Output representations

Inputs and outputs of networks

- Input:

  - Tensor $\mathbf{x}$

- Output:

  - Tensor $\mathbf{o}$
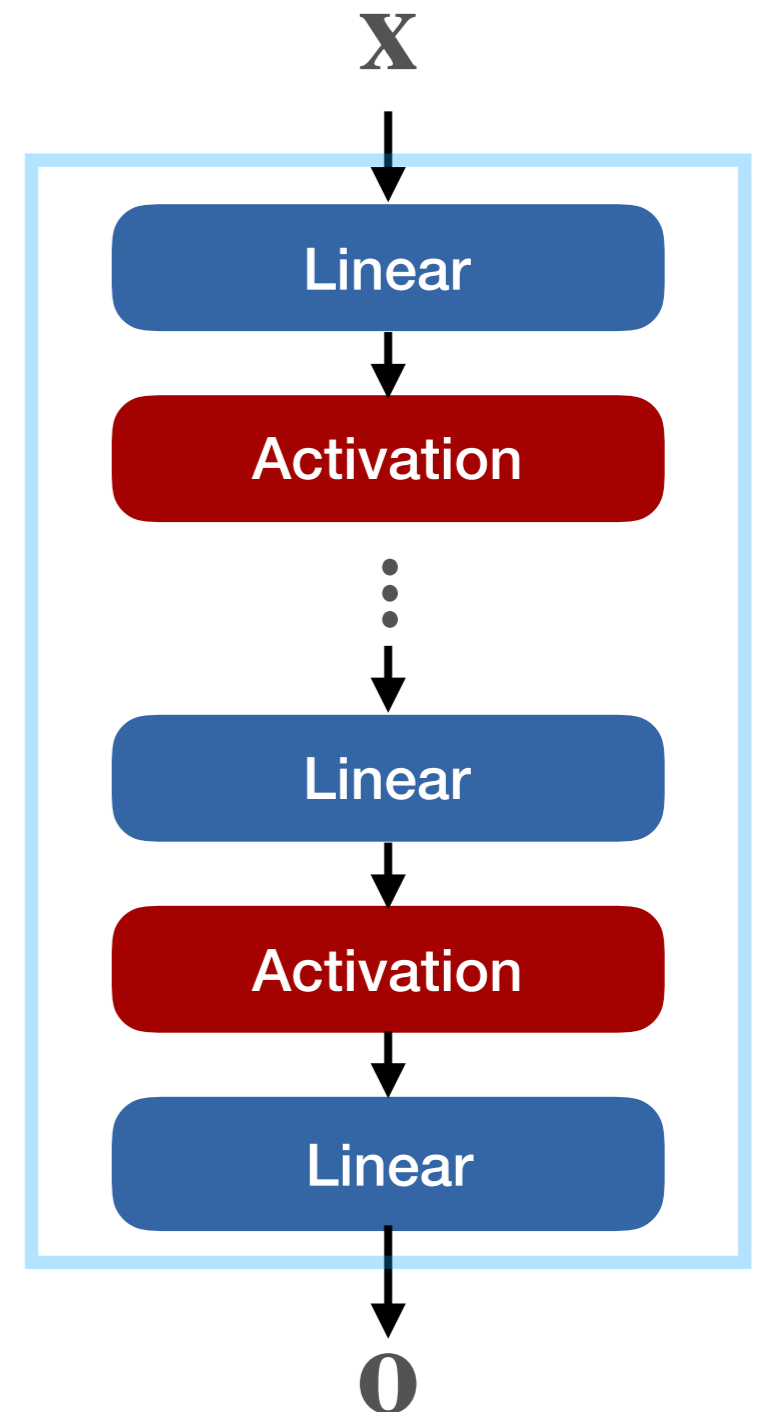
Regression

- vanilla tensor $\hat{\mathbf{y}} = \mathbf{0}$

Positive regression

- Option 1: ReLU

  - $\hat{\mathbf{y}} = \max(\mathbf{o}, 0)$

- Option 2: Soft ReLU

  - $\hat{\mathbf{y}} = \log(1 + e^{\mathbf{o}})$

Binary Classification

- Option 1: Thresholding

  - $\hat{\mathbf{y}} = \mathbf{o} > 0$

- Option 2: Logistic Regression

  - $p(1) = \sigma(\mathbf{o})$

General Classification

- Output more values, one per class

- Option 1: argmax

  - $\hat{y} = \text{argmax}_i \mathbf{o}_i$

- Option 2: softmax

  - $p(y) = \text{softmax}(\mathbf{o})_y$

Output representations in practice

- Do not add into model

- Always output raw values