ML for readmission reduction, DRG classification and resource allocation

Natalia Serna
Quantil
Alvaro J. Riascos Villegas
Universidad de los Andes y Quantil

Febrero, 2017

Contenido

Introduction

- Data driven decisions for reducing readmissions for heart failure: General methodology and case study. PLOS ONE. 2014.

Introduction

- Data driven decisions for reducing readmissions for heart failure: General methodology and case study. PLOS ONE. 2014.

Contenido

Data driven decisions for reducing readmissions for heart failure

- Motivation: More than 12 billions USD spent in preventable readmissions.
- Data: 793 hospital visits for heart failure.
- Objective I: Construct a classifier to predict readmissions within 30 days of discharge.
- Objective II: Introduce a decision problem, post discharge intervention costs vrs. readmission, and evaluate cost effectiveness.
- Results: Using out of sample 379 cases they report: Readmission mean cost is $13,000 USD. A post discharge plan reduces 30-day hospitalizations by 35 %. If the post discharge plan costs $1,214 then this ML guided decision problem would reduce readmissions by 18,2 % and costs by 3,8 %

Data driven decisions for reducing readmissions for heart failure

- Motivation: More than 12 billions USD spent in preventable readmissions.
- Data: 793 hospital visits for heart failure.
- Objective I: Construct a classifier to predict readmissions within 30 days of discharge.
- Objective II: Introduce a decision problem, post discharge intervention costs vrs. readmission, and evaluate cost effectiveness.
- Results: Using out of sample 379 cases they report: Readmission mean cost is $\$13,000$ USD. A post discharge plan reduces 30-day hospitalizations by 35 %. If the post discharge plan costs $\$1,214$ then this ML guided decision problem would reduce readmissions by 18, 2 % and costs by 3, 8 %

Data driven decisions for reducing readmissions for heart failure

- Motivation: More than 12 billions USD spent in preventable readmissions.
- Data: 793 hospital visits for heart failure.
- Objective I: Construct a classifier to predict readmissions within 30 days of discharge.
- Objective II: Introduce a decision problem, post discharge intervention costs vrs. readmission, and evaluate cost effectiveness.
- Results: Using out of sample 379 cases they report: Readmission mean cost is $\$13,000$ USD. A post discharge plan reduces 30-day hospitalizations by 35 %. If the post discharge plan costs $\$1,214$ then this ML guided decision problem would reduce readmissions by 18, 2 % and costs by 3, 8 %

Data driven decisions for reducing readmissions for heart failure

- Motivation: More than 12 billions USD spent in preventable readmissions.
- Data: 793 hospital visits for heart failure.
- Objective I: Construct a classifier to predict readmissions within 30 days of discharge.
- Objective II: Introduce a decision problem, post discharge intervention costs vrs. readmission, and evaluate cost effectiveness.
- Results: Using out of sample 379 cases they report: Readmission mean cost is $\$13,000$ USD. A post discharge plan reduces 30-day hospitalizations by 35 %. If the post discharge plan costs $\$1,214$ then this ML guided decision problem would reduce readmissions by 18, 2 % and costs by 3, 8 %

Data driven decisions for reducing readmissions for heart failure

- Motivation: More than 12 billions USD spent in preventable readmissions.
- Data: 793 hospital visits for heart failure.
- Objective I: Construct a classifier to predict readmissions within 30 days of discharge.
- Objective II: Introduce a decision problem, post discharge intervention costs vrs. readmission, and evaluate cost effectiveness.
- Results: Using out of sample 379 cases they report: Readmission mean cost is $\$13,000$ USD. A post discharge plan reduces 30-day hospitalizations by $35\,\%$. If the post discharge plan costs $\$1,214$ then this ML guided decision problem would reduce readmissions by $18,2\,\%$ and costs by $3,8\,\%$

Methodology

- ML methodology:
- LASSO type logistic regression was used to select the most important variables using cross validation.
- Compared with LACE (index of readmissions) using ROC and reclassification analysis. See supplementary information.

Methodology

- ML methodology:
- LASSO type logistic regression was used to select the most important variables using cross validation.
- Compared with LACE (index of readmissions) using ROC and reclassification analysis. See supplementary information.

Methodology

- ML methodology:
- LASSO type logistic regression was used to select the most important variables using cross validation.
- Compared with LACE (index of readmissions) using ROC and reclassification analysis. See supplementary information.

Methodology

- Decision methodology:
- Cost of intervention and readmission the same for all patients: $C_{intervene}$, $C_{readmit}$.
- Efficacy of intervention is a priori the same $P_{success}$.
- Without intervention expected cost of readmission is $C_0(p) = p \times C_{readmit}$.
- With intervention is: $C_1(p) = C_{intervene} + p(1 - p_{success}) \times C_{readmit}$.
- For $p \geq p^* = \frac{C_{intervene}}{p_{success} C_{readmit}}$, $C_0(p) > C_1(p)$ so the agent should be intervened.

Methodology

- Decision methodology:
- Cost of intervention and readmission the same for all patients: $C_{intervene}$, $C_{readmit}$.
- Efficacy of intervention is a priori the same $P_{success}$.
- Without intervention expected cost of readmission is $C_0(p) = p \times C_{readmit}$.
- With intervention is: $C_1(p) = C_{intervene} + p(1 - p_{success}) \times C_{readmit}$.
- For $p \geq p^* = \frac{C_{intervene}}{p_{success} C_{readmit}}$, $C_0(p) > C_1(p)$ so the agent should be intervened.

Methodology

- Decision methodology:
- Cost of intervention and readmission the same for all patients: $C_{intervene}$, $C_{readmit}$.
- Efficacy of intervention is a priori the same $P_{success}$.
- Without intervention expected cost of readmission is $C_0(p) = p \times C_{readmit}$.
- With intervention is: $C_1(p) = C_{intervene} + p(1 - p_{success}) \times C_{readmit}$.
- For $p \geq p^* = \frac{C_{intervene}}{p_{success} C_{readmit}}$, $C_0(p) > C_1(p)$ so the agent should be intervened.

Methodology

- Decision methodology:
- Cost of intervention and readmission the same for all patients: $C_{intervene}$, $C_{readmit}$.
- Efficacy of intervention is a priori the same $P_{success}$.
- Without intervention expected cost of readmission is $C_0(p) = p \times C_{readmit}$.
- With intervention is: $C_1(p) = C_{intervene} + p(1 - p_{success}) \times C_{readmit}$.
- For $p \geq p^* = \frac{C_{intervene}}{p_{success} C_{readmit}}$, $C_0(p) > C_1(p)$ so the agent should be intervened.

Methodology

- Decision methodology:
- Cost of intervention and readmission the same for all patients: $C_{intervene}$, $C_{readmit}$.
- Efficacy of intervention is a priori the same $P_{success}$.
- Without intervention expected cost of readmission is $C_0(p) = p \times C_{readmit}$.
- With intervention is: $C_1(p) = C_{intervene} + p(1 - p_{success}) \times C_{readmit}$.
- For $p \geq p^* = \frac{C_{intervene}}{p_{success} C_{readmit}}$, $C_0(p) > C_1(p)$ so the agent should be intervened.

Methodology

- Decision methodology:
- Cost of intervention and readmission the same for all patients: $C_{intervene}$, $C_{readmit}$.
- Efficacy of intervention is a priori the same $P_{success}$.
- Without intervention expected cost of readmission is $C_0(p) = p \times C_{readmit}$.
- With intervention is: $C_1(p) = C_{intervene} + p(1 - p_{success}) \times C_{readmit}$.
- For $p \geq p^* = \frac{C_{intervene}}{p_{success} C_{readmit}}$, $C_0(p) > C_1(p)$ so the agent should be intervened.

Results

- LACE: AUC 0,59 %
- Logistic LASSO: 0,66 %.
- Cross validation training AUC mean is 0,69 %
- Significant readmissions to other hospitals. Removing this patients improves AUC 0,71 %.
- Best model selects 253 out of 3, 300.

Results

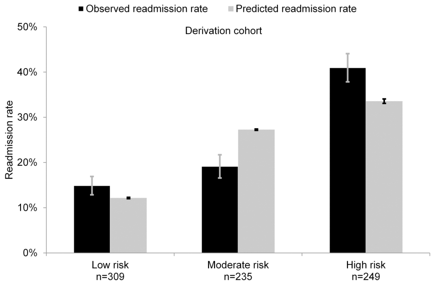
- LACE: AUC 0,59 %
- Logistic LASSO: 0,66 %.
- Cross validation training AUC mean is 0,69 %
- Significant readmissions to other hospitals. Removing this patients improves AUC 0,71 %.
- Best model selects 253 out of 3, 300.

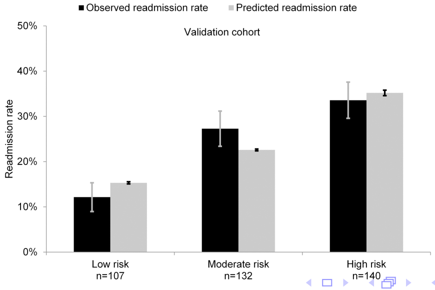
Results

- LACE: AUC 0,59 %
- Logistic LASSO: 0,66 %.
- Cross validation training AUC mean is 0,69 %
- Significant readmissions to other hospitals. Removing this patients improves AUC 0,71 %.
- Best model selects 253 out of $3,300$.

Results

- Calibración:





Results

- Variables are clustered and classified according to their evidential support.

Reclassification matrix

		Classifier			
		Low risk	**Moderate risk**	**High risk**	**Total reclassified (%)**
LACE	**Low risk (%)**	35.8	36.2	28.0	**64.2**
	Moderate risk (%)	16.8	33.6	49.6	**66.4**
	High risk (%)	0.0	23.5	76.5	**23.5**

doi:10.1371/journal.pone.0109264.t001

Results

- Top supportive: increases odds

Top supportive evidence

Variable class	Variable description	Log Odds Ratio	Log Odds Ratio Standard Error[1]
Lab Results	Lymphocyte % is low	0.0128	0.0027
Patterns of Engagement	Patient was admitted in past 6 months	0.0112	0.0031
Lab Results	BUN is high	0.0038	0.0012
Lab Results	Glucose level random is elevated	0.003	0.0012
Lab Results	Monocyte absolute is low	0.0028	0.0012
Other Diagnoses	History of nondependent abuse of drugs (ICD9 305.x)	0.0018	0.001
Other Diagnoses	History of chronic airway obstruction, not elsewhere classified (ICD9 496.x)	0.0017	0.0008
Other Diagnoses	History of gastrointestinal hemorrhage (ICD9 578.x)	0.0014	0.0007
Lab Results	AST is elevated	0.0013	0.0006
Other Diagnoses	History of cardiomyopathy (ICD9 425.x)	0.001	0.0006
Lab Results	Magnesium is low	0.001	0.0006
Lab Results	INR is elevated	0.0009	0.0004
Patterns of Engagement	Patient has been in isolated room in hospital	0.0009	0.0006
Lab Results	BNP is high	0.0007	0.0005

These are variables that receive positive log-odds ratio with the largest magnitude.
1. Obtained from sample standard error for cross-validation odds ratios
doi:10.1371/journal.pone.0109264.t002

Data driven decisions for reducing readmissions for heart failure

- Top supportive: decreases odds

Top disconfirming evidence

Variable class	Variable description	Log Odds Ratio	Log Odds Ratio Standard Error[1]
Patterns of Engagement	Number of emergency room visits during past 6 months <2	−0.0607	0.0035
Lab results	Hematocrit % is normal	−0.0442	0.0043
Lab results	BNP is normal	−0.044	0.0049
Lab results	Alkaline phosphatase is normal	−0.0428	0.0033
Lab results	Chloride is normal	−0.0428	0.0042
Cardiac medications	Patient is not on digoxin therapy	−0.0396	0.0039
Lab results	MCHC % is low	−0.0387	0.0039
Changes in lab results	TSH variation during current visit is low	−0.0343	0.003
Changes in lab results	CO2 variation during current visit is low	−0.0318	0.0039
Changes in lab results	RDW variation during current visit is low	−0.0308	0.0038
Changes in lab results	MCV variation during current visit is low	−0.0306	0.0036

These are variables that receive negative log-odds ratio with largest magnitude.
1. Obtained from sample standard error for cross-validation odds ratios
doi:10.1371/journal.pone.0109264.t003

Savings of different decision rules

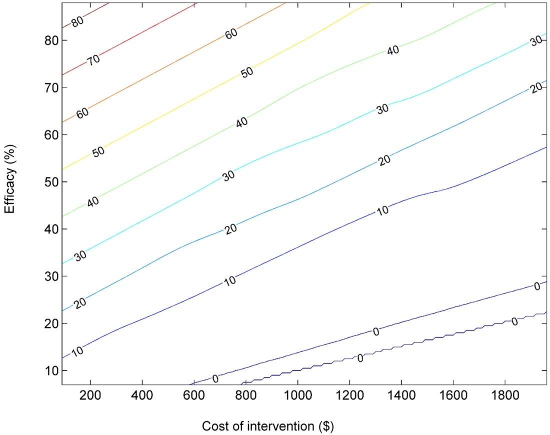
on of savings achieved and readmissions prevented for post-discharge programs with different costs and efficacies.

	Savings or losses				Readmissions prevented		
Efficacy	Patient-specific analysis and classifier	Patient-specific analysis and LACE	Intervention applied to all patients	Best uniform policy	Patient-specific analysis and classifier	Patient-specific analysis and LACE	Best uniform policy
25%	16.2%	15.9%	16.2%	16.2%	25.0%	24.5%	25.0%
35%	26.2%	26.2%	26.2%	26.2%	35.0%	35.0%	35.0%
25%	5.4%	1.3%	1.5%	1.5%	17.4%	2.9%	25.0%
35%	13.2%	9.1%	11.5%	11.5%	31.4%	22.2%	35.0%
25%	0.7%	0.0%	−13.2%	0.0%	5.2%	0.0%	0.0%
35%	3.8%	0.5%	−3.2%	0.0%	18.2%	2.6%	0.0%
25%	0.3%	0.0%	−27.8%	0.0%	0.8%	0.0%	0.0%
35%	0.8%	0.0%	−17.8%	0.0%	7.3%	0.0%	0.0%

ed: patient specific analysis using the classifier, patient-specific analysis using LACE, best uniform policy. For comparison of savings an additional column demonstrates the policy that applies intervention
l to loss (or savings).

0109264.t004

Savings of different decision rules



- Savings of decision analysis over no intervention.