Zhengming Wang

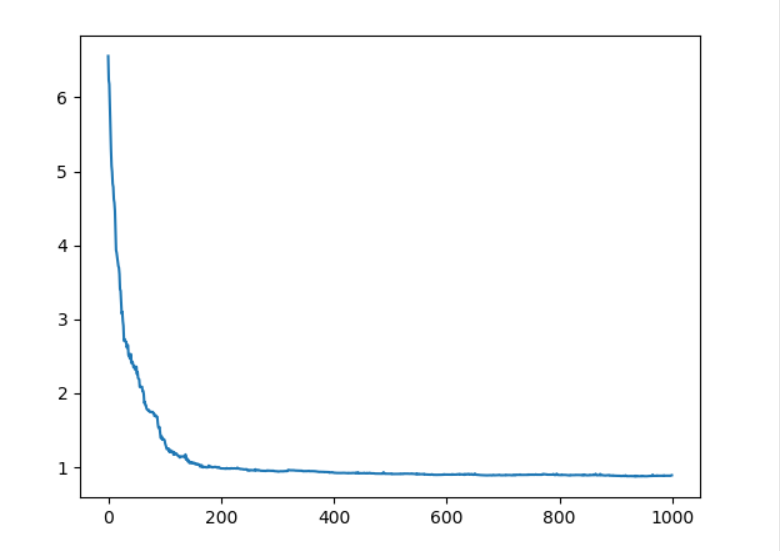
CS1699 – Deep Learning

Question 1

* First example: x = [1 1 1], y = [0 0]
* Weights are w(1)10 = w(1)20 = w(1)11 = w(1)21 = w(1)12 = w(1)22 = w(2)10 = w(2)20 = w(2)11 = w(2)21 = w(2)12 = w(2)22 = 0.05
* Activations are z1 = 0.5374, z2= 0.5374, y1=0.5259, y2=0.5259
* Compute Errors:
  + δy1 = y1 \* (1 - y1) \* (y1-ytrue) = 0.5259 \* (1 – 0.5259) \* (0.5259 – 0) = 0.1311
  + δy2 = y2 \* (1 - y2) \* (y2-ytrue) = 0.5259 \* (1 – 0.5259) \* (0.5259 – 0) = 0.1311
  + δz1= z1 \* (1 - z1) \* (w211\*δy1 + w221\*δy2) = 0.5374 \* (1 – 0.5374) \* (0.05 \* 0.1311 + 0.05 \* 0.1311) = 0.0033
  + δz2= 1z2 \* (1 - z2) \* (w212\*δy1 + w222\*δy2) = 0.5374 \* (1 – 0.5374) \* (0.05 \* 0.1311 + 0.05 \* 0.1311) = 0.0033
* Update weights
  + w(2)10= w(2)10 – 0.3 \* δy1 \* z0 = 0.05 – 0.3 \*0.1311 \* 1 = 0.0107
  + w(2)20= w(2)20 – 0.3 \* δy2 \* z0 = 0.05 – 0.3 \*0.1311 \* 1 = 0.0107
  + w(2)11= w(2)11 – 0.3 \* δy1 \* z1 = 0.05 – 0.3 \* 0.1311 \* 0.5374 = 0.0289
  + w(2)21= w(2)21 – 0.3 \* δy2 \* z1 = 0.05 – 0.3 \* 0.1311 \* 0.5374 = 0.0289
  + w(2)12= w(2)12 – 0.3 \* δy1 \* z2 = 0.05 – 0.3 \* 0.1311 \* 0.5374 = 0.0289
  + w(2)22= w(2)22 – 0.3 \* δy2 \* z2 = 0.05 – 0.3 \* 0.1311 \* 0.5374 = 0.0289
  + w(1)10= w(1)10 – 0.3 \* δz1 \* x0 = 0.05 – 0.3 \* 0.0033 \* 1 = 0.0490
  + w(1)20= w(1)20 – 0.3 \* δz2 \* x0 = 0.05 – 0.3 \* 0.0033 \* 1 = 0.0490
  + w(1)11= w(1)11 – 0.3 \* δz1 \* x1 = 0.05 – 0.3 \* 0.0033 \* 1 = 0.0490
  + w(1)21= w(1)21 – 0.3 \* δz2 \* x1 = 0.05 – 0.3 \* 0.0033 \* 1 = 0.0490
  + w(1)12= w(1)12 – 0.3 \* δz1 \* x2 = 0.05 - 0.3 \* 0.0033 \* 1 = 0.0490
  + w(1)22= w(1)22 – 0.3 \* δz2 \* x2 = 0.05 – 0.3 \*0.0033 \* 1 = 0.0490

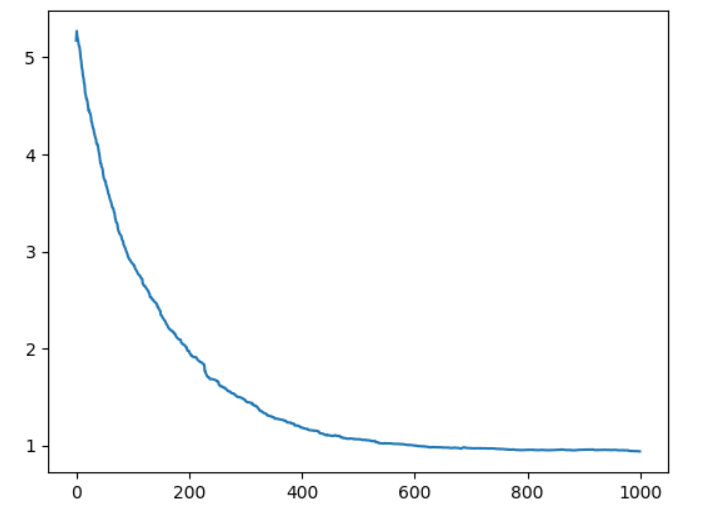
Learning Rate: 0.09

Error Over Time: 0.845196



Learning Rate: 0.009

Error Over Time: 0.91709



Learning Rate: 0.05

Error Over Time: 0.948737

