SMART HEALTH PREDICTION FOR AVOIDING FUTURE HEALTH RISK BY USING MACHINE LEARNING

[1]Prof. S. S. Ponde, [2]Ms. Swati V. Padghan

[1]Guide, [2]Student
Computer Science and Engineering (Master of Technology)
Deogiri Institute of Engineering and Management Studies Aurangabad, Maharashtra, India.

***Abstract*: Artificial intelligence (AI) aims to mimic human cognitive functions. It is bringing a paradigm shift to healthcare, powered by increasing availability of healthcare data and rapid progress of analytics techniques. We survey the current status of AI applications in healthcare and discuss its future. AI can be applied to various types of healthcare data (structured and unstructured). Popular AI techniques include machine learning methods for structured data, such as the classical support vector machine and neural network, and the modern deep learning, as well as natural language processing for unstructured data. Major disease areas that use AI tools include cancer, neurology and cardiology. We then review in more details the AI applications, in the major areas of early detection and diagnosis, as well as outcome prediction and prognosis evaluation.**

***Keywords:* Ai, Machine learning, Health care**

I. Introduction:

Numerous social insurance associations (doctor's facilities, medicinal focuses) in China are occupied in serving individuals with best-exertion social insurance benefit. These days, individuals give careful consideration on their physical conditions. They need higher quality and more customized medicinal services benefit. In any case, with the restriction of number of talented specialists and doctors, most medicinal services associations can't address the issue of open. Step by step instructions to give higher quality social insurance to more individuals with restricted labor turns into a key issue. The social insurance condition is for the most part seen as being 'data rich' yet 'information poor'. Doctor's facility data frameworks normally produce enormous measure of information which appears as numbers, content, outlines and pictures. There is a parcel of concealed data in this information immaculate. Information mining what's more, prescient investigation expects to uncover examples and principles by applying propelled information examination systems on a substantial arrangement of information for expressive and prescient purposes. Information mining is appropriate for preparing vast datasets from healing facility data framework and discovering relations among information highlights. It takes just a couple of scientists to examine information from doctor's facility data frameworks, and give enormous medicinal learning which can be used to help clinical basic leadership. Likewise, we could utilize information mining to give a self-benefit human services framework, which can serve bunches of individuals in the meantime. Oneself administration human services framework is of incredible hugeness to take care of the issue of unevenness between constrained therapeutic assets and requests.

II. Literature Survey:

1) Author Name- Neill DB, 2013 (Using artificial intelligence to improve hospital inpatient care. IEEE Intell System 2013;

Description-Electronic health records have become more available due to the guidelines of the Health
Information Technology for Economic and Clinical Health (HITECH) Act, which offers incentives to healthcare providers to adopt EHR to advance clinical processes and improve outcomes. Meanwhile, health insurance providers and non profits such as the Health Care Cost Institute 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.

Another new approach that might improve patient care focuses on statistical machine learning methods for
Detecting anomalous patterns in massive quantities of healthcare data. We recently developed a variety of machine learning methods based on *fast subset scanning* to detect patterns in massive datasets, efficiently identifying
Subsets of data records and attributes that are collectively anomalous or that maximize some measure of interest, such as a likelihood ratio statistic. In the patient care setting, our primary focus is to detect anomalous patterns of care that influence patient outcomes.

2) Author Name- Kolker E, Özdemir V, Kolker E (How Healthcare can refocus on its Super-Customers (Patients, $n=1$) and Customers (Doctors and Nurses) by Leveraging Lessons from Amazon, Uber, and Watson.

Description-The super-customers of Big Data and innovations, such as the ability to track one's physical activities and personal genomics, have been, by and large, patients and the healthy future patients. Contrary to that, the originally intended customers of Big Data doctors, nurses, and allied healthcare
Providers have lagged behind relative to the super customers. With patients as super-customers and sitting in the

driver seat as users, there are valuable lessons to be learned from innovative industries.

The lessons learned from successful industries and their businesses, such as Amazon, Apple, Boeing, IBM, Facebook, Google, Mercedes, Microsoft, Samsung, Starbucks, Toyota, and Uber, are instructive in that they share a common theme: consistent, year-after-year accommodation of the needs, pain points, and wants of their customers worldwide. Of all the industries, we think healthcare ought to be the most concerned with providing high-quality service to its super-customers as health is a common denominator in all countries and societies

3) Author Name- Dilsizian SE, Siegel EL.2014 (Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. Curr Cardiol Rep201 new and important scientific information cannot be processed and stored by a single human brain. Physicians learn thousands of different diseases in medical school and are expected to remember and apply a substantial subset of these in daily practice. But it is impossible for an individual physician to keep current on the broad spectrum of new data and discoveries and to reliably recall and utilize that information at all relevant time points. This is part of a major challenge in medical imaging, where real-time errors are estimated to average between 3 % and 5 % and constitute nearly 75 % of medical malpractice claims. Graber et al. estimated that approximately 75 % of diagnostic errors were related to "cognitive factors" Diagnostic errors outnumber other medical errors by 2- to 4-fold and represent nearly 40 % of total ambulatory malpractice claims.

Effective, evidence-based medical practice requires that physicians be familiar with the most recent guidelines and appropriate use criteria. Because of the exponentially growing amount of information in peer-review journals, textbooks, periodicals, consensus panels, and other sources, it is impossible for health care practitioners to keep up with more than a small fraction of relevant literature. Adherence to guidelines and evidence-based medicine may be made even more complex by the variability in "standards of practice" across different communities and states, a variability that complicates the concept of a "gold standard" for diagnosis and treatment of certain illnesses. Advanced computer systems, such as IBM's "Watson" technology, could assist by providing the most upto-date evidence-based information to inform proper patient care decisions. This information could combine data from a specific patient's history with data from large numbers of other patients with similar disease manifestations.

4) Author Name- Patel VL, Shortliffe EH, Stefanelli M, et al.2009 (The coming of age of artificial intelligence in medicine. Artif Intell Med 2009;

Description- The general AI research community was fascinated by the applications being developed in the medical world, noting that significant new AI methods were emerging as AIM researchers struggled with challenging biomedical problems. In fact, by 1978, the leading journal in the field (Artificial Intelligence, Elsevier, Amsterdam) had devoted a special issue solely to AIM research papers. Over the next decade, the community continued to grow, and with the formation of the American Association for Artificial Intelligence in 1980, a special subgroup on medical applications (AAAI-M) was created. It was against this background that Ted Shortliffe was asked to address the June 1991 conference of the organization that had become known as Artificial Intelligence in Medicine Europe (AIME), held in Maastricht.

The Netherlands. By that time the field was in the midst of ''AI winter'', although the introduction of personal computers and high-performance workstations was enabling new types of AIM research and new models for technology dissemination. In that talk, he attempted to look back on the progress of AI in medicine to date, and to anticipate the major challenges for the decade ahead.

5) Author Name- Karakülah G, Dicle O, Koşaner O, et al (Computer based extraction of phenoptypic features of human congenital anomalies from the digital literature with natural language processing techniques. Stud Health Technology Inform 2014; 205:570–4. [PubMed]Jha S, Topol EJ. Adapting to Artificial Intelligence: radiologists and pathologists as information specialists. JAMA 2016;

Description-Natural language processing is techniques have gained a wide use in obtaining and evaluating the information especially in life sciences due to the excessive increase of information in recent years. Besides, the automatized extraction of clinical information, such as signs, symptoms, medications and/or observations, from scientific documents and free-text formatted medical records using NLP should be considered crucial for the development of the CDSSs. Here we aimed at developing a computational strategy to extract the phenotypic features, which characterize HCAs from the case reports in the literature via text processing and NLP methods. In addition to this, by using the extracted information, we created an initial framework of an information base for a potential CDSS in the diagnosis of HCAs.

Natural Language Toolkit modules and the scripts written in Python were utilized in all text processing steps on the corpus. At the first step, the string. lowercase() function was used for standardizing the characters in the corpus. In the following preprocessing step, the sentence segments for each abstract were determined with Punkt algorithm. From these segmented sentences in each abstract, the white spaces were removed via tokenization in the third pre-processing steps.

III. Comparison:

Ref No.	Paper Name	Description	Advantage	Disadvantage
1	Using artificial intelligence to improve hospital inpatient care 2013	Overtreatment, 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	High availability to store electronics health care data using methodology and analysis technique	poor execution of care, and failure to adopt best practices
2	How Healthcare can refocus on its Super Customers (Patients, n=1) and Customers (Doctors and Nurses) by Leveraging Lessons from Amazon, Uber, and Watson.	The current bottlenecks in healthcare rest greatly in our ability to deeper understand and effectively communicate with its uber customers and customers within the broader social ecosystems surrounding	Ability to deeper understand and effectively communicate with its patient and patient within the broader social ecosystems surrounding	Poor implementation of the health care technique and high cost
3	Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment	The act of drug is at an intersection,with synchronous increments in Persistent volume, a blast in the sum and multifaceted nature of their apeutic and logical learning, and the change to electronic therapeutic records.	Increased demands on health care providers create greater risk for diagnostic and therapeutic errors.	Complex to handling big data and analysis database.
4	FThe coming of age of artificial intelligence VIII in medicine	The adolescence of AI in medicine: will the eld come of age in the 90s Artificial Intelligence in Medicine. In this article, the discussants react on medical AI research during the ub sequent years.	Ability to represent and utilize expert knowledge in symbolic form.	Complex design database difficult to understand implementation.
5	Computer based extraction of phenoptypic features of human congenital anomalies from the digital literature with natural language processing techniques	The mimicking of human cognition by computers was once a fable in science action but is becoming reality in medicine. The combination of big data and artificial intelligence, referred to by some as the fourth industrial revolution.	accurate diagnosis, great importance to provide the decision maker with decision support by presenting the literature knowledge	Based on natural language processing and limitation on dataset.

IV. Proposed System

Artificial intelligence (AI) aims to mimic human cognitive functions. It is bringing a paradigm shift to healthcare, powered by increasing availability of healthcare data and rapid progress of analytics techniques. We survey the current status of AI applications in healthcare and discuss its future. AI can be applied to various types of healthcare data (structured and unstructured). Popular AI techniques include machine learning methods for structured data, such as the classical support vector machine and neural network, and the modern deep learning, as well as natural language processing for unstructured data. Major disease areas that use AI tools

include cancer, neurology and cardiology. We then review in more details the AI applications, in the major areas of early detection and diagnosis, as well as outcome prediction and prognosis evaluation.

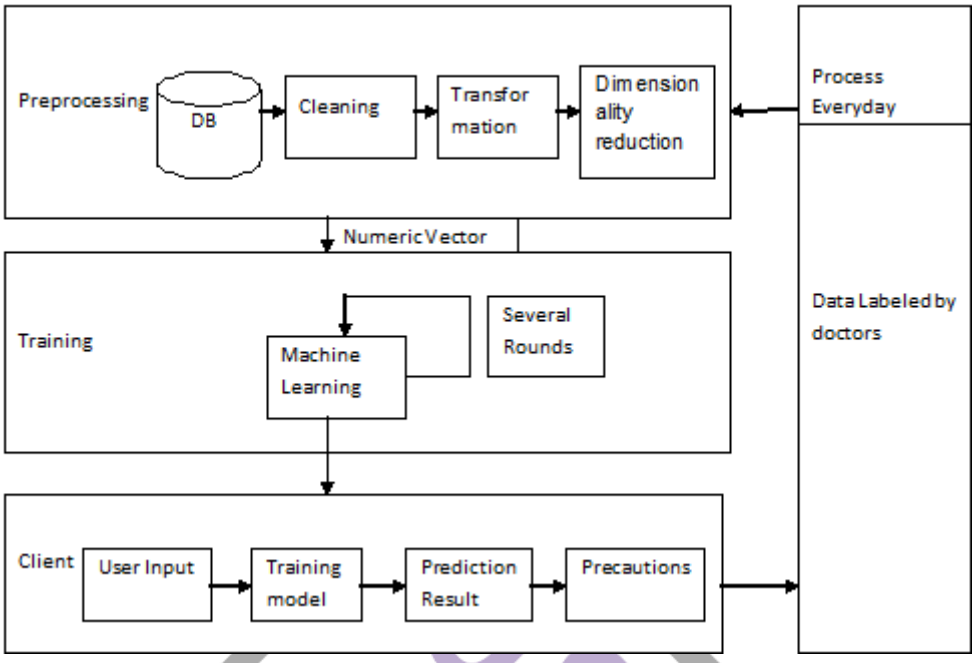

Figure 1 Proposed System

There are likewise numerous medicinal services administration applications in view of information mining. Information mining applications can be created to all the more likely recognize and track interminable infection states what's more, high-chance patients, outline proper mediations, and lessen the quantity of healing center armations and cases. An application anticipating the danger of in-healing center mortality in tumor patients with non terminal infection is created utilizing neural system. Another framework use neural system to foresee the aura in youngsters displaying to the crisis stay with bronchiolitis. Likewise, strategic relapse models are utilized to look at doctor's facility profiles in view of hazard balanced demise with 30 long periods of non-heart medical procedure.

Algorithm:

It is a classification technique based on Bayes Theorem with an assumption of independence among predictors. In easy terms, A naive bayes classifier assumes that the presence of a particular function in a class is unrelated to the presence of another feature. Naive Bayes Algorithm is a method that enables to construct classifiers. Classifiers are the models that classify the trouble instances and supply them class labels which are represented as vectors of predictors or feature values. It is based at the Bayes Theorem. It is referred to as naive Bayes because it assumes that the value of a feature is independent of the other feature i.e. Converting the value of a feature could no longer have an effect on the value of the other feature. It is likewise called as idiot Bayes because of the same cause. This algorithm works efficaciously for large data sets, as a result fine ideal for real-time predictions.

Naive Bayes classifier is the fast, accurate and reliable algorithm. Naive Bayes classifiers have high accuracy and speed on large datasets. Naive Bayes classifier assumes that the effect of a particular feature in a class is independent of other features. The name naive is utilized in light of the fact that it expect the highlights that go into the model is free of one another. That is changing the estimation of one component, does not legitimately impact or change the estimation of any of different highlights utilized in the calculation.

Naive Bayes algorithm Steps:

Step 1: Convert the data set into a frequency table

Step 2: Create Likelihood table by ending the probabilities

Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

Bayes Theorem is stated as:

$$P(h|d) = (P(d|h) * P(h)) / P(d)$$

Where,

P(h|d) is the probability of hypothesis h given the data d. This is called the posterior probability.
P(d|h) is the probability of data d given that the hypothesis h was true.
P(h) is the probability of hypothesis h being true (regardless of the data). This is called the prior probability of h.
P(d) is the probability of the data (regardless of the hypothesis).

You can see that we are inquisitive about calculating the posterior probability of P(h|d) from the prior probability p(h) with P(D) and P(d|h).
After calculating the posterior possibility for a number of one of a kind hypotheses, you can choose the hypothesis with the best possibility. This is the maximum likely hypothesis and can formally be called the most a posterior (MAP) hypothesis.

This can be written as:

$$MAP(h) = max(P(h|d)) \ or$$
$$MAP(h) = max((P(d|h) * P(h)) / P(d)) \ or$$
$$MAP(h) = max(P(d|h) * P(h))$$

The P(d) is a normalizing term which allows us to calculate the possibility. We can drop it when we're interested by the most probable hypothesis as it is consistent and simplest used to normalize.
Back to type, if we've got a fair wide variety of instances in every elegance in our training data, then the chance of every class (e.g. P(h)) might be same. Again, this would be a consistent time period in our equation and we should drop it so that we become with:

$$MAP(h) = max(P(d|h))$$

V. Result Analysis

GUI Home Page

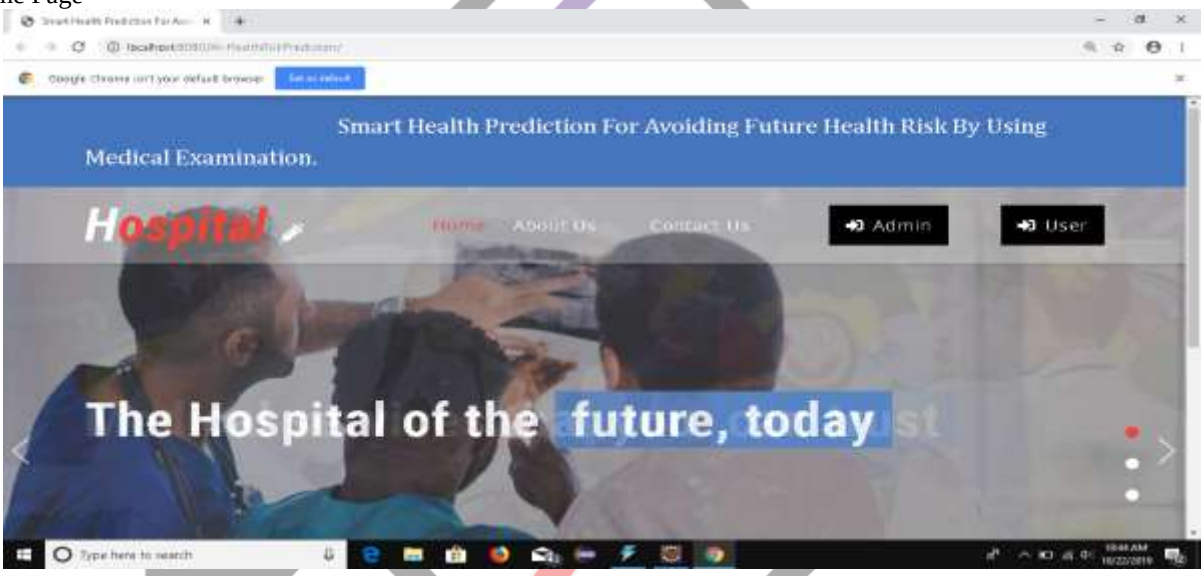

Figure 2 GUI Home Page

Login Page

Figure 3 Login Page

Symptoms Page

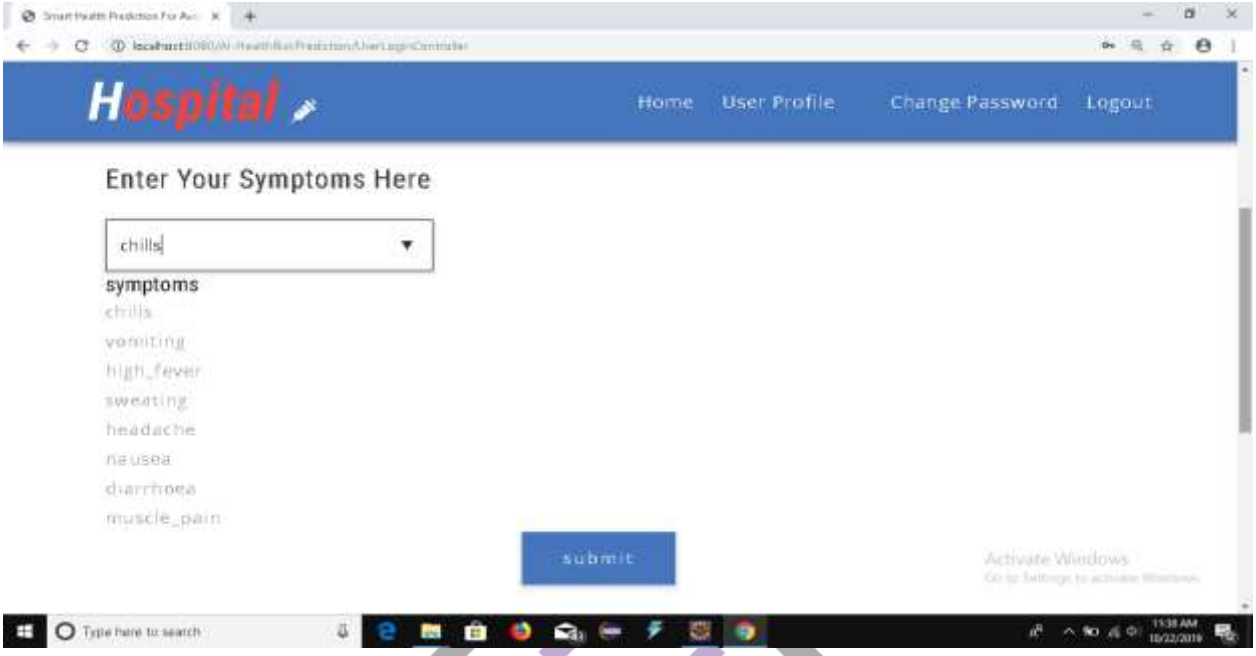

Figure 4 Symptoms Page

Disease Prediction

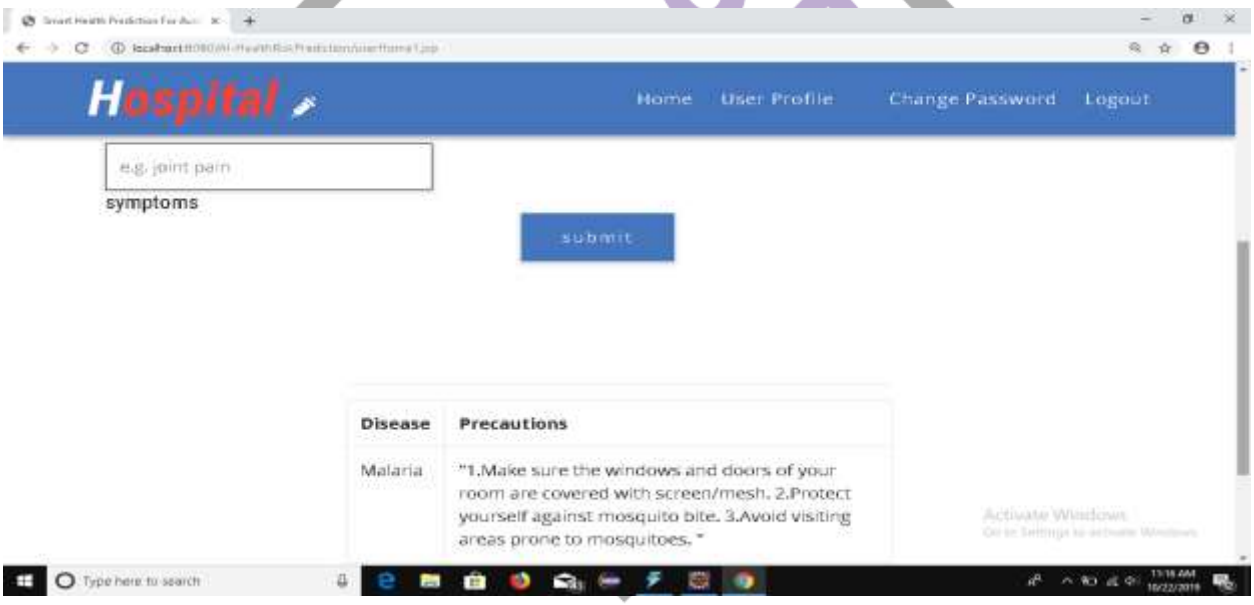

Figure 5 Disease Prediction

Precautions Prediction Page

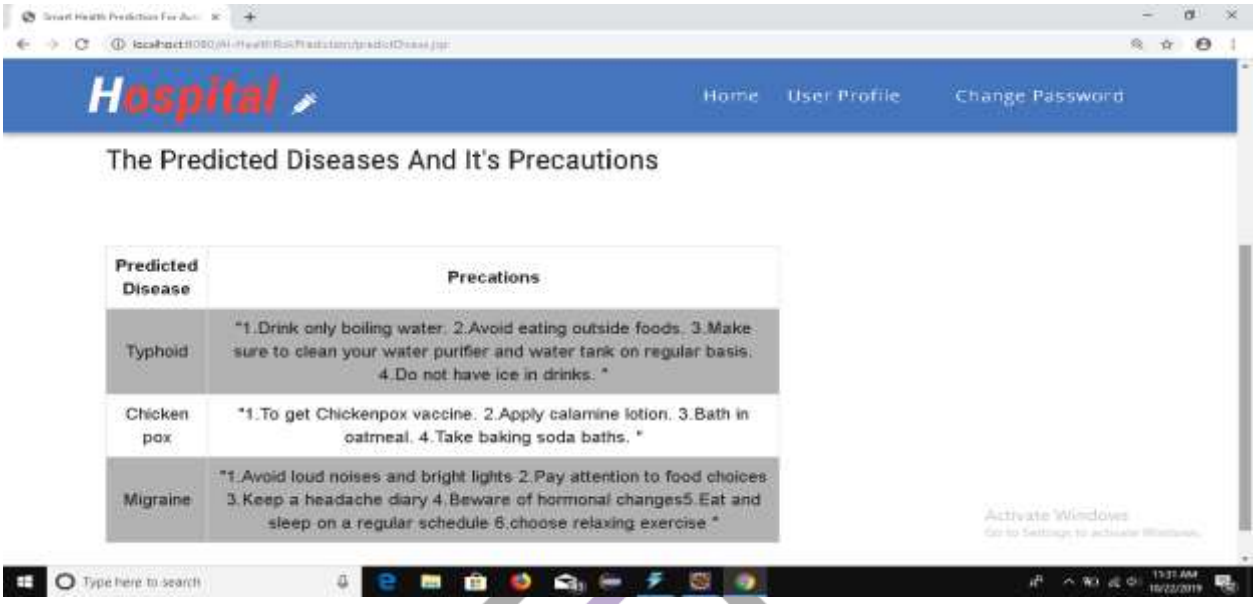

Figure 6 Precautions Prediction Page

Disease Accuracy Graph

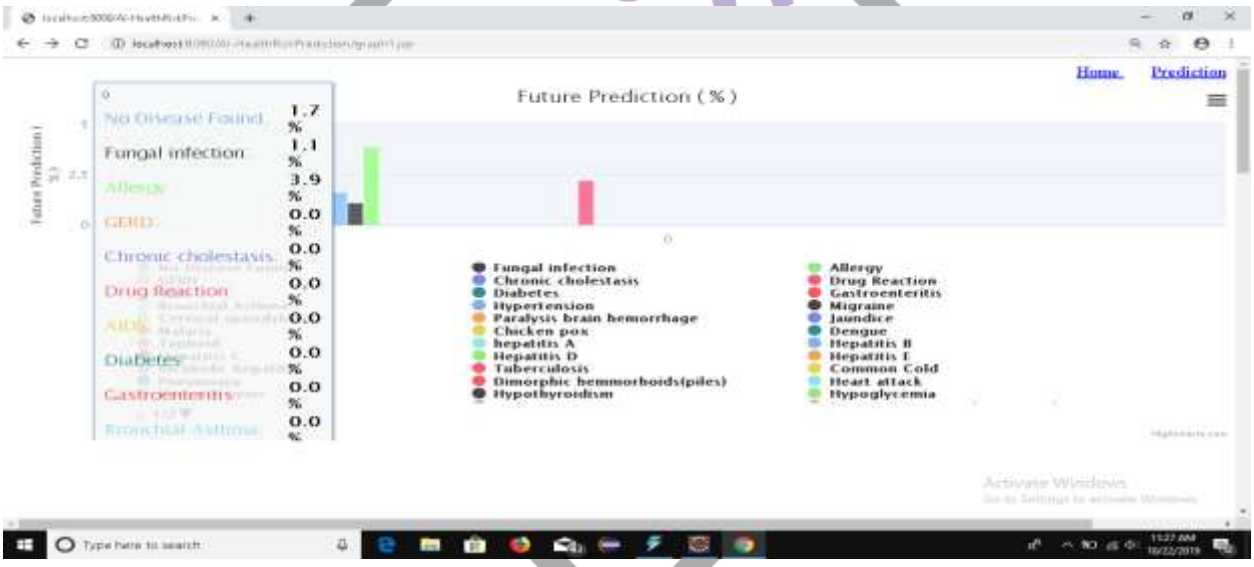

Figure 7 Disease Accuracy Graph

Confusion matrix based on aspect

	Predicted as aspect	**Not predicted as aspect**
Wrong aspect	TP	TN
Correct aspect	FP	FN

Confusion matrix based on dataset

	Positive	**Negative**
True	247	31
False	16	10

True Positive =247

False Positive =16

True Negative = 31

False Negative =10

Accuracy = (TP+TN) / (TP+FP+TN+FN)

Accuracy = 91.46%

Precision = (TP) / (TP+FP) = 94%

Recall = (TP) / (TP+FN) = 96%

Classifier	Accuracy
Random Forest	90.01%
Naive Bayes	89.06%

VI. Conclusion:

Reviewed the motivation of using AI in healthcare, presented the various healthcare data that AI has analyzed and surveyed the major disease types that AI has been deployed. We then discussed in details the two major categories of AI devices: ML and NLP. For ML, we focused on the two most popular classical techniques: RF and neural network, as well as the modern deep learning technique. We then surveyed the three major categories of AI applications.

A successful AI system must possess the ML component for handling structured data (images, EP data, genetic data) and the NLP component for mining unstructured texts. The sophisticated algorithms then need to be trained through healthcare data before the system can assist physicians with disease diagnosis and treatment suggestions.

References

[1] Neill DB. Using artificial intelligence to improve hospital inpatient care. IEEE Intell Syst 2013;
[2] Kolker E, Özdemir V, Kolker E. How Healthcare can refocus on its Super-Customers (Patients, *n* =1) and Customers (Doctors and Nurses) by Leveraging Lessons from Amazon, Uber, and Watson. OMICS2016;
[3] Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. Curr Cardiol
[4] Patel VL, Shortliffe EH, Stefanelli M, et al. The coming of age of artificial intelligence in medicine. Artif Intell Med 2009;
[5] Karakülah G, Dicle O, Koşaner O, et al. Computer based extraction of phenoptypic features of human congenital anomalies from the digital literature with natural language processing techniques. Stud Health Technol Inform 2014;205:570–4. [PubMed]Jha S, Topol EJ. Adapting to Artificial Intelligence: radiologists and pathologists as information special
[6] Srour M, Mazer B, Shevell MI. Analysis of clinical features predicting etiologic yield in the assessment of global developmental delay. Pediatrics. 2006;
[7] Demner-Fushman D, Chapman WW, McDonald CJ. What can natural language processing do for clinical decision support?. J Biomed Inform. 2009;
[8] National Library of Medicine. Medical Subject Headings [Internet]. (cited: 2013 November 23)
[9] Natural Language Toolkit [Internet]. (cited: 2013 November 28)
[10] Kiss T, Strunk J. Unsupervised Multilingual Sentence Boundary Detection. Comput Linguist. 2006;
[11] Fátima E. Sanz de León, Pamela Nicking Craig Rhodes† Prof. S. Yunkap Kwankam, Promoting appropriate ehealth technologies in the developing world: The sharing ehealth intellectual property for development (shipd) initiative of the world health organization.
[12] Mykowiecka A, Marciniak M, Kupść A, Rule-based information extraction from patients' clinical data, Journal of Biomedical Informatics, 2009.