Joint Tech Internship Community Program

Assignment – 1

BY
RAHUL K

Neural Network

Neural Network is a program in machine learning or a model that is similar to humans brains which makes the decision based on how our brains neurons work for analyzing and make decision .Neural Network also have the neurons. For MNIST we'll use neural network with one input and output layer and one or more hidden layers.

Neurons

It is the basic unit in Neural Network .It takes the input data and processes it and give it to the next layers .In MNIST each image pixel is given as input to the neuron in input layer

Layer

A set of neurons in Neural Network. They are mainly classified into 3 Input Layer, Hidden Layer, Output Layer

Input Layer

Layer in which data is taken as input. MNIST have 784 neurons (28 X 28 pixel)

Hidden Layer

Hidden Layers provide the non-linear relationship between the features and the labels

Output Layer

Output layer provides us the predictions For MNIST, it has 10 neurons (one for each digit).

Convolutional Layer

A Layer in Convolutional Neural Network in which convolutional filter passes along the input matrix it is used in feature extraction. In MNIST , convolutional layers can be used to extract features from the digit images

Convolutional Neural Network(CNN)

A Neural Network in which one of the layer is Convolutional Layer and in CNN one or more following pairs of below layer will be present

- Convolutional Layer
- Pooling Layer
- Dense Layer

CNN is mainly used for image classification problems for datasets like MNIST

Recurrent Neural Network(RNN)

RNN is a artificial neural network that uses temporal and time series data .It is used for Language translation, speech recognition , image captioning etc.It is not that much used MNIST dataset.

Activation Function

Activation function is used for activating a neuron it is the summation of product of the inputs and weights, bias. The main purpose is to introduce non linearity in output of a neuron . Common activation functions are ReLU, Sigmoid, and Tanh.

ReLU

Rectified Linear Unit(ReLU) is the most commonly used activation function it mainly is used in hidden layers of neural network. It involves less mathematical operations. In MNIST, ReLU is used in hidden layers to ensure the model can learn non-linear patterns.

Sigmoid

The shape of this activation function is **'s'** shaped. This function exists between 0 to 1.If the output of the model is probability based one then we have to use sigmoid. Used in the output layer for binary classification tasks, but not typically used in MNIST digit classification.

Tanh

It is similar to Sigmoid but it ranges from -1 to +1.The advantage of using this is the -ve inputs will be mapped strongly to -ve and zero inputs to zero. Can be used in hidden layers to normalize the output to the range [-1, 1] for MNIST models.

Softmax

Softmax activation function gives the vector of probabilities as output for the given input. It mainly used for multiclassed classifications. Used in the output layer of an MNIST model to provide the probabilities for each digit (0-9).

Forward Propagation

Forward Propagation is process in neural network where given input data passed through input , hidden, output layers . For MNIST, this means processing the pixel values through the layers to predict the digit.

Back Propagation

This takes the error rate of Forward Propagation and uses the loss for fine tuning the neural network in the direction opposite of Forward Propagation. MNIST, backpropagation adjusts the weights to minimize the difference between the predicted digit and the actual digit.

Loss Function

Loss Function compares the target and predicted values measures how well the neural network models the training data .We try to minimize the loss function.
For MNIST, a common loss function is cross-entropy loss.

Cost Function

Cost function is about comparing the predicted value and the actual value. The Loss function is for training example but the cost function is for entire training set for mini batch gradient descent. In MNIST, the cost function evaluates the overall performance of the neural network across all training examples.

Gradient Descent

Gradient decent is an optimizing algorithm for machine learning and neural networks. It trains machine learning models by minimizing errors between predicted and actual values. It is used to minimize the cost function by iteratively adjusting the network's weights in the direction that reduces the error. Used in MNIST to find the optimal weights that reduce the prediction error.

Learning Rate

A hyperparameter that determines the step size during the gradient descent optimization. It controls how much the model's weights are updated with respect to the loss gradient. In MNIST, a learning rate of 0.01 might be used to balance the speed and accuracy of learning.

Batch Size

Batch Size is a hyperparameter which represents the number of training sample that is passed to the neural network in a single pass. In MNIST, a batch size of 32 means 32 images are processed at a time during training.

Epochs

Epochs is the number of times the learning algorithm will work through the entire dataset. For MNIST, training might involve 10 epochs, meaning the entire dataset is processed 10 times.

Overfitting

When a model performs well for a training dataset and perform poor in the testing data then it is termed as overfitting . An MNIST model that achieves 99% accuracy on training data but only 80% on test data is overfitting.

Underfitting

When a model not learns and performs well for a training data and also it was not able to generalize the testing data then it is termed as Underfitting. An MNIST model that achieves 60% accuracy on both training and test data is underfitting, indicating it is too simple.

Training Set

Training Set is the set of data in the dataset that is used for the training of the model. Often 80% dataset is used as a training set . For MNIST, 48,000 images are used as the training set

Validation Set

Validation set is the set of remaining data that is available in the dataset which is used for validation of the model. For MNIST, 12,000 images might be used as the validation set.

Testing Set

Testing Set is the set of data that is used for testing the model. For MNIST, 10,000 images are used as the test set.

Cross-Validation

Cross-validation is a technique for validating the model efficiency by training it on the subset of input data and testing on previously unseen subset of the input data. In MNIST, 5-fold cross-validation means dividing the dataset into 5 parts, training on 4 parts, and testing on the remaining part, rotating until each part has been tested.

Hyperparameters

Parameters that are set manually for the training the data for effective results Example: Batch Size, Epoch ,etc.. In MNIST, hyperparameters include learning rate, batch size, and the number of layers.

Model Parameters

Model Parameters are learned during the training process itself which includes the weights, bias etc.. In MNIST, the weights and biases of the neurons are model parameters.

Regularization

Regularization in deep learning is a set of techniques that help neural networks avoid overfitting and improve their ability to generalize to new data. In MNIST, L2 regularization might be used to add a penalty for large weights to the loss function.

Dropout

Drop out is one of the regularization technique where randomly selected neurons are ignored during training. It prevents the network from becoming too reliant on specific neurons. In MNIST, dropout might be applied with a 0.5 probability, meaning half the neurons are randomly dropped during each training step.

Weight Initialization

Weight initialization is a procedure to set the weights of a neural network to small random values that define the starting point for the optimization (learning or training) of the neural network model. In MNIST, weights might be initialized using the He initialization method to ensure faster convergence.

Normalization

Normalization is a data preparation technique that is frequently used in machine learning. The process of transforming the columns in a dataset to the same scale is referred to as normalization. It scales input data to have a mean of zero and a standard deviation of one. For MNIST, pixel values might be normalized to the range [0, 1] by dividing each value by 255.

Standardization

Standardization is a data preprocessing technique used in statistics and machine learning to transform the features of dataset so that they have a mean of 0 and a standard deviation of 1

and it is a another term of normalization. For MNIST, pixel values might be standardized by subtracting the mean and dividing by the standard deviation of the dataset.