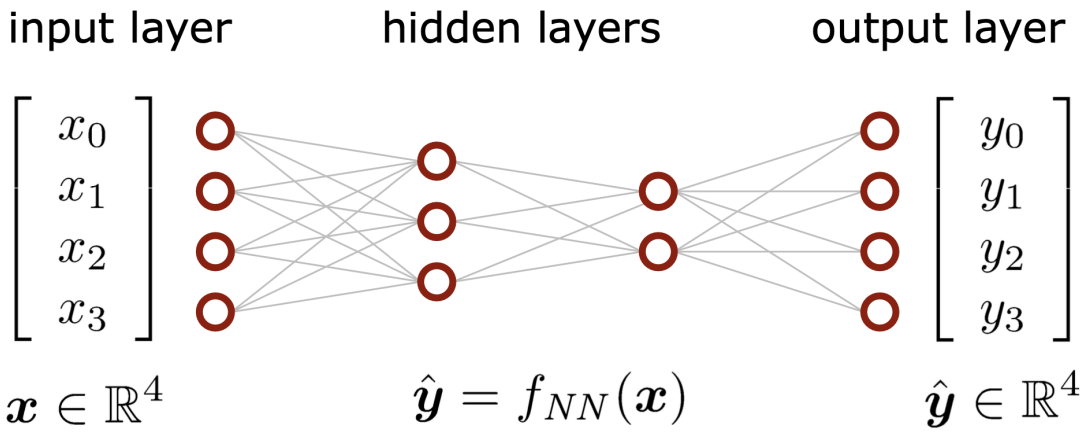
Week 5 - Introduction to ML

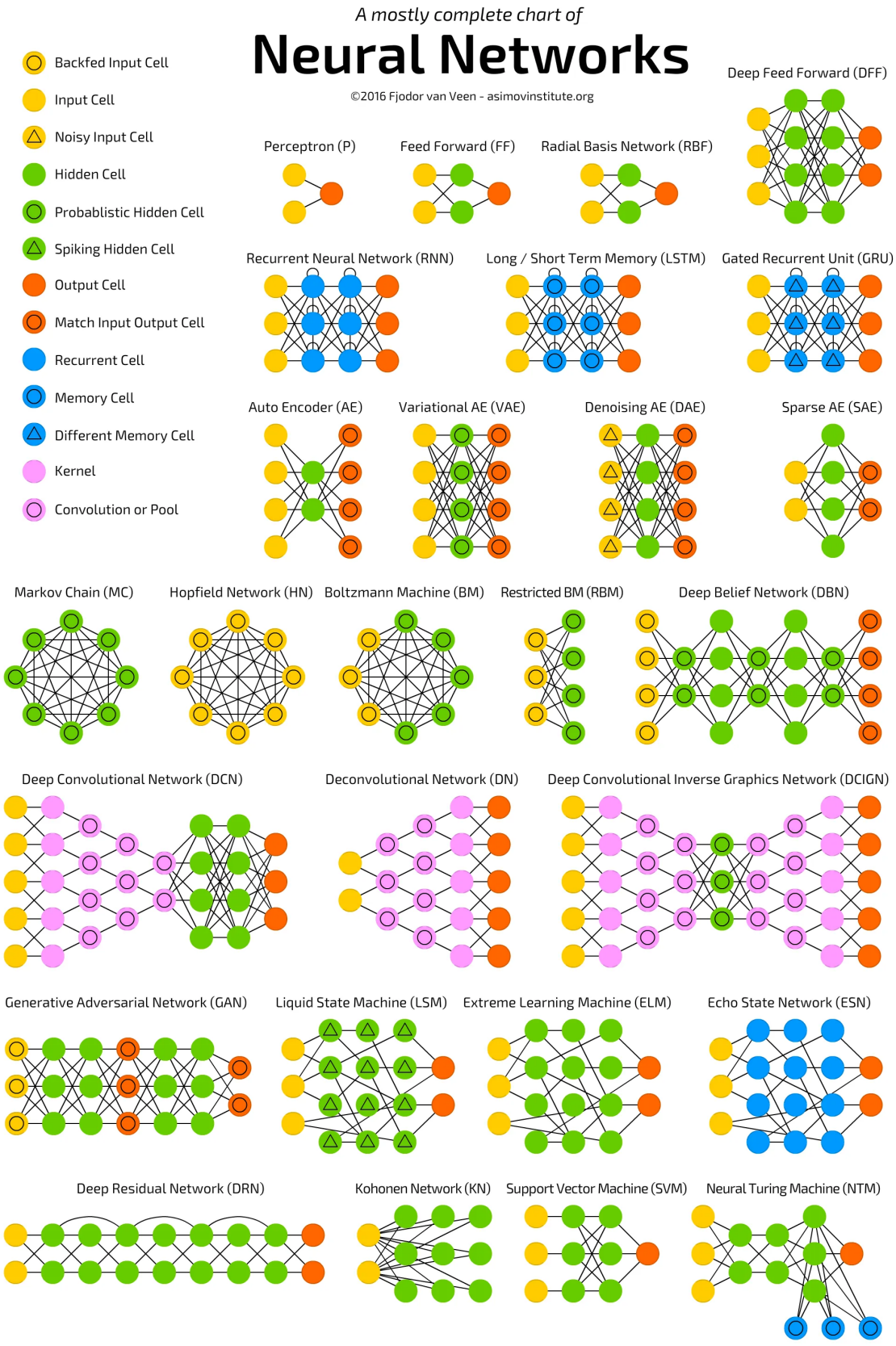
Introduction to ML

- **Machine Learning**
  - **overview**
    - Given some training data (and a proposed model), ML training finds optimal model parameters such that the prediction error is minimized
    - Training: minimize an error/loss function F(w) which depends upon the ML model and continuous w and the training data,

$$w^* = \arg\min_{w' \in W} F(w')$$

    - Typically, *w* is just D-dimensional Euclidean space R^D, but special problems (such as clustering) have other optimization spaces
    - Then use the trained parameters in the model, "$\star$, to make predictions about new unseen test data
    - ML algorithms usually categorized according to availability of labelled data: supervised, unsupervised, self-supervised, transfer learning
  - **Neural Networks are connected elementary (computing) units**
    - Multi-layer perceptron or MLP is often seen as the “vanilla” neural network


input layer
hidden layers
output layer
x_0
x_1
x_2
x_3
y_0
y_1
y_2
y_3
$\boldsymbol{x} \in \mathbb{R}^4$
$\hat{\boldsymbol{y}} = f_{NN}(\boldsymbol{x})$
$\hat{\boldsymbol{y}} \in \mathbb{R}^4$


A mostly complete chart of
Neural Networks
©2016 Fjodor van Veen - asimovinstitute.org
Backfed Input Cell
Input Cell
Noisy Input Cell
Hidden Cell
Probablistic Hidden Cell
Spiking Hidden Cell
Output Cell
Match Input Output Cell
Recurrent Cell
Memory Cell
Different Memory Cell
Kernel
Convolution or Pool
Perceptron (P)
Feed Forward (FF)
Radial Basis Network (RBF)
Deep Feed Forward (DFF)
Recurrent Neural Network (RNN)
Long / Short Term Memory (LSTM)
Gated Recurrent Unit (GRU)
Auto Encoder (AE)
Variational AE (VAE)
Denoising AE (DAE)
Sparse AE (SAE)
Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)
Deep Convolutional Network (DCN)
Deconvolutional Network (DN)
Deep Convolutional Inverse Graphics Network (DCIGN)
Generative Adversarial Network (GAN)
Liquid State Machine (LSM)
Extreme Learning Machine (ELM)
Echo State Network (ESN)
Deep Residual Network (DRN)
Kohonen Network (KN)
Support Vector Machine (SVM)
Neural Turing Machine (NTM)