Volume 5, Issue 9, September 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Data Collaboration for Insurance and Health Sciences: An Approach to Artificial Intelligence

Damodar Padubidri Bhat
Bachelor of Engineering (2014)
Department of Electronics and Communication Engineering
Sri Jayachamarajendra College of Engineering, Mysore, Karnataka, India

Abstract*—Data collaboration means visualizing data from all of your different data sources, and getting this data to the right people, in the right format, in time to use it in making effective decisions. Today's operations are swamped with data from multiple sources from every layer of the business. Data available in the insurance domain is powerful to build solutions to the field of health science like implementing health policies for the issues highlighted by the data of insurance sector and vice versa. This paper talks about the possibilities for insurance and health care sectors through Artificial Intelligence or Intelligent systems.

Keywords*—Insurance, Health Science, Artificial Intelligence (AI), Clusters, Data Extraction, In-memory database

I. INTRODUCTION

To have a collaboration of data from the field of Insurance to Health Science/Care, the most fundamental aspect is the accessibility of data. Insurance as we know is the equitable transfer of risk of a loss, from one entity to another for exchange of money. Health Care too is related monetary value. Apart from all the monetary aspects of these fields, one thing that draws attention is the amount of data present in these sectors. Data always has an important story to tell. It relies on us to give them a clear and convincing voice. To communicate effectively with data, you need to tell a story with it. While data relies on logic and reasoning, decisions are often made based on emotion. Merging logic and emotion can be a powerful combination to drive action from your insights.The challenge in making this happen is the problem of disparate data. At the base of the data structure, there's the real-time layer: instrumentation, sensors, PLCs, etc., all pumping out streams of information. At the top level, there's the business layer where we find things like financial applications.

Insurance as we know is undergoing a pervasive change from the traditional premium methodology to pay per performance initiatives. This is achieved through data collaboration of customer performance and probability of occurrence of an event. Health care too has been prevalent in the area of data collaboration. This area of data collaboration will enhance the decision making and policy frameworks in the sector. The data collaboration system can also perform some alarming functions, monitoring correlations between tags. Currently these industries are lacking this facility as the data collected or the historic data is not in a *clustered* manner.

Clusters

Cluster, as we define is the collection of data or observation regarding an attribute of a system. There can be a single or multiple clusters involved with a system. Clusters pertaining to insurance may be the type of insurance people avail, group insurances, claim causes, amount of claims for a particular reason. Similarly, health sector pertains to data relating to disease in a particular area, type of schemes the patient avail in medical claims, number of patients in a particular area etc.

Current framework of Insurance

Advances in artificial intelligence, also known as cognitive computing, are starting to cause a seismic shift in the professions. The eventual result is the eradication of many positions and the changing – usually lesser – roles for the "survivors" of this transformation. All the professions such as investment advisors and accountants will be impacted. Life insurance agents will also be severely affected. While this paradigm shift is going to take years and is dependent on technological innovation, coupled with the speed of complementary social change, it is an eventuality. The ability to source and construct life insurance portfolios, facilitate underwriting, and monitor policies can all be accomplished by the robo-life agent. Such an approach would often prove to be both substantially more efficient, a way to provide superior solutions, and considerably less expensive. It is these critical reasons, the vast majority of life insurance agent of today will, in time, become a relic of a previous generation.

There will be strong and determined opposition to this industry transformation. Certainly, many of today's life insurance agents will do everything in their power to fight back. They will likely slow down the process, somewhat. Moreover, many of the life insurance carriers will also push back for this evolution of the distribution system will severely and

© 2015, IJARCSSE All Rights Reserved

detrimentally impact some of them resulting in consolidation. Nevertheless, advances in cognitive computing will ultimately make this industry transformation a fait accompli.

It is important to note, that even as today's life insurance agents succumb, robo-life insurance agents will predominantly not directly replace them. People can certainly buy life insurance direct, but that is not having a meaningful effect on the sale of life insurance by agents. As the saying goes, "Life insurance is sold, not bought." What will likely happen is that other professionals – primarily attorneys and secondarily accountants – will incorporate the services of robo-life insurance agent into their practices. Instead of taking a commission, they will take a dramatically lower fee. The significant cost savings will be passed onto the purchasers. It is also important to keep in mind that the traditional business models of attorneys and accountants will also be upended by artificial intelligence.

None of this is going to happen quickly. However, it will occur incrementally, and when it does occur, the life insurance agent of today will pretty much become an anachronism. This will certainly be the case as the commission structure that supports agent-based distribution of life insurance is eradicated.

Currentframework in Health Science

The Indian healthcare system faces many challenges, including skyrocketing costs, high rates of drug-resistant and hospital-acquired infections, and failures of care delivery leading topreventable adverse health events. Overtreatments, poor execution of care, and failure to adopt best practices for preventive care and patient safety have huge and directly measurable impacts on both healthcare costs and patient outcomes.On the other hand, both the increasing availability of electronic health data and the on-going development of methodological approaches to analyse these data suggest the potential for the use of artificial intelligence and machine learning methods to improve the quality and lower the cost of patient care.

Electronic health records (EHR) have become more available. Meanwhile, health insurance providers and non-profits have committed to providing health insurance claims data with the goal of reducing costs while improving the quality and availability of coverage. Such sources provide detailed, time-stamped, and highly multivariate data for a large patient population, enabling the use of *AI* techniques to connect care practices and outcomes. However, the data's size and complexity—as well as the variability in content and format between different providers, data types, and care settings—create huge challenges. Additionally, the potential danger of the violation of patients' privacy has significant moral and legal ramifications, requiring extreme care in the use of health data.

Recent advances in AI for Patient Care

Recent Advances in AI for Patient Care Clinical decision support systems (CDSS) were one of the first successful applications of AI, focusing primarily on the diagnosis of a patient's condition given his symptoms and demographic information. Work on CDSS for medical diagnosis began in the early 1970s with Mycin-arule-based expert system for identifying bacteria that cause severe infections and recommending antibiotics to treat these infections. David Heckerman and his colleagues developed Pathfinder, which used Bayesian networks (a graphical model that encodes probabilistic relationships among variables of interest) to help pathologists more accurately diagnose lymph-node diseases. AI has also been useful for computer-aided detection of conspicuous structures (such as tumours or polyps) in medical images. Such approaches assist in the screening of mammography images, as well as the diagnosis of various forms of cancer, coronary artery disease, and congenital heart defects.

More recent advances in machine learning and AI build predictive models and make real-time inferences from a large patient population for purposes including alerts, stratifying risk, and predicting the length of stay. Several of these approaches focus on critical care, using physiological data that are routinely recorded in intensive care units. New-borns data of first 3 hours of life is collected to develop a hierarchical Bayesian model, the time seriestopic model.

II. METHODOLOGY AND VISUAL REPRESENTATIONS

Artificial intelligence is still in the very early stages of development–in so many ways, it can't match our own intelligence–and computers certainly can't replace doctors at the bedside. But today's machines are capable of crunching vast amounts of data and identifying patterns that humans can't. Artificial intelligence–essentially the complex algorithms that analyse this data–can be a tool to take full advantage of electronic medical records, transforming them from mere e-filing cabinets into full-fledged doctors' aides that can deliver clinically relevant, high-quality data in real time.

AI system for Data Collaboration and Extraction

Collaborative Decision Making with Data Science requires a team of humans and machines to work together. This requires humans and machines to work as a team. We all know, great teams are not just set of great individuals but require great collaboration and enhancing each other's strength.

This is true for every single activity in data science and analytics. As a data scientist, one has to be able to classify aspects that require brute force and others that require intuition and experience. Additionally, while designing decision support system one has to respect human intuition and judgement, along with the calculation power of machines. This will require data scientists to design systems that promote appropriate human and machine collaborations rather than systems skewed in either direction. Even decision makers need to be aware of this aspect that great decisions are made in collaboration with machines.

© 2015, IJARCSSE All Rights Reserved

The AI system mainly requires the following elements.

1. Access to data from the required sources with validity.
2. Algorithms for data extraction based on the requirement.
3. Data Collaboration Applications/Tools
4. Visual representations of the Data.

1. Access to data from the required sources with validity.

Even a passing understanding of information technology, we know that computers can recognize text and spoken words. We love-to-hate automated phone systems that use voice recognition and computer generated voices, but in truth, there are some that work pretty well and the performance of such programs is improving at a steady pace. We are also increasing our expectations of computers and handheld 'smart' devices when it comes to spoken commands and searches. We can thank tools such as Google's voice search option and apps like Dragon Dictate and Dragon Search from Nuance Communications for pushing the envelope in this area. However, for a computer to have the ability to make the necessary connections between words and their meanings in the context of what is being said or read is truly difficult. This is especially true when the knowledge domain becomes wider than the needs of an airline reservation system or searching for a restaurant in an unfamiliar location. These search engines have enabled unstructured data access through efficient algorithms. Intelligent systems for unstructured search are complex and high in terms of cost. When we consider insurance and health sciences as domain, the data basically are structured. But since the data related to insurance and health sciences/care include personal information, the organizations will decline the idea of AI system use for Insurances and Health Sciences. Insufficient data cannot derive appropriate results as the results for these domains are critical and extremely fragile.

2. Algorithms for data extraction based on the requirement

Data Extraction from the World Wide Web is a well-known, non-solved, and a critical problem when complex information systems are designed. These problems are related to the extraction, management and reuse of the huge amount of Web data available. These data have usually a high heterogeneity, volatility and low quality (i.e. format and content mistakes), so it is quite hard to build reliable systems.The integration of information needs from flexible and scalable mechanisms to obtain the necessary data from available sources. However, if these sources are not relational-based, or no design has been previously made by an expert (i.e. a database designer) it is usually quite hard to build, and maintain those mechanisms.Web Data Extraction (WDE) is a well-known and non-solved problem. It is related to the extraction, management and reuse of a huge amount of Web data available. These data have usually a high heterogeneity, volatility and low quality. One popular approach to address this problem is related to the concept of Wrappers. The Wrappers are specialized programs that automatically extract data from documents and convert the information stored into a structured format. Three main functions need to be implemented in a Web Wrapper. First, it must be able to download HTML pages from a website. Second, it must search for, recognize, and extract the specified data. Third, it has to save this data in a suitably structured format to enable further manipulation. The second function, related to a pattern-matching problem, is a hot topic for a wide number of research areas such as Artificial Intelligence (AI), Information Extraction (IE), Databases, Intelligent Agents (IA) or Semantic Web (SW). These areas have provided interesting results, and advances in several topics such as machine learning algorithms, automatic (or semi-automatic) Web-based data extraction tools, ontologies management(browsing, interaction, extraction), Web services (orchestration, composition), Service Oriented Architectures, etc. Due to the complex features of the Web, there is still an important open work in this area

For Insurance and Health Sciences, it is not possible to extract the data from World Wide Web as the data is confidential and is not available on the web. Here data extraction can be done only through mutual agreements between the organizations to share data to solve a chronic problem related to everyone in the domain. However, if there is mutual acceptance of the idea for AI implementation, the system required for data extraction should be fast enough to process the amount of data that is derived from this domain. In-memory database is preferred for this implementation as it has fastest inserts and extraction algorithms.

3. Data Collaboration Applications/Tools

Tools are essential to collaboration among team members, enabling the facilitation, automation, and control of the entire development process. Adequate tool support is especially needed in global software engineering because distance aggravates coordination and control problems, directly or indirectly, through its negative effects on communication. Following are some of the tools used for data collaboration.

A. Version Control Systems
B. Trackers
C. Build Tools
D. Modellers
E. Knowledge Centres
F. Communication Tools

In an industries point of view the tools built have analysed few keys issues mentioned below.

1. Organizations often lack visibility and control over how they share information.

© 2015, IJARCSSE All Rights Reserved

2. Employees are using consumer-grade file sharing without IT or business oversight.
3. The accidental mishandling of information and data happens every day.
4. Securing the perimeter and infrastructure does not ensure content security.
5. Regulatory issues around content security are real and evolving.

AI can solve these issues in Insurance and Health Sciences domain if the tools concentrate towards the issues mentioned and hence generating considerable increase in revenue and avoiding chronic problems.

4. Visual representations of the Data

Graphs, diagrams, drawings, sonographs, and x-rays are commonly used in contemporary science in the process of research, in communicating results, and depict phenomena that cannot be seen: structures too small to see with visible light (electron micro-graphs), relations among properties (graphs), steps in a mechanism (diagrams).Visual representations, like written or spoken sentences and numerical formulas, are external objects that function as symbols. Following are the representative graphs of Data of Insurance and Health Sciences.

Fig. 1 represents the data of an Insurance company which depicts the type of Insurance availed ina particular location. This will help the industry to analyse the type of claims and also help them regulate the policy effectively.

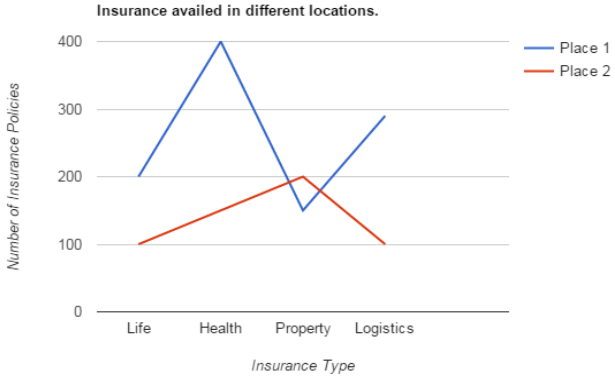


Fig.1 Insurance availed in different locations (Number of Insurance Policies vs. Insurance Type)

Fig.2 represents the types of claims that are seen in City-A. this data will help the insurance agencies to bring changes in the policies and implement pay for performance for motor insurances. Data representation should be simple so that every person analysing the representations understand the concepts easily.

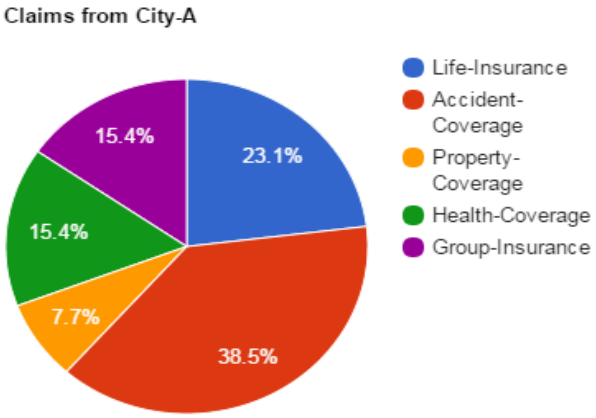
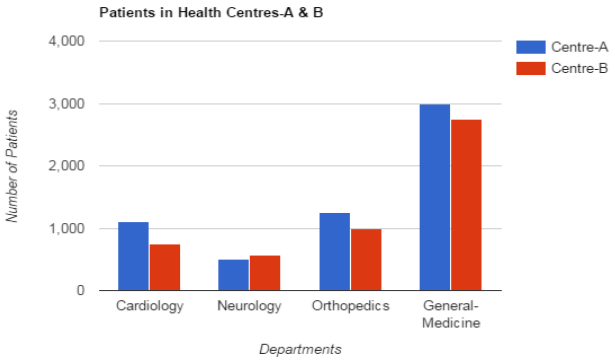


Fig.2 Claims of City-A

Fig. 3 represents the data collected in a health centre which depicts the number of patients admitted to the centre for various reasons. This data will help the insurance agencies to analyse their health policies in that region and also implement better policies for the applicants.



© 2015, IJARCSSE All Rights Reserved

III. ADVANTAGES OF AI IN INSURANCE AND HEALTH SCIENCE

Artificial Intelligence (AI) techniques can be used to design and build intelligent "agents" that can accomplish specific tasks efficiently. AIR is pioneering the application of AI techniques to solve portfolio optimization problems for the insurance industry using a branch of AI known as Reinforcement Learning (RL). RL methodologies are commonly used in the field of robotics, but they are also being adapted and applied to address large-scale and complex optimization problems. Data collaboration using AI is best suited to get results graphically and also easy understanding of issues and scenarios.

Few Technical Advantages are,

1. Improving Algorithms
2. Designing New Domain Modules
3. Modifiable systems for re-new projects
4. Increased Parallel Power, Increased Memory

IV. CONCLUSION

Reinforcement Learning techniques, by virtue of their ability to adapt to a stochastic environment, have the potential to advance the insurance portfolio optimization task by delivering superior solutions in the face of uncertainty. Tested against the commonly used traditional Algorithm in optimizing a book of policy groups, AI can prove to be a defining change in Insurance. Data Collaboration can enhance the understanding of data and also bring about a drastic change through solutions for chronic problems. Health Sciences too will be greatly advantaged by the solutions AI can provide through data collaboration. Ultimately, the people involved in delivering insurance services will be freed from mundane and error prone tasks as more intelligent systems are deployed. They will not be tied up in workaday search and retrieval of information tasks that may not be of the Highest quality, and be able to work with the correct information to more quickly complete their tasks.

REFERENCES

[1] Carnegie Mellon University-Using Artificial Intelligence to Improve Hospital Inpatient Care by Daniel B. Neill, H.J. Heinz III College.

[2] Artificial Intelligence Is Now Telling Doctors How to Treat You, byDANIELA Hernandez,Kaiser For Health News.

[3] Automatic Web Data Extraction Based On Genetic Algorithms And Regular Expressions By David F. Barrero And David Camacho And Maria D. R-Moreno.

[4] Enterprise Collaboration: Avoiding The Productivity And Control Trade-Of By Marcia Kaufman, Coo And Principal Analyst Daniel Kirsch, Senior Analyst.

© 2015, IJARCSSE All Rights Reserved