ML Solution

7.1

The correct activation functions are the ones that can satisfy the range of the output variables.

Heading is from -pi to pi and distance travelled is from 0 to inf.

Now lets look at the ranges of the activation functions in options:

1) ReLU has range 0 to inf.
2) Sigmoid has range 0 to 1.
3) Leaky ReLU has range -inf to inf.
4) Tanh has range -1 to 1.
5) ELU has range -alpha to inf
6) $f(x) = x^2$ has range 0 to inf

Therefore Leaky ReLU and ELU are the correct answers.

7.2 and 7.3

Let us first define notations:

x_{inp} is input vector.
h is output of hidden layer.
s is output of sigmoid layer.
y is output of output layer.
These are the weights and biases provided:

$$\mathbf{W_1} = \begin{bmatrix} w^1_{(0,0)} & w^1_{(0,1)} & w^1_{(0,2)} \\ w^1_{(1,0)} & w^1_{(1,1)} & w^1_{(1,2)} \\ w^1_{(2,0)} & w^1_{(2,1)} & w^1_{(2,2)} \\ w^1_{(3,0)} & w^1_{(3,1)} & w^1_{(3,2)} \end{bmatrix}$$

$$\mathbf{b_1} = \begin{bmatrix} b^1_{(0)} & b^1_{(1)} & b^1_{(2)} \end{bmatrix}$$

$$\mathbf{W_2} = \begin{bmatrix} w^2_{(0,0)} & w^2_{(0,1)} \\ w^2_{(1,0)} & w^2_{(1,1)} \\ w^2_{(2,0)} & w^2_{(2,1)} \end{bmatrix}$$

$$\mathbf{b_2} = \begin{bmatrix} b^2_{(0)} & b^2_{(1)} \end{bmatrix}$$

$$\mathbf{W_1} = \begin{bmatrix} -0.4 & -0.1 & -0.2 \\ 0.2 & 0.1 & 0.1 \\ -0.4 & -0.8 & -0.1 \\ -0.3 & 0.3 & -0.9 \end{bmatrix}$$

$$\mathbf{b_1} = \begin{bmatrix} -0.4 & -0.3 & -0.2 \end{bmatrix}$$

$$\mathbf{W_2} = \begin{bmatrix} 0.2 & -0.3 \\ 0.2 & -0.2 \\ 0.3 & 0.2 \end{bmatrix}$$

$$\mathbf{b_2} = \begin{bmatrix} 0.4 & -0.1 \end{bmatrix}$$

$$MSE = \frac{1}{2N} \sum_{i=1}^{N} (y - y')^2$$

Now, Lets do forward propagation.

$$h = x_{inp} w_1^T + b1$$

$$Sigmoid(x) = \frac{1}{1 + e^{-x}}$$

Therefore,

$$s = Sigmoid(h)$$

$$y = s w_2^T + b2$$

This is the method of doing forward propagation and is common to both 7.2 and 7.3

Now, if we want to do backpropagation, we need to first find the Error or Loss.

In our case, the loss function is MSE and we have only 1 sample. So $N = 1$.

Therefore,

$$Loss = \frac{1}{2} * (y_{target} - y)^2$$

Now, let us find the derivative of Loss with respect to y. [shorthand is ∂y]

$$\partial y = \frac{\partial L}{\partial y} = y - y_{target}$$

Similarly,

$$\frac{\partial y}{\partial s} = W_2$$

$$\frac{\partial Sigmoid(h)}{\partial h} = Sigmoid(h) * (1 - Sigmoid(h))$$

and,

$$\frac{\partial h}{\partial W_1} = x_{inp}$$

Therefore, now that we have all the derivatives:

$$\partial s = \partial L W_2$$

$$\partial h = \frac{\partial Sigmoid(h)}{\partial h} * \partial s$$

$$\partial w_1 = x_{inp} \partial h$$

These are the final answers that you will get after doing the required calculations mentioned above:

Answer to 7.2

Distance travelled for input $[0.3, -0.4, 1.0, 1.2] = 0.558054$

Heading of Distance travelled for input $[0.8, 0.80.8, 0.8] = -0.189464$

Answer to 7.3

$$\partial \mathbf{W_1} = \begin{bmatrix} 0.011629 & 0.009739 & 0.00179 \\ 0.017443 & 0.014608 & 0.002685 \\ 0.005814 & 0.00487 & 0.000895 \\ 0.017443 & 0.014608 & 0.002685 \end{bmatrix}$$