International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No.04, March 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

MACHINE LEARNING IN THE INTERNET OF THINGS – STANDARDIZING IOT FOR BETTER LEARNING

Rashid Ashraf Malik[1], Asif Iqbal Kawoosa[2], Ovais Shafi Zargar[3]

[1]RashidAshraf Malik is Assistant Professor in Computer Applications,
Amar Sigh College Srinagar,(India)
[2]Asif Iqbal Kawoosa is Assistant Professor in Computer Applications,
Amar Sigh College Srinagar, (India)
[3]Ovais Shafi Zargar is Assistant Professor in Computer Applications,
Amar Sigh College Srinagar, (India)

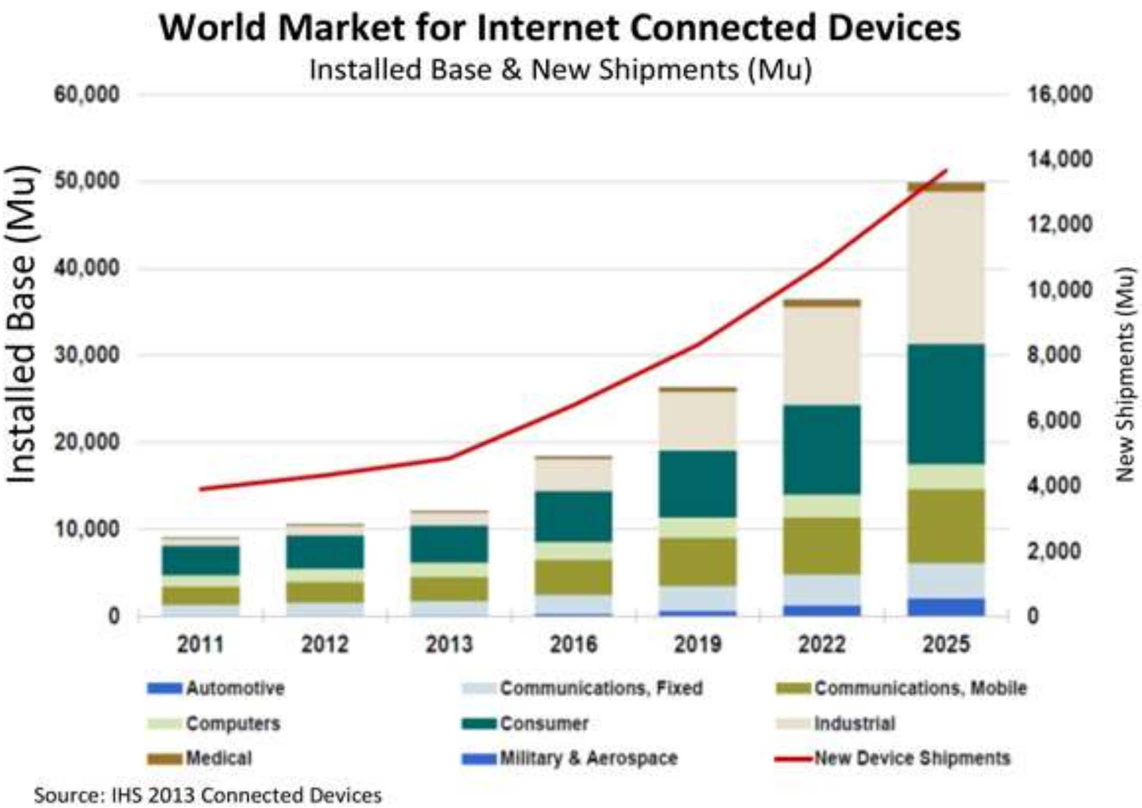
ABSTRACT

The prodigious scale of the machine driven environment has necessitated that machines shall not rely predominantly on the humans to take decisions. The pervasive computing has reduced machine to the size of a 30mm diameter like Bluetooth Low Energy Beacon, and the Internet of Things (IoT) is making these extremely small devices to connect and communicate with each other. Millions of such devices are implanted every day in everything, thus making it difficult for the humans to configure, deploy and refactor each device individually. Machine Learning is one area that enables these devices to learn, and take decisions without being explicitly programmed each time. One of the biggest challenges in machine learning, particularly amongst the miniature computing environment is the interoperability issues of these devices. IoT is throwing up millions of such devices in the computing environment at a very fast scale, thus, sometimes compromising the standards in terms of protocols, software stack, communication, bandwidth, data processing. Therefore, inhibiting the patterns to build up, and proving to be detrimental in taking decision for the machine learning. The aim of this paper is to put forward a mechanism that will help to do away with the interoperability issues and standardize the communication, especially at the software level. A standard is envisaged in this paper, that all the devices located in a particular cluster of the IoT or any connected machine environment shall have to comply with. The futuristic plan is also to develop a python module that can be run by the communication manager to check the device compatibility viz-viz the communication environment. This module will also be able to leverage the power of TensorFlow- An open source machine learning framework and help the users to do specific tasks in TensorFlow, which is otherwise massive workflow if it is done directly via TensorFlow. The Python module will give an option to call TensorFlow API discretely, and will also provide NoSQL type backup facility for data analysis.

Keywords: IoT, Machine Learning, Pervasive Computing, TensorFlow, API, Apache SINGA, Caffe

I.INTRODUCTION

Machine Learning(ML) and the Internet of Things(IOT) are the predominant technological jargons we get to hear day in and day out. The success and usability of both these technologies have paved a way to easily integrate our lives with the inevitable technological ecosystem. Given all the hype and buzz around machine learning and IoT, it can be difficult to cut through the noise and understand where the actual value lies, even though these technologies have penetrated deep into our work places, home, travel, entertainment, healthcare, education etc, and have benefitted us immensely. Therefore, there is a need to integrated ML and IoT to reap benefits, and create an enabling environment for the millions of devices that are connecting humans to the machine, and vice versa. The data models being used in Machine Learning(ML) that are typical of traditional data analytics are often static and of limited use in addressing fast-changing and unstructured data. When it comes to IoT, it's often necessary to identify correlations between dozens of sensor inputs and external factors that are rapidly producing millions of data points. One of the particular problem in configuring IoT clusters as of now is the plenitude of human action required to actually keep the IoT infrastructure working, particularly in decision making. This is where machine learning can be put in place and allow billions of sensors and other computing devices of IoT cluster to learn, and subsequently, take decisions on their own.



Gartner estimates that there will be 50 billion IoT devices by 2020. This represents an enormous opportunity for connecting everything, and enable data flow between devices of any form factor.

II.DEFINING THE IOT STACK

The IoT represents the connected world, and includes myriad devices designed to enrich our life experiences. It covers virtually every element of the Internet and networking market segments (which is why Cisco likes to refer to the IoT as "The Internet of Everything"). The full range of use cases for IoT devices can be categorized into three main areas:

2.1 ***Traditional Internet:*** This is the Internet we all know. It will be the traffic backbone for almost all new devices and many already-connected devices. PCs, smartphones, and tablets are part of the IoT.

2.2 ***Industrial Internet:*** This is where most of the machine-to-machine (M2M) exchanges will take place and where sensors and other data transmission and receiving devices will reside. Here, embedded systems will be able to operate without user interaction, in everything from traffic lights and power grids, to city infrastructure and industrial machinery.

2.3 ***Consumer Internet:*** This covers things like wearables and smart devices, where much of the growth will take place as new products emerge—smart refrigerators, advanced watches, and wrist devices, as well as visual devices (e.g., Google Glass and Oculus Rift). The IoT will bring connectivity, communication, and data gathering to existing devices, such as cars and home appliances, and allow for massive data collection without any human interaction. For example, the IoT-enabled car of the future will feature dozens of sensors, providing drivers with an active and useful health/performance assessment at all times. Instead of useless "check engine" lights that leave you guessing at what may be wrong, cars of the future will alert drivers when the tire pressure is low or the muffler is about to fail before anything actually fails or stops working. Or take the IoT-enabled refrigerator of the future, which will alert users of food spoilage, food recalls, or if a certain food staple is running low, such as milk or eggs. Flat-panel manufacturers are also working to replace the solid door so that users can see inside the refrigerator without opening it, and/or receive advertisements for sales at the local market and other Internet information displayed right on the refrigerator door.

III.THE UNDERLYING TECHNOLOGY AND INHERITED PROBLEMS

The IoT is fundamentally based upon following set of technologies and consequently inherits all the problems that these base technologies may exhibit.

3.1 ***Extreme miniaturization:*** Chips have shrunk over the years, but to make the IoT work, we needed something even smaller than what's found in a smartphone. With process technology down to 14 nm and heading to 10 nm, chips are getting to the point of not being easily seen. Early in 2014, Freescale introduced an IoT-ready chip that could fit inside the divot of a golf ball.

Problem Inherited and Recommended Solution: With this reduced chip size and the demand of IoT devices being extremely small, has to larger extend introduced a speed mismatch scenario between large form factor devices and small form factor devices. We recommend that all the devices that are being put in the IoT shall necessarily satisfy a minimum computing speed in terms of processing and storage. In future, we also aim to build a python module to test all the configurations of a device, before it can participate in the IoT cluster.

3.2 ***IPv6:*** IPv6 has 340 billion addresses; therefore, essentially, they should never run out. While as IPv4 provides 4.3 billion addresses, and many are unavailable to the public.

Problem Inherited and Recommended Solution: There is no check whether a device is communicating over the network either using IPv4 or IPv6. Let all the devices in a particular cluster comply with the same IP addressing scheme, either all the devices on IPV4 or IPV6. But ultimately all the devices will have to move to IPV6, especially when we talk of communication between IoT devices of multiple clusters.

3.3 ***Wi-Fi and LTE ubiquity***: For the IoT to be wireless, high-speed LTE will need to cover much of the nation, and Wi-Fi in homes will need to reach critical mass. Today, Wi-Fi modems are a standard feature on cable modems and LTE covers most of the nation. Verizon Wireless claims to cover 97% of America, while AT&T claims about 300 million of the nation's 320 million POPs are covered by its LTE network.

3.4 ***Scaled-out data centers:*** With 50 billion devices expected to constitute the IoT, the amount of data will be massive. The CSO of machine-data specialist Splunk told Gigaom, a research and new organization, that the average smart home will generate about 1 GB of data per week. This will put tremendous strain on data centers and will require far bigger data pipes and edge servers to handle and act upon all of the data coming in. The data center of the future will be much more active than traditional data centers, which often just serve as big data stores.

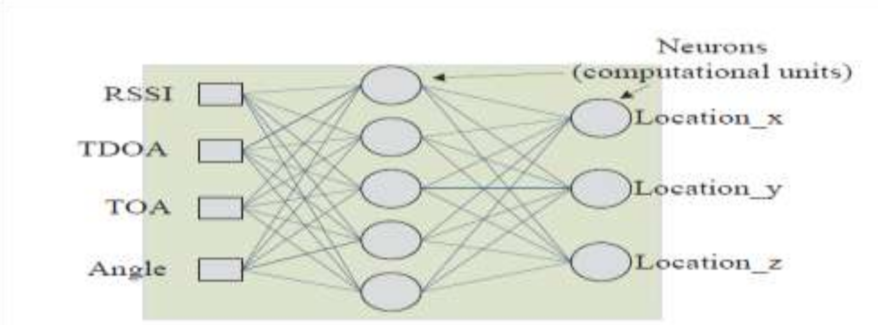
IV.THE BIG IMPEDIMENT IN IOT

With the billions of devices participating in IoT ecosystem, the architectural neutrality and the reduced human interference, or action is quite a tedious task. But if we are able to bring IoT ecosystem on a platform-independent terms, and enable IoT devices to look after themselves in terms of software updates, routing, data sharing, security and data analytics without human interference, or highly minimized human interference will make the transition easier from a connected world of fewer devices to the connected world of everything. So we are looking for solutions on two fronts. One to induce the capability among IoT devices to learn, and take decisions through machine learning. Second to make IoT platform independent by complying with the standards and protocols suggested.

V.MACHINE LEARNING AND DIFFERENT ALGORITHMS IN IOT

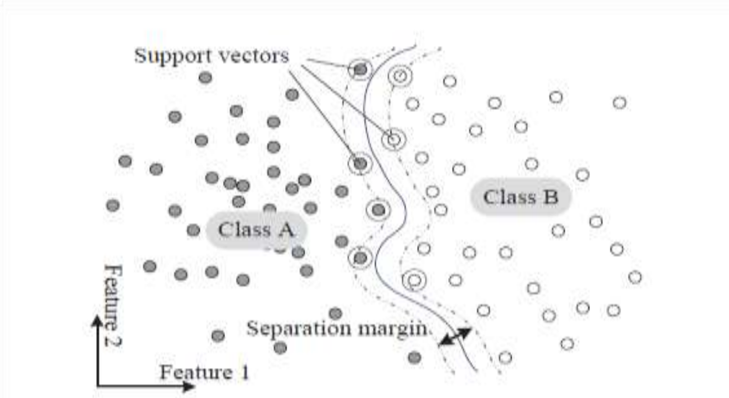
Machine learning (ML) was introduced in the late 1950's as a technique for artificial intelligence (AI). Over time, its focus evolved and shifted more to algorithms which are computationally viable and robust. In the last decade, machine learning techniques have been used extensively for a wide range of tasks including classification, regression and density estimation in a variety of application areas such as bioinformatics, speech recognition, spam detection, computer vision, fraud detection and advertising networks. In order to make IoT more efficient and scalable, following three machines learning algorithms can be implemented keeping in view the topology and computing ecosystem of the participatory devices in IoT.

Neural Networks: For example, sensor node localization problem (i.e., determining node's geographical position) can be resolved using neural networks. Node localization can be based on propagating angle and distance measurements of the received signals from anchor nodes. Such measurements may include received signal strength indicator (RSSI), time of arrival (TOA), and time difference of arrival (TDOA) as illustrated in Figure below. After several training, the neurons can compute the location of the node. This will help us to identify the node location exactly, and machines can detect the node location and target decisions accordingly.



Support Vector Machines(SVM)

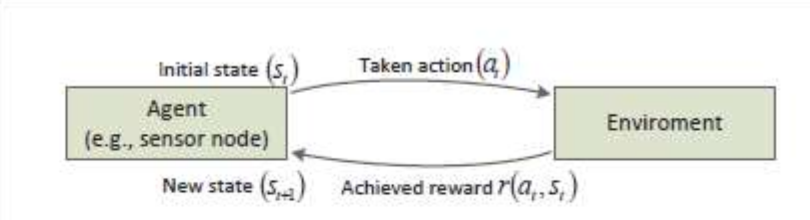
It is a machine learning algorithm that learns to classify data points using labeled training samples Basically, the problem is to classify those nodes into two parts. These parts are separated by as wide as possible margins (i.e., separation gaps), and new reading will be classified based on which side of the gaps they fall on as shown in Figure below. An SVM algorithm, which includes optimizing a quadratic function with linear constraints (that is, the problem of constructing a set of hyperplanes), provides an alternative method to the multi-layer neural network with nonconvex and unconstrained optimization problem. This will help us to find nodes in the vast array or pool of IoT devices, and consume data and perform analytics accordingly.



Reinforcement Learning

Reinforcement learning enables an agent (e.g., a sensor node in case of IoT) to learn by keeping trying and gaining experience, just like humans. As shown in Figure 5, an agent regularly updates its achieved rewards based on the taken action at a given state. The future total reward (i.e., the Q-value) of performing an action at a given state is computed using following Equation.

Q (st+1, at+1) = Q (st, at) + γ (r (st, at) + Q (st, at)) (1)



VI.TENSORFLOW

All the above algorithms can be designed in TensorFlow, which is an open source software library for numerical computation using data flow diagrams. We intend to write a python module in the future course of time that will take care of 2 important things as far as machine learning in IoT devices is concerned. First is web service which will integrate with respective methods in TensorFlow via API call. This will reduce the burden of learning TensorFlow to the users, and they can directly use our web service to implement machine learning algorithms in IoT devices. Second is the protocol testing of the IoT devices that is being put in the cluster to make sure it is complying with all the standards of making IoT platform independent. The algorithm that is being currently tested in our module is CART, which is discussed below.

VII.SUGGESTED TEST ALGORITHM FOR OUR PYTHON MODULE:

As a testing algorithm, we have implemented a classification and regression trees (CART). the input space is partitioned into axis-aligned cuboid regions Rk, and then a separate classification or regression model is assigned to each region in order to predict a label for the data points which fall into that region, or IoT Cluster. The classification task takes the form of

$$p\left(t=c|k\right)=\frac{1}{|R_k|}\sum_{i\in R_k}1\left(t_i=c\right),$$

$$y=\arg\max_c p\left(t=c|x\right)=\arg\max_c p\left(t=c|k\right)$$

This equation means *x* will be labeled by the most common (mode) label in its corresponding region. To formulate the regression task, we denote the value of the output vector by *t* and the predicted output vector for *x* by *y*. The regression task is expressed as.

$$y=\frac{1}{|R_k|}\sum_{i\in R_k}t_i$$

i.e., the output vector for *x* will be the mean of the output vector of data points in it's corresponding region.

ALGORITHM (Algorithm for training CART)

Input: labeled training data set D={x_i,y_i)} N

i=1

Output: Classification or regression tree.

FITTREE(0,D,node)

function FITTREE (depth,R,node)

if the task is classification then

node.prediction := most common label in R

else

node.prediction := mean of the output vector of the data points in R

end

(i*,z*,R_L,R_R) :=SPLIT(R)

If worth splitting and stopping criteria is not met then

node.test := xi < z*

node.left := FITTREE (depth+1,R_L,node)

node.right := FITTREE (depth+1,R_R,node)

end

return node.

VIII.CONCLUSION

IoT consists of a vast number of devices with varieties that are connected to each other and transmit huge amounts of data. The aim of this paper is not to confine machine learning in IoT to the theoretical concept, but actually create an enabling environment through easy to use software, libraries, web services etc that we want to create and distribute for free. During the course of development we may not confine only to TensorFlow, but also explore other frameworks like Apache SINGA, Microsoft Cognitive Toolkit, Apache Mahout, Caffe. In fact we are right now using some of the Apache SINGA libraries to generate assemblies for standard IoT devices.

REFERENCES

[1.] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A. H. Byers, Big Data: the Next Frontier for Innovation, Competition, and Productivity.
[2.] S. Bin, L. Yuan, W. XiaoyiResearch on data mining models for the internet of things
[3.] 2010 International Conference on Image Analysis and Signal Processing, IEEE (2010), pp. 127-132
[4.] A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi Internet of things for smart cities
[5.] IEEE Internet Things J., 1 (1) (2014), pp. 22-32

[6.] A. ShethComputing for human experience: semantics-empowered sensors, services, and social computing on the ubiquitous web

[7.] IEEE Internet Comput., 14 (1) (2010), pp. 88-91

[8.] https://www.leverege.com/blogpost/machine-learning-applications-in-iot

[9.] F. Bonomi, R. Milito, J. Zhu, S. Addepalli Fog computing and its role in the internet of things

[10.] Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, ACM (2012), pp. 13-16

[11.] A. Papageorgiou, M. Zahn, E. Kovacs**Efficient auto-configuration of energy-related parameters in cloud-based iot platforms**

[12.] Cloud Networking (CloudNet), 2014 IEEE 3rd International Conference on, IEEE (2014), pp. 236-241

[13.] H. Zhao, C. Huang **A data processing algorithm in epc internet of things**

[14.] A. Sheth, C. Henson, S.S. Sahoo**Semantic sensor web**

[15.] IEEE Internet Comput., 12 (4) (2008), pp. 78-83