# Assignment 4

The due date for submitting this assignment has passed. **Due on 2018-02-21, 23:59 IST.**

## Submitted assignment

1) Given that $E_n$ and $D_n$ are encoder and decoder blocks of $n^{th}$ autoencoder in SAE (Stacked autoencoder). What is the sequence of blocks during end-to-end training of SAE (with $n = 2$)  

```
Input → $E_1$ → $D_1$ → $E_2$ → $D_2$ → Input
```

1 point

- Any of the above

**No, the answer is incorrect.**

Score: 0

**Accepted Answers:**

```
Input → $E_1$ → $E_2$ → $D_2$ → $D_1$ → Input
```

2) Given a trained SAE ($n = 2$), how should the blocks be arranged for weight refinement (classification task)  

```
Input → $E_1$ → $E_2$ → $D_2$ → $D_1$ → $D_1$ → Input
```

1 point

- Any of the above

**No, the answer is incorrect.**

Score: 0

**Accepted Answers:**

```
Input → $E_1$ → $E_2$ → $D_2$ → Logistic regression
```

3) Given input $x$ and linear autoencoder (no bias) with random weights ($W$ for encoder and $W'$ for decoder), what mathematical form is minimized to achieve optimal weights

\[
|x - (W' \cdot W \cdot x)|
\]

1 point

- $|x - (W' \cdot N_{ij}(W \cdot x))|$
- $|x - N_{ij}(W' \cdot N_{ij}(W \cdot x))|$
- None of the above
4) Given an linear autoencoder which encodes input $x$ to $z$. For learning hierarchically high-level representation what should be the learning arrangement of second linear autoencoder (with weights $W_2$ and $W'_2$)

- $|x - (W'_2 \cdot W_2 \cdot x)|$
- $|z - (W'_2 \cdot W_2 \cdot z)|$
- $|z - (W'_2 \cdot W_1 \cdot x)|$
- None of the above

No, the answer is incorrect.  
Score: 0  
Accepted Answers:  
$|z - (W'_2 \cdot W_2 \cdot z)|$

5) In a de-noising autoencoder, noise is added to input $x$ for

- Avoiding overfitting
- Robust feature extraction
- Data augmentation
- All of the above

No, the answer is incorrect.  
Score: 0  
Accepted Answers:  
All of the above

6) In a linear autoencoder (without regularizer), if hidden layer perceptrons are equal to input layer perceptron then encoder and decoder weights are indulged to learn

- Optimal representations
- Identity matrix
- Sparse representations
- None of the above

No, the answer is incorrect.  
Score: 0  
Accepted Answers:  
Identity matrix

7) The role of regularizer in cost function is to

- Avoid overfitting
- Induce sparsity
- Simpler hypothesis
- All of the above

No, the answer is incorrect.  
Score: 0  
Accepted Answers:  
All of the above

8) Given feature vector $X$ and corresponding label ($y$), logistic regression relates $X$ and $y$ in the form of ($B$ is parameters to be learned)
$log\left( \frac{y}{1-y} \right) = BX$

$y = BX$

$y = \frac{1}{1+e^{BX}}$

None of the above

**No, the answer is incorrect.**

**Score: 0**

**Accepted Answers:**

$log\left( \frac{y}{1-y} \right) = BX$

9) Given feature vector $X$ (with dimension $j$), corresponding label $y$ (binary class) and weights ($b_1, b_2, \ldots, b_k$) of logistic regressor, $z$ be output (expected) of logistic regressor. What is the loss function ($L$) and gradient computed to correct the $b_k$ based on chain rule

- $ylog(z)$ and $(y - z)x_k$
- $y - \log(z)$ and $\frac{\partial \log(z)}{\partial y}$
- $ylog(z) + (1 - y)(1 - \log(z))$ and $\frac{\partial L}{\partial z} \cdot \frac{\partial \log(z)}{\partial y}$
- $ylog(z) + (1 - y)(1 - \log(z))$ and $(y + z)x_k$

**No, the answer is incorrect.**

**Score: 0**

**Accepted Answers:**

$ylog(z) + (1 - y)(1 - \log(z))$ and $\frac{\partial L}{\partial z} \cdot \frac{\partial \log(z)}{\partial y}$

10) What are the advantages of initializing MLP with pretrained autoencoder weights ($1$ point)

(i) Faster Convergence
(ii) Avoid overfitting
(iii) Simpler hypothesis

- (i) and (ii)
- (ii) and (iii)
- (i) and (iii)
- All of the above

**No, the answer is incorrect.**

**Score: 0**

**Accepted Answers:**

(i) and (ii)