Machine Learning in Healthcare

ND HIMSS Spring 2017 Conference

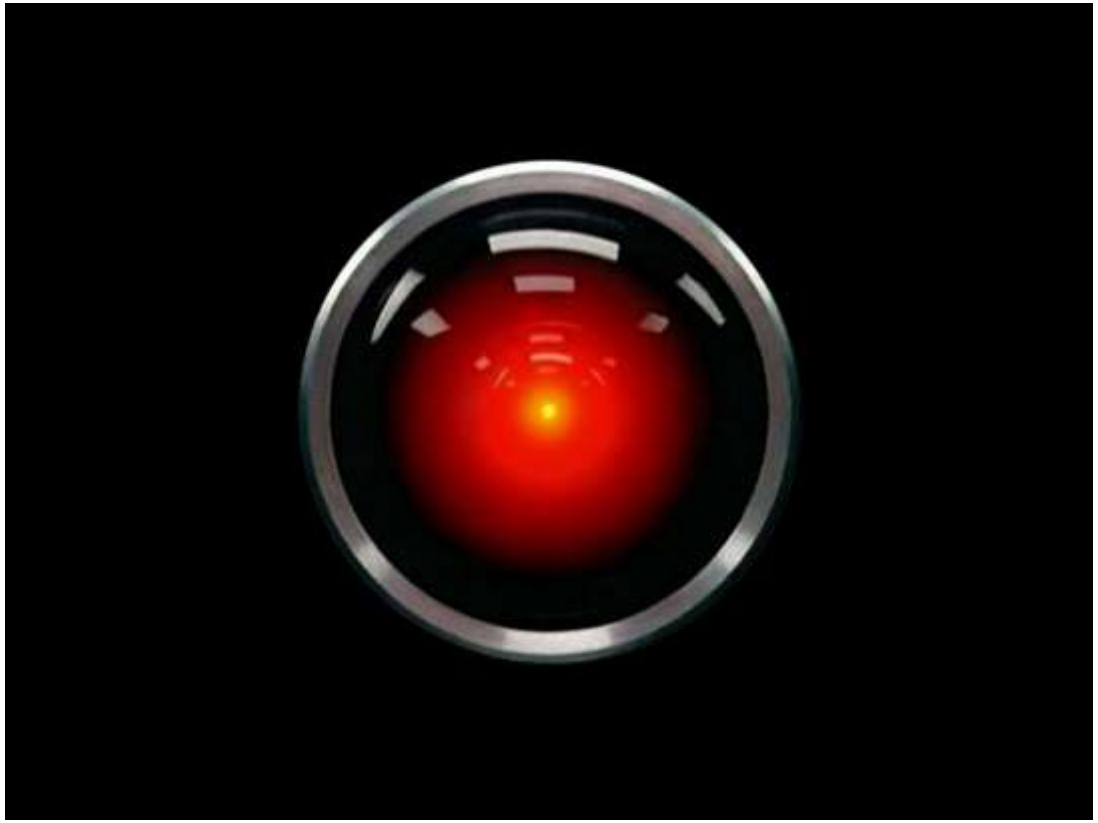
Fargo, ND

April 12, 2017

HealthCatalyst
ignite outcomes improvement

AI Quiz

AI Quiz



AI Quiz

Objectives

Learn some buzzwords

Why Bother?

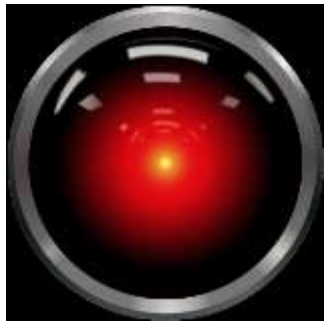
How to build a predictive model

Examine real-world predictive models

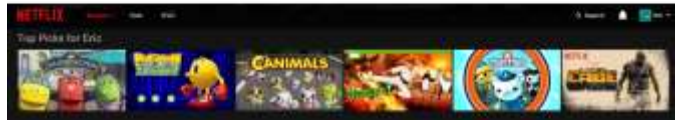
Getting Buy-In from Clinicians

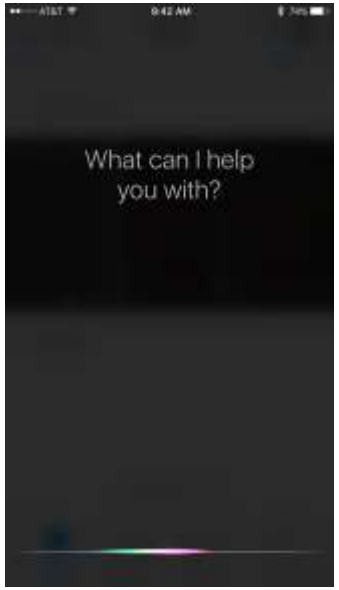
AI: Artificial Intelligence

General Artificial Intelligence



"Narrow" Artificial Intelligence





Machine Learning

Machine learning explores the study and construction of algorithms that can learn from and make predictions on data.

https://en.wikipedia.org/wiki/Machine_learning

Predictive analytics, or making predictions based on past data, is one of the artificial intelligence tasks that machine learning can solve.

I'm still confused…

Artificial Intelligence tries to replicate the capabilities of the human mind.

Machine Learning uses complex math to solve difficult problems.

Predictive Analytics, from the standpoint of healthcare or business, is one of the most important activities that is enabled by Machine Learning.

Predictive Models and **Risk Models** are the products of Predictive Analytics.

Why bother?

Classic Approaches

Mortality prediction

The Charlson Index was introduced in 1987 in the Journal of Chronic Disease as mortality risk score.

Readmission prediction

The LACE Index was introduced in the Canadian Medical Association Journal in 2010 to predict early death or unplanned readmission after discharge.

Shortcomings…

Using the LACE index to predict hospital readmissions in congestive heart failure patients

By Wang et. al, BMC Cardiovascular Disorders , 2014

CONCLUSION: The LACE Index may not accurately predict unplanned readmissions within 30 days from hospital discharge in CHF patients. The LACE high risk index may have utility as a screening tool to predict high risk ED revisits after hospital discharge.

Predicting readmissions: poor performance of the LACE index in an older UK population

By Cotter et al., Age Aging , 2012

CONCLUSION: The LACE Index is a poor tool for predicting 30-day readmission in older UK inpatients. the absence of a simple predictive model may limit the benefit of readmission avoidance strategies.

Limitations

Most standard models are trained with data from a broad, general population.

Most standard models are based upon data elements that are available through billing or claims data.

Advantages of building models

Trained on data from your environment.

Trained on data from your patients.

Answers your specific questions.

When should I build a model?

Trying to differentiate outcomes for complex cohorts

Predict infrequent events

Prioritize attention of limited resources to very frequent events

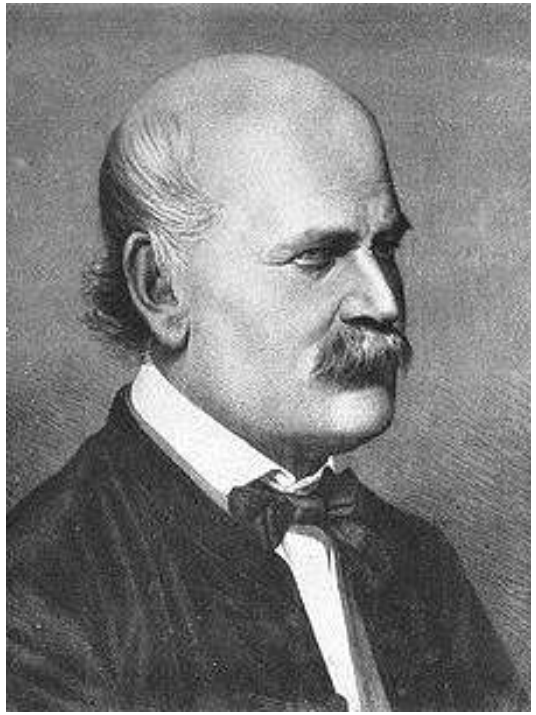
Predict outcomes as the result of modified behaviors

Incorporate features unlikely to be available to “standard” models

- Socio-economic data
- Geo-location data

Let’s Try It

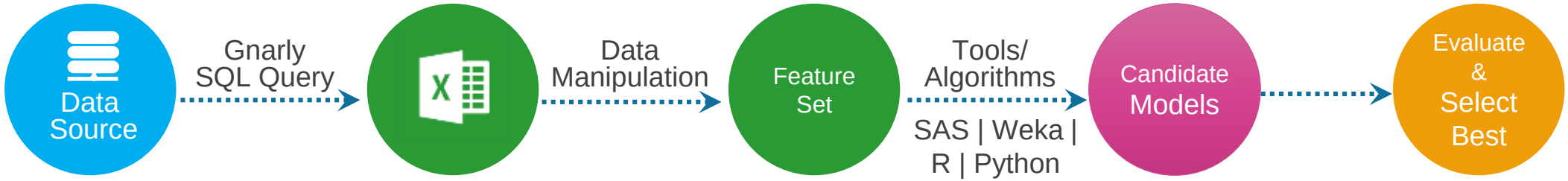
Let's Build a Predictive Model



Steps to build a model

1. Determine event of interest.
2. Determine our population.
3. Decide upon “features.”
4. Build feature sets.
5. Run through various algorithms: Train and Test.
6. Select the best model.

Typical Workflow for Building a Predictive Model



Features

Delivery Date	Delivery Location	Humour Temperament	Blood Letting	Physician Type	Hand Washing	Died
1/1/1844	Clinic 1	Sanguine	Yes	Physician	Yes	No
1/1/1844	Clinic 1	Melancholy	No	Physician	No	Yes
1/1/1844	Clinic 1	Balanced	No	Physician	No	No
1/1/1844	Clinic 2	Choleric	No	Midwife	Yes	No
1/1/1844	Clinic 2	Phlegmatic	No	Midwife	Yes	No

Training and Testing

Most records will be used to “train” or create the models.

The remaining records will be used to test, or determine the accuracy, of each model.

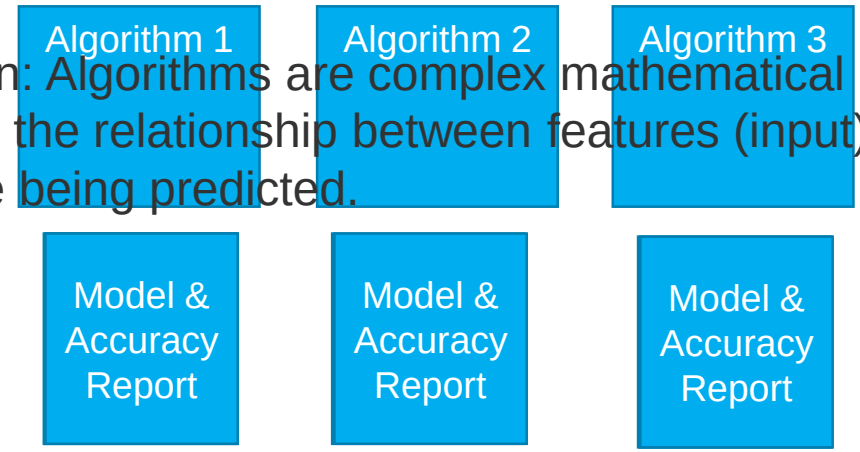
Developing a Predictive Model

Features (i.e. age, comorbidities, polypharmacy)

Definition: Simply put, a feature is an input to a machine learning model

Algorithms (i.e. Lasso, Random Forest, k-means)

Definition: Algorithms are complex mathematical processes that discover the relationship between features (input) and the outcome being predicted.



Result:

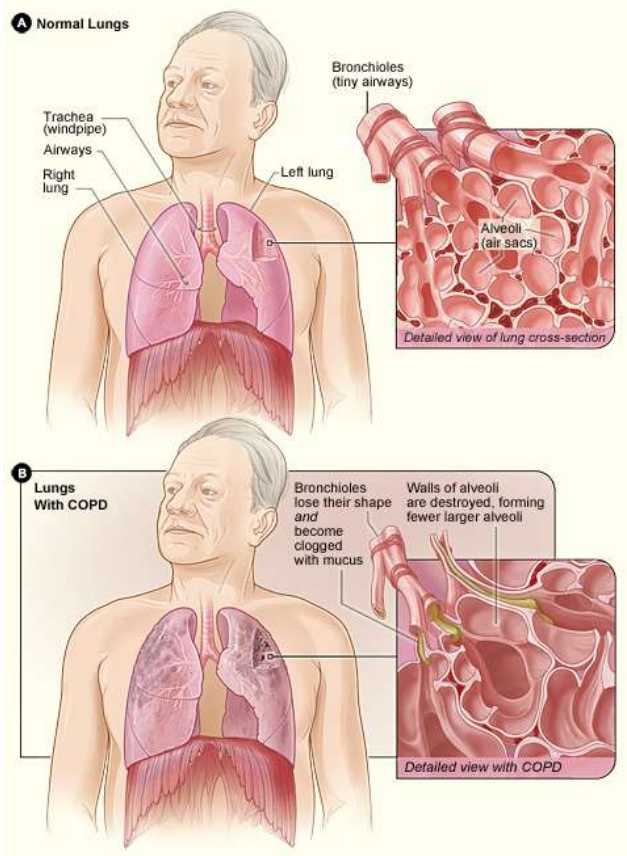
- Handful of best (most predictive) features
- Best algorithm that computes the relationships between input features to generate prediction
- Performance report summarizing best ‘model’

Dr. Semmelweis's Model

When Delivery Location = Clinic 1 and Hand Washing = No, women are 3 times more likely to die. Humours are not predictive, and blood letting correlates slightly with death.

Real World Models

Real World Use Case: COPD Readmissions



From nih.gov

COPD Readmission Challenge

Can we develop a model to help Pulmonary Navigators identify which COPD patients are most likely to experience an exacerbation that would lead to a readmission?

COPD Model Example

Total number of respiratory disease index admissions: 90,312

Total number of features: 29

Final number of features used: 19

COPD Model Example

Variable importance

Health Catalyst used the random forest's Gini impurity index along with area under the ROC curve (AUROC) maximization to determine which variables to incorporate into the final model. By removing variables below 0.1 on this index that did not have an impact on the AUROC (once removed), fourteen input variables were shown to account for the most significant impact on readmission prediction. The final variables are detailed below.

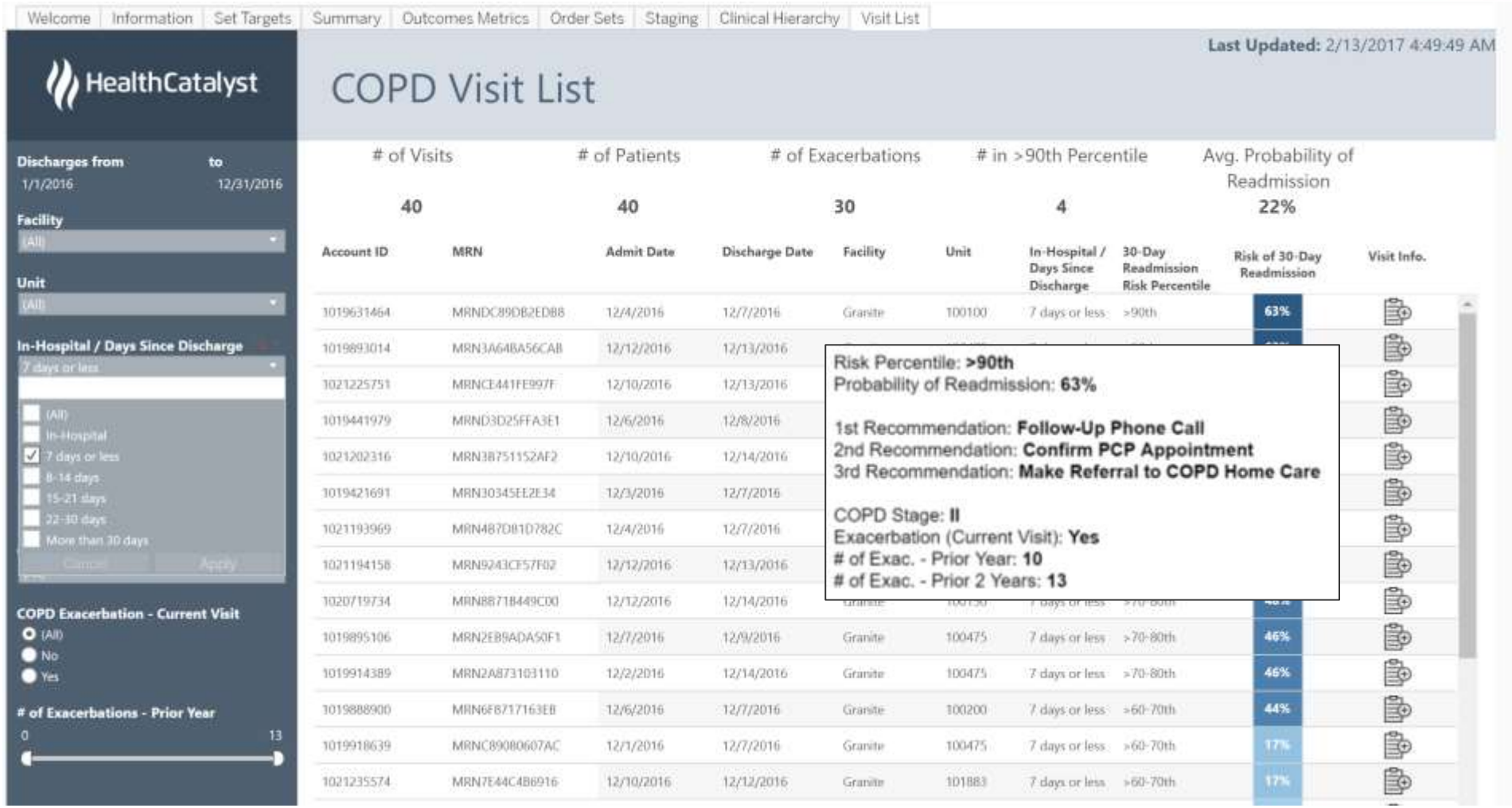
Variables considered

- ~~Acute exacerbation~~
- Admitted through ED
- Age
- Arrive Date
- ~~Barriers to medication~~
- BMI
- COPD stage
- County
- Facility
- Financial class
- ~~Gender~~
- History of dementia
- History of depression
- ~~History of diabetes~~
- History of heart failure
- ~~History of psychoses~~
- ~~History of renal disease~~
- ~~History of vascular disease~~
- Living arrangement
- Palliative care
- ~~Passed through ICU~~
- Powerplan utilization
- ~~Primary care physician~~
- Prior ED encounters
- Prior inpatient admissions
- ~~Resp. discharge medications~~
- Self-reported disability
- Smoking status
- Walking limitation

**Crossed-out variables were considered but didn't increase the model's accuracy.*

HealthCatalyst
ignite outcomes improvement

COPD Readmissions



Note: Data is from de-identified data set and in some places fabricated in order to show a reasonable representation of actual trends and observations from production data. All names, addresses, and other PHI are fabricated.

Likelihood of No Shows

Defining the Model

The question	data used
Which patients are likely to no-show to their scheduled appointments?	Total number of appointments: ~10 million Total number of input variables: 30 Final number of input variables used: 14

Variable importance

Health Catalyst used the random forest Gini impurity index to determine which variables to incorporate into the final model. By removing those below 0.1 on this index, fourteen input variables were shown to account for the most significant impact on no-show prediction. The final variables are detailed below:

Variables considered

- ~~Appointment Department Region~~
- Appointment Duration
- Appointment Time
- ~~Appointment Type~~
- ~~Benefit Plan~~
- ~~Cost Center~~
- Day of Week
- Days from Scheduled to Appointment
- Days to Next Holiday
- ~~Department~~
- ~~Distance in Miles (patient zip to location zip)~~
- Ethnic Group
- Financial Class
- ~~Gender~~
- ~~Language~~
- ~~Location ID~~
- Month
- Operating Hours Group
- Patient Age
- ~~Payor~~
- ~~Primary Financial Class~~
- ~~Provider ID~~
- ~~Provider Specialty~~
- ~~Provider Type~~
- Race
- ~~Recency~~
- Running Cancellations
- Running No Shows
- ~~Service Line~~
- Week of Year

**Crossed-out variables were considered but didn't increase the model's accuracy.*

Choice of model

A Random Forest model was used to calculate the relative impact of the above variables in respect to the appointment status of no-show. The model was created using 200 trees. The performance was as follows:

- Model performance: AUROC (c-statistic): 0.88
- Example cut point of .065: True-positive rate (Sensitivity): 0.800; False-positive rate (1-Sensitivity): 0.218

Deploying the Model in the Patient Access Application

This model has been deployed into an output table named SAM.PracticeManagement.NoShowOutput using the above logic. A probability score for each appointment is calculated and appended to an output table each time on the same schedule as the Patient Access subject area mart (SAM) refresh. This output table joins to the Patient Access SAM using AppointmentID.

References

Alaeddini A. *Probabilistic Models for Patient Scheduling.* [master's thesis]. Detroit, MI: Wayne State University; 2011.

Huang Y, Hanauer DA. Patient no-show predictive model development using multiple data sources for an effective overbooking approach. *Appl Clin Inform.* 2014;5(3):836-60.

Likelihood of No Shows

CLABSI

The question	data used
For patients with a central line, what is their risk of CLABSI over the course of the encounter?	Total number of central-line cases: 71,019 Total number of input variables: 23 Final number of input variables used: 16

Variable importance

Two-sample t-tests of input variables were used against the CLABSI result label to determine which variables should be included in the final model. Sixteen input variables accounted for the most significant impact on CLABSI prediction. Including additional input variables beyond these sixteen did not materially improve the model accuracy for this data set. The final variables are detailed below.

Variables considered

AgeInDays	HistoryCLABSI	LineDays	LineDaysPort
~~AlteplaseAdministered~~*	~~HistoryHIV~~	LineDaysFemoral	LineDaysTotal
CHGBathingNonCompliant	~~HistoryImmunodeficiency~~	LineDaysInternalJugular	LineDaysTunneled
DaysBeforePlacement	~~HistoryLeukemia~~	~~LineDaysMultiLumen~~	ParenteralNutrition
DaysSinceAdmit	~~HistoryLymphoma~~	LineDaysNonTunneled	RoutineBathingNonCompliant
Gender	~~HistoryNeutropenia~~	LineDaysPICC	

**Crossed-out variables were considered but didn't increase the model's accuracy.*

Choice of model

A Random Forest model was used to calculate the relative impact of the above variables in respect to the labeled outcome of patient infection. The model was created using the Gini criterion and 2,000 trees. The performance was as follows:

- Model performance: AUROC (c-statistic): 0.870
- Example cut point: True-positive rate (Sensitivity): 0.816; False-positive rate (1-Sensitivity): 0.169

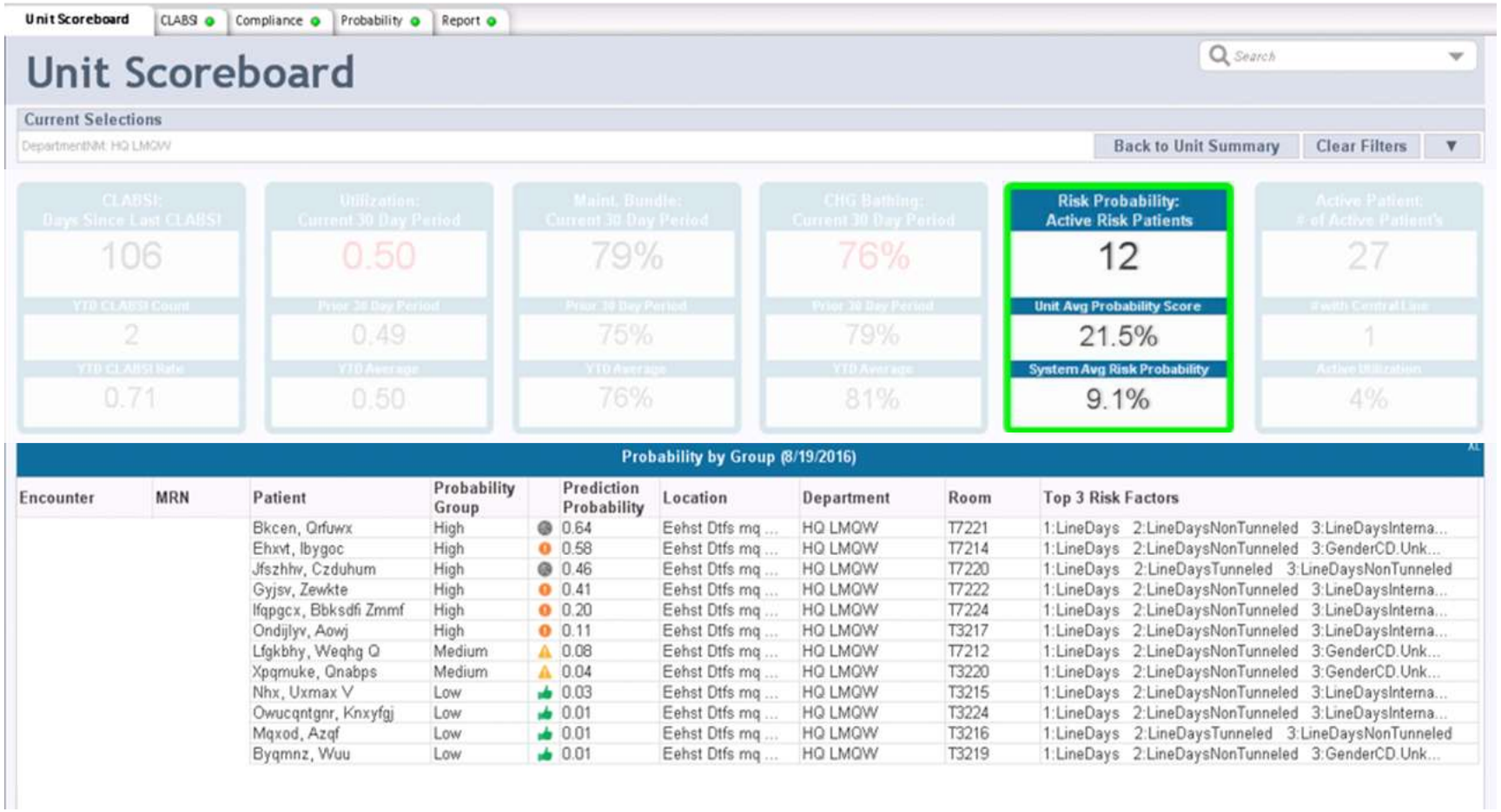
Deploying the Model in the CLABSI Application

This model has been deployed directly into the CLABSI subject area mart (SAM) using the above logic. A risk score for any patients that receive a central line is calculated and appended to an output table each time the SAM refreshes.

Reference

CLABSI Toolkit – Preventing Central-Line Associated Bloodstream Infections: Useful Tools, An International Perspective. The Joint Commission. Published May, 2012. Accessed August 8, 2016. http://www.jointcommission.org/CLABSIToolkit.

CLABSI



Encounter	MRN	Patient	Probability Group	Prediction Probability	Location	Department	Room	Top 3 Risk Factors
		Bkcen, Qrfuwx	High	0.64	Eehst Dtfs mq ...	HQ LMQW	T7221	1:LineDays 2:LineDaysNonTunneled 3:LineDaysInterna...
		Ehxvt, Ibygoc	High	0.58	Eehst Dtfs mq ...	HQ LMQW	T7214	1:LineDays 2:LineDaysNonTunneled 3:GenderCD.Unk...
		Jfszhhv, Czduhum	High	0.46	Eehst Dtfs mq ...	HQ LMQW	T7220	1:LineDays 2:LineDaysTunneled 3:LineDaysNonTunneled
		Gyjsv, Zewkte	High	0.41	Eehst Dtfs mq ...	HQ LMQW	T7222	1:LineDays 2:LineDaysNonTunneled 3:LineDaysInterna...
		Ifqpgcx, Bbksdfi Zmmf	High	0.20	Eehst Dtfs mq ...	HQ LMQW	T7224	1:LineDays 2:LineDaysNonTunneled 3:LineDaysInterna...
		Ondijlyv, Aowj	High	0.11	Eehst Dtfs mq ...	HQ LMQW	T3217	1:LineDays 2:LineDaysNonTunneled 3:LineDaysInterna...
		Lfgkbhy, Weqhg Q	Medium	0.08	Eehst Dtfs mq ...	HQ LMQW	T7212	1:LineDays 2:LineDaysNonTunneled 3:GenderCD.Unk...
		Xpqmuke, Qnabps	Medium	0.04	Eehst Dtfs mq ...	HQ LMQW	T3220	1:LineDays 2:LineDaysNonTunneled 3:GenderCD.Unk...
		Nhx, Uxmax V	Low	0.03	Eehst Dtfs mq ...	HQ LMQW	T3215	1:LineDays 2:LineDaysNonTunneled 3:LineDaysInterna...
		Owucqntgnr, Knxyfgj	Low	0.01	Eehst Dtfs mq ...	HQ LMQW	T3224	1:LineDays 2:LineDaysNonTunneled 3:LineDaysInterna...
		Mqxod, Azqf	Low	0.01	Eehst Dtfs mq ...	HQ LMQW	T3216	1:LineDays 2:LineDaysTunneled 3:LineDaysNonTunneled
		Byqmnz, Wuu	Low	0.01	Eehst Dtfs mq ...	HQ LMQW	T3219	1:LineDays 2:LineDaysNonTunneled 3:GenderCD.Unk...

Get Buy-In

Getting Buy-In from Clinicians

“My patients are sicker.”

“You have a FALSE POSITIVE rate of what?”

Tips for Getting Buy In from Clinicians

#1 Clinicians need to understand the model

If you cannot explain the algorithm, do not use it. Use a simpler algorithm that you can explain.

Tips for Getting Buy In from Clinicians

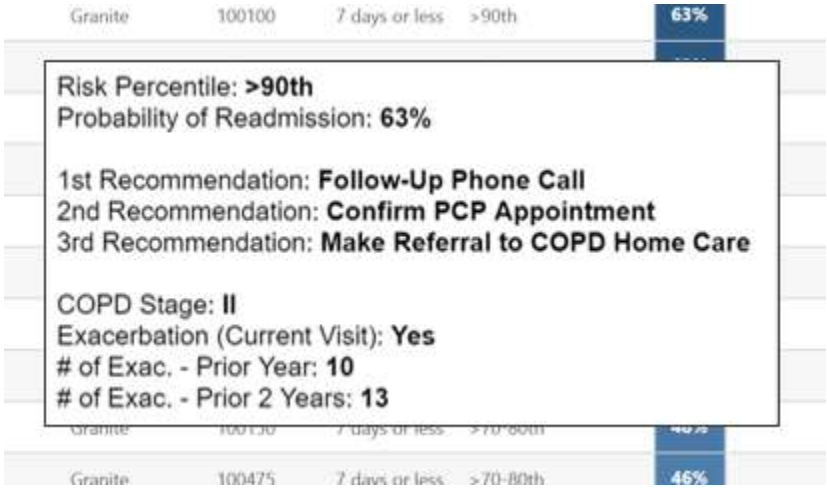
#2 Build a “model performance report”

Documentation for any interested stakeholder to learn about the model:

- Why was it created?
- What features were tried? Which were used?
- What algorithm was used?
- How accurate is the model?

Tips for Getting Buy In from Clinicians

#3 Provide details to end users



Tips for Getting Buy In from Clinicians

#4 It's just a suggestion

"Suggestive Analytics" may be a better term than "Predictive Analytics" to demonstrate that we are not trying to replace human judgement.

Review

Useful vocabulary for discussing predictive analytics

Usefulness of custom predictive models

The steps to build a predictive model

Examples of how predictive analytics has been deployed in the wild

Tips for getting buy-in from clinicians

Getting Started

You Need Smart People!

Data Scientist

- Formulates hypotheses about features driving a predictive model (with clinical input)
- Tries various algorithms to determine best approach for prediction
- Assesses model output and accuracy and operationalizes the best approach

Data Architect (Engineer)

- Finds and provisions source data
- Leverages definitions in analytics environment
- **Feature engineering**

Machine Learning Engineer

- Develops software to automate machine learning workflow
- Requires data science knowledge
- Requires knowledge of software engineering best practices
- **A rare find!**

HealthCatalyst
ignite outcomes improvement

healthcare.ai Open Source Software

Our open-source machine learning software product

Automates key tasks in developing models, or customizing existing models using local data

Makes deployment in an analytics environment easy and ‘production quality’

Machine Learning in Healthcare: Now for Everyone

Healthcare.ai is a community with education and open source technology tools focused on increasing the national adoption of machine learning in healthcare

WEEKLY BROADCASTS GET STARTED

Machine learning for healthcare just got a whole lot easier

The healthcare.ai packages are designed to streamline healthcare machine learning. They do this by including functionality specific to healthcare, as well as simplifying the workflow of creating and deploying models.

Learn more about machine learning via the healthcare.ai community by reading and subscribing to our weekly blogs, viewing our weekly YouTube live event broadcasts, and engaging our data science team

LEARN & ENGAGE WITH OUR EXPERTS

Getting Started in R and RStudio

Hosted by Levi Thatcher Director of Data Science, Health Catalyst

Thursday, February 23, 2017 3:00pm - 60min

WATCH LIVE

Machine Learning Broadcast Topic TBD