

Assignment # 1

Deep Learning

Gradient Descent:

Gradient Descent is an optimization technique employed to reduce the cost function in machine learning and deep learning. Its main goal is to identify the optimal parameters (weights) for a model, resulting in the minimal cost and thus the best model performance. This process involves backpropagation, which uses the chain rule.

There are various types of gradient descent learning algorithms are commonly used in machine learning and optimization

1. **Batch Gradient Descent:**

Batch Gradient Descent, also known as vanilla gradient descent, is the simplest form of gradient descent. In this method, the entire training dataset is used to compute the gradients of the cost function with respect to the model parameters in each iteration. Although this can be computationally demanding for large datasets, it guarantees convergence to a local minimum of the cost function.

2. **Stochastic Gradient Descent (SGD):**

Stochastic Gradient Descent (SGD) is a variation of gradient descent that updates the model parameters for each individual training example in the dataset. Unlike batch gradient descent, which computes the gradients using the entire dataset, SGD updates the parameters based on a single, randomly selected training example at a time. This approach can lead to faster convergence due to more frequent updates, but it may also cause more fluctuations in the cost function due to the randomness of the updates.

3. **Mini-batch Gradient Descent:**

Mini-batch Gradient Descent strikes a balance between batch gradient descent and stochastic gradient descent. It computes gradients using a small random subset of the training dataset, typically containing between 10 and 1000 examples, known as a mini-batch. This approach reduces the computational cost compared to batch gradient descent and decreases the variance of updates compared to SGD. Mini-batch gradient descent is widely used in deep learning due to its effective balance between convergence speed and stability.

Validation set:

A validation set is a subset of the dataset used in machine learning to evaluate the performance of a model during training. It is distinct from the training set, which is used to train the model, and the test set, which is used to assess the final performance of the model after training. Purpose of validation set in machine learning is to avoid overfitting and underfitting

Validation Loss:

Validation loss is a measure of how well a machine learning model performs on a validation set during the training process. It represents the error or discrepancy between the predicted outputs by the model and the actual outputs (ground truth) for the data in the validation set. The validation loss is used to evaluate the model's ability to generalize to new, unseen data.