 95.3% classification accuracy with large number of training samples.

3. Naive Bayes classifier:

A Naive Bayes classifier is a basic classifier in view of applying Bayesian hypothesis with solid and frail autonomy supposition. In the least complex way a Naive Bayes classifier expect that the nearness or nonattendance of a specific element of a class has no connection with the nearness or nonappearance of some other highlights given in a similar class variable. There is no structure learning strategy is required in innocent Bayes classifier so it is anything but difficult to build as contrast with different classifiers.

In 2013 Jun Zhang et al. [2] utilizes arrange web movement by Aggregating Correlated Naive Bayes Prediction and get high exactness with this approach. They proposed new (sack of-stream) BoF-based activity characterization system is to total the Naive Bayes (NB) forecasts of the connected stream. They proposed another approach of grouping to use the data among the associated movement streams created by the activity.

Prior, conditional, and joint probability for Random variables

- Prior Probability: $P(X)$
- Conditional Probability: $(PX_1 \mid PX_2)$, $(PX_2 \mid PX_1)$
- Joint Probablity: $X=(X_1,X_2)$, $P(X)=P(X_1,X_2)$
- Relationship: $P(X_1,X_2)$:$(PX_1 \mid PX_2)$ $P(X_1)$, $(PX_2 \mid PX_1)$ $P(X_1)$
- Independence: $P(X_1 \mid X_2)= P(X_2)$, $(PX_1 \mid PX_2) = P(X_1)$, $P(X_1,X_2)=P(X_1)P(X_2)$

$$P(C \mid \mathbf{X}) = \frac{P(\mathbf{X} \mid C)P(C)}{P(\mathbf{X})} \quad Posterior = \frac{Likelihood \times Prior}{Evidence}$$

In the approach of classification there are two steps, in a first step the single naïve Bayes predictor generates the posteriori class-conditional probabilities or each flow and in a second step the aggregated predictor aggregates the flow predictions to determine the final class for BoFs. Random Forest provided the best result as compare to all other algorithms, it gives 99.8% classification accuracy and Decision Stump provide lowest training time is 0.05 seconds but they uses less number of data samples.

4. C4.5 Decision tree classifier:

C4.5 is a prevalent choice tree Machine Learning calculation used to create Univarate choice tree. C4.5 is an improvement of Iterative Dichotomize 3 (ID3) calculation which is utilized to discover basic choice trees. C4.5 is likewise called a Statistical Classifier in view of its great capacity of arrangement. C4.5 settles on choice trees from an arrangement of preparing information tests, with the assistance of data entropy idea. The preparation informational collection contains of a more prominent number of preparing tests which are portrayed by various characteristics and it additionally comprises of the objective class. C4.5 chooses a specific trait of the information at every hub of the tree which is utilized to part its arrangement of information tests into subsets in some class. It depends on the measure of standardized data pick up that is gotten by choosing a trait for part the information. The quality with the most astounding standardized data pick up is picked and settled on a choice. From that point forward, the C4.5 calculation rehashes a similar activity on the littler subsets. C4.5 has made different enhancements to ID3 like it can deal with both consistent traits and discrete properties, it can deal with preparing information with missing quality esteems, it can likewise deal with characteristics with varying expenses and so forth. In 2012 Dong Shi et al. [8] they used to order and recognize the system with both managed and unsupervised learning methods. They utilize two sorts of dataset full highlights based and streamlined highlights based. Here investigation result demonstrates that the directed ML calculations give better outcome with include diminishment calculations as contrast with unsupervised ML calculations. Reenactment result closes 99% order precision with C4.5 calculation. In 2011 Kuldeep Singh et al. [11] they utilize five machine learning calculation (MLP, RBF, C4.5, Bayes Net, Naïve Bayes) as classifier to characterize the constant web activity order alongside utilizing diverse element determination calculations which are Correlation based FS, Consistency based FS and Principal Components Analysis based FS calculations. In this the connection based FS is utilized to distinguish and evacuate repetitive and unimportant highlights as could reasonably be expected. It utilizes an assessment methodology that looks at the helpfulness of individual component alongside the level of between relationship among the highlights. The consistency based FS is utilized to assess the subset of highlights at the same time and select ideal subset. The Principal Components Analysis (PCA) based FS maps the information focuses from a high dimensional space to a low dimensional space while keeping all the pertinent direct structure unaltered. In this paper C4.5 ML calculation gives the best outcome in all above it gives more than 90% characterization precision. The Net Flow gathering channel, pre-preparing and order are done in an online way as opposed to disconnected way; creator revealed 3.5 % mistakes on a normal.

5. C5.0 Machine Learning Algorithm

The C5.0 algorithm and its predecessor C4.5 described in [26] attempt to predict a dependent attribute by finding optimal value ranges of an independent set of attributes. At each stage of iteration, the algorithm aims to minimize information entropy by finding a single attribute that best separates different classes from each other. The process continues until the whole sample space is split into a decision tree isolating each class. Hence, in a sample space comprising application flow classes, if training data is given by pre-classified samples given by vector , each sample flow may consist of a -dimensional vector, where represents independent attributes which are used to identify the class in which falls:

$$S=\{f_1,f_2,f_3,\ldots f_n\}$$
$$F=\{f_1,f_2,f_3,\ldots f_z\}$$

6. Radial Basis Function Neural Network

Spiral premise work (RBF) systems have three layers engineering: an info layer, a concealed layer with a non-straight RBF work it an actuation work and a direct yield layer. Spiral Basis Function (RBF) is a multilayer sustain forward fake neural system which utilizes outspread premise capacities at each shrouded layer neuron. The yield pick up of this RBF neural system is a weighted straight superposition of all these premise capacities. The fundamental model of RBF neural system is appeared in Fig. 3. In this system, weights for input-concealed layer interconnections are settled, while the weights for shrouded yield layer interconnections are trainable. Following information - yield mapping capacity 1 as:

$$Y(x) = \sum_{n=0}^{m} W_i\, U\left(\left|\left|X - X_i\right|\right|\right) \qquad (1)$$

Where, U () premise capacity of shrouded layer which is connected at every neuron of concealed layer, M premise capacities comprising of the Euclidean separation between connected information sources X, Y(x) yield mapping capacity.

In 2013 Mussab M. Hassan et al. [5] utilizes half and half measurable activity classifier to order the P2P (shared) movement. Here additionally the works in two stages, initially disconnected heuristics learning corpus age and second is online measurable grouping, In this initial segment, Heuristic order the activity stream and second part machine learning calculation are utilized to characterize arrange movement. They apply 64 ML calculations to order movement and find that RBF ML calculations give great outcome. Regulated machine learning calculation gives the best outcome in rush hour gridlock grouping as contrast with unsupervised machine learning calculations.

B. Unsupervised (Clustering) Methods:

Clustering is an unsupervised machine learning approach, in which produces cluster samples according to the similarity of flow feature values. The main objective of clustering is to group the packets that have similar patterns. In clustering instances having similar properties can be put into the same group. There are three conditions are made when grouping the packets, which are as follows:

• If group is exclusive then packets can be put into a single group.
• If packets having the properties of multiple groups then packets can be put into many groups.
• If the group can be probabilistic then the packet can belong to a group with a fixed probability.

1. *DBSCAN based Approach*

DBSCAN (Density-based spatial clustering of applications with noise) is a data clustering algorithm. It is a density based clustering algorithm; it finds the number of clusters starting from the estimated density distribution of the corresponding nodes. There are two input parameters here, first is epsilon (Eps) and second is minimum number of points (minPts). Epsilon (Eps) is the space around a particular point object that is used to determine its Eps-neighborhood for a given point p and minPts is the minimum number of points within its eps-neighborhood. The concept of DBSCAN based on two parameter density-reachablity and density-connectivity, which formed the clusters in DBSCAN algorithm. Density-reachablity, a point p is density reachable from a point q in respect of Eps and minPts if there is a all points like p1,p2,p3…….pn are reachable from point q i.e. p1=q, p2=q……pn=q is called density reachablity. Density-connected, a point p is density connected to a point q if both points are density reachable from an object point o. In 2013 ShezadShaikh et al. [3] they classify network flows using DBSCAN algorithm. In this proposed method, they performed two operations first is clustering and the second is classification. In clustering, the large dataset is divided into small sets of similar data.In this research, they got 87% overall classification accuracy.

2. *Expectation Maximization based (Autoclass) Approach*

It is an iterative technique for looking amplifies probability parameters and produces bunches. There are chiefly two stages in desire expansion technique, first is Expectation step and the second one is Maximization step. In initial step evaluate that what parameter is utilizing arbitrary numbers and in a moment step the utilizations mean and change to re-assess the parameter, this procedure consistently continues till then they came to with a nearby amplify and this procedure is rehashed. In 2006 Jeffrey Erman and Anirban et al connected desire augmentation (EM) based grouping calculations for web movement characterization and furthermore arrange the web activity by utilizing a credulous Bayes order approach and relative investigation of results demonstrates that EM approach give better outcome as contrast with innocent Bayes. The similar outcomes demonstrate that 91% characterization exactness.

3. *K-Means based Approach*

K-Means grouping calculation is an apportioned based calculation; it divided objects of a dataset into K disjoint subsets. It boosts the homogeneity of the group and limits the square-blunder where square-mistake figured as the separation between each protest and the inside or mean of a bunch. The focuses of K-bunch are at first picked haphazardly and after that dataset parceled into closest group. K-Means iteratively figures new focuses and bunches individually and this procedure proceeds until the point that the groups are balanced out. In 2006 Jeffery Erman et al. [14] utilizes two machine learning calculations (K-Mean and DBSCAN) to order the web movement. Execution of the two calculations contrasted and Auto class (EM) calculation and they find that K-Means and DBSCAN calculations perform well when contrasted with EM calculation; result detailed that with K-Means grouping calculation gives 85% order exactness.

Performance of various method

Paper Ref No	year	Classification Method used	Feature selection algorithm	Classification Accuracy	Dataset
26	2016	C5.0 ML classifier	N/A	96.67 %	Compuer networks and communications
12	2011	RBF,C4.5,MLP, Bayes Net, Naïve Bayes	Correlation Based, Consistency Based and Principal Components Analysis	93.66 %	Proprietary Hand Classified Traces
16	2010	Bayesian Network, C4.5 and MLP	N/A	Above 95%	National Academic Network of Turkey (ULAKNET)
17	2009	C4.5	N/A	Approx. 95%	UTM campus network
6	2012	RBF,C4.5,MLP,Bayes Net, Naïve Bayes	CFS and CON algorithm	96.6%	Captured from an educational institution
14	2011	Feed-forward Neural	FCBF(Fast Correlation-Based	95%	Proprietary Hand Classified Traces

		Network	Filter) Algorithm		
21	2007	Bayesian trained neiral network	N/A	95%	Captured at Genome campus
12	2011	RBF,C4.5,MLP, Bayes Net, Naïve Bayes	Correlation Based, Consistency Based and Principal Components Analysis	93.66 %	Proprietary Hand Classified Traces

Table 1: Performance of various methods

V. CONCLUSION

In heterogeneity of networking systems, it is imperative to embrace machine learning techniques in the networking domain for potential breakthroughs. However, it is not easy for networking researchers to take it into practice due to the lack of machine learning related experiences and insufficient directions. In this article, we present a basic workflow to provide researchers with a practical guideline to explore new machine learning paradigms for future networking research. For a deeper comprehension, we summarize the latest advances in machine learning for networking, which covers multiple important network techniques, including measurement, prediction and scheduling. Moreover, numerous issues are still open and we shed light on the opportunities that need further research effort from both the networking and machine learning perspectives.

REFERENCES:

[1] YiboXue and et al. Traffic Classification: Issues and Challenges; International conference on computing, networking and communication (ICNC); IEEE, 2013; p. 545-549.

[2] Jun Zhang and et al. Internet Traffic Classification by Aggregating Correlated Naive Bayes Predictions; IEEE transactions on information forensics and security; vol. 8,;2013;p. 5-15.

[3] Shezad Shaikh1and et al. Implementation of DBSCAN Algorithm for Internet Traffic Classification; International Journal of Computer Science and Information Technology Research (IJCSITR); 2013; p. 25-32.

[4] Mussab M. Hassan and Muhammad N. Marsono. A Hybrid Heuristics-Statistical Peer-to-peer Traffic Classifier; International conference on computer system and industrial information (ICCSII); 2013; IEEE; p. 1-6.

[5] JaspreetKaur and et al. Internet Traffic Classification for Education Institutions Using Machine Learning; International Journal of Intelligent Systems and Applications (IJISA); 2012; MECS; vol. 4; p. 37-45.

[6] S. Agrawal and et al. Machine Learning Classifier for Internet Traffic from Academic Perspective; International Conference on Recent Advances and Future Trends in Information Technology (RAFIT); 2012; IJCA; p. 4-9.

[7] HamzaAwadHamza and et al. Taxonomy of Machine Learning Algorithms to classify realtime Interactive applications; International Journal of Computer Networks and Wireless Communications (IJCNWC); 2012; IRACST; p. 69-73.

[8] DONG Shi and et al. The Study of Network Traffic Identification Based on Machine Learning Algorithm; Fourth International Conference on Computational Intelligence and Communication Networks; 2012; IEEE; p. 205-208.

[9] Bin Hu and Yi Shen. Machine Learning Based Network Traffic Classification: A survey; Journal of Information & Computational Science; 2012; p. 3161- 3170.

[10] LiTing Hu and LiJun Zhang. Real-time Internet Traffic Identification Based on Decision Tree; World Automation Congress (WAC); 2012; IEEE; p. 1-3.

[11] Kuldeep Singh and S. Agrawal. Performance Evaluation of Five Machine Learning Algorithms and Three Feature Selection Algorithms for IP Traffic Classification; Evolution in Networks and Computer Communication; ; IJCA; 2011; p. 25-32.

[12] Y. Sun et al., "CS2P: Improving Video Bitrate Selection and Adaptation with Data-Driven Throughput Prediction," Proc. SIGCOMM 2016, ACM, pp. 272–85.

[13] Yu Wang and et al. Internet Traffic Classification Using Machine Learning: A Token-based Approach; Fourteenth International Conference on Computational Science and Engineering (CSE); 2011;IEEE; p. 285-289.

[14] Jeffrey Erman and et al. Internet Traffic Identification using Machine Learning; Global Telecommunications Conference(GLOBECOM);IEEE; 2006; p. 1-6.