Eugene Kim
2022848518

# YONSEI VISION 48HR ASSIGNMENT

## Task:

Create a deep learning model that is robust to noisy labels

## Methodology:

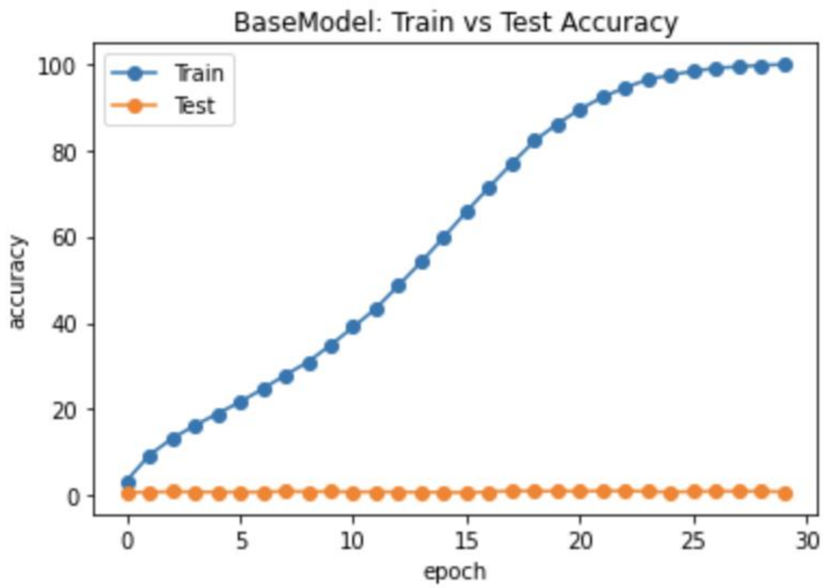
The goal is to separate noisy and true label to decrease the chance of selecting noisy labeled data when training. The separation of noisy and true label is done by training a separate model (supervisor model) with a portion of clean data. This clean data can be obtained by individually examining a portion of the data. I have obtained this data by using a subset of the true CIFAR100 dataset. By having a trustworthy model, we can then use the model to feed in the noisy data and find the cross-entropy loss of each datum. If the cross-entropy loss is small, we will give a larger weight in selecting this datum when training the final model (student model). With the collection of these weights, we can create a probability density function for selecting training batches. The newly sampled data is added to a separate CSV file and used for student model training. The accuracy of label separation depends on the accuracy of the supervisor model.

## Model Description:

The three models created for this project were the base, student, and supervisor model. All three of these models are the same standard ResNet with 8 basic blocks in total. Same model architecture is used for consistency. Each basic block consists of two convolutional layers with batch normalization. The Adam optimizer with a weight decay of 1e-4 was used for L2 regularization. The base model is trained on the noisy data with no prior cleaning. The supervisor model is trained on a portion of the clean CIFAR100 dataset (5000 points). The student model is trained on the weighted data obtained by the supervisor model. To determine the effectiveness of this methodology, the student model's accuracy performance is compared to the base model's accuracy performance.

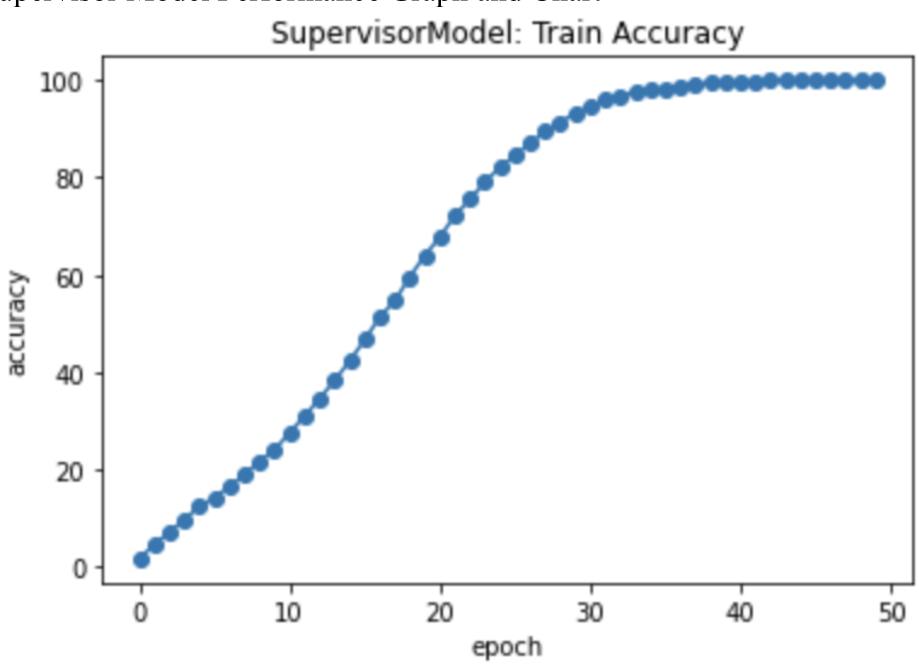
Experiment Results:

Base Model Performance Graph and Chart



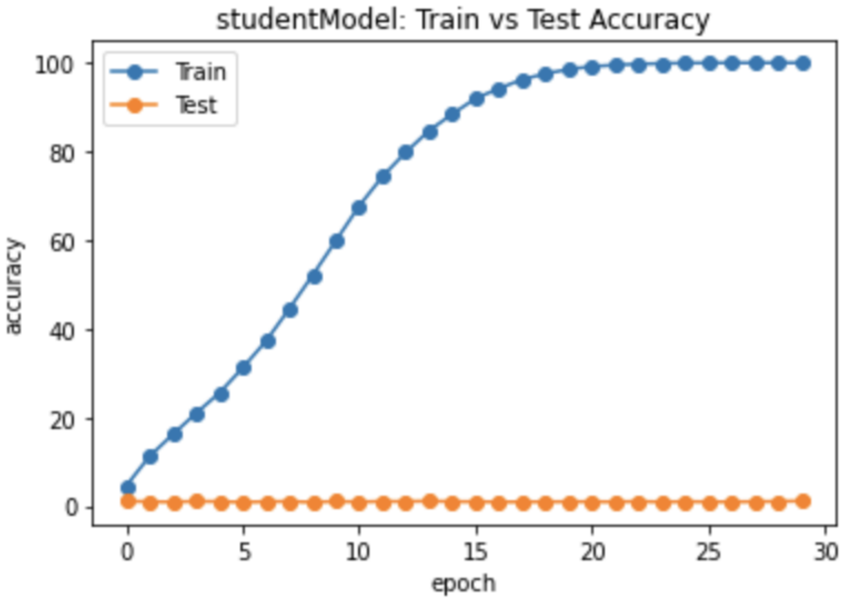

| | train_accuracy | test_accuracy |
|---|---|---|
| **0** | 3.4150 | 0.780156 |
| **2** | 13.1950 | 0.920184 |
| **4** | 18.9450 | 0.820164 |
| **6** | 24.7100 | 0.820164 |
| **8** | 30.9700 | 0.890178 |
| **10** | 39.0275 | 0.850170 |
| **12** | 48.7175 | 0.830166 |
| **14** | 59.8900 | 0.780156 |
| **16** | 71.4775 | 0.860172 |
| **18** | 82.1800 | 1.020204 |
| **20** | 89.4775 | 1.000200 |
| **22** | 94.4975 | 1.110222 |
| **24** | 97.3500 | 0.810162 |
| **26** | 98.9500 | 0.980196 |
| **28** | 99.6125 | 0.980196 |

Supervisor Model Performance Graph and Chart




| | train_accuracy |
|---|---|
| **0** | 1.72 |
| **2** | 7.16 |
| **4** | 12.38 |
| **6** | 16.52 |
| **8** | 21.74 |
| **10** | 27.62 |
| **12** | 34.64 |
| **14** | 42.34 |
| **16** | 51.34 |
| **18** | 59.60 |
| **20** | 67.78 |
| **22** | 75.52 |
| **24** | 82.10 |
| **26** | 87.04 |
| **28** | 91.36 |

Student Model Performance Graph and Chart




| | train_accuracy | test_accuracy |
|---|---|---|
| **0** | 4.5400 | 1.100220 |
| **2** | 16.1400 | 0.910182 |
| **4** | 25.5025 | 1.000200 |
| **6** | 37.2250 | 0.980196 |
| **8** | 52.0725 | 0.850170 |
| **10** | 67.7200 | 0.920184 |
| **12** | 79.8500 | 1.010202 |
| **14** | 88.4700 | 0.990198 |
| **16** | 94.2225 | 0.910182 |
| **18** | 97.5450 | 0.950190 |
| **20** | 99.1625 | 0.970194 |
| **22** | 99.7200 | 1.040208 |
| **24** | 99.9500 | 1.000200 |
| **26** | 99.9900 | 0.940188 |
| **28** | 99.9950 | 1.030206 |

Max test accuracy of base model: 1.140228
Max test accuracy of student model: 1.10222
Mean test accuracy of base model: 0.90985
Mean test accuracy of student model: 0.98753

The results indicate that there is an improvement of test accuracy for the student model compared to the base model. Furthermore, the student model had a faster training convergence rate compared to the base model.

Discussion:

The accuracy scores of the test sets for the base and student models were extremely low due to the lack of training and model complexity. However, any increase in accuracy for the student model is a positive indicator that this methodology is effective. Another positive effect from the student model is the faster training convergence rate. By having less noise in the data, the model was able to learn faster. It is also important to note that the supervisor model plays a significant role in the student model's accuracy. Therefore, I would like to train the supervisor model with more clean data to have better weighted data in the future. In terms of the model architecture, I would like to train a deeper network such as the ResNet50 or ResNet101 to compare the results between the base and student models if I had more time. Furthermore, I would like to experiment more with weight decay and dropout rates to increase generalization. In terms of the methodology, I would like to create different forms of PDFs such as giving weights in an exponential manner instead of a linear manner. This methodology was influenced by the MentorNet (Lu Jiang, etc.), Mixup (Hongyi, Zhang), and MentorMix (Lu Jiang, etc.) research papers. MentorNet and MentorMix both use a method called curriculum learning with a mentor network that supervises the learning of the student network.