ECE 449 Machine Learning (ZJU-UIUC Institute)

Small Project

deadline: 2021.11.14

1 Overview

Welcome to ECE449 Small Project!
This small project aims to familarize you with the pipeline of a machine learning problem. In this project, we will provide you a framework of solution to a well-known problem(CIFAR-10 classification). Your task is complete three functions and answer the questions.

2 Cifar-10 classification

The CIFAR-10 dataset (Canadian Institute For Advanced Research) is a collection of images that are commonly used to train machine learning and computer vision algorithms. It is one of the most widely used datasets for machine learning research. The CIFAR-10 dataset contains 60,000 32x32 color images in 10 different classes. The 10 different classes represent airplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks. There are 6,000 images of each class.

An implementation framework on kaggle is provided and the link is below, which can be edit and run directly on the cloud. We hope you can read the whole content and understand each step of a complete project. Then answer the following questions.

https://www.kaggle.com/ruizhechen/cifar-10-small-project

In the notebook, there are three incomplete parts: transforms, class Net(nn.Module) and loss&optimizer. We hope you can complete them according to the instructions.

3 Questions

Preprocessing

3.1 load data

What is the dimension(shape) of training set and test set?

What is the ratio of training set to validation set?

What is the function of class **CIFAR10_from_array**?

3.2 data augmentation

What kinds of data augmentation methods do you use in this project?

What is the benefits of data augmentation?

3.3 dataset normalization

How is the dataset normalized in this project?

What is the benefit of normalization?

Train

3.4 validation

What validation method is used in this project? Describe briefly how to realize.

3.5 build model

Put a screenshot of the structure of your network.

What are the input and output of the model?

3.6 loss&optimizer

What kind of loss and optimizer do you use in this project? Please formulate the equation of loss calculation.

Test

3.7 confusion matrix

How to understand the meaning of confusion matrix?