A standards based prescription for healthcare analytics

Joydeep Das - COO, KloudData, Inc.
August 16, 2016
Agenda

• Healthcare Industry Demands

• KloudKare® Healthcare Information Framework

• Use Cases - Scenarios / Challenges / Models / Lessons

• Prescription for Architecture / Process / Technology / People
o High growth, **business software solutions** company

o ‘**Trusted Partner**’ to customers and alliances

o **Customers:** 100+ in USA, South Asia; 15+ university/county/private hospitals

o **Industries:** Healthcare / Manufacturing / Public Sector

o **Domains:** Enterprise / Analytics / Mobility / UI-UX / Cloud

o **Locations:** Fremont, CA / Maryland, MD / India / Thailand

o **Team Strength:** 385+
Healthcare Industry Demands

**Mergers/Consolidations** – larger Accountable Care Organizations / Payer (+) Provider Networks

**Pay For Performance** – financial rewards for achieving/exceeding quality and cost measures

**Population Health Management** – improve outcomes of a group by analyzing at-risk patients

**Continuous Care** – track patient health status over time and share information b/w stakeholders

**Resource Optimization** – improve costs and utilization across high value human and physical assets
KloudKare® – Healthcare Information Framework

Standards Based Solution Framework
For Healthcare Providers
Use Case Scenarios – Advanced Analytics Focus Areas

- Higher bed and caregiver utilization based on length of stay (LOS)
- Optimal use of key physical assets based on real-time traceability as a service (TaaS)
- Better population health management based on patient risk stratification
- Improved value based care based on patient engagement
Use Case – Higher resource utilization based on LOS

- Higher bed and caregiver utilization based on length of stay (LOS)
- Optimal use of key physical assets based on real-time traceability as a service (TaaS)
- Better population health management based on patient risk stratification
- Improved value-based care based on patient engagement
LOS - Why Analyze And Predict?

Over staffing puts a hospital's financial viability at risk

Under staffing has an adverse impact on patients and staff

Resource optimization is a hard problem with significant impact on pay for performance (costs/quality)

Extremely long waiting hours at the Emergency

Avoid over-crowding and turning away patients in need for care
LOS Prediction – Challenges Faced

- **Complex Datasets** – Multiple data sources / representations
- **Variables** – Numerous under consideration
- **Multitude Of Tools** – Varying model building tools preferences of data scientists
- **Model Evolution** – New variables e.g. disease codes introduced periodically
- **Large Data Sets** – Consolidated and Historical
LOS Prediction – Solution Selection

Decision Tree Based Model:
- Decent model for complex decision making
- Allows automatic feature selection and variable selection
- Easy to interpret results enabling faster impact analysis
  - Parametric non-linear relations don’t affect accuracy / performance
  - Produces good models even with non normalized / sparse inputs
- Deployable as a robust / scalable / standardized format

Standards Based Processing (PMML) :
- Creation from multiple modeling tools
- In-Database processing for speed accuracy
- JDBC access from Visualization Tools
- Flexible model versioning and evolution

Big Data / MPP :
- Handles large data sets quickly and cost effectively
- Integration from variety of data sources / types (including multiple EMRs)
 ✓ Dynamic parameterization for ad-hoc analysis
 ✓ In-Database (Hadoop) Scoring
   • Add-Ons:
     • Readmission Predictions (Work in Progress)
     • Resource Recommendations (Work In Progress)
Use Case Scenarios – Population Health Management (PHM)

- Higher bed and caregiver utilization based on length of stay
- Optimal use of key physical assets based on real-time traceability as a service (TaaS)
- Better population health management based on patient risk stratification
- Improved value-based care based on patient engagement
PHM - Why Analyze and Predict?

Frequency/Scale of Health Crisis in the Community

Population Health impacts delivery of care to a group of individuals with similar healthcare needs

Healthcare cost per service

Patient Self-Management

Patient Experience
PHM Custom Project – Infant Mortality (Midwest State)

Challenges
• Significantly lagged other US states and developed countries in infant mortality rates
• Complex data integration required from many data sources – state, federal, local
• Many variables for model building - environmental pollutants, maternal health, crime, socio-economic, ...

Goals
• Reduce the mortality rates to 7.4 deaths per 1,000 births by 2015, and 6.89 in 2016
• Create one-time report for Governor’s office that will allow funding appropriate programs to achieve results

Implementation
• Highly flexible master data to cleanse / fuse data sets intelligently with automation for speed / quality
  • R + Python based solution to build data integration and data set for model
  • Graph and Regression Models to Predict Mortality Rates and Study relations between variables

Lessons
• Research project – we got away with multiple tools (R, Python, etc)
  • Two data scientists / One-time project / Non-production data
• Operational Population Health Management Analytics (Scoring Patient Risks) will demand standards and scale
Use Case Scenarios – TaaS and Patient Engagement

Higher bed and caregiver utilization based on length of stay

Optimal use of key physical assets based on real-time traceability as a service (TaaS)

Better population health management based on patient risk stratification

Improved value based care based on patient engagement
Summary View – TaaS and Patient Engagement Analytics

- **Challenges:** large/high velocity (RFID) data from many assets
- **Models:** survival models (correlations, regressions)
- **Operational:** periodic changes/evolutions
- **Lessons:** Plan for standards / scale / velocity / new data

- **Challenges:** multi-source data including from patients (mobile)
- **Models:** nearest neighbor models to assess compliance
- **Operational:** interactive recommendations / alerts
- **Lessons:** Plan for standards / data bursts / variety / new data
Prescription - Architecture

Source Tier

Processing Tier
- Model Building Tools
- Web Services
- MPP Platforms
- PMML

Visualization Tier
- Visualization Tools
- BI Metadata Layer

ERP / EMR / IoT / Big Data (Standards Based Healthcare Information Framework)
Prescription – Technology / Process / People

Subject Matter Expert

- Identify Use Case
- Evaluation of Results
- Model Validation
- Variable Selection

Ensemble Models & Self Learning Algorithms

- Data Exploration
- Data Preparation
- Model Building
- Data Transformation

Taxonomies & Rule-based Methods

- Unsupervised Learning & Probabilistic Reasoning

Subject Matter Expert

Ensemble Models & Self Learning Algorithms

Model Building

Data Transformation

Variable Selection

Copyright © 2016, KloudData Inc. All Rights Reserved.
THANK YOU