Zehra Gunindi

X	x1	x2	x3	x4	x5	x6
	Dsep	learning	is	especially	it's	cool
Inputs	1	0	0	1	0	1
	0	1	1	0	1	0
	1	1	0	0	1	1
	0	1	0	1	0	1
	1	0	1	0	1	0
	0	1	1	1	0	1

W_Q

1	1	0	0	1	0
0	0	0	1	0	0
0	1	0	1	0	1
0	0	1	0	1	1

$$Q = X \cdot W_Q$$

	q1	q2	q3	q4	q5	q6
Query	2	1	2	1	2	1
	0	1	0	1	0	1
	0	3	2	2	1	2
	2	2	2	1	2	2

X	x1	x2	x3	x4	x5	x6
	Dsep	learning	is	especially	it's	cool
Inputs	1	0	0	1	0	1
	0	1	1	0	1	0
	1	1	0	0	1	1
	0	1	0	1	0	1
	1	0	1	0	1	0
	0	1	1	1	0	1

W_K

0	1	1	0	0	0
0	0	0	0	1	0
1	1	0	0	0	-1
0	0	0	0	0	0

$$K = X \cdot W_K$$

	k1	k2	k3	k4	k5	k6
Key	1	2	1	0	2	1
	1	0	1	0	1	0
	1	0	0	0	1	0
	0	0	0	0	0	0

X	x1	x2	x3	x4	x5	x6
	Dsep	learning	is	especially	it's	cool
Inputs	1	0	0	1	0	1
	0	1	1	0	1	0
	1	1	0	0	1	1
	0	1	0	1	0	1
	1	0	1	0	1	0
	0	1	1	1	0	1

W_V

1	0	1	0	0	1
0	1	0	1	1	0
1	-1	0	0	1	0
0	1	1	0	-1	1

$$V = X \cdot W_V$$

	v1	v2	v3	v4	v5	v6
Value	2	2	1	2	1	3
	1	2	2	1	2	1
	2	-1	0	1	0	1
	0	3	1	1	1	2

	q1	q2	q3	q4	q5	q6
Query	2	1	2	1	2	1
	0	1	0	1	0	1
	0	3	2	2	1	2
	2	2	2	1	2	2

MatMul1

K^T

k1	1	1	1	0
k2	2	0	0	0
k3	1	1	0	0
k4	0	0	0	0
k5	2	1	1	0
k6	1	0	0	0

$$q\_to\_k\_similarity\_scores = matmul(Q, K^T)$$

	2	5	4	4	3	4
	4	2	4	2	4	2
	2	2	2	2	2	2
	0	0	0	0	0	0
	4	6	6	5	5	5
	2	1	2	1	2	1

dk = 6

Scale

	0.82	2.04	1.63	1.63	1.22	1.63
	1.63	0.82	1.63	0.82	1.63	0.82
	0.82	0.82	0.82	0.82	0.82	0.82
	0.00	0.00	0.00	0.00	0.00	0.00
	1.63	2.45	2.45	2.04	2.04	2.04
	0.82	0.41	0.82	0.41	0.82	0.41

Softmax

	2.3	7.7	5.1	5.1	3.4	5.1
	5.1	2.3	5.1	2.3	5.1	2.3
	2.3	2.3	2.3	2.3	2.3	2.3
	1.0	1.0	1.0	1.0	1.0	1.0
	5.1	11.6	11.6	7.7	7.7	7.7
	2.3	1.5	2.3	1.5	2.3	1.5

Attention Weight

	0.13	0.29	0.19	0.26	0.16	0.26
	0.28	0.09	0.19	0.11	0.24	0.11
	0.13	0.09	0.08	0.11	0.10	0.11
	0.06	0.04	0.04	0.05	0.05	0.05
	0.28	0.44	0.42	0.39	0.35	0.39
	0.13	0.06	0.08	0.08	0.10	0.08
Sum	1.00	1.00	1.00	1.00	1.00	1.00

MatMul2 (Z = V * A)

	5.0	3.7	4.3	4.0	4.7	4.0
	3.8	4.6	4.2	4.4	4.0	4.4
	2.3	2.3	2.3	2.3	2.3	2.3
	1.0	1.0	1.0	1.0	1.0	1.0
	9.0	7.6	8.3	7.9	8.6	7.9
	1.9	2.1	2.0	2.1	2.0	2.1

Inspired by the Excel made by Prof. Tom Yeh and https://stats.stackexchange.com/q/537197