CAP 5510: Introduction to Bioinformatics
CGS 5166: Bioinformatics Tools
Giri NARASIMHAN

www.cis.fiu.edu/~giri/teach/BioinfF18.html
Machine Learning
Support Vector Machines

- Supervised Statistical Learning Method for:
  - Classification
  - Regression

- Simplest Version:
  - **Training:** Present series of labeled examples (e.g., gene expressions of tumor vs. normal cells)
  - **Validation:** Step to fine-tune hyperparameters
  - **Prediction:** Predict labels of new examples.
Applications

- **Text Categorization & Information Filtering**
  - 12,902 Reuters Stories, 118 categories (91% !!)

- **Image Recognition**
  - Face Detection, tumor anomalies, defective parts in assembly line, etc.

- **Gene Expression Analysis**

- **Protein Homology Detection**
Neural Networks
Artificial Neurons

![Diagram of an artificial neuron](image)

- **Inputs**: $x_1, x_2, \ldots, x_m$
- **Weights**: $w_0, w_1, w_2, \ldots, w_m$
- **Net input function**: $\sum w_i x_i$
- **Activation function**: $f(\sum w_i x_i)$

output
Neural Networks

two-layered NN (perceptron)

input layer

hidden layer(s)

output layer
Types of Neural Networks

asimovinstitute.org
RNNs and CNNs

Recurrent Neural Network (RNN)

Deep Convolutional Network (DCN)
CNNs
Autoencoders & Deep NNs

Auto Encoder (AE)
NNs for handwriting
Kohonen’s Self-Organizing Maps (SOMs)

Corpus of 12,088 newsgroup comp.ai.neural-nets contributions
Issues with NNs

- Training Sets
- Initial Weights
- Overfitting