Assigment # 1

3-aug-24

Gradient Descent

Gradient Descent is a fundamental optimization algorithm used extensively in machine learning and deep learning. It helps models learn and improve by minimizing the loss function. Here's a quick breakdown:

1. **Initialization:** Start with initial parameters (e.g., weights in a neural network).
2. **Compute Gradient:** Calculate the gradient of the loss function, which tells us the direction of the steepest ascent.
3. **Update Parameters:** Move parameters in the opposite direction of the gradient, scaled by a learning rate.
4. **Iterate:** Repeat until the model converges (loss function changes very little) or after a set number of iterations.

Types of Gradient Descent:
- **Batch Gradient Descent:** Uses the entire dataset for each update. Accurate but slow for large datasets.
- **Stochastic Gradient Descent (SGD):** Uses one data point per update. Faster but noisier.
- **Mini-batch Gradient Descent:** Uses a subset of the data, balancing speed and accuracy.

Key Points:
- Learning Rate: Critical hyperparameter; too high can overshoot, and too low can be slow.
- Variants: Different types of gradient descent offer trade-offs in speed and accuracy.
- Convergence: Dependent on the learning rate and the shape of the loss function.

Gradient Descent is essential for optimizing models and making them more effective. It's a simple yet powerful tool in the machine learning toolbox.

Validation Set & Validation Loss

In machine learning, the data is often split into three sets: training, validation, and test sets.

- **Training Set**: Used to train the model.
- **Validation Set**: Used to tune hyperparameters and make decisions about model architecture. It helps in assessing how the model generalizes to unseen data during the training phase.
- **Test Set**: Used to evaluate the performance of the model after the training phase is complete.

Validation Loss

Validation loss is the loss calculated on the validation set. It is used to monitor the model's performance on unseen data and to prevent overfitting. During training, the model's performance on the training set might improve, but it might start to perform poorly on the validation set, indicating overfitting.

- **Early Stopping**: One of the techniques to prevent overfitting is early stopping, which stops training when the validation loss stops improving.
- **Hyperparameter Tuning**: Validation loss helps in tuning hyperparameters by comparing the performance of different models and selecting the best one.

In summary, validation loss is a crucial metric for evaluating and tuning the model during the training process, ensuring that the model generalizes well to new data.

Muhammad jawad khan

DS Batch 5 ned university – Deep Learning

Instructor : Sir Sajid