Generalized Inverse Classification

SDM '17

Michael T. Lash[1], Qihang Lin[2], W. Nick Street[2], Jennifer G. Robinson[3], and Jeffrey Ohlmann[2]

[1]Department of Computer Science, [2]Deparment of Management Sciences, [3]Department of Epidemiology

www.michaeltlash.com

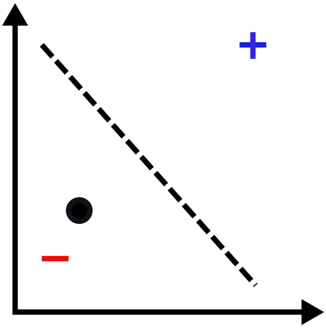
The University of Iowa

What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.

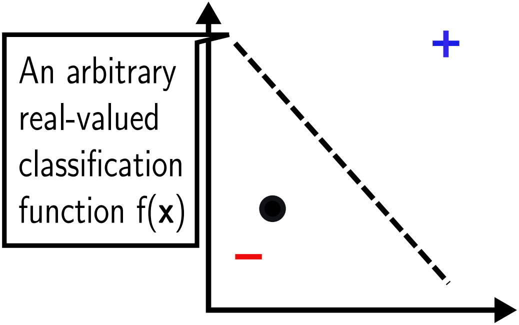
What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.



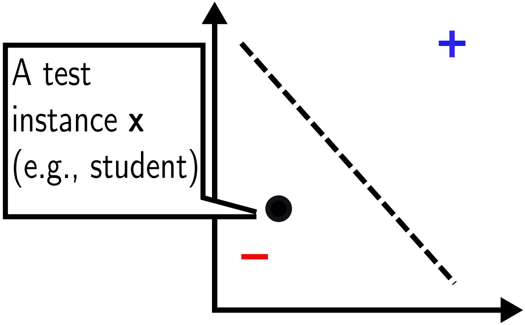
What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.



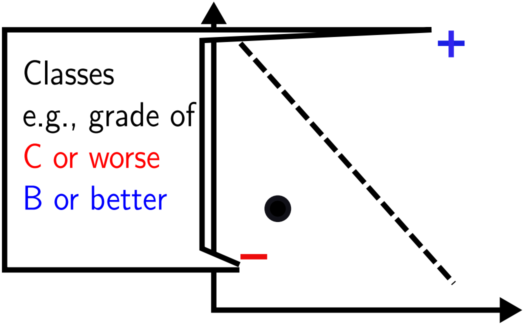
What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.



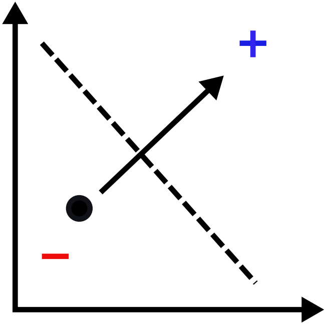
What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.



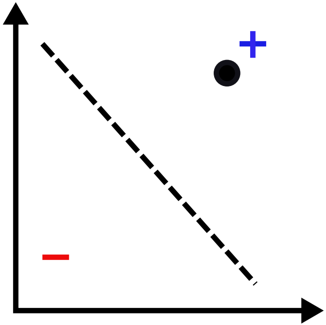
What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.



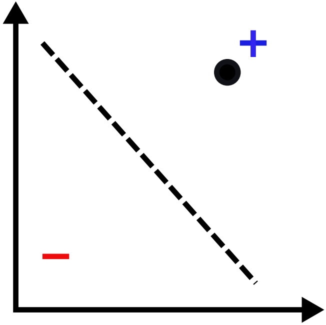
What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.



What is *inverse classification*?

- The process of making meaningful perturbations to a test instance such that the probability of a desirable outcome is maximized.



What about the meaningful part of the definition?

Meaningful Perturbations

Well...lets visit some past work!

- Michael T. Lash, Qihang Lin, W. Nick Street, and Jennifer G. Robinson, "A budget-constrained inverse classification framework for smooth classifiers", *arXiv preprint arXiv:1605.09068*, submitted.

Meaningful Perturbations

$$\min_{\mathbf{x}} f(\mathbf{x})$$

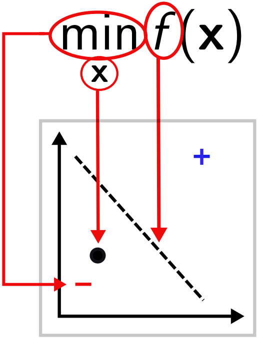
Begin with a basic formulation.

Meaningful Perturbations

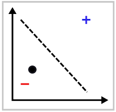
$$\min_{\mathbf{x}} f(\mathbf{x})$$

[For now] A differentiable classification function e.g., logistic regression, SVM, ANN, etc.

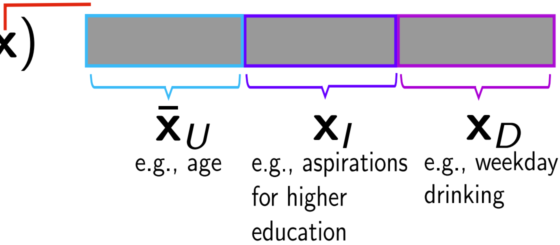
Meaningful Perturbations



Meaningful Perturbations

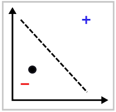


$$\min_{\mathbf{x}} f(\mathbf{x})$$

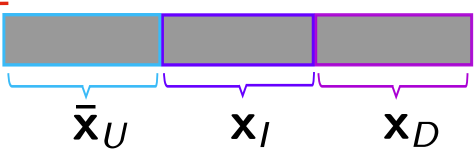


Segment features.

Meaningful Perturbations



$$\min_{\mathbf{x}} f(\mathbf{x})$$



$$H(\bar{\mathbf{x}}_U, \mathbf{x}_D)$$

Some regressor

$$H : \mathbb{R}^{|U|+|D|} \rightarrow \mathbb{R}^{|I|}$$

Estimate indirectly changeable.

Meaningful Perturbations

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

$$\mathrm{l}_i \leq x_i \leq u_i \text{ for } i \in D$$

Update objective function. Add constraints.

Meaningful Perturbations

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

$$l_i \leq x_i \leq u_i \text{ for } i \in D$$

$$\phi(\mathbf{z}) = \sum_{i \in D} c_i^+ (z_i)_+ + c_i^- (z_i)_-$$

Cost-change function

Meaningful Perturbations

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

Budget

$$\mathrm{l}_i \leq x_i \leq u_i \text{ for } i \in D$$

$$\phi(\mathbf{z}) = \sum_{i \in D} c_i^+ (z_i)_+ + c_i^- (z_i)_-$$

Meaningful Perturbations

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

$$\mathrm{l}_i \leq x_i \leq u_i \text{ for } i \in D$$

Bounds

1. Relax assumptions about $f(\cdot)$.

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

$$l_i \leq x_i \leq u_i \text{ for } i \in D$$

Assume:

- f($\mathbf{z}$) differentiable
- $\|\nabla f(\mathbf{z}) - \nabla f(\mathbf{z'})\| \leq L \|\nabla f(\mathbf{z}) - \nabla f(\mathbf{z'})\|$ $\mathbf{z}, \mathbf{z'} \in \mathbb{R}^{|D|}$

1. Relax assumptions about $f(\cdot)$.

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$
$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$
$$l_i \leq x_i \leq u_i \text{ for } i \in D$$

Assume:

- f($\mathbf{z}$) differentiable
- $\|\nabla f(\mathbf{z}) - \nabla f(\mathbf{z'})\| \leq L \|\nabla f(\mathbf{z}) - \nabla f(\mathbf{z'})\|$ $\mathbf{z}, \mathbf{z'} \in \mathbb{R}^{|D|}$

1. Relax assumptions about $f(\cdot)$.

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

$$l_i \leq x_i \leq u_i \text{ for } i \in D$$

Assume:

- f:$\mathbb{R}^p \to \mathbb{R}$

Generalized inverse classification

Main Contributions

1. Relax assumptions about $f(\cdot)$.
2. Quadratic cost-change function.

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

$$\text{l}_i \leq x_i \leq u_i \text{ for } i \in D$$

Main Contributions

1. Relax assumptions about $f(\cdot)$.
2. Quadratic cost-change function.

$$\min_{\mathbf{x}_D} f\left(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D\right)$$
$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$
$$l_i \leq x_i \leq u_i \text{ for } i \in D$$

$$\phi(\mathbf{z}) = \sum_{i \in D} c_i^+ (z_i)_+ + c_i^- (z_i)_-$$

1. Relax assumptions about $f(\cdot)$.
2. Quadratic cost-change function.

$$\min_{\mathbf{x}_D} f(\bar{\mathbf{x}}_U, H(\bar{\mathbf{x}}_U, \mathbf{x}_D), \mathbf{x}_D)$$

$$\text{s.t. } \phi(\mathbf{x}_D - \bar{\mathbf{x}}_D) \leq B$$

$$l_i \leq x_i \leq u_i \text{ for } i \in D$$

$$\phi(\mathbf{z}) = \sum_{i \in D} c_i^+ (z_i)_+^2 + c_i^- z(_i)_-^2$$

Main Contributions

1. Relax assumptions about $f(\cdot)$.
2. Quadratic cost-change function.
3. Three real-valued heuristic optimization methods and two sensitivity analysis-based optimization methods.
    * Projection operator to maintain feasibility.

Optimization Methodology

Heuristic

- Hill Climbing + Local Search (HC+LS)
- Genetic Algorithm (GA)
- Genetic Algorithm + Local Search (GA+LS)

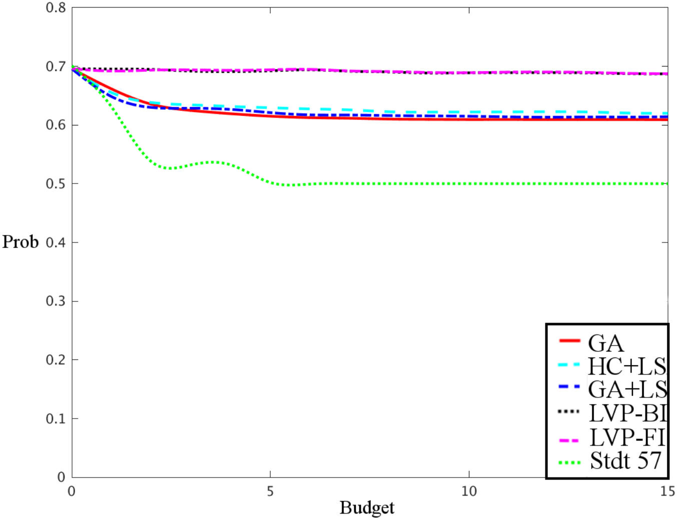
Sensitivity Analysis

- Local Variable Perturbation – First Improvement (LVP-FI)
- Local Variable Perturbation – Best Improvement (LVP-BI)

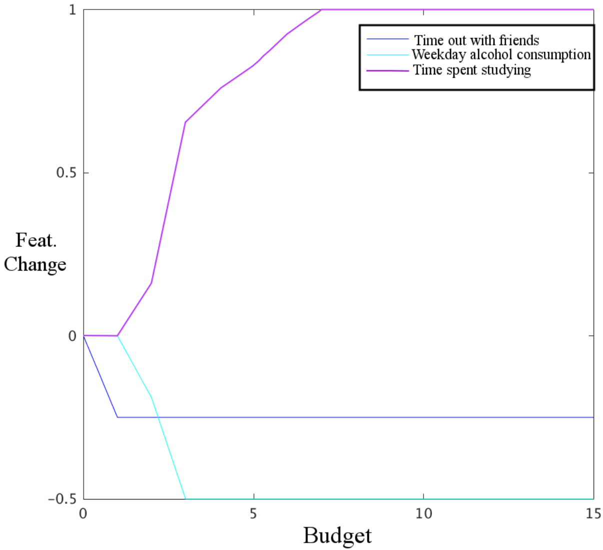
Experiment Decisions and Data

- $f(\cdot)$: Random forest
- $H(\cdot)$: Kernel regression
- Dataset 1: Student Performance (UCI Machine Learning Repository).
- Dataset 2: ARIC
- One f for optimization, separate f for heldout evaluation.

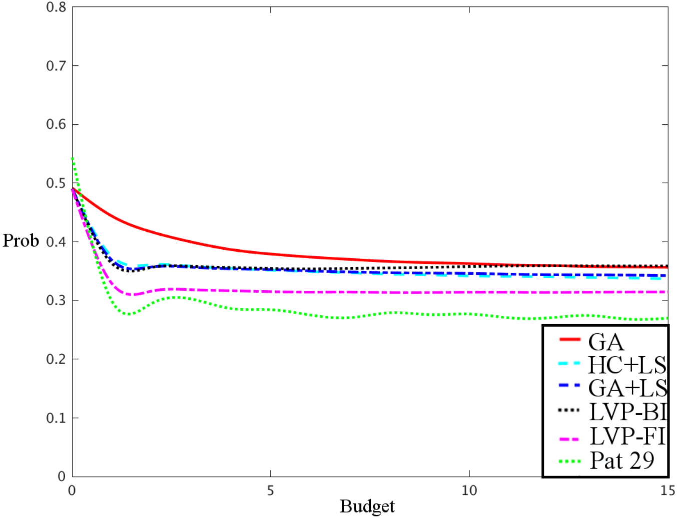
Results: Student Performance



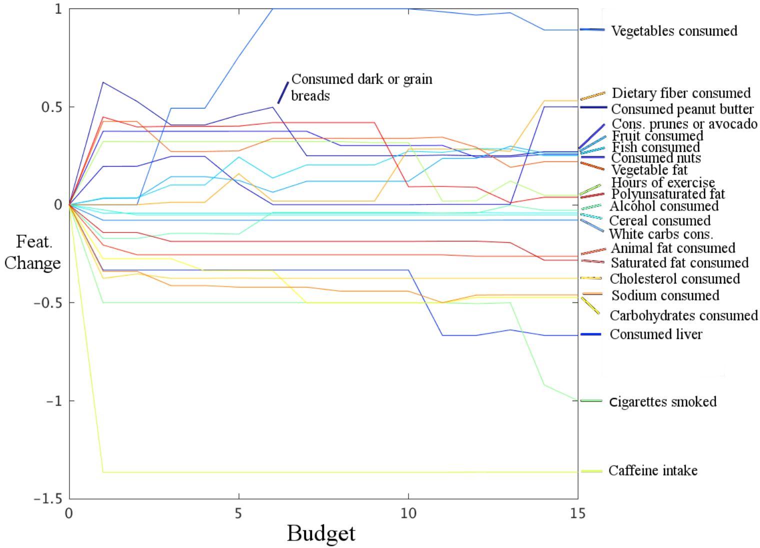
Results: Student Performance



Results: ARIC



Results: ARIC



Need sparsity constraints

Conclusions

- Generalized Inverse Classification: can use virtually any learned f (as shown by experiments w/ Random Forest classifier).
- Our proposed methods were successful, although this varied by dataset.

Generalized Inverse Classification

SDM '17

Michael T. Lash[1], Qihang Lin[2], W. Nick Street[2], Jennifer G. Robinson[3], and Jeffrey Ohlmann[2]

[1]Department of Computer Science, [2]Deparment of Management Sciences, [3]Department of Epidemiology

www.michaeltlash.com

The University of Iowa

Causality and Inverse Classification

Causality and Inverse Classification

- Yes! ....

Causality and Inverse Classification

- Yes! ....
- What we're doing:
  1. Imposing our own causal structure (DAG).
  2. We're not taking the usual counterfactual approach.

Causality and Inverse Classification

- Yes! ....
- What we're doing:
    1. Imposing our own causal structure (DAG).
    2. We're not taking the usual counterfactual approach.
- Future work will focus on incorporating causal methodology...