

Hao-Hu-Li-Liu: Visualization and Forecasting on Finance Data

Summary	This report used several PCA variations to reduce the finance data with 1000+ features to 2D, and tried to predict the results.
Strength	The performance of various methods are clearly demonstrated and explained with theoretical supports.
Weakness	“Forecasting” part is not well explained and I don’t understand which output is it predicting.
Writing	4. Generally speaking, the report is written in a clear and scientific manner. However, there are variables that are not explained, sentences that are awkward to read, and figures with very small font sizes.
Technical	3. Data are properly processed, various methods are appropriately applied, and their performances are somewhat explained. Just the forecasting part is confusing and seems not completed.
Overall	3. Overall speaking, this report has satisfactory technical and writing standards.
Confidence	3. I read the report carefully and think the results are correct, though I didn’t run their codes.

Li-Chen-Teng-Sheng: NIPS Conference Papers 1987-2015 Data Set

Summary	This report used K-means to cluster the vocabulary in NIPS papers and HMM to predict the probability of a word being used in its cluster.
Strength	Workflow is clearly explained. Visualisation of the result is amazing.
Weakness	The technical part and relevant codes are not presented.
Writing	5. Generally speaking, the report is written in a clear and scientific manner. Scatter plots and sunburst charts look very beautiful (look even better when using the interactive HTML version).
Technical	4. Though in this poster, some technical contents (algorithms, maths...) are not included, I believe a lot has been done behind the scene, judging from their results.
Overall	5. Overall speaking, this poster is perfectly finished, and demonstrated their strong capabilities.
Confidence	3. I read the report carefully and think the results are correct, though I didn’t run their codes.

Ma-Tang-Ruan-Huang: Ancestry Prediction via Dimensionality Reduction Techniques on SNPs Data

Summary	This report used several PCA variations to reduce the SNP data with 100k+ features to 2D, and verified these methods in an independent cohort.
Strength	The performance of various methods are clearly demonstrated. To benchmark these algorithms, a metric “ARI” in introduced.
Weakness	The reason why some methods perform better than others are not well explained.
Writing	4. The poster is neat and well organised. Scatter plots informative enough for readers to reach a conclusion.
Technical	4. The results obtained by applying these methods are convincing and a suitable metric is used to quantify the accuracy of different algorithms. However, the difference of these methods are not elaborated.
Overall	4. This report has good technical contents and writing standards. I believe metric they used to benchmark different methods is a good choice.
Confidence	3. I read the report carefully and think the results are correct, though I didn’t run their codes.

Xu-Zhuang-Zhou: Visualization and Dimensionality Reduction Techniques for US Crime Data

Summary	This report used several PCA variations to reduce and visualise the USA crime data with 24 features to 2D/3D.
Strength	The workflow to process the data and performance of various methods are clearly demonstrated.
Weakness	The way they convert crime type to a 4-point scale is not convincing. If “stealth” and “rape” are both very often, I don’t think they should belong to the same category. In addition, the results seems chaotic and not easy to separate the clusters (might be due to the above-mentioned reason).

Writing	3. The poster is neat and well organised. Language is fluent and suitable for the academic context.
Technical	3. The results obtained by applying these methods are not very promising. Visually, it's difficult to distinguish one cluster from another, as they highly overlap with each other. This problem isn't better in 3D. In addition, no in-depth discussion is presented on why this happened.
Overall	3. This report lacks technical contents the results presented are not too much informative.
Confidence	3. I read the report carefully and think the results are correct, though I didn't run their codes.

Yan-Yan-Lai: The statistic explanation of the US crime data

Summary	This report used PCA, Isomer, and UMAP to reduce and visualise the USA crime data with 41 features to 5.
Strength	Data is preprocessed with engineering common sense. Different methods are explained briefly and their performances are evaluated.
Weakness	For the first few figures, cities are labeled instead of the crime types. In my opinion, it would be more interesting to study that for a particular crime type, which factors are more important.
Writing	4. The report is well written in an academic style. But some latex functions are not properly used, like “” and “.”.
Technical	4. The results from PCA and Isomap are not satisfactory (just as the previous report that we reviewed), however, UMAP performs surprisingly well. It successfully categorises the data into clusters.
Overall	4. Overall, the quality of this report is quite good (much better than the previous one).
Confidence	3. I read the report carefully and think the results are correct, though I didn't run their codes.