# Greedy Deep Transform Learning

Authors: Jyoti Maggu, Angshul Majumdar
Email: jyotim@iiitd.ac.in

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

# Contents

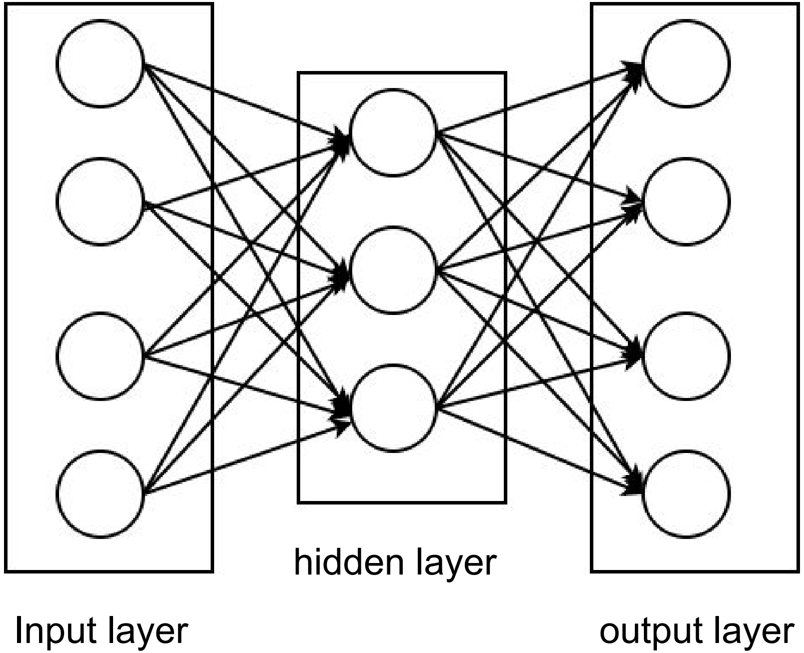
# Deep Transform Learning

- A new tool for Deep Learning.
- Stack one transform after another.
- Learning is done in greedy fashion.

# Deep Representation Learning

1. **Stacked AutoEncoder**

AutoEncoder:

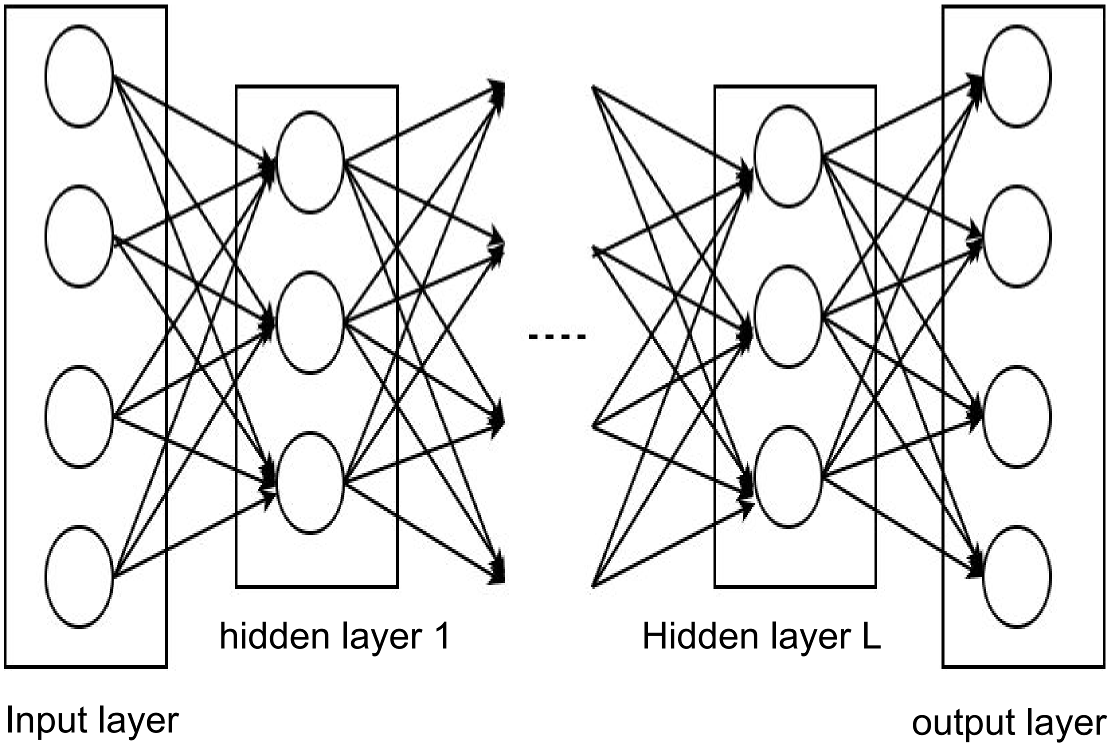


# Deep Representation Learning

1. **Stacked AutoEncoder**

- Nesting one AE inside another.
- Solved using greedy paradigm.
- Used generally for classification.

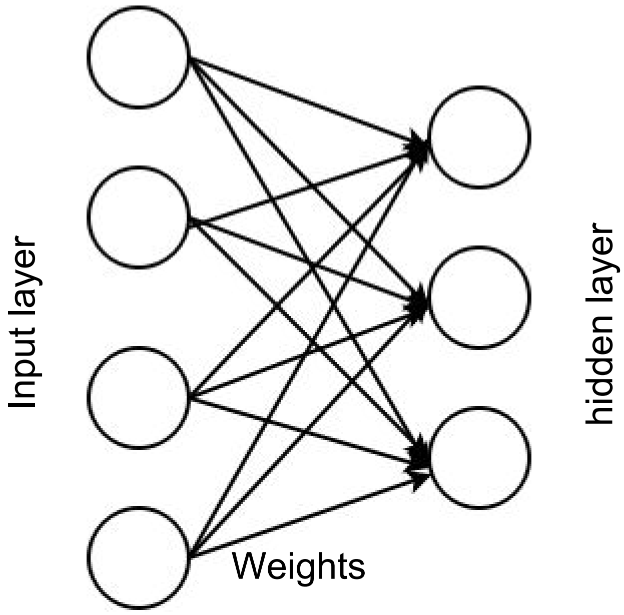
Nesting one AE into another



# Deep Representation Learning

1. Stacked AutoEncoder
2. **Deep Belief Network**
   - Undirected Graph Model.
   - Information content is preserved by cosine similarity between projection of data and learnt features.
   - Probabilistic Formulation.

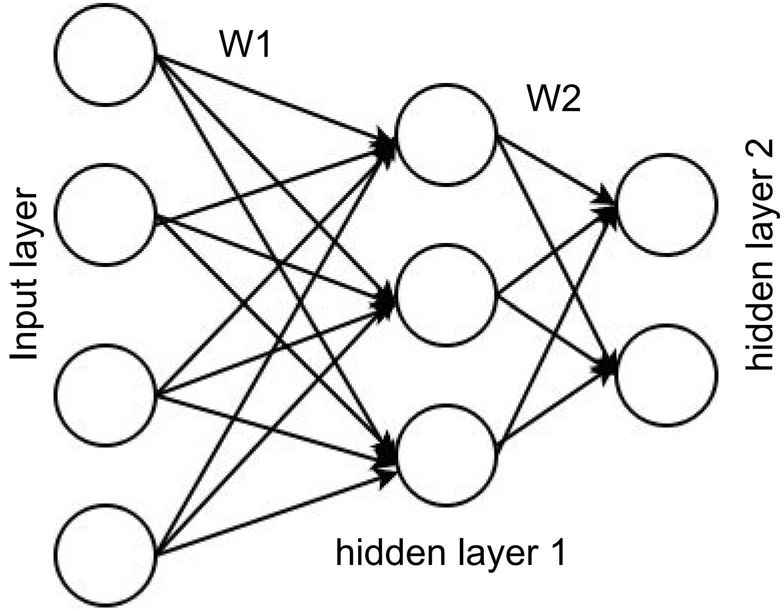
Restricted Boltzmann Machine



# Deep Representation Learning

1. Stacked AutoEncoder
2. **Deep Belief Network**

Deep Belief Network: Stacking one RBM into another.



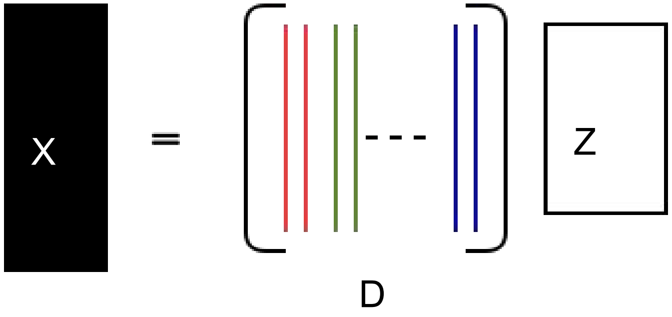
# Deep Representation Learning

1. Stacked AutoEncoder
2. Deep Belief Network
3. **Deep Dictionary Learning**

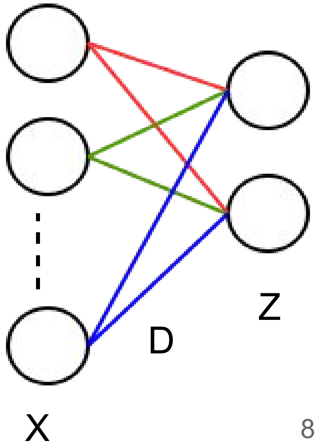
$$\min_{D,Z} ||X - DZ||_F^2 + \lambda ||Z||_1$$

## Dictionary Learning

- It learns basis for representing data.
- Columns of dictionary (Atoms) are connections between input and representation layer.
- X = D Z




NN Representation

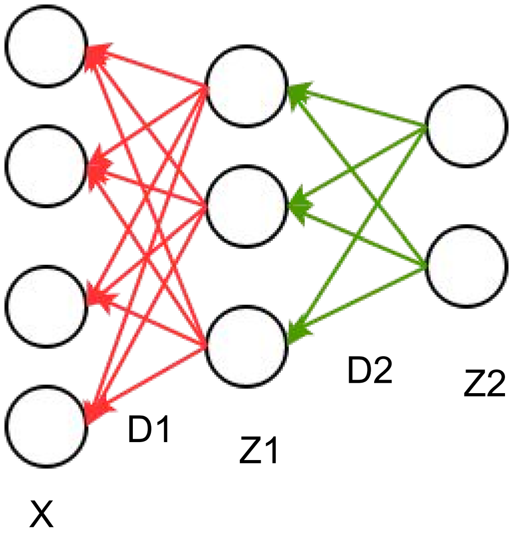


# Deep Representation Learning

1. Stacked AutoEncoder
2. Deep Belief Network
3. **Deep Dictionary Learning**

Deep Dictionary Learning

- Stacking one layer after another.
- Z1 = D2 Z2
- X= D1 Z1; X= D1 D2 Z2



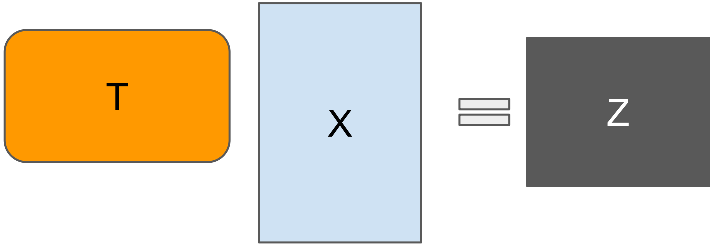
# Deep Representation Learning

1. Stacked AutoEncoder
2. Deep Belief Network
3. Deep Dictionary Learning
4. **Deep Transform Learning**

$$\min_{T,Z} ||TX-Z||_F^2 \quad + \quad \mu||Z||_0$$

# Transform Learning

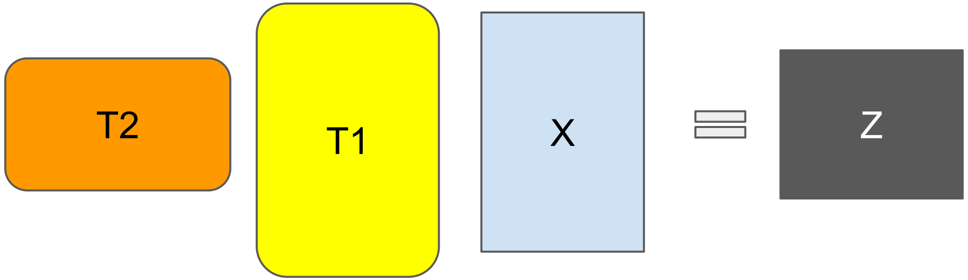
- While Dictionary Learning is a synthesis formulation, Transform Learning is its analysis equivalent.
- It learns an transform T such that it operates on the data X to generate the coefficients Z.



# Deep Representation Learning

1. Stacked AutoEncoder
2. Deep Belief Network
3. Deep Dictionary Learning
4. **Deep Transform Learning**

## Deep Transform Learning



# Problem Statement

- Given data from different classes; classify them accurately.
- Compare with existing techniques: SAE, DBN, DDL.
- Reduce train time and test feature generation time.

# Solution: Transform Learning

$$\min_{T,Z} ||TX - Z||_F^2 + \mu ||Z||_0$$

- But this leads to degenerate solution; trivial would be T=0, Z=0.
- So, to avoid this; new formulation becomes:

$$\min_{T,Z} ||TX - Z||_F^2 + \mu ||Z||_1 + \lambda(\varepsilon \; ||T||_F^2 - \; logdet \; T)$$

- This can be solved by alternating minimization iteratively.

$$Z \leftarrow \min_{Z} ||TX - Z||_F^2 + \mu ||Z||_0$$

$$T \leftarrow \min_{T} ||TX - Z||_F^2 + \lambda \; (\varepsilon \; ||T||_F^2 - \; logdet \; T)$$

# Greedy Deep Transform Learning

- Deeper representations are learnt by stacking one transform after another.
- The learning is done in a greedy fashion.

$$T_N(\phi...(T_2(\phi(T_1X)))) = Z$$

- By substituting, $T_{N-1}Z_{N-2} = \phi^{-1}(Z_{N-1})$ and so on, till $T_1X = \phi^{-1}(Z_1)$

- $$\min_{T_1Z_1} ||T_1X - Z_1||_F^2 + \lambda(||T_1||_F^2 - logdet\ T_1)$$

$$Z_1 = T_1X$$

$$\min_{T_2Z_2} ||T_2Z_1 - Z_2||_F^2 + \lambda(||T_2||_F^2 - logdet\ T_2)$$

# Experimental Results:

1. Accuracy improves by using Deep architectures.

Datasets used: MNIST, CIFAR-10 and SVHN

| Dataset | Level 1 | Level 2 | Level 3 |
|---|---|---|---|
| MNIST | 97.27 | 97.66 | **97.94** |
| CIFAR-10 | 81.12 | 81.89 | **82.60** |
| SVHN | 91.97 | 92.68 | **93.00** |

# Experimental Results:

1. Accuracy improves by using Deep architectures.
2. Results with Nearest Neighbours:

Datasets used: MNIST, CIFAR-10 and SVHN

Results with NN classifier using features from DTL:

| Dataset | Proposed | SDAE | DBN | DDL |
|---|---|---|---|---|
| MNIST | **97.94** | 97.33 | 97.05 | 97.75 |
| CIFAR-10 | **82.60** | 78.62 | 73.96 | 81.09 |
| SVHN | **93.00** | 91.11 | 88.29 | 92.26 |

# Experimental Results:

1. Accuracy improves by using Deep architectures.

2. Results with Nearest Neighbours:

3. Results with SRC:

Datasets used: MNIST, CIFAR-10 and SVHN

Results with SRC classifier using features from DTL:

| Dataset | Proposed | SDAE | DBN | DDL |
|---|---|---|---|---|
| MNIST | **98.96** | 98.33 | 98.43 | 98.81 |
| CIFAR-10 | **85.06** | 79.32 | 75.02 | 83.75 |
| SVHN | **94.55** | 92.05 | 90.11 | 93.62 |

# Experimental Results:

1. Accuracy improves by using Deep architectures.

2. Results with Nearest Neighbours:

3. Results with SRC:

4. Results with SVM:

Datasets used: MNIST, CIFAR-10 and SVHN

Results with SVM classifier using features from DTL:

| Dataset | Proposed | SDAE | DBN | DDL |
| --- | --- | --- | --- | --- |
| MNIST | **98.94** | 97.05 | 98.44 | 98.64 |
| CIFAR-10 | **85.55** | 78.90 | 74.30 | 84.96 |
| SVHN | **95.42** | 92.60 | 89.70 | 93.81 |

# Experimental Results:

1. Accuracy improves by using Deep architectures.
2. Results with Nearest Neighbours:
3. Results with SRC:
4. Results with SVM:
5. Feature generation time:

Time in seconds:

| Mode | Proposed | SDAE | DBN | DDL |
|---|---|---|---|---|
| Training | **25** | 120408 | 30071 | 107 |
| Testing | **50** | 61 | **50** | 79 |

# Conclusion

- Deep Transform Learning(proposed) outperforms DDL, SDAE and DBN in terms of accuracy.
- Features generated by DTL are good representations because all classifiers KNN, SRC and SVM are able to accurately classify test data.
- Train and Test time is less with proposed technique.

# Future Work

- Incorporate stochastic regularization techniques into DTL framework.
- Compare regularized DTL with advanced regularized tools like sparse AutoEncoder, Contractive AutoEncoder, sparse DBN.
- Making supervised DTL framework.

# Questions are welcome.