1 point

1. Choose the correct statements about MLP

- [ ] MLP can have only 1 hidden layer
- [ ] The first hidden layer contains predictions for your task
- [ ] MLP with a linear activation function is better than a linear model
- [x] We can train MLP with SGD
- [x] A hidden layer of MLP automatically learns new helpful features for the task

---

1 point

2. Apply a chain rule to calculate $\frac{\partial a}{\partial x}$ where $a(x, y) = sin(xy) \cdot e^x$.

Here is an example of the syntax: sin(x*y)*exp(x)

Preview

$$(y \cos(xy) + \sin(xy))\, e^x$$

exp(x)* (y*cos(y*x) + sin(y*x))

---

1 point

3. Choose the correct statements about backpropagation

- [x] It is an efficient way to apply a chain rule
- [x] It is the way to train modern neural networks
- [ ] It is done in one pass
- [ ] You can use non-differentiable loss to train your MLP

---

1 point

4. What is the time complexity of backpropagation algorithm w.r.t. number of edges N in the computational graph?

- (●) $O(N)$
- ( ) $O(N^2)$
- ( ) $O(N!)$
- ( ) $O(log(N))$