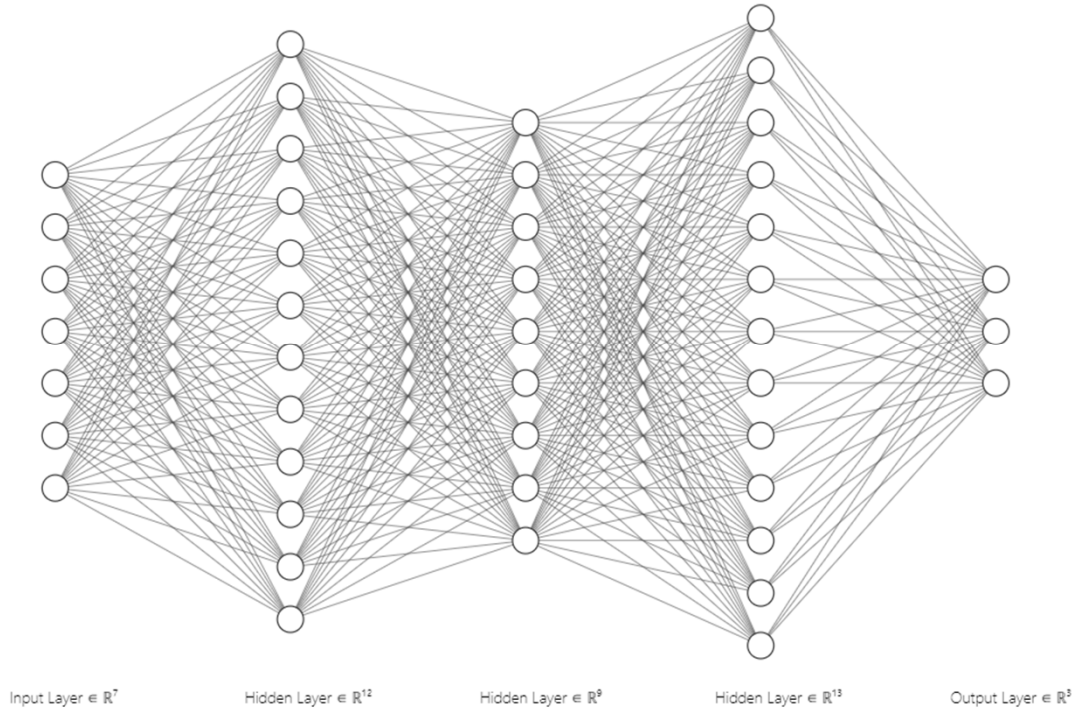
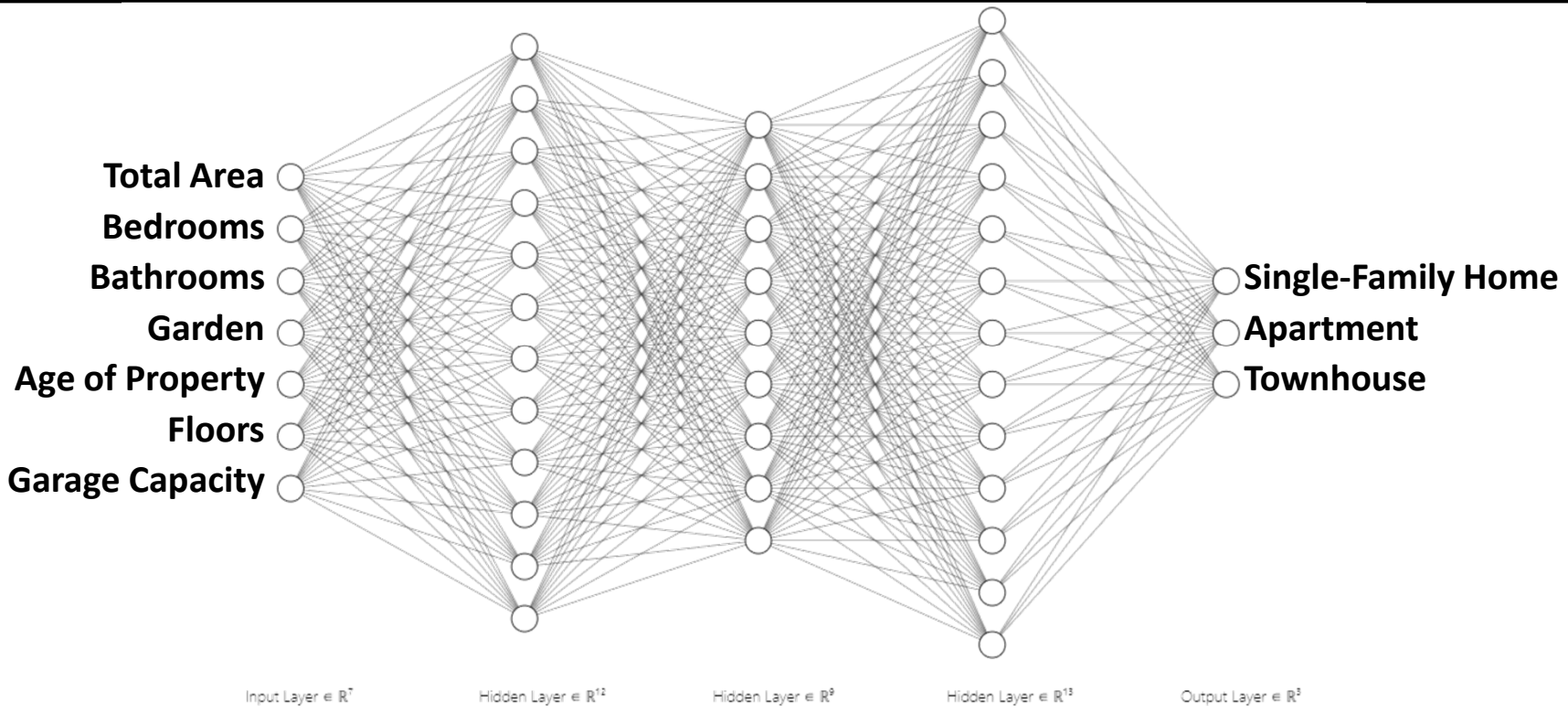
Main Components



What is The Objective of an ANN?

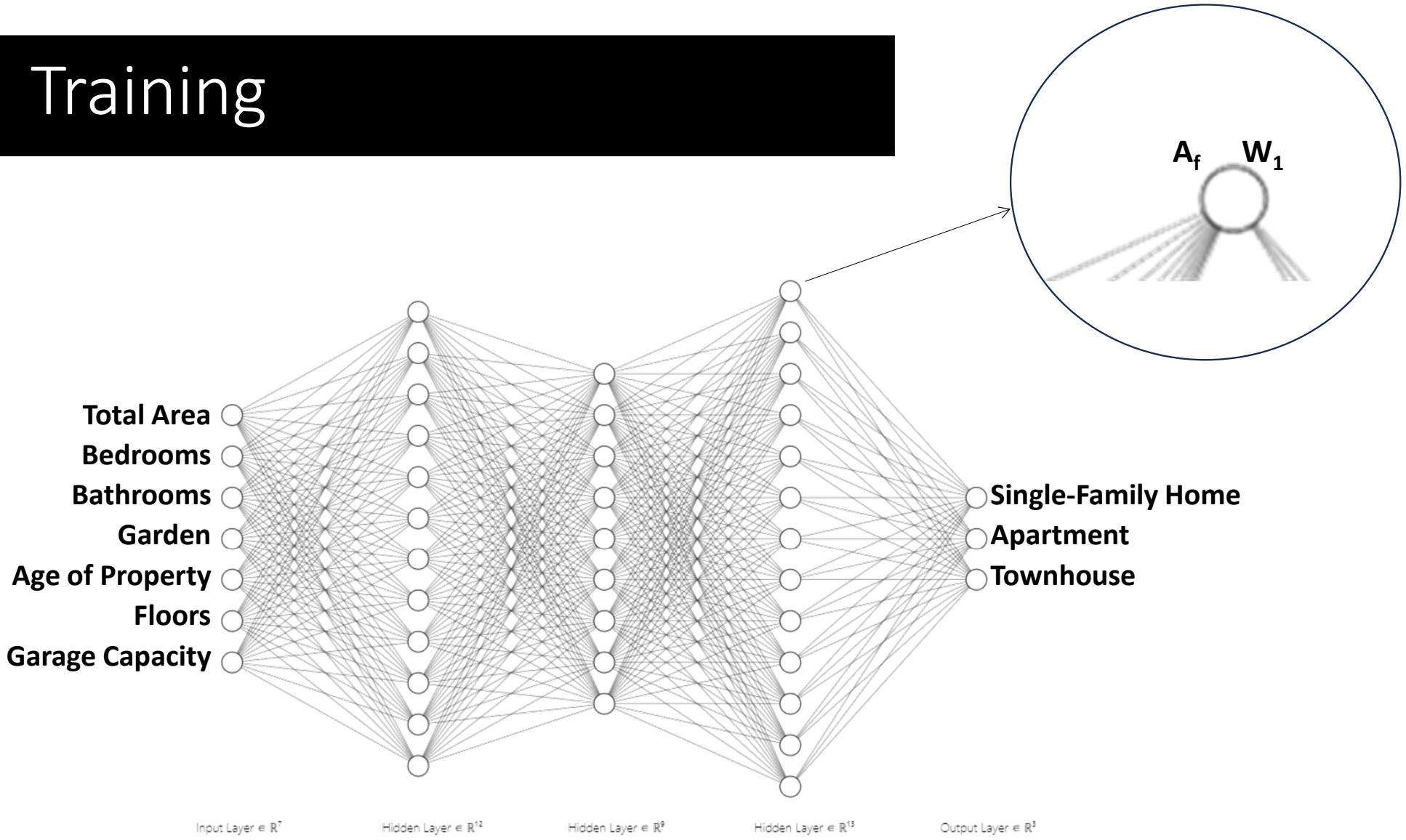


Given an input, produce the desired output.

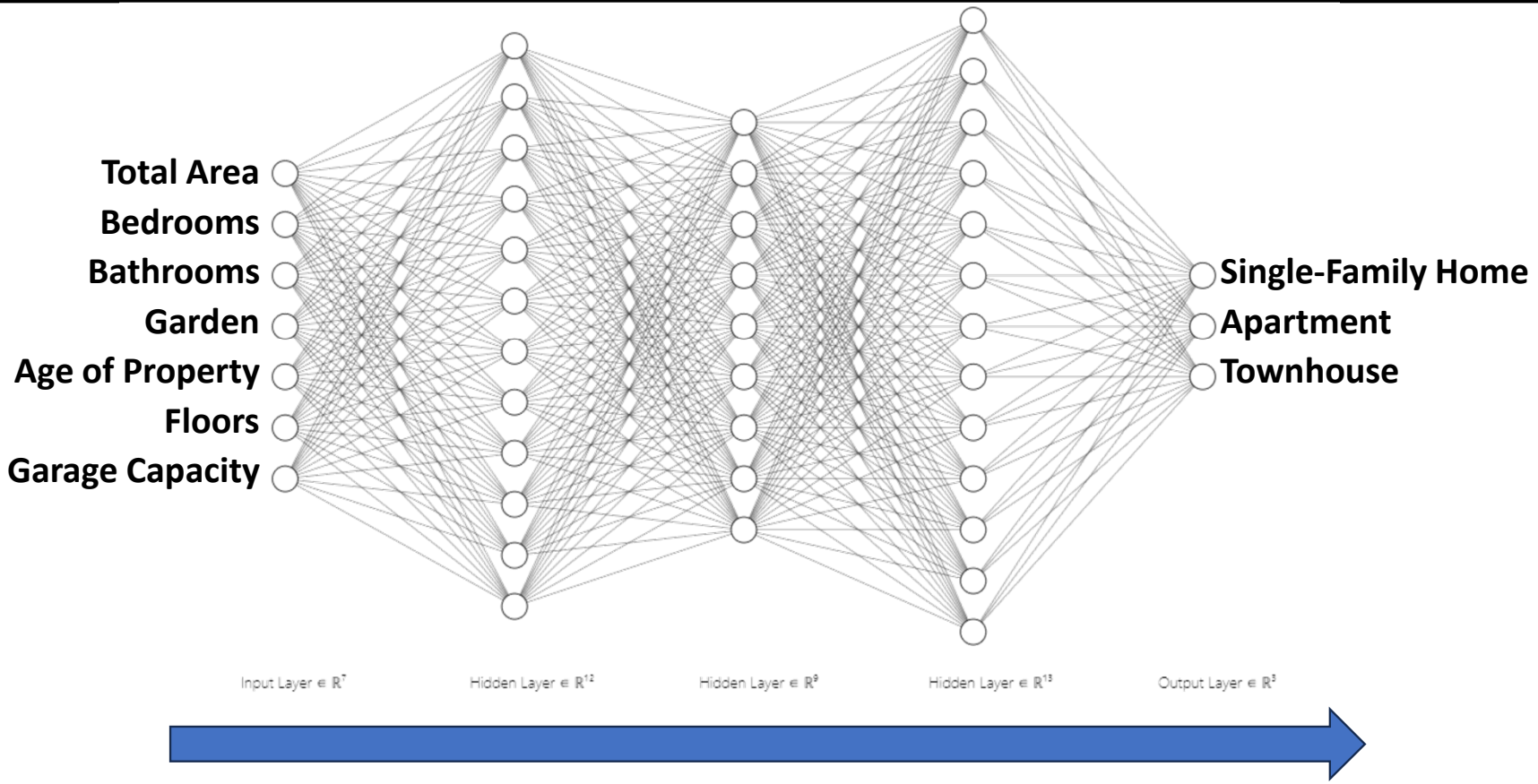
We Need to Train the Network!

Total Area (m^2)	Bedrooms	Bathrooms	Garden	Age (years)	Floors	Garage Capacity	Class
200	3	2	Yes	10	1	2	Single-Family Home
90	2	1	No	5	10	1	Apartment
150	3	3	No	7	3	1	Townhouse
...	...	...	...	...	...	...	...

Training

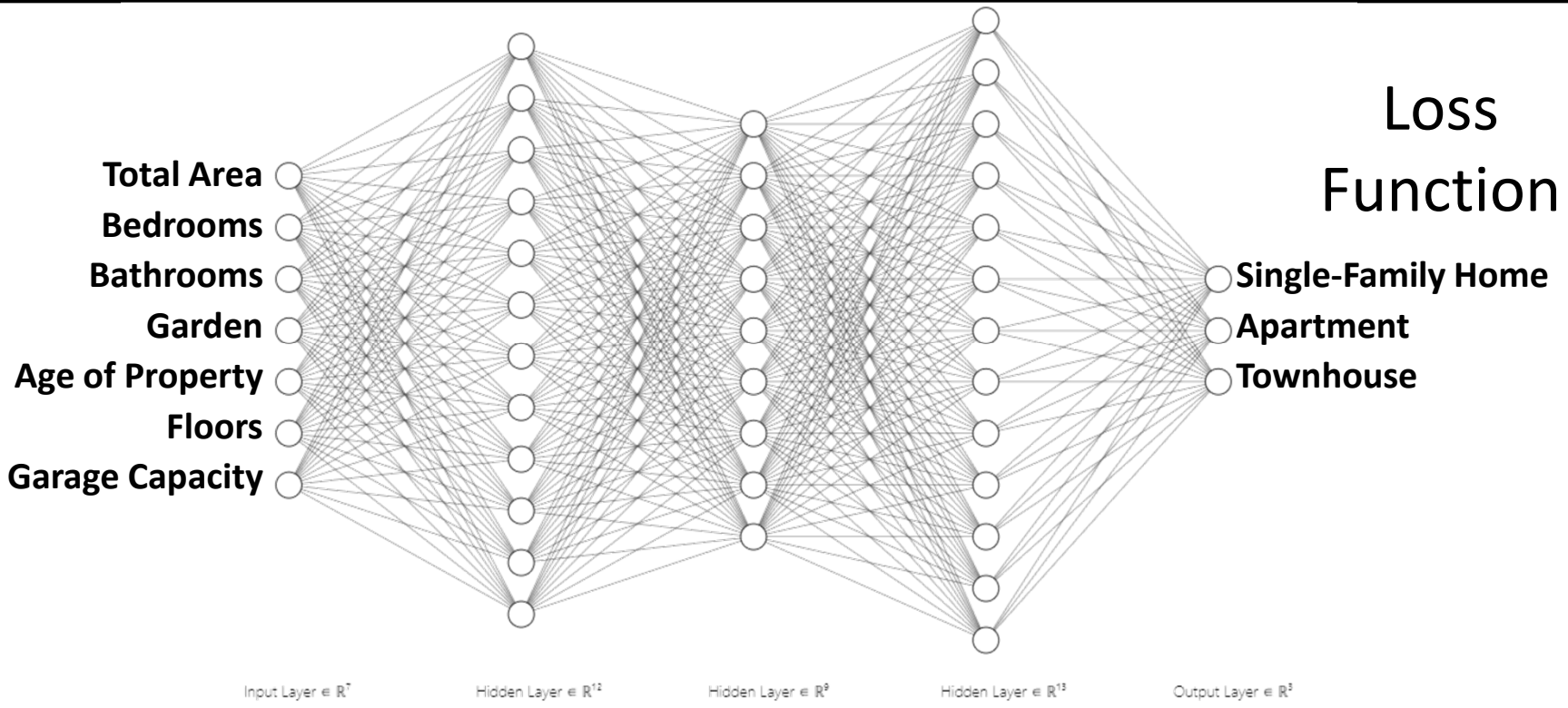

A_f
W_1
Total Area
Bedrooms
Bathrooms
Garden
Age of Property
Floors
Garage Capacity
Single-Family Home
Apartment
Townhouse
Input Layer ∈ $\mathbb{R}^7$
Hidden Layer ∈ $\mathbb{R}^{12}$
Hidden Layer ∈ $\mathbb{R}^9$
Hidden Layer ∈ $\mathbb{R}^{13}$
Output Layer ∈ $\mathbb{R}^3$

What is The Objective of an ANN?



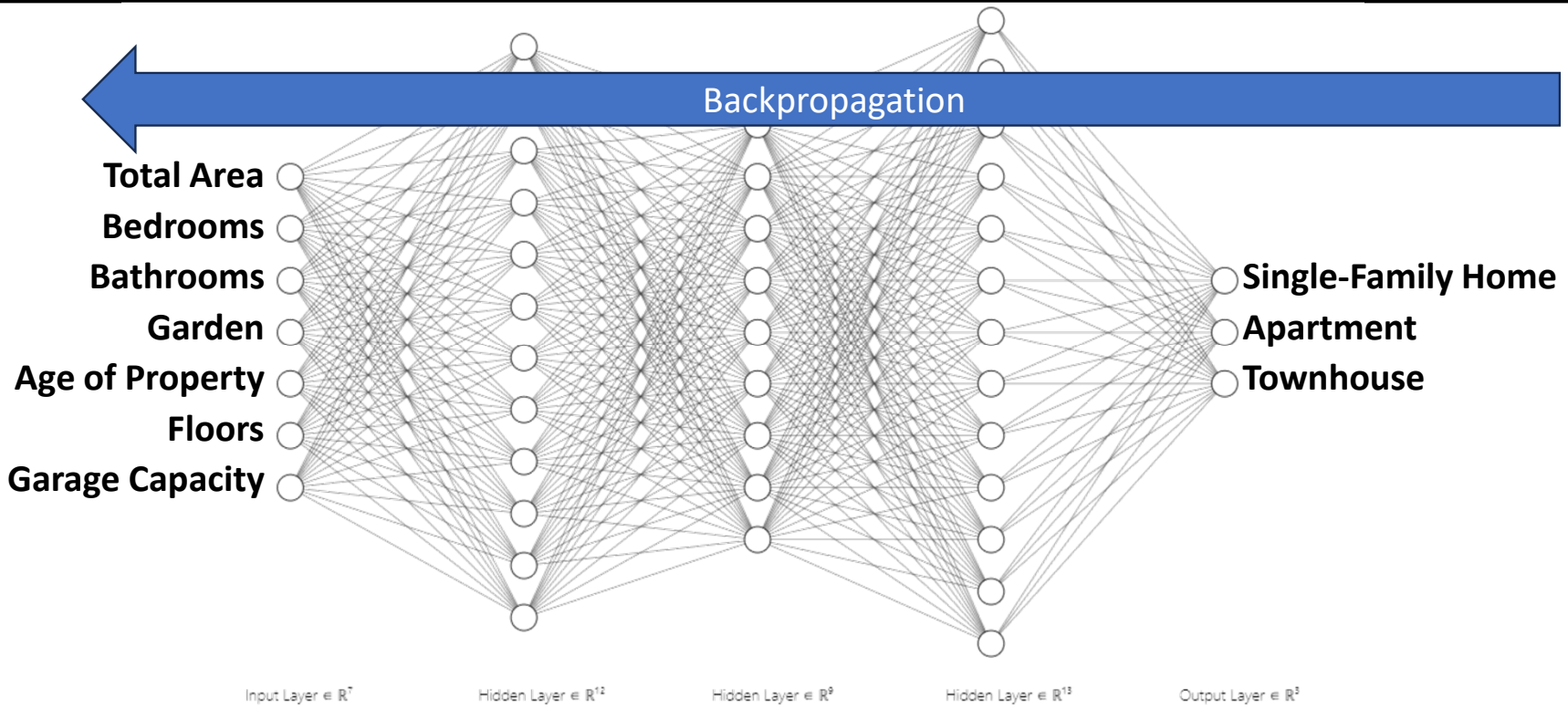
Data Flows Through the Network

What is The Objective of an ANN?



Loss Function

What is The Objective of an ANN?



The Process is Repeated Many Times…

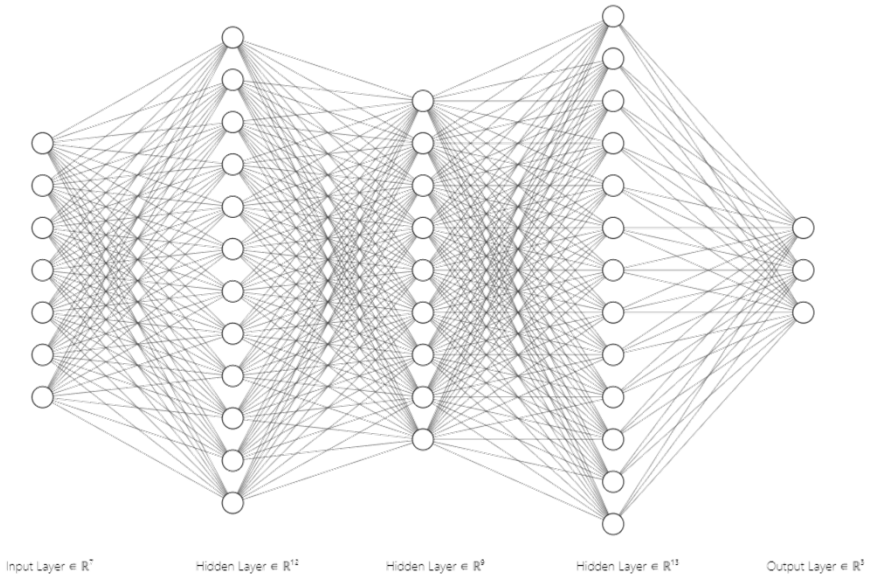
Weights are adjusted

The output tends to improve

The loss decreases

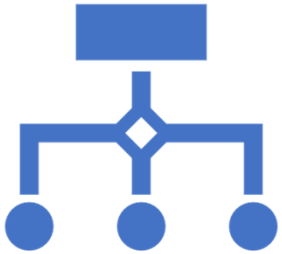
ANN vs Deep Learning

- Deep Learning is an ANN with 3 or more hidden layers



Concepts

Epochs: The number of times all records will pass through the network.



Batch Size: After how many records pass through the network will the weights be adjusted?

Gradient Descent

Determines the direction and magnitude of the weight adjustments.

Gradient Descent aids in optimizing the training process.

Adam is an advanced implementation of Gradient Descent."

Learning Rate

The learning rate determines the step size at each iteration while moving towards a minimum in the loss function.

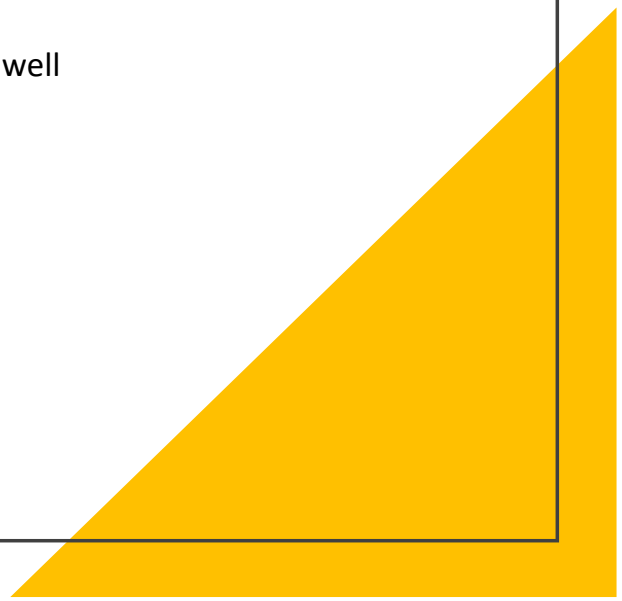
It influences how quickly or slowly a neural network updates its weights during training.

A smaller learning rate might converge slowly, while a larger one might overshoot the optimal solution.

Overfitting

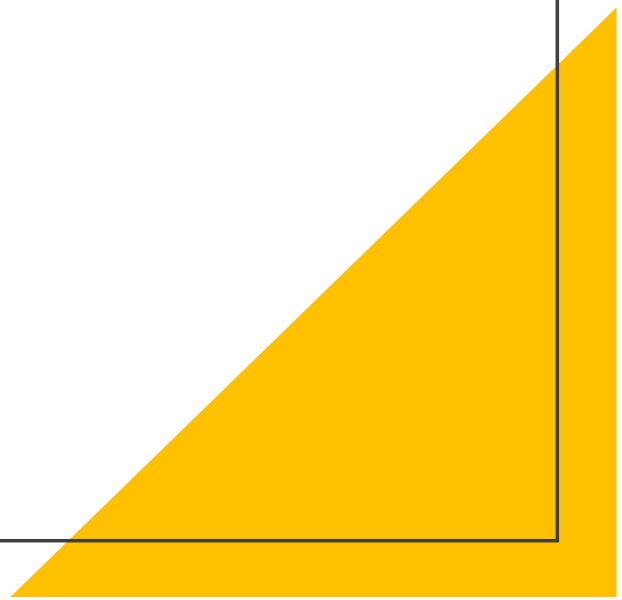
- Overfitting occurs when a model learns the training data too closely, including its noise and outliers, causing it to perform poorly on unseen or new data.
- It essentially means the model is too complex and captures patterns that don't generalize well beyond the training dataset.

- Regularization:
  - Early Stopping
  - Data Augmentation
  - L1 and L2
  - Dropout



Activation and Loss Functions

- There are many different activation functions:
  - Threshold
  - Sigmoid
  - ReLU
  - Tanh
- There are many different loss functions
  - MSE
    RMSE
  - Cross-Entropy



HyperParameter

- A hyperparameter is a parameter of the ANN that you set before training or using a pre-trained model. It influences the output as well as the time the ANN takes to run.