EE 516: Mathematical Foundations of Machine Learning	**Winter 2023**

Course Project

Due: March 24, 2023, 11:59PM PT

Instructor Name: John Lipor

Project Description

You will work in teams of three to form an in-depth "reading group" similar to those you might take part in during your academic career. As a group, you will select one of groups of three papers listed below. While all members should read each paper, each member will choose one paper to read carefully. You will then compare the theoretical and practical merits of each approach as a group and describe how the papers relate to the course material. Note that I do not expect you to understand every theoretical detail of your paper, but you should get a high-level picture of what the authors were able to prove and think about what tools they used to prove their results. You will then choose one paper to implement as a group in the programming language of your choice. You must test your algorithm on synthetic data and at least one real/benchmark dataset, which need not be one from the paper. You will turn in a final report at the end of the quarter (template will be provided).

Grading

The grading breakdown is as follows, with a topic-by-topic breakdown below. See the template for details on what is expected from each section.

Item	**Percentage**
Algorithm description & relation to course material (individual)	40%
Comparison of algorithms (group)	20%
Algorithm implementation & testing (group)	30%
Overall report presentation (group)	10%

Algorithm description & relation to course material: Graded on a scale of 0-40 as follows.

- 10 points: Problem description & formulation
- 10 points: Algorithm description
- 10 points: High-level description of theoretical results
- 10 points: Relation to course material

Comparison of algorithms: Graded on a scale of 0-20 as follows.

- 8 points: Comparison of theoretical guarantees
- 8 points: Comparison of empirical performance (computation time, synthetic data, and real data)
- 4 points: Overall critical thinking

Algorithm implementation & testing: Grade on a scale of 0-30 as follows.

- 20 points: Algorithm implementation
- 10 points: Results on synthetic data
- 10 points: Results on real/benchmark data

Overall report presentation: Grade on a scale of 0-10 as follows.

- 3 points: Template followed
- 3 points: Plots legible and labeled
- 4 points: Organization of group sections

Topics & Papers

Your group must choose one of the topics below unless you convince me that you have a relevant trio of papers not found here. Multiple groups may choose the same topic. I have also included optional survey articles that may be helpful in understanding the assigned papers. Survey articles do **not** count as one of your three papers.

Subspace Clustering

References [1, 2, 3]. Survey [4].

Robust Subspace Recovery

References [5, 6, 7]. Survey [8].

Nonlinear Dimensionality Reduction

References [9, 10, 11, 12] (choose three). Survey [13].

Incremental Low-Rank Methods

References [14, 15, 16]. Survey [17].

Robust PCA

References [18, 19, 20]. Survey [21].

Dictionary Learning

References [22, 23, 24]. Survey [25].

Low-Rank Matrix Completion

References [15, 26, 27]. Survey [28].

References

[1] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 35, pp. 2765–2781, Nov. 2013.

[2] J. Lipor, D. Hong, Y. S. Tan, and L. Balzano, "Subspace clustering using ensembles of k-subspaces," *arXiv preprint arXiv:1709.04744*, 2017.

[3] C. Lane, B. Haeffele, and R. Vidal, "Adaptive online k-subspaces with cooperative re-initialization," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2019, pp. 0–0.

[4] R. Vidal, "Subspace clustering," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 52–68, 2011.

[5] G. Lerman and T. Maunu, "Fast, robust and non-convex subspace recovery," *arXiv preprint arXiv:1406.6145*, 2014.

[6] M. Rahmani and G. K. Atia, "Coherence pursuit: Fast, simple, and robust principal component analysis," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6260–6275, 2016.

[7] A. Gitlin, B. Tao, L. Balzano, and J. Lipor, "Improving k-subspaces via coherence pursuit," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 6, pp. 1575–1588, 2018.

[8] G. Lerman and T. Maunu, "An overview of robust subspace recovery," *Proceedings of the IEEE*, vol. 106, no. 8, pp. 1380–1410, 2018.

[9] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[10] L. K. Saul and S. T. Roweis, "An introduction to locally linear embedding," *unpublished. Available at: http://www. cs. toronto. edu/˜ roweis/lle/publications. html*, 2000.

[11] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[12] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[13] L. Van Der Maaten, E. Postma, and J. Van den Herik, "Dimensionality reduction: a comparative," *J Mach Learn Res*, vol. 10, no. 66-71, p. 13, 2009.

[14] M. Brand, "Fast low-rank modifications of the thin singular value decomposition," *Linear algebra and its applications*, vol. 415, no. 1, pp. 20–30, 2006.

[15] L. Balzano, R. Nowak, and B. Recht, "Online identification and tracking of subspaces from highly incomplete information," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on.* IEEE, 2010, pp. 704–711.

[16] R. Kennedy, L. Balzano, S. J. Wright, and C. J. Taylor, "Online algorithms for factorization-based structure from motion," *Computer Vision and Image Understanding*, vol. 150, pp. 139–152, 2016.

[17] L. Balzano, Y. Chi, and Y. M. Lu, "Streaming pca and subspace tracking: The missing data case," *Proceedings of the IEEE*, vol. 106, no. 8, pp. 1293–1310, 2018.

[18] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, p. 11, 2011.

[19] T. Zhou and D. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," in *International conference on machine learning.* Omnipress, 2011.

[20] H. Xu, C. Caramanis, and S. Sanghavi, "Robust pca via outlier pursuit," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3047–3064, 2012.

[21] S. Ma and N. S. Aybat, "Efficient optimization algorithms for robust principal component analysis and its variants," *Proceedings of the IEEE*, vol. 106, no. 8, pp. 1411–1426, 2018.

[22] J. Mairal, F. Bach, and J. Ponce, "Task-driven dictionary learning," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 4, pp. 791–804, 2012.

[23] J. Sulam, B. Ophir, M. Zibulevsky, and M. Elad, "Trainlets: Dictionary learning in high dimensions," *IEEE Transactions on Signal Processing*, vol. 64, no. 12, pp. 3180–3193, 2016.

[24] Y. Zhai, Z. Yang, Z. Liao, J. Wright, and Y. Ma, "Complete dictionary learning via l4-norm maximization over the orthogonal group," *Journal of Machine Learning Research*, vol. 21, no. 165, pp. 1–68, 2020.

[25] I. Tosic and P. Frossard, "Dictionary learning," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 27–38, 2011.

[26] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*. ACM, 2013, pp. 665–674.

[27] P. Giampouras, R. Vidal, A. Rontogiannis, and B. Haeffele, "A novel variational form of the schatten-p quasi-norm," *arXiv preprint arXiv:2010.13927*, 2020.

[28] M. A. Davenport and J. Romberg, "An overview of low-rank matrix recovery from incomplete observations," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 4, pp. 608–622, 2016.