Model Optimization

Once your model is working, it's time to optimize the model's quality. Follow the steps below.

Add Useful Features

You can improve model performance by adding features that encode information not yet encoded by your existing features. You can find linear correlations between individual features and labels by using correlation matrices. To detect nonlinear correlations between features and labels, you must train the model with and without the feature, or combination of features, and check for an increase in model quality. You must justify the feature's inclusion by an increase in model quality.

Tune Hyperparameters

You found values of hyperparameters that make your model work. However, these hyperparameter values can still be tuned. You can tune the values manually by trial and error, but manual tuning is time consuming. Instead, consider using an automated hyperparameter tuning service, such as Cloud ML Hyperparameter Tuning (https://cloud.google.com/ml-engine/docs/tensorflow/hyperparameter-tuning-overview).

Tune Model Depth and Width

While debugging your model, you only increased model depth and width. In contrast, during model optimization, you either increase or decrease depth and width depending on your goals. If your model quality is adequate, then try reducing overfitting and training time by decreasing depth and width. Specifically, try halving the width at each successive layer. Since your model quality will also decrease, you need to balance quality with overfitting and training time.

Conversely, if you need higher model quality, then try increasing depth and width. For an example, see this Neural Network Playground (https://developers.google.com/machine-learning/crash-course/introduction-to-neural-networks/playground-exercises)
exercise. Remember that increases in depth and width are practically limited by

accompanying increases in training time and overfitting. To understand overfitting, see Generalization: Peril of Overfitting (/machine-learning/crash-course/generalization/peril-of-overfitting).

Since the depth and width are hyperparameters, you can use hyperparameter tuning to optimize depth and width.

Previous
← Check Your Understanding (/machine-learning/testing-debugging/metrics/check-your-understanding)

Next
Programming Exercise → (/machine-learning/testing-debugging/common/programming-exercise)