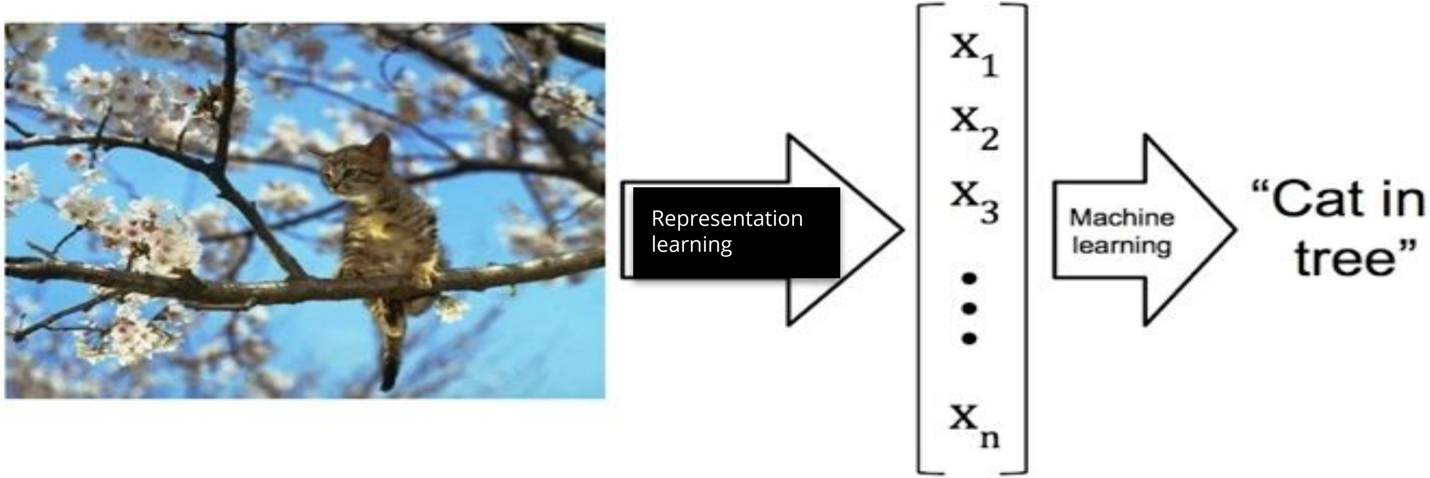
Multi view representation learning

B21PV02

Representation learning



Why representation learning?

- To model the input data => i.e., “understanding” the input
- Makes the machine learning task easier
  - Fewer examples will be needed for the prediction/classification task if relevant aspects of the input are already modeled.
  - Rich abstract representation of the data can make the prediction task simpler to be modeled.

Properties of the representations

- Smoothness
  - Small change in the input => small change in it’s representation
- Compactness
  - Remove redundancy in the input
- Abstraction
  - Recognize high-level properties shared between superficially dissimilar inputs.

Our main focus is on the abstraction methods using CCA

Correlated Representations

X1 : first view of the data

X2 : second view of the data

Now, to learn representations from this two-view data,

Learn the mappings f1 and f2, such that

$$\mathrm{corr}(f_1(x_1), f_2(x_2)) = \frac{\mathrm{cov}(f_1(x_1), f_2(x_2))}{\sqrt{\mathrm{var}(f_1(x_1)) \cdot \mathrm{var}(f_2(x_2))}}$$

the correlation between the learned representations is maximized.

Correlated representations

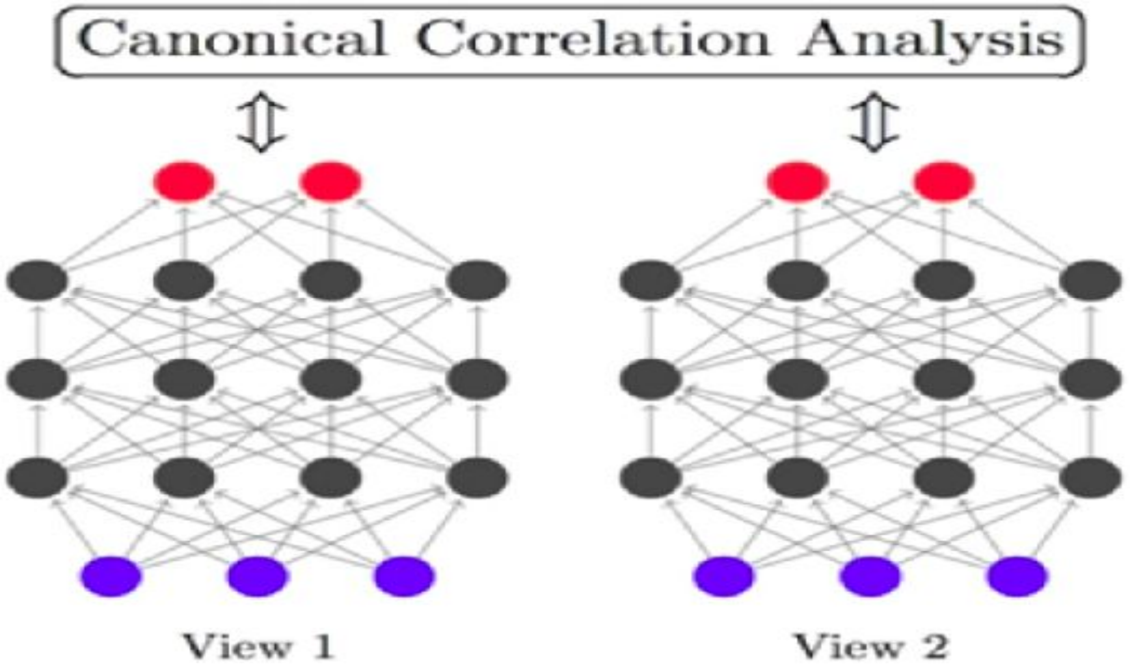
The correlated representations learned can

- Partially reconstruct information in the other view during the test time
- Remove noise that is uncorrelated across views.

CCA Methods for Representation Learning

- CCA
- Kernel CCA (kCCA)
- Deep CCA
- Generalized CCA
- Deep Generalized CCA (DGCCA)

Deep CCA - Two view non linear representation learning



Our implementation of DCCA

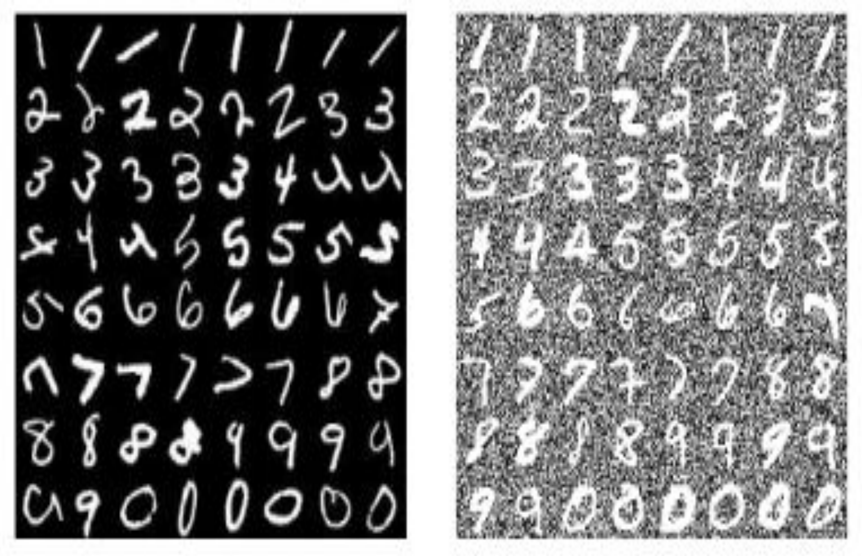
Dataset split : 60K/10K

Optimizer : RMSprop

Activation function : Sigmoid

Classifier used : SVM

Accuracy obtained : 96.64% on test data



Noisy MNIST dataset

Observations

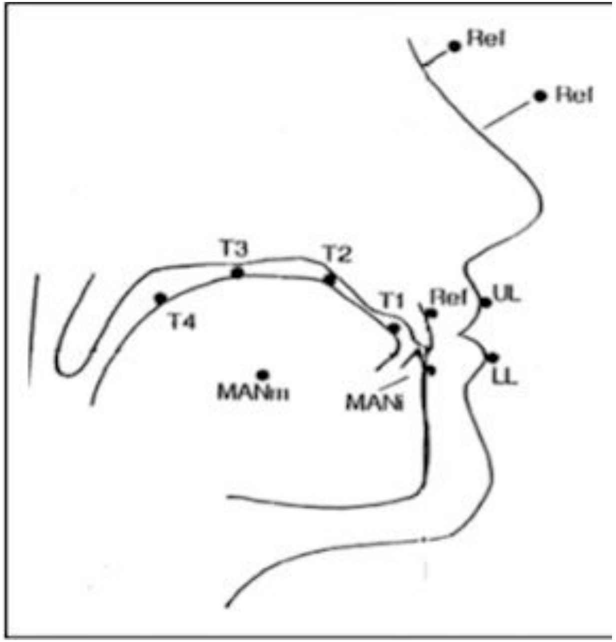
- DCCA learns deep representation mappings that uncover latent relationships in two views of data.
- It is a parametric method, unlike KCCA.
- Deeper representations can outperform shallow ones (even when the number of parameters is fixed).
- DCCA detects far more correlation than CCA or KCCA.

Application - Speech data

- Acoustic view
- Articulatory view

DCCA Acoustic features can model articulatory phenomena, improving the classifier performance.

Speech data - Wisconsin XRMB database



Extending to multi-view data (with more than two views)

- Initial Attempt : GCCA (multi-view linear representation learning)
  - Finding a shared representation G of J different views of the data

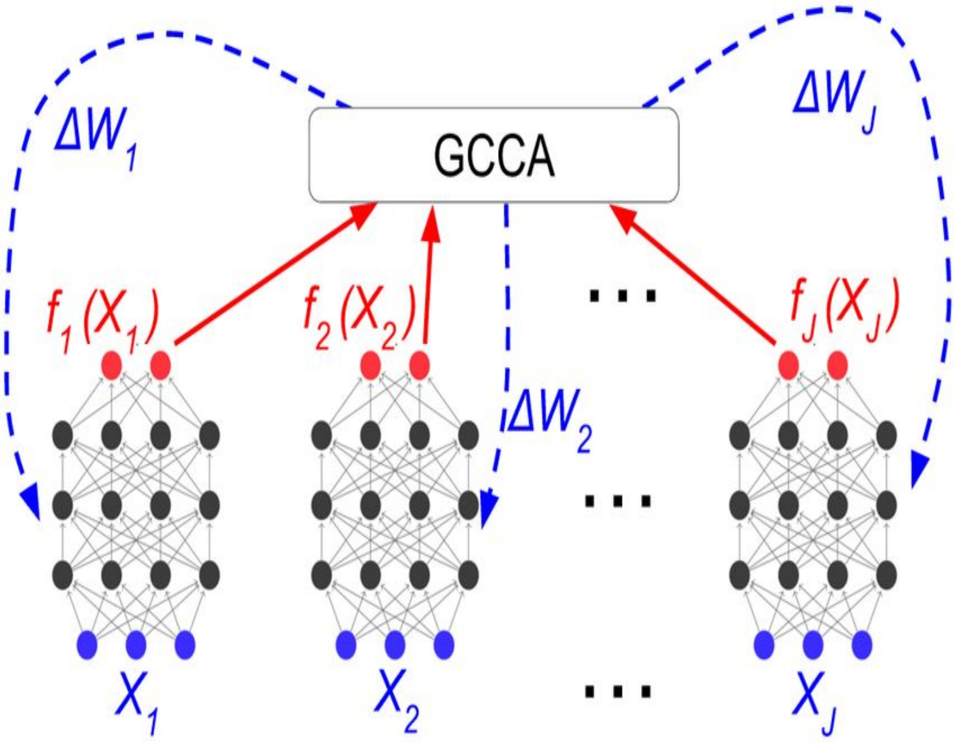
- Deep GCCA : multi-view non linear representation learning
  - Learning a non-linear representation for each view in order to maximize the correlation between the learnt representations across different views.

DGCCA - Deep Generalized CCA

Training

- Forward pass : Passing the input vectors in each view through multiple layers of non linear transformations.
- Backpropagation : Backpropagate the gradient of the GCCA objective wrt the network parameters to tune each view's network.

This is in a way similar to DCCA



Our Next Steps

- Multi view learning with DGCCA (more than two view data and non-linear representations)
- Applying it on a real-world problem
- Coming up with any potential improvements

References

- Benton, Adrian, et al. "Deep generalized canonical correlation analysis." arXiv preprint arXiv:1702.02519 (2017).
- Guo, Chenfeng, and Dongrui Wu. "Canonical Correlation Analysis (CCA) Based Multi-View Learning: An Overview." arXiv preprint arXiv:1907.01693 (2019).
- Weenink, David. "Canonical correlation analysis." Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam. Vol. 25. Amsterdam: University of Amsterdam, 2003.
- Andrew, Galen, et al. "Deep canonical correlation analysis." International conference on machine learning. PMLR, 2013.

Thank You