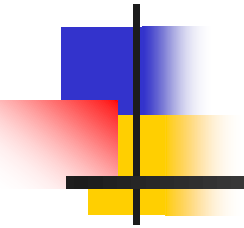
Business Intelligence and Decision Support Systems (9th Ed., Prentice Hall)



Chapter 6:

Artificial Neural Networks for Data Mining

Learning Objectives

- Understand the concept and definitions of artificial neural networks (ANN)
- Know the similarities and differences between biological and artificial neural networks
- Learn the different types of neural network architectures
- Learn the advantages and limitations of ANN
- Understand how backpropagation learning works in feedforward neural networks

Learning Objectives

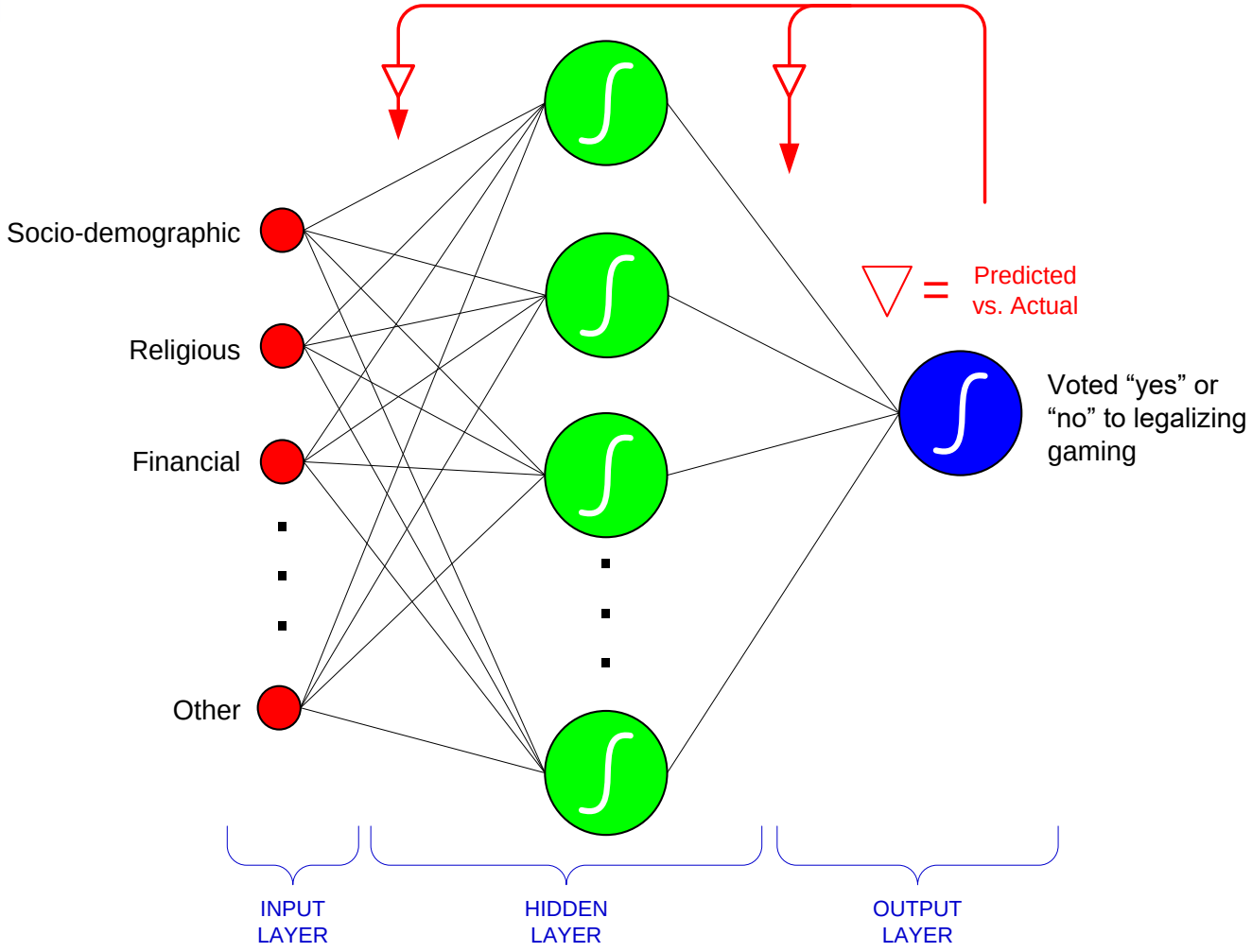
- Understand the step-by-step process of how to use neural networks
- Appreciate the wide variety of applications of neural networks; solving problem types of
  - Classification
  - Regression
  - Clustering
  - Association
  - Optimization

Opening Vignette:

"Predicting Gambling Referenda with Neural Networks"

- Decision situation
- Proposed solution
- Results
- Answer and discuss the case questions

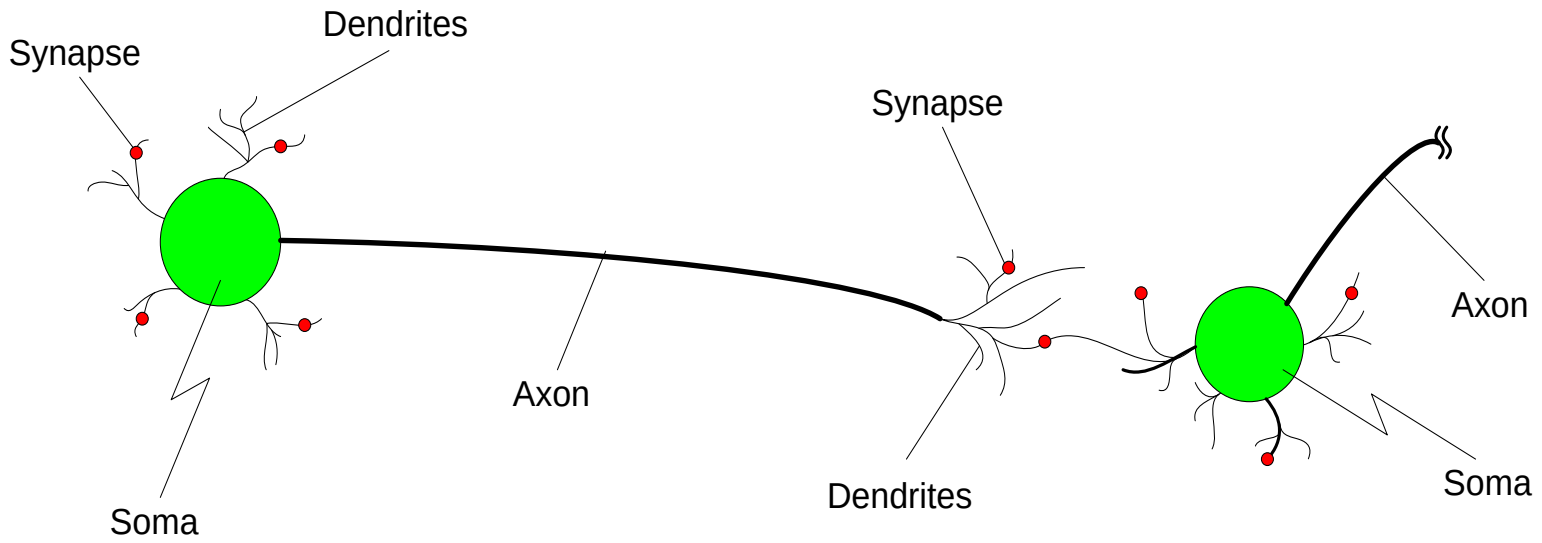
Opening Vignette: Predicting Gambling Referenda...



Neural Network Concepts

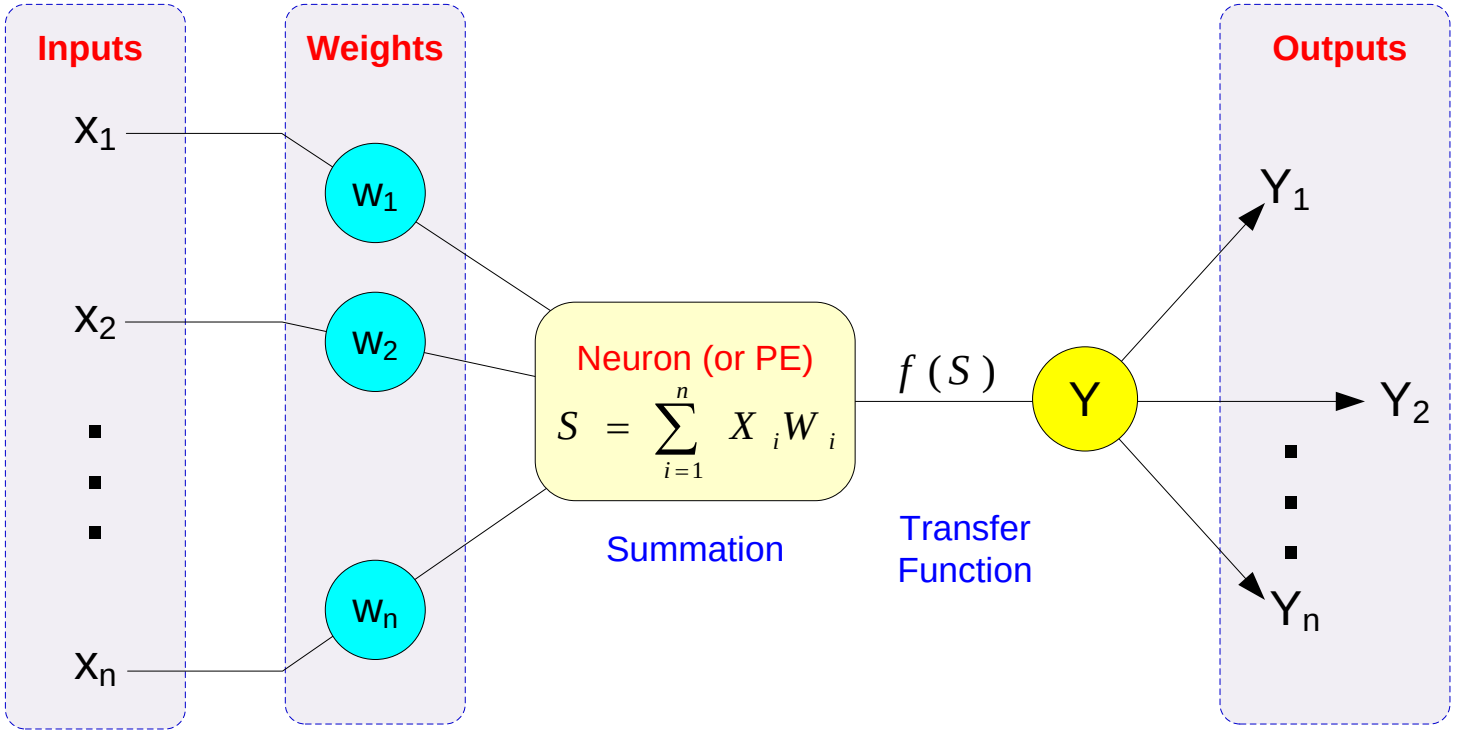
- Neural networks (NN): a brain metaphor for information processing
- Neural computing
- Artificial neural network (ANN)
- Many uses for ANN for
  - pattern recognition, forecasting, prediction, and classification
- Many application areas
  - finance, marketing, manufacturing, operations, information systems, and so on

Biological Neural Networks



- Two interconnected brain cells (neurons)

Processing Information in ANN



- A single neuron (processing element – PE) with inputs and outputs

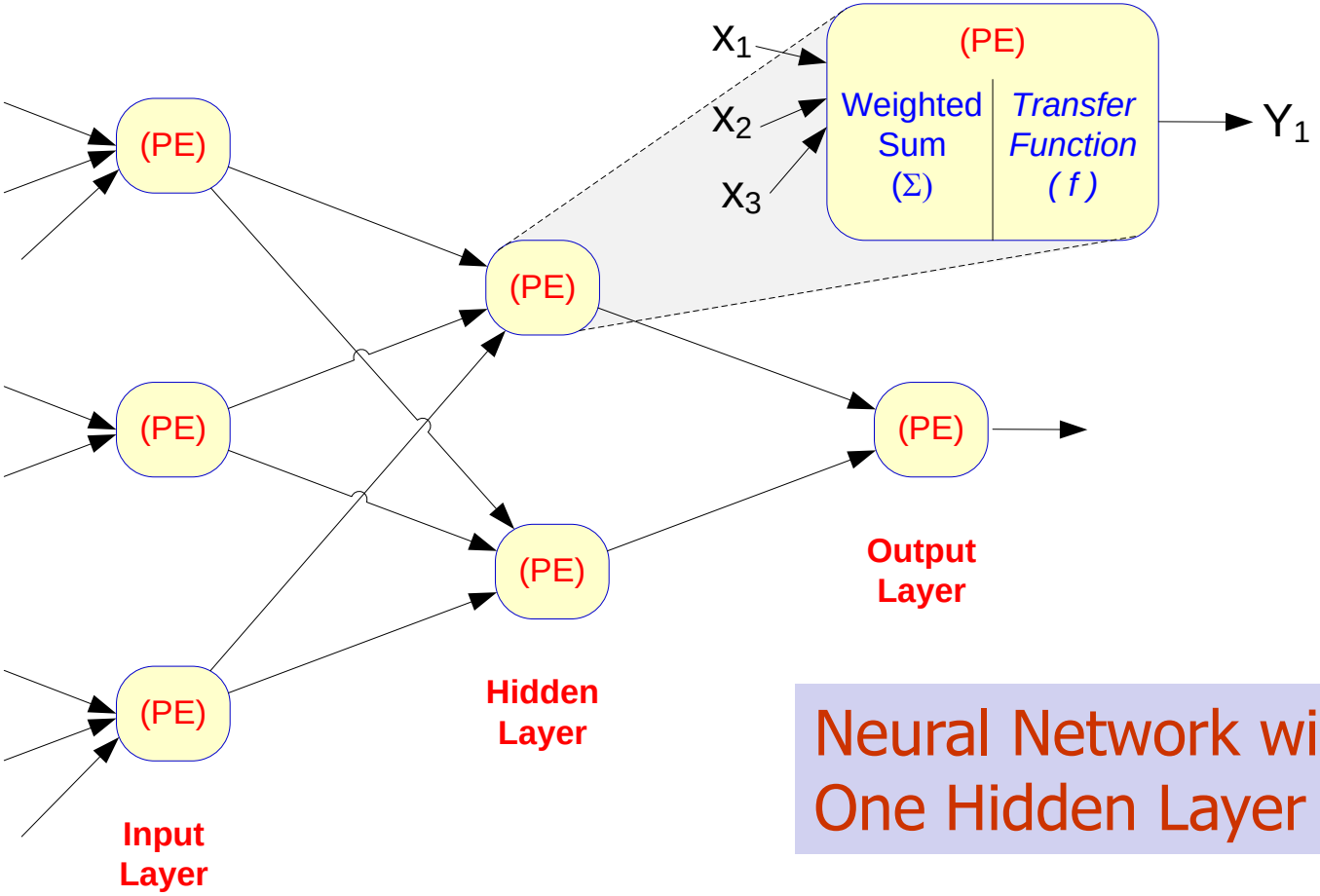
Biology Analogy

Biological	versus	Artificial	NNs
Soma		Node	
Dendrites		Input	
Axon		Output	
Synapse		Weight	
Slow		Fast	
Many neurons (10^9)		Few neurons (~100s)	

Elements of ANN

- Processing element (PE)
- Network architecture
  - Hidden layers
  - Parallel processing
- Network information processing
  - Inputs
  - Outputs
  - Connection weights
  - Summation function

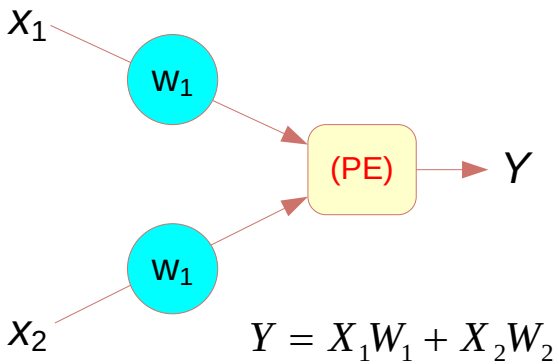
Elements of ANN



Neural Network with One Hidden Layer

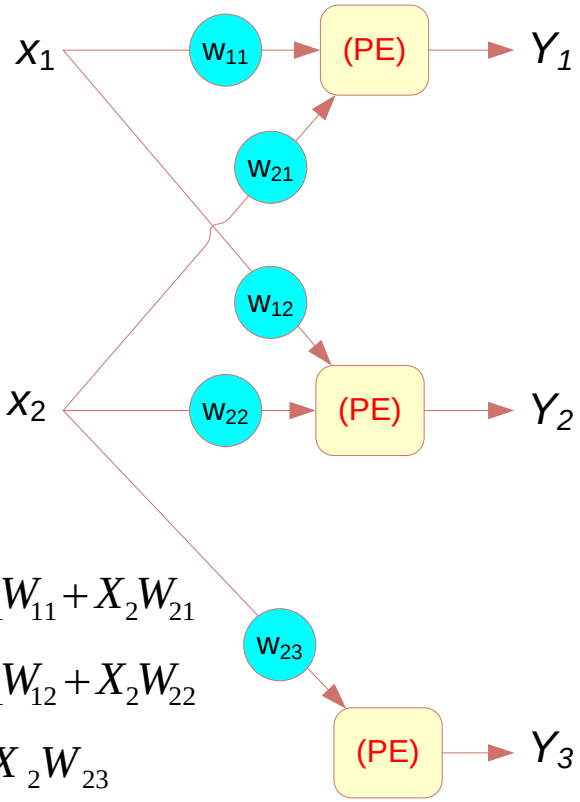
Elements of ANN

(a) Single neuron



PE: Processing Element (or neuron)

(b) Multiple neurons

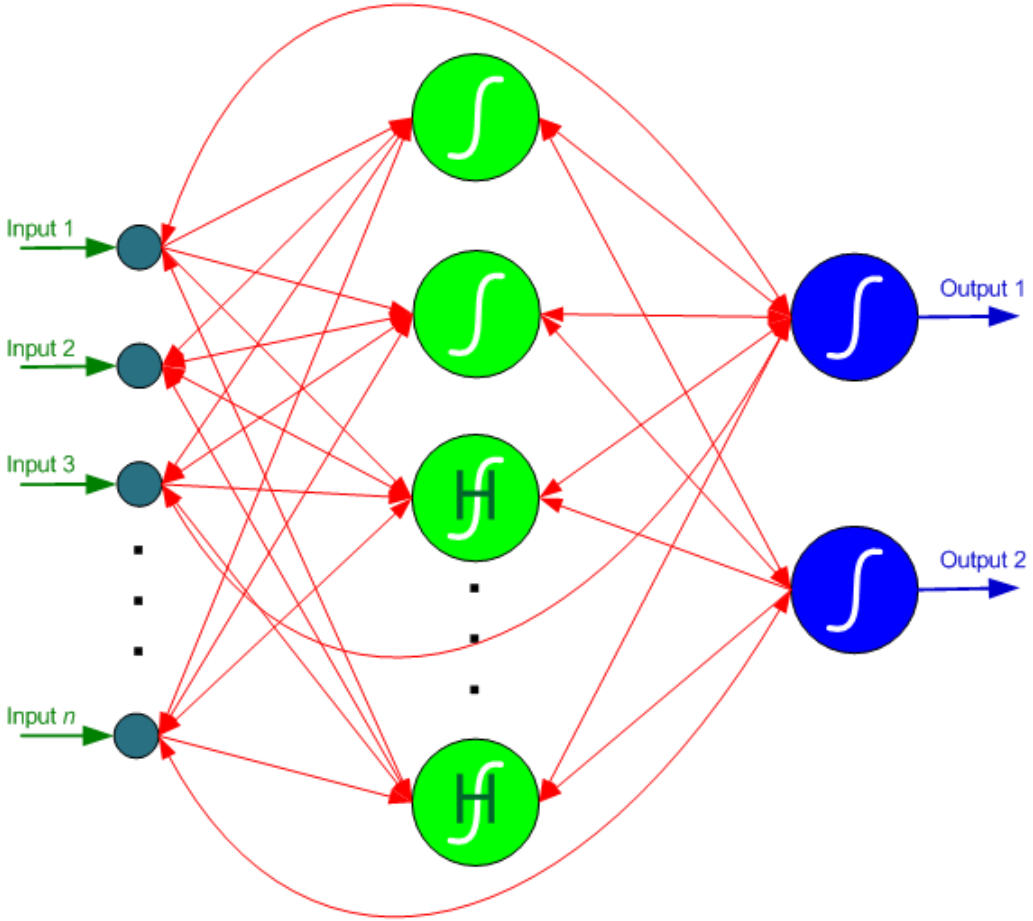


Summation Function for a Single Neuron (a) and Several Neurons (b)

Neural Network Architectures

- Several ANN architectures exist
  - Feedforward
  - Recurrent
  - Associative memory
  - Probabilistic
  - Self-organizing feature maps
  - Hopfield networks
  - … many more …

Neural Network Architectures
Recurrent Neural Networks



*H: indicates a "hidden" neuron without a target output

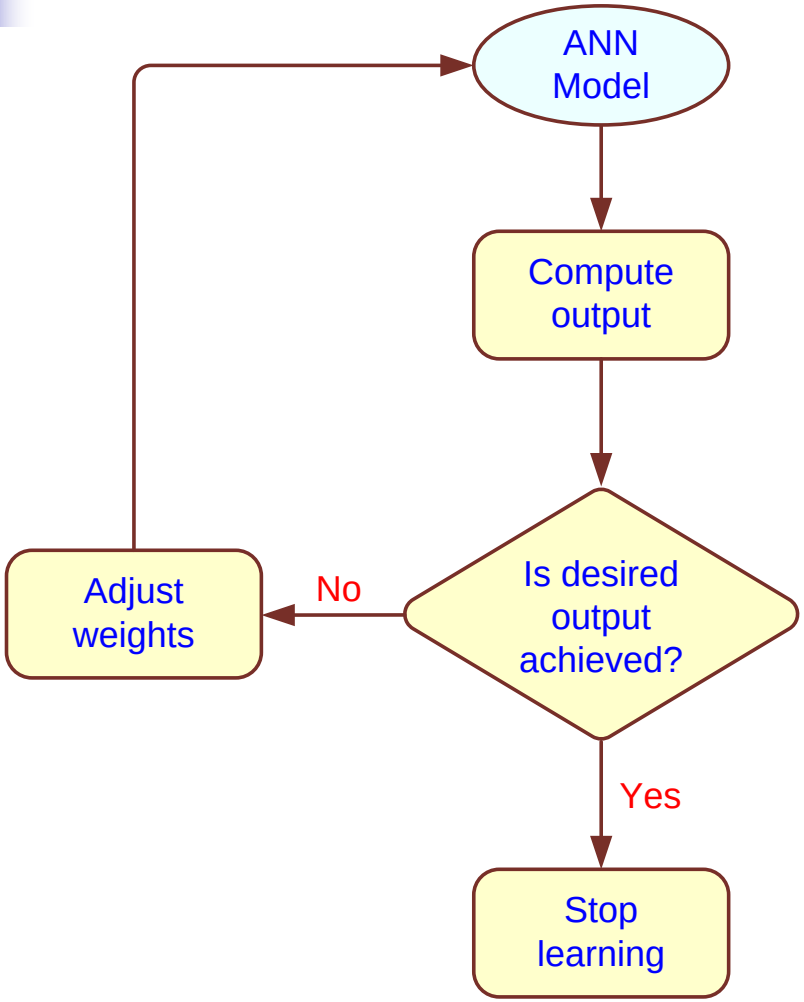
Neural Network Architectures

- Architecture of a neural network is driven by the task it is intended to address
  - Classification, regression, clustering, general optimization, association, ….
- Most popular architecture: Feedforward, multi-layered perceptron with backpropagation learning algorithm
  - Used for both classification and regression type problems

Learning in ANN

- A process by which a neural network learns the underlying relationship between input and outputs, or just among the inputs
- Supervised learning
  - For prediction type problems
  - E.g., backpropagation
- Unsupervised learning
  - For clustering type problems
  - Self-organizing
  - E.g., adaptive resonance theory

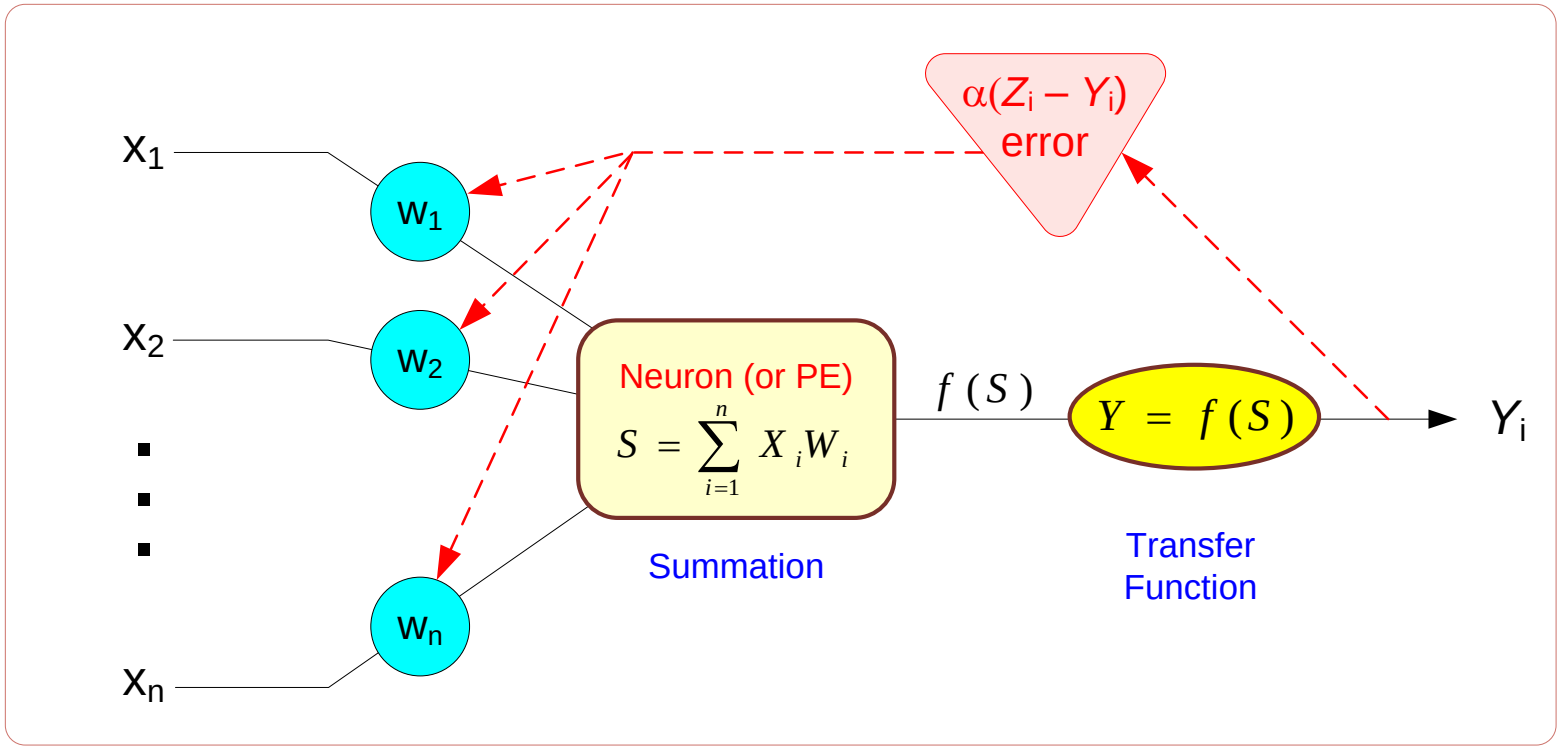
A Supervised Learning Process



Three-step process:

1. Compute temporary outputs
2. Compare outputs with desired targets
3. Adjust the weights and repeat the process

Backpropagation Learning

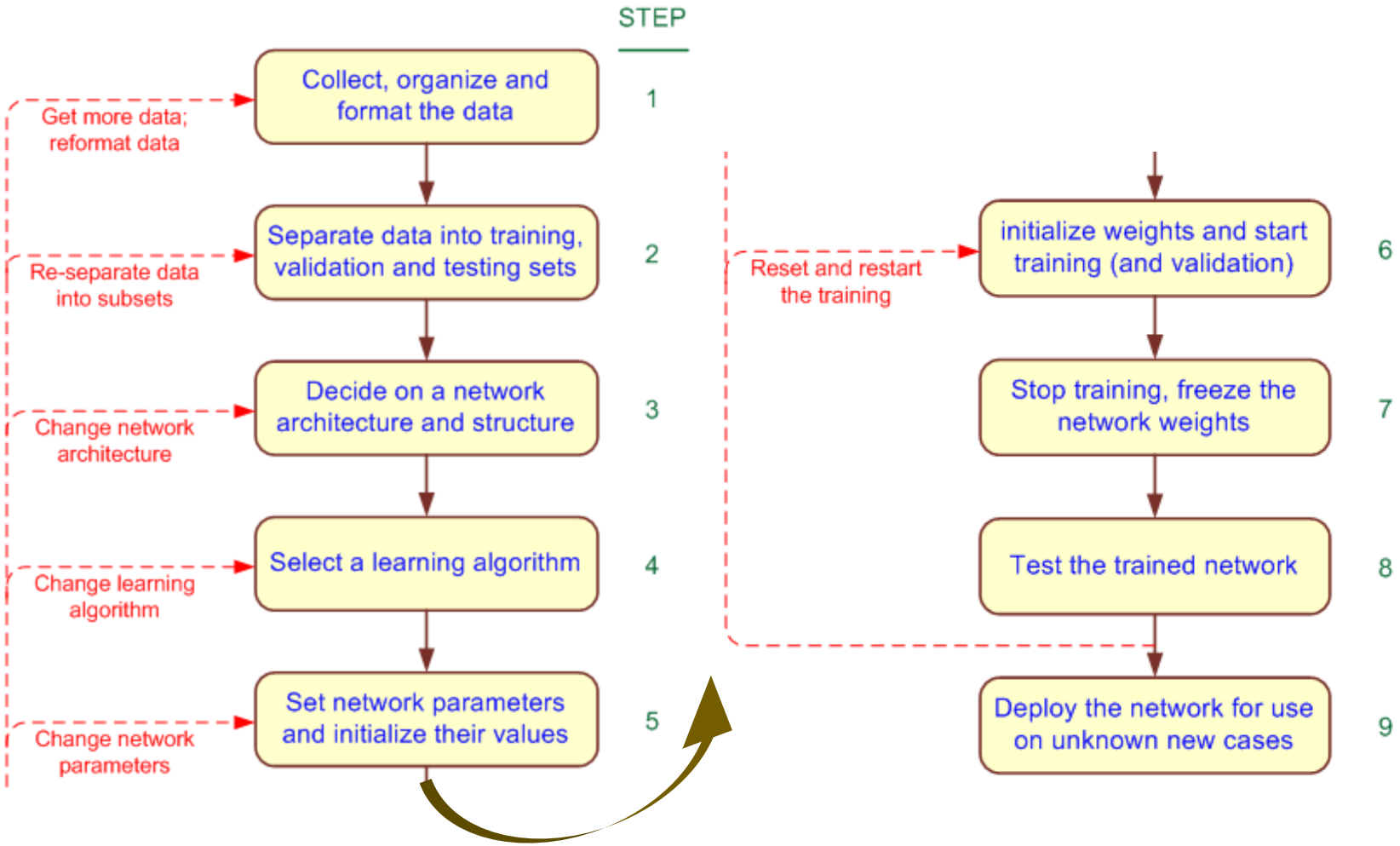


- Backpropagation of Error for a Single Neuron

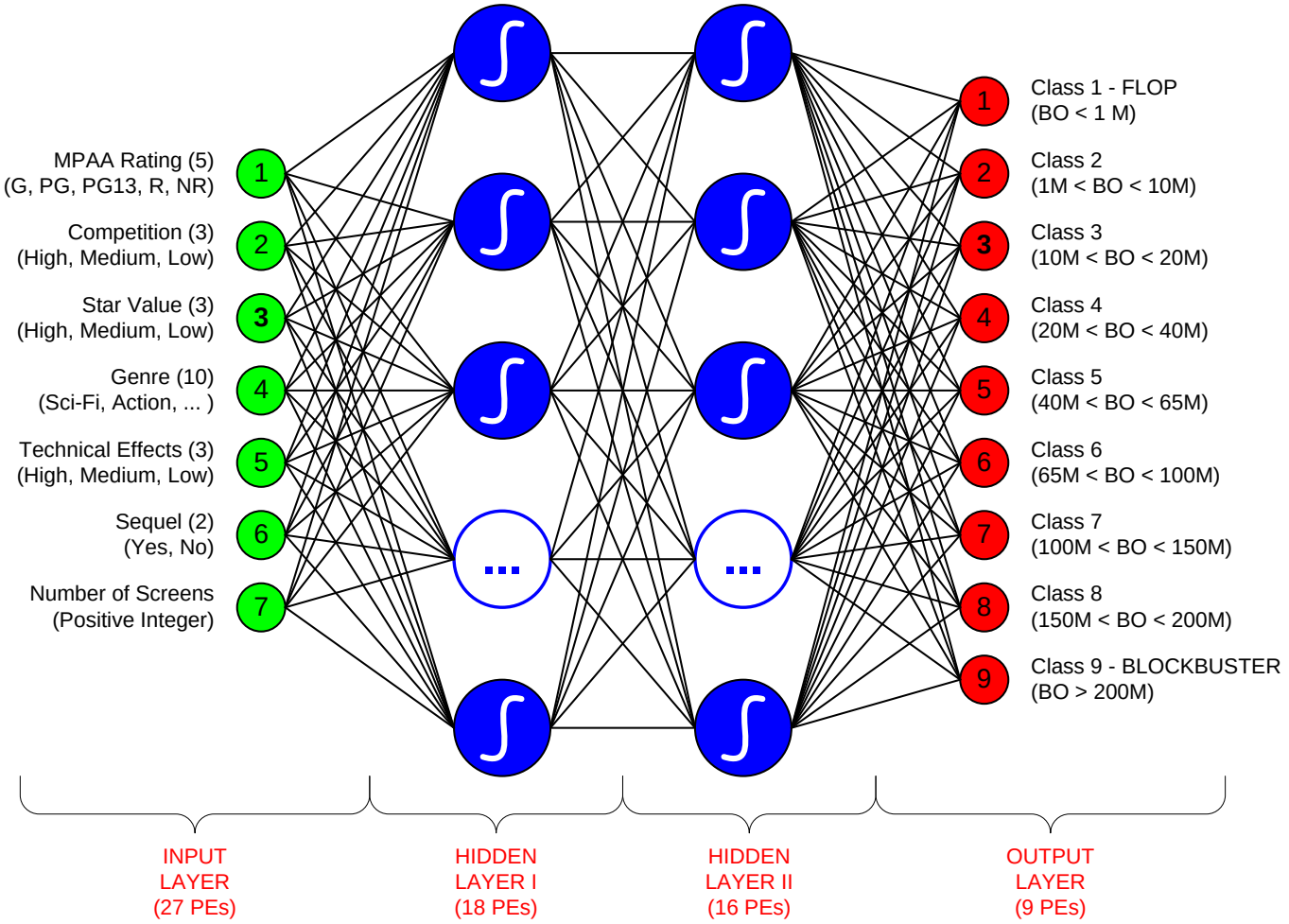
Backpropagation Learning

- The learning algorithm procedure:
  1. Initialize weights with random values and set other network parameters
  2. Read in the inputs and the desired outputs
  3. Compute the actual output (by working forward through the layers)
  4. Compute the error (difference between the actual and desired output)
  5. Change the weights by working backward through the hidden layers
  6. Repeat steps 2-5 until weights stabilize

Development Process of an ANN



An MLP ANN Structure for the Box-Office Prediction Problem



Testing a Trained ANN Model

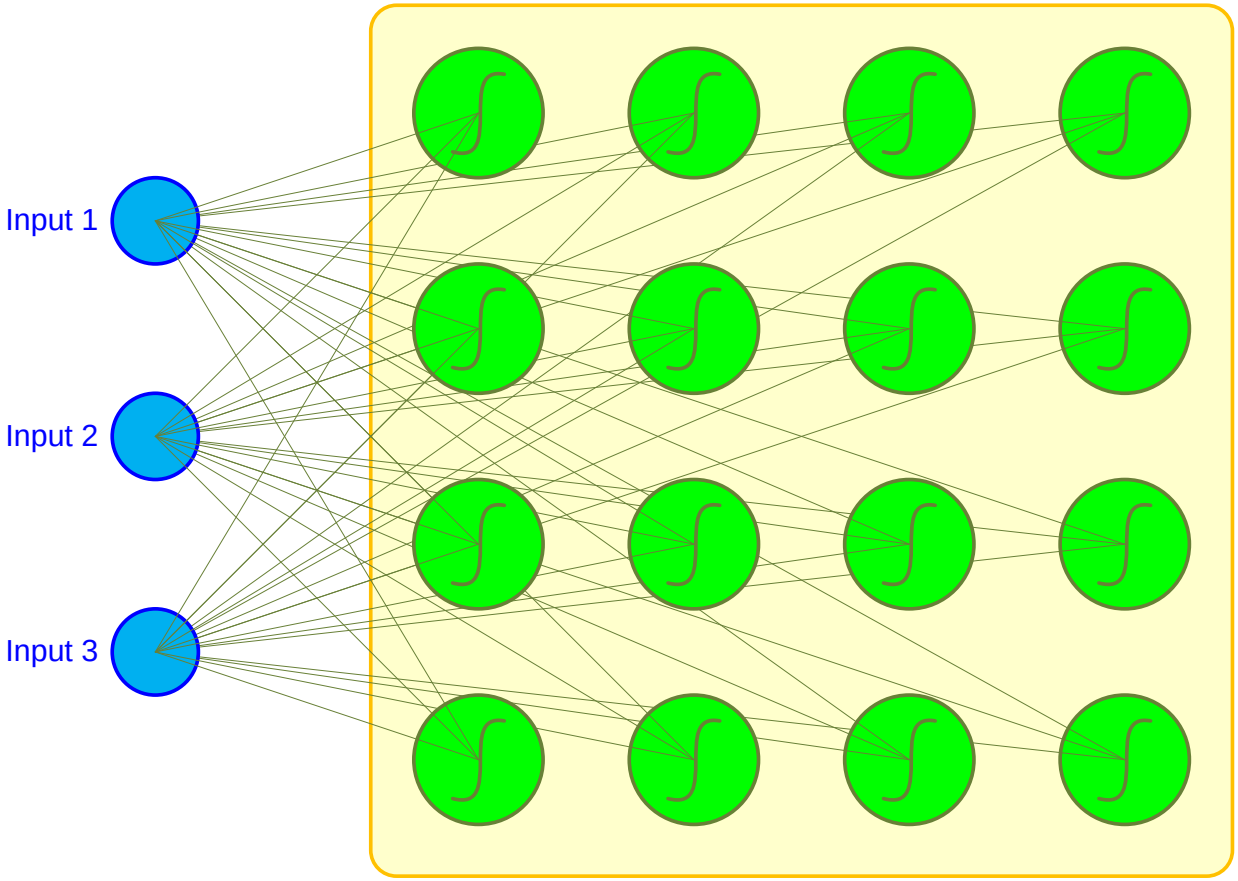
- Data is split into three parts
  - Training (~60%)
  - Validation (~20%)
  - Testing (~20%)

- *k*-fold cross validation
  - Less bias
  - Time consuming

Sensitivity Analysis on ANN Models

- A common criticism for ANN: The lack of expandability
- The black-box syndrome!
- Answer: sensitivity analysis
  - Conducted on a trained ANN
  - The inputs are perturbed while the relative change on the output is measured/recorded
  - Results illustrates the relative importance of input variables

Other Popular ANN Paradigms Self Organizing Maps (SOM)



- First introduced by the Finnish Professor Teuvo Kohonen
- Applies to clustering type problems

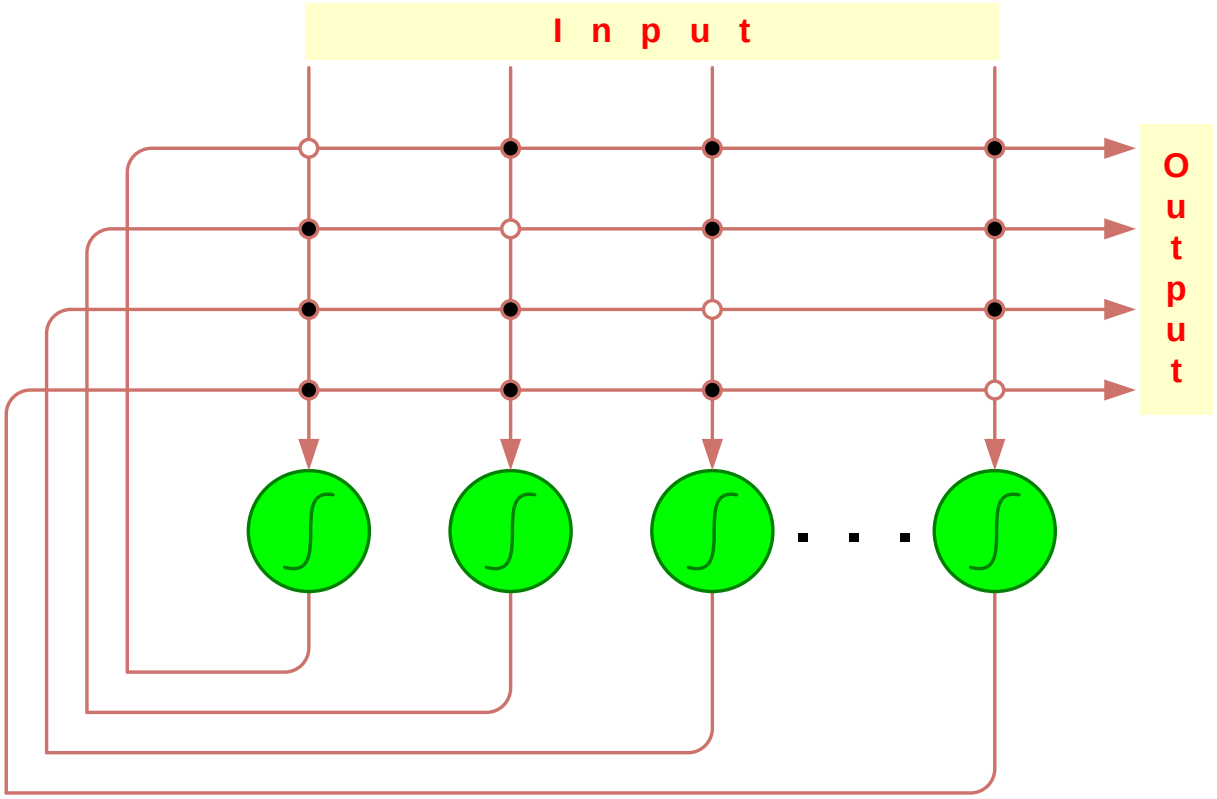
Other Popular ANN Paradigms Self Organizing Maps (SOM)

- SOM Algorithm –
  1. Initialize each node's weights
  2. Present a randomly selected input vector to the lattice
  3. Determine most resembling (winning) node
  4. Determine the neighboring nodes
  5. Adjusted the winning and neighboring nodes (make them more like the input vector)
  6. Repeat steps 2-5 for until a stopping criteria is reached

Other Popular ANN Paradigms Self Organizing Maps (SOM)

- Applications of SOM
  - Customer segmentation
  - Bibliographic classification
  - Image-browsing systems
  - Medical diagnosis
  - Interpretation of seismic activity
  - Speech recognition
  - Data compression
  - Environmental modeling, many more ...

Other Popular ANN Paradigms
Hopfield Networks



- First introduced by John Hopfield
- Highly interconnected neurons
- Applies to solving complex computational problems (e.g., optimization problems)

Applications Types of ANN

- Classification
  - Feedforward networks (MLP), radial basis function, and probabilistic NN
- Regression
  - Feedforward networks (MLP), radial basis function
- Clustering
  - Adaptive Resonance Theory (ART) and SOM
- Association
  - Hopfield networks
- Provide examples for each type?

Advantages of ANN

- Able to deal with (identify/model) highly nonlinear relationships
- Not prone to restricting normality and/or independence assumptions
- Can handle variety of problem types
- Usually provides better results (prediction and/or clustering) compared to its statistical counterparts
- Handles both numerical and categorical variables (transformation needed!)

Disadvantages of ANN

- They are deemed to be black-box solutions, lacking expandability
- It is hard to find optimal values for large number of network parameters
  - Optimal design is still an art: requires expertise and extensive experimentation
- It is hard to handle large number of variables (especially the rich nominal attributes)
- Training may take a long time for large datasets; which may require case sampling

ANN Software

- Standalone ANN software tool
  - NeuroSolutions
  - BrainMaker
  - NeuralWare
  - NeuroShell, … for more (see pcai.com) …
- Part of a data mining software suit
  - PASW (formerly SPSS Clementine)
  - SAS Enterprise Miner
  - Statistica Data Miner, … many more …