

---

**1. (2 points)** The perceptron algorithm was used for training of a single neuron with a hard threshold for binary classification tasks. Which is an assumption used for this algorithm to converge?

- **A.** The data can be represented using linear models.
- **B.** The learning rate has to be equal to 1.
- **C.** There exists a single line (or hyper-plane) that separates the points from both classes.
- **D.** The data is linearly separable.
- **E.** The initial value for the weight vector  $w$  has to be set to zero.

*Answer(s) submitted:*

- D

(correct)

---

**2. (2 points)** Which regularization penalty encourage sparsity?

- **A.** L1 penalty.
- **B.** L2 penalty.
- **C.** Square penalty.
- **D.** Lp penalty for  $p \geq 2$ .

*Answer(s) submitted:*

- A

(correct)

---

**3. (2 points)** When performing dropout, what component from the network is dropped randomly?

- **A.** Neurons in a layer

- **B.** Inputs for neuron
- **C.** Biases in the neuron
- **D.** Weight matrices in the neurons
- **E.** Inputs for the network

*Answer(s) submitted:*

- A

(correct)

---

**4. (2 points)** For the variational adversarial generation strategy, did this approach ensure finding an adversarial sample for every possible input?

- **A.** Yes
- **B.** No
- **C.** Sometimes
- **D.** For some family of inputs

*Answer(s) submitted:*

- A

(correct)

---

**5. (2 points)** For the differential adversarial generation strategy, did this approach ensure finding an adversarial sample for every possible input?

- **A.** Yes
- **B.** No
- **C.** Sometimes
- **D.** For some family of inputs

*Answer(s) submitted:*

- B

(correct)