A Roar Towards IoT Big data Analytics: A Transfer Learning Context Framework for Medical Internet of Things

[1]P. Ilakkiya [2]Dr. B. Indrani,

[1]Research Scholar, Department of Computer Science, Madurai Kamaraj University, Madurai, Tamil Nadu, India

[2]Assistant Professor, Department of Computer Science, DDE, Madurai Kamaraj University, Madurai, Tamil Nadu, India

Abstract

The main objective of this paper is to implement a Machine Learning algorithm for medical IoT big data analytics. Machine Learning (ML) is one of the groups of modern programming methodologies in Artificial Intelligence. It can learn linear and non-linear data representation in huge complex datasets. ML can explore hidden information of the data successfully and report the data variations. Various problems of AI are solved using ML mechanisms in several fields like time-series data, computer vision, and NLP. One of the recent industries generating voluminous increasingly, day by day is the medical industry. Conventional AI methods have been used in earlier research works that are slow in data analytics, not cost-effective and the prediction accuracy is also less. Hence, this paper aimed to implement and experiment the Transfer Learning Method for analyzing the bigdata generated from Medical Internet of Things (IoT). The transfer Learning method uses knowledge of the source-domain to analyze medical bigdata. From the experimental results, it is identified that the proposed Transfer Learning method outperforms than conventional AI methods.

Keywords: Machine Learning, IoT, Bigdata Analytics, Medical IoT, Prediction.

Introduction

One of the future process of internet evolution is IoT, where it communicating any kind of physical devices equipped with computational methods for improving the communication capabilities. It provides seamless communication with various levels of internet. All the data generated from various IoT devices are persisted in the cloud databases since the volume of the data is high. Since the volume of the data is high, it needs to develop a data analytical method along with cutting edge technologies for manipulating the data which is used for data mining concepts. Comparing with all the technologies the concepts of bigdata analytics is implemented using recent technologies can provide better efficiencies in data extraction.

In recent days IoT and Big data are two different technologies buzzing the world but are necessary in computing industries. A drive is installed to adopt the bigdata into various organizations where it makes use of bigdata analysis method tremendously in the recent years. So, most of the enterprises are forced to use the bigdata analytical methods in the past few years. Parallel to this, IoT have flickered the entire computing industries by showing what a fully interconnected world can offer us. Hence, both IoT and bigdata evolved separately, both technologies become interconnected over the time. Also, the association between the bigdata and IoT shows a convergence of two different technologies which are aligning the technologies in the best possible method. Since, both IoT and bigdata provides more reasons for excitement, combining these two technologies increases the anticipation. This paper provides how IoT and bigdata are organized to obtain the entire benefits IoT based bigdata generation.

Figure-1. IoT and Big Data [1]

IoT denotes the set of all physical devices interconnected in the internet. The word “thing” in “IoT” represents a device assigned with an IP address or a person. A “thing” monitor, gathers and transfers data through internet without any manual help of the human or embedded technology. All the things interconnected in the internet communicate with the external environment or internal states to take a proper decision. A large set of structured or unstructured data measured in peta or giga bytes is called as bigdata and by analysing the bigdata it is able to obtain the visions of the business trend.

IoT Related to Bigdata

Comparing with the data generated through the overall computing industries, nearly 4.4 trillion gigabytes of data is generated in the year of 2020 through IoT. It is not easy to comprehend the bigdata. Since, the number of devices is increasing rapidly in the IoT-internet, the amount of data generated in the year of 2020 is also increasing. During this year, more than 10 billion of sensors are connected in the internet. Also, the IoT devices are monitor, gather, analyze, share and transmit the data in real time industries. Without the data it is not able to obtain the functionalities and capabilities of the IoT devices. When extracting and analyzing the data, IoT devices become the major source of the data, the role of IoT comes into the picture. For analyzing the bigdata, the bigdata analytical method is the one of the main keys for IoT–bigdata for increasing the decision-making abilities.

Learning, processing, and storing a large amount of data obtained from various real time application using different technologies is the role of IoT-Bigdata. The following steps illustrates the IoT-Bigdata processing.

- A greater number of IoT devices are interconnected in the internet based real time applications.
- All the IoT devices generate large amount of (or time series) unstructured data and gathered as bigdata.
- The generated data highly depends on four different factors such as volume, variety, velocity, and veracity.
- Hence the devices are connected in a distributed network, a shared-distributed database is used for storing the bigdata.
- Biggest analytical tools like Hadoop, Spark and MapReduce are used for analyzing the IoT bigdata.
- Finally, a common set of reports are generated for analyzing the data.

Because of the IoT-bigdata is unstructured data, it is essential to analyse the data based on large set of queries to obtain fast and best decisions. Though the need of IoT-bigdata is compulsory. IoT and bigdata are highly interdependent, help and impact each other. When the number of IoT devices increases then it will increase the business demand in terms of bigdata process. For example, the increasing number of IoT devices generates a large amount of data where it requires a high storage space. Hence, the demand of storage needs to change the storage space regarding infrastructure for storing the bigdata. Also, the IoT and bigdata based applications accelerate the scope of the work in both directions. Hence, IoT and bigdata are different technologies which carry inter-dependency and need further development.

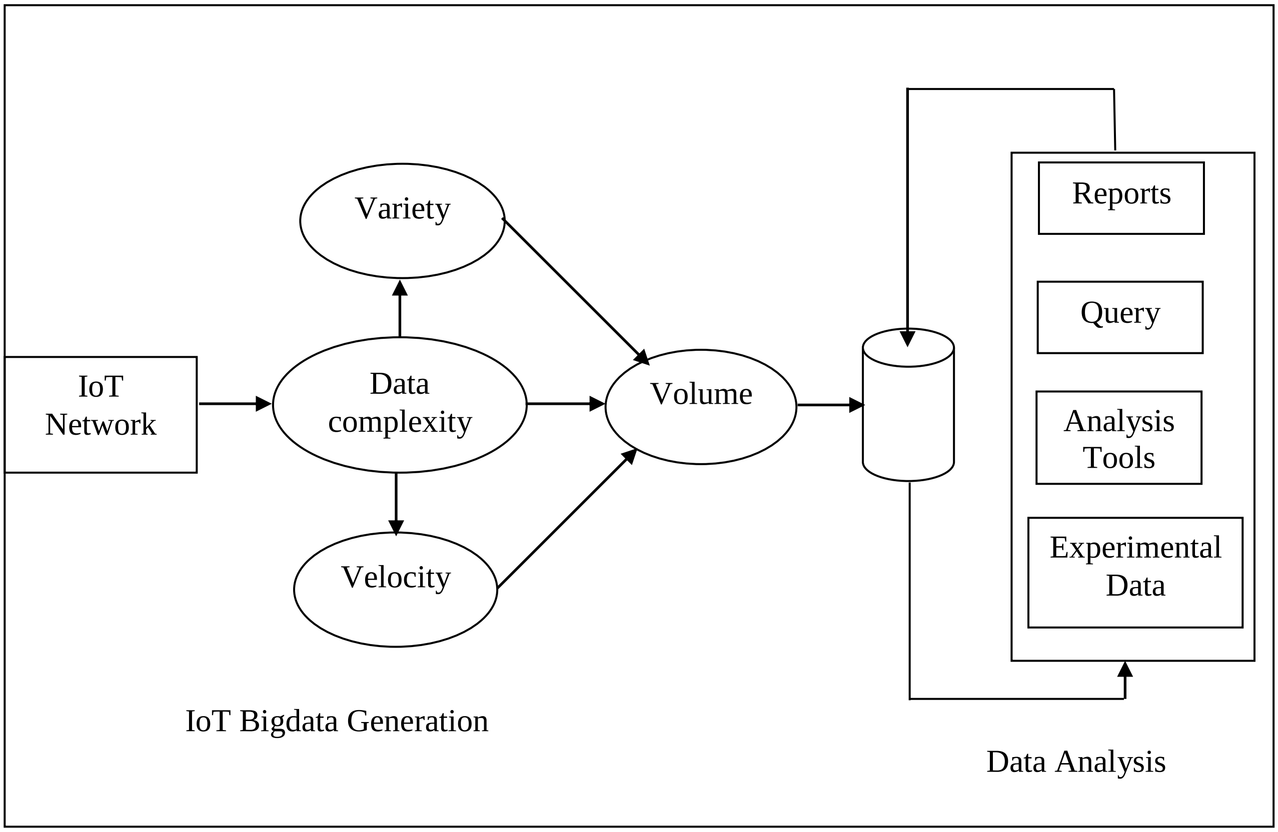

Figure-2. IoT-Bigdata Process

Managing and analyzing IoT-Bigdata leads to increase the ROI in every business, shaping the future process of the e-healthcare system, more benefits can be given to manufacturing industries, provides increased self-service processes and increase the demand of edge-computing in transport industry. It is highly essential and having highest priority in working on real-time data. Since IoT and Bigdata are increasing parallel to each other there is a high demand for deploying novel technologies is high. By integrating IoT and bigdata where it provides more benefits to the companies in terms of:

- Examine
- Reveal trends
- Find unseen patterns
- Find hidden correlations
- Reveal new information

Though various machine learning algorithms are existing in the data science industry, transfer learning is most recommended one for analysing the data obtained from multiple domains. Also, from the IoT-bigdata management and analysation, it is easy to identify how the businesses are influenced. From the result, it is able to correct the business and the organization to obtain an enhanced data management and take correct decisions. At each stage of the business process the IoT-bigdata provides more business. Transfer learning methods are popular and necessary because it can solve the problems by interconnecting multiple similar domains. It learns the similar data from different domain in terms of use-cases, similar input, time interval and assigned activities and provides a target domain. By integrating and converging IoT and Bigdata can provide more opportunities and real time applications in all the directions. Including this, it has the ability to revolutionize various point of views of our IT environment. So, recent computing industries are expecting innovative technologies and professionals to deploy IoT and Big-data analytics more advancements. Hence this paper aimed to utilise the transfer learning methods for managing and manipulating data obtained from IoT and bigdata domains.

Contribution of the Paper

The main objective of this research work is to handle the huge volume of data generated from various IoT devices interconnected with the cloud network. The entire data is dealt with the bigdata analytical method for preprocessing and persisting in the cloud storage where it helps to improve the datamining accuracy. Managing and manipulating the data is very difficult in IoT based cloud networks. Hence this research work initially focused on managing and manipulating the bigdata generated using IoT devices interconnected in cloud network. The data generated from the IoT devices and bigdata processes are interconnected together for managing the IoT-Bigdata. A detailed literature review can provide the importance of the proposed work and its contribution.

Literature Review

This review of the literature describes various earlier methods focused on managing IoT – Bigdata. Some of the methods helps to transform the data from raw format into understanding format. Few of the methods have been proposed for managing and classifying the data in accordance to the data nature and application.

Related to cloud data some of the authors have been concentrated on securing the private data. For example, **[2]** proposed method for IT enabled e-service for easy resource accessing from various places remotely. It is highly useful for a greater number of internet related business and transactions. One of the popular applications is online healthcare where it maintains the sensitive data and records about the patients. The author implemented security algorithm for safe the patient data and information. **[3]** proposed a Flexible and Efficient BioPNGA algorithm as a tightened security system for secret data sharing in cloud. The proposed method involves a biometric authentication system by public-key method. **[4]** proposed a password-based security system for online transaction. It accepted a crypted PIN number of authenticating people in order to access the data. The proposed method has been created as a protocol which uses Time Hash Function.

Initially IoT objects are different kinds of sensors, RFID devices, and various electronic and electrical devices connected in internet. All these devices are capable of generating and transmitting high amount of data continuously, discussed by the authors in [**5**]. In largescale industries trillion number of IoT devices are connected in internet [**6**]. Those devices can generate structured and unstructured data depending on the application. Authors in [**7**] described that the authors in IoT-Bigdata Management is also considered as a distributed system since it processes on unstructured data. In [**8**] the authors presented a detailed survey about various methods involved in IoT-Bigdata Management and analysis. A Cognitive Oriented IoT-Bigdata Framework [**9**] is implemented for providing a better architecture, set of defined layers to explore the knowledge based on the data generated using IoT-Bigdata.

One of the key researches in recent computing world is IoT and Bigdata management. Various activities are included in IoT-Bigdata management are data collection, integrating data, cleaning data, providing data storage, processing on data, analyze the data, and visualize the outputs [**7-10**]. It has been obtained by design and implement various kinds of data management tools and platforms which includes various activities related to IoT-Bigdata based knowledge discovery and extraction [**11**]. Author in [**12**] proposed a cognitive IoT framework for efficient knowledge creation and better decision making. To provide better and efficient decision making, author in [**13**] provide an information management system which involves data collection, extraction and analysis on time-based data. From the above discussion it is identified that, in recent days, due to voluminous data and different structure of the data, real time IoT and Bigdata industries require an effective, innovative and better algorithm.

Limitation and Motivation

One of the linear and non-linear learning methods for IoT bigdata representation is deep learning, can perform on large scale and complex bigdata. To explore the hidden features of various sources of data to obtain the variations successfully is possible. Deep learning methods are well configured to solve various AI problems like NLP, time series and computer vision, it is treated as an effective AI method for correlating various problems under similar domains using knowledge extraction from the domain sources. In addition to that, the training cost of the learning process is decreased. Similar problems are found in the IoT bigdata. To analyze and improve the efficiency of the IoT bigdata analytics, this paper motivated to use transfer learning method in the filed of IoT with Sensors. Where the sensors are,

- differ and various kinds of sensors collect different types of information from the environment.
- restricted in power and provide erroneous reading while the battery runs low.
- gathers similar information used to monitor similar events and activities.
- deployed in different environment and the system parameters can revolutionize IoT industries by sensor fusion.

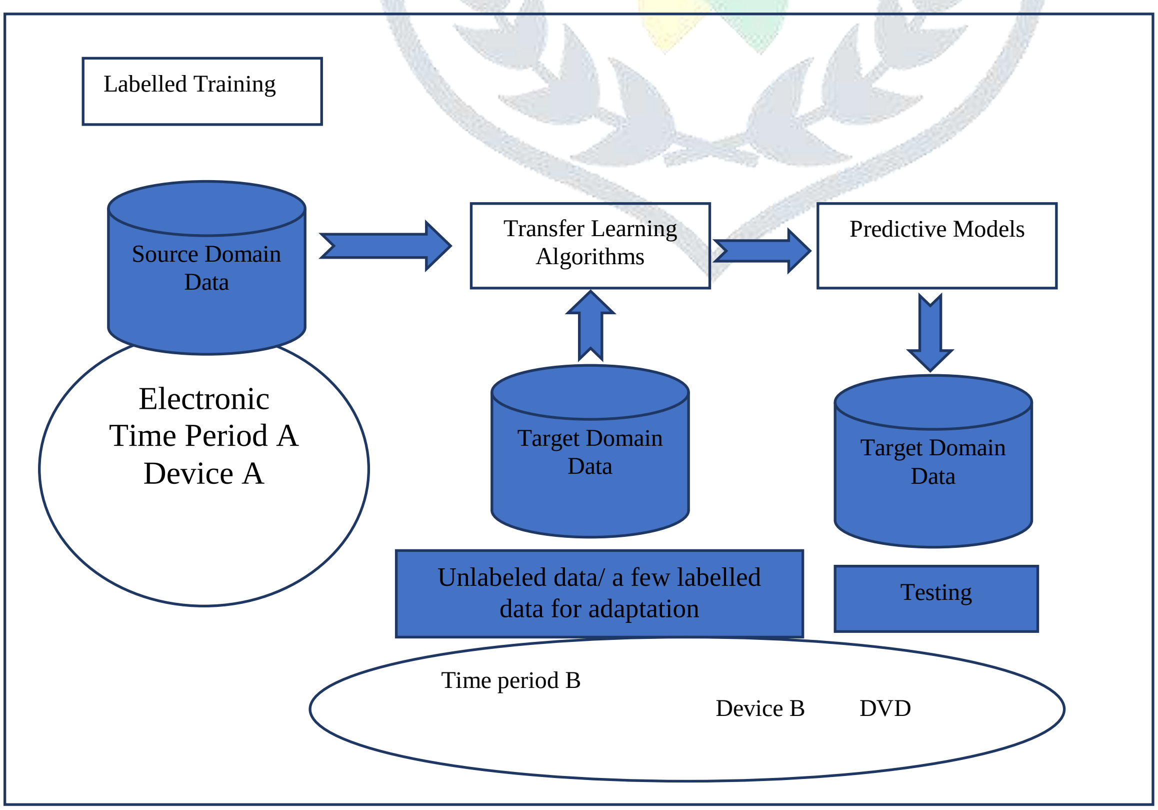

Figure-3. Transfer Learning across Domains on different time-periods Source

Transfer Learning Method

Training a machine learning model can be used as a reference for any unlabeled data available from related domains. To create this kind of trained data, recent programming methods bring an advanced transfer learning model which can corelate information from one domain to other similar domains. It also helps to predict the result from any related dataset and improves the prediction accuracy in the current problems. Two problems are connected by transfer learning methods which focus on learning the relevant parts of the training data. Transfer learning is highly applicable when the size of the historical data is less or data is generated in uncertain environment. Transfer learning methods are used for reinforcement learning, classification, and regression problems. The set of all data given in different domains are represented in the same feature values and do same data distribution. Training and testing various data obtained from different domains are represented in the feature space as:

$$\boldsymbol{X_S \neq X_T, or\ P_S(x) \neq P_T(x)}$$

Different conditional distributions or different label spaces are represented as

$$\boldsymbol{Y_S \neq Y_T, or\ f_S \neq f_T(P_S(y|x) \neq P_T(y|x))}$$

Some of the popular transfer learning methods are,

- Instance-based methods
- Parameter/Model based methods
- Relational Methods
- Feature-based methods

Instance based methods

Instance-based method processes on the overlapping features among the source and target data, where the source data and target data has more overlapping features, it is depicted in Figure-4. It shows that the set of features are gathered from source and target data to map them. The set of all data entities $\boldsymbol{X}$ is taken from source data and target data with overlapping features are shown in Figure-4. Based on this the similar data is analyzed and classified.

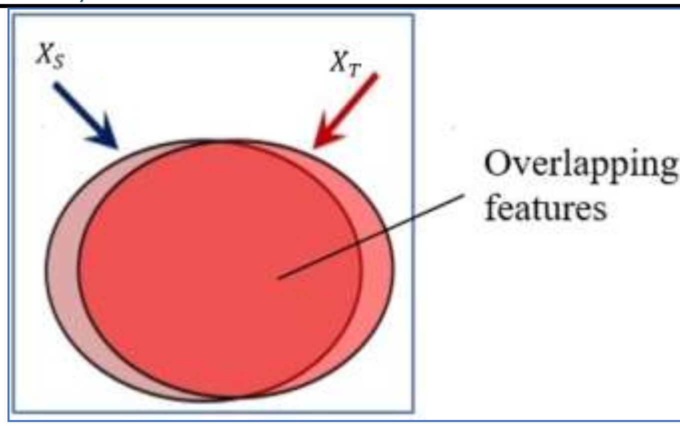

Figure-4. Instance Based Method

There are four different cases are used in the Instance based approach. In the first case unlabeled target data is analyzed and gives less accuracy is classification. Because the data do not have predefined label. In the second case some labelled data is used for analyzing the entire data. Some of the labelled data is taken from the source dataset and reused for target data analysis after recalculating the weight values. In the parameter/model and relational methods are quite difficult since they can do process on datasets where the source and target datasets must be related in terms of parameters and association among the datasets. IoT devices used in real time applications are different and application-specific. So, the data generated from various IoT devices are different can not be correlated based on the parameters and association.

Feature Based Method

From the overall methodologies, it has been noticed that instance-based method is not efficient in terms of time and cost. The classification accuracy is high if both source and target dataset has some overlapping features. The overlapping features existing in one domain supports the other domain for analyzation and classification, which is shown in Figure-5.

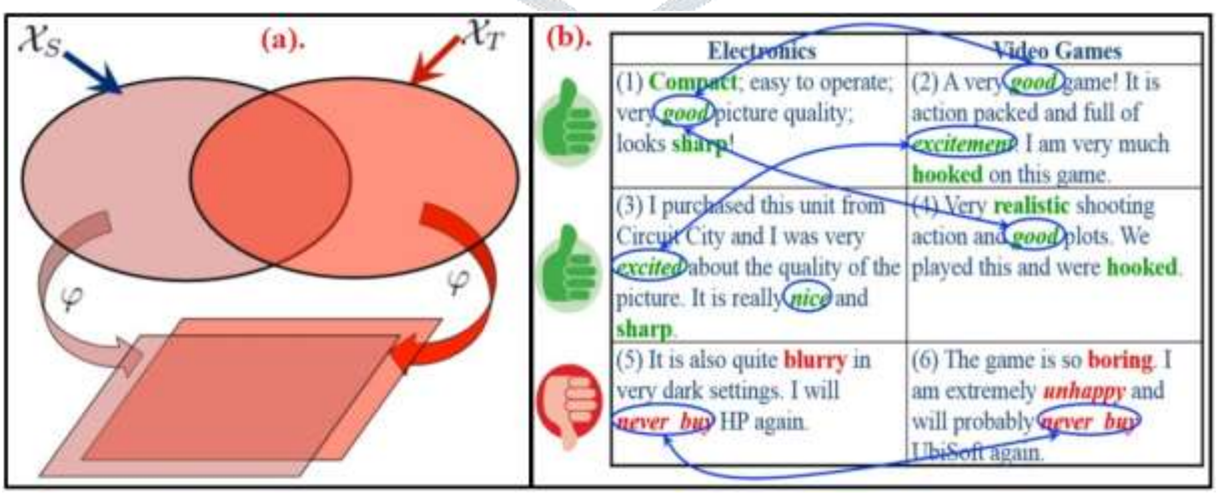

Figure-5. Feature Based Method, (**a).** Feature Extraction from Different Domains, (**b**). Feature Mapping

The feature set (φ) have learned from both source and target data is by knowledge base system or by general learning method for transformation. Where, X_S and X_T are the dataset obtained from source dataset and target dataset. For example, while considering two domains like electronics and video games, a set of features like good, sharp, compact hooked, real, boring and blur are the overlapping features which interconnects both domains. Figure-5 shows the application specific feature mapping for interconnecting two different data sources. For normal level and range of features common data analysis processes are used. Latent features are learned and analyzed using PCA and LCA method. But for higher-level feature set sparse coding and deep learning methods are preferred. In this paper, one of the deep learning methods, transfer learning method is used for managing and analyzing IoT – bigdata. IoT - bigdata are learned for mapping the data. It maps the source and target data in terms of latent space traversed by the features are used to diminish the domain variations and preserve the originality of the data structure and it can be written as:

$$\min_{\varphi} Dist\ (\varphi(X_S)), \varphi(X_T) + \lambda\Omega(\varphi)$$

s.t. conditions on $(\varphi(X_S))$, **and** $(\varphi(X_T))$

Based on the lesser distance the source and target data are correlated. The maximum mean discrepancy of the data analysis can be written for X_S and X_T is follows:

$$X_S = \{X_{S_i}\}_{i=1}^{ns}, X_T = \{X_{T_i}\}_{i=1}^{nT}$$

derived from $P_S(x)$ and $P_T(x)$ respectively. The similarity between the source and target data are calculated by finding the distance as:

$$Dist(P(X_S), P(X_T)) = \left\| \frac{1}{ns} \sum_{i=1}^{ns} \Phi(X_{S_i}) - \frac{1}{nT} \sum_{j=1}^{nT} \Phi(X_{T_i}) \right\|$$

If the value of $Dist(P(X_S), P(X_T))$ is very less then the data from source and target data are considered as similar data.

Transfer learning is mainly used for interconnecting or interchanging two different business knowledge from one geolocation to another geolocation and helps to extend the existing business into new business according to the market conditions. It also helps to transfer the economical, investment and financial knowledge from one sector into other sectors over the world based on the similar economy, technological advancements and GDP. It helps to map the political data across world and supports voting process. A novel concept used for Transfer knowledge among different features spaces is called as transfer learning.

In case of medical IoT – bigdata, the dataset and the feature set values are calculated as:

Source domain (Medical IoT) specific features, e.g. health-condition, disease, abnormal, and mild, moderate and severe.

Target domain (Medical diagnosis) specific features, e.g., disease, abnormal, moderate and severe.

Domain independent features (pivot features), e.g., disease, health-condition.

The pivot features are identified by calculating the term-frequency on all the domains, obtain the mutual information between the features and their classes, and obtain the mutual information between the features and their data domains. The pivot features are used for obtaining the learning the structural correspondence and aligning the spectral features.

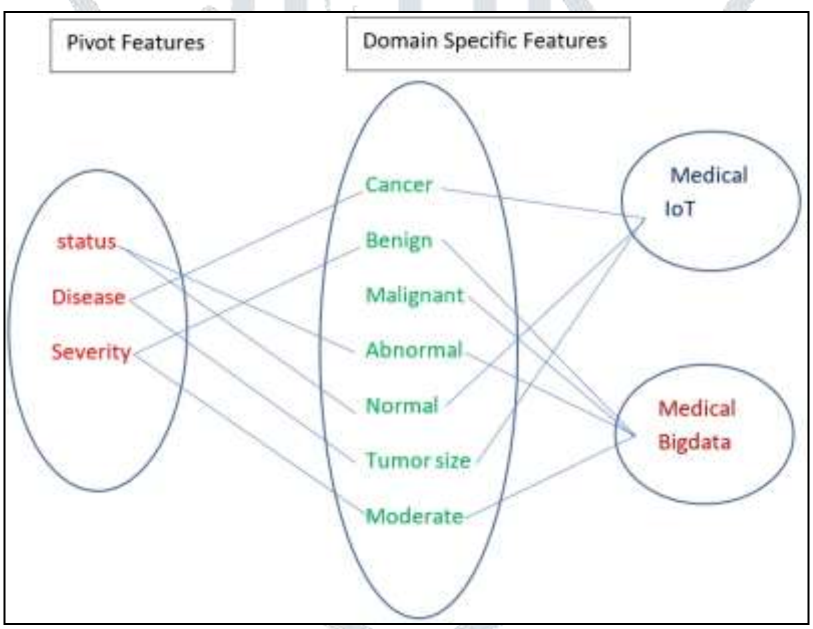

Figure-6. Feature Mapping

Feature mapping for medical IoT-Bigdata is shown in Figure-6. When a context provides appropriate information depending on a task is termed as context aware system. The proposed system achieves active context awareness, through Transfer Learning Context Framework (TLCF) (see Figure-7) which improves the conclusion made from the patient's medical analysis and health. The flow of the TLCF model for healthcare system is given in Figure-7. In every healthcare industry, various kinds of sensors and/or IoT devices are used for monitoring and capturing the patient or the environmental data. The monitored data is applied into data-acquisition method. To create a reasoning model, a context model is applied to gather contextual information from every data and feed into machine learning algorithm learning and extracting the features. Based on the feature values, a final decision is created on the patient data.

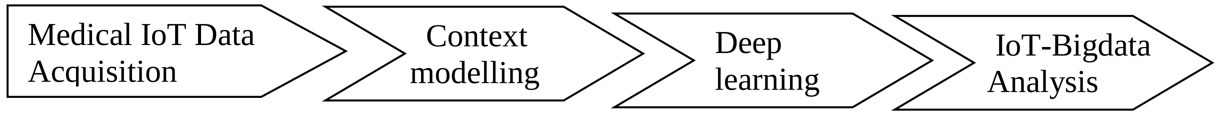

Figure -7. Workflow Transfer Learning Context Framework

IoT data are collected and construed into contexts for context modelling as shown in Figure-7. TLCF is used for healthcare decision supports and it delivers patients with a reasoning mechanism to enhance health services. Let us consider a cloud-based health care environment, with 'h' number of hospitals in a set $\boldsymbol{H = \{H_1, H_2, \ldots, H_h\}}$, $\forall \boldsymbol{h} \in \boldsymbol{H}$ as stated in Figure 3.19. Let different departments be connected by one hospital and for simplicity, it is implicated that similar and identical number of departments are present in each hospital. Let $\boldsymbol{DP = \{DP_1, DP_2, \ldots, DP_{dp}\}}$, $\forall \boldsymbol{dp} \in \boldsymbol{DP}$ be the collection of $\boldsymbol{dp}$ number of departments connected with every hospital.

Every department nearby is combined with various number of doctors out-patients and BAN patients, that are the basis for generating the huge data. It is to be distinguished that out-patients are the patients who attend a hospital for treatment without residing there. BAN patients are the chronic patients who have been fixed with smart body sensors to examine their health conditions regularly.

Let, $\boldsymbol{do}$ represents number of doctors available in a set $\boldsymbol{Do^x_{ij}}$, where $\boldsymbol{j = \{1, 2, .., do\}}$ in the $\boldsymbol{i^{th}}$ department of $\boldsymbol{x^{th}}$ hospital $\forall \boldsymbol{i} \in \boldsymbol{DP}$ and$\forall \boldsymbol{x} \in \boldsymbol{H}$. Hence, $\boldsymbol{Do^x_{ij} = \{Do^x_{1d} \cup Do^x_{2d} \cup \ldots \cup Do^x_{dpdo}\}}$$\forall \boldsymbol{i} \in \boldsymbol{DP}$ and $\forall \boldsymbol{x} \in \boldsymbol{H}$. For example, $\boldsymbol{Do^5_{12}}$ says the doctor-2 from department 1 in hospital 5.

Let $\boldsymbol{P^x_{ij}}$ be the set of patients in which $\boldsymbol{j = \{1, 2, \ldots p\}}$ in the $\boldsymbol{i^{th}}$ department of $\boldsymbol{x^{th}}$ hospital, $\forall \boldsymbol{i} \in \boldsymbol{DP}$ and $\forall \boldsymbol{x} \in \boldsymbol{H}$. Thus P represents number of patients present in the i^{th} department of $\boldsymbol{x^{th}}$ hospital. Thus, $\boldsymbol{P^x_{ij} = \{P^x_{1p} \cup P^x_{2p} \cup \ldots P^x_{mp}\}}$,$\forall \boldsymbol{i} \in \boldsymbol{DP}, \forall \boldsymbol{x} \in \boldsymbol{H}$. For example, $\boldsymbol{P^2_{34}}$ denotes the patient-4, that is associated to the department-3 in hospital -2. It is considered that patients with BANs are also hospitalized, that could be neither a patient nor a BAN at a time. Likely, $\boldsymbol{b}$ represents the number of BANs appeared in a set $\boldsymbol{B^x_{ij}}$, where $\boldsymbol{B^x_{ij} = \{B^x_{1b}, B^x_{2b}, \ldots, B^x_{8b}\}}$, $\forall \boldsymbol{i} \in \boldsymbol{DP}$ and $\forall \boldsymbol{x} \in \boldsymbol{H}$ and various number of BAN are accessible in several departments within the hospital. For example, $\boldsymbol{B^2_{13}}$ denotes BAN-3 is associated to the department-1 in hospital-2.

In the proposed model, a window based chronological data set and monitoring model is used to improve the quality of patient monitoring. Let $\boldsymbol{T = \{0, 1, 2, \ldots t\}}$ be a constant time frame that is partitioned into $\boldsymbol{W}$ number of windows, in which every window contains $\boldsymbol{Z}$ units of time interval. Every time interval could be deemed as a minute, an hour, a week, a month or a year based on the application going to be handled. Consequently $\boldsymbol{Do^x_{ij}(w)}$,$\boldsymbol{P^x_{ij}(w)}$, and $\boldsymbol{B^x_{ij}(w)}$denotes the amount of data produced from BAN, doctors and patients correspondingly in every window w. The gathered data inside window w are piled up in various cloud data centers as shown in Figure-8. Let $\{ DC_1, DC_2, \ldots, DC_n\}$ be the N numbers of geo-dispersed data center positioned in the cloud, where $n \in N$. These data centers are joined through M numbers of gateways $G = \{GW_1, GW_2, \ldots, GW_m\}$ where $m \in M$. In our structure, H represents number of those hospitals that are joined by those N numbers of geo-dispersed data center via M number of gateways.

The overall data is collected from the data centre which is available in the cloud. Initially the data gathered from various medical devices interconnected with the healthcare network. The data is initially processed and corrected to remove the noise. After correction the data is converted into data packets. Each data packet is created and stored in a predefined format where it has a field called as criticality bit field. Based on the criticality bit, the data packet is transmitted with highest priority assignment. Since, emergency patient's information is transmitted with highest priority to get appropriate treatment suggestion from the medical experts interconnected in the same network. Hence it improves the efficiency of medical data analysis and treatment prediction in the cloud.

Speed and accuracy of the mining process is increased by clustering method. Two types of clustering methods are used in the stage, are inter-cluster and intra-cluster. Intra-clustering method cluster the data within a hospital H_i where it classifies the patient data based on the patient ID, body area network ID, and classify the data based on the nature of the disease. The entire patient data is clustered and arranged in such a manner where predicting any data is very speedy and accurate. In order to globalize the data using cloud data centres, the overall data persisted in the cloud data centres is clustered and it is called as inter-cluster process. The inter cluster process cluster the entire data comes from various hospitals interconnected into cloud. Inter clustering is applied on the results of the intra-cluster. Hence the efficiency of the inter-clustering is improved in terms of reduced time complexity, increased speed and improved accuracy.

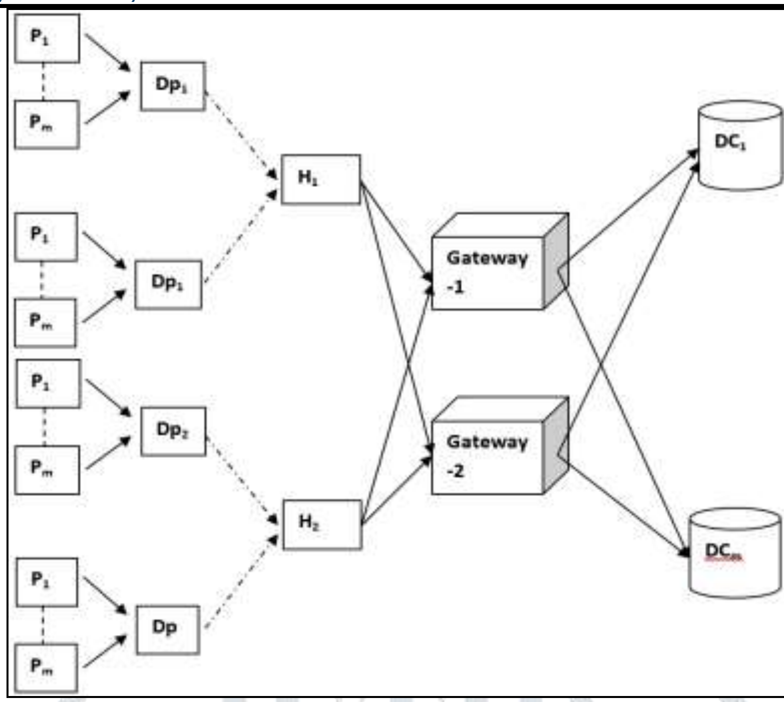

Figure-8. Healthcare Monitoring, Data Collection and Transmitting

The conditions of the patient are predicted by validating the medical data (E.g. ECG) and compared with the well-defined threshold values. The patient health condition is well-defined as normal, mild, moderate and severe by medical experts or laboratory experts and available in the medical industry. Each condition has a threshold value whereas it can be used for comparing with the recorded values. The patient data is collected in periodical manner and compared. In order to compare effectively, the recorded data is divided using widow system.

IoT-Bigdata Management

The window volume could be customized based on health conditions. Consider this example, let A be the blood pressure data of a patient, that is verified on the time framework $\boldsymbol{T = \{1, 2, 3, \ldots. t\}}$, with window frames $\boldsymbol{G = \{W_1, W_2, \ldots, W_m\}}$, as shown in Figure-9, where the interval of every window $\boldsymbol{W_i}$ holds 3units. In second window **(W_2),** the normal blood pressure data (**Avg.**$\boldsymbol{W_2}$) of three time slots (4, 5, 6) is verified as 140, that is higher than the vital condition **(if (Avg.W_w> 125)).** Those verified data by time interval are gathered by using our proposed data acquirement scheme and is conveyed to the data centre in the cloud for storage and analysis.

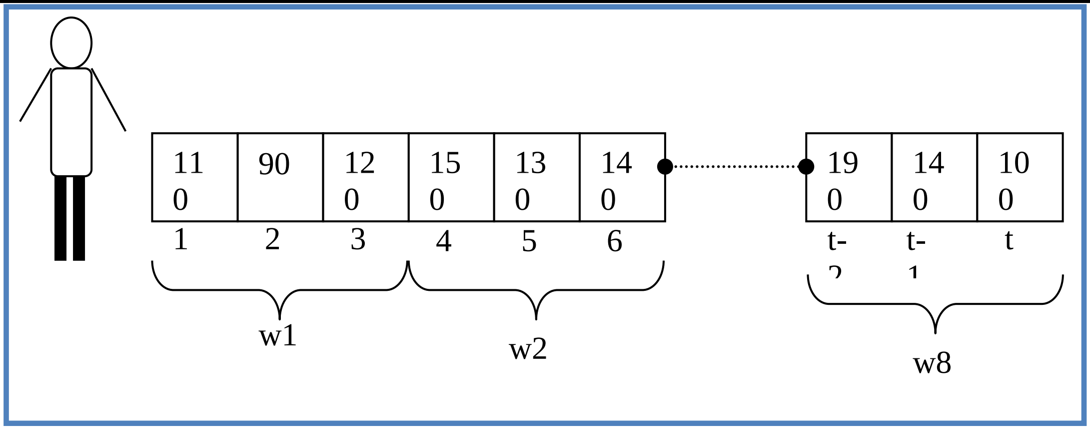

Figure-9. Data Collection in Windows Based Model

In huge data healthcare surroundings, the Electronic Health Record (EHR), 3D imaging, physiological data, medical, radiology images, genomic sequencing and billing data are the basis of huge data which illustrate the volume. Synchronized and urgent situation monitoring is the basis of streaming data that illustrates the velocity of data. Although the majority of the stages believe the patients physiological data as the huge data, we incorporate the visiting rate of the patients to the hospitals in our huge data processing models. To transmit the priority-based data, SDN-WSN applies hierarchical bit analysis in the data packet. Here the algorithm is applied to change the priority based on the critical bit value. If the critical bit value is “0”, then the data packet transmission is normal, that is as such in the queue the data transmitted. If it finds the critical bit is “1”, then that particular data packet is transmitted with high priority. After transmitting the data with highest priority, other data packets will be transmitted.

Experimental Results and Discussion

This section presents the detailed information about the experimental results obtained from the simulation / experiment carried out in this study. Based on the experimental results the performance of the study is evaluated. From the comparison it can also able to find out the overall quality of service of the proposed method. In this paper, applications such as healthcare is considered as indoor and outdoor monitoring applications respectively. In healthcare monitoring the target nodes (patients, bed, rooms, and equipment) are located inside a building or campus (so considered as Indoor). It is required to use the remote server and an IoT network in order to estimate the accuracy of the TLCF. As primary stage, the patient important signs are gathered from Bio Harness sensor (IoT device). For example, some of the parameters are considered as biomedical parameters, such as: heart rate, oxygen level, blood pressure, body temperature. These parameters are noted with the time duration, ambience and movement of the patient.

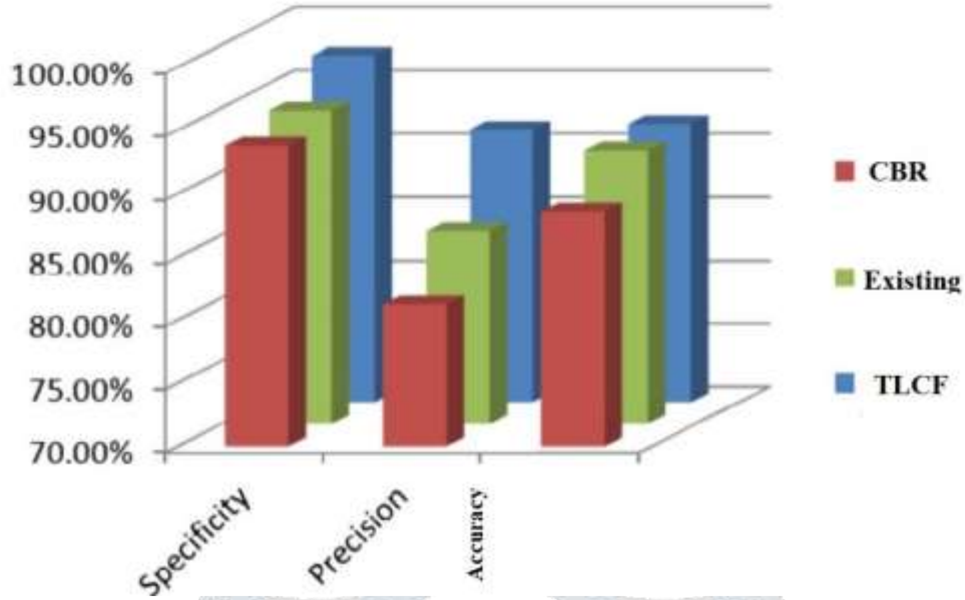

Figure-10. Comparison of three approaches

The wireless connection between the IoT network and user application is created employing Bluetooth, and the application associated to the home gateway which communicates data to the remote healthcare server. The implementation of the TLCF based reasoning engine is executed to provide reliable prediction.

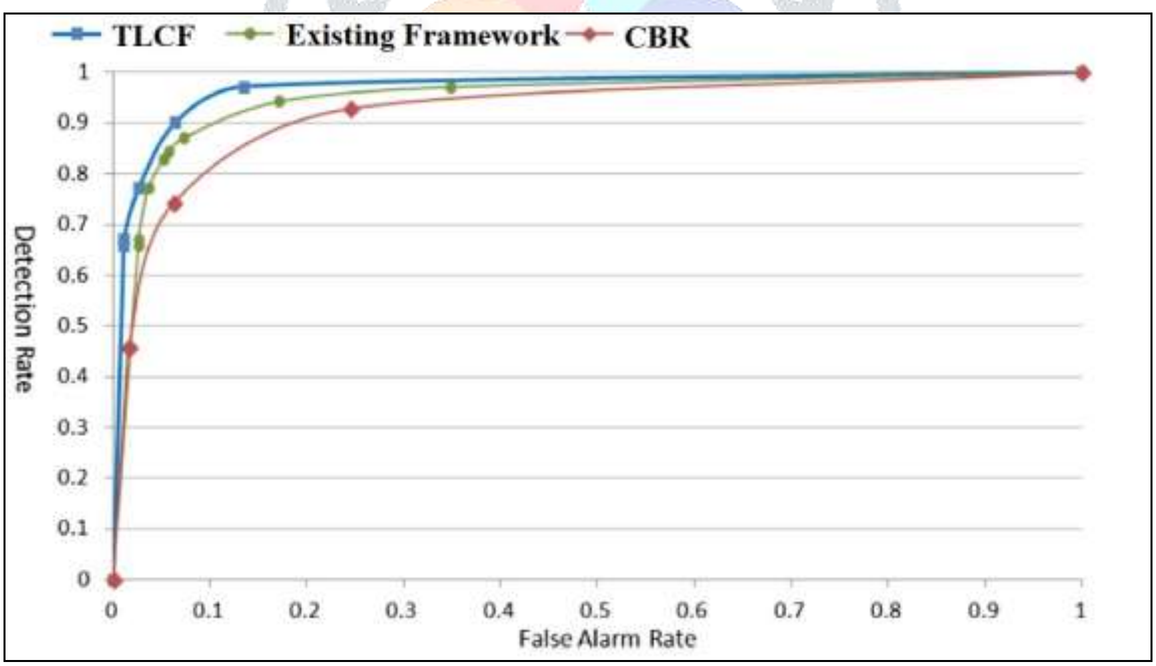

Figure-11. ROC For TCLF

In order to compare the proposed TCLF, we employed a case base which consists of 262 cases where 70 are abnormal and 192 are normal. We compared the anticipated reasoning approach in-terms of accuracy with straightforward Case Based Reasoning (CBR) approach and developing CBR approach employing dynamic weights in case retrieval. The proposed approach achieves better performance of 97 in specificity, 91% in precision and 92 and 92% in accuracy whereas normal CBR approach yields 93% in specificity, 81% in precision and 88% in accuracy.

To find accuracy for anomaly detection, two class prediction, namely, abnormal or normal labelled as negative or positive as outcomes were considered with the true and false value. TP (True Positive) represents real condition is abnormal and the prediction is also abnormal, while TN (True Negative) represents when prediction and real conditions are normal. The false alarm rate (FP), indicates if real condition is normal while FN prediction is normal and real condition is abnormal. The case of abnormality could be determined with the help of true positive and true negative value which could be represented as,

$$CA = TP + \frac{TN}{TP} + TN + FP + FN \times 100\%$$

Detection rate: This uses a method of confusion matrix which is represented as,

Detection rate = TP/TP+FN $\times$ 100%

False alarm rate: the false indication of the normal condition which is represented as,

$$False\ alarm\ rate = \frac{FP}{FP} + TN \times 100\%$$

We employed Receiver Operating Characteristic (ROC) from signal detection theory by determining the TP and FP rate. As a result, the ROC curve has been obtained as shown in Figure-11. It is evident that our proposed approach achieves better prediction method at upper left at coordinate (0, 1) which is considered as perfect classification. The Decision Tree classifier achieved the best accuracy. As shown in the Figure-10 the proposed TCLF obtained highest accuracy than the existing methods.

Performance Evaluation of IoT Network

Initially, from the simulation it is necessary to identify the IoT devices are efficient in monitoring the environment. To do that, the amount of data gathered in each round of operation (that is by changing the number of nodes, the simulation is executed) is calculated and verified in a fixed interval of time. So, based on the number of nodes the amount of data gathered from the network is calculated. The obtained results are depicted in Figure-12. In order to provide a performance comparison, the existing algorithm is also simulated in NS2 and the results are verified. The number of nodes deployed in the network in existing as well as in the proposed is uniformly like 100, 200, 300, 400 and 500 in five rounds of network operations. Instead of verifying the data gathering efficiency for each individual node, for set of nodes are applied in each round where it reduces the time and computational complexity of the algorithm. From Figure-12, it is noticed that the proposed method gathers more amount of data than the existing. Comparing with all the applications, in the agriculture environment the amount of data gathered is high than healthcare and healthcare is high than the existing environment.

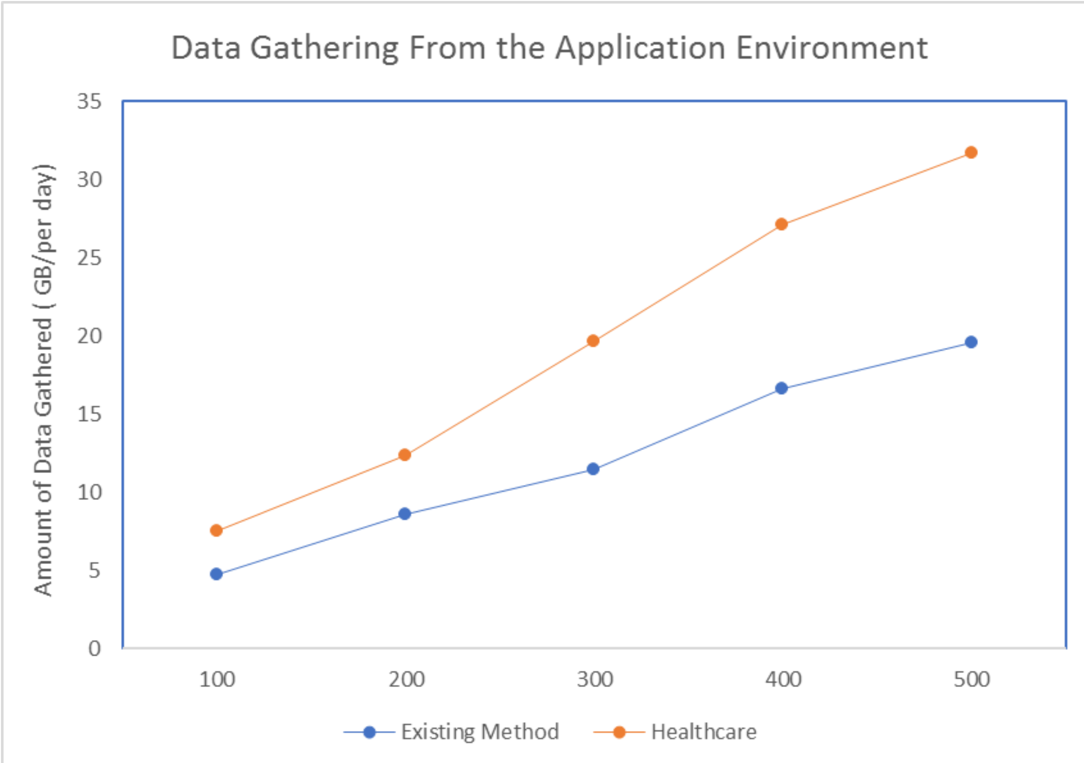

Figure-12. Amount of Data Gathering in a Day

Another factor needs to be considered here is the amount of data transmitted from the data gathered. It determines the efficiency of the IoT nodes and the framework.

In order to verify this data transmission in each round is calculated and the result is given in Figure-13. The average amount of data is collected from the environment is 25.588 GB and the amount of data transmitted 25.106 GB. From the experiment 97% of the collected data is transmitted to the server, and it denotes that the proposed agriculture monitoring is better IoT-network monitoring system than the healthcare monitoring system. Because the amount of data transmitted in other environment is higher than the data collected. Figure-12 shows the comparisons of the ability of the medical IoT application regarding the amount of data gathering from healthcare environment, whereas Figure-13 shows the comparison of the ability of the Medical IoT application regarding the amount of data transmission from healthcare environment into database server.

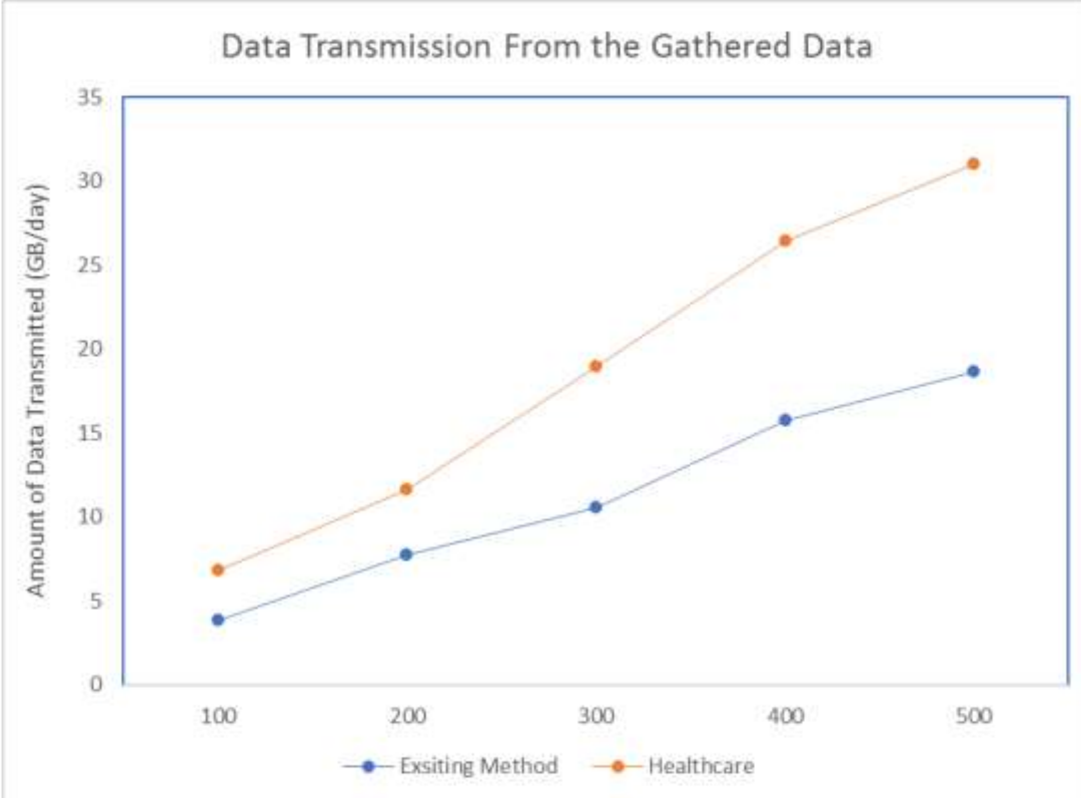

Figure-13. Amount of Data Transmitted in Day

Then it will be verified is there any data loss or leakage in the application. To do that, from the gathered data, how much amount of data is transferred to the corresponding destination or server is calculated. Figure-13. shows the amount of data transmitted in the environment from the gathered data. The amount of data transmitted in lesser than the gathered data is comparatively less. Among the three applications, the applications used in the proposed research work obtained high rate of data transmission than the existing. Also, the amount of data transmission in medical is high than the other application. It can also be cross verified by calculating the data loss by comparing the data received and the data transmitted and difference between data gathered and data transmitted. The packet delivery ratio is calculated by computing the amount of data received at the destination. Based on this the PDR is calculated in each round of execution and the obtained results are given in Table-1. From the Table-1, it is identified that the PDR obtained using the proposed TLCF is better than the existing approach. Whereas, our proposed indoor application environment (healthcare) obtained high PDR than the outdoor application, it is because of, in outdoor environment the dynamic behaviour of the IoT nodes can be affected by natural disasters like rain, temperature and wind.

Table 5.2 Packet Delivery Ratio (PDR)

Number of Nodes	Existing Method (%)	Healthcare (%)
100	88	98.76
200	86.5	97.96
300	83.21	96.99
400	80.97	95.91
500	78.45	94.93
	83.43	96.91

From the obtained PDR, it is considered that the medical IoT -bigdata application is better in data transmission in IoT-network.

Summary

This paper provides a detailed information about IoT and Bigdata analysis with necessary of transfer learning in deep learning. Transfer learning integrates IoT network and Bigdata analysis together to provide a contextual framework. Transfer learning extracts the features from the dataset generated from the IoT devices/network and analyze it using bigdata analysis methods. Data gathering from IoT is carried out initially and analyze the data using transfer learning method for data transmission. In order to transmit the data a window system is applied over the time of data generation. Amount of data generation, transmission and time factors are calculated for evaluating the performance of bigdata manipulation and data transmission rate in terms of PDR is calculated for IoT network performance analysis. From the results it is identified that the proposed TCLF is better for IoT-Bigdata analysis in medical industry.

Conclusion

The main objective of this paper is to provide a novel method for managing and transmitting IoT bigdata in medical industry. To do that, a medical IoT network is created in a simulation and data gathering process is carried out. To incorporate the IoT data and bigdata analytics, one of the popular deep learning methods, Transfer Learning method is used. A simulation is used for IoT network performance analysis and MATLAB programming is developed for bigdata analysis. From both simulation and implementation, the performance of IoT-Bigdata analysis framework is calculated and evaluated. From the results, it is concluded that the proposed Transfer Learning Context Framework is better for medical industry.

In future, it is extended by integrating security for preserving the medical data.

References

1. https://www.whizlabs.com/blog/iot-and-big-data/
2. M. Amutha Prabakar and B. Indrani and M. Karthigai Veni, (2017) "Provably Secure Two-Factor Authentication Scheme for E-Health Using Smart Card",IACR Cryptology ePrint Archive, Vol.2017, PP-501-512.
3. Indrani, B., & Veni, M.K. (2016). A Flexible and Efficient Algorithm for Generating Prime Numbers using Biometric Identity (Bio-PNGA).
4. B. Indrani, M.A. Prabakar (2015), "An efficient password based authentication scheme using dynamic identity and smart card", International Journal of Applied Engineering Research, Vol. 10, No. 9, PP. 24452-24462.
5. Ma, Meng, Ping Wang, and Chao-Hsien Chu (2013), "Data Management for Internet of Things: Challenges, Approaches and Opportunities", Green Cloud Computing and Communications, IEEE.
6. Henson, Cory, Payam Barnaghi , and Amit Sheth (2013), "From Data to Actinable Knowledge : Bigdata Challenges in the Web of Things", IEEE Internet Systems, Vol. 28, No. 6, PP. 6-11.
7. Peng, Zhen, Zhao Jingling, and Liao Qing, (2012), " Message Oriented Middleware Data processing model in Internet of Things", Computer Science and Network Technology, IEEE -2012.
8. Nilamadhab Mishra, Chung-Chih Lin, Hsien-Tsung Chang, (2014), "A Cognitive Oriented Framework For IoT Big-data Management Prospective", IEEE-2014.
9. Nilamadhab Mishra, Chung-ChihLin, and Hsien-TsungChang, (2014), "A Cognitive Adopted Framework for IoT Big-Data Management and Knowledge Discovery Prospective", International Journal of Distributed Sensor Networks, Hindawi Publications.
10. J. Zhou, L. Hu, F. Wang, H. Lu, and K. Zhao, "An efficient multidimensional fusion algorithm for iot data based on partitioning," Tsinghua Science and Technology, vol. 18, no. 4, pp. 369–378,2013.
11. M. Mozumdar, A. Shahbazian, and N.-Q. Ton, "A big data correlation orchestrator for Internet of Things," in Proceedings of The IEEE World Forum on Internet of Things (WF-IoT'14), pp. 304–308, IEEE, Seoul, Republic of Korea, March2014.
12. Z.Ding, X.Gao, J.Xu, and H.Wu, "IOT-Statistic DB: a general statistical database cluster mechanism for big data analysis in the internet of things," in Proceedings of the IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, pp. 535–543, IEEE, August2013.
13. C. Wang, M. Daneshmand, M. Dohler, X. Mao, R. Q. Hu, and H. Wang, "Guest Editorial—special issue on internet of things (IoT): architecture, protocols and services", IEEE Sensors Journal, vol.13, no.10, pp.3505–3508, 2013.

14. A.J.Stankovic, " Research directions for the internet of things," IEEE Internet of Things Journal, vol.1, no.1, pp.3–9,2014.
15. Q.Wu, G.Ding, Y.Xuetal, "Cognitive internet of things: a new paradigm beyond connection", IEEE Internet of Things Journal, vol.1,no.2,pp.129–143,2014.