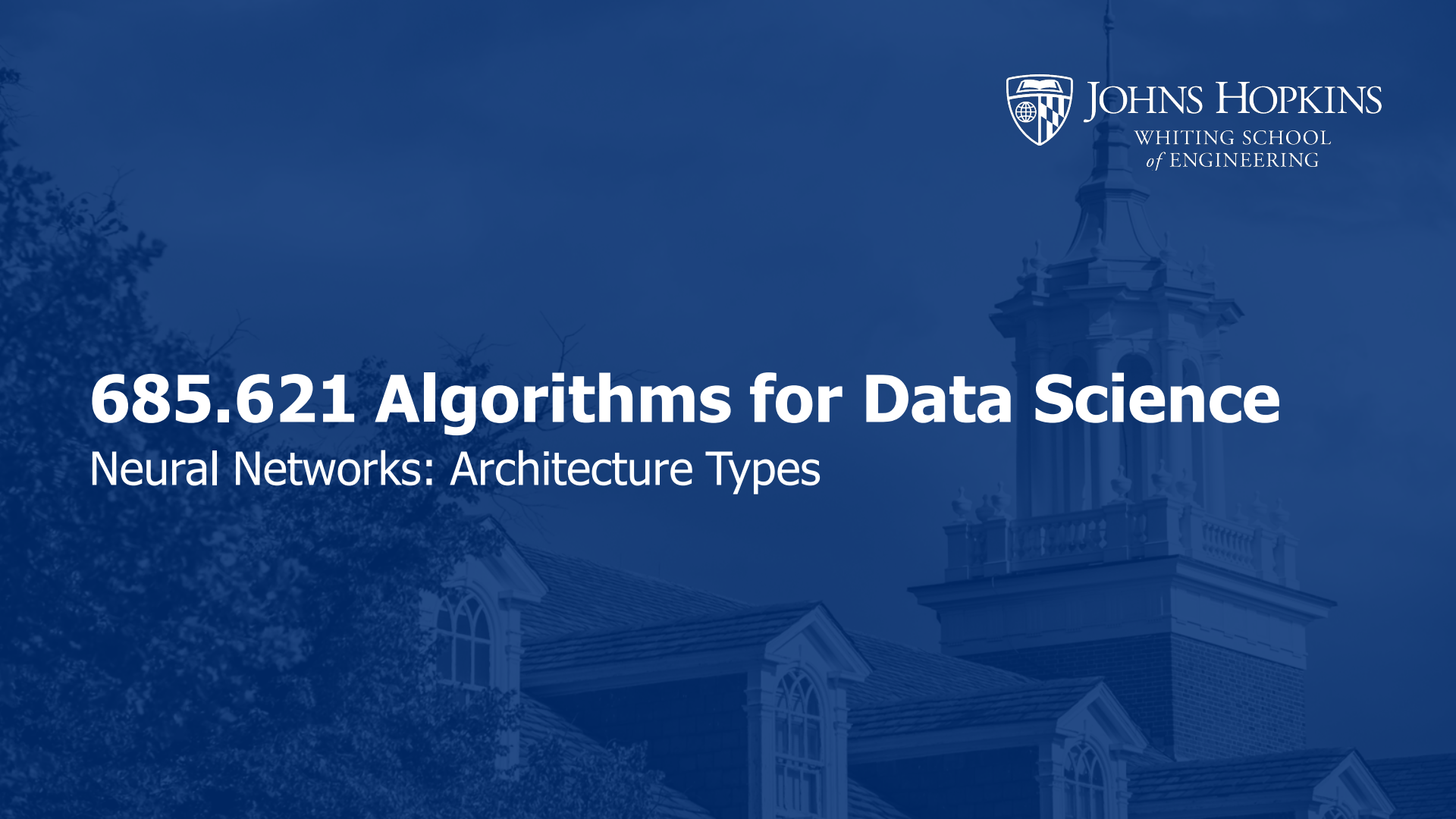
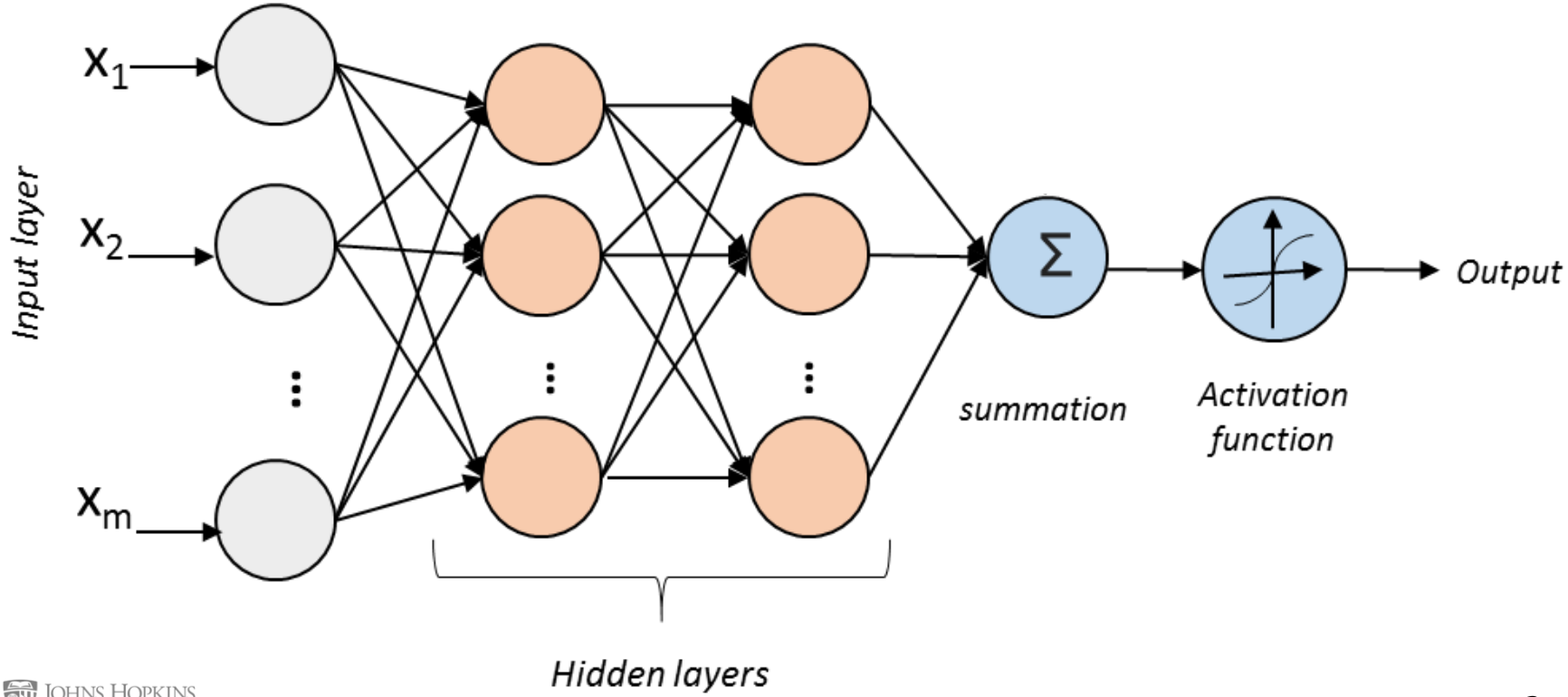


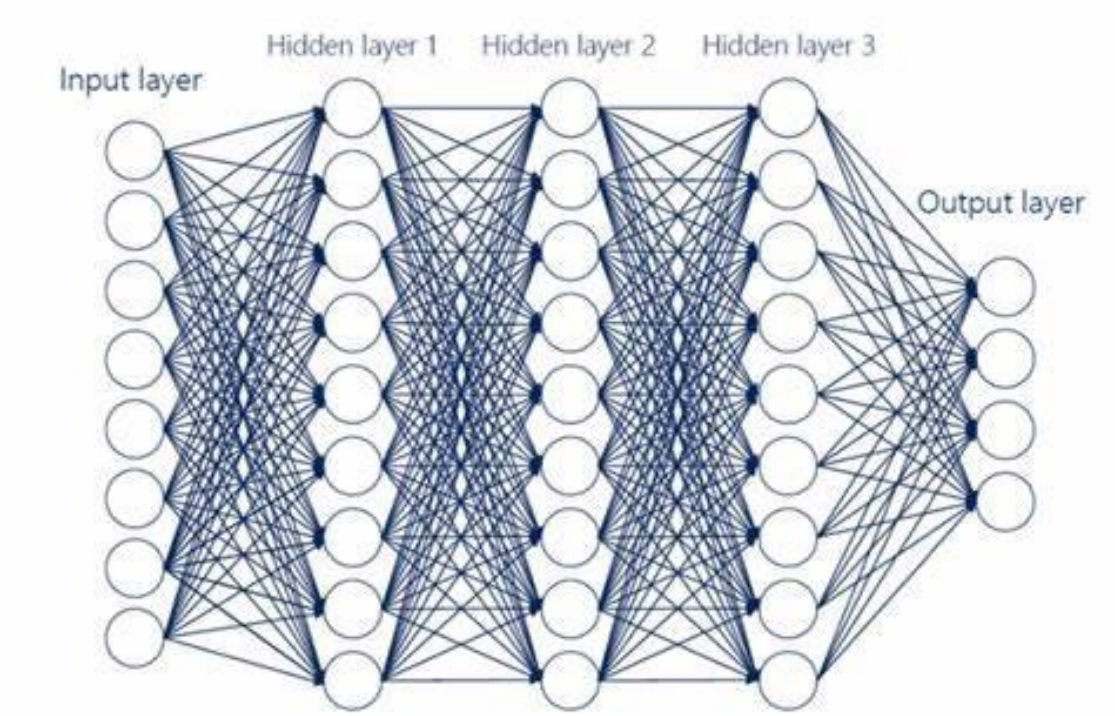
685.621 Algorithms for Data Science

Neural Networks: Architecture Types

Multi-Layer Perceptron (MLP)

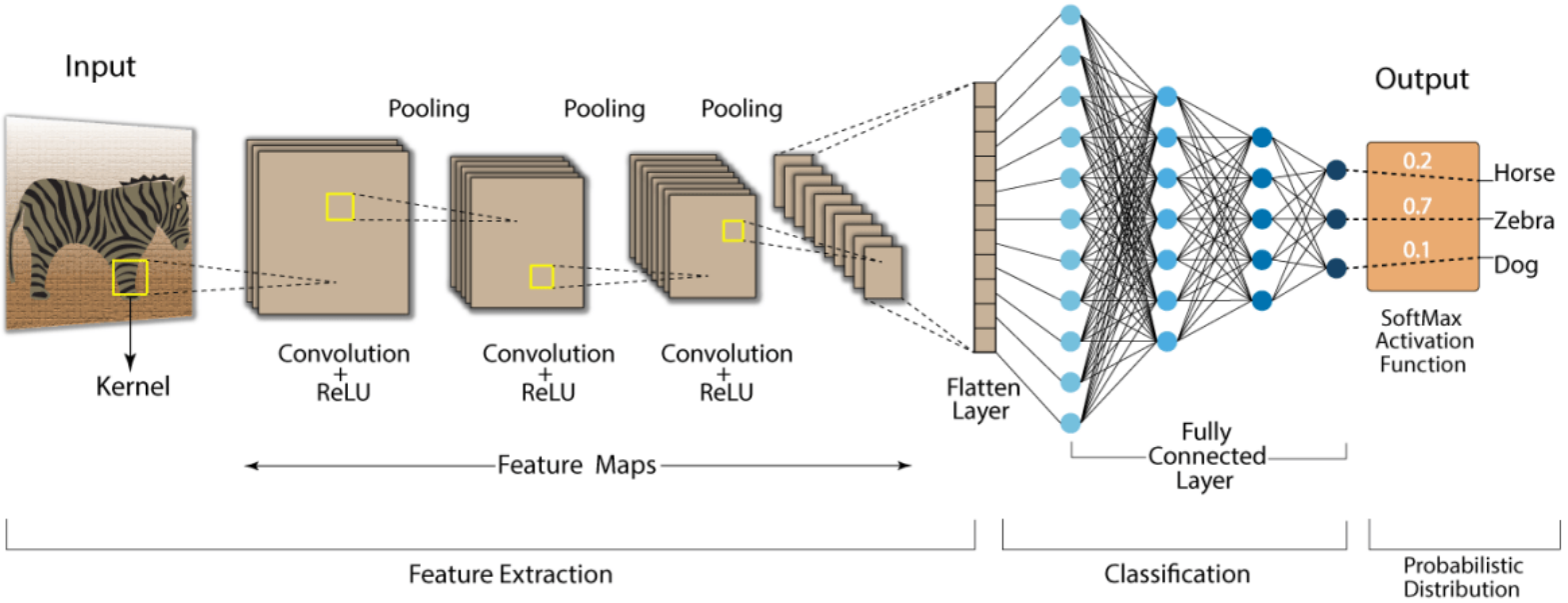


Deep Neural Networks (DNN)



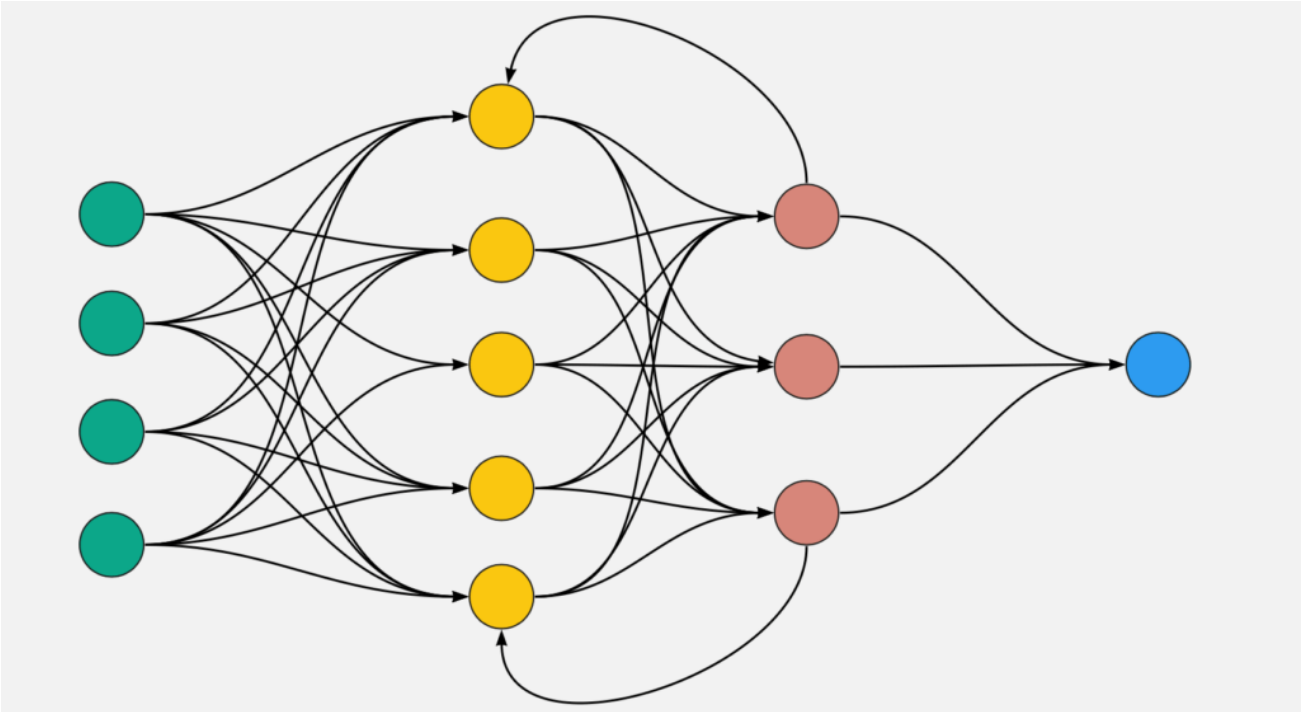
Merenda, Massimo & Porcaro, Carlo & Iero, Demetrio. (2020). Edge Machine Learning for AI-Enabled IoT Devices: A Review. Sensors. 20. 2533. 10.3390/s20092533.

Convolutional Neural Networks (CNN)



[All About Convolutional Neural Networks Cnns By Savindi Wijenayaka]

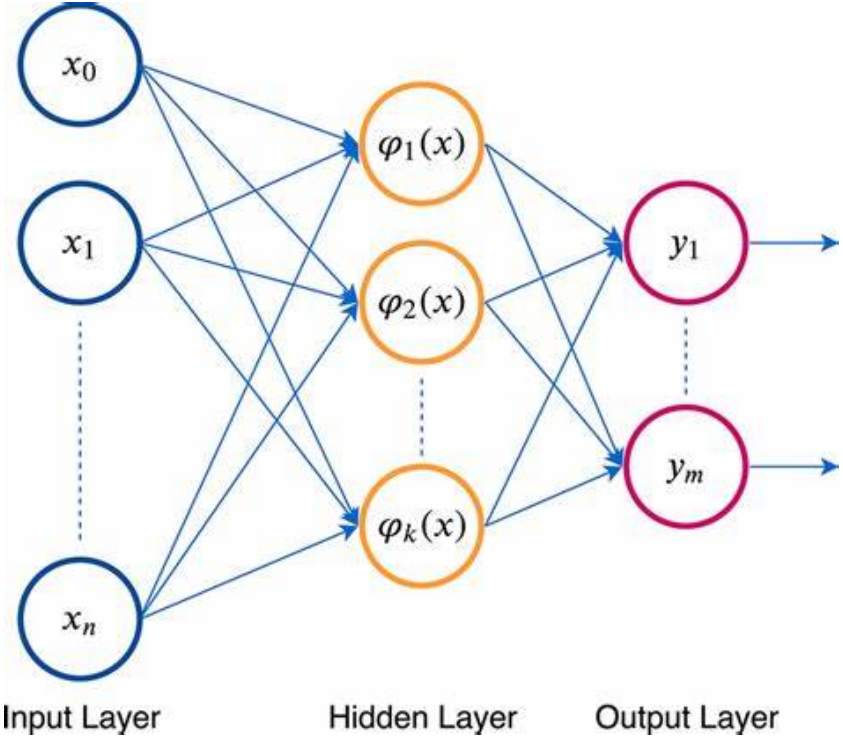
Recurrent Neural Networks (RNN)



How Recurrent Neural Network Works

Radial Basis Function (RBF) Network

- **Radial Basis Function** Networks take a different approach by using radial basis functions, typically **Gaussians**, as activation functions in the hidden layer.
- Each hidden neuron is '**activated**' based on how close the input is to its **center**.
- In practice, **center points** are chosen via clustering algorithms like **k-means**.



Dutta, Sagar & Basu, Banani & Talukdar, Fazal. (2020). Cascaded neural network based small array synthesis with robustness to noise. International Journal of RF and Microwave Computer-Aided Engineering. 31. 10.1002/mmce.22485.

Architectures Summary Slide

Architecture	Learning Type	Best For	Activation Type	Strengths	Limitations
MLP	Global	Tabular data, basic classification & regression	ReLU, Tanh, Sigmoid	General-purpose, universal approximation, easy to implement	Poor scaling with high-dimensional data, no locality
DNN	Global Hierarchical	Complex pattern recognition (text, images, speech)	ReLU	Learns hierarchical features, high expressive power	Risk of vanishing gradients, longer training time
CNN	Local to Global	Images, spatial data, video processing	ReLU	Spatial locality, parameter sharing, efficient for images	Not ideal for sequential or non-spatial data
RNN	Sequential	Time series, natural language, sequences	Tanh, ReLU, Sigmoid	Captures temporal dependencies, sequence-aware	Vanishing gradient over long sequences, slower training
RBF	Local	Small-scale function approximation, pattern recognition	Gaussian (Radial)	Fast convergence, localized learning, interpretable	Poor scalability, sensitive to center selection