

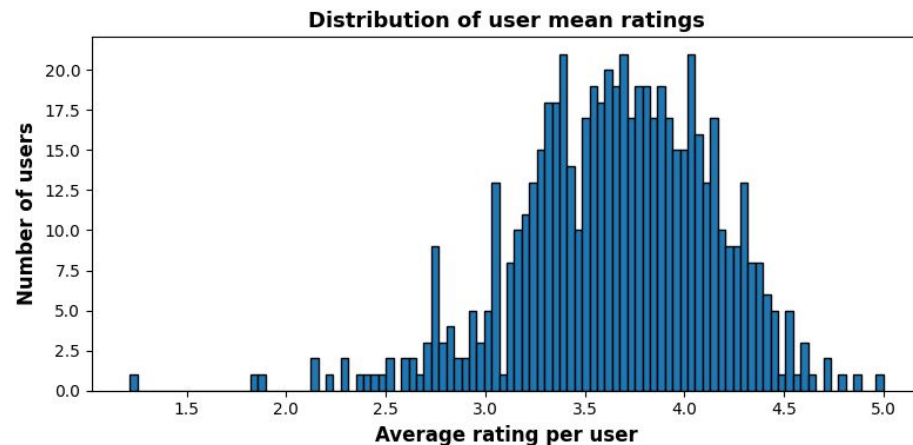
Data Science Lab

Collaborative Filtering

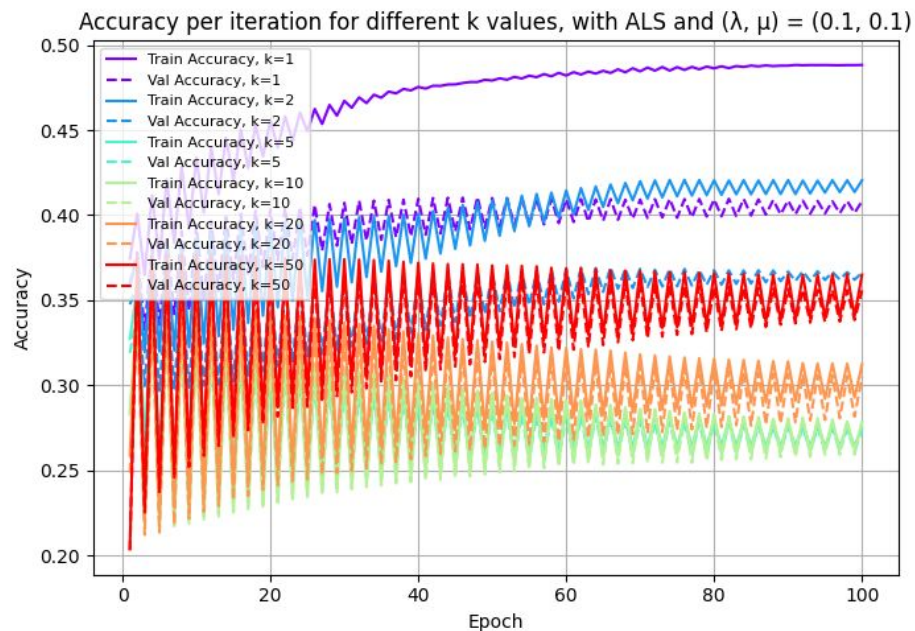
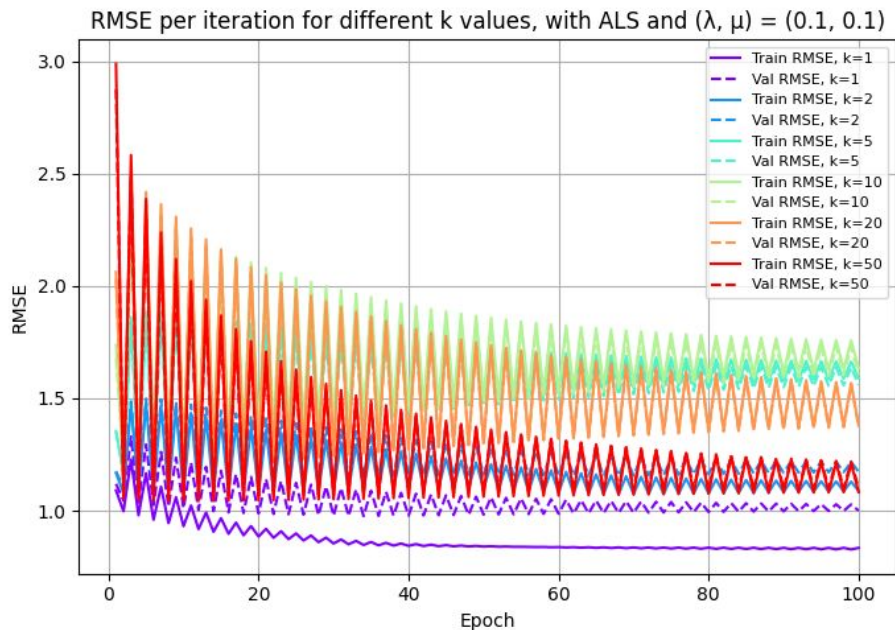
Louis Carron
Filomène Roquefort
Jules Roques

Implementation :

- Pre-Processing Normalization / User-Movie Bias
- MF with GD and ALS (show learning curves)
U and I initialization: Gaussian
- Iterative PCA
- Kernel PCA
- Parameters tuning
- Metrics: Loss, RMSE, Accuracy

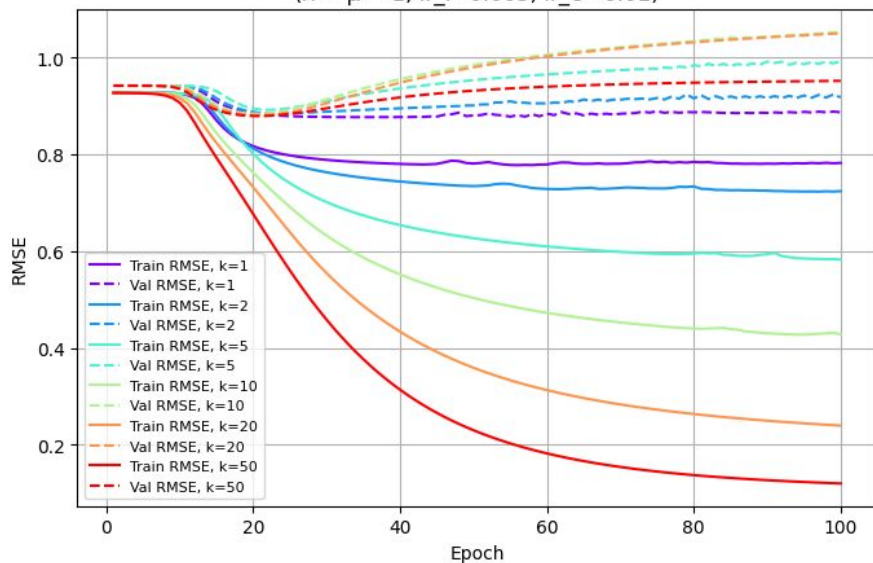


Matrix Factorization : Alternating Least Squares Parameters Tuning

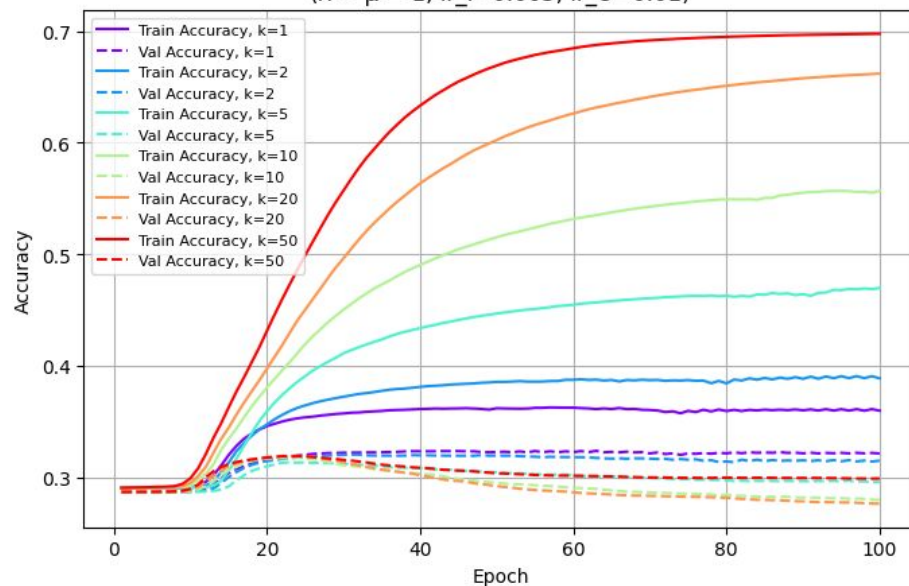


Matrix Factorization : Gradient Descent Parameters Tuning

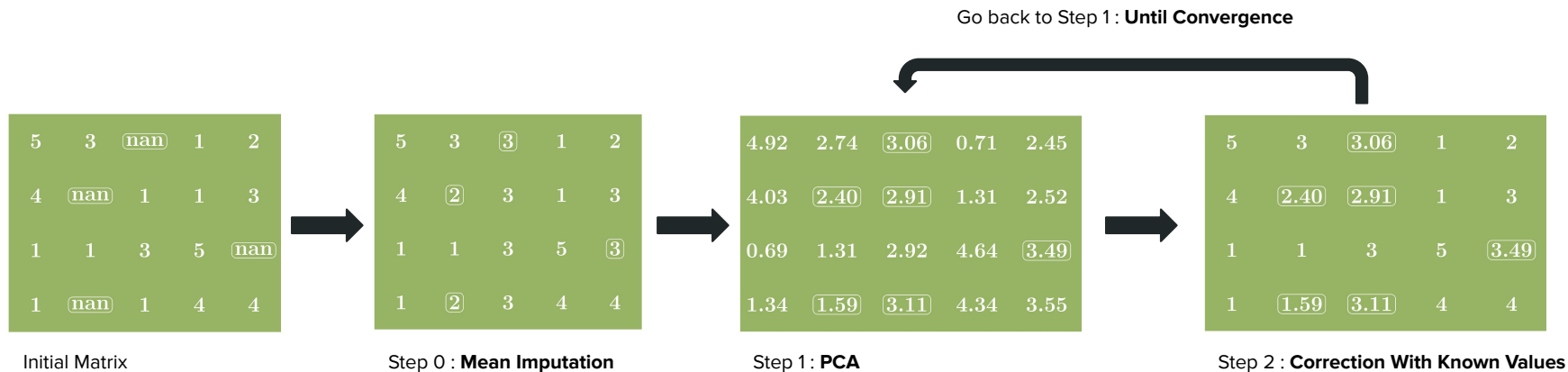
RMSE per iteration for different k values, with GD and normalisation,
($\lambda = \mu = 1$, $lr_L=0.005$, $lr_U=0.01$)



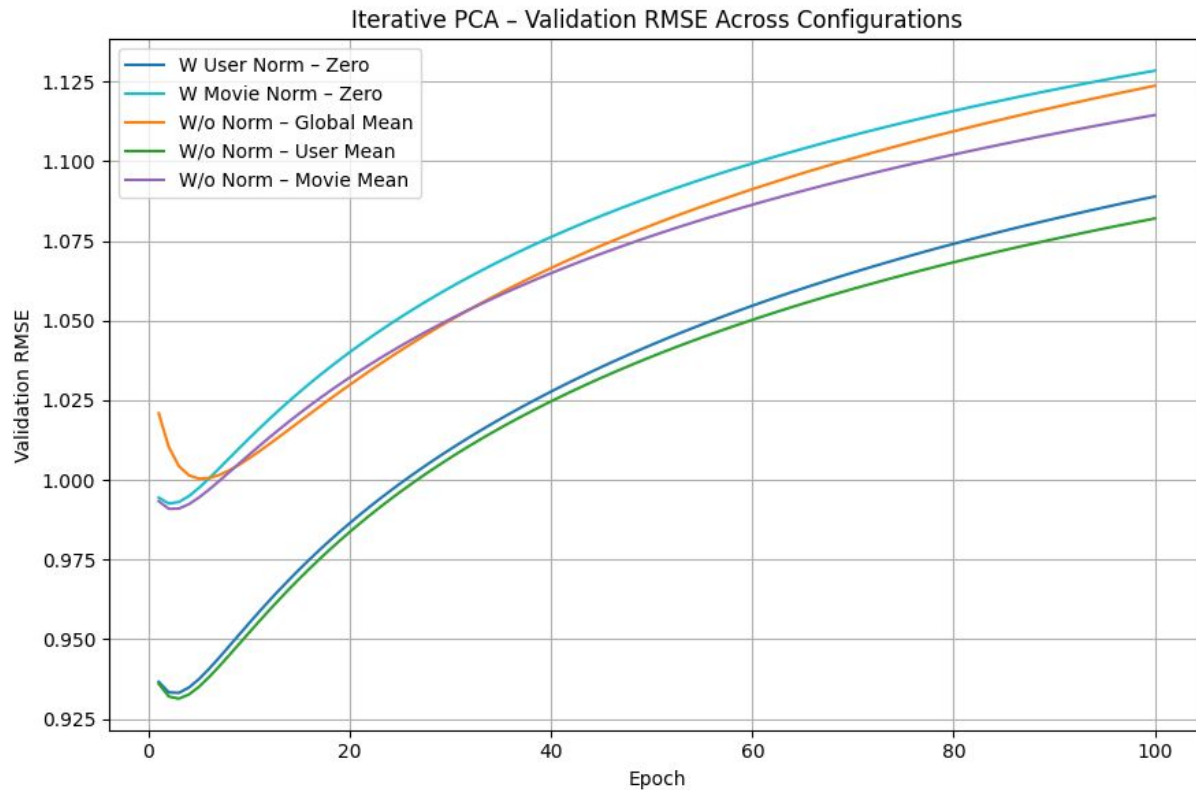
Accuracy per iteration for different k values, with GD and normalisation,
($\lambda = \mu = 1$, $lr_L=0.005$, $lr_U=0.01$)



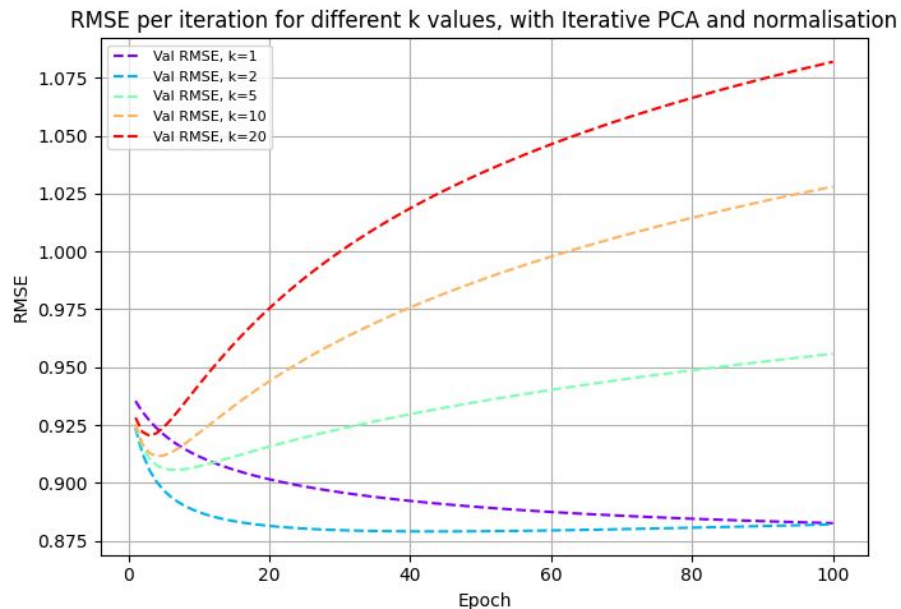
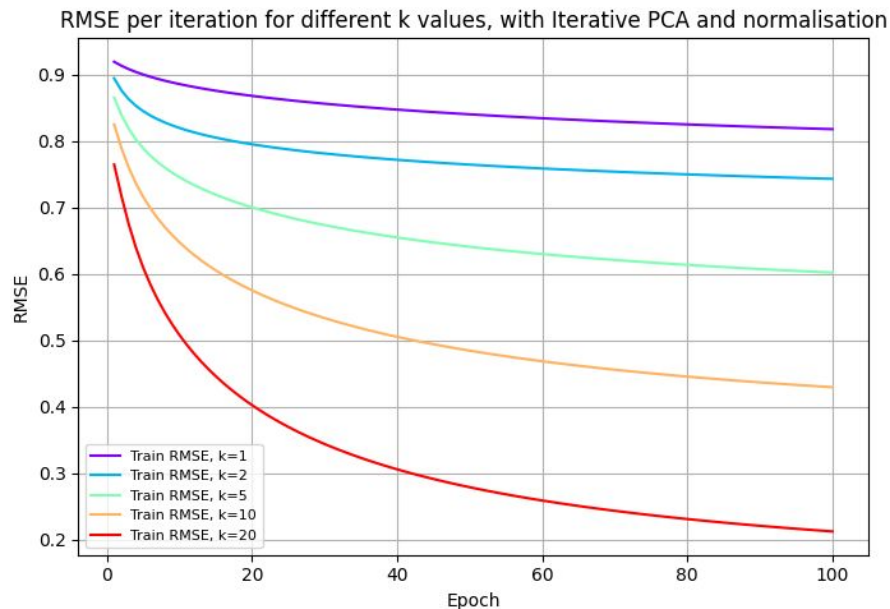
Iterative PCA : Algorithm



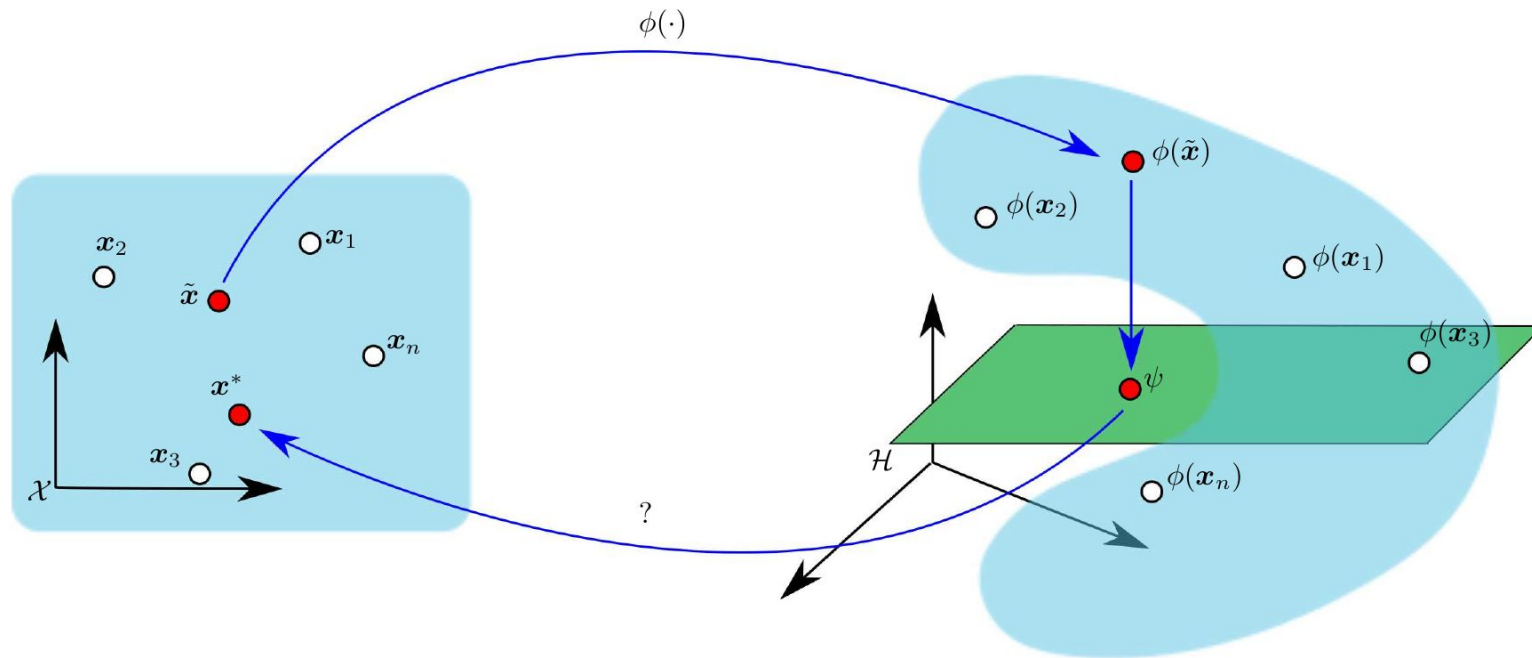
Iterative PCA : Dealing with Bias



Iterative PCA : Parameters Tuning



Iterative Kernel PCA : Intuitive Interpretation

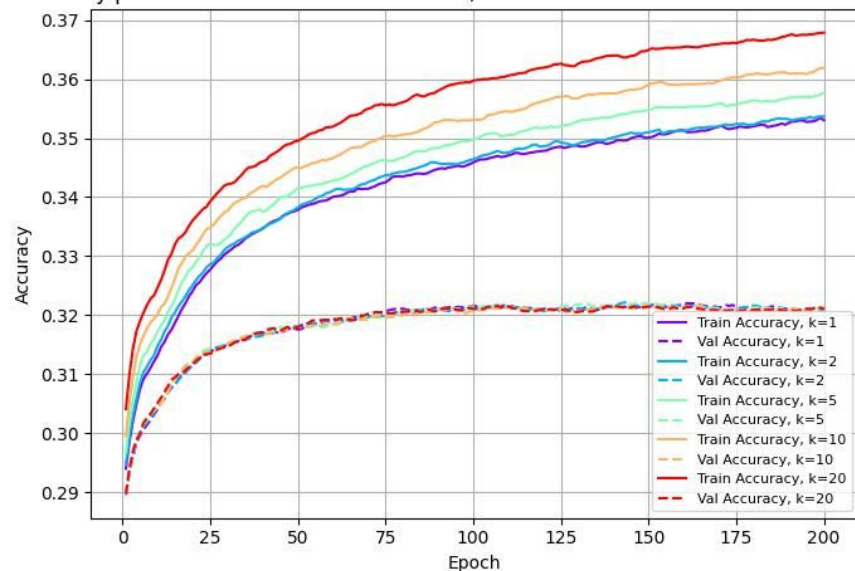


The **pre-image** problem in kernel-based machine learning

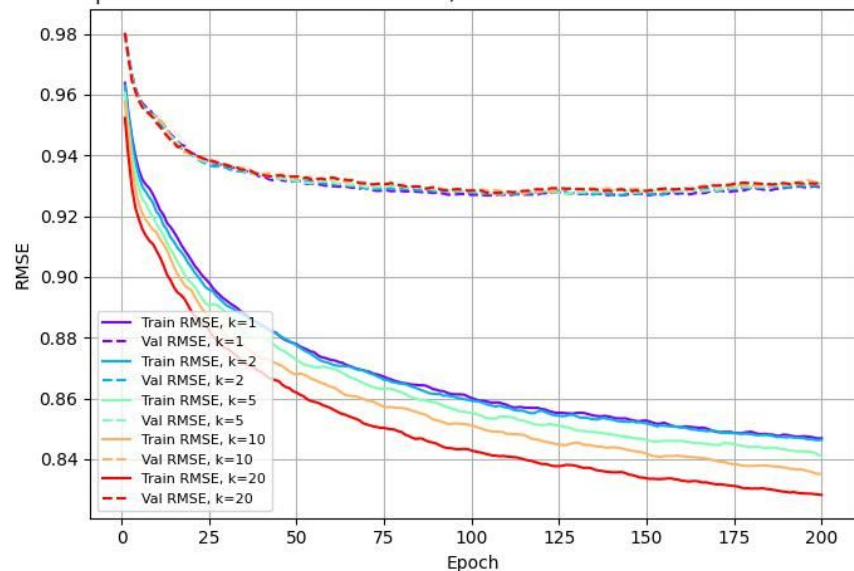
Paul Honeine, and Cédric Richard

Iterative Kernel PCA : Experimental Results

Accuracy per iteration for different k values, with Kernel Iterative PCA and normalisation



RMSE per iteration for different k values, with Kernel Iterative PCA and normalisation



Conclusion :

Table 1: Performance comparison of all implemented models after cross-validation.

Model	Best RMSE	Exact Accuracy (%)	Exec. Time (s)
Matrix Factorization (GD)	0.876	0.322	0.53
Matrix Factorization (ALS)	0.886	0.315	1.43
Iterative PCA	0.879	0.318	29.16
Iterative Kernel PCA (RBF)	0.927	0.322	3.86

- Impact of Normalisation
- Better Performances with Natively Sparse Methods

