**Step 0:** Read input and output

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X |  |  |  | Wh |  |  | bh |  |  | Hidden i/p |  |  | Hidden activations |  |  | wout | bout | output | Y | E |
| 1 | 0 | 1 | 0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 1 |  |
| 1 | 0 | 1 | 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 1 |  |
| 0 | 1 | 0 | 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 0 |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Step 1:** Initialize weights and biases with random values

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X |  |  |  | Wh |  |  | bh |  |  | Hidden i/p |  |  | Hidden activations |  |  | wout | bout | output | Y | E |
| 1 | 0 | 1 | 0 | 0.0602 | 0.7694 | 0.6673 | 0.5062 | 0.6823 | 0.9351 |  |  |  |  |  |  | 0.6615 | 0.0363 |  | 1 |  |
| 1 | 0 | 1 | 1 | 0.6918 | 0.0209 | 0.3092 |  |  |  |  |  |  |  |  |  | 0.6214 |  |  | 1 |  |
| 0 | 1 | 0 | 1 | 0.8209 | 0.5585 | 0.5298 |  |  |  |  |  |  |  |  |  | 0.3703 |  |  | 0 |  |
|  |  |  |  | 0.2188 | 0.8437 | 0.7764 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Step 2:** Calculate hidden layer input

hidden\_layer\_input= matrix\_dot\_product(X,wh) + bh

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X |  |  |  | Wh |  |  | bh |  |  | Hidden i/p |  |  | Hidden activations |  |  | wout | bout | output | Y | E |
| 1 | 0 | 1 | 0 | 0.0602 | 0.7694 | 0.6673 | 0.5062 | 0.6823 | 0.9351 | 1.3873 | 2.0102 | 2.1322 |  |  |  | 0.6615 | 0.0363 |  | 1 |  |
| 1 | 0 | 1 | 1 | 0.6918 | 0.0209 | 0.3092 |  |  |  | 1.6061 | 2.8539 | 2.9086 |  |  |  | 0.6214 |  |  | 1 |  |
| 0 | 1 | 0 | 1 | 0.8209 | 0.5585 | 0.5298 |  |  |  | 1.4168 | 1.5469 | 2.0207 |  |  |  | 0.3703 |  |  | 0 |  |
|  |  |  |  | 0.2188 | 0.8437 | 0.7764 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Step 3:** Perform non-linear transformation on hidden linear input  
*hiddenlayer\_activations = sigmoid(hidden\_layer\_input)*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X |  |  |  | Wh |  |  | bh |  |  | Hidden i/p |  |  | Hidden activations |  |  | wout | bout | output | Y | E |
| 1 | 0 | 1 | 0 | 0.0602 | 0.7694 | 0.6673 | 0.5062 | 0.6823 | 0.9351 | 1.3873 | 2.0102 | 2.1322 | 0.8002 | 0.8819 | 0.8819 | 0.6615 | 0.0363 |  | 1 |  |
| 1 | 0 | 1 | 1 | 0.6918 | 0.0209 | 0.3092 |  |  |  | 1.6061 | 2.8539 | 2.9086 | 0.8329 | 0.9455 | 0.9483 | 0.6214 |  |  | 1 |  |
| 0 | 1 | 0 | 1 | 0.8209 | 0.5585 | 0.5298 |  |  |  | 1.4168 | 1.5469 | 2.0207 | 0.8048 | 0.8245 | 0.8830 | 0.3703 |  |  | 0 |  |
|  |  |  |  | 0.2188 | 0.8437 | 0.7764 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Step 4:** Perform linear and non-linear transformation of hidden layer activation at output layer

output\_layer\_input = matrix\_dot\_product (hiddenlayer\_activations \* wout ) + bout  
*output = sigmoid(output\_layer\_input)*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X |  |  |  | Wh |  |  | bh |  |  | Hidden i/p |  |  | Hidden activations |  |  | wout | bout | output | Y | E |
| 1 | 0 | 1 | 0 | 0.0602 | 0.7694 | 0.6673 | 0.5062 | 0.6823 | 0.9351 | 1.3873 | 2.0102 | 2.1322 | 0.8002 | 0.8819 | 0.8819 | 0.6615 | 0.0363 |  | 1 | 0.1908 |
| 1 | 0 | 1 | 1 | 0.6918 | 0.0209 | 0.3092 |  |  |  | 1.6061 | 2.8539 | 2.9086 | 0.8329 | 0.9455 | 0.9483 | 0.6214 |  |  | 1 | 0.1786 |
| 0 | 1 | 0 | 1 | 0.8209 | 0.5585 | 0.5298 |  |  |  | 1.4168 | 1.5469 | 2.0207 | 0.8048 | 0.8245 | 0.8830 | 0.3703 |  |  | 0 | -0.8035 |
|  |  |  |  | 0.2188 | 0.8437 | 0.7764 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Step 5:** Calculate gradient of Error(E) at output layer  
*E = y-output*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X |  |  |  | Wh |  |  | bh |  |  | Hidden i/p |  |  | Hidden activations |  |  | wout | bout | output | Y | E |
| 1 | 0 | 1 | 0 | 0.0602 | 0.7694 | 0.6673 | 0.5062 | 0.6823 | 0.9351 | 1.3873 | 2.0102 | 2.1322 | 0.8002 | 0.8819 | 0.8819 | 0.6615 | 0.0363 | 0.8092 | 1 | 0.1908 |
| 1 | 0 | 1 | 1 | 0.6918 | 0.0209 | 0.3092 |  |  |  | 1.6061 | 2.8539 | 2.9086 | 0.8329 | 0.9455 | 0.9483 | 0.6214 |  | 0.8092 | 1 | 0.1786 |
| 0 | 1 | 0 | 1 | 0.8209 | 0.5585 | 0.5298 |  |  |  | 1.4168 | 1.5469 | 2.0207 | 0.8048 | 0.8245 | 0.8830 | 0.3703 |  | 0.8035 | 0 | -0.8035 |
|  |  |  |  | 0.2188 | 0.8437 | 0.7764 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Step 6:** Compute slope at output and hidden layer  
*Slope\_output\_layer= derivatives\_sigmoid(output)*  
*Slope\_hidden\_layer = derivatives\_sigmoid(hiddenlayer\_activations)*

|  |  |
| --- | --- |
| Learning rate | 0.1 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| delta hidden lyr |  |  | slope hidden lyr |  |  | error at hidden lyr |  |  | delta o/p | slope o/p | E |
|  |  |  | 0.1599 | 0.1042 | 0.0948 |  |  |  |  | 0.1544 | 0.1908 |
|  |  |  | 0.1392 | 0.1392 | 0.1392 |  |  |  |  | 0.1467 | 0.1786 |
|  |  |  | 0.1571 | 0.1447 | 0.1033 |  |  |  |  | 0.1579 | -0.8035 |

**Step 7:** Compute delta at output layer : *d\_output = E \* slope\_output\_layer\*lr*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| delta hidden lyr |  |  | slope hidden lyr |  |  | error at hidden lyr |  |  | delta o/p | slope o/p | E |
|  |  |  | 0.1599 | 0.1042 | 0.0948 |  |  |  | 0.0295 | 0.1544 | 0.1908 |
|  |  |  | 0.1392 | 0.1392 | 0.1392 |  |  |  | 0.0262 | 0.1467 | 0.1786 |
|  |  |  | 0.1571 | 0.1447 | 0.1033 |  |  |  | -0.1269 | 0.1579 | -0.8035 |

**Step 8:** Calculate Error at hidden layer

Error\_at\_hidden\_layer = matrix\_dot\_product(d\_output, wout.Transpose)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| delta hidden lyr |  |  | slope hidden lyr |  |  | error at hidden lyr |  |  | delta o/p | slope o/p | E |
|  |  |  | 0.1599 | 0.1042 | 0.0948 | 0.0195 | 0.0183 | 0.0109 | 0.0295 | 0.1544 | 0.1908 |
|  |  |  | 0.1392 | 0.1392 | 0.1392 | 0.0109 | 0.0163 | 0.0097 | 0.0262 | 0.1467 | 0.1786 |
|  |  |  | 0.1571 | 0.1447 | 0.1033 | -0.0839 | -0.0789 | -0.0470 | -0.1269 | 0.1579 | -0.8035 |

**Step 9:** Compute delta at hidden layer

*d\_hiddenlayer = Error\_at\_hidden\_layer \* slope\_hidden\_layer*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| delta hidden lyr |  |  | slope hidden lyr |  |  | error at hidden lyr |  |  | delta o/p | slope o/p | E |
| 0.0031 | 0.0019 | 0.0010 | 0.1599 | 0.1042 | 0.0948 | 0.0195 | 0.0183 | 0.0109 | 0.0295 | 0.1544 | 0.1908 |
| 0.0024 | 0.0008 | 0.0005 | 0.1392 | 0.1392 | 0.1392 | 0.0109 | 0.0163 | 0.0097 | 0.0262 | 0.1467 | 0.1786 |
| -0.0132 | -0.0132 | -0.0132 | 0.1571 | 0.1447 | 0.1033 | -0.0839 | -0.0789 | -0.0470 | -0.1269 | 0.1579 | -0.8035 |

**Step 10:** Update weights and biases at both output and hidden layer

wout = wout + matrix\_dot\_product(hiddenlayer\_activations.Transpose, d\_output)\*learning\_rate  
wh = wh+ matrix\_dot\_product(X.Transpose,d\_hiddenlayer)\*learning\_rate

*bh = bh + sum(d\_hiddenlayer, axis=0) \* learning\_rate  
bout = bout + sum(d\_output, axis=0)\*learning\_rate*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X |  |  |  | Wh |  |  | bh |  |  | Hidden i/p |  |  | Hidden activations |  |  | wout | bout | output | Y | E |
| 1 | 0 | 1 | 0 | 0.0608 | 0.7697 | 0.6675 | 0.5054 | 0.6814 | 0.9348 | 1.3873 | 2.0102 | 2.1322 | 0.8002 | 0.8819 | 0.8819 | 0.6558 | 0.0292 | 0.8092 | 1 | 0.1908 |
| 1 | 0 | 1 | 1 | 0.6905 | 0.0198 | 0.3087 |  |  |  | 1.6061 | 2.8539 | 2.9086 | 0.8329 | 0.9455 | 0.9483 | 0.6160 |  | 0.8092 | 1 | 0.1786 |
| 0 | 1 | 0 | 1 | 0.3087 | 0.5588 | 0.5300 |  |  |  | 1.4168 | 1.5469 | 2.0207 | 0.8048 | 0.8245 | 0.8830 | 0.3642 |  | 0.8035 | 0 | -0.8035 |
|  |  |  |  | 0.2177 | 0.8426 | 0.7760 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |