The Role of Machine Learning in Internet-of-Things (IoT) 
Research: A Review

Aneri M. Desai  
Dept. of Information technology  
SVM Institute of technology  
Bharuch, India

Rutvij H. Jhaveri  
Dept. of Computer Engineering  
SVM Institute of technology  
Bharuch, India

ABSTRACT
Rapid developments in hardware and software based connected devices with communication technologies introduce Internet-of-things (IoT). In recent year, IoT gained enormous popularity and used vastly in a variety of applications. IoT represents the new forms of communication such as between human and things, and between two things. At present era IoT concepts are highly influence by creating a new dimension in the internet world. Intelligent processing and analysis of big data is the key to developing smart IoT applications. Such applications are logistic, transportation, agriculture, healthcare, and environment. Increasing the use of internet and external factor of environment are caused dynamic change in IoT system. The use of smart environments in the delivery of a pervasive care is the research topic that has witnessed increasing interest in recent years. These smart environments aim to deliver pervasive care through ubiquitous sensing by monitoring the occupant's activities. In order to provide smarter environment, their need to be implement IoT with machine learning. In recent year, machine learning technique have been used widely because of its technologies such that identification, extraction, classification, regression and forecasting. Machine learning exploring historical data from camera and sensors and perform techniques which improve the lifespan of network. In this paper, we build survey on existing research work carried out for various applications of machine learning to IoT. We summarize techniques and tools for IoT and also their benefits and limitations. Using a summarized data we also look for different future challenges. The aim of this paper is to survey different IoT technologies to assist the people to live in a smart environment.

General Terms
Internet-of-Things (IoT), Ambient Intelligence, Smart Environment

Keywords
Internet-of-Things (IoT), Machine Learning, Intelligent Processing, Smart Environment

1. INTRODUCTION
During the last decade, the population of countries has increased and so, the people’s life expectancy has also increased in the countries. The current demographic trends lead to an unbalanced life so it’s creating a serious social challenge. Hence the providing and maintaining intelligent system with quality of service to the people is one of the most important issues to be addressed [1]. Rapid developments in hardware and software based connected devices with communication technologies introduce Internet-of-things (IoT). In present era, growing the use of internet and wireless communication were making IoT concept worth. Because of growing the use of internet and communication, IoT gains the most significant research attention. IoT leads two words such that internet and things. Hence Internet leads the future of internet means devices connected to wireless network and things connected to wireless smart sensors and human invention because of that IoT represents an interconnected devices with network for make system intelligent or smart [2]. Basically, IoT will empower to developments in RFID tags, smart sensors, internet protocol, and communication technologies [3].

![Fig 1: Usage of IoT in India](image)

Figure 1 represents the usage of IoT in india at last five years. We show that graph of IoT increased with time. In order to provide smarter environment, their will need to be implementing IoT with machine learning. Now learning is process of gathering knowledge by observing and monitoring activates. Machine learning is process of making new facts and through using analysis and development with learning process [4].
Figure 2 represent the usage of machine learning in India at last five years. We show that graph of machine learning increased with time. IoT implement communication or interaction between number of devices which produce and exchange data between real world objects. In Present era, mostly application are based on human or machine interaction and communication. Because of that monitoring and extracting information from different type of sensors and camera information are high level information so that there will be need to represent this information or data into machine or human understandable form. For there is need a machine learning mechanisms needed to implement with IoT [5].

Various organizations have developed different technologies of Machine to IoT application for its verification and use in real world applications. Hence, researchers have focused on improvisation of machine learning to IoT applications by addressing various design issues. In this paper, we will discuss recent research work and open challenges that would assist researchers to carry out further research on machine learning to IoT applications. We also defined a limitation of recent research work carry out. Also defined how general machine learning with IoT system can perform an operation.

Figure 3: General Operation of Machine Learning with IoT System.
Figure 3 shows the general operation of machine learning with IoT. We show that operation of system divided into three categories such that IoT applications, data storage, and machine learning. In first category, IoT applications are connected with different types of sensors and camera. In this paper, we review different type of IoT applications with machine learning such that agriculture, security, industry, health monitoring, traffic analysis, social networking. We review that in this applications collecting and monitoring of data are performed through sensors and camera. After collecting of data, system performs analysis operation and makes big data. In second category, there was problem how to store big data. Because of that in second category used a cloud and database for store big amount of data. In third category, machine learning operations are performed on big data which stored in data storage. In machine learning operations, initially data loaded from data storage. After that, extraction and analysis are performed on that data. Using of this data operations of development and testing model are performed. Based on this model parameters future prediction of system will done using different machine learning technologies.

The remainder of the paper is organized as follows. Section 2 surveys the recent research works carried out on machine learning to IoT applications. Section 3 provides conclusions based on the study.

2. RELATED WORK

IoT plays an important role in demographic aging by providing an intelligent system to the people. In recent years, many researchers have introduced different methods and protocols related to improving the quality of life using machine learning to IoT applications based system. Given below, shows the different methods and protocols which have been used to improve the quality of life. The same is summarized in table-1.

Meidan et al. [6] proposed a scheme which classifies and identifies the IoT devices in network using supervised learning method. The aim of this scheme is to provide a security and integrity to the network and its devices to ensuring that IoT devices connected to which devices in network. It uses a supervised machine learning algorithm which identifies IoT and non-IoT devices in traffic environment. In this technique, collects trained data set of network traffic, and using that dataset calculate two parameters such that threshold value and session sequence size. Using these parameters it identifies IoT and non-IoT devices in network traffic.

Ferdowsi et al. [7] proposed a deep learning based technologies for dynamic IoT watermarking. In this scheme, it identifies cyber attackers in network using watermarking with deep learning. It uses long short-term memory (LSTM) algorithm based on deep learning which detect attackers and find injection of data and eavesdropping in IoT devices. Signals identification of IoT devices using watermarking did two ways such that static and dynamic. In static way identification of signals and attackers were done with spread spectrum watermarking. In this method if an attacker starts to store data of network then attackers gains the knowledge of static bit stream of signals. While in dynamic identification of attackers and authentication of signals were done using deep learning. LSTM algorithm is based on deep recurrent neural networks (RNNs) which store information for long amount of time and learn the dependency of given sequence. Based on that algorithm generates bit streams dynamically which solve the problem of eavesdropping attack.

Kumar et al. [8] proposed a scheme used for blind people to navigate freely indoor as well as outdoor environments. This proposed scheme, implement as intelligent, self contained navigation and face recognition. This scheme is useful for blind people to providing them navigation in environment by avoiding obstacles in path and also find person in front of them. This proposed scheme, works as two main module such that intelligent navigation module and face recognition module. Intelligent navigation module worked on android based smart phone connected with ultrasonic sensors. Using smart phone with google map or gprs system it navigates blind people. Text to Speech (TTS) system used for gives audio based response to navigate blind people. In this scheme, also used ultrasonic sensors with Bluetooth which perform as microcontroller where ultrasonic sensor collects data and convert into speech information. After that information are transferred to mobile phone using Bluetooth. Second module known as face recognizes which utilize artificial neural network. Artificial neural network used as learning mechanism which recognizes images of person based on past data. This system communicates with smart phone camera for extraction of face feature and this feature individually stored in data as training set. Using this training set artificial neural network can learn face feature and recognize face of person. Based on that, system will gives audio response to blind people.

Truongi et al. [9] proposed a scheme which presents design of IoT system containing environmental data to cloud storage. Using of this data with machine learning algorithm it predicts future environmental condition for detection and prevention of fungal in local crop filed. IoT system stores environmental data such as air temperature, air humidity, rain fall, wind speed. Machine learning algorithm used a support vector machine regression (SVMr) to extract raw data and predicts air temperature, air humidity, rain fall, wind speed. Based on these prediction parameters it identifies the presence of harmful fungal diseases in local crop filed. In short this scheme represents the design of IoT device in terms of hardware and software which collects present and past environmental data and predicates environment condition. Based on that prediction, it identifies fungal in local crop fields.

Joshi et al. [10] proposed a scheme which used for small scale farming using IoT. In this scheme sensors are used to sense properties of soil and environment. In this scheme also used image processing for detection and prevention of crop or weeds growth. Bayesian statistics algorithm used as machine learning algorithm which estimate how much care needed to weeds. Image processing takes different images of weeds and perform object detection algorithm using background subtraction. In this scheme using machine learning techniques with sensors collects different images of weeds and soil to decide how much care will needed in small scale frame. Basically proposed method describes, how individual can maintain his own farm in small garden.

Jaishetty et al. [11] proposed a Predictive approach for agriculture monitoring and controlling. In this scheme sensors are used to sense properties of soil and environment. Proposed scheme collects data from physical sensing device with cloud. Based on collect data machine learning tells the current state of environment and crop and at same time machine learning also perform predictive analysis and give future state of environmental parameters like temperature value, rain value and light value for crop. Scheme also used light weight communication protocol (MQIT) which has control over
irrigation. Remote monitoring solution also used for data monitoring, visualization and understanding and predictive analysis to frame for given scope.

P. Kumar et al. [12] proposed an IoT based health monitoring system. In this system three steps are performed such that collection of data, transfer collected data to database system that means storing of data, and based on this past data perform prediction using logistic regression. In this scheme first IoT based sensor devices collects data from human body. In second scheme, it used apache Hbase with amazon s3 service used for storing of data. Based on this stored data perform logical regression and predicate clinical parameter to decide heat debases for human. To perform prediction apache mahout based machine learning libraries used with logical regression. Blood pressures, body temperature and heart rate are such predicting parameters.

Bichindaritz et al. [13] proposed a scheme which is machine learning classification based multilevel stress identification. In this scheme ECG signals are used with machine learning for identification of different level of stress. Identification of stress is performed on car driver at different traffic conditions. Stress level identified into three categories such that low level, medium level and high level. Here low levels stress means initial rest, medium level stress means at highway and high level stress means at cities. ECG signals of stress were obtained from MIT_BIH physio Net Multiparameters database. In this database stored different dataset with raw data such that time stamp, ECG, electromyogram (EMG), foot galvanic skin response (GSR), hand GSR, intermittent heart rate (IHR), marker, and respiration. Based on this dataset, machine learning classification is used for feature extraction. Using this feature extraction it predicates the level of stress of driver and prevent driver from unwanted condition.

Chen et al. [14] proposed a smart clothing based health monitoring system. In this scheme patients have to wear smart clothes. Limitation of wearable devices is uncomfortable for patient which introduces smart clothes system. Smart clothes contain different small size sensors to measure different parameters such that pulse sensor, body temperature sensor, electrocardiography (ECC) sensor, myocardial sensor, blood oxygen sensor, electroencephalographic (EEG) sensor. Main component of schemes are body sensor, data processing module, storage module, wireless communication module and warning module. In this scheme, choosing of smart clothes material is important. Advantage of smart cloth system is motion capturing. In this scheme used cloud system to store large amount of data and also used machine learning technique for prediction of disease and health status. After getting of health status it transfers status to smart phone via bluetooth and wifi.

Luo et al. [15] proposed a web QoS prediction using kernel machine learning scheme. It is data driven scheme where it predicates missing web QoS value based on user and item matrix using kernel least mean square (KLMS) technique. In this scheme matrix prepared using public data obtained from real deployed system. To find relevant missing value Pearson correlation coefficient (PCC) used which analysis hidden correlation and predicates next missing value. At initial all data were identified. After using KLMS find correlation between values. Based on this correlation gives rank and make small set of data and predicates missing value.

F Firouzi et al. [16] proposed machine learning based smarter healthcare scheme. Aim of this scheme provides reliable, adaptive, safe and efficient personal healthcare services to patient. This scheme includes wearable sensors, big data storage and big data analytics. To stored patient history used cloud storage. After that using machine learning techniques it analysis that big data and based on this analysis it predicates health status of patient. After predicting, if there will be need of alarm triggering then system fire alarm triggering and give personal care to patient.

Ilapakurti et al. [17] proposed an IoT framework for connected dairy. This framework supports both offline and online dairy of IoT. This framework connected to sensors using bluetooth to retrieve sensors data. So basically, mobile interface periodically connected with sensors using bluetooth and read sensors data. These sensors are attached to the cow’s cattle which measures temperature, humidity and heat stress. Here information of sensors only collects when edge processing command given by mobile application. After that data are stored in cloud database. This data can be now representing in cubic from. Apply regression to that data to find dependency between known and unknown value. Now, using machine learning techniques performs forecasting to predict parameters of cows which connected to dairy.

Ghavat et al. [18] proposed a scheme which develops a new protocol for address smart home for living. This protocol knows as wellness sensors network. Make home as smart home means provide safe environment for well-being of inhabitants. This protocol mainly based on sensing technology. Using sensing technology make home smart means providing intelligence to home. Local database server used for storing sensing data. After collecting this data perform filtering of data to remove unwanted and repeated data. Data has been collected now analyzed of data perform using machine learning technique and generate activities and behavior pattern.

Elias et al. [19] proposed a scheme which addresses the design and implementation of new wildlife monitoring system with IoT, machine learning and image processing. Here image processing and machine learning used for image recognition, analysis and classification. Data has been collected now analyzed of data perform filtering of data to remove unwanted and repeated data. Data has been collected now analyzed of data perform using machine learning technique and generate activities and behavior pattern.

Norouzzadeh et al. [20] proposed a scheme which addresses location and behavior of animals in the wild with using of deep learning. In this scheme camera-trap are used for collecting images of animals in wild. After collecting of images perform filtration of data using machine learning. In this scheme deep learning neural network are used for automatic data extraction from camera –trap images. This scheme based on real time wealth information of vast numbers of animals in wild. This automatic extraction reduces the cost of system. Automatic extraction performs with human gives labels to images. Because of that count of number of human label also extract information.

Anjomshoa et al. [21] proposed a scheme which is used for identification of social behavior using machine learning. In this scheme an intelligent system makes which enable to identify user continuously on social network. Using behavior pattern of user on social network system identifies user. This
scheme is basically behavior of users at five different social networking applications. Using smart phone sensors and interaction of users make feature set of that users which includes data usage of user, starting and ending time of session, number of session at different time. Data collected from different session stores in private cloud based server. Now training is performed continuously on sliding window of data which generating pattern of user behavior. Training procedure perform on basically two factor such as social activity rate and sociability factor rate. Verification of user performs using two machine learning algorithm such that supervise learning and unsupervised learning mechanism.

Soundarya et al. [22] proposed a Criminal behavior analysis using Genetic Weighted K-Means Clustering. Main purpose of this scheme is predicting and identifies the criminal behavior of users on social network based on their past records and behaviors. For identify and analysis of criminal past behavior used generic weight K-means clustering algorithm based on machine learning. So basically in this scheme behavior of criminals identify using machine learning and data mining. After that classification of behavior also needed so that using negative selection algorithm classification of data performed. In these scheme collections of data performed at two social network sites such as face book and weblog. After classification of data performed comparative analysis manger compare both site data and based on that using machine learning predicates future and gives new direction and advices to youngyuth.

Zhang et al. [23] proposed a scheme which used for identification and estimation of social behavior of users using deep reinforcement learning. In this scheme identify user’s pattern to study social behavior based on considering social characteristic in pervasive social networking. In this scheme Pattern refer as market and recognize social behavior of user in market from economical prospective and identify user’s pattern on social networking. Based on this pattern, it handles the representative market model using deep learning reinforcement algorithm. Based on this social behavior it identifies prices in market, political decision, such as tax.

Chin et al. [24] proposed a scheme which defines the relation between artificial intelligence, big Data and internet –of-Things (IoT) to provide more services to smart city. In this scheme different city of parameters are calculated such that rain fall distribution and temperature of data collected using deep learning. In this scheme define correlation between weather and short cycling journeys. After collecting of data using machine learning classification dataset are tested. Based on thus classification of dataset predicts classification of weather and cycling.

Devi et al. [25] proposed a scheme which is machine learning based congestion prediction algorithm. In this scheme provides smart city roads using machine learning with IoT. Using camera and sensors perform analyzing of traffic flow at city roads. After analyzing of traffic, prediction of congestion performed using machine learning. In this scheme traffic flow analysis data used as knowledge and predicates traffic congestion. Linear regression is part of machine learning which used for prediction of traffic congestion. In this scheme first test data are collected after that average speed of vehicle compare with threshold value if value is increase then road not congestion otherwise define that path is congested with traffic. Based on that second training set is making and apply linear regression for prediction of traffic congestion.

3. CONCLUSIONS
During this study, we discovered that most of the projects focus on a technical area, and not on user needs. Machine learning to IoT applications has now been a popular research area for researchers due to their applications in smart environment. On one side, their intelligent processing and big data analysis nature self proves to be advantageous in many critical applications, on the other side it also raises several critical issues to be addressed such scalability, cost, time elapse, and battery of sensors, handling of multiple sensors data and many more. IoT becomes an integral part of internet application with many issues which require different solutions.

![Number of Articles](image)

**Fig 4: Category wise Percentage of Articles Cited in the Paper**

In this paper, we review different machine learning mechanisms with IoT which attempt to reduce the burden of people in their daily life to providing intelligent system. The papers studied and reviewed, Categories of applications, are shown in Figure 4. After analyzing various schemes, we conclude that different schemes work differently in distinct situations and each has its own limitations. A scheme may prove to be more advantageous than others in one situation, but may not reach to the expectations in other situation. This paper throws light on various aspects of machine learning to IoT applications which can be helpful for researchers wishing to explore this area. We draw the conclusion that there is a big room for further research on IoT by incorporating effective other techniques. In future scope , we may be focus on more machine learning techniques with IoT and its applications.

4. REFERENCES
[3] A. Al-fuqaha, M. Guizani, M. Mohammadi, M. Aledhari,


5. APPENDIX

Table 1. Literature Review

<table>
<thead>
<tr>
<th>Title</th>
<th>Author, Publisher and year</th>
<th>Methodology</th>
<th>Objective</th>
<th>Performance Metrics</th>
<th>Limitation</th>
<th>Future Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligent face recognition and navigation system using neural learning for smart security in Internet of Things002E</td>
<td>Kumar et al.Springer (2017) [8]</td>
<td>Face recognition and intelligent navigation system.</td>
<td>Identification of face for blind people and quick audio messages to blind people for easily navigation.</td>
<td>Facial expressions</td>
<td>Method is focuses only on static face recognition.</td>
<td>Use of IoT devices for recognizing the face reactions.</td>
</tr>
<tr>
<td>Where’s the Bear?: Automating Wildlife Image Processing Using IoT and Edge Cloud Systems.</td>
<td>Elias et al. ACM(2017) [19]</td>
<td>Image processing and machine learning based animal identification.</td>
<td>Address design and implementation of new wildlife monitoring system.</td>
<td>Identification of animal time and bandwidth reduce.</td>
<td>To store and transfer of images to private to public cloud storage is difficult task.</td>
<td>Develop a notification system that uses real-time feedback from image analysis to alert authorities of poachers and to alert visitors (hikers, students, and researchers) to potential animal presence.</td>
</tr>
<tr>
<td>Social Behaviour metrics for Personalized Devices in the Internet of Things Era</td>
<td>Anjomshoa et al. IEEE(2016) [21]</td>
<td>Social behaviour identification using machine learning.</td>
<td>Continues verification of user in smart device using machine learning.</td>
<td>Social activity rate, and Sociability factor</td>
<td>Possible that participating people have similar research domain because of that behaviour of people cannot give accurate result.</td>
<td>Energy efficient system is main concern for mobile application so focus on that direction.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Understanding and Personalising Smart City Services Using Machine Learning, the Internet-of-Things and Big Data.</td>
<td>Chin et al. IEEE(2017) [24]</td>
<td>Providing smart city services using machine learning.</td>
<td>Investigate relationship between weather and short cycling journeys and identify correlation weather and big data to provide services to smart city.</td>
<td>Rani fall distribution analysis.</td>
<td>Weather based data maintaining is difficult task.</td>
<td>Other services are also provided to make city smarter.</td>
</tr>
</tbody>
</table>