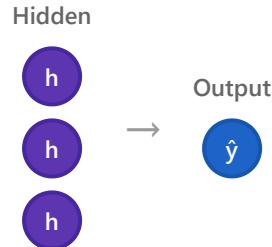


Output Layer Design: Regression vs Classification

Regression



Linear / No Activation

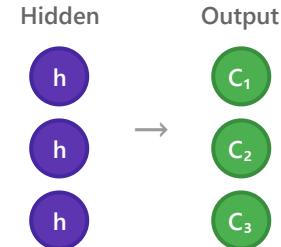
Output Units: 1 (or n)

Activation: None (linear)

Loss Function: MSE, MAE

Output Range: $(-\infty, +\infty)$

Classification



Sigmoid / Softmax

Output Units: K classes

Activation: Softmax

Loss Function: Cross-Entropy

Output Range: [0, 1] (probs)

Regression Details

Classification Details

- **Task:** Predict continuous values
- **Examples:** House price, temperature, stock price
- **Loss:** $MSE = (1/n)\sum(\hat{y} - y)^2$
- **No constraint** on output values

- **Task:** Predict discrete classes
- **Examples:** Image labels, sentiment, diagnosis
- **Loss:** $CE = -\sum y \log(\hat{y})$
- **Probabilities** sum to 1.0

Regression Example

Input: [bedroom=3, sqft=2000]

Output: \$450,000 (continuous)

Classification Example

Input: [image pixels]

Output: [cat: 0.8, dog: 0.15, bird: 0.05]