CENG499

Homework 1

Part 1

Berke Sina Ahlatcı

2468502

- $a_{ik}^{(0)} = a_{ik}^{(0)} - \alpha \frac{\partial SSE(y, O_0^{(2)})}{\partial a_{ik}^{(0)}}$

  1. $\Delta a_{ik}^{(0)} = -\alpha \sum_k \frac{\partial S}{\partial O_0^{(2)}} \frac{\partial O_0^{(2)}}{\partial O_k^{(1)}} \frac{\partial O_k^{(1)}}{\partial a_{ik}^{(0)}}$
  2. Replacing S with $(y - O_0^{(2)})^2$
  3. Addressing the outcome of $\frac{\partial O_0^{(2)}}{\partial O_k^{(1)}}$ as t_k and the derivative of the weighted sum with respect to a_{ik} as z_k
  4. $\Delta a_{ik}^{(0)} = -\alpha \sum_k -(2y - 2O_0^{(2)}) * t_k * O_k^{(1)}(1 - O_k^{(1)}) z_k$
  5. $a_{ik}^{(0)} = a_{ik}^{(0)} + \alpha \sum_k (2y - 2O_0^{(2)}) * t_k * O_k^{(1)}(1 - O_k^{(1)}) z_k$

- $a_{ik}^{(1)} = a_{ik}^{(1)} - \alpha \frac{\partial SSE(y, O_0^{(2)})}{\partial a_{k0}^{(1)}}$

  1. $\Delta a_{ik}^{(1)} = -\alpha \sum_k \frac{\partial S}{\partial O_0^{(2)}} \frac{\partial O_0^{(2)}}{\partial a_{ik}^{(1)}}$
  2. Replacing S with $(y - O_0^{(2)})^2$
  3. Addressing the derivative of the weighted sum with respect to a_{ik} as z_k
  4. $\Delta a_{ik}^{(1)} = -\alpha \sum_k -(2y - 2O_0^{(2)}) O_k^{(2)}(1 - O_k^{(2)}) z_k$
  5. $a_{ik}^{(1)} = a_{ik}^{(1)} \alpha \sum_k (2y - 2O_0^{(2)}) O_k^{(2)}(1 - O_k^{(2)}) z_k$

- $a_{ik}^{(0)} = a_{ik}^{(0)} - \alpha \frac{\partial CE([l_0, l_1, l_2], O^{(2)} = [O_0^{(2)}, O_1^{(2)}, O_2^{(2)})}{\partial a_{ik}^{(0)}}$

  1. $\Delta a_{ik}^{(0)} = -\alpha \sum_{k,i} \frac{\partial CE}{\partial O_i^{(2)}} \frac{\partial O_i^{(2)}}{\partial O_k^{(1)}} \frac{\partial O_k^{(1)}}{\partial a_{ik}^{(0)}}$
  2. Replacing CE with the cross-entropy function.
  3. Addressing the outcome of $\frac{\partial O_i^{(2)}}{\partial O_k^{(2)}}$ as t_{ik} and the derivative of the weighted sum with respect to a_{ik} as z_k
  4. $\Delta a_{ik}^{(0)} = -\sum_{k,i} -\frac{l_i}{O_i^{(2)}} * t_{ik} * O_k^{(1)}(1 - O_k^{(1)}) z_k$
  5. $a_{ik}^{(0)} = a_{ik}^{(0)} + \sum_{k,i} \frac{l_i}{O_i^{(2)}} * t_{ik} * O_k^{(1)}(1 - O_k^{(1)}) z_k$

- $a_{ik}^{(1)} = a_{ik}^{(1)} - \alpha \frac{\partial CE([l_0, l_1, l_2], O^{(2)} = [O_0^{(2)}, O_1^{(2)}, O_2^{(2)})}{\partial a_{ik}^{(1)}}$

  1. $\Delta a_{ik}^{(1)} = -\alpha \sum_{k,i} \frac{\partial CE}{\partial O_i^{(2)}} \frac{\partial O_i^{(2)}}{\partial a_{ik}^{(1)}}$
  2. Replacing CE with the cross-entropy function.
  3. Addressing the derivative of the weighted sum with respect to a_{ik} as z_k
  4. $\Delta a_{ik}^{(1)} = -\sum_{k,i} -\frac{l_i}{O_i^{(2)}} O_i^{(2)}(1 - O_i^{(2)}) z_k$
  5. $a_{ik}^{(1)} = a_{ik}^{(1)} + \sum_{k,i} \frac{l_i}{O_i^{(2)}} O_i^{(2)}(1 - O_i^{(2)}) z_k$