

PQP Report - Autoencoders and Resumes

Griffin Bishop, Sam Pridotkas, Leo Grande, Harry Sadoyan

Autoencoders

An autoencoder is an unsupervised deep learning model used for dimensionality reduction. It uses a number of hidden layers and an activation function to attempt to encode the data into a lower dimensional space, and then decode it again. The accuracy is measured by how closely the output resembles the original input. (Figure 1)

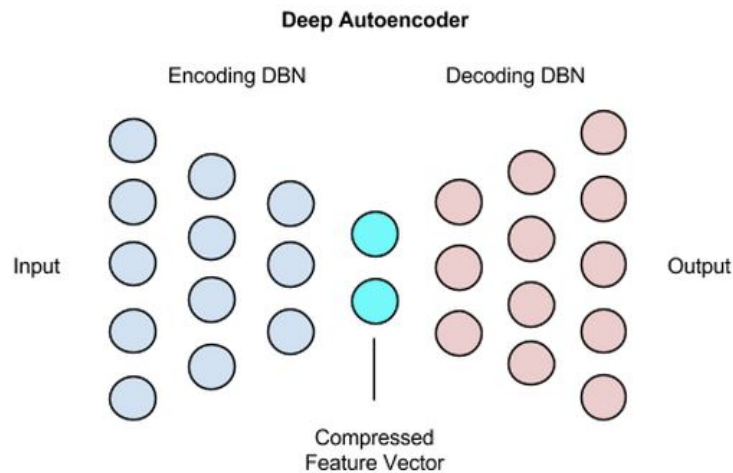


Figure 1: Autoencoder inner structure

(According to the Universal Approximation Theorem, increasing the number of hidden layers should improve accuracy up to a certain asymptote. Since autoencoders are an unsupervised learning method we don't need to worry about overfitting.)

Difference between PCA and autoencoders

1. Both PCA and autoencoders are dimensionality reduction techniques
 - a. They both map to lower dimensions, but neither are feature reduction methods
2. PCA is a linear technique while autoencoders can be (and usually are) non-linear. An autoencoder with a linear activation function and one hidden layer is effectively identical to PCA
3. When dealing with non-linear data, the autoencoder is expected to outperform PCA. Figure 2 shows a comparison of an autoencoder and PCA for the MNIST dataset.

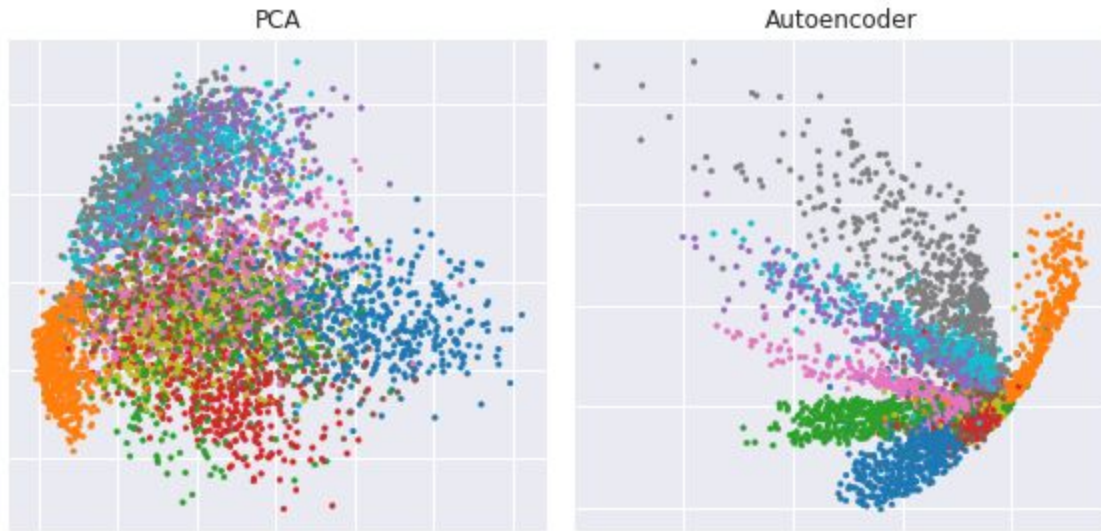


Figure 2: PCA vs Autoencoder for MNIST

Resumes Versus Movie Reviews:

We trained a random forest classifier using 20 principal components to distinguish movie industry resumes from movie reviews. The model performed with an accuracy of 99.8%

N-gram tf-idf

We trained a random forest classifier to distinguish positive movie reviews from negative ones. With 20 principal components, the model performs with an accuracy of 66%. After training the model using 3-gram tf-idf and 120 principal components, the accuracy improved to 78.88%

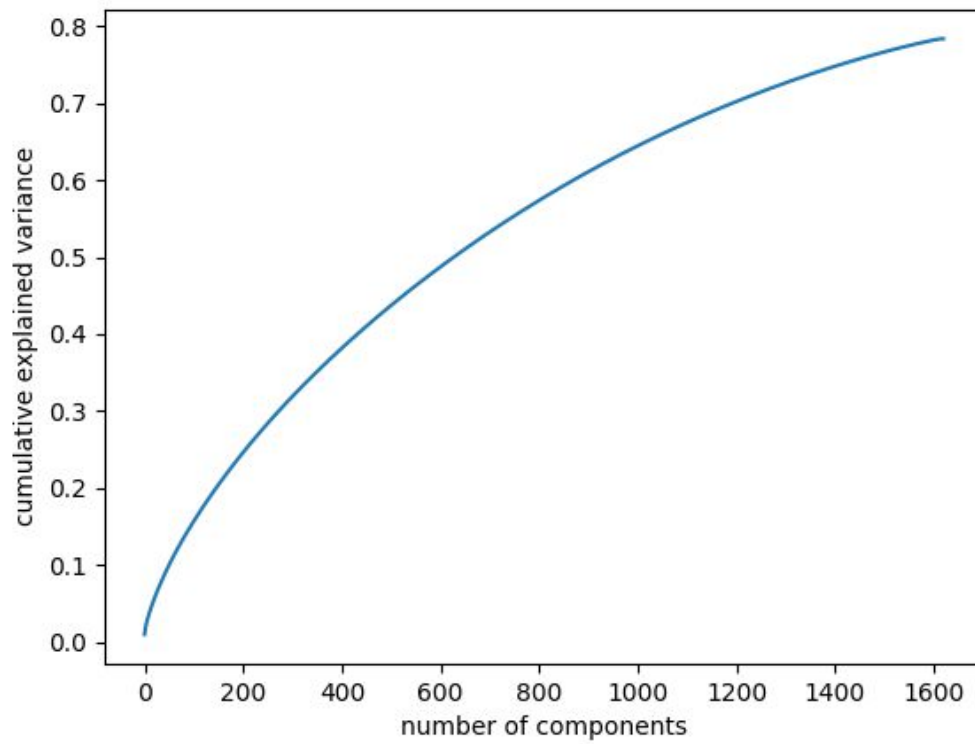


Figure 3: Cumulative explained variance of the first 1600 components

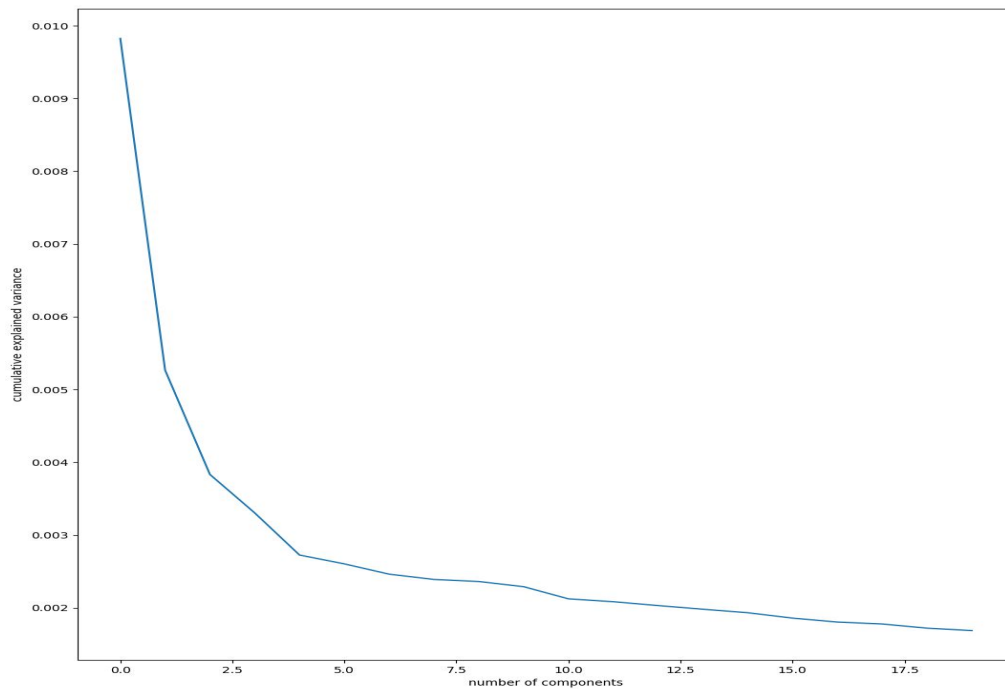


Figure 4: Individual explained variance of first 20 components