Homework-8

Q1:

Consider a deep neural net applied to decide between the following four categories:

cat, tiger, human face, lion

The neural net uses a softmax unit at the output layer. Consider the case where the values fed into the output layer are:

|  |  |
| --- | --- |
| cat | 0.5 |
| tiger | 0.8 |
| human face | -3 |
| lion | 0.6 |

The softmax converts these values into a probability vector.

1. Compute the probability vector.

|  |  |  |  |
| --- | --- | --- | --- |
| categories | real value (V) | q = e^V | p = q / Z |
| cat | 0.5 | 1.64872 | 0.28693 |
| tiger | 0.8 | 2.22554 | 0.38731 |
| human face | -3 | 0.04979 | 0.00866 |
| lion | 0.6 | 1.82212 | 0.31710 |

(use

in the calculation.)

Q (0.28693, 0.38731, 0.00866, 0.31710)

2. Which outcome is the most likely?

Tiger. Tiger’s probability is the maximum which is 0.38731.

3. Which outcome is the least likely?

Human face. Human face’s probability is the minimum which is 0.00866.

4. What is the result of cross-entropy cost function if the target output is lion?

cat, tiger, human face, lion => P (0, 0, 0, 1)

Q2:

In the table below cases 3, 4 are distributions, and cases 1, 2 can be converted into distributions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| case | A | B | C | D |
| 1 | 1 | -2 | 3 | -4 |
| 2 | 1 | 2 | -3 | 0 |
| 3 | 1 | 0 | 0 | 0 |
| 4 | 1/4 | 1/4 | 1/4 | 1/4 |

1. Use cross entropy to determine which distribution among 1, 2, 3 is most similar to 4. Show your computations.

Firstly, use softmax to convert cases 1, 2 into distributions.

|  |  |  |  |
| --- | --- | --- | --- |
| case 1 | real value (V) | q = e^V | p = q / Z |
| 1 | 1 | 2.71828 | 0.11841 |
| 2 | -2 | 0.13534 | 0.00590 |
| 3 | 3 | 20.08554 | 0.87490 |
| 4 | -4 | 0.01832 | 0.00080 |

Z = 22.95747, Q1 (0.11841, 0.00590, 0.87490, 0.00080)

|  |  |  |  |
| --- | --- | --- | --- |
| case 2 | real value (V) | q = e^V | p = q / Z |
| 1 | 1 | 2.71828 | 0.24364 |
| 2 | 2 | 7.38906 | 0.66227 |
| 3 | -3 | 0.04979 | 0.00446 |
| 4 | 0 | 1.00000 | 0.08963 |

Z = 11.15712, Q2 (0.24364, 0.66227, 0.00446, 0.08963)

∴

Q1 (0.11841, 0.00590, 0.87490, 0.00080)

Q2 (0.24364, 0.66227, 0.00446, 0.08963)

Q3 (1, 0, 0, 0)

Q4 (0.25, 0.25, 0.25, 0.25)

For sum of distribution is 1,

H (Q4, Q1) = 0.25 \* (log(1/0.11841) + log(1/0.00590) + log(1/0.87490) + log(0.00080))

= 0.25 \* (0.92661 + 2.22914 + 0.05804 + 3.09691)

= 1.57768

H (Q4, Q2) = 0.25 \* (log(1/0.24364) + log(1/0.66227) + log(1/0.00446) + log(1/0.08963))

= 0.25 \* (0.61325 + 0.17896 + 2.35067 + 1.04755)

= 1.04761

H (Q4, Q3) = 0.25 \* (log(1/1) + log(∞) + log(∞) + log(∞))

= 0.25 \* ∞

= ∞

∵H(Q4, Q2) < H(Q4, Q1) < H(Q4, Q3)

∴case 2 is most similar to case 4.

2. Use cross entropy to determine which distribution among 1, 2, 4 is most similar to 3, Show your computations.

Q1 (0.11841, 0.00590, 0.87490, 0.00080)

Q2 (0.24364, 0.66227, 0.00446, 0.08963)

Q3 (1, 0, 0, 0)

Q4 (0.25, 0.25, 0.25, 0.25)

H (Q3, Q1) = 1 \* log(1/0.11841)

= 0.92661

H (Q3, Q2) = 1 \* log(1/0.24364)

= 0.61325

H (Q3, Q4) = 1 \* log(1/0.25)

= 0.60206

∵H(Q3, Q4) < H(Q3, Q2) < H(Q3, Q1)

∴case 4 is most similar to case 3.