Machine Learning Final Project

CIFAR-10 classification with Deep and Machine learning methods

At first, I perform experiments with ML based methods, and I considered following preprocess for dataset:

- **PCA**

  Principal Component Analysis(PCA) is a way for dimension reduction that works in an **unsupervised** manner. In order to find proper number of eigenvectors in PCA algorithm, I used Proportion of Variance(POV) formula.

  $$POV = \frac{\sum_{i=1}^{d} \lambda_i}{\sum_{i=1}^{D} \lambda_i}$$

  In above formula d is number of eigenvectors, λ is eigenvalue, and POV is the amount of variance that explain with selected d components. I considered $POV = 0.9$, and found $d = 99$. So after performing PCA on dataset, the new shape for dataset is $60000 * 99$.

- **LDA**

  Linear Discriminant Analysis(LDA) is a way for dimension reduction that works in a **supervised** manner. LDA maximize distance between projected class means, while also achieving a small variance within each class. After performing LDA on dataset, the new shape for dataset is $60000 * 9$.

- **PLS**

  Partial Least Squares(PLS) is a way for dimension reduction and is based on NIPALS algorithm proposed by Herman Wold's paper [1] and works in a **supervised** manner. As proved by Barker and Rayens [2] PLS is to be preferred over PCA when discrimination is the goal and dimension reduction is needed.

- **Feature Selection**

  After compute variance of each feature, features are sorted based on their variance and first one hundred features that has most variance has selected. After feature selection, the new shape for dataset is $60000 * 100$

Results:

Table 1 Machine Learning methods results

Classifier / Feature	1-NN	5-NN	9-NN	Logistic Regression	AdaBoost	SVM
Raw data	35.39	33.98	33.98	40.24	33.08	14.75
PCA	19.88	20.58	20.91	24.83	23.61	15.5
LDA	36.52	43.13	45.67	49.99	47.13	**50.68**
PLS	30.54	34.02	36.12	34.06	29.75	18.67
PLS-LDA	29.96	34.09	36.13	33.97	30.88	41.45
PCA-LDA	17.33	19.85	22.19	25.74	25.43	28.21
Feature selection	19.68	20.34	20.67	15.95	20.35	11.39

As results show, LDA preprocces achieved best accuracy among other preprocces, and features selected with feature selection preprocess achieved poor result so it's better to try different approach to select features.

Then I evaluate some developed deep neural network to perform classification [3-5], and the following results achieved.

Table 2 Deep Learning methods results

Model / Metric	Accuracy	Description
A simple small Deep CNN [3]	76.2	
A medium Deep CNN [4]	81.8	
A large Deep CNN [5]	86.9	
DenseNet [6]	87	Trained for 20 epoches

According to results in table 2, as we use deeper network, we get higher accuracy, but the time for training networks will increase.

Conclusion

Since the CIFAR-10 dataset is large, and a deep network need large data, deep learnig methods get better result than machine learning methods on classifying dataset.

Refrences

[1] Wold H. Estimation of principal components and related models by iterative least squares

[2] Barker, M., Rayens, W., *Partial least squares for discrimination ,* Journal of Chemometrics, 17(3), 166–173, 2003

[3] https://keras.io/examples/cifar10_cnn/

[4] https://bit.ly/30MIUJF

[5] https://bit.ly/2GoPVbS

[6] Gao Huang, "Densely Connected Convolutional Networks" Cornell University