Model performance evaluation proposal:

Motivation:

1. RegenHD(BaseChangeHD) currently works best under low hyper dimension, so to show its effectiveness, we need to compare the accuracy between it and the baseline in low dimension.

2. This makes the model hard to converge, especially when the base regeneration frequency is high. Even when we stop regenerating, the training may still not converge (equally hard compared to baseline).

(shs, 5 iteration/regeneration, D = 2000, regen rate = 0.2, did not converge in 200 iteration)

3. The model usually reaches near-optimal test accuracy before converging. Therefore, convergence may not be necessary to get a good accuracy.

🡺 Instead of waiting for convergence, we can limit the number of train iterations and still get a good result.

Variables:

update frequency,

iterations trained (during regen, post regen),

effective dimensions,

regen rate (drop rate)

**Evaluation method 1:**

Fix dimension D(=200, 500, 1k, ), selects a set of regen rate and update frequency

Train the model **to a certain effective dimension**, say D'

Compare accuracy and iterations among:

1) baseline model with Dimension = D,

2) regenHD model(s),

3) baseline model with Dimension = D', and

~~4) possibly figure out the dimension of a baseline model that has the same accuracy as the regenHD model. Say its dimension is D''. (Maybe not)~~

Chart for each dataset may look like: (Config includes (regen rate, update frequency))

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Config \ D’ | 1000 | 2000 | 3000 | 4000 |
| 1)Baseline – D | N/A | (One entry because dimension is fixed to D) | | | |
| 2)Regen model | (0.1, 1-iteration) | Accuracy, D’’ |  |  |  |
| (0.1, 5-iteration) |  |  |  |  |
| (0.2, 1-iteration) |  |  |  |  |
| ….. |  |  |  |  |
| 3)Baseline – D’ | N/A |  |  |  |  |

Prediction of outcome: (to show)

1. **Accuracy: 3) >= 2) = 4) >> 1)**
2. Iteration: 1) and 2) will not converge, so a maximum iteration will be reached. 3) Should converge when D' is high enough, so there is a cut off line there (refer to table).(But **D’ baseline train slower** ) 4) usually D'' should be smaller than D', so it will converge slower.

**(Evaluation method 2:**

Fix dimension D(=200), **fix max iterations** (to 200, 300, 400 …. maybe)

Energy efficiency: max iteration should be kind of linear to D’

Train the model **over a set of regen rate and update frequency**,

Compare accuracy and iterations among:

1) baseline model with Dimension = D,

2) regenHD model(s), and maybe

3) baseline model with Dimension = D' (From baseline accuracy table)

Chart for each dataset may look like: (Config is (regen rate, update frequency))

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Config \ Iteration | 100 | 200 | 300 | 400 |
| 1)Baseline – D | N/A |  |  |  |  |
| 2)Regen model | (0.1, 1-iteration) | (Accuracy, D’) |  |  |  |
| (0.1, 5-iteration) |  |  |  |  |
| (0.2, 1-iteration) |  |  |  |  |
| ….. |  |  |  |  |

Should be companied by the baseline accuracy table.

Prediction of outcome:

1. Accuracy: 3) >= 2) >> 1); accuracy may saturate on large number of iterations.
2. Effective dimension for each cell: easily calculated

**Evaluation method 3: (might be good complement for others)**

Fix regen rate, update frequency, and max iterations.

Train the model **on different initial dimensions (D).**

Compare accuracy and iterations among:

1) baseline model with Dimension = D,

2) regenHD model(s),

)