Conditional Anomaly Detection Methods for Patient-Management Alert Systems

Michal Valko, Gregory Cooper, Amy Seybert,
Shyam Visweswaran, Melissa Saul, Milos Hauskrecht

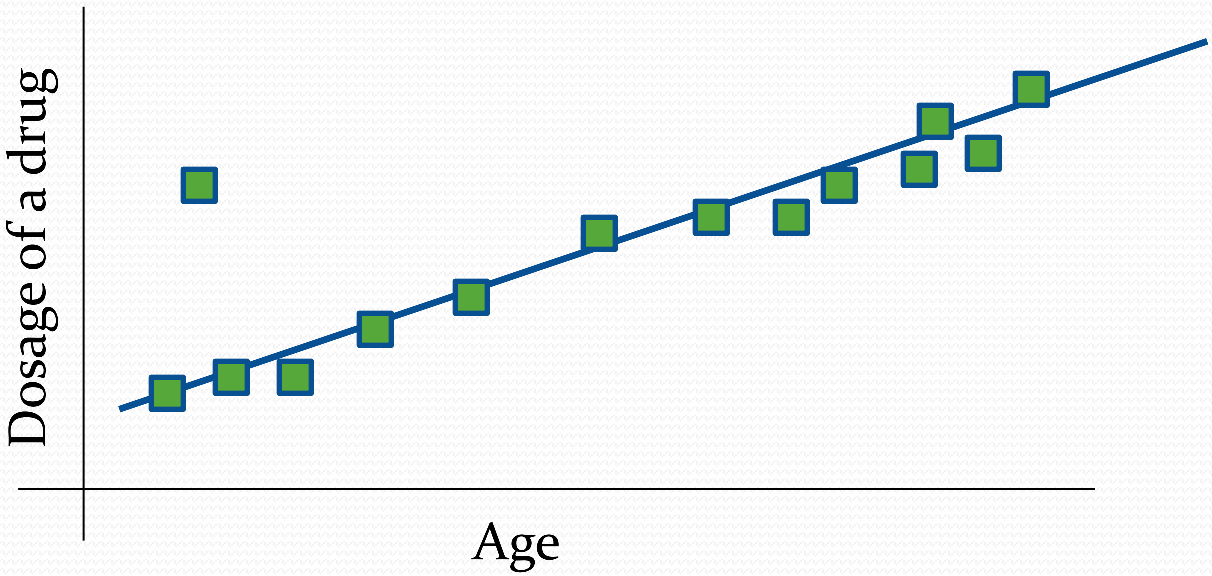
University of Pittsburgh

Funded by the National Library of Medicine: R21-LM009102 and R01–LM06696
and National Science Foundation: IIS–0325581

Anomaly Detection

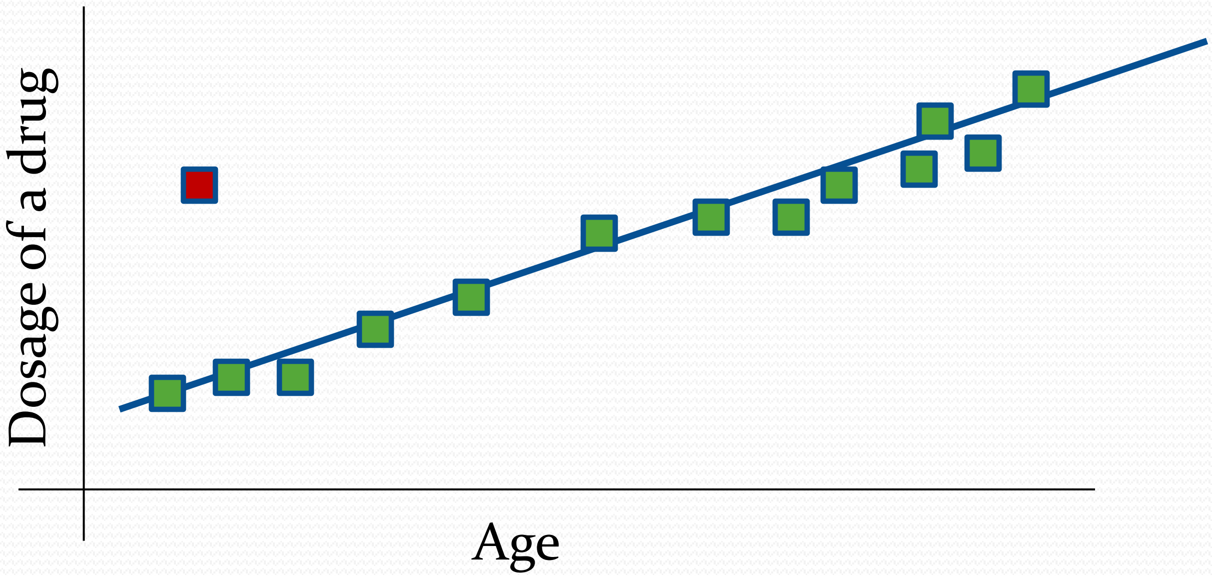
- Goal: Identify unusual patterns in data.
- Methods: from statistics and machine learning
- Contribution: conditional anomaly detection framework
- Application: medical error detection

Conditional Anomaly



- Patient electronic records have: demographics, conditions, labs, medications administered, procedures performed,…

Conditional Anomaly



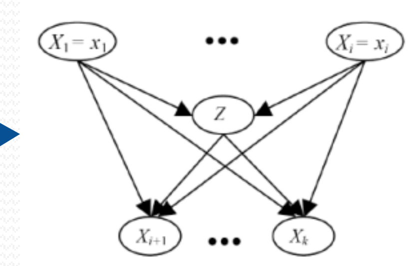
Assumption: Anomalies correspond to medical errors

"Medical errors account for 200 000 preventable deaths a year. "

(HealthGrades study, Wall Street Journal, July 27th 2004)

Medical Database

Group of similar patients



Model

d(Decisions | Description , Model) < α ?

Anomaly Call

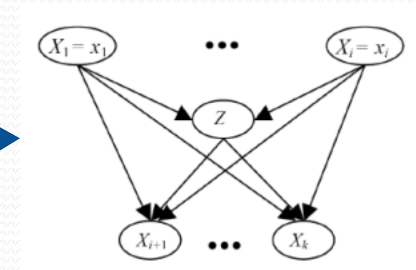
projection

Current patient record

Description (Contex) + Decision(s)

Medical Database

Group of similar patients



Model

d(Decisions | Description , Model) < α ?

Anomaly Call

Current patient record

$\underbrace{\text{Description (Contex)}}$ + $\underbrace{\text{Decision(s)}}$

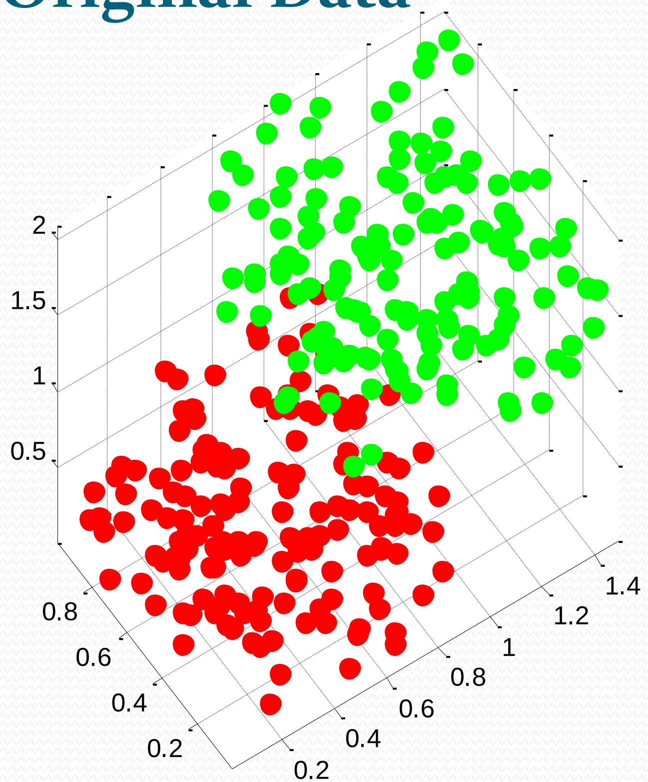
Selecting Similar Patients

- All other patients in the database
- Select only the closest patients
- What is a good distance metric?
  - Euclidean, Mahalanobis …
    - don't take into the account the decision variables

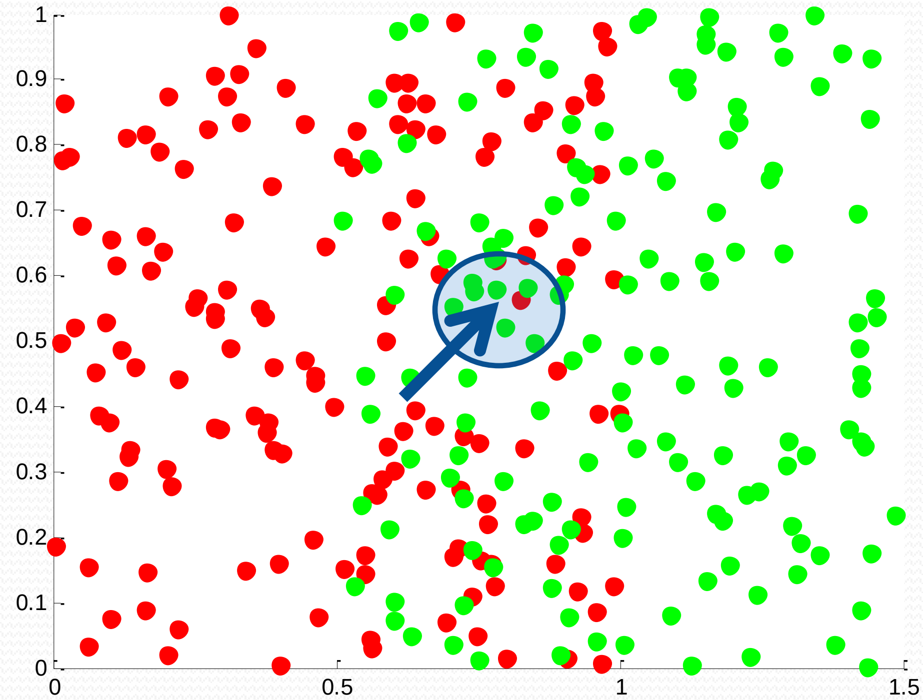
- Learn the metric which puts patients with the similar decisions closer together.
  - NCA (Goldberger et al. NIPS2004)

Neighborhood Component Analysis

Original Data



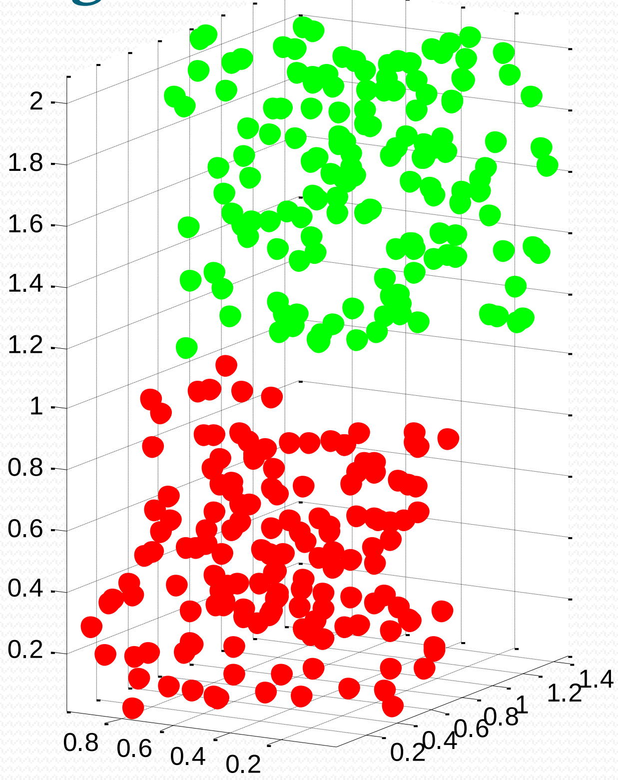
Initial Linear Projection



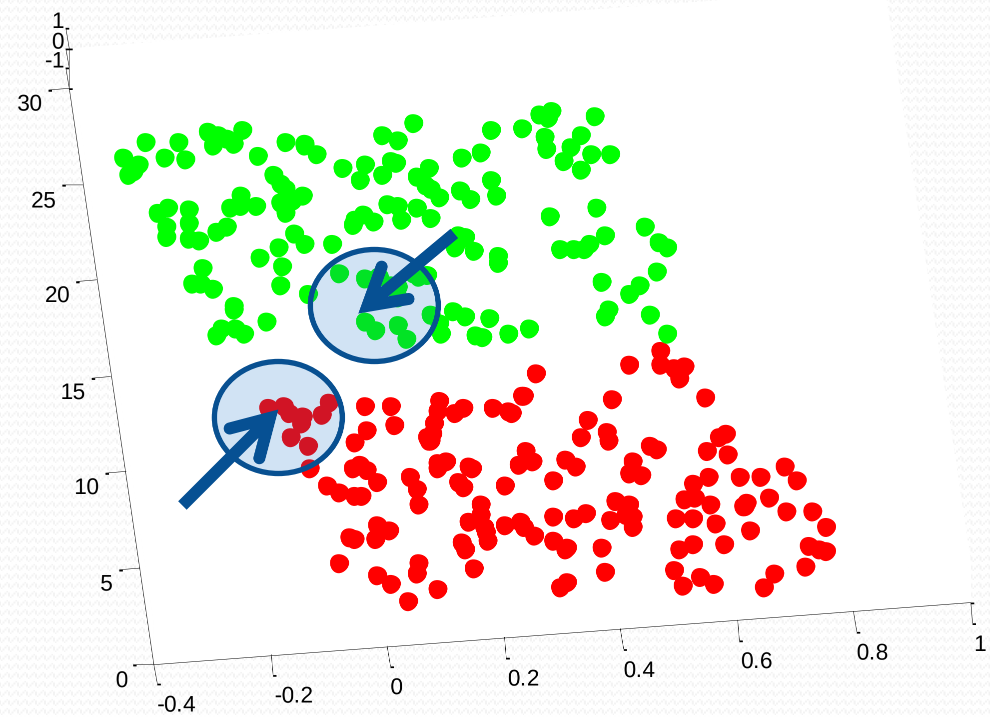
Goldberger et al. NIPS2004

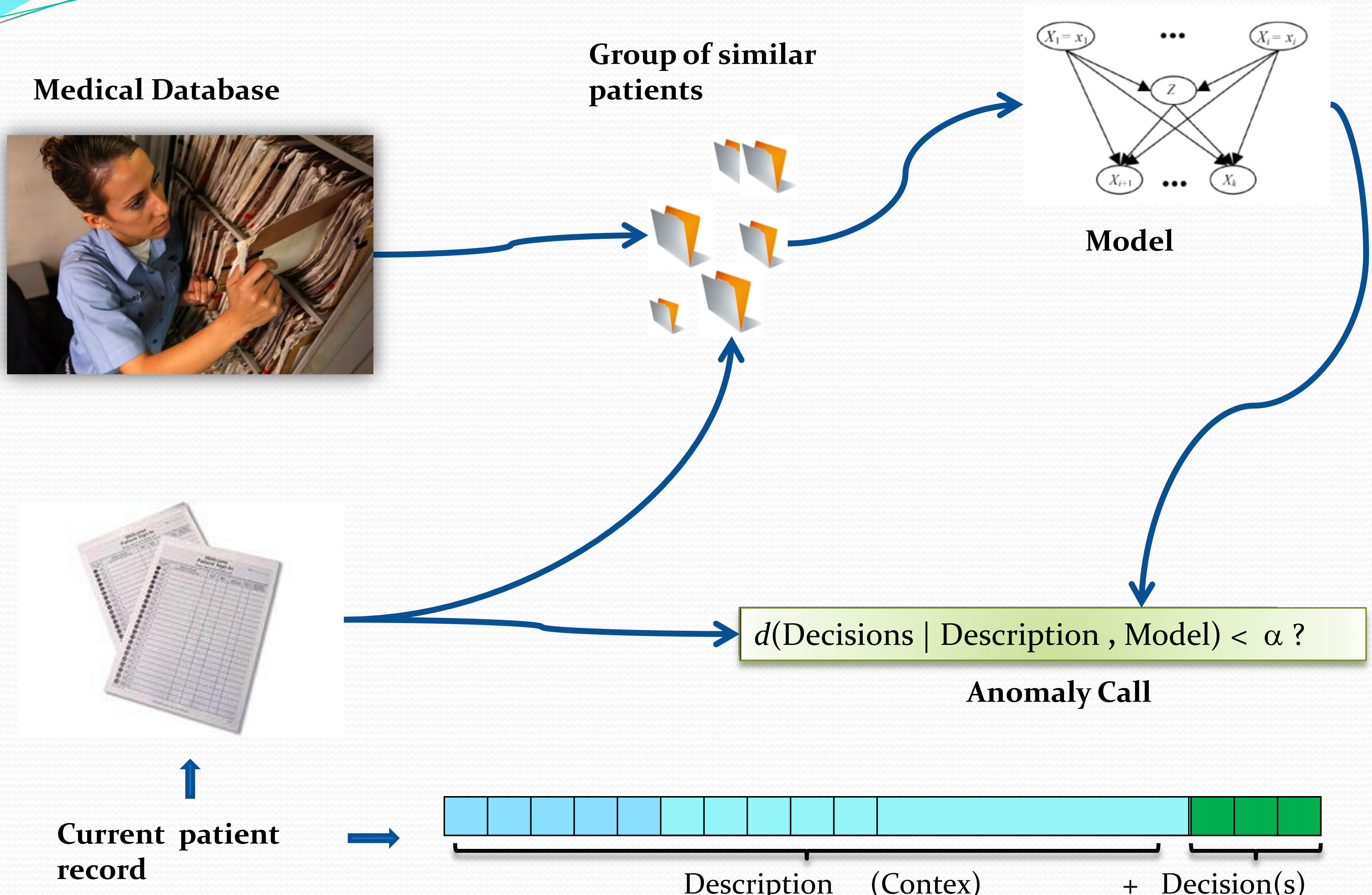
Neighborhood Component Analysis

Original Data



Learned Linear Projection

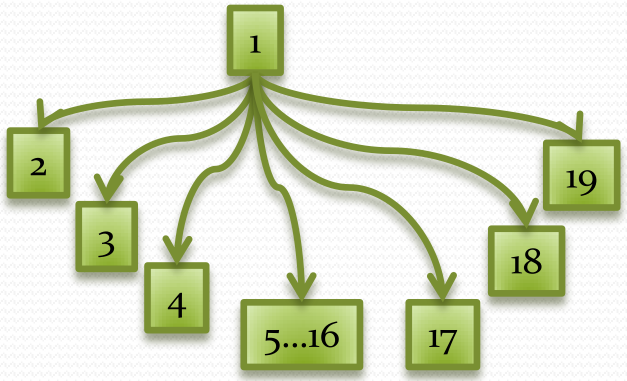



Medical Database
Group of similar patients
$X_1 = x_1$
$X_i = x_i$
Z
X_{i+1}
X_k
Model
d(Decisions | Description , Model) < α ?
Anomaly Call
Current patient record
Description (Contex)
+ Decision(s)

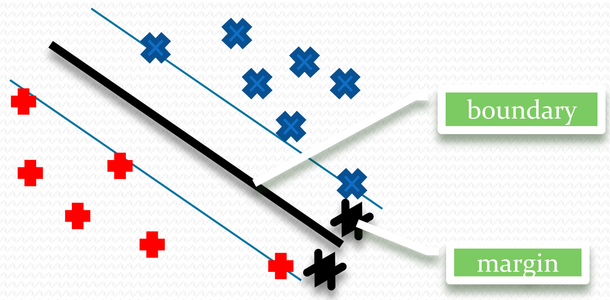
Learn Model/Build Projections

- Naïve Bayes Network

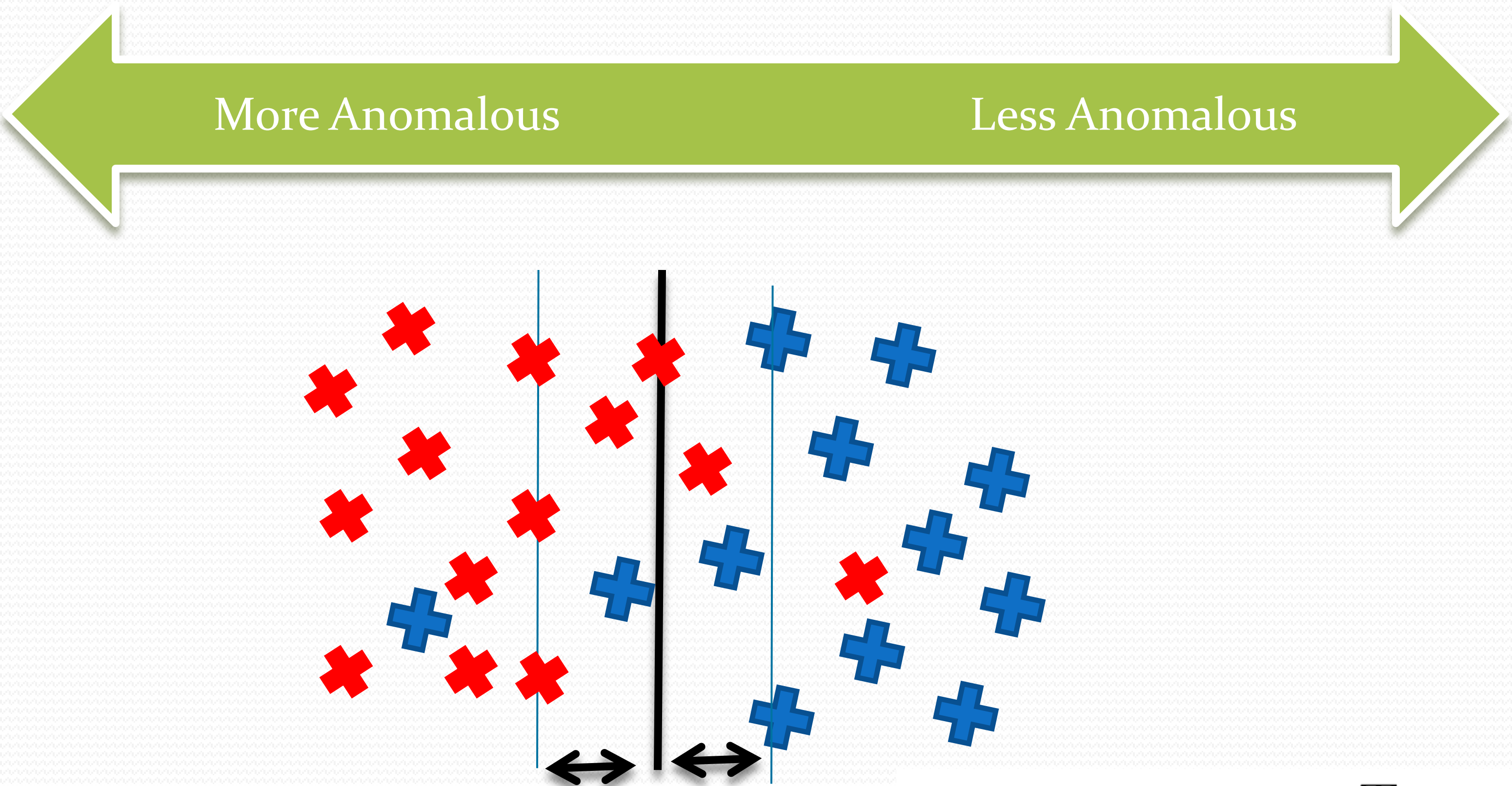
$$d(y|\mathbf{x}) = P(y|\mathbf{x})$$



- Support Vector Machines projections



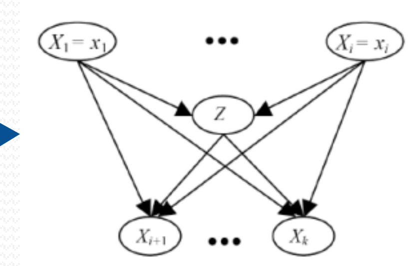
Support Vector Machines projections



$$d(y|\mathbf{x}) = y(\mathbf{w}^T\mathbf{x} + w_0)$$

Medical Database

Group of similar patients

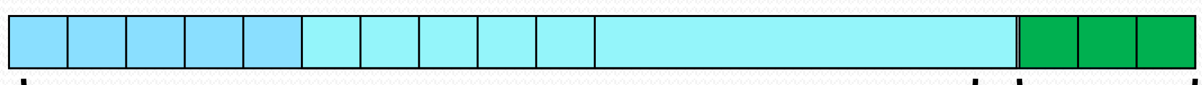


Model

$d(\text{Decisions} \mid \text{Description}, \text{Model}) < \alpha$?

Anomaly Call

Current patient record



Description (Contex) + Decision(s)

Experiments

PORT

- PORT dataset (Kapoor 1996)
- Patients diagnosed with the community acquired **pneumonia**

	Target attributes
X_1	Hospitalization
	Prediction attributes
	Demographic factors
X_2	Age > 50
X_3	Gender (male = true, female = false)
	Coexisting illnesses
X_4	Congestive heart failure
X_5	Cerebrovascular disease
X_6	Neoplastic disease
X_7	Renal disease
X_8	Liver disease
	Physical-examination findings
X_9	Pulse $\geq$ 125 / min
X_{10}	Respiratory rate $\geq$ 30 / min
X_{11}	Systolic blood pressure < 90 mm Hg
X_{12}	Temperature < 35 °C or $\geq$ 40 °C
	Laboratory and radiographic findings
X_{13}	Blood urea nitrogen $\geq$ 30 mg / dl
X_{14}	Glucose $\geq$ 250 mg / dl
X_{15}	Hematocrit < 30%
X_{16}	Sodium < 130 mmol / l
X_{17}	Partial pressure of arterial oxygen < 60 mm Hg
X_{18}	Arterial pH < 7.35
X_{19}	Pleural effusion

Experiments

PORT

- 2287 patient cases
- 19 binary attributes
- 100 evaluated by the panel of three physicians
- 23 anomalies

	Target attributes
X_1	Hospitalization
	Prediction attributes
	Demographic factors
X_2	Age > 50
X_3	Gender (male = true, female = false)
	Coexisting illnesses
X_4	Congestive heart failure
X_5	Cerebrovascular disease
X_6	Neoplastic disease
X_7	Renal disease
X_8	Liver disease
	Physical-examination findings
X_9	Pulse ≥ 125 / min
X_{10}	Respiratory rate ≥ 30 / min
X_{11}	Systolic blood pressure < 90 mm Hg
X_{12}	Temperature $< 35\,^{\circ}$C or $\geq 40\,^{\circ}$C
	Laboratory and radiographic findings
X_{13}	Blood urea nitrogen ≥ 30 mg / dl
X_{14}	Glucose ≥ 250 mg / dl
X_{15}	Hematocrit $< 30\%$
X_{16}	Sodium < 130 mmol / l
X_{17}	Partial pressure of arterial oxygen < 60 mm Hg
X_{18}	Arterial pH < 7.35
X_{19}	Pleural effusion

Experiments

PORT

- Goal: Detect whether the decision of hospitalization is *anomalous*, conditioning on the description variables

	Target attributes
X_1	Hospitalization
	Prediction attributes
	Demographic factors
X_2	Age > 50
X_3	Gender (male = true, female = false)
	Coexisting illnesses
X_4	Congestive heart failure
X_5	Cerebrovascular disease
X_6	Neoplastic disease
X_7	Renal disease
X_8	Liver disease
	Physical-examination findings
X_9	Pulse ≥ 125 / min
X_{10}	Respiratory rate ≥ 30 / min
X_{11}	Systolic blood pressure < 90 mm Hg
X_{12}	Temperature $< 35\,^{\circ}$C or $\geq 40\,^{\circ}$C
	Laboratory and radiographic findings
X_{13}	Blood urea nitrogen ≥ 30 mg / dl
X_{14}	Glucose ≥ 250 mg / dl
X_{15}	Hematocrit $< 30\%$
X_{16}	Sodium < 130 mmol / l
X_{17}	Partial pressure of arterial oxygen < 60 mm Hg
X_{18}	Arterial pH < 7.35
X_{19}	Pleural effusion

Experiments

HIT

- Post-surgical cardiac patients treated at University of Pittsburgh Medical Center
- In risk of Heparin Induced Thrombocytopenia
- HPF4 is a specific test to confirm HIT

HPF4 lab test ordered

Platelet count
Previous Platelet count
Platelet drop from the last value
Platelet % drop from the last value
Platelet drop from the last value
Platelet % drop from the first value
Platelet drop from the nadir value
Platelet % drop from the nadir value

…
Hemoglobin count
Previous Hemoglobin count
Hemoglobin drop from the last value
Hemoglobin % drop from the last value
Hemoglobin drop from the last value

…
Patient on Heparin
Time on Heparin
Transfusion in last 28 hours.

Experiments

HIT

- 4273 patients
- 45767 patients states
- unbalanced (only 271 positive)
- 45 attributes
- 60 evaluated by the pharmacy expert
- 28 anomalies

HPF4 lab test ordered

Platelet count
Previous Platelet count
Platelet drop from the last value
Platelet % drop from the last value
Platelet drop from the last value
Platelet % drop from the first value
Platelet drop from the nadir value
Platelet % drop from the nadir value
...
Hemoglobin count
Previous Hemoglobin count
Hemoglobin drop from the last value
Hemoglobin % drop from the last value
Hemoglobin drop from the last value
...
Patient on Heparin
Time on Heparin
Transfusion in last 28 hours.

Experiments

HIT

- Goal: : Detect whether the order of HP Factor 4 lab test was *anomalous*, conditioning on the description variables

HPF4 lab test ordered

Platelet count
Previous Platelet count
Platelet drop from the last value
Platelet % drop from the last value
Platelet drop from the last value
Platelet % drop from the first value
Platelet drop from the nadir value
Platelet % drop from the nadir value
…
Hemoglobin count
Previous Hemoglobin count
Hemoglobin drop from the last value
Hemoglobin % drop from the last value
Hemoglobin drop from the last value
…
Patient on Heparin
Time on Heparin
Transfusion in last 28 hours.

Evaluation

- Algorithm catches many anomalies
  - high sensitivity
- Algorithm's predictions are accurate
  - high specificity

- Combine sensitivity and specificity for various detection thresholds to ROC
- Specificity >= 95%
  (at most 1 false alarm in 20 normal cases)

Results

PORT dataset

metric	model	#cases	area
any	NB	2286	11.6 %
metric	**model**	**#cases**	**area**
NCA	NB	40	16.8 %
Mahalanobis	NB	40	17.6 %
RCA	NB	40	17.6 %
Euclidean	NB	40	16.4 %

BASELINE

metric	model	#cases	area
any	SVM	2286	12.1 %
metric	**model**	**#cases**	**area**
NCA	SVM	40	19.0 %
Mahalanobis	SVM	40	11.9 %
RCA	SVM	40	10.4 %
Euclidean	SVM	40	11.2 %

BEST

Results

HIT dataset

metric	model	#cases	area
any	NB	45766	3.0 %
metric	**model**	**#cases**	**area**
NCA	NB	100	30.7 %
Mahalanobis	NB	100	16.2 %
RCA	NB	100	16.2 %
Euclidean	NB	100	12.0 %

BASELINE

BEST

metric	model	#cases	area
any	SVM	45766	21.9 %
metric	**model**	**#cases**	**area**
NCA	SVM	100	30.4 %
Mahalanobis	SVM	100	18.6 %
RCA	SVM	100	18.6 %
Euclidean	SVM	100	28.9 %

Conclusion

- Conditional anomaly framework
  - Can discover potential medical errors
- Selection of closest patients
  - Models tuned to the individual patient
- Metric learning
  - Lowers the influence of irrelevant data

Current/Future Work

- Automatic population size selection
- Multiple decisions
- Ability to work with **many** more features
- Sensible temporal abstractions
- More scalable metric learning methods

Results

MODEL	METRIC	SELECTION	RESULT	
Naïve Bayes	any	ALL	11.6%	BASELINE
	Euclidean	CLOSEST 40	16.4%	
	Learned Metric	CLOSEST 40	16.8%	
Probability from the Distance Metric	Euclidean	ALL	8.0%	
	Euclidean	CLOSEST 40	8.0%	
	Learned Metric	ALL	18.0%	
	Learned Metric	CLOSEST 40	20.2%	BEST

Conclusion: Two-fold improvement over baseline.

Neighborhood Component Analysis

$$p_{ij} = \frac{\exp(-\|Ax_i - Ax_j\|^2)}{\sum_{k \neq i} \exp(-\|Ax_i - Ax_k\|^2)} \quad , \quad p_{ii} = 0$$

$$f(A) = \sum_i \sum_{j \in C_i} p_{ij}$$

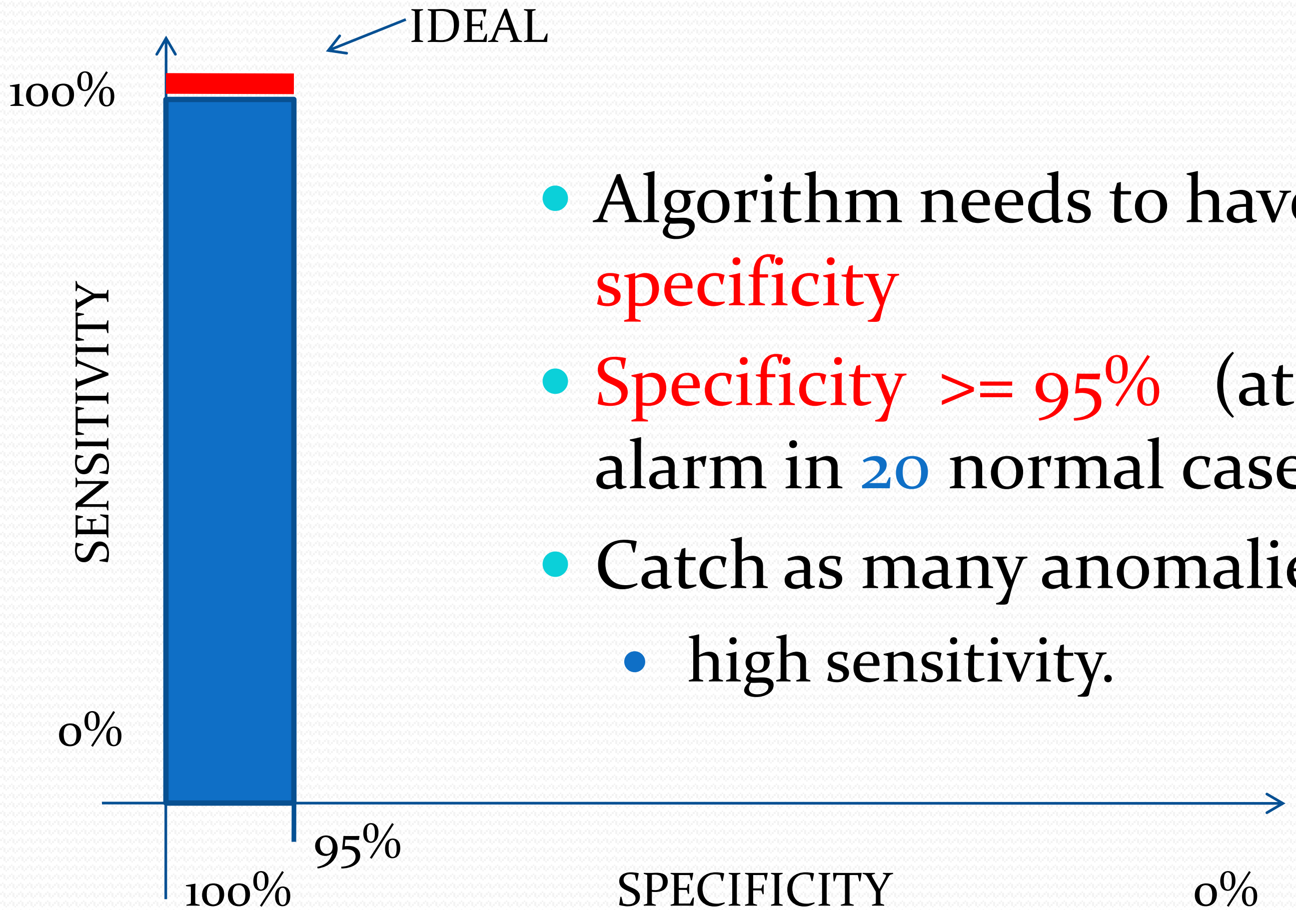
Goldberger et al. NIPS2004

Neighborhood Component Analysis

$$\|Ax_i - Ax_j\|^2$$

$$\sum_{j \in C_i} p_{ij}$$

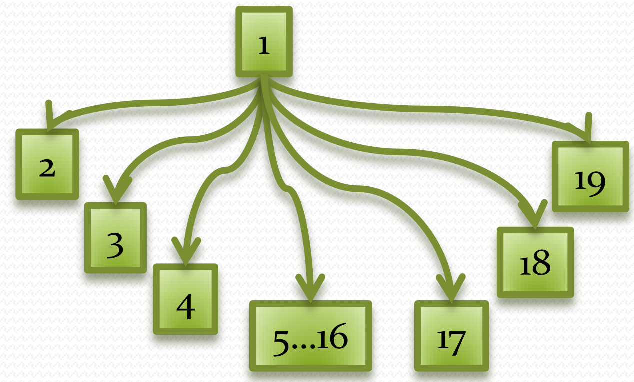
Evaluation



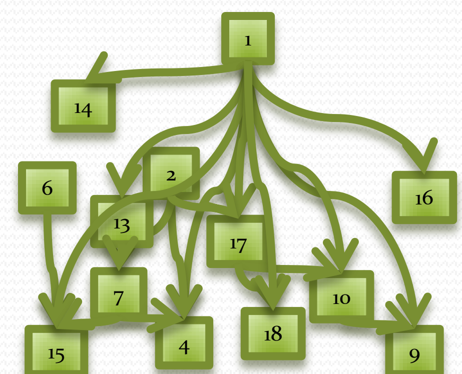
- Algorithm needs to have high specificity
- Specificity >= 95% (at most 1 false alarm in 20 normal cases)
- Catch as many anomalies
  - high sensitivity.

Learn Probabilistic Model

- Bayesian Network with
  Fixed structure



- Learn the Bayesian Network structure
  and parameters from the data



Eaton & Murphy, UAI 2007

Results

MODEL	METRIC	SELECTION	RESULT
Naïve Bayes	any	ALL	11.6%
	Euclidean	CLOSEST 40	16.4%
	Learned Metric	CLOSEST 40	16.8%
Probability from the Distance Metric	Euclidean	ALL	8.0%
	Euclidean	CLOSEST 40	8.0%
	Learned Metric	ALL	18.0%
	Learned Metric	CLOSEST 40	20.2%
Learn Bayes Network Structure and Parameters	any	ALL	13.8%
	Euclidean	CLOSEST 40	17.8%
	Learned Metric	CLOSEST 40	26.4%

BASELINE

BEST

Conclusion: Two-fold improvement over baseline.

Results

MODEL	METRIC	SELECTION	RESULT	
Naïve Bayes	any	ALL	11.6%	BASELINE
	Euclidean	CLOSEST 40	16.4%	
	Learned Metric	CLOSEST 40	16.8%	
Learn Bayes Network Structure and Parameters	any	ALL	13.8%	
	Euclidean	CLOSEST 40	17.8%	
	Learned Metric	CLOSEST 40	26.4%	BEST

Conclusion: Two-fold improvement over baseline.

Results

HIT dataset

metric	**model**	**AU–ROC**	**AU–PR**
any	NB	57.8 %	50.9 %
metric	**model**	**AU–ROC**	**AU–PR**
NCA	NB	90.6 %	90.8 %
Mahalanobis	NB	84.9 %	80.5 %
RCA	NB	84.9 %	80.5 %
Euclidean	NB	85.3 %	78.9 %

metric	**model**	**AU–ROC**	**AU–PR**
any	SVM	87.3 %	86.6 %
metric	**model**	**AU–ROC**	**AU–PR**
NCA	SVM	90.8 %	90.6 %
Mahalanobis	SVM	87.6 %	82.9 %
RCA	SVM	87.6 %	82.9 %
Euclidean	SVM	90.4 %	90.8 %