Framework for optimal use of production data in NIR calibrations

Jacob Søgaard Larsen PhD Student

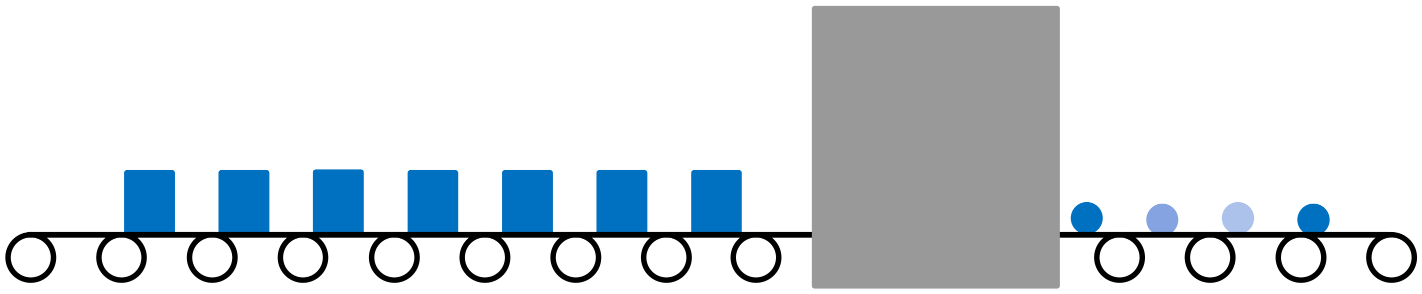
Main Supervisor: Line Clemmensen (DTU Compute)
Co-supervisors: Bjarne Ersbøll (DTU Compute)
Anders Stockmarr (DTU Compute)
Anders Larsen (Q-Interline)
Thomas Skov (KU-FOOD)

DTU Compute
Department of Applied Mathematics and Computer Science

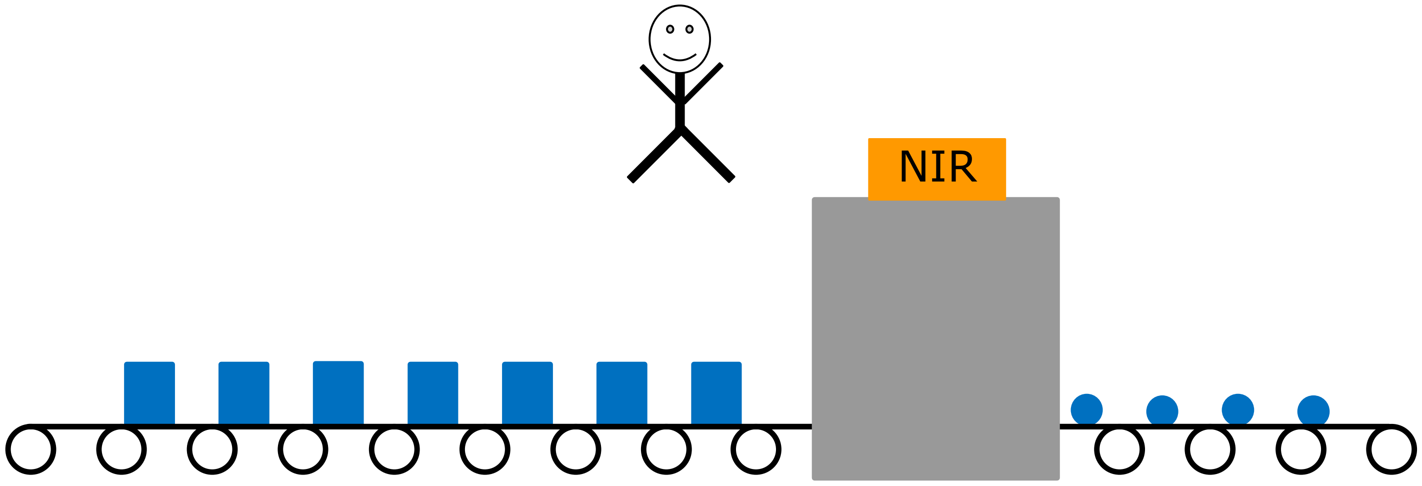
Agenda

- Motivation
- Linear Joint Trained Framework
- What happens?
- Example - Transfer of calibration
- Future perspectives

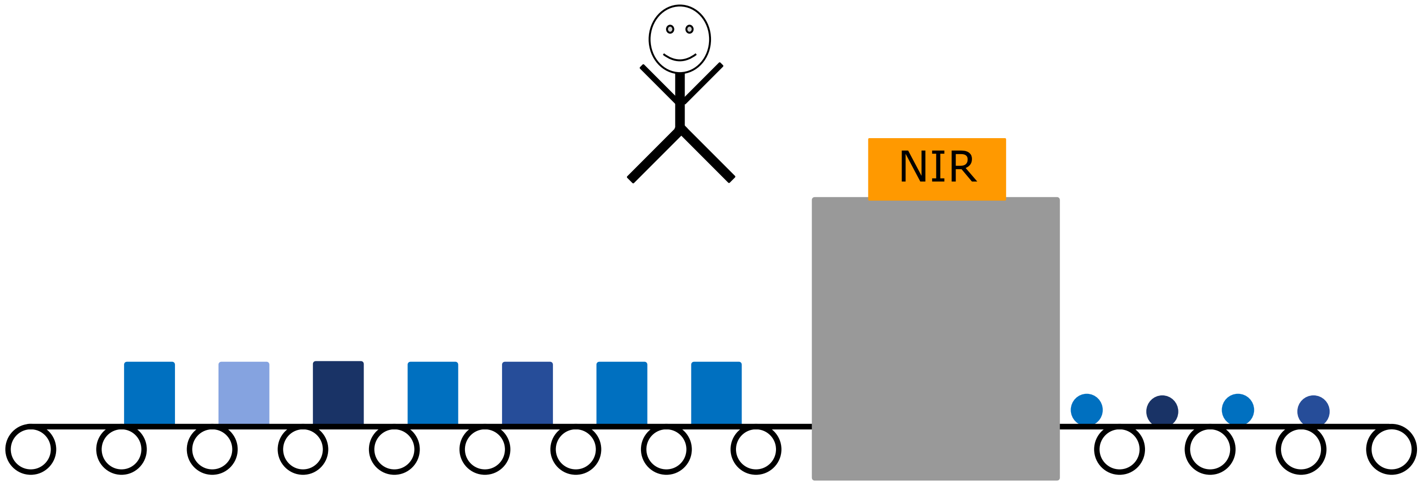
Motivation



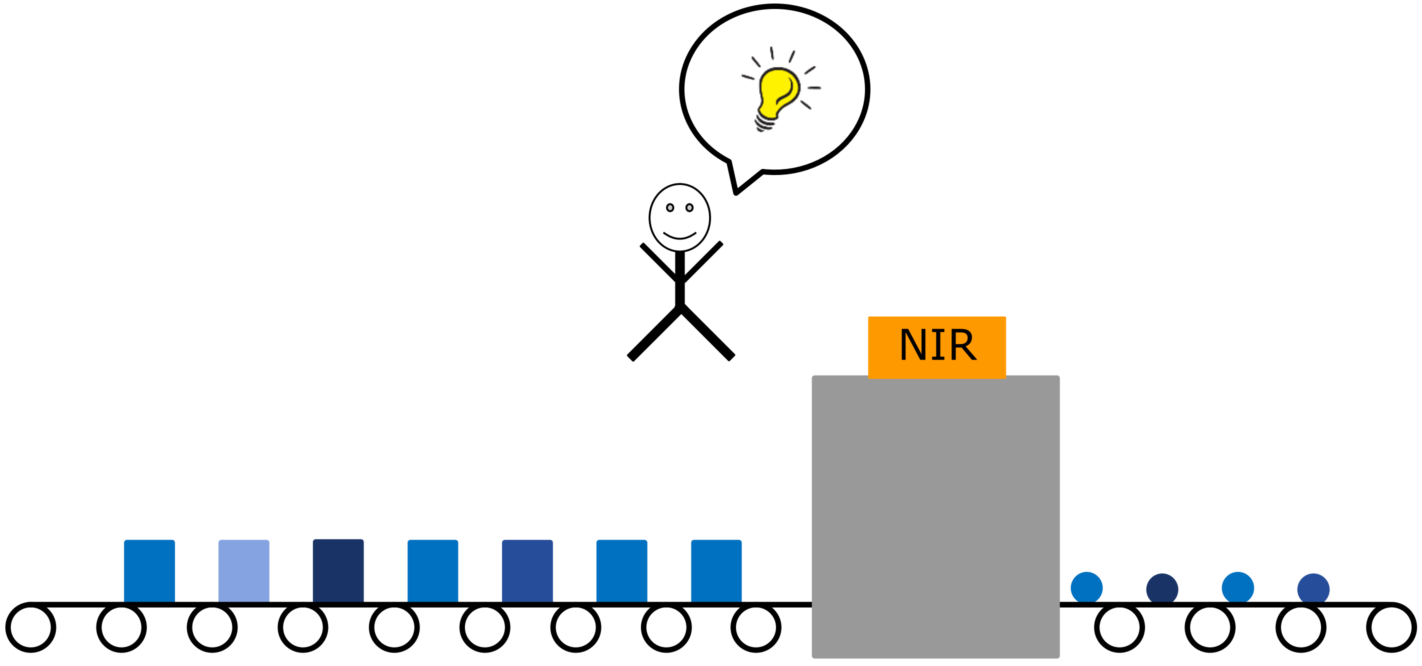
Motivation



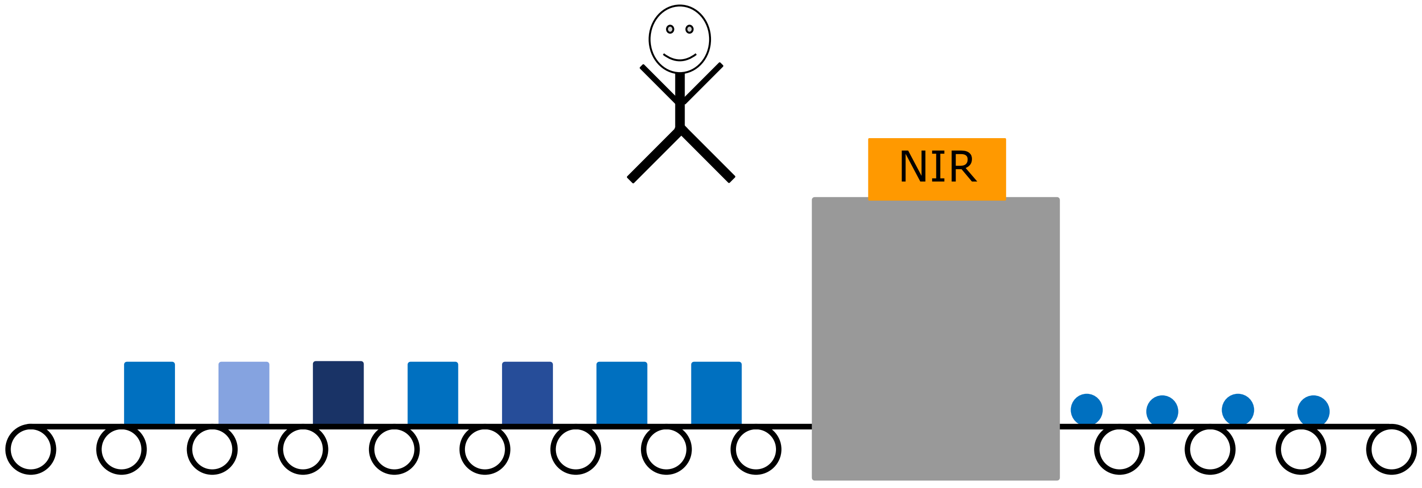
Motivation



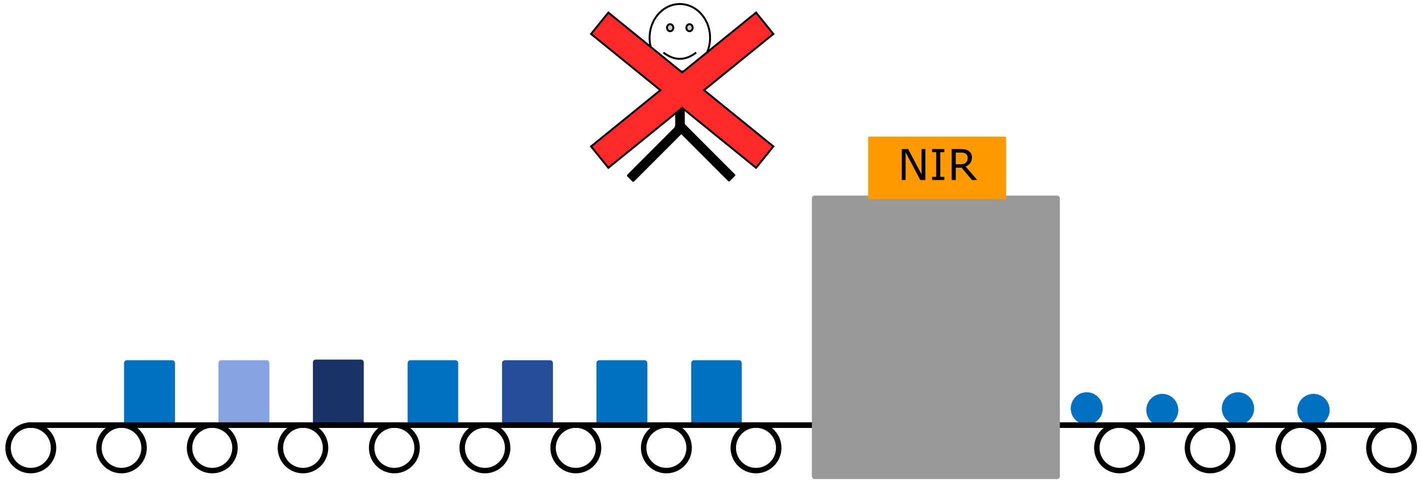
Motivation



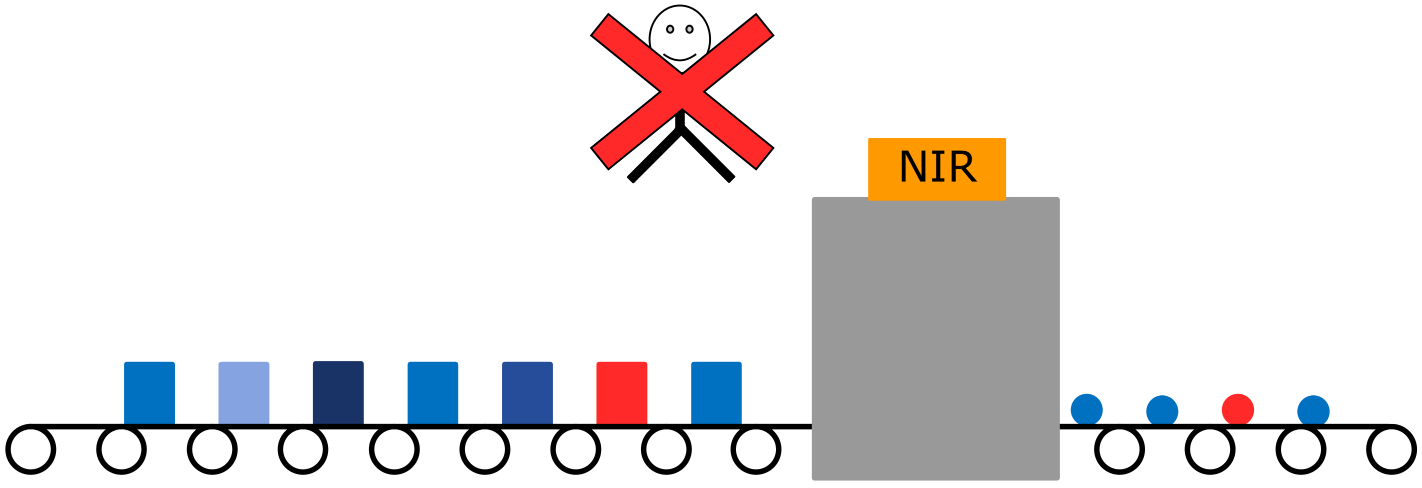
Motivation



Motivation



Motivation



What causes Covariate Shift

- Changing raw materials
  - Change of supplier
  - Supplier changes production
  - Weather
  - Etc.

- Change of recipe
  - Raw materials
  - Mechanical

- New instrument/sensor

- Any unseen variation not present in calibration data

Linear Joint Trained Framework

Labelled data: $X_L = \begin{bmatrix} x_1^T \\ \vdots \\ x_{n_L}^T \end{bmatrix}$ $Y_L = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$ (assume X_L to have column mean 0)

Unlabelled data: $X_U = \begin{bmatrix} x_{n_L+1}^T \\ \vdots \\ x_N^T \end{bmatrix}$ $X_U X_U^T = U_U \mathrm{D}_U U_U^T$ $X_U^{(\gamma_2)} = \sqrt{\gamma_2}(D_U + \gamma_2 I)^{-\frac{1}{2}} U_U^T X_U$

Define: $X = \begin{bmatrix} X_L \\ \sqrt{\gamma_1} X_U^{(\gamma_2)} \end{bmatrix}$ $Y = \begin{bmatrix} Y_L \\ \mathbf{1}\bar{Y}_L \end{bmatrix}$

Optimization problem: $\beta_{JT} = \underset{\beta}{\operatorname{argmin}} \|X\beta - Y\|_2^2 = \underset{\beta}{\operatorname{argmin}} \|X_L\beta - Y_L\|_2^2 + \gamma_1 \left\| X_U^{(\gamma_2)} \beta \right\|_2^2$

(Ryan, Culp 2015)

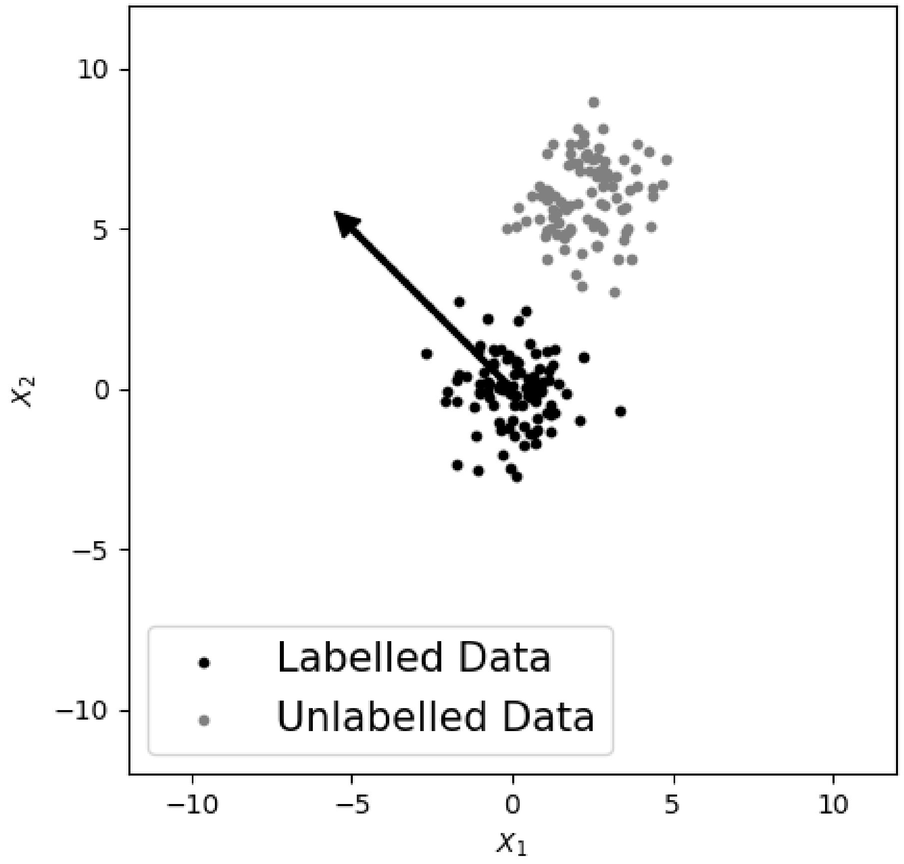
Ridge Regression

Optimization problem: $\beta_{RR} = \underset{\beta}{\operatorname{argmin}} \|X_L\beta - Y_L\|_2^2 + \lambda\|\beta\|_2^2$

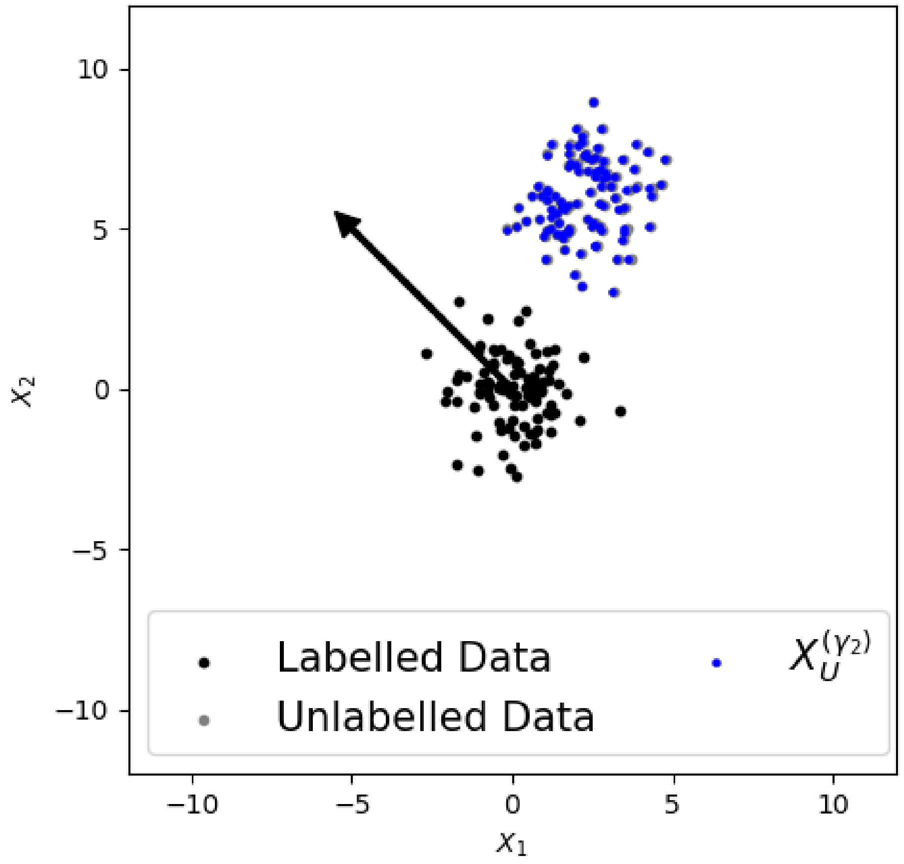
Equivalent to: $X = \begin{bmatrix} X_L \\ \sqrt{\lambda} I \end{bmatrix}$ $\quad Y = \begin{bmatrix} Y_L \\ \mathbf{1}\bar{Y}_L \end{bmatrix}$

$$\beta_{RR} = \underset{\beta}{\operatorname{argmin}} \|X\beta - Y\|_2^2$$

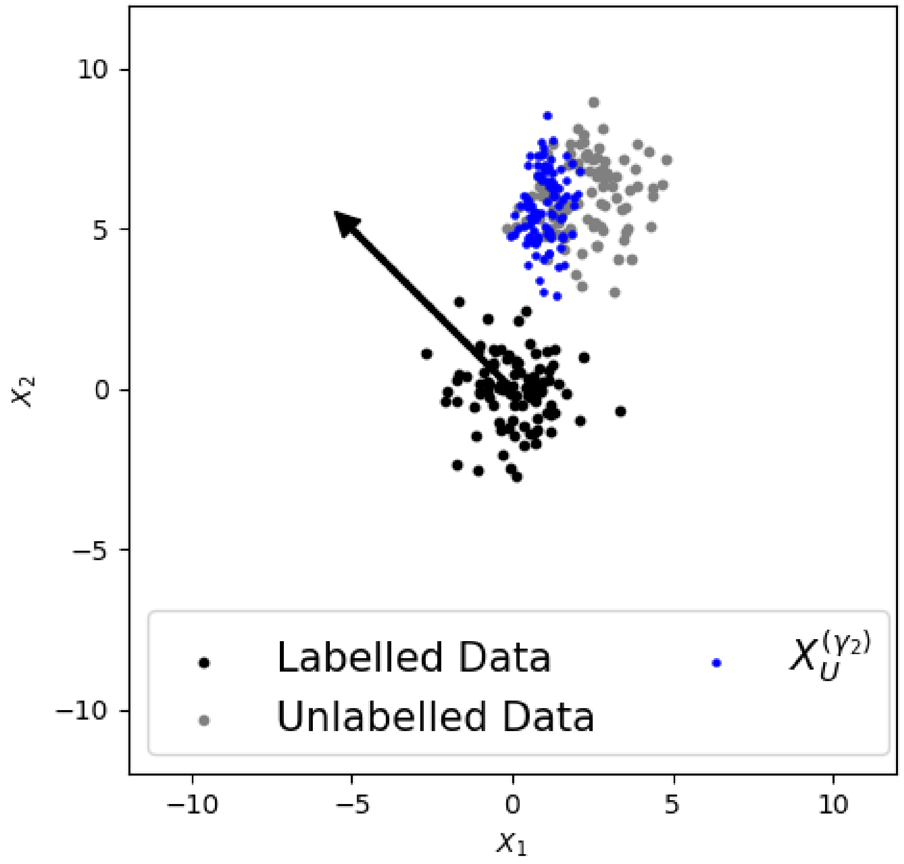
Example



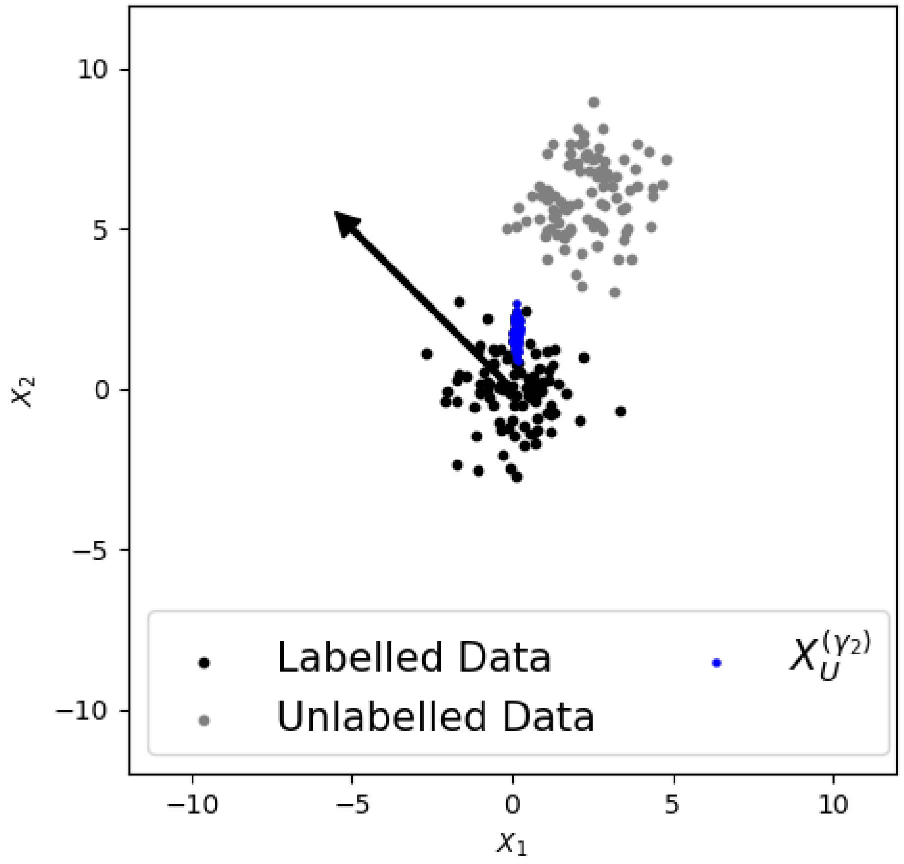
Example $(\gamma_2 = 10^3)$



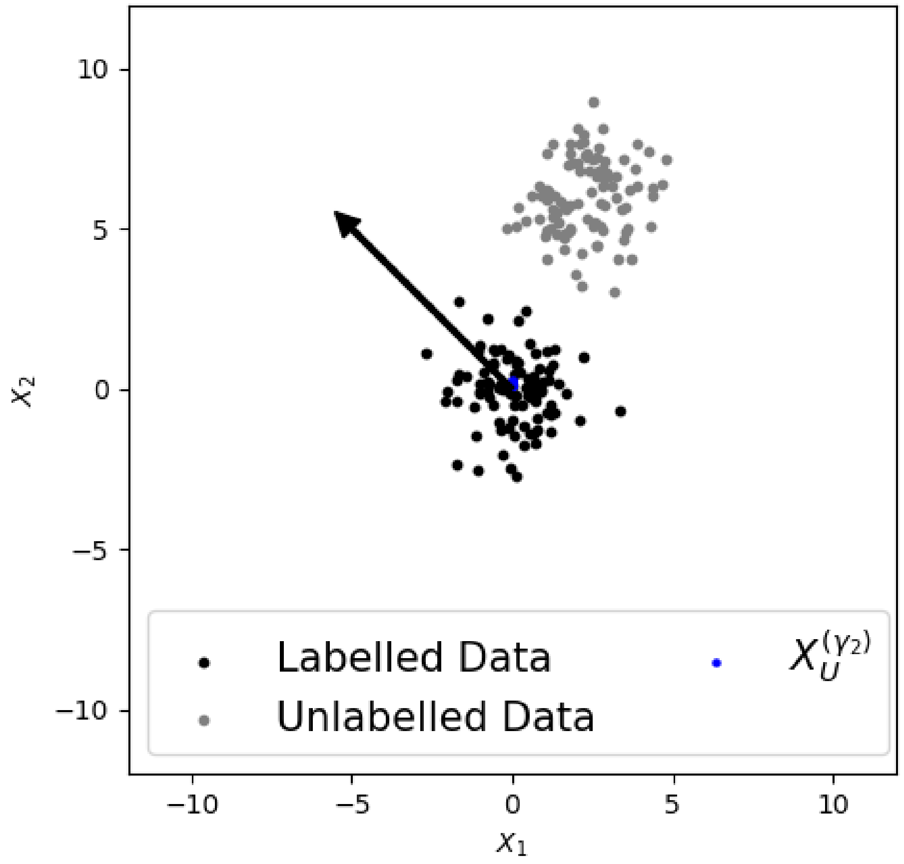
Example $(\gamma_2 = 10^1)$



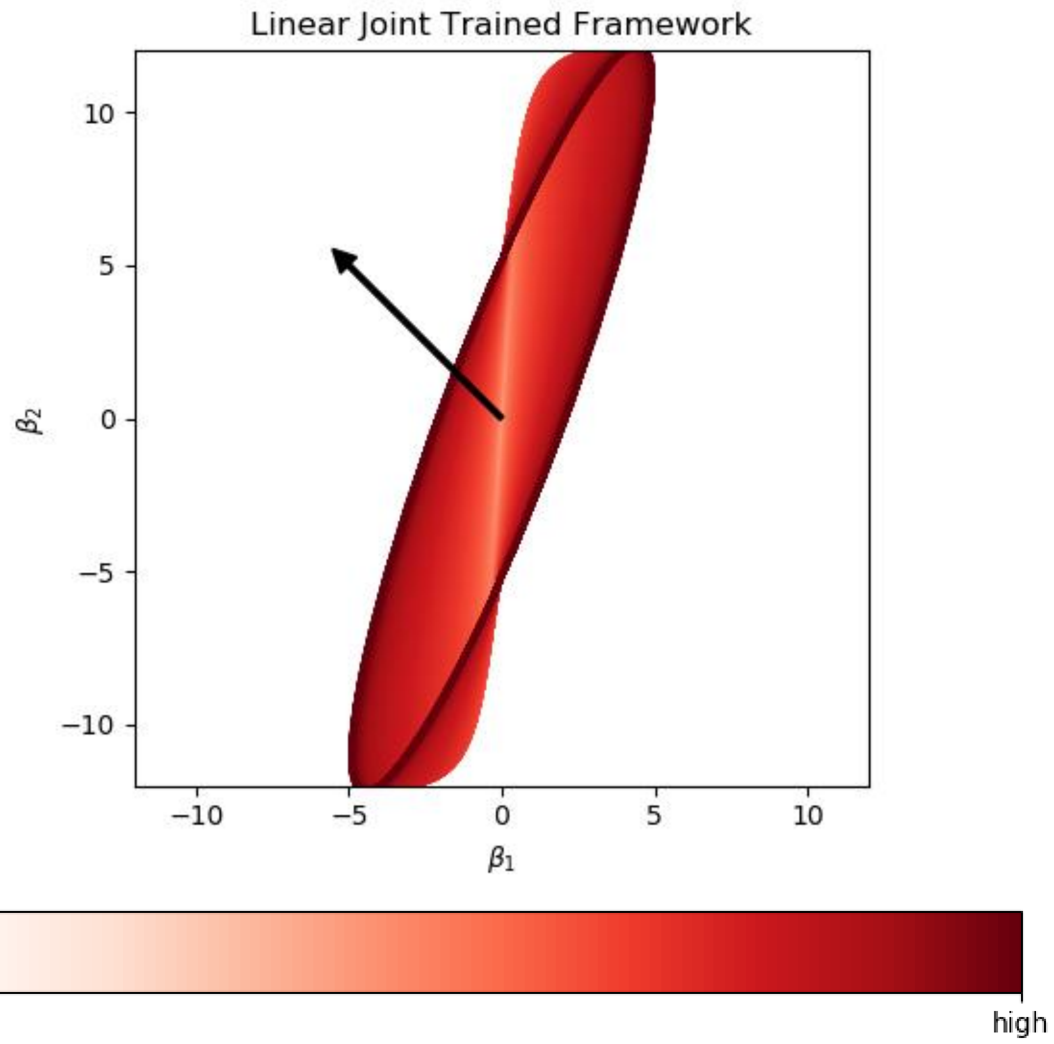
Example $(\gamma_2 = 10^{-1})$



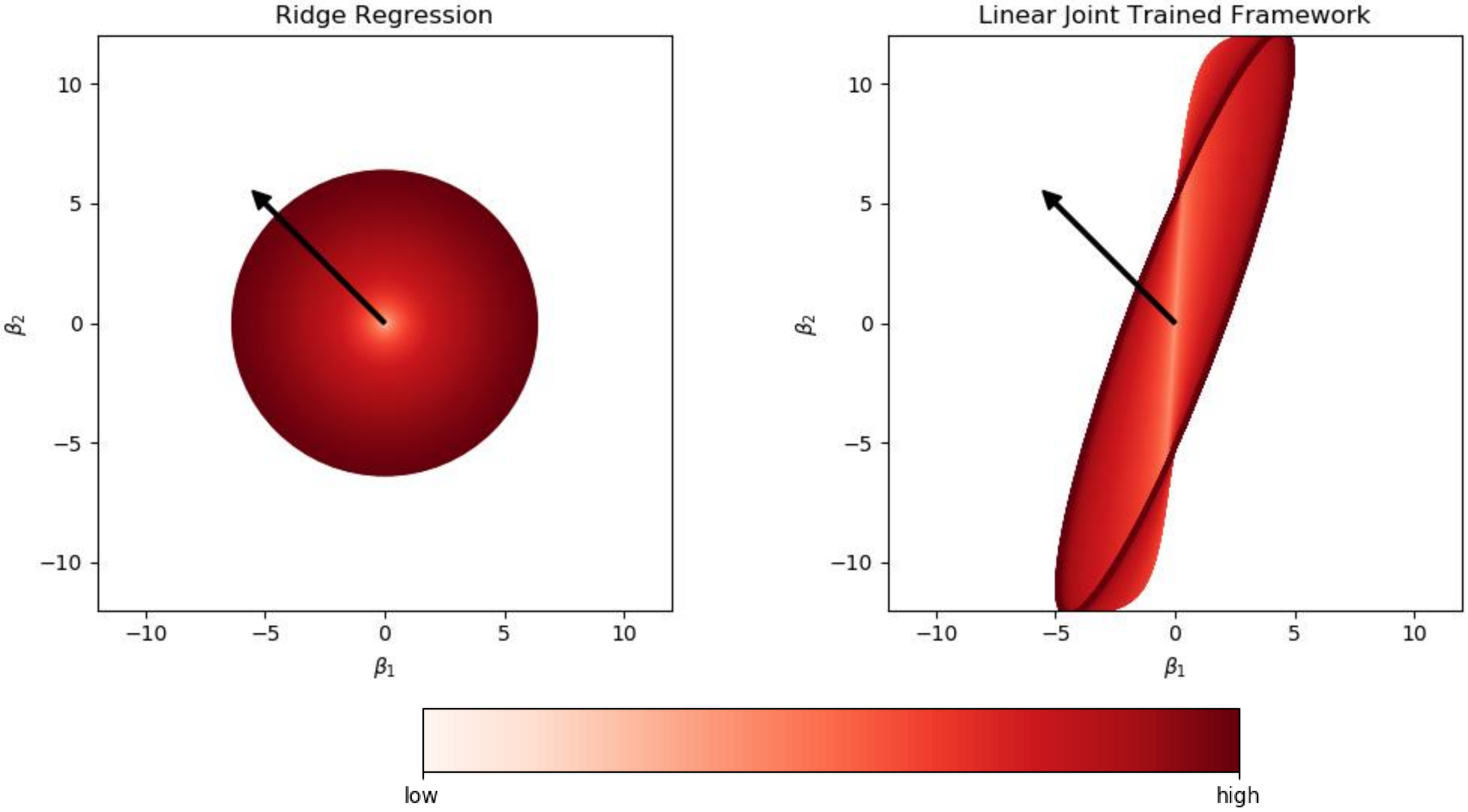
Example $(\gamma_2 = 10^{-3})$



What happens?



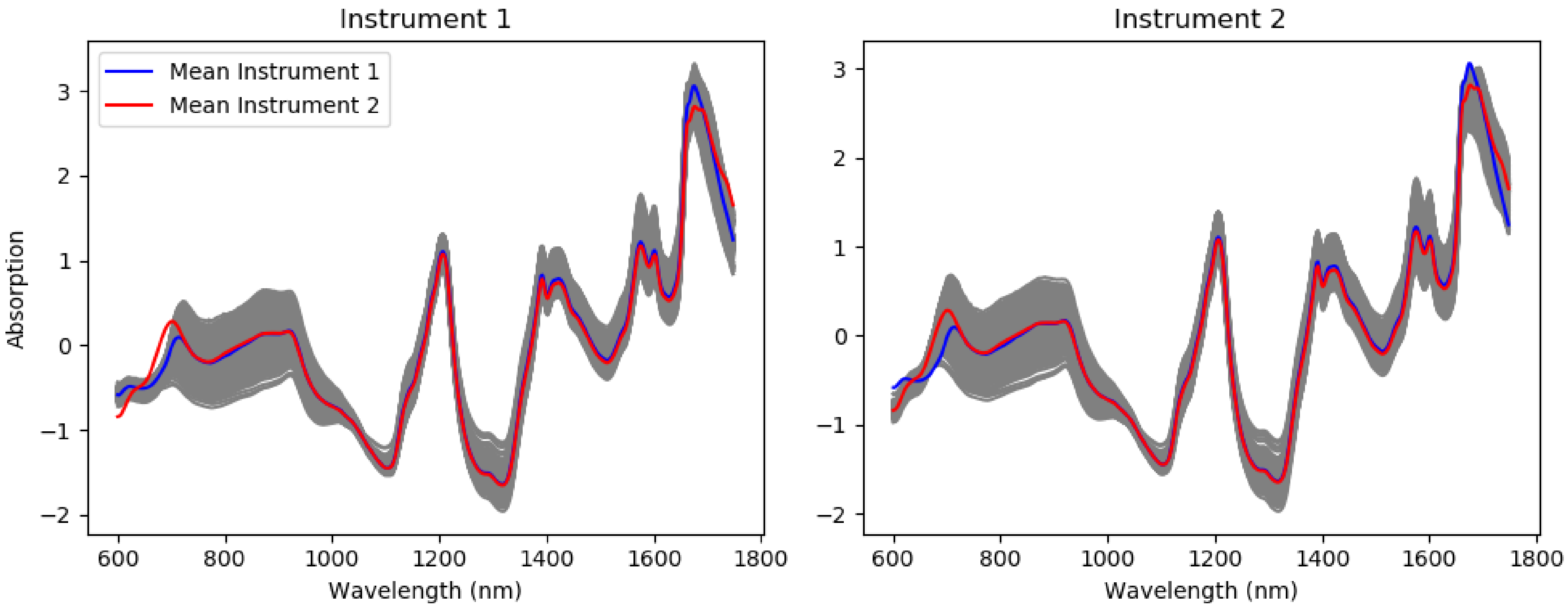
Comparison to Ridge Regression



Data

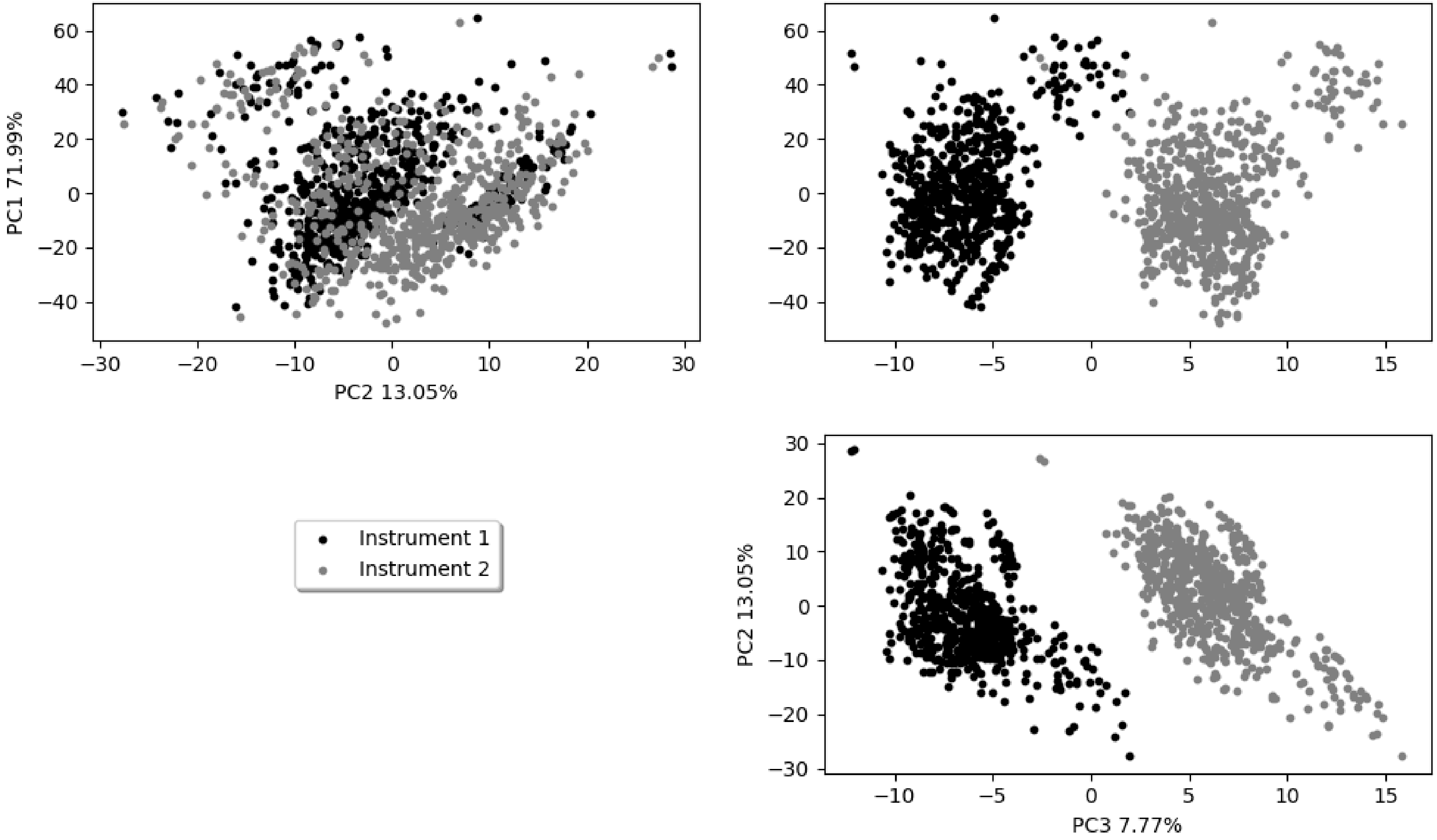
- IDRC 2002 Shootout (D.W. Hopkins 2003)
- 655 Pharmaceutical tablets
- Each tablet is measured on two different NIR Spectrometers
- Weight percentage of active ingredient

IDRC 2002

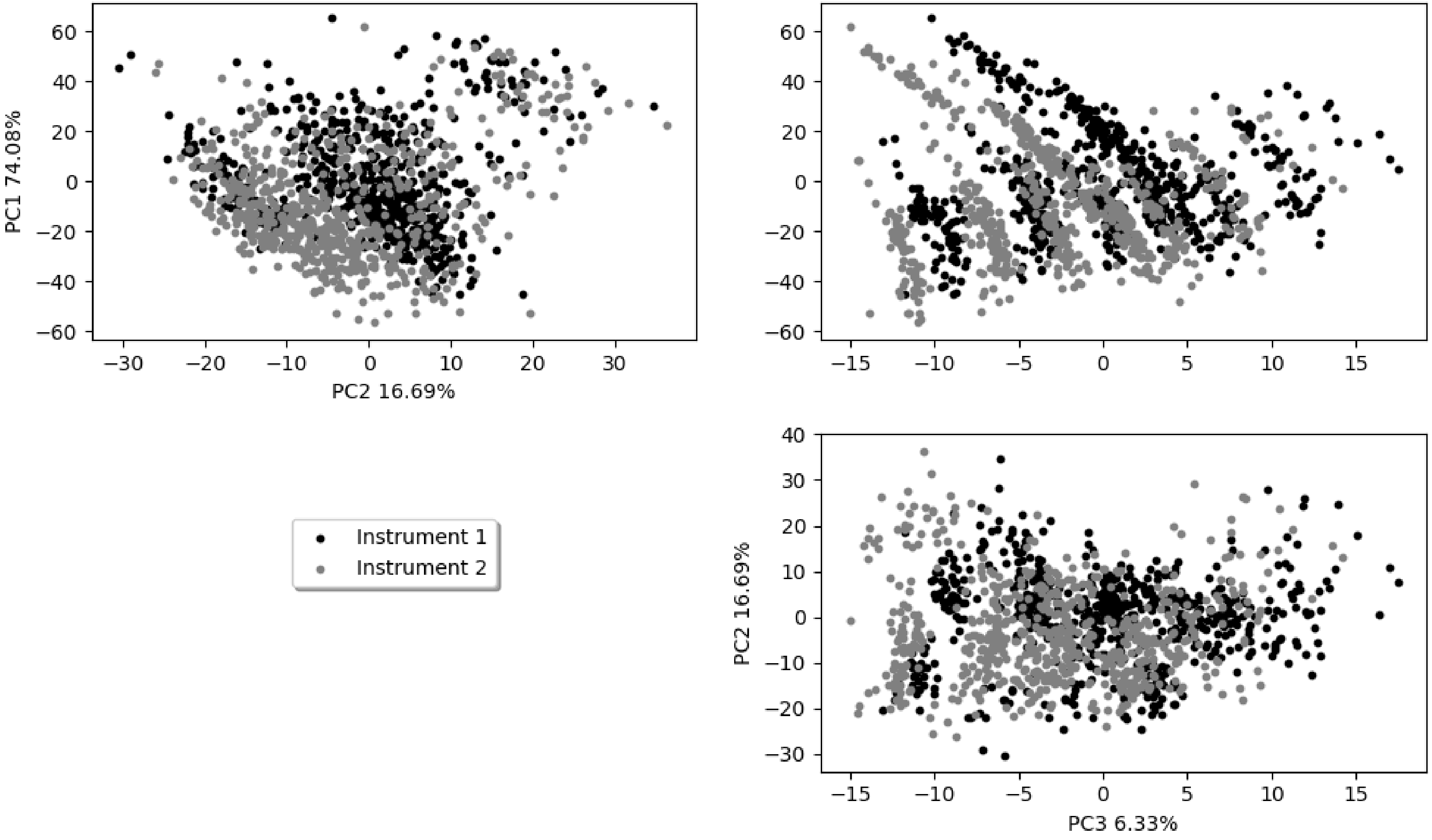


Instrument 1
Instrument 2
Mean Instrument 1
Mean Instrument 2
Absorption
Wavelength (nm)
Wavelength (nm)

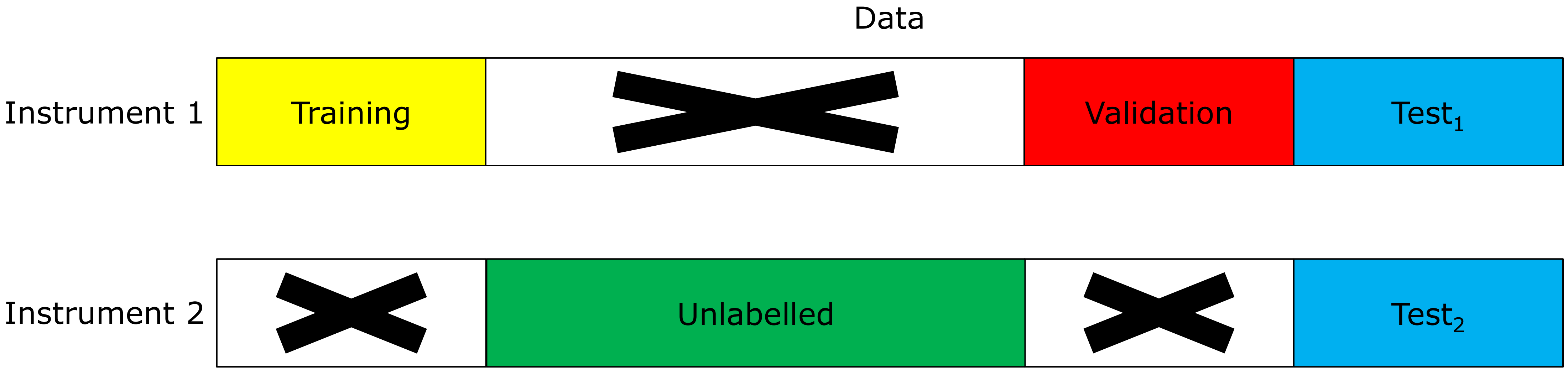
IDRC 2002



IDRC 2002



Experimental setup

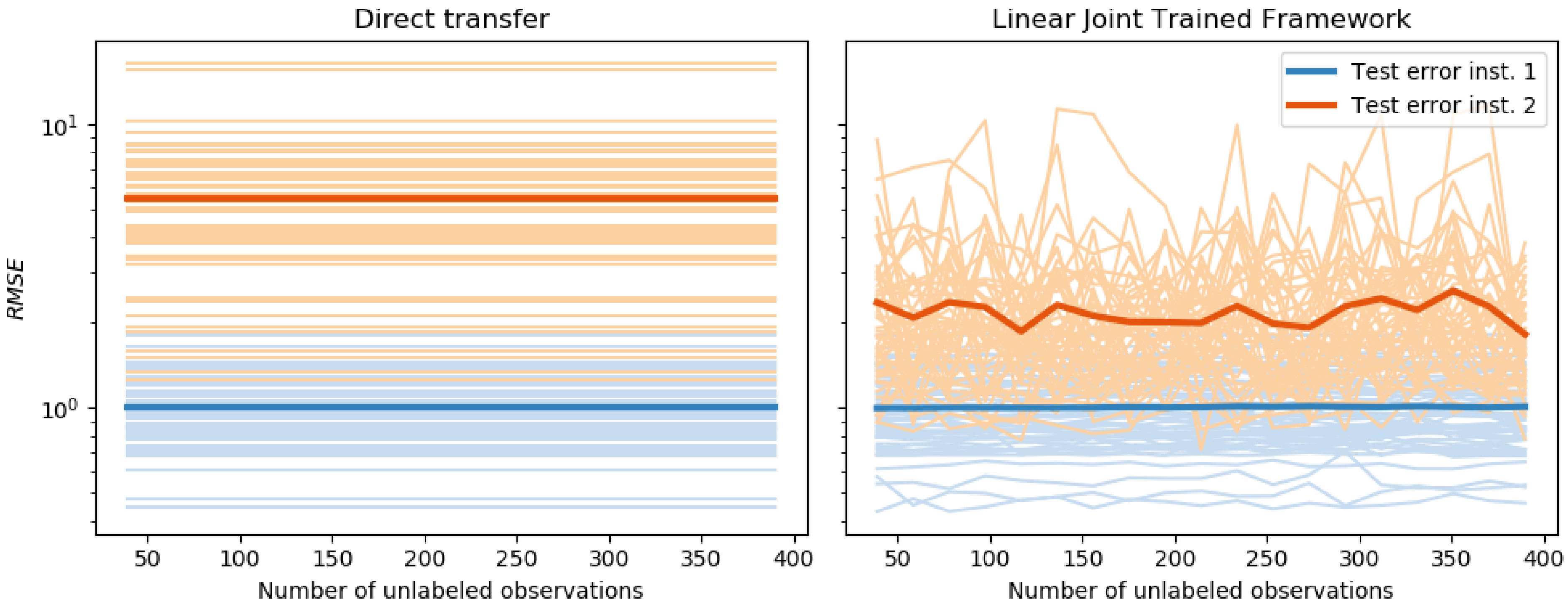


Performance Metric

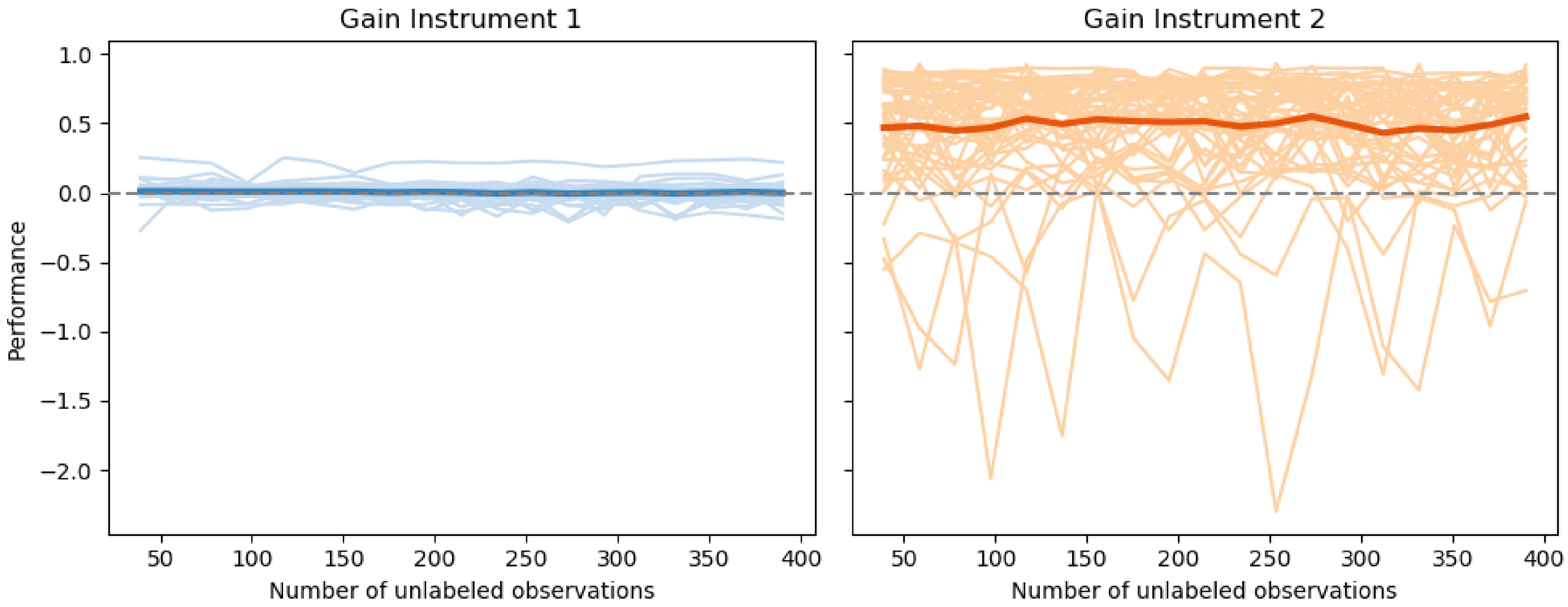
$$\mathrm{RMSE_i} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}\,, \qquad i \in \{Naive, JT\}$$

$$\text{Performance} = \frac{RMSE_{Naive} - RMSE_{JT}}{RMSE_{Naive}}$$

Performance when transferring model



Gain when transferring model



Future Perspectives - my work

- Improve Linear Joint Trained Framework
- Online detect and model Covariate Shift
- Online selection of tuning parameters for Linear Joint Trained Framework

References

- K. J. Ryan and M. V. Culp. On Semi-Supervised Linear Regression in Covariate Shift Problems, Journal of Machine Learning Research 16 (2015) 3183-3217
- D.W. Hopkins. Shoot-out 2002: Transfer of Calibration for Content of Active in a Pharmaceutical Tablet, NIR News 14 (2003) 10–13.