IEOR E4650 Business Analytics

Session 11: Quality of Predictions and Healthcare Analytics

Spring 2018

Prof. Adam Elmachtoub

Outline

- The challenge of readmissions reductions at Tahoe Healthcare
- Classification
- Performance of a classifier
- Economic tradeoffs in classification

Case Study: Tahoe Healthcare System

- Case study uses real, but anonymized, data
- Operates 14 hospitals in the Pacific Northwest
- 18% of total revenues are from Medicare reimbursement for the three HRRP conditions
- Management is concerned about the impact of the new HRRP rules on reimbursement revenues
- CareTracker, a new program the clinical staff has piloted with AMI patients, has proved effective at reducing readmissions through a combination of patient education and post-discharge monitoring:
  - Cost/patient: $1,200
  - Reduces readmission risk by 40%
  - Reimbursement penalty/readmit: $8,000

Should CareTracker be deployed?

Hospital Readmissions

New England Journal of Medicine Study (2009)

- Approximately 20% of hospitalized Medicare patients are readmitted within 30 days; 34% are readmitted within 90 days
- Estimated cost to the US healthcare system: $17 billion

2010 Affordable Care Act established a Hospital Readmissions Reduction Program (HRRP)

- Medicare payments to hospitals are reduced for excess readmissions
- Three conditions: acute myocardial infraction (AMI), heart failure (HF), and pneumonia
- Based on 30-day, risk-adjusted readmissions rate
- 3-year rolling horizon measure

Source: Medicare and Medicaid: Readmissions Reduction Program (p.112)
http://tinyurl.com/9kqtob2

Medicare Readmissions Stats

TABLE 5–3

Hospital readmissions for seven conditions make up almost 30 percent of spending on readmissions

Condition	Type of hospital admission	Number of admissions with readmissions	Readmission rate	Average Medicare payment for readmission	Total spending on readmissions
Heart failure	Medical	90,273	12.5%	$6,531	$590,000,000
COPD	Medical	52,327	10.7	6,587	345,000,000
Pneumonia	Medical	74,419	9.5	7,165	533,000,000
AMI	Medical	20,866	13.4	6,535	136,000,000
CABG	Surgical	18,554	13.5	8,136	151,000,000
PTCA	Surgical	44,293	10.0	8,109	359,000,000
Other vascular	Surgical	18,029	11.7	10,091	182,000,000
Total for seven conditions		318,760			$2,296,000,000
Total DRGs		1,134,483			$7,980,000,000
Percent of total		28.1%			28.8%

Note: COPD (chronic obstructive pulmonary disease), AMI (acute myocardial infarction), CABG (coronary artery bypass graft), PTCA (percutaneous transluminal coronary angioplasty), DRG (diagnosis related group). Analysis is for readmissions within 15 days of discharge from the initial inpatient stay. Readmissions are identified using 3M's software that defines potentially preventable readmissions.

Source: 3M analysis of 2005 Medicare discharge claims data.

Report to Congress: Promoting Greater Efficiency in Medicare
http://www.caretransitions.org/documents/MedPAC%20report.pdf

Interventions to Reduce Readmissions

- During hospitalization
  - Tailored patient care
  - Communication with PCP, family and home care
  - Patient education
- At discharge
  - Discharge planning
  - Patient/caregiver education
  - Transition coaching
  - Schedule and prepare follow-up appointments
- Post-discharge
  - Home nursing visits
  - Phone follow-up checks
  - Tele-health monitoring

Interventions are labor-intensive and costly

Status Quo: Don't Use CareTracker

	Treatment		
Outcome	1	0	Total
1	0	998	998
0	0	3,384	3,384
Total	0	4,382	4,382

Cost matrix

	Treatment	
Outcome	1	0
1		\$8000
0		\$0

- Outcome (in data set): Readmitted = 1, Not readmitted = 0
- Treatment (hypothetical): CareTracker = 1, No CareTracker = 0

Total Cost: (998 × \$8000)+ (3,384 × \$0) = \$7,984,000

Should CareTracker Be Used On All AMI Patients?

	Treatment		
Outcome	1	0	Total
1	998	0	998
0	3,384	0	3,384
Total	4,382	0	4,382

Cost matrix

	Treatment	
Outcome	1	0
1	\$6000	\$8000
0	\$1200	\$0

Where the \$6000 comes from for the "readmitted / CareTracker" patients:

- Caretracker cost: \$1,200
- Readmit cost: \$4,800 (= 0.6 × \$8,000)
- Total cost: \$6,000 (= \$1,200 + \$4,800)

Total Cost:
(998 × \$6000) + (0 × \$8000) + (3,384 × \$1200) + (0 × \$0) = \$10,048,800

Net change in cost: \$7,984,000 − \$10,048,800 = -\$2,064,800

Universal application of CareTracker is not beneficial

What Else Could Be Done?

Idea: Rather than applying CareTracker to every AMI patient, only apply it to those "at risk" for readmission

How?

- How do we go about assessing readmission risk?
- What threshold on risk do we use to classify patients as "at risk" and target them for CareTracker?

This requires a classification prediction: the ability to declare (yes/no) whether a particular patient is at risk and should be given special treatment

What is the Potential Value to be Captured?

Suppose we had perfect foresight and could apply CareTracker only to patients we knew would be readmitted

	Treatment		
Outcome	1	0	Total
1	998	0	998
0	0	3,384	3,384
Total	998	3,384	4,382

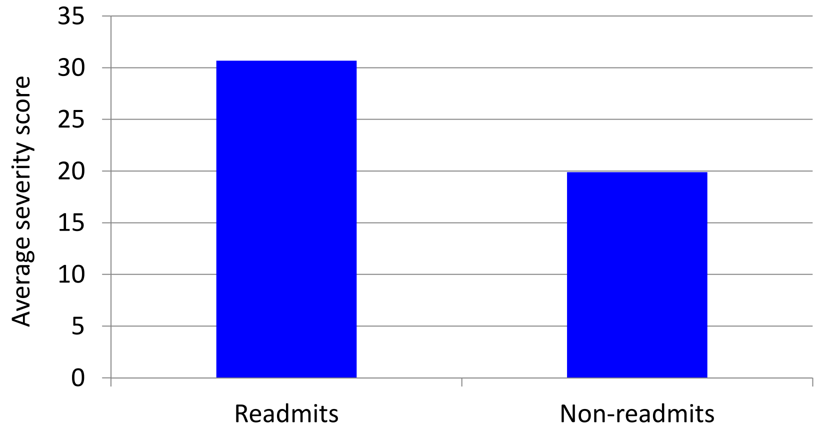
Cost matrix

	Treatment	
Outcome	1	0
1	\$6000	\$8000
0	\$1200	\$0

Total Cost: (998 × \$6000) + (3,384 × \$0) = \$5,988,000
Net change in cost: \$7,984,000 − \$5,988,000 = \$1,996,000

- The maximum reduction in cost possible from CareTracker is: \$1,996,000
- This provides a benchmark for evaluating improvements
- How to capture some of this value?

One Simple Idea for Classification



- Patients who are readmitted have higher average illness severity scores when initially admitted
- Maybe we should target only patients with high severity scores for enrollment in CareTracker?

What severity score would you recommend we use for classifying patients?

Performance of the Severity Classifier

Classification steps

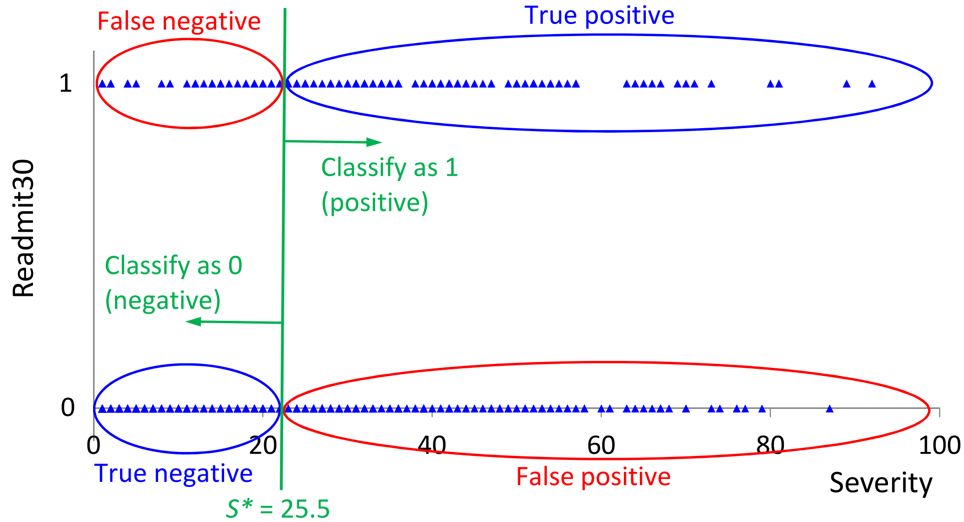
1. Pick a threshold on severity score, S^*
2. If patient has severity score $S > S^*$, then classify as"high risk" for readmission and enroll them in CareTracker
3. If patient has severity score $S \leq S^*$, then classify them as "low risk" and do not enroll them in CareTracker

How well does this classifier perform?

Issues

1. What are the types of classification errors?
2. How do they depend on the threshold?
3. How can we pick the optimal threshold?

Severity Score Classifier: Classification Outcomes



Four possible outcomes

1. We predict readmit and are correct (true positive)
2. We predict readmit and are incorrect (false positive)
3. We predict no readmit and are correct (true negative)
4. We predict no readmit and are incorrect (false negative)

Evaluating a Classifier: Confusion Matrix

There are four possible outcomes of any (binary) classification:

		Predicted 1	Predicted 0	
Actual	1	**True positive**	**False negative**	Row sum: total actual positives (e.g., readmits)
	0	**False positive**	**True negative**	Row sum: total actual negatives (e.g., non-readmits)
		Column sum: total predicted positives (e.g., predicted readmits)	Column sum: total predicted negatives (e.g., predicted non-readmits)	

1 - Positive: readmitted

0 - Negative: not readmitted

Error Rates

		Predicted 1	Predicted 0
Actual	1	**True positive**	**False negative**
	0	**False positive**	**True negative**

How likely are we to make an error of some type in classifying?

$$\text{Total error rate} = \frac{\#\text{ False positives} + \#\text{ False negatives}}{\text{Total number of outcomes}}$$

How likely are we to misclassify a negative as a positive?

$$\text{False positive rate} = \frac{\#\text{ False positives}}{\text{Total number of actual negatives}}$$

How likely are we to correctly classify an observation as positive?

$$\begin{array}{c}\text{True positive rate}\\ \text{(sensitivity)}\end{array} = \frac{\#\text{ True positives}}{\text{Total number of actual positives}}$$

Severity Classifier with $S^* = 25.5$

Confusion matrix:

	Predicted		
Actual	1	0	Total
1	546	452	998
0	1,041	2,343	3,384
Total	1,587	2,795	4,382

How likely are we to make an error of some type in classifying?

Total error rate = 34.1% ((1,041+452)/4,382)

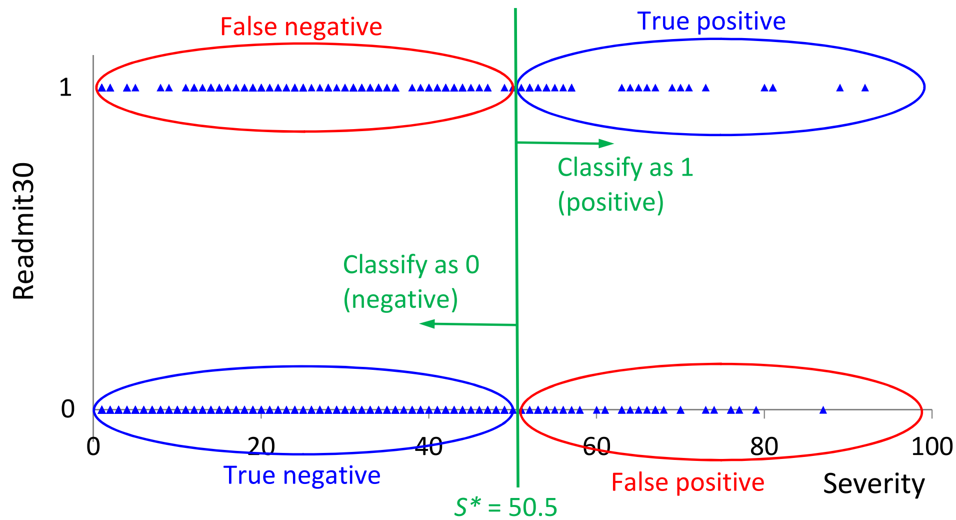
How likely are we to misclassify a negative as a positive?

False positive rate = 30.8% (1,041/3,384)

How likely are we to correctly classify an observation as positive?

True positive rate = 54.7% (546/998)

Comparison of $S^* = 25.5$ to $S^* = 50.5$



What happens when S^* is increased from 25.5 to 50.5?

Severity Classifier for $S^* = 50.5$

Confusion matrix for $S^* = 25.5$

	Predicted		
Actual	1	0	Total
1	546	452	998
0	1,041	2,343	3,384
Total	1,587	2,795	4,382

Confusion matrix for $S^* = 50.5$

	Predicted		
Actual	1	0	Total
1	172	826	998
0	194	3,190	3,384
Total	366	4,016	4,382

Total error rate has gone down:

$$\text{Total error rate} = 23.3\% \quad ((194+826)/4{,}382)$$

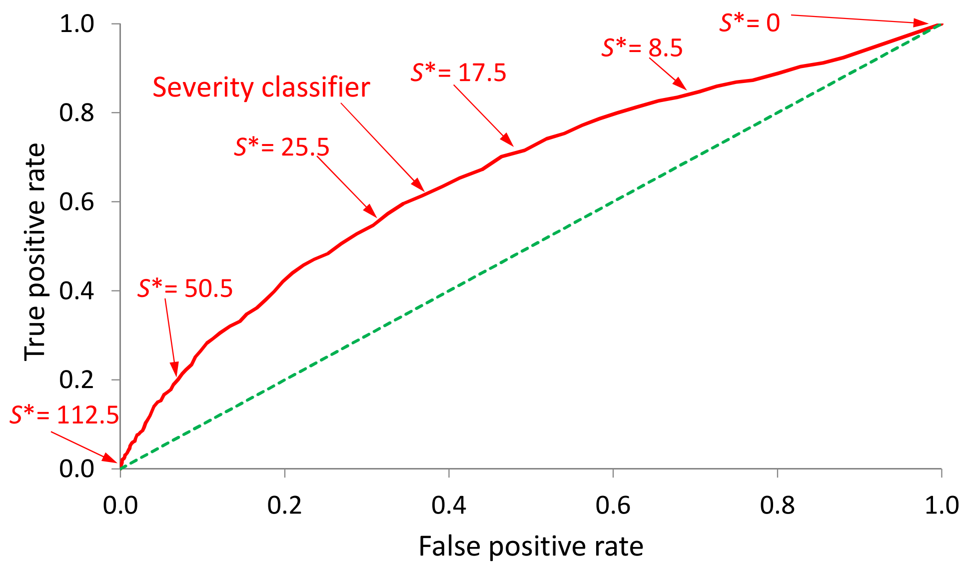
False positives went down (good), so we are treating fewer patients who do not need intervention:

$$\text{False positive rate} = 5.7\% \quad (194/3{,}384)$$

True positives went down (bad), which means we are intervening less often with patients who are at risk:

$$\text{True positive rate} = 17.2\% \quad (172/998)$$

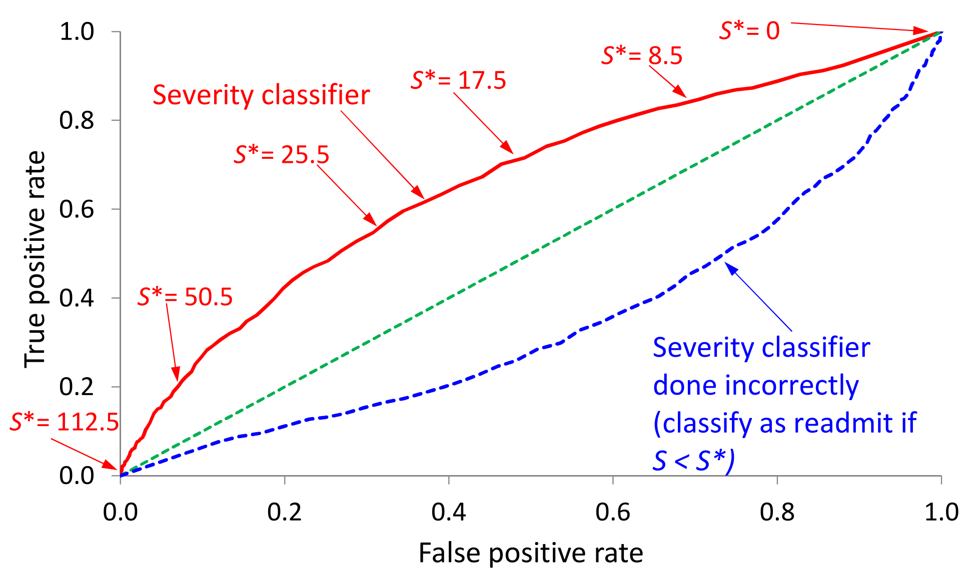
Receiver Operating Characteristic (ROC) Curve



ROC curve: plot of the true positive rate versus the false positive rate for a range of severity thresholds S^*

Green line (45 degrees): Random (zero intelligence) classifier

Receiver Operating Characteristic (ROC) Curve



What would the ROC curve look like if classification was done incorrectly (if $S < S^*$, then classify as readmit)?

Receiver Operating Characteristic (ROC) Curve

- ROC curve captures the trade-off between two different types of errors (false positives, false negatives)
- Changing the threshold will only decrease one type of error at the expense of increasing another type of error
- How can we decide which threshold to select?

Net Benefit for $S^* = 25.5$

Classification rule: If severity score $>$ 25.5, then predict readmission and give CareTracker treatment; otherwise no CareTracker treatment

Confusion matrix

	Predicted		
Actual	1	0	Total
1	546	452	998
0	1,041	2,343	3,384
Total	1,587	2,795	4,382

Cost matrix

	Predicted	
Actual	1	0
1	\$6000	\$8000
0	\$1200	\$0

Status quo (no Caretracker):	\$7,984,000
Cost for $S^* = 25.5$:	\$8,141,200
Net benefit:	−\$157,200

Detailed cost calculation for $S^* = 25.5$:

(546 × \$6000) + (452 × \$8000) + (1,041 × \$1200) + (2,343 × \$0) = \$8,141,200

Net Benefit for $S^* = 50.5$

Classification rule: If severity score > 50.5, then predict readmission and give CareTracker treatment; otherwise no CareTracker treatment

Confusion matrix

	Predicted		
Actual	1	0	Total
1	172	826	998
0	194	3,190	3,384
Total	366	4,016	4,382

Cost matrix

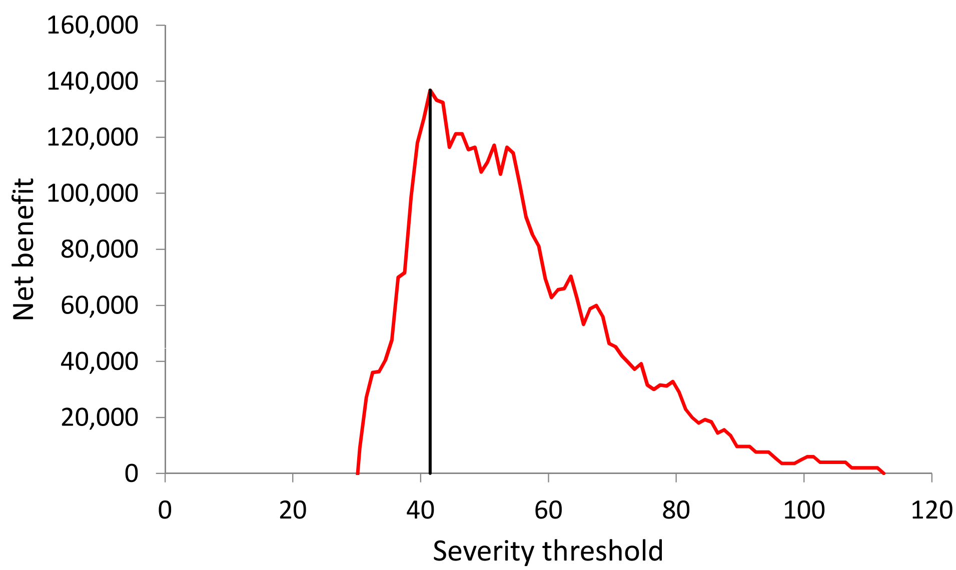
	Predicted	
Actual	1	0
1	\$6000	\$8000
0	\$1200	\$0

Status quo (no Caretracker):	\$7,984,000
Cost for $S^* = 50.5$:	\$7,872,800
Net benefit:	\$111,200

Detailed cost calculation for $S^* = 50.5$:

$(172 \times \$6000) + (826 \times \$8000) + (194 \times \$1200) + (3{,}190 \times \$0) = \$7{,}872{,}800$

Optimizing over the Severity Threshold S^*



$S^* = 41.5$ provides the maximum net benefit of \$136,800

Classification has turned CareTracker from a losing to a winning proposition!

Can we do better?

Data Dictionary

Variable	Interpretation
age	age of patient
female	1 if female, 0 if male
flu_season	1 if admitted during flu season, 0 otherwise
ed_admit	1 if admitted in ER, 0 otherwise
severity_score	severity of illness score
comorbidity_score	severity of pre-existing conditions
readmit30	1 if readmitted in 30 days, 0 otherwise

Can More Information Improve the Classification?

age	female	flu_season	ed_admit	severity	comorbidity	readmit30
100	1	1	1	38	112	0
83	1	0	1	8	109	1
74	0	1	0	1	80	0
66	1	1	1	25	4	0
68	1	1	1	25	32	0
80	1	0	1	29	172	0
71	1	0	1	31	271	1
72	0	0	0	47	221	1
69	1	0	1	44	193	0
65	1	0	0	10	130	0
74	0	0	1	11	38	0
72	0	0	0	5	8	0
67	0	0	1	12	52	0
75	0	0	1	41	75	0
67	1	1	1	45	73	1
82	1	0	0	1	26	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Logistic Regression

- Severity is only one of the variables that might be predictive of readmissions
- Logistic regression provides a prediction of the probability of readmission using all of these factors

LogisticRegTrain	Constant	age	female	flu_ season	ed_ admit	severity score	comorbidity score
Coefficients	−4.016	0.002	0.190	0.743	−0.159	0.027	0.016
Std error	0.410	0.005	0.082	0.082	0.115	0.002	0.001
p-value	0.000	0.733	0.021	0.000	0.165	0.000	0.000
Num valid obs	4382						

$$\text{Probability of readmit} = \frac{\exp(w)}{1+\exp(w)} = \frac{e^w}{1+e^w}$$

$$\text{with } w = -4.016 + 0.002 \times \text{age} + 0.190 \times \text{female} + 0.743 \times \text{flu\_season} \\ - 0.159 \times \text{ed\_admit} + 0.027 \times \text{severity} + 0.016 \times \text{comorbidity}$$

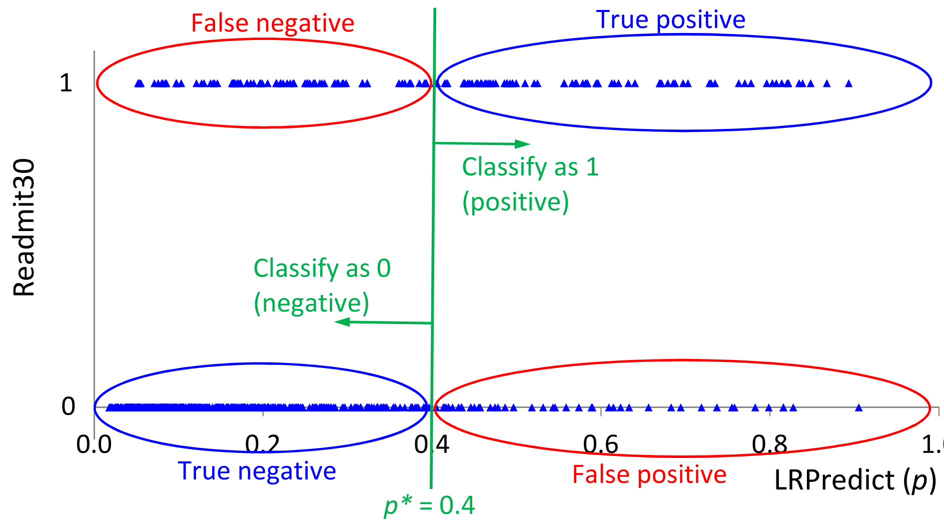
Converting Logistic Probabilities to Classification

	A	B	C	D	E	F	G	H	I
1	age	female	flu_season	ed_admit	severity score	comorbidity score	readmit30	LR Predict	classification
2	100	1	1	1	38	112	0	0.43	1
3	83	1	0	1	8	109	1	0.13	0
4	74	0	1	0	1	80	0	0.14	0
5	66	1	1	1	25	4	0	0.08	0
6	68	1	1	1	25	32	0	0.13	0
7	80	1	0	1	29	172	0	0.41	1
8	71	1	0	1	31	271	1	0.78	1
9	72	0	0	0	47	221	1	0.70	1
10	69	1	0	1	44	193	0	0.59	1
11	65	1	0	0	10	130	0	0.20	0
12	74	0	0	1	11	38	0	0.04	0
13	72	0	0	0	5	8	0	0.03	0
14	67	0	0	1	12	52	0	0.05	0
15	75	0	0	1	41	75	0	0.15	0
16	67	1	1	1	45	73	1	0.32	0

Classification: Compare the predicted probability p to a threshold probability p^*

- If $p > p^*$, classify patient as readmit (1), i.e., receive CareTracker
- If $p \leq p^*$, classify patient as no readmit (0) i.e., don't receive CareTracker
- Spreadsheet shows results for $p^* = 0.40$

Logistic Classifier with $p^* = 0.40$



Compared to the severity classifier (with $S^* = 40.5$), the logistic classifier accounts for additional information and is based on a threshold on the probability of readmission.

Net Benefit for Logistic Classifier with $p^* = 0.40$

Classification rule: If $p^* > 0.40$, then predict readmission and give CareTracker treatment; otherwise no CareTracker treatment

Confusion matrix

	Predicted		
Actual	1	0	Total
1	442	556	998
0	324	3,060	3,384
Total	766	3,616	4,382

Cost matrix

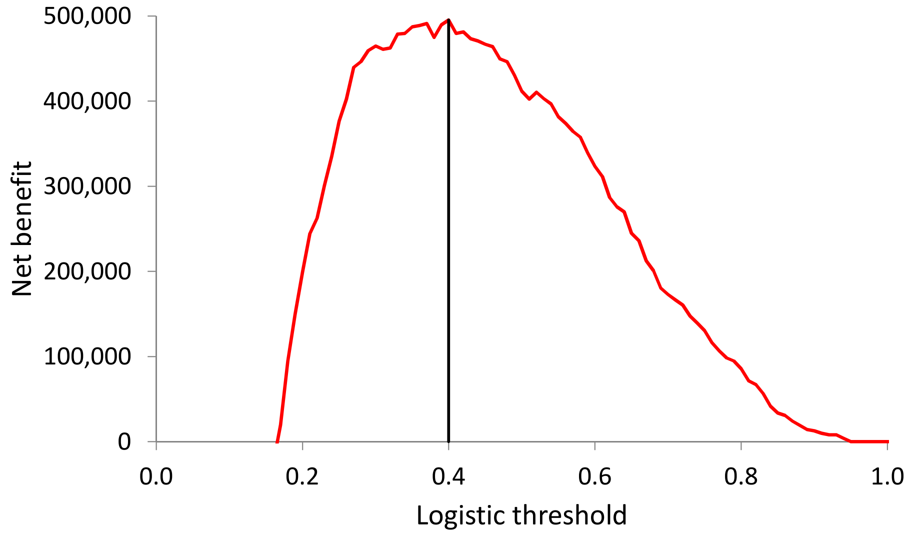
	Predicted	
Actual	1	0
1	\$6000	\$8000
0	\$1200	\$0

Status quo (no Caretracker):	\$7,984,000
Cost for $p^* = 0.4$:	\$7,488,800
Net benefit:	\$495,200

Detailed cost calculation for $p^* = 0.40$:

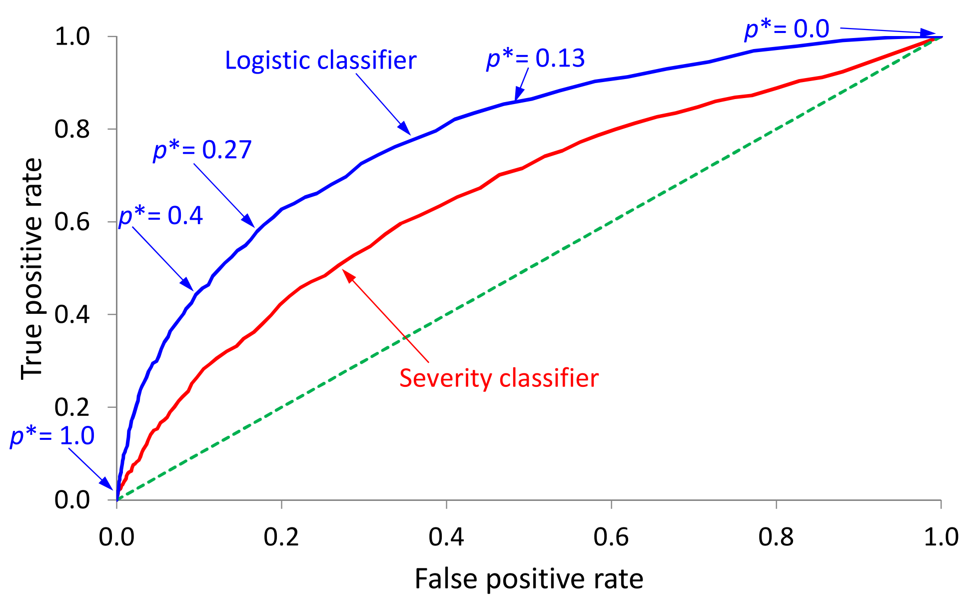
(442 $\times$ \$6000) + (556 $\times$ \$8000) + (324 $\times$ \$1200) + (3,060 $\times$ \$0) = \$7,488,800

Optimizing over the Logistic Probability Threshold p^*



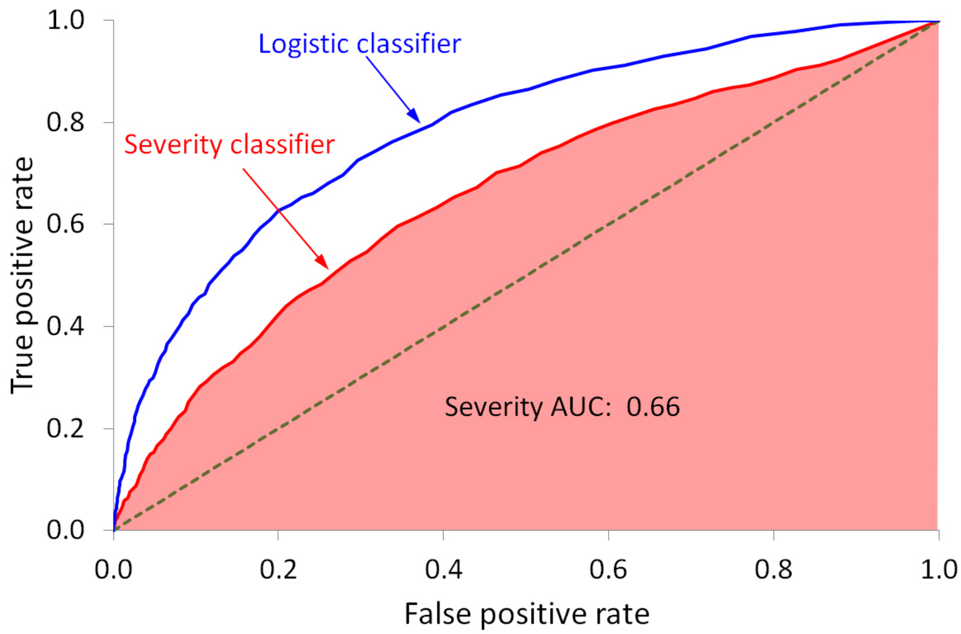
- Logistic threshold $p^* = 0.40$ provides the maximum net benefit of $495,200
- Captures 25% (495,200/1,976,000) of the potential value relative to perfect foresight
- Reduces readmission cost by 6% (459,200/7,984,000)

Logistic and Severity Classifiers: ROC Curves

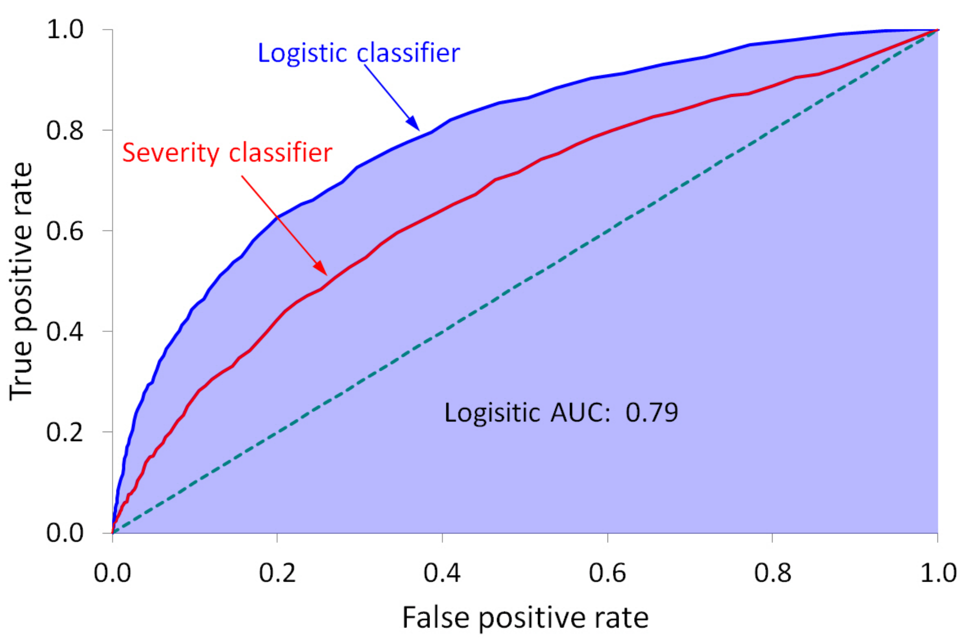


Logistic classifier provides a provides a better tradeoff between true positives (TP) and false positives (FP)

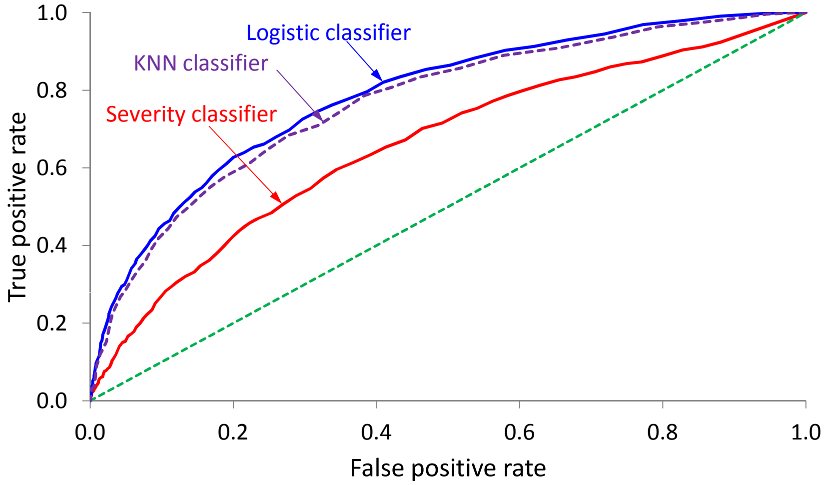
Area Under the Curve (AUC): Severity Classifier



Area Under the Curve (AUC): Logistic Classifier



KNN Classifier: ROC and AUC



What is done in practice?

Classifier	AUC	Net Benefit	Threshold (S^* or p^*)
Severity	0.66	136,800	40.5
Logistic	0.79	495,200	0.40
KNN	0.77	455,200	0.33

KNN uses $k = 64$ nearest neighbors

Microsoft Amalga System

READMISSIONS MANAGEMENT

Reducing Readmissions

Background

With Medicare deadlines for readmissions improvements looming in 2012, it's critical to start addressing readmission risks today. Solutions that help you count last month's readmissions are no longer sufficient. You need to know which patients in your hospital today are at risk for being readmitted within 30 days of discharge, so you can take action and address these risks before they walk out your door.

Microsoft Amalga helps you proactively identify inpatient and Emergency Department (ED) patients at risk for readmissions and helps you take action to avoid preventable readmissions, reduce costs, and deliver higher quality care today and tomorrow.

Overview

By using predictive modeling technologies, Amalga helps you reduce preventable readmissions by enabling you to:

- Effectively define and monitor patient groups across your enterprise
- Use data collected in Amalga to predict readmission probability based on hospital experience and historical data
- Proactively manage at-risk patients throughout their stay and at discharge
- Assess patterns in key indicators to identify and address root causes of readmissions

Features & Benefits

With Amalga, you can:

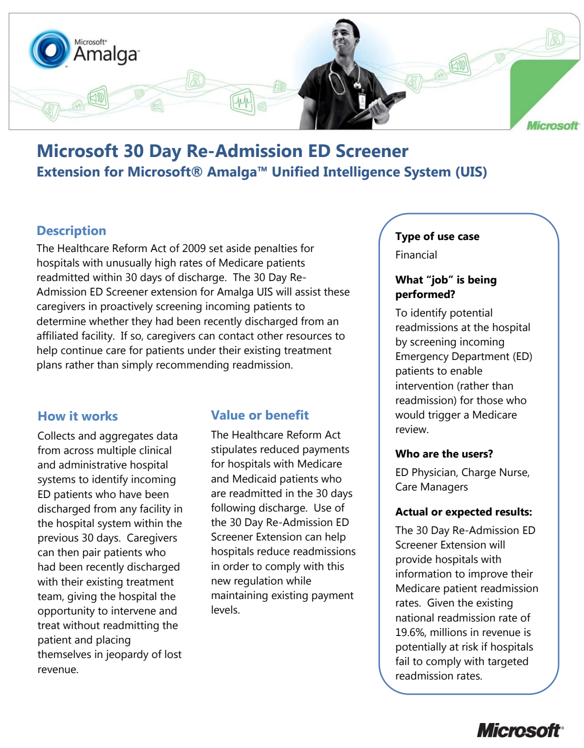
- Actively identify and track patient groups
- Integrate disparate systems and then identify patient cohorts by key attributes and characteristics
- Use predictive modeling technologies to help identify patients at risk for readmissions
- Analyze readmission patterns and monitor 30-day inpatient and 72-hour ED readmissions
- See simplified reports so you can address root causes sooner

Microsoft Amalga, Vergence, and Microsoft HealthVault together create the connected health intelligence platform. The platform enables health enterprises to streamline operations, coordinate care, and engage patients.

whatsnextinhealth.com/readmissions

Microsoft Amalga

Microsoft



Microsoft Amalga

Microsoft

Microsoft 30 Day Re-Admission ED Screener

Extension for Microsoft® Amalga™ Unified Intelligence System (UIS)

Description

The Healthcare Reform Act of 2009 set aside penalties for hospitals with unusually high rates of Medicare patients readmitted within 30 days of discharge. The 30 Day Re-Admission ED Screener extension for Amalga UIS will assist these caregivers in proactively screening incoming patients to determine whether they had been recently discharged from an affiliated facility. If so, caregivers can contact other resources to help continue care for patients under their existing treatment plans rather than simply recommending readmission.

How it works

Collects and aggregates data from across multiple clinical and administrative hospital systems to identify incoming ED patients who have been discharged from any facility in the hospital system within the previous 30 days. Caregivers can then pair patients who had been recently discharged with their existing treatment team, giving the hospital the opportunity to intervene and treat without readmitting the patient and placing themselves in jeopardy of lost revenue.

Value or benefit

The Healthcare Reform Act stipulates reduced payments for hospitals with Medicare and Medicaid patients who are readmitted in the 30 days following discharge. Use of the 30 Day Re-Admission ED Screener Extension can help hospitals reduce readmissions in order to comply with this new regulation while maintaining existing payment levels.

Type of use case

Financial

What "job" is being performed?

To identify potential readmissions at the hospital by screening incoming Emergency Department (ED) patients to enable intervention (rather than readmission) for those who would trigger a Medicare review.

Who are the users?

ED Physician, Charge Nurse, Care Managers

Actual or expected results:

The 30 Day Re-Admission ED Screener Extension will provide hospitals with information to improve their Medicare patient readmission rates. Given the existing national readmission rate of 19.6%, millions in revenue is potentially at risk if hospitals fail to comply with targeted readmission rates.

Microsoft

Microsoft Amalga System Data

ACCOUNT	ADMITDTTM	DISCHARGEDTTM	AGE	SEX	PROB_NUM_%	FACTOR
	12/03/2010 14:57	12/08/2010 18:03	62	F	37.9	Num past 6m visits = 6 to 10 /
	12/08/2010 18:45	12/08/2010 18:45	74	M	32.72	stayed <1 day in the hospital / P
	11/16/2010 16:14	12/08/2010 18:50	48	M	30.83	Patient had dx = Chronic renal fa
	12/02/2010 13:49	12/08/2010 18:14	68	M	29.05	Patient had dx = Disorders of flu
	12/01/2010 05:26	12/08/2010 18:55	44	M	28.54	
	12/01/2010 19:08	12/08/2010 18:13	61	M	27.36	Patient had dx = Acute renal fail
	11/30/2010 21:50	12/08/2010 18:52	70	M	18.05	Patient had dx = Other personal
	12/08/2010 08:51	12/08/2010 18:45	68	M	16.57	stayed <1 day in the hospital
	12/03/2010 20:32	12/08/2010 17:50	80	M	16.18	Patient had dx = Disorders of flu
	12/01/2010 01:13	12/08/2010 18:06	79	M	15.52	
	12/08/2010 18:39	12/08/2010 18:39	22	F	14.53	stayed <1 day in the hospital / Av
	12/08/2010 19:01	12/08/2010 19:01	25	F	14.42	stayed <1 day in the hospital / P
	12/08/2010 18:05	12/08/2010 18:05	24	M	14.39	stayed <1 day in the hospital
	12/08/2010 18:26	12/08/2010 18:26	53	F	13.59	stayed <1 day in the hospital / 44

Data

- Age, gender, demographics
- Diagnosis codes
- Lab results
- Medications
- Vital signs
- Procedures
- Treatment history

Amalga system uses up to 30,000 variables

Webinar: Machine Learning and Predictive Modeling in Healthcare
http://www.youtube.com/watch?v=K6Hlxa1WEMg

Summary

- Classification predictions are widely used
  - Credit card fraud: is a given transaction fraudulent?
  - Anti-terrorism: is someone crossing the border a potential terrorist?
  - Spam detection: is a given email spam?
- Types of classification errors
  - False positive: treating patients who do not need the treatment
  - False negative: not treating patients who need the treatment
- Economic tradeoffs in classification
  - Different costs for each type of error leads to different optimal classification thresholds
- Performance measures of classifiers
  - ROC curves: shows the tradoff between the true and false positive rates
  - AUC: area under the ROC curve (roughly analogous to R^2 in a regression)