Let's include an **activation function** into the equation and explain how it would affect the calculation.

Updated Model with Activation Function:

Suppose we introduce an activation function $f(\cdot)$ after the weighted sum y, before generating the prediction $\hat{y}$. The equations change as follows:

1. **Weighted Sum**:

$$y = w_1 \cdot \text{area} + w_2 \cdot \text{bedrooms} + b$$

2. **Activation Function**:

$$\hat{y} = f(y)$$

   Here, $f(y)$ could be any activation function like:

   - **ReLU**: $f(y) = \max(0, y)$
   - **Sigmoid**: $f(y) = \frac{1}{1+e^{-y}}$
   - **Tanh**: $f(y) = \frac{e^{y}-e^{-y}}{e^{y}+e^{-y}}$

3. **Error**:

$$e = y_{\text{true}} - \hat{y}$$

4. **Loss**:

$$\text{Loss} = e^2$$

Backpropagation with Activation Function:

To compute the gradient with respect to w_1 (i.e., $\frac{\partial L}{\partial w_1}$), we use the chain rule:

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial e} \cdot \frac{\partial e}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial y} \cdot \frac{\partial y}{\partial w_1}$$

Step-by-Step:

1. **Gradient of Loss w.r.t Error**:

$$\frac{\partial L}{\partial e} = 2e$$

2. **Gradient of Error w.r.t Prediction**:

$$\frac{\partial e}{\partial \hat{y}} = -1$$

3. **Gradient of Prediction w.r.t Weighted Sum** (depends on activation function):

$$\frac{\partial \hat{y}}{\partial y} = f'(y)$$

   For different activation functions:

   - **ReLU**: $f'(y) = 1$ if $y > 0$, else 0
   - **Sigmoid**: $f'(y) = f(y) \cdot (1 - f(y))$
   - **Tanh**: $f'(y) = 1 - f(y)^2$

4. **Gradient of Weighted Sum w.r.t Weight**:

$$\frac{\partial y}{\partial w_1} = \text{area}$$

5. **Final Gradient**:

$$\frac{\partial L}{\partial w_1} = 2e \cdot (-1) \cdot f'(y) \cdot \text{area}$$

Summary Equation:

$$\frac{\partial L}{\partial w_1} = -2e \cdot f'(y) \cdot \text{area}$$