Medium

Prateek Kumar

2024-12-27

```
library(torch)
```

```
## Warning: package 'torch' was built under R version 4.3.3
```

```
library(magrittr)
```

```
## Warning: package 'magrittr' was built under R version 4.3.3
```

```
spam_url <- "https://hastie.su.domains/ElemStatLearn/datasets/spam.data"
spam_data <- read.table(spam_url, header = FALSE)
x_data <- as.matrix(spam_data[, -ncol(spam_data)])
y_data <- as.numeric(spam_data[, ncol(spam_data)]) + 1
```

```
x_tensor <- torch_tensor(x_data, dtype = torch_float())
y_tensor <- torch_tensor(y_data, dtype = torch_long())
```

```
spam_dataset <- dataset(
  name = "spam_dataset",

  initialize = function(x, y) {
    self$x <- x
    self$y <- y
  },

  .getbatch = function(index) {
    list(
      x = self$x[index, ],
      y = self$y[index]
    )
  },

  .length = function() {
    self$y$size(1)
  }
)
```

```
dataset <- spam_dataset(x=x_tensor, y=y_tensor)
batch_size <- 32
dataloader <- dataloader(dataset, batch_size = batch_size, shuffle = TRUE)
```

```
net <- nn_module(
  "SpamNet",
  initialize = function() {
    self$fc1 <- nn_linear(ncol(x_data), 64)
    self$fc2 <- nn_linear(64, 2)  # Binary classification
  },
  forward = function(x) {
    x %>%
      self$fc1() %>%
      nnf_relu() %>%
      self$fc2() %>%
      nnf_log_softmax(dim = 1)
  }
)

model <- net()
optimizer <- optim_sgd(model$parameters, lr = 0.01)
```

```
num_epochs <- 10
for (epoch in 1:num_epochs) {
  losses <- c()

  coro::loop(for (batch in dataloader) {
    optimizer$zero_grad()
    output <- model(batch$x)
    loss <- nnf_nll_loss(output, batch$y)
    loss$backward()
    optimizer$step()
    losses <- c(losses, loss$item())
  })

  cat(sprintf("Epoch %d: Loss = %.4f\n", epoch, mean(losses)))
}
```

```
## Epoch 1: Loss = 15.2835
## Epoch 2: Loss = 3.4516
## Epoch 3: Loss = 3.4838
## Epoch 4: Loss = 3.5370
## Epoch 5: Loss = 3.4524
## Epoch 6: Loss = 3.4494
## Epoch 7: Loss = 3.4467
## Epoch 8: Loss = 3.4456
## Epoch 9: Loss = 3.4494
## Epoch 10: Loss = 3.4516
```