AI based HealthCare Platform for Real Time, Predictive and Prescriptive Analytics using Reactive Programming

Jagreet Kaur¹, Dr. Kulwinder Singh Mann²
¹Data Scientist, XenonStack, IT Park, Chandigarh, India.
²Professor and Head of the IT Department, GNDEC, Ludhiana, India

Abstract. AI in Healthcare needed to bring real, actionable insights and Individualized insights in real time for patients and Doctors to support treatment decisions., We need a Patient Centred Platform for integrating EHR Data, Patient Data, Prescriptions, Monitoring, Clinical research and Data. This paper proposes a generic architecture for enabling AI based healthcare analytics Platform by using open sources Technologies Apache beam, Apache Flink Apache Spark, Apache NiFi, Kafka, Tachyon, Gluster FS, NoSQL- Elasticsearch, Cassandra. This paper will show the importance of applying AI based predictive and prescriptive analytics techniques in Health sector. The system will be able to extract useful knowledge that helps in decision making and medical monitoring in real-time through an intelligent process analysis and big data processing.

1. Introduction
As with the change in the time and advancements in the technology, there is a need to make a systematic change to health systems to improve the quality, efficiency, and effectiveness of patient care. Chronic diseases like heart disease, stroke, cancer, and diabetes are considered as the most common, expensive, and preventable health problems but due to the poor health care systems, patients can’t able to take good care of the problems.

The strategic aim of value based Health Care is to ensure that everyone can use the health services they are needed for their good health and well-being. The focus on value-based care corresponds to an increased concentration on patient-centric care. By focussing on technologies and healthcare processes on patient results, a doctor, hospitals, and health insurance need to work with each other to personalize care that is effective, transparent in its delivery and billing, and measured based on satisfaction of the patient.

Suppose a patient is suffering from ache or pain and need to visit a physician. After listening to patient symptoms, Physician inputs them into the computer, which helps to show the latest research that is needed to know about how to diagnose and treat the problem. Patient has an MRI or an x-ray and a computer helps to detect a problem that could be too small for a human to see. Finally, a computer monitors at patient medical records and family history and compares both with the most recent research to suggest a treatment for a patient that is specifically personalized to his needs.

Nowadays, Patients need an affordable and high quality of healthcare. According to the Health Data, most of the information is not in structured, relational format. About 80% information is in an unstructured format. Due to the limited structured data about the patient’s health conditions is available and patients have only very limited opportunities to actively involved in the process. It is very difficult to use this massive unstructured amount of data from different sources to take the right decision for the right patient at the right time by the doctors. This will slow down the personalized
care to the patient. So, it is necessary to develop a new strategy or system to care for patients which involves the health of patients while decreasing the cost of care[1].

These services range from clinical care to the public services for individual patients that are helpful for the health of whole populations. There is a need to improve the healthcare quality and coordination so that outcomes are consistent with current professional knowledge. The cost of treatment for the problems should be reduced so that each and every patient get their personalized treatment at lower cost.

Adoption of Electronic Health Records and systematic collection of data by health care providers were predicted to improve the efficiency and quality of patient care. Machine learning is improving diagnostics, predicting outcomes, and just beginning to scratch the surface of personalized care. Human surgeons will not be replaced by machines in the future, but AI can positively help surgeons to make better clinical decisions or can even replace human decision in certain areas of healthcare (eg. radiology). The growing availability of healthcare data and speedy development of big data analytic methods have made possible the current successful applications of AI in healthcare [2].

Data analytics has become increasingly important in almost every region of the economy. Health care involves a diverse set of public and private data collection systems having different sources of Data. The volume of data has been increasing exponentially in the past era as health care contributors turned to EHRs, digitized laboratory slides, and high-resolution radiology videos and images. There are petabytes of data stored in databases of health insurance companies and the trillions of data points streaming from sensors like activity trackers and various other continuous monitoring devices. As shown in fig1. Text or natural language data resides in many fields including several fields such as output from medical devices, physician notes, nursing notes, surgical notes, radiology notes, pathology reports, admission notes, clinical data, Genomic data, Behaviour data etc. These fields may have valuable information about the patient including diagnosis, history, family history, complaints, statistics, and opinions, demographics, medical history, allergies, laboratory test results etc.

**Figure1. Health care data sources**

To transform the current healthcare system into a preventive, active and value-based system, the interoperability, exchange, and sharing of health data are needed. One of the major challenges is data integration i.e. to integrate the data that is obtained for each patient into one system, as that will allow for fast data analysis, and give physicians all the information they need to perfectly treat their patients [3]. However, most of the data is encrypted and access is restricted due to the patient privacy and many medical devices are not interoperable[4]. So, once a single database can be established, ML can solve the major problems. Machine learning provides a way to find reason and patterns about data, which enables physicians to move to personalized care known as precision medicine. There are many ways as how machine learning algorithms can be used in healthcare, but all of them depend on having sufficient amount of data and permission to use it. Previously, alerts and advise for a medical treatment have been developed based on physicians studies, and data into their software. However, that can restrict the accuracy of that data because it might come from different populations. On the other hand, Machine learning nan be refined using data that is available in that particular environment i.e. anonymous patient information from a hospital and the
area it serves. Physicians can use machine learning to predict hospital readmission for chronically ill patients. The patients that are at risk of being readmitted are identified that makes possible for providers to offer improved post-discharge support. By lowering the rate of readmission, it can improves the lives of those most at risk.

The rest of this paper is organized as follows: Section II gives an overview of Interdependency of Cloud Computing with Big Data Technologies. Section III describes about the Bigdata in terms of Healthcare. Section IV the related work. Section V describes the platform design in which steps needed to do analytics will be discussed. Section VI describes the architectural design. Section VII gives an overview about artificial intelligence that helps the system to understand from its experience and reactive machine learning.

### 1. Need of Health care system

- **Reduce readmissions.** Health care systems can reduce readmissions in an efficient and patient-centred manner. Clinicians can receive daily guidance as to which patients are readmitted and how they reduce that risk.

- **Prevent hospital acquired infections (HAIs).** Health systems can reduce HAIs, such as central-line associated bloodstream infections (CLABSIs)—40 percent of CLABSI patients die—by predicting which patients with a central line will develop a CLABSI. Clinicians can monitor high-risk patients and intervene to reduce that risk by focusing on patient-specific risk factors.

- **Reduce hospital Length-of-Stay (LOS).** Health systems can reduce LOS and improve other outcomes like patient satisfaction by identifying patients that are likely to have an increased LOS and then ensure that best practices are followed.

- **Predict chronic disease.** Machine learning can help hospital systems identify patients with undiagnosed or misdiagnosed chronic disease, predict the likelihood that patients will develop chronic disease, and present patient-specific prevention interventions.

- **Predict no-shows.** Health systems can create accurate predictive models to assess, with each scheduled appointment, the risk of a no-show, ultimately improving patient care and the efficient use of resources [5].

### 2. Interdependency of Cloud Computing with Big Data Technologies

Healthcare is moving towards the digital platform, more patient-oriented and data-driven. Availability of data, irrespective of the location of the patient and the clinician, has become the key to both patient satisfaction and improved clinical outcomes. Cloud technologies can significantly facilitate this trend. Healthcare functionality can be improved by cloud-based healthcare IT systems that provide the potential for broad interoperability and integration. Healthcare cloud services are Internet-based and use standard protocols, so connecting them to other systems and applications is very simple. The ability to share information easily and securely is a critical capability, and cloud services are good enablers for this[6].

Healthcare data needs to be shared across various settings and geographies which further burden the healthcare provider and the patient causing a significant delay in treatment and loss of time. Cloud services can permit remote access to applications and data[7]. The Internet connectivity enables access at anytime from anywhere. Cloud caters to all these requirements thus providing the healthcare organizations an incredible opportunity to improve services to their customers, the patients, to share information more easily than ever before, and improve operational efficiency at the same time [8].

Cloud computing has special features for physicians and patients that help to reduce the burden of heavy investments and to utilize outsourced resources, software, hardware, automated resource management, parallel computing, and utility computing[9]. The objectives of such usage include improvement in patient care, continuous feeds of physiological data are analyzed, financial processes are optimized and resource utilization.
2.1. Big Challenges in Health care system

- No unified format: As there is no standard method adopted in hospitals, labs to generate reports etc., they generate reports in their own formats so it is a challenge to integrate different formats and process the data.
- Large volume of data: If we take the case of image data like X-rays etc., a large size of data is generated for a single patient and there is a need to efficiently handle such data and extract useful information.
- Real time data handling: Need to capture the patient’s behavioral data through several sensors, social interactions and communications.
- Integration and sharing of data: Efficient data integration and sharing rely on standardized coding schemes and terminologies. Currently, standardized coding systems are used for high-level information coding only. A lot of information is not coded properly. In order to data integration and sharing, coding schemes, as well as data models need to be improved and standardized[10].

2.2. Real-Time Analytics & Predictive Analytics: Knowledge Discovery

To take the Right decision for the Right patient at the right time by the doctors, Doctors/Physicians need knowledgeable Data, Real-Time Analytics, and Predictive Modelling to improve the health care [11]. Real-time healthcare analytics can help to improve the quality of care, cost, and meet various requirements by streamlining and automating the process of collecting and measuring vast amounts of healthcare data. Too often today, there is a need to handle multi-sourced and multi-typed of data, including unstructured and semi-structured data and relational databases like SQL aren’t capable of handle this. Also, need to analyze terabytes or petabytes of data having response times of sub seconds and it is difficult for a single server to handle this. Last, batch processes are needed for some jobs. But many times, we need to analyze rapidly changing, multi-structured data in real time. It is very difficult to do Extract, Transform and Load(ETL) processes to cleanse data at that time. Real-time analytics tools handle these issue by bringing dissimilar information from many sources into the one place. Real-time healthcare analytics can help to improve the quality of care, reduce costs, and help to meet regulatory requirements by automating the process of collecting and measuring huge amounts of healthcare data. It can specify vision into historical, current, and predictive disease and treatment of unaddressed conditions that need the physician’s attention so that problems can be addressed before the patient is readmitted and the hospital is penalized.

Knowledge discovery and Data mining are two distinct terms. Knowledge Discovery is a kind of process used for extracting the useful information from the database and Data Mining is one of the steps involved in this process. Data Mining is the algorithm used for extracting the patterns from the data.

Knowledge Discovery involves various steps such as Data Cleaning, Data Integration, Data Selection, Data Transformation, Data Mining, Pattern Evaluation, Knowledge Presentation. Knowledge Discovery is a process that has total focus on deriving the useful information from the database, interpretation of storage mechanism of data, implementation of optimum algorithms and visualization of results. This process gives more importance on finding the understandable patterns of data that further used for grasping useful information.

Knowledgeable data is basically the knowledge extracted from information. It is essential for healthcare organizations to effectively manage both internal knowledge and externally generated knowledge to provide the best possible health care facilities to the patients. It increases decision-making capabilities of Physicians [12]. Predictive modeling is a technique to predict in advance a disease of a patient so that diagnose of the problem can be done in advance. Predictive Modelling benefit to large Healthcare providers to
discover the treatment for personalized care managing large patient populations. Predictive analyses can be done by inputting the EHR data into a data model that was processed by various analytic techniques and analyze the results. The rich sources of data along with analytics capabilities have potential for an increased understanding mechanisms of disease and better health care.

3. BigData in terms of Healthcare sector

Big Data refers to the vast quantities of data generated from the Internet and digitization of all sorts of information, including health records[13]. New big data technologies hold commitment for centralizing and analyzing these digital data in order to discover trends and make predictions. Big data is commonly characterized using a number of V’s. The first three are volume, velocity, and variety. Volume refers to the vast amounts of data that is generated every second, minutes, hour, and day in our digitized world. Variety refers to the ever increasing different forms of data that can come in such as text, images, voice. Velocity refers to the speed at which data is being generated and the pace at which data moves from one point to the next. Volume, variety, and velocity are the three main dimensions that characterize big data and describe its challenges[14]. More Vs has been introduced to the big data community as we discover new challenges and ways to define big data are veracity, visualization, viability, vulnerability, volatility, and value. Veracity refers to whether the data is trustable. Visualization refers to the way by which data is presented to the user. Viablity refers to the relevance of the data. Vulnerability refers to the security of the data. Volatility refers to the change of data with the time Last, Value refers to check the data for the meaningful results.

3.1. Big Data trends in Healthcare Sector

Evidence-based medicine should be given, which involves a use of all clinical data and factoring that into clinical and advanced analytics. To capture and bring all information of patient together is difficult but it gives a complete view for understanding into care coordination and population health management, and patient engagement and outreach.

The improvement of the quality of care can be focussed if the various extents of health data are combined in the automated health data analysis. The data dimensions include (a) the clinical data that describes the health status and history of a patient, (b) the clinical and administrative clinical process data, (c) the information about diseases as well as related population data, and (d) the awareness about changes. If the data analysis is limited to only one data feature, for example, the administrative and financial data, it will not be possible to identify new standards for personalized treatments. So, the integration of the various heterogeneous datasets and exchange of health data is an important requirement of big health data applications.

![Figure 2 Data Platform](image)

3.2. Key Elements of Healthcare Sector

Integrating Data: To understand every possible feature about customers, citizens, and clients to apply advanced analysis and computation to modify existing strategies, there is a need to integrate a large set of data and business analytics. In a similar manner, To identify new patterns and to improve healthcare by finding the right treatment for an individual’s, there is a need to link heterogeneous data sets securely.
Generating New Knowledge: The primitive uses of Big Data were to generate new visions through predictive analytics. In addition to the clinical and administrative data, integrate patient data and their environment that helps in better predictions and right treatments to the right patients. The predictions may also help in identifying the areas that need improvement such as treatments, early identification of worse health states, readmissions, etc.

Transforming Knowledge into Practice: Although the standardization of data collection and fresh analytical approaches to the big data revolution in healthcare are important, it's their practical application that will get it across the line.

The understandings that are obtained from big data have the potential to reform multiple extents of healthcare like the outcomes from comparisons, the effectiveness of various treatments, and the predictive models that can be used for diagnosing, treating, and delivering care.

4. Related Works

Now days, Healthcare analytics is a new trend in the field of analytics. Due to the advancements in technologies in Health care sector, there is a major breakthrough in data collection. To improve the processing of predictable healthcare system, we have proposed a series of Big Data Health Care System. There are many techniques proposed in order to proficiently process large volume of medical record. Akshay Raul, Atharva Patil, Prem Raheja, Rupali Sawant proposed “Knowledge Discovery, Analysis and Prediction in Healthcare using Data Mining And Analytics”[16]. In this paper, Author proposed a Healthcare system that will be used to create a public awareness about alternative drugs for a specific medicine, availability of that alternative medicine in an area. The Proposed system helps the patients to find an alternate medicine which is prescribed by the doctor. Aditi Bansal and Priyanka Ghare proposed “Healthcare Data Analysis using Dynamic Slot Allocation in Hadoop”. In this paper HealthCare System is analysis using Hadoop using Dynamic Hadoop Slot Allocation (DHSA) method[17]. This paper proposed a framework which focus on improving the performance of MapReduce workloads and maintain the system. DHSA will focuses on the maximum utilization of slots by allocating map (or reduce) slots to map and reduce tasks dynamically. Van-Dai Ta, Chuan-Ming Liu, Goodwill Wandile Nkabinde proposed “Big Data Stream Computing in Healthcare Real-Time Analytics”[18]. In this paper, Author proposed a generic architecture for big data healthcare analytic by using open sources, including Hadoop, Apache Storm, Kafka and NoSQL Cassandra. Wullianallur Raghupathi and Viju Raghupathi has proposed “Big data analytics in healthcare: promise and Potential”[19]. In this paper author proposed the potential of big data analytics in healthcare. The paper provides an overview of big data analytics for healthcare practitioners and researchers. Big data analytics in healthcare is growing into a promising field for providing vision from very large data sets and improving results while reducing costs. This paper proposes a generic architecture for big data healthcare analytic by using open sources, including Hadoop, Apache Storm, Kafka and NoSQL Cassandra.

However, all these works either consider the techniques to extract features form specific healthcare data sources or only focus on batch-oriented task to compute which has higher latency computing. In our proposed framework, data sources are supposed constantly coming with high rate, variety of formats, and high volume. Stream computing of real data and use of cloud computing in serving layer will enhance the results of healthcare analytics.

5. Platform Design

This demonstrates the major components of the Healthcare Analytical platform and the types of services that is needed to build a model.

Data Ingest: The first problem that everyone gets into is data ingest. The device data is coming from anywhere and of any type, it will be represented in a standard manner to analyze.

In this part of the application, solution of these problems is discussed:

How to ingest data from medical devices?

How to transform the device data into a format that can be analyzed in a Streams application?

What are the common data schema types when analyzing medical device data?
Data Preparation: Next, there is a need to prepare the data to analyze. Some of the common problems need to remove include Deduplication, Resampling, Normalization, Cleaning, Noise reduction etc[20].

Data Processing: Data processing can be done by using two types like Base Analytics and Aggregated analytics.

Base Analytics: The platform should provide a set of basic analytics of data. These analytics can be used as building blocks for the complex analytics and prediction models. Here are some of the things we can do like Simple vital analytics (calculating rolling average, raising alerts when vitals going beyond normal range), Analytics of ECG, EEG, ICP Waveforms etc.

Aggregated Analytics: This is an area where we combine and aggregate results from the base analytics to form more sophisticated analytic rules. For example, for septic shock detection, the user should be able to describe a rule like this:
- if temperature is > 37 degree Celsius or < 33 degrees Celsius
- and if heart rate is > 90
- and respiratory rate is > 19 or PaCO2 < 31 mm HG
- and WBC > 13,000/mm3, or < 5000/mm3, or > 9% band
- then raise an alert for early septic shock detection.

Patient Data Correlation / EMR Integration: For more complex analytics, we may need to use EMR data or doctor’s notes. For example, we may need to retrieve patient’s medical history. Or need to retrieve doctor’s notes that include some of the doctor’s observations[21]. These types of data can be ingested from existing hospital infrastructure using the HL7 / FHIR protocol.

Central Monitoring Dashboard: As part of the platform, there is a need to create a simple dashboard. This will help users to visualize their data and validate their analytic results. The dashboard can be web-based, mobile.

Alert and Notification Framework: When an important event occurs, we need to be able to notify and alert the right people to help the patient. In this part of the framework, we need to deliver notification to the right people based on alert types and patient information, alerts can be delivered via email, text messaging, etc. or should be displayed onto a dashboard.

6. Healthcare Platform Architecture

We require to collect the data, process the data, storage of data, and finally serve the data for analysis, machine learning, and dashboards. In this section, we propose an architecture including the advantages of batch and stream computing to enhance big data computing in healthcare. It can also deal with a large data set by providing cost reduction along low latency and better health care conditions. The fig 2. shows the detailed architecture of Healthcare Platform.

Data sources: As there are diverse sources of Health data available now a days like data from medical devices, Social media, EMR, Back- office-systems, Pharmacy, Medical images and HL7 events
Data ingestion: While developing the system, the first step needed is to integrate the data i.e. collect the data from various data sources. Data Ingestion can be from multiple sources, so need a unified platform where we can manage all our data sources. Here the data can be generated from two types of sources, one is real time data and other is batch data. To ingest data from real time data sources, Apache NiFi and Kafka API can be used.

- **Apache NiFi**: Apache NiFi is a distributed, scalable, fault-tolerant workflow automation platform. So, we will be having unified view on our Web UI from where we can monitor all our different data sources from where data ingestion is taking place[22].

- **Apache Kafka**: Apache Kafka is a Distributed Messaging System and disk based cache. It can store the data that is collected by Apache NiFi. It provides unified, high-throughput, low latency and works on publish-subscribe model.

To ingest data in batches, NFS (Network File system) can be used. Data from Batch files can be ingested using Apache NiFi and Kafka API also but to process older data that was stored on distributed servers can only be accessed through NFS. NFS protocol provides remote access to shared disks across networks. An NFS-enabled server can share directories and files with clients, allowing users and programs to access files on remote systems as if they were stored locally. An NFS-mounted cluster allows easy data ingestion of data sources such as files, images... from other machines leveraging standard Linux commands, utilities, applications, and scripts.Gluster FS is free and open source software that can be used for distributed storage of images and videos and used for data analysis whereas Tachyon provide in memory storage with fast data processing. It is also used to process the data at memory speed.

Data Processing: After ingestion of data from various sources, next step is to process the data i.e. to clean the data, normalize the data, remove all outliers, data mining, etc. For data processing, Apache spark and apache apex will be used[23].

Apache spark is a next generation distributed parallel processing framework having a rich set of APIs for machine learning, GraphX for graph processing[24]. Spark is much faster than MapReduce for iterative algorithms because Spark tries to keep things in memory whereas MapReduce involves more reading and writing from disk[25].

Apache Gearpump is a lightweight real-time big data streaming engine. It is a simple pump that consists of two gears. It is difficult to process stream data because of low latency assurance, infinite data, Out-of-order data, • Correctness requirement, Cheap to update user logics but Apache Gearpump solve all these problems as it provides good user interface, control the flow of data, out-of-order processing.

Apache Apex is an engine for processing streaming data. Some other engines that can also fulfill the requirement of stream processing are Apache storm, Apache flink[26]. But Apache Apex has a built-in support for fault-tolerance, scalability and focus on operability.

Apache Flink is an open source platform for distributed stream and batch data processing. Flink has an API that perform operations on streams of data. At the time of submission of a job, these operations are turned into a graph by Flink,. It also supports an API that are compatible with storm.

Data Storage: As already discussed, Health care data consists of a variety of data like images, doctor notes, lab reports, insurance data etc. Different types of storage system are also needed like cloud to store the data. To store the data Gluster FS, Tachyon, Cassandra, Titan DB and Elasticsearch can be used. Gluster FS is free and open source software that can be used for distributed storage of images and videos and data analysis. Tachyon is used to process the data at memory speed. Also provide in memory storage with fast data processing. Apache Cassandra is a free and open-source distributed database management system used to handle large amounts of data like sensor data and claim data. It is an open source project developed by Facebook Titan dB is a database used to store graph and link analysis. Elastic search is a search engine used for indexing.

7. AI Based Healthcare System

AI refers to ‘Artificial Intelligence’ which means making machines capable of performing intelligent tasks like human beings. AI performs automated tasks using intelligence.

Artificial intelligence in healthcare is a new research area that combines representational and computing techniques with the visions of expert physicians to produce tools for improving health care.
It is the study of ideas which allows computers to do the things that make people seem intelligent. The central goals of Artificial Intelligence are to make computers more practical and to understand the principles which make intelligence possible.

![Diagram](image)

**Figure 4. AI enabled Healthcare System**

There is need to make system AI enabled so that machine can automatically predict and prescribe the results from its own experience. To make system AI enabled, Natural Language processing, Knowledge representation, Automated reasoning, and Machine Learning should be in the system as shown in fig 3. NLP helps to make computer/machines as intelligent as human beings in understanding language.

### 7.1. Use Cases of AI based Healthcare System Architecture

One of the most effective use of AI based Healthcare system is to optimize the clinical process. AI based mobile app ask patients about their symptoms and provides an easy to understand information about their Health. The system uses natural Language processing to provide a rich experience and machine learning algorithms to create a map of patient condition and give a personalized experience[27].

AI is helpful in Monitoring patient vitals in real time. In a hospital, nurses do rounds and visit each and every bed after every few hours to manually measure and record patient vital signs. But the patient’s condition may drop between the time of planned visits. This means that physicians often respond to problems reactively, in situations where earlier arrival may have improved the patient’s wellbeing.

Latest wireless sensors can capture and transmit patient’s vitals far more often and these measurements can stream using some various new technologies. The real-time alerts can be generated and Physicians can use these signals and respond more rapidly to unexpected changes[28].

The treatment and prevention of dangerous diseases depends on detecting the symptoms at the right time. In various cases, early diagnosis can result in complete cure. Contrariwise, a late or wrong diagnosis can have damaging results. It is difficult for humans to make reliable decisions. AI algorithms can rapidly ingest billions of samples in short order and collect useful patterns. Machine learning algorithms can be used to make the knowledge bases used by expert systems and predictive analysis of data. Hence, AI based Healthcare system is helpful for patients as well as doctors.

### 7.2. Integrating Big Data, Analytics, Artificial Intelligence, and Machine Learning in Healthcare

With enormous amounts of computational power, machines can analyze large sets of data points and apply correlation modelling in a predictive way and in real time. Big Data technology has the potential to influence machine learning capabilities accurately and real-time decision-making capability helps in improving overall operating efficiency and reducing unnecessary costs.

In healthcare, machine learning plays a big role by understanding different parameters and correlate them with diseases. Machine learning is defined as a process in which computers use machine learning algorithms to analyse large set of data that is represented in non-linear ways, identify patterns, and make predictions that can be tested and confirmed.

At a high level, there is supervised machine learning, in which outcomes will be predicted from available data and information from previous outcomes. Further, in unsupervised machine learning, unknown patterns of outcomes will be predicted from data.

Machine learning can be used in healthcare by increasing efficiencies, saving money and saving lives. CT scan data can be analyze and cross-applied to patient records to see who is most at risk for a particular disease.
for disease. Physicians can predict post-discharge outcomes to reduce readmissions and enhance patient flow. They can make curative diagnostics faster, more accurate, and more accessible. Machine learning, deep learning, and cognitive computing are various steps towards a high degree of artificial intelligence, but they are different things. Deep learning is a subset of machine learning. It uses artificial neural networks that simulate human brain connections that become “trained” with the time to provide answers to questions with nearly 100 percent accuracy. AI is basically the intelligence that how we make machines intelligent, whereas machine learning is the process of implementing the compute methods that support it. AI algorithms can predict post-discharge results, reducing readmissions and optimizing patient flow. In other words, physicians can use these solutions to make medical diagnostics faster, more precise and manageable. For a caregiver, identifying a pattern is the first step which further solves the complicated process of treatment. Here, artificial intelligence, machine learning will do the same type of thing. Companion diagnostics will help to find the gaps in our current data resources and help us move into truly personalized medicine.

7.3. Anomaly detection to discover unusual patient management
Anomaly detection is used to identify unusual data that diverge from the majority of samples in a dataset. Recently, due to the advancement in data collection and electronic health record, anomaly detection techniques are also helpful in the health care sector. Anomalies are divided into three categories like point anomaly, contextual anomaly and collective anomaly.
To detect an unusual behavior of data, there are many sensors attached to the patient, used to monitor various physiological parameters, and transmit the data to an analytical device which must analyze the collected data, and raise alarms to the physicians only when the patient health degrades.

7.4. Reactive Machine Learning in Healthcare
All types of Healthcare can be divided into two categories: proactive and reactive.
Reactive healthcare includes reacting to an adversarial disease, injury, disorder or symptom. If a patient suffering with a fever and body aches than he may react by visiting the doctor. Depending on the doctor’s diagnosis, he or she may recommend you with antibiotics to help to body fight with the infection. Both the patient and the doctor are reacting to the symptoms[29].
Proactive healthcare differs from reactive healthcare in the sense that action need to be taken before the visibility of the symptoms. A patient can take a proactive approach towards his health by boosting the immune system with vitamin C, antioxidants, and by drinking plenty of fluids without waiting to detect the problem.
Healthcare is said to be reactive if it satisfies the four basic principles of reactive system i.e. responsive, resilient, elastic and message driven. Responsive healthcare system refers to proper results of the medicines prescribed by the physicians at all circumstances. Resilient health care system requires the engagement of all the system’s components, elastic refers to the healthcare system should be scalable and message driven architecture refers to the request-response messages for exchanging patient details etc. across different hospitals to provide the best of in-Patient and Out-Patient service integration within and across healthcare system.

8. Conclusions
This paper provides a generalized framework for personalized healthcare which influences the advantages of remote monitoring, cloud computing, big data and reactive machine learning. It provides systematic approach to support fast growing data of people with severe diseases. It also simplifies the task of physicians by not overwhelming them with false alerts. The proposed architecture can support for Artificial intelligence bssed healthcare analytics by providing batch and stream computing, extendable storage solution and query management. To achieve more efficient result from healthcare, there is still a need to handle the health care data that is growing time by time at high rate, and larger scale with tons of inconsistent data sources. A distributed system should be arranged to interchange data among labs, hospital systems, and clinical centres. Previous analytics tools have become transparent and friendly. But, when they emerged with open source development tools and platforms, it will become very complex, need complex programming and need the
application of a variety of Refine, KNIME and data mining can be used to improve the efficiency of data analytics.

9. References
[1] R. Weil, ’’Big Data In Health: A New Era For Research And Patient Care Alan R. Weil’’, Health Affair, Vol. 33, N° 7, pp 1110, 2014.
[2] A. Alexandru, D. Coardos, Big Data in Healthcare and Medical Applications in Romania, IEEE International Conference on Automation, Quality and Testing, Robotics, THETA 20th edition, 2016.
[3] T., Huang, L., Lan, ’’Promises and Challenges of Big Data Computing in Health Sciences’’, Big Data Research vol. 2, pp 2-11, 2015.
[4] K. Roney, If Interoperability is the Future of Healthcare, What's the Delay?, Becker's Hospital Review, 2012.
[5] Yichuan Wang, LeeAnn Kung, Terry Anthony Byrd, “Big data Analytics: Understanding its capabilities and potential benefits for healthcare organizations”, in Press Article,” Technology forecasting & social change”, Science Direct, Elsevier, 2016.
[6] Jonathan Northover, Brian McVeigh, Sharat Krishnagiri. Healthcare in the cloud: the opportunity and the challenge. MLD.
[7] CSCC, Impact of Cloud Computing on Healthcare, pp. 1-18, 2012.
[8] I.A.T. Hashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, S.U. Khan, The rise of “big data” on cloud computing: Review and open research issues, Information Systems, Vol. 47, pp. 98-115, 2015.
[9] P. Piletic, How is cloud and big data changing healthcare in 2016.
[10] Priyanka Jain, Sanjay Ojha, “Significance of Big Data Analytics”, International Journal of Software and Web Sciences (IJSWS), 2015
[11] Basma Boukenze, Hajar Mousannif, Abdelkrim Haqiq, “A Conception of a Predictive Analytics Platform in Healthcare Sector by Using Data Mining Techniques and Hadoop” in Proc. Conf. International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 8, August 2016.
[12] Akshay Raul, Atharva Patil, Prem Raheja, Rupali Sawant, “Knowledge Discovery, Analysis and Prediction in Healthcare using Data Mining and Analytics”, in Proc. Conference, 2nd International Conference on Next Generation Computing Technologies (NGCT-2016) Dehradun, India 14-16 October 2016.
[13] Frost & Sullivan, Drowning in big data? Reducing information technology complexities and costs for healthcare organizations, 2012.
[14] B. Feldman, E.M. Martin, T. Skotnes, Big Data in Healthcare Hype and Hope, October 2012. Dr. Bonnie 360, 2012.
[15] Akshay Raul; Atharva Patil; Prem Raheja; Rupali Sawant, “ Knowledge discovery, analysis and prediction in healthcare using data mining and analytics, in Proc. 2nd International Conference on Next Generation Computing Technologies (NGCT),IEEE, 2016.
[16] M. Roopa, Dr.S. Manju Priya, “A Review of Big Data Analytics in Healthcare” in Proc. Conf. International Journal for Scientific Research & Development| Sp. Issue – Data Mining 2015.
[17] Aditi Bansal and Priyanka Ghare, “Healthcare Data Analysis using Dynamic Slot Allocation in Hadoop” in Proc. Conf. International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-3 Issue-5, November 2014.

[18] Van-Dai Ta, Chuan-Ming Liu, Goodwill Wandile Nkabinde, “Big Data Stream Computing in Healthcare Real-Time Analytics”, in Proc. Conf. International Conference on Cloud Computing and Big Data Analysis, IEEE, 2016.

[19] J. Archenaa and E.A.Mary Anita, “A Survey of Big Data Analytics in Healthcare and Government” in Proc. Conf. 2nd International Symposium on Big Data and Cloud Computing (ISBCC’15), pp-408 – 413, 2015.

[20] Gualtieri, M; Curran,R, "The Forrester Wave™: Big Data Streaming Analytics, Q1 2016," Forrestr.com, Cambridge MA, 2016.

[21] Wullianallur Raghupathi and Viju Raghupathi, “Big Data analytics in Health care: Promise and Potential”, A review paper in Health Information Science and Systems, vol. 2, no. I, Feb. 2014, p. 3.

[22] Rotsnarani Suthy, Mrutyunjaya Panda, “Big Data Analysis using Hadoop: A Survey” in Proc. Conf. International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 7, July 2015.

[23] Mukesh Borana 1, Manish Giri 2, Sarang kamble, “Healthcare Data Analysis using Hadoop”, in Proc. Conf. International Research Journal of Engineering and Technology (IRJET), Volume: 02 Issue: 07, pp-583, Oct-2015.

[24] nifi.apache.org, “nifi.apache.org/minifi,” 18 November 2016.

[25] Dalia AbdulHadi AbdulAmeer, “Medical Data Mining: Health Care Knowledge Discovery Framework Based on Clinical Big Data Analysis” in Proc. Conf. International Journal of Scientific and Research Publications, Volume 5, Issue 7, July 2015.

[26] Meng et al, "MLib: Machine Learning in Apache Spark," Journal of Machine Learning Research 17, San francisco, 2016.

[27] Telmo da Silva Morais, “Survey on Frameworks for Distributed Computing: Hadoop, Spark and Storm” in Proc. Conf. Doctoral Symposium in Informatics Engineering, 2015.

[28] R. Evans, "Apache Storm, a Hands on Tutorial," IEEE International Conference on Cloud Engineering, Tempe, AZ, 2015, pp. 2-2, 2015.

[29] Harsh Pathak, Manas Rathi, “Introduction to Real-Time Processing in Apache Apex”, in Proc. Conf. International Journal of Research in Advent Technology, 2016.

[30] Steve G. Peters, James D. Buntrock,”'Big Data and the Electronic Health Record’”, Ambulatory Care Manage , Vol. 37, No. 3, pp. 206–210, 2014.

[31] Engineer Bainomugisha, Andoni Lombide Carreton, Tom van Cutsem, Stijn Mostinckx, and Wolfgang de Meuter. A survey on reactive programming. ACM Computing Surveys (CSUR), 45(4):52, 2013.