UPN Intelligent Systems Course Final Grade Project 2020

Premise

Train a NN to determine if student will approve a course, based on this rules:

- Consider three test results.
- The student will have almost two approved units to pass the course.
- To get the final result, there's no math calculation. Just take the mentioned previously.
- Results should be between 0 to 20.

```
TaskLocalRNG()
```

```
begin
    using Flux ✓
    using Plots ✓
    using Random ✓
    using Metrics ✓
    using MLUtils ✓
    using Statistics ✓
    using ProgressLogging ✓
    Random.seed!(9)
end
```

Generating grades data

```
generatedata (generic function with 1 method)
```

```
generatedata() = rand(0:20, 3) |> Vector{Int32}
```

```
checkcriteria (generic function with 1 method)
```

```
checkcriteria(t) = (map(x -> x>11, t) |> sum > 1) |> Int32
```

```
n_student = 100000
```

```
n_student = 100_000
```

```
generated_data =
```

```
[[13, 8, 0], [6, 17, 5], [4, 17, 18], [0, 15, 7], [3, 11, 2], [19, 12, 8], [9, 18, 19],
```

```
generated_data = map(_ -> generatedata(), 1:n_student)
```

```
data_x =
3×100000 Matrix{Int32}:
 13   6   4   0   3  19   9   5   7  20  3  …   7  15  17  20  20   4   0  10   1  13
  8  17  17  15  11  12  18   6  18   3  6      14   2  20   7   4   8  20  19  12   7
  0   5  18   7   2   8  19  15  14   1  5      17  11   8   8   8  13  19  18  17  14
```

```
data_x = reduce(hcat, generated_data)
```

```
data_y =
1×100000 Matrix{Int32}:
 0  0  1  0  0  1  1  0  1  0  0  0  0  0  …  0  0  0  1  0  1  0  0  0  1  1  1  1
```

```
data_y = generated_data' .|> checkcriteria
```

Preparing data

```
norm_data_x =
3×100000 Matrix{Float32}:
  0.494644  -0.660265  -0.990239  -1.65019    …  -0.000316775  -1.4852     0.494644
 -0.331437   1.15444    1.15444    0.824248       1.48464       0.328954  -0.496535
 -1.65127   -0.824478   1.32518   -0.493761       1.32518       1.15983    0.66375
```

```
norm_data_x = Flux.normalise(data_x) |> Matrix{Float32}
```

```
((3×80000 view(::Matrix{Float32}, :, 1:80000) with eltype Float32:
     0.494644  -0.660265  -0.990239  -1.65019    …   0.989605  -0.660265  -0.000316775
    -0.331437   1.15444    1.15444    0.824248       0.824248   0.163856  -0.826731
    -1.65127   -0.824478   1.32518   -0.493761      -0.163044  -1.32055   -1.48591
```

```
train_data, test_data = MLUtils.splitobs((norm_data_x, data_y); at=0.8)
```

```
train_data_loader =
80-element DataLoader(::Tuple{SubArray{Float32, 2, Matrix{Float32}, Tuple{Base.Slice{B
  with first element:
  (3×1000 Matrix{Float32}, 1×1000 Matrix{Int32},)
```

```
train_data_loader = DataLoader(train_data, batchsize=1000)
```

```
test_data_loader =
20000-element DataLoader(::Tuple{SubArray{Float32, 2, Matrix{Float32}, Tuple{Base.Slic
  with first element:
  (3×1 Matrix{Float32}, 1×1 Matrix{Int32},)
```

```
test_data_loader = DataLoader(test_data)
```

Defining model

```
Adam(0.05, (0.9, 0.999), 1.0e-8, IdDict())
```

```
begin
    model = Chain(
        Dense(3 => 8, relu),
        Dense(8 => 1, sigmoid)
    ) |> gpu
    optimizer = Adam(0.05)
end
```

```
Chain(
  Dense(3 => 8, relu),                  # 32 parameters
  Dense(8 => 1, σ),                     # 9 parameters
)                   # Total: 4 arrays, 41 parameters, 420 bytes.
```

```
model
```

```
loss (generic function with 1 method)
```

```
loss(x, y) = Flux.binarycrossentropy(model(x), y) |> gpu
```

Training phase

```
epochs = 15
```

```
epochs = 15
```

```
begin
    train_losses = Float32[]
    test_losses = Float32[]

    @progress for e in 1:epochs
        for d in train_data_loader
            gs = gradient(Flux.params(model)) do
                l = loss(d...)
            end
            Flux.update!(optimizer, Flux.params(model), gs)
        end
        push!(train_losses, loss(train_data_loader.data...))
        push!(test_losses, loss(test_data_loader.data...))
    end
end
```

100%

Testing phase

```
accuracy (generic function with 1 method)
```

```
accuracy(y, Y) = mean(Y .== y)
```

```
begin
    test_pred = Int32[]
    test_truth = test_data_loader.data[2]

    for (x, y) in test_data_loader
        pred = round(Int32, model(x)[1])
        append!(test_pred, pred)
    end
end
```

```
incorrect_predictions = 5
```

```
incorrect_predictions = test_truth .- test_pred' |> sum |> abs
```

```
@info "Correct predictions: $(size(test_pred)[1] - incorrect_predictions) -
Incorrect: $incorrect_predictions"
```

```
Correct predictions: 19995 - Incorrect: 5
```

```
mae = 0.00035
```

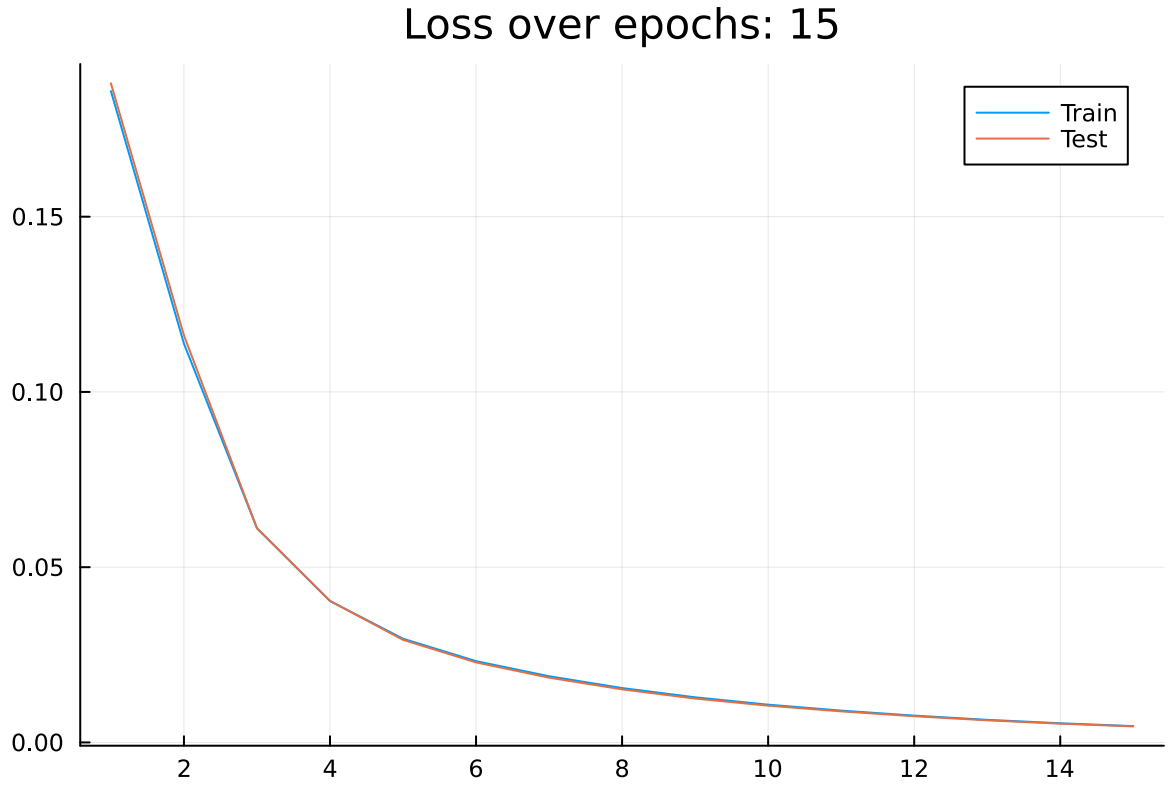
```
mae = Metrics.mae(test_pred', test_truth)
```

```
mse = 0.00035
```

```
mse = Metrics.mse(test_pred', test_truth)
```

```
acc = 0.99965
```

```
acc = accuracy(test_pred', test_truth)
```




```
begin
    plot(1:epochs, train_losses, title="Loss over epochs: $epochs", label="Train")
    plot!(1:epochs, test_losses, label="Test")
end
```