CSE 234 Fall 2021 Surprise Quiz 3

Full Name: _____________________________  PID: _____________________________

INSTRUCTIONS:

1. This quiz is for **10min** and **10pts**. It is *open* book/notes/electronics.

2. Mark/write your answers using a dark pen or pencil. Ensure your writing is clear and legible.

**Q1. [2pts]** Which of the following model building steps will likely *not raise* the bias component of test error when training a DL model?

(A) Reducing the number of layers  
(B) Strengthening regularization  
(C) Reducing number of neurons per layer  
(D) Raising the learning rate

**Q2. [2pts]** Which of the following distributed SGD approaches communicates model updates only once per shard during an epoch?

(A) Parameter Server  
(B) Horovod  
(C) MOP  
(D) None of the others

**Q3. [2 x 3pts]** Suppose you run Hyperband for hyperparameter tuning of an ML classifier with the following knobs: $R = 16, \eta = 4$. It yields the following brackets in the same table format from the paper as explained in class:

<table>
<thead>
<tr>
<th></th>
<th>$s = 2$</th>
<th>$s = 1$</th>
<th>$s = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n_1$</td>
<td>$n_1$</td>
<td>$n_1$</td>
</tr>
<tr>
<td>$i$</td>
<td>$n_i$</td>
<td>$n_i$</td>
<td>$n_i$</td>
</tr>
<tr>
<td>0</td>
<td>16</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

1. How many configs got killed in the bracket with $s = 2$?

2. How many configs ran for at least 1 epoch?
CSE 234 Fall 2021 Surprise Quiz 3

Full Name : PID :

INSTRUCTIONS:

1. This quiz is for **10min** and **10pts**. It is open book/notes/electronics.

2. Mark/write your answers using a dark pen or pencil. Ensure your writing is clear and legible.

**Q1. [2pts]** Which of the following model building steps will likely *not raise* the bias component of test error when training a DL model?

(A) Reducing the number of layers  (B) Strengthening regularization

(C) Reducing number of neurons per layer  (D) Raising the learning rate

**Q2. [2pts]** Which of the following distributed SGD approaches communicates model updates only once per shard during an epoch?

(A) Parameter Server  (B) Horovod  (C) MOP  (D) None of the others

**Q3. [2 x 3pts]** Suppose you run Hyperband for hyperparameter tuning of an ML classifier with the following knobs: \( R = 16; \eta = 4 \). It yields the following brackets in the same table format from the paper as explained in class:

<table>
<thead>
<tr>
<th>i</th>
<th>s = 2</th>
<th>s = 1</th>
<th>s = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n _i )</td>
<td>( n _i )</td>
<td>( n _i )</td>
</tr>
<tr>
<td>0</td>
<td>16</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

1. How many configs got killed in the bracket with \( s = 1 \)?

2. How many configs ran for at least 5 epochs?
CSE 234 Fall 2021 Surprise Quiz 3

INSTRUCTIONS:

1. This quiz is for 10 min and 10 pts. It is open book/notes/electronics.

2. Mark/write your answers using a dark pen or pencil. Ensure your writing is clear and legible.

Q1. [2 pts] Which of the following model building steps will likely not raise the bias component of test error when training a DL model?

(A) Reducing the number of layers  
(B) Strengthening regularization  
(C) Reducing number of neurons per layer  
(D) Raising the learning rate

Q2. [2 pts] Which of the following distributed SGD approaches communicates model updates only once per shard during an epoch?

(A) Parameter Server  
(B) Horovod  
(C) MOP  
(D) None of the others

Q3. [2 x 3 pts] Suppose you run Hyperband for hyperparameter tuning of an ML classifier with the following knobs: \( R = 16; \eta = 4 \). It yields the following brackets in the same table format from the paper as explained in class:

<table>
<thead>
<tr>
<th>i</th>
<th>s = 2</th>
<th>s = 1</th>
<th>s = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n_i )</td>
<td>( \bar{n} )</td>
<td>( n_i )</td>
</tr>
<tr>
<td>0</td>
<td>16</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

1. How many configs got killed in the bracket with \( s = 0 \)?

2. How many configs ran for at least 2 epochs?