

Week16

Uncertainty Guided Continual Learning with Bayesian Neural Networks

Name: Yushuo Wang

Which samples to be replayed

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7599% ,0.9474% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8177% ,0.8911% ,0.8349% ,0.0000% ,0.0000% ,  
    ,0.7927% ,0.8625% ,0.7370% ,0.9385% ,0.0000% ,  
    ,0.7776% ,0.8828% ,0.7427% ,0.8958% ,0.9005% ,  
ACC: 0.8399%  
  
BWT : -0.06%
```

Choosing 1% **Median** variance samples



```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7719% ,0.9432% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7776% ,0.9068% ,0.8161% ,0.0000% ,0.0000% ,  
    ,0.7792% ,0.8813% ,0.7484% ,0.9302% ,0.0000% ,  
    ,0.7547% ,0.8724% ,0.7078% ,0.8609% ,0.8901% ,  
ACC: 0.8172%  
  
BWT : -0.07%
```

Choosing 1% **Random** variance samples

10 epochs

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7099% ,0.9516% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7932% ,0.8469% ,0.8245% ,0.0000% ,0.0000% ,  
    ,0.7807% ,0.8427% ,0.7396% ,0.9505% ,0.0000% ,  
    ,0.7547% ,0.8672% ,0.7484% ,0.8776% ,0.9380% ,  
ACC: 0.8372%  
  
BWT : -0.07%
```

Choosing 1% **Median** variance samples



```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7828% ,0.9479% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7589% ,0.8906% ,0.8318% ,0.0000% ,0.0000% ,  
    ,0.7615% ,0.8719% ,0.6979% ,0.9344% ,0.0000% ,  
    ,0.7786% ,0.8714% ,0.7167% ,0.8583% ,0.9286% ,  
ACC: 0.8307%  
  
BWT : -0.07%
```

Choosing 1% **Random** variance samples

20 epochs

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8438% ,0.9302% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8401% ,0.9214% ,0.7760% ,0.0000% ,0.0000% ,  
    ,0.8281% ,0.8995% ,0.7661% ,0.9438% ,0.0000% ,  
    ,0.8245% ,0.9120% ,0.7557% ,0.9026% ,0.9401% ,  
ACC: 0.8670%  
  
BWT : -0.02%  
ACC: 0.8673%  
  
BWT : -0.02%
```

Choosing top 5% **Median** variance samples



```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8302% ,0.9484% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8578% ,0.9172% ,0.7818% ,0.0000% ,0.0000% ,  
    ,0.8307% ,0.8948% ,0.7557% ,0.9240% ,0.0000% ,  
    ,0.8406% ,0.8964% ,0.7354% ,0.8849% ,0.8906% ,  
ACC: 0.8496%  
  
BWT : -0.03%  
ACC: 0.8741%  
  
BWT : -0.02%
```

Choosing 5% **Random** variance samples

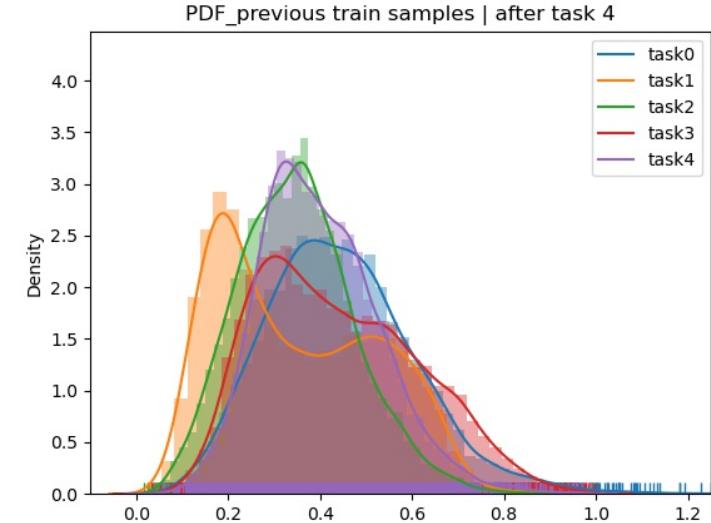
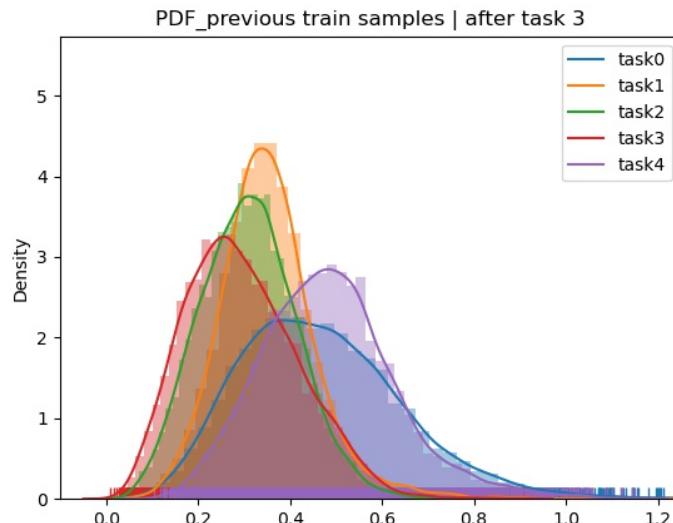
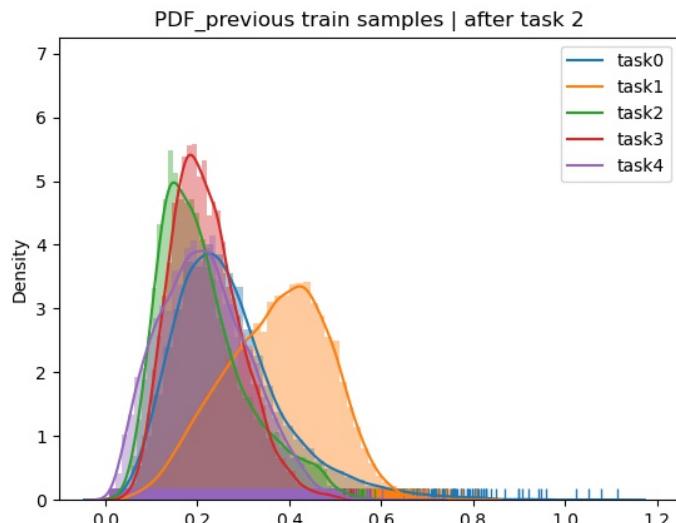
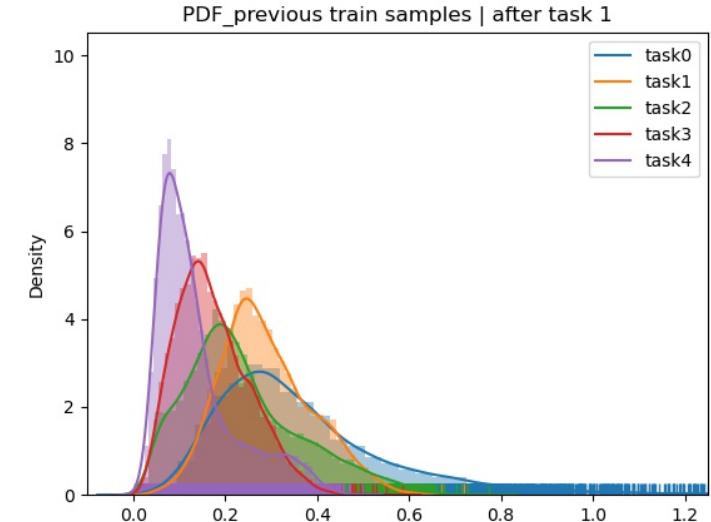
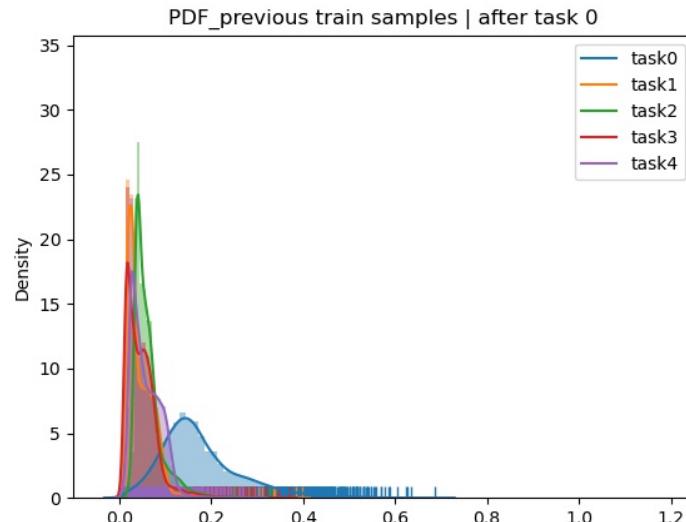
20 epochs

Random - not stable

Variance of Training samples

```
Accuracies =  
[0.8531%, 0.0000%, 0.0000%, 0.0000%, 0.0000%,  
0.8302%, 0.9484%, 0.0000%, 0.0000%, 0.0000%,  
0.8578%, 0.9172%, 0.7818%, 0.0000%, 0.0000%,  
0.8307%, 0.8948%, 0.7557%, 0.9240%, 0.0000%,  
0.8406%, 0.8964%, 0.7354%, 0.8849%, 0.8906%]  
ACC: 0.8496%  
  
BWT : -0.03%
```

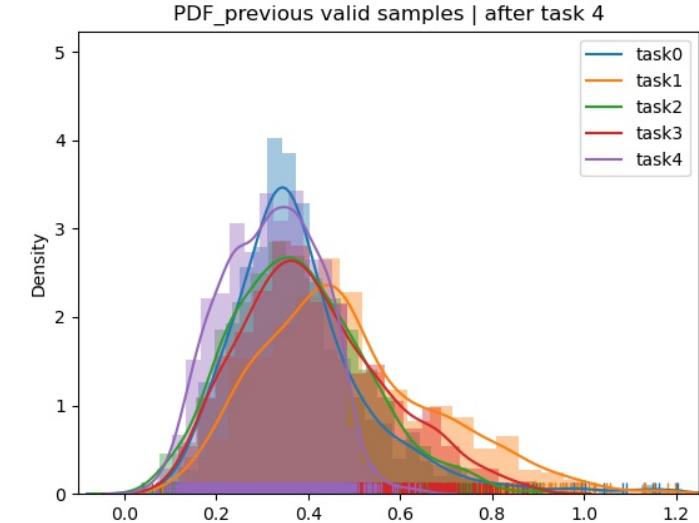
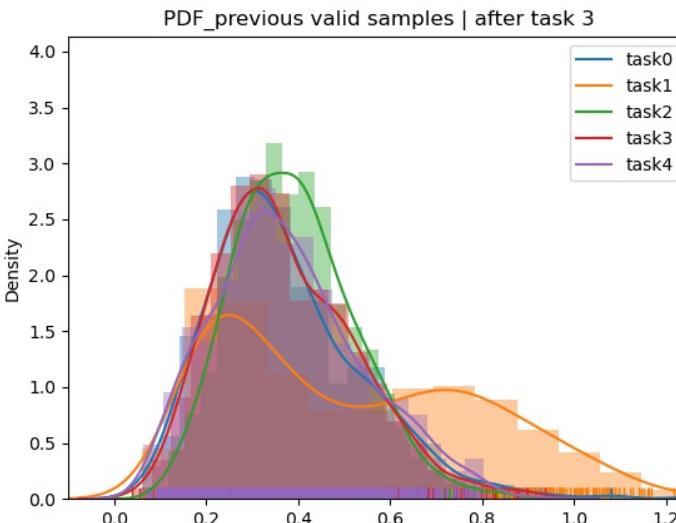
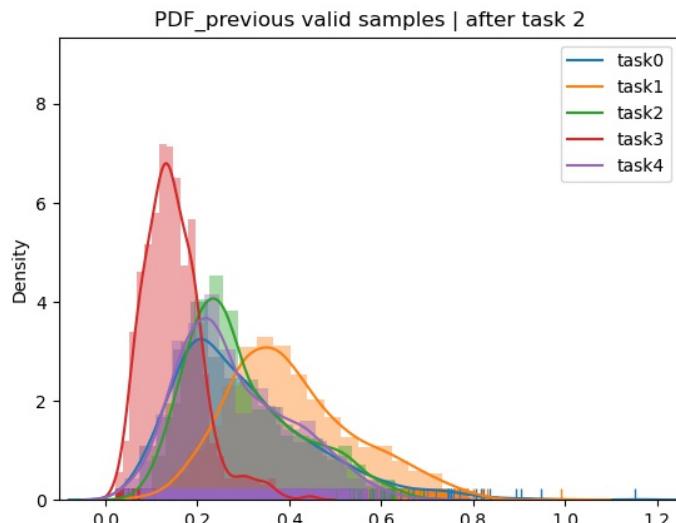
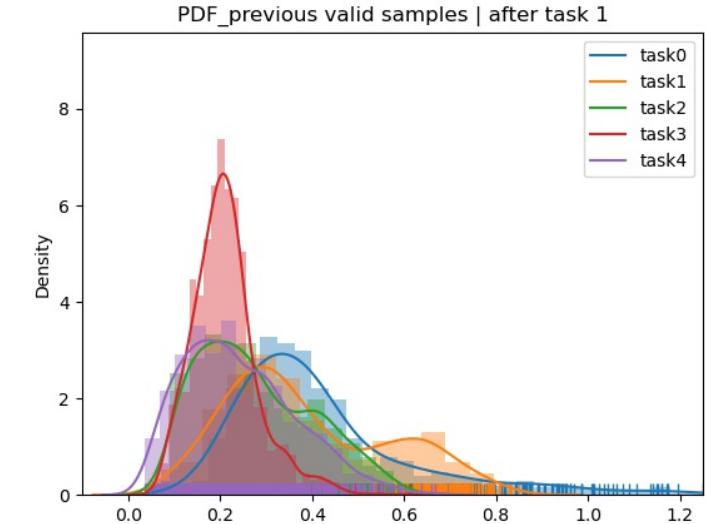
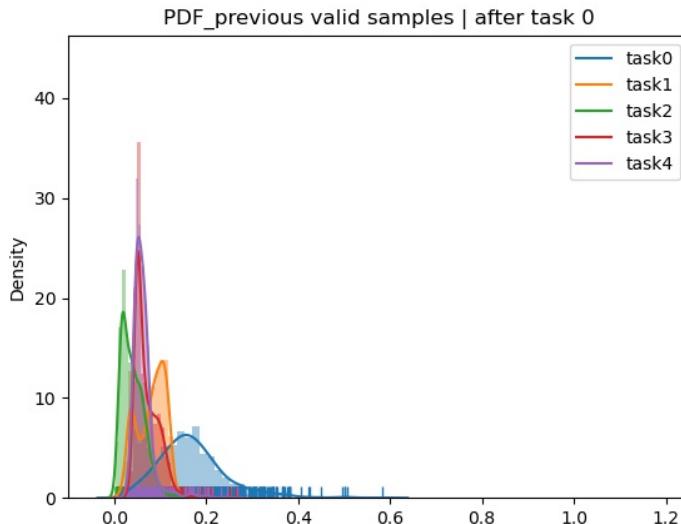
Choosing 5% **Random** variance samples
20 epochs



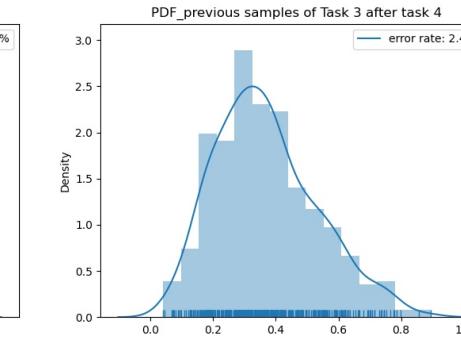
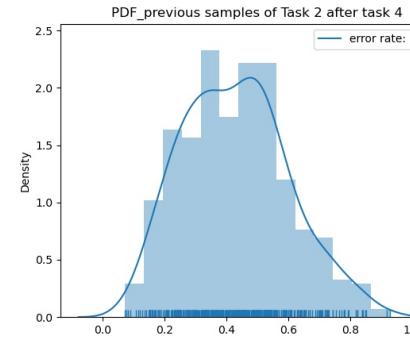
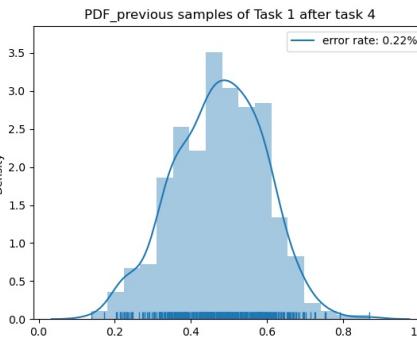
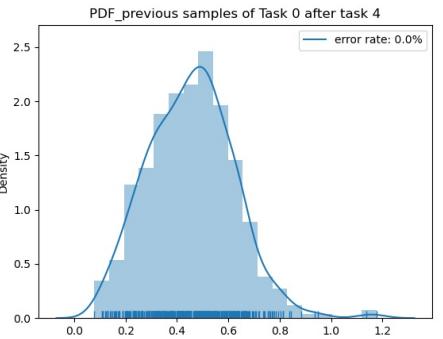
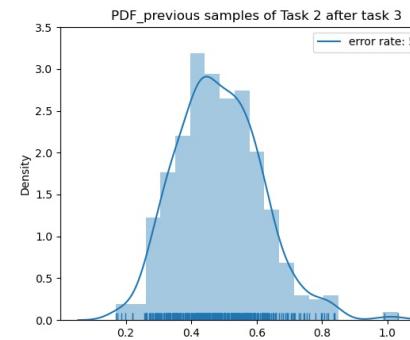
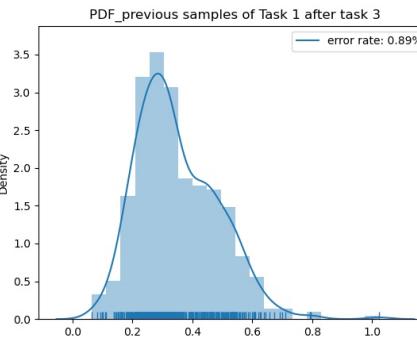
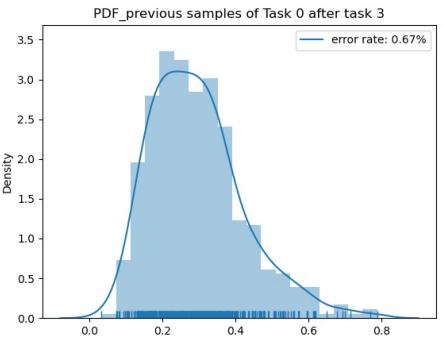
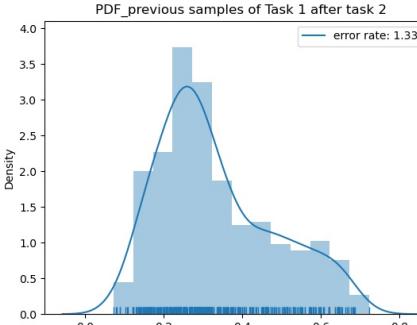
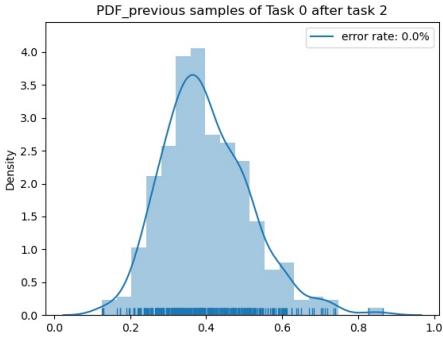
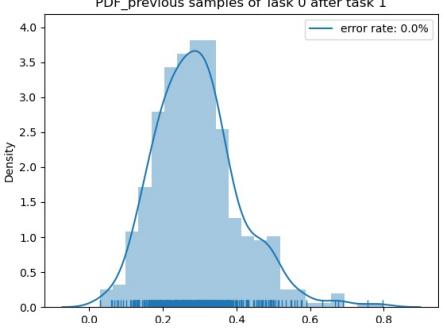
Variance of Valid samples

```
Accuracies =  
, 0.8531%, 0.0000%, 0.0000%, 0.0000%, 0.0000%,  
, 0.8302%, 0.9484%, 0.0000%, 0.0000%, 0.0000%;  
, 0.8578%, 0.9172%, 0.7818%, 0.0000%, 0.0000%;  
, 0.8307%, 0.8948%, 0.7557%, 0.9240%, 0.0000%;  
, 0.8406%, 0.8964%, 0.7354%, 0.8849%, 0.8906%;  
ACC: 0.8496%  
  
BWT : -0.03%
```

Choosing 5% Random variance samples
20 epochs



Training from task 1 to 4



How the variance of choosing samples changes

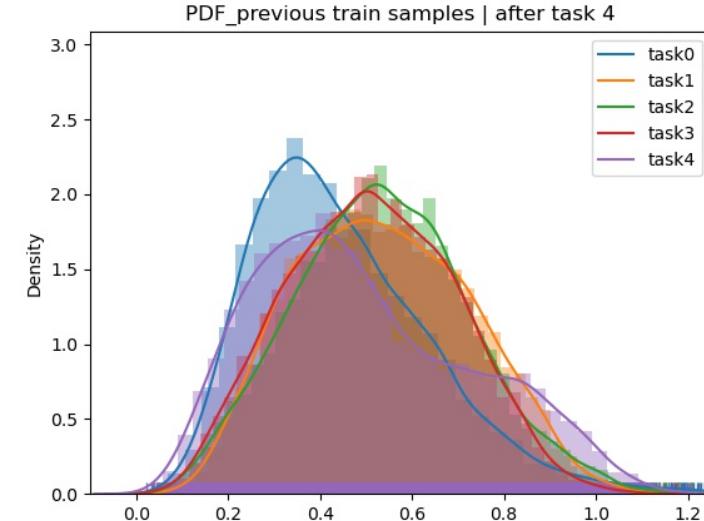
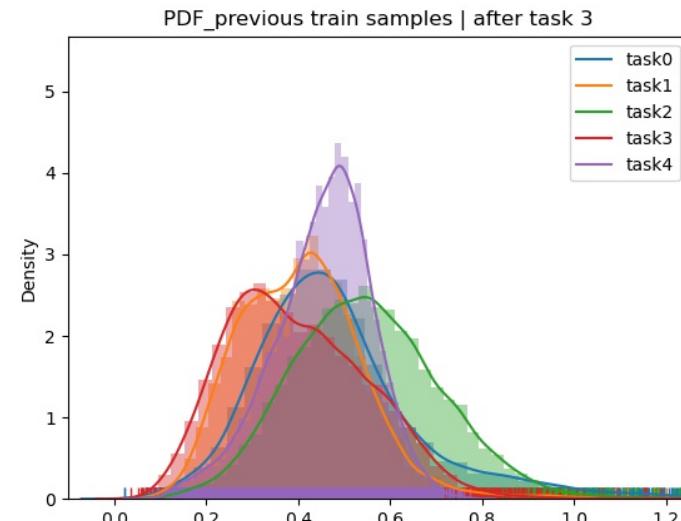
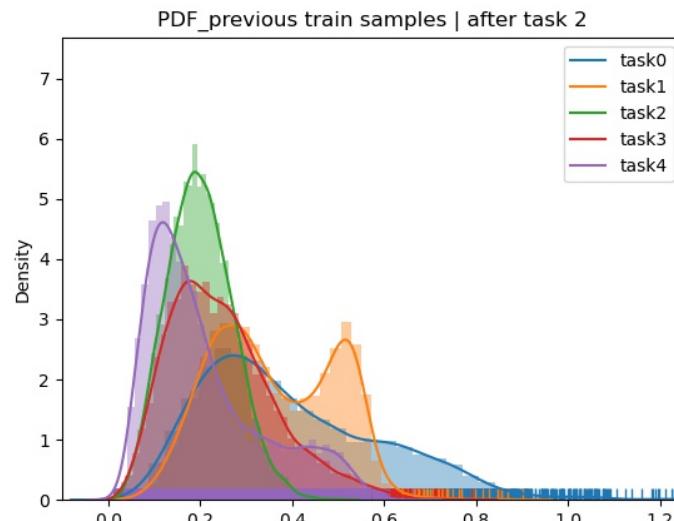
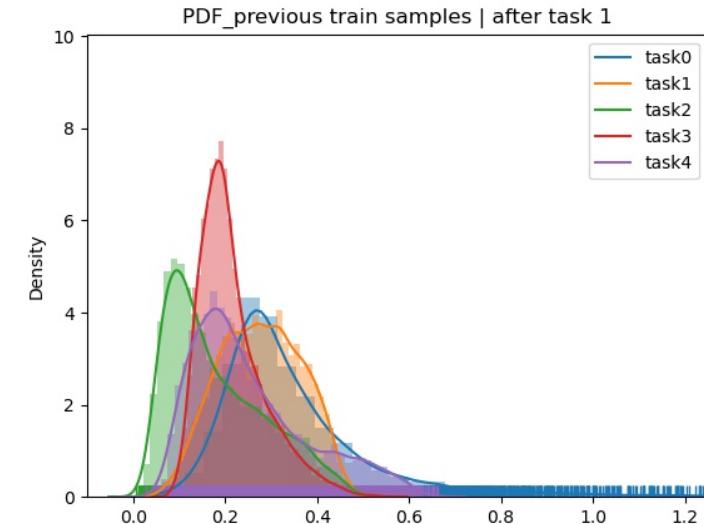
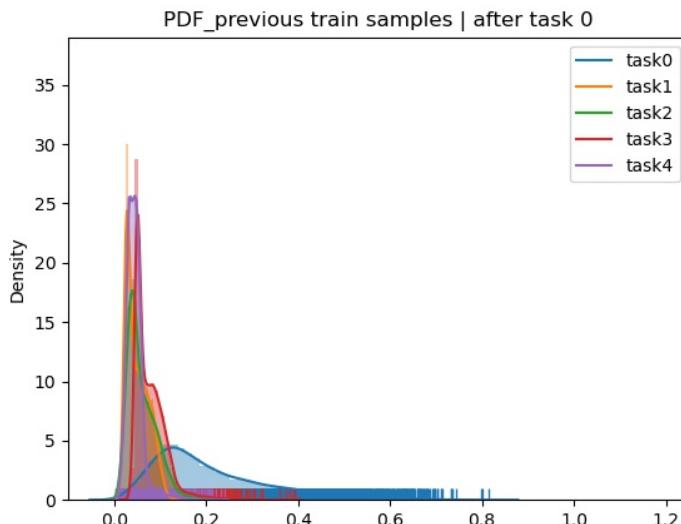
```
Accuracies =
, 0.8531% , 0.0000% , 0.0000% , 0.0000% , 0.0000% ,
, 0.8302% , 0.9484% , 0.0000% , 0.0000% , 0.0000% ,
, 0.8578% , 0.9172% , 0.7818% , 0.0000% , 0.0000% ,
, 0.8307% , 0.8948% , 0.7557% , 0.9240% , 0.0000% ,
, 0.8406% , 0.8964% , 0.7354% , 0.8849% , 0.8906% ,
ACC: 0.8496%
BWT : -0.03%
```

Choosing 5% Random variance samples
20 epochs

Variance of Training samples

```
Accuracies =  
    , 0.8531% , 0.0000% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.8438% , 0.9302% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.8401% , 0.9214% , 0.7760% , 0.0000% , 0.0000% ,  
    , 0.8281% , 0.8995% , 0.7661% , 0.9438% , 0.0000% ,  
    , 0.8245% , 0.9120% , 0.7557% , 0.9026% , 0.9401% ,  
ACC: 0.8670%  
  
BWT : -0.02%
```

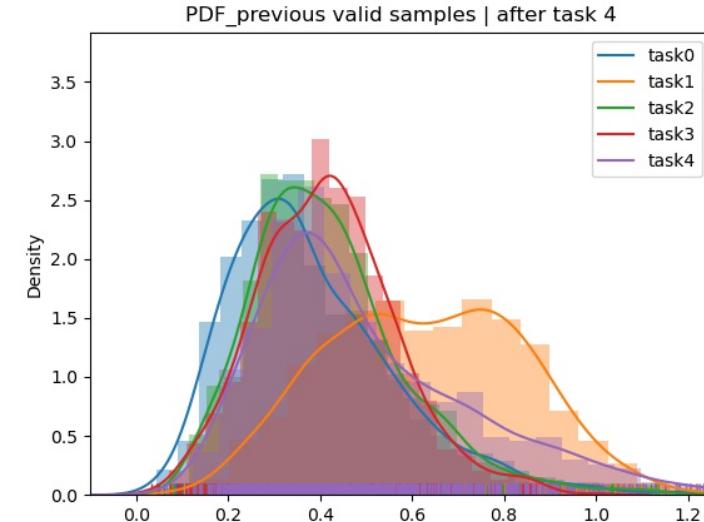
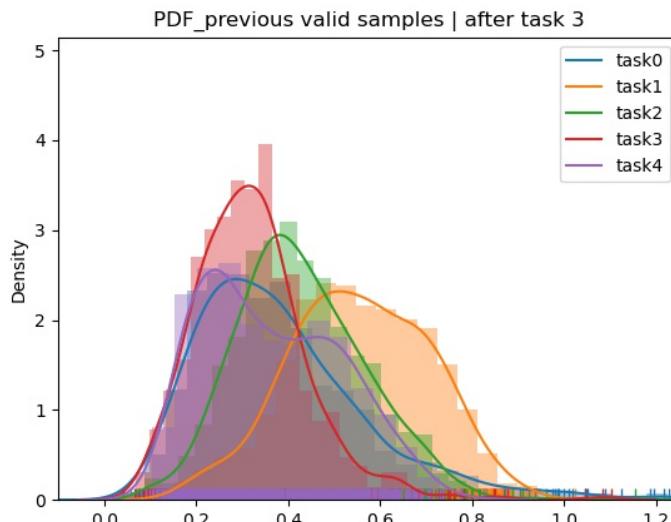
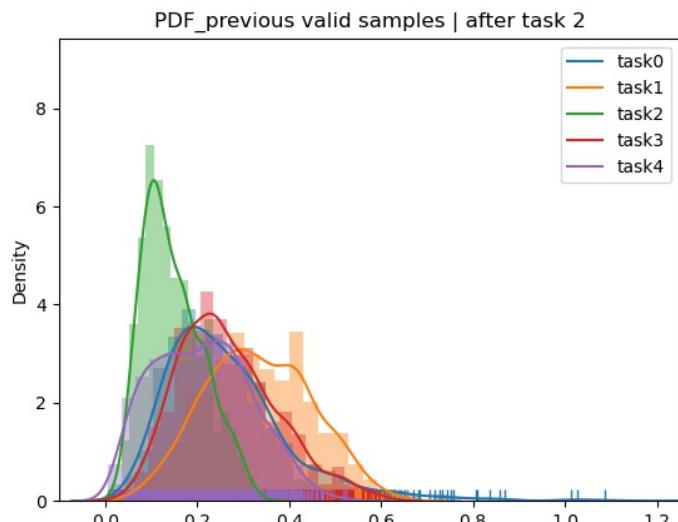
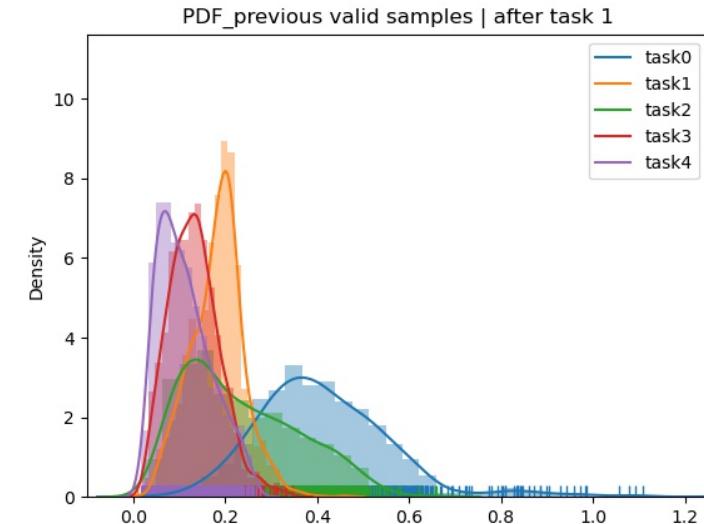
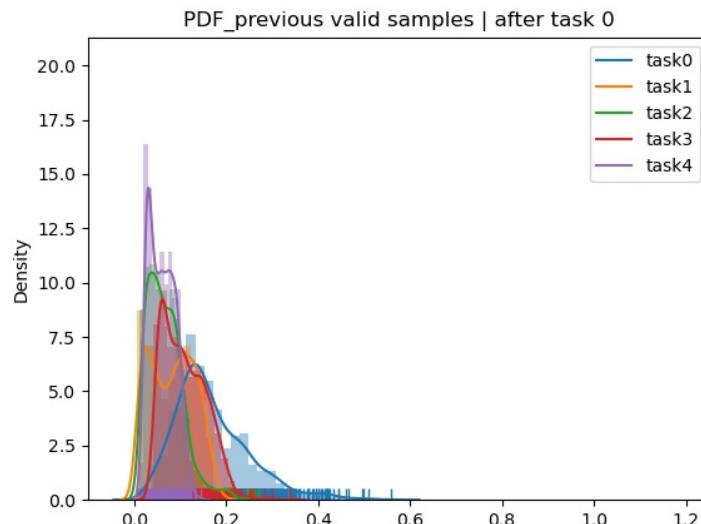
Choosing 5% **Interval** variance samples
20 epochs



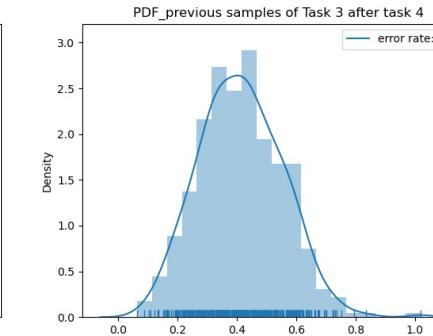
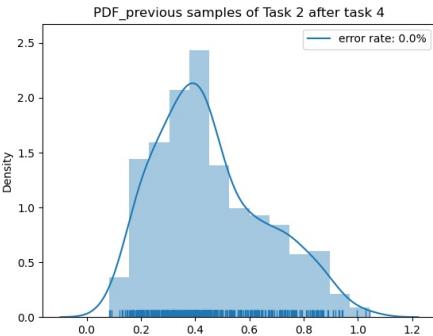
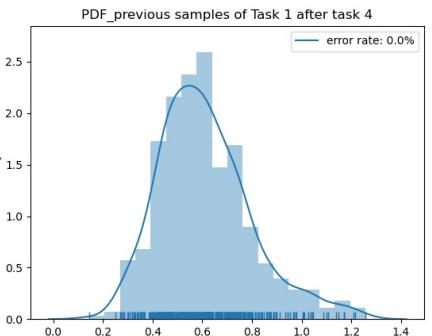
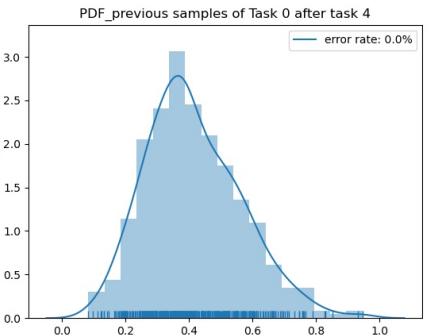
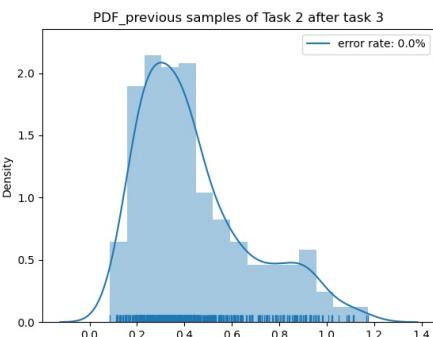
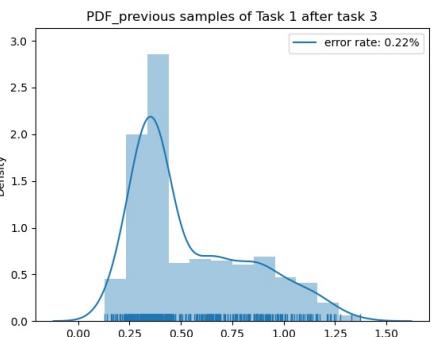
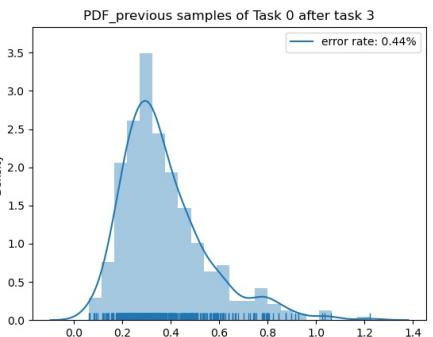
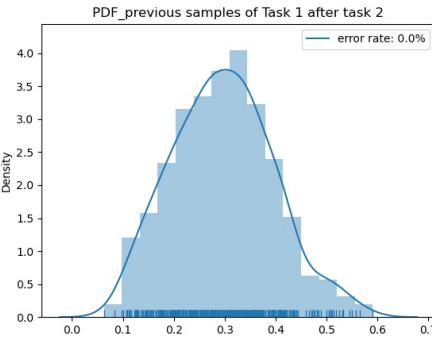
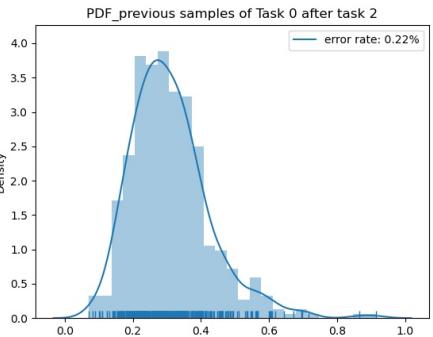
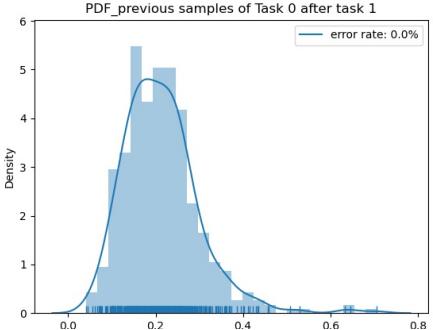
Variance of Valid samples

```
Accuracies =  
    , 0.8531% , 0.0000% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.8438% , 0.9302% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.8401% , 0.9214% , 0.7760% , 0.0000% , 0.0000% ,  
    , 0.8281% , 0.8995% , 0.7661% , 0.9438% , 0.0000% ,  
    , 0.8245% , 0.9120% , 0.7557% , 0.9026% , 0.9401% ,  
ACC: 0.8670%  
  
BWT : -0.02%
```

Choosing 5% Interval variance samples
20 epochs



Training from task 1 to 4



How the variance of choosing samples changes

```
Accuracies =
, 0.8531%, 0.0000%, 0.0000%, 0.0000%, 0.0000%,
, 0.8438%, 0.9302%, 0.0000%, 0.0000%, 0.0000%,
, 0.8401%, 0.9214%, 0.7760%, 0.0000%, 0.0000%,
, 0.8281%, 0.8995%, 0.7661%, 0.9438%, 0.0000%,
, 0.8245%, 0.9120%, 0.7557%, 0.9026%, 0.9401%,
ACC: 0.8670%
```

BWT : -0.02%

Choosing 5% Interval variance samples
20 epochs

Store the Best Model

When do we stop the training ?

- Catastrophic forgetting

Store the Best Model

Criterion:

- loss[current]*
- loss[current] + loss[previous]
- acc[current]
(previous samples: go through only)
- acc[current] + acc[previous]
- acc + loss

```
*****
Task 0 (cifar10-2)
*****
| Epoch 1, time=1331.1ms/ 30.2ms | Train: loss=5.951, acc= 79.3% | Valid: loss=5.811, acc= 80.6% | *
| Epoch 2, time=1323.5ms/ 30.4ms | Train: loss=4.737, acc= 76.6% | Valid: loss=4.693, acc= 74.8% | *
| Epoch 3, time=1325.0ms/ 30.4ms | Train: loss=5.197, acc= 83.4% | Valid: loss=5.201, acc= 81.1% |
| Epoch 4, time=1324.9ms/ 30.3ms | Train: loss=5.307, acc= 86.2% | Valid: loss=5.241, acc= 84.8% |
| Epoch 5, time=1324.6ms/ 30.1ms | Train: loss=5.866, acc= 88.0% | Valid: loss=6.199, acc= 84.5% |
| Epoch 6, time=1325.7ms/ 30.0ms | Train: loss=5.305, acc= 87.8% | Valid: loss=5.663, acc= 85.3% |
| Epoch 7, time=1327.2ms/ 30.4ms | Train: loss=4.894, acc= 91.1% | Valid: loss=5.237, acc= 88.3% |
| Epoch 8, time=1324.9ms/ 30.4ms | Train: loss=5.224, acc= 89.7% | Valid: loss=5.871, acc= 86.3% |
| Epoch 9, time=1323.1ms/ 30.2ms | Train: loss=3.888, acc= 90.3% | Valid: loss=4.392, acc= 84.7% | *
| Epoch 10, time=1326.7ms/ 30.4ms | Train: loss=5.487, acc= 92.0% | Valid: loss=6.392, acc= 87.6% |

>>> Test on task 0 - cifar10-2      : loss=444425.000, acc=85.312% <<
Save at ../checkpoints/cifar_ucb
*****
Task 1 (cifar10-0)
*****
| Epoch 1, time=1322.9ms/ 30.1ms | Train: loss=4.000, acc= 92.8% | Valid: loss=4.622, acc= 90.7% | *
| Epoch 2, time=1321.6ms/ 30.2ms | Train: loss=2.277, acc= 95.6% | Valid: loss=3.231, acc= 92.6% | *
| Epoch 3, time=1319.3ms/ 30.1ms | Train: loss=2.288, acc= 96.6% | Valid: loss=3.455, acc= 94.2% |
| Epoch 4, time=1319.9ms/ 30.1ms | Train: loss=2.915, acc= 98.3% | Valid: loss=4.302, acc= 94.6% |
| Epoch 5, time=1320.6ms/ 30.1ms | Train: loss=3.929, acc= 98.1% | Valid: loss=5.075, acc= 94.4% |
| Epoch 6, time=1320.0ms/ 30.2ms | Train: loss=3.081, acc= 98.4% | Valid: loss=4.546, acc= 94.8% |
| Epoch 7, time=1318.5ms/ 30.1ms | Train: loss=3.610, acc= 98.2% | Valid: loss=5.627, acc= 93.1% |
| Epoch 8, time=1320.0ms/ 30.2ms | Train: loss=4.538, acc= 98.9% | Valid: loss=6.713, acc= 94.2% |
| Epoch 9, time=1319.5ms/ 30.1ms | Train: loss=4.142, acc= 99.6% | Valid: loss=5.440, acc= 95.1% |
| Epoch 10, time=1318.8ms/ 30.2ms | Train: loss=2.756, acc= 99.5% | Valid: loss=5.222, acc= 94.5% |

>>> Test on task 0 - cifar10-2      : loss=1158664.000, acc=75.990% <<
>>> Test on task 1 - cifar10-0      : loss=292657.656, acc=94.740% <<
```

* original UCB

Store the Best Model

Criterion:

```
*****
Task 0 (cifar10-2)
*****
| Epoch 1, time=1321.1ms/ 30.2ms | Train: loss=5.951, acc= 79.3% | Valid: loss=5.811, acc= 80.6% | *
| Epoch 2, time=1323.5ms/ 30.4ms | Train: loss=4.737, acc= 76.6% | Valid: loss=4.693, acc= 74.8% | *
| Epoch 3, time=1323.0ms/ 30.4ms | Train: loss=5.197, acc= 83.4% | Valid: loss=5.201, acc= 81.1% |
| Epoch 4, time=1324.9ms/ 30.3ms | Train: loss=5.307, acc= 86.2% | Valid: loss=5.241, acc= 84.8% |
| Epoch 5, time=1324.6ms/ 30.1ms | Train: loss=5.866, acc= 88.0% | Valid: loss=6.199, acc= 84.5% |
| Epoch 6, time=1325.7ms/ 30.0ms | Train: loss=5.305, acc= 87.8% | Valid: loss=5.663, acc= 85.3% |
| Epoch 7, time=1327.2ms/ 30.4ms | Train: loss=4.894, acc= 91.1% | Valid: loss=5.237, acc= 88.3% |
| Epoch 8, time=1324.9ms/ 30.4ms | Train: loss=5.224, acc= 89.7% | Valid: loss=5.871, acc= 86.3% |
| Epoch 9, time=1323.1ms/ 30.2ms | Train: loss=3.888, acc= 90.3% | Valid: loss=4.392, acc= 84.7% | *
| Epoch 10, time=1326.7ms/ 30.4ms | Train: loss=5.487, acc= 92.0% | Valid: loss=6.392, acc= 87.6% |

>>> Test on task 0 - cifar10-2 : loss=444425.000, acc=85.312% <<
Save at ..../checkpoints/cifar_uct
*****
Task 1 (cifar10-0)
*****
| Epoch 1, time=1322.9ms/ 30.1ms | Train: loss=4.000, acc= 92.8% | Valid: loss=4.622, acc= 90.7% | *
| Epoch 2, time=1321.6ms/ 30.2ms | Train: loss=2.277, acc= 95.6% | Valid: loss=3.231, acc= 92.6% | *
| Epoch 3, time=1319.3ms/ 30.1ms | Train: loss=2.288, acc= 96.6% | Valid: loss=3.455, acc= 94.2% |
| Epoch 4, time=1319.9ms/ 30.1ms | Train: loss=2.915, acc= 98.3% | Valid: loss=4.302, acc= 94.6% |
| Epoch 5, time=1320.6ms/ 30.1ms | Train: loss=3.979, acc= 98.1% | Valid: loss=5.075, acc= 94.4% |
| Epoch 6, time=1320.0ms/ 30.2ms | Train: loss=3.081, acc= 98.4% | Valid: loss=4.546, acc= 94.8% |
| Epoch 7, time=1318.5ms/ 30.1ms | Train: loss=3.610, acc= 98.2% | Valid: loss=5.627, acc= 93.1% |
| Epoch 8, time=1320.0ms/ 30.2ms | Train: loss=4.538, acc= 98.9% | Valid: loss=6.713, acc= 94.2% |
| Epoch 9, time=1319.5ms/ 30.1ms | Train: loss=4.142, acc= 99.6% | Valid: loss=5.440, acc= 95.1% |
| Epoch 10, time=1318.8ms/ 30.2ms | Train: loss=2.756, acc= 99.5% | Valid: loss=5.222, acc= 94.5% |

>>> Test on task 0 - cifar10-2 : loss=1158664.000, acc=75.990% <<
>>> Test on task 1 - cifar10-0 : loss=292657.656, acc=94.740% <<
```

loss[current]

```
*****
Task 0 (cifar10-2)
*****
| Epoch 1, time=1324.4ms/ 30.1ms | Train: loss=5.951, acc= 79.3% | Valid: loss=5.811, acc= 80.6% | @
| Epoch 2, time=1322.9ms/ 30.2ms | Train: loss=4.737, acc= 76.6% | Valid: loss=4.693, acc= 74.8% | @
| Epoch 3, time=1318.6ms/ 30.4ms | Train: loss=5.197, acc= 83.4% | Valid: loss=5.201, acc= 81.1% | @
| Epoch 4, time=1321.5ms/ 30.1ms | Train: loss=5.307, acc= 86.2% | Valid: loss=5.241, acc= 84.8% | @
| Epoch 5, time=1318.7ms/ 30.5ms | Train: loss=5.866, acc= 88.0% | Valid: loss=6.199, acc= 84.5% | @
| Epoch 6, time=1321.5ms/ 30.2ms | Train: loss=5.305, acc= 87.8% | Valid: loss=5.663, acc= 85.3% | @
| Epoch 7, time=1319.6ms/ 30.4ms | Train: loss=4.894, acc= 91.1% | Valid: loss=5.237, acc= 88.3% | @
| Epoch 8, time=1320.9ms/ 30.5ms | Train: loss=5.224, acc= 89.7% | Valid: loss=5.871, acc= 86.3% | @
| Epoch 9, time=1317.9ms/ 30.3ms | Train: loss=3.888, acc= 90.3% | Valid: loss=4.392, acc= 84.7% | @
| Epoch 10, time=1318.7ms/ 30.3ms | Train: loss=5.487, acc= 92.0% | Valid: loss=6.392, acc= 87.6% | @

Wrong Predicted:
tensor(8947.) tensor(8913.) tensor(8784.)
```

>>> Test on task 0 - cifar10-2 : loss=5.309, acc=88.073% <<

>>> Test on task 0 - cifar10-2 : loss=5.309, acc=88.073% <<

>>> Test on task 1 - cifar10-0 : loss=11.904, acc=95.365% <<

acc[current] + acc[previous]

```
*****
Task 0 (cifar10-2)
*****
| Epoch 1, time=1301.1ms/ 30.3ms | Train: loss=5.951, acc= 79.3% | Valid: loss=5.811, acc= 80.6% | @
| Epoch 2, time=1297.9ms/ 30.1ms | Train: loss=4.737, acc= 76.6% | Valid: loss=4.693, acc= 74.8% | @
| Epoch 3, time=1295.6ms/ 30.3ms | Train: loss=5.197, acc= 83.4% | Valid: loss=5.201, acc= 81.1% | @
| Epoch 4, time=1297.3ms/ 30.1ms | Train: loss=5.307, acc= 86.2% | Valid: loss=5.241, acc= 84.8% | @
| Epoch 5, time=1296.6ms/ 30.2ms | Train: loss=5.866, acc= 88.0% | Valid: loss=6.199, acc= 84.5% | @
| Epoch 6, time=1295.9ms/ 30.4ms | Train: loss=5.305, acc= 87.8% | Valid: loss=5.663, acc= 85.3% | @
| Epoch 7, time=1295.5ms/ 30.0ms | Train: loss=4.894, acc= 91.1% | Valid: loss=5.237, acc= 88.3% | @
| Epoch 8, time=1296.5ms/ 30.0ms | Train: loss=5.224, acc= 89.7% | Valid: loss=5.871, acc= 86.3% | @
| Epoch 9, time=1295.0ms/ 30.0ms | Train: loss=3.888, acc= 90.3% | Valid: loss=4.392, acc= 84.7% | @
| Epoch 10, time=1294.6ms/ 30.4ms | Train: loss=5.487, acc= 92.0% | Valid: loss=6.392, acc= 87.6% | @
| Epoch 11, time=1294.1ms/ 30.3ms | Train: loss=4.755, acc= 94.8% | Valid: loss=5.674, acc= 87.7% | @
| Epoch 12, time=1294.9ms/ 30.0ms | Train: loss=4.282, acc= 95.3% | Valid: loss=4.890, acc= 88.4% | @
| Epoch 13, time=1294.9ms/ 30.5ms | Train: loss=4.301, acc= 97.3% | Valid: loss=5.92, acc= 90.0% | @
| Epoch 14, time=1294.8ms/ 30.2ms | Train: loss=3.854, acc= 97.0% | Valid: loss=5.001, acc= 89.1% | @
| Epoch 15, time=1295.5ms/ 30.4ms | Train: loss=4.329, acc= 97.7% | Valid: loss=5.961, acc= 90.3% | @
| Epoch 16, time=1296.1ms/ 30.1ms | Train: loss=4.813, acc= 96.0% | Valid: loss=7.259, acc= 88.4% | @
| Epoch 17, time=1295.3ms/ 30.3ms | Train: loss=3.941, acc= 96.8% | Valid: loss=5.768, acc= 88.6% | @
| Epoch 18, time=1295.0ms/ 30.2ms | Train: loss=4.118, acc= 98.8% | Valid: loss=6.401, acc= 88.8% | @
| Epoch 19, time=1295.7ms/ 30.1ms | Train: loss=3.486, acc= 99.5% | Valid: loss=5.215, acc= 89.3% | @
| Epoch 20, time=1295.5ms/ 30.1ms | Train: loss=6.284, acc= 88.802% <<

>>> Test on task 0 - cifar10-2 : loss=6.284, acc=88.802% <<
Save at ..../checkpoints/cifar_uct

*****
Task 1 (cifar10-0)
*****
| Epoch 1, time=1291.5ms/ 30.2ms | Train: loss=5.174, acc= 94.2% | Valid: loss=5.488, acc= 92.9% | @
| Epoch 2, time=1292.3ms/ 30.3ms | Train: loss=4.410, acc= 96.7% | Valid: loss=6.211, acc= 94.9% | @
| Epoch 3, time=1291.8ms/ 30.3ms | Train: loss=2.724, acc= 96.7% | Valid: loss=4.355, acc= 93.6% | @
| Epoch 4, time=1290.0ms/ 30.2ms | Train: loss=2.013, acc= 98.7% | Valid: loss=3.807, acc= 94.9% | @
| Epoch 5, time=1291.9ms/ 30.3ms | Train: loss=2.777, acc= 98.3% | Valid: loss=4.814, acc= 94.4% | @
| Epoch 6, time=1291.9ms/ 30.2ms | Train: loss=1.678, acc= 98.8% | Valid: loss=3.978, acc= 94.0% | @
| Epoch 7, time=1293.1ms/ 30.2ms | Train: loss=2.903, acc= 99.6% | Valid: loss=5.583, acc= 95.2% | @
| Epoch 8, time=1290.7ms/ 30.2ms | Train: loss=2.897, acc= 99.8% | Valid: loss=5.429, acc= 95.8% | @
| Epoch 9, time=1289.9ms/ 30.2ms | Train: loss=3.410, acc= 99.9% | Valid: loss=6.780, acc= 95.5% | @
| Epoch 10, time=1291.1ms/ 30.5ms | Train: loss=4.215, acc= 99.8% | Valid: loss=8.066, acc= 95.1% | @
| Epoch 11, time=1292.1ms/ 30.3ms | Train: loss=6.303, acc= 99.8% | Valid: loss=10.358, acc= 94.3% | @
| Epoch 12, time=1290.5ms/ 30.2ms | Train: loss=4.547, acc= 99.8% | Valid: loss=7.015, acc= 95.4% | @
| Epoch 13, time=1292.2ms/ 30.2ms | Train: loss=6.448, acc= 99.8% | Valid: loss=9.014, acc= 94.1% | @
| Epoch 14, time=1291.5ms/ 30.2ms | Train: loss=6.231, acc= 99.8% | Valid: loss=8.237, acc= 95.3% | @
| Epoch 15, time=1291.1ms/ 30.5ms | Train: loss=4.492, acc= 100.0% | Valid: loss=8.090, acc= 95.6% | @
| Epoch 16, time=1290.5ms/ 30.3ms | Train: loss=5.118, acc= 100.0% | Valid: loss=8.492, acc= 95.6% | @
| Epoch 17, time=1290.0ms/ 30.2ms | Train: loss=7.478, acc= 99.9% | Valid: loss=10.366, acc= 95.8% | @
| Epoch 18, time=1290.5ms/ 30.3ms | Train: loss=4.064, acc= 100.0% | Valid: loss=7.362, acc= 96.2% | @
| Epoch 19, time=1289.1ms/ 30.3ms | Train: loss=3.156, acc= 100.0% | Valid: loss=6.466, acc= 95.8% | @
| Epoch 20, time=1290.7ms/ 30.2ms | Train: loss=4.553, acc= 100.0% | Valid: loss=8.224, acc= 95.4% | @

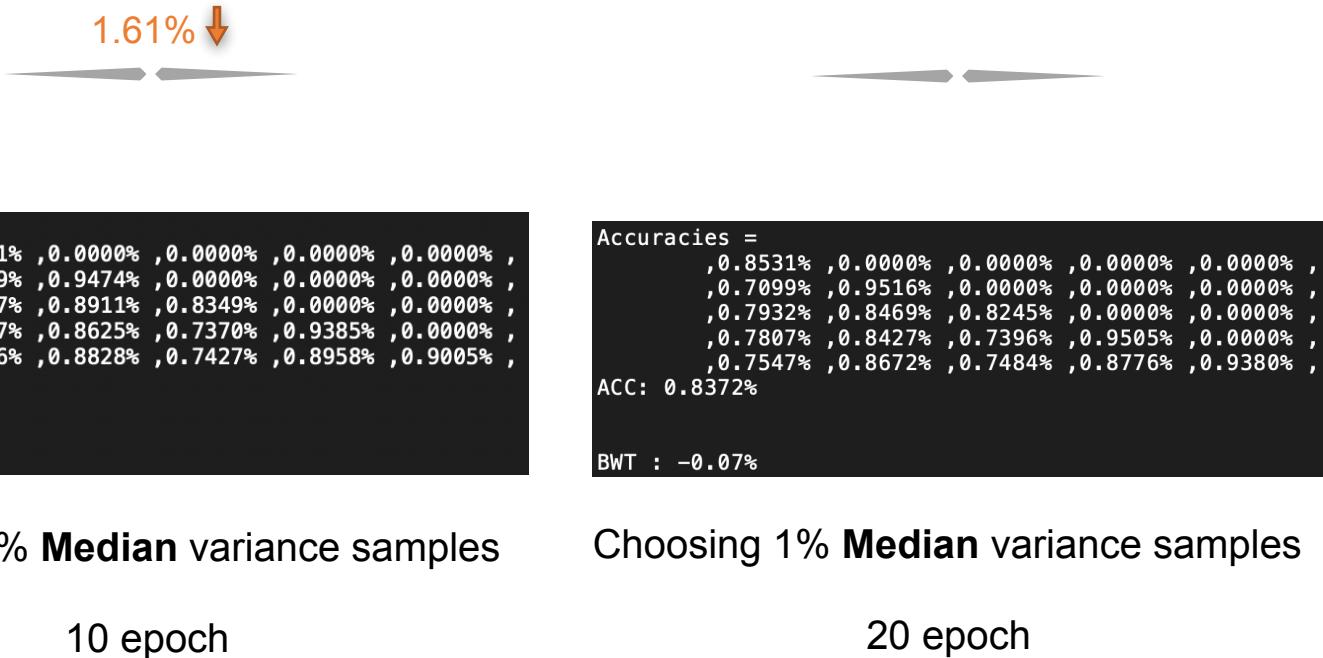
>>> Test on task 0 - cifar10-2 : loss=27.979, acc=76.042% <<
>>> Test on task 1 - cifar10-0 : loss=7.046, acc=95.729% <<
Save at ..../checkpoints/cifar_uct
```

acc[current]

Store the Best Model - all acc

```
Accuracies =  
    ,0.8807% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7584% ,0.9594% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7651% ,0.8766% ,0.8443% ,0.0000% ,0.0000% ,  
    ,0.7546% ,0.8615% ,0.7505% ,0.9585% ,0.0000% ,  
    ,0.7490% ,0.8490% ,0.7208% ,0.8568% ,0.9432% ,  
ACC: 0.8238%  
  
BWT : -0.09%
```

1.61% ↓



```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7599% ,0.9474% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8177% ,0.8911% ,0.8349% ,0.0000% ,0.0000% ,  
    ,0.7927% ,0.8625% ,0.7370% ,0.9385% ,0.0000% ,  
    ,0.7776% ,0.8828% ,0.7427% ,0.8958% ,0.9005% ,  
ACC: 0.8399%  
  
BWT : -0.06%
```

Choosing 1% **Median** variance samples

10 epoch

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7099% ,0.9516% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7932% ,0.8469% ,0.8245% ,0.0000% ,0.0000% ,  
    ,0.7807% ,0.8427% ,0.7396% ,0.9505% ,0.0000% ,  
    ,0.7547% ,0.8672% ,0.7484% ,0.8776% ,0.9380% ,  
ACC: 0.8372%  
  
BWT : -0.07%
```

Choosing 1% **Median** variance samples

20 epoch

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8438% ,0.9302% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8401% ,0.9214% ,0.7760% ,0.0000% ,0.0000% ,  
    ,0.8281% ,0.8995% ,0.7661% ,0.9438% ,0.0000% ,  
    ,0.8245% ,0.9120% ,0.7557% ,0.9026% ,0.9401% ,  
ACC: 0.8670%  
  
AC( BWT : -0.02%  
  
BWT : -0.02%
```

Choosing top 5% **Median** variance samples

Store the Best Model - currant acc

```
Accuracies =  
    ,0.8807% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7563% ,0.9474% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7245% ,0.8760% ,0.8568% ,0.0000% ,0.0000% ,  
    ,0.7552% ,0.8177% ,0.7755% ,0.9578% ,0.0000% ,  
    ,0.7365% ,0.8557% ,0.7620% ,0.9016% ,0.9385% ,  
ACC: 0.8389%  
  
BWT : -0.08%
```

```
Accuracies =  
    ,0.8880% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7604% ,0.9573% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7849% ,0.8687% ,0.8458% ,0.0000% ,0.0000% ,  
    ,0.7167% ,0.8604% ,0.7719% ,0.9646% ,0.0000% ,  
    ,0.7312% ,0.8719% ,0.7557% ,0.8359% ,0.9432% ,  
ACC: 0.8276%  
  
BWT : -0.09%
```

0.10% ↓

0.96% ↓

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7599% ,0.9474% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8177% ,0.8911% ,0.8349% ,0.0000% ,0.0000% ,  
    ,0.7927% ,0.8625% ,0.7370% ,0.9385% ,0.0000% ,  
    ,0.7776% ,0.8828% ,0.7427% ,0.8958% ,0.9005% ,  
ACC: 0.8399%  
  
BWT : -0.06%
```

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7099% ,0.9516% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.7932% ,0.8469% ,0.8245% ,0.0000% ,0.0000% ,  
    ,0.7807% ,0.8427% ,0.7396% ,0.9505% ,0.0000% ,  
    ,0.7547% ,0.8672% ,0.7484% ,0.8776% ,0.9380% ,  
ACC: 0.8372%  
  
BWT : -0.07%
```

Choosing 1% **Median** variance samples

10 epoch

```
Accuracies =  
    ,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8438% ,0.9302% ,0.0000% ,0.0000% ,0.0000% ,  
    ,0.8401% ,0.9214% ,0.7760% ,0.0000% ,0.0000% ,  
    ,0.8281% ,0.8995% ,0.7661% ,0.9438% ,0.0000% ,  
    ,0.8245% ,0.9120% ,0.7557% ,0.9026% ,0.9401% ,  
ACC: 0.8670%  
  
AC( BWT : -0.02%  
  
BWT : -0.02%
```

Choosing 1% **Median** variance samples

20 epoch

Choosing top 5% **Median** variance samples

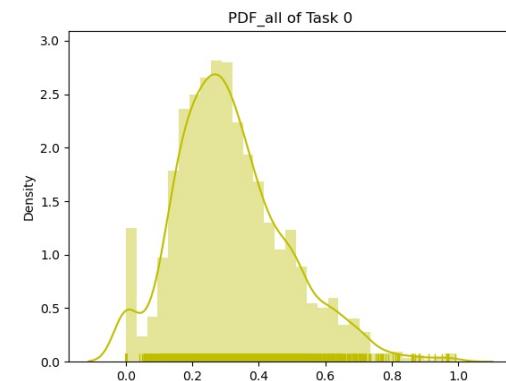
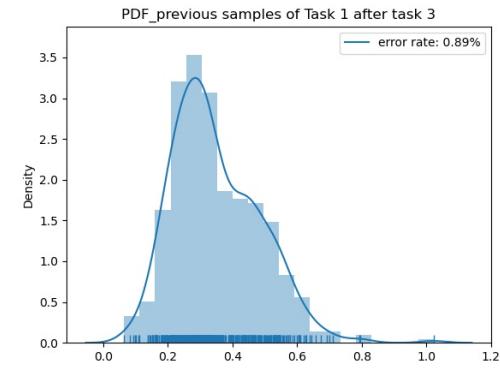
Thoughts

Random - not stable

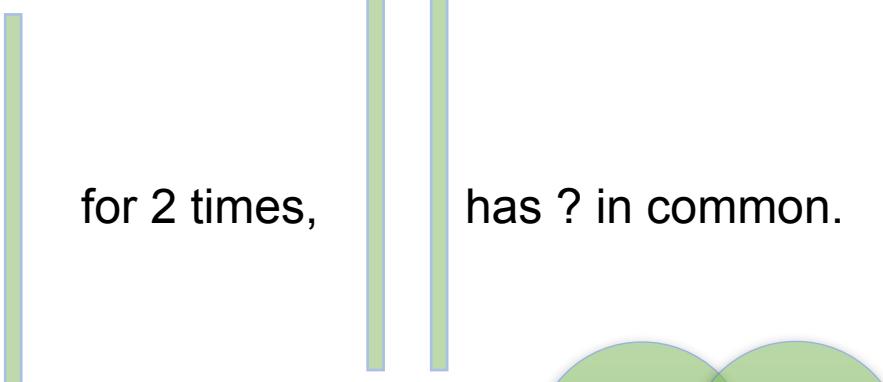
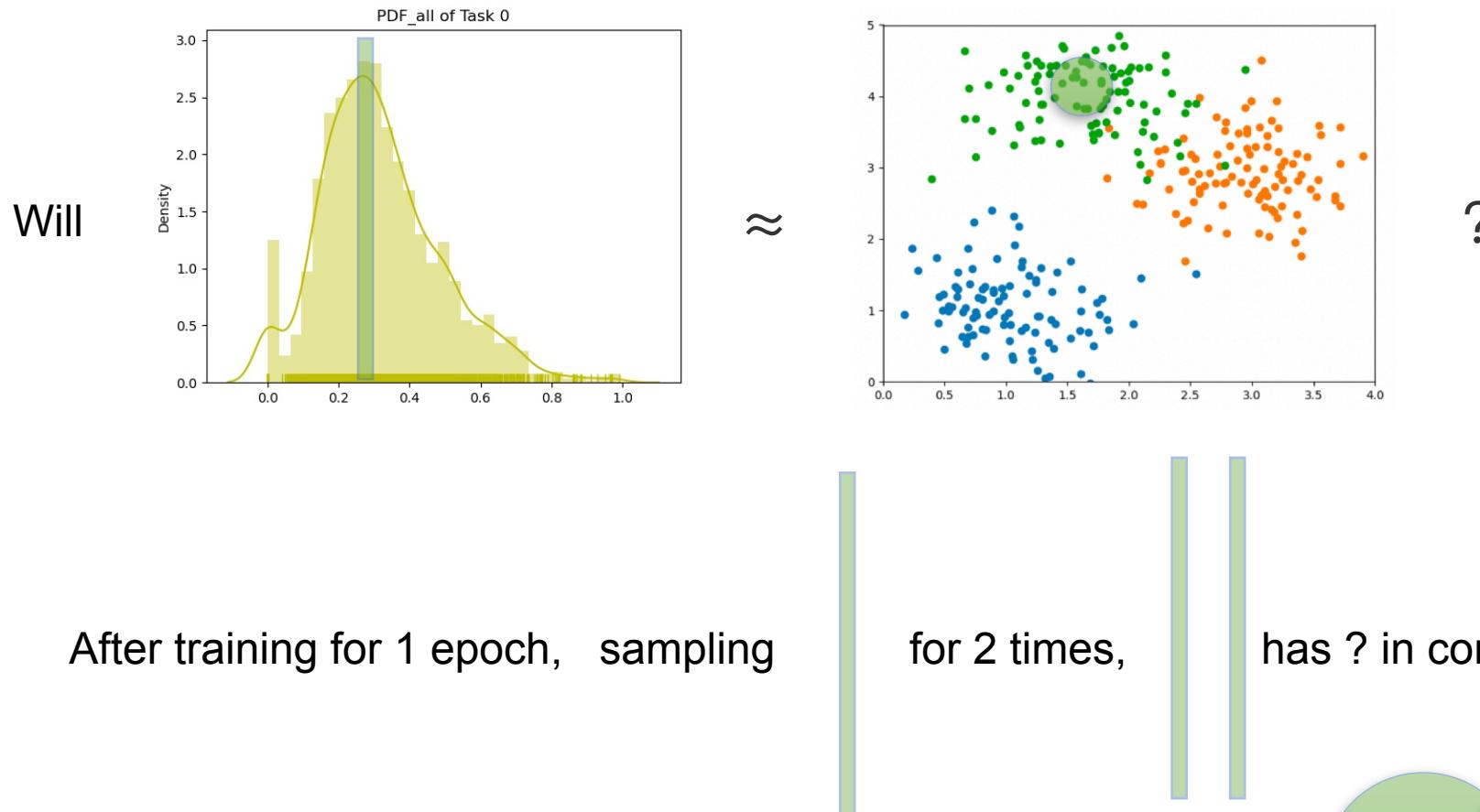
No 0 variance - for the replay samples

Optimizer:

Parameter Freezing - UCB



How do model thinks about the same samples?



Variances - Robust Enough?

After training for 1 epoch,



for 2 times,



has 2.3% in common.

After training for 1 epoch,



for 2 times,



has 70.0% in common.

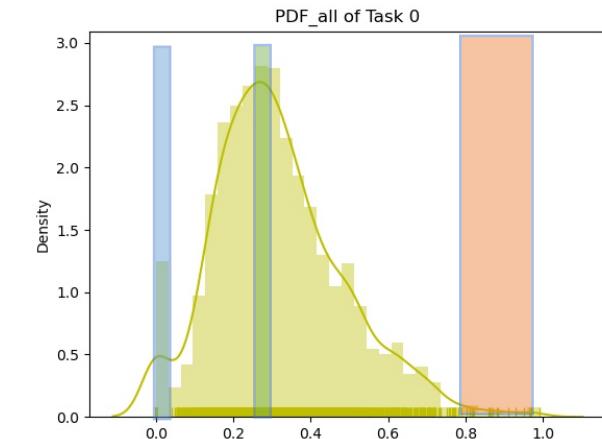
After training for 1 epoch,



for 2 times,



has 37.8% in common.

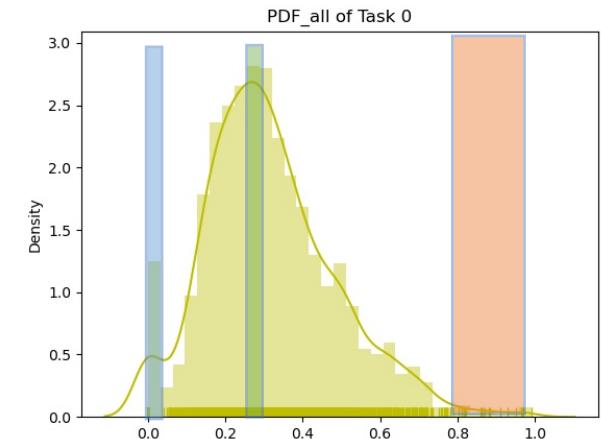


Variances - Robust Enough?

After training for 10 epoch,
for 2 times, has 4.2% in common.

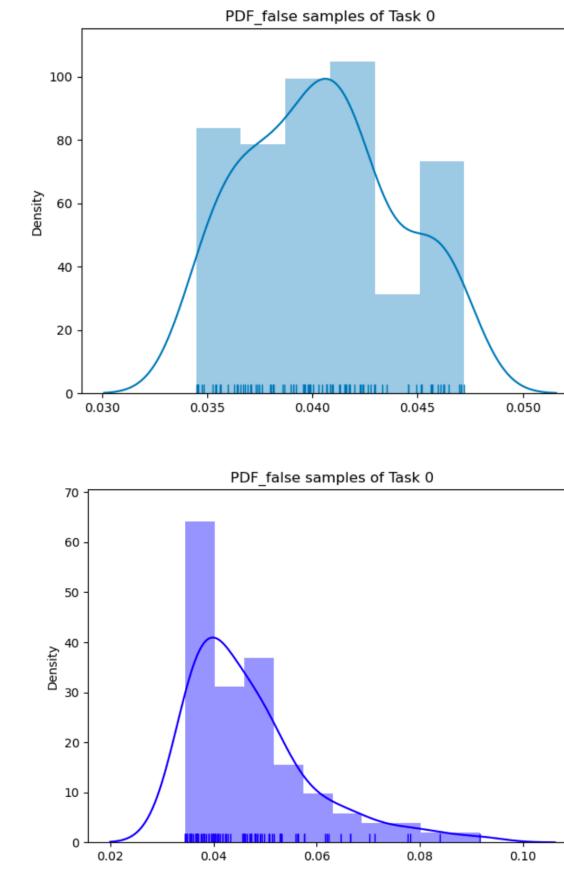
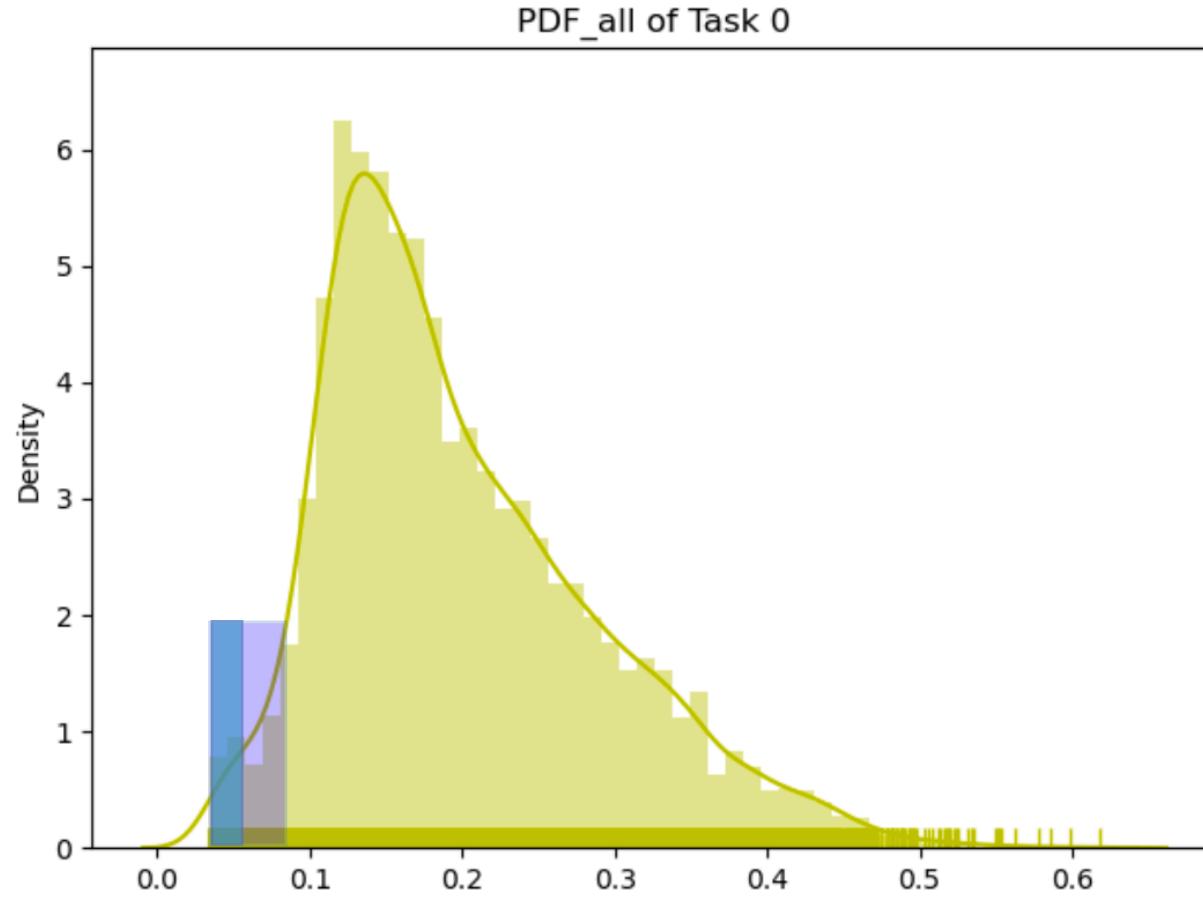
After training for 10 epoch,
for 2 times, has 62.0% in common.

After training for 10 epoch,
for 2 times, has 40.0% in common.

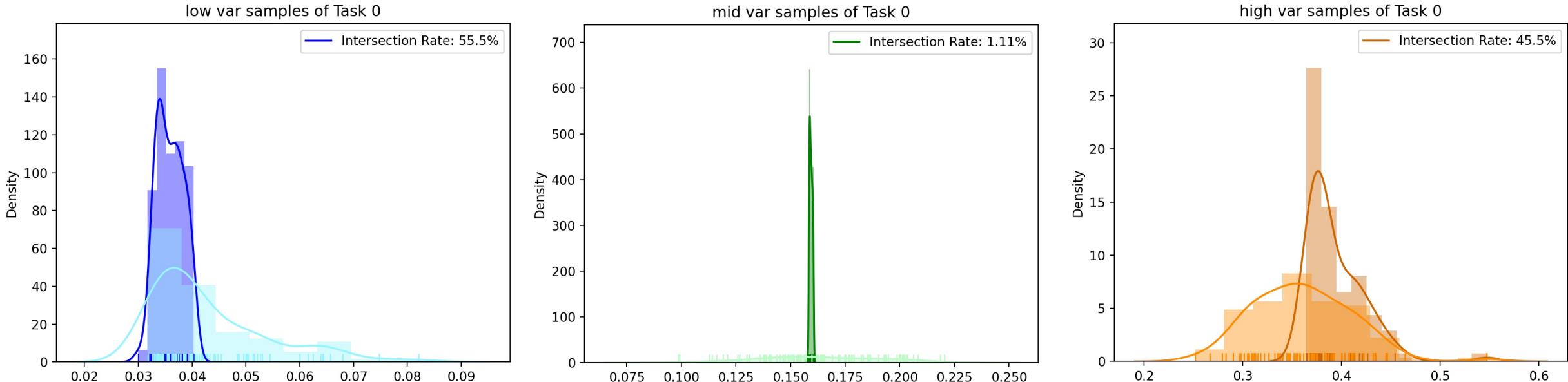


= 1%

Variances - Robust Enough?

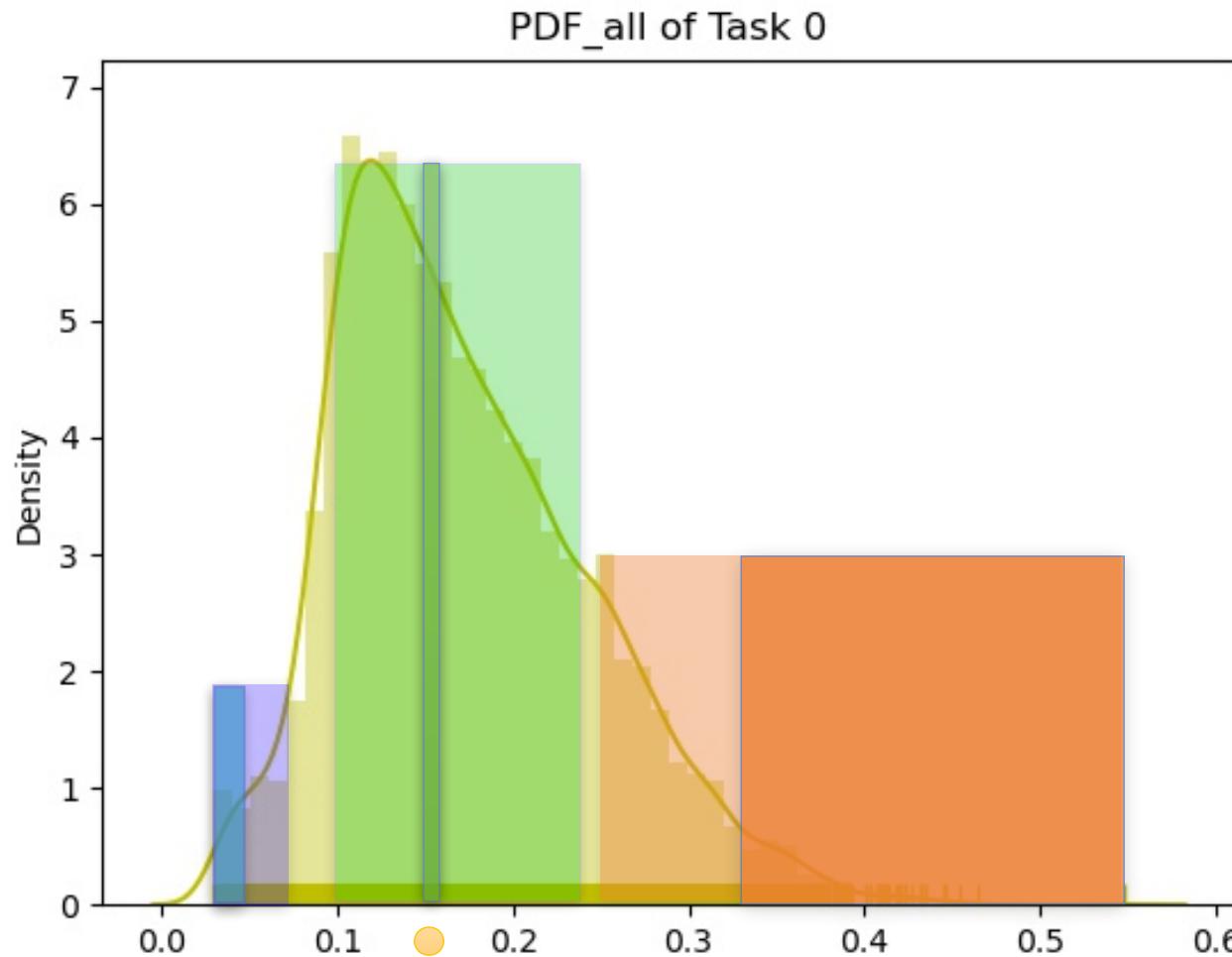


Variances - Robust Enough?



Middle variance samples are not specified, even in same epoch.

Variances - Robust Enough?



How the boundary changes when sampling twice.

Conclusions

Zero variance samples = correctly predicted



High variance only = solve the unsolved problem - confusing

Average variance only = solve the unsolved problem + majority

Low variance only = remind the model of previous tasks

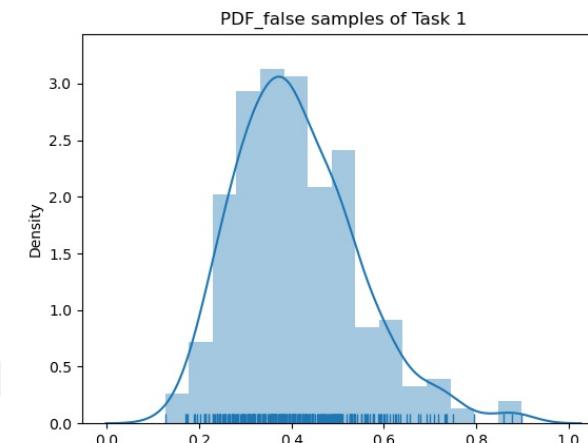
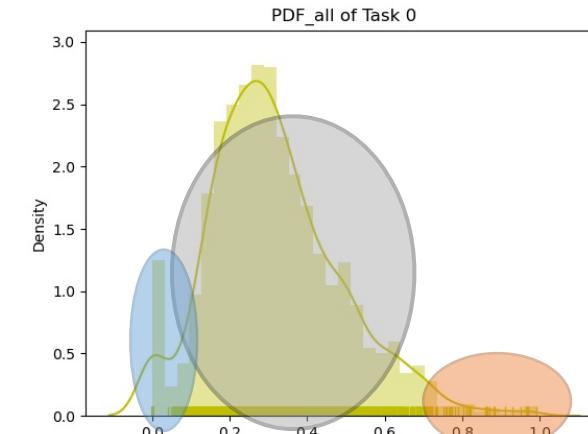
Low variance + Average variance = good



High variance samples: hard but not representative

Average variance samples: majority but we cannot have them all

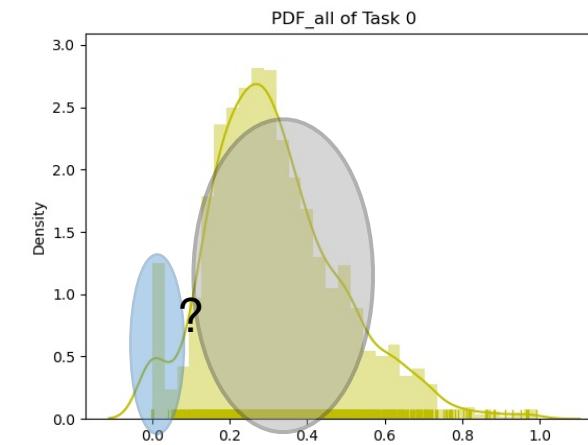
Low variance samples: easy & representative



Future

Extremely small amount of data to replay:

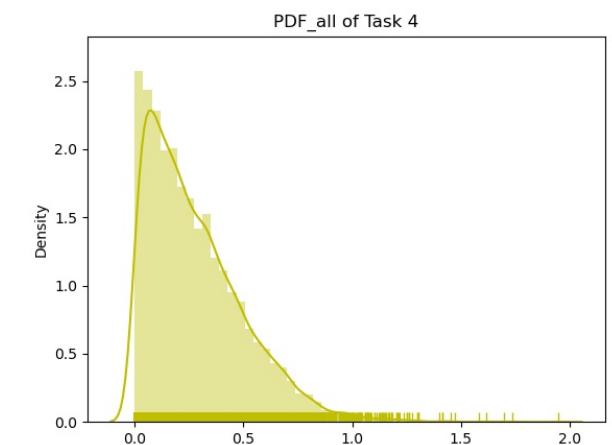
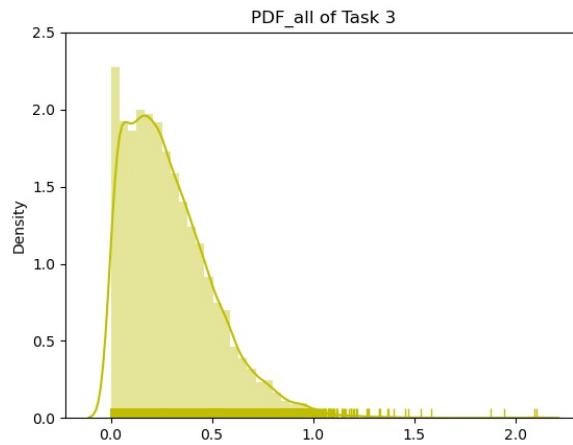
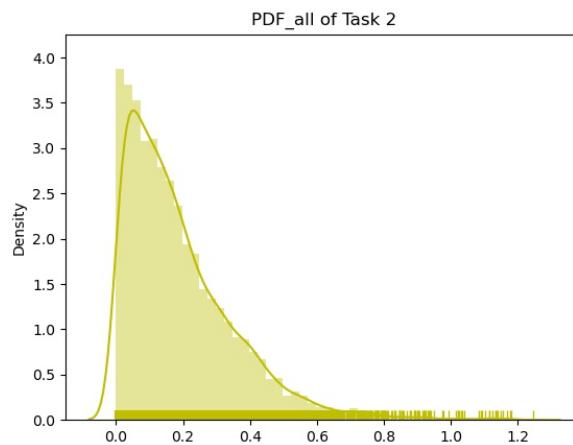
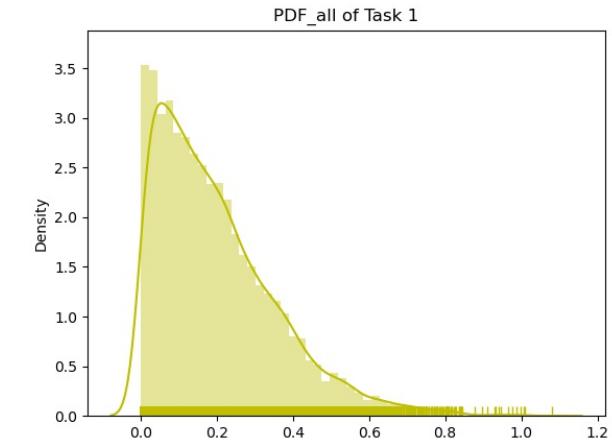
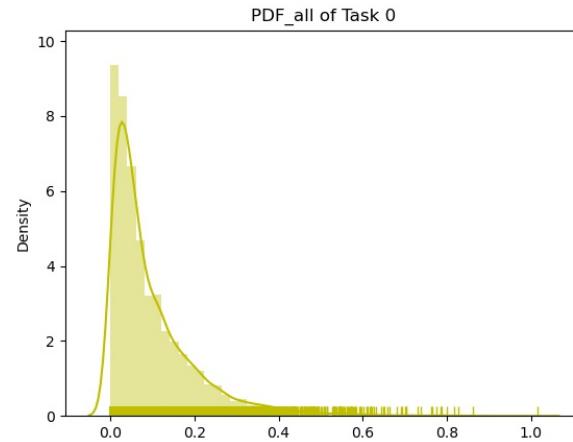
Low variance or Average variance ?



For single epoch

```
Accuracies =  
    , 0.8021% , 0.0000% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.6906% , 0.8786% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.7469% , 0.8385% , 0.7651% , 0.0000% , 0.0000% ,  
    , 0.7375% , 0.8104% , 0.7000% , 0.8896% , 0.0000% ,  
    , 0.7250% , 0.7948% , 0.6755% , 0.8740% , 0.8786% ,  
ACC: 0.7896%  
  
BWT : -0.05%
```

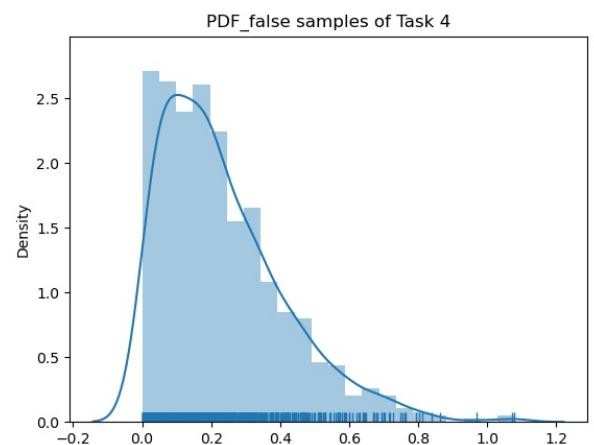
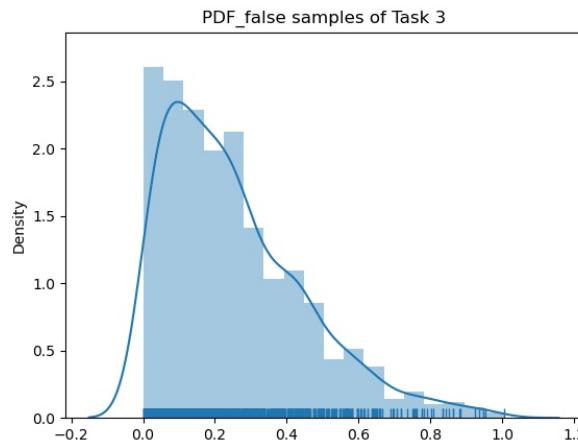
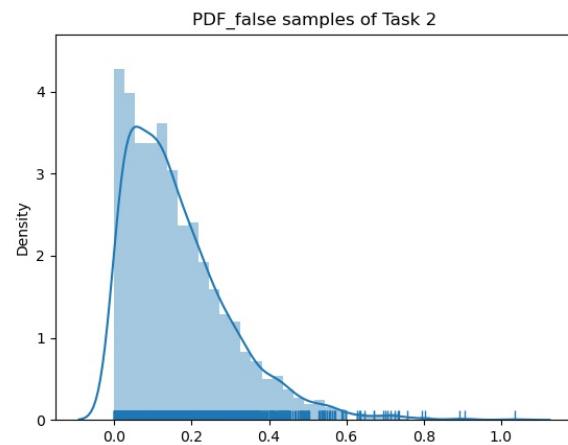
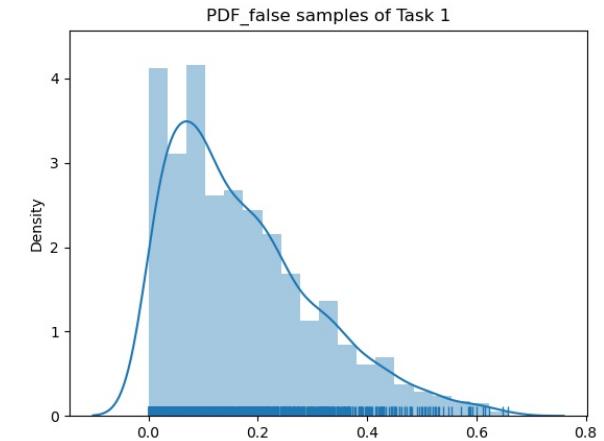
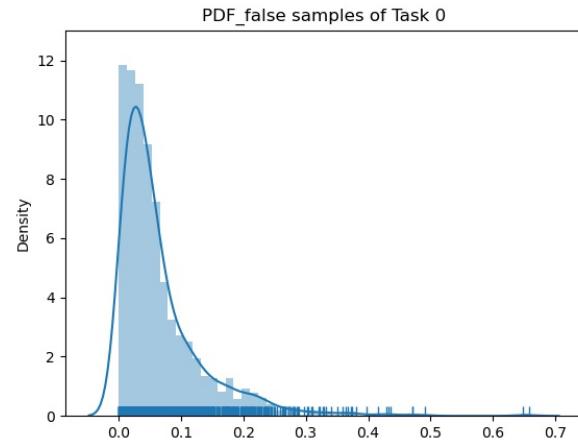
1% low 1epoch 2samples



For single epoch

```
Accuracies =  
    , 0.8021% , 0.0000% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.6906% , 0.8786% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.7469% , 0.8385% , 0.7651% , 0.0000% , 0.0000% ,  
    , 0.7375% , 0.8104% , 0.7000% , 0.8896% , 0.0000% ,  
    , 0.7250% , 0.7948% , 0.6755% , 0.8740% , 0.8786% ,  
ACC: 0.7896%  
  
BWT : -0.05%
```

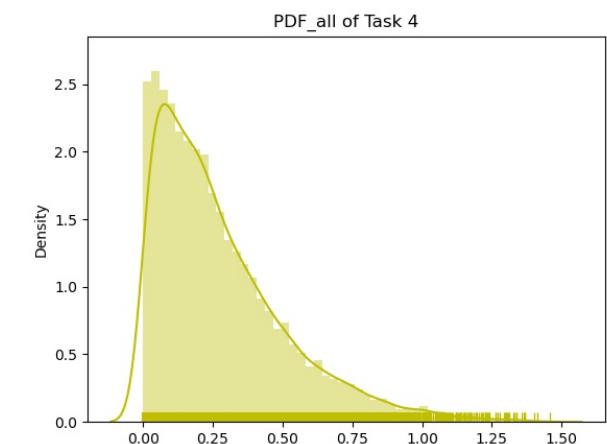
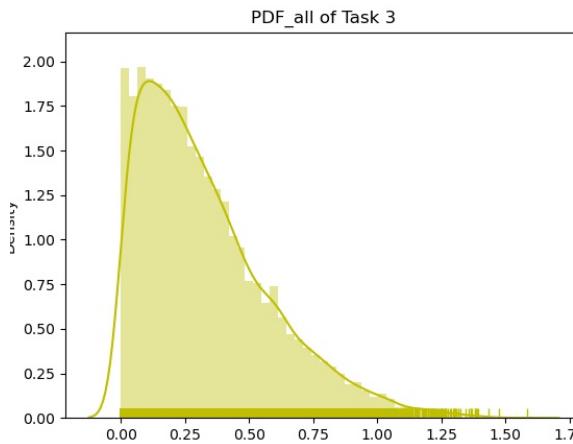
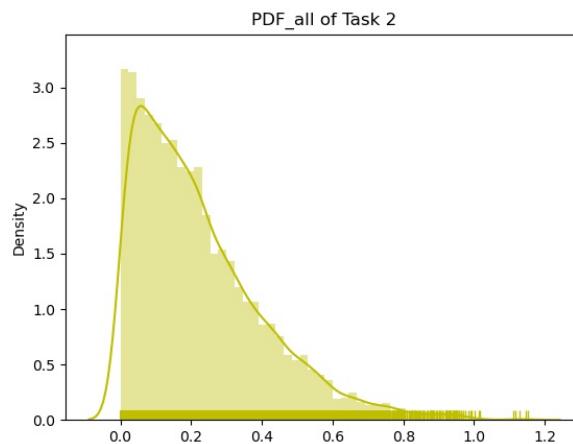
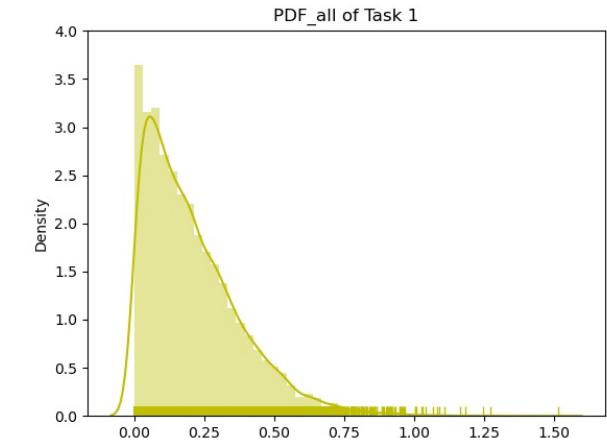
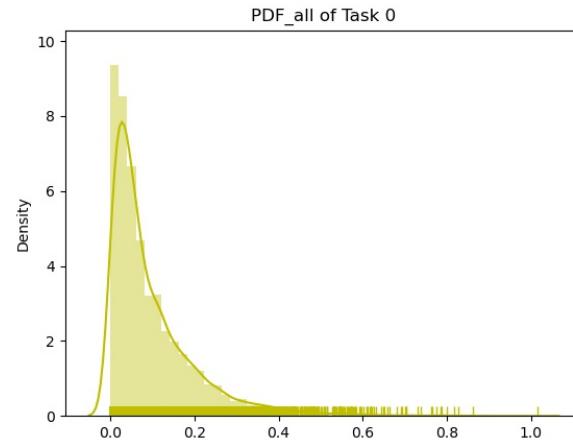
1% low 1epoch 2samples



For single epoch

```
Accuracies =  
    , 0.8021% , 0.0000% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.7234% , 0.7719% , 0.0000% , 0.0000% , 0.0000% ,  
    , 0.6922% , 0.8740% , 0.7495% , 0.0000% , 0.0000% ,  
    , 0.7188% , 0.8255% , 0.7333% , 0.8708% , 0.0000% ,  
    , 0.6943% , 0.8344% , 0.7094% , 0.8536% , 0.8531% ,  
ACC: 0.7890%  
  
BWT : -0.02%
```

1% random 1epoch 2samples



Thank you very much!!



Q&A

Which samples to be replayed

```
Accuracies =  
,0.8714% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
,0.7823% ,0.9531% ,0.0000% ,0.0000% ,0.0000% ,  
,0.8031% ,0.8531% ,0.8401% ,0.0000% ,0.0000% ,  
,0.7698% ,0.8542% ,0.7344% ,0.9630% ,0.0000% ,  
,0.7177% ,0.8573% ,0.6797% ,0.8958% ,0.9401% ,  
ACC: 0.8181%  
  
BWT : -0.10%
```

Choosing 1% **Sub-random** variance samples



```
Accuracies =  
,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
,0.7719% ,0.9432% ,0.0000% ,0.0000% ,0.0000% ,  
,0.7776% ,0.9068% ,0.8161% ,0.0000% ,0.0000% ,  
,0.7792% ,0.8813% ,0.7484% ,0.9302% ,0.0000% ,  
,0.7547% ,0.8724% ,0.7078% ,0.8609% ,0.8901% ,  
ACC: 0.8172%  
  
BWT : -0.07%
```

Choosing 1% **Random** variance samples

10 epochs

```
Accuracies =  
,0.8911% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
,0.8161% ,0.9469% ,0.0000% ,0.0000% ,0.0000% ,  
,0.7870% ,0.8667% ,0.8297% ,0.0000% ,0.0000% ,  
,0.7786% ,0.8693% ,0.7609% ,0.9599% ,0.0000% ,  
,0.7667% ,0.8766% ,0.7260% ,0.8698% ,0.9359% ,  
ACC: 0.8350%  
  
BWT : -0.08%
```

Choosing 5% **Sub-random** variance samples

```
Accuracies =  
,0.8531% ,0.0000% ,0.0000% ,0.0000% ,0.0000% ,  
,0.8302% ,0.9484% ,0.0000% ,0.0000% ,0.0000% ,  
,0.8578% ,0.9172% ,0.7818% ,0.0000% ,0.0000% ,  
,0.8307% ,0.8948% ,0.7557% ,0.9240% ,0.0000% ,  
,0.8406% ,0.8964% ,0.7354% ,0.8849% ,0.8906% ,  
ACC: 0.8496%  
  
BWT : -0.03%
```

Choosing 5% **Random** variance samples

20 epochs