Improving Quality of Care leveraging Advanced Analytics

CMC ltd has partnered with UC Irvine Health to implement Big Data solutions using Hadoop technology to implement the Big Data platform (Saritor). CMC implemented advanced machine learning algorithms to recognize implicit patterns in large EMR datasets that included structured and unstructured data to build models and predictive analytics capabilities to address specific use cases of interest on the platform. One such model and algorithm was to predict the 30-day readmission of Patients at the associated hospitals augmented with Patient Scorecards and Advanced Visualization.

Technology Stack

HortonWorks Hadoop Ecosystem on Redhat Linux that includes HBase for Big Data, MongoDB for indexing data files, R / Weka as a modeling Workbench, Apache Mahout for scale out Data mining, Tableau for Data Visualization, Solr and Lucene for large content indexing capabilities, Spring ecosystem for Web application development.

Scope of Work

- Building Saritor as the central enterprise data store and refinery leveraging Big Data technologies that involved consolidation of 20+ years of Patient EHR data from mainframe legacy systems
- Enable Clinicians and Researchers to explore and navigate the large datasets using a Google like NLP enabled search bar that would bring in the matching results instantly using advanced visualization techniques
- Enhancing the big data store using continuous streaming feeds from current EHR, HL7 datasets
- Implement unsupervised and supervised machine-learning algorithms
- Integrate information from various medical devices to capture data in realtime

Results

- The legacy data of 1.2 million patients, contained in 9 million patient medical EHR records was successfully ingested into the Saritor Hadoop Distributed File System. Savings include $500K annually to UC Irvine Health.
- For clinicians in patient care, complete patient records were retrieved via a web browser.
- For researchers, the drag and drop query and visualization tool allowed for the visualization of the deidentified patient datasets
- Algorithms for predicting 30 day readmits were built into Mahout for real time monitoring surveillance aided with Patient Scorecards

Currently we are working with UCI to implement other use cases such as Home Monitoring, In-Hospital Monitoring, Real-time Sepsis Detection and many others to improve the Quality of Healthcare.