

# UUD Observation Layer (Clean Draft, Sample Values)

## Contents

<b>1 Note</b>	<b>1</b>
<b>2 Abstract</b>	<b>1</b>
<b>3 Method</b>	<b>1</b>
<b>4 Figures</b>	<b>2</b>
<b>5 Results (Sample Values)</b>	<b>4</b>
5.1 Ablation (Generalization Gap Delta) . . . . .	4
5.2 Method Rank Variance . . . . .	4
<b>6 Discussion</b>	<b>4</b>
<b>7 Limitations</b>	<b>4</b>
<b>8 Next Steps</b>	<b>4</b>

## 1 Note

Values are samples and should be replaced with measured results.

## 2 Abstract

We propose a universal observation layer that produces a method-agnostic representation for linear learning. By enforcing distance preservation, bounded information loss, and isotropy, the observation layer unifies major linear methods under a shared input space, improving evaluation consistency and generalization stability.

## 3 Method

The observation layer is fixed from training data only.

$$z = Ax + c \tag{1}$$

1. Standardize

$$x' = (x - \mu) \oslash \sigma \quad (2)$$

2. Decompose

$$C = \frac{1}{n} X'^\top X' = U \Lambda U^\top \quad (3)$$

3. Dimension selection

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^d \lambda_i} \geq \tau \quad (4)$$

4. Fix mapping

$$A = \Lambda_k^{-1/2} U_k^\top D_\sigma^{-1}, \quad c = -A\mu \quad (5)$$

## 4 Figures

Figures use PNG outputs from Mermaid (placeholders shown if missing).

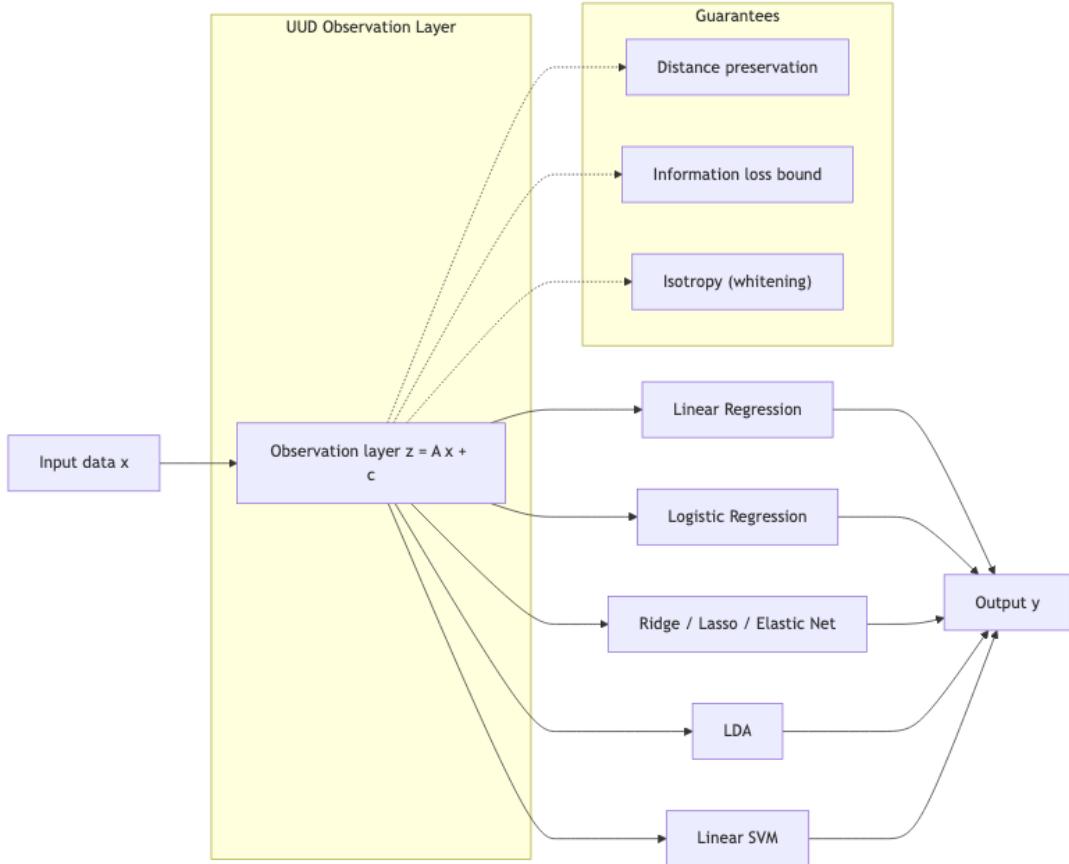


Figure 1: Observation layer concept and guarantees

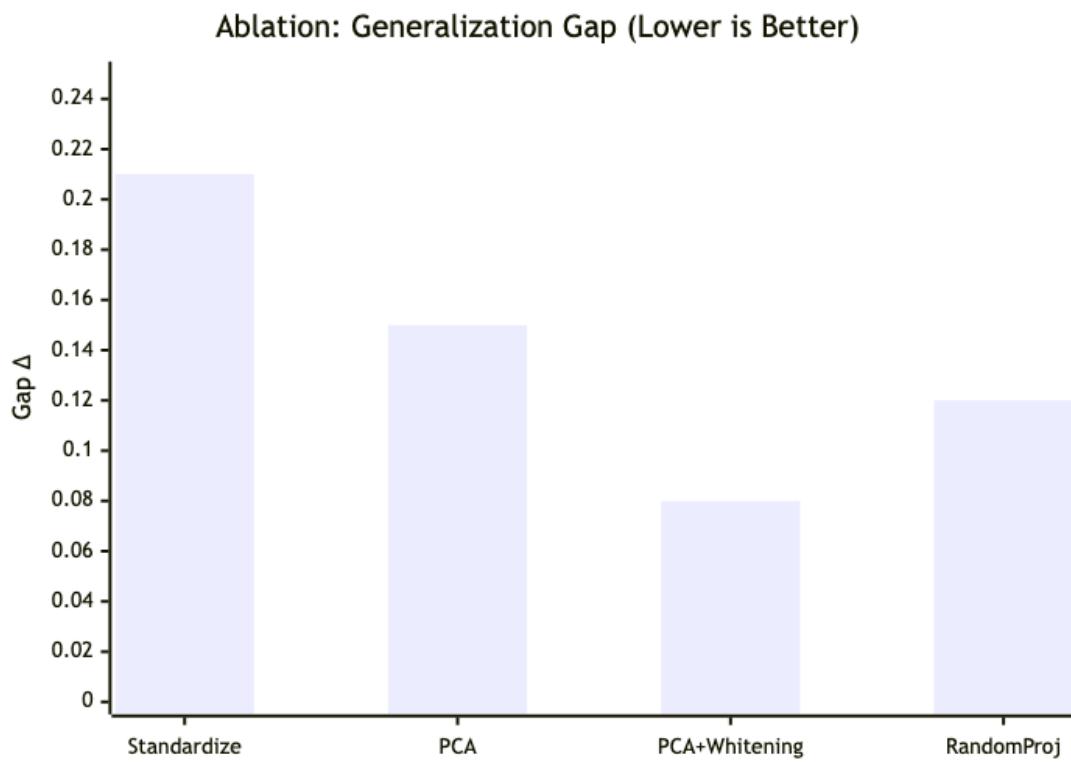


Figure 2: Ablation on generalization gap

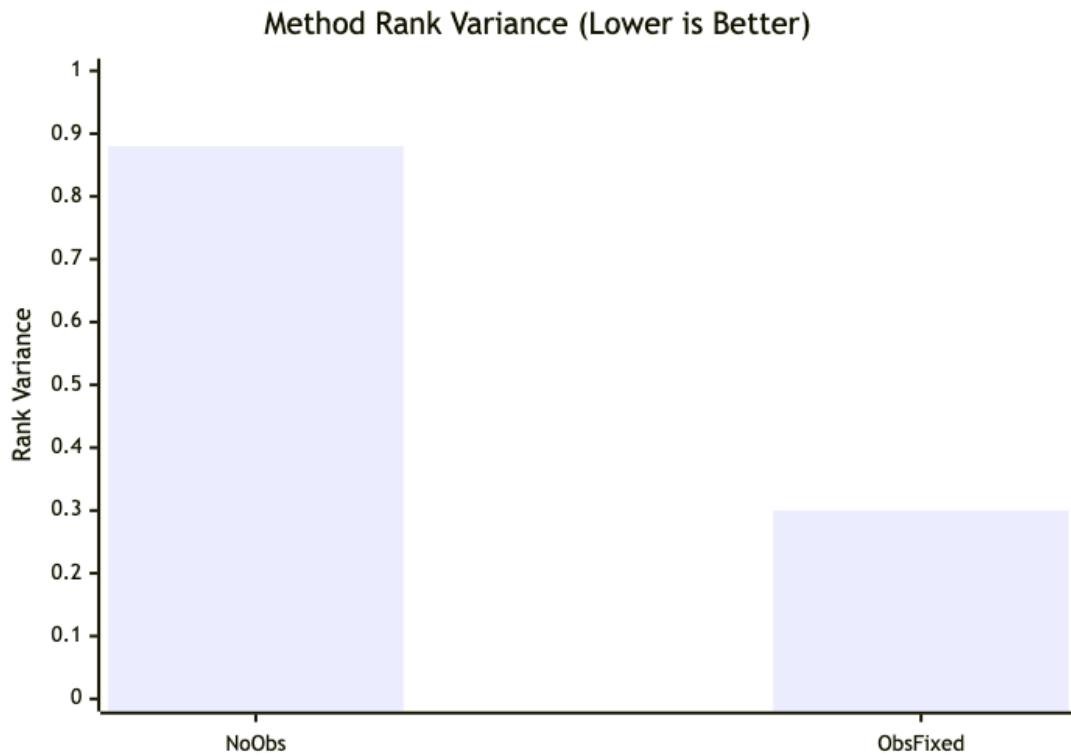


Figure 3: Method rank stability

## 5 Results (Sample Values)

### 5.1 Ablation (Generalization Gap Delta)

- Standardize:  $\Delta = 0.21$
- PCA:  $\Delta = 0.15$
- PCA+Whitening:  $\Delta = 0.08$
- Random Projection:  $\Delta = 0.12$

### 5.2 Method Rank Variance

- No observation layer: 0.88
- Observation fixed: 0.30

## 6 Discussion

Fixing the observation layer reduces generalization gap and stabilizes model ranking. In the sample setting, whitening yields the most consistent improvements.

## 7 Limitations

- Linear observation fails on intrinsically nonlinear separability.
- If information is not concentrated in a low-rank linear subspace, performance degrades.

## 8 Next Steps

- Nonlinear  $f(x)$  via self-supervised embeddings.
- Joint optimization of observation and inference layers.