www.ijiarec.com

ISSN:2348-2079

Volume-6 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Big Data Analytics in Healthcare

[1]Mrs. K E.Eswari,M.C.A.,M.Phil.,M.E., Associate Professor/MCA
[2]Mr. M.Mugilan,.II MCA, [3]Ms. A.Praveena, II MCA,
[4]Ms. G.Gowthini, II MCA, [5]Ms. S.Nivedhika II MCA
Department of MCA, Nandha Engineering College(Autonomus),Erode-52.
E.Mail-ID: meddymugilan@gmail.com, praveenagowtham284@gmail.com,
gowthinigunasekaran96@gmail.com, nivethika7623@gmail.com

***Abstract*—the rapidly expanding field of big data analytics has started to play a pivotal role in the evolution of healthcare practices and research. It has provided tools to accumulate, manage, analyze and assimilate large volumes of disparate, structured and unstructured data produced by current healthcare systems. Big data analytics has been recently applied towards aiding the process of care delivery and disease exploration. However, the adoption rate and research development in this space is still hindered by some fundamental problems inherent within the big data paradigm. In this paper, we discuss some of these major challenges with a focus on upcoming and promising areas of medical research; image, signal and genomics based analytics. Recent research which targets utilization of large volumes of medical data while combining multi modal data from disparate sources are discussed. Potential areas of research within this field which have ability to provide meaningful impact on healthcare delivery are also examined.**

I. INTRODUCTION

Big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it's not the amount of data that's important. It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves[1].

While the term "big data" is relatively new, the act of gathering and storing large amounts of information for eventual analysis is ages old. The concept gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the three Vs:

Volume: Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data. In the past, storing it would've been a problem – but newtechnologies (such as Hadoop) [4] have eased the burden.

*Velocity***:** Data streams in at an unprecedented speed and must be dealt with in a timely manner. data in near-real time.

*Variety***:** Data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

At SAS, we consider two additional dimensions when it comes to big data:

*Variability***:** In addition to the increasing velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Is something trending in social media? Daily, seasonal and event-triggered peak data loads can be challenging to manage. Even more so with unstructured data.

*Complexity:*Today's data comes from multiple sources, which makes it difficult to link, match, cleanse and transform data across systems. However, it's necessary to connect and correlate relationships, hierarchies and multiple data linkages or your data can quickly spiral out of control.

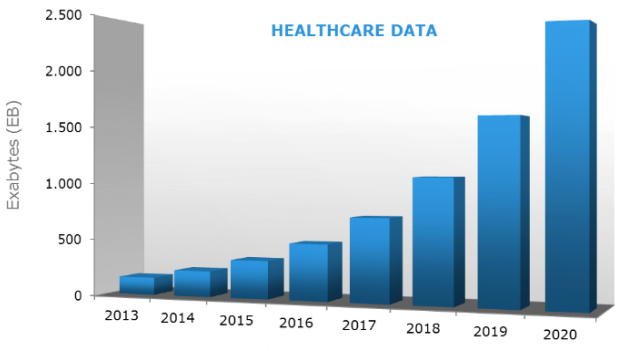

Fig.1. Healthcare Data

II. RELATEDWORK

Big data affects organizations across practically every industry. See how each industry can benefit from this onslaught of information.

A. Banking

With large amounts of information streaming in from countless sources, banks are faced with finding new and innovative ways to manage big data. While it's important to understand customers and boost their satisfaction, it's equally important to minimize risk and fraud while maintaining regulatory compliance. Big data brings big insights, but it also requires financial institutions to stay one step ahead of the game with advanced analytics.

B.Education

Educators armed with data-driven insight can make a significant impact on school systems, students and curriculums. By analyzing big data, they can identify at-risk students, make sure students are making adequate progress, and can implement a better system for evaluation and support of teachers and principals.

C.Government

When government agencies are able to harness and apply analytics to their big data, they gain significant ground when it comes to managing utilities, running agencies, dealing with traffic congestion or preventing crime. But while there are many advantages to big data, governments must also address issues of transparency and privacy.

D.Health Care

Patient records. Treatment plans. Prescription information. When it comes to health care, everything needs to be done quickly, accurately – and, in some cases, with enough transparency to satisfy stringent industry regulations. When big data is managed effectively, health care providers can uncover hidden insights that improve patient care.

E.Manufacturing

Armed with insight that big data can provide, manufacturers can boost quality and output while minimizing waste – processes that are key in today's highly competitive market. More and more manufacturers are working in an analytics-based culture, which means they can solve problems faster and make more agile business decisions.

III. WHY BIG DATA IN HEALTHCARE IS SO NEEDED

There's a huge need for big data in healthcare as well, due to rising costs in nations like the United States. As a McKinsey report states, "After more than 20 years of steady increases, healthcare expenses now represent 17.6 percent of GDP nearly $600 billion more than the expected benchmark for a nation of the United State's size and wealth."

In other words, healthcare costs are much higher than they should be, and they have been rising for the past 20 years. Clearly, we are in need of some smart, data-driven thinking in this area. And current incentives are changing as well: many insurance companies are switching from fee-for-service plans (which reward using expensive and sometimes unnecessary treatments and treating large amounts of patients quickly) to plans that prioritize patient outcomes.

As the authors of the popular Freakonomics books have argued, financial incentives matter – and incentives that prioritize patient's health over treating large amounts of patients are a good thing. Why does this matter for big data?

Well, in the previous scheme, healthcare providers had no direct incentive to share patient information with one another, which had made it harder to utilize the power of big data. Now that more of them are getting paid based on patient outcomes, they have a financial incentive to share data that can be used to improve the lives of patients while cutting costs for insurance companies.

Finally, physician decisions are becoming more and more evidence-based, meaning that they rely on large swathes of research and clinical data as opposed to solely their schooling and professional opinion. As in many other industries, data gathering and management is getting bigger, and professionals need help in the matter. This new treatment attitude means there is a greater demand for big data analytics in healthcare facilities than ever before, and the rise of SaaS business intelligence tools is also answering that need.

One of the biggest hurdles standing in the way to use big data in healthcare is how medical data is spread across many sources governed by different states, hospitals, and administrative departments. Integration of these data sources would require developing a new infrastructure where all data providers collaborate with each other.

Equally important is implementing new data analysis tools and strategies. Healthcare needs to catch up with other industries that have already moved from standard regression-based methods to more future-oriented like predictive analytics, machine learning, and graph analytics.

However, there are some glorious instances where healthcare doesn't lag behind, such as EHRs (especially in the US.) So,

even if these services are not your cup of tea, you are a potential patient, and so you should care about new healthcare analytics applications. Besides, it's good to take a look around sometimes and see how other industries cope with big data. They can inspire you to adapt and adopt some good ideas.

Fig.2. Big Data in Health Care

IV. APPLICATIONS OF BIG DATA ANALYTICS IN HEALTHCARE THAT CAN SAVE PEOPLE

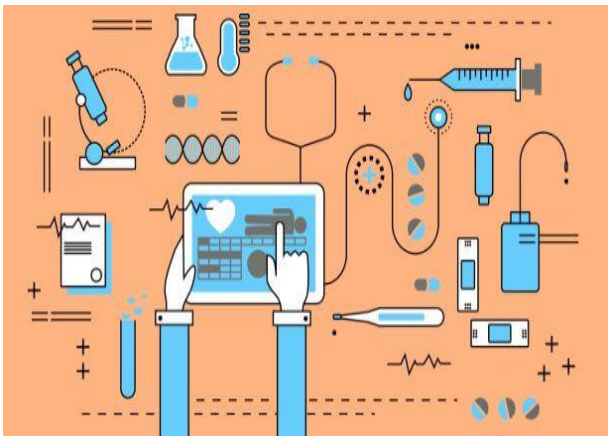

Fig 3: Big Data Analytics in Healthcare

Big Data has changed the way we manage, analyze and leverage data in any industry. One of the most promising areas where big data can be applied to make a change is healthcare. Healthcare analytics have the potential to reduce costs of treatment, predict outbreaks of epidemics, avoid preventable diseases and improve the quality of life in general. Average human lifespan is increasing along world population, which poses new challenges to today's treatment delivery methods. Healthcare professionals, just like business entrepreneurs, are capable of collecting massive amounts of data and look for best strategies to use these numbers. In this article, we would like to address the need of big data in healthcare: why and how can it help? What are the obstacles to its adoption? We will then provide you with 9 big data examples in healthcare that already exist and that we benefit from.

A. Big Data Is Saving Lives In Paris

Fig 4:Big Data Is Saving Lives In Paris

For our first example of big data in healthcare, we will look at one classic problem that any shift manager faces: how many people do I put on staff at any given time period? If you put on too many workers, you run the risk of having unnecessary labor costs add up. Too few workers, you can have poor customer service outcomes – which can be fatal for patients in industries like healthcare.

Big data is helping to solve this problem, at least at a few hospitals in Paris. A Forbes article details how four hospitals which are part of the Assistance Publique-Hôpitaux de Paris have been using data from a variety of sources to come up with daily and hourly predictions of how many patients are expected to be at each hospital.

One of the key data sets is 10 years' worth of hospital admissions records, which data scientists crunched using "time series analysis" techniques. These analyses allowed the researchers to see relevant patterns in admission rates. Then, they could use machine learning to find the most accurate algorithms that predicted future admissions trends.

Summing up the product of all this work, Forbes states: "The result is a web browser-based interface designed to be used by doctors, nurses and hospital administration staff – untrained in data science – to forecast visit and admission rates for the next 15 days. Extra staff can be drafted in when high numbers of visitors are expected, leading to reduced waiting times for patients and better quality of care."

B.Electronic Health Records (EHRs)

It's the most widespread application of big data in healthcare. Every patient has his own digital record which includes demographics, medical history, allergies, laboratory

test results etc. Records are shared via secure information systems and are available for healthcare providers from both public and private sector. Every record is comprised of one modifiable file, which means that doctors can implement changes over time with no paperwork and no danger of data replication.

EHRs can also trigger warnings and reminders when a patient should get a new lab test or track prescriptions to see if a patient has been following doctors' orders.

Although EHR are a great idea, many countries still struggle to fully implement them. U.S. has made a major leap with 94% of hospitals adopting EHRs according to this HITECH research, but the EU still lags behind. However, an ambitious directive drafted by European Commission is supposed to change it: by 2020 centralized European health record system should become a reality.

Kaiser Permanente is leading the way in the U.S., and could provide a model for the EU to follow. They've fully implemented a system called HealthConnect that shares data across all of their facilities and makes it easier to use EHRs. A McKinsey report on big data in healthcare states that "The integrated system has improved outcomes in cardiovascular disease and achieved an estimated $1 billion in savings from reduced office visits and lab tests."

C. Real-time Alerting

Other examples of big data in healthcare share one crucial functionality – real-time alerting. In hospitals, Clinical Decision Support (CDS) software analyzes medical data on the spot, providing health practitioners with advice as they make prescriptive decisions.However, doctors want patients to stay away from hospitals to avoid costly in-house treatments. Personal analytics devices, already trending as business intelligence buzzwords last year, have the potential to become part of a new healthcare delivery strategy. Wearables will collect patients' health data continuously and send this data to the cloud.Additionally, this information will be accessed to the database on the state of health of the general public, which will allow doctors to compare this data in socioeconomic context and modify the delivery strategies accordingly.

Healthcare institutions and care managers will use sophisticated tools to monitor this massive data stream and react every time the results will be disturbing.For example, if patient's blood pressure increases alarmingly, the system will send an alert in real time to the doctor who will then take action to reach the patient and administer measures to lower the pressure.Another example is that of Asthma polis, which has started to use inhalers with GPS-enabled trackers in order to identify asthma trends both on an individual level and looking at larger populations. This data is being used in conjunction with data from the CDC in order to develop better treatment plans for asthmatics.

D.Big Data is Helping To Prevent Opioid Abuse In The US

Our fourth example of big data analytics in healthcare is tackling a serious problem in the US. Here's a sobering fact: as of this year, overdoses from misused opioids have caused more accidental deaths in the U.S. than road accidents, which were previously the most common cause of accidental death.

Fig 5:Big Data is Helping To Prevent Opioid Abuse In The US

Big data influencer Bernard Marr writes about the problem in a Forbes article. The situation has gotten so dire that Canada has declared opioid abuse to be a "national health crisis," and President Obama earmarked $1.1 billion dollars for developing solutions to the issue while he was in office.

Once again, an application of big data analytics in healthcare might be the answer everyone is looking for: data scientists at Blue Cross Blue Shield have started working with big data experts at Fuzzy Logix to tackle the problem. Using years of insurance and pharmacy data, Fuzzy Logix analysts have been able to identify 742 risk factors that predict with a high degree of accuracy whether someone is at risk for abusing opioids.

As Blue Cross Blue Shield data scientist Brandon Cosley states in the Forbes piece: "It's not like one thing – 'he went to the doctor too much' – is predictive … it's like 'well you hit a threshold of going to the doctor and you have certain types of conditions and you go to more than one doctor and live in a certain zip code…' Those things add up."

To be fair, reaching out to people identified as "high risk" and preventing them from developing a drug issue is a delicate undertaking. However, this project still offers a lot of hope towards mitigating an issue which is destroying the lives of many people and costing the healthcare system a lot of money.

E. Predictive Analytics in Healthcare

We have already recognized predictive analytics as one of the biggest business intelligence trend two years in a row, but the potential applications reach far beyond business and much

further in the future. Optum Labs, an US research collaborative, has collected EHRs of over 30 million patients to create a database for predictive analytics tools that will improve the delivery of care.

The goal is to help doctors make big data-informed decisions within seconds and improve patients' treatment. This is particularly useful in case of patients with complex medical histories, suffering from multiple conditions. New tools would also be able to predict, for example, who is at risk of diabetes, and thereby be advised to make use of additional screenings or weight management.

F.Using Health Data For Informed Strategic Planning

The use of big data in healthcare allows for strategic planning thanks to better insights into people's motivations. Care mangers can analyze check-up results among people in different demographic groups and identify what factors discourage people from taking up treatment.

University of Florida made use of Google Maps and free public health data to prepare heat maps targeted at multiple issues, such as population growth and chronic diseases. Subsequently, academics compared this data with the availability of medical services in most heated areas. The insights gleaned from this allowed them to review their healthcare delivery strategy and add more care units to most problematic areas.

G. Big Data Just Might Cure Cancer

Another interesting example of the use of big data in healthcare is the Cancer Moonshot program. Before the end of his second term, President Obama came up with this program that had the goal of accomplishing 10 years worth of progress towards curing cancer in half that time.

This bold goal led to a panel which gave recommendations that included some big data use cases.

Medical researchers can use large amounts of data on treatment plans and recovery rates of cancer patients in order to find trends and treatments that have the highest rates of success in the real world.

For example, researchers can examine tumor samples in biobanks that are linked up with patient treatment records. Using this data, researchers can see things like how certain mutations and cancer proteins interact with different treatments and find trends that will lead to better patient outcomes.

This data can also lead to unexpected benefits, such as finding that Desipramine, which is an anti-depressant, has the ability to help cure certain types of lung cancer.

However, in order to make these kinds of insights more available, patient databases from different institutions such as hospitals, universities, and nonprofits need to be linked up.

Then, for example, researchers could access patient biopsy reports from other institutions. Another potential use case would be genetically sequencing cancer tissue samples from clinical trial patients and making these data available to the wider cancer database.

But, there are a lot of obstacles in the way, including:

- Incompatible data systems. This is perhaps the biggest technical challenge, as making these data sets able to interface with each other is quite a feat.
- Patient confidentiality issues. There are differing laws state by state which govern what patient information can be released with or without consent, and all of these would have to be navigated.
- Simply put, institutions which have put a lot of time and money into developing their own cancer dataset may not be eager to share with others, even though it could lead to a cure much more quickly.

However, as an article by Fast Company states, there are precedents to navigating these types of problems: "…the U.S. National Institutes of Health (NIH) has hooked up with a half-dozen hospitals and universities to form the Undiagnosed Disease Network, which pools data on super-rare conditions (like those with just a half-dozen sufferers), for which every patient record is a treasure to researchers."

Hopefully, Obama's panel will be able to navigate the many roadblocks in the way and accelerate progress towards curing cancer using the strength of big data.

H.Telemedicine

Telemedicine has been present on the healthcare services market for over 40 years, but only today, with the arrival of online video conferences, smartphones, wireless devices, and wearables, has it been able to come into full bloom. The term refers to delivery of remote clinical services using technology.

It is used for primary consultations and initial diagnosis, remote patient monitoring, and medical education for health professionals. Some more specific uses include telesurgery – doctors can perform operations with the use of robots and high-speed real-time data delivery without physically being in the same location with a patient.

Clinicians use telemedicine to provide personalized treatment plans and prevent hospitalization or re-admission. Such use of big data in healthcare can be linked to the use of predictive analytics as seen previously. It allows clinicians to

predict acute medical events in advance and prevent deterioration of patient's conditions.

By keeping patients away from hospitals, telemedicine helps to reduce costs of healthcare and improve the quality of service. Patients can avoid waiting lines and doctors don't waste time for unnecessary consultations and paperwork. Telemedicine also improves the availability of healthcare as patients' state can be monitored and consulted anywhere and anytime.

I. Big Data Is Helping To Prevent Unnecessary ER Visits

Saving time, money and energy using big data analytics in healthcare is necessary. What if we told you that over the course of 3 years, one woman visited the ER more than 900 times. This woman's issues were exacerbated by the lack of shared medical records between local emergency rooms, increasing the cost to taxpayers and hospitals, and making it harder for this woman

Fig 6:Big Data Is Helping To Prevent Unnecessary ER Visits

"Everybody meant well. But she was being referred to three different substance abuse clinics and two different mental health clinics, and she had two case management workers both working on housing. It was not only bad for the patient, it was also a waste of precious resources for both hospitals."

In order to prevent future situations like this from happening, Alameda county hospitals came together to create a program called PreManage ED, which shares patient records between emergency departments.

This system lets ER staff know things like:

- If the patient they are treating has already had certain tests done at other hospitals, and what the results of those tests are
- If the patient in question already has a case manager at another hospital, preventing unnecessary assignments
- What advice has already been given to the patient, so that a coherent message to the patient can be maintained by healthcare providers.

V. CONCLUSION

Due to rapid enhancement in big data prediction and analysis healthcare domain has got a valuable attention from recent few years. Existing traditional machine learning approaches take much time in computation when data set volume increase. Similarly, in health care domain huge data is collected from different sources and researchers always try to make problem simpler to patient. In big data, we can find remarkable and hidden information that can help in understanding the nature of problem more deeply.

REFERENCES

[1] Andrew McAfee, Erik Brynjolfsson, Thomas HDavenport, DJ Patil, and Dominic Barton. Big data:the management revolution. Harvard Bus Rev, 90(10):60_68, 2012.

[2] Cli_ord Lynch. Big data: How do your data grow? Nature,455(7209):28_29, 2008.

[3] Adam Jacobs. The pathologies of big data. Communications of the ACM, 52(8):36_44, 2009.

[4] Paul Zikopoulos, Chris Eaton, et al. Understanding big data: Analytics for enterprise class hadoop and streaming data. McGraw-Hill Osborne Media, 2011.

[5] James Manyika, Michael Chui, Brad Brown, et al. Big data: The next frontier for innovation, competition, and productivity. 2011.

[6]S. Suthaharan, "Big data classification: Problems and challenges in network intrusion prediction with machine learning," SIGMETRICS Perform. Eval.Rev., vol. 41, no. 4, pp. 70–73, 2014.

[7] B. P. RAO, "of the notes: Brief notes on big data: Acursory look," 2015.

[8] R. Hermon and P. A. Williams, "Big data in healthcare: What is it used for?" 2014.

[9] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: promise and potential," Health Information Science and Systems, vol. 2, no. 1, p. 1, 2014.

[10] O. H. Fang, N. Mustapha, and M. N. Sulaiman,"Integrating biological information for feature selection in microarray data classification," in Second International Conference on Computer Engineering and Applications (ICCEA), vol. 2.IEEE, 2010, pp. 330–334.

[11] P. Groves, B. Kayyali, D. Knott, and S. Van Kuiken, "The big datarevolution in healthcare," McKinseyQuarterly,vol.2,2013