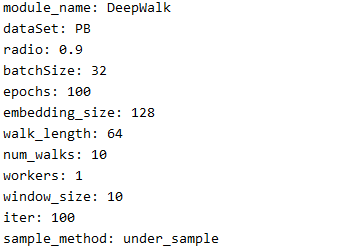
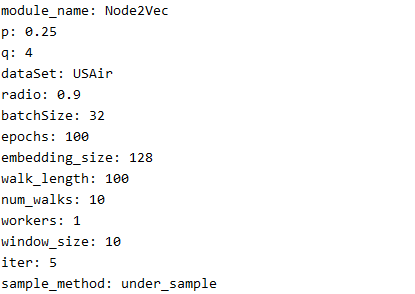
## 启发类算法相关实验

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | PB | USAir | Bio-CE-GT | hamster | Yeast |
| AA | 0.500 | 0.52 | 0.518 | 0.503 | 0.513 |
| ACT | 0.801 | 0.860 | 0.770 | 0.762 | 0.809 |
| CN | 0.582 | 0.579 | 0.453 | 0.421 | 0.529 |
| Jaccard | 0.664 | 0.662 | 0.624 | 0.592 | 0.649 |
| Katz | 0.700 | 0.710 | 0.722 | 0.714 | 0.729 |
| RW | 0.700 | 0.710 | 0.727 | 0.719 | 0.734 |
| RWR | 0.710 | 0.712 | 0.730 | 0.725 | 0.737 |

## 随机游走相关实验



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| PB | 1466 | 206 | 1466 | 206 | 0.877 | 0.877 | 0.877 |
| USAir | 157 | 33 | 180 | 56 | 0.791 | 0.826 | 0.791 |
| Bio-CE-GT | 271 | 32 | 292 | 53 | 0.869 | 0.894 | 0.869 |
| hamster | 866 | 335 | 919 | 388 | 0.712 | 0.721 | 0.712 |
| Yeast | 1152 | 33 | 1137 | 18 | 0.978 | 0.972 | 0.978 |



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| PB | 1522 | 248 | 1424 | 150 | 0.881 | 0.877 | 0.881 |
| USAir | 170 | 38 | 175 | 43 | 0.810 | 0.817 | 0.810 |
| Bio-CE-GT | 260 | 18 | 306 | 64 | 0.873 | 0.935 | 0.873 |
| hamster | 1196 | 94 | 1160 | 58 | 0.939 | 0.927 | 0.939 |
| Yeast | 1153 | 20 | 1150 | 17 | 0.984 | 0.982 | 0.984 |

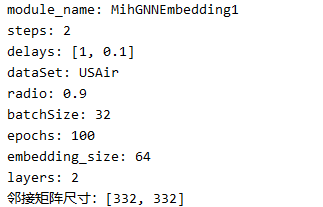
## 嵌入模型1（MihGNNEmbedding1）

模型数学表达为：



并且在传播过程中只针对正样本进行邻居节点的汇聚工程，对于负样本直接取嵌入状态；

由于对正负样本进行了分别处理，那么在预测过程中对输入的预测样本使用正样本的处理过程还是负样本的处理过程呢？



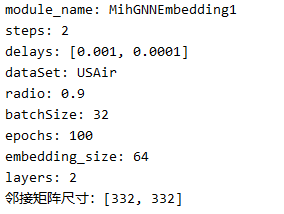
使用正样本的处理过程：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） |
| USAir | 212 | 210 | 3 | 1 | 0.504 | 0.502 |

使用负样本的处理过程：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） |
| USAir | 0 | 0 | 213 | 213 | 0.500 | 0.000 |

如过重新设定每一介邻居的影响因子例如：



此模型会产生退化：



使用正样本的处理过程：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） |
| USAir | 178 | 52 | 161 | 35 | 0.796 | 0.774 |

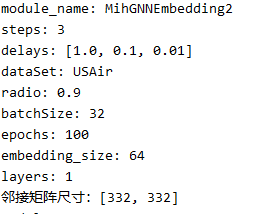
使用负样本的处理过程：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） |
| USAir | 149 | 43 | 170 | 64 | 0.749 | 0.776 |

## 嵌入模型2（MihGNNEmbedding2，使用子图）

模型数学表达：





|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 3 | 0 | 213 | 210 | 0.507 | 1.0 |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

使用交叉熵crossEntropy作为损失函数的实验结果：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 182 | 22 | 191 | 31 | 0.875 | 0.892 |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

## 嵌入模型3（MihGNNEmbedding4）

引入注意力机制（单头注意力），数学表达为：

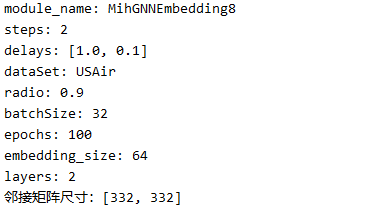


实验结果：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 189 | 17 | 196 | 24 | 0.904 | 0.917 | 0.904 |
| Bio-CE-GT | 278 | 44 | 280 | 46 | 0.861 | 0.863 | 0.861 |
| hamster | 1019 | 140 | 1114 | 235 | 0.850 | 0.879 | 0.850 |
| Yeast | 943 | 98 | 1072 | 227 | 0.861 | 0.906 | 0.861 |
| PB | 1449 | 248 | 1424 | 223 | 0.859 | 0.854 | 0.859 |

## 嵌入模型4（MihGNNEmbedding8&MihGNNEmbedding10）

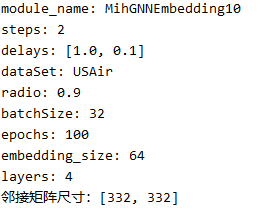




实验结果：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 209 | 69 | 144 | 4 | 0.829 | 0.752 | 0.925 |
| Bio-CE-GT | 310 | 93 | 231 | 14 | 0.835 | 0.769 | 0.936 |
| hamster | 1157 | 373 | 881 | 97 | 0.813 | 0.756 | 0.916 |
| Yeast | 1082 | 276 | 894 | 88 | 0.844 | 0.797 | 0.948 |
| PB | 1531 | 470 | 1202 | 141 | 0.817 | 0.765 | 0.909 |

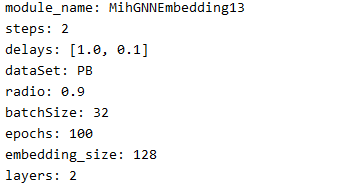
使用交叉熵crossEntropy作为损失函数：



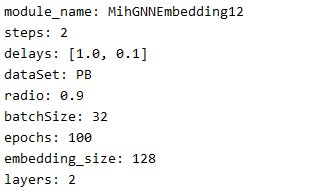
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 184 | 14 | 199 | 29 | 0.899 | 0.929 | 0.899 |
| Bio-CE-GT | 292 | 26 | 298 | 32 | 0.910 | 0.918 | 0.910 |
| hamster | 1128 | 101 | 1153 | 126 | 0.909 | 0.918 | 0.909 |
| Yeast | 1030 | 63 | 1107 | 140 | 0.913 | 0.942 | 0.913 |
| PB | 1467 | 223 | 1449 | 205 | 0.872 | 0.868 | 0.872 |

## 嵌入模型5（MihGNNEmbedding12&MihGNNEmbedding13）



实验结果：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 190 | 20 | 193 | 23 | 0.899 | 0.905 | 0.951 |
| Bio-CE-GT | 272 | 14 | 310 | 52 | 0.898 | 0.951 | 0.974 |
| hamster | 1052 | 107 | 1147 | 202 | 0.877 | 0.908 | 0.950 |
| Yeast | 953 | 12 | 1158 | 217 | 0.902 | 0.988 | 0.982 |
| Router | 564 | 67 | 559 | 62 | 0.897 | 0.893 | 0.897 |
| PB | 1424 | 240 | 1432 | 248 | 0.854 | 0.855 | 0.933 |

使用交叉熵作为损失函数：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 187 | 11 | 202 | 26 | 0.913 | 0.944 | 0.913 |
| Bio-CE-GT | 289 | 26 | 298 | 35 | 0.906 | 0.917 | 0.906 |
| hamster | 1106 | 92 | 1162 | 148 | 0.904 | 0.923 | 0.904 |
| Yeast | 1062 | 53 | 1117 | 108 | 0.931 | 0.952 | 0.931 |
| PB | 1499 | 233 | 1439 | 173 | 0.879 | 0.865 | 0.879 |

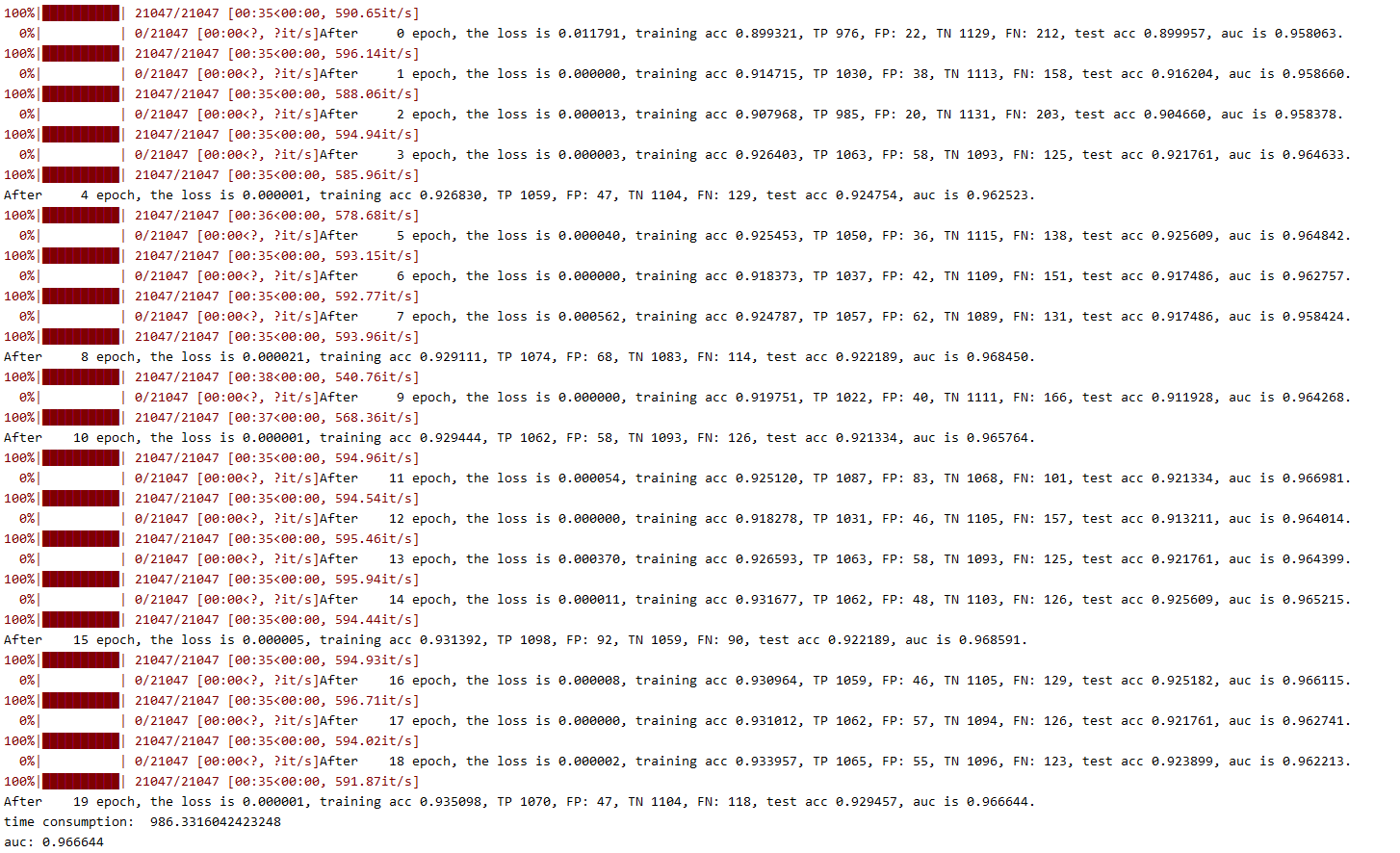
## 嵌入模型6（MihGNNEmbedding9）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 195 | 23 | 190 | 18 | 0.904 | 0.894 | 0.904 |
| Bio-CE-GT | 282 | 36 | 288 | 42 | 0.880 | 0.887 | 0.880 |
| hamster | 1092 | 180 | 1074 | 162 | 0.864 | 0.858 | 0.864 |
| Yeast | 1041 | 90 | 1080 | 129 | 0.906 | 0.920 | 0.906 |
| PB | 1491 | 238 | 1434 | 181 | 0.875 | 0.862 | 0.875 |

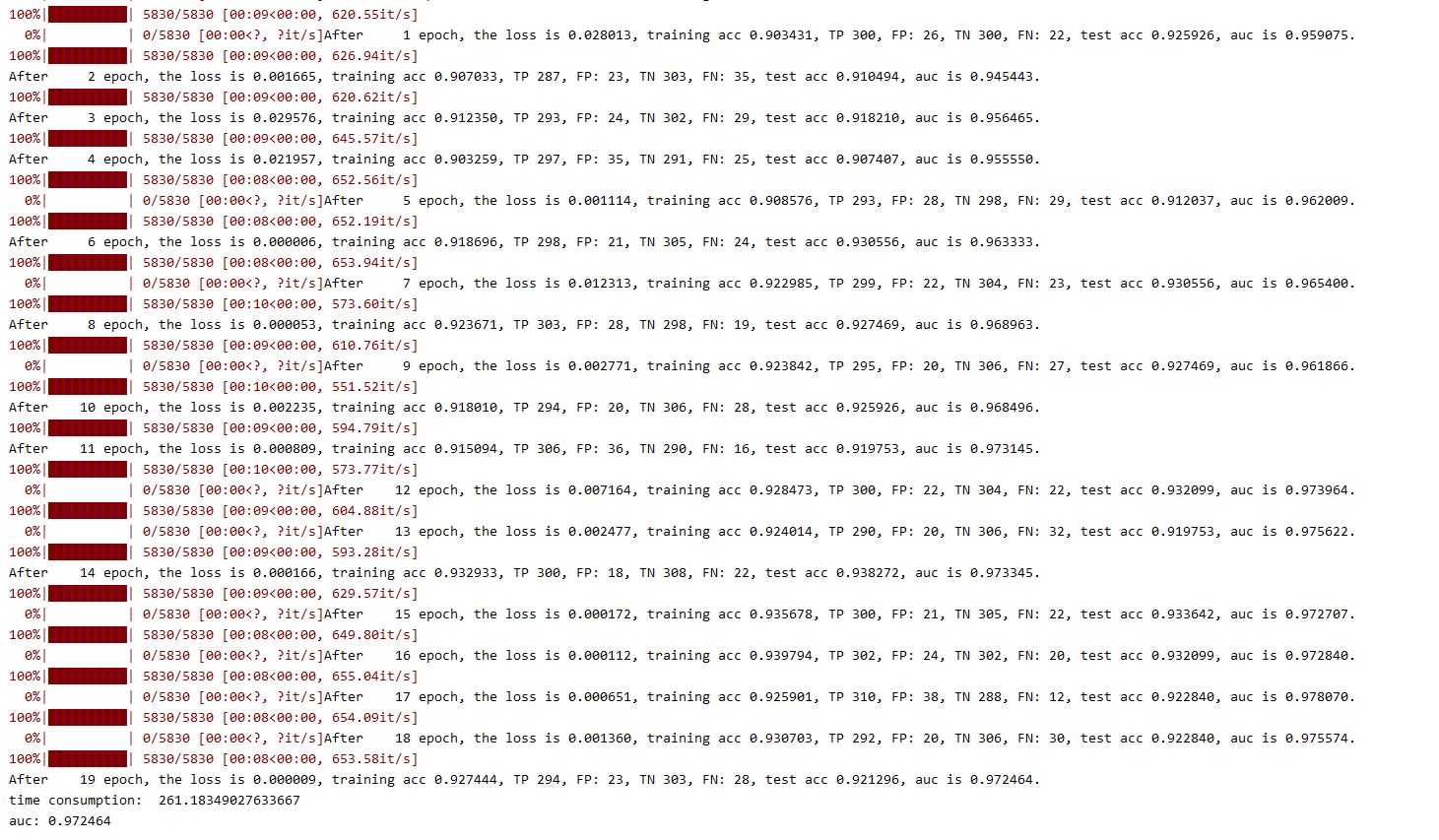
## SEAL

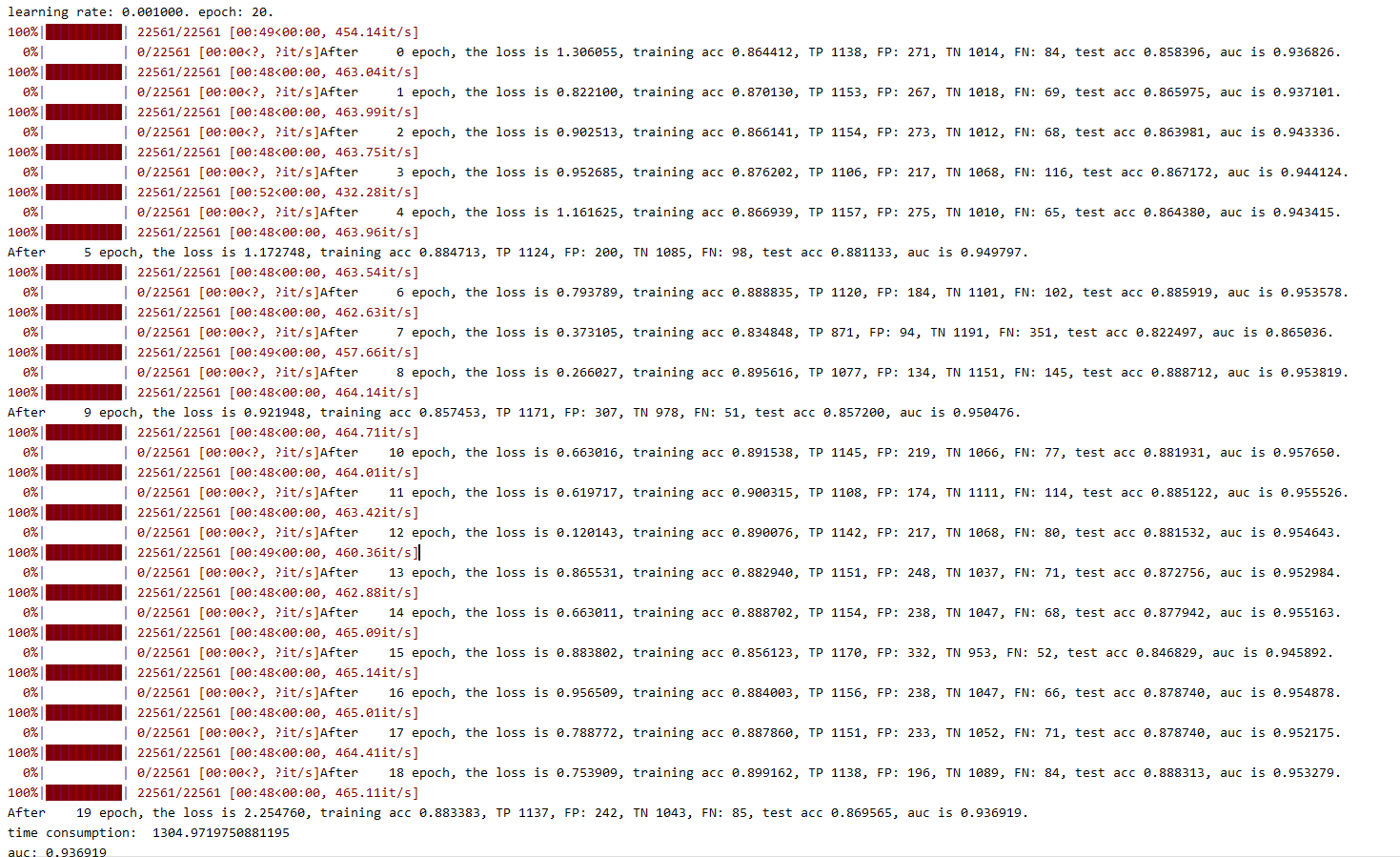
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 数据集 | TP | FP | TN | FN | AP（准确率） | AC（精确率） | AUC |
| USAir | 213 | 34 | 170 | 9 | 0.899 | 0.862 | 0.954 |
| Bio-CE-GT | 294 | 23 | 303 | 28 | 0.922 | 0.927 | 0.973 |
| hamster | 1137 | 242 | 1043 | 85 | 0.870 | 0.823 | 0.937 |
| Yeast | 1070 | 47 | 1104 | 118 | 0.929 | 0.958 | 0.967 |
| PB | 1632 | 230 | 1697 | 246 | 0.875 | 0.876 | 0.943 |

Yeast

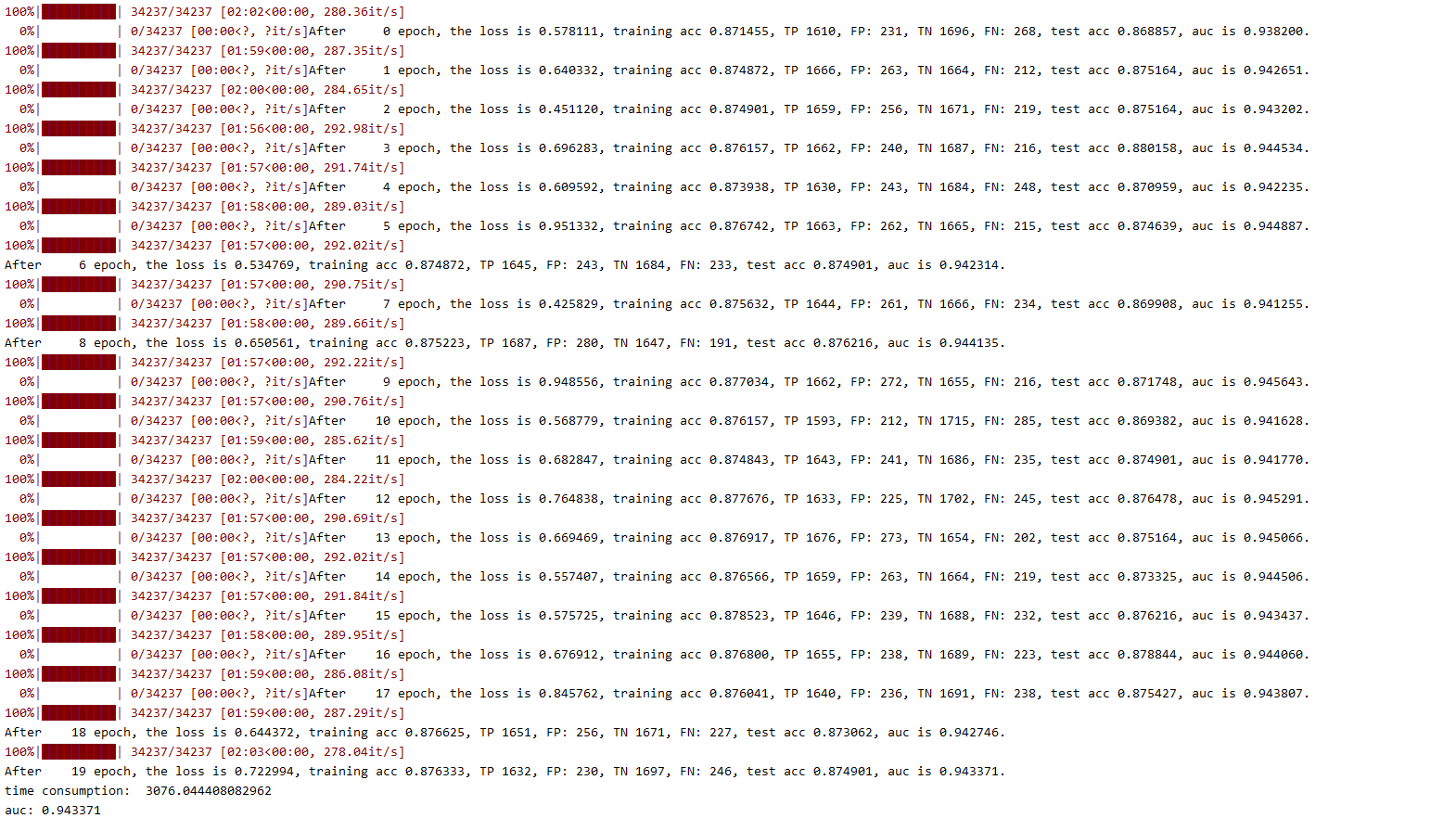


Bio-CE-GT:

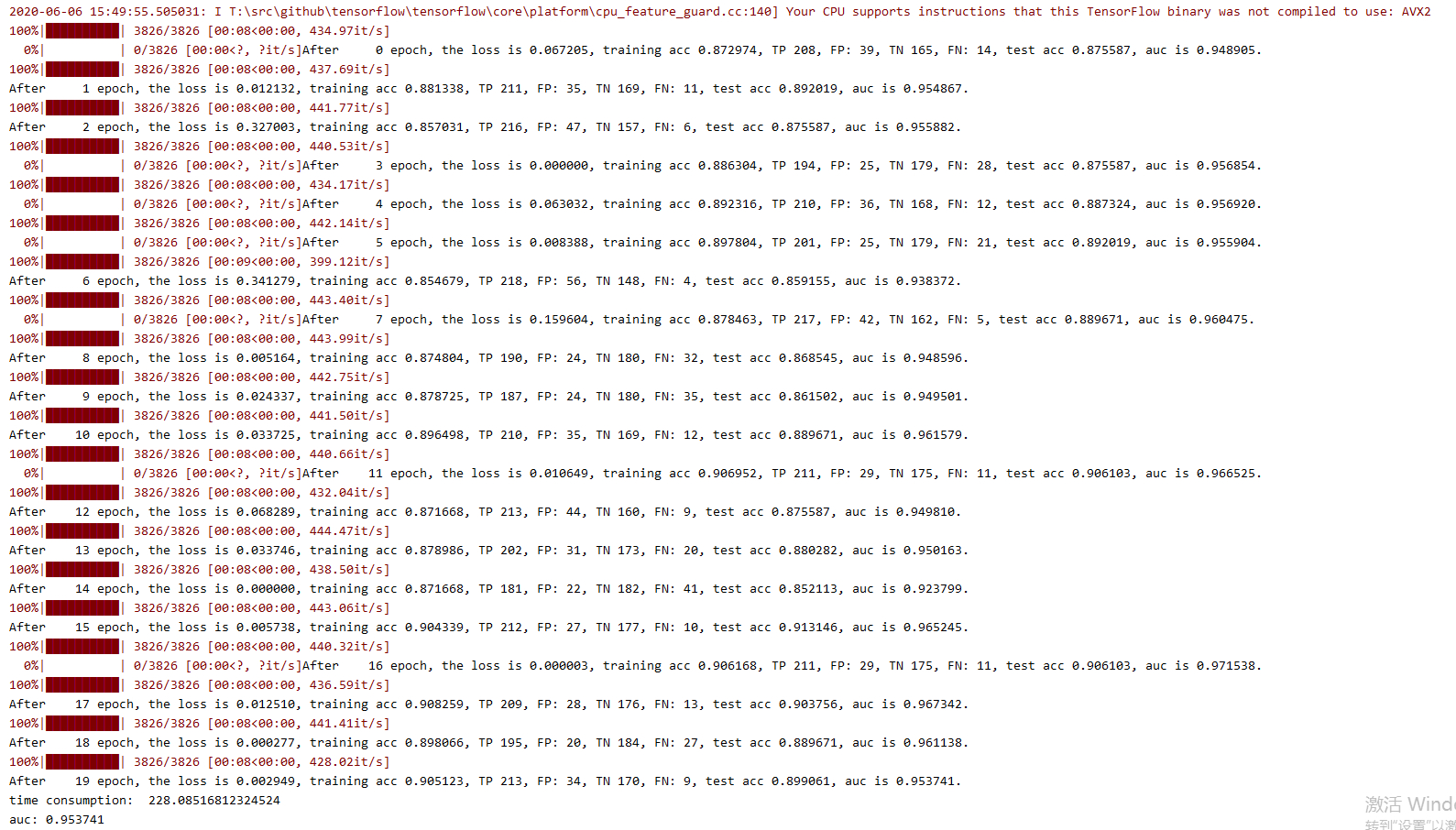


Hamster

PB:



USAir:



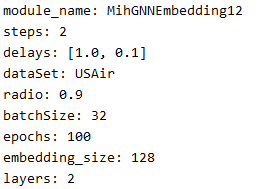
## 使用全数据进行的相关实验

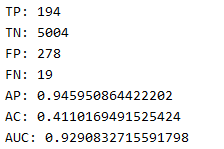
在CrossEntropyLoss中进行如下设置：

nn.CrossEntropyLoss(weight = self.weight, reduction = 'sum')

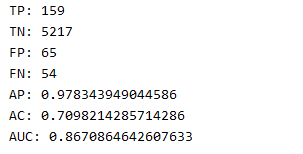
其中weight为[正样本数：负样本数]

### USAir + MihGNNEmbedding12

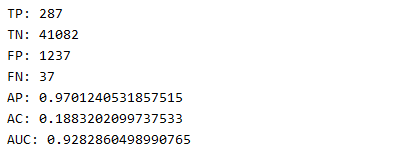




不设置CrossEntropyLoss的权重：



### Bio-CE-GT + MihGNNEmbedding12



不设置CrossEntropyLoss的权重：

