CSIC 5011 Final Project: Fashion-MNIST Classification based on Manifold Learning

FAN Junyi, XIAN Zhuozhi {jfanap, zxian}@connect.ust.hk

Department of Mathematics, HKUST

1. Introduction

In this project, we first applied several manifold learning methods to find the lower-dimensional embedding of the high-dimensional Fashion-MNIST dataset. Next we conducted the classification for the types of the fashion based on the embedding results. Moreover, we estimated the intrinsic dimension of the dataset by using residual variance.

2. Data Description

Zalando's Fashion-MNIST dataset consists of 60,000 training images and 10,000 test images, labeled for 10 distinct types. The images are in grayscale with the size 28-by-28. For this project, we selected 10,000 images for our experiments.

3. Methodology

To embed the high dimensional Fashion-MNIST data into the lower dimesional space, the following methods for dimensioanlity reduction or nonlinear manifold learning are utilized:

- Principal Component Analysis (PCA)
- Multidimensional Scaling (MDS)
- Isometric Maps method (ISOMAP) is an extension of classical MDS, which seeks to preserve the geodesic distances between points [5].
- Locally Linear Embedding (LLE) attempts to project the data points into a low-dimensional space that best preserves the local geometric structures constructed in the neighbourhoods of the data points [4].
- Modified Locally Linear Embedding (MLLE) is an extension of LLE, which uses multiple weight vectors projected from orthogonal complement of local PCA to construct the weight matrix [8].
- Hessian Locally Linear Embedding (Hessian LLE) is an variation of LLE, which tends to construct the approximative Hessian operator on the manifold on which the data resides so that the coordinate representation is obtained from the near-zero eigenvectors of the operator [3].
- Laplacian Eigenmaps computes a low-diemsional representation of the data set that optimally preserves local neighbourhood information by using the notion of the Laplacian of the graph [1].
- Local Tangent Space Alignment (LTSA) uses the tangent space near a data point to represent the local geometry and then aligns those tangent spaces to construct the global coordinate system for the nonlinear manifold [9].
- Diffusion Maps constructs the weight matrix of the graph of the data that approximates the heat diffusion operator. [2].
- t-Distributed Stochastic Neighbor Embedding (t-SNE) is a method for visualizing high-dimensional data by giving each datapoint a location in a two or three-dimensional map, which overcomes the croweing problem in SNE [7].

4. Results for embedding

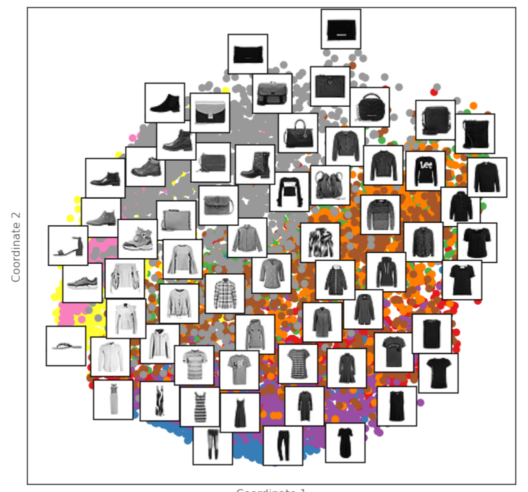

Fig. 1: PCA

Fig. 2: MDS

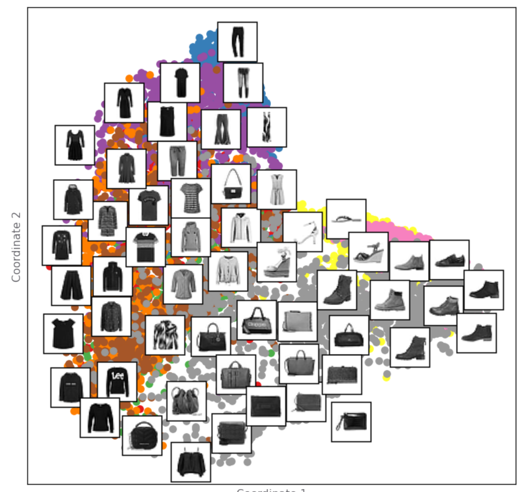

Fig. 3: ISOMAP

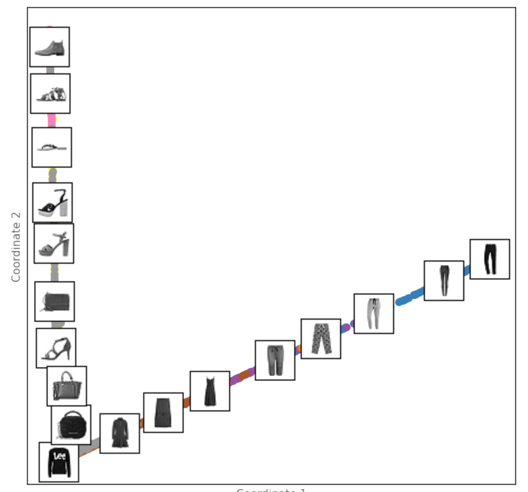

Fig. 4: LLE

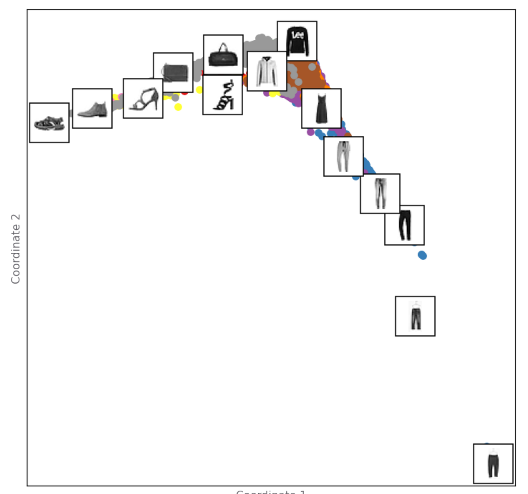

Fig. 5: MLLE

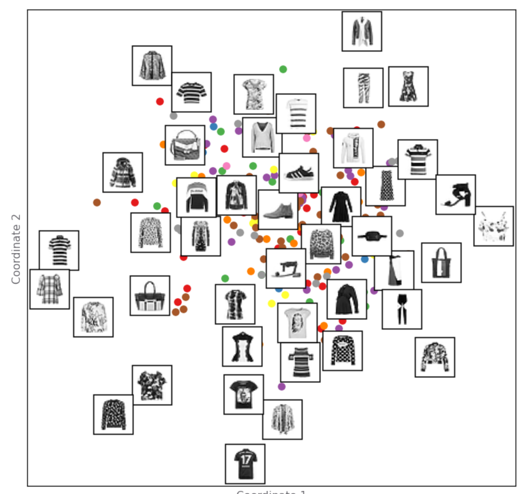

Fig. 6: Hessian LLE

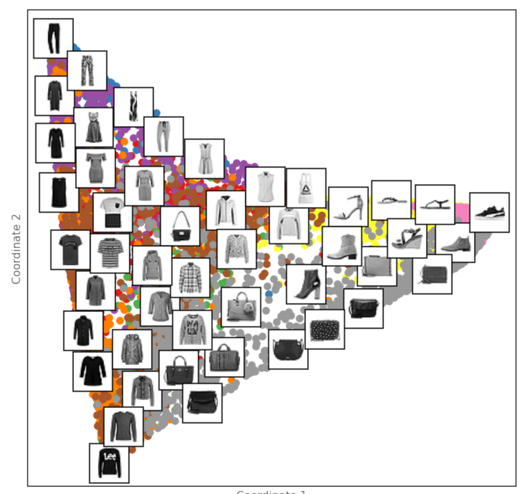

Fig. 7: Laplacian Eigenmaps

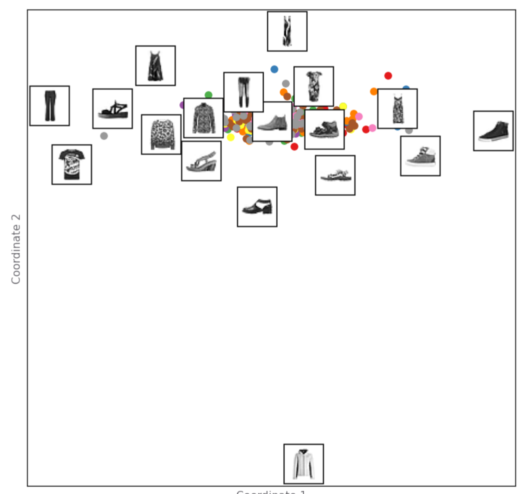

Fig. 8: LTSA

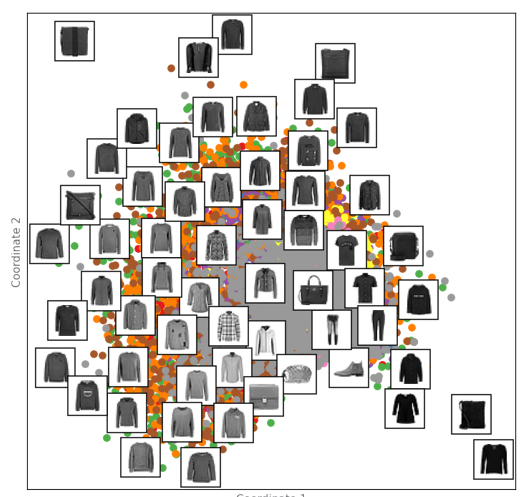

Fig. 9: Diffusion Maps

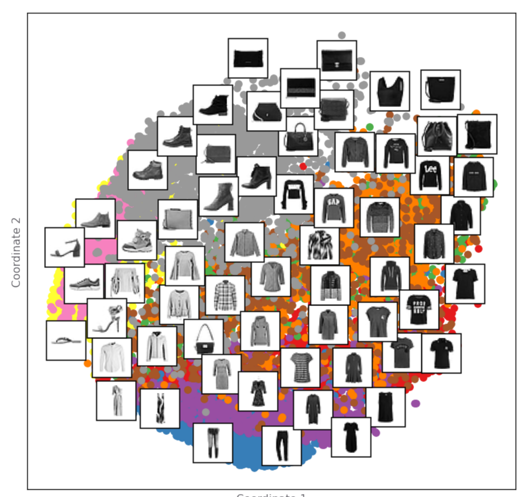

Fig. 10: t-SNE

The 2D-embedding of Fashion-MINST dataset using PCA, MDS, ISOMAP, LLE, MLLE, Hessian LLE, Laplacian LLE, LTSA, Diffusion Maps and t-SNE are shown in Fig. 1-10, respectively. Hessian LLE, LTSA and Diffusion Maps have bad embedding results, since the boundaries between different types of the fashion in Fig. 6,8,9 are not clear. For other methods, there exists clear boundaries between the clothes, pants, shoes and bags in the corresponding figures, especially in Fig. 4,5, hence they all perform well for embedding, while LLE and MLLE work best.

5. Results for classification

Method for embedding	Accuracy via Random Forest	Accuracy via SVM
PCA	49.5%	54.5%
MDS	53.4%	57.3%
ISOMAP	56.8%	59.7%
LLE	67.4%	50.2%
MLLE	64.3%	61.6%
Hessian LLE	10.3%	10.3%
Laplacian Eigenmaps	55.4%	58.0%
LTSA	10.9%	10.6%
Diffusion Maps	14.1%	16.4%
t-SNE	59.2%	56.1%
None	84.7%	84.6%

Tab. 1: Classification Accuracy

Based on the 2D-embedding results, we chose random forest and SVM as the classifiers to identify the types of the fashion in the dataset. The 5-fold accuracy for different combinations of embedding methods and classifiers are listed in Tab. 1, which show that LLE + Random Forest has the best performance.

6. Further Analysis

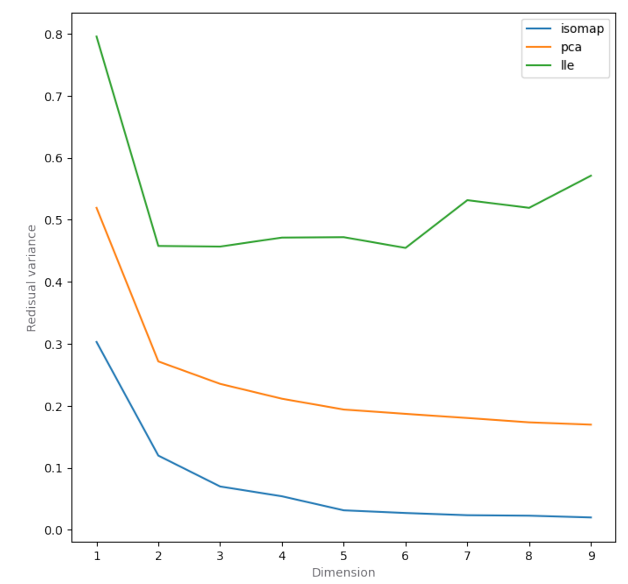

Fig. 11: Residual Variance

The quality of geometric structure holding for an embedding can be accessed by residual variance R_v, which is calculated by $R_v = 1 - R^2$, where R is the linear correlation for all the data points of original graph distances and embedded euclidean distances [6]. Then the intrinsic dimension of the data can be estimated from the figures about the residual variances for different embedding dimensions, shown in Fig. 11. Observing Fig. 11, for PCA, ISOMAP and LLE, the corresponding estimations of the intrinsic dimension are 2.

7. Conclusion

For embedding part, we performed different manifold learning methods (e.g. ISOMAP, LLE and Diffusion Maps) to embed the high dimensional Fashion-MNIST dataset into two dimensional space and visualize it. For classification part, we classified the types of the fashion by using different combinations of embedding methods and classifiers. Furthermore, we calculated residual variances for different embedding dimensions to estimate the intrinsic dimension of the dataset.

8. References

[1] Mikhail Belkin and Partha Niyogi. "Laplacian eigenmaps for dimensionality reduction and data representation". In: *Neural computation* 15.6 (2003), pp. 1373–1396.

[2] Ronald R Coifman and Stéphane Lafon. "Diffusion maps". In: *Applied and computational harmonic analysis* 21.1 (2006), pp. 5–30.

[3] David L Donoho and Carrie Grimes. "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data". In: *Proceedings of the National Academy of Sciences* 100.10 (2003), pp. 5591–5596.

[4] Sam T Roweis and Lawrence K Saul. "Nonlinear dimensionality reduction by locally linear embedding". In: *science* 290.5500 (2000), pp. 2323–2326.

[5] Joshua B Tenenbaum, Vin De Silva, and John C Langford. "A global geometric framework for nonlinear dimensionality reduction". In: *science* 290.5500 (2000), pp. 2319–2323.

[6] Sanket Thakre, Vishnu Harshith, and Anand K Kanjarla. "Intrinsic Dimensionality of Microstructure Data". In: *Integrating Materials and Manufacturing Innovation* 10.1 (2021), pp. 44–57.

[7] Laurens Van der Maaten and Geoffrey Hinton. "Visualizing data using t-SNE." In: *Journal of machine learning research* 9.11 (2008).

[8] Zhenyue Zhang and Jing Wang. "MLLE: Modified locally linear embedding using multiple weights". In: *Advances in neural information processing systems*. 2007, pp. 1593–1600.

[9] Zhenyue Zhang and Hongyuan Zha. "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment". In: *SIAM journal on scientific computing* 26.1 (2004), pp. 313–338.

9. Contribution

- **Coding** & **Presentation**: FAN Junyi
- **Poster** & **Presentation**: XIAN Zhuozhi