# (Version:4)

CONFIDENTIAL

# HealthShop – A Machine Learning-driven Comparison Shopping Service

Introduction:

Healthcare costs have risen 65% above general CPI on an average every year (Source: 6/29/2015, Forbes). Private payers (Employer-sponsored) and public payers (Medicare, Medicaid) – in total accounting for 82% of covered lives - pay for the bulk of health spending in America.

The question is how can payers rein in runaway health care costs? One way is by addressing preventable significant “wasteful” spending. If we let patients to comparison shop for health services, patients are highly likely to self-select towards lower cost care solutions that already exist in the market today.

HealthShop has identified several wasteful spending targets. Let us illustrate with one example. If a patient gets surgery done at an Ambulatory Surgery Center (ASC) as compared to a hospital, the average cost savings is between $363–$1,000 (Source: May, 2015, U of Louisville, U of Minnesota study, Health Affairs). Yet, many patients are unaware of this savings accruing to themselves and their sponsors when evaluating care providers.

By making it easy for patients to comparison shop, we will save money for payers and patients. By applying our expertise in machine learning, healthcare domain knowledge, expertise in web technologies and successful track record in founding and running successful companies, we intend to build a great new business.

General Description of the Product:

Our product is a software-as-a-service website. Employers or public/private payers subscribe to this on a subscription-fee basis. They in turn allow free access to their members. Authenticated members shop for health services by first searching using free-text, voice dictation or by uploading any medical document that describes their medical condition. To illustrate by way of an example: a patient enters a search phrase, “I need some surgery to fix a retinal tear in my left eye”. HealthShop in real time (while the patient is still in the middle of typing the phrase) presents her with suggested search results. She then selects one result from the suggested list. HealthShop then returns search results as an ordered list containing the following information: Health condition name; description; provider team details (list of facilities, named physicians etc); cost range; patient portion of cost etc. The consumer at this point apply filters and sort criteria not dissimilar to ones we are used to seeing on an Amazon shopping page. The consumer then selects the top two results that suits them the most (e.g. the one with the highest quality rating but costing right in the middle of the lot and a second one with a medium quality rating but costing significantly less) to compare further. The consumer eventually narrows down to one pathway forward. At this point she is presented with a set of “next steps.” In one possible pathway, the patient is recommended to call the lead provider in order to begin the process of actually scheduling the first appointment.

We intend to start with two targeted savings objectives. Future plan is to include additional savings targets. Our starting solution will be an information shopping tool. In future releases, we will enable automated first appointment bookings and more.

Technical Approach:

The technical foundation for the HealthShop system comprises of a) a Care Bundle Recommendation System (CBR) and b) a Search Phrase Recommendation system (SPR) and c) a Web front end for authenticated patient access to the health shopping cart. We will cover CBR and SPR in more detail below.

Care Bundle Recommendation System (CBR):

The data-driven competitive shopping service involves data that reflect the interplay between primarily three major entities (the care “**receiver”,** most often the patient; the care “**provider”,** most often physicians/hospitals and the “**payer**” of the services i.e. insurance/government/employer/patient).

The utility functions for these three actors are divergent.

Receiver: Wants to maximize successful outcome, lower “out of pocket” spend and achieve convenience goals.

Provider: Wants to deliver the best care to the receiver while optimizing (maximized revenue minus minimized expense) profits against available resources.

Payer: Wants to minimize total spend by reducing wasteful expenditure while delivering the minimum set of health services that meets regulatory requirements and member (citizen or employee) health needs.

To be able to derive insights and drive user experiences, we need to model the dynamic interaction between receivers, providers and payers. The dynamism is captured in the different data assets that get generated over time. Let us illustrate some of the data that we plan to assimilate or leverage from our existing assets.

1. Receiver: A receiver has demographics data, primary location, medical history, out of pocket expense history and quality/satisfaction goals. Most of this data can be readily acquired from payer’s existing systems with the exception of satisfaction data. This data be inferred initially and instrumented subsequently by us.
2. Provider: Each provider would have a set of physicians with their bonafides captured as verbose text. Additionally, one can have receiver-based reviews. The provider has attributes such as location, parking availability, information on nearby rehab facilities etc. Provider services catalog such as in-patient, outpatient services, laboratory, radiology services etc. Each of these are conditioned on the disease, diagnoses and the associated procedural codes. Geographical, demographical and performance information about providers. This information includes details of provider organizations as well as individual providers. Details include names, accreditations, certifications, skills, experience, quality scores etc.
3. Payer: The different insurance companies can be evaluated quantifiably by the different plans they offer and the geographies they cater to. One can build a platform for receivers and providers both providing feedback. A derived quality might be available based on how long a receiver has been with the same payer as well as industry ratings such as “star” ratings.

In our informed judgment, the above scenario is a great match for applying neural network techniques applicable to complex dynamic multi-layer complex network (MLCN). MLCNs have been modelled in the context of Event Based Social Networks (EBSNs) in the last two years. HealthShop’s data scientist, Sayan Pathak is a co-author of a recently published article, “Deep Learning driven Venue Recommender for Event-based Social Networks” in a peer-reviewed journal “Transactions on Knowledge and Data Engineering”. Directly leveraging the MLCN work done in that paper, we have developed an approach using deep learning techniques that allow us to leverage diverse range of information in a common framework. One of the key challenge is dealing with data sparsity. We have developed an approach using deep transfer learning technique to fill-in the gaps by leveraging different complementary data corpus. We have further developed a technique that uses recurrent deep learning networks to model temporal and structural dynamics in the graph. Temporal dynamics involves interaction between entities over time and structural dynamics involve the association between actors changing over time e.g., a patient choosing a different physician for taking care of their health or even a certain health condition or a payer dropping a certain provider or adding a new provider.

The aforementioned techniques have shown dramatic improvement in our ability to do predictive modeling and recommendation.

CBR System Summary: Given a patient’s profile and disease condition, CBR will provide range of costs, locations and different rating attributes. The API will allow the slice and dice of this information similar by applying filters, sorts etc. This will be leveraged the front-end web application.

Search Phrase Recommendation system (SPR):

The SPR system will be built with data that we acquire from expert physicians, covered procedure documentation from providers and payers and end users. The search phrase recommendation system will leverage well established text search technologies from the web domain. These including deep neural nets for embedding phrases into a common latent space where semantically similar terms (query from the user and the phrases in our document corpus) will have very small distance while mismatches shall have a large distance. Using the distance between query and document we shall rank relevant documents.

Another key technology that is gaining prominence is question-answering system. Using this technology, we shall enable the user to express their query in colloquial spoken English (our first language) and the system shall generate a natural language response. This is key for HealthShop as a user may not be familiar with the right medical terminology but is likely to have a more lay person's expression.

Combining these two technologies with the deep recommender we described under the CBR heading we believe we have a comprehensive solution to a problem that is in great need in the healthcare market.

Team:

Our team is comprised of two core people.

Prem Urali, a successful entrepreneur who has founded two successful companies: Commercia founded in 1999 and acquired by Microsoft in 2000 and HealthUnity founded in 2004 and acquired by ZeOmega in 2015. In both cases, he was the CEO from founding till successful exits. He is also our chief product and business architect. He holds an MS in Computer Engineering from Iowa State University and an MBA from The Wharton School.

Our cofounder is Sayan Pathak, a highly accomplished machine learning expert. He has been doing machine learning in the field of health care vision application, neuroscience, on-line advertising and speech recognition in the industry for the past 20 years. He is a member of faculty at the Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur and at the Department of Bioengineering, University of Washington, Seattle. He has been a Principal Investigator on several National Institutes of Health Small Business Innovation Research grants and commercialized FDA-approved machine learning based technology in the field of Radiology. Sayan has several peer-reviewed publications, awarded multiple times and issued and pending patents. He has been a presenter at various research, academic and industry conferences. He teaches an online course on Machine Learning on EdX.