Congratulations! You passed!
TO PASS 70% or higher

Keep Learning

GRADE
100%

Phase 3. Project 1

LATEST SUBMISSION GRADE
100%

1. What learning phenomena is the team observing now? 1 / 1 point

- (•) Overfitting
- ( ) Convergence
- ( ) Underfitting
- ( ) Generalization

Correct
Overfitting is when the model begins to memorize the training set and thus achieves very high performance on the training set. Overfitted models fail to generalize, leading to poor performance on the validation set.

Phase 3. Project 1
Graded Quiz • 30 min
Due Mar 29, 2:59 PM CST

- [x] Stronger data augmentation

Correct
Stronger data augmentation makes the training samples more difficult to memorize, which may force the model to learn broader patterns in the training dataset.

- [x] Early stopping

Correct
Early stopping means stopping the model before it memorizes the training dataset. When the model is stopped, it is often before it has memorized the training set, but after it has learned generalizable patterns.

- [x] Dropout

Correct
Dropout randomly sets parameters in the model to zero, which thus requires the model to build in redundancies throughout its layers. The expenditure of model complexity on redundancy encourages the model to learn broader patterns in the training dataset.

- [x] Weight decay (L2 regularization)

Correct
Weight decay encourages the model to be more simple (i.e. induces a penalty in the loss function if the model has too many non-zero weights), thus encouraging the model to learn broader patterns in the training dataset.

- [ ] Increasing the number of model parameters

Phase 3. Project 1
Graded Quiz • 30 min
Due Mar 29, 2:59 PM CST

- (•) An added penalty in the loss function that discourages models from becoming overly complex
- ( ) An added penalty in the loss function that encourages semantic clustering in feature space
- ( ) An added penalty in the loss function that ensures the model is well calibrated
- ( ) An added penalty in the loss function that mitigates class imbalance

Correct
Weight decay adds an extra number to the loss function that grows as the parameter values increase. Weight decay encourages the model to use as few non-zero parameters as possible, thus making it less complex.

4. What does dropout do? 1 / 1 point

- ( ) Dropout randomly removes layers in the network during training in order to prevent overreliance on any one layer
- ( ) Dropout randomly removes neurons in the network during training in order to improve the rate of convergence
- ( ) Dropout randomly removes layers in the network during training in order to improve the rate of convergence
- (•) Dropout randomly removes neurons in the network during training in order to prevent overreliance on any one neuron

Correct
Dropout encourages the model to build in redundancies within itself, which often leads to better generalization.

Phase 3. Project 1
Graded Quiz • 30 min
Due Mar 29, 2:59 PM CST

Correct
Dropout probability is chosen outside of the optimization (training) process, and is therefore a hyperparameter.

- [x] Weight decay strength

Correct
Weight decay strength is chosen outside of the optimization (training) process, and is therefore a hyperparameter.

- [ ] Model weights
- [x] Learning rate

Correct
Learning rate is chosen outside of the optimization (training) process, and is therefore a hyperparameter.

Phase 3. Project 1
Graded Quiz • 30 min
Due Mar 29, 2:59 PM CST

- ( ) AUROC is a poor metric for performance because it can only be used in multi-class settings
- ( ) AUROC is a poor metric for performance because they have a predetermined threshold in mind
- (•) Accuracy is a poor metric for performance because of the high class imbalance
- ( ) Accuracy is a poor metric for performance because of the small number of samples in the test set

Correct
In a situation with high class imbalance, it is possible for the model to predict the majority class (in this case COVID-negative) every time and produce the results much better than random guessing. In this case, the model is predicting negative for every sample, allowing it to achieve 90% accuracy (since 90% of samples are negative cases).

7. Further analysis shows that the model is predicting that every patient is COVID-negative. What can be done to mitigate this effect? **Check all that apply**. 1 / 1 point

- [ ] Lowering the learning rate to improve convergence
- [ ] Using dropout during training to improve performance on the test set
- [x] Undersampling COVID-positive exams during training

Correct
Undersampling COVID-negative exams during training will allow the model to see COVID-positive exams more frequently, thus allowing it to learn more about the positive cases.

- [x] Upweighting COVID-positive exams loss during training

Correct