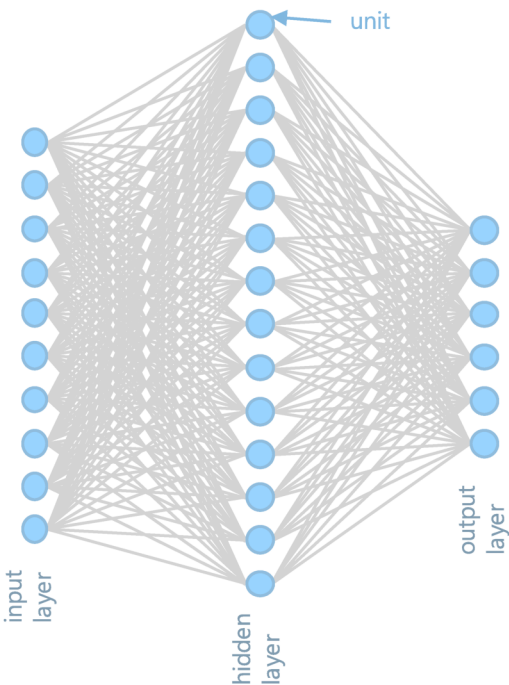


## Multi-Layer Perceptron



## Some practical differences

	ML Perceptrons	Decision Trees/DAGs
locality of computation	All "neurons" involved	<b>Conditional computation</b>
data transformation	<b>Projections, non-linearities</b>	Identity, copy
training algorithm	<b>Gradient descent, back-prop</b>	Greedy, randomized
structure	fixed ahead of time	learned and optimized
output	soft-max probabilities	histograms, probabilities
...	...	...

## Decision Tree

